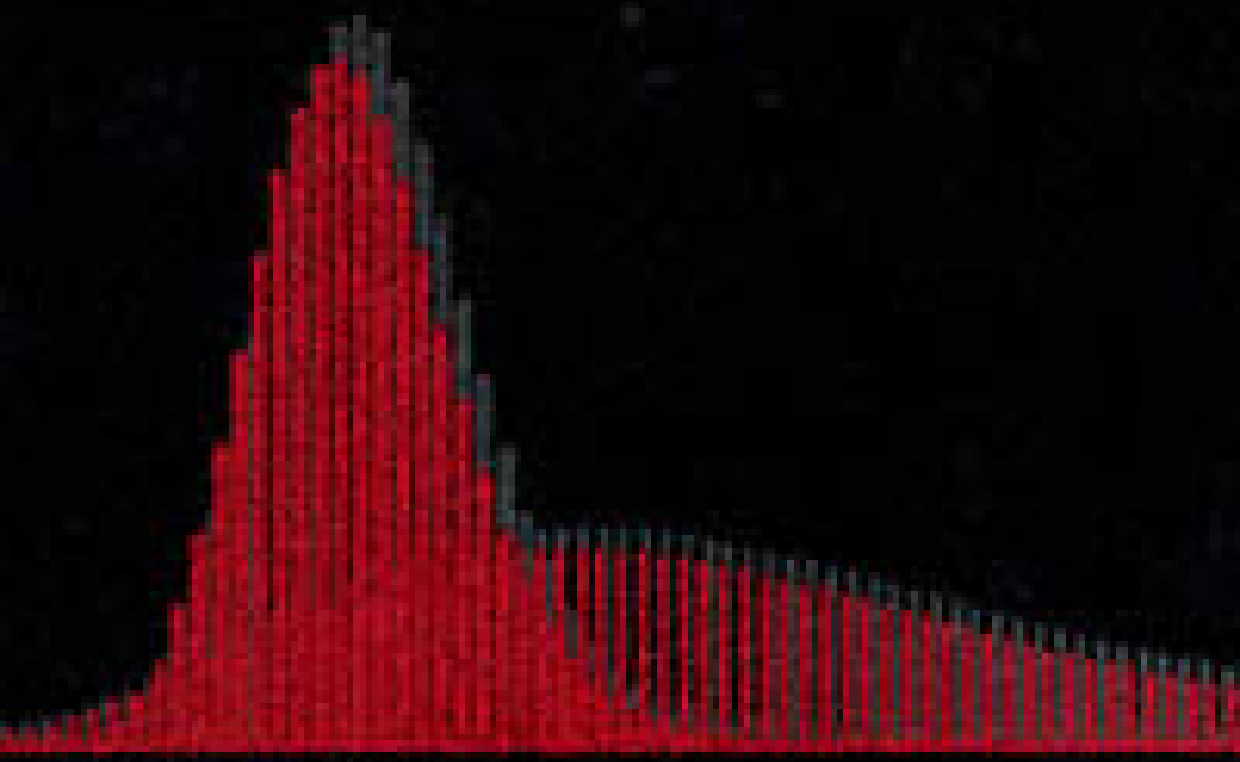
Quality and Reliability 55

Reliability Engineering and Risk Analysis

A Practical Guide



Mohammad Modarres
Mark Kaminskiy
Vasiliy Krivtsov

Reliability Engineering and Risk Analysis

QUALITY AND RELIABILITY

A Series Edited by

EDWARD G. SCHILLING
Coordinating Editor
Center for Quality and Applied Statistics
Rochester Institute of Technology
Rochester, New York

RICHARD S. BINGHAM, JR.
Associate Editor for
Quality Management
Consultant
Brooksville, Florida

LARRY RABINOWITZ
Associate Editor for
Statistical Methods
College of William and Mary
Williamsburg, Virginia

THOMAS WITT
Associate Editor for
Statistical Quality Control
Rochester Institute of Technology
Rochester, New York

1. Designing for Minimal Maintenance Expense: The Practical Application of Reliability and Maintainability, *Marvin A. Moss*
2. Quality Control for Profit: Second Edition, Revised and Expanded, *Ronald H. Lester, Norbert L. Enrick, and Harry E. Mottley, Jr.*
3. QCPAC: Statistical Quality Control on the IBM PC, *Steven M. Zimmerman and Leo M. Conrad*
4. Quality by Experimental Design, *Thomas B. Barker*
5. Applications of Quality Control in the Service Industry, *A. C. Rosander*
6. Integrated Product Testing and Evaluating: A Systems Approach to Improve Reliability and Quality, Revised Edition, *Harold L. Gilmore and Herbert C. Schwartz*
7. Quality Management Handbook, *edited by Loren Walsh, Ralph Wurster, and Raymond J. Kimber*
8. Statistical Process Control: A Guide for Implementation, *Roger W. Berger and Thomas Hart*

9. Quality Circles: Selected Readings, *edited by Roger W. Berger and David L. Shores*
10. Quality and Productivity for Bankers and Financial Managers, *William J. Latzko*
11. Poor-Quality Cost, *H. James Harrington*
12. Human Resources Management, *edited by Jill P. Kern, John J. Riley, and Louis N. Jones*
13. The Good and the Bad News About Quality, *Edward M. Schrock and Henry L. Lefevre*
14. Engineering Design for Producibility and Reliability, *John W. Priest*
15. Statistical Process Control in Automated Manufacturing, *J. Bert Keats and Norma Faris Hubele*
16. Automated Inspection and Quality Assurance, *Stanley L. Robinson and Richard K. Miller*
17. Defect Prevention: Use of Simple Statistical Tools, *Victor E. Kane*
18. Defect Prevention: Use of Simple Statistical Tools, Solutions Manual, *Victor E. Kane*
19. Purchasing and Quality, *Max McRobb*
20. Specification Writing and Management, *Max McRobb*
21. Quality Function Deployment: A Practitioner's Approach, *James L. Bossert*
22. The Quality Promise, *Lester Jay Wollschlaeger*
23. Statistical Process Control in Manufacturing, *edited by J. Bert Keats and Douglas C. Montgomery*
24. Total Manufacturing Assurance, *Douglas C. Brauer and John Cesarone*
25. Deming's 14 Points Applied to Services, *A. C. Rosander*
26. Evaluation and Control of Measurements, *John Mandel*
27. Achieving Excellence in Business: A Practical Guide to the Total Quality Transformation Process, *Kenneth E. Ebel*
28. Statistical Methods for the Process Industries, *William H. McNeese and Robert A. Klein*
29. Quality Engineering Handbook, *edited by Thomas Pyzdek and Roger W. Berger*
30. Managing for World-Class Quality: A Primer for Executives and Managers, *Edwin S. Shecter*
31. A Leader's Journey to Quality, *Dana M. Cound*
32. ISO 9000: Preparing for Registration, *James L. Lamprecht*
33. Statistical Problem Solving, *Wendell E. Carr*
34. Quality Control for Profit: Gaining the Competitive Edge. Third Edition, Revised and Expanded, *Ronald H. Lester, Norbert L. Enrick, and Harry E. Mottley, Jr.*
35. Probability and Its Applications for Engineers, *David H. Evans*
36. An Introduction to Quality Control for the Apparel Industry, *Pradip V. Mehta*
37. Total Engineering Quality Management, *Ronald J. Cottman*
38. Ensuring Software Reliability, *Ann Marie Neufelder*
39. Guidelines for Laboratory Quality Auditing, *Donald C. Singer and Ronald P. Upton*
40. Implementing the ISO 9000 Series, *James L. Lamprecht*
41. Reliability Improvement with Design of Experiments, *Lloyd W. Condra*
42. The Next Phase of Total Quality Management: TQM II and the Focus on Profitability, *Robert E. Stein*

43. Quality by Experimental Design: Second Edition, Revised and Expanded, *Thomas B. Barker*
44. Quality Planning, Control, and Improvement in Research and Development, *edited by George W. Roberts*
45. Understanding ISO 9000 and Implementing the Basics to Quality, *D. H. Stamatis*
46. Applying TQM to Product Design and Development, *Marvin A. Moss*
47. Statistical Applications in Process Control, *edited by J. Bert Keats and Douglas C. Montgomery*
48. How to Achieve ISO 9000 Registration Economically and Efficiently, *Gurmeet Naroola and Robert Mac Connell*
49. QS-9000 Implementation and Registration, *Gurmeet Naroola*
50. The Theory of Constraints: Applications in Quality and Manufacturing: Second Edition, Revised and Expanded, *Robert E. Stein*
51. Guide to Preparing the Corporate Quality Manual, *Bernard Froman*
52. TQM Engineering Handbook, *D. H. Stamatis*
53. Quality Management Handbook: Second Edition, Revised and Expanded, *edited by Raymond J. Kimber, Robert W. Grenier, and John Jourdan Heldt*
54. Multivariate Quality Control: Theory and Applications, *Camil Fuchs and Ron S. Kenett*
55. Reliability Engineering and Risk Analysis: A Practical Guide, *Mohammad Modarres, Mark Kaminskiy, and Vasiliy Krivtsov*

ADDITIONAL VOLUMES IN PREPARATION

Reliability Engineering and Risk Analysis

A Practical Guide

Mohammad Modarres
University of Maryland
College Park, Maryland

Mark Kaminskiy
QUALCOMM, Inc.
San Diego, California

Vasiliy Krivtsov
Ford Motor Company
Dearborn, Michigan

MARCEL DEKKER, INC. NEW YORK • BASEL

Library of Congress Cataloging-in-Publication Data
Modarres, M. (Mohammad)
Reliability engineering and risk analysis / Mohammad Modarres, Mark Kaminskiy, Vasiliy Krivtsov.
p. cm. — (Quality and reliability ; 55)
Includes bibliographical references (p.).
ISBN 0-8247-2000-8 (alk. paper)
1. Reliability (Engineering) 2. Risk assessment. I. Kaminskiy, Mark. II. Krivtsov, Vasiliy. III. Title.
IV. Series.
TA169.M627 1999
620N00452—dc21 99-26668
CIP

This book is printed on acid-free paper.

Headquarters
Marcel Dekker, Inc.
270 Madison Avenue, New York, NY 10016
tel: 212-696-9000; fax: 212-685-4540

World Wide Web http://www.dekker.com

The publisher offers discounts on this book when ordered in bulk quantities. For more information, write to Special Sales/Professional Marketing at the headquarters address above.

Current printing (last digit):
10 9 8 7 6 5 4 3 2 1

PRINTED IN THE UNITED STATES OF AMERICA

About the Series

The genesis of modern methods of quality and reliability will be found in a sample memo dated May 16, 1924, in which Walter A. Shewhart proposed the control chart for the analysis of inspection data. This led to a broadening of the concept of inspection from emphasis on detection and correction of defective material to control of quality through analysis and prevention of quality problems. Subsequent concern for product performance in the hands of the user stimulated development of the systems and techniques of reliability. Emphasis on the consumer as the ultimate judge of quality serves as the catalyst to bring about the integration of the methodology of quality with that of reliability. Thus, the innovations that came out of the control chart spawned a philosophy of control of quality and reliability that has come to include not only the methodology of the statistical sciences and engineering, but also the use of appropriate management methods together with various motivational procedures in a concerted effort dedicated to quality improvement.

This series is intended to provide a vehicle to foster interaction of the elements of the modern approach to quality, including statistical applications, quality and reliability engineering, management, and motivational aspects. It is a forum in which the subject matter of these various areas can be brought together to allow for effective integration of appropriate techniques. This will promote the true benefit of each, which can be achieved only through their interaction. In this sense, the whole of quality and reliability is greater than the sum of its parts, as each element augments the others.

The contributors to this series have been encouraged to discuss fundamental concepts as well as methodology, technology, and procedures at the leading edge of the discipline. Thus, new concepts are placed in proper perspective in these evolving disciplines. The series is intended for those in manufacturing,

engineering, and marketing and management, as well as the consuming public, all of whom have an interest and stake in the products and services that are the lifeblood of the economic system.

The modern approach to quality and reliability concerns excellence: excellence when the product is designed, excellence when the product is made, excellence as the product is used, and excellence throughout its lifetime. But excellence does not result without effort, and products and services of superior quality and reliability require an appropriate combination of statistical, engineering, management, and motivational effort. This effort can be directed for maximum benefit only in light of timely knowledge of approaches and methods that have been developed and are available in these areas of expertise. Within the volumes of this series, the reader will find the means to create, control, correct, and improve quality and reliability in ways that are cost effective, that enhance productivity, and that create a motivational atmosphere that is harmonious and constructive. It is dedicated to that end and to the readers whose study of quality and reliability will lead to greater understanding of their products, their processes, their workplaces, and themselves.

Edward G. Schilling

Preface

This book provides a practical and comprehensive overview of reliability and risk analysis techniques. It is written both for engineering students at the undergraduate and graduate levels and for practicing engineers. The book concentrates mainly on reliability analysis. In addition, elementary performance and risk analysis techniques have also been presented. Since reliability analysis is a multi-disciplinary subject, the scope is not limited to any one engineering discipline; rather, the material is applicable to most engineering disciplines. The contents of the book are based primarily on the materials used in three courses at the senior and undergraduate and graduate levels at the University of Maryland, College Park. These courses have been offered over the past 15 years. This book has benefited greatly from the contribution of many talented students who actively participated in gathering and updating practical and useful materials. Therefore, the book presents a large number of examples to clarify technical subjects. Additionally, there are many end-of-chapter exercises that are based mainly on the prior exams and homework problem sets of the reliability and risk analysis courses at the University of Maryland.

The emphasis of the book is the introduction and explanation of methods and techniques used in reliability and risk studies, and discussion of their use and limitations rather than detailed derivations. These methods and techniques cover a wide range of topics that arise in routine reliability engineering and risk analysis activities. The book assumes that the readers have little or no background in probability and statistics. Thus, following an introductory chapter (Chapter 1) that defines reliability, availability, and risk analysis, Chapter 2 provides a detailed review of probability and statistics essential to understanding the reliability methods discussed in the book.

We have developed a software tool (called RARE—Reliability and Risk Evaluator) using Microsoft's Excel™ that automates the techniques and methods

discussed in this book. This software is educational in nature and is intended to help students practice applications of these important methods and to reduce the computational burden. Furthermore, it is a finalized software tool that can be used to analyze a wide variety of real-world reliability data. The RARE software has a simple interface that allows users to easily perform their calculations. Also, the results of the calculations are summarized in a useful and simple graphical and/or tabular format for report generation. The reliability methods for which an application routine has been developed in RARE have been clearly identified in the book. Appendix D contains a user's manual for RARE.

We have structured the book so that reliability methods applied to component reliability are described first (Chapter 3). This is because components are the most basic building blocks of engineering systems. The techniques discussed in Chapter 3 provide a comprehensive overview of the state-of-the-art in transforming basic field data into estimates of component reliability. Chapter 4 describes these analytical methods in the context of a more complex engineering unit; that is, a system containing many interacting components. Chapter 4 introduces new analysis methods using the results of the component reliability analysis described in Chapter 3 to calculate estimates of the reliability of the whole system that is composed of these components. The material in Chapters 1 through 4 are appropriate for an advanced undergraduate course in reliability engineering, or an introductory graduate level course.

Chapters 3 and 4 assume that the components (or systems) are "replaceable." That is, upon a failure, the component is replaced with a new one. However, many components are "repairable." That is, upon a failure they are repaired and placed back into service. In this case, availability as a metric becomes the key measure of performance. The techniques for availability and reliability analysis of repairable components and systems are discussed in Chapter 5. This chapter also explains the corresponding use of the analytical methods discussed in Chapters 3 and 4 when performing availability analysis of components and engineering systems.

Chapter 6 discusses a number of important methods frequently used in modeling reliability, availability, and risk problems. For example, in Section 6.2, we discuss the concept of uncertainty, sources of uncertainty, parameter and model uncertainty, and probabilistic methods for quantifying and propagating parameter uncertainties in engineering systems (or models). Similar to Chapter 6, Chapter 7 describes special topics related to reliability and availability data analysis. For example Section 7.1 describes accelerated life testing methods. Examples clarifying uses of modeling and data analysis methods and their shortcomings are also presented in Chapters 6 and 7.

In Chapter 8, we discuss the method of risk analysis. A number of the analytical methods explained in the preceding chapters have been integrated to

perform risk assessment or risk management. Recently, probabilistic risk assessment (PRA) has been a major topic of interest in light of hazards imposed by many engineering designs and processes. Steps involving performance of a PRA are discussed in this chapter.

A complete solution set booklet has been developed by W. M. Webb and M. Modarres. This booklet may be provided to educators and industrial users by sending a written request to the publisher.

The authors are especially thankful to Dr. Daniel Young for writing the discussions on Electrical Failure Mechanisms in Section 1.3. The book could have not been completed without the help and corrections of our students and colleagues at the University of Maryland. It would be difficult to name all, but some names to mention include: Drs. L. Chen, H. Dezfuli, Y. Guan, H. Hadavi, K. Hsueh, N. Kececi, D. Koo, A. Mosleh, J. Ruiz, C. Smidts, and J. N. Wang. We would also like to acknowledge J. Case of Ford Motor Company for his review and valuable suggestions on the FMEA and reliability growth sections of the book. Special thanks goes to Y. S. Hu for his unfailing technical and organizational support without which this work would have not been possible. Finally, the editorial help of D. Grimsman and typing and graphical support of W. M. Webb are highly appreciated.

Mohammad Modarres
Mark Kaminskiy
Vasiliy Krivtsov

This page intentionally left blank

Contents

About the Series Edward G. Schilling iii
Preface v

1 Reliability Analysis in Perspective **1**
1.1 Why Study Reliability? 1
1.2 Failure Models 2
1.3 Failure Mechanisms 4
1.4 Performance Measures 9
1.5 Definition of Reliability 14
1.6 Definition of Availability 16
1.7 Definition of Risk 17
References 18

2 Basic Reliability Mathematics: Review of Probability and Statistics **21**
2.1 Introduction 21
2.2 Elements of Probability 21
2.2.1 Sets and Boolean Algebra 21
2.2.2 Basic Laws of Probability 26
2.2.3 Bayes' Theorem 33
2.3 Probability Distributions 38
2.3.1 Random Variable 39
2.3.2 Some Basic Discrete Distributions 39
2.3.3 Some Basic Continuous Distributions 47
2.3.4 Joint and Marginal Distributions 61

2.4 Basic Characteristics of Random Variables 65
2.5 Estimation and Hypothesis Testing 73
2.5.1 Point Estimation 74
2.5.2 Interval Estimation and Hypothesis Testing 78
2.6 Frequency Tables and Histograms 81
2.7 Goodness-of-Fit Tests 83
2.7.1 Chi-Square Test 83
2.7.2 Kolmogorov Test 87
2.8 Regression Analysis 92
Exercises 97
References 103

3 Elements of Component Reliability 105
3.1 Concept of Reliability 105
3.1.1 Reliability Function 106
3.1.2 Failure Rate 107
3.2 Common Distributions in Component Reliability 115
3.2.1 Exponential Distribution 115
3.2.2 Weibull Distribution 116
3.2.3 Gamma Distribution 118
3.2.4 Normal Distribution 120
3.2.5 Lognormal Distribution 120
3.2.6 Extreme Value Distributions 121
3.3 Component Reliability Model 127
3.3.1 Graphical Nonparametric Procedures 127
3.3.2 Probability Plotting 133
3.3.3 Total-Time-on-Test Plots 141
3.4 Classical Parametric Estimation 144
3.4.1 Exponential Distribution Point Estimation 147
3.4.2 Exponential Distribution Interval Estimation 150
3.4.3 Lognormal Distribution 154
3.4.4 Weibull Distribution 155
3.4.5 Binomial Distribution 156
3.5 Classical Nonparametric Distribution Estimation 158
3.5.1 Confidence Intervals for Cumulative Distribution Function and Reliability Function for Complete and Singly Censored Data 158
3.5.2 Confidence Intervals for Cumulative Distribution Function and Reliability Function for Multiply Censored Data 162
3.6 Bayesian Estimation Procedures 164

3.6.1 Estimation of the Parameter λ of Exponential Distribution 166
3.6.2 Bayesian Estimation of the Parameter of Binomial Distribution 173
3.7 Methods of Generic Failure Rate Determination 185
Exercises 186
References 194

4 System Reliability Analysis 197
4.1 Reliability Block Diagram Method 198
4.1.1 Series System 198
4.1.2 Parallel Systems 200
4.1.3 Standby Redundant Systems 203
4.1.4 Load-Sharing Systems 207
4.1.5 Complex Systems 209
4.2 Fault Tree and Success Tree Methods 215
4.2.1 Fault Tree Method 215
4.2.2 Evaluation of Logic Trees 219
4.2.3 Success Tree Method 232
4.3 Event Tree Method 235
4.3.1 Construction of Event Trees 235
4.3.2 Evaluation of Event Trees 237
4.4 Master Logic Diagram 238
4.5 Failure Mode and Effect Analysis 248
4.5.1 Types of FMEA 249
4.5.2 FMEA/FMECA Procedure 250
4.5.3 FMEA Implementation 250
4.5.4 FMECA Procedure: Criticality Analysis 262
Exercises 268
References 279

5 Reliability and Availability of Repairable Items 281
5.1 Repairable System Reliability 282
5.1.1 Basic Random Processes Used as Probabilistic Models of Repairable Systems 282
5.1.2 Statistical Data Analysis for Repairable Systems 289
5.1.3 Data Analysis for the HPP 290
5.1.4 Data Analysis for NHPP 295
5.2 Availability of Repairable Systems 306
5.2.1 Instantaneous (Point) Availability 307
5.2.2 Limiting Point Availability 310
5.2.3 Average Availability 311

5.3 Use of Markovian Methods for Determining System Availability 312
5.4 Use of System Analysis Techniques in the Availability Calculations of Complex Systems 319
Exercises 327
References 330

6 Selected Topics in Reliability Modeling 333
6.1 Stress-Strength Analysis 333
6.2 Software Reliability Analysis 338
6.2.1 Introduction 338
6.2.2 Software Reliability Models 339
6.2.3 Software Life Cycle Models 345
6.3 Human Reliability 346
6.3.1 Human Reliability Analysis Process 347
6.3.2 HRA Models 352
6.3.3 Human Reliability Data 359
6.4 Measures of Importance 360
6.4.1 Birnbaum Measure of Importance 360
6.4.2 Criticality Importance 362
6.4.3 Fussell–Vesely Importance 363
6.4.4 Risk Reduction Worth Importance 364
6.4.5 Risk Achievement Worth Importance 365
6.4.6 Practical Aspects of Importance Measures 369
6.5 Reliability-Centered Maintenance 370
6.5.1 History and Current Procedures 370
6.5.2 Optimal Preventive Maintenance Scheduling 374
6.6 Reliability Growth 376
6.6.1 Graphical Method 377
6.6.2 Duane Method 377
6.6.3 Army Material Systems Analysis Activity (AMSAA) Method 381
Exercises 383
References 385

7 Selected Topics in Reliability Data Analysis 389
7.1 Accelerated Life Testing 389
7.1.1 Basic Accelerated Life Notions 389
7.1.2 Some Popular AL (Reliability) Models 393
7.1.3 Accelerated Life Data Analysis 394

7.1.4 Accelerated Life Model for Time-Dependent Stress 401
7.1.5 Exploratory Data Analysis for Time-Dependent Stress 405
7.2 Analysis of Dependent Failures 408
7.2.1 Single Parameter Models 412
7.2.2 Multiple Parameter Models 415
7.2.3 Data Analysis for Common Cause Failures 419
7.3 Uncertainty Analysis 421
7.3.1 Types of Uncertainty 423
7.3.2 Uncertainty Propagation Methods 425
7.3.3 System Reliability Confidence Limits Based on Component Failure Data 430
7.3.4 Maximus Method 432
7.3.5 Graphic Representation of Uncertainty 441
7.4 Use of Expert Opinion for Estimating Reliability Parameters 442
7.4.1 Geometric Averaging Technique 445
7.4.2 Bayesian Approach 446
7.4.3 Statistical Evidence on the Accuracy of Expert Estimates 447
7.5 Probabilistic Failure Analysis 448
7.5.1 Detecting Trends in Observed Failure Events 450
7.5.2 Failure Rate and Failure Probability Estimation for Data with No Trend 450
7.5.3 Failure Rate and Failure Probability Estimation for Data with Trend 451
7.5.4 Evaluation of Statistical Data 451
7.5.5 Root-Cause Analysis 453
Exercises 456
References 457

8 Risk Analysis 461
8.1 Risk Perception and Acceptability 461
8.1.1 Risk Perception 461
8.1.2 Risk Acceptability 462
8.2 Determination of Risk Values 465
8.3 Formalization of Risk Assessment 468
8.4 Steps in Conducting a Probabilistic Risk Assessment 470
8.4.1 Methodology Definition 470
8.4.2 Familiarization and Information Assembly 471
8.4.3 Identification of Initiating Events 472
8.4.4 Sequence or Scenario Development 475
8.4.5 System Analysis 476
8.4.6 Internal Events External to the Facility 476

8.4.7 External Events 477
8.4.8 Dependent Failure Considerations 477
8.4.9 Failure Data Analysis 478
8.4.10 Quantification 479
8.5 A Simple Example of Risk Analysis 480
8.6 Precursor Analysis 494
8.6.1 Introduction 494
8.6.2 Basic Methodology 495
8.6.3 Categorization and Selection of Precursor Events 496
8.6.4 Properties of Precursor Estimator for the Occurrence Rate of Hazard Exposure Events and Its Interpretation 497
8.6.5 Applications of Precursor Analysis 500
8.6.6 Differences Between Precursor Analysis and Probabilistic Risk Assessments 502
References 503

Appendix A: Statistical Tables 505
Table A.1 Standard Normal Distribution Table 506
Table A.2 Percentiles of the t Distribution 507
Table A.3 Percentiles of the χ^2 Distribution 508
Table A.4 Critical Values D_n^{γ} for the Kolmogorov Goodness-of-Fit Test 509
Table A.5a Percentage Points of the F-Distribution (90th Percentile Values of the F-Distribution) 510
Table A.5b Percentage Points of the F-Distribution (95th Percentile Values of the F-Distribution) 511
Table A.5c Percentage Points of the F-Distribution (99th Percentile Values of the F-Distribution) 512

Appendix B: Generic Failure Data 513
Table B.1 Generic Failure Data for Mechanical Items 514

Appendix C: Software for Reliability and Risk Analyses 519
Table C.1 Selected PC-Based Software for Logic (Boolean-Based) Analysis 520
Table C.2 Capabilities of Other PC-Based Software 523

Appendix D: Reliability Analysis and Risk Evaluator (RARE) Quick User's Manual **525**
D.1 Introduction 526
D.2 RARE Installation 526
D.2.1 Hardware and Software Requirements 526
D 2.2 Installation Procedure 528
D.3 Disclaimer 528
D.4 Running RARE Programs 528
D.4.1 Main Controls Program 528
D.4.2 Goodness of Fit Program 529
D.4.3 Nonparametric Estimation Program 530
D.4.4 Sample Size Estimation Program 531
D.4.5 Distribution Program 532
D.4.6 Exponential Distribution Estimation Program 533
D.4.7 Interval Estimation Program 534
D.4.8 Bayesian Analysis Program 535
D.4.9 Repairable System Analysis Program 536

Index **539**

This page intentionally left blank

Reliability Engineering and Risk Analysis

This page intentionally left blank

1

Reliability Analysis in Perspective

1.1 WHY STUDY RELIABILITY?

Engineering systems, components and devices are not perfect. A perfect design is one that remains operational and attains system's objective without failure during a preselected life. This is the deterministic view of an engineering system. This view is idealistic, impractical, and economically infeasible. Even if technical knowledge is not a limiting factor in designing, manufacturing, constructing and operating a perfect design, the cost of development, testing, materials and engineering analysis may far exceed economic prospects for such a system. Therefore, practical and economical limitations dictate the use of not-so-perfect designs. Designers, manufacturers and end users, however, strive to minimize the occurrence and recurrence of failures. In order to minimize failures in engineering systems, the designer must understand "why" and "how" failures occur. This would help them prevent failures. In order to maximize system performance and efficiently use resources, it is also important to know how often such failures may occur. This involves *predicting* the occurrence of failures.

The prevention of failures and the process of understanding why and how failures occur involves appreciation of the physics of failure. Failure mechanisms are the means by which failures occur. To effectively minimize the occurrence of failures, the designer should have an excellent knowledge of failure mechanisms which may be inherently associated with the design, or can be introduced from outside of the system (e.g., by users or maintainers). When failure mechanisms are known and appropriately considered in design, manufacturing, construction, production and operation, they can be "minimized" or the system can be "protected" against them through careful engineering and economic analysis. This is generally a deterministic reliability analysis process.

All potential failures in a design are generally not known or well understood. Accordingly, the prediction of failures is inherently a probabilistic problem.

Therefore, reliability analysis is also a probabilistic process. This book deals with the reliability analyses involving prediction of failures and deals with it probabilistically. However, a brief review of failure mechanisms and failure prevention issues is presented in Sections 1.2 and 1.3.

1.2 FAILURE MODELS

Failures are the result of the existence of source *challenges* and conditions occurring in a particular scenario. The system has an inherent *capacity* to withstand such challenges, which capacity may be reduced by specific internal or external conditions. When challenges surpass the capacity of the system, a failure may occur.

Specific models use different definitions and metrics for capacity and challenge. "Adverse Conditions" generated artificially or naturally, internally or externally, may increase or induce challenges to the system, and/or reduce the capacity of the item to withstand challenges.

Figure 1.1 depicts elements of a framework to construct failure models. Several simple failure models, discussed by Dasgupta and Pecht (1991), are consistent with the framework presented in Figure 1.1. A summary of these models have been provided below.

Stress-Strength Model. The item (e.g., a system barrier or device) fails if and only if the challenge (i.e., stress) exceeds the capacity (i.e., strength). The *stress* represents an aggregate of the challenges and external conditions. This failure model may depend on environmental conditions or the occurrence of critical events, rather than the mere passage of time or cycles. Strength is often treated as a random variable representing effect of all conditions affecting the strength, or lack of knowledge about the item's strength (e.g., the item's capability, mechanical strength, and dexterity). Two examples of this model are: (a) a steel bar in axial tension, and (b) a transistor with a voltage applied across the emitter-collector.

Damage-Endurance Model. This model is similar to the stress-strength model, but the scenario of interest is that *stress* causes damage that accumulates irreversibly, as in corrosion, wear, embrittlement, and fatigue. The aggregate of challenges and external conditions leads to the metric represented as cumulative damage. The cumulative damage may not degrade performance. The item fails when and only when the cumulative damage exceeds the endurance, (i.e., the damage accumulates until the endurance of the item is reached). As such, an item's capacity is measured by its tolerance of damage endurance. Accumulated damage does not disappear when the *stresses* are removed, although sometimes treatments such as annealing are possible. Endurance is often treated as a random variable.

Similar to the stress-strength model, endurance is an aggregate measure for effects of challenges and external conditions on the item's capability to withstand cumulative stresses.

Challenge-Response Model. This model closely resembles the framework shown in Fig. 1.1. An element of the system may have failed, but only when the element is challenged (needed) does it cause the system to fail. A common consumer example is the emergency brake of a car. Most computer program (software) failures are of this type. Telephone switching systems also resemble this failure model. This failure model depends on when critical events happen in the environment, rather than the mere passage of time or cycles.

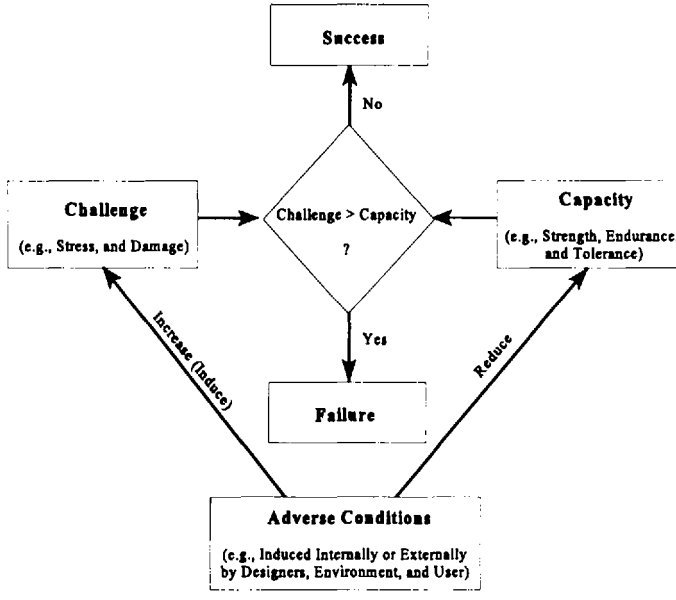


Figure 1.1 Framework for modeling failure.

Tolerance-Requirements Model. A system performance characteristic is satisfactory if and only if it falls within acceptable tolerance limits. Examples of this are a copier machine, and a measuring instrument where gradual degradation eventually results in a user deciding that performance quality is unacceptable.

In the models discussed above, challenges are caused by failure-inducing agents. Two of the most important failure inducing agents are "stress" and "time." Stress can be created due to mechanical, thermal, electrical, chemical, and radiation-induced forces, for example, by turning on and off a standby-component. Passage of time, on the other hand gives more opportunity for stress to be accumulated (or accumulative of damage).

A comprehensive consideration of reliability requires analysis of the two failure-inducing agents of stress and time. Both time and stress may be analyzed deterministically (e.g., by identifying the sources of stress), or probabilistically (e.g., by treating stress and time as random variables). In either case, it is necessary to understand why and how such stresses lead to a failure. This requires studying physics of failure in general and failure mechanisms in particular.

The main body of this book addresses the probabilistic treatment of time and stress as agents of failure. Equally important, however, is understanding the failure mechanisms and the physics of failure. In the next section, we discuss a brief summary of these mechanisms. For further readings on physics of failure, we refer the reader to a number of other books and articles. For example, Dasgupta and Pecht (1991), Pecht (1991), Collins (1993), and Amerasekera and Campbell (1987).

It is important to differentiate between stress-inducing and stress-increasing mechanisms. Stress-induced mechanisms elevate the stresses applied to an item indirectly, by motivating or persuading creation of stress. On the other hand, stress-increasing mechanisms directly cause added stress. For example, the failure mechanism "impact" may deform an item leading to elevated stress due to added forces applied from adjacent items. Therefore, the added stress is not a direct cause of impact, but impact has caused a condition (deformation) that has led to additional stress. Similarly, the failure mechanism impact may cause direct stresses due to the forces applied to the item itself. Note that some failure mechanisms may be considered both "stress-induced" and "stress-increased" depending on the way the added stress has been established. Strength-reduced mechanism cause the capacity of an item to withstand normal stresses to be decreased, resulting in a failure.

1.3 FAILURE MECHANISMS

Failure mechanisms are physical processes whose occurrence either leads to or is caused by stress, and may deteriorate the capacity (e.g., strength or endurance) of an item. Since failure mechanisms for mechanical and electronic/electrical equipment are somewhat different, these mechanisms are discussed separately.

Mechanical Failure Mechanisms can be divided into three classes; stress-induced, strength-reduced, and stress-increased. Stress-induced mechanisms refer to mechanisms that cause or are the result of localized stress (permanent or

temporary). For example, elastic deformation may be the result of a force applied on the item that causes deformation (elastic), that disappears when the applied force is removed. Strength-reducing mechanisms are those that lead (indirectly) to a reduction of the item's strength or endurance to withstand stress or damage. For example, radiation may cause material embrittlement, thus reducing the materials capacity to withstand cracks or other damage. Stress-increasing mechanisms are those whose direct effect is an increase in the applied stress. For example, fatigue could cause direct, permanent stress in an item. Table 1.1 shows a breakdown of each class of mechanism.

Table 1.1 Categorization of Failure Mechanisms

Stress-induced failure mechanisms	Strength-reduced failure mechanisms	Stress-increased failure mechanisms
Brittle fracture	Wear	Fatigue
Buckling	Corrosion	Radiation
Yield	Cracking	Thermal-shock
Impact	Diffusion	Impact
Ductile fracture	Creep	Fretting
Elastic deformation	Radiation damage	
	Fretting	

Table 1.2 summarizes the cause, effect and physical processes involving common mechanical failure mechanisms.

Electrical Failure Mechanisms tend to be more complicated than those in purely mechanical systems. This is caused by the complexity of electrical items (e.g., devices) themselves. In integrated circuits, a typical electrical device, such as a resistor, capacitor, or transistor, is manufactured on a single crystalline chip of silicon, with multiple layers of various metals, oxides, nitrides, and organics on the surface, deposited in a controlled manner. Often a single electrical device is composed of several million elements, compounding any reliability problem present at the single element level. Once the electrical device is manufactured, it must be packaged, with electrical connections to the outside world. These connections, and the packaging, are as vital to the proper operation of the device as the electrical elements themselves.

Failure mechanisms for electrical devices are usually divided into three types: *electrical stress*, *intrinsic*, and *extrinsic* failure mechanisms. These are discussed below.

Electrical stress failure occurs when an electrical device is subjected to voltage levels higher than design constraints, damaging the device and degading electrical

Table 1.2 Mechanical Failure Mechanisms

Mechanism	Causes	Effect	Description
Buckling	Compressive load application Dimensions of the items	Item deflects greatly Possible complete loss of load carrying ability	When load applied to items such as struts, columns, plates, or thin walled cylinders reaches a critical value a sudden major change in geometry, such as bowing, winking, or bending occurs
Corrosion	Chemical action on the surface of the item Contact between two dissimilar metals in electrical contacts (galvanic corrosion) Improper welding of certain copper, chromium, nickle, aluminum, magnesium, and zinc alloys Abrasive or viscid flow of chemicals over the surface of an item Collapsing of buckles and cavities adjacent to pressure walls Living organisms in contact with the item High stress in a chemically active environment (stress-corrosion) causing cracking	Reduction in strength Cracking Fracture Geometry changes	Undesired deterioration of the item as a result of chemical or electrochemical interaction with the environment. Corrosion closely interacts with other mechanisms such as cracking, wear, and fatigue.
Impact	Sudden load from dropping an item or having been struck	Localized stresses Deformation Fracture	Failure occurs by the interaction of generated dynamic or abrupt loads, that result in large local stresses and strains
Fatigue	Fluctuating force (loads)	Cracking leading to deformation and fracture	Application of fluctuating normal loads (far below the yield point) causing pitting, cracking. Fatigue is a progressive failure phenomenon that initiates and propagates cracks.

Wear	Solid surfaces in rubbing contact Particles (sometimes removed from the surface) entrapped between rubbing surfaces Corrosive environment near rubbing contacts and loose particles entrapped between rubbing surfaces	Cumulative change in dimensions Deformation and strength reduction	Wear is not a single process. It can be a complex combination of local shearing, plowing, welding, tearing, causing gradual removal of discrete particles from contacting surfaces in motion. Particles entrapped between mating surfaces. Corrosion often interacts with wear processes and changes the character of the surfaces.
Creep	Loading, usually at high temperature, leading to gradual plastic deformation	Deformation of item Rupture	Plastic deformation in an item accrues over a period of time under stress until the accumulated dimensional changes interfere with the item's ability to properly function.
Thermal shock	Rapid cooling, heating Large differential temperature	Yield Fracture Embrittlement	Thermal gradients in an item causing major differential thermal strains which exceed the ability of the material to withstand without yielding or fracture.
Yield	Large static force Operational load or motion	Geometry changes Deformation Break	Plastic deformation in an item occurs by operational loads or motion.
Radiation damage	Nuclear Radiation	Changes in material property Loss of ductility	Radiation causes rigidity and loss of ductility. Polymers are more susceptible than metals. In metals, radiation reduces ductility resulting in other failure mechanisms.

Table 1.3 Electrical Stress Failure Mechanisms

Mechanism	Causes	Effects	Description
Electrical overstress (EOS)	Improper application of handling	Localized melting Gate oxide breakdown	Device is subjected to voltages higher than design constraints
Electrostatic discharge (ESD)	Common static charge buildup	Localized melting Gate oxide breakdown	Contact with static charge buildup during device fabrication or later handling results in high voltage discharge into device

characteristics. This failure mechanism is often a result of human error. Also known as electrical overstress, uncontrolled currents in the electrical device can cause resistive heating or localized melting at critical circuit points, which usually results in catastrophic failure but has also been known to cause latent damage. Electrostatic discharge is one common way of imparting large, undesirable currents into an electrical device.

Intrinsic failure mechanisms are related to the electrical element itself. Most failure mechanisms related to the semiconductor chip and electrically active layers grown on its surface are in this category. Intrinsic failures are related to the basic electrical activity of the device and usually result from poor manufacturing or design procedures. Intrinsic failures cause both reliability and manufacturing yield problems. Common intrinsic failure mechanisms are gate oxide breakdown, ionic contamination, surface charge spreading, and hot electrons.

Extrinsic failure mechanisms are external failure mechanisms for electrical devices which stem from problems with the device packaging and interconnections. Most extrinsic failure mechanisms are mechanical in nature. Often deficiencies in the electronic device and packaging manufacturing process cause these mechanisms to occur, though operating environment has a strong effect on the failure rate also. In recent years semiconductor technology has reached a high level of maturity, with a corresponding high level of control over intrinsic failure mechanisms. As a result, extrinsic failures have become more critical to the reliability of the latest generation of electronic devices.

Many electrical failure mechanisms are interrelated. Often, a partial failure due to one mechanism will ultimately manifest as another. For example, oxide breakdown may be caused by poor oxide processing during manufacturing, but it may also be exasperated by electrostatic discharge, damaging an otherwise intact oxide layer. Corrosion and ionic contamination may be initiated when a packaging failure allows unwanted chemical species to contact the electronic devices, and then failure can occur through trapping, piping, or surface charge spreading. Many intrinsic failure mechanisms may be initiated through an extrinsic problem: once the package of an electrical device is damaged there are a variety of intrinsic failure mechanisms which may manifest themselves in the chip itself.

Tables 1.3–1.5 summarize the cause, effect, and physical processes involving common electrical stress, intrinsic, and extrinsic failures mechanisms.

1.4 PERFORMANCE MEASURES

Overall performance of an item (component, device, product, subsystem, or system) results from implementation of various programs that ultimately improve the performance of the item. Historically, these programs have been installed

Table 1.4 Intrinsic Failure Mechanisms

Mechanism	Causes	Effects	Description
Gate oxide breakdown	EOS ESD Poor gate oxide processing	Degradation in current-voltage (I-V) characteristics.	Oxide layer which separates gate metal from semiconductor is damaged or degrades with time
Ionic contamination	Undesired ionic species are introduced into semiconductor	Degradation in I-V characteristics Increase in threshold voltage	Undesired chemical species can be introduced to device through human contact, processing materials, improper packaging, etc.
Surface charge spreading	Ionic contamination or excess surface moisture	Short circuiting between devices Threshold voltage shifts, or parasitic formation	Undesired formation of conductive pathways on surfaces alters electrical characteristic of device
Slow trapping	Poor interface quality	Threshold voltage shifts	Defects at gate oxide interface trap electrons, producing undesired electric fields
Hot electrons	High electric fields in conduction channel	Threshold voltage shifts	High electric fields create electrons with sufficient energy to enter oxide
Piping	Crystal defects Phosphorous or gold diffusion	Electrical shorts in emitter or collector	Diffusion along crystal defects in the silicon during device fabrication cause electrical shorts

Mechanism	Causes	Effects	Description
Packaging failures	Most mechanical failure mechanisms can cause electrical device packaging failures	Usually increased resistance or open circuits	See section on mechanical failure mechanisms (Sec 1.3)
Corrosion	Moisture, dc operating voltages, and Na or Cl ionic species	Open circuits	The combination of moisture, dc operating voltages, and ionic catalysts causes electrochemical movement of material, usually the metallization
Electromigration	High current densities Poor device processing	Open circuits	High electron velocities become sufficient to impact and move atoms, resulting in altered metallization geometry and, eventually, open circuits
Contact migration	Uncontrolled material diffusion	Open or short circuits	Poor interface control cause metallization to diffuse into the semiconductor. Often this occurs in the form of metallic "spikes"
Microcracks	Poorly processed oxide steps	Open circuits	Formation of a metallization path on top of a sharp oxide step results in a break in the metal or a weakened area prone to further damage
Stress migration	High mechanical stress in electrical device	Short circuits	Metal migration occurs to relieve high mechanical stress in device
Bonding failures	Poor bond control	Open circuits	Electrical contact to device package (bonds) are areas of high mechanical instability, and can separate if processing is not strictly controlled
Die attachment failures	Poor die attach integrity or corrosion	Hot spots Parametric shifts in circuit Corrosion mechanical failures	Corrosion or poor fabrication causes voids in die attach, or partial or complete de-adhesion
Particulate contamination	Poor manufacturing and chip breakage	Short circuits	Conductive particles may be sealed in a hermetic package or may be generated through chip breakage
Radiation	Trace radioactive elements in device or external radiation source	Can cause various degrading effects	High energy radiation can create hot electron-hole pairs that can interfere with and degrade device performance

through a trial-and-error approach. For example, they are sometimes established based on empirical evidence gathered during investigation of failures. An example of such programs is a root-cause failure analysis program. It is worthwhile, at this point, to understand why a well established reliability analysis and engineering program can influence the performance of today's items. For this reason, let us first define what constitutes the performance of an item.

The performance of an item can be described by four elements:

Capability or the item's ability to satisfy functional requirements;
Efficiency or the item's ability to effectively and easily realize objectives;
Reliability or the item's ability to start and continue to operate;
Availability or the item's ability to quickly become operational following a failure.

It is evident that the first two measures are influenced by the design, construction, production or manufacturing of the item. Capability and efficiency reflect the levels to which the item is designed and built. For example, the designer ensures that design levels are adequate to meet the functional requirements of a product. On the other hand, reliability is an operations related issue and is influenced by the item's potential to remain operational. In a repairable item, the ease with which the item is maintained, repaired, and returned to operation is measured by its maintainability. Based on the above definitions it would be possible to have an item that is highly reliable, but does not achieve a high performance. Examples include items that do not fully meet their stated design objectives. Humans play a major role in the design, construction, production, operation, and maintenance of the item. This common role can significantly influence the values of the four performance measures. The role of humans is often determined by various programs and activities that support the four elements of performance, proper implementation of which leads to a *quality* item.

To put all of these factors in perspective, consider the development of a high-performance product in an integrated framework. For this purpose, let us consider the so-called diamond tree conceptually shown in Fig. 1.2. In this tree, the top goal is high-performance during the life cycle of an item and is hierarchically decomposed into various goals, functions, activities, programs, and organizations. By looking down from the top of this structure, one can describe *how* various goals and subgoals are achieved, and by looking up, one can identify *why* a goal or function is necessary. Figure 1.2 shows only typical goals, but also reflects the general goals involved in designing and operating a high-performance item. For a more detailed description of the diamond tree the readers are referred to Hunt and Modarres (1985).

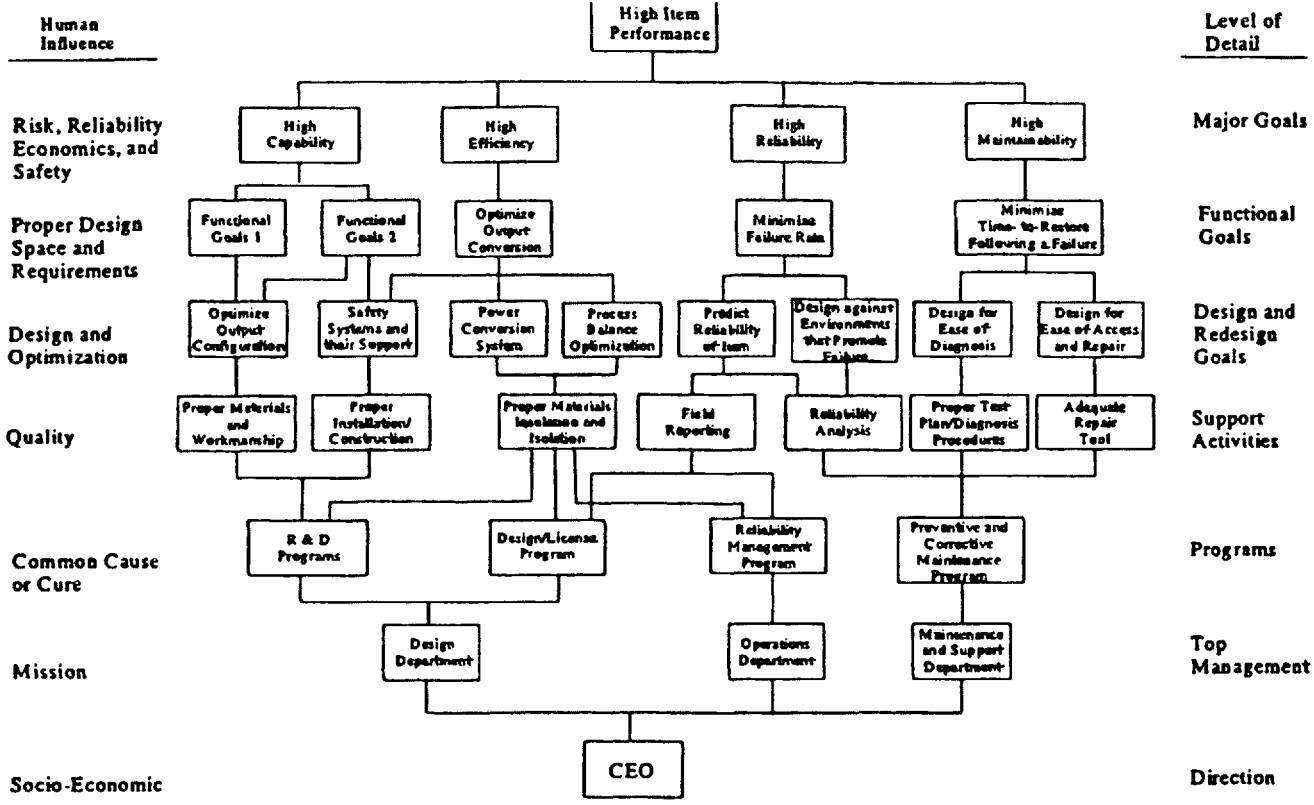


Figure 1.2 A conceptual diamond tree representation for achieving high performance.

Reliability and availability play a key role in the overall framework shown in Fig. 1.2. Of the four elements of performance, we are mainly interested in reliability and availability. In this book, only the repairable aspects of a maintainable system are of interest to us. Therefore, we will only discuss reliability and availability as two important measures of performance. A more detailed look at the goals of improving reliability, in an integrated manner, would yield a better perspective on the role of reliability and availability analysis as shown by the hierarchy depicted in Fig. 1.3. From this, one can put into a proper context the role of reliability and availability analysis.

Clearly, reliability is an important element in achieving high-performance since it directly and significantly influences the item's performance and ultimately its life-cycle cost and economics. Poor reliability directly causes increased warranty costs, liabilities, recalls, and repair costs. Poor quality would also lead to poor performance. Therefore, a high quality design, production, manufacturing, and operation program leads to low failure frequencies, effective maintenance and repair, and ultimately high performance.

In this book we are also interested in risk analysis. However, risk associated with an item is not a direct indicator of performance. Risk is the item's potential to cause a loss (e.g., loss of other systems, loss to humans, environmental damage, or economic loss). However, a quantitative measure of risk can be an important metric for identifying and highlighting items that are risk-significant (i.e., they may be associated with a potentially significant loss). This metric, however, is useful to set adequate performance levels for risk-significant items. Conversely, performance may highly influence an item's risk. For example, a highly reliable item is expected to fail less frequently resulting in small risk. On the other hand, risk of an item may be an indicator for items that should attain a high performance. Accordingly, risk and performance of an item synergistically influence each other. This concept is depicted in Fig. 1.4.

1.5 DEFINITION OF RELIABILITY

As we discussed earlier, reliability has two connotations. One is probabilistic in nature; the other is deterministic. In this book, we generally deal with the probabilistic aspect. Let us first define what we mean by reliability. The most widely accepted definition of *reliability* is the ability of an item (product, system, . . . etc.) to operate under designated operating conditions for a designated period of time or number of cycles. The *ability* of an item can be designated through a probability (the probabilistic connotation), or can be designated deterministically. The deterministic approach, as indicated in Section 1.1, deals with understanding how and why an item fails, and how it can be designed and tested to prevent such

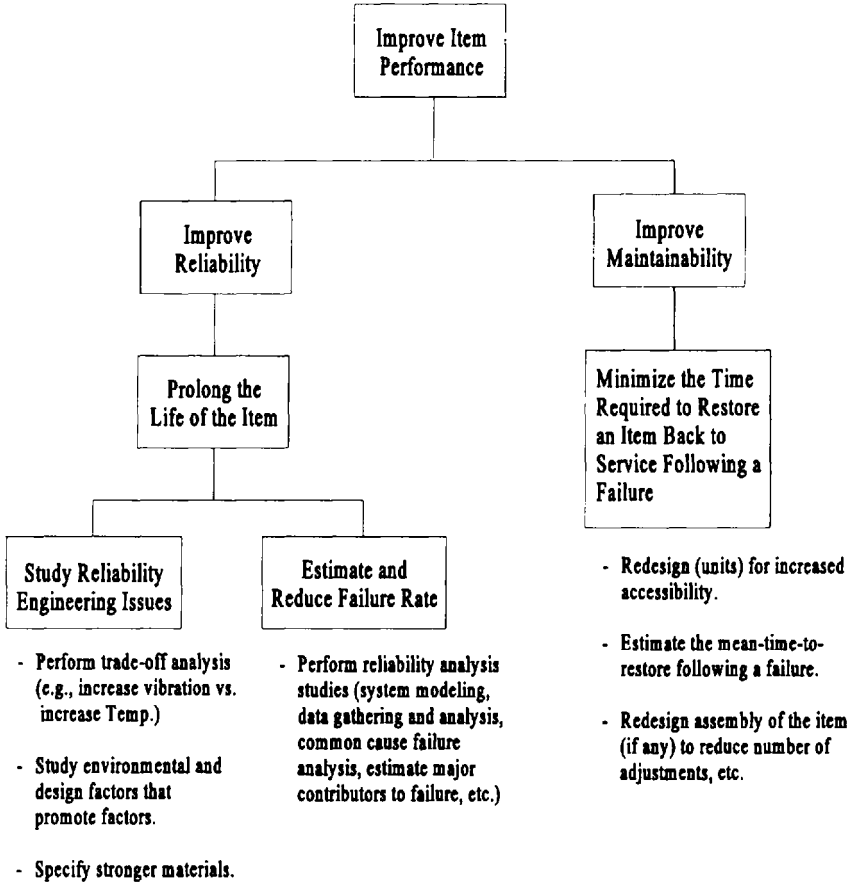


Figure 1.3 A conceptual hierarchy for improving performance.

failures from occurrence or recurrence. This includes such analyses as deterministic analysis and review of field failure reports, understanding physics of failure, the role and degree of test and inspection, performing redesign, or performing reconfiguration. In practice, this is an important aspect of reliability analysis.

The probabilistic treatment of an item's reliability according to the definition above can be summarized by

$$R(t) = \Pr\left(T \geq t \mid c_1, c_2, \cdots\right) \tag{1.1}$$

where

t = the designated period of time or cycles for the item's operation (mission time),
T = time to failure or cycle to failure of the item,
$R(t)$ = reliability of the item, and
$c_1, c_2, \ldots$ = designated conditions, such as environmental conditions.

Often, in practice, $c_1, c_2, \ldots$ are implicitly considered in the probabilistic reliability analysis and thus Eq. (1.1) reduces to

$$R(t) = \Pr(T > t) \tag{1.2}$$

Expressions (1.1) and (1.2) are discussed further in Chapter 3.

1.6 DEFINITION OF AVAILABILITY

Availability analysis is performed to verify that an item has a satisfactory probability of being operational, so it can achieve its intended objective. In Fig. 1.3, an item's availability can be considered as combination of its reliability and maintainability. Accordingly, when no maintenance o4433r repair is performed (e.g., in nonrepairable items), reliability can be considered as instantaneous availability.

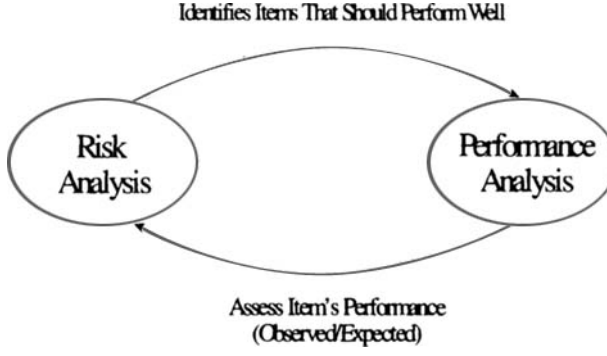


Figure 1.4 Synergistic effects between risk and performance of an item.

Mathematically, the availability of an item is a measure of the fraction of time that the item is in operating condition in relation to total or calendar time. There are several measures of availability, namely, inherent availability, achieved availability, and operational availability. For further definition of these availability measures, see Ireson and Coombs (1988). Here, we describe inherent availability, which is the most common definition used in the literature.

A more formal definition of *availability* is the probability that an item, when used under stated conditions in an ideal support environment (i.e., ideal spare parts, personnel, diagnosis equipment, procedures, etc.), will be operational at a given time. Based on this definition, the average availability of an item during an interval of time T can be expressed by

$$A = \frac{u}{u + d} \tag{1.3}$$

where

u = uptime during time T,
d = downtime during time T,
$T = u + d$.

Time-dependent expressions of availability and measures of availability for different types of equipment are discussed in more detail in Chapter 5. The mathematics and methods for reliability analysis discussed in this book are also equally applicable to availability analysis.

1.7 DEFINITION OF RISK

Risk can be viewed both qualitatively and quantitatively. Qualitatively speaking, when there is a source of danger (hazard), and when there are no safeguards against exposure of the hazard, then there is a possibility of loss or injury. This possibility is referred to as risk. The loss or injury could result from business, social, or military activities; operation of equipment; investment; etc. *Risk* can be formally defined as the potential of loss (e.g., material, human, or environment, losses) resulting from exposure to a hazard.

In complex engineering systems, there are often safeguards against exposure of hazards. The higher the level of safeguards, the lower the risk. This also underlines the importance of highly reliable safeguard systems and shows the roles of and relationship between reliability analysis and risk analysis.

In this book, we are concerned with quantitative risk analysis. Since quantitative risk analysis involves estimation of the degree or probability of loss,

risk analysis is fundamentally intertwined with the concept of probability of occurrence of hazards. Risk analysis consists of answers to the following questions (see Kaplan and Garrick (1981)):

1. What can go wrong that could lead to an outcome of hazard exposure?
2. How likely is this to happen?
3. If it happens, what consequences are expected?

To answer question 1, a list of outcomes (or scenarios of events leading to the outcome) should be defined. The likelihood of these scenarios should be estimated (answer to question 2), and the consequence of each scenario should be described (answer to question 3). Therefore, risk can be defined, quantitatively, as the following set of triplets:

$$R = < S_i, P_i, C_i > \qquad i = 1, 2, \ldots, n, \tag{1.4}$$

where

S_i = is a scenario of events that lead to hazard exposure,
P_i = is the likelihood of scenario i, and
C_i = is the consequence (or evaluation measure) of scenario i, e.g., a measure of the degree of damage or loss.

Since Eq. (1.4) involves estimation of the likelihood of occurrence of events (e.g., failure of safeguard systems), most of the methods described in Chapters 2 through 7 become relevant. However, we have specifically devoted Chapter 8 to a detailed, quantitative description of these methods as applied to risk analysis.

REFERENCES

Amerasekera, E. A., and Campbell, D. S., "*Failure Mechanisms in Semiconductor Devices*," John Wiley and Sons, 1987.

Collins, J. A., "*Failure of Materials in Mechanical Design, Analysis, Prediction, and Prevention*," (2nd ed.). John Wiley and Sons, 1993.

Dasgupta, A., and Pecht, M., "*Materials Failure Mechanisms and Damage Models*," IEEE Transactions on Reliability, vol. 40, No. 5, 1991.

Hunt, R. N., and Modarres, M., "*A Use of Goal Tree Methodology to Evaluate Institutional Practices and Their Effect on Power Plant Hardware Performance*," American Nuclear Society Topical Meeting on Probabilistic Safety Methods and Applications, San Francisco, CA, 1985.

Ireson, W. G., and Coombs, C. F. eds., "*Handbook of Reliability Engineering and Management*," McGraw-Hill, New York, NY, 1988.

Kaplan, S., and Garrick, J., "*On the Quantitative Definition of Risk*," Risk Analysis, vol. 1, No.1, 1981.

Pecht, M., eds., "*Handbook of Electronic Package Design*," CALCE Center for Electronic Packaging, University of Maryland, College Park, MD, Marcel Dekker, Inc., New York, NY, 1991.

This page intentionally left blank

2

Basic Reliability Mathematics: Review of Probability and Statistics

2.1 INTRODUCTION

In this chapter, we discuss the elements of mathematical theory that are relevant to the study of reliability of physical objects. We begin with a presentation of basic concepts of probability. Then we briefly consider some fundamental concepts of statistics that are used in reliability data analysis.

2.2 ELEMENTS OF PROBABILITY

Probability is a concept that people use formally and casually every day. The weather forecasts are probabilistic in nature. People use probability in their casual conversations to show their perception of the likely occurrence or nonoccurrence of particular events. Odds are given for the outcome of sport events, and are used in gambling.

Formal use of probability concepts is widespread in science, for example, astronomy, biology, and engineering. In this chapter, we discuss the formal application of probability theory in the field of reliability engineering.

2.2.1 Sets and Boolean Algebra

To perform operations associated with probability, it is often necessary to use sets. A *set* is a collection of items or elements, each with some specific characteristics. A set that includes all items of interest is referred to as a *universal set*, denoted by Ω. A *subset* refers to a collection of items that belong to a universal set. For example, if set Ω represents the collection of all pumps in a power plant, then the collection of electrically driven pumps is a subset E of Ω. Graphically, the relationship between subsets and sets can be illustrated through Venn diagrams. The Venn diagram in Fig. 2.1 shows the universal set Ω by a rectangle, and subsets E_1 and E_2 by circles. It can also be seen that E_2 is a subset of E_1. The relationship

between subsets E_1 and E_2 and the universal set can be symbolized by $E_2 \subset E_1 \subset \Omega$.

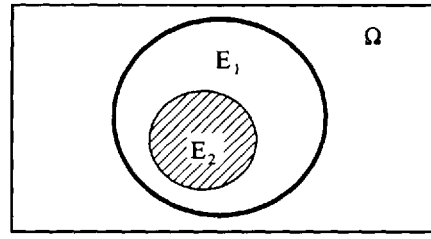


Figure 2.1 Venn diagram.

The *complement* of a set E, denoted by, $\bar{E}$ and called *E not*, is the set of all items (or more specifically events) in the universal set that do not belong to set E. In Fig. 2.1, the nonshaded area outside of the set E_2 bounded by the rectangle represents $\bar{E}_2$. It is clear that sets E_2 and $\bar{E}_2$ together comprise Ω.

The *union* of two sets, E_1 and E_2, is a set that contains all items that belong to E_1 or E_2. The union is symbolized either by $E_1 \cup E_2$ or $E_1 + E_2$, and is read E_1 or E_2. That is, the set $E_1 \cup E_2$ represents all elements that are in E_1, E_2 or both E_1 and E_2. The shaded area in Fig. 2.2 shows the union of sets E_1 and E_2.

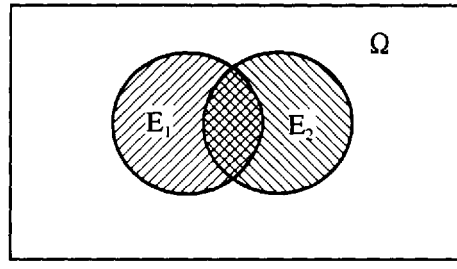


Figure 2.2 Union of two sets, E_1 and E_2

Suppose E_1 and E_2 represent positive odd and even numbers between 1 and 10, respectively. Then

$$E_1 = \{ 1, 3, 5, 7, 9 \}$$
$$E_2 = \{ 2, 4, 6, 8, 10 \}$$

The union of these two sets is:

$$E_1 \cup E_2 = \{ 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 \}$$

or, if $E_1 = \{ x, y, z \}$ and $E_2 = \{ x, t, z \}$, then

$$E_1 \cup E_2 = \{ x, y, z, t \}.$$

Note that element x is in both sets E_1 and E_2.

The *intersection* of two sets, E_1 and E_2, is the set of items that are common to both E_1 and E_2. This set is symbolized by $E_1 \cap E_2$ or $E_1 \cdot E_2$, and is read E_1 and E_2. In Fig. 2.3, the shaded area represents the intersection of E_1 and E_2.

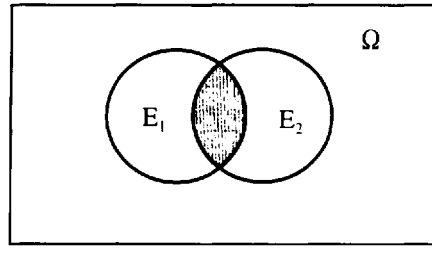


Figure 2.3 Intersection of two sets, E_1 and E_2.

Suppose E_1 is a set of manufactured devices that operate for $t > 0$ but fail prior to 1000 hours of operation. If set E_2 represents a set of devices that operate between 500 and 2000 hours, then $E_1 \cap E_2$ can be obtained as follows:

$$E_1 = \{ t \mid 0 < t < 1000 \}$$
$$E_2 = \{ t \mid 500 < t < 2000 \}$$
$$E_1 \cap E_2 = \{ t \mid 500 < t < 1000 \}$$

Also, if sets $E_1 = \{ x, y, z \}$ and $E_2 = \{ x, t, z \}$, then

$$E_1 \cap E_2 = \{ x \}$$

Note that the first two sets in this example represent "continuous" elements, and the second two sets represent "discrete" elements. This concept will be discussed in more detail further in this chapter.

A *null* or *empty* set, $\emptyset$, refers to a set that contains no items. One can easily see that the complement of a universal set is a null set, and vice versa. That is,

$$\begin{aligned} \overline{\Omega} &= \emptyset \\ \Omega &= \overline{\emptyset} \end{aligned} \tag{2.1}$$

Two sets, E_1 and E_2, are termed *mutually exclusive* or *disjoint* when $E_1 \cap E_2 = \emptyset$. In this case, there are no elements common to E_1 and E_2. Two mutually exclusive sets are illustrated in Fig. 2.4.

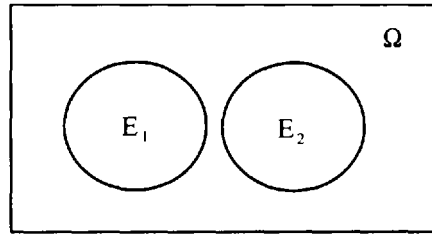


Figure 2.4 Mutually exclusive sets, E_1 and E_2.

From the discussions thus far, as well as from the examination of the Venn diagram, the following conclusions can be drawn:

The intersection of set E and a null set is a null set:

$$E \cap \emptyset = \emptyset \tag{2.2}$$

The union of set E and a null set is E:

$$E \cup \emptyset = E \tag{2.3}$$

The intersection of set E and the complement of E is a null set:

$$E \cap \bar{E} = \emptyset \tag{2.4}$$

The intersection of set E and a universal set is E:

$$E \cap \Omega = E \tag{2.5}$$

The union of set E and a universal set is the universal set:

$$E \cup \Omega = \Omega \tag{2.6}$$

The complement of the complement of set E is E:

$$\bar{\bar{E}} = E \tag{2.7}$$

The union of two identical sets E is E:

$$E \cup E = E \tag{2.8}$$

The intersection of two identical sets E is E:

$$E \cap E = E \tag{2.9}$$

Example 2.1

Simplify the following expression:

$$E_1 \cap E_1 \cup \emptyset$$

Solution:

Since $E_1 \cap E_1 = E_1$, then the expression reduces to $E_1 \cup \emptyset = E_1$

Boolean algebra provides a means of evaluating sets. The rules are fairly simple. The sets of axioms in Table 2.1 provide all the major relations of interest in Boolean algebra including some of the expressions discussed in Eqs. (2.1) through (2.9).

Table 2.1 Laws of Boolean Algebra

$X \cap Y = Y \cap X$ $X \cup Y = Y \cup X$	Commutative Law
$X \cap (Y \cap Z) = (X \cap Y) \cap Z$ $X \cup (Y \cup Z) = (X \cup Y) \cup Z$	Associative Law
$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$	Distributive Law
$X \cap X = X$ $X \cup X = X$	Idempotent Law
$X \cap (X \cup Y) = X$ $X \cup (X \cap Y) = X$	Absorption Law
$X \cap \bar{X} = \phi$ $X \cup \bar{X} = \Omega$ $\bar{\bar{X}} = X$	Complementation Law
$\overline{(X \cap Y)} = \bar{X} \cup \bar{Y}$ $\overline{(X \cup Y)} = \bar{X} \cap \bar{Y}$	De Morgan's Theorem

Example 2.2

Simplify the following Boolean expression:

$$\overline{[(A \cdot B) + (A \cdot \bar{B}) + (\bar{A} \cdot \bar{B})]}$$ De Morgan's Theorem

$$= \overline{(A \cdot B)} \cdot \overline{(A \cdot \bar{B})} \cdot \overline{(\bar{A} \cdot \bar{B})}$$ De Morgan's Theorem

$$= (\bar{A} + \bar{B}) \cdot (\bar{A} + \bar{\bar{B}}) \cdot (\bar{\bar{A}} + \bar{\bar{B}})$$ Complementation Law

$$= (\bar{A} + \bar{B}) \cdot (\bar{A} + B) \cdot (A + B)$$ Distributive Law

$$= [\bar{A} \cdot (A + B)] + [\bar{B} \cdot (A + B)] \cdot (A + B)$$ Distributive Law

$$= [\bar{A} + (B \cdot \bar{A})] + [(\bar{B} \cdot \bar{A}) + (\bar{B} \cdot B)] \cdot (A + B)$$ Absorption Law

$$= [\bar{A} + (B \cdot \bar{A})] \cdot (A + B)$$ Distributive Law

$$= \bar{A} \cdot (A + B)$$ Absorption Law

$$= \bar{A} \cdot B.$$

2.2.2 Basic Laws of Probability

In probability theory, the elements that comprise a set are outcomes of an experiment. Thus, the universal set Ω represents the mutually exclusive listing of all possible outcomes of the experiment and is referred to as the *sample space* of the experiment. In examining the outcomes of rolling a die, the sample space is $S = \{1,2,3,4,5,6\}$. This sample space consists of six items (elements) or *sample points*. In probability concepts, a combination of several sample points is called an *event*. An event is, therefore, a subset of the sample space. For example, the event of "an odd outcome when rolling a die" represents a subset containing sample points 1, 3, and 5.

Associated with any event E of a sample space S is a probability shown by $\Pr(E)$ and obtained from the following equation:

$$\Pr(E) = \frac{m(E)}{m(S)} \tag{2.10}$$

where $m(.)$ denotes the number of elements in the set (.) and refers to the *size* of the set.

The probability of getting an odd number when tossing a die is determined by using $m(\textit{odd outcomes}) = 3$ and $m(\textit{sample space}) = 6$. In this case, $\Pr(\textit{odd outcomes}) = 3/6 = 0.5$.

Note that Eq. (2.10) represents a comparison of the relative size of the subset

represented by the event E to the sample space S. This is true when all sample points are equally likely to be the outcome. When all sample points are not equally likely to be the outcome, the sample points may be weighted according to their relative frequency of occurrence over many trials or according to expert judgment.

At this point, it is important that the readers appreciate some intuitive differences between three major conceptual interpretations of probability described below.

Classical Interpretation of Probability (Equally Likely Concept)

In this interpretation, the probability of an event E can be obtained from Eq. (2.10), provided that the sample space contains N equally likely and different outcomes, i.e., $m(S) = N$, n of which have an outcome (event) E, i.e., $m(E) = n$. Thus $\Pr(E) = n/N$. This definition is often inadequate for engineering applications. For example, if failures of a pump to start in a process plant are observed, it is unknown whether all failures are equally likely to occur. Nor is it clear if the whole spectrum of possible events is observed. That case is not similar to rolling a perfect die, with each side having an equal probability of 1/6 at any time in the future.

Frequency Interpretation of Probability

In this interpretation, the limitation on the lack of knowledge about the overall sample space is remedied by defining the probability as the limit of n/N as N becomes large. Therefore, $\Pr(E) = \lim_{N \to \infty} (n/N)$. Thus if we have observed 2000 starts of a pump in which 20 failed, and if we assume that 2000 is a large number, then the probability of the pump failure to start is 20/2000 = 0.01.

The frequency interpretation is the most widely used classical definition today. However, some argue that because it does not cover cases in which little or no experience (or evidence) is available, nor cases where estimates concerning the observations are intuitive, a broader definition is required. This has led to the third interpretation of probability.

Subjective Interpretation of Probability

In this interpretation, $\Pr(E)$ is a measure of the degree of belief one holds in a specified event E. To better understand this interpretation, consider the probability of improving a system by making a design change. The designer believes that such a change results in a performance improvement in one out of three missions in which the system is used. It would be difficult to describe this problem through the first two interpretations. That is, the classical interpretation is inadequate since there is no reason to believe that performance is as likely to improve as to not improve. The frequency interpretation is not applicable because no historical data

exist to show how often a design change resulted in improving the system. Thus, the subjective interpretation provides a broad definition of the probability concept.

Calculus of Probability

The basic rules used to combine and treat the probability of an event are not affected by the interpretations discussed above; we can proceed without adopting any of them. (There is much dispute among probability scholars regarding these interpretations. Readers are referred to Cox (1946) for further discussions of this subject.)

In general, the axioms of probability can be defined for a sample space S as follows:

1. $\Pr(E) > 0$, for every event E such that $E \subset S$,
2. $\Pr(E_1 \cup E_2 \cup \cdots \cup E_n) = \Pr(E_1) + \Pr(E_2) + \cdots + \Pr(E_n)$, where the events $E_1, E_2, \ldots, E_n$ are such that no two have a sample point in common,
3. $\Pr(S) = 1$.

It is important to understand the concept of independent events before attempting to multiply and add probabilities. Two events are *independent* if the occurrence or nonoccurrence of one does not depend on or change the probability of the occurrence of the other. Mathematically, this can be expressed by

$$\Pr(E_1 \mid E_2) = \Pr(E_1) \tag{2.11}$$

where $\Pr(E_1 \mid E_2)$ reads "the probability of E_1, given the E_2 has occurred." To better illustrate, let's consider the result of a test on 200 manufactured identical parts. It is observed that 23 parts fail to meet the length limitation imposed by the designer, and 18 fail to meet the height limitation. Additionally, 7 parts fail to meet both length and height limitations. Therefore, 152 parts meet both of the specified requirements. Let E_1 represent the event that a part does not meet the specified length, and E_2 represent the event that the part does not meet the specified height. According to Eq. (2.10), $\Pr(E_1) = (7 + 23)/200 = 0.15$, and $\Pr(E_2) = (18 + 7)/200 = 0.125$. Furthermore, among 25 parts (7 + 18) that have at least event E_2, 7 parts also have event E_1. Thus, $\Pr(E_1 \mid E_2) = 7/25 = 0.28$. Since $\Pr(E_1 \mid E_2) \neq \Pr(E_1)$, events E_1 and E_2 are dependent.

We shall now discuss the rules for evaluating the probability of simultaneous occurrence of two or more events, that is, $\Pr(E_1 \cap E_2)$. For this purpose, we recognize two facts. First, when E_1 and E_2 are independent, the probability that both E_1 and E_2 occur simultaneously is simply the multiplication of the probabilities that E_1 and E_2 occur individually. That is,

$$\Pr(E_1 \cap E_2) = \Pr(E_1) \cdot \Pr(E_2)$$

Second, when $E_1$and E_2 are dependent, the probability that both $E_1$and E_2 occur simultaneously is obtained from the following expressions:

$$\Pr(E_1 \cap E_2) = \Pr(E_1) \cdot \Pr(E_2 \mid E_1) \tag{2.12}$$

We will elaborate further on Eq. (2.12) when we discuss Bayes' Theorem. It is easy to see that when $E_1$and E_2 are independent, and Eq. (2.11) is applied, Eq. (2.12) reduces to

$$\Pr(E_1 \cap E_2) = \Pr(E_1) \cdot \Pr(E_2)$$

In general, the probability of joint occurrence of n independent events E_1, E_2, . . . , E_n is the product of their individual probabilities.
That is,

$$\Pr(E_1 \cap E_2 \cap \cdots \cap E_n) = \Pr(E_1) \cdot \Pr(E_2) \cdots \Pr(E_n) = \prod_{i=1}^{n} \Pr(E_i) \tag{2.13}$$

The probability of joint occurrence of n dependent events E_1, E_2, . . ., E_n is obtained from

$$\Pr(E_1 \cap E_2 \cap \cdots \cap E_n) = \Pr(E_1) \cdot \Pr(E_2 \mid E_1) \cdot \Pr(E_3 \mid E_1 \cap E_2) \cdots \cdots \Pr(E_n \mid E_1 \cap E_2 \cap \cdots \cap E_{n-1}) \tag{2.14}$$

where $\Pr(E_3 \mid E_1 \cap E_2 \cap \ldots)$ denotes the conditional probability of E_3 given the occurrence of both $E_1$and $E_2$and so on.

Example 2.3

Suppose that Vendor 1 provides 40% and Vendor 2 provides 60% of electronic devices used in a computer. It is further known that 2.5% of Vendor 1's supplies are defective, and only 1% of Vendor 2's supplies are defective. What is the probability that a unit is both defective and supplied by Vendor 1? What is the same probability for Vendor 2?

Solution:

E_1= the event that a device is from Vendor 1,
E_2 = the event that a device is from Vendor 2,
D = the event that a device is defective,
$D \mid E_1$ = the event that a device is known to be from Vendor 1 and defective,
$D \mid E_2$ = the event that a device is known to be from Vendor 2 and defective.

Then,

$$\Pr(E_1) = 0.40,\ \Pr(E_2) = 0.60,\ \Pr(D \mid E_1) = 0.025, \text{ and } \Pr(D \mid E_2) = 0.01$$

From (2.14), the probability that a defective device is from Vendor 1 is

$$\Pr(E_1 \cap D) = \Pr(E_1) \cap \Pr(D \mid E_1) = (0.4)\,(0.025) = 0.01$$

Similarly,

$$\Pr(E_2 \cap D) = 0.006$$

The evaluation of the probability of union of two events depends on whether or not these events are mutually exclusive. To illustrate this point, let's consider the 200 electronic parts that we discussed earlier. The union of two events E_1 and E_2 includes those parts that do not meet the length requirement, or the height requirement, or both. That is, a total of 23 + 18 + 7 = 48. Thus,

$$\Pr(E_1 \cup E_2) = 48/200 = 0.24$$

In other words, 24% of the parts do not meet one or both of the requirements. We can easily see that $\Pr(E_1 \cup E_2) \neq \Pr(E_1) + \Pr(E_2)$, since $0.24 \neq 0.125 + 0.15$. The reason for this inequality is the fact that the two events E_1 and E_2 are not mutually exclusive. In turn, $\Pr(E_1)$ will include the probability of inclusive events $E_1 \cap E_2$, and $\Pr(E_2)$ will also include events $E_1 \cap E_2$. Thus, joint events are counted twice in the expression $\Pr(E_1) + \Pr(E_2)$. Therefore, $\Pr(E_1 \cap E_2)$ must be subtracted from this expression. This description, which can also be seen in a Venn diagram, leads to the following expression for evaluating the probability of the union of two events that are not mutually exclusive:

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1 \cap E_2) \tag{2.15}$$

Since $\Pr(E_1 \cap E_2) = 7/200 = 0.035$, then $\Pr(E_1 \cup E_2) = 0.125 + 0.15 - 0.035 = 0.24$, which is what we expect to get. From (2.15) one can easily infer that if E_1 and E_2 are mutually exclusive, then $\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2)$. If events E_1 and E_2 are dependent, then by using (2.12), we can write (2.15) in the following form:

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1) \cdot \Pr(E_2 \mid E_1) \tag{2.16}$$

Equation (2.15) for two events can be logically extended to n events.

$$\begin{aligned}\Pr(E_1 \cup E_2 \cup \cdots \cup E_n) = {} & \left[\Pr(E_1) + \Pr(E_2) + \cdots + \Pr(E_n)\right] \\ & - \left[\Pr(E_1 \cap E_2) + \Pr(E_1 \cap E_3) + \cdots + \Pr(E_{n-1} \cap E_n)\right] \\ & + \left[\Pr(E_1 \cap E_2 \cap E_3) + \Pr(E_1 \cap E_2 \cap E_4) + \cdots\right] \\ & - \cdots \\ & - (-1)^{n+1}\left[E_1 \cap E_2 \cap \cdots \cap E_n\right]\end{aligned} \tag{2.17}$$

Equation (2.17) consists of 2^{n-1} terms. If events $E_1, E_2, \ldots, E_n$ are mutually exclusive, then

$$\Pr(E_1 \cup E_2 \cup \cdots \cup E_n) = \Pr(E_1) + \Pr(E_2) + \cdots + \Pr(E_n) \tag{2.18}$$

When events $E_1, E_2, \ldots, E_n$ are not mutually exclusive, a useful method known as a *rare event approximation* can be used. In this approximation, (2.18) is used if all $\Pr(E_i)$ are small e.g., $\Pr(E_i) < (50n)^{-1}$.

Example 2.4

Determine the maximum error in the right hand side of (2.17) if (2.18) is used instead of (2.17). Find this error for $n = 2, 3, 4$, and assume $\Pr(E_i) < (50n)^{-1}$.

Solution:

For maximum error assume $\Pr(E_i) = (50\,n)^{-1}$

For $n = 2$, using (2.17),

$$\Pr(E_1 \cup E_2) = \frac{2}{50 \times 2} - \left(\frac{1}{50 \times 2}\right)^2 = 0.01900$$

Using (2.18),

$$\Pr(E_1 \cup E_2) = \frac{2}{50 \times 2} = 0.02000$$

$$\left|\max \%\ \text{Error}\right| = \left|\frac{0.01990 - 0.02000}{0.01990} \times 100\right| = 0.50\,\%$$

For $n = 3$, using (2.18),

$$\Pr(E_1 \cup E_2 \cup E_3) = \frac{3}{50 \times 3} - 3\left(\frac{1}{50 \times 3}\right)^2 + \left(\frac{1}{50 \times 3}\right)^3 = 0.01987$$

$$\Pr(E_1 \cup E_2 \cup E_3) = \frac{3}{50 \times 3} = 0.02000$$

$$|\max \% \text{ Error}| = 0.65\%$$

Similarly for $n = 4$,

$$|\max \% \text{ Error}| = 0.76\%$$

For dependent events, (2.17) can also be expanded to the form of Eq. (2.16) by using (2.14). If all events are independent, then according to (2.13), (2.15) can be further simplified to

$$\Pr(E_1 \cup E_2) = \Pr(E_1) + \Pr(E_2) - \Pr(E_1) \cdot \Pr(E_2) \tag{2.19}$$

Equation (2.19) can be algebraically reformatted to the easier form of

$$\Pr(E_1 \cup E_2) = 1 - (1 - \Pr(E_1)) \cdot (1 - \Pr(E_2)) \tag{2.20}$$

Equation (2.19) can be expanded in the case of n independent events to

$$\begin{aligned} &\Pr(E_1 \cup E_2 \cup \cdots \cup E_n) \\ &= 1 - (1 - \Pr(E_1)) \cdot (1 - \Pr(E_2)) \cdots (1 - \Pr(E_n)) \end{aligned} \tag{2.21}$$

Example 2.5

A particular type of a computer chip is manufactured by three different suppliers. It is known that 5% of chips from Supplier 1, 3% from Supplier 2, and 8% from Supplier 3 are defective. If one chip is selected from each supplier, what is the probability that at least one of the chips is defective?

Solution:

D_1 = the event that a chip from Supplier 1 is defective,
D_2 = the event that a chip from Supplier 2 is defective,
D_3 = the event that a chip from Supplier 3 is defective.

$D_1 \cup D_2 \cup D_3$ is the event that at least one chip from Supplier 1, Supplier 2, or Supplier 3 is defective. Since the occurrence of events D_1, D_2, and D_3 are independent, we can use (2.21) to determine the probability of $D_1 \cup D_2 \cup D_3$. Thus,

$$\Pr(D_1 \cup D_2 \cup D_3) = 1 - (1 - 0.05)(1 - 0.03)(1 - 0.08) = 0.152$$

In probability evaluations, it is sometimes necessary to evaluate the probability of the complement of an event, that is, $\Pr(\bar{E})$. To obtain this value, let's begin with (2.10) and recognize that the probability of all events in sample space S is 1. The sample space can also be expressed by event E and its complement $\bar{E}$. That is

$$\Pr(S) = 1 = \Pr(E \cup \bar{E})$$

Since E and $\bar{E}$ are mutually exclusive, $\Pr(E \cup \bar{E}) = \Pr(E) + \Pr(\bar{E})$. Thus, $\Pr(E) + \Pr(\bar{E}) = 1$.
By rearrangement, it follows that

$$\Pr(\bar{E}) = 1 - \Pr(E) \tag{2.22}$$

It is important to emphasize the difference between independent events and mutually exclusive events, since these two concepts are sometimes confused. In fact, two events that are mutually exclusive are not independent. Since two mutually exclusive events E_1 and E_2 have no intersection, that is, $E_1 \cap E_2 = \emptyset$, then $\Pr(E_1 \cap E_2) = \Pr(E_1) \cdot \Pr(E_2 \mid E_1) = 0$. This means that $\Pr(E_2 \mid E_1) = 0$, since $\Pr(E_1) \neq 0$. For two independent events, we expect to have $\Pr(E_2 \mid E_1) = \Pr(E_2)$, which is not zero except for the trivial case of $\Pr(E_2) = 0$. This indicates that two mutually exclusive events are indeed dependent.

2.2.3 Bayes' Theorem

An important theorem known as *Bayes' Theorem* follows directly from the concept of conditional probability, a form of which is described in (2.12). For example, three forms of Eq. (2.12) for events A and E are

$$\Pr(A \cap E) = \Pr(A) \cdot \Pr(E \mid A)$$

and

$$\Pr(A \cap E) = \Pr(E) \cdot \Pr(A \mid E)$$

therefore

$$\Pr(A) \cdot \Pr(E \mid A) = \Pr(E) \cdot \Pr(A \mid E)$$

By solving for $\Pr(A \mid E)$, it follows that

$$\Pr(A \mid E) = \frac{\Pr(A) \cdot \Pr(E \mid A)}{\Pr(E)} \tag{2.23}$$

This equation is known as Bayes' Theorem.

It is easy to prove that if event E depends on some previous events that can occur in one of the n different ways $A_1, A_2, \ldots, A_n$, then (2.23) can be generalized to

$$\Pr(A_j \mid E) = \frac{\Pr(A_j)\Pr(E \mid A_j)}{\sum_{i=1}^{n} \Pr(A_i)\Pr(E \mid A_i)} \tag{2.24}$$

The right-hand side of the Bayes' equation consists of two terms: $\Pr(A_j)$, called the *prior probability*, and

$$\frac{\Pr(E \mid A_j)}{\sum_{i=1}^{n} \Pr(E \mid A_i)\Pr(A_i)}$$

the *relative likelihood* or the factor by which the prior probability is revised based on evidential observations (e.g., limited failure observations). $\Pr(A_j \mid E)$ is called the *posterior probability*, that is, given event E, the probability of event A_i can be updated (from prior probability $\Pr(A_j)$). Clearly, when more evidence (in the form of events E) becomes available, $\Pr(A_j \mid E)$ can be further updated. Bayes' Theorem provides a means of changing one's knowledge about an event in light of new evidence related to the event. We return to this topic and its application in failure data evaluation in Chapter 3. For further studies about Bayes' Theorem, refer to Lindley (1965).

Example 2.6

Two experts are inquired about the expected reliability of the product at the end of its useful life. One expert assesses the reliability as 0.98 and the other gives 0.60. Assume both experts are considered equally credible. That is, there is a 50/50 chance that each expert is correct. Calculate the probability that experts are correct in the following situations:

a. The product is tested for useful life and does not fail (i.e., the outcome of the test is a success).
b. Two tests are performed and both are successful.
c. One test is performed and it results in failure.
d. Two tests are performed. One is successful, but the other results in failure.

Solution

a. Denote:

A_1 = event that expert 1 is correct,
A_2 = event that expert 2 is correct,
B = event that the test is successful.

Then

$$\Pr(A_1 \mid B) = \frac{\Pr(A_1)\Pr(B \mid A_1)}{\Pr(A_1)\Pr(B \mid A_1) + \Pr(A_2)\Pr(B \mid A_2)}$$

$$\Pr(A_1 \mid B) = \frac{0.5 \times 0.98}{0.5 \times 0.98 + 0.5 \times 0.60} = 0.62$$

$$\Pr(A_2 \mid B) = 1 - 0.62 = 0.38$$

It is clear that because of the successful test the credibility of expert A_1 estimate rises in light of this test, and credibility of expert A_2 declines.

b. Denote:

B_1 = event that test 1 is successful,
B_2 = event that test 2 is successful.

Then

$$\Pr(A_1 \mid B_1 \cap B_2) = \frac{\Pr(A_1)\Pr(B_1 \cap B_2 \mid A_1)}{\Pr(A_1)\Pr(B_1 \cap B_2 \mid A_1) + \Pr(A_2)\Pr(B_1 \cap B_2 \mid A_2)}$$

$$\Pr(A_1 \mid B_1 \cap B_2) = \frac{0.5 \times (0.98)^2}{0.5 \times (0.98)^2 + 0.5 \times (0.6)^2} = 0.73$$

$$\Pr(A_2 \mid B_1 \cap B_2) = 1 - 0.73 = 0.27$$

As expected, the credibility of expert 1 is further increased.

c. Denote:

$\overline{B}$ = event that test is a failure.

Then

$$\Pr(\bar{B} \mid A_1) = 0.02, \qquad \Pr(\bar{B} \mid A_2) = 0.40$$

$$\Pr(A_1 \mid \bar{B}) = \frac{\Pr(A_1)\Pr(\bar{B} \mid A_1)}{\Pr(A_1)\Pr(\bar{B} \mid A_1) + \Pr(A_2)\Pr(\bar{B} \mid A_2)}$$

$$\Pr(A_1 \mid \bar{B}) = \frac{0.5 \times 0.02}{0.5 \times 0.02 + 0.5 \times 0.40} = 0.05$$

$$\Pr(A_2 \mid \bar{B}) = 1 - 0.05 = 0.95$$

The credibility of expert 2 estimation substantially rises due to the originally pessimistic reliability value that he or she had provided.

d. Denote:

B_1 = event that the first test is successful,

$\overline{B_2}$ = event that the second test is a failure.

Then

$$\Pr(A_1 \mid B_1 \cap \bar{B}_2) = \frac{\Pr(A_1)\Pr(B_1 \cap \bar{B}_2 \mid A_1)}{\Pr(A_1)\Pr(B_1 \cap \bar{B}_2 \mid A_1) + \Pr(A_2)\Pr(B_1 \cap \bar{B}_2 \mid A_2)}$$

$$\Pr(A_1 \mid B_1 \cap \bar{B}_2) = \frac{(0.5)(0.98 \times 0.02)}{(0.5)(0.98 \times 0.02) + (0.5)(0.6 \times 0.4)} = 0.07$$

$$\Pr(A_2 \mid B_1 \cap \bar{B}_2) = 1 - 0.07 = 0.93$$

Clearly the results fall between the results of b. and c.

Example 2.7

Suppose that 70% of an inventory of memory chips used by a computer manufacturer comes from Vendor 1 and 30% from Vendor 2, and that 99% of chips from Vendor 1 and 88% of chips from Vendor 2 are not defective. If a chip

from the manufacturer's inventory is selected and is defective, what is the probability that the chip was made by Vendor 1. What is the probability of selecting a defective chip (irrespective of the vendor)?

Solution:

Let

A_1 = event that a chip is supplied by Vendor 1,
A_2 = event that a chip is supplied by Vendor 2,
E = event that a chip is defective,
$E \mid A_1$ = event that a chip known to be made by Vendor 1 is defective,
$A_1 \mid E$ = event that a chip known to be defective is made by Vendor 1.

Thus,

$$\Pr(A_1) = 0.7,\ \Pr(A_2) = 0.3,\ \Pr(E \mid A_1) = 1 - 0.99 = 0.01,\ \Pr(E \mid A_2) = 1 - 0.88 = 0.12$$

Using (2.24),

$$\Pr(A_1 \mid E) = \frac{0.7 \cdot 0.01}{0.7 \cdot 0.01 + 0.3 \cdot 0.12} = 0.163$$

Thus, the prior probability that Vendor 1 was the supplier 0.7, is changed to a posterior probability of 0.163 in light of evidence that the chosen unit is defective. From the denominator of (2.24),

$$\Pr(E) = \sum_{i=1}^{n} \Pr(A_i) \Pr(E \mid A_i) = 0.043$$

Example 2.8

A passenger air bag (PAB) disable switch is used to deactivate the PAB in cases when the passenger seat of a car is not occupied. This saves the PAB from being wasted when the car gets into a frontal collision. The switch itself is an expensive component, so its feasibility needs to be justified based on the probability of a passenger seat being occupied, given a collision happened. Available data show that a commercial van driver usually has a passenger 30% of the time. Besides, the expert opinion analysis indicates that the driver is 40% less likely to get into a collision with a passenger than without one. Given a frontal collision has occurred, what is the probability of the passenger seat in a commercial van being occupied?

Solution:

Denote:

$\Pr(P)$ = probability of a passenger seat being occupied,

$\Pr(\bar{P})$ = probability of a passenger seat not being occupied,

$\Pr(C)$ = probability of getting into a collision,

$\Pr(C \mid P)$ = probability of getting into a collision, given the passenger seat is occupied,

$\Pr(P \mid C)$ = probability of the passenger seat being occupied, given a collision happened.

Using the Bayes' theorem, the probability in question can be found as:

$$\Pr(P \mid C) = \frac{\Pr(C \mid P)\Pr(P)}{\Pr(C \mid P)\Pr(P) + \Pr(C \mid \bar{P})\Pr(\bar{P})}$$

According to the statement of the problem, the probabilities of getting into a collision with and without a passenger are related to each other in the following manner:

$$\Pr(C \mid P) = 0.4 \Pr(C \mid \bar{P})$$

Keeping this in mind, as well as that $\Pr(P) = 1 - \Pr(\bar{P})$, one finally gets:

$$\Pr(P \mid C) = \frac{\Pr(P)}{\frac{1}{0.4}(1 - \Pr(P)) + \Pr(P)} = \frac{0.3}{\frac{1}{0.4}(1 - 0.3) + 0.3} = 0.146$$

Therefore, in one hundred similar collisions, 15 are expected to have the passenger seat occupied.

2.3 PROBABILITY DISTRIBUTIONS

In this section, we concentrate on basic probability distributions that are used in mathematical theory of reliability and reliability data analysis. In Chapter 3, we also discuss applications of certain probability distributions in reliability analysis. A fundamental aspect in describing probability distributions is the concept of a *random variable* (r.v.). We begin with this concept and then continue with the basics of probability distributions applied in reliability analysis.

2.3.1 Random Variable

Let's consider an experiment with a number of possible outcomes. If the occurrence of each outcome is governed by chance (random outcome), then possible outcomes may be assigned a numerical value.

An upper case letter (e.g., X, Y) is used to represent a random variable (r.v.), and a lower case letter is used to determine the numerical value that the r.v. can take. For example, if r.v. X represents the number of system breakdowns during a given period of time t (e.g., number of breakdowns per year) in a process plant, then x_i shows the actual number of observed breakdowns.

Random variables can be divided into two classes, namely, *discrete* and *continuous*. A r.v. is said to be discrete if its sample space is countable, such as the number of system breakdowns in a given period of time. A r.v. is said to be continuous if it can take on a continuum of values. That is, it takes on values from an interval(s) as opposed to a specific countable number. Continuous r.v.s are a result of measured variables as opposed to counted data. For example, the operation of several light bulbs can be modeled by a r.v. T, which takes on a continuous survival time t for each light bulb. Clearly, time t is not countable.

2.3.2 Some Basic Discrete Distributions

Consider a discrete random variable, X. The probability distribution for a discrete r.v. is usually denoted by the symbol $\Pr(x_i)$, where x_i is one of the values that r.v. X takes on. Let r.v. X have the sample space S designating the countable realizations of X which can be expressed as $S = \{x_1, x_2, \ldots, x_k\}$ where k is a finite or infinite number. The discrete probability distribution for this space is then a function $\Pr(x_i)$, such that

$$\Pr(x_i) \geq 0, \quad i = 1, 2, \ldots, k,$$

and

$$\sum_{i=1}^{k} \Pr(x_i) = 1 \tag{2.25}$$

Discrete Uniform Distribution

Suppose that all possible k outcomes of an experiment are equally likely. Thus, for the sample space $S = \{x_1, x_2, \ldots, x_k\}$ one can write

$$\Pr(x_i) = p = 1/k, \qquad \textit{for } i = 1, 2, \ldots, k \tag{2.26}$$

A traditional model example for this distribution is rolling a die. If r.v. X describes the numbered 1 to 6 faces, then the discrete number of outcomes is $k = 6$. Thus, $\Pr(x_i) = 1/6$, $x_i = 1, 2, \ldots, 6$.

Binomial Distribution

Consider a random trial having two possible outcomes, for instance, success, with probability p, and failure with probability $1 - p$. Consider a series of n independent trials with these outcomes. Let r.v. X denote the total number of successes. Since the number is a nonnegative integer, the sample space is $S = \{0, 1, 2, \ldots, n\}$. The probability distribution of r.v. X is given by the binomial distribution:

$$\Pr(x) = \binom{n}{x} p^x (1 - p)^{n-x}, \quad x = 0, 1, 2, \ldots, n \tag{2.27}$$

which gives the probability that a known event or outcome occurs exactly x times out of n trials. In (2.27), x is the number of times that a given outcome has occurred. The parameter p indicates the probability that a given outcome will occur. The symbol $\binom{n}{x}$ denotes the total number of ways that a given outcome can occur without regard to the order of occurrence. By definition,

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \tag{2.28}$$

where $n! = n(n-1)(n-2), \ldots, 1$, and $0! = 1$.

In the following examples, the binomial probability is treated (in the framework of classical statistical inference approach) as a constant nonrandom quantity. The situations where this probability or other distribution parameters are treated as random are considered in the framework of Bayes' approach discussed in Chapter 3.

Example 2.9

A random sample of 15 valves is observed. From past experience, it is known that the probability of a given failure within 500 hours following maintenance is 0.18. Calculate the probability that these valves will experience 0, 1, 2, . . . , 15, independent failures within 500 hours following their maintenance.

Solution:

Here the r.v. X designates the failure of a valve that can take on values of 0, 1, 2, . . . , 15, and $p = 0.18$. Using (2.27),

x	$\Pr(x_i)$	x	$\Pr(x_i)$	x	$\Pr(x_i)$	x	$\Pr(x_i)$
0	5.10×10^{-2}	4	1.61×10^{-1}	8	2.90×10^{-3}	12	3.15×10^{-7}
1	1.68×10^{-1}	5	7.80×10^{-2}	9	4.66×10^{-4}	13	1.60×10^{-8}
2	2.58×10^{-1}	6	5.40×10^{-2}	10	5.66×10^{-5}	14	4.95×10^{-10}
3	2.45×10^{-1}	7	1.40×10^{-2}	11	5.26×10^{-6}	15	6.75×10^{-12}
							$\sum \Pr(x_i) = 1.00$

Example 2.10

In a process plant, there are two identical diesel generators for emergency ac needs. One of these diesels is sufficient to provide the needed emergency ac. Operational history indicates that there is one failure on 100 demands for each of these diesels.

a. What is the probability that at a given time of demand both diesel generators will fail?
b. If, on the average, there are 12 test related demands per year for emergency ac, what is the probability of at least one failure for diesel A in a given year? (Assume diesel A is demanded first.)
c. What is the probability that for the case described in b., both diesels A and B will fail on demand simultaneously at least one time in a given year?
d. What is the probability in c. of exactly one simultaneous failure in a given year?

Solution:

a. $q = 1/100 = 0.01$ $p = 99/100 = 0.99$

Assume A and B are independent. (See Chapter 7 for dependent failures treatments.)

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B) = 0.01 \cdot 0.01 = 0.0001$$

That is, there is a 1/10,000 chance that both A and B will fail on a given demand.

b. Using the binomial distribution one can find

$$\Pr(X = 0) = \binom{12}{0} (0.01)^0(0.99)^{12} = 0.886$$

which is the probability to observe no failure in 12 trials (demands). Therefore, the probability of at least one failure per diesel generator in a year

$$\Pr(X \geq 1) = 1 - \Pr(X < 1) = 1 - \Pr(X = 0) = 1 - 0.886 = 0.114$$

c. Using the results obtained in a., the probabilities of simultaneous failure and nonfailure of generators A and B are $q = 0.0001$ and $p = 0.9999$, respectively. Then, similarly to the previous case

$$\Pr(Y = 12) = \binom{12}{12} (0.9999)^{12}(0.0001)^0 = 0.9988$$

which is the probability of 12 successes for the system composed of both A *and* B on 12 demands (trials). Therefore, the probability of at least 1 failure of both A and B in a year

$$\Pr(X \geq 1) = 1 - \Pr(X < 1) = 1 - \Pr(X = 0)$$

$$= 1 - \Pr(Y = 12) = 1 - 0.9988 = 0.0012$$

d. For exactly *one simultaneous* failure in a given year,

$$\Pr(X = 1) = \binom{12}{1} (0.0001)^1(0.9999)^{11} = 0.00119$$

Hypergeometric Distribution

The hypergeometric distribution is the only distribution associated with a finite population, considered in this book. Let us have a finite population of N

items among which there are D items of interest, for example, N identical components, among which D components are defective. The probability to find x ($x \leq D$) objects of interest within a sample (without replacements) of n ($n \leq N$) items is given by the hypergeometric distribution:

$$\Pr(x; N, D, n) = \frac{\binom{D}{x}\binom{N-D}{n-x}}{\binom{N}{n}} \tag{2.29}$$

where

$$x = 0, 1, 2, \ldots n; \quad x \leq D; \quad n - x \leq N - D \tag{2.30}$$

The hypergeometric distribution is commonly used in statistical quality control and acceptance-rejection test practice. This distribution approaches the binomial one with parameters $p = D/N$ and n, when the ratio, n/N, becomes small.

Example 2.11

A manufacturer has a stockpile of 286 computer units. It is known that 121 of the units are more reliable than the other units. If a random sample of four computer units is selected without replacement, what is the probability that no units, two units and all four units are from high reliability units?

Solution:

Use (2.29) with

x = number of high reliability units in the sample,
n = number of units in the selected sample,
$n - x$ = number of nonhigh (low) reliability units in the sample,
N = number of units in the stockpile,
D = number of high reliability units in the stockpile,
$N - D$ = number of nonhigh (low) reliability units in the stockpile.

Possible values of x are $0 \leq x \leq 4$. The results of calculations are given in the table below

x	Pr(x)
0	$\dfrac{\binom{121}{0}\binom{286-121}{4-0}}{\binom{286}{4}} = 0.109$
2	$\dfrac{\binom{121}{2}\binom{286-121}{4-2}}{\binom{286}{4}} = 0.360$
4	$\dfrac{\binom{121}{4}\binom{286-121}{4-4}}{\binom{286}{4}} = 0.031$

Poisson Distribution

This model assumes that events of interest are evenly dispersed at random in a time or space domain, with some constant intensity, λ. For example, r.v. X can represent the number of failures observed at a process plant per year (time domain), or the number of buses arriving at a given station per hour (time domain), if they arrive randomly and independently in time. It can also represent the number of cracks per unit area of a metal sheet (space domain). It is clear that a r.v. X following the Poisson distribution is, in a sense, a number of random events, so that it takes on only integer values. If r.v. X follows the Poisson distribution, then

$$\Pr(x) = \frac{\rho^x \exp(-\rho)}{x!}, \quad \text{for } \rho > 0, \quad x = 0, 1, 2, \ldots \tag{2.31}$$

where ρ is the only parameter of the distribution, which is also its mean. For example if X is a number of events observed in a nonrandom time interval, t, then $\rho = \lambda t$, where λ is the so-called *rate* (time domain) or *intensity* (space domain) *of occurrence* of Poisson events.

Example 2.12

A nuclear plant receives its electric power from a utility grid outside of the plant. From past experience, it is known that loss of grid power occurs at a rate of once a year. What is the probability that over a period of 3 years no power outage will occur? That at least two, power outages will occur?

Solution:

Denote, $\lambda = 1/\text{year}$, $t = 3$ years, $\rho = 1 \times 3 = 3$.
Using (2.31) find

$$\Pr(X = 0) = \frac{3^0 \exp(-3)}{0!} = 0.050$$

$$\Pr(X = 1) = \frac{3^1 \exp(-3)}{1!} = 0149$$

$$\Pr(X \geq 2) = 1 - \Pr(x \leq 1) = 1 - \Pr(X = 0) - \Pr(X = 1)$$

$$= 1 - 0.050 - 0.149 = 0.801$$

Example 2.13

Inspection of a high pressure pipe reveals that on average, two pits per meter of pipe has occurred during its service. If the hypothesis that this pitting intensity is constant and the same for all similar pipes is true, what is the probability that there are fewer than five pits in a 10-meter long pipe branch? What is the probability that there are five or more pits?

Solution:

Denote,

$$\lambda = 2, t = 10, \rho = 2 \times 10 = 20$$

Based on (2.31) the probabilities of interest can be found as

$$\Pr(X < 5) = \Pr(X=4) + \Pr(X=3) + \Pr(X=2) + \Pr(X=1) + \Pr(X=0)$$

$$\Pr(X < 5) = \frac{20^4 \exp(-20)}{4!} + \frac{20^3 \exp(-20)}{3!} + \frac{20^2 \exp(-20)}{2!} + \frac{20^1 \exp(-20)}{1!} + \frac{20^0 \exp(-20)}{0!} = 1.69 \times 10^{-5}$$

$$\Pr(X \geq 5) = 1 - 1.69 \times 10^{-5} = 0.99983$$

The Poisson distribution can be used as an approximation to the binomial distribution when the parameter, p, of the binomial distribution is small (e.g., when $p \leq 0.1$) and parameter n is large. In this case, the parameter of the Poisson distribution, ρ, is substituted by np in (2.31). Besides, it should be noted that as ρ increases, the Poisson distribution approaches the Normal distribution with mean and variance of ρ. This asymptotical property is used as the normal approximation for the Poisson distribution.

Example 2.14

A radar system uses 650 similar electronic devices. Each device has a failure rate of 0.00015 per month. If all these devices operate independently, what is the probability that there are no failures in a given year?

Solution:

The average number of failures of a device per year

$$\rho' = 0.00015 \times 12 = 0.0018$$

The average number of failures of a radar system per year

$$\rho = n\rho' = 0.0018 \times 650 = 1.17$$

Finally, the probability of 0 failures per year, according to (2.31) is given by

$$\Pr(X = 0) = \frac{1.17^0 \exp(-1.17)}{0!} = 0.31$$

Geometric Distribution

Consider a series of binomial trials with probability of success, p. Introduce a r.v., X, equal to the length of a series (number of trials) of successes before the first failure is observed. The distribution of r.v. X is given by the geometrical distribution:

$$\Pr(x) = p(1 - p)^{x-1}, \quad x = 1, 2, \ldots \tag{2.32}$$

The term $(1 - p)^{x-1}$ is the probability that the failure will not occur in the first $(x - 1)$ trials. When multiplied by p, it accounts for the probability of the failure in the xth trial.

Example 2.15

In a nuclear power plant diesel generators are used to provide emergency electric power to the safety systems. It is known that 1 out of 52 tests performed on a diesel generator, results in a failure. What is the probability that the failure occurs at the 10th test?

Solution:

Using (2.32) with $x = 10$ and $p = 1/52 = 0.0192$, yields

$$\Pr(x = 10) = (0.0192)(1 - 0.0192)^9 = 0.016$$

The books by Johnson and Kotz (1969), Hahn and Shapiro (1967), and Nelson (1982) are good references for other discrete probability distributions.

2.3.3 Some Basic Continuous Distributions

In this section, we present certain continuous probability distributions that are fundamental to reliability analysis. A continuous r.v. X has a probability of zero of assuming one of the exact values of its possible outcomes. For example, if r.v. T represents the time interval within which a given emergency action is performed by a pilot, then the probability that a given pilot will perform this emergency action, for example, in *exactly* 2 minutes is equal to zero. In this situation, it is appropriate to introduce the probability associated with a small range of values that the r.v. can take on. For example, one can determine $\Pr(t_1 < T < t_2)$, i.e., the probability that the pilot would perform the emergency action sometime between 1.5 and 2.5 minutes. To define probability that a r.v. assumes a value less than

given, one can introduce the so-called *cumulative distribution function* (cdf), $F(t)$, of a continuous r.v. T as

$$F(t) = \Pr(T \le t) \tag{2.33}$$

Similarly, the cdf of a discrete r.v. X is defined as

$$F(x_i) = \Pr(X \le x_i) = \sum_{\text{all } x_i \le x} \Pr(x_i) \tag{2.34}$$

For a continuous r.v., X, a *probability density function* (pdf), $f(t)$, is defined as:

$$f(t) = \frac{dF(t)}{dt}$$

It is obvious that the cdf of a r.v. t, $F(t)$, can be expressed in terms of its pdf, $f(t)$, as:

$$F(t) = \Pr(T \le t) = \int_0^t f(\xi)\, d\xi \tag{2.35}$$

$f(t)$dt is called the *probability element,* which is the probability associated with a small interval dt of a continuous r.v. T.

The cumulative distribution function of any continuous r.v. must satisfy the following conditions:

$$F(-\infty) = 0 \tag{2.36}$$

$$F(\infty) = \int_{-\infty}^{\infty} f(\xi)\, d\xi = 1 \tag{2.37}$$

$$0 \le F(t) = \int_{-\infty}^{t} f(\xi)\, d\xi \le 1 \tag{2.38}$$

and

$$F(x_1) = \int_{-\infty}^{x_1} f(\xi)\, d\xi \le F(x_2) = \int_{-\infty}^{x_2} f(\xi)\, d\xi, \quad \text{if } x_1 \le x_2 \tag{2.39}$$

The cumulative distribution function is used to determine the probability that a r.v., T, falls in an interval (a, b):

$$\Pr(a < T < b) = \int_{-\infty}^{b} f(\xi)\,d\xi - \int_{-\infty}^{a} f(\xi)\,dx = \int_{a}^{b} f(\xi)\,d\xi \tag{2.40}$$

or:

$$\Pr(a < T < b) = F(b) - F(a)$$

For a discrete r.v. X, the cdf is defined in a similar way, so that the analogous equations are

$$\Pr(c < X < d) = \sum_{x_i \le d} p(x_i) - \sum_{x_i \le c} p(x_i) = F(d) - F(c)$$

It is obvious that the pdf $f(t)$ of a continuous r.v. T must have the following properties:

$$f(T) \ge 0 \quad \text{for all } t, \qquad \int_{-\infty}^{\infty} f(t)\,dt = 1$$

and

$$\int_{t_1}^{t_2} f(t)\,dt = \Pr(t_1 < T < t_2)$$

Example 2.16

Let the r.v. T have the pdf

$$f(t) = \begin{cases} \dfrac{t^2}{a}, & 0 < t < 6 \\ 0, & \text{otherwise} \end{cases}$$

What is the value of parameter a? Find $\Pr(1 < T < 3)$.

Solution:

$$\int_{-\infty}^{\infty} f(t)\,dt = \int_{t_1}^{t_2} f(t)\,dt = 1$$

$$\int_{t_1}^{t_2} f(t)\,dt = \int_{0}^{6} \frac{t^2}{a}\,dt = \frac{t^3}{3a}\Big|_0^6 = \frac{216}{3a} - 0 = 1, \quad \text{then } a = \frac{216}{3} = 72$$

$$\Pr(1 < T < 3) = \int_{1}^{3} \frac{t^2}{72}\,dt = \frac{t^3}{216}\Big|_1^3 = \frac{27}{216} - \frac{1}{216} \approx 0.12$$

Normal Distribution

Perhaps the most well known and important continuous probability distribution is the Normal distribution (sometimes called Gaussian distribution). A normal pdf has the symmetric bell-shaped curve shown in Fig. 2.5, called the normal curve. In 1733, De Moivre developed the mathematical representation of the normal pdf, as follows:

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] \quad -\infty < t < \infty, \quad -\infty < \mu < \infty \tag{2.41}$$

where μ and σ are the parameters of the distribution, $\sigma > 0$.

From (2.41) it is evident that once μ and σ are specified, the normal curve can be determined. We will see later in Chapter 3 that the parameter μ, which is referred to as the mean, and the parameter σ, which is referred to as the standard deviation, have special statistical meaning.

According to (2.41), the probability that the r. v. T takes on a value between abscissas $t = t_1$ and $t = t_2$ is given by

$$\Pr(t_1 < T < t_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{t_1}^{t_2} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] dt \tag{2.42}$$

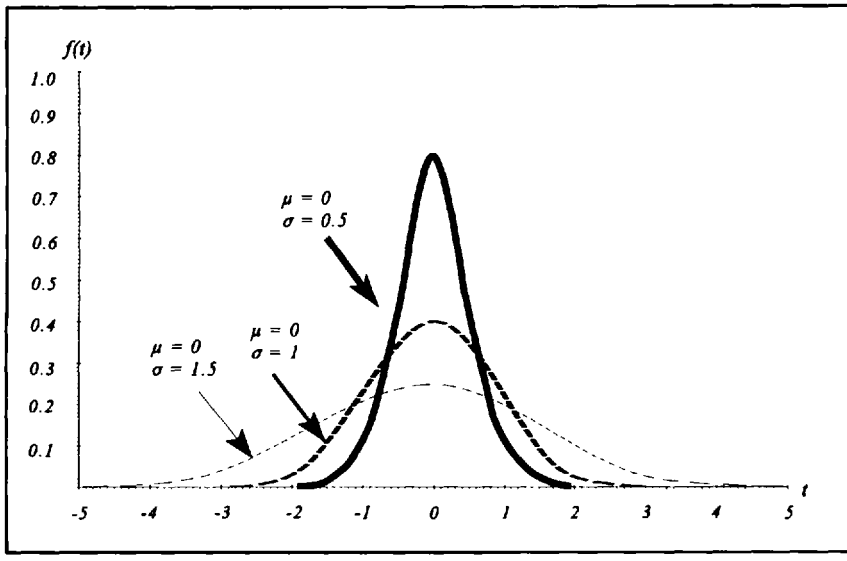


Figure 2.5 Normal distribution.

The integral cannot be evaluated in a closed form, so the numerical integration and tabulation of normal cdf are required. However, it would be impractical to provide a separate table for every conceivable value of μ and σ. One way to get around this difficulty is to use the transformation of the normal pdf to the so-called standard normal pdf which has a mean of zero (μ = 0) and a standard deviation of 1 (σ = 1). This can be achieved by means of the r.v. transformation Z, such that

$$Z = \frac{T - \mu}{\sigma} \tag{2.43}$$

That is, whenever r.v. T takes on a value t, the corresponding value of r.v. Z is given by $z = (t - \mu)/\sigma$. Therefore, if T takes on values $t = t_1$ or $t = t_2$, the r.v. Z takes on values $z_1 = (t_1 - \mu)\sigma$, and $z_2 = (t_2 - \mu)/\sigma$. Based on this transformation, we can write

$$\Pr(t_1 < T < t_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{t_1}^{t_2} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] dt$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{z_1}^{z_2} \exp\left(-\frac{Z^2}{2}\right) dZ \qquad (2.44)$$

$$= \Pr(z_1 < Z < z_2)$$

where Z, also, has the normal pdf with a mean of zero and a standard deviation of 1. Since the standard normal pdf is characterized by fixed mean and standard deviation, only one table is necessary to provide the areas under the normal pdf curves. Table A.1 (see Appendix A) presents the area under the standard normal curve corresponding to $\Pr(a < Z < \infty)$.

Example 2.17

A manufacturer states that his light bulbs have a mean life of 1700 hours and a standard deviation of 280 hours. Assuming the light bulb lives are normally distributed, calculate the probability that a given light bulb will last less than 1000 hours.

Solution:

First, the corresponding Z value is calculated as

$$Z = \frac{1000 - 1700}{280} = -2.5$$

Notice that the lower tail of a normal pdf goes to $-\infty$, so that the formal solution is given by

$$\Pr(-\infty < T < 1000) = \Pr(-\infty < Z < -2.5) = 0.0062$$

which can be represented as

$$\Pr(-\infty < T < 1000) = \Pr(-\infty < T \leq 0) + \Pr(0 < T < 1000)$$

The first term in the right side does not have a physical meaning and it is negligibly small. For the problem considered this probability is

$$\Pr(-\infty < T < 0) = \Pr(-\infty < Z = -1700/280 = -6.07) = 6.42 \times 10^{-10}$$

which can be considered as negligible. So, finally, one can write

$$\Pr(-\infty < T < 1000) \approx \Pr(0 < T < 1000) \approx \Pr(-\infty < Z < -2.5) = 0.0062$$

Lognormal Distribution

A positively defined random variable is said to be lognormally distributed if its logarithm is normally distributed. The lognormal distribution has considerable applications in engineering. One major application of this distribution is to represent a random variable that is the result of multiplication of many independent random variables.

If T is a normally distributed r.v., the transformation $Y = \exp(T)$ transforms the normal pdf representing r.v. T with mean μ_t and standard deviation σ_t to a lognormal pdf, $f(y)$, which is given by

$$f(y) = \frac{1}{\sigma_t\, y\sqrt{2\pi}} \exp\left[-\frac{(\ln y - \mu_t)^2}{2\sigma_t^2}\right], \quad y > 0, \quad -\infty < \mu_t < \infty, \quad \sigma_t > 0 \qquad (2.45)$$

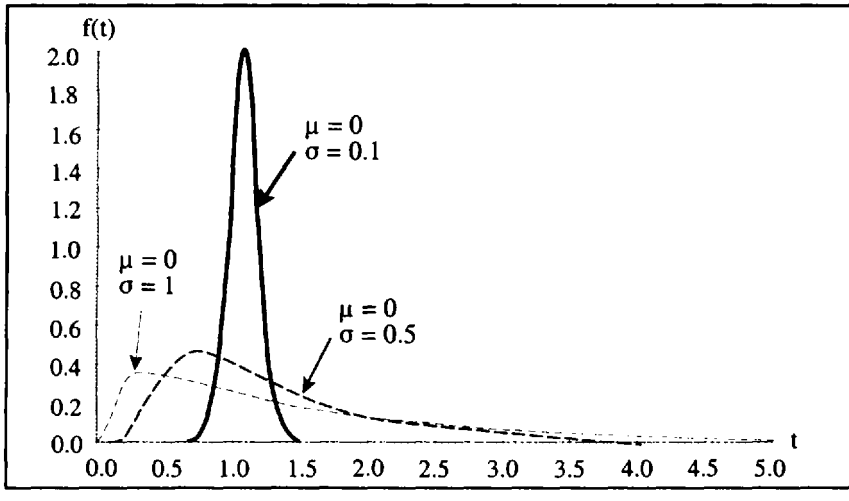


Figure 2.6 Lognormal distribution.

Figure 2.6 shows the pdf's of the lognormal distribution for different values of μ_t and σ_t.

The area under the lognormal pdf curve $f(y)$ between two points, y_1 and y_2, which is equal to the probability that r.v. Y takes a value between y_1 and y_2, can be determined using a procedure similar to that outlined for the Normal distribution. Since logarithm is a monotonous function and ln y is normally distributed, the standard normal r.v. with

$$z_1 = \frac{\ln y_1 - \mu_t}{\sigma_t}$$

and (2.46)

$$z_2 = \frac{\ln y_2 - \mu_t}{\sigma_t}$$

provides the necessary transformation to calculate the probability as follows

$$\begin{aligned} \Pr(y_1 < Y < y_2) &= \Pr(\ln y_1 < \ln Y < \ln y_2) \\ &= \Pr(\ln y_1 < T < \ln y_2) \\ &= \Pr(z_1 < Z < z_2) \end{aligned}$$

If μ_t and σ_t are not known, but μ_y and σ_y are known, the following equations can be used to obtain μ_t and σ_t:

$$\mu_t = \ln\left(\frac{\mu_y}{\sqrt{1 + \frac{\sigma_y^2}{\mu_y^2}}} \right) \tag{2.47}$$

$$\sigma_t = \ln \sqrt{1 + \frac{\sigma_y^2}{\mu_y^2}} \tag{2.48}$$

From (2.47) and (2.48), μ_y and σ_y can also be determined in terms of μ_t and σ_t.

$$\mu_y = \exp\left(\mu_t + \frac{\sigma_t^2}{2}\right) \tag{2.49}$$

$$\sigma_y = \sqrt{\exp\left(\sigma_t^2\right) - 1} \cdot \mu_y \tag{2.50}$$

Exponential Distribution

This distribution was historically the first distribution used as a model of time-to-failure distribution, and it is still the most widely used in reliability problems. The distribution has one-parameter pdf given by

$$\begin{aligned} f(t) &= \lambda \exp(-\lambda t) \quad \lambda, t > 0 \\ &= 0 \qquad\qquad\quad t \le 0 \end{aligned} \tag{2.51}$$

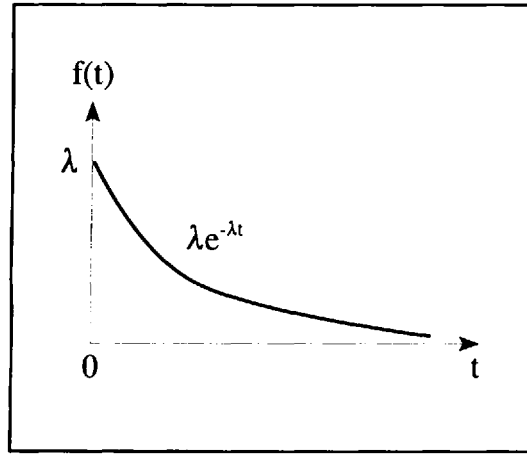


Figure 2.7 Exponential distribution.

Figure 2.7 illustrates the exponential pdf. In reliability engineering applications, the parameter λ is referred to as the *failure rate*. This notion is introduced in

Chapter 3. The exponential distribution is closely associated with the Poisson distribution. Consider the following test. A unit is placed on test at $t = 0$. When the unit fails it is instantaneously replaced by an identical new one, which, in turn, is instantaneously replaced on its failure by another identical new unit, etc. The test is terminated at nonrandom time T. It can be shown that if the number of failures during the test is distributed according to the Poisson distribution with the mean of λT, then the time between successive failures (including the time to the first failure) has the exponential distribution with parameter λ. The test considered is an example of, the so-called, Homogeneous Poisson Process. We will elaborate more on the Homogeneous Poisson Process in Chapter 5.

Example 2.18

A system has a constant failure rate of 0.001 failures per hour. What is the probability that this system will fail before $t = 1000$ hours? Determine the probability that it will work for at least 1000 hours.

Solution:

Calculate the cdf for the exponential distribution at $t = 1000$ hours

$$\Pr(t < 1000) = \int_0^{1000} \lambda \exp(-\lambda t)\, dt = -\exp(-\lambda t) \Big|_0^{1000}$$
$$= 1 - \exp(-1) = 0.632$$

Therefore,

$$\Pr(t > 1000) = 1 - \Pr(t < 1000) = 0.368$$

Weibull Distribution

This distribution is widely used to represent the time to failure or life duration of components as well as systems. The continuous r.v. T representing the time to failure follows a Weibull distribution if its pdf is given by

$$f(t) = \begin{cases} \dfrac{\beta t^{\beta-1}}{\alpha^{\beta}} \exp\left[-\left(\dfrac{t}{\alpha}\right)^{\beta}\right], & t, \alpha, \beta > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.52}$$

Figure 2.8 shows the Weibull pdf's with various values of parameters of α and β. A careful inspection of these graphs reveals that the parameter β determines the shape of the distribution pdf. Therefore, β is referred to as the *shape parameter*. The parameter α, on the other hand, controls the scale of the distribution. For this reason, α is referred to as the *scale parameter*. In the case when $\beta = 1$, the Weibull distribution is reduced to the exponential distribution with $\lambda = 1/\alpha$, so the exponential distribution is a particular case of the Weibull distribution. For the values of $\beta > 1$, the distribution becomes bell-shaped with some skew. We will elaborate on this distribution and its use in reliability analysis, further in Chapter 3.

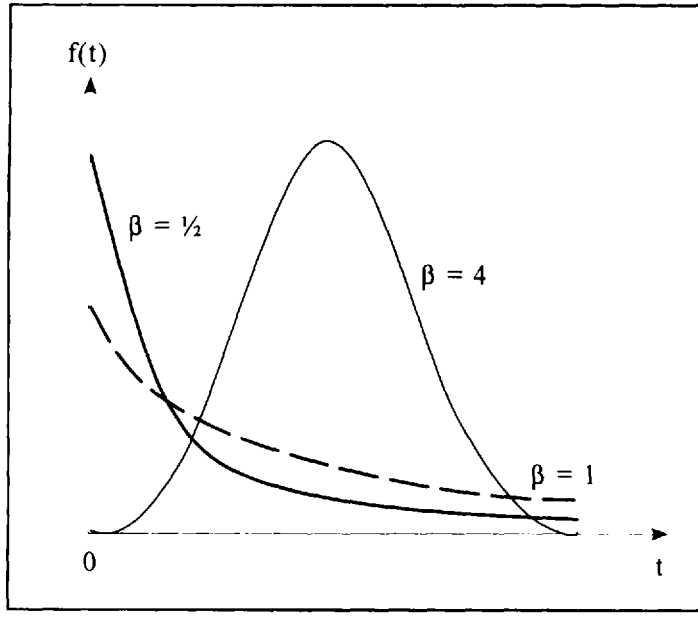


Figure 2.8 Weibull distribution.

Example 2.19

For the Weibull distribution with the shape parameter β and the scale parameter α, find the cdf.

Solution:

$$F(t) = \Pr(T \le t) = \int_0^t \frac{\beta}{\alpha}\left(\frac{\tau}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{\tau}{\alpha}\right)^{\beta}\right] d\tau = 1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$

Gamma Distribution

The gamma distribution can be thought of as a generalization of the exponential distribution. For example, if the time T_i between successive failures of a system has the exponential distribution, then a r.v. T such that $T = T_1 + T_2 + \ldots + T_n$ follows the gamma distribution. In the given context, T represents the cumulative time to the nth failure.

A different way to interpret this distribution is to consider a situation in which a system is subjected to shocks occurring according to the Poisson process (with parameter λ). If the system fails after receiving n shocks, then the time-to-failure of such a system follows a gamma distribution.

The pdf of the gamma distribution with parameters β and α is given by

$$f(t) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\beta}\right), \quad \alpha, \beta, t > 0 \tag{2.53}$$

where $\Gamma(\alpha)$ denotes the so-called *gamma function* defined as

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x}\, dx \tag{2.54}$$

Note that when α is a positive integer, $\Gamma(\alpha) = (\alpha - 1)!$, but in general the parameter α is not necessarily an integer.

The mean and the variance of gamma distribution are

$$E(t) = \frac{\alpha}{\beta}$$

$$\text{var}(t) = \frac{\alpha}{\beta^2}$$

The parameter α is referred to as the shape parameter and the parameter β is referred to as the scale parameter. It is clear that if $\alpha = 1$, (2.53) is reduced to the exponential distribution. Another important special case of the gamma distribution is the case when $\beta = 2$ and $\alpha = n/2$, where n is a positive integer, referred to as the number of *degrees of freedom*. This one parameter distribution is known as the Chi-square distribution. This distribution is widely used in reliability data analysis.

Chapter 3 provides more information about applications of the gamma distribution in reliability analysis. Figure 2.9 shows the gamma distribution pdf curves for some values of α and β.

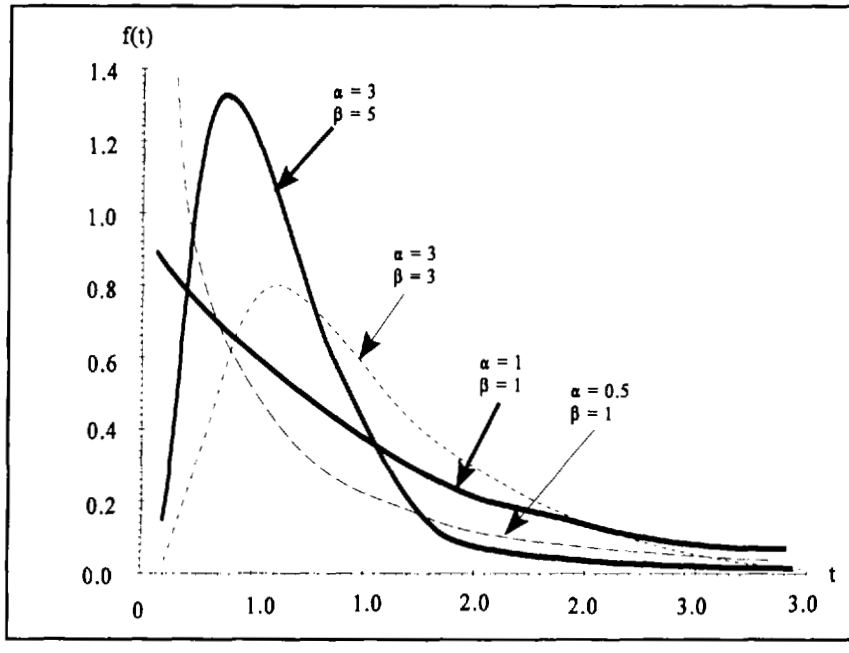


Figure 2.9 Gamma distribution.

Beta Distribution

The beta distribution is a useful model for random variables that are distributed in a finite interval. The pdf of the standard beta distribution is defined over the interval (0,1) as:

$$f(t;\ \alpha,\ \beta) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, t^{\alpha-1}(1-t)^{\beta-1}, & 0 \le t \le 1,\ \alpha > 0,\ \beta > 0 \\ 0, & \text{otherwise} \end{cases} \tag{2.55a}$$

Similar to the gamma distribution, the cdf of the beta distribution cannot be written in closed form. It is expressed in terms of, the so-called incomplete beta function, $I_t(\alpha, \beta)$, i.e.,

$$f(t;\ \alpha,\ \beta) = \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \displaystyle\int_0^t x^{\alpha-1}(1-x)^{\beta-1}\,dx, & 0 \le t \le 1,\ \alpha > 0,\ \beta > 0 \\ 0, & t < 0 \\ 1, & t > 1 \end{cases} \tag{2.55b}$$

The mean value and the variance of the beta distribution are

$$E(t) = \frac{\alpha}{\alpha+\beta}$$

$$\text{var}(t) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{2.56}$$

For the special case of $\alpha = \beta = 1$, the beta distribution reduces to the standard uniform distribution. Practically, the distribution is not used as a time-to-failure distribution. The beta distribution is widely used as an auxiliary distribution in nonparametric classical statistical distribution estimation, as well as a prior distribution in the Bayesian statistical inference. These special applications of beta distribution are discussed in Chapter 3. Figure 2.10 shows the beta distribution pdf curves for some selected values of α and β.

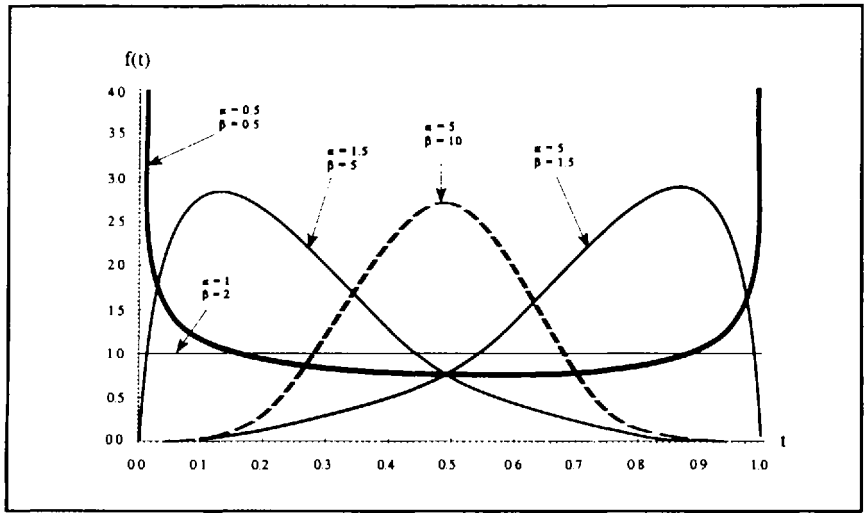


Figure 2.10 Beta distribution.

2.3.4 Joint and Marginal Distributions

Thus far, we have discussed distribution functions that are related to one-dimensional sample spaces. There exist, however, situations in which more than one r.v. is simultaneously measured and recorded. For example, in a study of human reliability in a control room situation, one can simultaneously estimate (1) the r.v. T representing time that various operators spend to fulfill an emergency action, and (2) the r.v. E representing the level of training that these various operators have had for performing these emergency actions. Since one expects E and T to have some relationships (e.g., more trained operators act faster than less trained ones), a joint distribution of both r.v. T and r.v. E can be used to express their mutual dispersion.

Let X and Y be two r.v.s (not necessarily independent). The probability density function for their simultaneous occurrence is denoted by $f(x, y)$ and it is called the *joint probability density function* of X and Y. If r.v. X and Y are discrete, the joint probability density function can be denoted by $\Pr(X = x, Y = y)$, or simply $\Pr(x, y)$. Thus, $\Pr(x, y)$ gives the probability that the outcomes x and y occur simultaneously. For example, if r.v. X represents the number of circuits of a given type in a process plant, and Y represents the number of failures of the

circuit in the most recent year, then Pr(7,1) is the probability that a randomly selected process plant has seven circuits and one of them failed once in the most recent year. The function $f(x, y)$ is a joint probability density function of continuous r.v. X and Y if

1. $f(x, y) \geq 0, \quad -\infty \leq x, y \leq \infty$, and
2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1.$

Similarly, the function $\Pr(x, y)$ is a joint probability function of the discrete random variable X and Y if

1. $\Pr(x, y) \geq 0$, for all values of x and y, and
2. $\sum_{x} \sum_{y} \Pr(x, y) = 1.$

The probability that two or more joint r.v.s fall within a specified subset of the sample space is given by

$$\Pr(x_1 < X \leq x_2, \quad y_1 < Y \leq y_2) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} f(x, y)\, dx\, dy$$

for continuous r.v.s., and by

$$\Pr(x_1 < X \leq x_2, \quad y_1 < Y \leq y_2) = \sum_{x_1 < x \leq x_2,\ y_1 < y \leq y_2} \Pr(x, y)$$

for discrete r.vs.

The *marginal probability density functions* of X and Y are defined respectively as

$$g(x) = \int_{-\infty}^{\infty} f(x, y)\, dy, \tag{2.57}$$

and

$$h(y) = \int_{-\infty}^{\infty} f(x, y)\, dx \tag{2.58}$$

for continuous r.v.s, and by

$$\Pr(x) = \sum_{y} Pr(x, y),$$

and

$$\Pr(y) = \sum_{x} Pr(x, y), \tag{2.59}$$

for the discrete r.v.s.

Using (2.12), the conditional probability of an event y, given event x, is

$$\Pr(Y = y \mid X = x) = \frac{\Pr(X = x \cap Y = y)}{\Pr(X = x)} = \frac{\Pr(x, y)}{\Pr(x)}, \quad \Pr(x) > 0 \tag{2.60}$$

where X and Y are discrete r.v.s. Similarly, one can extend the same concept to continuous r.v.s X and Y and write

$$f(x \mid y) = \frac{f(x, y)}{h(y)}, \quad h(y) > 0$$

or (2.61)

$$f(y \mid x) = \frac{f(x, y)}{g(x)}, \quad g(x) > 0$$

where (2.60) and (2.61) are called the *conditional probability density functions* of discrete and continuous r.v.s, respectively. The conditional probability density functions have the same properties as any other pdf. Similar to (2.11), if r.v. X and r.v. Y are independent, then $f(x \mid y) = f(x)$ for continuous r.v.s, and $\Pr(x \mid y) = \Pr(x)$ for discrete r.v.s. This would lead to the conclusion that for independent r.v.s X and Y,

$$f(x, y) = g(x) \cdot h(y)$$

if X and Y are continuous, and

$$\Pr(x, y) = \Pr(x) \cdot \Pr(y) \tag{2.62}$$

if X and Y are discrete.

Equation (2.62) can be expanded to a more general case as

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1) f_2(x_2), \ldots, f_n(x_n) \tag{2.63}$$

where $f(x_1, x_2, \ldots, x_n)$ is a joint probability density function of r.v.s $X_1, X_2, \ldots, X_n$, and $f_1(x_1), f_2(x_2), \ldots, f_n(x_n)$ are marginal probability density functions of $X_1, X_2, \ldots, X_n$ respectively.

Example 2.20

Let r.v. T_1 represent the time (in minutes) that a machinery operator spends to locate and correct a routine problem, and let r.v. T_2 represent the length of time (in minutes) that he needs to spend reading procedures for correcting the problem. If r.v.s T_1 and T_2 are represented by the joint probability function

$$f(t_1, t_2) = \begin{cases} c(t_1^{1/3} + t_2^{1/5}), & \text{when } 60 > t_1 > 0, \quad 10 < t_2 > 0, \quad \text{and} \\ 0, & \text{otherwise} \end{cases}$$

Find:

a. The value of c.
b. The probability that an operator will be able to take care of the problem in less than 10 minutes. Assume that the operator in this accident should spend less than 2 minutes to read the necessary procedures.
c. Whether r.v. X and r.v. Y are independent.

Solution:

a. $$\Pr(t_1 < 60, t_2 < 10) = \int_0^{t_1 = 10} \int_0^{t_2 = 60} c(t_1^{1/3} + t_2^{1/5})\, dt_1\, dt_2 = 1,$$

$$c = 3.92E-4$$

b. $$\Pr(t_1 < 10, t_2 < 2) = \int_0^{t_1 = 10} \int_0^{t_2 = 2} 3.92E-4\left(t_1^{1/3} + t_2^{1/5}\right)\left(176.17 + 60t_2^{1/5}\right) dt_2\, dt_1 = 0.02$$

c. $$f(t_2) = \int_0^{60} f(t_1, t_2)\, dt_1 = 3.92 \times 10^{-4}\,(176.17 + 60t_2^{1/5})$$

Similarly,

$$f(t_1) = \int_0^{t_2 = 10} f(t_1, t_2)\, dt_2 = 3.92E-4\,(10t_2^{1/3} + 13.21)$$

Since $f(t_1, t_2) \neq f(t_1)\, x\, f(t_2)$, t_1 and t_2 are not independent.

2.4 BASIC CHARACTERISTICS OF RANDOM VARIABLES

In this section, we introduce some other basic characteristics of random variables which are widely used in reliability engineering.

The *expectation* or *expected value* of r.v. X is a characteristic applied to continuous as well as discrete r.v.s. Consider a discrete r.v. X that takes on values x_i with corresponding probabilities $\Pr(x_i)$. The expected value of X denoted by $E(X)$ is defined as

$$E(X) = \sum_i x_i \Pr(x_i) \tag{2.64}$$

Analogously, if T is a continuous r.v. with a pdf $f(t)$, then the expectation of T is defined as

$$E(T) = \int_{-\infty}^{\infty} t f(t)\, dt \tag{2.65}$$

$E(X)$ [or $E(T)$] is a widely used concept in statistics known as the *mean*, or in mechanics known as the center of mass, and sometimes denoted by μ. $E(X)$ is also referred to as the *first moment about the origin*. In general, the kth moment about the origin (ordinary moment) is defined as

$$E(T^k) = \int_{-\infty}^{\infty} t^k f(t)\, dt \tag{2.66}$$

for all integer $k \geq 1$.

In general, one can obtain the expected value of any real-value function of a r.v. In the case of a discrete distribution, $\Pr(x_i)$, the expected value of function $g(X)$ is defined as

$$E[g(X)] = \sum_{i=1}^{k} g(x_i) \Pr(x_i) \tag{2.67}$$

Similarly, for a continuous r.v. T, the expected value of $g(T)$ is defined as:

$$E[g(T)] = \int_{-\infty}^{\infty} g(t) f(t)\, dt \tag{2.68}$$

Example 2.21

Determine (a) the first and (b) the second moments about origin for the Poisson distribution.

Solution:

(a)
$$E(X) = \sum_{x=0}^{\infty} \frac{x \exp(-\rho)\rho^x}{x!} = \sum_{x=1}^{\infty} \frac{x \exp(-\rho)\rho^x}{x!}$$
$$= \rho \sum_{x=1}^{\infty} \frac{\exp(-\rho)\rho^{x-1}}{(x-1)!}$$

Using the substitution $y = x - 1$ the sum can be written as

$$E(X) = \rho \sum_{y=0}^{\infty} \frac{\exp(-\rho)\rho^y}{y!}$$

According to (2.25)

$$\sum_{y=0}^{\infty} \frac{\exp(-\rho)\rho^y}{y!} = 1$$

Thus, $E(X) = \rho$.

(b) Using (2.67), we can write

$$E(X^2) = \sum_{x=0}^{\infty} \frac{x^2 \exp(-\rho)\rho^x}{x!} = \sum_{x=1}^{\infty} \frac{x^2 \exp(-\rho)\rho^x}{x!}$$
$$= \rho \sum_{x=1}^{\infty} \frac{\exp(-\rho)\rho^{x-1}}{(x-1)!}$$

Let $y = x - 1$ [as in (a)]. Then

$$E(X^2) = \rho \sum_{x=0}^{\infty} (y+1) \frac{\exp(-\rho)\rho^y}{y!}$$

$$= \rho \sum_{x=0}^{\infty} \frac{y \exp(-\rho)\rho^y}{y!} + \rho \sum_{x=0}^{\infty} \frac{\exp(-\rho)\rho^y}{y!}$$

Since

$$\sum_{y=0}^{\infty} \frac{y \exp(-\rho)\rho^y}{y!} = \rho, \quad \text{and} \quad \sum_{y=0}^{\infty} \frac{\exp(-\rho)\rho^y}{y!} = 1,$$

$$E(X^2) = \rho^2 + \rho$$

Example 2.22

Find $E(T)$ and $E(T^2)$ if T is an exponential r.v. with parameter λ.

Solution:

Using (2.65) we get

$$E(T) = \lambda \int_0^{\infty} t \exp(-\lambda t)\, dt = -t \exp(-\lambda t) \Big|_0^{\infty}$$

$$+ \int_0^{\infty} \exp(-\lambda t)\, dt = 0 - \frac{\exp(-\lambda t)}{\lambda} \Big|_0^{\infty} = \frac{1}{\lambda}$$

Similarly, by using (2.68)

$$E(T^2) = \int_0^{\infty} t^2 \lambda \exp(-\lambda t)\, dt = \frac{2}{\lambda^2}$$

A measure of dispersion or variation of r.v. about its mean is called the

variance, and is denoted by var(X) or $\sigma^2(X)$. The variance is also referred to as the *second moment about the mean* (which is analogous to the moment of inertia in mechanics), sometimes it is referred to as the *central moment*, and is defined as

$$\text{var}(X) \equiv \sigma^2(X) = E\left[(X - \mu)^2\right] \tag{2.69}$$

where σ is known as the *standard deviation*, σ^2 is known as the *variance*. In general, the kth moment about the mean is defined (similar to (2.66)) as

$$E\left[(X - \mu)^k\right], \quad \text{for all integer } k > 0 \tag{2.70}$$

Table 2.2 represents useful simple algebra associated with expectations. The rules given in Table 2.3 can be applied to discrete as well as continuous r.v.s.

Table 2.2 The Algebra of Expectations

1. $E(aX) = aE(X)$, a = constant
2. $E(a) = a$, a = constant
3. $E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$
4. $E[X \pm Y] = E[X] \pm E[Y]$
5. $E[X \cdot Y] = E[X] \cdot E[Y]$, if X and Y are independent

One useful method of determining the moments about the origin of a distribution is the use of the Laplace transform. Suppose the Laplace transform of pdf $f(t)$ is $F(S)$, then

$$F(S) = \int_0^\infty f(t) \exp(-St)\, dt \tag{2.71}$$

and

$$\frac{-dF(S)}{dS} = \int_0^\infty t f(t) \exp(-St)\, dt \tag{2.72}$$

Since for $S = 0$ the right-hand side of (2.72) reduces to the expectation $E(T)$, then

$$E(T) = -\left[\frac{dF(S)}{dS}\right]_{S=0}$$

In general, it is possible to show that

$$E(T^k) = \left[(-1)^k \frac{d^k F(S)}{dS^k}\right]_{S=0} \tag{2.73}$$

Expression (2.73) is useful to determine moments of pdfs whose Laplace transforms are known or can be easily derived.

Example 2.23

Using the results of Example 2.22, find the variance var(X) for the exponential distribution.

Solution:

From Table 2.2 and (2.69),

$$\text{var}(T) = E(T - \mu)^2 = E(T^2 + \mu^2 - 2\mu T) = E(T^2) + E(\mu^2) - E(2\mu T)$$

Since $E(T^2) = 2/\lambda^2$, and $E(\mu^2) = \mu^2 = [E(T)]^2 = 1/\lambda^2$, then

$$E(2\mu T) = 2\mu E(T) = 2\mu^2 = 2/\lambda^2, \text{ and}$$

$$\text{var}(T) = 2/\lambda^2 + 1/\lambda^2 - 2/\lambda^2 = 1/\lambda^2$$

The concept of expectation equally applies to joint probability distributions. The expectation of a real-value function h of discrete r.v.s $X_1, X_2, \ldots, X_n$ is

$$E\left[h(X_1, X_2, \ldots, X_n)\right] = \sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} h(x_1, x_2, \ldots, x_n) \Pr(x_1, x_2, \ldots, x_n) \tag{2.74}$$

where $\Pr(x_1, x_2, \ldots, x_n)$ is the discrete joint pdf of r.v.s X_i. When dealing with continuous r.v.s, the summation terms in (2.74) are replaced with integrals

$$E[h(X_1, X_2, \ldots, X_n)]$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} h(x_1, x_2, \ldots, x_n) f(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n$$

where $f(x_1, x_2, \ldots, x_n)$ is the continuous joint pdf of r.v.s X_i. In the case of a bivariate distribution with two r.v.s X_1 and X_2, the expectation of the function

$$h(X_1, X_2) = [X_1 - E(X_1)][X_2 - E(X_2)]$$

is called the *covariance* of r.v.s X_1 and X_2, and is denoted by $\text{cov}(X_1, X_2)$. Using Table 2.2 , it is easy to show that

$$\text{cov}(X_1, X_2) = E(X_1 \cdot X_2) - E(X_1)E(X_2) \tag{2.75}$$

A common measure of determining the linear relation between two r.v.s is a *correlation coefficient*, which carries information about two aspects of the relationship:

1. Strength, measured on a scale from 0 to 1; and
2. Direction, indicated by the plus or minus sign.

Denoted by $\rho(X_1, X_2)$, the correlation coefficient between r.v.s X_1 and X_2 is defined as

$$\rho(X_1, X_2) = \frac{\text{cov}(X_1, X_2)}{\sqrt{\text{var}(X_1)\,\text{var}(X_2)}} \tag{2.76}$$

Clearly, if X_1 and X_2 are independent, then from (2.75), $\text{cov}(X_1, X_2) = 0$, and from (2.76), $\rho(x_1, x_2) = 0$. For a linear function of several r.v.s, the expectation and variance are given by

$$E\left(\sum_{i=1}^{n} a_i X_i \right) = \sum_{i=1}^{n} a_i E(X_i) \tag{2.77}$$

$$\text{var}\left(\sum_{i=1}^{n} a_i X_i \right) = \sum_{i=1}^{n} a_i^2 \,\text{var}(X_i) + 2 \sum_{i=1}^{n-1} \sum_{j=1}^{n} a_i a_j \,\text{cov}(X_i, X_j) \tag{2.78}$$

In cases where r.v.s are independent, (2.78) becomes simplified to

$$\text{var}\left(\sum_{i=1}^{n} a_i X_i \right) = \sum_{i=1}^{n} a_i^2 \, \text{var}(X_i) \tag{2.79}$$

Example 2.24

Find the correlation coefficient between r.v.s T_1 and T_2 (see Example 2.20).

$$f(t_1, t_2) = \begin{cases} 3.92E\text{-}4\left(t_1^{1/3}, t_2^{1/5}\right), & 60 > t_1 > 0, \quad 10 > t_2 > 0 \\ 0, & \text{otherwise} \end{cases}$$

Solution:

From Example 2.20, part c,

$$f(t_1) = 3.92E\text{-}4\,[\,10\,t_1^{1/3} + 13.2\,]$$

$$f(t_2) = 3.92E\text{-}4\,[\,176.17 + 60 t_2^{1/5}\,],$$

$$E(t_1) = \int_0^{60} t_1 f(t_1)\, dt_1$$

$$= 3.92E\text{-}4\left[\frac{10(3)}{7} (60)^{7/3} + 13.2\left(\frac{1}{2} \right) (60)^2 \right] = 23.8$$

$$E(t_2) = \int_0^{10} t_2 f(t_2)\, dt_1$$

$$= 3.92E\text{-}4\left[176.2\left(\frac{1}{2} \right)(10)^2 + 60\left(\frac{5}{11} \right)(10)^{11/5} \right] = 5.1$$

$$E(t_1 \cdot t_2) = \int_0^{t_2 = 10} \int_0^{t_1 = 60} 3.92E\text{-}4\; t_1\, t_2 \left[t_1^{1/3} + t_2^{1/5} \right] dt_1\, dt_2 = 169$$

thus,

$$\begin{aligned}\operatorname{cov}(t_1, t_2) &= E(t_1 \cdot t_2) - E(t_1)E(t_2) \\ &= 169 - (23.8)(5.1) = 46.6\end{aligned}$$

Similarly,

$$E(t_1^2) = \int_0^{60} 3.92E-4\left[10t_1^{7/3} + 13.2t_1^2\right]dt_1 = 1367.0$$

$$E(t_2^2) = \int_0^{10} 3.92E-4\left[176.2t_2^2 + 60t_2^{11/5}\right]dt_2 = 34.7$$

$$\operatorname{var}(t_1) = E(t_1^2) - [E(t_1)]^2 = 1367.0 - (23.8)^2 = 800.6$$

$$\operatorname{var}(t_2) = E(t_2^2) - [E(t_2)]^2 = 34.7 - (5.1)^2 = 8.7$$

$$\rho(t_1, t_2) = \frac{\operatorname{cov}(t_1, t_2)}{\sqrt{[\operatorname{var}(t_1)\operatorname{var}(t_2)]}} = \frac{46.6}{\sqrt{800.6 \times 8.7}} = 0.56$$

This indicates that there is a somewhat strong positive correlation between the time the operator spends to solve a machinery problem and the length of time he or she needs to spend on reading the problem correction related procedures.

Example 2.25

Two identical pumps are needed in a process plant to provide a sufficient cooling flow. The flow out of each pump is known to be normally distributed with a mean of 540 gpm and a standard deviation of 65 gpm. Calculate

a. The distribution of the resulting (total) flow from both pumps,
b. The probability that the resulting flow is less than 1000 gpm.

Solution:

a. If M_1 and M_2 are the flows from each pump, then the total flow is M

$= M_1 + M_2$. Since each of the r.v.s M_1 and M_2 are normally distributed, it can be shown that, if M_1 and M_2 are independent, M is also normally distributed. From (2.77), the mean of r.v. M is given by

$$\mu_M = E(M) = E(M_1) + E(M_2) = 540 + 540 = 1080 \text{ gpm}$$

Because M_1 and M_2 are assumed to be independent, then using (2.78) the variance of r.v. M is obtained as

$$\text{var}(M) = \text{var}(M_1) + \text{var}(M_2) = (65)^2 + (65)^2 = 8450, \text{ and}$$

$$\sigma(M) = 91.9 \text{ gpm}$$

b. Using standard normal distribution transformation (2.43)

$$z = \frac{1000 - 1080}{91.9} = -0.87$$

This corresponds to

$$\Pr(M \leq 1000) = \Pr(Z \leq -0.87) = 0.19$$

2.5 ESTIMATION AND HYPOTHESIS TESTING

Reliability and performance data obtained from special tests, experiments or practical use of a product provide a basis for performing statistical inference about underlying distribution. Each observed value is considered as a *realization* (or *observation*) of some hypothetical r.v., that is, a value that the r.v., say X, can take on. For example, the number of pump failures following a demand in a large plant can be considered as realization of some r.v.

A set of observations from a distribution is called a sample. The number of observations in a sample is called the *sample size*. In the framework of classical statistics, a sample is usually composed of random independently and identically distributed observations. From a practical point of view this assumption means that elements of a given sample are obtained independently and under the same conditions.

To check the applicability of a given distribution (for example, binomial distribution in the pump failure case) and to estimate the parameters of the

distribution, one needs to use special statistical procedures known as *hypothesis testing* and *estimation* which are briefly considered below.

2.5.1 Point Estimation

Point and interval estimation are the two basic kinds of estimation procedures considered in statistics. Point estimation provides a single number obtained on the basis of data set (a sample) which represents a parameter of the distribution function or other characteristic of the underlying distribution of interest. As opposed to the interval estimation, the point estimation does not provide any information about its accuracy. Interval estimation is expressed in terms of *confidence intervals*. The confidence interval includes the true value of the parameter with a specified confidence probability.

Suppose, we are interested in estimating a single-parameter distribution $F(X,\theta)$ based on a random sample $x_1, \ldots, x_n$. Let $t(x_1, \ldots, x_n)$ be a single-valued (simple) function of $x_1, \ldots, x_n$. It is obvious that $t(x_1, \ldots, x_n)$ is also a r.v., which is referred to as a *statistic*. A point estimate is obtained by using an appropriate statistic and calculating its value based on the sample data. The statistic (as a function) is called the *estimator*, meanwhile its numerical value is called the *estimate*.

Consider the basic properties of point estimators. An estimator $t(x_1, \ldots, x_n)$ is said to be an *unbiased* estimator for θ if its expectation coincides with the value of the parameter of interest θ, i.e., $E[t(x_1, \ldots, x_n)] = \theta$ for any value of θ. Thus, the bias is the difference between the expected value of an estimate and the true parameter value itself. It is obvious that the smaller the bias, the better the estimator is.

Another desirable property of an estimator $t(x_1, \ldots, x_n)$ is the property of *consistency*. An estimator t is said to be consistent if, for every $\epsilon > 0$,

$$\lim_{n \to \infty} P\left(\left| t(x_1, \ldots, x_n) - \theta \right| < \epsilon\right) = 1 \tag{2.80}$$

This property implies that as the sample size n increases, the estimator $t(x_1, \ldots, x_n)$ gets closer to the true value of θ. In some situations several unbiased estimators can be found. A possible procedure for selecting the best one among the unbiased estimators can be based on choosing one having the least variance. An unbiased estimator t of θ, having minimum variance among all unbiased estimators of θ, is called *efficient*.

Another estimation property is *sufficiency*. An estimator $t(x_1, \ldots, x_n)$ is said to be a sufficient statistic for the parameter θ if it contains all the information about θ that is in the sample $x_1, \ldots, x_n$. In other words the sample $x_1, \ldots, x_n$ can

be replaced by $t(x_1, \ldots, x_n)$ without loss of any information about the parameter of interest θ.

Several methods of estimation are considered in mathematical statistics. In the following section, two of the most common methods, i.e., method of moments and method of maximum likelihood, are briefly discussed.

Method of Moments

In the previous section the mean and the variance of a continuous r.v. X were defined as the expected value of X and expected value of $(x - \mu)^2$, respectively. Quite naturally, one can define the *sample mean* and *sample variance* as the respective expected values of a sample of size n from the distribution of X, namely, $x_1, \ldots, x_n$, as follows:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \tag{2.81}$$

and

$$S^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{2.82}$$

so that, $\bar{x}$ and S^2, can be used as the point estimates of the distribution mean, μ, and variance, σ^2. It should be mentioned that estimator of variance (2.82) is biased, since $\bar{x}$ is estimated from the same sample. However, it can be shown that this bias can be removed by multiplying it by $n/(n - 1)$:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{2.83}$$

Generalizing the examples considered, it can be said that the method of moments is an estimation procedure based on empirically estimated (or *sample*) moments of the random variable. According to this procedure, the sample moments are equated to the corresponding distribution moments. The solutions of the equations obtained provide the estimators of the distribution parameters.

Example 2.26

A sample of eight manufactured shafts is taken from a plant lot. The diameters of the shafts are 1.01, 1.08, 1.05, 1.01, 1.00, 1.02, 0.99, and 1.02 inches. Find the sample mean and variance.

Solution:

From (2.81), $\bar{x} = 1.0225$
From (2.83), $S^2 = 0.0085$

Maximum Likelihood Method

This method is one of the most widely used methods of estimation. Consider a continuous r.v., X, with probability density function $f(X,\theta_0)$, where θ_0 is a parameter. Let us have a sample $x_1, \ldots, x_n$ of size n from the distribution of r.v. X. Under the maximum likelihood approach, the estimate of θ is found as the value of θ_0 which delivers the highest (or *most likely*) probability density of observing the particular set $x_1, \ldots, x_n$. The likelihood of obtaining this particular set of sample values is proportional to the joint probability density function $f(x,\theta_0)$ calculated at the sample points $x_1, \ldots, x_n$. The likelihood function for a continuous distribution is introduced as

$$L(x_1, \ldots, x_n; \theta_0) = f(x_1, \theta_0) f(x_2, \theta_0), \ldots, f(x_n, \theta_0) \tag{2.84}$$

Generally speaking, the definition of the likelihood function is based on the probability (for a discrete random variable) or the probability density function (for continuous random variable) of the joint occurrence of n events, $X = x_1, \ldots, X = x_n$. The maximum likelihood estimate, $\hat{\theta}_0$ is chosen as one that maximizes the likelihood function, $L(x_1, \ldots, x_n; \theta_0)$, with respect to θ_0.

The standard way to find a maximum of a parameter is to calculate the first derivative with respect to this parameter and equate it to zero. This yields the equation:

$$\frac{\partial L(x_1, \ldots, x_n; \theta_0)}{\partial \theta_0} = 0 \tag{2.85}$$

from which the maximum likelihood estimate $\hat{\theta}_0$ can be obtained.

Due to the multiplicative form of the likelihood function, it turns out, in many cases, to be more convenient to maximize the logarithm of the likelihood function instead, i.e., to solve the following equation:

$$\frac{\partial \log L(x_1, \ldots, x_n; \theta_0)}{\partial \theta_0} = 0 \tag{2.86}$$

Because the logarithm is monotonous transformation, the estimate of θ_0 obtained from this equation is the same as that obtained from (2.85). For some cases equations (2.85) or (2.86) can be solved analytically, for other cases they have to be solved numerically.

Under some general conditions, the maximum likelihood estimates are consistent, asymptotically efficient, and asymptotically normal.

Example 2.27

Consider a sample $t_1, \ldots, t_n$ of n times to failure of a component whose time to failure is assumed to be exponentially distributed with parameter λ (the failure rate). Find the maximum likelihood estimator for λ.

Solution:

Using (2.84) and (2.86) one can get

$$L(t,\lambda) = \prod_{i=1}^{n} \lambda \exp(-\lambda t_i) = \lambda^n \exp\left[-\lambda\left(\sum_{i=1}^{n} t_i\right)\right]$$

$$\ln L = n(\ln \lambda) - \lambda \sum_{i=1}^{n} t_i$$

$$\frac{\partial \ln L}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} t_i = 0$$

$$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} t_i}$$

Recalling the second order condition $\partial^2 \ln L / \partial^2 \lambda^2 = -n/\lambda^2 < 0$, it is clear that the estimate $\hat{\lambda}$ is indeed the maximum likelihood estimate for the problem considered. Recalling Example 2.26 it is worth mentioning that the estimate can also be obtained using the method of moments.

 See the software supplement for the automated ML estimation of the parameters for most of the distributions discussed in this chapter.

2.5.2 Interval Estimation and Hypothesis Testing

A two-sided confidence interval for an unknown distribution parameter θ of continuous r.v. X, based on a sample $x_1, \ldots, x_n$ of size n from the distribution of X is introduced in the following way. Consider two statistics $\theta_l(x_1, \ldots, x_n)$ and $\theta_u(x_1, \ldots, x_n)$ chosen in such a way that the probability that parameter θ_0 lies in an interval $[\theta_l, \theta_u]$ is

$$\Pr\left[\theta_l(x_1, \ldots, x_n) < \theta_0 < \theta_u(x_1, \ldots, x_n)\right] = 1 - \alpha \tag{2.87}$$

The *random* interval $[l,u]$ is called a $100(1 - \alpha)\%$ *confidence interval* for the parameter θ_0. The endpoints l and u are referred to as the $100(1 - \alpha)\%$ upper and lower confidence limits of θ_0; $(1 - \alpha)$ is called the *confidence coefficient* or *confidence level*. The most commonly used values for α are 0.10, 0.05, and 0.01. In the case when $\theta_0 > \theta_l$ with the probability of 1, θ_u is called the *one-sided upper confidence limit* for θ_0. In the case when $\theta_0 < \theta_u$ with probability of 1, θ_l is the *one-sided lower confidence limit* for θ_0. A $100(1 - \alpha)\%$ confidence interval for an unknown parameter θ_0 is interpreted as follows: if a series of repetitive experiments (tests) yields random samples from the same distribution and the same confidence interval is calculated for each sample, then $100(1 - \alpha)\%$ of the constructed intervals will, *in the long run*, contain the true value of θ_0.

Consider a typical example illustrating the basic idea of confidence limits construction. Consider a procedure for constructing confidence intervals for the mean of a normal distribution with known variance. Let $x_1, x_2, \ldots, x_n$, be a random sample from the normal distribution, $N(\mu, \sigma^2)$, in which μ is unknown, and σ^2 is assumed to be known. It can be shown that the sample mean $\bar{X}$ (as a statistic) has the normal distribution $N(\mu, \sigma^2/n)$. Thus, $(\bar{X} - \mu)/\sqrt{n}/\sigma$ has the standard normal distribution. Using this distribution one can write

$$\Pr\left(-z_{1-\alpha/2} \le \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \le z_{1-\alpha/2} \right) = 1 - \alpha \tag{2.88}$$

where $z_{1-\alpha/2}$ is the $100(1 - \alpha/2)$th percentile of the standard normal distribution, which can be obtained from Table A.1. After simple algebraic transformations, the inequalities inside the parentheses of Equation (2.88) can be rewritten as

$$\Pr\left(\bar{X} - z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha \tag{2.89}$$

Equation (2.89) provides the symmetric $(1 - \alpha)$ confidence interval of interest.

Generally, a two-sided confidence interval is wider for a higher confidence level $(1 - \alpha)$. As the sample size n increases, the confidence interval becomes shorter for the same confidence coefficient $(1 - \alpha)$.

In the case when σ^2 is unknown, and it is estimated using (2.83), the respective confidence interval is given by

$$\Pr\left(\bar{x} - t_{\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2} \frac{S}{\sqrt{n}} \right) = 1 - \alpha \tag{2.90}$$

where $t_{\alpha/2}$ is the percentile of t-student distribution with $(n - 1)$ degrees of freedom. Values of t_α for different numbers of degrees of freedom are given in Table A.2. Confidence intervals for σ^2 for a normal distribution can be obtained as

$$\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \tag{2.91}$$

where $\chi^2_{1-\alpha}(n - 1)$ is the percentile of χ^2 distribution with $(n - 1)$ degrees of freedom which are given in Table A.3.

Hypothesis Testing

Interval estimation and hypothesis testing may be viewed as mutually inverse procedures. Let us consider a r.v. X with a known probability density function $f(x; \theta)$. Using a random sample from this distribution one can obtain a point estimate $\hat{\theta}$ of the parameter θ. Let θ have a hypothesized value of $\theta = \theta_0$. Under these quite realistic conditions, the following question can be raised: Is the $\hat{\theta}$ estimate compatible with the hypothesized value θ_0? In terms of *statistical hypothesis testing* the statement $\theta = \theta_0$ is called the *null hypothesis,* which is denoted by H_0. For the case considered it is written as

$$H_0: \theta = \theta_0$$

The null hypothesis is always tested against an *alternative hypothesis,* denoted by H_1, which for the case considered might be the statement $\theta \neq \theta_0$, which is written as

$$H_1: \theta \neq \theta_0$$

The null and alternative hypotheses are also classified as *simple* (or *exact* when

they specify exact parameter values) and *composite* (or *inexact* when they specify an interval of parameter values). In the considered example, H_0 is simple and H_1 is composite. An example of a simple alternative hypothesis might be H_1: $\theta = \theta^*$.

For testing statistical hypotheses *test statistics* are used. In many situations the test statistic is the point estimator of the unknown distribution. In this case, as in the case of the interval estimation, one has to obtain the distribution of the test statistic used.

Recall the example considered above. Let $x_1, x_2, \ldots, x_n$, be a random sample from the normal distribution, $N(\mu, \sigma^2)$, in which μ is an unknown parameter, and σ^2 is assumed to be known. One has to test the simple null hypothesis

$$H_0: \mu = \mu^*$$

against the composite alternative

$$H_0: \mu \neq \mu^*$$

As the test statistic, use the same (2.81) sample mean, $\bar{x}$, which has the normal distribution $N(\mu, \sigma^2/n)$. Having the value of the test statistic $\bar{x}$, one can construct the confidence interval using (2.89) and see whether the value of μ^* falls inside the interval. This is the test of the null hypothesis. If the confidence interval includes μ^*, the null hypothesis is not rejected at *significance level* α.

In terms of hypothesis testing, the confidence interval considered is called the *acceptance region*, the upper and the lower limits of the acceptance region are called the *critical values*, while the significance level α is referred to as a probability of *type I error*. In making a decision about whether or not to reject the null hypothesis, it is possible to commit the following errors:

reject H_0 when it is true (type I error),
do not reject H_0 when it is false (type II error).

The probability of the *type II error* is designated by β. These situations are traditionally represented by the following table:

Decision	State of Nature (True Situation)	
	H_0 is true	H_0 is false
Reject H_0	Type I error	No error
Do not reject H_0	No error	Type II error

It is clear that increasing the acceptance region, decreases α , and simultaneously results in increasing β. The traditional approach to this problem is to keep the probability of type I error α at a low level (0.01, 0.05, or 0.10) and to minimize the probability of a *type II error* as much as possible. The probability of not making of *type II error* is referred to as the *power of the test.* Examples of a special class of hypothesis testing are considered in Section 2.7.

2.6 FREQUENCY TABLES AND HISTOGRAMS

When studying distributions, it becomes convenient to start with some preliminary procedures useful for data editing and detecting outliers by constructing *empirical distributions* and *histograms*. Such preliminary data analysis procedures might be useful themselves (the data speak for themselves), as well as they may be used for other elaborate analyses (the goodness-of-fit testing, for instance).

To illustrate some of these procedures, consider the data set composed of observed times-to-failure of 100 identical devices that are placed on a life test. The observed time-to-failure data are given in Table 2.3.

The measure of interest here is the probability associated with each interval of time-to-failure. This can be obtained using (2.10) i.e., by dividing each interval frequency by the total number of devices tested. Sometimes it is important in reliability estimation to indicate how well a set of observed data fits a known distribution, i.e., to determine whether a hypothesis that the data originate from a known distribution is true. For this purpose, it is necessary to calculate the expected frequencies of failures from the known distribution, and compare them with the observed frequencies. Several methods exist to determine the adequacy of such a fit. We discuss these methods further in Section 2.7.

Example 2.28

Consider the time-to-failure data for a device in Table 2.3. It is believed that data come from an exponential distribution with parameter $\lambda = 0.005$ 1/hr. Determine the expected frequencies.

Solution:

The probability that a r.v. T takes values between 0 and 100 hours according to (2.40) is

$$\Pr(0 < T < 100) = \int_0^{100} 5 \times 10^{-3} \exp(-5 \times 10^{-3} t)\, dt$$

$$= \left[1 - \exp(-5 \times 10^{-3} t)\right] \Big|_0^{1000} = 0.393$$

By multiplying this probability by the total number of devices observed (100), we will be able to determine the expected frequency. The expected frequency here would be $0.393 \times 100 = 39.3$ for the 0–100 interval. The results for the rest of the intervals are shown in Table 2.3.

Table 2.3 Frequency Table

Class interval	Observed frequency	Expected frequency
0–100	35	39.3
100–200	26	23.8
200–300	11	14.5
300–400	12	8.8
400–500	6	5.3
500–600	3	3.2
600–700	4	2.0
700–800	2	1.2
800–900	0	0.7
900–1000	1	0.4

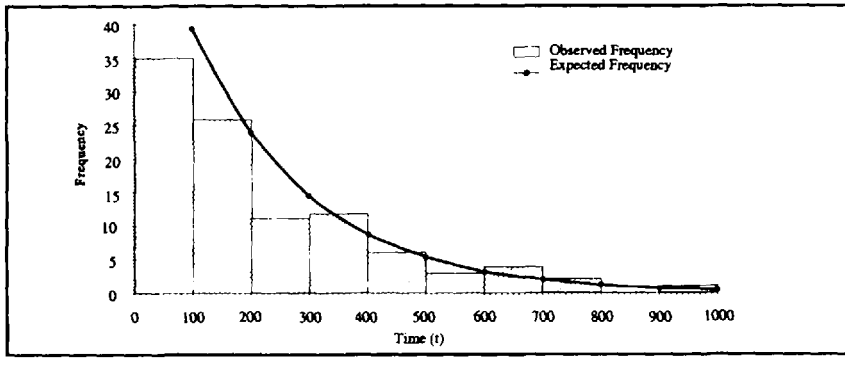


Figure 2.11 Observed frequencies (histogram) and expected frequencies in Example 2.28.

A comparison of the observed and expected frequencies would reveal differences as great as 4.6. Figure 2.11 illustrates the respective graphic representation and its comparison to the exponential distribution with $\lambda = 5 \times 10^{-3}$. The graphic representation of empirical data is commonly known as a histogram.

 See the software supplement for the automated evaluation of expected frequencies for most of the distributions discussed in this chapter.

2.7 GOODNESS-OF-FIT TESTS

Consider the problem of determining whether a sample belongs to a hypothesized theoretical distribution. For this purpose, we need to perform a test that estimates the adequacy of a fit by determining the difference between the frequency of occurrence of a r.v. characterized by an observed sample, and the expected frequencies obtained from the hypothesized distribution. For this purpose, the so-called goodness-of-fit tests are used.

Below we briefly consider two procedures often used as goodness-of-fit tests: the Chi-square and Kolmogorov goodness-of-fit tests.

2.7.1 Chi-Square Test

As the name implies, this test is based on a statistic that has an approximate Chi-square distribution. To perform this test, an observed sample taken from the population representing a r.v. X must be split into k ($k \geq 5$) nonoverlapping intervals (the lower limit for the first interval can be $-\infty$, as well as the upper limit for the last interval can be $+\infty$). The assumed (hypothesized) distribution model is then used to determine the probabilities p_i that the r.v. X would fall into each interval i ($i = 1, \ldots, k$). This process was described to some extent in Section 2.6. By multiplying p_i by the sample size n, we get the expected frequency for each interval. Denote the expected frequency as e_i. It is obvious that $e_i = np_i$. If the observed frequency for each interval i of the sample is denoted by o_i, then the magnitude of differences between e_i and o_i can characterize the adequacy of the fit.

The Chi-square test uses the statistic X^2 which is defined as

$$W = X^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} \tag{2.92}$$

The X^2 statistic approximately follows the chi-square distribution (mentioned in Section 2.3). If the observed frequencies o_i differ considerably from the expected frequencies e_i, then W will be large and the fit is considered to be poor. A good fit would obviously lead to not rejecting the hypothesized distribution, whereas a poor fit leads to the rejection. It is important to note that one can only fail to support the hypothesis, so the person rejects it rather than positively affirm its truth. Therefore, the hypothesis is either *rejected* or *not rejected* as opposed to accepted or not accepted. The test can be summarized as follows:

STEP 1. Choose a hypothesized distribution for the given sample.
STEP 2. Select a specified significance level of the test denoted by α.
STEP 3. Define the rejection region $R \geq \chi^2_{1-\alpha}(k - m - 1)$, where $\chi^2_{1-\alpha}(k - m - 1)$ is the $(1 - \alpha)100$ percentile of the Chi-square distribution with k-m-1 degrees of freedom (the percentiles are given in Table A.3), k is the number of intervals, and m is the number of parameters estimated from the sample. If the parameters of the distribution were estimated without using the given sample, then $m = 0$.
STEP 4. Calculate the value of the Chi-square statistic, W, from (2.92).
STEP 5. If $W > R$, reject the hypothesized distribution; otherwise do not reject the distribution.

It is important at this point to specify the role of α in the chi-square test. Suppose the calculated value of W in (2.92) exceeds the 95th percentile, $\chi^2_{0.95}(\bullet)$ given in Table A.3. This indicates that chances are lower than 1 in 20 that the observed data are from the hypothesized distribution. In this case, the model should be rejected (by not rejecting the model, one makes the type II error discussed above). On the other hand, if the calculated value of W is smaller than $\chi^2_{0.95}(\bullet)$, chances are greater than 1 in 20 that the observed data match the hypothesized distribution model. In this case, the model should not be rejected (by rejecting the model, one makes the type I error discussed above).

One instructive step in Chi-square testing is to compare the observed data with the expected frequencies to note which classes (intervals) contributed most to the value of W. This sometimes could help to indicate the nature of deviations.

Example 2.29

The number of parts ordered per week by a maintenance department in a manufacturing plant is believed to follow a Poisson distribution. Use a chi-square goodness-of-fit test to determine the adequacy of the Poisson distribution. Use the following data (see Figure 2.12).

Solution:

No. of parts per week (x)	Observed frequency (o_i)	Expected frequency (e_i)	X^2 Statistic $(o_i - e_i)^2/e_i$
0	18	15.783	0.311
1	18	18.818	0.036
2	8	11.219	0.923
3	5	4.459	0.066
4	2	1.329	0.339
5	1	0.317	1.472
Total	52	52	3.147

Since under the Poisson distribution model, events occur at a constant rate, then a natural estimate of ρ is

$$\hat{\rho} = \frac{\text{No. of parts used}}{\text{No. of weeks}} = \frac{62}{52} = 1.19 \text{ parts/week}$$

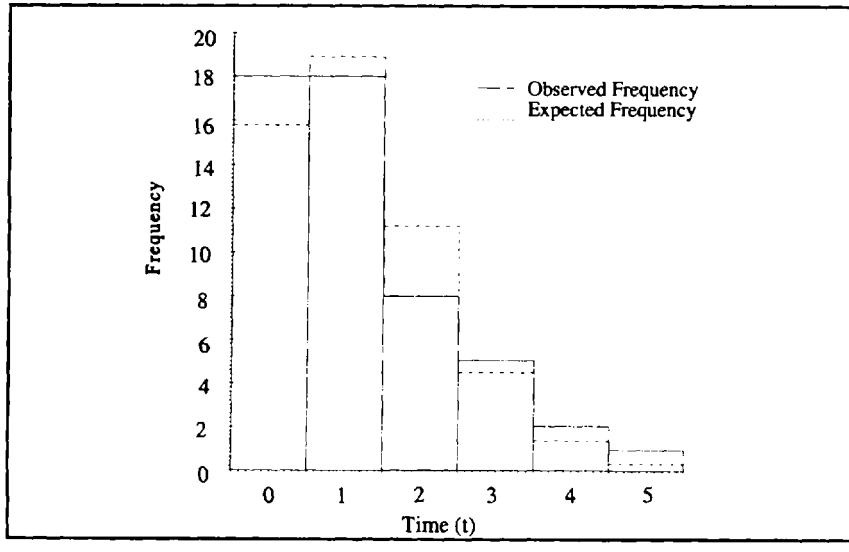


Figure 2.12 Observed and expected frequencies in Example 2.29.

From the Poisson distribution,

$$\Pr(X = x_i) = \frac{\rho \exp(-\rho)}{x_i!}$$

Using $\hat{\rho} = 1.2$, one gets $\Pr(X = 0) = 0.301$. Therefore, $e_i = 0.301 \times 52 = 15.7$. Other expected frequencies are calculated in the same way. Since we obtained one parameter (ρ) from the sample, $m = 1$. Therefore, $R = \chi^2_{0.95}(6 - 1 - 1) = 9.49$, from Table A.3. Since $W = 3.147 < R$, there is no reason to reject the hypothesis that the data are from a Poisson distribution.

Example 2.30

Table 2.4 shows the accumulated mileage for a sample of 100 automobiles after 2 years in service. The mileage accumulation pattern is believed to follow a normal distribution. Use the chi-square test to check this hypothesis at 0.05 significance level (see Figure 2.13).

Table 2.4 The Accumulated Mileage of 100 Passenger Vans after 2 Years in Service

32797	38071	16768	26713	25754	37603	39485	15261	45283	41064
47119	35589	43154	35390	32677	26830	25056	20269	16651	27812
33532	44264	22418	40902	29180	25210	28127	14318	27300	28433
55627	20588	14525	22456	28185	16946	29015	19938	36837	36531
11538	25746	52448	35138	22374	30368	10539	32231	21075	45554
34107	28109	28968	27837	41267	24571	41821	44404	27836	8734
26704	29807	32628	28219	33703	43665	49436	32176	47590	32914
9979	16735	31388	21293	36258	55269	37752	42911	21248	28172
10014	28688	26252	31084	30935	29760	43939	18318	21757	26208
22159	22532	31565	27037	49432	17438	27322	37623	17861	24993

Solution:

Using (2.81) and (2.83), find that the estimates of mean and standard deviation for the hypothesized normal distribution are equal to 30011 miles and 10472 miles, respectively. Group data from Table 2.4 to calculate the observed frequencies. Use the equation of normal pdf (2.41) to find the expected frequencies.

Grouped Data			Estimated	X^2
Interval		Frequency	Frequency	Statistic
Start	End	(o_i)	(e_i)	$(o_i-e_i)^2/e_i$
1E-08	13417	5	5.4479	0.0368
13417	18104	9	7.1245	0.4937
18104	22791	13	11.7500	0.1330
22791	27478	14	15.9143	0.2303
27478	32165	20	17.7014	0.2985
32165	36852	15	16.1699	0.0846
36852	41539	8	12.1305	1.4064
41539	46226	9	7.4733	0.3119
46226	> 46226	7	6.2882	0.0806
Total		*100*	*100.0000*	*3.0758*

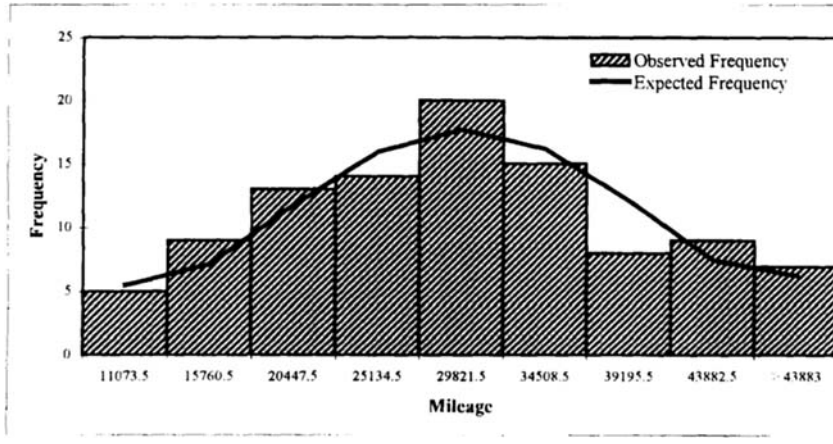


Figure 2.13 Observed and expected frequencies in Example 2.30.

Since both of the distribution parameters were estimated from the given sample, then $m = 2$. The critical chi-square value of the statistic is, therefore, $\chi^2_{0.95}(10 - 2 - 1) = 14.1$. This is higher than the test statistic $W = 3.748$, therefore, there is no reason to reject the hypothesis about the normal distribution at 0.05 significance level.

See the software supplement for the automated Chi-square test.

2.7.2 Kolmogorov Test

In the framework of this test, the individual sample components are treated without clustering them into intervals. Similar to the Chi-square test, a

hypothesized cumulative distribution function, cdf, is compared with its estimate known as *empirical* (or *sample*) *cumulative distribution function*.

A sample cdf is defined for an ordered sample $t_{(1)} < t_{(2)} < t_{(3)} < \ldots < t_{(n)}$ as

$$S_n(t) = \begin{cases} 0 & -\infty < t < t_{(1)} \\ \dfrac{i}{n} & t_{(i)} \le t < t_{(i+1)} \quad i = 1, \ldots, n-1 \\ 1 & t_{(n)} \le t < \infty \end{cases} \tag{2.93}$$

Statistic $K - S$ used in the Kolmogorov test to measure the maximum difference between $S_n(t)$ and a hypothesized cdf, $F(t)$, is introduced as

$$K - S = \max_i \left[\left| F(t_i) - S_n(t_i) \right|, \left| F(t_i) - S_n(t_{i-1}) \right| \right] \tag{2.94}$$

Similar to the chi-square test, the following steps compose the test:

STEP 1. Choose a hypothesized cumulative distribution $F(T)$ for the given sample.
STEP 2. Select a specified significance level of the test, α..
STEP 3. Define the rejection region $R > D_n(\alpha)$, where $D_n(\alpha)$ can be obtained from Table A.4.
STEP 4. If $K - S > D_n(\alpha)$, reject the hypothesized distribution and conclude that $F(t)$ does not fit the data; otherwise, do not reject the hypothesis.

Example 2.31

Time to failure of an electronic device is measured in a life test. The failure times are 254, 586, 809, 862, 1381, 1923, 2542, and 4211 hours. Is the exponential distribution with $\lambda = 5 \times 10^{-4}$ an adequate representation of this sample (see Figure 2.14)?

Solution:

For an exponential distribution with $\lambda = 5 \times 10^{-4}$, we get $F_n(t) = 1 - \exp(-5 \times 10^{-4} t)$. For $\alpha = 0.05$, $D_8(0.05) = 0.457$. Thus, the rejection area is $R > 0.457$.

Time to failure	Empirical cdf			Fitted cdf	$K - S$ Statistic	
t	i	$S_n(t_i)$	$S_n(t_{i-1})$	$F_n(t_i)$	$\lvert F_n(t_i) - S_n(t_i)\rvert$	$\lvert F_n(t_i) - S_n(t_{i-1})\rvert$
254	1	0.125	0.000	0.119	0.006	0.119
586	2	0.250	0.125	0.254	0.004	0.129
809	3	0.375	0.250	0.333	0.042	0.083
862	4	0.500	0.375	0.350	0.150	0.025
1381	5	0.625	0.500	0.499	0.126	0.001
1923	6	0.750	0.625	0.618	0.132	0.007
2542	7	0.875	0.750	0.719	**0.156**	0.031
4211	8	1.000	0.875	0.878	0.122	0.003

Since $K - S = 0.156 < 0.457$, we should not reject the hypothesized exponential distribution model.

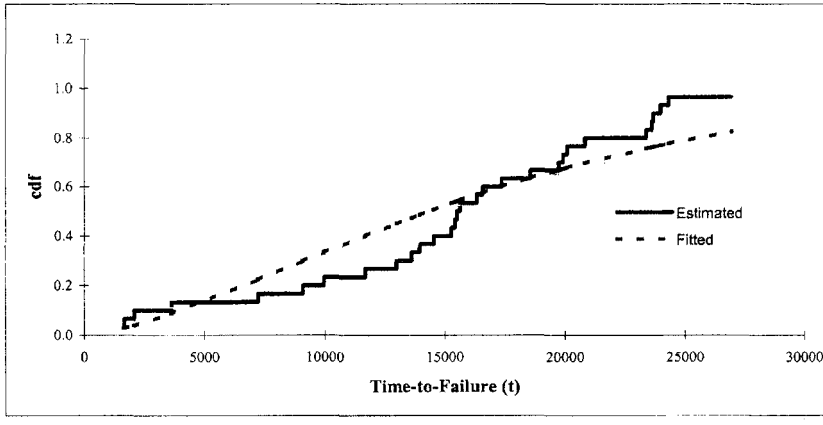


Figure 2.14 Empirical and fitted cdf in Example 2.31.

Table 2.5 Wearout Time of Automobile Brake Pads

Time to failure		Empirical cdf		Fitted cdf	K - S Statistic	
t	i	$S_n(t_i)$	$S_n(t_{i-1})$	$F_n(t_i)$	$\lvert F_n(t_i) - S_n(t_i)\rvert$	$\lvert F_n(t_i) - S_n(t_{i-1})\rvert$
1643	1	0.033	0.000	0.028	0.005	0.028
1664	2	0.067	0.033	0.029	0.038	0.004
2083	3	0.100	0.067	0.040	0.060	0.027
3625	4	0.133	0.100	0.088	0.045	0.012
7230	5	0.167	0.133	0.224	0.057	0.091
9095	6	0.200	0.167	0.299	0.099	0.132
9968	7	0.233	0.200	0.334	0.100	0.134
11689	8	0.267	0.233	0.401	0.135	0.168
12989	9	0.300	0.267	0.451	0.151	0.184
13622	10	0.333	0.300	0.474	0.141	0.174
13953	11	0.367	0.333	0.486	0.119	0.153
14527	12	0.400	0.367	0.506	0.106	0.140
15263	13	0.433	0.400	0.532	0.099	0.132
15428	14	0.467	0.433	0.538	0.071	0.104
15503	15	0.500	0.467	0.540	0.040	0.073
15629	16	0.533	0.500	0.544	0.011	0.044
16342	17	0.567	0.533	0.568	0.001	0.035
16584	18	0.600	0.567	0.576	0.024	0.009
17374	19	0.633	0.600	0.601	0.033	0.001
18571	20	0.667	0.633	0.637	0.030	0.003
19739	21	0.700	0.667	0.670	0.030	0.003
19936	22	0.733	0.700	0.675	0.058	0.025
20102	23	0.767	0.733	0.679	0.087	0.054
20832	24	0.800	0.767	0.698	0.102	0.068
23378	25	0.833	0.800	0.758	0.075	0.042
23612	26	0.867	0.833	0.763	0.103	0.070
23678	27	0.900	0.867	0.765	0.135	0.102
23971	28	0.933	0.900	0.771	0.163	0.129
24341	29	0.967	0.933	0.778	**0.188**	0.155
26964	30	1.000	0.967	0.826	0.174	0.140

Example 2.32

The wearout time (to failure) of automobile brake pads shown in Table 2.5 is believed to follow a Weibull distribution with parameters: α = 18,400 miles and β = 1.5. Use the Kolmogorov-Smirnov test to check this hypothesis at 0.1 significance level.

Solution:

Use (2.52) and (2.94) to compute the expected and empirical cdf, respectively (see Figure 2.15).

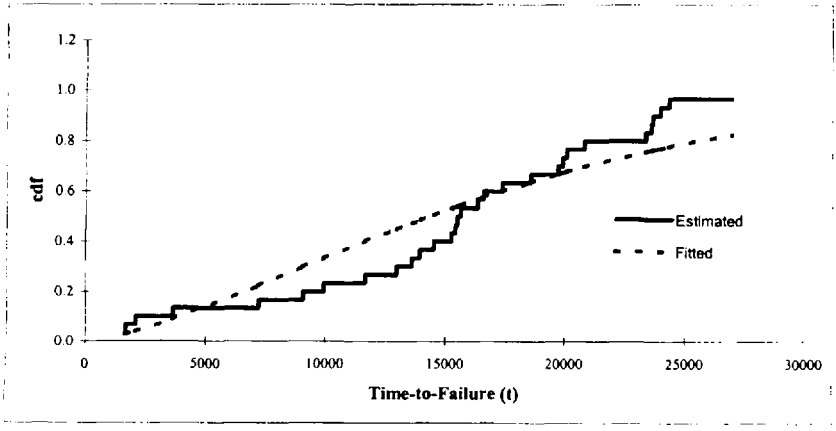


Figure 2.15 Empirical and fitted cdf in Example 2.32.

The $K - S$ statistic of the given data set is equal to 0.188, which is lower than $D_{30}(0.1) = 0.218$. This means that we do not reject the null hypothesis at 0.1 significance level.

See the software supplement for the automated Kolmogorov test calculation.

2.8 REGRESSION ANALYSIS

In Section 2.7, we mainly dealt with one or two random variables. However, reliability and risk assessment problems often require relationships among several random variables or between random and nonrandom variables. For example, time-to-failure of electrical generator can depend on its age, environmental temperature, and power capacity. In this case we can consider the time-to-failure as a random variable Y, which is a function of the variables x_1 (age), x_2 (temperature), and x_3 (power capacity).

In *regression analysis* one refers to Y as the *dependent variable* and to x_1, x_2, . . ., x_k as the *independent variables*, *explanatory variables* or *factors*. Generally speaking, independent variables $x_1,\ldots, x_k$ might be random or nonrandom variables whose values are known or chosen by the experimenter (in the case of the, so-called, *Design of Experiments* (DoE)). The conditional expectation of Y for any given values of $x_1,\ldots, x_k$, $E(Y \mid x_1,\ldots, x_k)$ is known as the *regression* of Y on $x_1,\ldots, x_k$. In other words, regression analysis estimates the average value for the dependent variable corresponding to each value of the independent variable.

In the case when the regression of Y is a linear function with respect to the independent variables $x_1, \ldots, x_k$, it can be written in the form

$$E(Y \mid x_1, \cdots, x_k) = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k \tag{2.95}$$

The coefficients $\beta_0, \beta_1, \ldots, \beta_k$ are called *regression coefficients* or *parameters*. When the expectation of Y is nonrandom, the relationship (2.95) is a deterministic one. The corresponding regression model for the random variable Y can be written in the following form:

$$Y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \epsilon \tag{2.96}$$

where ϵ is the *random error*, assumed to be independent (for all combinations of x considered) r.v. distributed with mean $E(\epsilon) = 0$ and finite variance σ^2. If ϵ is normally distributed, one deals with the *normal* regression.

Simple Linear Regression

Consider the regression model for the simple deterministic relationship

$$Y = \beta_0 + \beta_1 x \tag{2.97}$$

Let us have n pairs of observations $(x_1, y_1), \ldots, (x_n, y_n)$. Also, assume that for any given value x, the dependent variable Y is related to the value of x by

$$Y = \beta_0 + \beta_1 x + \epsilon \tag{2.98}$$

where ϵ is normally distributed with mean 0 and variance σ^2. The r.v. Y has, for a given x, normal distribution with mean $\beta_0 + \beta_1 x$ and variance σ^2. Also suppose that for any given values $x_1, \ldots, x_n$, random variables $Y_1, \ldots, Y_n$ are independent. For the above n pairs of observations the joint pdf of $y_1, \ldots, y_n$ is given by

$$f_n(y \mid x, \beta_0, \beta_1, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \right] \tag{2.99}$$

Function (2.99) is the likelihood function (discussed in Section 2.5) for the parameters β_0 and β_1. Maximizing this function with respect to β_0 and β_1 reduces the problem to minimizing the sum of squares

$$S(\beta_0, \beta_1) = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

with respect to β_0 and β_1.

Thus, the maximum likelihood estimation of the parameters β_0 and β_1 is the estimation by the *method of least squares*. The values of β_0 and β_1 minimizing $S(\beta_0,\beta_1)$ are those for which the derivatives

$$\frac{\partial S(\beta_0, \beta_1)}{\partial \beta_0} = 0, \qquad \frac{\partial S(\beta_0, \beta_1)}{\partial \beta_1} = 0 \tag{2.100}$$

The solution of the above equations yields the least squares estimates of the parameters β_0 and β_1 (denoted $\hat{\beta}_0$ and $\hat{\beta}_1$) as

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}, \qquad \hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}) y_i}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \tag{2.101}$$

where

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i, \qquad \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Note that the estimates are linear functions of the observations y_i, they are also unbiased and have the minimum variance among all unbiased estimates.

The estimate of the dependent variable variance σ^2 can be found as

$$S^2 = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{(n-2)} \tag{2.102}$$

where

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i \tag{2.103}$$

are predicted by the regression model values for the dependent variable, $(n-2)$ is the number of degrees of freedom (2 is the number of the estimated parameters of the model). The estimate of variance of Y (2.102) is also called the *residual variance* and it is used as a measure of accuracy of model fitting as well. The positive square root of S^2 in (2.102) is called the *standard error of the estimate of* Y and the numerator in (2.102) is called the *residual sum of squares*. For more detailed discussion on reliability applications of regression analysis see Lawless (1982).

Example 2.33

An electronic device was tested under the elevated temperatures of 50° C, 60° C, and 70° C. The test results as times to failure for samples of ten items in hours are given in Table 2.6 below (see Figure 2.16).

This is an example of *Accelerated Life Testing* discussed in Chapter 7. Assuming the logarithm of time-to-failure t follows the normal distribution with the mean given by the Arrhenius model, i.e.,

$$E(\ln t) = A + \frac{B}{T}$$

where $T = t°C + 273$ is the absolute temperature, find the estimates of parameters A and B.

Solution:

The equation above can be easily transformed to the simple linear regression (2.97)

$$Y = \beta_0 + \beta_1 x$$

Table 2.6 Times to Failure of an Electronic Component Under Different Temperatures

Time-to-Failure (t)			Ln(t)		
50°C	60°C	70°C	50°C	60°C	70°C
1950	607	44	7.5756	6.4085	3.7842
3418	644	53	8.1368	6.4677	3.9703
4750	675	82	8.4659	6.5147	4.4067
5090	758	88	8.5350	6.6307	4.4773
7588	1047	123	8.9343	6.9537	4.8122
10890	1330	189	9.2956	7.1929	5.2417
11601	1369	204	9.3588	7.2218	5.3181
15288	1884	243	9.6348	7.5412	5.4931
19024	2068	317	9.8535	7.6343	5.7589
22700	2931	322	10.0301	7.9831	5.7746
$E[\ln(t)]$			8.9820	7.0549	4.9037

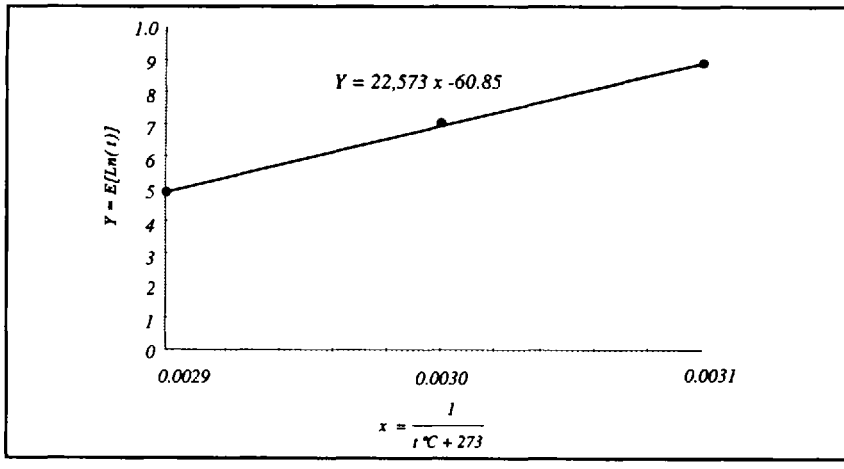


Figure 2.16 Regression line in Example 2.33.

by using transformations $Y = \ln t$, $x = 1/T$, $\beta_0 = A$, $\beta_1 = B$. Accordingly, from the data in Table 2.6:

$Y = E[\ln(t)]$	t C	$x = \frac{1}{t°C + 273}$
8.9820	50	0.0031
7.0549	60	0.0030
4.9037	70	0.0029

Using the transformed data above and (2.101), one can find the estimates of parameters A and B as

$$\hat{A} = \exp(-60.85) = 3.76E-27\, h^{-1}, \qquad \hat{B} = 22573\,°\text{K}$$

Example 2.34

The test data of a new automotive component are expressed in miles-to-failure, while the field (warranty) data on a similar component are represented in months (in service) to failure. To complete the comparative analysis of the new and old component reliability behavior, one needs to translate the mile dependent failure data into the time dependent failure data. Use data in *ex2_34.txt* file of the software supplement to find the miles-to-months in service (MIS) correlation.

Solution:

The problem can be solved by establishing miles-to-MIS regression based on historical patterns of mileage accumulation by the given automobile model. The independent variable x in Figure 2.17 represents the age of a car expressed in months in service. The dependent variable Y represents the accumulated mileage by a given month in service. The straight line is the linear regression of Y on x. The slope ($\beta_1 = 2178.2$) and the intercept ($\beta_0 = 2607.9$) of the regression line are obtained as the least square estimates given by (2.101). The correlation coefficient (2.76) for the given data set is 0.57, which indicates that there is a positive correlation between Miles-to-failure and MIS-to-failure. Therefore, the mileage dependent failure data of the new component can be translated into the time dependent data using the following equation:

$$\text{(Miles-to-failure)} = 2178.2\text{(MIS-to-failure)} + 2607.9$$

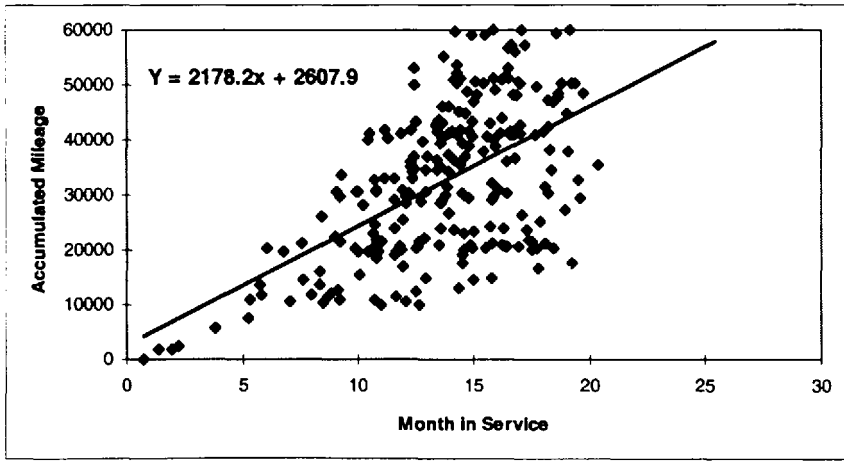


Figure 2.17 Regression line in Example 2.34.

EXERCISES

2.1 Simplify the following Boolean functions:

a. $\overline{\overline{(A \cap B \cup C)} \cap \bar{B}}$
b. $(A \cup B) \cap (\bar{A} \cup \bar{B} \cap \bar{A})$
c. $\overline{A \cap B \cap B \cap \bar{C}} \cap \bar{B}$

2.2 Reduce the following Boolean function:

$$A \cap B \cap \overline{(C \cup (\bar{C} \cup A) \cup \bar{B})}$$

2.3 Simplify the following Boolean expressions:

a) $\overline{\overline{[(A \cap B) \cup C]} \cap \bar{B}}$
b) $[(A \cup B) \cap \bar{A}] \cup (\bar{B} \cap \bar{A})$

2.4 Reduce Boolean function

$$G = (A \cup B \cup C) \cap \overline{(A \cap \bar{B} \cap \bar{C})} \cap \bar{C}$$

If $\Pr(A) = \Pr(B) = \Pr(C) = 0.9$, what is $\Pr(G)$?

2.5 Simplify the following Boolean equations:

a. $(A \cup B \cup C) \cap \overline{(A \cap \bar{B} \cap \bar{C})} \cap \bar{C}$
b. $(A \cup B) \cap \bar{B}$

2.6 Reduce the following Boolean equation:

$$\overline{(A \cup \overline{(B \cap C)})} \cap \overline{(B \cup \overline{(D \cap A)})}$$

2.7 Use both equations (2.17) and (2.21) to find the reliability $\Pr(s)$. Which equation is preferred for numerical solution?

$$\Pr(s) = \Pr(E_1 \cup E_2 \cup E_3), \quad \Pr(E_1) = 0.8,$$

$$\Pr(E_2) = 0.9, \quad \Pr(E_3) = 0.95$$

2.8 A stockpile of 40 relays contain 8 defective relays. If five relays are selected at random and the number of defective relays is known to be greater than two, what is the probability that exactly four relays are defective?

2.9 Given that $P = 0.006$ is the probability of an engine failure on a flight between two cities, find the probability of:

a. No engine failure in 1000 flights
b. At least one failure in 1000 flights
c. At least two failures in 1000 flights

2.10 A random sample of 10 resistors is to be tested. From past experience, it is known that the probability of a given resistor being defective is 0.08. Let X be the number of defective resistors.

a. What kind of distribution function would be recommended for modeling the random variable X?
b. According to the distribution function in (a), what is the probability that in the sample of 10 resistors, there are more than 1 defective resistors in the sample?

2.11 How many different license plates can be made if each consists of three numbers and three letters, and no number or letter can appear more than once on a single plate?

2.12 The consumption of maneuvering jet fuel in a satellite is known to be normally distributed with a mean of 10,000 hours and a standard deviation of 1000 hours. What is the probability of being able to maneuver the satellite for the duration of a 1-year mission?

2.13 Suppose a process produces electronic components, 20% of which are defective. Find the distribution of x, the number of defective components in a sample size of five. Given that the sample contains at least three defective components, find the probability that four components are defective.

2.14 If the heights of 300 students are normally distributed, with a mean of 68 inches and standard deviation of 3 inches, how many students have:

a) heights of more than 70 inches?
b) heights between 67 and 68 inches?

2.15 Assume that 1% of a certain type of resistor are bad when purchased. What is the probability that a circuit with 10 resistors has exactly 1 bad resistor?

2.16 Between the hours of 2 and 4 p.m. the average number of phone calls per minute coming into an office is two and one-half. Find the probability that during a particular minute, there will be more than five phone calls.

2.17 A guard works between 5 p.m. and 12 midnight; he sleeps an average of 1 hour before 9 p.m., and 1.5 hours between 9 and 12. An inspector finds him asleep, what is the probability that this happens before 9 p.m.?

2.18 The number of system breakdowns occurring with a constant rate in a given length of time has a mean value of two breakdowns. What is the probability that in the same length of time, two breakdowns will occur?

2.19 An electronic assembly consists of two subsystems, A and B. Each assembly is given one preliminary checkout test. Records on 100 preliminary checkout tests show that subsystem A failed 10 times. Subsystem B alone failed 15 times. Both subsystems A and B failed together five times.

a) What is the probability of A failing, given that B has failed.
b) What is the probability that A alone fails.

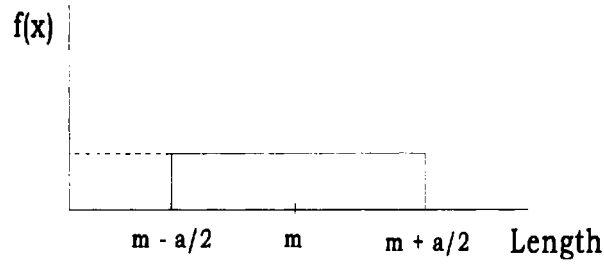
2.20 A presidential election poll shows one candidate leading with 60% of the vote. If the poll is taken from 200 random voters throughout the U.S., what is the probability that the candidate will get less than 50% of the votes in the election? (Assume the 200 voters sampled are true representatives of the voting profile.)

2.21 A newspaper article reports that a New York medical team has introduced a new male contraceptive method. The effectiveness of this method was tested using a number of couples over a period of 5 years. The following statistics are obtained:

Year	Total number of times the method was employed	Number of unwanted pregnancies
1	8200	19
2	10,100	18
3	2120	1
4	6120	9
5	18,130	30

a. Estimate the mean probability of an unwanted pregnancy per use. hat is the standard deviation of the estimate?
b. What are the 95% upper and lower confidence limits of the mean and standard deviation?

2.22 Suppose the lengths of the individual links of a chain distribute themselves with a uniform distribution, shown below.



a. What is the height of the rectangle?
b. Find the cumulative pdf for the above distribution. Make a sketch of the distribution and label the axes.
c. If numerous chains are made from two such links hooked together, what is the pdf of two-link chains ?
d. Consider a 100-link chain. What is the probability that the length of the chain will be less than 100.5 m if $a = 0.1$ m?

2.23 If $f(x,y) = \frac{1}{2}xy^2 + \frac{1}{2}yx^2$, $0 < x < 1$, $0 < y < 2$:

a. Show that $f(x, y)$ is a joint probability density function.
b. Find $\Pr(x > y)$, $\Pr(y > x)$, $\Pr(x = y)$.

2.24 A company is studying the feasibility of buying an elevator for a building under construction. One proposal is a 10-passenger elevator that, on average, would arrive in the lobby once per minute. The company rejects this proposal because it expects an average of five passengers per minute to use the elevator.

a. Support the proposal by calculating the probability that in any given minute, the elevator does not show up, and 10 or more passengers arrive.
b. Determine the probability that the elevator arrives only once in a 5-minute period.

2.25 The frequency distribution of time to establish the root causes of a failure by a group of experts is observed and given below.

Time (hr)	Frequency
45 - 55	7
55 - 65	18
65 - 75	35
75 - 85	28
85 - 95	12

Test whether a normal distribution with known $\sigma = 10$ is an appropriate model for these data.

2.26 A random number generator yields the following sample of 50 digits:

Digit	0	1	2	3	4	5	6	7	8	9
Frequency	4	8	8	4	10	3	2	2	4	5

Is there any reason to doubt the digits are uniformly distributed? (Use the Chi-square goodness-of-fit test.)

2.27 A set of 40 high-efficiency pumps is tested, all of the pumps fail ($F = 40$) after 400 pump-hours ($T = 400$). It is believed that the time to failure of the pumps follows an exponential distribution. Using the following table and the goodness-of-fit method, determine if the exponential distribution is a good choice.

Time interval (hour)	Number of observed failures
0 - 2	6
2 - 6	12
6 - 10	7
10 - 15	6
15 - 25	7
25 - 100	2
	Total = 40

2.28 Use Eq. (2.73) and calculate mean and variance of a Weibull distribution.

2.29 Consider the following repair times

Repair time (y)	0-4	4-24	24-72	72-300	300-5400
No. observed frequency	17	41	12	7	9

Use the Chi-square goodness-of-fit test to determine the adequacy of a lognormal distribution:

a. For 5% level of significance.
b. For 1% level of significance.

2.30 Consider the following time to failure data with the ranked value of t_i. Test the hypothesis that the data fit a normal distribution. (Use the Kolmogorov test for this purpose.)

Event	1	2	3	4	5	6	7	8	9	10
Time to failure (hr)	10.3	12.4	13.7	13.9	14.1	14.2	14.4	15.0	15.9	16.1

2.31 If a device has a cycle-to-failure, t, which follows an exponential distribution with $\lambda = 0.003$ failures/cycle.

a. Determine the mean-cycle-to-failure for this device.
b. If the device is used in a space experiment and is known to have survived for 300 cycles, what is the probability that it will fail sometimes after 1000 cycles?

REFERENCES

1. Cox, R .T., "*Probability, Frequency and Reasonable Expectation*," American Journal of Physics., 14:1, 1946.

2. Hahn, G. J. and Shapiro, S. S., "*Statistical Models in Engineering,*" John Wiley and Sons, New York, 1967.
3. Hill, H. E. and Prane, J. W., "*Applied Techniques in Statistics for Selected Industries: Coatings, Paints and Pigments,*" John Wiley and Sons, New York, 1984.
4. Johnson N. L. and Kotz, S., "*Distribution in Statistics,*" 2 Volumes, John Wiley and Sons, New York, 1970.
5. Lindley, D. V., "*Introduction to Probability and Statistics from a Bayesian Viewpoint,*" 2 Volumes., Cambridge Press, Cambridge, 1965.
6. Nelson, W., "*Applied Life Data Analysis,*" John Wiley and Sons, New York, 1982.
7. Lawless, J. F., "*Statistical Models and Methods for Life Time Data,*" John Wiley and Sons, New York, 1982.

3
Elements of Component Reliability

In this chapter, we discuss the basic elements of component reliability estimation. The discussion centers primarily around the classical frequency approach to component reliability. However, we also present some aspects of component reliability analysis based on Bayesian approach.

We start with a formal definition of reliability and define commonly used terms and metrics. These formal definitions are not necessarily limited to reliability of an actual component; rather, they encompass a broad group of physical items (i.e., components, subsystems, systems, etc.), which are considered as components in the framework of reliability formalism. We then focus on some important aspects of component reliability analysis in the rest of this chapter.

3.1 CONCEPT OF RELIABILITY

Reliability has many connotations. In general, it refers to an item's ability to successfully perform an intended function. The better the item performs its intended function, the more reliable it is. Formally, reliability is viewed as both an *engineering* and a *probabilistic* notion. Indeed, both of these views form the fundamental basis for reliability studies. The reliability engineering notion deals with those design and analysis activities that extend an item's life by controlling its potential failure modes. Examples include designing stronger and more durable elements, parrying harmful environmental conditions, minimizing loads and stresses applied to an item during its use, and providing a preventive maintenance program to minimize the occurrence of failures.

To quantitatively measure the reliability of an item, we use a probabilistic metric, which treats reliability as a probability of the successful achievement of an item's intended function. The formal probabilistic definition of reliability given in Section 1.5, is its mathematical representation. The right-hand side of (1.1) denotes the probability that a specified failure time T exceeds a specified mission time t given that stress conditions $c_1, c_2, \ldots$ are met.

Practically, r.v. T represents *time-to-failure* of an item, and stress conditions $c_1, c_2, \ldots$ represent conditions (e.g., design-related conditions) that are specified, a priori, for successful performance of the item. Other representations of r.v include *number of cycles-to-failure*, or *miles-to-failure* and so on. In the remainder of this book, we consider mainly time-to-failure representation, although the same treatment equally applies to other representations. Conditions $c_1, c_2, \ldots$ are often implicitly considered; therefore, (1.1) is written in a simplified form of (1.2). We use (1.2) in the remainder of this book except for the section on accelerated life testing in Chapter 7.

3.1.1 Reliability Function

Let's start with the formal definition given by expression (1.1). Furthermore, let $f(t)$ denote a pdf representing the r.v. T. According to (2.33), the probability of failure of the item as a function of time is defined by

$$\Pr(T \le t) = \int_0^t f(\theta)\, d\theta = F(t), \quad \text{for } t \ge 0 \tag{3.1}$$

where $F(t)$ denotes the probability that the item will fail sometime up to time t. According to our formalism expressed in (1.2), (3.1) is the *unreliability* of the item. Formally, we can call $F(t)$ (which is the time-to-failure cdf) the *unreliability function*. Conversely, we can define the *reliability function* (a.k.a, *the survivor* or *survivorship function)* as

$$R(t) = 1 - F(t) = \int_t^{\infty} f(\tau)\, d\tau \tag{3.2}$$

The *p-level quantile* of a continuous r.v., T, with cdf, $F(t)$, is defined as the value t_p, such that $F(t_p) = p\,;\ 0 < p < 1$.

The *median* is defined as the quantile of the level of $p = 0.5$. Similar to the mean it is used as a location parameter. A quantile is often referred to as "100p percent point," or "100pth *percentile*." In reliability the 100pth percentile of time-to-failure is the point at which the probability of an item failure is equal to p. For example, the, so-called, B_{10} life of mechanical components, frequently quoted by manufacturers, is the time by which 10% of the components are expected to fail. The most popular percentiles used in reliability are 1, 5, 10, and 50 percentiles.

Provided we have the pdf, $f(t)$, we can get $R(t)$. Basic characteristics of time-to-failure distribution and basic reliability measures can be expressed in terms of pdf, $f(t)$, cdf, $F(t)$, or reliability function, $R(t)$. The *mean time-to-failure* (MTTF), for example, illustrates the expected time during which the item will perform its function successfully (sometimes called *expected life*). According to (2.65),

$$\text{MTTF} = E(t) = \int_0^\infty t f(t)\, dt \tag{3.3}$$

If $\lim_{t \to \infty} t f(t) = 0$, then, integrating by parts, it is easy to get another form of (3.3) given by

$$E(t) = \int_0^\infty R(t)\, dt \tag{3.4}$$

It is important to make a distinction, at this point, between MTTF and the *mean time between failures* (MTBF). Obviously, the former metric is associated with nonrepairable components, whereas the latter is related to the repairable components. In the case of MTBF, the pdf in (3.3) can be the pdf of time between the first failure and the second failure, the second failure and the third failures etc. If we have surveillance and the item is completely renewed through replacement, maintenance, or repair, the MTTF coincides with MTBF. Theoretically, it means that the renewal process is assumed to be perfect. That is, the item that goes through repair or maintenance is assumed to exhibit characteristics of a new item. In practice this may not be true. In this case, one needs to determine the MTBF for the item for each renewal cycle (each, ith time-between-failures interval). However, the approach based on the *as-good-as-new* assumption can be quite adequate for many reliability considerations. In Chapter 5 the topic of MTTF and MTBF will be revisited.

Let $R(t)$ be the reliability function of an item at time t. The probability that the item will survive for time τ, given that it has survived for time t, is called the *conditional reliability function*, and is given by

$$R(\tau \mid t) = \frac{R(t + \tau)}{R(t)} \tag{3.5}$$

Therefore, the conditional probability of failure during the same interval is

$$F(\tau \mid t) = 1 - R(\tau \mid t) \tag{3.6}$$

3.1.2 Failure Rate

The *failure rate,* or *hazard rate,* $h(t)$, is introduced as

$$h(t) = \lim_{\tau \to 0} \frac{1}{\tau}\left(\frac{F(t + \tau) - F(t)}{R(t)} \right) = \frac{f(t)}{R(t)} \tag{3.7}$$

so, it is evident that $h(t)$ is the time-to-failure conditional pdf. The failure rate can also be expressed in terms of the reliability function as

$$h(t) = -\frac{d}{dt}[\ln R(t)] \tag{3.8}$$

so that

$$R(t) = \exp\left[-\int_0^t h(x)\,dx\right] \tag{3.9}$$

The integral of the failure rate in the exponent is known as the *cumulative failure rate*, or *cumulative hazard function*, $H(t)$:

$$H(t) = \int_0^t h(x)\,dx \tag{3.10}$$

As mentioned above, the failure rate can be defined as the conditional pdf of the component time-to-failure, given the component has survived to time t. The expected value associated with such pdf is referred to as the *residual* MTTF.

Example 3.1

A device time-to-failure follows the exponential distribution. If the device has survived up to time t, determine its residual MTTF.

Solution:

According to (3.3),

$$\text{MTTF} = \frac{\int_0^\infty \tau f(t+\tau)\,d\tau}{\int_t^\infty f(\tau)\,d\tau} = \frac{\int_0^\infty \tau\lambda e^{-\lambda(t+\tau)}\,d\tau}{\int_t^\infty \lambda e^{-\lambda T}\,d\tau} = \frac{e^{-\lambda t}\int_0^\infty \tau\lambda e^{-\lambda T}\,d\tau}{e^{\lambda t}} = \frac{1}{\lambda}$$

Let us introduce another useful reliability measure related to failure rate. For a given time interval, t, the average f ailure rate, $\langle h(t)\rangle$, is given by

$$\langle h(t) \rangle = \frac{1}{t} \int_0^t h(x)\, dx \tag{3.11}$$

or

$$\langle h(t) \rangle = -\frac{\log R(t)}{t} \tag{3.12}$$

therefore,

$$\langle h(t) \rangle = \frac{H(t)}{t} \tag{3.13}$$

If the time interval, t, is equal to a given percentile, t_p, then

$$\langle h(t_p) \rangle = -\frac{\log(1-p)}{t_p} \tag{3.14}$$

Hazard rate is an important function in reliability analysis since it shows changes in the probability of failure over the lifetime of a component. In practice, $h(t)$ often exhibits a bathtub shape and it is referred to as a *bathtub curve*. A bathtub curve is shown in Figure 3.1.

Generally, a bathtub curve can be divided into three regions. The, so-called, *burn-in* early failure region exhibits a *decreasing failure rate*, characterized by early failures attributable to defects in design, manufacturing, or construction. A time-to-failure distribution having a decreasing failure rate is referred to as a distribution belonging to the class of decreasing failure rate (DFR) distribution.

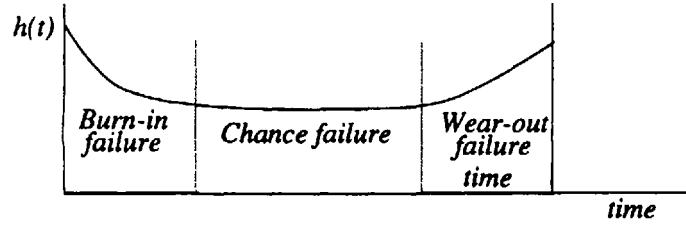


Figure 3.1 Typical bathtub curve.

Analogously, a time-to-failure distribution having a decreasing average failure rate is referred to as a distribution belonging to the class of *decreasing failure rate average* (DFRA) distribution.

The, so-called, *chance failure region* of the bathtub curve exhibits a reasonably *constant failure rate*, is characterized by random failures of the component. In this period, many mechanisms of failure due to complex underlying physical, chemical, or nuclear phenomena give rise to this approximately constant failure rate. The third region, called *wear-out region*, which exhibits an *increasing failure rate*, is characterized mainly by complex aging phenomena. Here the component deteriorates (e.g., due to accumulated fatigue) and is more vulnerable to outside shocks. It is helpful to note that these three regions can be radically different for different types of components. Figure 3.2 and Figure 3.3 show typical bathtub curves for mechanical and electrical devices, respectively. It is evident that electrical devices can exhibit a relatively larger chance failure period. Figure 3.4 shows the effect of various levels of stress on a device. It is clear that as stress level increases, the chance failure region decreases and, premature wear-out occurs. Therefore, it is important to minimize stress factors such as harsh operating environment, to maximize reliability.

Similar to DFR and DFRA distribution, the *increasing failure rate* (IFR) and *increasing failure rate average* (IFRA) distributions are considered in the framework of mathematical theory of reliability (Barlow and Proschan (1981)).

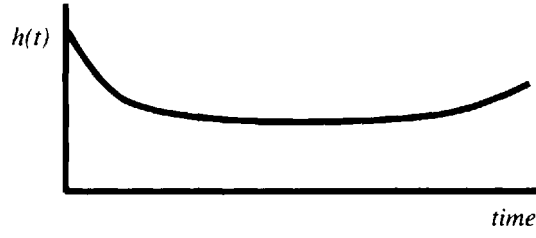


Figure 3.2 A typical bathtub curve for mechanical devices.

Table 3.1 lists the cdf's (unreliability functions) and hazard rate functions for important pdfs.

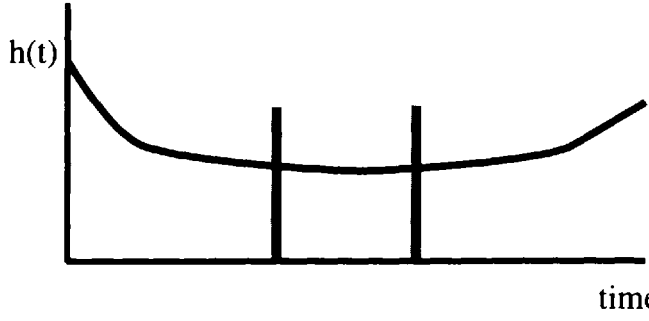


Figure 3.3 Bathtub curve for typical mechanical devices.

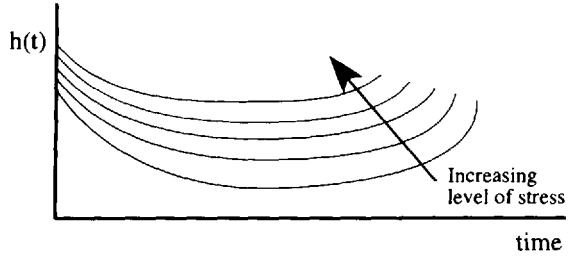
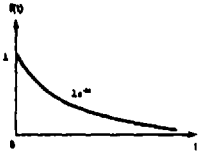
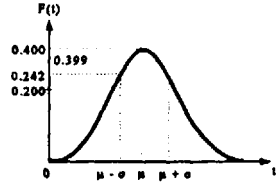
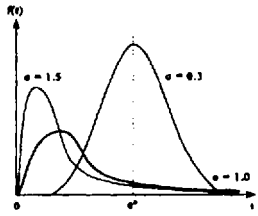
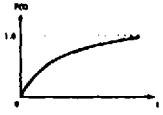
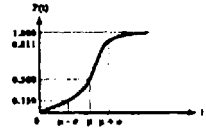
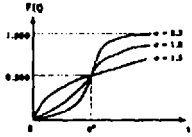
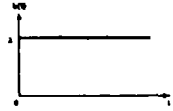
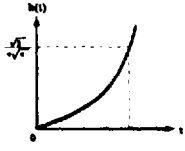
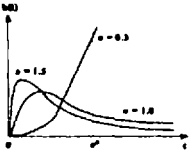
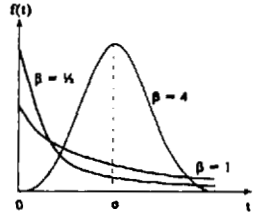
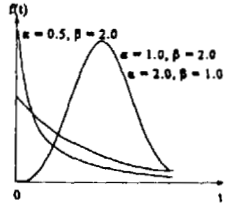
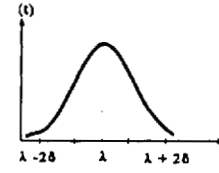
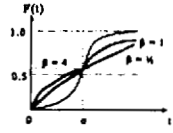
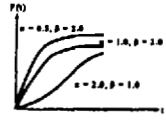
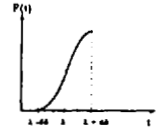
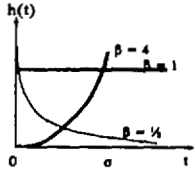
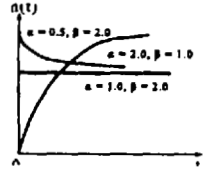
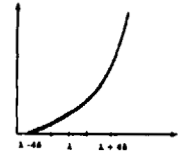


Figure 3.4 Effect of stress on a typical bathtub curve.

Table 3.1 Important Time-to-Failure Distributions and their Characteristics

Distribution characteristic	Exponential distribution	Normal distribution	Lognormal distribution
pdf, $f(t)$	$\lambda \exp(-\lambda t)$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right]$	$\frac{1}{\sigma_t t\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_t^2}(\ln t - \mu)^2\right]$
cdf, $F(t)$	$1-\exp(-\lambda t)$	$\frac{1}{\sigma\sqrt{2\pi}} \int_0^t \exp\left[-\frac{(\theta-\mu)^2}{2\sigma^2}\right] d\theta$	$\frac{1}{\sigma_t\sqrt{2\pi}} \int_0^t \frac{1}{\theta} \exp\left[-\frac{1}{2}\frac{(\ln\theta-\mu_t)^2}{\sigma_t^2}\right] d\theta$
Instantaneous failure rate, $h(t)$	λ	$\frac{f(t)}{1-F(t)}$	$\frac{f(t)}{1-F(t)}$
Mean time to failure (MTTF)	$1/\lambda$	μ	$\exp\left[\left(\mu_t + \frac{1}{2}\sigma_t^2\right)\right]$
$f(t)$			
$F(t)$			
$h(t)$			
Major applications in component reliability	Life distribution of complex nonrepairable systems. Life distribution "burn-in" of some components.	Life distribution of high stress components. Stress-strength analysis Tolerance analysis.	Size distribution of breaks (in pipes, etc.). Life distribution of some transistors. Prior parameter distribution in Bayesian analysis.

Distribution characteristic	Weibull distribution	Gamma distribution	Smallest extreme value distribution
pdf, $f(t)$	$\frac{\beta(t)^{\beta-1}}{\alpha^{\beta}} \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$	$\frac{1}{\beta^{\alpha}\Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\beta}\right)$	$\frac{1}{\delta} \exp\left[\frac{1}{\delta}(t-\lambda) - \exp\left(\frac{t-\lambda}{\delta}\right)\right]$
cdf, $F(t)$	$1 - \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$	$\frac{\int_0^t y^{\alpha-1}\exp(-y/\beta)\,dy}{\beta^{\alpha}\Gamma(\alpha)}$	$1 - \exp\left[-\exp\left(\frac{t-\lambda}{\delta}\right)\right]$
Instantaneous failure rate, $h(t)$	$\frac{\beta}{\alpha}\exp\left(\frac{t}{\alpha}\right)^{\beta-1}$	$\frac{t^{\alpha-1}\exp\left(-\frac{t}{\beta}\right)}{\beta^{\alpha}\left[\Gamma(\alpha) - \int_0^t y^{\alpha-1}\exp(-y/\beta)\,dy\right]}$	$\frac{1}{\delta}\exp\left[\frac{t-\lambda}{\delta}\right]$
Mean time to failure (MTTF)	$\alpha\Gamma\left(\frac{1+\beta}{\beta}\right)$	$\beta\alpha$	$\lambda - 0.5776\delta$
$f(t)$			
$F(t)$			
$h(t)$			
Major applications in component reliability	Corrosion resistance. Life distribution of many basic components, such as capacitors, relays ball bearings, and certain motors.	Distributions of time between recalibration or maintenance of components. Time to failure of system with standby components. Prior distribution in Bayes' estimation.	Distribution of breaking strength of some components Breakdown voltage of capacitors. Extreme natural phenomena, such as temperature and rainfall minima.

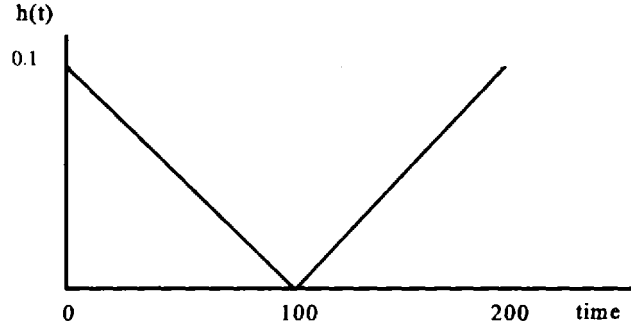
Example 3.2

Failure rate $h(t)$ of a device is approximated by

$$\begin{aligned} h(t) &= +0.1 - 0.001t, \quad 0 \le t \le 100 \\ &= -0.1 + 0.001t, \quad t > 100 \end{aligned}$$

shown in the figure below. Find the pdf and the reliability function for $t \le 200$.

Solution:



For $0 \le t \le 100$,

$$\int_0^t h(\theta)\,d\theta = \int_0^t (0.1 - 0.001\theta)\,d\theta = \left[0.1\theta - \frac{0.001}{2}\theta^2\right]\Bigg|_0^t = 0.1t - 0.0005t^2$$

thus

$$R(t) = \exp\left(-0.1t + 0.0005t^2\right)$$

Using (3.7), one gets

$$f(t) = (0.1 - 0.001t)\exp\left(-0.1t + 0.0005t^2\right)$$

Note that $R(100) = \exp(-5)$, so the solution of the problem for $t > 100$ is of academic interest only.

For $t > 100$,

$$h(t) = -0.1 + 0.001t$$

Accordingly,

$$R(t) = R(100)\exp\left(\int_{100}^{t}(0.1 - 0.001\theta)\,d\theta\right)$$

$$= R(100)\exp\left(0.1t - 0.0005t^2 - 5\right)$$

$$f(t) = (-0.1 + 0.0001t)\,R(100)\exp\left(\int_{100}^{t}(0.1 - 0.001\theta)\,d\theta\right)$$

$$f(t) = (-0.1 + 0.0001t)\,R(100)\exp\left(0.1t - 0.0005t^2 - 5\right)$$

3.2 COMMON DISTRIBUTIONS IN COMPONENT RELIABILITY

Table 3.1 displays some basic reliability characteristics of exponential, normal, lognormal, Weibull, gamma and smallest extreme value distributions, which are commonly used as time-to-failure distribution models for components. Some other characteristics of each of these distributions are further discussed in this section.

3.2.1 Exponential Distribution

The exponential distribution is the most commonly used distribution in reliability analysis. This can be attributed primarily to its simplicity and the fact that it gives the simple, constant hazard rate model, corresponding to a situation that is often realistic. In the context of the bathtub curve, this distribution can simply represent the chance failure region. It is evident that for components whose chance failure region is long, in comparison with the other two regions, this distribution might be adequate. This is often the case for electrical components and mechanical components, especially in certain applications, when new components are screened and only those that are determined to have passed (the burn-in period) are used. For such components, exponential distribution is a reasonable choice. In general, exponential distribution is considered as a good model for representing systems and complex, nonredundant components consisting of many interacting parts.

In Section 2.3 we noted that the exponential distribution can be introduced using the Homogeneous Poisson Process (HPP). Now let's assume that each failure in this process is caused by a random shock, and the number of shocks

occurring in a time interval of length t is described by a Poisson distribution with the mean number of shocks equal to λt. Then, the random number of shocks, n, occurring in the interval $[0, t]$ is given by

$$\Pr[X = n] = \frac{\exp(-\lambda t)(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \ldots, \quad \lambda, t > 0$$

where λ is the rate at which the shocks occur. Since based on this model, the first shock causes component failure, then the component is functioning only when no shocks occur, i.e., $n = 0$. Thus, one can write

$$R(t; \lambda) = \Pr[X = 0] = \exp(-\lambda t) \tag{3.15}$$

Using relationship (3.2), the exponential pdf obviously can be obtained as

$$f(t) = \lambda \exp(-\lambda t) \tag{3.16}$$

Let us now consider one of the most interesting properties of the exponential distribution: *a failure process represented by the exponential distribution has no memory*. Consider the law of conditional probability and assume that an item has survived after operating for a time t. The probability that the item will fail sometime between t and $t + \Delta t$ is

$$\Pr(t \le T \le t + \Delta t \mid T > t) = \frac{\exp(-\lambda t) - \exp(-\lambda(t + \Delta t))}{\exp(-\lambda t)}$$

$$= 1 - \exp(-\lambda \Delta t)$$

which is independent of t. In other words, the component that has worked up to time t has no memory of its past. This property can also be easily described by the shock model. That is, at any point along time t, the rate at which fatal shocks occur is the same regardless of whether any shock has occurred up to time t.

3.2.2 Weibull Distribution

The Weibull distribution has a wide range of applications in reliability analysis. This distribution covers a variety of shapes. Due to its flexibility for describing hazard rates, all three regions of the bathtub curve can be represented by the Weibull distribution. It is possible to show that the Weibull distribution is appropriate for a system or complex component composed of a number of components or parts whose failure is governed by the most severe defect of its components or parts (the, so-called, *weakest link model*). The pdf of the Weibull distribution is given by

$$f(t) = \frac{\beta(t)^{\beta-1}}{\alpha^{\beta}} \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right], \quad \alpha, \beta > 0, \quad t > 0 \tag{3.17}$$

Using (3.7), the failure rate, $h(t)$, can be derived as

$$h(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1}, \quad \alpha, \beta > 0, \quad t > 0 \tag{3.18}$$

Sometimes the transformation $\lambda = 1/\alpha^{\beta}$ is used. In this case (3.18) is transformed to $h(t) = \lambda\beta t^{\beta-1}$. This form will be used later in Chapter 5.

Parameters α and β of the Weibull distribution are referred to as the *scale* and *shape* parameters, respectively. If $0 < \beta < 1$ in (3.18), the Weibull distribution is a DFR distribution which can be used to describe burn-in (early) type failure behavior. For $\beta = 1$, the Weibull distribution reduces to the exponential distribution. If $\beta > 1$, the Weibull distribution can be used as a model for the wearout region of the bathtub curve (as an IFR distribution). Main applications of the Weibull include:

Corrosion resistance studies.

Time to failure of many types of hardware, including capacitors, relays, electron tubes, germanium transistors, photo conductive cells, ball bearings, and certain motors.

Time-to-failure of basic elements of a system (components, parts, etc.), although the time-to-failure of the system itself can be better represented by the exponential distribution.

In some cases, a parameter called *location parameter* is used in the Weibull distribution to account for a period of guaranteed (failure free) life. The failure rate is represented by

$$h(t) = \frac{\beta}{\alpha}\left(\frac{t-\theta}{\alpha}\right)^{\beta-1}, \quad \beta, \alpha > 0, \quad 0 < \theta < t < \infty \tag{3.19}$$

Accordingly, the pdf and reliability function become

$$f(t) = \frac{\beta}{\alpha}\left(\frac{t-\theta}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{t-\theta}{\alpha}\right)^{\beta}\right], \quad t \geq \theta \tag{3.20}$$

and

$$R(t) = \exp\left[-\left(\frac{t-\theta}{\alpha}\right)^{\beta}\right], \quad t \geq \theta \tag{3.21}$$

3.2.3 Gamma Distribution

The gamma distribution was introduced in Section 2.3 as a generalization of the exponential distribution. Recalling the simple shock model considered in Section 3.2.1, one can expand this model for the case when a component fails after being subjected to k successive random shocks arriving according to the homogeneous Poisson process. Time-to-failure distribution of the component in this case follows the gamma distribution.

Examples of its application include the distribution of times between recalibration of an instrument that needs recalibration after k uses; time between maintenance of items that require maintenance after k uses; and time-to-failure of a system with standby components, having the same exponential time-to-failure distribution.

The pdf of the gamma distribution has two parameters, α and β, and it was given in Chapter 2 by (2.53)

$$f(t) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} t^{\alpha-1} \exp\left(-\frac{t}{\beta}\right), \quad \alpha, \beta, t \geq 0$$

The mean value and the variance of the gamma distribution are respectively:

$$\begin{aligned} E(T) &= \alpha\beta \\ \sigma^2(T) &= \alpha\beta^2 \end{aligned} \tag{3.22}$$

The gamma cdf and reliability function, in general, do not have closed forms. In the case when the shape parameter α is an integer, the gamma distribution is known as *the Erlangian* distribution. In this case the reliability and failure rate functions can be expressed in terms of Poisson distribution as:

$$R(t) = \sum_{\varrho=\eta}^{\alpha-\Theta} \frac{\left(\frac{t}{\beta}\right)^{k} \exp\left(-\frac{t}{\beta}\right)}{k!} \tag{3.23}$$

$$R(t) = \sum_{k=0}^{\alpha-1} \frac{\left(\frac{t}{\beta}\right)^k \exp\left(-\frac{t}{\beta}\right)}{k!} \tag{3.23a}$$

$$h(t) = \frac{t^{\alpha-1}}{\beta^{\alpha}\Gamma(\alpha)\sum_{k=0}^{\alpha-1}\frac{(t/\beta)^k}{k!}} \tag{3.24}$$

Accordingly, α shows the number of "shocks" required before a failure occurs, and β represents the mean time to occurrence of a shock.

The gamma distribution is a DFR distribution for $\alpha < 1$, a constant failure rate for $\alpha = 1$, and an IFR distribution for $\alpha > 1$. Thus, the gamma distribution can represent each of three regions of the bathtub curve.

Example 3.3

The mean time to adjustment of an engine in a fighter plane is $M = 100$ hours. (Assume time to adjustment follows the exponential distribution.) Suppose there is a rule to replace certain parts of the engine after three consecutive adjustments.

a. What is the distribution of the time-to-replace?
b. What is the probability that a given engine does not require part replacement for at least 200 hours?
c. What is the mean time-to-replace?

Solution:

a. Use gamma distribution for T with $\alpha = 3$, $\beta = 100$.

b. $$R = \sum_{k=0}^{2} \frac{\left(\frac{t}{100}\right)^k \exp\left(-\frac{t}{100}\right)}{k!}$$

$$= \frac{\left(\frac{200}{100}\right)^0 \exp(-2)}{0!} + \frac{(2)^1 \exp(-2)}{1!} + \frac{(2)^2 \exp(-2)}{2!}$$

$$= 0.135 + 0.271 + 0.271 = 0.677$$

c. Mean time to replace = $E(T) = \alpha\beta = 3 \times 100 = 300$ hours.

3.2.4 Normal Distribution

The normal distribution is a basic distribution of statistics. The popularity of this distribution in reliability engineering can be explained by the, so-called, *Central Limit Theorem*. In engineering terms, according to this theorem, the sum of the large number, *n*, of independent random variables approaches the normal distribution. This distribution is an appropriate model for many practical engineering situations, e.g., it can be used as distribution of diameters of manufactured shafts. Since a normally distributed random variable can take on a value from $(-\infty, \infty)$ range, it has limited applications in reliability-type problems that involve time-to-failure estimations, because "time" cannot take on negative values. However, for cases where the mean μ is positive and is larger than σ by several folds, the probability that the r.v. T takes negative values can be negligible. For those cases where the probability that r.v. T takes negative values is not negligible, the respective truncated normal distribution can be used, see Johnson and Kotz (1970).

The normal pdf was introduced in Chapter 2 by (2. 41) as

$$f(t) = \frac{1}{\sigma_t \sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_t^2} (t - \mu_t)^2 \right], \quad -\infty < t < \infty, \quad -\infty < \mu < \infty$$

where μ is the MTTF and σ is the standard deviation of failure time. The normal distribution failure rate is always a monotonically increasing function of time *t*, so, the normal distribution is an IFR distribution.Thus, the normal distribution can be used as a model representing the wear-out region of the bathtub curve. Normal distribution is also a widely-used model representing stress and/or strength in the framework of the, so-called, *stress-strength* reliability models, which are time independent reliability models (see *Stress-Strength Analysis* in Chapter 6 [Section 6.1]).

3.2.5 Lognormal Distribution

The lognormal distribution is widely used in reliability engineering. The lognormal distribution represents the distribution of a r.v. whose logarithm follows the normal distribution. This model is particularly suitable for failure processes that are the result of many small multiplicative errors. Specific applications of this distribution include time to failure of components due to fatigue cracks (Mann et al., 1974; Provan, 1987). Other applications of the lognormal distribution are associated with failures attributed to maintenance activities. The distribution is also used as a model representing the distribution of particle sizes observed in breakage processes and the life distribution of some electronic components. In

Bayesian reliability analysis the lognormal distribution is a popular model to represent the, so-called, prior distributions. We discuss this topic further in Section 3.6.

The lognormal distribution is a two-parameter distribution. For a r.v. T, the lognormal pdf is

$$f(t) = \frac{1}{\sigma_t t\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma_t^2}(\ln t - \mu_t)^2\right], \qquad (3.25)$$

$$0 \le t < \infty, \quad -\infty < \mu_t < \infty, \quad \sigma > 0$$

where $\mu_t = E(\ln t)$ and $\sigma_t^2 = \text{var}(\ln t)$. The failure rate for the lognormal distribution initially increases over time and then decreases. The rate of increase and decrease depends on the values of the parameters μ_t and σ_t. In general, this distribution is appropriate for representing time to failure for a component whose early failures (or processes resulting in failures) dominate its overall failure behavior.

3.2.6 Extreme Value Distributions

The extreme value distributions are considered in the framework of the Extreme Value Theory. Basic applications of this theory are associated with distributions of extreme loads in structural and maritime engineering (distributions of extreme winds, earthquakes, floods, ocean waves, etc.), other reliability engineering problems, as well as in environmental contamination studies.

Some Basic Concepts and Definitions

Let $x_1, x_2, \ldots, x_n$ be a sample of independently and identically distributed random variables (e.g., representing environmental contamination concentration values) with cdf, $F(X)$. In extreme value theory, the distribution, $F(X)$, is called *parent* distribution. Rearrange the sample in increasing value order so that $x_{1:n} < x_{2:n} < \cdots < x_{n:n}$. The statistics $x_{i:n}$ ($i = 1, 2, \ldots, n$) obtained are called the *order statistics*.

It can be shown that the cdf, $F_{r:n}(x)$, of rth order statistic, $X_{r:n}$, can be expressed in terms of the binomial distribution as

$$F_{r:n}(X) = \sum_{k=r}^{n} \binom{n}{k} F^k(X)[1 - F(X)]^{n-k} \qquad (3.26)$$

Clearly, the maximum statistic of a sample of size n is the last order statistic ($r = n$), so its cdf can be written, using (3.26) as

$$F_{n:n}(X) = F^n(X) \qquad (3.27)$$

The distribution obtained is called the *distribution of maxima*. Getting ahead of our consideration, we will mention that this is (if X is the time-to-failure) the time-to-failure distribution of a parallel system (considered in Chapter 4) composed of n identical components.

The *distribution of minima* of a random sample of size n can be obtained from (3.26), as a particular case, when $r = 1$, i.e.,

$$F_{1:n}(X) = 1 - [1 - F(X)]^n \tag{3.28}$$

Similar to the case considered above this is the time-to-failure distribution of a series system (considered in Chapter 4) composed of n identical components.

Example 3.4 (Castillo (1988))

One of the main engineering concerns in the design of a nuclear power plant is the estimation of the probability distribution of the distances of possible earthquakes to the tentative location of the plant. Due to the presence of a fault in the area surrounding this location, it has been established that the epicenter of an earthquake can occur, equally likely, at any point within the 50 km radius of the fault. If the plant location is aligned with the fault and its closest extreme is 200 km away, then the distance between the epicenter and the plant can be assumed to be distributed uniformly between 200 and 250 km. Find the probability of having minimum distances less than 210 km.

Solution:

The cdf of the distance to the closest earthquake in series of 5 and 10 earthquakes is given by

$$F_{1:n}(x) = 1 - \left(1 - \frac{x - 200}{50}\right)^n; \quad n = 5, 10$$

The probability of having minimum distances less than 210 Km which is equivalent to the standard uniform distribution U(0, 1) value of (210 - 200)/50, are 0.672 and 0.893 for the series of 5 and 10 earthquakes, respectively.

Order Statistics from Samples of Random Size

Previous considerations were associated with a fixed sample size. There are many practical situations where the sample size is random and one is interested in the extreme statistics. For example, the number of defects having random sizes in material can be itself random.

Denote the distribution of sample size, n, by

$$p_i = \Pr\left[n = n_i\right]$$

Let the statistic of interest, x, have, for a fixed sample size n, pdf $f(x,n)$ and cdf $F(x,n)$. Using the total probability rule, the pdf, $g(x)$, and cdf, $G(x)$, of the statistic x can be written as:

$$\begin{aligned} g(x) &= \sum_i p_i f\left(x, n_i\right) \\ G(x) &= \sum_i p_i F\left(x, n_i\right) \end{aligned} \tag{3.29}$$

An example of this model application to the problem of nuclear power plant core damage frequency estimation is considered in Chapter 8.

When the sample size n goes to infinity, the distributions of maxima and minima have the following limits

$$\lim_{n \to \infty} F^n(x) = 0, \quad F(x) < 1 \tag{3.30}$$

$$\lim_{n \to \infty} \left\{1 - [1 - F(x)]^n\right\} = 1, \quad F(x) \leq 1 \tag{3.31}$$

These limits show that the limit distributions take on only values 0 and 1. It is said that such distribution *degenerates*. To avoid this degeneration the following linear transformation is used:

$$Y = a_n + b_n x \tag{3.32}$$

where a_n and b_n are the constants chosen to get the following *limit distributions:*

for maxima:

$$\lim_{n \to \infty} H_n(a_n + b_n x) = \lim_{n \to \infty} F^n(a_n + b_n x) = H(x) \quad \text{for all } x \tag{3.33}$$

for minima:

$$\lim_{n \to \infty} L_n(c_n + c_n x) = \lim_{n \to \infty} 1 - \left[1 - F\left(c_n + c_n x\right)\right]^n = L(x) \quad \text{for all } x \tag{3.34}$$

which do not degenerate.

Asymptotic Distributions of Maxima and Minima

We discussed how to get the distribution of maxima and minima from a given parent distribution in the case of finite samples of fixed and random sample sizes. The *asymptotic* distributions of extreme values are used in the following situations (Castillo (1988)):

1. The sample size increases to infinity.
2. The parent distribution is unknown.
3. The sample size is large but unknown.

Three Types of Limit Distributions

The fundamental result of Extreme Value Theory consists in existence of only three feasible types of limit distributions for maxima, H, and three similar feasible types of limit distribution for minima, L, (Frechet, 1927; Fisher and Tippet, 1928; Gnedenko, 1941). This result is given by the following theorems.

There are only three types of nondegenerated distributions for maxima, $H(x)$, satisfying the condition (3.33) (the cdf's are given in the standard forms, i.e, with scale parameter equal to 1):

Type I (the Gumbel distribution):

$$H_1(x) = \exp\left[-\exp\left(-\frac{x}{\delta}\right)\right], \quad \delta > 0, \quad -\infty < x < \infty \tag{3.35}$$

Type II (the Freshet distribution):

$$H_2(x) = \begin{cases} \exp\left[-\left(\frac{x}{\delta}\right)^{-\gamma}\right] & \delta > 0, \ \gamma > 0, \ x > 0 \\ 0 & x \le 0 \end{cases} \tag{3.36}$$

Type III (the Weibull distribution):

$$H_3(x) = \begin{cases} 1 & x > 0, \\ \exp\left[-\left(-\frac{x}{\delta}\right)^{\gamma}\right] & \delta > 0, \ \gamma > 0, \ x \le 0 \end{cases} \tag{3.37}$$

The similar theorem for distributions of minima states that there are only three types of nondegenerated distributions for minima, $L(x)$, satisfying the condition (3.34). These are:

Type I (the Gumbel distribution):

$$L_1(x) = 1 - \exp\left[-\exp\left(\frac{x}{\delta}\right)\right], \quad \delta > 0, \quad -\infty < x < \infty \tag{3.38}$$

Type II (the Freshet distribution):

$$L_2(x) = \begin{cases} 1 - \exp\left[-\left(-\frac{x}{\delta}\right)^{\gamma}\right], & \delta > 0, \quad \gamma > 0, \quad x < 0 \\ 1 & x > 0 \end{cases} \tag{3.39}$$

Type III (the Weibull distribution):

$$L_3(x) = \begin{cases} 0 & x < 0 \\ 1 - \exp\left(-\frac{x}{\delta}\right)^{\gamma}, & \gamma > 0, \quad x > 0 \end{cases} \tag{3.40}$$

Currently, the following three extreme value distributions are widely used in reliability engineering: the Weibull distribution for minima (discussed in Section 3.2.2), the Gumbel distribution for minima and the Gumbel distribution for maxima. The last two distributions are also referred to as *the smallest extreme value distribution* and *the largest extreme value distribution* (Nelson (1982)).

Similar to the three parameter Weibull distribution (3.20), the smallest extreme value distribution and the largest extreme value distribution are sometimes used as two parameter distributions. In this case their pdf's take on the following forms.

The pdf of the smallest extreme value distribution is given by

$$f(t) = \frac{1}{\delta}\exp\left[\frac{1}{\delta}(t - \lambda) - \exp\left(\frac{t - \lambda}{\delta}\right)\right],$$
$$-\infty < \lambda < \infty, \quad \delta > 0, \quad -\infty < t < \infty \tag{3.41}$$

The parameter λ is called the *location* parameter and can take on any value. The parameter δ is called the *scale* parameter and is always positive. The failure rate for the smallest extreme value distribution is

$$h(t) = \frac{1}{\delta}\exp\left(\frac{t-\lambda}{\delta}\right) \tag{3.42}$$

which is an increasing function of time, so that the smallest extreme value distribution is IFR distribution, which can be used as a model for component failures due to aging. In this model, the component's wear-out period is characterized by an exponentially increasing failure rate. Clearly, negative values of t are not meaningful when it is representing time to failure.

The Weibull distribution and the smallest extreme value distribution are closely related to each other. It can be shown that if a r.v. X follows the Weibull distribution with pdf (3.17), the transformed r.v. $T = \ln(X)$ follows the smallest extreme value distribution with parameters

$$\begin{aligned} \lambda &= \ln(\alpha) \\ \delta &= \frac{1}{\beta} \end{aligned} \tag{3.43}$$

The two parameter largest extreme value pdf is given by

$$f(t) = \frac{1}{\delta}\exp\left[-\frac{1}{\delta}(t-\lambda) - \exp\left(-\frac{t-\lambda}{\delta}\right)\right], \quad -\infty < t < \infty, \quad \delta > 0, \quad -\infty < \lambda < \infty \tag{3.44}$$

The largest extreme value distribution, though not very useful for component failure behavior modeling, is useful for estimating natural extreme phenomena.

For further discussions regarding the extreme value distributions, see Castillo (1988), Gumbel (1958), and Johnson and Kotz (1970).

Example 3.5

The maximum demand for electric power at any given time during a year is directly related to extreme weather conditions. An electric utility has determined that the distribution of maximum power demands can be presented by the largest

extreme value distribution with $\lambda = 1200$ (MW) and $\delta = 480$ (MW). Determine the probability (per year) that the demand will exceed the utility's maximum installed power of 3000 (MW).

Solution:

Since this is the largest extreme value distribution, we should integrate (3.26) from 3000 to ∞.

$$\Pr(t > 3000) = \int_{3000}^{\infty} f(t)\,dt = 1 - \exp\left[-\exp\left(\frac{-(t-\lambda)}{\delta}\right)\right]$$

Since

$$\frac{t-\lambda}{\delta} = \frac{3000 - 1200}{480} = 3.75$$

then

$$\Pr(t > 3000) = 0.023$$

3.3 COMPONENT RELIABILITY MODEL

In the previous section, we discussed several distribution models useful for reliability analysis of components. A *probability model* is referred to a mathematical expression that describes in terms of probabilities that a r.v. is spread over its range. It is necessary at this point to discuss how field and test data can support the selection of a probability model for reliability analysis. In this section, we consider several procedures for selecting and estimation the models using observed failure data. These methods can be divided into two groups: nonparametric methods (which do not need a particular distribution function) and parametric methods (which are based on a selected distribution function). We discuss each of these methods in more detail. Besides, some graphic exploratory data analysis procedures are considered in Section 7.3.

3.3.1 Graphical Nonparametric Procedures

The nonparametric approach, in principle, attempts to directly estimate the reliability characteristic of an item (e.g., the pdf, reliability, and hazard rates) from

a sample. The shape of these functions, however, is often used as an indication of the most appropriate parametric distribution representation. Thus, such procedures can be considered as tools for exploratory (preliminary) data analysis. It is important to mention that failure data from a maintained item can be used as the sample only if after the maintenance the item can be assumed to be *as good as new*. Then each failure time can be considered as a sample observation independent of the previously observed failure times. Therefore, n observed times to failure of such a maintained component is equivalent to putting n independent new components under test.

Small Samples

Suppose n times to failure make a small sample (e.g., $n < 25$). Let the data be ordered such that $t_1 \le t_2 \le \cdots \le t_n$. Blom (1958) introduced the following nonparametric estimators for the reliability functions of interest:

$$\hat{h}(t_i) = \frac{1}{(n - i - 0.625)(t_{i+1} - t_i)}, \quad i = 1, 2, \ldots, n-1 \tag{3.45}$$

$$\hat{R}(t_i) = \frac{n - i + 0.625}{(n + 0.25)}, \quad i = 1, 2, \ldots, n \tag{3.46}$$

and

$$\hat{f}(t_i) = \frac{1}{(n + 0.25)(t_{i+1} - t_i)}, \quad i = 1, 2, \ldots, n-1 \tag{3.47}$$

Although there are other estimators besides those above, Kimbal (1960) concludes that estimators (3.45)–(3.47) have good properties and recommends their use. One should keep in mind that 0.625 and 0.25 are correction terms of a minor importance, which assumes a small bias and a small mean square error for the Weibull distribution estimation (Kapur and Lamberson (1977)).

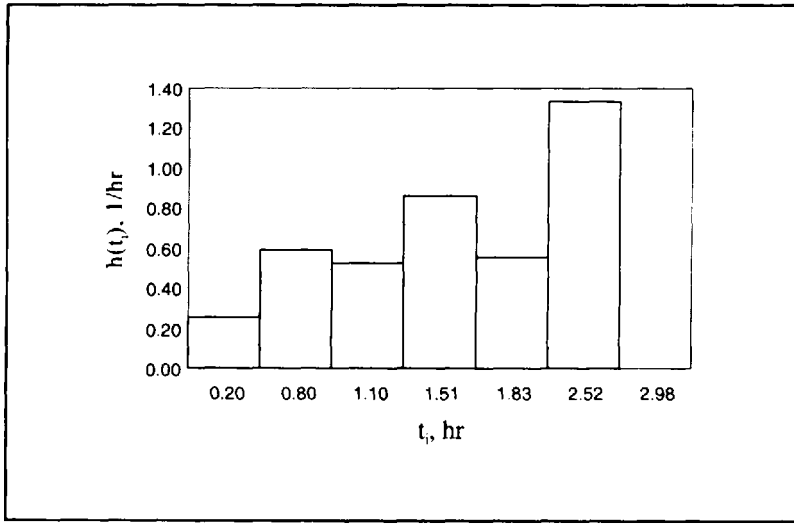
Example 3.6

A high-pressure pump in a process plant has the failure times t_i (shown in the following table). Plot $\hat{h}(t)$, $\hat{R}(t)$, $\hat{f}(t)$ and discuss the results.

Solution:

i	t_i	$t_{i+1} - t_i$	$\hat{h}(t)$	$\hat{R}(t)$	$\hat{f}(t)$
1	0.20	0.60	0.25	0.91	0.23
2	0.80	0.30	0.59	0.78	0.46
3	1.10	0.41	0.53	0.64	0.34
4	1.51	0.32	0.86	0.50	0.43
5	1.83	0.69	0.55	0.36	0.20
6	2.52	0.42	1.34	0.22	0.30
7	2.98			0.09	

From the histogram below, one can conclude that the failure rate is reasonably constant over the operating period of the component, with an increase toward the end. A point of caution: although a constant hazard rate might be concluded, several other tests and additional observations may be needed to support the conclusion. Additionally, the histogram is only a representative of the case under study. An extension of the result to future times or other cases (e.g., other high-pressure pumps) may not be accurate.



Large Samples

Suppose n times to failure makes a large sample. Suppose further that the sample is grouped into a number of equal time-to-failure increments, Δt. According to the definition of reliability, a nonparametric estimate of the reliability function is

$$\hat{R}(t_i) = \frac{N_s(t_i)}{N} \tag{3.48}$$

where $N_s(t_i)$ represents the number of surviving components. Note that estimator (3.48) is absolutely compatible with the empirical (sample) cdf (2.93) introduced in Section (2.5). Time t_i is usually taken to be the upper endpoint of each interval. Similarly, the pdf is estimated by

$$\hat{f}(t_i) = \frac{N_f(t_i)}{N\Delta t} \tag{3.49}$$

where $N_f(t_i)$ is the number of failures observed in the interval $(t_i, t_i + \Delta t)$. Finally, using (3.48) and (3.49), one gets

$$\hat{h}(t_i) = \frac{N_f(t_i)}{N_s(t_i)\Delta t} \tag{3.50}$$

It is clear that for $i = 1$, $N_s(t_i) = N$; and for $i > 1$, $N_s(t_i) = N_s(t_{i-1}) - N_f(t_i)$. Equation (3.50) gives an estimate of average failure rate for the interval $(t_i, t_i + \Delta t)$. When $N_s(t_i) \to \infty$ and $\Delta t \to 0$, estimate (3.50) approaches the true hazard rate $h(t)$.

In (3.50), $N_f(t_i)/N_s(t_i)$ is the estimate of probability that the component will fail in the interval $(t_i, t_i + \Delta t)$, since $N_s(t_i)$ represents the number of components functioning at t_i. Dividing this quantity by Δt, the estimate of failure rate (probability of failure per unit of time for interval Δt) is obtained. It should be noted that the accuracy of this estimate depends on Δt. Therefore, if smaller Δts are used, we would, theoretically, expect to obtain a better estimation. However, a drawback of using smaller Δts is the decrease in the amount of data for each interval to estimate $\hat{R}(t)$ and $\hat{f}(t)$. Therefore, selecting Δt requires consideration of both of these opposing factors.

Example 3.7

Times to failure (in h) for an electrical device are obtained during three stages of the component's life. The first stage is believed to be associated with an

infant mortality of the component; the second stage represents chance failures; and the third stage represents the wear-out period. Plot the failure rate for this component, using the data provided below.

Solution:

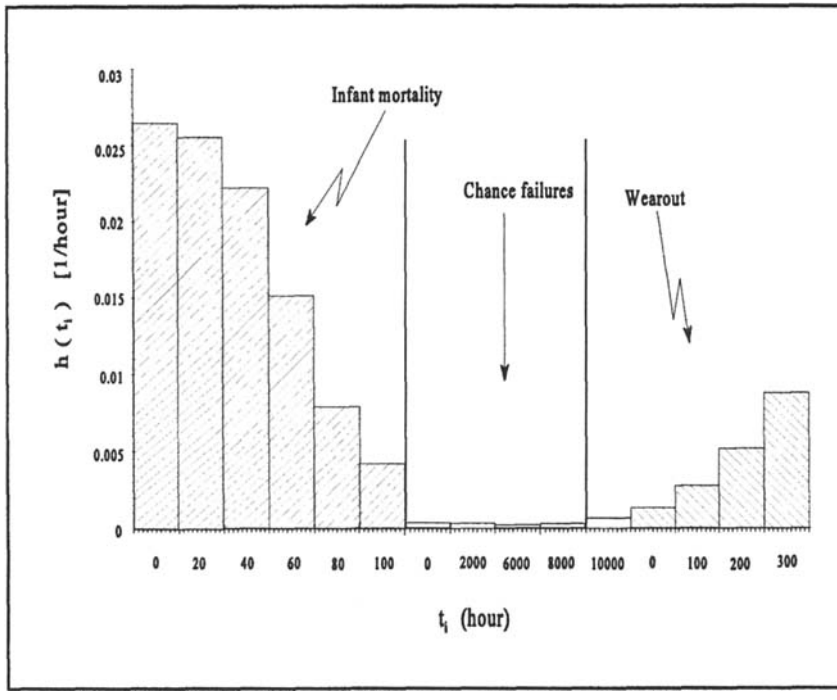
Use Equations (3.48)–(3.50) to calculate the empirical hazard rate, reliability and pdf.

Given Data			Calculated Data			
Interval, t_i		Frequency				
Beginning	End	$N_f(t_i)$	$N_s(t_i)$	$h(t_i)$	$R(t_i)$	$f(t_i)$
Infant Mortality Stage						
0	20	79	71	0.02633	1.00000	0.02633
20	40	37	34	0.02606	0.47333	0.01233
40	60	15	19	0.02206	0.22667	0.00500
60	80	6	13	0.01579	0.12667	0.00200
80	100	2	11	0.00769	0.08667	0.00067
100	120	1	10	0.00455	0.07333	0.00033
120	> 120	10	0	0.05000	0.06667	0.00333
Total 150						

Given Data			Calculated Data			
Interval, t_i		Frequency				
Beginning	End	$N_f(t_i)$	$N_s(t_i)$	$h(t_i)$	$R(t_i)$	$f(t_i)$
Chance Failure Stage						
0	2000	211	289	0.00021	1.00000	0.00021
2000	4000	142	147	0.00025	0.57800	0.00014
4000	6000	67	80	0.00023	0.29400	0.00007
6000	8000	28	52	0.00018	0.16000	0.00003
8000	10000	21	31	0.00020	0.10400	0.00002
10000	> 10000	31	0	0.00050	0.06200	0.00003
Total 500						

Given Data			Calculated Data			
Interval, t_i		Frequency				
Beginning	End	$N_f(t_i)$	$N_s(t_i)$	$h(t_i)$	$R(t_i)$	$f(t_i)$
Wearout Stage						
0	100	34	266	0.00113	1.00000	0.00113
100	200	74	192	0.00278	0.88667	0.00247
200	300	110	82	0.00573	0.64000	0.00367
300	> 300	82	0	0.01000	0.27333	0.00273
		Total300				

The graph below plots the estimated hazard rate functions for the three observation periods (please note that the three periods are combined on the same x-axis).



 See the software supplement for the automated graphical nonparametric estimation for small and large samples.

3.3.2 Probability Plotting

Probability plotting is a simple graphical method of displaying and analyzing observed data. The data are plotted on special probability papers in a way that a transformed cdf would be a straight line. Each type of distribution has its own probability paper. If a set of data is hypothesized to originate from a known distribution, the graph can be used to conclude whether or not the hypothesis might be rejected. From the plotted line, one can also roughly estimate the parameters of the hypothesized distribution. Probability plotting is often used in reliability analysis to test the appropriateness of using known distributions to present a set of observed data. This method is used because it provides simple and visual representation of the data. This approach is an informal, qualitative decision making method. However, the goodness-of-fit tests discussed in Chapter 2 is a formal quantitative method. Therefore, the plotting method should be used with care and preferably as an exploratory data analysis procedure. In the following we will discuss some factors that should be considered when the probability plotting is used.

We are going to briefly discuss the probability papers for the basic distributions considered in this book. The reader is referred to Nelson (1979), Nelson (1982), Martz and Waller (1982) and Kececioglu (1991) for further discussion regarding other distributions and various plotting techniques.

Exponential Distribution Probability Plotting

Taking the logarithm of the expression for the reliability function of the exponential distribution (3.15) one gets

$$\ln R(t) = -\lambda t \tag{3.51}$$

If $R(t)$ is plotted as a function of time, t, on semilogarithmic plotting paper, according to (3.51) the resulting plot will be a straight line with the slope of $(-\lambda)$.

Consider the following n times to failure observed from a life test: $t_1 \le t_2 \le \cdots \le t_n$. According to (3.51), an estimate of the reliability $R(t_i)$ can be made for each t_i. A crude nonparametric estimate of $R(t_i)$ is clearly $1 - i/n$ (recall (3.48) and (2.93)). However, as it was noted in Section 3.3.1, statistic (3.46) provides better estimation for $R(t)$ for the Weibull distribution (recall that the exponential distribution is a particular case of the Weibull one).

Graphically, the y-axis shows $R(t_i)$ and the x-axis shows t_i. The resulting points should reasonably fall on a straight line if these data can be described by the exponential distribution. Since the slope of $\ln R(t)$ vs. t is negative, it is also possible to plot $\ln(1/R(t))$ vs. t in which the slope is positive. In practice one may use the so-called Kimbal estimater for R(t) and plot $(n - i + 0.625)/(n + 0.25)$ against t_i using semi-log paper. Other, appropriate estimators of $R(t_i)$ include the, so-called, mean rank, $(n - i + 1)/(n+1)$ and the median rank, $(n - i + 0.7)/(n + 0.4)$ (Kapur and Lamberson, 1977).

It is also possible to estimate the MTTF from the plotted graph. For this purpose, at the level of $R = 0.368$ (or $1/R = e \approx 2.718$), a line parallel to the x-axis is drawn. At the intersection of this line and the fitted line, another line vertical to the x-axis is drawn. The value of t read on the x-axis is an estimate of MTTF, and its inverse is $\hat{\lambda}$. Exponential plot is a particular case of the Weibull distribution, and therefore, the Weibull paper may be used to determine whether or not exponential distribution is a good fit.

Example 3.8

Nine times to failure of a diesel generator are recorded as 31.3, 45.9, 78.3, 22.1, 2.3, 4.8, 8.1, 11.3, and 17.3 days. If the diesel is restored to "as good as new" after each failure, determine whether the data represent the exponential distribution. Find $\hat{\lambda}$ and $\hat{R}$(193 hours).

Solution:

First arrange the data in increasing order and then calculate the corresponding $\hat{R}(t_i)$.

i	t	$\frac{n - i + 0.625}{n + 0.25}$	$\frac{n + 0.25}{n - i + 0.625}$
1	2.3	0.93	1.07
2	4.8	0.82	1.21
3	8.1	0.72	1.40
4	11.3	0.61	1.64
5	17.3	0.50	2.00
6	22.1	0.39	2.55
7	31.3	0.28	3.53
8	45.9	0.18	5.69
9	78.3	0.07	14.80

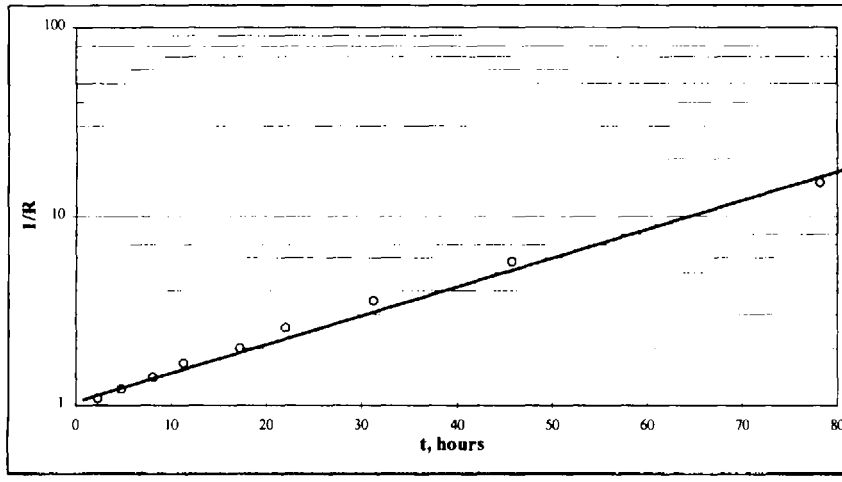


Figure 3.5 Exponential probability plot in Example 3.8.

Figure 3.5 shows a plot of the above data on logarithmic paper.

$$\hat{\lambda} = \frac{\ln 7 - \ln 3}{48.6 - 28} = 0.041 \text{ failures/day} = 1.71E-3 \text{ failures/hour}$$

$$\hat{R}(193) = \exp[-(1.71E-3)(193)] = 0.72$$

Weibull Distribution Probability Plotting

Similar to plots of the exponential distribution, plots of the Weibull distribution require special probability papers. If the observed data form a reasonably straight line, the Weibull distribution can be considered as a competing model. Recalling the expression for the Weibull cdf from Table 3.1 or from Example 2.17, one can get the following relationships for the respective reliability function

$$\frac{1}{R(t)} = \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$
$$\ln\left[\ln\left(\frac{1}{R(t)}\right)\right] = \beta \ln t - \beta \ln \alpha \tag{3.52}$$

This linear (in ln$t$) relationship provides the basis for the Weibull plots. It is evident that ln(ln[1/$R(t)$]) plots as a straight line against ln$t$ with slope β and y-intercept of ($-\beta \ln \alpha$). Accordingly, the values of the Weibull parameters α and β can be obtained from the y-intercept and the slope of the graph, respectively.

As mentioned before, several estimators of $R(t)$ can be used. The most recommended estimator is (3.47) (identical to that used in exponential plots). The corresponding plotting procedure is simple. On the special Weibull paper (see Figure 3.6), t_i is plotted in the logarithmic x-axis and the estimate of $F(t) = [(i - 0.375)/(n + 0.25)] \times 100$ is plotted on the y-axis (often labeled % failure). The third scale shown in Figure 3.6 is for ln ln ($1/R(t)$), but it is more convenient to use the estimate of $F(t)$.

The degree to which the plotted data fall on a straight line determines the conformance of the data to the Weibull distribution. If the data give reasonably good plot as a straight line, the Weibull distribution is a reasonable fit, and the shape parameter β and the scale parameter α can be roughly estimated. If a line is drawn parallel to the plotted straight line from the center of a small circle⊕ (sometimes called ORIGIN) until it crosses the SMALL BETA ESTIMATOR axis, the value of β can be obtained. To find α, draw a horizontal line from the 63.2% cdf level until it intersects the fitted straight line. From this point, draw a vertical line until it intersects with the x-axis, and read the value of parameter α at this intersection.

Example 3.9

Time to failure of a device is assumed to follow the Weibull distribution. Ten of these devices are put on reliability test. The times to failure (in hours) are: 89, 132, 202, 263, 321, 362, 421, 473, 575, and 663. If the Weibull distribution is a correct choice to model the data, what are the parameters of this distribution? What is the reliability of the device at 1000 hours?

Solution:

i	1	2	3	4	5	6	7	8	9	10
t_i	89	132	202	263	321	362	421	473	575	663
$\frac{i - 0.375}{n + 0.25} \times 100$	6.10	15.85	25.61	35.37	45.12	54.88	64.46	74.39	84.15	93.90

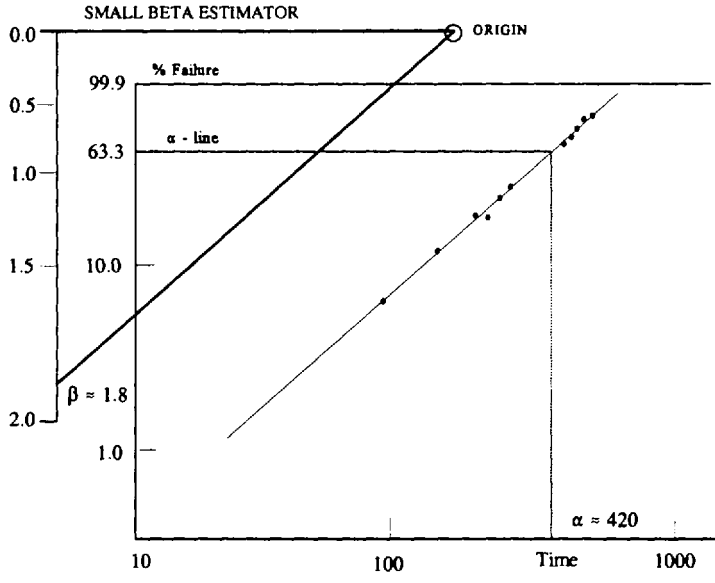


Figure 3.6 Weibull probability plot.

Figure 3.6 shows the fitted line on the Weibull probability paper. Clearly the fitting is reasonably good. The graphical estimate of β is approximately 1.8, and the estimate of α is approximately 420 hours. Percent failure at 1000 hours is about 99.1%; the reliability is about 0.9% [$R(t = 1000) = 0.009$].

In cases where the data do not fall on a straight line but are concave or convex in shape, it is possible to find a *location parameter* θ (i.e., to try using the three-parameter Weibull distribution (3.20) introduced in Section 3.2.2) that might "straighten out" these points. For this procedure, see Kececioglu (1991) and Nelson (1979).

If the failure data are grouped, the class midpoints t'_i (rather than t_i) should be used for plotting, where $t'_i = (t_{i-1} + t_i)/2$. One can also use class endpoints

instead of midpoints. Recent studies suggest that the Weibull parameters obtained by using class endpoints in the plots are better matched with those of the maximum likelihood estimation method.

Normal and Lognormal Distribution Probability Plotting

The same special probability papers can be used for both normal and lognormal plots. On the x-axis, t_i (for normal) and $\ln(t_i)$ (for lognormal) are plotted, while on the y-axis, the $F(t_i)$ estimated by the same $(i - 0.375)/(n + 0.25)$ is plotted (or other plotting points). It is easy to show that normal cdf can be linearized using the following transformation

$$\Phi^{-1}\left[F(t_i)\right] = \frac{1}{\sigma} t_i - \frac{\mu}{\sigma} \tag{3.53}$$

where $\Phi^{-1}(.)$ is the inverse of the standard normal cdf. Some lognormal papers are logarithmic on the x-axis, in which case t_i can be directly expressed. In the case of lognormal distribution, t_i in (3.53) is replaced by $\ln(t_i)$. If the plotted data fall on a straight line, a normal or lognormal distribution might be conformed. To estimate the mean parameter μ, the value of 50% is marked on the x-axis and a line parallel to the y-axis is drawn until intersection with the plotted straight line. From the intersection, a horizontal line to the x-axis is drawn. Its intersection with the y-axis gives the estimate of parameter μ (mean or median of the normal distribution and the median of the lognormal distribution). Similarly, if the corresponding y-axis intersection for the 84% value is selected from the x-axis, the parameter σ can be estimated for the normal distribution as $\sigma \approx t_{84\%} - t_{50\%}$ or $\sigma \approx t_{84\%} - \mu$. For the lognormal distribution, $\sigma \approx \ln t_{84\%} - \mu_t$.

Example 3.10

The time it takes for a thermocouple to drift upward or downward to an unacceptable level is measured and recorded in a process plant. (See the following table.) Determine whether the drifting time can be modeled by a normal distribution.

Solution:

Figure 3.7 shows that the data conform the normal distribution, with $\mu = t_{50\%} = 17.25$ months and $t_{84\%} = 20.75$ months. Therefore the estimate of $\sigma \approx 20.75 - 17.25 = 3.5$ months.

i	t_i (months)	$\frac{i - 0.375}{n + 0.25} \times 100$
1	11.2	4.39
2	12.8	11.40
3	14.4	18.42
4	15.1	25.44
5	16.2	32.46
6	16.3	39.47
7	17.0	46.49
8	17.2	53.50
9	18.1	60.53
10	18.9	67.54
11	19.3	74.56
12	20.0	81.58
13	21.8	88.60
14	22.7	95.61

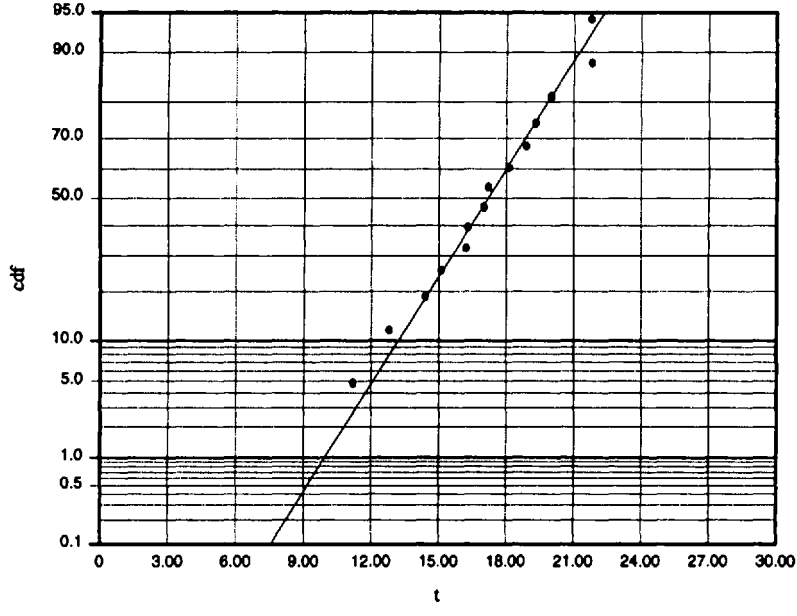


Figure 3.7 Normal distribution plot in Example 3.10.

Example 3.11

Five components are placed on fatigue crack tests. The failure times are given in the table below. Determine the conformance of the data with a lognormal distribution. Estimate the parameters of the lognormal distribution.

Solution:

i	t_i (months)	$\frac{i - 0.375}{n + 0.25} \times 100$
1	363	11.90
2	1115	30.95
3	1982	50.00
4	4241	69.05
5	9738	88.10

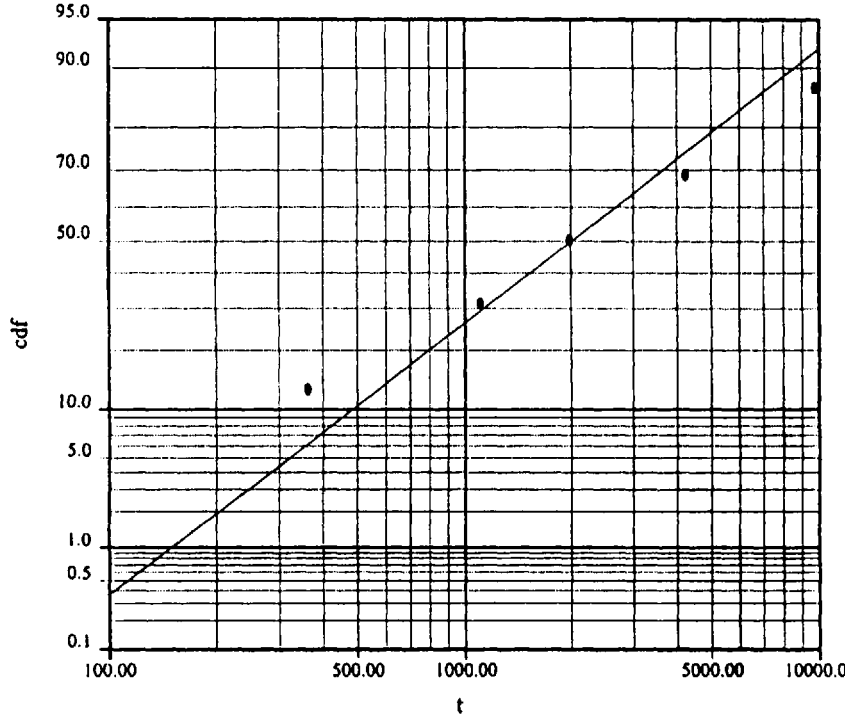


Figure 3.8 Lognormal distribution plot for Example 3.11.

From the probability plot in Figure 3.8,

$$\hat{\mu}_t = \ln(\hat{\mu}_s) = \ln(2000) = 7.61$$
$$\hat{\sigma}_t \approx \ln(t_{84\%}) - \hat{\mu}_t = \ln(6700) - 7.61 \approx 1.12$$

It should be noted that with the present level of the computer support in reliability data analysis, the term "probability paper" is only used to refer to the graphical method of parameter estimation, while the paper itself has become obsolete. It takes a simple electronic spreadsheet to program the above equations of cdf linearization and estimate the distribution parameters by the least square method. A modern reliability engineer does no longer have to use a ruler and an eyeball judgement to analyze reliability data.

 See the software supplement for the "electronic" probability paper for the exponential, Weibull, normal, and lognormal distributions.

3.3.3 Total-Time-on-Test Plots

The total-time-on-test plot is a graphical procedure helping to determine whether the underlying distribution exhibits an increasing failure rate, a constant failure rate (the distribution is exponential), or a decreasing failure rate. This procedure is discussed in detail by Barlow and Campo (1975) and Barlow (1978). Additionally, Davis (1952) discussed the use of this method for optimal replacement policy problems.Although it is possible to treat grouped failure data using this method, we discuss its use for ungrouped failure data only. Consider the observed failure times of n components such that $t_1 \le t_2 \le \cdots \le t_n$. If the number of survivors to time t_i is denoted by $N_s(t_i)$ then the survival probability (the reliability function) can be estimated as

$$R(t_i) = \frac{N_s(t_i)}{n}$$

It is clear that $N_s(t_i)$ is a step-wise function. The *total time on test* to age t_i denoted by $T(t_i)$ is obtained from

$$T(t_i) = \int_0^{t_i} N_s(t)\,dt \tag{3.54}$$

Equation (3.54) can be expressed in a more tractable form:

$$\int_0^{t_i} N_s(t)\,dt = nt_1 + (n-1)(t_2 - t_1) + \cdots + (n-i+1)(t_i - t_{i-1}) \quad (3.55)$$

since n components have survived up to time t_1 (time to the first failure), $(n-1)$ of the components survived during the period between t_1 and t_2, and so on. The, so-called, *scaled total time on test* at time t_i is defined as

$$\tilde{T}(t_i) = \frac{\int_0^{t_i} N_s(t)\,dt}{\int_0^{t_n} N_s(t)\,dt} \quad (3.56)$$

It can be shown that for the exponential distribution, $\tilde{T}(t_i) = i/n$. A graphical representation of total time on test is formed by plotting i/n on the x-axis and $\tilde{T}(t_i)$ on the y-axis. Its deviation from the reference line $\tilde{T}(t_i) = i/n$ is then assessed. If the data fall on the $\tilde{T}(t_i) = i/n$ line, the exponential distribution can be assumed. If the plot is concave over most of the graph, there is a possibility of an increasing failure rate. If the plot is convex over most of the graph, then it indicates a possibility of a decreasing failure rate. If the plot does reasonably fall on a straight line and does not reveal concavity or convexity, one can assume the exponential distribution.

The total-time-on-test plot, similar to other graphical procedures, should not be used as the only test for determining model adequacy. This is of particular importance when only a small sample is available. The total-time-on-test plots are simple to carry out and provide a good alternative to more elaborate hazard and probability plots. These plots are scale invariant and, unlike probability plots, no special plotting papers are needed.

Example 3.12

In a nuclear power plant, the times to failure (in hours) of the feedwater pumps are recorded and given in the table below. Use the total-time-on-test plot to draw a conclusion about the time dependence of the failure rate function.

Solution:

Using Equations (3.54)–(3.56) the following results are received.

i	t_i	$(n - i + 1) \cdot (t_i - t_{i-1})$	$T(t_i)$	i/n	$\tilde{T}(t_i)$
1	1400	14,000	14,000	0.1	0.14
2	3500	18,900	32,900	0.2	0.33
3	5900	19,200	52,100	0.3	0.53
4	7600	11,900	64,000	0.4	0.65
5	8600	6000	70,000	0.5	0.71
6	9000	2000	72,000	0.6	0.73
7	11,600	10,400	82,400	0.7	0.83
8	12,000	1200	83,600	0.8	0.84
9	19,100	14,200	97,800	0.9	0.99
10	20,400	1300	99,100	1.0	1.00

Figure 3.9 indicates a mild tendency toward an increasing failure rate over the observation period.

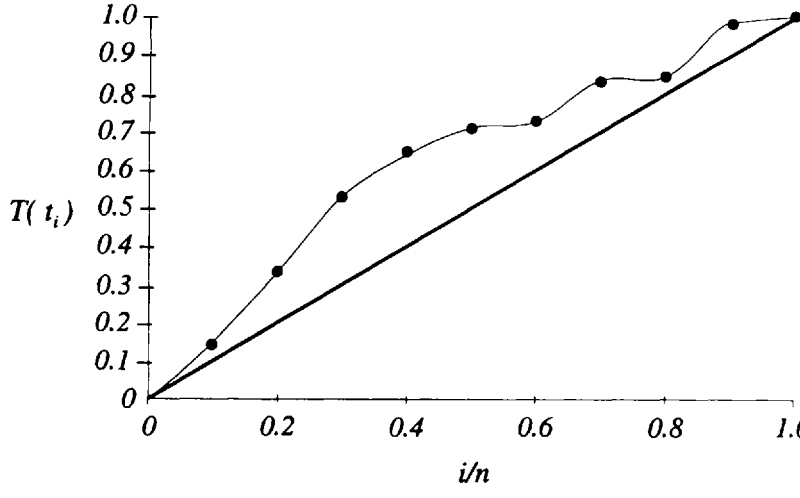


Figure 3.9 Total-time-on-test plot for Example 3.12.

 See the software supplement for the automated total time on test plot.

3.4 CLASSICAL PARAMETRIC ESTIMATION

This section deals with statistical methods for estimating reliability model parameters, such as λ of the exponential distribution, μ and σ of the normal and lognormal distribution, p of the binomial distribution, and α and β of the Weibull distribution. The objective is to find a point estimate and a confidence interval for the parameters of interest. We briefly discussed this topic in Chapter 2 for estimating parameters and confidence intervals associated with a normal distribution. In this section we expand the discussion to include estimation of parameters of other distributions useful to reliability analysis.

It is important to realize why we need to consider confidence intervals in the estimation process. In essence, this need stems from the fact that we have a limited amount of information (e.g., on times-to-failure), and thus we cannot state our estimation with certainty. Therefore, the confidence interval is highly influenced by the amount of data available. Of course other factors, such as diversity and accuracy of the data sources and adequacy of the selected model can also influence the state of our uncertainty regarding the estimated parameters. When discussing the goodness-of-fit tests, we deal with uncertainty due to the adequacy of the model by using the concept of levels of significance. However, uncertainty due to diversity and accuracy of the data sources is a much more difficult issue to deal with.

The methods of parameter estimation discussed in this section are more formal and accurate methods of determining distribution parameters than the methods described previously (such as the plotting methods).

Estimation of parameters of time-to-failure or failure on demand distribution can be based on field data as well as on data obtained from special life (reliability) tests. In life testing, a sample of components from a hypothesized population of such components is placed on test using the environment in which the components are expected to function, and their times to failure are recorded. In general, two major types of tests are performed. The first is testing *with replacement* of the failed items, and the second is testing *without replacement.* The test with replacement is sometimes called *monitored.*

Samples of times-to-failure or times-between-failures (later, the term *time-to-failure* will be used wherever it does not result in a loss of generality) are seldom *complete* samples. A complete sample is the one in which all items have failed during a test for a given observation period, and all the failure times are known (distinct). The likelihood function for a complete sample was introduced in Section 2.5. In the following sections, the likelihood functions for some types of censoring are discussed. Modern products are usually so reliable that a complete

sample is a rarity, even in accelerated life testing. Thus, as a rule, reliability data are incomplete.

Left and Right Censoring

Let N be the number of items in a sample, and assume that all units of the sample are tested simultaneously. If, during the test period, T, only r units have failed, the failure times being known, and the failed items are not replaced, the sample is called *singly censored on the right at T.* In this case, the only information we have about N-r unfailed units is that their failure times are greater than the duration of the test, T. Formally, an observation is called "right censored at T," if the exact value of observation is not known, but is known that it is greater than or equal to T (Lawless (1982)).

If a distinct failure time for an item is not known, but it is known that it is less than a given value, the failure time is called *left censored.* This type of censoring practically never appears in reliability data collection and so it is not discussed. If the only information available is that an item failed in an interval (for example, between successive inspections), the respective data are called *grouped* or *interval* data. Such data were considered in Section 3.3.

It is important to understand the way (sometimes the term *mechanism* is used) in which censored data are obtained. The basic discrimination is associated with *random* and *nonrandom* censoring, the simplest cases of which are discussed below.

Type I Censoring

Consider the situation of right censoring. If the test is terminated at a given nonrandom time, T, the number of failures, r, observed during the test period will be a random variable. These censored data are *type I* or *time right singly censored* data, and the corresponding test is sometimes called *time-terminated.* For the general case, a type I censoring is considered under the following scheme of observations.

Let each unit in a sample of n units be observed during different periods of time $L_1, L_2, \ldots, L_n$. The time-to-failure of an individual unit, t_i, is considered as a distinct value, if it is less than the corresponding time period, i.e., if $t_i < L_i$. Otherwise, t_i is considered to be the *time-to-censoring*, which indicates that the time-to-failure of the ith item is greater than L_i. This is the case of *type I multiply censored* data; the case considered above is its particular case, when $L_1 = L_2 = \cdots = L_n = T$. The type I multiply censored data are quite common in reliability testing. For example, a test may start with a sample size of n but at some given times $L_1, L_2, \ldots, L_k$ ($k < n$) the prescribed numbers of units can be deleted from (or placed on) the test.

Another example of multiply censored data is the miles-to-failure information on a fleet of vehicles, which are observed for a given period of time. The mileage corresponding to failures of a particular component (e.g., alternator) would be considered as the distinct miles-to-failure. At the same time, the failure mileage of other components (e.g. battery, connectors, power distribution box, etc.) may be considered as the miles-to-censoring.

Type II Censoring

A test may also be terminated when a nonrandom number of failures (say r) specified in advance, have been observed. In this case, the duration of the test number is a random variable. This situation is known as *type II right censoring*.

It is clear that under the type II censoring only the r smallest times to failure $t_{(1)} < t_{(2)} < \cdots < t_{(r)}$ out of sample of N times to failure are observed as distinct ones. The times-to-failure $t_{(i)}$ ($i = 1, 2, \ldots, r$) are considered (as in the previous case of the type I censoring) as identically distributed r.v.s.

Types of Reliability Tests

When n components are placed on a life test, whether with replacement or not, it is sometimes necessary, due to the long life of certain components, to terminate the test and perform the reliability analysis based on the observed data up to the time of termination. Accordingly to the above discussion, there are two basic types of possible life test termination. The first type is *time terminated* (which results in the type I right censored data), and the second is *failure terminated* (resulting in type II right censored data). In the time-terminated life test, n units are placed on a test and the test is terminated after a predetermined time elapsed. The number of components failed during the test time and the corresponding time to failure of each component are recorded. In the failure-terminated life tests, n units are placed on test and the test is terminated when a predetermined number of component failures have occurred. The time to failure of each failed component including the time of the last failure occurred are recorded.

Type I and type II life tests can be performed with replacement or without replacement. Therefore, four types of life test experiments are possible. Each of these is discussed in more detail in the remainder of this section.

Random Censoring

The random censoring is typically used in reliability data analysis when there are several failure modes, which must be estimated separately. The times to failure due to each failure mode are considered as random variables having different distributions, while the whole system is considered as a *competing risks* (or *series*) system.

An example of this type of censoring could be a car dealer collecting failure data during warranty service of a given car model. The dealer and the car manufacturer are interested in reliability estimation of the car model. A car in the sample under service can fail not only due to a generic failure cause, but also due to accidents or driver errors. In the latter cases, time-to-accident or time-to-human error must be treated as time to censoring. Note that the situation might be opposite. For example, some organizations (say, an insurance company) might be interested in studying psychology of drivers for the same car model, so they need to estimate an accident rate for the given model. In this case, from the data analysis point of view, the time to generic failure becomes time-to-censoring, meanwhile the time-to-accident becomes time-to-failure. Such dualism of censored data is also important for reliability data banks development. Reliability estimation for this type of censoring is discussed in Section 3.5.

3.4.1 Exponential Distribution Point Estimation

Type I Life Test with Replacement

Suppose n components are placed on test with a replacement (i.e., monitored), and the test is terminated after a specified time t_0. The accumulated (by both failed and unfailed components) time on test, T, (in hours or other time units), is given by:

$$T = nt_0 \tag{3.57}$$

Time, T, is also called the *total time on test*. Equation (3.57) shows that at each time instant, from the beginning of the test up to time t_0, exactly n components have been on test. Accordingly, if r failures have been observed up to time t_0, then assuming the exponential distribution, the maximum likelihood point estimate of the failure rate of the component, can be found in the way similar to one considered in Example 2.25, i.e., as

$$\hat{\lambda} = \frac{r}{T}$$

The corresponding estimate of the mean-time-to-failure is given by

$$\hat{\text{MTTF}} = \frac{T}{r} \tag{3.58}$$

The number of units tested during the test, n', is

$$n' = n + r \tag{3.59}$$

Type I Life Test Without Replacement

Suppose n components are placed on test without replacement, and the test is terminated after a specified time t_0 during which r failures have occurred. The total time on test, T, for the failed and survived components is

$$T = \sum_{i=1}^{r} t_i + (n - r)t_0 \tag{3.60}$$

where $\sum_{i=1}^{r} t_i$ rpresents the accumulated time on test of the r failed components (r is random here), and $(n - r)t_0$ is the accumulated time on test of the surviving components. Using (3.60), the failure rate and MTTF estimates can be obtained from (3.58) and (3.59), respectively. Since no replacement has taken place, the total number of components tested during the test is $n' = n$.

Type II Life Test with Replacement

Consider a situation when n components are being tested with replacement (i.e., monitored), and a component is replaced with an identical component as soon as it fails (except for the last failure). If the test is terminated after a time t_r when the rth failure has occurred (i.e., r is specified (nonrandom) but t_r is random), then the total time on test, T, associated with failed and unfailed components, is given by

$$T = nt_r \tag{3.61}$$

Note that t_r, unlike t_0, is a variable in this case. If the time-to-failure follows the exponential distribution, λ is estimated as

$$\hat{\lambda} = \frac{r}{T} \tag{3.62}$$

and the respective estimate of MTTF is

$$\widehat{\text{MTTF}} = \frac{T}{r} \tag{3.63}$$

The total number of units tested, n', is

$$n' = n + r - 1 \tag{3.64}$$

where (r - 1) is the total number of failed *and* replaced components. All failed components are replaced except the last one, because the test is terminated when the last component fails (i.e., the rth failure has been observed).

Type II Life Test Without Replacement

Consider another situation when n components are being tested without replacement, i.e., when a failure occurs, the failed component is not replaced by a new one. The test is terminated at time t_r when the rth failure has occurred (i.e., r is specified but t_r is random). The total time on test of both failed and unfailed components is obtained from

$$T = \sum_{i=1}^{r} t_i + (n - r)\, t_r \tag{3.65}$$

where $\sum_{i=1}^{r} t_i$ is the accumulated time contribution from the failed components, and $(n - r)t_r$ is the accumulated time contribution from the survived components. Accordingly, the failure rate and MTTF estimates for the exponentially distributed time-to-failure can be obtained using (3.65), (3.62), and (3.63). It should also be noted that the total number of units in this test is

$$n' = n \tag{3.66}$$

since no components are being replaced.

Example 3.12

Ten light bulbs are placed under life test. The test is terminated at $t_0 = 850$ hours. Eight components fail before 850 hours have elapsed. Determine the accumulated component hours and an estimate of the failure rate and MTTF for the following situation:

a. The components are replaced when they fail.
b. The components are not replaced when they fail.
c. Repeat a. and b., assuming the test is terminated when the eighth component fails.

The failure times obtained are: 183, 318, 412, 432, 553, 680, 689, and 748.

Solution:

a. Type I test:

Using (3.56), T = 10 × 850 = 8500 component hours.

$\hat{\lambda} = 8/8500 = 9.4 \times 10^{-4}$ hr^{-1}, or from (3.58), MTTF = 1062.5 hrs.

b. Type I test:

Using (3.60), $\sum_{i=1}^{\infty} t_i = 4015$, $(n - r)\, t_0 = (10 - 8)850 = 1700$.

Thus, $T = 4015 + 1700 = 5715$ component-hours.

$\hat{\lambda} = 8/5715 = 1.4 \times 10^{-3}$ hr^{-1}.

From (3.58), MTTF = 714.4 hrs.

c. Type II test:

Here, t_r is the time-to-the-eighth failure, which is 748.

Using (3.61), $T = 10 \times 748 = 7480$ component-hours.

From (3.62), $\hat{\lambda} = 8/7480 = 1.1 \times 10^{-3}$ hr^{-1}.

From (3.63), MTTF = 935 hours.

Using (3.65), $\sum_{i=1}^{\infty} T_i = 4015$,

$(n - r)\, T_r = (10 - 8)748 = 1496$.

Thus, $T = 4015 + 1496 = 5511$ component-hours.

From (3.62), $\hat{\lambda} = 8/5511 = 1.5 \times 10^{-3}$ hr^{-1}.

From (3.63), MTTF = 688.8 hours.

A simple comparison of results shows that although the same set of data is used, the effect of the type of the test and the effect of the replacement of the failed units can have a reasonable effect on the estimated parameters.

3.4.2 Exponential Distribution Interval Estimation

In Example 2.26 and in Section (3.4.1), we discussed the maximum likelihood estimator for the parameter λ (failure rate, or hazard rate) of the exponential distribution. This point estimator is $\hat{\lambda} = r/T$, where r is the number of failures observed and T is the total time on test. Epstein (1960) has shown that if

the time-to-failure is exponentially distributed with parameter λ, the quantity $2r\lambda / \hat{\lambda} = 2\lambda T$ has the Chi-square distribution with $2r$ degrees of freedom for the type II censored data (failure-terminated test). Based on this, one can construct the corresponding confidence intervals. Because uncensored data can be considered as the particular case of the type II right censored data (when $r = n$), the same procedure is applicable to complete (uncensored) sample.

Using the distribution of $2r\lambda / \hat{\lambda}$, one can write

$$\Pr\left[\chi^2_{\alpha/2}(2r) \le \frac{2r\lambda}{\hat{\lambda}} \le \chi^2_{1-\alpha/2}(2r)\right] = 1 - \alpha \tag{3.67}$$

By rearranging and using $\hat{\lambda} = r/T$, the two-sided confidence interval for the true value of λ can be obtained as

$$\Pr\left[\frac{\chi^2_{\alpha/2}(2r)}{2T} \le \lambda \le \frac{\chi^2_{1-\alpha/2}(2r)}{2T}\right] = 1 - \alpha \tag{3.68}$$

The corresponding upper confidence limit (the one-sided confidence interval) obviously is

$$\Pr\left[0 \le \lambda \le \frac{\chi^2_{1-\alpha}(2r)}{2T}\right] = 1 - \alpha \tag{3.69}$$

Accordingly, confidence intervals for MTTF and $R(t)$ at a time $t = t_0$ can also be obtained as one-sided and two-sided confidence intervals from (3.68) and (3.69). The results are summarized in Table 3.2 (Martz and Waller (1982)).

Table 3.2 $100(1 - \alpha)\%$ Confidence Limits on λ, MTTF, and $R(t_0)$

Parameter	Type I (Time terminated test)			
	One-sided confidence limits		Two-sided confidence limits	
	Lower limit	Upper limit	Lower limit	Upper limit
λ	0	$\frac{\chi^2_{1-\alpha}(2r+2)}{2T}$	$\frac{\chi^2_{\alpha/2}(2r)}{2T}$	$\frac{\chi^2_{1-\alpha/2}(2r+2)}{2T}$
MTTF	$\frac{2T}{\chi^2_{1-\alpha}(2r+2)}$	∞	$\frac{2T}{\chi^2_{1-\alpha/2}(2r+2)}$	$\frac{2T}{\chi^2_{\alpha/2}(2r)}$
$R(t_0)$	$\exp\left[-\frac{\chi^2_{1-\alpha}(2r+2)}{2T}t_0\right]$	1	$\exp\left[-\frac{\chi^2_{1-\alpha/2}(2r+2)}{2T}t_0\right]$	$\exp\left[-\frac{\chi^2_{\alpha/2}(2r)}{2T}t_0\right]$

Parameter	Type II (Failure terminated test)			
	One-sided confidence limits		Two-sided confidence limits	
	Lower limit	Upper limit	Lower limit	Upper limit
λ	0	$\frac{\chi^2_{1-\alpha}(2r)}{2T}$	$\frac{\chi^2_{\alpha/2}(2r)}{2T}$	$\frac{\chi^2_{1-\alpha/2}(2r)}{2T}$
MTTF	$\frac{2T}{\chi^2_{1-\alpha}(2r)}$	∞	$\frac{2T}{\chi^2_{1-\alpha/2}(2r)}$	$\frac{2T}{\chi^2_{\alpha/2}(2r)}$
$R(t_0)$	$\exp\left[-\frac{\chi^2_{1-\alpha}(2r)}{2T}t_0\right]$	1	$\exp\left[-\frac{\chi^2_{1-\alpha/2}(2r)}{2T}t_0\right]$	$\exp\left[-\frac{\chi^2_{\alpha/2}(2r)}{2T}t_0\right]$

As opposed to type II censored data, the corresponding exact confidence limits for type I censored data are not available. The approximate two-sided confidence interval for failure rate, λ, for the type I (time-terminated test) data usually are constructed as:

$$\Pr\left[\frac{\chi^2_{\alpha/2}(2r)}{2T} \le \lambda \le \frac{\chi^2_{1-\alpha/2}(2r+2)}{2T}\right] = 1-\alpha \tag{3.70}$$

The respective upper confidence limit (a one-sided confidence interval) is given by

$$\Pr\left[0 \le \lambda \le \frac{\chi^2_{1-\alpha}(2r+2)}{2T}\right] = 1-\alpha \tag{3.71}$$

If no failure is observed during a test, the formal estimation gives $\hat{\lambda} = 0$, or MTTF = ∞. This cannot realistically be true; since we may have had a small or limited test. Had the test been continued, eventually a failure would have been observed. An upper confidence estimate for λ can be obtained for $r = 0$. However, the lower confidence limit cannot be obtained with $r = 0$. It is possible to relax this limitation by conservatively assuming that a failure occurs exactly at the end of the observation period. Then $r = 1$ can be used to evaluate the lower limit for two-sided confidence interval. This conservative modification, although sometimes used to allow a complete statistical analysis, lacks firm statistical basis. Welker and Lipow (1974) have shown methods to determine approximate nonzero point estimates in these cases.

Example 3.13

Twenty-five units are placed on a reliability test that lasts 500 hours. In this test, eight failures occur at 75, 115, 192, 258, 312, 389, 410, and 496 hours. The failed units are replaced. Find $\hat{\lambda}$, one-sided and two-sided confidence intervals for λ and MTTF at the 90% confidence level; one-sided and two-sided 90% confidence intervals on reliability at $t_0 = 1000$ hours.

Solution:

This is a type I test. The accumulated time T is obtained from (3.56)

$$T = 25 \times 500 = 12{,}500 \text{ hours.}$$

The point estimate of failure rate is

$$\hat{\lambda} = 8/12{,}500 = 6.4E-4 \text{ hr}^{-1}$$

One-sided confidence interval for λ is

$$0 \le \lambda \le \frac{\chi^2(2 \times 8 + 2)}{2 \times 12{,}500}$$

From Table A.3,

$$\chi^2_{0.9}(18) = 25.99, \quad 0 \le \lambda \le 1.04E-3 \text{ hr}^{-1}$$

Two-sided confidence interval for λ is

$$\frac{\chi^2_{0.05}(2 \times 8)}{2 \times 12{,}500} \le \lambda \le \frac{\chi^2_{0.95}(2 \times 8 + 2)}{2 \times 12{,}500}$$

From Table (A.3),

$$\chi^2_{0.05}(16) = 7.96, \quad \text{and} \quad \chi^2_{0.95}(18) = 28.87$$

Thus

$$3.18E-4 \text{ hr}^{-1} \le \lambda \le 1.15E-3 \text{ hr}^{-1}$$

One-sided 90% confidence interval for $R(1000)$ are

$$\exp[(-1.04E-3)(1000)] \le R(1000) \le 1$$

or

$$0.35 \le R(1000) \le 1$$

Two-sided 90% confidence interval for $R(t)$ is

$$\exp[(-1.15E-3)(1000)] \le R(1000) \le \exp[(-3.18E-4)(1000)]$$

or

$$0.32 \le R(1000) \le 0.73$$

 See the software supplement for the automated point and interval estimation of the exponential distribution parameters.

3.4.3 Lognormal Distribution

The lognormal distribution is commonly used to represent occurrence of certain events in time. For example, a r.v. representing the length of time interval required for repair of hardware follows a lognormal distribution. Because the lognormal distribution has two parameters, parameter estimation poses a more challenging problem than for the exponential distribution. Taking the natural logarithm of data, the analysis is reduced to the case of the normal distribution, so that the point estimates for the two parameters of the lognormal distribution for a complete sample of size n can be obtained from

$$\hat{\mu}_t = \sum_{i=1}^{n} \frac{\ln(t_i)}{n} \tag{3.72}$$

$$\hat{\sigma}_t^2 = \frac{\sum_{i=1}^{n} \left[\ln(t_i) - \hat{\mu}\right]^2}{n-1}. \tag{3.73}$$

The confidence interval for μ_t is given by

$$\Pr\left[\hat{\mu}_t - \frac{(\hat{\sigma}_t)\, t_{\alpha/2}}{\sqrt{n}} \le \mu_t \le \hat{\mu}_t + \frac{(\hat{\sigma}_t)\, t_{\alpha/2}}{\sqrt{n}}\right] = 1 - \alpha \tag{3.74}$$

The respective confidence interval for σ_t^2 is :

$$\Pr\left[\frac{(n-1)\hat{\sigma}_t^2}{\chi_{1-\alpha/2}(n-1)} \le \sigma_t^2 \le \frac{(n-1)\hat{\sigma}_t^2}{\chi_{\alpha/2}(n-1)}\right] = 1-\alpha \qquad (3.75)$$

In the case of censored data, the corresponding statistical estimation turns out to be much more complicated, readers are referred to (Nelson (1982)) and (Lawless (1982)).

3.4.4 Weibull Distribution

The Weibull distribution can be used for the data which are assumed to be from an increasing, decreasing, or constant failure rate distributions. Similar to the (log)normal distribution, it is a two-parameter distribution and its estimation, even in the case of complete (uncensored) data, is not a trivial problem.

It can be easily shown that, in the situation when all n units placed on test or under observation have failed, the maximum likelihood estimates of β and α parameters of the Weibull distribution (3.17) can be obtained as a solution of the following system of nonlinear equations:

$$\frac{\sum_{i=1}^{n}(t_i)^{\beta}\ln t_i}{\sum_{i=1}^{n}(t_i)^{\beta}} - \frac{1}{\beta} = \frac{1}{n}\sum_{i=1}^{n}\ln t_i \qquad (3.76)$$

and

$$\hat{\alpha} = \left(\frac{\sum_{i=1}^{n}\ln(t_i)^{\beta}}{n}\right)^{1/\beta}$$

This system can be solved using an appropriate numerical procedure. The corresponding confidence estimation also is not trivial. Readers are referred to Bain (1978), Mann et al. (1974), Nelson (1982), Leemis (1995) for further discussions. In Chapter 5, we will discuss another form of estimating the Weibull distribution parameters.

Example 3.14

Using the data given in Example 3.9, obtain the maximum likelihood estimators for the parameters of a Weibull distribution.

Solution:

Using the numerical procedure, system (3.76) is solved, which result in the following estimates $\hat{\beta} = 2.1$, $\hat{\alpha} = 396$ hours. Comparison of these results with the plot from Example 3.9 is reasonable, but it points out the approximate nature of these data analysis methods, and emphasizes the need for using more than one of the methods discussed.

 See the software supplement for the automated maximum likelihood estimation of the exponential, normal, lognomal and Weibull distribution parameters.

3.4.5 Binomial Distribution

When the data are in the form of failures occurring on demand, i.e., X failures observed in n trials, there is a constant probability of failure (or success), and the binomial distribution can be used as an appropriate model. This is often the situation for standby components (or systems). For instance, a redundant pump is demanded for operation n times in a given period of test or observation.

The best estimator for p is given by the obvious formula:

$$\hat{p} = \frac{x}{n} \tag{3.77}$$

The lower and upper confidence limits for p can be found, using the, so-called, Clopper–Pearson procedure, see (Nelson (1982)):.

$$p_l = \left\{ 1 + (n - x + 1) x^{-1} F_{1-\alpha/2}[2n - 2x + 2; 2x] \right\}^{-1} \tag{3.78}$$

$$p_u = \left\{ 1 + (n - x) \left\{ (x + 1) F_{1-\alpha/2} [2x + 2; 2n - 2x] \right\}^{-1} \right\}^{-1} \tag{3.79}$$

where $F_{(1-\alpha/2)}(f_1; f_2)$ is the $(1 - \alpha/2)$ quantile (or the 100 $(1 - \alpha/2)$ percentiles) of the F-distribution with f_1 degrees of freedom for the numerator, and f_2 degrees of freedom for the denominator. Table A.5 contains some percentiles of F-distribution. As mentioned in Chapter 2, the Poisson distribution can be used as an approximation to the binomial distribution when the parameter, p, of the

binomial distribution is small and parameter n is large, e.g., $x < n/10$, which means that approximate confidence limits can be constructed using (3.70) with $r = x$ and $T = n$.

Example 3.15

An emergency pump in a nuclear power plant is in a standby mode. A total of 563 start tests for the pump and only 3 failures have been observed. No degradation in the pump's physical characteristics or changes in operating environment have been observed. Find the 90% confidence interval for the probability of failure per demand.

Solution:

Denote $n = 563$, $x = 3$. Using (3.80–3.82), find $\hat{p} = 3/563 = 0.0053$.

$$p_l = \{1 + (563 - 3 + 1)/3\, F_{0.95}(2 \times 563 - 2 \times 3 + 2;\ 2 \times 3)\}^{-1} = 0.0014$$

Where $F_{0.95}(1122;6) = 3.67$ from Table A.5.

Similarly,

$$p_u = \{1 + (563 - 3)\{(3 + 1)F_{0.95}(2 \times 3 + 2;\ 2 \times 563 - 2 \times 3\}^{-1}\}^{-1} = 0.0137$$

Therefore,

$$\Pr(0.0014 \le p \le 0.0137) = 90\%$$

Example 3.16

In a commercial nuclear plant, the performance of the emergency diesel generators has been observed for about 5 years. During this time, there have been 35 real demands with 4 observed failures. Find the 90% confidence limits and point estimate for the probability of failure per demand. What would the error be if we used (3.70) instead of (3.78) and (3.79) to solve this problem?

Solution:

Here, $x = 4$ and $n = 35$. Using (3.77),

$$\hat{p} = \frac{4}{35} = 0.114$$

To find lower and upper limits, use (3.78) and (3.79). Thus,

$$p_l = \{1 + [(35 - 4 + 1)/4]\, F_{0.95}(2 \times 35 - 2 \times 4 + 2;\ 2 \times 4)\}^{-1} = 0.04$$

$$p_u = \{1 + (35 - 4)\{(4 + 1)F_{0.95}(2 \times 4 + 2;\ 2 \times 35 - 2 \times 4\}^{-1}\}^{-1} = 0.243$$

If we used (3.70),

$$p_l = \frac{\chi^2_{0.05}(8)}{2 \times 35} = \frac{2.733}{70} = 0.039$$

$$p_u = \frac{\chi^2_{0.95}(10)}{2 \times 35} = \frac{18.31}{70} = 0.262$$

The error due to this approximation is

$$\text{Lower limit error} = \frac{|0.04 - 0.039|}{0.04} \times 100 = 2.5\%$$

$$\text{Upper limit error} = \frac{|0.243 - 0.262|}{0.243} \times 100 = 7.8\%$$

Note that this is not a negligible error, and (3.70) should not be used. Since $x > n/10$, equation (3.70) is clearly not a good approximation.

 See the software supplement for the automated interval estimation of the exponential, normal, lognomal and binomial distribution parameters.

3.5 CLASSICAL NONPARAMETRIC DISTRIBUTION ESTIMATION

Based on the considerations given in Section 3.1, one can state that any reliability measure or index can be expressed in terms of time-to-failure cumulative distribution function (cdf) or reliability function. Thus, the problem of estimation of these functions is of great importance. The commonly used estimate of cdf is the *empirical (or sample) distribution function* (edf) introduced for uncensored data in Chapter 2, (see (2.93)). In this section we consider some other nonparametric point and confidence distribution estimation procedures applicable for censored data.

3.5.1 Confidence Intervals for Cumulative Distribution Function and Reliability Function for Complete and Singly Censored Data

The construction of an edf requires a complete sample. It can also be done for the right censored samples for the failure times which are less than the last time to failure observed ($t < t_{(r)}$). The edf is a random function, since it depends on the sample units. For any given point, t, the edf, $S_n(t)$, is the fraction of sample items failed before t.

The edf is, in a sense, the estimate of the probability, p, in a binomial trial, and this probability is $p = F(t)$. Note that it is easy to show that the maximum likelihood estimator of the binomial parameter p coincides with $S_n(t)$, and $S_n(t)$ is a consistent estimator of the cdf, $F(t)$.

Using relationship (3.2) between the cdf and reliability function, and edf (2.93), it is easy to get the respective estimate of the reliability function. This estimate, called *empirical (or sample) reliability function*, is given by

$$R_n(t) = \begin{cases} 1, & 0 < t < t_{(1)} \\ 1 - \dfrac{i}{n}, & t_{(i)} \le t < t_{(i+1)} \quad i = 1, \cdots, n-1 \\ 0, & t_{(n)} \le t < \infty \end{cases} \tag{3.80}$$

where $t_{(1)} < t_{(2)} < \cdots < t_{(n)}$ are the ordered sample data (the so-called, order statistics).

It is clear that the mean number of failures observed during time, t, is $E(r) = pn = F(t)n$, so that the mean value of the fraction of sample items failed before t, is $E(r/n) = p = F(t)$ and the variance of this fraction is given by

$$\operatorname{var}\left(\frac{r}{n}\right) = \frac{p(1-p)}{n} = \frac{F(t)[1 - F(t)]}{n} \tag{3.81}$$

For practical problems considered, (3.81) is used with replacing $F(t)$ with $S_n(t)$. As the sample size, n, increases the binomial distribution can be approximated by a normal distribution (consistent with the discussion in Chapter 2) with the same mean and variance (i.e., $\mu = np$, $\sigma^2 = np(1-p)$), which provides reasonable results if np and $n(1-p)$ are both greater or equal to 5. Using this approximation, the following $100(1-\alpha)\%$ confidence interval for the unknown cdf, $F(t)$, at any given point t can be constructed as:

$$\begin{aligned} & S_n(t) - z_{1-\alpha/2}\left(\frac{S_n(t)[1 - S_n(t)]}{n}\right)^{1/2} \\ & \le F(t) \le S_n(t) + z_{1-\alpha/2}\left(\frac{S_n(t)[1 - S_n(t)]}{n}\right)^{1/2} \end{aligned} \tag{3.82}$$

where z_α is the quantile of level α of the standard normal distribution.

The corresponding estimates for the reliability (survivor) function can be obtained using (3.2), $R_n(t) = 1 - S_n(t)$.

Example 3.17

Using the data from Example 3.9 find the point nonparametric estimate and 95% confidence interval for the cdf, $F(t)$, for $t = 350$ hours.

Solution:

Using (2.93) the point estimate for $F(350)$ is $S_n(350) = 5/10$ (note that we have five observations out of ten which are less than 350 hours).

The respective approximate 95% confidence interval based on (3.82) is

$$0.5 - 1.96\left(\frac{0.5(1-0.5)}{10}\right)^{1/2} \le F(350) \le 0.5 + 1.96\left(\frac{0.5(1-0.5)}{10}\right)^{1/2}$$

Therefore, $\Pr(0.1900 < F(350) < 0.6099) = 0.95$.

Using a complete or right censored sample from an unknown cdf, one can also get the strict *confidence intervals for the unknown cdf, $F(t)$*. This can be done using the same Clopper–Pearson procedure for constructing the confidence intervals for a binomial parameter p, i.e., using (3.78) and (3.79). These limits can also be expressed in more compact form in terms of the, so-called, incomplete beta function as follows.

The lower confidence limit, $F_l(t)$, at the point t where $S_n(t) = r/n$ ($r = 0, 1, 2, \ldots, n$), is the largest value of p that satisfies the following inequality

$$I_p(r, n - r + 1) \le \frac{\alpha}{2} \tag{3.83}$$

and the upper confidence limit, $F_u(t)$, at the same point is the smallest p satisfying the inequality

$$I_{1-p}(n - r, r + 1) \le \frac{\alpha}{2} \tag{3.84}$$

where $I_x(\alpha, \beta)$ is the incomplete beta function, which was introduced in Chapter 2 as the cdf of the beta distribution (2.55). The incomplete beta function is difficult to tabulate, however its numerical approximation is available within any statistical package.

 See the software supplement for the automated construction of confidence intervals for an unknown cdf using both (3.78–3.79) and (3.83–3.84).

Example 3.18

For the data from Example 3.17 find the strict 95% confidence interval for the cdf, $F(t)$, for $t = 350$ hours, using (3.83) and (3.84).

Solution:

Using (3.83) the lower confidence limit is found from

$$I_p(5,10 - 5 + 1) \leq 0.025 \text{ as } 0.1871$$

and, using (3.84), the upper confidence limit is found from

$$I_{1-p}(5, 6) \leq 0.025 \text{ as } 0.8131$$

Therefore, the strict confidence interval is

$$\Pr(0.1871 < F(350) < 0.8131) = 0.95$$

which is reasonably close to the approximate interval obtained in the previous example.

Another typical reliability estimation problem, which can be solved using this nonparametric approach, is to estimate the lower confidence limit for the reliability function, using the same type of data. This can be done using (3.84), in which $1 - p = 1 - F(t)$ is replaced by the reliability function, $R(t)$. Accordingly, one gets

$$I_R(n - r, r + 1) \leq \alpha. \tag{3.85}$$

This procedure is illustrated by the following example.

Example 3.19

Twenty-two identical components were tested during 1000 hours, and no failure was observed. Find the lower confidence limit, $R_l(t)$, for the reliability function at $t = 1000$ hours and for confidence probability $1 - \alpha = 0.90$.

Solution:

We need to find the largest value of R satisfying (3.85). For the problem considered $\alpha = 0.1$, $r = 0$, and $n = 22$, therefore,

$$I_R(22, 1) \leq 0.1$$

Solving for R, one gets $R_l \approx 0.90$.

Another possible application of (3.85) is associated with reliability demonstration tests, when the lower $1-\alpha$ confidence limit for reliability, R_l, is given together with acceptable number of failures, r, during the test duration. The problem is to find the sample size, n, to be tested with the number of failures not exceeding r, so that quantities n, r, R_l would satisfy (3.85). It should be noted that for given R_l, r, and α relationship (3.85) cannot be satisfied for any sample size n. For given values of R_l and α, the minimum sample size, for which (3.85) can be satisfied, corresponds to $r = 0$. This is illustrated by the following example.

Example 3.20

The reliability test on an automotive component has to demonstrate the lower limit of reliability of 86% with 90% confidence. Under assumption that no failure is tolerated, what sample size has to be tested to satisfy the above requirements?

Solution:

With $r = 0$, (3.85) can be rewritten in the following simple form:

$$(R_l)^n = 1-\alpha,$$

from which the required sample size is found as

$$n = \frac{\ln(1-\alpha)}{\ln(R_l)} = \frac{\ln(1-0.9)}{\ln(0.86)} \approx 15$$

 See the software supplement for the automated evaluation of the sample size in reliability demonstration for any values of R_l, α, and r.

3.5.2 Confidence Intervals for Cumulative Distribution Function and Reliability Function for Multiply Censored Data

The point and confidence estimation considered so far do not apply to multiply censored data. For such samples, the, so-called, *Kaplan–Meier* or *product-limit* estimate, which is the MLE of the cdf, can be used.

Let's have a sample of n times to failure, among which only k failure times are distinct times to failure. Denote these ordered times (order statistics) as: $t_{(1)} \le$

$t_{(2)} \leq \ldots \leq t_{(k)}$, and let $t_{(0)}$ be equal to zero, i.e, $t_{(0)} = 0$. Let n_j be the number of items under observation just before $t_{(j)}$. Assume that the time-to-failure cdf is continuous, so that there is only one failure at every $t_{(i)}$. Then, $n_{j+1} = n_j - 1$. Under these conditions, the product limit estimate is given by:

$$S_n(t) = 1 - R_n(t) =$$

$$\begin{cases} 0 & 0 \leq t < t_{(1)} \\ 1 - \prod_{j=1}^{i} \dfrac{n_j - 1}{n_j} & t_{(i)} \leq t < t_{(i+1)}, \quad i = 1, \cdots, m-1 \\ 1 & t \geq t_{(m)} \end{cases} \tag{3.86}$$

where integer $m = k$, if $k < n$, and $m = n$, if $k = n$.

It is possible to show that for uncensored (complete) data samples, the product limit estimate coincides with the edf given by (2.93). In general case (including discrete distribution, censored, or grouped data), the Kaplan–Meier estimate is given by

$$S_n(t) = 1 - R_n(t) =$$

$$\begin{cases} 0 & 0 \leq t < t_{(1)} \\ 1 - \prod_{j=1}^{i} \dfrac{n_j - 1}{n_j} & t_{(i)} \leq t < t_{(i+1)}, \quad i = 1, \cdots, m-1 \\ 1 & t \geq t_{(m)} \end{cases} \tag{3.87}$$

where d_j is the number of failures at $t_{(j)}$.

For estimation of variance of S_n (or R_n) the, so-called, Greenwood's formula (Lawless (1982)) is used:

$$\widehat{\operatorname{var}}\left[S_n(t)\right] = \widehat{\operatorname{var}}\left[R_n(t)\right] = \sum_{j:\, t_{(j)} < t} \frac{d_j}{n_j\left(n_j - d_j\right)} \tag{3.88}$$

Combining the product-limit estimate and Greenwood's formula into equations similar to (3.82), one can construct the corresponding approximate confidence limits for the reliability of cdf of interest.

Example 3.21

The table below shows censored test data of a mechanical component. Find the Kaplan–Meier estimate of the time-to-failure cdf.

i	Ordered failure time, $t_{(i)}$, or time to censoring, $t_{(i)}$*	$S_n[t_{(i)}]$
0	0	0
1	31.7	1 - 15/16 = 0.059
2	39.2	1 - (15/16) (14/15) = 0.125
3	57.2	1 - (15/16) (14/15) (13/14) = 0.187
4	65.0*	0.187
5	65.8	1 (13/16) [(13 1 1)/12] = 1 0.745 = 0.255
6	70	1 0.745 (10/11) = 0.323
7	75*	0.323
8	75.2*	0.323
9	87.5*	0.323
10	88.3*	0.323
11	94.2*	0.323
12	101.7*	0.323
13	105.8	0.492
14	109.2*	0.492
15	110	0.746
16	130 *	0.746

 See the software supplement for the automated estimation of the reliability function for multiply censored data.

3.6 BAYESIAN ESTIMATION PROCEDURES

In Section 3.5, we discussed the importance of quantifying uncertainties associated with the estimates of various distribution parameters. In the preceding sections, we also discussed formal statistical methods for quantifying the uncertainties. Namely, we discussed the concept of confidence intervals, which, in essence, is a statistical treatment of available information. It is evident that the more data are available, the more confident and accurate the statistical inference is. For this reason, the statistical approach is sometimes called the *frequentist* method of treatment. In the

framework of Bayesian approach, the parameters of interest are treated as random variables, the true values of which are unknown.Thus, a distribution can be assigned to represent the parameter; the mean (or for some cases the median) of the distribution can be used as an estimate of the parameter of interest. The pdf of a parameter in Bayesian terms can be obtained from a *prior* and *posterior* pdf. In practice, however, the prior pdf is used to represent the relevant prior knowledge, including subjective judgment regarding the characteristics of the parameter and its distribution. When the prior knowledge is combined with other relevant information (often statistics obtained from tests and observations), a posterior distribution is obtained, which better represents the parameter of interest. Since the selection of the prior and the determination of the posterior often involve subjective judgements, the Bayesian estimation is sometimes called the *subjectivist* approach to parameter estimation.

The basic concept of the Bayesian estimation was discussed in Chapter 2. In essence, the Bayes' theorem can be written in one of three forms: discrete, continuous, or mixed. Martz and Waller (1982) have significantly elaborated on the concept of Bayesian technique and its application to reliability analysis. The discrete form of Bayes' theorem was discussed in Section 2.2.3. The continuous and mixed forms which are the common forms used for parameter estimation in reliability and risk analysis are briefly discussed below.

Let θ be a parameter of interest. It can be a parameter of a time-to-failure distribution or a reliability index, such as mean time-to-failure, failure rate, etc. Suppose parameter θ is a continuous r.v., so that the prior and posterior distributions of θ can be represented by continuous pdfs. Let $h(\theta)$ be a continuous prior pdf of θ, and let $l(\theta|t)$ be the likelihood function based on sample data, t, then the posterior pdf of θ is given by

$$f(\theta \mid t) = \frac{h(\theta)\, l(\theta \mid t)}{\int_{-\infty}^{\infty} h(\theta)\, l(\theta \mid t)\, d\theta} \tag{3.89}$$

Relationship (3.89) is the Bayes' theorem for a continuous r.v. The Bayesian inference includes the following three stages:

Constructing the likelihood function based on the distribution of interest and type of data available (complete samples, censored data, grouped data, etc.)

Quantification of the prior information about the parameter of interest in the form of a prior distribution

Choosing a loss function, $L(\theta, \hat{\theta})$, which is a measure of discrepancy between the true value of the parameter θ and its estimate $\hat{\theta}$. The most popular lost function are the quadratic (squared-error) loss function

$$L(\theta, \hat{\theta}) = (\theta - \hat{\theta})^0 \tag{3.90}$$

which is equal to the square of the distance between θ and its estimate $\hat{\theta}$. Also, the loss function of the following form can be used

$$L(\theta, \hat{\theta}) = |\theta - \hat{\theta}| \tag{3.91}$$

Calculation of the posterior distribution using Bayes' theorem
Obtaining point and interval estimates.

A point Bayesian estimate is the estimate which minimizes the so-called Bayesian risk, $G(\hat{\theta})$, which is introduced as the expected (mean) value of the loss function with respect to the posterior distribution, i.e.,

$$G(\hat{\theta}) = \int_{-\infty}^{\infty} L(\theta, \hat{\theta}) f(\theta \mid t)\, d\theta \tag{3.92}$$

The estimate $\hat{\theta}$ is chosen as one minimizing the mean of the loss function (3.92), i.e.,

$$\hat{\theta} = \arg \min_{-\infty < \theta < \infty} G(\theta)$$

If loss function (3.90) is used, the corresponding Bayes' point estimate is the posterior mean of θ, i.e.,

$$\hat{\theta} = \int_{-\infty}^{\infty} \theta f(\theta \mid t)\, d\theta \tag{3.93}$$

If loss function (3.91) is chosen, the corresponding Bayes' point estimate is the median of the posterior distribution of θ.

The Bayes' analog of classical confidence interval is known as Bayes' probability interval. For constructing Bayes' probability interval, the following obvious relationship based on the posterior distribution, is used:

$$\Pr(\theta_l < \theta \leq \theta_u) = 1 - \alpha \tag{3.94}$$

Similar to the classical estimation, the Bayesian estimation procedures can be divided into parametric and nonparametric ones. The following are examples of the parametric Bayesian estimation.

3.6.1 Estimation of the Parameter λ of Exponential Distribution

Consider a test of N units which results in r distinct times to failure $t_{(1)} < t_{(2)} < \cdots < t_{(r)}$ and $N - r$ times to censoring $t_{c1}, t_{c2}, \ldots, t_{c(N-r)}$, so that the total time on test, T, is

$$T = \sum_{i=1}^{r} t_{(i)} + \sum_{i=1}^{N-r} t_{ci} \tag{3.95}$$

A time-to-failure is supposed to have the exponential distribution. The problem is to estimate the parameter λ of the exponential distribution.

Suppose a gamma distribution is used as the prior distribution of parameter λ. This distribution was already discussed in the present chapter as well as in Chapter 2. The pdf of the gamma distribution is given by (2.53) as a time-to-failure distribution. Rewrite the pdf as a function of λ, which is now being considered as a r.v.:

$$h(\lambda; \delta, \rho) = \frac{1}{\Gamma(\delta)} \rho^{\delta} \lambda^{\delta-1} e^{-\rho\lambda}, \quad \lambda > 0, \quad \rho \geq 0, \quad \delta \geq 0 \tag{3.96}$$

In the Bayesian context, the parameters δ and ρ, as the parameters of prior distribution, are sometimes called *hyperparameters*. Selection of the hyperparameters is discussed later, but for the time being, suppose that these parameters are known. Also, suppose that the quadratic loss function (3.90) is used.

For the available data and the exponential time-to-failure distribution, the likelihood function can be written as

$$\begin{aligned} l(\lambda \mid t) &= \prod_{i=1}^{r} \lambda e^{-\lambda t_{(i)}} \prod_{j=1}^{N-r} e^{-\lambda t_{cj}} \\ &= \lambda^{r} e^{-\lambda T} \end{aligned} \tag{3.97}$$

where T is the total time on test given by (3.95).

Using the prior distribution (3.96), the likelihood function (3.97) and the Bayes' theorem in the form (3.89), one can find the posterior pdf of the parameter λ, as:

$$f(\lambda \mid T) = \frac{e^{-\lambda(T+\rho)} \lambda^{r+\delta-1}}{\int_0^{\infty} \lambda^{r+\delta-1} e^{-\lambda(T+\rho)} d\lambda} \tag{3.98}$$

Recalling the definition of the gamma function (2.54), it is easy to show that the integral in the denominator of (3.98) is

$$\int_0^{\infty} \lambda^{r+\delta-1} e^{-\lambda(T+\rho)} d\lambda = \frac{\Gamma(\delta + r)}{(\rho + T)^{\delta+r}} \tag{3.99}$$

Finally, the posterior pdf of λ can be written as

$$f(\lambda \mid T) = \frac{(\rho + T)^{\delta + r}}{\Gamma(\delta + r)} \lambda^{r + \delta - 1} e^{-\lambda(T + \rho)} \tag{3.100}$$

Comparing with the prior pdf(3.96), one can conclude that the posterior pdf (3.100) is also a gamma distribution with parameters $\rho' = r + \delta$, and $\lambda' = T + \rho$. Prior distributions which result in posterior distributions of the same family are referred to as *conjugate prior distributions.*

Keeping in mind that the quadratic loss function is used, the point Bayesian estimate of λ is the mean of the posterior gamma distribution with parameters ρ' and λ', so that the point Bayesian estimate, λ_B, is

$$\lambda_B = \frac{\rho'}{\lambda'} = \frac{r + \delta}{T + \rho} \tag{3.101}$$

The corresponding probability intervals can be obtained using (3.94). For example, the $100(1 - \alpha)$ level upper one-sided Bayes' probability interval for λ can be obtained from the following equation based on the posterior distribution (3.100).The same upper one-sided probability interval for λ can

$$\Pr(\lambda < \lambda_u) = 1 - \alpha \tag{3.102}$$

be expressed in a more convenient form similar to the classical confidence interval, i.e., in terms of the Chi-square distribution, as:

$$\Pr\left[2\lambda(\rho + T) < \chi^2_{1-\alpha}[2(\delta + r)]\right] = 1 - \alpha$$

and

$$\lambda_u = \frac{\chi^2_{1-\alpha}[2(\delta + r)]}{2(\rho + T)} \tag{3.103}$$

Note that, contrary to the classical estimation, the number of degrees of freedom, $2(\delta + r)$, for the Bayes' confidence limits is not necessarily integer.

The gamma distribution was chosen as the prior distribution for the purpose of simplicity and performance. Now let's consider the reliability interpretation of the Bayes' estimate obtained.

Selecting Parameters of Prior Distribution

The point Bayes' estimate λ_B (3.101) can be interpreted as follows. The parameter δ can be considered as a prior (fictitious) number of failures "observed" during a prior (fictitious) test, having ρ as the total time on test. So, intuitively, one would choose the prior estimate of λ as the ratio δ/ρ, which coincides with the *mean value* (recall (3.22)) of the prior gamma distribution used (3.96).

The corresponding practical situation is quite opposite—usually one has a prior estimate of λ, meanwhile the parameters δ and ρ must be found. Having the prior point estimate λ_p, one can only estimate the ratio $\delta/\rho = \lambda_p$. For estimating these parameters separately, some additional information about the degree of belief or accuracy of this prior estimate is required. Since variance of the gamma distribution is δ/ρ^2, the coefficient of variation of the prior distributionis $1/\delta^{1/2}$ (the coefficient of variation is the ratio of standard deviation to mean). The coefficient of variation can be used as a measure of relative accuracy of the prior point estimate of λ_p. It is further discussed in Section 7.2.

Thus, having the prior point estimate, λ_p, and the relative error of this estimate, one can estimate the corresponding parameters of the prior gamma distribution. To get a feeling about the scale of these errors, consider the following numerical example. Let the prior point estimate λ_p be 0.01 (in some arbitrary units). The corresponding values of the coefficient of variation, expressed in percent, for different values of the parameters δ and ρ, are given in Table 3.3.

The approach is a very simple and convenient way of expressing prior information (e.g., knowledge or degree of belief expressed by experts) in terms of the gamma prior distribution. Another approach to selecting parameters of the prior gamma distribution is based on the quantile (or percentile) estimation of

Table 3.3 Parameters and Coefficients of Variation of Gamma Prior Distribution with Mean Equal to 0.01

Shape Parameter, δ, [prior (fictitious) number of failures]	Scale Parameter, ρ, [prior (fictitious) total time on test]	Coefficient of Variation, %
1	100	100
5	500	45
10	1000	32
100	10000	10

the prior distribution (Kapur and Lamberson (1977), Martz and Waller (1982)). Let $H(\lambda, \rho, \delta)$ be the cdf of the prior gamma distribution. Recalling the definition of a quantile of level p, λ_p, one can write

$$H(\lambda_p, \rho, \delta) = \int_0^{\lambda_p} h(x)\,dx = p \tag{3.104}$$

where $h(\lambda)$ is given by (3.96).

Having a pair of quantiles, say, λ_{p1} and λ_{p2} of levels p_1 and p_2 one gets two equations (3.104) with two unknowns. These equations uniquely determine the parameters of gamma distribution. Practically, the procedure is as follows.

1. Expert specifies the values p_1, λ_{p1}, p_2, and λ_{p2} such that

$$H(\lambda_{p1}, \rho, \delta) = \Pr(\lambda < \lambda_{p1}) = p_1$$

$$H(\lambda_{p2}, \rho, \delta) = \Pr(\lambda < \lambda_{p2}) = p_2$$

   For example, he/she specifies that there is 90% probability that parameter λ is less than 0.1 and 50% probability that λ is less than 0.01.
2. A numerical procedure is used to solve the system of equations in 1. to find the values of the parameters δ and ρ.

Usually the value of p_1 is chosen as 0.90 or 0.95, and the value of p_2 as 0.5 (the median value). In (Martz and Waller (1982)) a special table and graphs are provided for solving the system of equations above.

The procedure discussed above is not limited by gamma distribution. This procedure is known as *the method of quantiles*, and is applied as an estimation method of distribution parameters, not necessarily in Bayes' context.

Example 3.22

A sample of identical units was tested. Six failures were observed during the test. The total time on test is 1440 hours. The time-to-failure distribution is assumed to be exponential.

The gamma distribution with the mean of 0.01 hr^{-1} and with the coefficient of variation of 30% was selected as a prior distribution to represent the parameter of interest, λ. Find the posterior point estimate and the upper 90% probability limit for λ. See Figure 3.10 for the prior and the posterior distribution of λ.

Solution:

The respective parameters of the distribution are $\delta = 11.1$ and $\rho = 1100$ hr. Using (3.101), the point posterior estimate (the mean of the posterior distribution) of the hazard rate is evaluated as

$$\lambda_B = \frac{11.1 + 6}{1110 + 1440} = 6.71E-3 \text{ hr}^{-1}$$

Using (3.103), get the 90% upper limit of one-sided Bayes' probability interval for λ as

$$\lambda_u = \frac{\chi^2_{0.9}(17.1)}{2(2550)} = 1.85E-3 \text{ hr}^{-1}$$

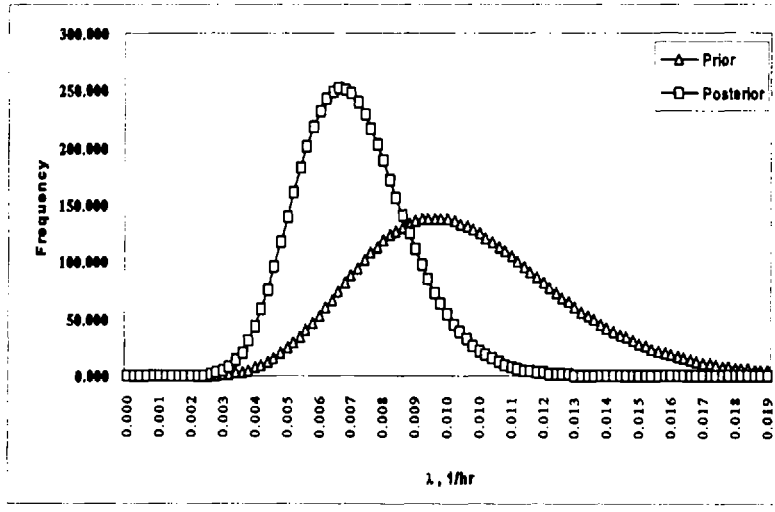


Figure 3.10 Prior and posterior distribution of λ in Example 3.22.

Uniform Prior Distribution

This prior distribution has a very simple form, so that it is easy to use as an expression of prior information. Consider the prior pdf in the form:

$$h(\lambda; a, b) = \begin{cases} \dfrac{1}{b-a}, & a < \lambda < b \\ 0, & \text{otherwise} \end{cases} \tag{3.105}$$

Using the likelihood function (3.97), the posterior pdf can be written as

$$f(\lambda \mid t) = \frac{\lambda^r e^{-\lambda T}}{\int_a^b \lambda^r e^{-\lambda T} d\lambda}, \quad a < \lambda < b \tag{3.106}$$

Substituting $y = \lambda T$ in the denominator of (3.106), one gets

$$\int_a^b \lambda^r e^{-\lambda T} d\lambda = \int_{aT}^{bT} \frac{y^r e^{-y} dy}{T^{r+1}}$$

Recall the definition of the incomplete gamma function

$$\Gamma(c, z) = \int_0^z y^{c-1} e^{-y} dy, \quad c > 0$$

in which $\Gamma(c, \infty) = \Gamma(c)$, where $\Gamma(c)$ is the ("complete") gamma function (see Equation (2.54) in Chapter 2). Accordingly the denominator in (3.106) can be written as

$$\int_a^b \lambda^r e^{-\lambda T} d\lambda = \frac{1}{T^{r+1}} \left[\Gamma(r+1, bT) - \Gamma(r+1, aT) \right]$$

and the posterior pdf (3.106) takes on the form

$$f(\lambda \mid t) = \frac{t^{r+1} \lambda^r e^{-\lambda T}}{\left[\Gamma(r+1, bT) - \Gamma(r+1, aT) \right]}, \quad a < \lambda < b \tag{3.107}$$

Note that the classical maximum likelihood point estimate, r/T, is the mode of the posterior pdf (3.107).The corresponding mean value (which is the Bayes' point estimator of λ for the case of the squared-error loss function) is given by (Martz and Waller (1982)):

$$\lambda_B = \frac{\Gamma(r+2, bT) - \Gamma(r+2, aT)}{T\left[\Gamma(r+1, bT) - \Gamma(r+1, aT) \right]} \tag{3.108}$$

The two-sided Bayes' probability interval can now be found in the standard way, i.e., by integrating (3.107), which results in solving the following equations with respect to λ_l and λ_u

$$\Pr(\lambda < \lambda_l) = \int_a^{\lambda_l} \frac{t^{r+1} \lambda^r e^{-\lambda T} d\lambda}{\Gamma(r+1, bT) - \Gamma(r+1, aT)}$$

$$= \frac{\Gamma(r+1, \lambda_l T) - \Gamma(r+1, aT)}{\Gamma(r+1, bT) - \Gamma(r+1, aT)} = \frac{\alpha}{2} \tag{3.109}$$

$$\Pr(\lambda > \lambda_u) = \frac{\Gamma(r+1, bT) - \Gamma(r+1, \lambda_u T)}{\Gamma(r+1, bT) - \Gamma(r+1, aT)} = \frac{\alpha}{2}$$

Example 3.23

An electronic component has the exponential time-to-failure distribution. The uniform prior distribution of λ is given by $a = 10^{-6}$ and $b = 5 \times 10^{-6}$ hr^{-1} in Equation (3.105). A life test of the component results in $r = 30$ failures in total time on test of $T = 10^7$ hours. Find the point estimate (mean) and 90% two-sided Bayes' probability interval for the failure rate λ.

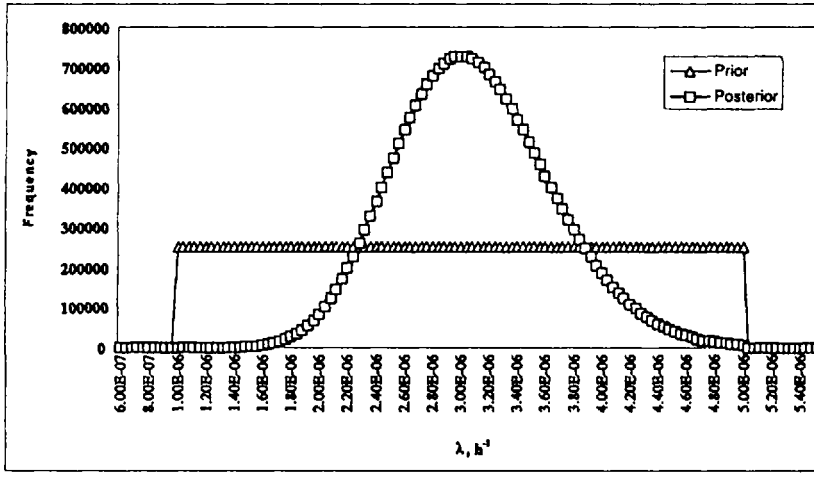


Figure 3.11 Prior and posterior distribution of λ in Example 3.23.

Solution:

Using (3.108–3.109), the point estimate $\lambda_B = 3.1 \times 10^{-6}$ 1/hr and the 90% two-sided Bayes' probability interval is $(2.24 \times 10^{-6} < \lambda < 4.04 \times 10^{-6})$ 1/hr. Figure 3.11 shows the prior and posterior distribution of λ.

3.6.2 Bayesian Estimation of the Parameter of Binomial Distribution

The binomial distribution plays an important role in reliability. Suppose that n identical units have been placed on test (without replacement of the failed units) for a specified time, t, and that the test yields r failures. The number of failures, r, can be considered as a discrete random variable having the binomial distribution with parameters n and $p(t)$, where $p(t)$ is the probability of failure of asingle unit during time t. As discussed in Section 3.5, $p(t)$, as a function of time, is the time to failure cumulative distribution function, as well as $1 - p(t)$ is the reliability or survivor function. A straightforward application of the binomial distribution is the modeling of a number of failures to start on demand for a redundant unit. The

probability of failure in this case might be considered as time independent. Thus, one should keep in mind two possible applications of the binomial distribution:

1. the survivor (reliability) function or time-to-failure cdf, and
2. the binomial distribution itself.

The maximum likelihood estimate of the parameter p is the ratio r/n, which is widely used as a classical estimate. To get a Bayesian estimation procedure for the reliability (survivor) function, let us consider p as the survivor probability in a single Bernoulli trial (so, now the "success" means surviving). If the number of units placed on test, n, is fixed in advance, the probability distribution of the number, x, of unfailed units during thetest (i.e., the number of "successes") is given by the binomial distribution with parameters n and x (see (2.27)):

$$f(x; n, p) = \frac{n!}{(n-x)!\,x!} p^{x}(1-x)^{n-x}$$

The corresponding likelihood function can be written as

$$l(p \mid x) = c p^{x}(1-x)^{n-x}$$

where c is a constant which does not depend on the parameter of interest, p. For any continuous prior distribution with pdf $h(p)$ the corresponding posterior pdf can be written as

$$f(p \mid x) = \frac{p^{x}(1-p)^{n-x} h(p)}{\int p^{x}(1-p)^{n-x} h(p)\,dp} \qquad (3.110)$$

Standard Uniform Prior Distribution

Consider the particular case of uniform distribution, $U(a,b)$, which in the Bayes' context represents "a state of total ignorance." While this seems to have little practical importance, nevertheless, it is interesting from the methodological point of view. For this case one can write

$$h(p) = \begin{cases} 1, & 0 < p \le 1 \\ 0, & \text{otherwise} \end{cases}$$

and

$$f(p|x) = \frac{p^{(x+1)-1}(1-p)^{(n-x+1)-1}}{\int_0^1 p^{(x+1)-1}(1-p)^{(n-x+1)-1}\,dp} \tag{3.111}$$

The integral in the denominator can be expressed as

$$\int_0^1 p^{(x+1)-1}(1-p)^{(n-x+1)}\,dp = \frac{\Gamma(x+1)\,\Gamma(n-x+1)}{\Gamma(n+2)}$$

So, the posterior cdf can be easily recognized as the pdf of the beta distribution, $f(p; x+1, n-x+1)$, which was introduced in Chapter 2. Recalling the expression for the mean value of the beta distribution (2.56), the point Bayes' estimate of p can be written as

$$p_B = \frac{x+1}{n+2} \tag{3.112}$$

Note that the estimate is different as compared with the respective classical estimate (x/n), but when the sample size increases the estimates are getting closer to each other.

Recalling that the cdf of the beta distribution is expressed in terms of the incomplete beta function (see Equation (3.85)), the $100(1-\alpha)\%$ two-sided Bayes' probability intervalfor p can be obtained by solving the following equations

$$\begin{aligned} \Pr(p < p_l) &= I_{p_l}(x+1, n-x+1) = \frac{\alpha}{2} \\ \Pr(p > p_u) &= I_{p_u}(x+1, n-x+1) = 1-\frac{\alpha}{2} \end{aligned} \tag{3.113}$$

It can be mentioned that the probability intervals above are very similar to the corresponding classical confidence intervals (3.83) and (3.84).

Example 3.24

Calculate the point estimate and the 95% two-sided Bayesian probability interval for the reliability of a new component based on the life test of 300 components, out of which 4 have failed. Suppose that for this component no historical information is available. Accordingly, its prior reliability estimate may be assumed to be uniformly distributed between 0 and 1.

Solution:

Using (3.112), find

$$R = 1 - p_B = 1 - \frac{4 + 1}{300 + 2} = 0.9834$$

Using (3.113), the 95% upper and lower limits are evaluated as 0.9663 and 0.9946, respectively.

It is interesting to compare the above results with classical ones. The point estimate of the reliability is $R = 1 - p = 1 - 4/300 = 0.9867$, and the 95% upper and lower limits, according to (3.81) and (3.82), are 0.9662 and 0.9964, respectively.

Truncated Standard Uniform Prior Distribution

Consider the following prior pdf of p

$$h(p \mid p_0, p_1) = \begin{cases} \dfrac{1}{p_1 - p_0}, & 0 \le p_0 < p < p_1 \le 1 \\ 0, & \text{otherwise} \end{cases} \tag{3.114}$$

The corresponding posterior pdf cannot be expressed in a closed form, but it can be written in terms of the incomplete beta function (Martz and Waller (1982)) as

$$f(p \mid x) = \frac{\dfrac{\Gamma(n+2)}{\Gamma(x+1)\Gamma(n-x+1)} p^{(x+1)-1}(1-p)^{(n-x+1)-1}}{I_{p_1}(x+1, n-x+1) - I_{p_0}(x+1, n-x+1)}, \tag{3.115}$$

$$0 \le p_0 < p \le 1$$

where the numerator is in form of a beta pdf with parameters $x + 1$ and $n - x + 1$. The same posterior pdf can be written in a simpler form which can be more convenient for straightforward point estimate calculations

$$f(p \mid x) = \frac{p^{(x+1)-1}(1-p)^{(n-x+1)-1}}{\displaystyle\int_{p_2}^{p_1} p^{(x+1)-1}(1-p)^{(n-x+1)-1}\, dp} \tag{3.116}$$

The posterior mean can be obtained in terms of the incomplete beta function as

$$p_B = \frac{x+1}{n+2}\left[\frac{I_{p_1}(x+2, n-x+1) - I_{p_0}(x+2, n-x+1)}{I_{p_1}(x+1, n-x+1) - I_{p_0}(x+1, n-x+1)}\right] \tag{3.117}$$

where the first multiplier coincides with the corresponding estimate for the case of the standard uniform prior, and the second one can be considered as a correction term associated with the truncated uniform prior distribution. The same estimate can be written in terms of the posterior pdf (3.116) as

$$p_B = \frac{\int_{p_0}^{p_1} p^{(x+1)}(1-p)^{(n-x+1)-1}\,dp}{\int_{p_0}^{p_1} p^{(x+1)-1}(1-p)^{(n-x+1)-1}\,dp} \tag{3.118}$$

Using the posterior pdf, the $100(1-\alpha)\%$ two-sided Bayes' probability interval for p can be obtained as solutions of the following equations

$$\begin{aligned} &I_{p_l}(x+1, n-x+1) \\ &= \left(1 - \frac{\alpha}{2}\right) I_{p_0}(x+1, n-x+1) + \frac{\alpha}{2} I_{p_1}(x+1, n-x+1) \\ &I_{p_u}(x+1, n-x+1) \\ &= \left(1 - \frac{\alpha}{2}\right) I_{p_1}(x+1, n-x+1) + \frac{\alpha}{2} I_{p_0}(x+1, n-x+1) \end{aligned} \tag{3.119}$$

Example 3.25

A new sensor installed on 500 vehicles was observed for 12 months in service (MIS) and 4 failures were recorded. The reliability of a similar sensor at 12 MIS has been known not to exceed 0.985, which can be expressed in terms of the uniform prior distribution with $p_0 = 0.985$, and $p_1 = 1$. Find the point estimate (posterior mean) and the 90% one-sided lower Bayes' probability interval for the reliability of the new sensor.

Solution:

According to (3.117), the posterior mean is 0.9926 and the 90% posterior lower limit from (3.119) is calculated as 0.9856. Figure 3.12 displays the prior and posterior distributions of $1 - p$.

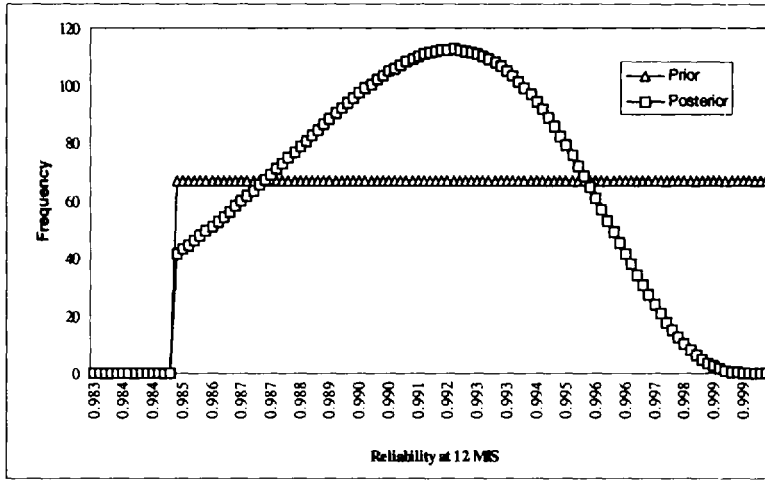


Figure 3.12 Prior and posterior distribution of $1 - p$ in Example 3.23.

Beta Prior Distribution

The most widely used prior distribution for the parameter, p, of the binomial distribution is the beta distribution which was introduced in Chapter 2. The pdf of the distribution can be written in the following convenient form:

$$h(b; x_0, n_0) = \begin{cases} \dfrac{\Gamma(n_0)}{\Gamma(x_0)\,\Gamma(n_0 - x_0)}\, p^{x_0 - 1}(1-p)^{n_0 - x_0 - 1} & 0 \le p \le 1 \\ 0, & \text{otherwise} \end{cases} \tag{3.120}$$

where $n_0 > x_0 > 0$.

The pdf provides a great variety of different shapes. It is important to note

that the standard uniform distribution is a particular case of the beta distribution. When x_0 is equal to one and n_0 is equal to two, (3.120) reduces to the standard uniform distribution. Moreover, the beta prior distribution turns out to be a conjugate prior distribution for the estimation of the parameter p of the binomial distribution of interest.

Considering the expression for the mean value of the beta distribution (2.56), it is clear that the prior mean is x_0/n_0, so that the parameters of the prior, x_0 and n_0, can be interpreted as a pseudo number of identical units survived (or failed) a pseudo test of n_0 units duringpseudo time t. Thus, while selecting the parameters of the prior distribution an expert can express his knowledge in terms of the pseudo test considered (i.e., in terms of x_0 and n_0). On the other hand, an expert can evaluate the prior mean, i.e., the ratio, x_0/n_0, and his/her degree of belief in terms of standard deviation or coefficient of variation of the prior distribution. For example, if the coefficient of variation is used, it can be treated as a measure of uncertainty (relative error) of prior assessment.

Let p_{Pr} be the prior mean and k be the coefficient of variation of the prior beta distribution. The corresponding parameters x_0 and n_0 can be found as a solution of the following equation system

$$\begin{aligned} p_{pr} &= \frac{x_0}{n_0}, \\ n_0 &= \frac{1 - p_{pr}}{k^2 p_{pr}} - 1 \end{aligned} \tag{3.121}$$

The prior distribution can also be estimated using test or field data collected for analogous products. In this case the parameters x_0 and n_0 are directly obtained from the tests or field data.

Example 3.26

Let the prior mean (point estimate) of the reliability function be chosen as $p_{Pr} = x_0/n_0 = 0.9$. Select the parameters x_0 and n_0.

Solution:

The choice of the parameters x_0 and n_0 can be (similar to one considered in Section 3.6.1) based on values of the coefficient of variation used as a measure of dispersion (accuracy) of the prior point estimate p_{Pr}. Some values of the coefficient of variation and the corresponding values of the parameters x_0 and n_0 for $p_{Pr} = x_0/n_0 = 0.9$ are given in the table below.

n_0	x_0	Coefficient of variation, %
1	0.9	23.6
9	10	10.0
90	100	3.3
900	1000	1.0

The posterior pdf is

$$f(p|x) = \frac{\Gamma(n+n_0)}{\Gamma(x+x_0)\Gamma(n+n_0-x-x_0)} p^{(x+x_0)-1}(1-p)^{(n+n_0-x-x_0)-1} \tag{3.122}$$

which is also a beta distribution pdf. The corresponding posterior mean is given by

$$p_B = \frac{x+x_0}{n+n_0} \tag{3.123}$$

Note that as n approaches infinity, the Bayesian estimate approaches the maximum likelihood estimate, x/n, (3.77). In other words, the classical inference tends to dominate the Bayes' inference as the amount of data increases.

One should also keep in mind that the prior distribution parameters can also be estimated based on prior data (data collected on similar equipment for example) which is straightforward using the respective sample size, n_0, and the number of failures observed, x_0.

It is easy to see that the corresponding $100(1-\alpha)\%$ two-sided Bayesian probability interval for p can be obtained as solutions of the following equations:

$$\begin{aligned} \Pr(p < p_l) &= I_{p_l}(x+x_0, n+n_0-x-x_0) = \frac{\alpha}{2} \\ \Pr(p > p_u) &= I_{p_u}(x+x_0, n+n_0-x-x_0) = 1-\frac{\alpha}{2} \end{aligned} \tag{3.124}$$

Example 3.27

A design engineer assesses the reliability of a new component at the end of its useful life (T = 10,000 hours) as 0.75 ± 0.19. A sample of 100 new components have been tested for 10,000 hours and 29 failures have been recorded. Given the test results, find the posterior mean and the 90% Bayesian probability interval for the component reliability, if the prior distribution of the component reliability is assumed to be a beta distribution.

Solution:

The prior mean is obviously 0.75 and the coefficient of variation is 0.19/0.75 = 0.25. Using (3.121), the parameters of the prior distribution are evaluated as x_0 = 3.15 and n_0 = 4.19. Thus, according to (3.123), the posterior point estimate of the new component reliability is $R(10{,}000) = (3.15 + 71)/(4.19 + 100) = 0.712$. According to (3.124), the 90% lower and upper confidence limits are 0.637 and 0.782, respectively. Figure 3.13 shows the prior and the posterior distributions of 1 - p.

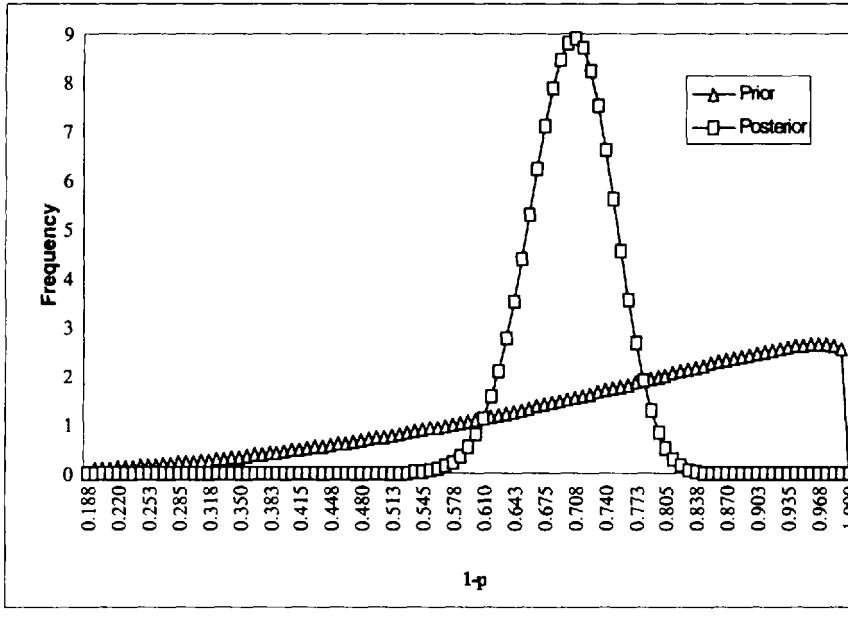


Figure 3.13 Prior and posterior distribution of 1 - p in Example 3.27.

Lognormal Prior Distribution

The following example illustrates the case when the prior distribution and the likelihood function do not result in a conjugate posterior distribution, and the posterior distribution obtained cannot be expressed in terms of standard function. This is the case when a numerical integration is required.

Example 3.28

The number of failures to start a diesel generator on demand has a binomial distribution with parameter p. The prior data on the performance of the similar diesel are obtained from field data, and p is assumed to follow the lognormal distribution with known parameters $\mu_s = 0.05$ and $\sigma_s = 0.04$ (the respective values of μ_t and σ_t are -3.22 and 0.51). A limited test of the diesel generators of interest shows that 8 failures are observed in 582 demands. Calculate the Bayesian point estimate of p (mean and median) and the 90th percentiles of p. Compare these results with corresponding values for the prior distribution.

Solution:

Since we are dealing with a demand failure, a binomial distribution best represents the observed data. The likelihood function is given by

$$\Pr(X \mid p) = \binom{582}{8} p^8 (1 - p)^{574}$$

and the prior pdf is

$$f(p) = \frac{1}{\sigma_t p \sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{\ln(p) - \mu_t}{\sigma_t} \right)^0 \right] \qquad p > 0$$

Using the initial data, the posterior pdf becomes

$$f(p \mid X) = \frac{p^7 (1 - p)^{574} \exp\left[-\frac{1}{2} \left(\frac{\ln(p) + 3.22}{0.51} \right)^2 \right]}{\int_0^1 p^7 (1 - p)^{574} \exp\left[-\frac{1}{2} \left(\frac{\ln(p) + 3.22}{0.51} \right)^2 dp \right]}.$$

Table 3.4 Results of a Numerical Integration in Example 3.28

i	Probability p_i	Prior pdf	Likelihood function	Prior* likelihood	Posterior pfd	Posterior cdf
0	1.23E - 08	0.00E + 00	1.54E - 28	0.00E + 00	0.00E + 00	0.00E + 00
1	1.78E - 03	3.68E 03	3.47E 03	1.28E 05	1.01E 03	1.80E 06
2	3.55E - 03	1.94E - 01	3.98E 02	7.72E 03	6.13E 01	1.09E 03
3	5.33E 03	1.15E + 00	1.08E - 01	1.24E 01	9.85E + 00	1.86E 02
4	7.11E 03	3.05E + 00	1.62E 01	4.93E 01	3.91E + 01	8.81E 02
5	8.89E - 03	5.53E + 00	1.76E - 01	9.72E 01	7.71E + 01	2.25E 01
6	1.07E 02	8.18E + 00	1.55E 01	1.27E + 00	1.01E + 02	4.04E 01
7	1.24E 02	1.07E + 01	1.19E - 01	1.27E + 00	1.01E + 02	5.83E 01
8	1.42E - 02	1.28E + 01	8.20E - 02	1.05E + 00	8.35E + 01	7.32E 01
9	1.60E 02	1.46E + 01	5.22E - 02	7.62E 01	6.05E + 01	8.39E 01
10	1.78E - 02	1.60E + 01	3.11E - 02	4.97E 01	3.94E + 01	9.09E 01
11	1.95E 02	1.69E + 01	1.76E - 02	2.99E - 01	2.37E + 01	9.51E 01
12	2.13E - 02	1.76E + 01	9.56E - 03	1.68E 01	1.33E + 01	9.75E 01
13	2.31E 02	1.79E + 01	5.00E - 03	8.96E 02	7.11E + 00	9.88E 01
14	2.49E 02	1.80E + 01	2.53E 03	4.57E 02	3.63E + 00	9.94E 01
15	2.67E - 02	1.79E + 01	1.25E - 03	2.24E 02	1.78E + 00	9.97E 01
16	2.84E 02	1.77E + 01	6.01E - 04	1.06E 02	8.44E 01	9.99E 01
17	3.02E - 02	1.73E + 01	2.83E - 04	4.90E 03	3.89E - 01	9.99E 01
18	3.20E - 02	1.69E + 01	1.31E - 04	2.21E 03	1.75E 01	1.00E + 00
19	3.38E - 02	1.64E + 01	5.93E - 05	9.70E 04	7.70E 02	1.00E + 00
20	3.55E - 02	1.58E + 01	2.65E - 05	4.18E 04	3.32E 02	1.00E + 00
21	3.73E - 02	1.52E + 01	1.17E - 05	1.77E - 04	1.41E 02	1.00E + 00
22	3.91E - 02	1.46E + 01	5.07E 06	7.39E - 05	5.87E 03	1.00E + 00
23	4.09E - 02	1.40E + 01	2.17E - 06	3.04E 05	2.41E - 03	1.00E + 00
24	4.27E 02	1.34E + 01	9.22E - 07	1.23E - 05	9.77E 04	1.00E + 00
25	4.44E - 02	1.27E + 01	3.87E - 07	4.93E 06	3.91E 04	1.00E + 00
...	...	...	...	...	...	..
91	1.62E - 01	5.19E 01	3.77E - 37	1.96E 37	1.55E - 35	1.00E + 00
92	1.63E 01	4.98E - 01	1.17E - 37	5.83E 38	4.63E 36	1.00E + 00
93	1.65E - 01	4.78E 01	3.62E - 38	1.73E 38	1.37E 36	1.00E + 00
94	1.67E 01	4.59E - 01	1.12E 38	5.13E - 39	4.07E 37	1.00E + 00
95	1.69E 01	4.41E - 01	3.43E - 39	1.51E 39	1.20E 37	1.00E + 00
96	1.71E 01	4.23E - 01	1.05E - 39	4.46E 40	3.54E 38	1.00E + 00
97	1.72E 01	4.07E 01	3.22E 40	1.31E 40	1.04E 38	1.00E + 00
98	1.74E 01	3.91E 01	9.79E - 41	3.83E 41	3.04E 39	1.00E + 00
99	1.76E 01	3.76E - 01	2.97E - 41	1.12E 41	8.86E 40	1.00E + 00
100	1.78E 01	3.61E 01	8.99E - 42	3.25E 42	2.58E 40	1.00E + 00
			Sum	7.09E + 00		

It is evident that the denominator cannot be expressed in a closed form, so a numerical integration must be applied. Table 3.4 shows results of a numerical integration used to find the posterior distribution. In this table the values of p_i are arbitrarily selected between 1.23 E-8 and 1.78 E-1. Then the numerator and denominator of the Posterior Pdf is calculated. The comparison of the prior and posterior is given below. Figure 3.14 displays the prior and the posterior distributions of p.

	Prior	Posterior
Mean	0.0516	0.0130
Median	0.0399	0.0121
5th Percentile	0.0123	0.0064
95th Percentile	0.1293	0.0197

The point estimate of the actual data using the classical inference is

$$\hat{p} = \frac{8}{582} = 0.0137$$

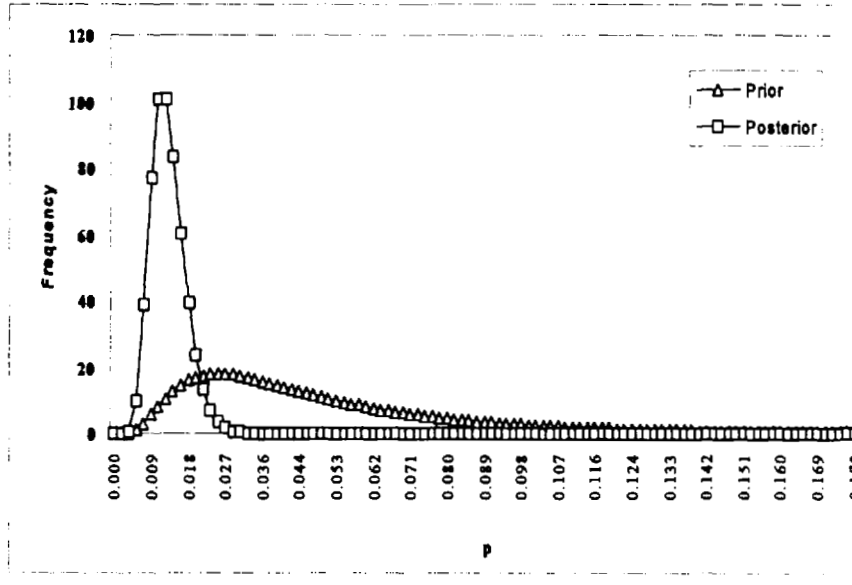


Figure 3.14 Prior and posterior distribution of p in Example 3.28.

 See the software supplement for the automated Bayesian estimation of both conjugate and nonconjugate distributions.

3.7 METHODS OF GENERIC FAILURE RATE DETERMINATION

Due to the lack of observed data, component reliability determination may requireuse of generic failure data adjusted for the various factors that influence the failure rate for the component under analysis. Generally, these factors are:

1. *Environmental Factors* — These factors affect the failure rate due to extreme mechanical, electrical, nuclear, and chemical environments. For example, a high-vibration environment, would lead to high stresses that promote failure of components.
2. *Design Factors* — These factors affect the failure rate due to the quality of material used and workmanship, material composition, functional requirements, geometry, and complexity.
3. *Operating Factors* — These factors affect the failure rate due to the applied stresses resulting from operation, testing, repair, and maintenance practices, etc.

To a lesser extent, the *age factor* is used to correct for early and wear-out periods, and *original factor* is used to correct for the accuracy of the data source (generic data). For example, obtaining data from observed failure records as opposed to expert judgement may affect the failure rate dependability.

Accordingly, the failure rate can be represented as

$$\lambda_a = \lambda_g K_E K_D K_O \ldots, \tag{3.125}$$

where λ_a is the actual failure rate and λ_g is the generic base failure rate, and K_E, K_D, and K_O are correction factors for the environment, design, and operation, respectively. It is possible to subdivide each of the correction factors to their contributing subfunctions accordingly. For example, $K_E = f(k_a, k_b, \ldots)$, when k_a and k_b are factors such as vibration level, moisture, and pH level. These factors may be different for different types of components.

This concept is used in the procedure specified in government contracts for determining the actual failure rate of electronic components. The procedure is summarized in MIL-HDBK-217. In this procedure, a base failure rate of the component is obtained from a table, and then they are multiplied by the applicable adjusting factors for each type of component. For example, the actual failure rate of a tantalum electrolytic capacitor is given by

$$\lambda_P = \lambda_b (\pi_E \cdot \pi_{SR} \cdot \pi_Q \cdot \pi_{CV}) \tag{3.126}$$

where λ_P is the actual component failure rate and λ_b is the base (or generic) failure rate, and the π factors are adjusting factors for the environment, series resistance, quality, and capacitance factors. Values of λ_b and the factors are given in MIL-HDBK-217 for many types of electrical and electronic components. Generally, λ_b is obtained from an empirical model called the Arrhenius model

$$\lambda_b = K \exp(-E/kT)$$

where: E = activation energy for the process, $k = 1.38 \times 0^{-23}$ J · K ^{1}, T = absolute temperature (°K), K = a constant.

The Arrhenius model forms the basis for a large portion of electronic components described in MIL-HDBK-217. However, care must be applied in using this database, especially because the data in this handbook are derived from repairable systems (and hence, apply to such systems). Also, application of the various adjusting factors can drastically affect the actual failure rates. Therefore, proper care must be applied to ensure correct use of the factors and to verify the adequacy of the factors suggested (Pecht (1995)). Also the appropriateness of the Arhenius model has been debated many times in the literature. The statistical procedures for fitting the Arrhenius model and other reliability models with explanatory factors are considered in the accelerated life testing section (see Chapter 7, Section 7.1).

For other types of components, many different generic sources of data are available. Among them are IEEE-500 (1984), Guidelines for Process Equipment Data (1989), Nuclear Power Plant, and Probability Risk Assessment (PRA) data sources. For example, Table B.1 (in Appendix B) shows a set of data obtained from NUREG/CR-4550 (1990).

EXERCISES

3.1 For a gamma distribution with the scale parameter of 400, and the shape parameter of 3.8, determine Pr(x < 200).

3.2 Time to failure of a relay follows a Weibull distribution with α = 10 years. β = 0.5.

Find the following:

a) Pr (failure after 1 year)
b) Pr (failure after 10 years)
c) The MTTF

3.3 The hazard rate of a device is $h(t) = 1/\sqrt{t}$. Find the following:

a) Probability density function
b) Reliability function
c) MTTF
d) Variance

3.4 Assume that 100 components are placed on test for 1000 hours. From previous testing, we believe that the hazard rate is constant, and the MTTF = 500 hours. Estimate the number of components that will fail in the time interval of 100 to 200 hours. How many components will fail if it is known that 15 components failed in $T < 100$ hours?

3.5 Assume that t, the random variable that denotes life in hours of a specified component, has a cumulative density function (cdf) of

$$F(t) = \begin{cases} 1 - \dfrac{100}{t}, & t \geq 100 \\ 0, & t < 100 \end{cases}$$

Determine the following:

a) pdf $f(t)$
b) Reliability function $R(t)$
c) MTTF

3.6 Show whether a uniform distribution represents an increasing failure rate, decreasing failure rate, or constant failure rate.

3.7 Consider the Rayleigh distribution:

$$f(t) = \frac{2t}{\alpha^2} \exp\left[-\frac{t^0}{\alpha^2}\right] \qquad t \geq 0, \quad \alpha > 0$$

a) Find the hazard rate $h(t)$ corresponding to this distribution.
b) Find the Reliability function $R(t)$.
c) Find the MTTF. $\left(\text{Notice: } \displaystyle\int_0^\infty \exp[-ax^2] = \frac{1}{2}\sqrt{\frac{\pi}{a}} \right)$
d) For which part of the bathtub curve is this distribution adequate?

3.8 Due to the aging process, the failure rate of a nonrepairable (i.e., replaceable) item is increasing according to $\lambda(t) = \lambda\beta t^{\beta-1}$. Assume that the value of λ and β are estimated as $\hat{\beta} = 1.62$ and $\hat{\lambda} = 1.2 \times 10^{-5}$ hour. Determine the probability that the item will fail sometime between 100 and 200 hours. Assume an operation beginning immediately after the onset of aging.

3.9 Suppose r.v. X has the exponential pdf $f(x) = \lambda \exp[-\lambda x]$, for $x > 0$, and $f(x) = 0$, for $x \leq 0$. Find $\Pr(x > a + b \mid x > a)$ given $a, b > 0$.

3.10 The following time to failure data are found when 158 transformer units are put under test. Use a nonparametric method to estimate $f(t)$, $h(t)$, and $R(t)$ of the transformers. No failures are observed prior to 1750 hours.

Age range (hr.)	No. of failures
1750 2250	17
2250 2750	54
2750 3250	27
3250 3750	17
3750 4250	19
4250 4750	24

3.11 A test was run on 10 electric motors under high temperature. The test was run for 60 hours, during which six motors failed. The failures occurred at the following times: 37.5, 46.0, 48.0, 51.5, 53.0, and 54.5 hours. We don't know whether an exponential distribution or a Weibull distribution model is better for representing these data. Use the plotting method as the main tool to discuss the appropriateness of these two models.

3.12 A test of 25 integrated circuits over 500 hours yields the following data:

Time interval	No. of failures in each interval
0 100	10
100 200	7
200 300	3
300 400	3
400 500	2

Plot the pdf, hazard rate, and reliability function for each interval of these integrated circuits using a nonparametric method.

3.13 Total test time of a device is 50,000 hours. The test is terminated after the first failure. If the pdf of the device time-to-failure is known to be exponentially distributed, what is the probability that the estimated failure rate is not greater than 4.6×10^{-5} (hrs^{-1}).

3.14 A manufacturer uses exponential distribution to model number "cycle- to-failure" of its products. In this case, r.v. T in the exponential pdf represents the number of cycles to failure. $\lambda = 0.003$ f/cycle.

a) What is the mean number of cycles to failure for this product?
b) If a component survives for 300 cycles, what is the probability that it will fail sometime after 500 cycles? Accordingly, if 1000 components have survived 300 cycles, how many would one expect to fail after 500 cycles?

3.15 The shaft diameters in a sample of 25 shafts are measured. The sample mean of diameter is 0.102 m, with a standard deviation of 0.005 m. What is the upper 95% confidence limit on the mean diameter of all shafts produced by this process, assuming the distribution of shaft diameters is normal?

3.16 The sample mean life of 10 car batteries is 102.5 months, with the standard deviation of 9.45 months. What are the 80% confidence limits for the mean and standard deviation of a pdf that represents these batteries?

3.17 The breaking strength X of 5 specimens of a rope of 1/4 inch diameter are 660, 460, 540, 580, and 550 lbs. Estimate the following:

a) The mean breaking strength by a 95% confidence level assuming normally distributed strength.
b) The point estimate of strength value at which only 5% of such specimens would be expected to break if $\bar{x}$ is assumed to be an unbiased estimate of the true mean, and s^2 is assumed to be the true standard deviation. (Assume x is normally distributed.)
c) The 90% confidence interval of the estimate of the standard deviation.

3.18 One hundred and twenty four devices are placed on an overstress test with failures occurring at the following times.

Time (hours)	Total no. of failures	Time (hours)	Total no. of failures
0.4	1	8.0	20
1.0	3	12.0	30
2.0	5	25.0	50
5.0	15		

a) Plot the data on Weibull probability paper.
b) Estimate the shape parameter.
c) Estimate the scale parameter.
d) What other distributions may also represent these failure data?

3.19 Seven pumps have failure times (in months) of 15.1, 10.7, 8.8, 11.3, 12.6, 14.4, and 8.7. (Assume an exponential distribution.)

a) Find a point estimate of the MTTF.
b) Estimate the reliability of a pump for $t = 12$ months.
c) Calculate the 95% two-sided interval of λ.

3.20 The average life of a certain type of small motor is 10 years, with a standard deviation of 2 years. The manufacturer replaces free of charge all motors that fail while under warranty. If the manufacturer is willing to replace only 3% of the motors that fail, what warranty period should be offered? Assume the time to failure of the motors follows a normal distribution.

3.21 A manufacturer claims that certain machine parts will have a mean diameter of 4 cm, with a standard deviation of 0.01 mm. The diameters of five parts are measured and found to be (in mm): 39.98, 40.01, 39.96, 40.03, and 40.02. Would you accept this claim with a 90% confidence level?

3.22 You are to design a life test experiment to estimate the failure rate of a new device. Your boss asks you to make sure that the 80% upper and lower limits of the estimate interval (two-sided) do not differ by more than a factor of 2. Due to cost constraints, the components will be tested until they fail. Determine how many components should be put on test.

3.23 For an experiment, 25 relays are allowed to run until the first failure occurred at $t = 15$ hours. At this point, the experimenters decide to continue the test for another 5 hours. No failures occur during this extended period, and the test is terminated. Using the 90% confidence level, determine the following:

a) Point estimate of MTTF.
b) Two-sided confidence interval for MTTF.
c) Two-sided confidence interval for reliability at $t = 25$ hours.

3.24 A locomotive control system fails 15 times out of the 96 times it is activated to function. Determine the following:

a) A point estimate for failure probability of the system.
b) 95% two-sided confidence intervals for the probability of failure. (Assume that after each failure, the system is repaired and put back in an as-good-as-new state.)

3.25 A sample of 10 measurements of a sphere diameter gives a mean of 4.38 inches, with a standard deviation of 0.06 inch. Find the 99% confidence limits of the actual mean and standard deviation.

3.26 The following sample of measurements is taken from a study of an industrial process, which is assumed to follow a normal distribution: 8.9, 9.8, 10.8, 10.7, 11.0, 8.0, and 10.8. For this sample, the 95% confidence error on estimating the mean (μ) is 2.2. What sample size should be taken if we want the 99% confidence error to be1.5, assuming the same sample variance?

3.27 Suppose the generic failure rate of a component corresponding to an exponential time to failure model is $\lambda_g = 10^{-3}$ (hr^{-1}) with a standard deviation of $\lambda_g/2$. Assume that ten components are closely observed for 1500 hours and one failure is observed. Using the Bayesian method, calculate the mean and variance of λ from the posterior distribution. Calculate the 90 percent lower confidence limit.

3.28 In the reactor safety study, the failure rate of a diesel generator can be described as having a lognormal distribution with the upper and lower 90% bounds of $3E-2$ and $3E-4$ respectively. If a given nuclear plant experiences 2 failures in 8760 hours of operation, determine the upper and lower 90% bounds given this plant experience. (Consider the reactor safety study values as prior information.)

3.29 Five measurements of the breaking strength of a computer board were recorded as 0.28, 0.30, 0.27, 0.33, 0.31 Kgf. Find the point estimate and the 99% confidence intervals for the actual mean breaking strength assuming the breaking strength is distributed exponentially.

3.30 The number of days in a 50-day period during which x failures of an assembly line is recorded as follows. Use a Chi - square goodness of fit test to determine whether a Poisson distribution is a good fit to these data. Perform the test at a 5% significance level.

Number of failures, x	0	1	2	3	4
Number of Days x failures observed	21	18	7	3	1

3.31 Fifty identical units of a manufactured product are tested for 300 hours, only one failure is observed (the failed unit is replaced with a good one).

a) Find an estimate of the failure rate of this unit.
b) Find the 90% confidence interval (two - sided) for the actual failure rate.

3.32 A mechanical life test of 18 circuit breakers of a new design was run to estimate the percentage failed by 10,000 cycle of operation. Breakers were inspected on a schedule, and it is known that failures occurred between certain inspections as shown,

Cycles (× 1000)	10-15	15-17.5	17.5-20	20-25	25-30	30+
Number of failures	2	3	1	1	2	9 survived

a) Make a Weibull plot of these data. Is this a good fit?
b) Graphically estimate percentage failing by 10,000 cycles.
c) Graphically estimate the Weibull distribution parameters.

3.33 Fifty-eight fans in service are supposed to have an exponential life distribution with an MTTF of 28,700 hours. Assuming that a failed fan is replaced with a new that does not fail, predict the number of such fans that will fail in 2000 hours.

3.34 A manufacturer tests 125 high-performance contacts and finds that 3 are defective.

a) Calculate the probability that a random contact is defective.
b) What is the 90% confidence interval for the estimated probability in (a)?

3.35 If the time-to-failure pdf of a component follows a linear model as follows,

$$f(t) = ct \qquad c < t < 10{,}000$$

$$= 0 \qquad \text{otherwise}$$

Determine:

a) Reliability function.
b) Failure rate function.

3.36 The cycle-to-failure T for a certain kind of component has the instantaneous failure rate $\lambda(t) = 2.5 \times 10^{-5} t^2$, ≥ 0 (cycles $^{-1}$). Find the MCTF (mean-cycle-to-failure), and the reliability of this component at 100 cycles.

3.37 The following data were collected by Frank Proschan in 1983. Operating hours to first failure of an engine cooling part in 13 aircrafts are:

Aircraft	1	2	3	4	5	6	7	8	9	10	11	12	13
Hours	194	413	90	74	55	23	97	50	359	50	130	487	102

a) Would these data support an increasing failure rate, decreasing failure rate or constant failure rate assumption?
b) Based on a graphic nonparametric analysis of these data, confirm the results obtained in part (a).

3.38 The following times-to-failure in hours were observed in an experiment where 14 units were tested until eight of them have failed:

80, 310, 350, 470, 650, 900, 1100, 1530

Assuming that the units have a constant failure rate, calculate a point estimate of the failure rate. Also calculate a 95% one-sided confidence interval of the failure rate.

3.39 A life test of 10 small motors with a newly designed insulator has been performed. The following data are obtained:

Motor No.	1	2	3	4	5	6	7	8	9	10
Failure Time (hr)	1175	1200	1400	1450	1580	1870	1930	2120	2180	2430

a) Make a Weibull plot of these data and estimate the parameters.
b) Estimate the motor reliability after 6 months of continuous operation.

3.40 Use the data in problem 3.39 to perform a total-time-on-test plot.

3.41 A company redesigns one of its compressors and wants to estimate reliability of the new product. Using past experience, the company believes that the reliability of the new compressor will be higher than 0.5 (for a given mission time). The company's testing of one new compressor showed that the product successfully achieved its mission.

a) Assuming a uniform prior distribution for the above reliability estimate, find the posterior estimate of reliability based on the test data.
b) If the company conducted another test, which resulted in another mission success, what would be the new estimate of the product reliability?

REFERENCES

Bain, L. J. , "*Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods,*" Marcel Dekker, New York, 1978.

Barlow, R. E., "*Analysis of Retrospective Failure Data using Computer Graphics,*" Proceedings of the 1978 Annual Reliability and Maintainability Symposium, pp. 113-116, 1978.

Barlow, R. E. and Campo R. A., "*Total Time on Test Processes and Applications to Failure Data Analysis, Reliability and Fault Tree Analysis,*" eds. Barlow, Fussell and Singpurwalla, SIAM, Philadelphia, pp. 451 481, 1975.

Barlow, R. E. and Proschan, F., "*Statistical Theory of Reliability and Life Testing: Probability Models,*" To Begin With, Silver Spring, MD, 1981.

Blom, G., "*Statistical Estimates and Transformed beta Variables,*" John Wiley and Sons, New York, 1958.

Castillo, E., "*Extreme Value Theory in Engineering*," Academy Press, San Diego, CA. 1988.

Davis, "*An Analysis of Some Failure Data*," J. Am. Stat. Assoc., 47, pp. 113-150, 1952.

Epstein, B., "*Estimation from Life Test Data*," Technometrics, 2, 447, 1960.

Fisher, R. A. and Tippet, L. H. C., "*Limiting Forms of the Frequency Distributions of the Largest or Smallest Member of a Sample*," Proc. Cambridge Philos. Soc., 24, pp. 180-190, 1928.

Frechet, M., "*Sur la loi de probabilite de l'ecart maximum*," Ann. Soc. Polon. Math, Cracow, 6, p. 93, 1927.

Gnedenko, B. V., "*Limit Theorems for the Maximal Term of a Variational Series*," Comptes Rendus de l'Academie des Sciences de l'URSS, 32, pp. 7-9, 1941.

Gumble, E. J., "*Statistics of Extremes*," Columbia University Press, New York, 1958.

Hahn, G. J. and S. S. Shapiro, "*Statistical Models in Engineering*," John Wiley and Sons, New York, NY, 1967.

IEEE Std. 500, "*Guide to the Collection and Presentation of Electrical, Electronic, Sensing Component and Mechanical Equipment Reliability Data for Nuclear Power Generating Stations*," IEEE Standards, New York, NY, 1984.

Johnson, N. L. and Kotz S., "*Distributions in Statistics*," John Wiley and Sons, New York, NY,1970.

Kapur, K. C. and Lamberson, L. R., "*Reliability in Engineering Design*," John Wiley and Sons, New York, NY, 1977.

Kececioglu, D., "*Reliability Engineering Handbook*," Prentice Hall, New Jersey, 1991.

Kimbal, "*On the Choice of Plotting Position on Reliability Paper*," J. Amer. Stat. Assoc. 55, pp. 546-560, 1960.

Lawless, J. F., "*Statistical Models and Methods for Lifetime Data*," John Wiley and Sons, New York, 1982.

Lewis, L. M., "*Reliability: Probabilistic Models and Statistical Methods*," Prentice-Hall, Englewood Cliffs, New Jersey, 1995.

Mann, N. R. E., Schafer, R. E. and Singpurwalla, N. D., "*Methods for Statistical Analysis of Reliability and Life Data*," John Wiley and Sons, New York, 1974.

Martz, H. F. and R. A. Waller, "*Bayesian Reliability Analysis*," John Wiley and Sons, New York, 1982.

MIL- HDBK-217F, Notice #2, "*Military Handbook, Reliability Prediction of Electronic Equipment*," 1995.

Center for Chemical Process Safety of the American Institute of Chemical Engineer, "*Guidelines for Process Equipment Data*," New York, 1989.

Nelson, W., "*Applied Life Data Analysis*," John Wiley and Sons, New York, 1982.

Nelson, W., "*How to Analyze Data with Simple Plots*," ASQC Basic Reference in Quality Control: Statistical Techniques, Am. Soc. Quality Control, Milwaukee, WI, 1979.

NUREG/CR-4450, "*Analysis of Core Damage Frequency From Internal Events*," Vol. 1, U.S. Nuclear Regulatory Commission, Washington, DC, 1990.

O'Connor, P. D. T., "*Practical Reliability Engineering*," 3rd ed., John Wiley and Sons, New York, 1996.

Pecht, M., "*Product Reliability, Maintainability, and Supportability Handbook*," CRC Press Inc., Boca Raton, FL, 1995.

Provan, J. W., "*Probabilistic Approaches to the Material-Related Reliability of Fracture-Sensitive Structures, in Probabilistic Fracture Mechanics and Reliability*," Provan, J. W., ed., Martinus Nijhoff Publishers, Dordrecht, The Netherlands, 1987.

Welker, E. L. and Lipow M., "*Estimating The Exponential Failure Rate Dormant Data with No Failure Events*," Proc. Rel. Maint. Symp., Vol. 1 (2), p. 1194, 1974.

4
System Reliability Analysis

Assessment of the reliability of a system from its basic elements is one of the most important aspects of reliability analysis. A system is a collection of items (subsystems, components, software, human operators, etc.) whose proper, coordinated operation leads to the proper functioning of the system. In reliability analysis, it is therefore important to model the relationship between various items as well as the reliability of the individual items to determine the reliability of the system as a whole. In Chapter 3, we elaborated on the reliability analysis at a basic item level (one for which enough information is available to predict its reliability). In this chapter, we discuss methods to model the relationship between system components, which allow us to determine overall system reliability.

The physical configuration of an item that belongs to a system is often used to model system reliability. In some cases, the manner in which an item fails is important for system failure and should be considered in the system reliability analysis. For example, in a system composed of two parallel electronic units, if a unit *fails* short, the system will fail, but for most other types of failures of the unit, the system will still be functional since the other unit works properly.

There are several system modeling schemes for reliability analysis. In this chapter we describe the following modeling schemes: *reliability block diagram*, which includes parallel, series, standby, shared load, and complex systems; *fault tree and success tree methods*, which include the method of construction and evaluation of the tree; *event tree method*, which includes modeling of multi-system designs and complex systems whose individual units should work in a chronological or approximately chronological manner to achieve a mission; *failure mode and effect analysis*; and *master logic diagram analysis*. We assume here that items composing a system are statistically independent (according to the definition provided in Chapter 2). In Chapter 7, we will elaborate on system reliability considerations when components are statistically dependent.

4.1 RELIABILITY BLOCK DIAGRAM METHOD

Reliability block diagrams are frequently used to model the effect of item failures on system performance. It often corresponds to the physical arrangement of items in the system. However, in certain cases, it may be different. For instance, when two resistors are in parallel, the system fails if one fails short. Therefore, the reliability block diagram of this system for the "fail short" mode of failure would be composed of two series blocks. However, for other modes of failure of one unit, such as "open" failure mode, the reliability block diagram is composed of two parallel blocks. In the remainder of this section, we discuss the reliability of the system for several types of the system functional configurations. A block represents one or a collection of some basic parts of the system for which reliability data are available.

4.1.1 Series System

A reliability block diagram is in a series configuration when failure of any one block (according to the failure mode of each item based on which the reliability block diagram is developed) results in the failure of the system. Accordingly, for functional success of a series system, all of its blocks (items) must successfully function during the intended mission time of the system. Figure 4.1 shows the reliability block diagram of a series system consisting of N blocks.

Figure 4.1 Series system reliability block diagram.

The reliability of the system in Figure 4.1 is the probability that all N blocks succeed during its intended mission time t. Thus, probabilistically, the system reliability $R_s(t)$ for independent blocks is obtained from

$$R_s(t) = R_1(t) \cdot R_2(t) \cdots R_N(t) = \prod_{i=1}^{n} R_i(t) \tag{4.1}$$

where $R_i(t)$ represents the reliability of the ith block. The hazard rate (instantaneous failure rate) for a series system is also a convenient expression. Since $H(t) = -d\{\ln R(t)\}/dt$, according to (4.1), the hazard rate of the system, $h_s(t)$ is

$$\lambda_s(t) = \frac{-d \ln \prod_{i=1}^{N} R_i(t)}{dt} = \sum_{i=1}^{N} \frac{-d \ln R_i(t)}{dt} = \sum_{i=1}^{N} \lambda_i(t) \tag{4.2}$$

Let's assume a constant hazard rate model for each block (e.g., assume an exponential time to failure for each block). Thus, $\lambda_i(t) = \lambda_i$. According to (4.2), the system failure rate is

$$\lambda_s = \sum_{i=1}^{N} \lambda_i \tag{4.3}$$

Expression (4.3) can also be easily obtained from (4.1) by using the constant failure rate reliability model for each block, $R_i(t) = \exp(-\lambda_i t)$.

$$R_s(t) = \prod_{i=1}^{N} \exp(-\lambda_i t) = \exp\left(-t \sum_{i=1}^{N} \lambda_i\right) = \exp(-\lambda_s t) \tag{4.4}$$

Using (4.2) and (4.3), the MTTF of the system can be obtained as follows:

$$\text{MTTF}_s = \frac{1}{\lambda_s} = \frac{1}{\sum_{i=1}^{N} \lambda_i} \tag{4.5}$$

Example 4.1

A system consists of three units whose reliability block diagram is in a series. The failure rate for each unit is constant as follows: $\lambda_1 = 4.0 \times 10^{-6}$ hr^{-1}, $\lambda_2 = 3.2 \times 10^{-6}$ hr^{-1}, and $\lambda_3 = 9.8 \times 10^{-6}$ hr^{-1}. Determine the following parameters of the system:

a. λ_s.
b. R_s (1000 hours).
c. MTTF_s.

Solution:

a. According to (4.3), $\lambda_s = 4.0 \times 10^{-6} + 3.2 \times 10^{-6} + 9.8 \times 10^{-6} = 1.7 \times 10^{-5}$ hr^{-1}.
b. $R_s(t) = \exp(-\lambda_s t) = \exp(-1.7 \times 10^{-5} \times 1000) = 0.983$, or unreliability of $\bar{R}(1000) = 0.017$.
c. According to (4.5), $\text{MTTF}_s = 1/\lambda_s = 1/(1.7 \times 10^{-5}) = 58{,}823.5$ hr.

4.1.2 Parallel Systems

In a parallel configuration, the failure of all blocks results in a system failure. Accordingly, success of only one block would be sufficient to guarantee the success of the system. Figure 4.2 shows a parallel system consisting of N blocks.

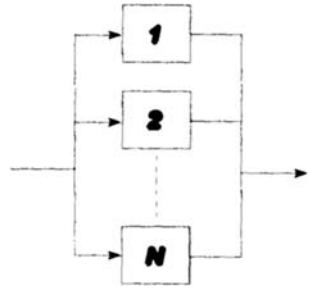


Figure 4.2 Parallel system block diagram.

For a set of N independent blocks,

$$F_s(t) = F_1(t) \cdot F_2(t) \cdots F_N(t) = \prod_{i=1}^{N} F_i(t) \tag{4.6}$$

Since $R_i(t) = 1 - F_i(t)$, then

$$R_s(t) = 1 - F_s(t) = 1 - \prod_{i=1}^{N} \left[1 - R_i(t)\right] \tag{4.7}$$

The system hazard rate can also be derived by using $h(t) = -d \ln R(t)/dt$.

For consideration of various characteristics of system reliability, let's analyze a special case where the failure rate is constant for each block (exponential time to failure model), and the system is composed of only two blocks. Since $R_i(t) = \exp(-\lambda_i t)$, then according to (4.7),

$$\begin{aligned} R_s &= 1 - [1 - \exp(-\lambda_1 t)]\,[1 - \exp(-\lambda_2 t)] \\ &= \exp(-\lambda_1 t) + \exp(-\lambda_2 t) - \exp[-(\lambda_1 + \lambda_2) t] \end{aligned} \tag{4.8}$$

Since $h_s(t) = \dfrac{f_s(t)}{R_s(t)}$ and $f_s(t) = \dfrac{d\left[R_s(t)\right]}{dt}$, then using (4.8),

$$f_s(t) = \lambda_1 \exp(-\lambda_1 t) + \lambda_2 \exp(-\lambda_2 t) - (\lambda_1 + \lambda_2) \exp[-(\lambda_1 + \lambda_2) t]$$

Thus,

$$h_s(t) = \frac{\lambda_1 \exp(-\lambda_1 t) + \lambda_2 \exp(-\lambda_2 t) - (\lambda_1 + \lambda_2) \exp[-(\lambda_1 + \lambda_2) t]}{\exp(-\lambda_1 t) + \exp(-\lambda_2 t) - \exp[-(\lambda_1 + \lambda_2) t]} \tag{4.9}$$

The MTTF of the system can also be obtained as

$$\begin{aligned} \text{MTTF}_S &= \int_0^\infty R_s(t)\, dt \\ &= \int_0^\infty \left[\exp(-\lambda_1 t) + \exp(-\lambda_2 t) - \exp[-(\lambda_1 + \lambda_2) t] \right] dt \\ &= \frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_1 + \lambda_2} \end{aligned} \tag{4.10}$$

Accordingly, one can use the binomial expansion to derive the MTTF for the system of N parallel blocks (units):

$$\begin{aligned} \text{MTTF}_S = &\left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \cdots + \frac{1}{\lambda_N} \right) \\ &- \left(\frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_1 + \lambda_3} + \cdots + \frac{1}{\lambda_{N-1} + \lambda_N} \right) \\ &+ \left(\frac{1}{\lambda_1 + \lambda_2 + \lambda_3} + \cdots + \frac{1}{\lambda_{N-2} + \lambda_{N-1} + \lambda_N} \right) \cdots \\ &+ \left(-1 \right)^{N+1} \frac{1}{\lambda_1 + \lambda_2 + \cdots + \lambda_N} \end{aligned} \tag{4.11}$$

In the special case where all units are identical with a constant failure rate λ (e.g., in an active redundant system), (4.7) simplifies to the following form:

$$R_s(t) = 1 - [1 - \exp(-\lambda t)]^N \tag{4.12}$$

and from (4.11),

$$\text{MTTF}_S = \text{MTTF} \left(1 + \frac{1}{2} + \cdots + \frac{1}{N} \right) \tag{4.13}$$

It can be seen from (4.13) that in the design of active redundant systems, the MTTF of the system exceeds the MTTF of an individual unit. However, the contribution to the MTTF of the system from the second unit, the third unit, and so on would have a diminishing return as N increases. That is, there would be an optimum number of parallel blocks (units) by which a designer can maximize the reliability and at the same time minimize the cost of the component in its life cycle.

Let's consider a more general structure of series and parallel systems: the so-called *K-out-of-N system*. In this type of system, if any combination of K units out of N independent units work, it guarantees the success of the system. For simplicity, assume that all units are identical (which, by the way, is often the case). The binomial distribution can easily represent the probability that the system functions:

$$\begin{aligned} R_s(t) &= \sum_{r=K}^{N} \binom{N}{r} [R(t)]^r [1 - R(t)]^{N-r} \\ &= \sum_{r=0}^{N-K} \binom{N}{r} [R(t)]^r [1 - R(t)]^{N-r} \end{aligned} \tag{4.14}$$

Example 4.2

A system is composed of the same units as in Example 4.1. However, these units are in parallel. Find the time-to-failure cdf (unreliability) and MTTFs of the system.

Solution:

According to (4.7),

$$R_s(t) = 1 - (1 - e^{-\lambda_1 t})(1 - e^{-\lambda_2 t})(1 - e^{-\lambda_3 t})$$

$$\begin{aligned} R_s(1000) &= (1 - e^{-4.0 \times 10^{-6} \times 1000})(1 - e^{-3.2 \times 10^{-6} \times 1000})(1 - e^{-9.8 \times 10^{-6} \times 1000}) \\ &= 1.25 \times 10^{-7} \end{aligned}$$

$$\begin{aligned} \text{MTTF}_s &= \left(\frac{1}{\lambda_1} + \frac{1}{\lambda_2} + \frac{1}{\lambda_3}\right) - \left(\frac{1}{\lambda_1 + \lambda_2} + \frac{1}{\lambda_1 + \lambda_3} + \frac{1}{\lambda_2 + \lambda_3}\right) \\ &\quad + \left(\frac{1}{\lambda_1 + \lambda_2 + \lambda_3}\right) = 4.35 \times 10^5 \text{ hours} \end{aligned}$$

Example 4.3

How many components should be used in an active redundancy design to achieve a reliability of 0.999 such that, for successful system operation, a minimum of two components is required? Assume a mission of $t = 720$ hours for a set of components that are identical and have a failure rate of 0.00015 hr^{-1}.

Solution:

For each component $R(t) = \exp(-\lambda t) = \exp(-0.00015 \times 720) = 0.8976$. According to (4.14),

$$0.999 = 1 - \sum_{r=0}^{1} \binom{N}{r} [0.8976]^r [0.1024]^{N-r}$$

$$= 1 - [0.1024]^N - N[0.8976][0.1024]^{N-1}$$

From the above equation, $N = 5$, which means that at least five components should be used to achieve the desired reliability over the specified mission time.

4.1.3 Standby Redundant Systems

A system is called a standby redundant system when some of its units remain idle until they are called for service by a sensing and switching device (SS). For simplicity, let's consider a situation where only one unit operates actively and the others are in standby, as shown in Figure 4.3.

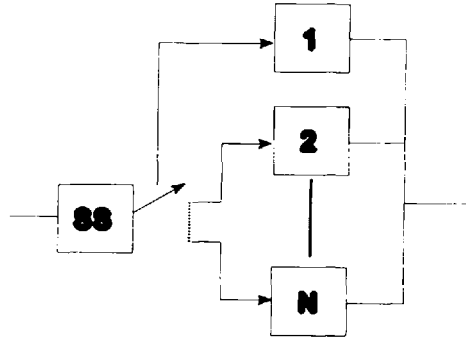


Figure 4.3 Standby redundant system.

In this configuration, unit 1 operates constantly until it fails. The sensing and switching device recognizes a unit failure in the system and switches to another unit. This process continues until all standby units have failed, in which case the system is considered failed. Since units 2 to N do not operate constantly (as is the case in active parallel systems), we would expect them to fail at a much slower rate. This is because the failure rate for components is usually lower when the components are operating than when they are idle or dormant.

It is clear that system reliability is totally dependent on the reliability of the sensing and switching device. The reliability of a redundant standby system is the reliability of unit 1 over the mission time t (i.e., the probability that it succeeds the whole mission time) plus the probability that unit 1 fails at time t_1 prior to t *and* the probability that the sensing and switching unit does not fail by t_1 *and* the probability that standby unit 2 does not fail by t_1 (in the standby mode) *and* the probability that standby unit 2 successfully functions for the remainder of the mission in an active operation mode, and so on.

Mathematically, the reliability function for a two block (unit) standby device according to this definition can be obtained as:

$$R_s(t) = R_1(t) + \int_0^{\infty} f_1(t_1)\,dt_1 \cdot R_{ss}(t_1) \cdot R'_2(t_1) \cdot R_2(t - t_1) \qquad (4.15)$$

where $f_1(t)$ is the pdf for the time to failure of unit 1, $R_{ss}(t_1)$ is the reliability of the sensing and switching device, $R'_2(t)$ is the reliability of unit 2 in the standby mode of operation, and $R_2(t\text{-}t_1)$ is the reliability of unit 2 after it started to operate at time t_1. Let's consider a case where time to failure of all units follows an Exponential distribution.

$$\begin{aligned} R_s(t) &= \exp(-\lambda_1 t) + \\ &\int_0^t \left[\lambda_1 \exp(-\lambda_1 t_1)\right]\left[\exp-(\lambda_{ss}\, t_1)\right] \\ &\left[\exp(-\lambda'_2 t_1)\right]\left\{\exp\left[-\lambda_2(t - t_1)\right]\right\} dt_1 \qquad (4.16) \\ &= \exp(-\lambda_1 t) + \frac{\lambda_1 \exp(-\lambda_2 t)}{\lambda_1 + \lambda_{ss} + \lambda'_2 - \lambda_2} \\ &\left\{1 - \exp\left[-(\lambda_1 + \lambda_{ss} + \lambda'_2 - \lambda_2)t\right]\right\} \end{aligned}$$

For the special case of perfect sensing and switching and no standby failures, $\lambda_{ss} = \lambda'_2 = 0$,

$$R_s(t) = \exp(-\lambda_1 t) + \frac{\lambda_1 \exp(-\lambda_2 t)}{\lambda_1 - \lambda_2} \left\{1 - \exp\left[-(\lambda_1 - \lambda_2)t\right]\right\}$$
$$= \exp(-\lambda_1 t) + \frac{\lambda_1}{\lambda_1 - \lambda_2} \left[\exp(-\lambda_2 t) - \exp(-\lambda_1 t)\right] \tag{4.17}$$

If the two units are identical, i.e., $\lambda_1 = \lambda_2 = \lambda$, then

$$R_s(t) = \exp(-\lambda t) + \lambda t \exp(-\lambda t) = (1 + \lambda t)\exp(-\lambda t) \tag{4.18}$$

In the case of perfect switching, a standby system possesses the same characteristic as the so called "shock model." That is one can assume that the Nth shock (i.e., the Nth unit failure) causes the system to fail. Thus, a gamma distribution can represent the time to failure of the system such that

$$R_s(t) = 1 - \int_0^t \frac{\lambda^N}{\Gamma(N)} x^{N-1} \exp(-\lambda x)\, dx$$
$$= \exp(-\lambda t)\left[1 + \lambda t + \frac{(\lambda t)^2}{2!} + \cdots + \frac{(\lambda t)^{N-1}}{(N-1)!}\right] \tag{4.19}$$

Accordingly, the MTTF of the above system is given by

$$\text{MTTF}_s = \frac{N}{\lambda} \tag{4.20}$$

which is N times the MTTF of a single unit. Expression (4.20) explains why high reliability can be achieved through a standby system when the switching is perfect and no failure occurs during standby.

When more than two units are in standby, the equation becomes somewhat difficult, but the concept is almost the same. For example, for three units with perfect switching,

$$R_s(t) = R_1(t) + \int_{t_1=0}^{t} f_1(t_1)\, dt_1 \cdot R_2(t - t_1)$$
$$+ \int_0^t f_1(t_1)\, dt_1 \int_0^{t-t_1} f_2(t_2)\, R_3(t - t_1 - t_2)\, dt_2 \tag{4.21}$$

If the sensing and switching devices are not perfect, appropriate terms should be added to (4.21) to account for their unreliability—similar to (4.15).

Example 4.4

Consider two identical independent units with $\lambda = 0.01$ hr^{-1}. Mission time $t = 24$ hours. Compare the reliability of a system made of these units if they are placed in:

a. Parallel configuration.
b. Series configuration.
c. Standby configuration with perfect switching.
d. Standby configuration with imperfect switching and standby failure rates of $\lambda_{sb} = 1 \times 10^{-6}$ and $\lambda' = 1 \times 10^{-5}$ hr^{-1} respectively.

Solution:

Let's assume an exponential time to failure model for each unit:

$$R(t) = \exp(-\lambda t) = \exp(-0.01 \times 24) = 0.7866$$

Then:

a. For the parallel system, using (4.12),

$$R_s(24) = 1 - (1 - 0.7866)^2 = 0.9544$$

b. For the series system, using (4.1),

$$R_s(24) = 0.7866 \times 0.7866 = 0.6187$$

c. For the standby system with perfect switches, using (4.18)

$$R_s(24) = (1 + 0.24)\exp(-0.01 \times 24) = 0.9755$$

d. For the standby system with imperfect switching and standby failure rate using (4.16),

$$R_s(24) = 0.7866 + \frac{(0.01)(0.7866)}{1.1 \times 10^{-5}}$$

$$\left[1 - \exp(-1.1 \times 10^{-5} \times 24)\right] = 0.9754$$

4.1.4 Load-Sharing Systems

A load-sharing system refers to a parallel system whose units equally share the system function. For example, if a set of two identical parallel pumps delivers x gpm of water to a reservoir, each pump delivers x/2 gpm. If a minimum of x gpm is required at all times, and one of the pumps fails at a given time t_o, then the other pump's speed should be increased to provide x gpm alone. Other examples of load sharing are multiple load-bearing units (such as those in a bridge), and load-sharing multi-unit electric power plants. In these cases, when one of the units fails, the others should carry its load. Since these other units would then be working under more stressful conditions, they would experience a higher rate of failure.

Load-sharing system reliability models can be divided in two groups—time-independent models and time-dependent ones. Note that most of the reliability models, discussed in this book are time-dependent. The time-independent reliability models are considered in the framework of, the so-called, Stress-Strength Analysis which is briefly discussed in Chapter 1. Historically first time-independent load-sharing system model was developed by Daniels (1945), and it is known as the Daniels model. This model was originally applied to textile strength problems and now it is also applied to composite materials.

To illustrate the basic ideas associated with these kinds of models, consider a simple parallel system composed of two identical components (Crowder, et al. (1991)). Let $F(s)$ be the time-independent failure probability for the component subjected to load (stress) s. Denote by $F_2(s)$ the failure probability for a parallel system of two identical blocks (units). The reliability function of the system, $R_2(s)$ is $1 - F_2(s)$. Initially, both components are subjected to an equal load s. When one unit fails, the nonfailed unit takes on the full load $2s$.

The probability of the system failure, $F_2(s)$, can be modeled as follows. Let A be the event when the first unit fails under load s and the second unit fails under load $2s$; let B be the event in which the second unit fails under load s *and* the first unit fails under load $2s$. Finally, let $A \cap B$ be the event that both units fail under load s

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$$

It is evident that

$$\Pr(A) = \Pr(B) = F(s)\,F(2s), \quad \Pr(A \cap B) = F^2(s)$$

hence

$$F_2(s) = 2F(s)\,F(2s) - F^2(s)$$

and

$$R_2(s) = 1 - 2F(s)\,F(2s) + F^2(s)$$

A similar equation for reliability of three component load-sharing system contains seven terms, and the problem gets more difficult as the number of components increases. For such situations different recursive procedures were developed (Crowder, et al. (1991)).

Now, consider a simple example of time-dependent load-sharing system model. Let's assume again that two components share a load (i.e., each component carries half the load), and the time-to-failure distribution for both components is $f_h(s,t)$. When one component fails (i.e., one component carries the full load), the time-to-failure distribution is $f_f(2s,t)$. Let's also assume that the corresponding reliability functions during full-load and half-load operation are $R_f(2s,t)$ and $R_h(s,t)$ respectively. The system will succeed if both components carry half the load, or if component 1 fails at time t_0 and component 2 carries a full load thereafter, or if component 2 fails at time t_0 and component 1 carries the full load thereafter. Accordingly, the system reliability function $R_s(t)$ can be obtained from (Kapur and Lamberson (1977))

$$R_s(t) = [R_h(s,t)]^2 + 2\int_0^t f_h(s,t_1)\,R_h(s,t_1)\,R_f(2s,\,t - t_1)\,dt_1 \qquad (4.22)$$

In (4.22), the first term shows the contribution from both components working successfully, with each carrying a half load; the second term represents the two equal probabilities that component 1 fails first and component 2 takes the full load at time t_0, or vice versa.

If there are switching or control mechanisms involved to shift the total load to the nonfailed component when one component fails, then similar to (4.15), the reliability of the switching mechanism can be incorporated into (4.22).

In the special situation where exponential time-to-failure models with failure rates λ_f and λ_h can be used for the two components under full and half loads, respectively, then (4.22) can be simplified to

$$R_s(t) = \exp(-2\lambda_h t) + \frac{2\lambda_h \exp(-\lambda_f t)}{(2\lambda_h - \lambda_f)} \left\{ \exp\left[-(2\lambda_h - \lambda_f)t \right] \right\} \qquad (4.23)$$

The reader is referred to (Crowder, et al. (1991)) for a review of more sophisticated time-dependent load-sharing models.

4.1.5 Complex Systems

Most practical systems are neither parallel, nor series, but exhibit some hybrid combination of the two. These systems are often referred to as *parallel-series system*. Figure 4.4 shows an example of such a system.

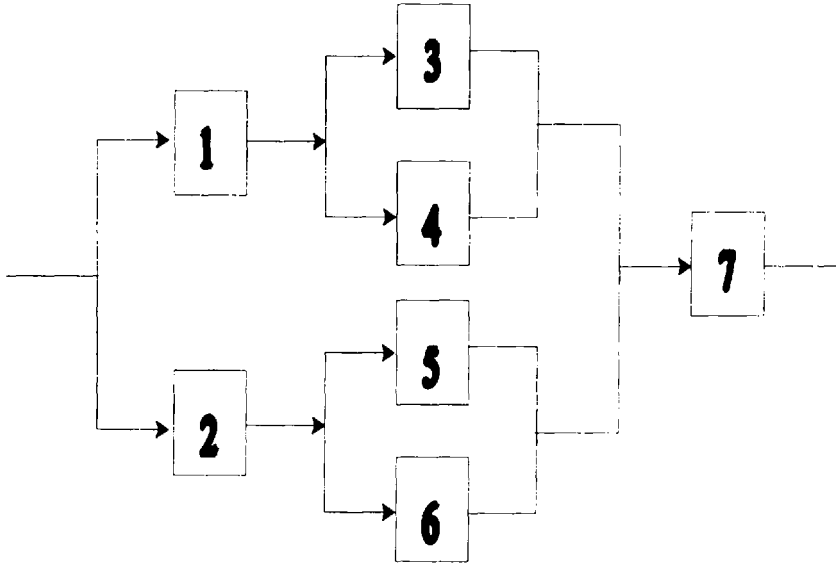


Figure 4.4 Complex parallel-series system.

Another type of complex system is one that is neither series nor parallel alone, nor parallel-series. Figure 4.5 shows an example of such a system.

A parallel-series system can be analyzed by dividing it into its basic parallel and series modules and then determining the reliability function for each module

separately. The process can be continued until a reliability function for the whole system is determined. For the analysis of all types of complex systems, Shooman (1990) describes several analytical methods for complex systems. These are the *inspection method, event space method, path-tracing method,* and *decomposition.* These methods are good only when there are not a lot of units in the system. For analysis of a large number of units, fault trees would be more appropriate. In the following, we discuss the decomposition and path-tracing methods.

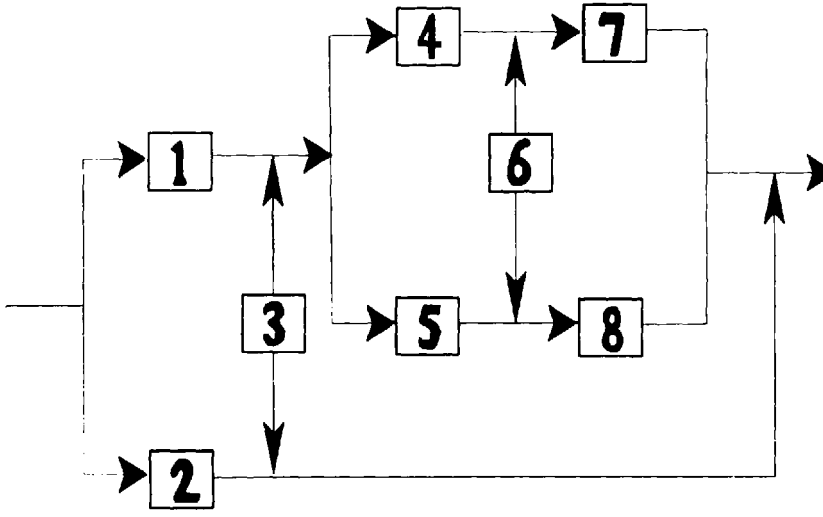


Figure 4.5 Complex nonparallel-series system.

The *decomposition method* relies on the conditional probability concept to decompose the system. The reliability of a system is equal to the reliability of the system given that a chosen unit (e.g., unit 3 in Figure 4.5) is good (i.e., working) times the reliability of unit 3, plus the reliability of the system given unit 3 is bad (i.e., failed) times the unreliability of unit 3.

$$R_s(t) = R_s(t \mid \text{unit 3 good}) \cdot R_3(t) + R_s(t \mid \text{unit 3 bad})[1 - R_3(t)] \quad (4.24)$$

If (4.24) is applied to all units that make the system a nonparallel series (such as units 3 and 6 in Figure 4.5), the system would reduce to a simple parallel-series

system. Thus, for Figure 4.5 and for the conditional reliability terms in (4.24), it follows that

$$R_s(t \mid \text{unit 3 good}) = R_s(t \mid \text{unit 6 good} \cap \text{unit 3 good})\, R_6(t) + R_s(t \mid \text{unit 6 bad} \cap \text{unit 3 bad})\,[1 - R_6(t)] \tag{4.25}$$

or

$$R_s(t \mid \text{unit 3 bad}) = R_s(t \mid \text{unit 6 good} \cap \text{unit 3 bad})\, R_6(t) + R_s(t \mid \text{unit 6 bad} \cap \text{unit 3 bad})\,[1 - R_6(t)] \tag{4.26}$$

Each of the conditional reliability terms in (4.25) and (4.26) represents a purely parallel-series system, the reliability determination of which is simple. For example, $R_s(t \mid \text{unit 6 good} \cap \text{unit 3 bad})$ corresponds to a reliability block diagram shown in Figure 4.6.

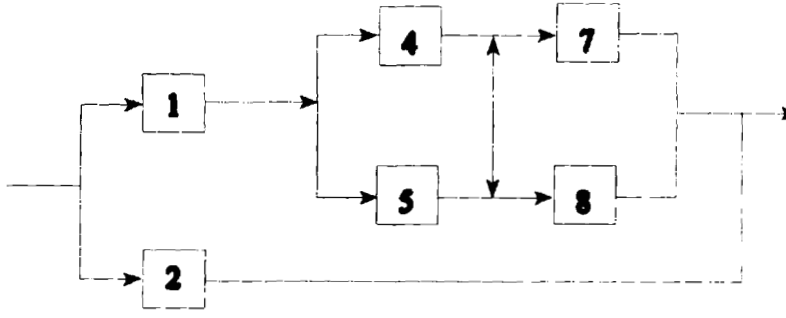


Figure 4.6 Representation of $R_s(t \mid \text{unit 6 good} \cap \text{unit 3 bad})$

The combination of (4.24) through (4.26) results in an expression for $R(s)$.

A more computationally intensive method for determining the reliability of a complex system involves the use of path set and cut set methods (path-tracing methods). A *path set* (or tie set) is a set of units that form a connection between input and output when traversed in the direction of the reliability block diagram

arrows. Thus, a path set merely represents a "path" through the graph. A *minimal path set* (or minimal tie set) is a path set containing the minimum number of units needed to guarantee a connection between the input and output points. For example, in Figure 4.5, path set $P_1 = (1, 3)$ is a minimal path set, but $P_2 = (1, 3, 6)$ is not since units 1 and 3 are sufficient to guarantee a path.

A *cut set* is a set of units that interrupt all possible connections between the input and output points. A *minimal cut set* is the smallest set of units needed to guarantee an interruption of flow. In practice, minimal cut sets show a combination of unit failures that cause a system to fail. For example, in Figure 4.5, the minimal path sets are: $P_1 = (2)$, $P_2 = (1, 3)$, $P_3 = (1, 4, 7)$, $P_4 = (1, 5, 8)$, $P_5 = (1, 4, 6, 8)$, $P_6 = (1, 5, 6, 7)$. The minimal cut sets are: $C_1 = (1, 2)$, $C_2 = (4, 5, 3, 2)$, $C_3 = (7, 8, 3, 2)$, $C_4 = (4, 6, 8, 3, 2)$, $C_5 = (5, 6, 7, 3, 2)$. If a system has m minimal path sets denoted by $P_1, P_2, \ldots, P_m$, then the system reliability is given by

$$R_s(t) = \Pr\left(P_1 \cup P_2 \cup \cdots \cup P_m\right) \tag{4.27}$$

where each path set P_i represents the event that units in the path set survive during the mission time t. This guarantees the success of the system. Since many path sets may exist, the union of all these sets gives all possible events for successful operation of the system. The probability of this union clearly represents the reliability of the system. It should be noted here that in practice, the path sets P_is are not disjointed. This poses a problem for determining the left-hand side of (4.27). In Section 4.2, we will explain formal methods to deal with this problem. However, an upper bound on the system reliability may be obtained by assuming that the P_is are highly disjointed. Thus,

$$R_s(t) \leq \Pr(P_1) + \Pr(P_2) + \cdots + \Pr(P_m) \tag{4.28}$$

Expression (4.28) yields better answers when we deal with small reliability values. Since this is not usually the case, (4.28) is not a good bound for use in practical applications.

Similarly, system reliability can be determined through minimal cut sets. If the system has n minimal cut sets denoted by $C_1, C_2, \ldots, C_n$, then the system reliability is obtained from

$$R_s(t) = 1 - \Pr\left(C_1 \cup C_2 \cup \cdots \cup C_n\right) \tag{4.29}$$

where C_i represents the event that units in the cut set fail sometime before the mission time t. This guarantees system failure. The Pr $(\cdot)$ term on the right hand

side of (4.29) shows the probability that at least one of all possible minimal cut sets exists before time t. Thus it represents the probability that the system fails sometimes before t. By subtracting this probability from 1, the reliability of the system is obtained. Similar to the union of path sets, the union of cut sets are not usually disjoint. Again, (4.29) can be written in the form of its lower bound, which is a much simpler expression given by

$$R_s(t) \geq 1 - \left[\Pr(C_1) + \Pr(C_2) + \cdots + \Pr(C_m) \right] \tag{4.30}$$

Notice that each element of a path set represents the *success* of a unit operation, whereas each element of a cut set represents the *failure* of a unit. Thus, for probabilistic evaluations, the reliability function of each unit should be used in connection with path set evaluations, i.e., (4.28), while the unreliability function should be used in connection with cut set evaluations, i.e., in (4.30).

The bounding technique used in (4.30), in practice, yields a much better representation of the reliability of the system than (4.28) because most engineering units have reliability greater than 0.9 over their mission time, making the use of (4.30) appropriate.

Example 4.5

Consider the reliability block diagram in Figure 4.5. Determine the lower bound of the system reliability function if the hazard rates of each unit are constant and are $\lambda_1, \lambda_2, \ldots, \lambda_5$.

Solution:

Using the system cut sets discussed earlier and (4.30),

$$R_s(t) \geq 1 - \left[\Pr(C_1) + \Pr(C_2) + \cdots + \Pr(C_5) \right]$$

assuming C_1 and C_2 are independent, and

$$\Pr(C_1) = \left[1 - \exp(-\lambda_1 t) \right] \left[1 - \exp(-\lambda_2 t) \right]$$

and so on. Therefore,

$$R_s(t) \geq 1 - \left[\left(1 - e^{-\lambda_1 t}\right)\left(1 - e^{-\lambda_2 t}\right) + \left(1 - e^{-\lambda_2 t}\right)\left(1 - e^{-\lambda_3 t}\right)\left(1 - e^{-\lambda_4 t}\right)\left(1 - e^{-\lambda_5 t}\right) + \left(1 - e^{-\lambda_2 t}\right)\left(1 - e^{-\lambda_3 t}\right)\left(1 - e^{-\lambda_7 t}\right)\left(1 - e^{-\lambda_8 t}\right) + \left(1 - e^{-\lambda_2 t}\right)\left(1 - e^{-\lambda_3 t}\right)\left(1 - e^{-\lambda_4 t}\right)\left(1 - e^{-\lambda_6 t}\right)\left(1 - e^{-\lambda_8 t}\right) + \left(1 - e^{-\lambda_2 t}\right)\left(1 - e^{-\lambda_3 t}\right)\left(1 - e^{-\lambda_5 t}\right)\left(1 - e^{-\lambda_6 t}\right)\left(1 - e^{-\lambda_7 t}\right)\right]$$

For some typical values of λ, the lower bound for $R_s(t)$ can be compared to the exact value of $R_s(t)$. Here, "exact" means the cut sets are not assumed disjoint. For example, Figure 4.7 shows the exact and the lower probability bound of system reliability for $\lambda_1 = 1 \times 10^{-6}$ hr^{-1}, $\lambda_2 = 1 \times 10^{-5}$ hr^{-1}, $\lambda_3 = 2 \times 10^{-5}$ hr^{-1}, and $\lambda_4 = \lambda_5 = \lambda_6 = \lambda_7 = \lambda_8 = 1 \times 10^{-4}$ hr^{-1}.

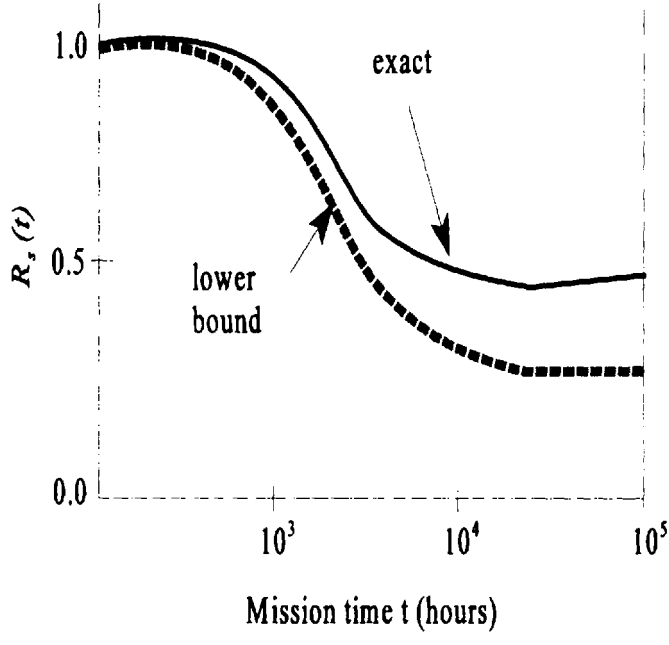


Figure 4.7 System reliability function in Example 4.5.

It is evident from Figure 4.7 that as time increases, the reliability of the system decreases (unit failure probability increases), causing (4.30) to yield a poor approximation. At this point, it is more appropriate to use (4.28). Again, notice that (4.28) and (4.30) assume the path sets and cut sets are disjoint.

In cases of very complex systems that have multiple failure modes for each unit and involved physical and operational interactions, the use of reliability block diagrams becomes difficult. The use of logic-based models such as fault tree and success tree analyses are more appropriate in this context. We will elaborate on this topic in the next section.

4.2 FAULT TREE AND SUCCESS TREE METHODS

The operation of a system can be considered from two opposite viewpoints: the various ways that a system fails, or the various ways that a system succeeds. Most of the construction and analysis methods used are, in principle, the same for both fault trees and success trees. First we will discuss the fault tree method, and then describe the success tree method.

4.2.1 Fault Tree Method

The fault tree approach is a deductive process by means of which an undesirable event, called the *top event*, is postulated, and the possible ways for this event to occur are systematically deduced. For example, a typical top event looks like "failure of control circuit A to send a signal when it should." The deduction process is performed so that the fault tree embodies all component failures that contribute to the occurrence of the top event. It is also possible to include individual failure modes of each component as well as human and software errors (and the relation between the two) during the system operation. The fault tree itself is a graphical representation of the various combinations of failures that lead to the occurrence of the top event.

A fault tree does not necessarily contain all possible failure modes of the components (or units) of the system. Only those failure modes which contribute to the existence occurrence of the top event are modeled. For example, consider a failed safe control circuit. If loss of the dc power to the circuit causes the circuit to open a contact, which in turn sends a signal to another system for operation, a top event of "control circuit fails to generate a safety signal" would not include the "failure of dc power source" as one of its events, even though the dc power source (e.g., batteries) is part of the control circuit. This is because the top event would not occur due to the loss of the dc power source.

The postulated fault events that appear on the fault tree structure may not be exhaustive. Only those events considered important can be included. However, it should be noted that the decision for inclusion of failure events is not arbitrary; it is influenced by the fault tree construction procedure, system design and operation, operating history, available failure data, and the experience of the analyst. At each intermediate point, the postulated events represent the *immediate*, *necessary*, and *sufficient* causes for the occurrence of the intermediate (or top) events.

The fault tree itself is a logical model, and, thus, represents the qualitative characterization of the system logic. There are, however, many quantitative algorithms to evaluate fault trees. For example, the concept of cut sets discussed earlier can also be applied to fault trees by using the Boolean algebra method. By using $\Pr(C_1 \cup C_2 \cdots \cup C_m)$, the probability of occurrence of the top event can be determined using (4.29).

To understand the symbology of logic trees, including fault trees, consider Figure 4.8. In essence, there are three types of symbols: *events*, *gates*, and *transfers*. Basic events, undeveloped events, condition events, and external events are sometimes referred to as *primary events*. When postulating events in the fault tree, it is important to include not only the undesired component states (e.g., applicable failure modes), but also the time when they occur.

To better understand the fault tree concept, let us consider the complex block diagram shown in Figure 4.4. Let us also assume that the block diagram models a circuit in which the arrows show the direction of current flow. A top event of "no current at point F" is selected, and all events that cause this top event are deductively postulated. Figure 4.9 shows the results.

As another example, consider the pumping system shown in Figure 4.10. Sufficient water is delivered from the water source T_1 when only one of the two pumps, P-1 or P-2, works. All the valves V-1 through V-5 are normally open. The sensing and control system S senses the demand for the pumping system and automatically starts both P-1 and P-2, (if one of the two pumps fails to start or fails during operation, the mission is still considered successful if the other pump functions properly). The two pumps and the sensing and control system use the same ac power source AC. Assume the water content in T_1 is sufficient and available, there are no human errors, and no failure in the pipe connections is considered important.

It is clear that the system's mission is to deliver sufficient water when needed. Therefore, the top event of the fault tree for this system should be "no water is delivered when needed." Figure 4.11 shows the fault tree for this example. In Figure 4.11, the failures of AC and S are shown with *undeveloped events*. This is because one can further expand the fault tree if one knows what makes up the failures of AC and S, in which case these events will be *intermediate events*.

PRIMARY EVENT SYMBOLS

BASIC EVENT - A basic event requiring no further development

CONDITIONING EVENT - Specific conditions or restrictions that apply to any logic gate (used primary with PRIORITY AND and INHIBIT gate)

UNDEVELOPED EVENT - An event which is not further developed either because it is of insufficient consequence or because information is unavailable

EXTERNAL EVENT - An event which is normally expected to occur

INTERMEDIATE EVENT SYMBOLS

INTERMEDIATE EVENT - An event that occurs because of one or more antecedent causes acting through logic gates

GATE SYMBOLS

AND - Output occurs if all of the input events occur.

OR - Output occurs if at least one of the input events occurs

EXCLUSIVE OR - Output occurs if exactly one of the input feventsoccurs

PRIORITY AND - Output occurs if all of the input events occur in a specific sequence (the sequence is represented by a CONDITIONING EVENT drawn to the right of the gate)

PRIORITY AND - Output occurs if all of the input events occur in a specific sequence (the sequence is represented by a CONDITIONING EVENT drawn to the right of the gate)

Not - OR - Output occurs if at least one of the input events does not occur

Not - AND - Output occurs if all of the input events do not occur

TRANSFER SYMBOLS

TRANSFER IN - Indicates that the tree is developed further at the occurrence of the corresponding TRANSFER OUT (e.g., on another page)

TRANSFER OUT - Indicates that this portion of the tree must be attached at the corresponding TRANSFER IN

Figure 4.8 Primary event, gate, and transfer symbols used in logic trees.

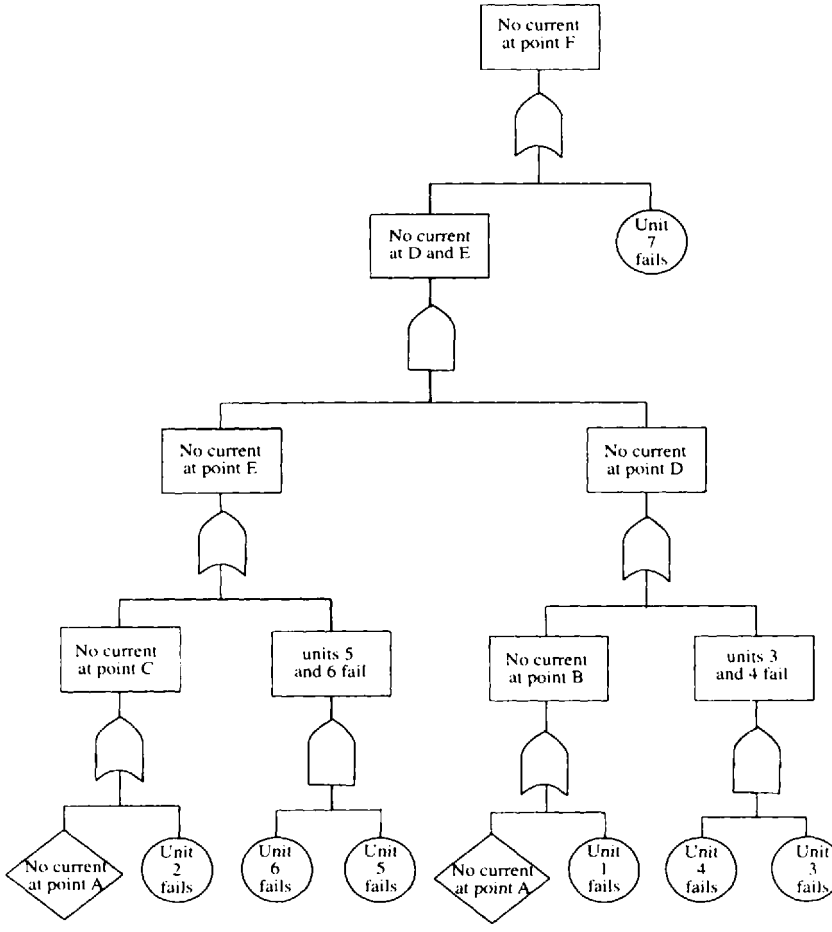


Figure 4.9 Fault tree for the complex parallel-series system in Figure 4.4.

However, since enough information (e.g., failure characteristics and probabilities) about these events is known, we have stopped their further development at this stage. Although the development of the fault tree in Figure 4.11 is based on a strict deductive procedure (i.e., systematic decomposition of failures starting from "sink" and deductively proceeding toward "source"), one can rearrange it to the more concise and compact equivalent form shown in Figure 4.12. While the development of the fault tree in Figure 4.11 requires only a minimum

understanding of the overall functionality and logic of the system, direct development of more compact versions requires a much better understanding of the overall system logic. If more complex logical relationships are required, other logical representations can be described by combining the two basic AND and OR gates. For example, the *K-out-of-N* and *exclusive OR* logics can be described, as shown in Figure 4.13.

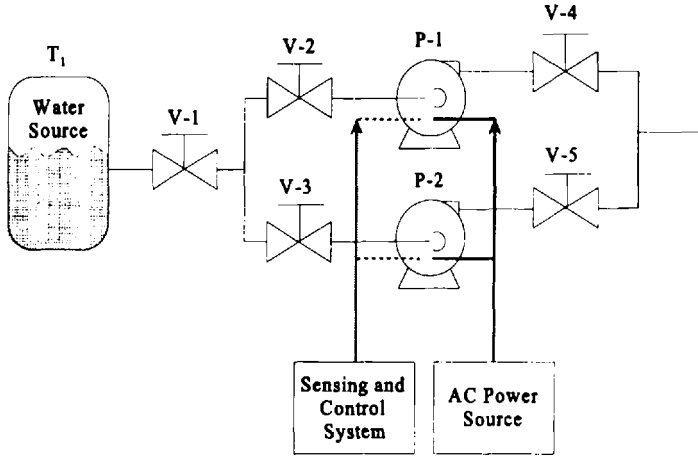


Figure 4.10 An example of a pumping system.

For a more detailed discussion of the construction and evaluation of fault trees, refer to Vesely et al. (1981).

4.2.2 Evaluation of Logic Trees

The evaluation of logic trees (e.g., fault trees, success trees, and master logic diagrams) involves two distinct aspects: *logical* or *qualitative evaluation* and *probabilistic* or *quantitative evaluation*. Qualitative evaluation involves the determination of the logic tree cut sets, path sets or logical evaluations to rearrange the tree logic for computational efficiency (similar to the rearrangement presented in Figure 4.12 for a fault tree). Determining the logic tree cut sets or path sets involves some straightforward Boolean manipulation of events that we describe

here. However, there are many types of logical rearrangements and evaluations, such as fault tree modularization, that are beyond the scope of this book. The reader is referred to Vesely et al. (1981) for a more detail discussion of this topic. In addition to the traditional Boolean analysis of logic trees, a combinatorial approach will also be discussed. This technique generates mutually exclusive cut or path sets.

Boolean Algebra Analysis of Logic Trees

The quantitative evaluation of logic trees involves the determination of the probability of the occurrence of the top event. Accordingly, unreliability or reliability associated with the top event can also be determined. The qualitative evaluation of logic trees through the generation of cut or path sets is conceptually very simple. The tree OR-gate logic represents the *union* of the input events. That is, all the input events must occur to cause the output event to occur. For example, an OR gate with two input, events A and B and the output event Q can be represented by its equivalent Boolean expression, $Q = A \cup B$. Either A or B or both must occur for the output event Q to occur. Instead of the union symbol $\cup$, the equivalent "+" symbol is often used in engineering applications. Thus, $Q = A + B$. Generally, for an OR gate with n inputs, $Q = A_1 + A_2 + \ldots + A_n$. The AND gate can be represented by the intersect logic. Therefore, the Boolean equivalent of an AND gate with two inputs A and B would be $Q = A \cap B$ (or $Q = A \cdot B$).

Determination of cut sets using the above expressions is possible through several algorithms. These algorithms include the *top-down* or *bottom-up* successive substitution method, the *modularization approach*, and *Monte Carlo* simulation. The *Fault Tree Handbook*, Vesely et al. (1981), describes the underlying principles of these qualitative evaluation algorithms. The most widely used and straightforward algorithm is the successive substitution method. In this approach, the equivalent Boolean representation of each gate in the logic tree is determined such that only primary events remain. Various Boolean algebra rules are applied to reduce the Boolean expression to its most compact form, which represents the minimal path or cut sets of the logic tree. The substitution process can proceed from the top of the tree to the bottom or vice versa. Depending on the logic tree and its complexity, either the former or the latter approach, or a combination of the two, can be used.

As an example, let's consider the fault tree shown in Figure 4.11. Clearly, each node represents a failure. The step-by-step, top-down Boolean substitution of the top event is presented below.

Step 1: $T = E_1 \cdot E_2$.

Step 2: $E_1 = E_3 + V_3 + V_5 + E_4$,
$E_2 = E_3 + V_4 + V_2 + E_5$, $T = E_3 + V_3 \cdot V_4 + V_3 \cdot V_2 + V_5 \cdot V_4 + V_5 \cdot V_2 + E_4 \cdot V_4 + E_4 \cdot V_2 + E_4 \cdot E_5 + V_3 \cdot E_5 + V_5 \cdot E_5$ (T has been reduced by using the Boolean identities $E_3 \cdot E_3 = E_3$, $E_3 + E_3 \cdot X = E_3$, and $E_3 + E_3 = E_3$.)

Step 3: $E_3 = T_1 + V_1$,
$E_4 = E_6 + P_2 + AC$,
$E_5 = E_6 + P_1 + AC$,
$T = T_1 + V_1 + AC + V_3 \cdot V_4 + V_3 \cdot V_2 + V_5 \cdot V_4 + V_5 \cdot V_2 + V_4 \cdot P_2 + P_2 \cdot V_2 + E_6 + P_2 \cdot P_1 + V_3 \cdot P_1 + V_5 \cdot P_1$. (Again, identities such as $AC + AC = AC$ and $E_6 + V_3 \cdot E_6 = E_6$ have been used to reduce T.)

Step 4: $E_6 = AC + S$,
$T = S + AC + T_1 + V_1 + V_3 \cdot V_4 + V_3 \cdot V_2 + V_5 \cdot V_4 + V_5 \cdot V_2 + V_4 \cdot P_2 + P_2 \cdot V_2 + P_2 \cdot P_1 + V_3 \cdot P_1 + V_5 \cdot P_1$.

The Boolean expression obtained in Step 4 represents four minimal cut sets with one element (cut set of size 1), and nine minimal cut sets with two elements (cut set of size 2). The size 1 cut sets are occurrence of failure events S, AC, T_1, V_1. The size 2 cut sets are events V_3 and V_4; V_3 and V_2; V_5 and V_4; V_5 and V_2; V_4 and P_2; P_2 and V_2; P_2 and P_1; V_3 and P_1; and V_5 and P_1. A simple examination of each cut set shows that its occurrence guarantees the occurrence of the top event (failure of the system). For example, the cut set V_5 and P_1, which represents simultaneous failure of valve V_2 and pump P_1, causes the two flow branches of the system to be lost, which in turn disables the system. The same substitution approach can be used to determine the path sets. In this case the events are success events representing adequate realization of describe functions.

It is clear from this fault tree example that the evaluation of a large logic tree by hand can be a formidable job. A number of computer based programs are available for the analysis of logic trees. Specter and Modarres (1996) elaborate on the important characteristics of these software programs. Appendix C describes some of the premier software tools in the market. Quantitative evaluation of the cut sets or path sets has already been discussed under the context of the reliability block diagram. For example, expression (4.29) forms the basis for quantitative evaluation of the cut sets.

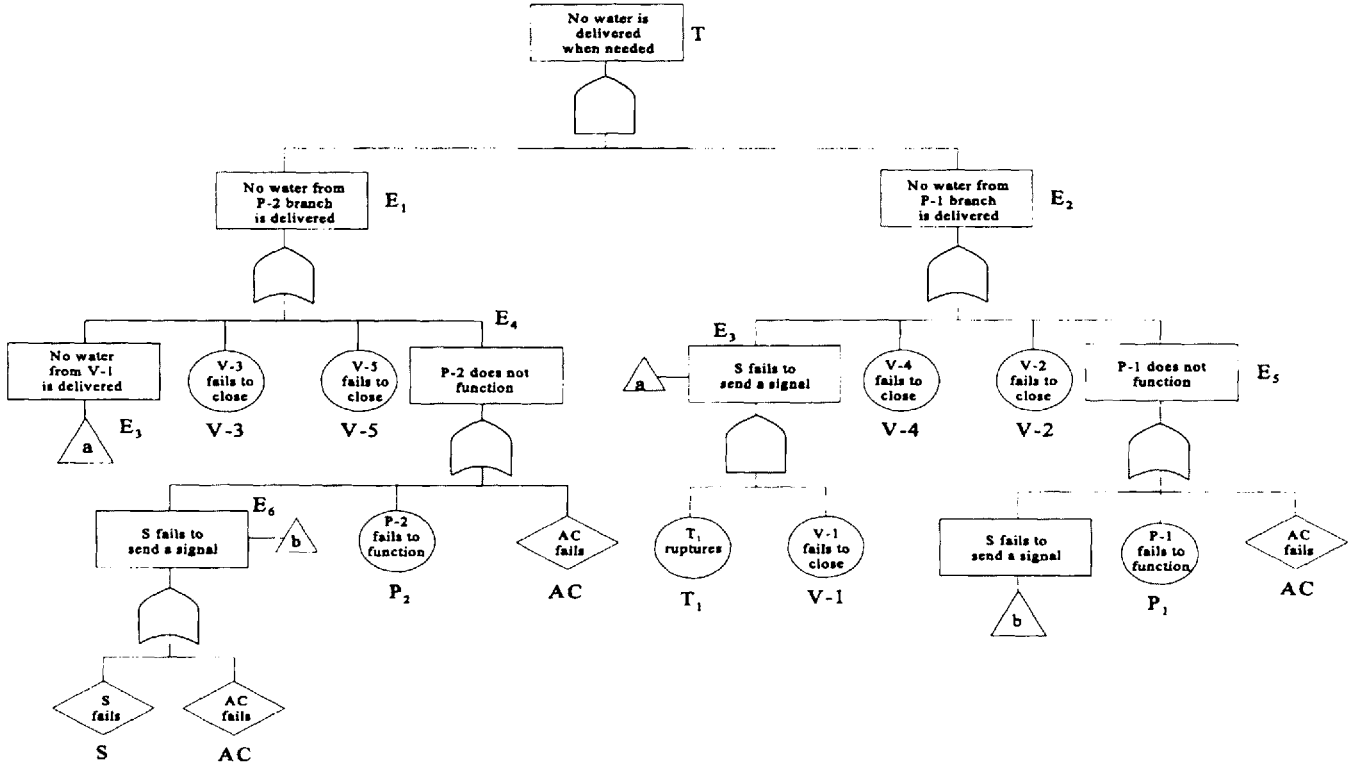


Figure 4.11 Fault tree for the pumping system in Figure 4.10.

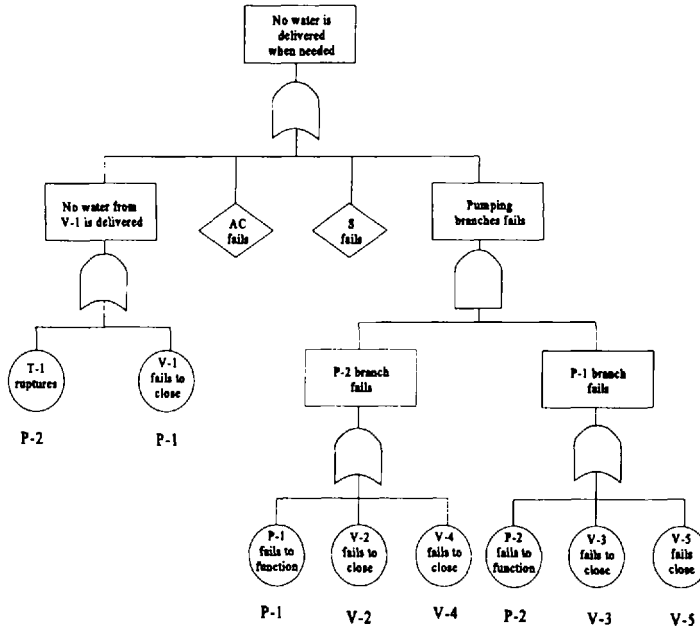


Figure 4.12 More compact form of the fault tree in Figure 4.11.

That is, the probability that the top event, T, occurs in a mission time t is

$$\Pr(T) = \Pr\left(C_1 \cup C_2 \cup \cdots \cup C_n\right) \tag{4.31}$$

Probability of the top event in a system reliability framework can be thought of as the unreliability of the system. To understand the complexities discussed earlier for the determination of Pr(T), let's consider the case where the following two cut sets are obtained:

$$C_1 = A \cdot B$$
$$C_2 = A \cdot C$$

Then,

$$\Pr(T) = \Pr(A \cdot B + A \cdot C) \tag{4.32}$$

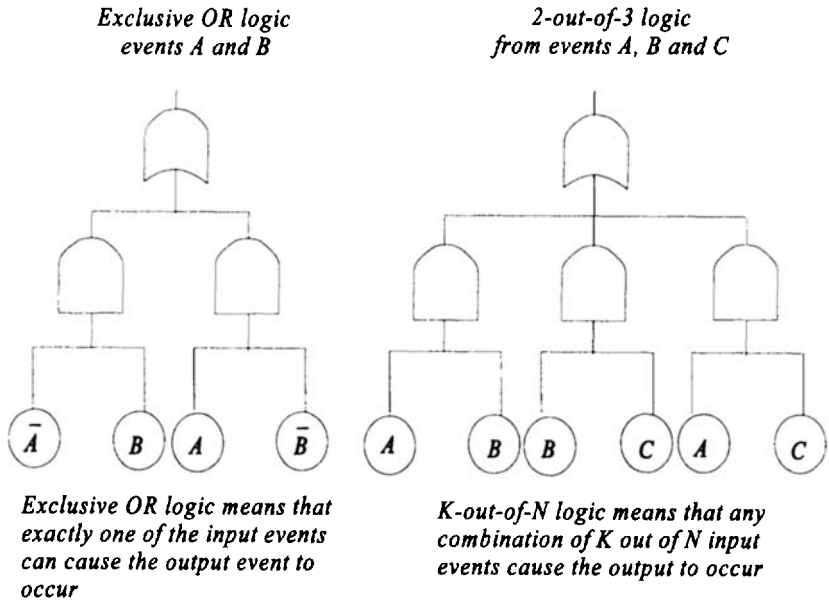


Figure 4.13 Exclusive OR and K-out-of-N logics.

According to (4.7),

$$\begin{aligned} \Pr(T) &= \Pr(A \cdot B) + \Pr(A \cdot C) - \Pr(A \cdot B \cdot A \cdot C) \\ &= \Pr(A \cdot B) + \Pr(A \cdot C) - \Pr(A \cdot B \cdot C) \end{aligned}$$

If A, B, and C are independent, then

$$\begin{aligned} \Pr(T) = {} & \Pr(A) \cdot \Pr(B) + \Pr(A) \cdot \Pr(C) \\ & - \Pr(A) \cdot \Pr(B) \cdot \Pr(C) \end{aligned} \tag{4.33}$$

The determination of the cross-product terms, such as $\Pr(A)\ \Pr(B)\ \Pr(C)$ in (4.33), poses a dilemma in the quantitative evaluation of cut sets, especially when the number of the cut sets is large. In general, there are 2^{n-1} of such terms in cut sets.

For example, in the 13 cut sets generated for the pumping example, there are 8191 such terms. For large logic trees, this can be a formidable job even for powerful mainframe computers.

Fortunately, when dealing with cut sets, evaluation of these cross product terms is often not necessary, and the bounding approach shown in (4.30) is quite adequate. As discussed earlier, this is true whenever we are dealing with small probabilities, which is often the case for probability of failure events. In these cases, e.g., in (4.33), $\Pr(A) \cdot \Pr(B) \cdot \Pr(C)$ is substantially smaller than $\Pr(A) \cdot \Pr(B)$ and $\Pr(A) \cdot \Pr(C)$. Thus the bounding result can also be used as an approximation of the true reliability or unreliability value of the system. This is often called the *rare event approximation*. Let's assume, that

$$\Pr(A) = \Pr(B) = \Pr(C) = 0.1$$

Then,

$$\Pr(A) \cdot \Pr(B) = \Pr(A) \cdot \Pr(C) = 0.01$$

and

$$\Pr(A) \cdot \Pr(B) \cdot \Pr(C) = 0.001$$

The latter is smaller than the former by an order of magnitude. Although $\Pr(T) = 0.019$, the rare event approximation yields $\Pr(T) \approx 0.02$. Obviously, the smaller the probabilities of the events, the better the approximation.

As another example, consider the simple block diagram shown in Figure 4.14 which represents a system that has three paths from point X to point Y.

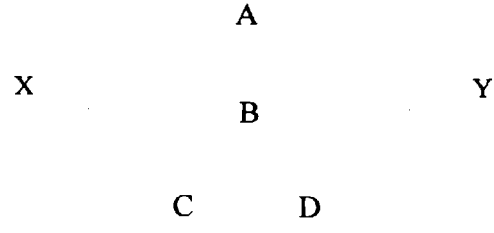


Figure 4.14 System block diagram.

The equivalent fault tree is shown in Figure 4.15. The equivalent Boolean substitution equations are:

$$T = A \cdot B \cdot G_1$$
$$G_1 = C + D$$
$$T = A \cdot B \cdot (C + D),$$
$$T = A \cdot B \cdot C + A \cdot B \cdot D$$

If the probability of events A, B, and C is 0.1, and the probability of event D is 0.2, the top event probability is evaluated as follows.

Using the rare event approximation discussed earlier,

$$\Pr(T) \approx \Pr(A) \cdot \Pr(B) \cdot \Pr(C) + \Pr(A) \cdot \Pr(B) \cdot \Pr(D)$$

therefore,

$$\Pr(T) \approx 0.1 \times 0.1 \times 0.1 + 0.1 \times .1 \times 0.2 = 0.003$$

Note that the terms $A \cdot B \cdot C$ and $A \cdot B \cdot D$ are not mutually exclusive and, therefore, the value of $\Pr(T)$ is approximate, since the rare event approximation has been used.

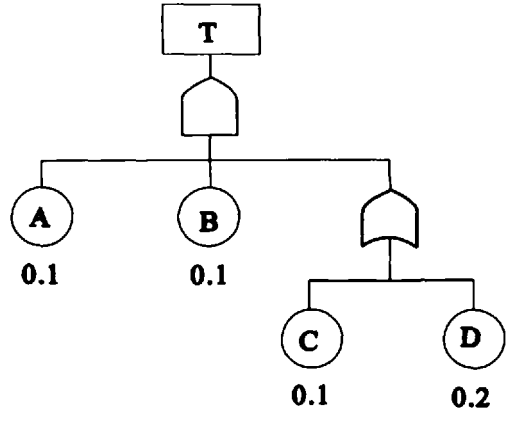


Figure 4.15 Fault Tree Representation of Figure 4.1.

When all events are independent, in order to calculate the exact failure probability, using minimal cut sets their cross product terms most also be included in calculation of Pr(T),

$$\Pr(T) \approx \Pr(A) \cdot \Pr(B) \cdot \Pr(C) + \Pr(A) \cdot \Pr(B) \cdot \Pr(D) - \Pr(A) \cdot \Pr(B) \cdot \Pr(C) \cdot \Pr(D)$$

Accordingly,

$$\Pr(T) \approx 0.1 \times 0.1 \times 0.1 + 0.1 \times 0.1 \times 0.2 - 0.1 \times 0.1 \times 0.1 \times 0.2 = 0.0028$$

Combinatorial Technique for Evaluation of Logic Trees

Unlike the substitution technique, which is based on Boolean reduction, the combinatorial method does not convert the tree logic into Boolean equations to generate cut or path sets. Rather, this method which is similar to the truth table approach relies on a combinatorial algorithm to exhaustively generate all probabilistically significant combinations of both "failure" and "success" events and subsequently propagate effect of each combination on the logic tree to determine the state of the top event. Because successes and failures are combined, all combinations are mutually exclusive. The quantification of logic trees based on the combinatorial method yields a more exact result.

To illustrate the combinatorial approach, consider the fault tree in Figure 4.15. All possible combinations of success or failure events should be generated. Because there are 4 events and 2 states (success or failure) for each event then there are $2^4 = 16$ possible system states (i.e., actual physical states). Some of these states constitute system operation (when top event T does not happen), and some states constitute failure (when top event T does happen). These 16 states are illustrated in Table 4.1. In this table, the subscript S refers to the nonoccurrence of an event (success), and subscript F is referred to the failure or occurrence of the event in the fault tree.

Only combinations 14, 15, and 16 lead to the occurrence of the top event T which results in system failure probability of $\Pr(T) = 0.0018 + 0.0008 + 0.0002 = 0.00028$. This is the exact value (provided that the events are independent). Clearly, this is consistent with the exact calculation by the Boolean reduction method. Note that sum of the probabilities of all possible combinations (16 of them in this case) is unity because the combinations are all mutually exclusive and cover all event space (i.e., the universal set). Combinations 14, 15, and 16 are mutually exclusive cut sets.

In order to visualize the difference between the results generated from the Boolean reduction and the combinatorial approach the Venn Diagram technique is helpful. Again consider the simple system in Figure 4.14 consisting of four events A, B, C, and D. The Boolean reduction process results in the minimal cut sets corresponding to system failure. These are, $A \cdot B \cdot C$ and $A \cdot B \cdot D$.

Table 4.1 Combinatorial Method of Evaluating Event Tree

Combination Number	Combination Definition (System states)	Probability of C_i	System Operation T
1	$A_S \cdot B_S \cdot C_S \cdot D_S$	0.5832	S
2	$A_S \cdot B_S \cdot C_S \cdot D_F$	0.1458	S
3	$A_S \cdot B_S \cdot C_F \cdot D_S$	0.0648	S
4	$A_S \cdot B_S \cdot C_F \cdot D_F$	0.0162	S
5	$A_S \cdot B_F \cdot C_S \cdot D_S$	0.0648	S
6	$A_S \cdot B_S \cdot C_S \cdot D_F$	0.0162	S
7	$A_S \cdot B_F \cdot C_F \cdot D_S$	0.0072	S
8	$A_S \cdot B_F \cdot C_F \cdot D_F$	0.0018	S
9	$A_F \cdot B_S \cdot C_S \cdot D_S$	0.0648	S
10	$A_F \cdot B_S \cdot C_S \cdot D_F$	0.0162	S
11	$A_F \cdot B_S \cdot C_F \cdot D_S$	0.0072	S
12	$A_F \cdot B_S \cdot C_F \cdot D_F$	0.0018	S
13	$A_F \cdot B_F \cdot C_S \cdot D_S$	0.0072	S
14	$A_F \cdot B_F \cdot C_S \cdot D_F$	0.0018	F
15	$A_F \cdot B_F \cdot C_F \cdot D_S$	0.0008	F
16	$A_F \cdot B_F \cdot C_F \cdot D_F$	0.0002	F
		$\sum C_{i=} 1.0000$	

S = Success; F = Failure.

The left side of Figure 4.16 represents a Venn diagram for the two cut sets above. Each cut set is represented by one shaded area. The two shaded areas are overlapping indicating that the cut sets are not mutually exclusive. Now consider how combinations 14, 15, and 16 are represented in the Venn diagram (right side of Figure 4.16). Again, each shaded area corresponds to a combination. In this case, there is no overlapping of the shaded areas. That is, the combinatorial approach generates mutually *exclusive* sets, and those sets that lead to system

failure are called *eventually exclusive* sets. Therefore, when the rare event approximation is used, the contributions generated by the combinatorial approach has no overlapping area and produces the exact probability. Since for size problems, usually the rare event approximation is the only practical choice, if the exact probabilities are desired, or failure probabilities are greater that 0.1, then the combinatorial approach is preferred.

A typical logic model may contain hundreds of events. For n events, there are 2^n combinations. Obviously, for a large n (e.g., $n > 20$), the generation of this large number of combinations is impractical; a more efficient method would be needed. An algorithm to generate combinations which probabilities exceed some cutoff limit (e.g., 10^{-7}) is proposed by Dezfuli, et al. (1994). The algorithm generates combinations that are referred to as probabilistically significant combinations.

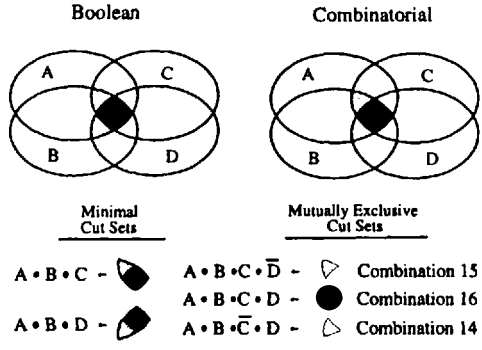


Figure 4.16 Boolean and combinatorial diagrams of events.

In this combinatorial algorithm, the total number of events is first determined. Each event has an associated probability of failure occurrence. A combination represents the status (i.e., failed or not failed) of every event in the entire logic diagram. The collection of all failed blocks within a combination is referred to as a "failure set" (FS). A failure set may have zero elements, meaning there is no failure events in the combination. This set is called the nil combination. The objective is to generate other probabilistically significant combinations. The following assumptions are made:

1. The failure events are independent.
2. The nil combination is a significant combination.

Given a combination C, the assumption of the independence implies that the probability of the combination is:

$$P_C = \prod_{i \in FS} P_i \prod_{i \notin FS} (1 - P_i) \tag{4.34}$$

here P_i is the probability of an individual failure event.

Consider the combination C', which differs from the combination C in a sense that an event j is added to its failure set (i.e., transition of a success event to a failure event). From the above results, it can be concluded that

$$P'_C = P_C \times \frac{P_j}{1 - P_j} \tag{4.35}$$

Note that adding a block j to the failed set increases the probability of a combination if $P_j > 0.5$, and decreases the probability of a combination if $P_j < 0.5$.

Consider also the combination C'', which differs from the combination C in that block j is replaced with block k (i.e., the replacement of a block in the failed set with another block). Therefore,

$$P''_C = P_C \times \frac{P_j}{1 - P_j} \times \frac{1 - P_k}{P_k} \tag{4.36}$$

This shows that replacing an event of a failed set in a combination with an event that has a lower failure probability results in a combination of lower probability, and replacing an event with an event that has a higher failure probability results in a combination of higher probability.

As such, the events are sorted in a decreasing order of probability. Each event is identified by its position in this ranking, such that $P_i > P_j$ when $i < j$. Each combination is identified by a *list* of the event it contains in the failed set. To make the correspondence between combinations and lists unique, the list must be in ascending rank order, which corresponds to decreasing probability order.

Now consider a list representing a combination. Define a *descendant* of the list to be a list with one extra event appended to the failed set. Since the list must be ordered, this extra event must have a higher rank (lower probability) than any events in the original list. If there is no such event, there is no descendants. The basis of the algorithm can be computerized easily as it is done in the REVEAL_W™ software, see Dezfuli, et al. (1994) and Appendix C. One should generate all descendants of the input list, and recursively generate all subsequent descendants. Since the algorithm begins with an empty list, it is clear that the

algorithm will generate all possible lists. Figure 4.17 illustrates this scheme for the simple case of four events.

To generate only significant lists, we first need to prove that if a list is not significant, its descendants are not significant. The nil set is significant. According to (4.35), at least one item of the list must have a probability lower that 0.5. Any failure event added to form the descendant would also have a probability lower than 0.5. Therefore, the probability of the descendant would be lower than that of the original set, and, therefore, cannot be significant.

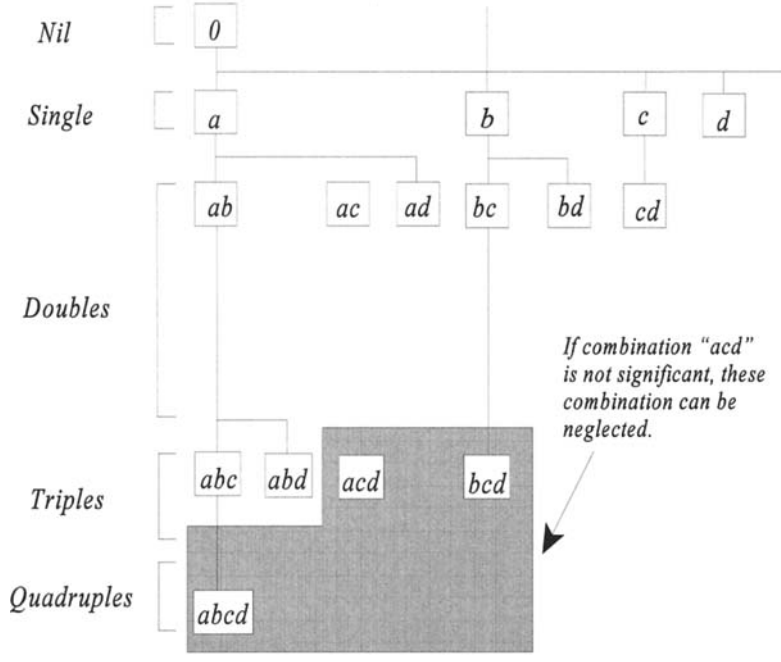


Figure 4.17 Computer algorithm for combinatorial approach.

The algorithm takes advantage of this property. The descendants are generated in an increasing rank (decreasing probability) order of the added events. Equation (4.36) shows that the probability of the generated combinations is also decreasing. Each list is checked to see whether it is significant. If it is not significant, the routine exits without any recursive operation and without generating any further

descendants of the original input list. Figure 4.17 shows the effect, if the state consisting of events *a, c,* and *d* is found to be insignificant; all the indicated combinations are immediately excluded from further consideration.

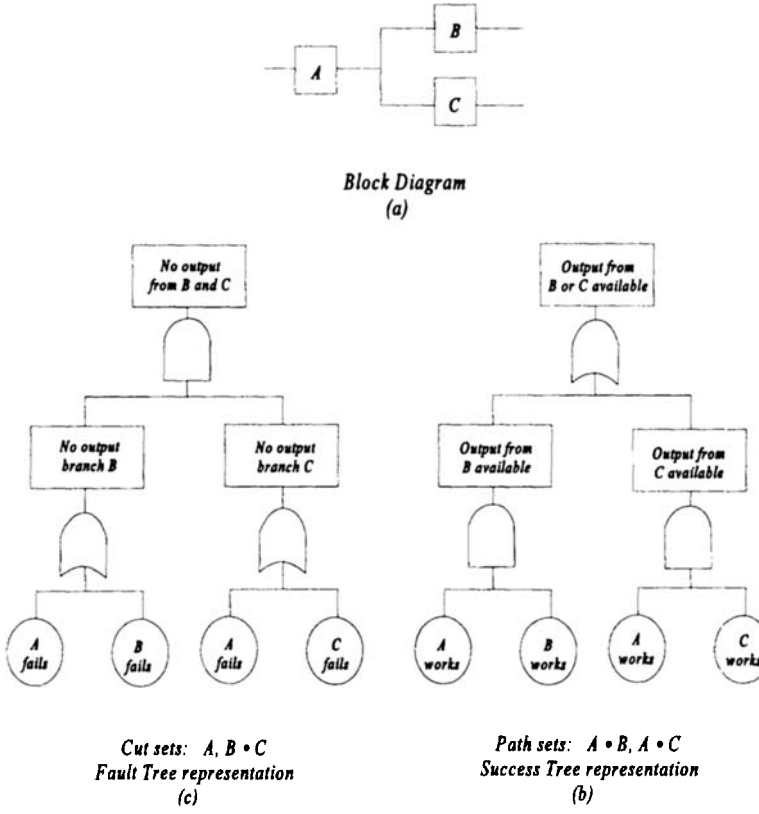


Figure 4.18 A correspondence between a fault and success trees.

4.2.3 Success Tree Method

The success tree method is conceptually the same as the fault tree method. By defining the *desirable* top event, all intermediate and primary events that

guarantee the occurrence of this desirable event are deductively postulated. Therefore, if the logical complement of the top event of a fault tree is used as the top event of a success tree, the Boolean structure represented by the fault tree is the Boolean complement of the success tree. Thus, the success tree, which shows the various combinations of success events that guarantee the occurrence of the top event, can be logically represented by path sets instead of cut sets.

To better understand this problem, consider the simple block diagram shown in Figure 4.18a. The fault tree for this system is shown in Figure 4.18b and the success tree in Figure 4.18c. Figure 4.19 shows an equivalent representation of Figure 4.18c.

By inspecting Figure 4.18b and Figure 4.18c, it is easy to see that changing the logic of one tree (changing AND gates to OR gates and vice versa) and changing all primary and intermediate events to their logical complements yields the other tree. This is also true for cut sets and path sets. That is, the logical complement of the cut sets of the fault tree yields the path sets of the equivalent success tree. This can easily be seen in Figure 4.18. The complement of cut sets is

$$\overline{A + B \cdot C} = \text{(apply De Morgan's Theorem)}$$

$$\overline{A} \cdot \overline{B \cdot C} = \text{(apply De Morgan's Theorem)}$$

$$\overline{A} \cdot (\overline{B} + \overline{C}) = \overline{A} \cdot \overline{B} + \overline{A} \cdot \overline{C}$$

which are the path sets.

Qualitative and quantitative evaluations of success paths are mechanistically the same as those of fault trees. For example, the top-down successive substitution of the gates and reduction of the resulting Boolean expression yield the minimal path sets. Accordingly, the use of (4.27), or its lower bound (4.28), allows to determine the top-event probability (in this case, reliability). As noted earlier, (4.27) poses a computational problem. In this context of using path sets, Wang and Modarres (1990) have described several options for efficiently dealing with this problem

A convenient way to reduce complex Boolean equations, especially the paths sets, is to use the following expressions:

$$\begin{aligned} \Pr(T) &= \Pr(P_1 \cup P_2 \cup \cdots \cup P_n) \\ &= \Pr(P_1) + \Pr(\bar{P}_1 \cap P_2) + \Pr(\bar{P}_1 \cap \bar{P}_2 \cap P_3) + \cdots \\ &\quad + \Pr(\bar{P}_1 \cap \bar{P}_2 \cap \cdots \cap \bar{P}_{n-1} \cap P_n) \end{aligned} \tag{4.37}$$

For further discussions in applying (4.37), see Fong and Buzacoot (1987).

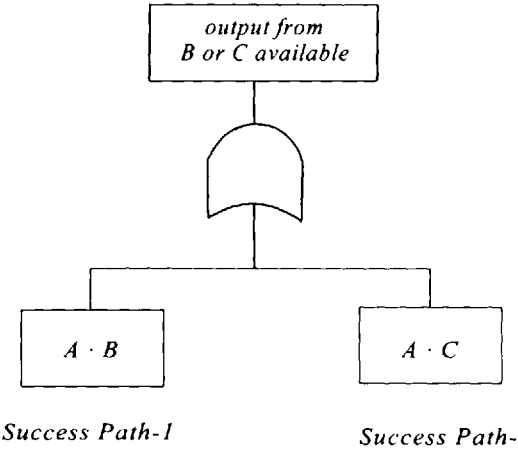


Figure 4.19 Equivalent Representation of a Success Tree in Figure 4.18(c)

The combinatorial approach discussed in section 4.2.2 is far superior for generating mutually exclusive path sets that assure a system's successful operation. For example, combinations 1–3 in Table 4.1 represent all mutually exclusive path sets for the system shown in Figure 4.14.

Success trees, as opposed to fault trees, provide a better understanding and display of how a system functions successfully. While this is important for designers and operators of complex systems, fault trees are more powerful for analyzing failures associated with systems and determining the causes of system failures. The minimal path sets of a system shows the system user how the system operates successfully. A collection of events in a minimal path set is sometimes referred to as a *success path*. A logical equivalent of a success tree can also be represented by using the top event as an output to an OR gate in which input to the gate would show the success paths. For example, Figure 4.19 shows the equivalent representation for the success tree in Figure 4.18c.

In complex systems, the type of representation given in Figure 4.19 is useful for efficient system operation.

4.3 EVENT TREE METHOD

If successful operation of a system depends on an approximately chronological, but discrete, operation of its units or subsystems (e.g., units should work in a defined sequence for operational success), then an event tree is appropriate. This may not always be the case for a simple system, but it is often the case for complex systems, such as nuclear power plants where the subsystems should work according to a given sequence of events to achieve a desirable outcome. Event trees are particularly useful in these situations.

4.3.1 Construction of Event Trees

Let's consider the event tree built for a nuclear power plant and shown in Figure 4.20. The event trees are horizontally built structures that start on the left, where the *initiating event* is modeled. This event describes a situation when a legitimate demand for the operation of a system(s) occurs. Development of the tree proceeds chronologically, with the demand on each unit (or subsystem) being postulated. The first unit demanded appears first, as shown on the top of the structure. In Figure 4.20, the events (referred to as event tree headings) are as follows:

RP = Operation of the reactor-protection system to shutdown the reactor
ECA = Injection of emergency coolant water by pump A
ECB = Injection of emergency coolant water by pump B
LHR = Long-term heat removal

At a branch point, the upper branches of an event shows the *success* of the event heading and the lower branch shows its failure. In Figure 4.20, following the occurrence of the initiating event A, RP needs to work (event B). If RP does not work, the overall system will fail (as shown by the lower branch of event B). If RP works, then it is important to know whether ECB functions or not. If ECB does not function, even though RP has worked, the overall system would still fail. However, if ECB functions properly, it is important for LHR to function. Successful operation of LHR leads the system to a successful operating state, and failure of LHR (event E) leads the overall system to a failed state. Likewise, if ECA functions, it is important that it be followed by a proper operation of LHR. If LHR fails, the overall system would be in a failed state. If LHR operates successfully, the overall system would be in a success state. It is obvious that operation of certain subsystems may not be necessarily dependent on the occurrence of some preceding events. For example, if ECA operates successfully it does not matter for the overall system success whether or not ECB operates.

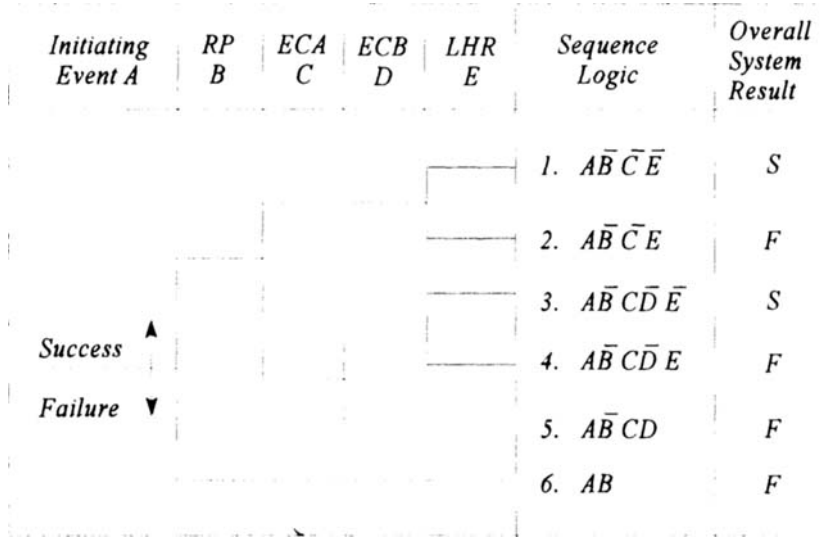


Figure 4.20 Example of an event tree.

The outcome of each of the sequences of events is determined by the analyst and shown at the end of each sequence. This outcome, in essence, describes the final outcome of each sequence, whether the overall system succeeds, fails, initially succeeds but fails at a later time, or vice versa. The logical representation of each sequence can also be shown in the form of a Boolean expression. For example, for sequence 5 in Figure 4.20, events A, C, and D have occurred, but event B has not occurred (shown by $\bar{B}$). Clearly, these sequences are mutually exclusive.

The event trees are usually developed in a binary format; i.e., the heading events are assumed to either occur or not occur. In cases where a spectrum of outcomes is possible, the branching process can proceed with more than two outcomes. In these cases, the qualitative representation of the event tree branches in a Boolean sense would not be possible.

The development of an event tree, although somewhat deductive, in principle, requires a good deal of inductive thinking by the analyst. To demonstrate this issue and further understand the concept of event tree development, let's consider the system shown in Figure 4.10. One can think of a situation where the sensing and control system device S initiates one of the two pumps. At the same time, the ac power source AC should always exist to allow S

and pumps P-1 and P-2 to operate. Thus, if we define three distinct events S, AC and pumping system PS for a sequence of events starting with the initiating event, an event tree that includes these three events can be constructed. Clearly if AC fails, both PS and S fail; if S fails, only PS fails. This would lead to placing AC as the first event tree heading followed by S and PS. This event tree is illustrated in Figure 4.21.

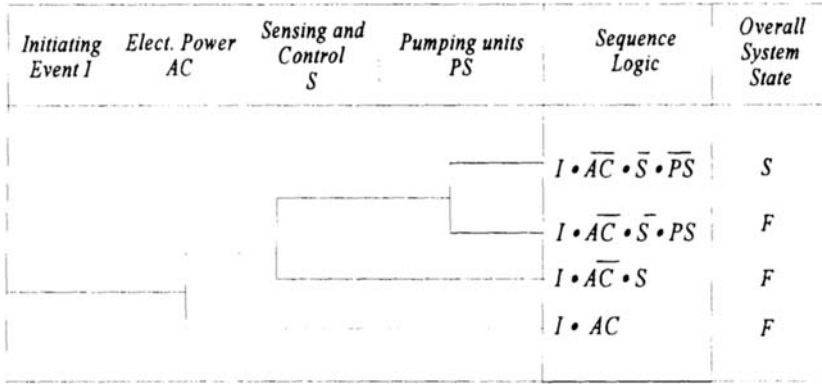


Figure 4.21 Event tree for the pumping system.

Events represent discrete states of the systems. The logic of these states can be modeled by fault trees. This way the event tree sequences and the logical combinations of events can be considered. This is a powerful aspect of the event tree technique. If the event tree headings represent complex subsystems or units, using a fault tree for each event tree heading can conveniently model the logic. Clearly, other system analysis models, such as reliability block diagrams and logical representations in terms of cut sets or path sets, can also be used.

4.3.2 Evaluation of Event Trees

Qualitative evaluation of event trees is straightforward. The logical representation of each event tree heading, and ultimately each event tree sequence, is obtained and then reduced through the use of Boolean algebra rules. For example, in sequence 5 of Figure 4.20, if events B, C, and D are represented by the following Boolean expressions, the reduced Boolean expression of the sequence can be obtained.

$$A = a$$
$$B = b + c \cdot d$$
$$C = e + d$$
$$D = c + e \cdot h$$

The simultaneous Boolean expression and reduction proceeds as follows:

$$A \cdot \bar{B} \cdot C \cdot D = a \cdot (\overline{b + c \cdot d}) \cdot (e + d) \cdot (c + e \cdot h)$$
$$= a \cdot (\bar{b} \cdot \bar{c} + \bar{b} \cdot \bar{d}) \cdot (e \cdot c + e \cdot h + d \cdot c)$$
$$= a \cdot \bar{b} \cdot \bar{c} \cdot e \cdot h + a \cdot \bar{b} \cdot c \cdot \bar{d} \cdot e$$

If an expression explaining all failed states is desired, the union of the reduced Boolean equations for each sequence that leads to failure should be obtained and reduced.

Quantitative evaluation of event trees is similar to the quantitative evaluation of fault trees. For example, to determine the probability associated with an A · $\bar{B}$ · C · D sequence, one would consider:

$$\Pr(A \cdot \bar{B} \cdot C \cdot D) = \Pr(a \cdot \bar{b} \cdot \bar{c} \cdot e \cdot h + a \cdot \bar{b} \cdot c \cdot \bar{d} \cdot e)$$
$$= \Pr(a \cdot \bar{b} \cdot \bar{c} \cdot e \cdot h) + \Pr(a \cdot \bar{b} \cdot c \cdot \bar{d} \cdot e)$$
$$= \Pr(a) \cdot [1 - \Pr(b)][1 - \Pr(c)]\Pr(e) \cdot \Pr(h) +$$
$$\Pr(a) \cdot [1 - \Pr(b)]\Pr(c) \cdot [1 - \Pr(d)]\Pr(e)$$

Since the two terms are disjoint, the above probability is exact. However, if the terms are not disjoint, the rare event approximation can be used here.

4.4 MASTER LOGIC DIAGRAM

For complex systems such as a nuclear power plant, modeling for reliability analysis or risk assessment may become very difficult. In complex systems, there are always several functionally separate subsystems that interact with each other, each of which can be modeled independently. However, it is necessary to find a logical representation of the overall system interactions with respect to the individual subsystems. The master logic diagram (MLD) is such a model.

Consider a functional block diagram of a complex system in which all of the functions modeled are necessary in one way or another to achieve a desired

objective. For example, in the context of a nuclear power plant, the independent functions of heat generation, normal heat transport, emergency heat transport, reactor shutdown, heat to mechanical conversion, and mechanical to electrical conversion collectively achieve the goal of safely generating electric power. Each of these functions, in turn, is achieved through the design and operating function from others. For example, emergency heat transport may require internal cooling, which is obtained from other so-called support functions.

The MLD clearly shows the interrelationships among the independent functions (or systems) and the independent support functions. The MLD (in success space) can show the manner in which various functions, subfunctions, and hardware interact to achieve the overall system objective. On the other hand, an MLD in a failure space can show the logical representation of the causes for failure of functions (or systems). The MLD (in success or failure space) can easily map the propagation of the effect of failures, i.e., establish the trajectories of event failure propagation.

In essence, the hierarchy of the MLD is displayed by the dependency matrix. For each function, subfunction, subsystem, and hardware item shown on the MLD, the effect of failure or success of all combinations of items is established and explicitly shown by •. Consider the MLD shown in a success space in Figure 4.22 [Modarres (1992)]. In this diagram, there are two major functions (or systems), F_1 and F_2.

Together, they achieve the system objective. Each of these functions, because of reliability concerns is further divided into two identical subfunctions, each of which can achieve the respective parent functions. This means that both subfunctions must be lost for F_1 or F_2 to be lost. Suppose the development of the subfunctions (or systems) can be represented by their respective hardware, which interface with other support functions (or support systems) S_1, S_2, and S_3. Support functions are those that help the main functions to be realized. For example, if a pump function is to "provide pressure," then functions "provide ac power," "cooling and lubrication," "activation and control" are called support functions. However, function (or system) S_1 can be divided into two independent subfunctions (or systems) ($S_{1\text{-}1}$ and $S_{1\text{-}2}$), so that each can interact independently with the subfunctions (or systems) of F_1 and F_2. The dependency matrix is established by reviewing the design specifications or operating manuals that describe the relationship between the items shown in the MLD, in which the dependencies are explicitly shown by •. For instance, the dependency matrix shows that failure of S_3 leads directly to failure of S_2 , which in turn results in failures of $F_{1\text{-}1}$, $F_{2\text{-}2}$ and $F_{2\text{-}1}$. This failure is highlighted on the MLD in Figure 4.22.

A key element in the development of an MLD is the assurance that the items for which the dependency matrix is developed (e.g., $S_{1\text{-}1}$, $S_{1\text{-}2}$, S_3, $F_{2\text{-}-}$, $F_{1\text{-}2}$, and

F_{2-2}) are all physically independent. "Physically independent" means they do not share any other system parts. Each element may have other dependencies, such as common cause failure (see Chapter 7). Sometimes it is difficult to distinguish *a priori* between main functions and supporting functions. In these situations, the dependency matrix can be developed irrespective of the main and supporting functions. Figure 4.23 shows an example of such a development. However, the main functions can be identified easily by examining the resulting MLD; they are those functions that appear, hierarchically, at the top of the MLD model and do not support other items.

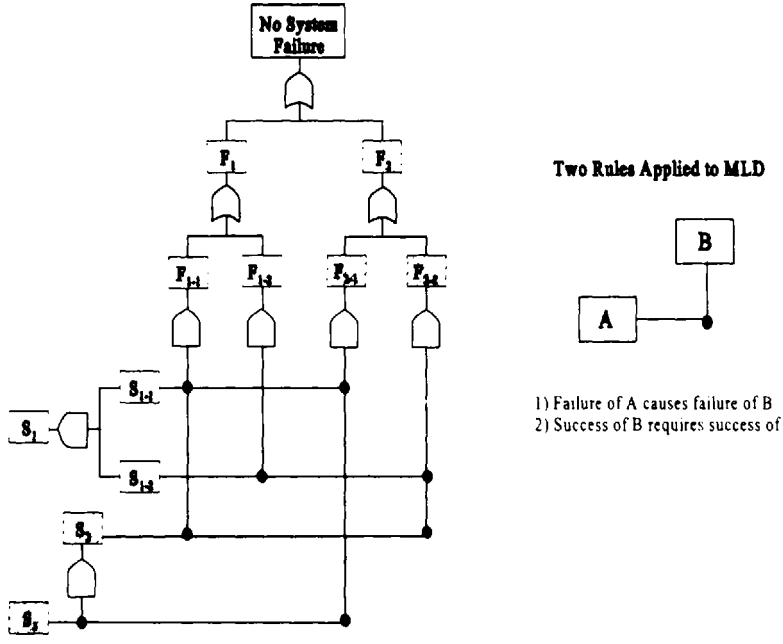


Figure 4.22 Master logic diagram showing the effect of failure of S_3.

The analysis of an MLD is straightforward. Using the combinatorial approach described earlier one must determine all possible 2^n combinations of failures of independent items (elements), map them onto the MLD and propagate their effects, using the MLD logic. The combinatorial approach discussed in Section 4.2.2 is the most appropriate method for that purpose, although the Boolean reduction method can also be applied. Table 4.2 shows the combinations for the example in Figure 4.22. For reliability calculations, one can combine those

end-state (effects) that lead to the system success. Suppose independent items (here, systems or subsystems) $S_{1\text{-}1}$, $S_{1\text{-}2}$, S_2, and S_3 have a failure probability of 0.01 for a given mission, and the probability of independent failure of $F_{1\text{-}1}$, $F_{1\text{-}2}$, $F_{2\text{-}1}$, and $F_{2\text{-}2}$ is also 0.01. Table 4.3 shows the resulting probability of the end-state effects. If needed, calculation of failure probabilities for the MLD items (e.g., subsystems) can proceed independent of the MLD, through one of the conventional system reliability analysis methods (e.g., fault free analysis).

Table 4.2 shows all possible mutually exclusive combinations of items modeled in the MLD with probability of failure greater than 1.0E–6, (i.e., $S_{1\text{-}1}$, $S_{1\text{-}2}$, S_2, S_3, $F_{1\text{-}1}$, $S_{1\text{-}2}$, and $F_{2\text{-}2}$.). Those combinations that lead to a failure of the system are mutually exclusive cut sets.

Table 4.2 Dominant Combinations of Failure and Their Respective Probabilities

Combination no. (i)	Failed items	Probability of failed items (and success of other items)	End State (actually failed and casually failed elements)
1	None	$9.3E-1$	Success
2	S_3	$9.3E-3$	$F_{1\ 1}, F_2, F_{2\ 1}, F_{2\ 2}, S_2, S_3$
3	$F_{2\ 2}, S_3$	$9.4E-5$	$F_{1\ 1}, F_2, F_{2\ 1}, F_{2\ 2}, S_2, S_3$
4	S_2, S_3	$9.4E-5$	$F_{1\ 1}, F_2, F_{2\ 1}, F_{2\ 2}, S_2, S_3$
5	$F_{1\ 1}, S_3$	$9.4E-5$	$F_{1\ 1}, F_2, F_{2\ 1}, F_{2\ 2}, S_2, S_3$
6	$F_{2\ 1}, S_3$	$9.4E-5$	$F_{1\ 1}, F_2, F_{2\ 1}, F_{2\ 2}, S_2, S_3$
7	$S_{1\ 2}$	$9.3E-3$	$F_{1\ 2}, F_{2\ 2}, S_{1\ 2}, S_1$
8	$F_{1\ 2}, S_{1\ 2}$	$9.4E-5$	$F_{1\ 2}, F_{2\ 2}, S_{1\ 2}, S_1$
9	$F_{2\ 2}, S_{1\ 2}$	$9.4E-5$	$F_{1\ 2}, F_{2\ 2}, S_{1\ 2}, S_1$
10	$S_{1\ 1}$	$9.3E-3$	$F_{1\ 2}, F_{2\ 2}, S_{1\ 2}, S_1$
11	$F_{2\ 1}, S_{1\ 1}$	$9.4E-5$	$F_{1\ 1}, F_{2\ 1}, S_{1\ 1}, S_1$
12	$F_{1\ 1}, S_{1\ 1}$	$9.4E-5$	$F_{1\ 1}, F_{2\ 1}, S_{1\ 1}, S_1$
13	S_2	$9.3E-3$	$F_{1\ 1}, F_{2\ 2}, S_2$
14	$F_{1\ 1}, S_2$	$9.4E-5$	$F_{1\ 1}, F_{2\ 2}, S_2$
15	$F_{2\ 2}, S_2$	$9.4E-5$	$F_{1\ 1}, F_{2\ 2}, S_2$
16	$F_{2\ 2}$	$9.3E-3$	$F_{2\ 2}$
17	$F_{2\ 1}$	$9.3E-3$	$F_{2\ 1}$
18	$F_{1\ 2}$	$9.3E-3$	$F_{1\ 2}$
19	$F_{1\ 1}$	$9.3E-3$	$F_{1\ 1}$
20	$S_{1\ 2}, S_3$	$9.4E-5$	$F_1, F_{1\ 1}, F_{1\ 2}, F_2, F_{2\ 1}, F_{2\ 2}, S_{1\ 2}, S_2, S_3, S_1$

Table 4.3 Combinations that Lead to Failure of the System

Combination no. (i)	Failure combination	Probability of state
1	$F_{1,2}, S_3$	9.4E - 5
2	$S_{1,1}, S_{1,2}$	9.4E - 5
3	$S_{1,2}, S_3$	9.4E - 5
4	$F_{2,1}, S_{1,2}, S_3$	9.5E - 7
5	$F_{2,2}, S_{1,1}, S_{1,2}$	9.5E - 7
6	$F_{2,1}, S_{1,2}, S_2$	9.5E - 7
7	$F_{2,1}, S_{1,1}, S_{1,2}$	9.5E - 7
8	$F_{1,1}, F_{2,1}, S_{1,2}$	9.5E - 7
9	$S_{1,1}, S_{1,2}, S_3$	9.5E - 7
10	$S_{1,2}, S_2, S_3$	9.5E - 7
	Total Probability of System Failure	2.9E - 4

Table 4.4 Combinations Leading to the System Failure When $S_{1,2}$ Is Known to Have Failed

Combination no. (i)	Failure combination	Probability of state
1	S_3	9.6E - 3
2	$S_{1,1}$	9.6E - 3
3	S_2, S_3	9.7E - 5
4	$F_{2,1}, S_3$	9.7E - 5
5	$F_{1,1}, F_{2,1}$	9.7E - 5
6	$F_{2,1}, S_2$	9.7E - 5
7	$F_{2,1}, S_{1,1}$	9.7E - 5
8	$S_{1,1}, S_2$	9.7E - 5
9	$F_{1,1}, S_3$	9.7E - 5
10	$S_{1,1}, S_3$	9.7E - 5
11	$S_{1,1}, S_{1,2}$	9.7E - 5
	Total Probability of System Failure	2.01E - 2

One may only select those combinations that lead to a complete failure of the system. The sum of the probabilities of occurrence of these combinations determines the failure probability of the system. If one selects the combinations that lead to the system's success, then the sum of the probabilities of occurrence of these combinations determines the reliability of the system. Table 4.3, for example, shows dominant combinations (those greater than 1.0E–7) that lead to the system's failure.

Another useful analysis that may be performed via MLD is the calculation of the conditional system probability of failure. In this case, a particular element of the system is set to failure, and all other combinations that lead to the system's failure may be identified. Table 4.4 shows all combinations within the MLD that lead to the system's failure, when element $S_{1\text{-}2}$ is set to failure.

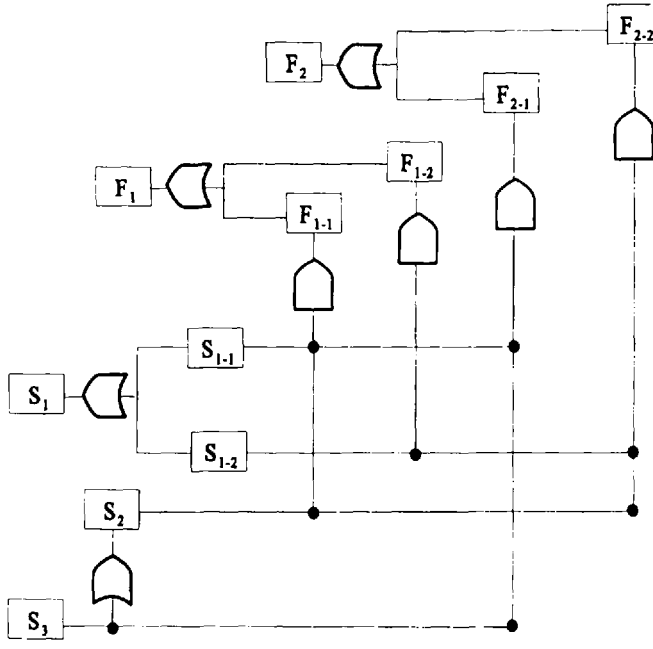


Figure 4.23 MLD with all system functions treated similarly.

Example 4.6

Consider the H-Coal process shown in Figure 4.24. In case of an emergency, a shutdown device (SDD) is used to shutdown the hydrogen flow. If the reactor temperature is too high, an emergency cooling system (ECS) is also needed to reduce the reactor temperature. To protect the process plant when the reactor temperature becomes too high, both ECS and SDD must succeed. The SDD and ECS are actuated by a control device. If the control device fails, the emergency cooling system will not be able to work. However, an operator can manually operate (OA) the shutdown device and terminate the hydrogen flow. The power for the SDD, ECS, and control device comes from an outside electric company (off-site power-OSP). The failure data for these systems are listed in Table 4.5. Draw an MLD and use it to find the probability of losing both the SDD and ECS.

Solution:

The MLD is shown in Figure 4.25. Important combinations of independent failures and their impacts on other components are listed in Table 4.6. The probability of losing both the ECS and SDD for each end state is calculated and listed in the third column of Table 4.7. Combinations that exceed 1×10^{-6} are included in Table 4.7. The combinations that could lead to failure of both SDD & ECS are shown in Table 4.7. Using (4.35), the probability of losing both systems is calculated as 4.99×10^{-3}.

Table 4.5 Failure Probability of Each System

System failure	Failure probability
OSP	2.0*E* - 2
OA	1.0*E* - 2
ACS	1.0*E* - 3
SDD	1.0*E* 3
ECS	1.0*E* - 3

Table 4.6 Leading Combination of Failure in the System

State no. (i)	Failed units	Probability*	End state
1	None	$9.94E-1$	Success
	OSP	$1.99E-3$	
	OSP, ECS	$1.99E-6$	
	OSP, SDD	$1.99E-6$	
2	OSP, ACS	$1.99E-6$	SDD, ECS, ACS, OSP
	OSP, ACS, SDD	$2.00E-9$	
	OSP, ECS, SDD	$2.00E-9$	
	OSP, ECS, ACS	$2.00E-9$	
3	ACS	$9.95E-4$	ECS, ACS
	ECS, ACS	$9.96E-7$	
4	SDD	$9.95E-4$	SDD
5	OA	$9.95E-4$	OA
6	ECS	$9.95E-4$	ECS
	OSP, OA	$1.99E-6$	
7	OSP, ACS, OA	$2.00E-9$	SDD, ECS, ACS
	OSP, ECS, OA	$2.00E-9$	OSP, OA
8	OSP, SDD, OA	$2.00E-9$	
	ECS, OA	$9.96E-7$	ECS, OA
9	ACS, OA	$9.96E-7$	SDD, ECS, ACS, OA
10	ACS	$9.96E-7$	SDD

*Includes probability of success of elements not affected.

Table 4.7 Probability of Losing Two Systems

Combination no.	Units failed	Probability*	Contribution to total prob. (%)
1	OSP	$1.99E-3$	39.9
2	ACS	$9.95E-4$	19.9
3	ECS	$9.95E-4$	19.9
4	SDD	$9.95E-4$	19.9
5	OSP and ECS	$1.99E-6$	Negligible
6	OSP and SDD	$1.99E-6$	Negligible
7	OSP and ACS	$1.99E-6$	Negligible
8	OSP and OA	$1.99E-6$	Negligible

*Includes probability of success of elements not affected.

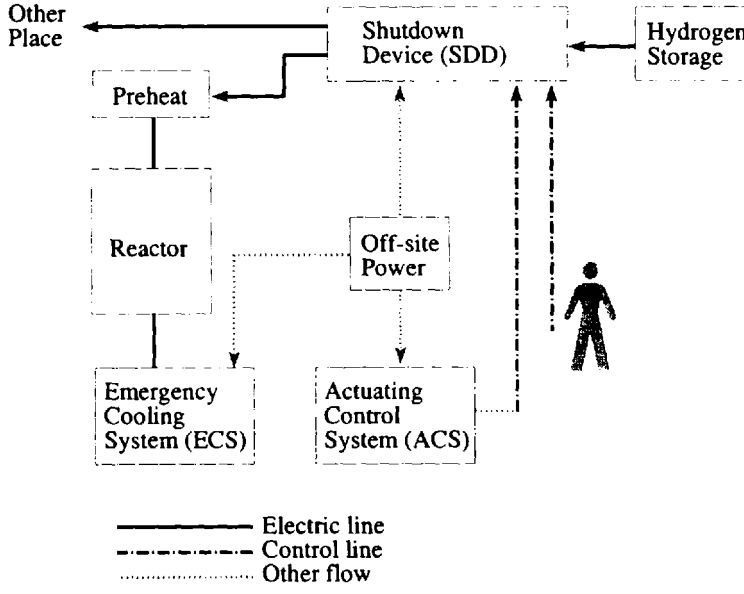


Figure 4.24 Simplified diagram of the safety systems..

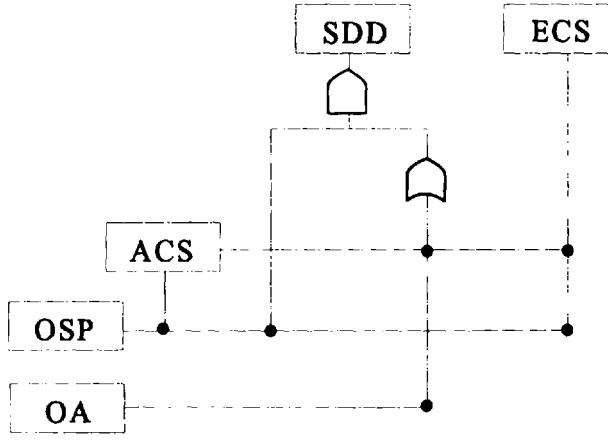


Figure 4.25 MLD for the safety system in Figure 4.24.

Example 4.7

The simple event tree shown in Figure 4.26 has 5 events (A, B, C, D, and E) which make up the headings of the event tree. The initiating event is labeled I.

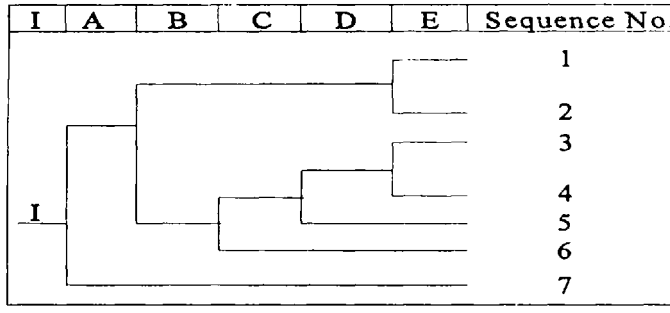


Figure 4.26 Simple event tree.

Consider sequence No. 5, which is highlighted with a bold line. The logical equivalent of the sequence is:

$$S_5 = I \cdot \overline{A} \cdot B \cdot \overline{C} \cdot D$$

where, S_5 is the 5th sequence and I is the initiating event. Develop an equivalent MLD representation of this event tree.

Solution:

Sequence 5 occurs when the expression $\overline{A} \cdot B \cdot \overline{C} \cdot D$ is true. Note that the above Boolean expression involves two failed elements (i.e., B and D). We can express these terms, in the success space, through the complement of $\overline{A} \cdot B \cdot \overline{C} \cdot D$, which is

$$\begin{aligned} \overline{\overline{A} \cdot B \cdot \overline{C} \cdot D} &= \overline{(\overline{A} \cdot \overline{C}) \cdot (B \cdot D)} \\ &= \overline{(\overline{A} \cdot \overline{C})} + \overline{(B \cdot D)} \\ &= \overline{(\overline{A} \cdot \overline{C})} + (\overline{B} + \overline{D}) \end{aligned}$$

The last expression represents every event in a success space (e.g., $\overline{A} \cdot \overline{B} \cdot \overline{C}$) and its equivalent MLD logic is shown in Figure 4.27.

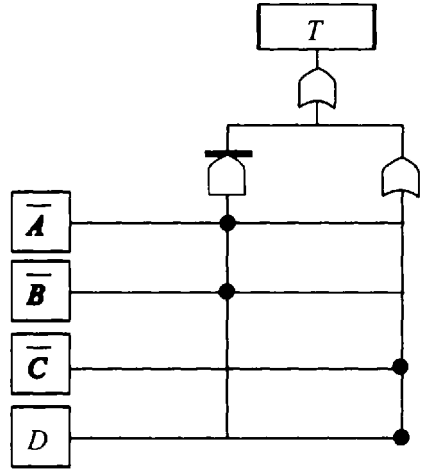


Figure 14.27 MLD equivalent of event tree shown in Figure 4.26.

4.5 FAILURE MODE AND EFFECT ANALYSIS

Failure mode and effect analysis (FMEA) is a powerful technique for reliability analysis. This method is inductive in nature. In practice, it is used in all aspects of system failure analysis from concept to implementation. The FMEA analysis describes inherent causes of events that lead to a system failure, determines their consequences, and devises methods to minimize their occurrence or recurrence.

The FMEA proceeds from one level or a combination of levels of abstraction, such as system functions, subsystems, or components. The analysis assumes that a failure has occurred. The potential *effect* of the failure is then postulated and its potential causes are identified. A *criticality* or the *risk priority number* (RPN) rating may also be determined for each failure mode and its resulting effect. The rating is normally based on the probability of the failure *occurrence*, the *severity* of its effect(s), and its *detectability*. Failures that score high in this rating represent areas of greatest risk, and their causes should be mitigated.

Although the FMEA is an essential reliability task for many types of system design and development processes, it provides very limited insight into probabilistic representation of system reliability. Another limitation is that FMEA is performed for only one failure at a time. This may not be adequate for systems in which multiple failure modes can occur, with reasonable likelihood, at the same time. (Deductive methods are very powerful for identifying these kind of failures.) However, FMEA provides valuable qualitative information about the system design and operation. An extension of FMEA is called Failure Mode and Effect Criticality Analysis (FMECA), which provides more quantitative treatment of failures.

The FMEA was first developed by the aerospace industry in the mid-sixties. The standard reference is US MIL-STD-1629A (1980). Since then, the method has been adopted by many other industries, which have modified it to meet their needs. For example, the automotive industry uses the FMEA refined by the Society of Automotive Engineers (SAE) recommended Practice J1739 (1994) of FMEA application. The methods of FMEA and FMECA are briefly discussed in this section. For more information, the readers are referred to the above mentioned publications.

4.5.1 Types of FMEA

Depending on the stage in product development, one may perform two types of FMEA (SAE Recommended Practice J1739 (1994)): *design FMEA* and *process FMEA*.

Design FMEA is used to evaluate the failure modes and their effects for a product before it is released to production and is normally applied at the subsystem and the component abstraction levels. The major objectives of a design FMEA are:

1. identify failure modes and rank them according to their effect on the product performance, thus establishing a priority system for design improvements;
2. identify design actions to eliminate potential failure modes or reduce the occurrence of the respective failures;
3. document the rationale behind product design changes and provide future reference for analyzing field concerns, evaluating new design changes and developing advanced designs.

Process FMEA is used to analyze manufacturing and assembly processes. The major objectives of a process FMEA are:

1. identify failure modes that can be associated with manufacturing or assembly process deficiencies;

2. identify highly critical process characteristics that may cause particular failure modes;
3. identify the sources of manufacturing/assembly process variation (equipment performance, material, operator, environment) and establish the strategy to reduce it.

4.5.2 FMEA/FMECA Procedure

Outlined below is a logical sequence of steps by means of which FMEA/FMECA is usually performed.

Define the system to be analyzed. Identify the system decomposition (indenture) level, which will be subject to analysis. Identify internal and interface system functions, restrains, develop failure definitions.

Construct a block diagram of the system. Depending on system complexity and the objectives of the analysis, consider at least one of these diagrams: structural (hardware), functional, combined, master logic diagram (MLD). (The latter method is considered in greater detail in Section 4.4.)

Identify all potential item failure modes and define their effects on the immediate function or item, on the system, and on the mission to be performed.

Evaluate each failure mode in terms of the worst potential consequence, which may result and assign a severity classification category.

Identify failure detection methods and compensating provision(s) for each failure mode.

Identify corrective design or other actions required to eliminate the failure or control the risk.

Document the analysis and identify the problems, which could not be corrected by design.

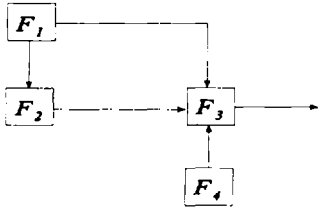
4.5.3 FMEA Implementation

FMEA for Aerospace Applications

The FMEA is usually performed using a tabular format. A worksheet implementation of a typical MIL-STD-1629A FMEA procedure is shown in Table 4.8. The major steps of the analysis are described below.

System Description and Block Diagrams. It is important to first describe the system in a manner that allows the FMEA to be performed efficiently and

understood by others. This description can be done in different levels of abstraction. For example, at the highest level (i.e., the functional level), the system can be represented by a *functional block diagram*. The functional block diagram is different from the *reliability block diagram* discussed earlier in this chapter. Functional block diagrams illustrate the operation, interrelationship, and interdependence of the functional entities of a system. For example, the pumping system of Figure 4.10 can be represented by its functional block diagram, as shown in Figure 4.28. In this figure, the components that support each system function are also described.



Function	Functional Description	Components Involved
F_1	Provide AC Power	AC
F_2	Sensing and Control	S
F_3	Provide Pumping	V-2,V-3, V-4 V-5 ,P-1, P-2
F_4	Maintain Source	T-1, V-1

Figure 4.28 Functional block diagram for the pumping system.

Item/Functional Identification. Provide the descriptive name and the nomenclature of the item under analysis. If the failures are postulated at a lower abstraction level, such levels should be shown. A fundamental item of current FMEA may be subject to a separate FMEA, which further decomposes this item into more basic parts. The lower the abstraction level, the greater the level of detail required for the analysis. This step provides necessary information for the *identification number, functional identification (nomenclature),* and *function* columns in the FMEA.

Failure Modes and Causes and Mission Phase/Operational Mode. The manner of failure of the function, subsystem, component, or part identified in the second column of the table is called the failure-mode and is listed in the *failure mode and causes* column of the FMEA table. The causes (a failure mode can have

more than one cause) of each failure mode should also be identified and listed in this column. The failure modes applicable to components and parts are often known *a priori*. Typical failure modes for electronic components are *open, short, corroded, drifting, misaligned*, etc. Some representative failure modes for mechanical components include: *deformed, cracked, fractured, sticking, leaking, and loosened*. However, depending on the specific system under analysis, the environmental design, and other factors, only certain failure modes may apply. This should be known and specified by the analyst.

Failure Effects. The consequences of each failure mode on the item's operation should be carefully examined and recorded in the column labeled *failure effects*. The effects can be distinguished at three levels: *local, next higher abstraction level*, and *end effect*. Local effects specifically show the impact of the postulated failure mode on the operation and function of the item under consideration. The consequence of each failure mode on the operation and functionality of an item under consideration is described as its local effect. It should be noted that sometimes no local effects can be described beyond the failure mode itself. However, the consequences of each postulated failure on the output of the item should be described along with second order effects. End-effect analysis describes the effect of postulated failure on the operation, function, and status of the next higher abstraction level and ultimately on the system itself. The end effect shown in this column may be the result of multiple failures. For example, the failure of a supporting subsystem in a system can be catastrophic if it occurs along with another local failure. These cases should be clearly recognized and discussed in the end-effect column.

Failure Detection Method. Failure detection features for each failure mode should be described. For example, previously known symptoms can be used based on the item's behavior pattern(s) indicating that a failure has occurred. The described symptom can cover the operation of a component under consideration (logical symptom) or can cover both the component and the overall system, or equipment evidence of failure.

Compensating Provision. A detected failure should be corrected so as to eliminate its propagation to the whole system so as to maximize reliability. Therefore, at each abstraction level provisions that will alleviate the effect of a malfunction or failure should be identified. These provisions include such items as: a) redundant elements for continued and safe operation, b) safety devices, and c) alternative modes of operation, such as backup and standby units. Any action that may require operator action, should be clearly described.

Severity. Severity classification is used to provide a qualitative indicator of the worst potential effect resulting from the failure mode. For the FMEA purposes, MIL-STD-1629A classifies severity levels in the following categories:

Table 4.8 US MIL-STD-1629A FMEA Worksheet Format

FAILURE MODE AND EFFECTS ANALYSIS

System _____
Indenture level _____
Reference drawing _____
Mission _____

Date _____
Sheet ____ of _____
Compiled by _____
Approved by _____

IDENTIFICATION NUMBER	ITEM/FUNCTIONAL IDENTIFICATION (NOMENCLATURE)	FUNCTION	FAILURE MODES AND CAUSES	MISSION PHASE/ OPERATIONAL MODE	FAILURE EFFECTS			FAILURE DETECTION METHOD	COMPENSATING PROVISIONS	SEVERITY CLASS	REMARKS
					LOCAL EFFECTS	NEXT HIGHER LEVEL	END EFFECT				

Effect	Rating	Criteria
Catastrophic	1	A failure mode that may cause death or complete mission loss.
Critical	2	A failure mode that may cause severe injury or major system degradation, damage, or reduction in mission performance.
Marginal	3	A failure that may cause minor injury or degradation in system or mission performance.
Minor	4	A failure that does not cause injury or system degradation but may result in system failure and unscheduled maintenance or repair.

Remarks. Any pertinent information, clarifying items, or notes should be entered in the column labeled *remarks*.

FMEA for Transportation Applications

The SAE J1739 FMEA procedure is, in principle, similar to the above reviewed MIL-STD-1629A FMEA. However, some definitions and ratings differ from those discussed so far. The key criteria for identifying and prioritizing potential design deficiencies here is the risk priority number defined as the product of the severity, occurrence and detection ratings. An example of a SAE J1739 FMEA format is shown in Table 4.9.

The content of the *Item/Function, Potential Failure Mode, Potential Effect(s) of Failure, Potential Cause(s)/Failure Mechanism(s)* and the *Recommended Actions* steps of this FMEA procedure is similar to the respective parts of the MIL-STD-1629A FMEA discussed above.

Severity is evaluated on a ten-grade scale as shown in the table below. Note that in contrast to the MIL-STD-1629A FMEA, a higher rating here corresponds to a higher severity (and, consequently, a higher RPN).

Effect	Rating	Criteria
Hazardous	10	Safety related failure modes causing noncompliance with government regulations without warning
Serious	9	Safety related failure modes causing noncompliance with government regulations with warning
Very high	8	Failure modes resulting in loss of primary vehicle/system/component function.
High	7	Failure modes resulting in a reduced level of vehicle/ system/component performance and customer dissatisfaction.
Moderate	6	Failure modes resulting in loss of function by comfort/convenience systems/components.
Low	5	Failure modes resulting in a reduced level of performance of comfort/convenience systems/ components.
Very low	4	Failure modes resulting in loss of fit and finish, squeak and rattle functions.
Minor	3	Failure modes resulting in partial loss of fit and finish, squeak and rattle functions.
Very minor	2	Failure modes resulting in minor loss of fit and finish, squeak and rattle functions.
None	1	No effect.

Occurrence is defined as the likelihood that a specific failure cause/mechanism will occur. The rating is based on the estimated or expected failure frequency as shown in the table below.

Likelihood of failure	Estimated or expected failure frequency	Rating
Very high (failure is almost inevitable)	> 1 in 2	10
	1 in 3	9
High (frequently repeated failures)	1 in 8	8
	1 in 20	7
Moderate (occasional failures)	1 in 80	6
	1 in 400	5
	1 in 2000	4
Low (rare failures)	1 in 15,000	3
	1 in 150,000	2
Remote (failures are unlikely)	< 1 in 150,000	1

Current Design Controls. Before the design is finalized and released to production, the engineer has a complete control over it in terms of possible design changes. Three types of *design control* are usually considered, those that: (1) prevent the failure cause/mechanism or mode from occurring or reduce their rate of occurrence, (2) detect the cause/mechanism and lead to corrective actions, or (3) detect the failure mode.

The preferred approach is to first use type 1 controls, if possible; second, use the type 2 controls; and third, use type 3 controls. The initial occurrence ranking are affected by the type 1 controls, provided they are integrated as a part of the design intent. The initial detection rankings are based on the type 2 or 3 controls, provided the prototypes and models being used are representative of design intent.

Detection is defined as the ability of the proposed type 2 design controls to detect a potential cause/mechanism, or the ability of the proposed type 3 design controls to detect the respective failure mode before the system/component is released to production.

Risk Priority Number is the product of the Severity, Occurrence and Detection ratings and is used to rank the order of potential design concerns. While the RPN is a major measure of design risk, special attention should be given to the high severity failure modes irrespective of the resultant RPN number.

Detection	Rating	Criteria
Uncertain	10	Design control will not and/or can not detect a potential cause/mechanism and subsequent failure mode.
Very remote	9	Very remote chance the design control will detect a potential cause/mechanism and subsequent failure mode.
Remote	8	Remote chance the design control will detect a potential cause/mechanism and subsequent failure mode.
Very low	7	Very low chance the design control will detect a potential cause/mechanism and subsequent failure mode.
Low	6	Low chance the design control will detect a potential cause/mechanism and subsequent failure mode.
Moderate	5	Moderate chance the design control will detect a potential cause/mechanism and subsequent failure mode.
Moderately high	4	Moderately high chance the design control will detect a potential cause/mechanism and subsequent failure mode.
High	3	High chance the design control will detect a potential cause/mechanism and subsequent failure mode.
Very high	2	Very high chance the design control will detect a potential cause/mechanism and subsequent failure mode.
Almost certain	1	The design control will almost certainly detect a potential cause/mechanism and subsequent failure mode.

Table 4.9 SAE J1739 FMEA Worksheet Format

Potential Failure Mode and Effects Analysis (Design FMEA)

System: _____ FMEA Number: _____

Subsystem: _____ Page ___ of ___

Component: _____ Design Responsibility: _____ Prepared by: _____

Model Year / Vehicle(s): _____ Key Date: _____ FMEA Date (Orig): _____ (Rev.) _____

Core Team: _____

Item / function	Potential failure mode	Potential effect(s) of failure	Severe	Potential cause(s) / failure mechanism(s)	Occur	Current design controls	Detect	RPN	Recommended actions	Responsibility and target completion date	Action Results				
											Actions taken	Severe	Occur	Detect	RPN

Table 4.10 FMEA in Example 4.8

Potential Failure Mode and Effects Analysis (Design FMEA)

_x_ System
___ Subsystem
___ Component: Generic Front Lighting System
Model Year / Vehicle(s): 2000/LITTLE TRUCKS
Core Team:

Design Responsibility: Electrical Engineering
Key Date:

FMEA Number:
Page _1_ of _5_
Prepared by:
FMEA Date (Orig): __97.02_____ (Rev.) _____

Item / function	Potential failure mode	Potential effect(s) of failure	Severe	Potential cause(s) / failure mechanism(s)	Occur	Current design controls	Detect	RPN	Recommended Actions	Responsibility and target completion date	Action Results: Actions taken	Sevee	Occur	Detect	RPN
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Provide Illumination for vehicle's line of travel, as defined by a. beam width b. intensity c. vertical aim d. horizontal aim	System does not provide adequate illumination including high beam and low beam.	Customer dissatisfaction and/or noncompliance with government regulation(s).	9	Inadequate reflector size	1	System analysis modeling vehicle integration testing	2	18							
				Defective bulb	5	Supplier bulb durability testing Lighting system durability testing Vehicle durability testing	3	135	Pursue CBA on high reliability bulb	John Doe 11/98	Engineering change to 104317 at var. cost penalty of $.45	9	1	3	27
				Defective wiring harness — bulb circuit (includes MPC and bulb connector)	2	Supplier bulb durability testing Lighting system durability testing Vehicle durability testing	3	54							

Table 4.10 Continued

Item / function	Potential failure mode	Potential effect(s) of failure	Severe	Potential cause(s) / failure mechanism(s)	Occur	Current design controls	Detect	RPN	Recommended Actions	Responsibility and target completion date	Action Results				
											Actions taken	Severe	Occur	Detect	RPN
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
				Inadequate vertical alignment setting specified	1	Specification review assembly drawing review	1	9							
				Inadequate horizontal alignment setting specified (includes tolerances)	2	Specification review assembly drawing review	1	18							
Provide directional (turn) signals	System does not provide adequate turn signal indication	Noncompliance with government regulation(s)	9	Incorrect reflector size	2	SAM — sys. anal. model vehicle integration testing	2	36							
				Defective bulb	1	Supplier bulb durability testing Lighting system durability testing Vehicle durability testing	3	27							
				Defective socket	2	Supplier bulb durability testing Lighting system durability testing Vehicle durability testing	3	54							

Provide a lighted indication of vehicle's position while parked	System does not provide adequate parking indication.	Noncompliance with government regulation(s).	9	Incorrect reflector geometry	3	System analysis modeling vehicle integration testing	2	54							
				Defective position bulb	3	Supplier bulb durability testing Lighting system durability testing Vehicle durability testing	3	81							
				Defective position socket	1	Supplier bulb durability testing Lighting system durability testing Vehicle durability testing	3	27							
				Defective wiring harness—position circuit	2	Supplier bulb durability testing Lighting system durability testing Vehicle durability testing	4	72							
Adjusts beam elevation when commanded by driver to compensate for load effects on vehicle attitude.	Driver is unable to adjust beam for load conditions, or control is inadequate.	Driver's ability to see the road may not be optimal and noncompliance with govt. regulations.	5	Defective alignment motor/shaft	1	Motor and shaft durability testing Lighting system durability testing Vehicle durability testing	5	25							
				Broken (molded) attachment points	1										

Action Results columns describe the implemented corrective actions along with the estimated reduction in Severity, Occurrence and Detection rating and the resultant RPN.

Example 4.8

Based on the functional block diagram of the vehicle generic front lighting system (see Figure 4.29) , develop a design FMEA on the system abstraction level.

Solution:

The FMEA of the vehicle generic front lighting system is shown in Table 4.10. As seen from the table, the highest RPN corresponds to the failure mode potentially caused by a defective light bulb. The corrective action of pursuing the CBA (cost-benefit analysis) on a more reliable bulb reduces the occurrence rating of this failure mode from 5 to 1, which, in turn, decreases the RPN to 27.

4.5.4 FMECA Procedure: Criticality Analysis

Criticality analysis is the combination of a probabilistic determination of a failure mode occurrence combined with the impact it has on the system mission success. Table 4.11 shows an example of a criticality analysis worksheet format. The criticality analysis part of this worksheet is explained below.

Failure Effect Probability β. The β value represents the conditional probability that the failure effect with the specified criticality classification will occur given that the failure mode occurs. For complex systems, β is difficult to determine unless a comprehensive logic model of the system (e.g., a fault tree or an MLD) exists. Therefore, in many cases, estimation of β becomes primarily a matter of judgement greatly driven by the analyst's prior experience. The general guidelines shown in Table 4.12 can be used for determining β.

Failure Mode Ratio α. The fraction of the item (component, part, etc.) failure rate, λ, related to the particular failure mode under consideration is evaluated and recorded in the *failure mode ratio* (α) column. The failure mode ratio is the probability that the item will fail in the identified mode of failure. If all potential failure modes of an item are listed, the sum of their corresponding α values should be equal to 1. The values of α should normally be available from a data source (e.g, MIL-STD-338). However, if not available, the values can be assessed based on the analyst's judgement.

Table 4.11 FMECA Worksheet Format

CRITICALITY ANALYSIS

System _____ Date _____
Indenture level _____ Sheet ____ of _____
Reference drawing _____ Compiled by _____
Mission _____ Approved by _____

IDENTIFICATION NUMBER	ITEM/ FUNCTIONAL IDENTIFICATION (NOMENCLATURE)	FUNCTION	FAILURE MODES AND CAUSES	MISSION PHASE/ OPERATIONAL MODE	SEVERITY CLASS	FAILURE PROBABILITY FAILURE RATE DATA SOURCE	FAILURE EFFECT PROBABILITY (β)	FAILURE MODE RATIO (α)	FAILURE RATE (λ - p)	OPERATING TIME (T)	FAILURE MODE CRIT # $C_m = \beta\alpha\lambda_p t$	ITEM CRIT # $C_r = \Sigma(C_m)$	REMARK

Table 4.12 Failure Effect Probabilities for Various Failure Effects

Failure effect	β value
Actual loss	1.00
Probable loss	$0.1 < \beta \leq 1.0$
Possible loss	$0 < \beta \leq 0.1$
No effect	0

Failure Rate λ. The generic or specific failure rate for each failure mode of the item should be obtained and recorded in the *failure rate (λ)* column. The estimates of λ can be obtained from the test or field data, or from generic sources of failure rates discussed in Section 3.7.

Operating Time T. The operating time, in hours, or the number of operating cycles of the item should be listed in the corresponding column.

Failure Mode Criticality Number C_r is used to rank each potential failure mode based on its occurrence and the consequence of its effect. For a particular severity classification, the C_r of an item is the sum of the failure mode criticality numbers C_m that have the same severity classification.

Thus,

$$C_r = \sum_{i=1}^{n} (C_m)_i$$

and

$$C_m = \beta \alpha \lambda T$$

where C_m is the criticality of an individual failure mode, and n is the number of failure modes of an item with the same severity classification.

Based on the criticality number, a so-called *criticality matrix* is usually developed to provide a visual way of identifying and comparing each failure mode to all other failures with respect to severity. Figure 4.30 shows an example of such a matrix. This matrix can also be used for a qualitative criticality analysis in a FMEA-type study. Along the vertical dimension of the matrix, the probability of occurrence level (subjectively estimated by the analyst in an FMEA study) or the criticality number C_r (calculated in an FMECA study) is entered. Along the

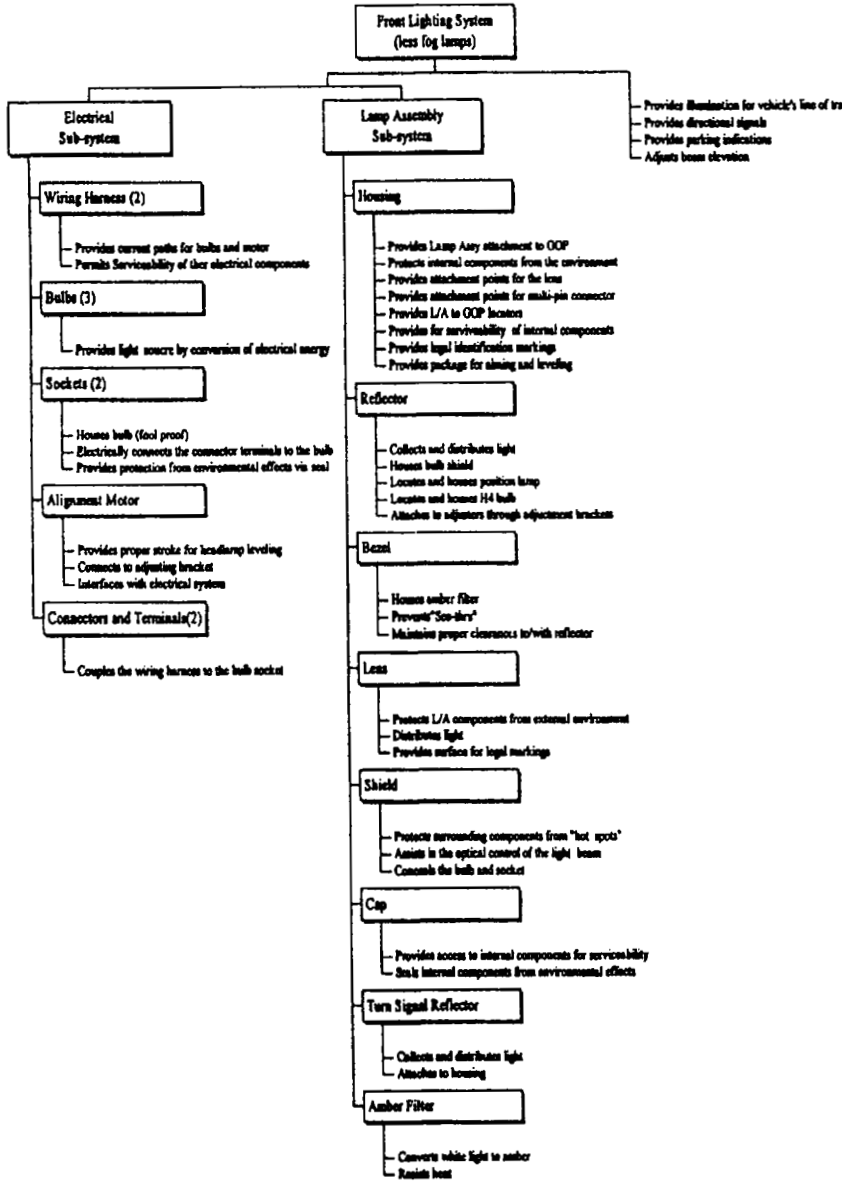


Figure 4.29 Hierarchic breakdown of the front lighting system.

Table 4.13 FMECA for the Amplifier System in Example 4.9

INDENT NO.	CKT NAME	FAILURE MODE	EFFECTS		SEVERITY CLASS	β	α	λ	MISSION TIME (HOURS)	CRITICALITY NO. Cm	REMARKS
			LOCAL	SYSTEM							
1.	A.	a) Open b) Short c) Other	Circuit A failure Both A & B circuit failure A lost	Degraded Failure Degraded	III II IV	0.069[1] 1.00 0.0093[2]	.90 .05 .05	1×10^{-4}	72	4.47×10 3.60×10^{-4} 3.35×10^{-5}	May cause secondary failure.
2.	B.	a) Open b) Short c) Other	Circuit B failure Both A & B circuit failure B lost	Degraded Failure Degraded	III II IV	0.069 1.00 0.0093	.90 .05 .05	1×10^{-4}	72	4.47×10 3.60×10^{-3} 3.35×10^{-5}	May cause secondary failure
$\Sigma\lambda_{II} = 1.8 \times 10^{-4}$, $\Sigma\lambda_{III} = 1 \times 10^{-4}$, $\Sigma\lambda_{IV} = 1 \times 10^{-4}$								$\Sigma\lambda =$ 2×10^{-4}	$C_t = \Sigma C_m = 16.21 \times 10^{-3}$		

Note: 1. Pr (System failure | A open) = β for "A" open mode of failure = 1 $R_A(72) = 1$ exp($1 \times 10^{-4} \times 72$) = 1 0.931 = 0.069

2. Assume failure rate doubles due to degradation: R_A = exp ($2 \times 10^{-4} \times 72$) = 0.866, then Pr (System failure | A degraded) = 1 [0.931 + 0.866 (0.931)(0.866)] = 1 0.99075 = 0.00925.

horizontal dimension of the matrix, the severity classification of an effect is entered. The severity increases from left to right. Each item on the FMEA or FMECA could be represented by one or more points on this matrix. If the item's failure modes correspond to more than one severity effect, each failure mode will correspond to a different point in the matrix. Clearly, those severities that fall in the upper-right quadrant of the matrix require immediate attention for reliability or design improvements.

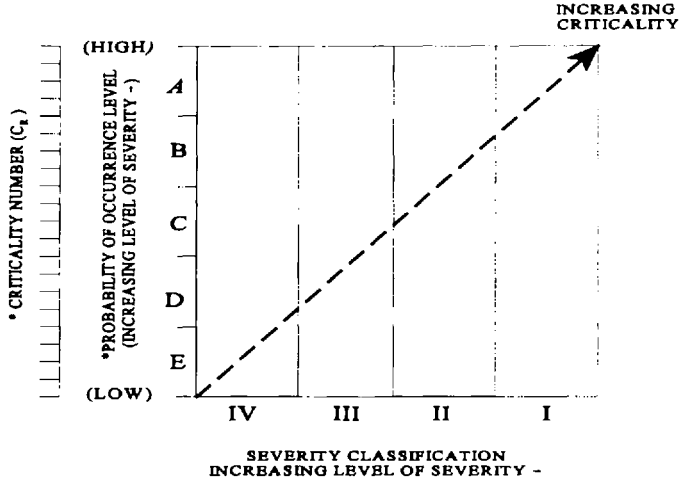


Figure 4.30 Example of Criticality Matrix. *Note: both criticality number (C_r) and probability of occurrence level are shown for convenience.

Example 4.9

Develop FMECA for a system of two amplifiers A and B in parallel configuration. In a given mission, this system should function for a period of 72 hours.

Solution:

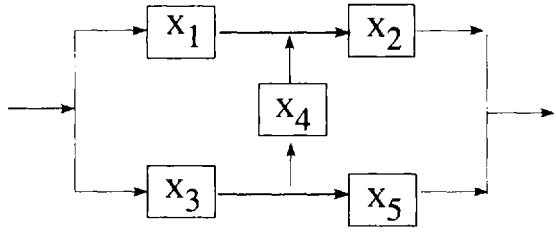
The summary of this analysis is displayed in Table 4.13. One can draw the following conclusions for this mission of the system.

1. The system will be expected to critically fail with a probability of 0.0036 + 0.0036 = 0.0072.
2. The system will experience a failure resulting in system degradation with a probability of $3.35E-5 \times 2 = 6.7E-5$.
3. The system will experience a critical failure due to "open" circuit failure mode with a probability of $4.47 \times 10^{-3} \times 2 = 8.94 \times 10^{-3}$.

The above approximate probabilities can only hold true if the product of α, β, λ, and T is small (e.g., < 0.1). Normally, criticality numbers are used as a measure of severity and not as a prediction of system reliability. Therefore, the most effective design would allocate more engineering resources to the areas with high criticality numbers, and on minimizing the Class I and Class II severity failure modes

EXERCISES

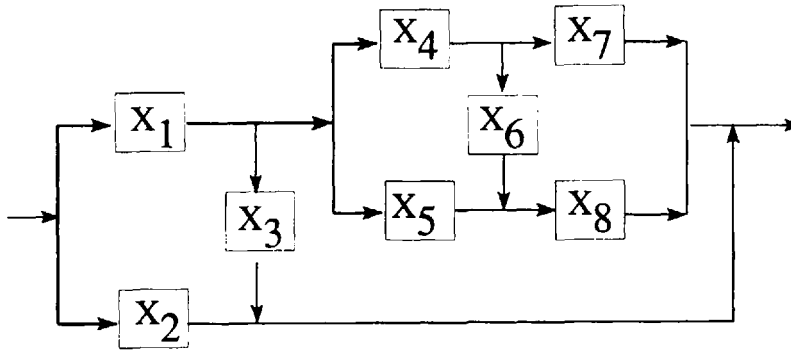
4.1 Consider the circuit below:



Assume the reliability of each unit $R(x_i) = \exp(-\lambda_i t)$, and $\lambda_i = 2.0E-4$ hr^{-1} for all i. Find the following:

a) Minimal path sets.
b) Minimal cut sets.
c) MTTF.
d) Reliability of the system at 1000 hours.

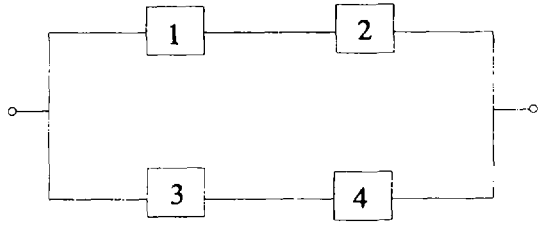
4.2 Consider the circuit below:



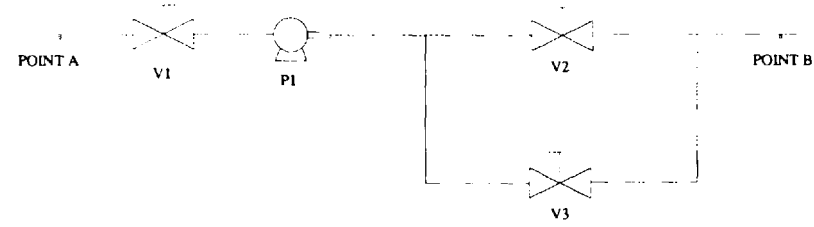
Find the following:

a) Minimal path sets.
b) Minimal cut sets.
c) Reliability of the system at 1000 hours.
d) Probability of failure at 1000 hours, using cut sets, to verify results from c.
e) Accuracy of the results of d) and/or c), using an approximate method.
f) MTTF of the system.

4.3 Calculate the reliability of the system shown in the figure below for a 1,000-hour mission. What is the MTTF for this system?

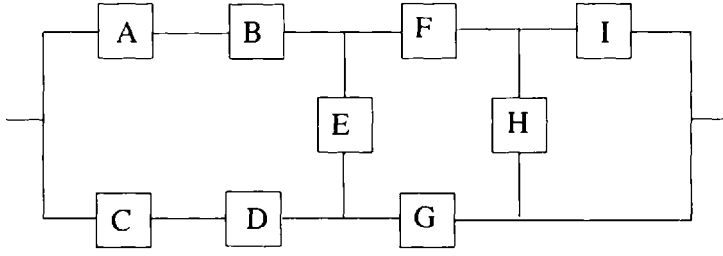


4.4 Consider the piping system shown in the figure below. The purpose of the system is to pump water from point *A* to point *B*. The time to failure of all the valves and the pump can be represented by the exponential distributions with failure rates λ_v and λ_p respectively.



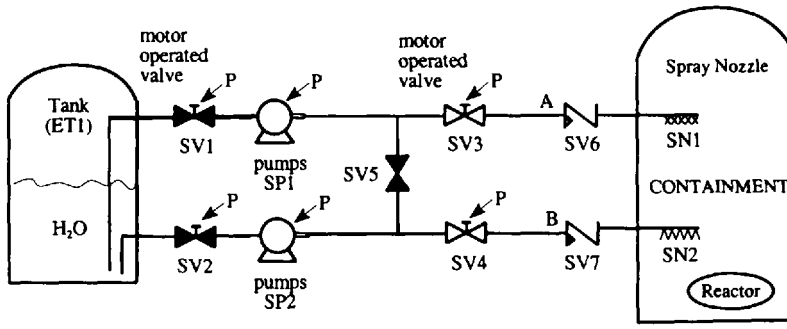
a) Calculate the reliability function of the system.
b) If $\lambda_v = 10^{-3} hr^{-1}$ and $\lambda_p = 2 \times 10^{-3} hr^{-1}$, and the system has survived for 10 hours, what is the probability that it will survive another 10 hours?

4.5 Estimate reliability of the system represented by the following reliability block diagram for a 2500-hour mission. Assume a failure rate of 10^{-6}/ hour for each unit.



4.6 A containment spray system is used to scrub and cool the atmosphere around a nuclear reactor during an accident. Develop a fault tree using "No H_2O spray" as the top event. Assume the following conditions:

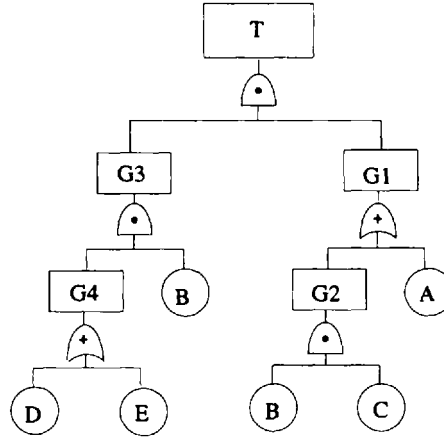
There are no secondary failures.
There is no test and maintenance.
There are no passive failures.
There are independent failures.
One of the two pumps and one of the two spray heads is sufficient to provide spray. (Only one train is enough.)
One of the valves sv_1 or sv_2 is opened after demand. However, sv_3 and sv_4 are always normally open.
Valve sv_5 is always in the closed position.



There is no human error.

SP_1, SP_2, sv_1, sv_2, sv_3, and sv_4 use the same power source P to operate.

4.7 Consider the fault tree below.

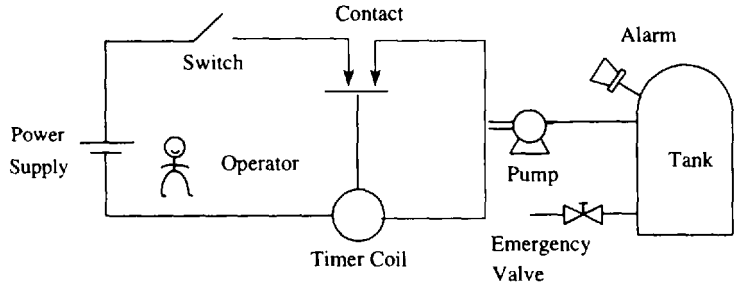


Find the following:

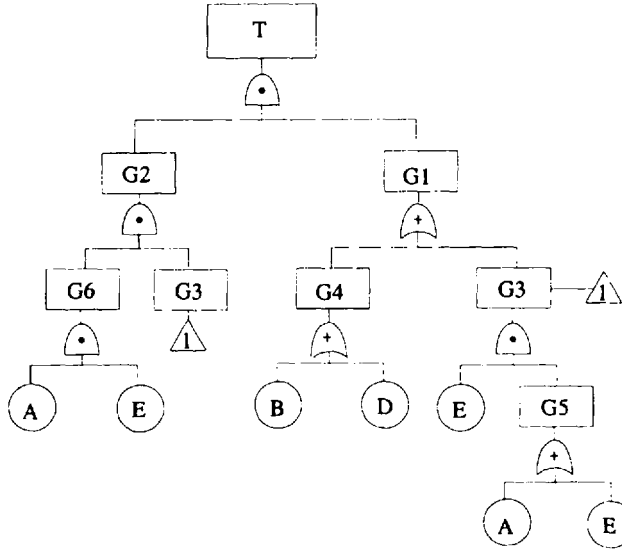
a) Minimal cut sets
b) Minimal path sets
c) Probability of the top event if the following probabilities apply:

$$\Pr(A) = \Pr(C) = \Pr(E) = 0.01$$
$$\Pr(B) = \Pr(D) = 0.0092$$

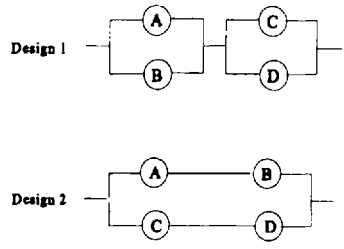
4.8 Consider the pumping system below. System cycles every hour. Ten minutes are required to fill the tank. Timer is set to open contact 10 minutes after switch is closed. Operator opens switch or the tank emergency valve if he/she notices an overpressure alarm. Develop a fault tree for this system with the top event "Tank ruptures."



4.9 Find the cut sets and path sets of the fault tree shown below using the top-down substitution method.



4.10 Compare Design 1 and Design 2 below.

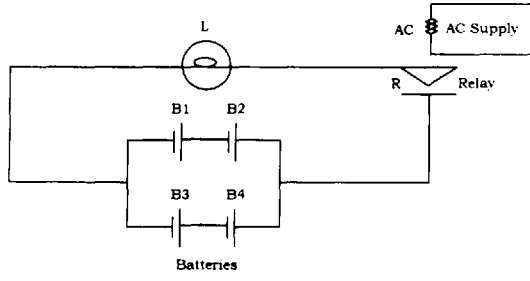


Failure rates (hr^{-1})

$$\lambda_A = 10^{-6} \quad \lambda_C = 10^{-3} \quad \lambda_B = 10^{-6} \quad \lambda_D = 10^{-3}$$

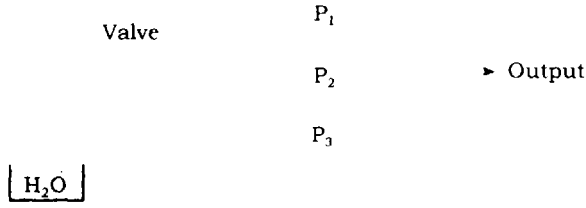
a) Assume that the components are nonrepairable. Which is the better design?
b) Assume that the system failure probability cannot exceed 10^{-2}. What is the operational life for Design 1 and for Design 2.

4.11 Consider the following electric circuit for providing emergency light during a blackout.



In this circuit, the relay is held open as long as ac power is available, and either of the four batteries is capable of supplying light power. Start with the top event "No Light When Needed."

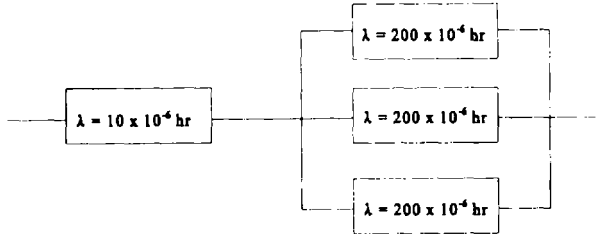
a) Draw a fault tree for this system.
b) Find the minimal cut sets of the system.
c) Find the minimal path sets of the system.



4.12 Consider the following pumping system consisting of three identical parallel pumps, and a valve in series. The pumps have a constant failure rate of λ_p (hr^{-1}) and the valve has a constant failure rate of λ_v (hr^{-1}) (accidental closure).

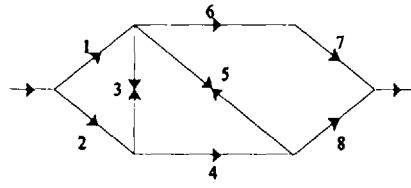
a) Develop an expression for the reliability of the system using the success tree. (Assume non-repairable components.)
b) Find the average reliability of the system over time period T.
c) Repeat questions (a) and (b) for the case that $\lambda t < 0.1$, and then approximate the reliability functions. Find the average reliability of the system when $\lambda = 0.001$ hr^{-1} and $T = 10$ hours.

4.13 In the following system, which uses active redundancy, what is the probability that there will be no failures in the first year of operation? Assume constant failure rates given below.



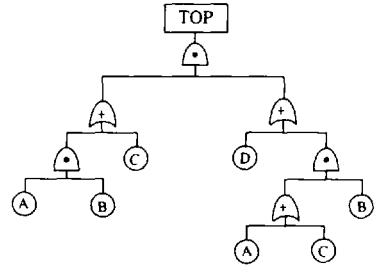
4.14 A filter system is composed of 30 elements, each with a failure rate of 2×10^{-4} (hr^{-1}). The system will operate satisfactorily with two elements failed. What is the probability that the system will operate satisfactorily for 1000 hours?

4.15 Consider the reliability diagram below.

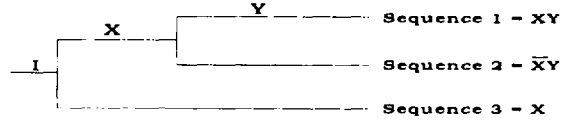


a) Find all minimal path sets.
b) Find all minimal cut sets.
c) Assuming each component has a reliability of 0.90 for a given mission time, compute the system reliability over mission time.

4.16 In the following fault tree, find all minimal cut sets and path sets. Assuming all component failure probabilities are 0.01, find the top event probability.



4.17 An event tree is used in reactor accident estimation as shown:



where sequence 1 is a success and sequences 2 and 3 are failures. The cut sets of system X and Y are

$$X = A \cdot B + A \cdot C + D$$

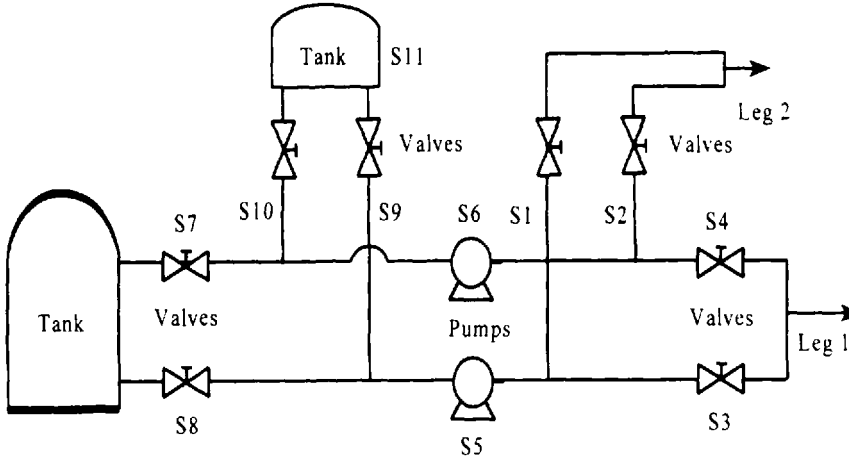
and

$$Y = B \cdot D + E + A.$$

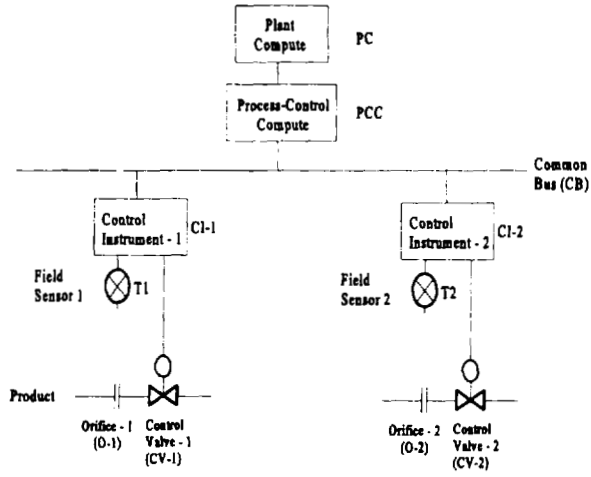
Find cut sets of sequence 2 and sequence 3.

4.18 In a cement production factory, a system such as the one shown below is used to provide cooling water to the o.zsutside of the furnace. Develop an MLD for this system.

System bounds:	$S_{12}, S_{11}, S_9, S_8, S_7, S_6, S_5, S_4, S_3, S_2, S_1$
Top event:	Cooling from legs 1 and 2
Not-allowed events:	Passive failures and external failures
Assumptions:	Only one of the pumps or legs is sufficient to provide the necessary cooling. Only one of the tanks is sufficient as a source.



4.19 Develop an MLD model of the following system. Assume the following:

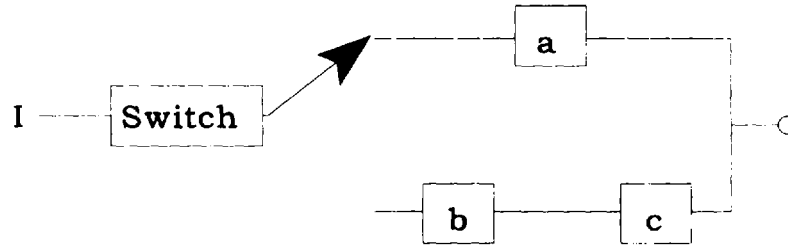


One of the two product lines is sufficient for success.
Control instruments feed the sensor values to the process-control computer, which calculates the position of the control valves.
The plant computer controls the process-control computer.

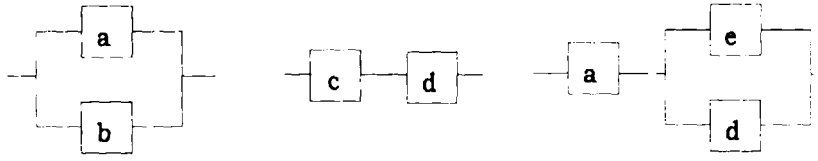
a) Develop a fault tree for the top event "inadequate product feed".
b) Find all the cut sets of the top event.
c) Find the probability of the top event.
d) Determine which components are critical to the design.

4.20 Perform a FMECA analysis for the system described in Exercise 4.19. Compare the results with part (d) of Exercise 4.19.

4.21 A standby system is shown below. Assume that components a, b, and c are identical with a constant failure rate of 1×10^{-3} (hr^{-1}) and a constant standby failure rate of 1×10^{-5} (hr^{-1}). The probability that the switch fails to operate if component "a" fails is 5×10^{-2}. Calculate the reliability of this system at $t = 200$ hours. (Note that either "a" or "b and c" is required for system operation.)



4.22 A super-computer requires subsystem "*A*" and either subsystem "*C*" or subsystem "B" to function so as to save some critical data following a sudden loss of power to the computer. Subsystems *A*, *B*, and *C* are configured as shown below:

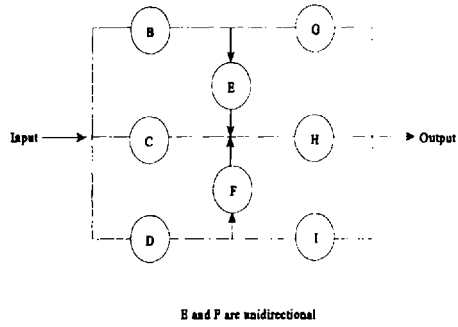


Subsystem A Subsystem B Subsystem C

Use the following component reliability values to determine the probability that critical data will not be saved following a loss of power: $R_a = 0.99$, $R_b = 0.98$, $R_c = 0.999$, $R_d = 0.998$, $R_e = 0.99$.

4.23 A system of two components arranged in parallel redundancy has been observed to fail on average, every 1000 hours. Data for the individual components which come from the field experience indicate a failure rate of about 0.01 per hour. Is there an inconsistency in component-level and system-level failure rates?

4.24 Consider the system shown below.



Develop a fault tree for this system with the top event of "No Output from the System." Calculate the reliability of this system for a 100 hour mission. Assume MTTF of 600 hours for all components.

REFERENCES

Crowder, M. J., Kimber A. C., Smith, R. L., and Sweeting, T. J., "*Statistical Analysis of Reliability Data,*" Chapman & Hall, London, New York, 1991.

Daniels, H. E., "*The Statistical Theory of the Strength of Bundles of Threads*". I. Proc. R.Soc., London, A183, 404–435, 1945.

Dezfuli, H., et al., "*Application of REVEAL_W to Risk-based Configuration Control,*" Reliability Engineering and System Safety J., 44(3), 1994.

Fong, C. C. and J. A. Buzacoot, "*An Algorithm for Symbolic Reliability Combination with Path-Sets or Cut-Sets,*" IEEE Trans. Reliability, 36 (1), 34–37, 1987.

Kapur, K., and Lamberson, L., "*Reliability in Engineering Design,*" Wiley, New York, 1977.

MIL-STD-1629A, *Procedure for Performing a Failure Mode, Effects, and Criticality Analysis*, Department of Defense, NTIS, Springfield, Virginia, 1980.

MIL-STD-338, "*Electronic Reliability Design Handbook,*" NTIS, Springfield, Virginia, 1980.

Modarres, M., "*Application of the Master Plant Logic Diagram in Risk-Based Evaluations,*" Amer. Nucl. Society Topical Mtg. on Risk Management, Boston, MA, 1992.

SAE Reference Manual J1739, "*Potential Failure Mode and Effect Analysis in Design and Manufacturing,*" 1994.

REVEAL_W User's Manual. Version 1.0, Scientech, Inc., Maryland, USA, 1994.

Shooman, M. L., "*Probabilistic Reliability: An Engineering Approach,*" 2nd ed., Kreiger, Melbourne, FL, 1990.

Vesely, W. E., Goldberg F., Roberts N. and Haasl D., "*Fault Tree Handbook*," NUREG-0492, U.S. Nuclear Regulatory Commission, Washington, D.C., 1981.

Specter, H., Modarres, M., "*Functional Specifications for a PRA Based Design Making Tool*," Empire State Electric Energy Research Corporation, EP 95-14, New York, NY, 1996.

Wang, J. and Modarres, M., "*REX : An Intelligent Decision and Analysis Aid for Reliability and Risk Studies*," Rel. Eng. & Syst. Safety J., 30, 185–239, 1990.

5

Reliability and Availability of Repairable Items

When we perform reliability studies, it is important to distinguish between repairable and nonrepairable items. The reliability analysis methods discussed in Chapters 3 and 4 are largely applicable to nonrepairable items. In this chapter, we examine repairable systems and discuss methods used to determine the failure characteristics of these systems, as well as the methods for predicting their reliability and availability.

Nonrepairable items are those that are discarded and replaced with new ones when they fail. For example, light bulbs, transistors, contacts, unmanned satellites, and small appliances are nonrepairable items. Reliability of a nonrepairable item is expressed in terms of its time-to-failure distribution, which can be represented by respective cdf, pdf, or hazard (failure) rate function, as was discussed in Chapter 3.

Repairable items, generally speaking, are not replaced following the occurrence of a failure; rather, they are repaired and put into operation again. On the other hand, if a nonrepairable item is a component of a repairable system, estimation of the distribution of a number of the component replacements over a given time interval is a problem considered in the framework of repairable systems reliability. In contrast to nonrepairable items, reliability problems associated with nonrepairable items are, basically, considered using different *random* (*stochastic*) *processes'* models, some of them are discussed below.

In this chapter, we are also interested in the notion of *availability*. Items can be repaired, and repair activities take time. The probability that an item (system) is up (functioning) can be measured by a probability value called availability, which shows the probability that the system is up. Conversely, the probability that the system is down is called *unavailability*.

We will start with probabilistic models and statistical methods that are used

in determining the failure characteristics of repairable items and their reliability. We will then define the concept of availability and explain availability evaluation methods for repairable items. Although the presentation of the material in this chapter focuses on system reliability and availability, the methods are equally applicable to components.

5.1 REPAIRABLE SYSTEM RELIABILITY

5.1.1 Basic Random Processes Used as Probabilistic Models of Repairable Systems

For the situations when the down time associated with preventive maintenance, repair or replacement actions is negligible, compared with the mean time between failures (MTBF), the, so-called, *point processes* are used as probabilistic models for respective failure processes. The point process can be informally defined as a model for randomly distributed events, having negligible duration. The following point processes are mainly used as probabilistic failure process models (Leemis (1995)): homogeneous Poisson process (HPP), renewal process (RP) and non-homogeneous Poisson process (NHPP). These processes will be discussed later in this section. For those situations when the respective down time is not negligible, compared with MTBF, the, so-called, alternating renewal process (ARP) is used.

Usually a point process is related to a single item (e.g., a system). A *sample path* (*trajectory*) or *realization* of a point process is the successive failure times of an item: $T_1, T_2, \ldots, T_k, \ldots$ (see Figure 5.1). We can also use the point process model for studying a group of identical items, if the number of items in the group is constant. We must also remember that the sampling scheme considered is "with instantaneous replacement."

A realization of a point process is expressed in terms of the *counting function*, $N(t)$, which is introduced as the number of failures, which occur during interval $(0,t)$ (Leemis (1995)), i.e., for $t > 0$

$$N(t) = \max \left(k \mid T_k \leq t \right) \tag{5.1}$$

It is clear that $N(t)$ is a random function. Denote the mean of $N(t)$ by $\Lambda(t)$, i.e., $E[N(t)] = \Lambda(t)$. A realization, $N(t)$, and the respective $\Lambda(t)$ are shown in Figure 5.1. $\Lambda(t)$ and its derivative, $\Lambda(t)' = \lambda(t)$, known as the *rate of occurrence of failures (ROCOF)* or the *intensity function*, are the basic characteristics of a point process. Sometimes the notation $v(t)$ is used instead of $\lambda(t)$. (Please note that notation $\lambda(t)$ can be misleading, it should not be confused with the hazard (failure) rate function for which the same notation is often used.) At this point, it is important to make clear the difference between the failure (hazard) rate function, $h(t)$, and ROCOF, $\lambda(t)$. As it was discussed in Chapter 3, the hazard (failure) rate

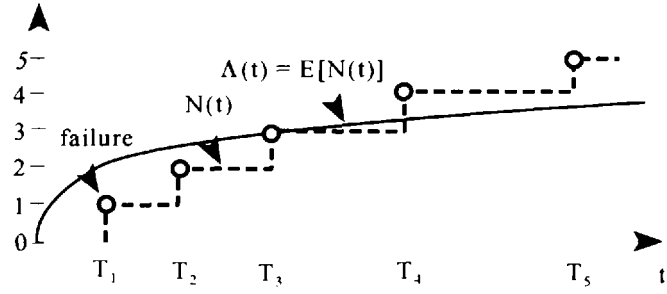


Figure 5.1 Geometric interpretation of f(t), N(t), and Λ(t) for a repairable system.

function, $h(t)$, is a characteristic of time-to-failure (or time-between-failures) distribution, while ROCOF, $\lambda(t)$, is a characteristic of point process.

To move forward, we should recall the sampling procedures associated with $f(t)$ and $h(t)$:

N items are tested to failure without replacement (so the number of items in a test is time dependent); or

an item is tested to failure with instantaneous replacement by the new item from the same population.

$$\int_{t_1}^{t_2} f(t)\,dt$$

From the standpoint of probabilistic interpretations, pdf, $f(t)$, is the unconditional pdf, so, the integralis the unconditional probability of failure in the interval (t_1, t_2). Meanwhile, the failure rate function, $h(t)$, is the conditional pdf, and integral

$$\int_{t_1}^{t_2} h(t)\,dt$$

is the conditional probability of failure in the interval (t_1, t_2).

Under the sampling procedure for a point process, one (or N) item(s) is (are) tested with instantaneous replacement by an item (not necessarily from the same population). The number of items under the test is always constant. The respective probabilistic interpretations of ROCOF, $\lambda(t)$, is given by the following equation

$$\int_{t_1}^{t_2} \lambda(t)\,dt = E\left[N(t_2, t_1)\right]$$

where $E[N(t_2,t_1)]$ is the mean number of failures which occur in the interval (t_1,t_2).

Now, we can summarize the time-dependent reliability behavior of repairable and nonrepairable items in terms of the hazard rate function and ROCOF (Leemis (1995)). The term *burn-in* is used for a nonrepairable item when its failure (hazard) rate function is decreasing in time, and the term *wear-out* is used when the failure (hazard) rate function is increasing. The life of nonrepairable item is described by the time-to-failure distribution of a single nonnegative random variable.

For repairable items the term *improvement* is used when its ROCOF is decreasing and the term *deterioration* is used when its ROCOF is increasing. The life of repairable items, generally speaking, cannot be described by a distribution of a single nonnegative random variable; in this case such characteristics as time between successive failures are used (the first and the second, the second and the third, and so on).

Now we discuss the basic point processes which are used in the modeling of repairable systems. Below, we briefly consider their main probabilistic properties and basic estimation procedures.

Homogeneous Poisson Process

Homogeneous Poisson process (HPP), with ROCOF λ, is defined as a point process having the following properties:

$N(0) = 0$,
the process has independent increments (i.e., the numbers of failures observed in nonoverlapping intervals are independent),
the number of failures, observed in any interval of length, t, has the Poisson distribution with mean λt.

The last property of the HPP is not only important for straightforward reliability applications, but also can be used for the hypothesis testing that a random process considered is the HPP. It is obvious that the HPP is stationary, i.e., λ is constant. Consider some other useful properties of the HPP.

Superposition of the HPPs

As it was mentioned earlier, HPP, RP, and NHPP are used for modeling the failure behavior of a *single* item. In many situations it is important to model the failure pattern of several identical items *simultaneously* (the items must be put in service or on a test at the same moment). The superposition of several point processes is the ordered sequence of all failures that occur in any of the individual point processes.

The superposition of several HPP processes with parameters $\lambda_1, \lambda_2, \ldots, \lambda_i$

is the HPP with $\lambda = \lambda_1 + \lambda_2 + \cdots + \lambda_k$. The well-known example is a series system, introduced in Chapter 4, with exponentially distributed elements.

Distribution of Intervals Between Failures

As it was shown in Section 3.2.1, under the HPP model, the distribution of intervals between successive failures is modeled by the exponential distribution with failure rate λ. The HPP is the only process for which the failure rate of time-between-failures distribution coincides with its ROCOF.

Let now T_{n0} be the time from an origin (test start) to the n_0 th failure, where n_0 is a fixed (nonrandom) integer. In this notation the time to the first failure is T_1. It is clear that T_{n0} is the sum of n_0 independent r.v.'s each exponentially distributed. As it was discussed in Section 3.4.2, the random variable, $2\lambda T_{n0}$, has the Chi-squared distribution with $2n_0$ degrees of freedom:

$$2\lambda T_{n0} = \chi^2_{2n0} \tag{5.2}$$

Later in this section, we will also be dealing with $\ln(T_{n0})$. Using relationship (5.2) one can write

$$\ln(T_{n0}) = -\ln(2\lambda) + \ln(\chi^2_{2n0})$$

This expression shows that one has to deal with log Chi-squared distribution, for which the following results of Bartlett and Kendall are available (Cox and Lewis (1978)). For the large samples the following (asymptotic) normal approximation for the log Chi-squared distribution can be used:

$$E(\ln T_{n_0}) \approx \ln\left(\frac{n_0}{\lambda}\right) - \frac{1}{\left(2n_0 - \frac{1}{3} + \frac{1}{16n_0}\right)}$$
$$\operatorname{var}(\ln T_{n_0}) \approx \frac{1}{n_0 - \frac{1}{2} + \frac{1}{10n_0}} \tag{5.3}$$

This approximation is used in the following as a basis for a trend analysis procedure (see Section 5.1.4 and Example 5.5).

Renewal Process

The renewal process (RP) retains all the properties related to the HPP, except for the last property. In the case of RP the number of failures observed in any interval of length t, generally speaking, does not have to follow the Poisson distribution. Therefore, the time-between-failures distribution of RP can be any continuous distribution. Thus, RP can be considered as a generalization of HPP for

the case when the time-between-failures is assumed to have any distribution (Leemis (1995)).

The RP based model is appropriate for the situations where an item is renewed to its original state (as a new one) upon failure. This model is not applicable in the case of a repairable system consisting of several components, if only a failed component is replaced upon failure.

The following classification of the RPs is based on the coefficient of variation, σ/μ, (standard deviation to mean ratio) of the time-between-failures distribution. A RP is called *underdispersed* (or conversely *overdispersed*) if the coefficient of variation of the time-between-failures distribution is less than (greater than) 1. It can be shown that if time-between-failures distribution is IFR (DFR), its coefficient of variation is less than (greater than) 1 (Barlow and Proschan (1981)), and so the corresponding RP is underdispersed (overdispersed). Recall that for the exponential distribution $\sigma/\mu = 1$. In opposite to the *overdispersed* RP and the HPP, for which any preventive action policy, formally, does not have any sense, different optimal preventive action schedules can be considered for the underdispersed renewal processes.

Many of the reliability applications of the HPP and the RPs are reduced to solving the following problems:

find the distribution of $T_n = t_1 + t_2 + \cdots + t_n$, the time to nth failure.
find the distribution of the number of failures by time t.

The simplest particular case of RP is the HPP (the exponential time-between-failure distribution). In general, all the problems are not easy to solve, nevertheless, for the distribution of T_n, the first two moments (the mean and variance of T_n) can be easily found as

$$E(T_n) = nE(t)$$

and

$$\mathrm{var}(T_n) = n\,\mathrm{var}(t)$$

Renewal Equation

Let $\Lambda(t) = E[N(t)]$, where $N(t)$ is given by (5.1). Function $\Lambda(t)$ is sometimes called the *renewal function*. It can be shown (see Hoyland and Rausand (1994)) that $\Lambda(t)$ satisfies the, so-called, *renewal equation*:

$$\Lambda(t) = F(t) + \int_0^t F(t-s)\,d\Lambda(s) \tag{5.4}$$

where $F(t)$ is the cdf of time-between-failures (t_is). By taking the derivative of

both sides of (5.4) with respect to t, one gets the following integral equation for ROCOF, $\lambda(t)$,

$$\lambda(t) = f(t) + \int_0^t f(t-s)\lambda(s)\,ds$$

where $f(t)$ is the pdf of F(t). The integral equation obtained can be solved using a Laplace transformation. The solutions for the exponential and gamma distributions can be obtained in closed form. For the Weibull distribution only the recursion procedures are available (Hoyland and Rausand (1994)). The possible numerical solutions for other distributions and different types of renewals can be obtained using Monte Carlo simulation. For more information see Kaminskiy and Krivtsov (1997).

The statistical estimation of cdf or pdf of time-between-failures distribution on the basis of ROCOF or $\Lambda(t)$ observations is difficult.

For the HPP

$$\frac{\Lambda(t)}{t} = \lambda$$

In general, the *elementary renewal theorem* states the following asymptotic property of the renewal function:

$$\lim_{t \to \infty} \frac{\Lambda(t)}{t} = \frac{1}{\text{MTTF}}$$

Some confidence limits for $\Lambda(t)$ are given in Hoyland and Rausand (1994). Contrary to the HPPs, the superposition of RPs, in general, is not a RP.

Example 5.1

Time-between-failure of a repairable unit is supposed to follow the Weibull distribution with scale parameter $\alpha = 100$ hours and shape parameter $\beta = 1.5$. Assuming that repairs are perfect, i.e., the unit is renewed to its original state upon a failure, assess the mean number of repairs during mission time $t = 1000$ hours.

Solution:

Use the elementary renewal theorem. The Weibull mean is given by (see Table 3.1)

$$\text{MTTF} = \alpha \cdot \Gamma\left(\frac{\beta + 1}{\beta}\right)$$

so, for the given values of α and β, MTTF = 90.27 hours. Thus, the mean number of repairs during mission time $t = 1000$ hours can be estimated as

$$\Lambda(1000) = \frac{1000}{90.27} = 11.08$$

Nonhomogeneous Poisson Process (NHPP)

The definition of the *Nonhomogeneous Poisson Process* (NHPP) retains all the properties related to the HPP, except for the last one. In the case of NHPP, λ is not constant, and the probability that exactly n failures occur in any interval, (t_1, t_2), has the Poisson distribution with the mean

$$\int_{t_1}^{t_2} \lambda(t)\, dt$$

Therefore,

$$\Pr\left[N(t_2) - N(t_1) = n\right] = \frac{\left(\int_{t_1}^{t_2} \lambda(t)\, dt\right)^n \exp\left(-\int_{t_1}^{t_2} \lambda(t)\, dt\right)}{n!} \tag{5.5}$$

for $n = 0, 1, 2, \ldots$ The function

$$\Lambda(t) = \int_0^t \lambda(\tau)\, d\tau$$

analogous to the renewal function is often called the *cumulative intensity function* (Leemis (1995)), while the ROCOF $\lambda(t)$ is called the *intensity function*.

Unlike the HPP or the RP, the NHPP is capable of modeling improving and deteriorating systems. If the intensity function (ROCOF) is decreasing, the system is improving, and if the intensity function is increasing, the system is deteriorating. If the intensity function is not changing with time, the process reduces to the HPP.

It should be noted that the NHPP retains the independent increment property, but the times between failures are neither exponentially distributed nor identically distributed.

The reliability function for the NHPP can be introduced for a given time interval (t_1, t_2) as the probability of survival over this interval, i.e.,

$$R(t_1, t_2) = \Pr\left[N(t_2) - N(t_1) = 0\right] = \frac{\left(\int_{t_1}^{t_2} \lambda(t)\, dt\right)^0 \exp\left(-\int_{t_1}^{t_2} \lambda(t)\, dt\right)}{0!} \tag{5.6}$$

$$= \exp\left(-\int_{t_1}^{t_2} \lambda(t)\, dt\right)$$

It is obvious that in the case of the HPP (where λ = const.) this function is reduced to the conditional reliability function (3.5) for the exponential distribution.

5.1.2 Statistical Data Analysis for Repairable Systems

From the discussion in the previous section, it is obvious that the HPP cases are the simplest cases for repairable equipment data analysis. For example, in such situations the procedures for exponential distribution estimation discussed in Chapter 3 (classical and Bayes') can be applied. The main underlying assumption for these procedures, when applied to repairable systems, is that rate of occurrence of failures (ROCOF), λ, is constant and will remain constant over all time intervals of interest. Therefore, the data should be tested for potential increasing or decreasing trends.

The use of the estimators for HPP are justified only after it has been proven that the ROCOF is reasonably constant, i.e., there is no evidence of an increasing or decreasing trend. An increasing trend is not necessarily due to random aging processes. Poor use of equipment, including poor testing, maintenance, and repair work, and out-of-spec (overstressed) operations, can lead to premature aging and be major contributions to increasing trends.

Figure 5.2 depicts three cases of occurrences of failure in a repairable system.

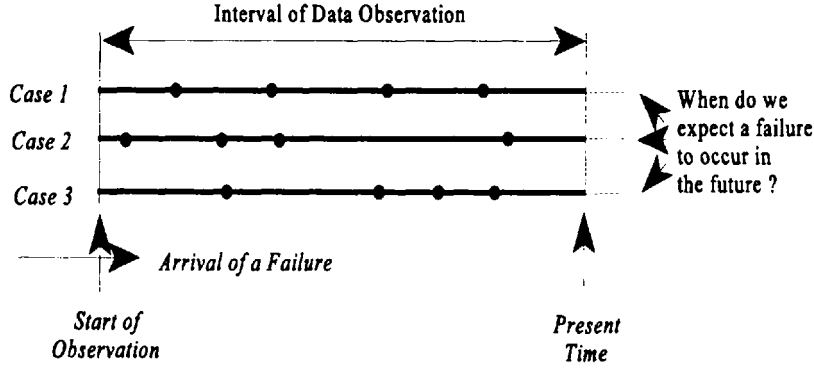


Figure 5.2 Three cases of failure occurrence.

The constant ROCOF estimators give the same point and confidence estimates for each of the three situations shown in Figure 5.2, since the number of failures and length of experience are the same for each. Clearly, Case 2 shows a decreasing failure rate, while Case 3 shows an increasing failure rate. We would therefore

expect that, given a fixed time interval in the future, the system, shown as Case 3, would be more likely to fail than the other two systems.

This shows the importance of considering trends in occurrence of failures when predicting system reliability.

According to Ascher (1984) and O'Connor (1991), the following points should be considered in failure rate trend analyses:

1. Failure of a component may be partial, and repair work done on a failed component may be imperfect. Therefore, the time periods between successive failures are not necessarily independent. This is a major source of trend in the failure rate.
2. Imperfect repairs performed following failures do not renew the system, i.e., the component will not be as good as new following maintenance or repair. The constant failure rate assumption holds only if the component is assumed to be as good as new; only then can the statistical inference methods using a constant ROCOF assumption be used.
3. Repairs made by adjusting, lubricating, or otherwise treating component parts that are wearing out provide only a small additional capability for further operation, and do not renew the component or system. These types of repair may result in a trend of an increasing ROCOF.
4. A component may fail more frequently due to aging and wearing out. In the remainder of this section, we provide a summary of a typical trend-analysis process, and discuss the subsequent calculation of unavailability estimates.

Several procedures may be used to check the HPP model assumptions. For example, the goodness-of-fit criteria discussed in Chapter 2 can be applied to testing the exponential distribution of times-between failures, or the Poisson distribution of the number of failures observed in equal length time intervals. Another useful procedure, discussed in the Chapter 3 is the total-time-on-test.

5.1.3 Data Analysis for the HPP

Procedures Based on the Poisson Distribution

Suppose that a failure process is observed for a predetermined time t_0 during which n failures have been recorded at times $t_1 < t_2, \ldots < t_n$, where, obviously, $t_n < t_0$. The process is assumed to follow a HPP.

The corresponding likelihood function can be written as

$$L = \lambda^n e^{\lambda t_0}$$

It is clear that, with t_0 fixed, the number of events, n, is a *sufficient statistic* (note that one does not need to know $t_1, t_2, \ldots, t_n$ to construct our likelihood function). Thus, the statistical inference can be based on the Poisson distribution of the number of events. As a point estimate of λ one usually takes n/t_0, which is the *unique unbiased estimate* based on the sufficient statistic.

A typical problem associated with repairable systems, in which the failure behavior follows the HPP, is to test for the null hypothesis $\lambda = \lambda_0$, (or the mean number of events, $\mu = \mu_0 = \lambda_0 t_0$) against the alternative $\lambda > \lambda_0$ ($\mu > \mu_0$). The alternative hypothesis has the exact level of significance, P_+, corresponding to the observed number of failures n, given by (Cox and Lewis (1968)):

$$P_+(n, \mu_0) = \Pr(N \geq n \mid \mu = \mu_0) = \sum_{r=n}^{\infty} \frac{\mu_0^r e^{-\mu_0}}{r!} \tag{5.7}$$

For the alternatives $\lambda < \lambda_0$ ($\mu < \mu_0$), the exact level of significance corresponding to an observed value n is given by

$$P_-(n, \mu_0) = \Pr(N \leq n \mid \mu = \mu_0) \tag{5.8}$$

If the two-sided alternatives are considered, the level of significance is defined to be

$$P(n, \mu_0) = 2\min\left[P_+(n, \mu_0), P_-(n, \mu_0)\right] \tag{5.9}$$

If the normal approximation to the Poisson distribution is used (see Section 2.3.2), the corresponding statistic, having the standard normal distribution, is

$$\frac{|n - \mu_0| - 0.5}{\sqrt{\mu_0}} \tag{5.10}$$

where 0.5 is a correction term.

Example 5.2

Twelve failures of a new repairable unit were observed during a three year period. From the past experience it is known that for similar units, the rate of occurrence of failures, λ_0, is 3.33 year^{-1}. Check the hypothesis that the rate of occurrence of failures of the new unit λ is equal to λ_0.

Solution:

Choose 5% significance level. Using Table A1, find the respective acceptance region for statistic (5.10) as interval (- 1.96, 1.96). Keeping in mind that $\mu_0 = \lambda_0 t = 3.33 \times 3 = 10$, calculate statistic (5.10):

$$\frac{|12 - 10| - 0.5}{\sqrt{10}} \approx 0.47$$

which is inside the acceptance region. Thus, the hypothesis that the rate of occurrence of failures of the new unit is equal to the rate of similar units, λ_0, is not rejected.

Another typical problem associated with repairable systems, which failure behavior can be modeled by the HPP is the comparison of two HPPs. Such problems can appear, for example, when two identical units are operated in different plants or by different personnel, and one is interested in the corresponding ROCOF comparison.

Assume that our data are the observations on two independent HPPs and the goal is to compare the corresponding rates of occurrence, λ_1 and λ_2.

Let the data collected be the numbers of failures n_1 and n_2, observed in nonrandom time intervals T_1 and T_2 correspondingly. The random numbers of events n_1 and n_2, can be considered as observed values of independent random variables with Poisson distributions having the means $\mu_1 = \lambda_1 T_1$ and $\mu_2 = \lambda_2 T_2$, so that, we can write

$$\Pr(N_1 = n_1, N_2 = n_2) = \frac{\exp(-\mu_1)\mu_1^{n_1}}{n_1!} \frac{\exp(-\mu_2)\mu_2^{n_2}}{n_2!} \qquad (5.11)$$

To compare the ROCOFs for the processes considered, one may use the following statistic (Cox and Lewis (1968))

$$\rho = \frac{\mu_2}{\mu_1} = \frac{T_2 \lambda_2}{T_1 \lambda_1}$$

Since the nonrandom time intervals T_1 and T_2 are known, inference about ρ is identical to the inference about the ratio λ_2 / λ_1. The inference about ρ can be done, based on the conditional distribution of N_2 (or N_1) given $N_2 + N_1 = n_2 + n_1$. This probability can be written as

$$\Pr\left(N_2 = n_2 \mid N_1 + N_2 = n_1 + n_2\right) = \frac{\Pr\left(N_1 = n_1,\ N_2 = n_2\right)}{\Pr\left(N_1 + N_2 = n_1 + n_2\right)}$$

$$= \frac{\dfrac{\mu_1^{n_1}\mu_2^{n_2}}{n_1!\,n_2!}\exp\left[-\left(\mu_1 + \mu_2\right)\right]}{\dfrac{\left(\mu_1 + \mu_2\right)^{n_1 + n_2}}{\left(n_1 + n_2\right)!}\exp\left[-\left(\mu_1 + \mu_2\right)\right]} \qquad (5.12)$$

$$= \left(\frac{n_1 + n_2}{n_1}\right)\theta^{n_2}\left(1 - \theta\right)^{n_1}$$

where $\theta = \rho/(1 + \rho)$.

In the case where $\lambda_1 = \lambda_2$, the probability (5.12) is binomial with parameter $T_1/(T_1 + T_2)$, and this parameter is 0.5 in an important particular case of equal length time intervals. Thus, exact procedures for the binomial distribution or its normal approximations can be used for making inference about ρ.

Example 5.3

In nuclear power plants, *Accident Sequence Precursors* are defined as "those operational events which constitute important elements of accidence sequences leading to severe core damage" (see Section 8.6). In Table 8.11, the annual cumulative numbers of precursors for the U.S. plants are given for the period of 1984–1993. The occurrence of precursors is assumed to follow an HPP. There were 32 events observed in 1984 and 39 in 1993. Test the hypothesis that the rate of occurrence of events (per year) is the same for the years given.

Solution:

For the data given $n_1 = 32$ and $n_2 = 39$. Because $T_1 = T_2 = 1$ year, our null hypothesis is $H_0 : \rho_0 = 1$, so that $\theta_0 = 0.5$. Using the normal approximation (similar to (5.10)), calculate the following statistic

$$\frac{\left|\, n_2 - n\theta_0 \,\right| - 0.5}{\sqrt{n\theta_0\left(1 - \theta_0\right)}}$$

where $n = n_1 + n_2$. Thus, one gets

$$\frac{\left|\, 39 - 71/2 \,\right| - 0.5}{\sqrt{71 \times 0.5 \times 0.5}} \approx 0.71$$

which is inside an acceptance region for any reasonable significance level, α. In other words the data do not show any significant change in the rate of precursor occurrence (H_0 is not rejected).

Procedures Based on the Exponential Distribution of Time Intervals

In Section 3.2.1 it was shown that under the HPP model, the intervals between successive failures have the exponential distribution. Therefore, data analysis procedures for the exponential distribution considered in Chapter 3 (classical as well as Bayes') can be used. Some special techniques applicable for the HPP are considered in the next section, where the data analysis for the HPP is treated as a particular case of data analysis for the NHPP.

Assume again that failure data are the observations from two independent HPPs and our goal is to compare the corresponding rates of occurrence (ROCOF), λ_1 and λ_2.

Let t_1 and t_2 be the times at which predetermined (nonrandom) numbers, n_1 and n_2 , of failures occur for the corresponding processes. It is clear that t_1 and t_2 can be considered as realizations (observed values) of independent random variables, T_1 and T_2, for which the quantity $2\lambda T$ has the Chi-squared distribution with $2n$ degrees of freedom (see Section 3.4.2) . We can introduce statistic

$$R = \left(\frac{2\lambda_2 T_2}{2n_2} \right) \Big/ \left(\frac{2\lambda_1 T_1}{2n_1} \right) \tag{5.13}$$

which follows the F distribution with $(2n_2, 2n_1)$ degrees of freedom (Cox and Lewis (1968)). Based on this statistic, the confidence intervals for the ratio λ_2/λ_1 can be written as:

$$\Pr\left(F_{1-\alpha/2} < R < F_{\alpha/2} \right) = 1 - \alpha,$$

$$\Pr\left(F_{1-\alpha/2} \frac{t_1 n_2}{t_2 n_1} < \frac{\lambda_2}{\lambda_1} < F_{\alpha/2} \frac{t_1 n_2}{t_2 n_1} \right) = 1 - \alpha$$

where F_α is the upper α quantile of the F distribution with $(2n_2, 2n_1)$ degrees of freedom. Substituting the observed values, t_1 and t_2, one gets the confidence interval corresponding to the confidence probability $1 - \alpha$ as

$$F_{1-\alpha} \frac{t_1 n_2}{t_2 n_1} < \frac{\lambda_2}{\lambda_1} < F_\alpha \frac{t_1 n_2}{t_2 n_1} \tag{5.14}$$

The corresponding null-hypothesis that $\lambda_2/\lambda_1 = r_0$ can be tested using the two tailed test for the statistic

$$\frac{r_0}{\left[\left(\frac{n_2}{t_2}\right) \Big/ \left(\frac{n_1}{t_1}\right)\right]} \tag{5.15}$$

having under H_0 the F distribution with $(2n_2, 2n_1)$ degrees of freedom (see Table A.5).

Example 5.4

The failure data on two identical items used at two different sites were collected. At the first site, observations continued till the eighth failure, which was observed at 1880 hours. At the second site observations continued till the twelfth failure, which was observed at 1654 hours. Assuming that the time-between-failure distributions of both items are exponential, check if the items are identical from a reliability standpoint, i.e., test the null hypothesis, H_0: $\lambda_1 = \lambda_2$

Solution:

Calculate statistic (5.15) for $r_0 = 1$

$$\frac{1}{\left[\left(\frac{12}{1654}\right) \Big/ \left(\frac{8}{1880}\right)\right]} = 0.586$$

Using 10% confidence level and Table A5, find the acceptance region as (0.48, 2.24). So, our null hypothesis is not rejected.

5.1.4 Data Analysis for NHPP

As it was mentioned above, the NHPP can be used to model improving and deteriorating systems: if the intensity function (ROCOF) is decreasing, the system is improving, and if the intensity function is increasing, the system is deteriorating. The problem of ROCOF trend analysis is of great importance simply because any preventive actions do not have any sense for the HPP due to the memoryless property of the respective exponential time-between-failure distribution.

Formally, we can test for trend, taking the null hypothesis of no trend, i.e., that the events form the HPP and applying a goodness-of-fit test for the

exponential distribution of the intervals between successive failures the Poisson distribution of the number of failures in the time intervals of constant (nonrandom) length. A simple graphical procedure based on this property is to plot the cumulative number of failures versus the cumulative time. Deviation from linearity indicates the presence of a trend.

These tests are not sensitive enough against the NHPP alternatives, so it is better to apply the following methods (Cox and Lewis (1968)) .

Regression Analysis of Time Intervals

Suppose one has a reasonably long series of failures and the problem is to examine any gradual trend in the rate of failure occurrence.

Choose an integer, l, which is recommended to be no less than 4, but such that no appreciable change in ROCOF arises during the interval of occurrence of l failures.

Let t_1 be the observed time from the start to the lth failure, t_2 be the time from the lth failure to the $2l$th failure, and so on. Finally, we have got a series of intervals $t_1, t_2, \ldots, t_r$. If the process considered is the HPP, using Equations (5.3) one can write:

$$\begin{aligned} E(\ln t_i) &= -\ln \lambda_i + c_l \\ \operatorname{var}(\ln t_i) &= v_l \end{aligned} \tag{5.16}$$

where c_l and v_l are known constants independent of λ, for example,

$$v_l = \frac{1}{l - 0.5}$$

and t_i ($i = 1, 2, \ldots$) are independently distributed. Assume that the observations are generated by a process satisfying all the conditions for a HPP, except that the ROCOF λ is *slowly* varying with time. Consider the approximation that

λ is a constant, λ_i, within the period covered by t_i, and that an independent variable z_i can be attached to each t_i such that in the case of simplest model,

$$\ln \lambda_i = \alpha + \beta z_i \tag{5.17}$$

For example, z_i might be

- the midpoint of the interval t_i, if λ is being considered as a function of time, t
- the value of any constant or, averaged over the interval t_i, independent variable, which could responsible for ROCOF variation.

Under the above assumptions, we obtain the following linear regression model:

$$E(\ln t_i) = -(\alpha' + \beta z_i)$$
$$\text{var}(\ln t_i) = v_i$$

where $\alpha' = \alpha - c_i$ and β are unknown parameters and v_i is a *known* constant. Using the standard regression procedures (as discussed in Section 2.8), one can

obtain the standard least-squares estimates of parameters α' and β
test approximately the null hypothesis $\beta = 0$ and obtain approximate confidence limits for β,
compare the residual variance with the respective theoretical value, v_i, to check the adequacy of the model.

One can include in the model considered above additional independent variables. For example, we can generalize model (5.17) to a loglinear polynomial model

$$\log \lambda_i = \alpha + \beta z_i + \gamma z_i^2 + \cdots$$

Another regression approach, performed in terms of counts of failures observed in successive equal time intervals, is considered in (Cox and Lewis (1968)). The regression procedures considered can also be performed in the framework of Bayesian approach to regression, given, for example, in (Judge, et al. (1988)). The Maximum Likelihood estimation for model (5.17) is considered by Lawless (1982), who also applied this model to failure data on a set of similar air-conditioning units.

Example 5.5

Consider the following data in the form of successive times between failures of a repairable item. Let t_1 be the observed time from the start to the 4th failure, t_2 be the time from the 4th failure to the 8th failure, and so on, and let z_i be the time at the center of the interval t_i. Using the data below, fit the simple linear regression model (5.17) and determine whether or not there is any trend in ROCOF.

Interval number, i	$\ln t_i$	z_i (in relative units)
1	0.151	0.581
2	0.157	1.748
3	0.275	2.991
4	-0.445	3.970
5	-0.983	4.478
6	-0.703	4.913

Solution:

Rewrite Equation (5.18) in the form:

$$E(\ln t_i) = -(\alpha' + \beta z_i) = \gamma_0 + \gamma z$$

$$\mathrm{var}(\ln t_i) = v_4$$

where $\alpha' = \alpha - c_4$, c_4, and v_4 are given by (5.3), i.e.,

$$c_4 = \ln 4 - \frac{1}{2 \cdot 4 - \frac{1}{3} + \frac{1}{16 \cdot 4}} = 1.256$$

$$v_4 = \frac{1}{4 - 0.5 + \frac{1}{10 \cdot 4}} = 0.284$$

Meanwhile, α and β are unknown parameters to be estimated. Using the standard least-squares estimates (2.101) for γ_0 and γ based on the data, obtain:

$$\hat{\gamma}_0 = 0.540, \qquad \hat{\gamma} = -0.256$$

Therefore,

$$\alpha = \alpha' + c_4 = -0.540 + 1.256 = 0.716$$

$$\beta = 0.256$$

Finally,

$$\lambda(t) = 100.761 + 0.256t$$

To check the adequacy of the ROCOF model obtained, we need to check the hypothesis that the theoretical variance $v_4 = 0.284$ (having infinite number of degrees of freedom) is not less than the residual variance which can be calculated using (2.102). The value of the residual variance is 0.114, and it has $6 - 2 = 4$ degrees of freedom. Using the significance level of 5% and the respective critical value from Table A.5, conclude that our hypothesis is not rejected, so the model obtained is adequate.

Maximum Likelihood Procedures

Under the NHPP model the intervals between successive events are independently distributed and the probability that, starting from time t_i, the next failure occurs in $(t_{i+1}, t_{i+1} + \Delta t)$ can be approximated by (Cox and Lewis (1968)):

$$\lambda(t_{i-1})\Delta t \exp\left(- \int_{t_i}^{t_{i-1}} \lambda(x)\, dx \right)$$

where the first multiplier is the probability of failure in $(t_{i+1}, t_{i+1} + \Delta t)$, and the second one is the probability of a failure-free operation in the interval (t_i, t_{i+1}).

If the data are the successive failure times, $t_1, t_2, \ldots, t_n$, $(t_1 < t_2 < \cdots < t_n)$ observed in the interval $(0, t_0)$, $t_0 > t_n$ (the data are type I censored), the likelihood function for any $\lambda(t)$ dependence, can be written as

$$\prod_{i=1}^{n} \lambda(t_i)\, e^{-\int_0^{t_1} \lambda(x)dx} e^{-\int_{t_1}^{t_2} \lambda(x)dx} \cdots e^{-\int_{t_{n-1}}^{t_n} \lambda(x)dx} e^{-\int_{t_n}^{t_0} \lambda(x)dx} = \prod_{i=1}^{n} \lambda(t_i)\, e^{-\int_0^{t_0} \lambda(x)dx} \tag{5.19}$$

The corresponding log-likelihood function is given by

$$l = \sum_{i=1} n \ln \lambda(t_i) - \int_0^{t_0} \lambda(x)\, dx \tag{5.20}$$

To avoid complicated notation, consider the case when ROCOF takes the simple form similar to (5.17), i.e.,

$$\lambda(t) = e^{\alpha + \beta t} \tag{5.21}$$

Note that the model above is, in some sense, more general than the linear one, $\lambda(t) = \alpha + \beta t$, which can be considered as a particular case of (5.21), when $\beta t << 1$.

Plugging (5.21) in (5.19) and (5.20) one gets

$$L_1(\alpha, \beta) = \exp\left[n\alpha + \beta \sum_{i=1}^{n} t_i - \frac{e^{\alpha}\left(e^{\beta t_0} - 1\right)}{\beta} \right] \tag{5.22}$$

$$l = \ln\left[L_1(\alpha, \beta)\right] = n\alpha + \beta \sum_{i=1}^{n} t_i - \frac{e^{\alpha}\left(e^{\beta t_0} - 1\right)}{\beta} \tag{5.23}$$

The conditional likelihood function can be found by dividing (5.22) by the marginal probability of observing n failures, which is given by the respective term of the Poisson distribution with mean

$$\int_0^{t_0} \lambda(x)\, dx = \frac{e^{\alpha}\left(e^{\beta t_0} - 1\right)}{\beta}$$

The conditional likelihood function is given by (Cox and Lewis (1968))

$$L_c = \frac{e^{n\alpha + \beta \sum_1^n t_i} \exp\left(-\frac{e^{\alpha} e^{\beta t_0} - 1}{\beta} \right)}{\frac{\left(e^{\alpha} e^{\beta t_0} - 1 \right)^n}{\beta^n n!} \exp\left(-\frac{e^{\alpha} e^{\beta t_0} - 1}{\beta} \right)} = \frac{n! \beta^n}{\left(e^{\beta t_0} - 1 \right)^n} e^{\left(\beta \sum_1^n t_i \right)} \tag{5.24}$$

Because $0 < t_1 < t_2 < \cdots < t_n < t_0$, the conditional likelihood function (5.24) is the pdf of an ordered sample of size n from the truncated exponential distribution having the pdf

$$f(t) = \frac{\beta}{e^{\beta t_0} - 1} e^{\beta t}, \qquad 0 \le t \le t_0, \qquad \beta \neq 0. \tag{5.25}$$

Thus, for any β the conditional pdf of $\Sigma\, t_i$ is the same as for the sum of n independent random variables having the pdf (5.25). It is easy to see that for $\beta = 0$, the pdf (5.25) becomes the uniform distribution over $(0, t_0)$.

Example 5.6

In a repairable system, the following eight failures have been observed at: 595, 905, 1100, 1250, 1405, 1595, 1850, and 1995 hours. Assume the observation ends at the time when the last failure is observed, and that the time to repair is negligible. Test whether these data exhibit a trend in a form of (5.21).

Solution:

Taking the derivative of (5.25) with respect to β and the derivative of (5.25) with respect of α, and equating them to zero, results in the following system of equations for maximum likelihood estimates of these parameters

$$\sum_{i=1}^{n} t_i + \frac{n}{\beta} - \frac{n t_n}{1 - e^{-\beta t_n}} = 0$$

$$e^{\alpha} = \frac{n \beta}{e^{-\beta t_n} - 1}$$

For the data given $n_t = 8$, $t_n = 1995$ hours, and $\Sigma t_i = 10{,}695$ hours. Solving these equations numerically, one gets the following trend model

$$\lambda(t) = e^{-6.8134 + 0.0011t}$$

Laplace's Test

Now we are going to use conditional pdf (5.25) to test the null hypothesis, $H_0 : \beta = 0$, against the alternative hypothesis $H_1: \beta \neq 0$. This test is known as the Laplace test (sometimes it is also called the Centroid test). As mentioned above, under the condition of $\beta = 0$, pdf (5.25) is reduced to the uniform distribution over $(0, t_0)$ and $S = \Sigma t_i$ has the distribution of the sum of n independent uniformly distributed random variables. Thus, one can use the distribution of the following statistic

$$U = \frac{S - \frac{n t_0}{2}}{\sqrt{t_0\left(\frac{n}{12}\right)}} = \frac{\frac{\sum_{i=1}^{n} t_i}{n} - \frac{t_0}{2}}{t_0\sqrt{\frac{1}{12n}}} \qquad (5.26)$$

which has approximately the standard normal distribution (Cox and Lewis (1978)).

If the alternative hypothesis is $H_1: \beta \neq 0$, then the large values of $|U|$ indicate an evidence against the null hypothesis. If the alternative hypothesis is $H_1: \beta > (<)$ 0, then the large values of U $(-U)$ provide evidence against the null hypothesis. In other words, if U is close to 0, there is no evidence of trend in the data, and the process is assumed to be stationary (i.e., an HPP). If $U < 0$, the trend is decreasing, i.e., the intervals between successive failures (interarrival values) are becoming larger. If $U > 0$, the trend is increasing. For the latter two situations, the process is not stationary (i.e., it is an NHPP).

If the data are failure terminated (type II censored) statistic (5.26) is replaced by

$$U = \frac{\frac{\sum_{i=1}^{n-1} t_i}{n-1} - \frac{t_n}{2}}{t_n\sqrt{\frac{1}{12(n-1)}}} \qquad (5.27)$$

Example 5.7

Consider the failure arrival data for a motor-operated rotovalve in a process system. This valve is normally in standby mode, and is demanded when

overheating occurs in the process. The only major failure mode is "failure to start upon demand." The arrival dates of this failure mode (in calendar time) are shown in the table below. Determine whether an increasing failure rate is justified. Assume that a total of 5256 demands occurred between January 1, 1970 and August 12, 1986, and that demands occur at a constant rate. The last failure occurred on August 12, 1986.

Date	Failure order number	Date	Failure order number
04-20-1970	1	05-04-1981	16
09-19-1970	2	05-05-1981	17
10-09-1975	3	08-31-1981	18
12-16-1974	4	09-04-1981	19
12-21-1975	5	12-02-1982	20
07-24-1977	6	03-23-1983	21
01-22-1978	7	12-16-1983	22
01-29-1978	8	03-28-1984	23
06-15-1978	9	06-06-1984	24
01-01-1979	10	07-19-1984	25
05-12-1979	11	06-23-1985	26
07-23-1979	12	07-01-1985	27
11-17-1979	13	01-08-1986	28
07-24-1980	14	04-18-1986	29
11-23-1980	15	08-12-1986	30

Solution:

Let's distribute the total number of demands (5256) over the period of observation. Let's also calculate the interarrival time of failures (in months), the interarrival of demands (number of demands between two successive failures), and the arrival demand. These values are shown in Table 5.1.

Since the observation ends at the last failure, the following results are obtained using (5.27):

$$\Sigma t_i = 95.898$$

$$n - 1 = 29$$

$$\frac{\sum t_i}{n-1} = 3307$$

$$\frac{t_n}{2} = 2628$$

$$U = \frac{3307 - 2628}{5256\sqrt{\frac{1}{12 \times 29}}} = 2.41$$

Table 5.1 Arrival and Interarrival for the Rotovalve

Date	Interarrival time (months)	Interarrival demand (days)	Arrival demand (days)
04-20-1970	4	104	104
09-19-1970	5	131	235
10-09-1975	62	1597	1832
12-16-1975	2	59	1891
12-21-1975	0	4	1895
07-24-1977	19	503	2398
01-22-1978	6	157	2555
01-29-1978	0	6	2561
06-15-1978	5	118	2679
01-01-1979	7	173	2852
05-12-1979	4	113	2966
07-23-1979	2	62	3028
11-17-1979	4	101	3129
07-24-1980	8	216	3345
11-23-1980	4	106	3451
05-04-1981	5	140	3591
05-05-1981	0	1	3592
08-31-1981	4	102	3694
09-04-1981	0	3	3697
12-02-1982	15	393	4090
03-23-1983	4	96	4186
12-16-1983	9	232	4418
03-28-1984	3	89	4507
06-06-1984	2	61	4568
07-19-1984	1	37	4605
06-23-1985	11	293	4898
07-01-1985	0	7	4905
01-08-1986	6	165	5070
04-18-1986	3	86	5157
08-11-1986	4	99	5256

To test the null hypothesis that there is no trend in the data, and the ROCOF, λ, of rotovalves is constant, we would use Table A.1 with $U = 2.41$. Therefore, we can reject the null hypothesis at the 5% significance level (the respective acceptance region is $(-1.96 + 1.96)$).

The existence of a trend in the data in Example 5.7 indicates that the interarrivals of rotovalve failures are not independently and identically distributed (IID) random variables, and thus the stationary process for evaluating reliability

of rotovalves is incorrect. Rather, these interarrival times can be described in terms of the NHPP.

Another form of $\lambda(t)$ considered by Bassin (1969, 1973) and Crow (1974) is

$$\lambda(t) = \lambda\beta t^{\beta-1} \tag{5.28}$$

Expression (5.28) has the same form as the failure (hazard) rate of nonrepairable items (3.18) for the Weibull distribution. Using (5.6), the reliability function of a repairable system having ROCOF (5.28) for an interval $(t, t + t_1)$ can be obtained as follows

$$R(t, t + t_1) = e^{-\lambda(t+t_1)^{\beta} - \lambda t^{\beta}} \tag{5.29}$$

Crow (1974) has shown that under the condition of a single system observed to its nth failure, the maximum likelihood estimator of β and λ can be obtained as:

$$\hat{\beta} = \frac{n}{\sum_{i=1}^{n-1} \ln \frac{t_n}{t_i}} \tag{5.30}$$

$$\hat{\lambda} = \frac{n}{t_n^{\hat{\beta}}} \tag{5.31}$$

The $1 - \alpha$ confidence limits for inferences on β and λ have been developed and discussed by Bain (1978).

Example 5.8

Using the information in Example 5.7, calculate the maximum likelihood estimator of β and λ. Also, plot the demand failure rate as a function of time from 1971 to 1999.

Solution:

Using (5.30) and (5.31), we can calculate $\hat{\beta}$ and $\hat{\lambda}$ as 1.59 and 3.71×10^{-5} respectively. Using $\hat{\beta}$ and $\hat{\lambda}$, the functional form of the demand failure rate can be obtained by using (5.28) as

$$\lambda(d) = 3.71 \times 10^{-5} \times 1.59 d^{0.59}$$

where d represents the demand number (time in days).

The plot of the demand failure rate (ROCOF of NHPP) as a function of calendar time for the rotovalve is shown in Figure 5.3. For comparison purposes,

the constant demand failure rate function (HPP case) is also shown. For the HPP, the point estimate of λ was obtained by dividing the number of failures by the number of demands. The upper and lower confidence intervals were obtained using the HPP assumption.

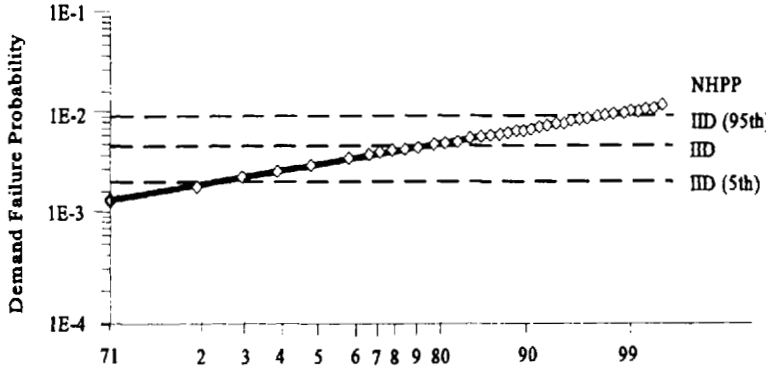


Figure 5.3 Comparison of NHPP and HPP models for rotovalve example.

Example 5.9

In a repairable system, the following six interarrival times between failures have been observed: 16, 32, 49, 60, 78, and 182 (in hours). Assume the observation ends at the time when the last failure is observed.

a. Test whether these data exhibit a trend. If so, estimate the trend model parameters as given in (5.28).
b. Find the probability that the interarrival time for the seventh failure will be greater than 200 hours?

Solution:

Use the Laplace's test to test the null hypothesis that there is no trend in the data at 10% significance level (the respective acceptance region is (-1.645 +1.645)). From (5.27) find

$$U = \frac{\frac{16 + (16 + 32) + \cdots}{5} - \frac{417}{2}}{417\sqrt{\frac{1}{12(5)}}} = -1.82$$

Notice that $t_n = 417$. The value of U obtained indicates that the NHPP can be applicable (H_0 is rejected) and the sign of U shows that the trend is decreasing.

Using (5.30) and (5.31), we can find

$$\hat{\beta} = \frac{6}{\ln \frac{417}{16} + \ln \frac{417}{16+32} + \cdots} = 0.712$$

$$\hat{\lambda} = \frac{6}{(417)^{0.712}} = 0.0817 \text{hr}^{-1}$$

Thus, $\hat{\lambda}(t) = 0.058\, t^{0.288}$. From (5.29) with t_1=200,

$$\Pr(\text{7th failure occurs within 200 hours}) = 1 - \exp[-\lambda(|(t_0 + t_1)^{\beta} - \lambda(t_0)^{\beta}|)] = 0.85.$$

The probability that the interarrival time is greater than 200 hours is 1 - 0.85 = 0.15.

Crow (1990) has expanded estimates (5.30) and (5.31) to include situations where data originate from multi-unit repairable systems.

 See the software supplement for the automated Laplace test and the NHPP estimation procedures.

5.2 AVAILABILITY OF REPAIRABLE SYSTEMS

We defined reliability as the probability that a component or system will perform its required function over a given time. The notion of availability is related to repairable (or maintained) items only. We define availability as the probability that a repairable system (or component) will function at time t, as it is supposed to, when called upon to do so. Respectively, the unavailability of a repairable item, $q(t)$ is defined as the probability that the item is in a failed state (down) at time t. There are several definitions of availability, the most common ones are as follows.

1. *Instantaneous (point) availability* of a repairable item at time t, $a(t)$, is the probability that the system (or component) is up at time t.

2. *Limiting availability*, a, is defined as the following limit of instantaneous availability, $a(t)$

$$a = \lim_{t \to \infty} a(t) \tag{5.32}$$

3. *Average availability*, $\bar{a}$ is defined for a fixed time interval, T, as

$$\bar{a} = \frac{1}{T} \int_0^T a(t)\, dt \tag{5.33}$$

4. The respective *limiting average availability* is defined as

$$\bar{a}_l = \lim_{T \to \infty} \frac{1}{T} \int_0^T a(t) dt \tag{5.34}$$

It should be noted that the limiting average availability has limited applications. We elaborate on each of the three definitions of availability in the remainder of this section.

If a component or system is nonrepairable, its availability coincides with its reliability function, $R(t)$, i.e.,

$$a(t) = R(t) = \exp\left[- \int_0^t \lambda(\tau)\, d\tau \right] \tag{5.35}$$

where $\lambda(t)$ is the failure (hazard) rate function. The unavailability, $q(t)$, is obviously, related to $a(t)$ as

$$q(t) = 1 - a(t) \tag{5.36}$$

From the modeling point of view, repairable systems can be divided into the following two groups:

1. Repairable systems for which failure is immediately detected (revealed faults).
2. Repairable systems for which failure is detected upon inspection (sometimes referred to as periodically inspected (tested) systems).

5.2.1 Instantaneous (Point) Availability

For the first group systems, it can be shown (see Section 5.3) that $a(t)$ and $q(t)$ are obtained from the following ordinary differential equations:

$$\begin{aligned} \frac{da(t)}{dt} &= -\lambda(t)\, a(t) + \mu(t)\, q(t) \\ \frac{dq(t)}{dt} &= \lambda(t)\, a(t) - \mu(t)\, q(t) \end{aligned} \tag{5.37}$$

where $\lambda(t)$ is the failure rate and $\mu(t)$ is the repair rate.

The most widely used models for availability are based on the exponential time-between-failure and repair time distribution. Based on (5.37) it can be shown (see Section 5.3) that in this case (no trend exists in the rate of occurrence of failure and repair), the point availability and unavailability of the system (or component) are given by

$$a(t) = \frac{\mu}{\lambda + \mu} + \frac{\lambda}{\lambda + \mu} \exp[-(\lambda + \mu)t]$$
$$q(t) = \frac{\lambda}{\lambda + \mu} - \frac{\lambda}{\lambda + \mu} \exp[-(\lambda + \mu)t] \tag{5.38}$$

Note that in (5.38), $\mu = 1/\tau$, where τ is the average time interval per repair (sometimes referred to as mean time-to-repair (MTTR)). Clearly, MTBF = $1/\lambda$ in this case.

For the second type of repairable systems mentioned above, the determination of availability is a difficult problem. Caldorela (1977) presents a form of $a(t)$ for cases where no trend in the failure rate exists, and the inspection interval (η), duration of inspection (θ), and duration of repair (τ) are fixed. In these cases,

$$a(t) = e^{-(t - m\eta)\lambda_{eff}} \cdot \left[1 - e^{-\left(\frac{t - m\eta}{\theta}\right)^q}\right]$$

where

$$\lambda_{eff} = \frac{\eta - \theta}{\eta}\left(\frac{n - \theta}{\eta} + \frac{2\tau}{\eta}\right) + 2\left[1 - \frac{\Gamma(1/q)}{q}\right]\frac{\theta}{\eta^2}$$
$$q = \ln[3 - \ln(\theta\lambda)] \tag{5.39}$$

and m is the inspection interval number ($1, 2, \ldots, n$). When $t > m\eta + \theta$, it is easy to show that

$$a(t) \approx \exp[-(t - m\eta)\lambda_{eff}].$$

Example 5.10

Find the unavailability, as a function of time, for a system that is inspected once a month. Duration of inspection is 1.5 hours. Any required repair takes an average of 19 hours. Assume the failure rate of the system is 3×10^{-6} hr^{-1}.

Solution:

Using (5.39), for $\theta = 1.5$, $\tau = 19$, $\eta = 720$, $\lambda = 3 \times 10^{-6}$, we can get the plot of $q(t)$ as shown in Figure 5.4.

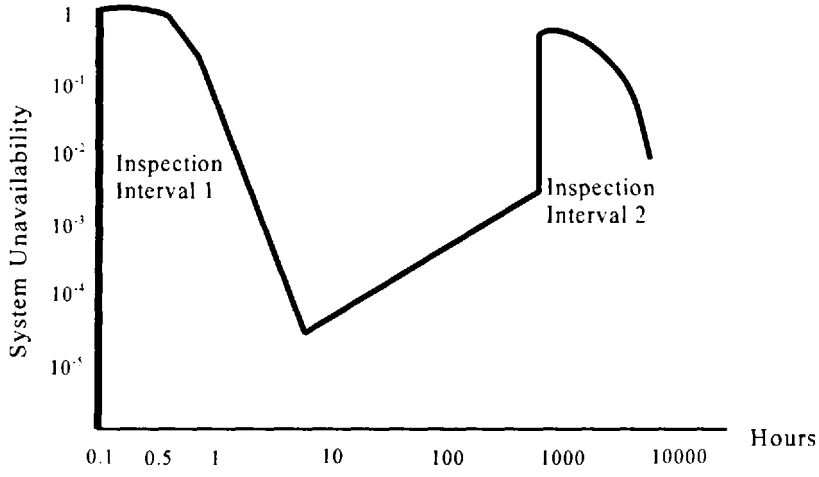


Figure 5.4 Unavailability of the system as a function of time.

For simplicity, the pointwise availability function can be represented in an approximate form. This simplifies availability calculations significantly. For example, for a periodically tested component, if the repair and test durations are very short compared with the operation time, and the test and repair are assumed perfect, one can neglect their contributions to unavailability of the system. This can be shown using Taylor expansion of the unavailability equation (see Lofgren (1985)). In this case for each test interval T, the availability and unavailability functions are

$$\begin{aligned} a(t) &\approx 1 - \lambda t \\ q(t) &\approx \lambda t \end{aligned} \tag{5.40}$$

The plot of the unavailability as a function of time, using (5.40), will take a shape similar to that in Figure 5.5. Clearly if the test and repair durations are long, one must include their effect.

Vesely and Goldberg (1981) have used the approximate pointwise unavailability functions for this case. The functions and their plot are shown in Figure 5.6. The average values of the approximate unavailability functions shown in Figures 5.5 and 5.6 are discussed in Section 5.2.3 and are presented in Table 5.2.

It should be noted that, due to random imperfection in test and repair activities, it is possible that a residual unavailability q would remain following a

test and/or repair. Thus, unlike the unavailability function shown in Figure 5.5, the unavailability function in Figure 5.6 exhibits a residual unavailability q_0 due to these random imperfections.

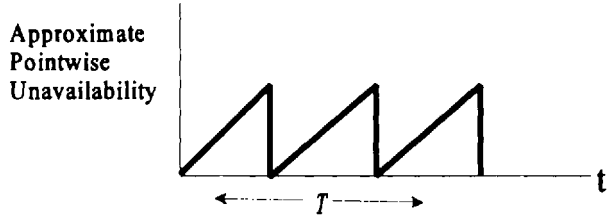
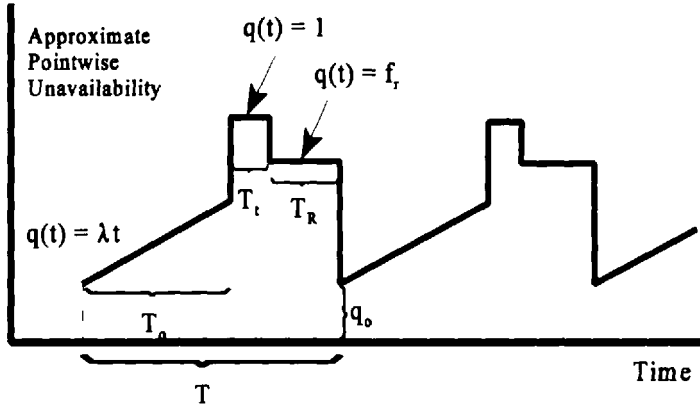


Figure 5.5 Approximate pointwise unavailability for a periodically tested item.

5.2.2 Limiting Point Availability

It is easy to see that some of the pointwise availability equations discussed in Section 5.2.1 have limiting values. For example, (5.38) has the following limiting value:



T = Test interval, T_R = Average repair time (hr),
T_t = Average test duration (hr), f_r = Frequency of repair,
q_0 = Residual unavailability.

Figure 5.6 Pointwise unavailability for a periodically tested item including test and repair outages.

$$a = \lim_{t \to \infty} a(t)$$

$$= \frac{\mu}{\lambda + \mu}$$

or its equivalent

$$a = \frac{\text{MTBF}}{\text{MTBF} + \text{MTTR}} \tag{5.41}$$

Equation (5.41) is sometimes referred to as the asymptotic availability of a repairable system with constant rate of occurrence of failure and repair.

5.2.3 Average Availability

According to its definition, average availability is a constant measure of availability over a period of time T. For noninspected items, T can take on any value (preferably, it should be about the mission length). For inspected items, T is normally the inspection (or test) interval or mission length T_m. Thus, for nonrepairable items, if the inspection interval is T, then the approximate expression for point availability with constant λ can be used. If we assume $\bar{a} \approx 1 - \lambda t$ (which might be applicable, if at least $\lambda t < 0.1$), then

$$a = \frac{1}{T} \int_0^T (1 - \lambda t)\, dt = 1 - \frac{\lambda T}{2} \tag{5.42}$$

Accordingly, for all types of systems, one can get such approximations for average availabilities. Vesely et al. (1981) have discussed the average unavailability for various types of systems. Table 5.2 shows these functions.

Table 5.2 Average Availability Functions

Type of item	Average unavailability	Average availability
Nonrepairable	$\frac{1}{2}\lambda T_m$	$1 - \frac{1}{2}\lambda T_m$
Repairable revealed fault	$\frac{\lambda\tau}{1 + \lambda\tau}$	$\frac{1}{1 + \lambda\tau}$
Repairable periodically tested	$\frac{1}{2}\lambda T_0 + f_r \frac{T_R}{T} + \frac{T_t}{T}$	$1 - \frac{1}{2}\lambda T_0 + f_r \frac{T_R}{T} + \frac{T_t}{T}$

λ = constant failure rate $(\text{hr})^{-1}$, T_m = mission length (hr), τ = average downtime or MTTR (hr), T = test interval (hr), T_R = average repair time (hr), T_t = average test duration (hr), f_r = frequency of repair per test intervals, T_0 = operating time (up time) = $T - T_R - T_t$.

Equations in Table 5.2 can also be applied to standby equipment, with λ representing the standby (or demand) failure rate, and the mission length or operating time being replaced by the time between two tests.

5.3 USE OF MARKOVIAN METHODS FOR DETERMINING SYSTEM AVAILABILITY

Markovian methods are useful tools for evaluating the availability of a system that has multiple states (e.g., up, down, and degraded). For example, consider a system with the states shown in Figure 5.7. In the framework of Markovian models, the transitions between various states are characterized by constant *transition rates* (these rates, generally speaking, may not necessarily be constant in practice).

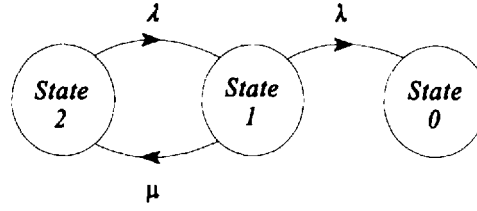


Figure 5.7 A Markovian model for a system with three discrete states.

Consider a system with a given number of discrete states, n. Introduce the following characteristics of the system:

$$\Pr_i(t) = \Pr(\text{the system is in state } i \text{ at time } t), \quad \sum_{i=1}^{n} \Pr_i(t) = 1$$

ρ_{ij} = transition rate from state i to state j, $(i, j = 1, 2, \ldots, n)$.

Because ρ_{ij} is constant, the random time the system is at state i until the transition to state j follows the exponential distribution with rate ρ_{ij}. Assuming that $\Pr_i(t)$ is differentiable, it is possible to show (Hoyland and Rausand (1994)) that

$$\frac{d\Pr_i(t)}{dt} = -\Pr_i(t)\left(\sum_{j(j\neq i)} \rho_{ij}\right) + \left(\sum_{j(j\neq i)} \rho_{ji} \Pr j(t)\right). \tag{5.43}$$

If a differential equation similar to (5.43) is written for each state, and the resulting set of differential equations is solved, one may obtain the time-dependent probability of each state. This can be seen better in the following example.

Example 5.11

Consider a system with constant failure rate λ and constant repair rate μ in a standby redundant configuration. When the system fails, its repair starts immediately, which puts it back into operation. The system has two states: state 0—when the system is down, and state 1—when the system is operating (Fig. 5.8).

a. Find the probabilities of these states.
b. Determine the availability of this system.

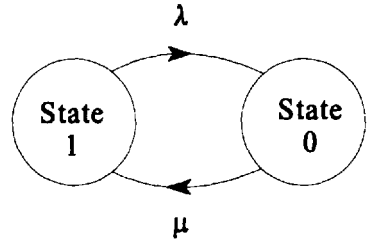


Figure 5.8 Markovian model for Example 5.11.

Solution:

Assuming that the system is functioning at time $t = 0$, i.e., $\Pr_1(0) = 1$ and $\Pr_0(0) = 0$, and using the governing differential equation (5.43) find

$$\begin{aligned} \frac{d\Pr_0(t)}{dt} &= +\lambda \Pr_1(t) - \mu \Pr_0(t) \\ \frac{d\Pr_1(t)}{dt} &= -\lambda \Pr_1(t) + \mu \Pr_0(t) \end{aligned} \tag{5.44}$$

For the above set of equations, matrix $A = \begin{pmatrix} \lambda & -\mu \\ -\lambda & \mu \end{pmatrix}$ is referred to as the *transition matrix.*

The above equations can be solved, for example, using the Laplace transformation. Below, we take the Laplace transform of both sides of the equations:

$$s\bar{P}_1(s) - 1 = \lambda \bar{P}_1(s) - \mu \bar{P}_0(s),$$

$$s\bar{P}_0(s) = -\lambda \bar{P}_1(s) + \mu \bar{P}_0(s).$$

The solution of the above system is given by

$$\bar{P}_0(s) = \frac{\lambda}{s(s+\lambda+\mu)}$$

$$\bar{P}_1(s) = \frac{s+\mu}{s(s+\lambda+\mu)}$$

Finding the respective inverse Laplace transform, it follows that availability $a(t)$ is obtained from

$$a(t) = \Pr_1(t) = L^{-1}\left\{\frac{s+\mu}{s(s+\lambda+\mu)}\right\} = \frac{\mu}{\lambda+\mu} + \frac{\lambda}{\lambda+\mu}\exp[-(\lambda+\mu)t]$$

which coincides with Equation (5.38) discussed in Section 5.2.
Accordingly, unavailability is

$$q(t) = \Pr_1(t) = 1 - a(t) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu}\exp[-(\lambda+\mu)t]$$

Example 5.12

A system that consists of two cooling units has the three states shown in the Markovian model in Figure 5.9. When one unit system fails, the other system takes over and repair on the first starts immediately. When both systems are down, there are two repair crews to simultaneously repair the two systems. The three states are as follows:

State 0, when both systems are down,
State 1, when one of the systems is operating and the other is down, and
State 2, when the first system is operating and the second is in standby (in an operating ready condition).

a. Determine the probability of each state.
b. Determine the availability of the entire system.

Solution:

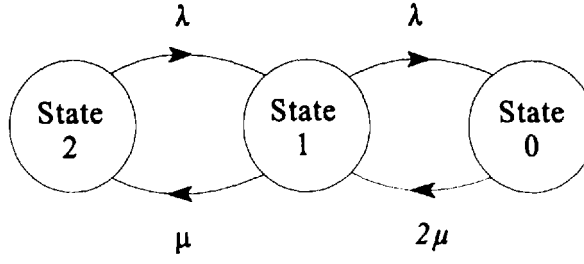


Figure 5.9 Markovian model in Example 5.12.

a. The governing differential equations are

$$\frac{d\Pr_2(t)}{dt} = -\lambda \Pr_2(t) + \mu \Pr_1(t),$$

$$\frac{d\Pr_1(t)}{dt} = +\lambda \Pr_2(t) - (\mu + \lambda) \Pr_1(t) + 2\mu \Pr_0(t),$$

$$\frac{d\Pr_0(t)}{dt} = +\lambda \Pr_1(t) + 2\mu \Pr_0(t).$$

Taking the Laplace transform of both sides of the equations yields the following:

$$sP_2(s) - \Pr_2(0) = -\lambda P_2(s) + \mu P_1(s),$$

$$sP_1(s) - \Pr_1(0) = +\lambda P_2(s) - (\mu + \lambda) P_1(s) + 2\mu P_0(s),$$

$$sP_0(s) - \Pr_0(0) = +\lambda P_1(s) - 2\mu P_0(s).$$

$\Pr_2(0) = 1$ and $\Pr_1(0) = \Pr_0(0) = 0$. Solving the above set of equations, $\Pr_i(s)$ can be calculated as

$$P_2(s) = \frac{1}{\lambda + s} + \frac{\mu\lambda(2\mu + s)}{s(s + \lambda)(s - k_1)(s - k_2)},$$

$$P_1(s) = \frac{\lambda(2\mu + s)}{s(s - k_1)(s - k_2)},$$

$$P_0(s) = \frac{\lambda^2}{s(s - k_1)(s - k_2)}.$$

where

$$k_1 = \frac{-2\lambda - 3\mu - \sqrt{4\lambda\mu + \mu^2}}{2},$$

$$k_2 = \frac{2\mu\lambda + \lambda^2 + 2\mu^2}{k_1}.$$

If the inverses of the above Laplace transforms are taken, the probability of each state can be determined as follows:

$$\Pr_2(t) = e^{-\lambda t} + G_1 e^{-\lambda t} + G_2 e^{k_1 t} + G_3 e^{k_2 t} + G_4$$

where

$$G_1 = \frac{\mu(\lambda + 2\mu)}{(\lambda + k_1)(\lambda + k_2)}, \qquad G_2 = \frac{\mu\lambda(2\mu + k_1)}{(k_1 + \lambda)(k_1 - k_2)(k_1)}$$

$$G_3 = \frac{\mu\lambda(2\mu + k_1)}{k_2(k_2 - k_1)(k_2 + \lambda)}, \qquad G_4 = \frac{2\mu^2}{2\mu\lambda + \lambda^2 + 2\mu^2}$$

And,

$$\Pr_1(t) = A_1 e^{k_1 t} + A_2 e^{k_2 t} + A_3$$

where

$$A_1 = \frac{2\mu\lambda}{(k_1 - k_2)k_1} + \frac{\lambda}{k_1 - k_2},$$

$$A_2 = \frac{\lambda(2\mu + k_2)}{(k_1 - k_2)(k_2)}, \qquad A_3 = \frac{2\mu\lambda}{2\mu\lambda + \lambda^2 + 2\mu^2}$$

And,

$$\Pr_0(t) = B_1 \exp[k_1 t] + B_2 \exp[k_2 t] + B_3,$$

where

$$B_1 = \frac{\lambda^2}{(k_1 - k_2)k_1} \qquad B_2 = \frac{\lambda^2}{(k_2 - k_1)(k_1)}$$

and

$$B_3 = \frac{\lambda^2}{2\mu\lambda + \lambda^2 + 2\mu^2}$$

b. The availability of the two units system, is $a(t) = \Pr_2(t) + \Pr_1(t)$, and the unavailability of the entire system is $q(t) = \Pr_0(t)$.

It is possible to simply find the limiting pointwise availability from the governing equations of the system. For this purpose, consider the Markovian transition diagram shown in Figure 5.10.

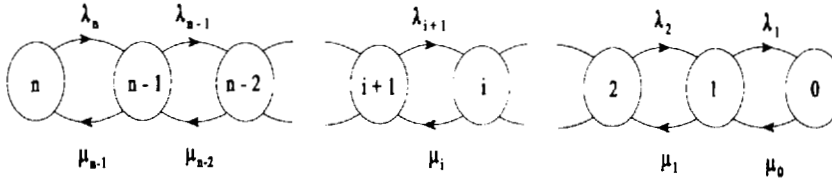


Figure 5.10 A Markovian transition diagram with n states.

It may be shown that

$$\Pr_{i-1}(\infty)\,\lambda_{i-1} = \mu_i \Pr_i(\infty)$$

or,

$$\Pr_i(\infty) = \frac{\lambda_{i-1} \times \lambda_{i-2} \cdots \lambda_n}{\mu_i \times \mu_{i+1} \cdots \mu_{n-1}} \Pr_n(\infty) \tag{5.45}$$

Since $\sum_{i=0}^{n} \Pr_i(\infty) = 1$, solving (5.45) for $\Pr_n(\infty)$ yields

$$\Pr_i(\infty) = \frac{1}{1 + \sum_{i=0}^{n-1} \dfrac{\lambda_{i-1} \times \lambda_{i-2} \cdots \lambda_n}{\mu_i \times \mu_{i-1} \cdots \mu_{n-1}}} \tag{5.46}$$

Accordingly, the system's limiting pointwise unavailability (and similarly its availability) can be obtained.

$$q = \Pr_i(\infty) = \frac{\sum_{i=1}^{n} P_i \lambda_i}{\sum_{i=1}^{n-1} P_i \mu_i} \Pr_n(\infty) \tag{5.47}$$

If the system is unavailable when it is in any of the states $(0, 1, \ldots, r-1)$, then

$$q = \sum_{i=0}^{r-1} \Pr_i(\infty) = \Pr_n(\infty) \sum_{i=1}^{r-1} \frac{\lambda_{i-1} \times \lambda_{i-2} \cdots \lambda_n}{\mu_i \times \mu_{i-1} \cdots \mu_{n-1}} \tag{5.48}$$

Example 5.13

For Example 5.12, determine the limiting pointwise unavailability from (5.47) and confirm it with the results obtained in that example.

Solution:

Since $\lambda_2 = \lambda_1 = \lambda$, $\mu_1 = \mu$, $\mu_0 = 2\mu$ from (5.45),

$$Pr_1(\infty) = \frac{2\mu}{\lambda} Pr_0(\infty)$$

and

$$Pr_2(\infty) = \frac{\mu}{\lambda} Pr_1(\infty) = \frac{2\mu^2}{\lambda^2} Pr_0(\infty)$$

Since

$$Pr_0(\infty) + Pr_1(\infty) + Pr_2(\infty) = 1$$

from (5.49),

$$q = Pr_0(\infty) = \frac{\lambda^2}{2\mu^2 + 2\mu\lambda + \lambda^2}$$

Accordingly,

$$a = Pr_1(\infty) + Pr_2(\infty) = \frac{2\mu^2 + 2\mu\lambda}{2\mu^2 + 2\mu\lambda + \lambda^2}$$

This can be verified from the solution for $Pr_0(t)$. Since k_1 and k_2 are negative, the exponential terms approach zero, then

$$Pr_0(\infty) = B_3 = \frac{\lambda^2}{2\mu^2 + 2\mu\lambda + \lambda^2}$$

Similarly,

$$Pr_1(\infty) = A_3$$

and

$$Pr_2(\infty) = G_4$$

Thus

$$a = \frac{2\mu^2}{2\mu^2 + 2\mu\lambda + \lambda^2} + \frac{2\mu\lambda}{2\mu^2 + 2\mu\lambda + \lambda^2}$$

Therefore, the results obtained in Examples 5.12 and 5.13 are consistent.

It is clear that if a trend exists in the parameters that characterize system availability (e.g., failure rate and repair rate), one cannot use the Markovian

method; only solutions of (5.43) with time dependent ρ can be used. Solving such equations may pose difficulty in systems with many states. However, with the emergence of efficient numerical algorithms and powerful computers, solutions to these equations are indeed possible.

5.4 USE OF SYSTEM ANALYSIS TECHNIQUES IN THE AVAILABILITY CALCULATIONS OF COMPLEX SYSTEMS

In Chapter 4, we discussed a number of methods for estimating the reliability of a system from the reliability of its individual components or units. The same concept applies here also. That is, one can use the availability (or unavailability) functions for each component of a complex system and use, for example, system cut sets to obtain system availability (or unavailability). The method of determining system availability in these cases is exactly similar to the system reliability estimation methods.

Example 5.14

Assume all components of the system shown in Figure 4.4 are repairable (revealed fault) with a failure rate of 10^{-3} (hour^{-1}) and a mean down time of 15 hours. Component 7 has a failure rate of 10^{-5} (hour^{-1}), with a mean downtime of 10 hours. Calculate the average system unavailability.

Solution:

The cut sets are (7), (1, 2), (1, 5, 6), (2, 3, 4), and (3, 4, 5, 6). The unavailability of component 1 through 6, according to Table 5.2, is

$$q_{1-6} = \frac{\lambda\tau}{1+\lambda\tau} = \frac{10^{-3} \times 15}{1 + 10^{-3} \times 15} = 9.85E - 3$$

Similarly,

$$q_7 = \frac{10^{-5} \times 10}{1 + 10^{-5} \times 10} = 9.99E - 5$$

Using the rare event approximation,

$$q_{sys} = q(\text{cut sets}) = q_7 + q_1 \cdot q_2 + q_1 \cdot q_5 \cdot q_6 + q_2 \cdot q_3 \cdot q_4 + q_3 \cdot q_4 \cdot q_5 \cdot q_6$$

Thus,

$$q_{sys} = 9.99 \times 10^{-5} + 9.70 \times 10^{-5} + 9.56 \times 10^{-7} + 9.56 \times 10^{-7} + 9.41 \times 10^{-9}$$
$$= 1.99 \times 10^{-4}$$

Example 5.15

The auxiliary feedwater system in a pressurized water reactor (PWR) plant is used for emergency cooling of steam generators. The simplified piping and instrument diagram (P&ID) of a typical system like this is shown in Figure 5.11a. The reliability block diagram in Figure 5.11b represents this P&ID. Calculate the system unavailability. Assume all of the components are in standby mode and are periodically tested with the following characteristics. (Characteristics are shown collectively for each block.)

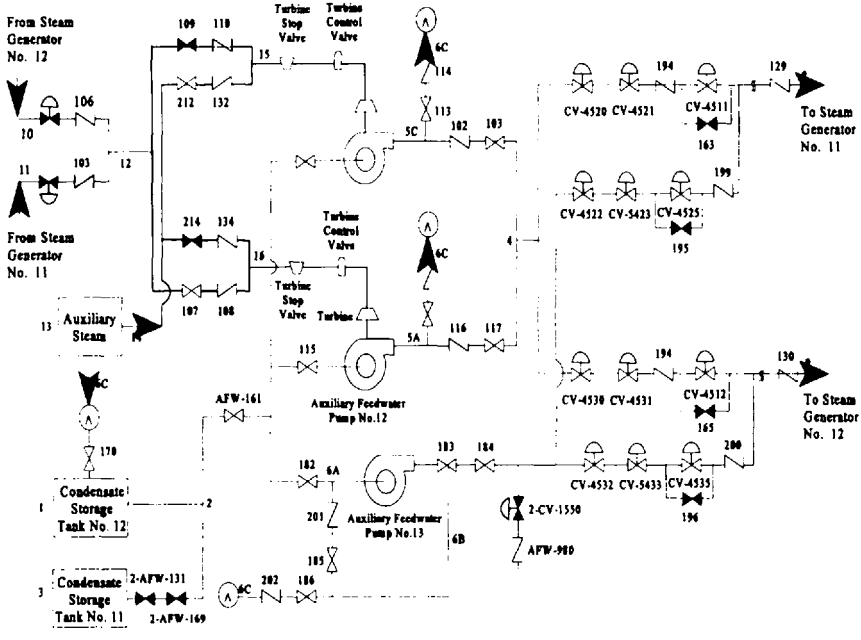


Figure 5.11a Auxiliary feedwater system simplified P&ID.

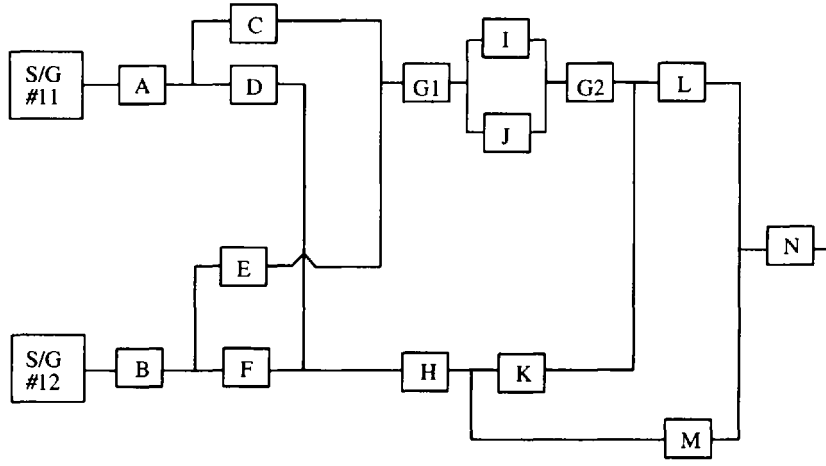


Figure 5.11b Simplified auxiliary feedwater system of a PWR.

Block name	Failure rate $(hours)^{-1}$	Frequence of repair	Average test duration (hours)	Average repair time (hours)	Test interval (hours)
A	1×10^{-7}	9.2×10^{-3}	0	5	720
B	1×10^{-7}	9.2×10^{-3}	0	5	720
C	1×10^{-6}	2.5×10^{-2}	0	10	720
D	1×10^{-6}	2.5×10^{-2}	0	10	720
E	1×10^{-6}	2.5×10^{-2}	0	10	720
F	1×10^{-6}	2.5×10^{-2}	0	10	720
$G(G_1$ and $G_2)$	1×10^{-7}	7.7×10^{-4}	0	15	720
H	1×10^{-7}	1.8×10^{-4}	0	24	720
I	1×10^{-4}	6.8×10^{-1}	2	36	720
J	1×10^{-4}	6.8×10^{-1}	2	36	720
K	1×10^{-5}	5.5×10^{-1}	2	24	720
L	1×10^{-7}	4.3×10^{-3}	0	10	720
M	1×10^{-4}	1.5×10^{-1}	0	10	720
N	1×10^{-7}	5.8×10^{-4}	0	5	720

Solution:

According to Table (5.2), we can calculate the unavailability of each block.

Block Name	Unavailability	Block Name	Unavailability
A	$1.0E-4$	H	$4.2E-4$
B	$1.0E-4$	I	$7.3E-4$
C	$7.0E-4$	J	$7.3E-4$
D	$7.0E-4$	K	$1.4E-4$
E	$7.0E-4$	L	$2.4E-4$
F	$7.0E-4$	M	$1.1E-1$
G(G_1 and G_2)	$5.2E-4$	N	$4.0E-5$

The cut sets of the block diagram in Figure 5.11b are as follows:

1) N	10) C E H	19) J I K M
2) L M	11) B D L	20) D F J I
3) H L	12) B D G	21) C E K M
4) G H	13) B C H	22) C D E H
5) A B	14) B C D	23) B D J I
6) H J I	15) A F L	24) B C K M
7) G K M	16) A E F	25) A F J I
8) D F L	17) A E H	26) A E K M
9) D G F	18) A G F	

Using the same procedure as the one used in Example 5.14 and rare event approximation, we can easily compute the average system unavailability as $q_{sys} = 7.49 \times 10^{-5}$

One important point to recognize in the availability estimation of redundant systems with periodically tested components is that components whose simultaneous failures cause the system to fail (i.e., sets of components in each cut set of the system) should be tested in a *staggered* manner. This way the system would not become totally unavailable during the testing and repair of its components. For example, consider a system of two parallel units, each of which is periodically tested and has a pointwise unavailability behavior that can be approximated by the model shown in Figure 5.6. If the components are not tested in a staggered manner, the system's pointwise unavailability exhibits the shape shown in Figure 5.12.

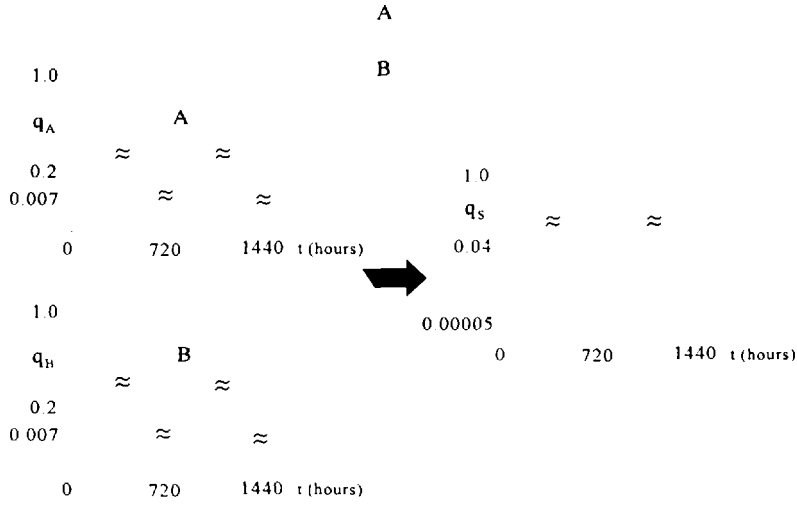


Figure 5.12 Unavailability of a parallel system using nonstaggered testing.

On the other hand, if the components are tested in a staggered manner, the system unavailability would exhibit the shape illustrated in Figure 5.13.

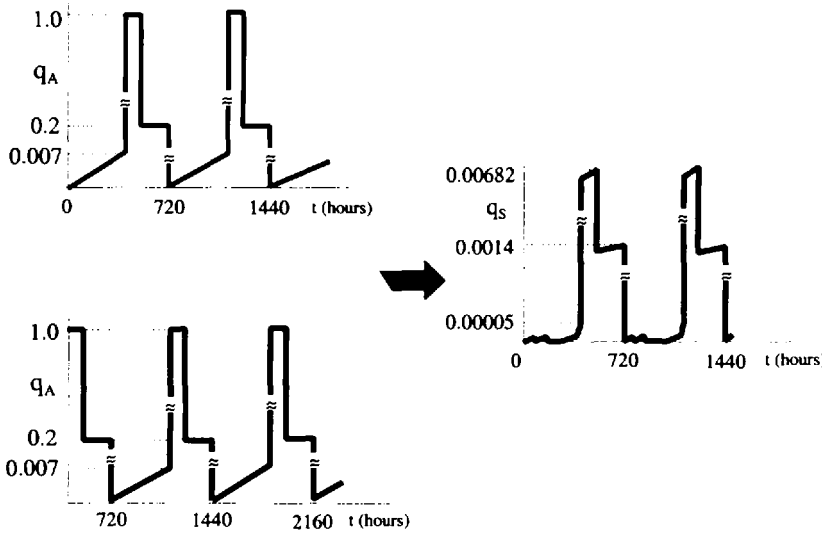


Figure 5.13 Unavailability of a parallel system using staggered testing.

Clearly, the average unavailability in the case of staggered testing is lower. This subject is discussed in more detail by Vesely and Goldberg (1981) and Ginzburg and Vesely (1990). Also, to minimize unavailability, one can find an optimum value for test interval as well as the optimum degree of staggering.

Modarres (1984) has suggested a simple method for estimating approximate average system unavailability of a series-parallel system having a single input node and single output node, and repairable (revealed fault) components. In this method, it is assumed that the components or blocks are independent and $\lambda_i\tau_i << 1$ for each component or block of the system, where λ_i is the constant failure rate (i.e., no failure rate trend is assumed), and τ_i is the component's mean downtime. In this method, series and parallel blocks of the system are systematically replaced with equivalent "super blocks." The equivalent failure rate (or occurrence rate) λ and mean downtime τ of the super blocks can be calculated from Table 5.3. Example 5.16 is an illustration of the application of this method.

Example 5.16

Consider the series-parallel system shown in Figure 5.14, with the component data shown in Table 5.4.

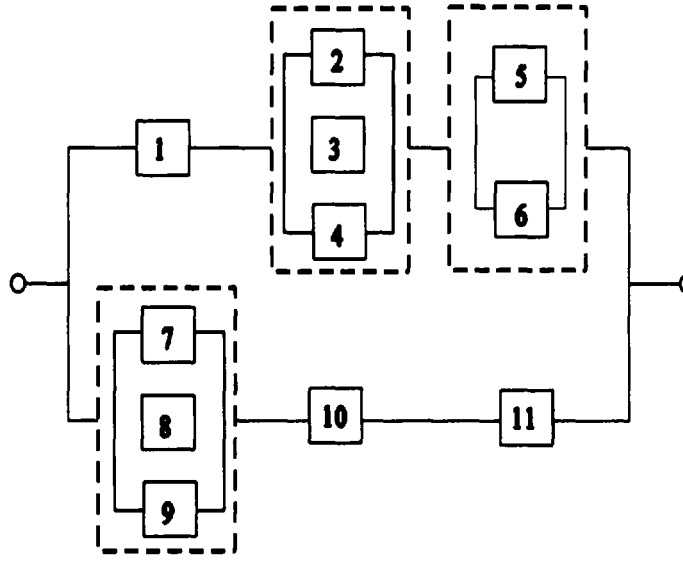


Figure 5.14 Sample series-parallel system.

This system is composed of two parallel blocks. Each block is composed of sub-block(s) and component(s).

a. Determine the approximate occurrence rate λ and mean downtime τ of this system.
b. Determine the approximate average unavailability of the system.

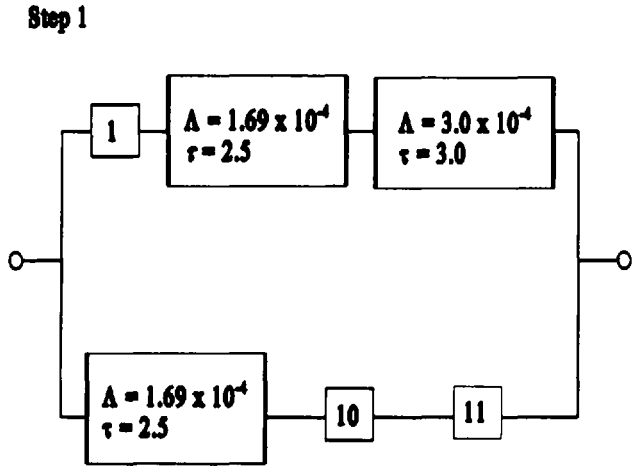


Figure 5.15a Step-by-step resolution of the system in Figure 5.13.

Solution:

Assuming independence between blocks and super-blocks:

a. The super-blocks are enclosed by dotted lines in Figure 5.14. First, all of the blocks are resolved and their equivalent λ and τ are obtained. Next, their equivalent λ and τ are determined. Finally, the whole system is resolved. Equations in Table 5.3 are applied to the system along with the failure data summarized in Table 5.4 to obtain λ and τ values. The steps are illustrated in Figures 5.15a and 5.15b.
b. The approximate unavailability of the system can be calculated using $q = \lambda\,\tau/(1 + \lambda\,\tau)$ from Table 5.2. Thus, $q = 2.9 \times 10^{-5} \times 2.3/(1 + 2.9 \times 10^{-5} \times 2.3) = 6.67 \times 10^{-5}$. This can be compared with the direct calculation method using the cut set concept (similar to Examples 5.14 and 5.15), which yields the average system unavailability of 6.57×10^{-5}. The

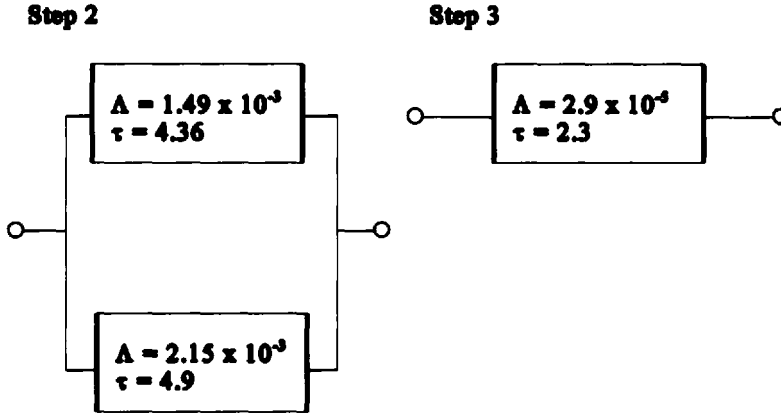


Figure 5.15b Step-by-step resolution of the system in Figure 5.13.

difference is due to the approximate nature of this approach and the assumption that the whole system's time to failure approximately follows an exponential distribution.

Table 5.3 Failure Characteristics for Parallel or Series

Type of block	Block failure characteristic	
	Occurrence rate Λ	Mean down time τ
Parallel	$\left[\prod_{i=1}^{n} \lambda_i \tau_i\right] \sum_{i=1}^{n} \frac{1}{\tau_i}$	$\frac{1}{\sum_{i=1}^{n} \frac{1}{\tau_i}}$
Series	$\left(1 - \sum_{i=1}^{n} \lambda_i \tau_i\right) \sum_{i=1}^{n} \lambda_i$	$\frac{\sum_{i=1}^{n} \lambda_i \tau_i}{\left(1 - \sum_{i=1}^{n} \lambda_i \tau_i\right)^2 \sum_{i=1}^{n} \lambda_i}$

Table 5.4 Summary of Failure Data for the Components Shown in Figure 5.13

Component serial number	Failure rate λ_i (per 1000 hour)	Mean downtime τ_i (hour)
1	1	5.0
2	10	7.5
3	10	7.5
4	10	7.5
5	5	6.0
6	5	6.0
7	10	7.5
8	10	7.5
9	10	7.5
10	10	5.0
11	10	5.0

EXERCISES

5.1 The following shows fire incidents during 6 equal time intervals of 22 chemical plants.

Time interval	1	2	3	4	5	6
No. of fires	6	8	16	6	11	11

Do you believe the fire incidents are time-dependent? Prove your answer.

5.2 A simplified schematic of the electric power system at a nuclear power plant is shown in the figure below.

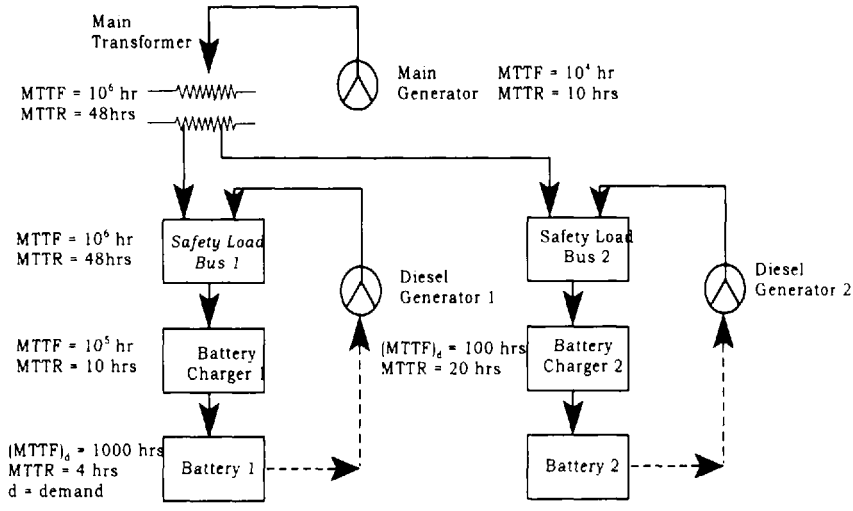
a. Draw a fault tree with the top event "Loss of Electric Power from Both Safety Load Buses."
b. Determine the unavailability of each event in the fault tree for 24 hours of operation.
c. Determine the top event probability.

Assume the following:

Either the main generator or one of the two diesel generators is sufficient.

One battery is required to start the corresponding diesel generator.

Normally, the main generator is used. If that is lost, one of the diesel generators provides the electric power on demand.



5.3 An operating system is repaired each time it has a failure and is put back into service as soon as possible (monitored system). During the first 10,000 hours of service, it fails five times and is out of service for repair during the following times:

1000–1050 hrs
3660–4000 hrs
4510–4540 hrs
6130–6170 hrs
8520–8560 hrs

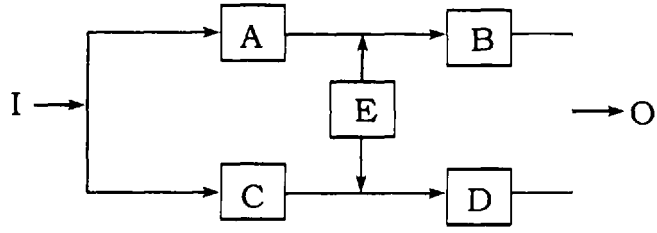
a) Is there a trend in the data?
b) What is the reliability of the system 100 hours after the system is put into operation? What is the asymptotic availability assuming no trends in λ and μ?
c) If the system has been operating for 10 hours without a failure, what is the probability that it will continue to operate for the next 10 hours without a failure?
d) What is the 80% confidence interval for the mean time to repairs ($\tau = 1/\mu$)?

5.4 The following cycle-to-failure data have been obtained from a repairable component. The test stopped when the 5th failure occured.

Repair no.	1	2	3	4	5
Cycle-to-failure (interarrival of cycles)	5010	6730	4031	3972	4197

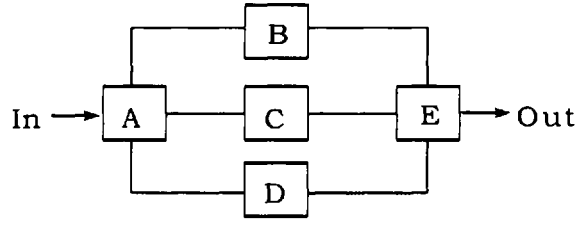
a) Is there any significant trend in these data?
b) Determine the rate of occurrence of failures.
c) What is the reliability of the component 1000 cycles after the 5^{th} failure is repaired?

5.5 Determine the limiting pointwise unavailability of the system shown below:



Assume that all components are identical and are repaired immediately after each experiences a failure. Rate of occurrence of the failure for each component is $\lambda = 0.001(\text{hour})^{-1}$, and mean-time-to-repair is 15 hours.

5.6 We are interested in unavailability of the system shown below:



The following information is available:

A and E are identical components with $\lambda_A = \lambda_E = 1 \times 10^{-5}$ /hr, $\mu_A = \mu_E = 0.1$/hr.

B, C, and D are identical periodically tested components with $\lambda_B = \lambda_C = \lambda_D = 1 \times 10^{-5}$ /hr. All test durations are equal ($\tau_t = 1$ hr), all frequency of repair per cycle are equal ($f = 0.25$), and all durations of repair are equal ($\tau_r = 15$ hr).

Given the above information, calculate unavailability of the system assuming that all components are independent.

REFERENCES

Ascher, H. and H. Feingold, "*Repairable Systems Reliability: Modeling and Inference, Misconception and Their Causes*," Marcel Dekker, New York, 1984.

Bain, L.J., "*Statistical Analysis of Reliability and Life-Testing Models Theory and Methods*," Marcel Dekker, New York, 1978.

Barlow, R.E. and Proschan, F., "*Statistical Theory of Reliability and Life Testing: Probability Models*," To Begin With, Silver Spring, MD, 1981.

Bassin, W.M., "*Increasing Hazard Functions and Overhaul Policy*," *ARMS IEEE-69C 8-R*, pp. 173–180, 1969.

Bassin, W.M., "*A Bayesian Optimal Overhaul Interval Model for the Weibull Restoration Process*," *J. Am. Stat. Soc.* 68, pp. 575–578, 1973.

Caldorela, G., "*Unavailability and Failure Intensity of Components*," Nuclear Engineering and Design J., 44, p. 147, 1977.

Cox, D.R. and P.A. Lewis, "*The Statistical Analysis of Series and Events*," Methuen, London, 1978.

Crow, L.H., "*Reliability Analysis for Complex Repairable Systems, Reliability and Biometry*," F. Proschan and R.J. Serfling, eds., SIAM, Philadelphia, 1974.

Crow, L.H., "*Evaluating the Reliability of Repairable Systems*," Proc. of Ann. Rel. and Maint. Symp., IEEE, Orlando, FL, 1990.

Ginzburg, T. and Vesely, W.E., "*FRANTIC-ABC User's Manual: Time-Dependent Reliability Analysis and Risk Based Evaluation of Technical Specifications*," Applied Biomathematics, Inc., Setauket, New York, 1990.

Hoyland, A., and Rausand, M., "*System Reliability Theory: Models and Statistical Methods*," John Wiley and Sons, New York, 1994.

Judge, G.G., Hill, R.C., Griffiths, W.E., Lutkepohl, H., and Lee, T.-Ch, "*Introduction to the Theory and Practice of Econometrics*," John Wiley and Sons, New York, 1980.

Kaminskiy, M., and Krivtsov, V., "*A Monte Carlo Approach to Warranty Repair Predictions*," SAE Technical Paper Series, #972582, SAE Aerospace International RMLS Conference, Dallas, TX, 1997.

Lawless, J.F., "*Statistical Models and Methods for Lifetime Data*," Wiley, New York, 1982.

Leemis, L.M., "*Reliability: Probabilistic Models and Statistical Methods*," Prentice-Hall, Englewood Cliffs, New Jersey, 1995.

Lofgren, E., "*Probabilistic Risk Assessment Course Documentation*," U.S. Nuclear Regulatory Commission, NUREG/CR-4350, Vol.5–System Reliability and Analysis Techniques, Washington, DC, 1985.

Modarres, M., "*A Method of Predicting Availability Characteristics of Series-Parallel Systems*," IEEE Transaction on Reliability, R-33, 4, pp. 309–312, 1984.

O'Connor, P., "*Practical Reliability Engineering*," 3rd edition, Wiley, New York, 1991.

Vesely, W.E., Goldberg, F.F., Powers, J.T., Dickey, J.M., Smith, J.M., and Hall, E. "*FRANTIC II—A Computer Code for Time-Dependent Unavailability Analysis*," U.S. Nuclear Regulatory Commission, NUREG/CR-1924, Washington, DC, 1981.

This page intentionally left blank

6
Selected Topics in Reliability Modeling

In this chapter, we will discuss a number of topics important to reliability modeling. These topics are not significantly related to each other, nor are they presented in a particular order. Some of the topics are still the subject of current research; the methods presented represent a summary of the state of the art.

6.1 STRESS-STRENGTH ANALYSIS

As discussed in Chapter 1, a failure occurs when the stress applied to an item exceeds its strength. The probability that no failure occurs is equal to the probability that the applied stress is less than the item's strength, i.e.,

$$R = \Pr(S > s) \tag{6.1}$$

where:

R is the reliability of the item,
s is the applied stress, and
S is the item's strength.

Examples of stress related failures include the following:

1. Misalignment of a journal bearing, lack of lubricants, or incorrect lubricants generating an internal load (mechanical or thermal stress) that causes the bearing to fail.
2. The voltage applied to a transistor gate is too high, causing a high temperature that melts the transistor's semiconductor material.
3. Cavitation causes pump failure, which in turn causes a violent vibration that ultimately breaks the rotor.

4. Lack of heat removal from a feed pump in a power plant results in overheating of the pump seals, causing the seals to break.
5. Thermal shock causing a pressurized vessel to experience fracture due to crack growth

Engineers need to ensure that the strength of an item exceeds the applied stress for all possible stress situations. Traditionally, in the deterministic design process, safety factors are used to cover the spectrum of possible applied stresses. This is generally a good engineering principle, but failures occur despite these safety factors. On the other hand, safety factors that are too stringent result in over design, high cost, and sometimes poor performance.

If the range of major stresses is known or can be estimated, a probabilistic approach can be used to address the problem. This approach eliminates over design, high cost and failures caused by stresses that are not considered early in the design. If the distribution of S and s can be estimated as $F(S)$ and $g(s)$, then

$$R = \int_0^\infty g(S)\, dg(s)$$

$$F = \int_0^\infty F(S) \left[\int_S^\infty f(s)\, ds \right] dS \tag{6.2}$$

Figure 6.1 shows typical relation between $F(S)$ and $g(s)$ distributions.

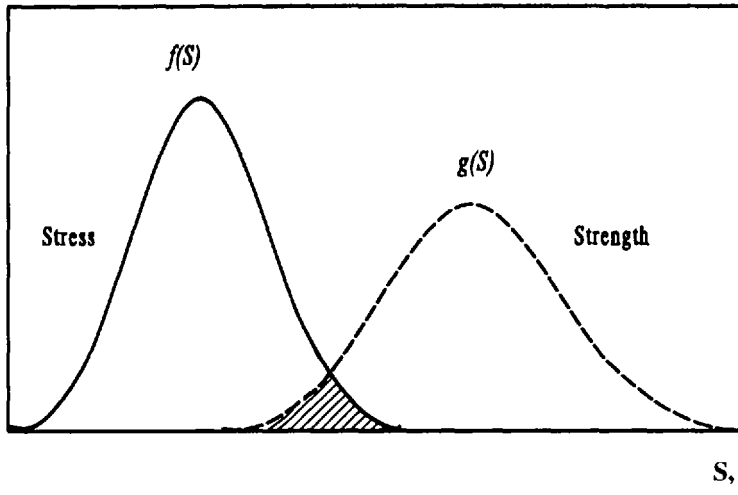


Figure 6.1 Stress-strength distributions.

The *Safety Margin* (SM) is defined as

$$\text{SM} = \frac{E(S) - E(s)}{\sqrt{\text{var}(S) + \text{var}(s)}} \tag{6.3}$$

The SM shows the relative difference between the mean values for stress and for strength. The larger the SM, the more reliable the item will be. Use of (6.3) is a more objective way of measuring the safety of items. It also allows for calculation of reliability and probability of failure as compared with the traditional deterministic approach using safety factors. However, good data on the variability of stress and strength are often not easily available. In these cases, engineering judgement can be used to obtain the distribution including engineering uncertainty. The section on expert judgement explains methods for doing this in more detail.

The distribution of stress is highly influenced by the way the item is used and the internal and external operating environments. The design determines the strength distribution, and the degree of quality control in manufacturing primarily influences the strength variation.

It is easy to show that for a normally distributed S and s,

$$R = \phi(\text{S M}) \tag{6.4}$$

where ϕ(SM) is the cumulative standard normal distribution with z = SM (see Table A.1).

Example 6.1

Consider the stress and strength of a beam in a structure represented by the following normal distributions:

$$\mu_S = 420 \text{ kg/cm}^2 \text{ and } \sigma_S = 32 \text{ kg/cm}^2$$
$$\mu_s = 310 \text{ kg/cm}^2 \text{ and } \sigma_s = 72 \text{ kg/cm}^2$$

What is the reliability of this structure?

Solution:

$$\text{SM} = \frac{420 - 310}{\sqrt{32^2 + 72^2}} = 1.4$$

with $z = 1.4$ and using Table A.1,

$$R = \phi(1.4) = 0.91$$

Example 6.2

A random variable representing the strength of a nuclear power plant containment building follows a lognormal distribution with the mean strength of 0.905 MPa, and standard deviation of 0.144 MPa. Four possible accident scenarios can lead to high pressure conditions inside the containment that may exceed its strength. The pressures cannot be calculated precisely, but can be represented as another random variable that follows a lognormal distribution.

a. For a given accident scenario that causes a mean pressure load inside the containment of 0.575 MPa with a standard deviation of 0.117 MPa, calculate the probability that the containment fails.

b. If the four scenarios are equally probable and each leads to high pressure conditions inside the containment with the following mean and standard deviations, calculate the probability that the containment fails.

μ_L(MPa)	0.575	0.639	0.706	0.646
σ_L(MPa)	0.117	0.063	0.122	0.061

c. If the containment strength distribution is divided into the following failure mode contributors with the mean failure pressure and standard deviation indicated, repeat part a.

Failure mode	Mean pressure, μ_S(MPa)	Standard deviation, σ_S(MPa)
Liner tear around personnel airlock	0.910	1.586*E* 3
Basemat shear	0.986	1.586*E* 3
Cylinder hoop membrane	1.089	9.653*E* - 4
Wall-basemat junction shear	1.131	1.586*E* - 3
Cylinder longitudinal membrane	1.241	1.034*E* 3
Dome membrane	1.806	9.653*E* - 4
Personnel air lock door buckling	1.241	1.655E 3

Solution:

If S is a normally distributed r.v. representing strength, and L is a normally distributed r.v. representing pressure stress (load), then the r.v., $Y = \ln(S) - \ln(L)$, is also normally distributed.

For the lognormal distribution with mean and standard deviation of μ_y and σ_y, the respective mean and standard deviation of the normal distribution, μ_t and σ_t, can be obtained using (2.47) and (2.48). Then:

a.
$$R = \Phi(SM) = \Phi\left(\frac{\mu_{St} - \mu_{Lt}}{\sqrt{\sigma_{St}^2 + \sigma_{Lt}^2}} \right)$$

$$= \Phi\left(\frac{-0.112 - (-0.574)}{\sqrt{0.158^2 + 0.201^2}} \right) = 0.9649$$

The probability of containment failure:

$$F = 1 - R = 0.0351$$

b. Because the four scenarios are "equally probable", then the system is equivalent to a series system, such that: $R = R_1 \times R_2 \times R_3 \times R_4$.

$$R_1 = \Phi(SM_1) = \Phi(1.81) = 0.9649$$
$$R_2 = \Phi(SM_2) = \Phi(1.83) = 0.9664$$
$$R_3 = \Phi(SM_3) = \Phi(1.07) = 0.8577$$
$$R_4 = \Phi(SM_4) = \Phi(1.79) = 0.9633$$

The probability of containment failure:

$$F = 1 - R = 1 - R_1 \times R_2 \times R_3 \times R_4 = 0.2296$$

c. Because each failure mode may cause a system failure, this case can be treated as a series system. Because we know the median of the lognormal distribution instead of the mean, it takes several algebra steps to solve for the respective means and standard deviations.

$$R_a = \Phi(SM_a) = \Phi(2.38) = 0.9913$$
$$R_b = \Phi(SM_b) = \Phi(2.78) = 0.9973$$
$$R_c = \Phi(SM_c) = \Phi(3.27) = 0.9995$$
$$R_d = \Phi(SM_d) = \Phi(3.46) = 0.9997$$
$$R_e = \Phi(SM_e) = \Phi(3.92) \approx 1$$
$$R_f = \Phi(SM_f) = \Phi(5.78) \approx 1$$
$$R_g = \Phi(SM_g) = \Phi(3.92) \approx 1$$

The probability of containment failure:

$$F = 1 - R = 1 - R_a \times R_b \times R_c \times R_d \times R_e \times R_f \times R_g = 0.0122$$

If both the stress and strength distributions are exponential with parameters λ_s and λ_S, the reliability can be estimated as:

$$R = \frac{\lambda_s}{\lambda_s + \lambda_S}$$

For more information about stress-strength methods in reliability analysis, the readers are referred to O'Connor (1991) and Kapur and Lamberson (1977).

6.2 SOFTWARE RELIABILITY ANALYSIS

6.2.1 Introduction

Many techniques have been developed for analyzing the reliability of physical systems. However, their extension to software has been problematic for two reasons. First, software faults are design faults, while faults in physical systems are equipment breakage or human error. Second, software systems are more complex than physical systems, so the same reliability analysis methods may be impractical to use.

Software has deterministic behavior, whereas hardware behavior is both deterministic and probabilistic. Indeed, once a set of inputs to the software has been selected, and provided that the computer and operating system with which the software will run is error free, the software will either fail or executes correctly. However, our knowledge of the inputs selected, of computer, of the operating system, and of the nature and position of the fault may be uncertain. One may, however, translate this uncertainty into probabilities. A software fault is a triggering event that causes software error. A software bug (error in the code) is an example of a fault.

Accordingly, we adopt a probabilistic definition for software reliability. Software reliability is the probability that the software product will not fail for a specified time under specified conditions. This probability is a function of the input to and use of the product, as well as a function of the existence of faults in the software. The inputs to the product will determine whether an existing fault is encountered or not.

Faults can be grouped as design faults, operational faults or transient faults. All software faults are design faults; however, hardware faults may occur in any of the three classes. Faults can also be classified by the source of the fault; software and hardware are two of the possible sources of the fault. Sources of

faults are: input data, system state, system topology, humans, environment, and unknown causes. For example, the source of many transient faults is unknown.

Failures in software are classified by mode and scope. A failure mode may be sudden or gradual; partial or complete. All four combinations of these are possible. The scope of failure describes the extent within the system of the effects of the failure. This may range from an internal failure, whose effect is confined to a single small portion of the system, to a pervasive failure, which affects much of the system, see Lawrence (1993).

Software, unlike hardware, is unique in that its failure modes are the result of design flaws, as opposed to any kind of internal physical mechanisms and external environmental conditions such as aging, for example see McDermid, (1991). As a result, traditional reliability techniques, which tend to focus on physical component failures rather than system design faults, have been unable to close the widening gap between the powerful capabilities of modern software systems and the levels of reliability that can be computed for them. The real problem of software reliability is one of managing complexity. There is a natural limitation on the complexity of hardware systems. With the introduction of digital computer systems, however, designers have been able to arbitrarily implement complex designs in software. The result is that the central assumption implicit in traditional reliability theory, that the design is correct and failures are the result of fallible components, is no longer valid.

In order to assess the reliability of a software, a software reliability model will be needed. In the remainder of the section details of classes of software reliability models, and two such models are discussed. Also discussed are the models used to assess software life cycle.

6.2.2 Software Reliability Models

Several software reliability models (SRMs) have been developed over the years. These techniques are variously referred to as "analyses" or "models," but there is a distinct difference between the two. An analysis (such as fault tree analysis) is carried out by creating a model (the fault tree) of a system, and then using that model to calculate properties of interest, such as reliability.

The standard reliability models such as fault tree analysis (FTA), event tree analysis (ETA), failure modes and effect analysis (FMEA), and Markov models discussed in this book are adequate for systems whose component remain unchanged for long periods of time. They are less flexible for systems that undergo frequent design changes. If, for example, the failure rate of a component is improved through design or system configuration changes, the reliability model must be re-evaluated. A reliability growth model (see Section 6.6) is more appropriate for these cases. In this model, a software is tested for a period of time, during which failures may occur. These failures lead to modification to the design

or manufacture of a component; the new version then goes back into test. This cycle is continued until design objectives are met. Software reliability growth is a very active research area today.

When these models are applied to software reliability one can group them into two main categories: predictive models and assessment models (Smidts (1996)). Predictive models typically address the reliability of software early in the design cycle. Different elements of a life cycle development of software is discussed later. Predictive models are developed to assess the risks associated with the development of software under a given set of requirements and for specified personnel before the project truly starts. Predictive software reliability models are few in number (Smidts (1996)), and as such in this section the predictive models are not discussed. Assessment models evaluate present and project future software reliability from failure data gathered when the integration of the software starts.

Classification

Most existing SRMs may be grouped into four categories:

1. Time between failure model
2. Fault seeding model
3. Input-domain based model
4. Failure count model

Each category of models is summarized as follows:

Time Between Failure Model. This category includes models that provide an estimate of the times between failures in a software. Key assumptions of this model are

independent time between successive failures,
equal probability of exposure of each fault,
embedded faults are independent of each other,
no new faults introduced during corrective actions.

Specific SRMs that estimate mean-time-between-failures are:

Jelinski–Moranda (1972) model,
Schick and Wolverton (1973) model,
Littlewood–Verrall's Bayesian model (Littlewood and Verrall (1973) and Littlewood (1979)),
Goel and Okumoto (1979) imperfect debugging model.

Fault Seeding Model. This category of SRMs includes models that assess the number of faults in the software at time zero via seeding extraneous faults. Key assumptions of this model are:

seeded faults are randomly distributed in the software,
indigenous and seeded faults have equal probabilities of being detected.

The specific SRM that falls into this category is Mills fault seeding model (Mills (1972)). In this model, an estimate of the number of defects remaining in a program can be obtained by a seeding process that assumes a homogeneous distribution of representative class of defects. The variables in this measure are: the number of seed faults introduced N_S, the number of intentional seed faults found n_s, and the number of faults found n_F that were not intentionally seeded.

Before seeding, a fault analysis is needed to determine the types of faults expected in the code and their relative frequency of occurrence. An independent monitor inserts into the code N_S faults that are representative of the expected indigenous faults. During testing, both seeded and unseeded faults are identified. The number of seeded and indigenous faults discovered permits an estimate of the number of faults remaining for the fault type considered. The measure cannot be computed unless some seeded faults are found. The maximum likelihood estimate of the unseeded faults is given by

$$\hat{N}_F = n_F N_S / n_S \tag{6.5}$$

Example 6.3

Forty faults of a given type are seeded into a code and, subsequently, 80 faults of that type are uncovered: 32 seeded and 48 unseeded. Calculate an estimate of unseeded faults. How many faults remain to be found?

Solution:

Using (6.5), $N_s = 60$, and the estimate of faults remaining is

$$N_F \text{ (remaining)} = N_F - n_F = 60 - 48 = 12$$

Input-Domain Based Model. This category of SRMs includes models that assess the reliability of software when the test cases are sampled randomly from a well-known operational distribution of software inputs. The reliability estimate is obtained from the number of observed failures during execution. Key assumptions of these models are:

input profile distribution is known,
random testing is used (input are selected randomly),
input domain can be partitioned into equivalence classes.

Specific models of this category are:

Nelson's model (Nelson (1978)),
Ramamoorthy and Bastani's model (Ramamoorthy and Bastani (1982)).

We will further elaborate on Nelson's model.

Nelson's Model. This model is typically used for systems with ultrahigh-reliability requirements, such as software used in nuclear power plants and are limited to about 1000 lines of code. The model is applied to the validation phase of the software (acceptance test) to estimate the reliability. Nelson defines the reliability of a software run n times (for n test cases) and which failed n_f times as

$$R = 1 - n_f/n \tag{6.6}$$

where n is the total number of test cases and n_f is the number of failures experienced out of these test cases.

Failure Count Model. This category of SRMs estimate the number of faults or failures experienced in specific intervals of time. Key assumptions of these models are:

test intervals are independent of each other,
testing intervals are homogeneously distributed,
number of faults detected during nonoverlapping intervals are independent of each other.

The SRMs that fall into this category are

Shooman's exponential model (1975),
Goel–Okumoto's nonhomogeneous Poisson process (1979),
Musa's execution time model (Musa et al. (1987)),
Goel's generalized nonhomogeneous Poisson process model (1983),
Musa–Okumoto's logarithmic Poisson execution time model (Musa et al. (1987)).

We will further elaborate on the Musa and Musa–Okumoto's models.

Musa Basic Execution Time Model (BETM) Model. This model (Musa (1975)) assumes that failures occur in the form of a nonhomogeneous Poisson process. The unit of failure intensity is the number of failures per central process unit (CPU) time. This relates failure events to the processor time used by the software. In the BETM, the reduction in the failure intensity function remains constant, irrespective of whether any failure is being fixed.

The failure intensity, as a function of number of failures experienced, is obtained from:

$$\lambda(\mu) = \lambda_0 \left(1 - \frac{\mu}{\nu_0} \right) \tag{6.7}$$

where

$\lambda(\mu)$ is the failure intensity (failures per CPU-hour),

λ_0 is the initial failure intensity at the start of execution,

μ is the expected number of failures experienced up to a given point in time,

ν_0 is the total number of failures.

The number of failures that should be fixed in order to move from a present failure intensity, to a target intensity, is given by

$$\Delta\mu = \frac{\nu_0}{\lambda_0} (\lambda_p - \lambda_F) \tag{6.8}$$

where p is the present failure intensity λ_F is the target (final) failure intensity. The execution time required to reach this objective is

$$\Delta\tau = \frac{\nu_0}{\lambda_0} \ln \left(\frac{\lambda_p}{\lambda_F} \right) \tag{6.9}$$

In these equations, ν_0 and λ_0 can be estimated in different ways, see Musa et al. (1987).

Musa–Okumoto Logarithmic Poisson Time Model (LPETM), (Musa et al. (1987)). According to the LPETM, the failure intensity is given by

$$\lambda(\mu) = \lambda_0 e^{-\theta\mu} \tag{6.10}$$

where θ is the failure intensity decay parameter and λ, μ, and λ_0 are the same as in the BETM. This model assumes that repair of the first failure has the greatest impact in reducing failure intensity and that the impact of each subsequent repair decreases exponentially.

In the LPETM, no estimate of ν_0 is needed. The expected number of failures that must occur to move from a present failure intensity of λ_p to a target intensity of λ_F is

$$\Delta\mu = \frac{1}{\theta} \ln \left(\frac{\lambda_p}{\lambda_F} \right) \tag{6.11}$$

The execution time to reach this objective is given by

$$\Delta\tau = \frac{1}{\theta}\left(\frac{1}{\lambda_F} - \frac{1}{\lambda_p}\right) \tag{6.12}$$

As we have seen, the execution time components of these models are characterized by two parameters. These are listed in Table 6.1.

Table 6.1 Execution Time Parameters

Parameter	Model	
	Basic	Logarithmic Poisson
Initial failure intensity	λ_0	λ_0
Failure intensity change		
Total failures	ν_0	—
Failure intensity decay parameter	—	θ

Example 6.4

Assume that a software will experience 200 failures in its lifetime. Suppose, it has now experienced 100 of them. The initial failure intensity was 20 failures/CPU-hour. Using BETM and LPETM calculate the current failure intensity (assume failure intensity decay parameter is 0.02/failure).

Solution:

For BETM,

$$\lambda(\mu) = \lambda_0\left(1 - \frac{\mu}{\nu_0}\right)$$

$$= 20\left(1 - \frac{100}{200}\right) = 10 \text{ failures per CPU-hour}$$

For LPETM,

$$\lambda(\mu) = \lambda_0 e^{-\theta\mu}$$

$$= 20\, e^{-(0.02)(100)} = 2.70 \text{ failures per CPU-hour}$$

The most common approach to software reliability analysis is testing. Testing is often performed by feeding random inputs into the software and

observing the output produced to discover incorrect behavior. Because of the extremely complex nature of today's modern computer systems, however, these techniques often result in the generation of an enormous number of test cases. For example, Petrella et al. (1991) discuss Ontario Hydro's validation testing of its Darlington Nuclear Generating Station's new computerized emergency reactor shutdown systems that required a minimum of 7000 separate tests to demonstrate 99.99% reliability at 50% confidence.

Software reliability growth models have not had a great impact so far in reducing the quantity and cost of software testing necessary to achieve a reasonable level of reliability.

6.2.3 Software Life Cycle Models

Many different life cycle models exist for developing software systems. These differ in the timing that various activities must be done in order to produce a high-quality software product. According to Boehm (1988) the following types of process models exist:

sequential models
loop models (waterfall models)
V-models (V stands for verification)
viewpoint models
spiral models

These models have different motivations, strengths, and weaknesses. Many reliability, performance, and safety problems can be resolved only by the careful design of a software product. These must be addressed early in the life cycle, no matter which life cycle model is used. The life cycle models generally require the same type of tasks to be carried out; they differ in the ordering of these tasks in time (Lawrence (1993)). We will further elaborate on the Waterfall model.

The Waterfall Model. This is a life cycle model for software development. The classical waterfall model of software development assumes that each phase of the life cycle can be completed before the next phase can start (Pressman (1987)). The model permits the developer to return to previous phases. For example, if a requirements error is discovered during the implementation phase (see Figure 6.2), the developer is expected to halt the development, return to the requirement phase, fix the problem, change the design accordingly, and then restart the implementation from the revised design. In practice, one may only stop the implementation affected by the newly discovered requirement. The waterfall model has been severely criticized as not being realistic to many software development situations (Lawrence (1993)). Despite all of these concerns it remains a useful model for situations where the requirements are known and stable before development begins, and where little change to requirements is anticipated (Lawrence (1993)).

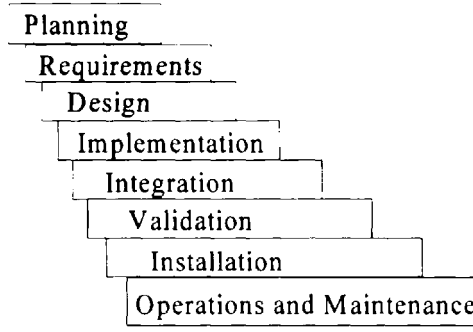


Figure 6.2 The waterfall model.

6.3 HUMAN RELIABILITY

It has long been recognized that human error has a substantial impact on the reliability of complex systems. Accidents at Three Mile Island and Chernobyl clearly show how human error can defeat engineered safeguards and play a dominant role in the progression of accidents. About 70% of aviation accidents are caused by human malfunctions, similar figures apply to the shipping and process industry. The reactor safety Study (1975) revealed that more than 60% of the potential accidents in the nuclear industry are related to human errors. In general, the human contribution to overall system performance is at least as important as that of hardware reliability.

To obtain a precise and accurate measure of system reliability, human error must be taken into account. Analysis of system designs, procedures, and postaccident reports shows that human error can be an immediate accident initiator or can play a dominant role in the progress of undesired events. Without incorporating human error probabilities, the results are incomplete and often underestimated.

To estimate human error probabilities (and, thus, human reliability), one needs to understand human behavior. However, human behavior is very difficult to model. Literature shows that there is not a strong consensus on the best way to capture all human actions and quantify human error probabilities. The assumptions, mechanisms, and approaches used by any one specific human model cannot be applied to all human activities. Current human models need further advancement, particularly in capturing and quantifying intentional human errors. Limitations and difficulties in current human reliability analysis (HRA) include the following:

1. Human behavior is a complex subject that cannot be described as a simple component or system. Human performance can be affected by social, environmental, psychological, organizational, and physical factors that are difficult to quantify.
2. Human actions cannot be considered to have binary success and failure states, as in hardware failure. Furthermore, the full range of human interactions have not been fully analyzed by HRA methods.
3. The most difficult problem with HRA is the lack of appropriate data on human behavior in extreme situations.

Human error may occur in any phase of the design, manufacturing, construction, and operation of a complex system. Design, manufacturing, and construction errors are also the cause of many types of errors during system operation. The most notable errors are dependent failures whose occurrence can cause loss of system redundancy. These may be discovered in manufacturing and construction, or during system operation. Normally, quality assurance programs are designed and implemented to minimize the occurrence of these types of human error.

In this book, we are concerned with human reliability during system operation, where human operators are expected to maintain, supervise, and control complex systems. In the remainder of this section, human reliability models are reviewed, and important models are described in some detail. Emphasis is on the basic ideas, advantages, and disadvantages of each model, and their applicability to different situations. Then, we describe the important area of data analysis in HRA. After the links between models and data are reviewed, the problems of human reliability data sources and respective data acquisition are addressed.

6.3.1 Human Reliability Analysis Process

A comprehensive method of evaluating human reliability is the method called systematic human action reliability procedure (SHARP) developed by Hannaman and Spurgin (1984). The SHARP defines seven steps to perform HRA. Each step consists of inputs, activities, rules, and outputs. The inputs are derived from prior steps, reliability studies, and other information sources, such as procedures and accident reports. The rules guide the activities which are needed to achieve the objectives of each step. The output is the product of the activities performed by analysts. The goals for each step are as follows:

1. Definition: Ensure that all different types of human interactions are considered.
2. Screening: Select the human interactions that are significant to system reliability.
3. Qualitative Analysis: Develop a detailed description of important human actions.

4. Representation: Select and apply techniques to model human errors in system logic structures, e.g., fault trees, event trees, MLD, or reliability block diagram.
5. Impact Assessment: Explore the impact of significant human actions identified in the preceding step on the system reliability model.
6. Quantification: Apply appropriate data to suitable human models to calculate probabilities for various interactions under consideration.
7. Documentation: Include all necessary information for the assessment to be understandable, reproducible, and traceable.

The relationships among these steps are shown in Figure 6.3. These steps in human reliability consideration are described in more detail below.

Step 1: Definition

The objective of Step 1 is to ensure that key human interactions are included in the human reliability assessment. Any human actions with a potentially significant impact on system reliability must be identified at this step to guarantee the completeness of the analysis.

Human activities can generally be classified in Figure 6.3.

Type 1: Before any challenge to a system, an operator can affect availability, reliability, and safety by restoring safeguard functions during testing and maintenance.

Type 2: By committing an error, an operator can initiate a challenge to the system causing the system to deviate from its normal operating envelope.

Type 3: By following procedures during the course of a challenge, an operator can operate redundant systems (or subsystems) and recover the systems to their normal operating envelope.

Type 4: By executing incorrect recovery plans, an operator can aggravate the situation or fail to terminate the challenge to the systems.

Type 5: By improvising, an operator can restore initially failed equipment to terminate a challenge.

As recommended by the SHARP, HRA should use the above classification and investigate the system to reveal possible human interactions. Analysts can use the above-mentioned characteristics for different types of activities. For example, Type 1 interactions generally involve components, whereas Type 3 and Type 4 interactions are mainly operating actions that can be considered at system level. Type 5 interactions are recovery actions that may affect both systems and components. Type 2 interactions can generally be avoided by confirming that human-induced errors are included as contributors to the probability of all possible challenges to the system. The output from this step can be used to revise and

enrich system reliability models, such as event trees and fault trees, to fully account for human interactions. This output will be used as the input to the next step.

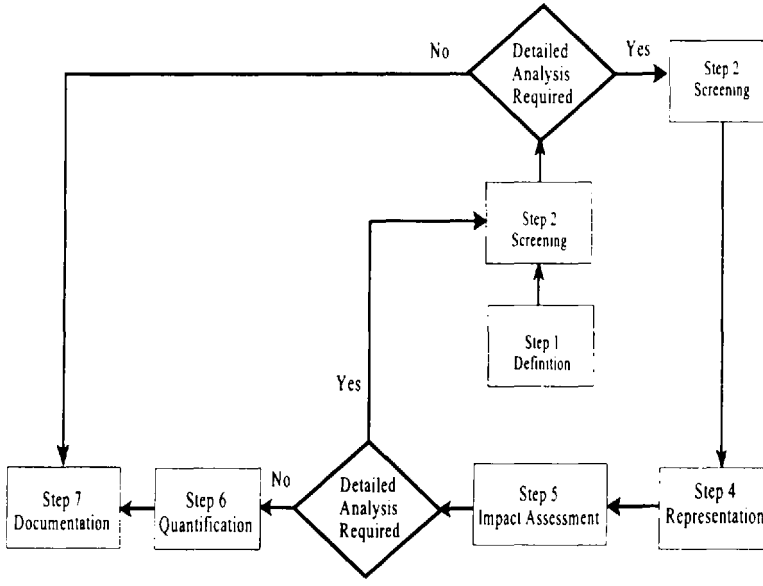


Figure 6.3 Systematic human action reliability procedure, Hannaman and Spurgin (1984).

Step 2: Screening

The objective of screening is to reduce the number of human interactions identified in Step 1 to those that might potentially challenge the safety of the system. This step provides the analysts with a chance to concentrate their efforts on key human interactions. This is generally done in a qualitative manner. The process is judgemental.

Step 3: Qualitative Analysis

To incorporate human errors into equipment failure modes, analysts need more information about each key human interaction identified in the previous steps to help in representing and quantifying these human actions. The two goals of qualitative analysis are:

1. Postulate what operators are likely to think and do, and what kind of actions they might take in a given situation, and
2. Postulate how an operator's performance may modify or trigger a challenge to the system.

This process of qualitative analysis may be broken down into four key stages.

1. Information gathering.
2. Prediction of operator performance and possible human error modes.
3. Validation of predictions.
4. Representation of output in a form appropriate for the required function.

In summary, the qualitative analysis step requires a thorough understanding of what performance-shaping factors (e.g., task characteristics, experience level, environmental stress, and social-technical factors) affect human performance. Based on this information, analysts can predict the range of plausible human action. The psychological model proposed by Rasmussen (1987) is a useful way of conceptualizing the nature of human cognitive activities. The full spectrum of possible human action following a misdiagnosis is typically very hard to recognize. Computer simulations of performance described by Woods et al. (1988) and Amendola et al. (1987) offer the potential to assist human reliability analysts in predicting the probability of human errors.

Step 4: Representation

To combine the HRA results with the system analysis models of Chapter 4, human error modes need to be transformed into appropriate representations. Representations are selected to indicate how human actions can affect the operation of a system.

Three basic representations have been used to delineate human interactions: the operator action tree (OAT) described by Wreathall (1981), the confusion matrix described by Potash et al. (1981), and the HRA event trees described by Swain and Guttman (1983). Figure 6.4 shows an example of OAT. The HRA tree is discussed in Section 6.3.2.

Step 5: Impact Assessment

Some human actions can introduce new impacts on the system response. This step provides an opportunity to evaluate the impact of the newly identified human actions on the system.

The human interactions represented in Step 4 are examined for their impact on challenges to the system, system reliability, and dependent failures. Screening techniques are applied to assess the importance of the impacts. Important human

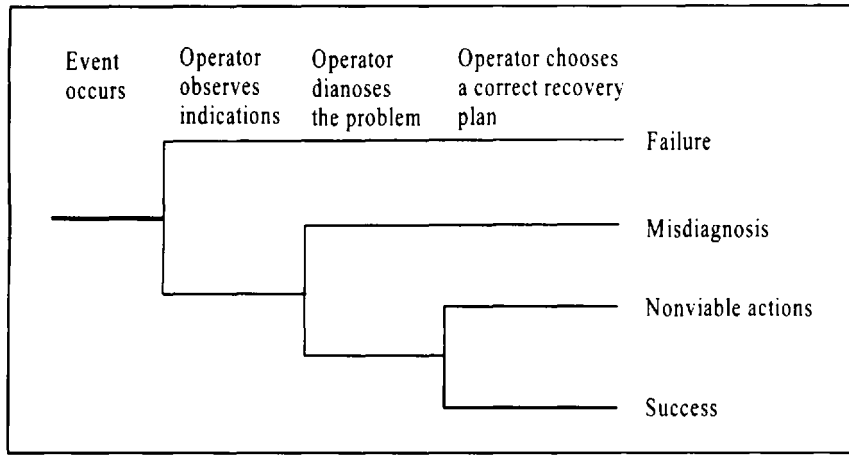


Figure 6.4 Operator action tree.

interactions are found, reviewed, and grouped into suitable categories. If the reexamination of human interactions identifies new human-induced challenges or behavior, the system analysis models (e.g., MLD, fault tree) are reconstructed to incorporate the results.

Step 6: Quantification

The purpose of this step is to assess the probabilities of success and failure for each human activity identified in the previous steps. In this step, analysts apply the most appropriate data or models to produce the final quantitative reliability analysis. Selection of the models should be based on the characteristics of each human interaction.

Guidance for choosing the appropriate data or models to be adopted is provided below.

For procedural tasks, the data from Swain and Guttman (1983) or equivalent can be applied.

For diagnostic tasks under time constraints, time-reliability curves from Hall et al. (1982) or the human cognitive reliability (HCR) model from Hannaman et al. (1984) can be used.

For situations where suitable data are not available, expert opinion approaches, such as paired comparison by Hunns and Daniels (1980) and the success likelihood index method by Embry et al. (1984) can be used.

For situations where multiple tasks are involved, the dependence rules

discussed by Swain and Guttman (1983) can be used to assess the quantitative impact.

Step 7: Documentation

The objective of Step 7 is to produce a traceable description of the process used to develop the quantitative assessments of human interactions. The assumptions, data sources, selected model and criteria for eliminating unimportant human interactions should be carefully documented. The human impact on the system should be stated clearly.

6.3.2 HRA Models

The HRA models can be classified into the following categories. Representative models in each are also summarized.

1. Simulation Methods
   a) Maintenance Personnel Performance Simulation (MAPPS)
   b) Cognitive Environment Simulation (CES)
2. Expert Judgement Methods
   a) Paired Comparison
   b) Direct Numerical Estimation (Absolute Probability Judgement)
   c) Success Likelihood Index Methodology (SLIM)
3. Analytical Methods
   a) Technique for Human Error Rate Prediction (THERP)
   b) Human Cognitive Reliability Correlation (HRC)
   c) Time Reliability Correlation (TRC)

We will briefly discuss each of these models. Human error is a complex subject. There is no single model that captures all important human errors and predicts their probabilities. Poucet (1988) reports the results of a comparison of the HRA models. He concludes that the methods could yield substantially different results, and presents their suggested use in different contexts.

Simulation Methods

These methods primarily rely on computer models that mimic human behavior under different conditions.

Maintenance Personnel Performance Simulation (MAPPS). MAPPS, developed by Siegel et al. (1984), is a computerized simulation model that provides human reliability estimation for testing and maintaining tasks. To perform the simulation, analysts must first find out the necessary tasks and substasks that individuals must perform. Environmental motivational tasks and

organizational variables that influence personnel performance reliability are input into the program. Using the Monte-Carlo simulation, the model can output the probability of success, time to completion, idle time, human load, and level of stress. The effects of a particular parameter or subtask performance can be investigated by changing the parameter and repeating the simulation.

The simulation output of task success is based on the difference between the ability of maintenance personnel and the difficulty of the subtask. The model used is

$$\Pr(\text{ success }) = \exp(y) / (1+ \exp(y)) \tag{6.13}$$

where $y > 0$ is the difference between personnel ability and task difficulty.

Cognitive Environment Simulation (CES). Woods (1988) has developed a model based on techniques from artificial intelligence (AI). The model is designed to simulate a limited resources problem solver in a dynamic, uncertain, and complex situation. The main focus is on the formation of intentions, situations and factors leading to intentional failures, forms of intentional failures, and the consequence of intentional failures.

Similar to the MAPPS model, the CES model is a simulation approach that mimics the human decision making process during an emergency condition. But CES is a deterministic approach, which means the program will always obtain the same results if the input is unchanged. The first step in CES is to identify the conditions leading to human intentional failures. CES provides numerous performance-adjusting factors to allow the analysts to test different working conditions. For example, analysts may change the number of people interacting with the system (e.g., the number of operators), the depth or breadth of working knowledge, or the human-machine interface. Human error prone points can be identified by running the CES for different conditions. The human failure probability is evaluated by knowing, a priori, the likelihood of occurrence of these error prone points.

In general, CES is not a human rate quantification model. It is primarily a tool to analyze the interaction between problem-solving resources and task demands.

Expert Judgement Methods

The primary reason for using expert judgement in HRA is that there often exist little or no relevant or useful human error data. Expert judgement is discussed in more detail in Section 6.4. There are two requirements for selecting experts:

they must have substantial expertise;
they must be able to accurately translate this expertise into probabilities.

Direct Numerical Estimation. For the direct numerical estimation method described by Stillwell et al. (1982), experts are asked to directly estimate the human error probabilities and the associated upper/lower bounds for each task. A consistency analysis might be taken to check for agreement among these judgements. Then, individual estimations are aggregated by either arithmetic or geometric average.

Paired Comparison. Paired comparison, described by Hunns and Daniels (1980), is a scaling technique based on the idea that judges are better at making simple comparative judgements than making absolute judgements. An interval scaling is used to indicate the relative likelihood of occurrence of each task. Saaty (1980) describes this general approach in the context of a decision analysis technique. The method is equally applicable to HRA.

Success Likelihood Index Methodology (SLIM). The success likelihood index methodology (SLIM) developed by Embry et al. (1984) is a structural method that uses expert opinion to estimate human error rates. The underlying assumption of SLIM is that the success likelihood of tasks for a given situation depends on the combination of effects from a small set of performance-shaping factors (PSFs) relevant to a group of tasks under consideration.

In this procedure, the experts are asked to assess the relative importance (weight) of each PSF with regard to its impact on the tasks of interest. An independent assessment is made to the level or the value of the PSFs in each task situation. After identifying and agreeing on the small set of PSFs, respective weights and ratings for each PSF are multiplied. These products are then summed to produce the success likelihood index (SLI), varying from 0 to 100 after normalization. This value indicates the expert's belief regarding the positive or negative effects of PSFs on task success.

The SLIM approach assumes that the functional relationship between success probability and SLI is exponential, i.e.,

$$\log [\Pr(\text{Operator success})] = (\text{SLI}) + b \qquad (6.14)$$

where a and b are empirically estimated constants. To calibrate a and b, at least two human tasks of known reliability must be used in (6.14), from which constants a and b are calculated.

This technique has been implemented as an interactive computer program. The first module, called multi-attribute utility decomposition (MAUD), analyzes a set of tasks to define their relative likelihood of success given the influence of PSFs. The second module, systematic approach to the reliability assessment of humans (SARAH), is then used to calibrate these relative success likelihoods to generate absolute human error probability. The SLIM technique has a good theoretical basis in decision theory. Once the initial database has been established with the SARAH module, evaluations can be performed rapidly. This method does

not require extensive decomposition of a task to an elemental level. For situations where no data are available, this approach enables HRA analysts to reasonably estimate human reliability. However, this method makes extensive use of expert judgement, which requires a team of experts to participate in the evaluation process. The resources required to set up the SLIM-MAUD database are generally greater than other techniques.

Analytical Methods

These methods generally use a model based on some key parameters that form the value of human reliabilities.

Technique for Human Error Rate Prediction (THERP). The oldest and most widely used HRA technique is the THERP analysis developed by Swain and Guttman (1983) and reported in the form of a handbook. The THERP approach uses conventional system reliability analysis modified to account for possible human error. Instead of generating equipment system states, THERP produces possible human task activities and the corresponding human error probabilities. THERP is carried out in the five steps described below.

1. Define system failures of interest

   From the information collected by examining system operation and analyzing system safety, analysts identify possible human interaction points and task characteristics, and their impact on the systems. Then, screening is performed to determine critical actions that require detailed analysis.

2. List and analyze related human actions

   The next step is to develop a detailed task analysis and human error analysis. The task analysis delineates the necessary task steps and the required human performance. The analyst then determines the errors that could possibly occur. The following human error categories are defined by THERP:
   Errors of omission (omit a step or the entire task).
   Errors of commission, including:
   Selection error (select the wrong control, choose the wrong procedures);
   Sequence error (actions carried out in the wrong order);
   Time error (actions carried out too early/too late);
   Qualitative error (action is done too little/too much).

At this stage, opportunities for human recovery actions (recovery from an abnormal event or failure) should be identified. Without considering recovery possibilities, overall human reliability might be dramatically underestimated.

The basic tool used to model tasks and task sequences is the HRA event tree. According to the time sequence or procedure order, the tree is built to represent possible alternative human actions. Therefore, if appropriate error probabilities of each subtask are known and the tree adequately depicts all human action sequences, the overall reliability of this task can be calculated. An example of an HRA event tree is shown in Figure 6.5.

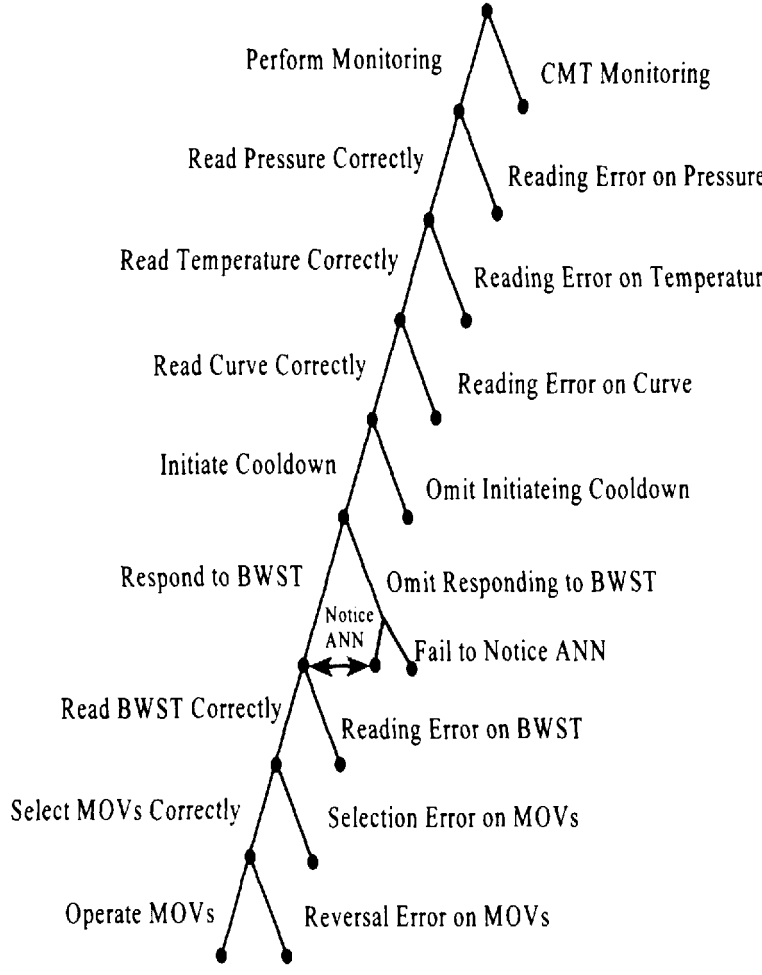


Figure 6.5 HRA event tree on operator actions during a small-break loss of coolant in nuclear plants. CMT, computer monitoring; ANN, annunciator; BWST, borated water storage tank; MOV, motor operated valve. (Hannaman and Spurgin (1984)).

3. Estimate relevant error probabilities

   As explained in the previous section, human error probabilities (HEPs) are required for the failure branches in the HRA event tree. Chapter 20 of Swain and Guttman (1983) provides the following information.

   Data tables containing nominal human error probabilities
   Performance models explaining how to account for PSFs to modify the nominal error data
   A simple model for converting independent failure probabilities into conditional failure probabilities

   In addition to the data source of THERP, analysts may use other data sources, such as the data from recorded incidents, trials from simulations, and subjective judgement data, if necessary.

4. Estimate effects of error on system failure events

   In the system reliability framework, the human error tasks are incorporated into the system model, such as a fault tree. Hence, the probabilities of undesired events can be evaluated and the contribution of human errors to system reliability or availability can be estimated.

5. Recommend changes to system design and recalculate system reliability

   A sensitivity analysis can be performed to identify dominant contributors to system unreliability. System performance can then be improved by reducing the sources of human error or redesigning the safeguard systems.

   THERP approach is very similar to the equipment reliability methods described in Chapter 4.The integration of human reliability analysis and equipment reliability analysis is straightforward using the THERP process. Therefore, it is easily understood by system analysts. Compared with the data for other models, the data for THERP are much more complete and easier to use. The handbook contains guidance for modifying the listed data for different environments. The dependencies among subtasks are formally modeled, although subjective. Conditional probabilities are used to account for this kind of task dependence.

   Very detailed THERP analysis can require a large amount of effort. In practice, by reducing the details of the THERP analysis to an appropriate level, the amount of work can be minimized. THERP is not appropriate for evaluating errors involving high-level decisions or

diagnostic tasks. In addition, THERP does not model underlying psychological causes of errors. Since it is not an ergonomic tool, this method cannot produce explicit recommendations for design improvement.

Human Cognitive Reliability (HCR) Correlation. During the development of SHARP, a need was identified to find a model to quantify the reliability of control room personnel responses to abnormal system operations. The HCR correlation, described by Hannaman et al. (1984), is essentially a normalized time-reliability correlation (described below) whose shape is determined by the available time, stress, human-machine interface, etc. Normalization is needed to reduce the number of curves required for a variety of situations. It was found that a set of three curves (skill-, rule-, and knowledge-based, developed by Rasmussen,1982) could represent all kinds of human decision behaviors. The application of HCR is straightforward. The HCR correlation curves can be developed for different situations from the results of simulator experiments. Therefore, the validity can be verified continuously. This approach also has the capability of accounting for cognitive and environmental PSFs.

Some of the disadvantages of the HCR correlation are:

The applicability of the HCR to all kinds of human activities is not verified.

The relationships of PSFs and nonresponse probabilities are not well addressed.

This approach does not explicitly address the details of human thinking processes. Thus, information about intentional failures cannot be obtained.

Time-Reliability Correlation (TRC). Hall et al. (1982) concentrated on the diagnosis and decision errors of nuclear power plant operators after the initiation of an accident. They criticized the behavioral approach used by THERP and suggested that a more holistic approach be taken to analyze decision errors.

The major assumption of TRC is that the time available for diagnosis of a system fault is the dominant factor in determining the probability of failure. In other words, the longer people take to think, the more unlikely they are to make mistakes. The available time for decision and diagnosis is delimited by the operator's first awareness of an abnormal situation and the initiation of the selected response. Because no data were available when the TRC was developed, an interim relationship was obtained by consulting psychologists and system analysts. Recent reports confirm that the available time is an important factor in correctly performing cognitive tasks. A typical TRC is shown in Figure 6.6.

Dougherty and Fragola (1988) is a good reference for TRC as well as other HRA methods. TRC is very easy and fast to use. However, TRC is still a premature approach. The exact relationship between time and reliability requires

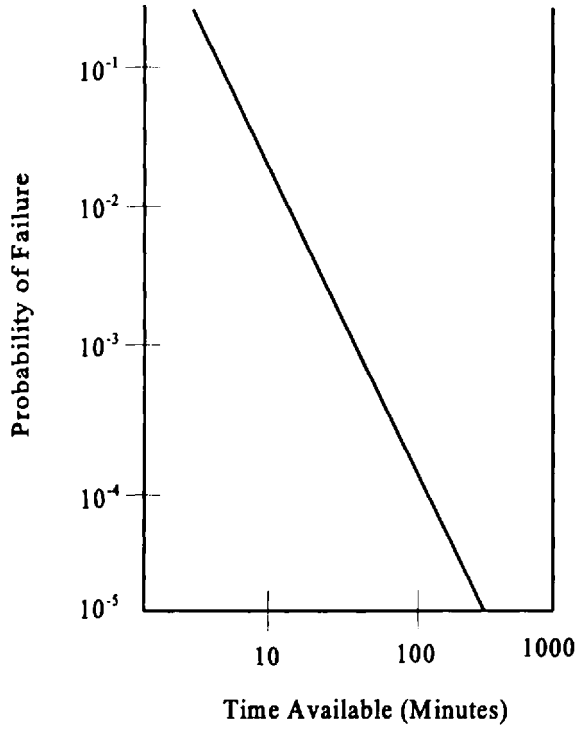


Figure 6.6 Time-reliability correlation for operators.

more experimental and actual observations. This approach overlooks other important PSFs, such as experience level, task complexity, etc. TRC focuses only on limited aspects of human performance in emergency conditions. The time available is the only variable in this model. Therefore, the estimation of the effect of this factor should be very accurate. However, TRC does not provide guidelines or information on how to reduce human error contributions.

6.3.3 Human Reliability Data

There is general agreement that a major problem for HRA is the scarcity of data on human performance that can be used to estimate human error rates and performance time. To estimate human error probabilities, one needs data on the relative frequency of the number of errors and/or the ratio of "near-misses" to total

number of attempts. Ideally, this information can be obtained from observing a large number of tasks performed in a given application. However, this is impractical for several reasons. First, error probabilities for many tasks, especially for rare emergency conditions, are very small. Therefore, it is very difficult to observe enough data within a reasonable amount of time to get statistically meaningful results. Second, possible penalties assessed against people who make errors in e.g., a nuclear power plant or in aircraft cockpit, discourages free reporting of all errors. Third, the costs of collecting and analyzing data could be unacceptably high. Moreover, estimation of performance times presents difficulties since data taken from different situations might not be applicable.

Data can be used to support HRA quantification in a variety of ways, e.g., to confirm expert judgement, develop human reliability data, or support development of an HRA model. Currently, available data sources can be divided into the following categories: (1) actual data, (2) simulator data, (3) interpretive information, and (4) expert judgement.

Psychological scaling techniques, such as paired comparisons, direct estimation, SLIM, and other structured expert judgement methods, are typically used to extrapolate error probabilities. In many instances, scarcity of relevant hard data makes expert judgement a very useful data source. This topic is discussed further in Chapter 7.

6.4 MEASURES OF IMPORTANCE

During the design reliability analysis, or risk assessment of a system, the specific components and their arrangement may render some to be more critical than others from the standpoint of their impact on the system reliability. For example, a series set of components within a system has a much higher importance (for failure) in a system, than the same set of components would have, if they were in parallel within the system. In this section, we describe five methods of measuring the importance of components: Birnbaum, Criticality, Fussell–Vesely, Risk-Reduction Worth, and Risk-Achievement Worth measures of importance. Usually, importance measures are used in the failure space, however, in the book their application in the success space has also been discussed.

6.4.1 Birnbaum Measure of Importance

Introduced by Birnbaum (1969), this measure of component importance, $I_i^{\mathrm{B}}(t)$, for success space (as described by Sharirli (1985)) is defined as

$$I_i^{\mathrm{B}}(t) = \frac{\partial R_s[R(t)]}{\partial R_i(t)} \tag{6.15}$$

where $R_s[R(t)]$ is reliability of the system as a function of the reliability of individual components, $R_i(t)$.

If, for a given component i, $I_i^B(t)$ is large, it means that a small change in the reliability of component i, $R_i(t)$, will result in a large change in the system reliability $R_s(t)$.

If system components are assumed to be independent, the Birnbaum measure of importance can be represented by (Hoyland and Rausand (1994)):

$$I_i^B(t) = R_s\left[R(t) \mid R_i(t) = 1\right] - R_s\left[R(t) \mid R_i(t) = 0\right] \qquad (6.16)$$

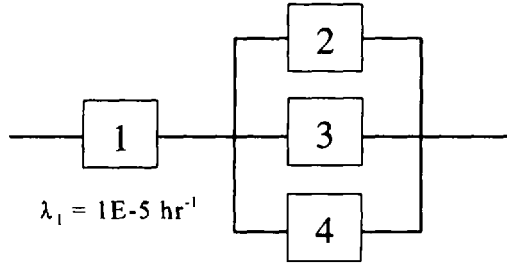
where $R_s[R(t)|R_i(t) = 1]$ and $R_s[R(t)|R_i(t) = 0]$ are the values of reliability function of the system with the reliability of component i set to 1 and 0, respectively.

Equation (6.15) and (6.16) are often used in conjunction with the unreliability, unavailability or risk function, $F_s[Q_i(t)]$, given in terms of individual component unreliability or unavailability $Q_i(t)$. In this case, (6.16) is replaced by

$$I_i^B = \frac{\partial F_s[Q(t)]}{\partial Q_i(t)} = F_s\left[Q(t)|Q_i(t)=1\right] - F_s\left[Q(t)|Q_i(t)=0\right] \qquad (6.17)$$

Example 6.5

Consider the system shown below. Determine the Birnbaum importance of each component at $t = 720$ hours. Assume an exponential time to failure.



$$\lambda_2 = \lambda_3 = \lambda_4 = 1\text{E-}4 \text{ hr}^{-1}$$

Solution:

$$R_1(t = 720) = 0.993 \; R_2(t = 720) = R_3(t = 720) = R_4(t = 720) = 0.487$$

The reliability function of the system is

$$R_s(R(t)) = R_1(t) \cdot \{1 - [1 - R_2(t)][1 - R_3(t)][1 - R_4(t)]\} = 0.859$$

Using (6.16),

$$I_1^B(t) = R_s[R(t)|R_1(t) = 1] - R_s[R(t)|R_1(t) = 0]$$
$$I_1^B(t) = \{1 - [1 - R_2(t)]\}[1 - R_3(t)][1 - R_4(t)]$$

therefore

$$I_1^B(t = 720) \approx 0.865$$
$$I_2^B(t) = R_1(t)[(1 - R_3(t))(1 - R_4(t))]$$
$$I_2^B(t = 720) \approx 0.26$$

Similarly,

$$I_3^B(t = 720) = I_4^B(t = 720) \approx 0.26$$

It can be concluded that the rate of improvement in component 1 has far more importance (impact) on system reliability than components 2, 3, and 4. For example, if the reliability of the parallel units increases by an order of magnitude, clearly the importance of components 2, 3, and 4 reduces (e.g., for $\lambda_2 = \lambda_3 = \lambda_4 = 10^{-4}$ / hr, $I_2^B = I_3^B = I_4^B \approx 0$, and $I_1^B = 1$). Similarly, if identical units are in parallel with component 1, the importance changes.

6.4.2 Criticality Importance

Birnbaum's importance for component i is independent of the reliability of component i itself. Therefore, I_i^B is not a function of $R_i(t)$. It is clear that it would be more difficult and costly to further improve the more reliable components than to improve the less reliable ones. From this, the criticality importance of component i is defined as

$$I_i^{CR}(t) = \frac{\partial R_s[(t)]}{\partial R_i(t)} \times \frac{R_i(t)}{R_s[R(t)]}$$

or

$$I_i^{CR}(t) = I_i^B \times \frac{R_i(t)}{R_s[R(t)]} \tag{6.18}$$

From (6.18), it is clear that the Birnbaum importance is corrected with respect to reliability of the individual components relative to the reliability of the whole system. Therefore, if the Birnbaum importance of a component is high, but the reliability of the component is low with respect to the reliability of the system, then criticality importance assigns a low importance to this component. Similarly, (6.18) can be represented by the unreliability or unavailability function

$$I_i^{CB}(t) = I_i^{B} \times \frac{Q_i(t)}{F_s[Q(t)]} \tag{6.19}$$

As such in Example 6.4,

$$I_1^{CB} = 0.865 \times \frac{0.993}{0.895} = 1$$

Since component 1 is more reliable, its contribution to reliability of the system (i.e., its criticality importance) increases. Whereas, components 2, 3, and 4 will have a less important contribution to the overall system reliability.

A subset of criticality importance measure is inspection importance measure (I_i^{W}). This measure is defined as the product of Birnbaum importance times the failure probability (unreliability or unavailability) of the component. Accordingly,

$$I(t)^{W} = I(t)^{B} \times Q_i(t) \tag{6.20}$$

This measure is used to prioritize operability test activities to ensure high component readiness and performance.

6.4.3 Fussell–Vesely Importance

In cases where component i contributes to system reliability, but is not necessarily critical, the Fussell–Vesely importance measure can be used. This measure, is introduced by W.E. Vesely and later applied by Fussell (1975), is in the form of

$$I_i^{FV}(t) = \frac{R_i[R(t)]}{R_s[R(t)]} \tag{6.21}$$

where $R_i[R(t)]$ is the contribution of component i to the reliability of the system. Similarly, using unreliability or unavailability functions,

$$I_i^{FV}(t) = \frac{F_i[Q(t)]}{F_s[Q(t)]} \tag{6.22}$$

where $F_i[Q(t)]$ denotes the probability that component i is contributing to system failure or system risk.

The Fussell–Vesely importance measure has been applied to system cut sets to determine the importance of individual cut sets to the failure probability of the whole system. For example, consider importance I_k of the kth cut set representing a system failure. In that case, (6.22) replaces

$$I_k^{FV}(t) = \frac{Q_k(t)}{Q_s(t)} \tag{6.23}$$

where

$Q_k(t)$ is the time dependent probability that minimal cut set k occurs, and

$Q_s(t)$ is the total time dependent probability that the system fails (due to all cut sets).

Generally, the minimal cut sets with the largest values of I_k are the most important ones. Equation (6.23) is equally applicable to mutually exclusive cut sets. Consequently, system improvements should initially be directed toward the minimal cut sets with the largest importance values.

If the probability of all minimal cut sets or mutually exclusive cut sets are known, then the following approximate expression can be used to find the importance of individual components.

$$I_i^{FV} \approx \frac{\sum_{j=1}^{m} Q_j(t)}{Q_s(t)} \tag{6.24}$$

where

$Q_j(t)$ is the probability that the jth cut set which contains component i is failed, and m is the number of minimal cut sets that contain component i

Expression (6.24) is an approximation; the situation of two minimal cut sets containing component i failing at the same time is neglected since its probability is very small.

6.4.4 Risk Reduction Worth Importance

The risk reduction worth (RRW) importance is a measure of the change in unreliability (unavailability, or risk) when an input variable (e.g., unavailability of component) is set to zero. That is by assuming that a component is "perfect" (or its failure probability is zero) and thus eliminating any postulated failure. This importance measure shows how much better the system can become as its components are improved.

This importance measure is used in failure domains although it can be equally used in the success domain too. The calculation may be done either as a ratio or as a difference. Accordingly, as a ratio

$$I_i^{RRW} = \frac{F_s[Q(t)]}{F_s[Q(t) \mid Q_i(t) = 0]} \tag{6.25}$$

and as a difference,

$$I_i^{RRW} = F_s[Q(t)] - F_s[Q(t) \mid Q_i(t) = 0] \tag{6.26}$$

where $F_s[Q(t) \mid Q_i(t) = 0]$ is the system unreliability (unavailability or risk) when unreliability (or unavailability) of component i is set to zero.

In practice, this measure is used to identify elements of the system (e.g., components) that are the best candidates for efforts leading to improving system reliability (risk or unavailability).

6.4.5 Risk Achievement Worth Importance

The risk achievement (increase) worth (RAW) importance is the inverse of risk reduction worth measure. The input variable (e.g., component unavailability) is set to one, and the effort of this change on system unreliability (unavailability or risk) is measured. Similar to risk reduction worth, the calculation may be done as a ratio or a difference. By setting component failure probability to one, RAW measures the increase in system failure probability assuming the worst case of failing the component. As a ratio RAW measure is

$$I_i^{RAW} = \frac{F_s[Q(t)]}{F_s[Q(t) \mid Q_i(t) = 1]} \tag{6.27}$$

and as a difference,

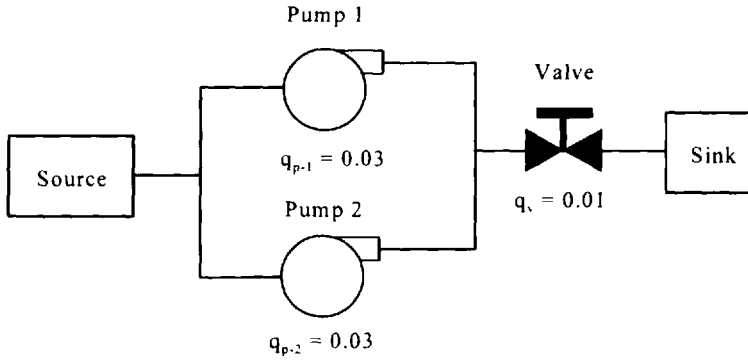
$$I_i^{RAW} = F_s\left[Q(t) \mid Q_i(t) = 1\right] - F_s[Q(t)] \tag{6.28}$$

where, $F_s[Q(t) \mid Q_i(t) = 1]$ is the system unreliability (unavailability, or risk), when unreliability (or unavailability) of component i is set to one.

The risk increase measure is useful for identifying elements of the system, which are the most crucial for making the system unreliable (unavailable or increasing the risk). Therefore, components with high I^{RAW} are the ones that will have the most impact, should their failure probability unexpectedly rise.

Example 6.6

Consider the water-pumping system below. Determine the Birnbaum, Criticality, and Fussell-Vesely importance measures of the valve (v), pump-1 ($p - 1$) and pump-2 ($p - 2$) using both reliability and unreliability versions of the importance measures.



Solution:

Because the component reliability, $R_{P-1} = R_{p-2} = 0.97$, $R_v = 0.99$.

The reliability function is

$$R_s[R(t)] = R_v \cdot [R_{p-1} + R_{p-2} - R_{p-1} \times R_{p-2}] = 0.989$$

Using the rare event approximation, the unreliability function is

$$F_s[Q(t)] = Q_{p-1} \times Q_{p-2} + Q_v = 0.011$$

1. Birnbaum's importance:

$$I_v^B = R_{p-1} + R_{p-2} - R_{p-1} \times R_{p-2} \approx 1$$

$$I_{p-1}^B = R_v - R_v \times R_{p-2} \approx 0.03$$

$$I_{p-2}^B = R_v - R_v \times R_{p-1} \approx 0.03$$

Using the unreliability function,

$$I_v^B \approx 1$$

$$I_{p-1}^B = Q_{p-2} \approx 0.03$$

$$I_{p-2}^B = Q_{p-1} \approx 0.03$$

2. Criticality Importance:

$$I_v^{CR} = 1 \times \frac{0.99}{0.989} \approx 1$$

$$I_{p-1}^{CR} = I_{p-2}^{CR} = 0.03 \times \frac{0.97}{0.989} \approx 0.029$$

Same criticality importance values are expected for the unreliability function. $R_v[R(t)]$ is obtained by retaining terms involving $R_v(t)$.

3. Fussell–Vesely Importance:

$$R_v[r(t)] = R_v[R(t)] \approx 0.989$$

$$R_{p-1}[r(t)] = R_v \times R_{p-1} - R_v \times R_{p-1} \times R_{p-2} \approx 0.029$$

$$R_{p-2}[R(t)] = R_v \times R_{p-2} - R_v \times R_{p-1} \times R_{p-2} \approx 0.029$$

$$I_v^{FV} = 1 \times \frac{0.989}{0.989} = 1$$

$$I_{p-1}^{FV} = I_{p-2}^{FV} = \frac{0.029}{0.989} \approx 0.029$$

Using the unreliability function,

$$F_v[R(t)] = Q_v = 0.01$$

$$F_{p-1}[Q(t)] = Q_{p-1} \times Q_{p-2} = 0.0001$$

$$F_{p-2}[Q(t)] = Q_{p-2} \times Q_{p-1} = 0.0009$$

Then,

$$I_v^{FV} = 1 \times \frac{0.01}{0.011} = 0.9$$

$$I_{p-1}^{FV} = I_{p-2}^{FV} = \frac{0.0009}{0.011} \approx 0.08$$

Example 6.7

Repeat Example 6.6 and calculate I^{RRW} and I^{RRW} for all components. Compare the results with I^{B}, I^{CR}, and I^{FV}.

Solution:

The unreliability function is $F_v[q(t)] = Q_{p\_1} \times Q_{p\_2} + Q_v = 0.011$

1. For RRW,

$$F_v\left[Q \mid Q_v = 0\right] = Q_{p\_1} \times Q_{p\_2} = 0.03 \times 0.03 = 0.0009$$

Therefore, for ratio measure,

$$I_v^{RRW} = \frac{0.011}{0.0009} = 12.2$$

For difference measure

$$I_v^{RRW} = 0.011 - 0.0009 = 0.01$$

Similarly for pumps as ratio,

$$I_{p\_1}^{RRW} = I_{p\_2}^{RRW} = \frac{0.011}{0.01} = 1.1$$

As difference,

$$I_{p\_1}^{RRW} = I_{p\_2}^{RRW} = 0.011 - 0.01 = 0.001$$

Note that in the ratio method, the larger numbers indicate increasing importance, whereas the reverse is true for the difference method. This is only a metric for identifying a component when its assured performance will highly affect system operation.

2. Similarly for RAW, the ratio method yields

$$I_v^{RAW} = \frac{1}{0.011} = 90.91$$

For the difference method,

$$I_v^{RAW} \approx 1$$

For the pumps, using the ratio method,

$$I_{p\_1}^{RAW} = I_{p\_2}^{RAW} = \frac{1 \times 0.03 + 0.01}{0.011} = 3.64$$

For the difference method,

$$I_{p\ 1}^{RAW} = I_{p\ 2}^{RAW} = (1 \times 0.03 + 0.01) - 0.011 = 0.029$$

The I_i^{RAW} shows importance of component i with respect to system unreliability when component i fails. Clearly by comparing the results to I^B, I^{CR}, and I^{FV} with I^{RAW} and I^{RRW}, the relative importance value measured by I_i^{RAW} is consistent. This is expected since all other measures are related to the degradation of the component. I^{RAW} is related to worth of improvement in component reliability.

6.4.6 Practical Aspects of Importance Measures

There are two principal factors that determine the importance of a component in a system: the structure (topology) of the system, and the reliability or unreliability of the components. Depending on the measure selected, one of the above may be pertinent. Also, depending on whether we use reliability or unreliability, some of these measures behave differently. In Example 6.6, this is seen in $I_{p\cdot1}^{FV}$ and $P_{p\cdot2}^{FV}$, where their importance in success space is almost 1 and in the failure space is 0.

The Birnbaum measure of importance completely depends on the structure of the system (e.g., whether the system is dominated by a parallel or series configuration). Therefore, it should only be used to determine the degree of redundancy and appropriateness of the system's logic.

The criticality importance is related to that of Birnbaum's. However, it is also affected by the reliability/unreliability of the components and the system. This measure allows for the evaluation of the importance of a component in light of its potential to improve system reliability. The effect of improvements on one component may result in changes in the importance of other components of the system.

The Fussell–Vesely measure of importance has been widely used in practice, mostly for measuring importance in the failure space using unreliability/unavailability functions. The measure is more influenced by the actual reliability/unreliability of the components and the systems as well as the logical structure of the system. Because of its simplicity, this measure has been widely used.

Generally, the importance of components should be used during design or evaluation of systems to determine which components or subsystems are important to the overall reliability of the system. Those with high importance could prove to be candidates for further improvement. In an operational context, items with high importance should be watched by the operators, since they are critical for the continuous operation of the system.

Some importance measures are calculated as dimensionless ratios, while others are absolute physical quantities or probabilities. The Birnbaum measure is an absolute measure, while the Fussell–Vesely is a relative one. Table 6.2 summarizes the importance measures discussed in this section.

It is widely felt that the relative measures (ratios) have the advantage of being more robust than the absolute measure: since many quantities appear in both the numerator and the denominator, it can be hoped that errors in their magnitudes will tend to divide out. This hope is realized in some models. On the other hand, either the denominator or the numerator in the relative measures may be dominated by terms that have nothing to do with the basic event of interest, so that errors or uncertainties in those terms may obscure the desired insights. It is felt that the risk achievement ratio and the risk reduction ratio are especially vulnerable to this kind of distortion.

The *absolute* measures have the advantage of providing an immediate sense of whether a given event is negligible on an absolute scale. The relative measures do not provide this information; the user must obtain system failure probability and perform some arithmetic in order to obtain this information.

A number of other measures of importance have been introduced, as well as computer program importance calculations. For more information, the readers are referred to Lambert (1975), Sharirli (1985), and NUREG/CR-4550 (1990).

6.5 RELIABILITY-CENTERED MAINTENANCE

6.5.1 History and Current Procedures

The reliability-centered maintenance (RCM) methodology is a systematic approach directed towards defining and developing applicable and effective failure management strategies. RCM finds its roots in the early 1960s. The initial development work was done by the North American civil aviation industry. It started when the airlines at that time noted that many of their maintenance practices were not only too expensive but also unsafe. This prompted the industry to put together a series of maintenance steering groups (MSG) to review everything they were doing to keep their aircrafts airborne. These groups consisted of representatives of the aircraft manufacturers, the airlines and the Federal Aviation Administration (FAA).

The first attempt at a rational, zero-based process for formulating maintenance strategies was promulgated by the air Transport Association in Washington, DC in 1968. The first attempt is now known as MSG-1. A refinement—now known as MSG-2—was promulgated in 1970.

In the mid-1970s, the U.S. Department of Defense became interested in the then state-of-the-art in aviation maintenance. They commissioned a report on the subject from the commercial aviation industry. This report written by Stanley

Nowlan and Howard Heap of United Airlines, was entitled "Reliability Centered Maintenance (RCM)." The report was published in 1978, and it is still the leading document in physical asset management.

RCM is a process used to decide what must be done to ensure that any item (e.g., system or process) continues to do its function.

Table 6.2 Interpretation of Importance Measures

Name	Definition	Interpretation	Comments
Birnbaum	Pr(coefficient of component *i*)	How often component *i* is needed to prevent system failure	Absolute measure; directly measures sensitivity of probability of system failure (or risk) to probability of component *i* failure
Criticality	Pr(coefficient of component *i*) × Pr(component *i* failure)/ Pr(system failure)	How often component *i* is needed to prevent system failure adjusted for relative probability of component i failure	Absolute measure, measures the sensitivity of system failure probability with respect to failure probability of component *i*
Fussell–Vesely	Pr[system failure (or risk) based on terms involving component *i*]/ Pr[total system failure (or risk)]	Fraction of system unavailability (or risk) involving failure of component *i*	Dimensionless, relative measure; reflects how much relative improvement is theoretically available from improving performance of component *i*. Denominator may contain some terms having nothing to do with component *i* operation
Risk reduction worth	Pr[system failure (or risk)]/ Pr[system failure given component *i* operates]	Shows relative improvements in Pr(system failure) realizable by improving component *i*; how much relative harm component *i* does, by not being perfect	Dimensionless, relative measure. Both the numerator and the denominator contain some terms having nothing to do with component *i* operation
Risk achievement worth	Pr[system failure given component *i* fails]/Pr[total system failure (or risk)]	How much relative good is done by component *i*; factor by which Pr[system failure] would increase with no credit for component *i*	Dimensionless, relative measure; both the numerator and the denominator contain some terms having nothing to do with component *i* operation

What users expect from their items is defined in terms of primary performance parameters such as output, throughput, speed, range, and carrying capacity. Where relevant, the RCM process also defines what users want in terms of risk (safety and environmental integrity), quality (precision, accuracy, consistency, and stability), control, comfort, containment, economy, customer service, and so on.

The next step in the RCM process is to identify ways in which the item can fail to live up to these expectations (failed states), followed by an FMEA, to identify all the events which are reasonably likely to cause each failed state.

Finally, the RCM process seeks to identify a suitable failure management policy for dealing with each failure mode in light of its consequences and technical characteristics. Failure management policy options include:

Predictive maintenance.
Preventive maintenance.
Failure-finding.
Change the design or configuration of the system.
Change the way the system is operated.
Run-to-failure.

The RCM process provides powerful rules for deciding whether any failure management policy is technically appropriate. It also provides adequate criteria for deciding how often routing tasks should be done. The RCM methodology involves a systematic and logical step-by-step consideration of:

1. The function(s) of a system or component.
2. The ways, the function(s) can be lost.
3. The importance of the function and its failure.
4. A priority-based consideration that identifies those failure management activities that both reduce failure potential and are cost-effective.

The key steps of this process include:

Definition of system boundaries. Boundaries must be clearly identified and clear explanation of the level of detail for the analysis be presented.

Determination of the functions of a system, its subsystems, or components. Each component within the system or subsystem may have one or more functions. These should be explained and inputs and outputs of functions across system boundaries must also be identified.

Determination of functional failures. A functional failure occurs when a system or subsystem fails to provide its required function.

Determination of dominant failure modes. One of the logical system analysis methods (e.g., fault tree or MLD) along with FMEA should be used to identify the modes that are the leading (high probability) causes for functional failures.

Determination of corrective actions and *optimal preventive maintenance schedules.* Applicable and effective course of action for each failure mode should be identified. This action may be to implement a preventive maintenance task, accept the likelihood of failure, or initiate redesign.

Integration of the results. The results of the failure management task along with other specifics of implementation are integrated into the maintenance plan.

From the above steps, it is clear that RCM methodology can be divided into two basic phases. First, the system and its boundaries are defined and then the system is decomposed to subsystems and components, and their functions are identified along with those failures that are likely to cause loss of the functions. Second, each of the functional failures is examined to determine the associated failure mode and to determine whether or not there are effective failure management strategies (or tasks) that eliminate or minimize occurrence of the failure mode identified.

For those failure modes for which an effective failure management task is specified, further definition is necessary. Each task should be labeled as either time-directed, condition-monitoring, or failure-finding. Time-directed tasks are generally applicable when the probability of failure increases with the time, that is the failure mode has a positive trend as discussed in Chapter 5. Time can be measured in several different ways, including actual run time or the number of startups (demands) or shutdowns of the component (with the given failure mode). Condition-monitoring tasks are generally applicable when one can efficiently correlate functional failures to detectable and measurable parameters of the system. For example, vibration of a pump can be measured to predict alignment problems. Failure-finding tasks are not preventive, but are intended to discover failures that are otherwise hidden. If no effective failure management task can be identified for a hidden failure, a scheduled functional failure-finding task may be devised.

In order to develop an optimal preventive maintenance program, an optimal schedule for such maintenance activities must be devised. In the following section, a reliability based technique for optimizing a preventive maintenance schedule is discussed.

6.5.2 Optimal Preventive Maintenance Scheduling

In this section we consider a simple example of optimal preventive action scheduling, which minimizes the average total cost of system functioning per unit time.

Denote the preventive maintenance interval by θ, the cost of failure which occurs during a system operation by c_1, and the cost of a preventive maintenance by c_2. Consider the problem of finding the optimal value of θ which minimizes the average total cost per unit time.

In order to find the mean length of the interval between two adjacent maintenance actions and the average cost of this interval per unit time assume that:

The interval between maintenance actions is constant and equals θ, if there is no failure, and all maintenance actions are preventive.

If a failure has occurred, it is assumed that a maintenance is instantly performed at a random time $\tau < \theta$. The mean length of this interval, Θ, is given by:

$$\Theta = E(\tau) = \int_0^{\theta} R(t)\,dt \tag{6.29}$$

where $R(t)$ is the reliability function of the component.
The average cost per unit time, $C(\theta)$, can be written as

$$C(\theta) = \frac{c_1 F(\theta) + c_2 R(\theta)}{\Theta} \tag{6.30}$$

where $F(Q) = 1 - R(Q)$ is the unreliability of the component.

The optimal value of θ can be fo
und using the first order condition (equating the first derivative to zero), i.e.,

$$\frac{dC(\theta)}{d\theta} = \frac{d}{d\theta}\left(\frac{c_1 F(\theta) + c_2 R(\theta)}{\Theta}\right) = 0 \tag{6.31}$$

which results in the following equation:

$$\lambda(\theta)\Theta - F(\theta) = \frac{c_2}{c_1 - c_2} \tag{6.32}$$

where $\lambda(\theta) = f(\theta)/R(\theta)$. For the practical applications it is better to rewrite this equation expressing the relative cost dependence:

$$\lambda(\theta)\Theta - F(\theta) = \frac{c_2/c_1}{1 - c_2/c_1} \tag{6.33}$$

In general, this equation can be solved only numerically.

The results of the numerical solution of (6.33) for the Weibull distribution with the shape parameter, β, equal to 2 (aging distribution), and the scale parameter, α, and some values of c_2/c_1 are given in the following table.

c_2/c_1	θ^*/α
0.9	5.078
0.8	2.274
0.7	1.529
0.6	1.243
0.5	1.080
0.4	0.968
0.3	0.885
0.2	0.816
0.1	0.759
0.05	0.733
0.025	0.720
0.01	0.713

Economic Benefit of Optimization

To estimate the economic benefit of the optimization considered, one needs to compare the average cost of failure per unit time without preventive maintenance, $E(c_1) = c_1/\text{MTTF}$, with the average cost of failure per unit time given by (6.30) when $\theta = \theta^*$ (i.e., under the optimal schedule) i.e., to calculate the ratio, ϵ:

$$\epsilon = \frac{\text{MTTF}}{\Theta^*}\left(F(\theta^*) + \frac{c_2}{c_1}R(\theta^*) \right) \tag{6.34}$$

The results of calculations for the example considered are given in the following table:

c_2/c_1	ϵ
0.9	1.000
0.8	1.000
0.7	1.000
0.6	0.993
0.5	0.966
0.4	0.922
0.3	0.862
0.2	0.783
0.1	0.690
0.05	0.644
0.025	0.605
0.01	0.589

As one could anticipate the less c_2/c_1 is, the greater the effect of optimization. For the values of $c_2/c_1 \approx 0.1$ it is about 30 cents per dollar.

6.6 RELIABILITY GROWTH

As the design cycle of a product progresses from concept to development, testing, and manufacturing, one expects that the implementation of design changes improves the product's reliability to achieve a design goal. Typically, a formal test analyze and fix (TAAF) program is implemented to discover design flaws and mitigate them. The gradual product improvement through the elimination of design deficiencies, which results in the increase of failure (inter)arrival times is known as *reliability growth*.

Generally speaking, reliability growth can be applicable to any level of a design decomposition, ranging from a component to a complete system. The (non-repairable) component level reliability growth can be readily established by comparing a chosen reliability metric for the consecutive design iterations or the product development milestone. For further reading on multiple comparisons of component level reliability data see Nelson (1982).

Most of the existing reliability growth models, however, are associated with repairable systems and, therefore, the basic reliability growth mathematics will be related to that considered in Section 5.1. The reliability growth methodology includes some new terms, that we have not formally defined thus far. The first is the *cumulative MTBF* (CMTBF), Θ_c, defined as the ratio of the total time on test t to the expected cumulative number of failures $E[N(t)]$.

$$\Theta_c = \frac{t}{E[N(t)]}$$

The second term is the *instantaneous MTBF* (IMTBF), Θ_i, defined as the inverse of the ROCOF function considered in Section 5.1.1.

$$\Theta_i = \frac{1}{\lambda(t)}$$

The difference between the two is that Θ_c is a function of ROCOF integrated over the interval (0, t), whereas Θ_i is the inverse of ROCOF at a given point in time *t*.

6.6.1 Graphical Method

One of the easiest and straightforward methods to assess reliability growth of a repairable system is to plot the cumulative number of failures versus cumulative time on test. The reliability growth is said to take place, if the TAAF based design changes lead to an increment drop in a cumulative number of failures as a function of total time on test (see Figure 6.7). Generally, the concave (convex) plot would indicate a reliability growth (deterioration), while a straight line would be an indication of no change in reliability behavior.

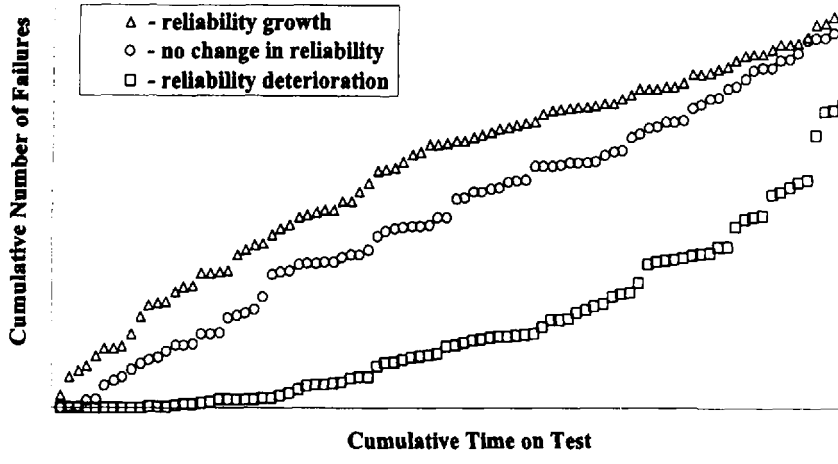


Figure 6.7 A nonparametric method of reliability growth evaluation.

6.6.2 Duane Method

Empirical studies conducted by Duane (1964) on a number of repairable systems have shown that the cumulative MTBF plotted against cumulative time on test in

a log-log space exhibit an almost linear relationship. Duane postulated a reliability growth model that expresses the CMTBF as a function of total time on test in the following form:

$$\Theta_c(t) = \frac{t}{N} = \frac{t^{1-\beta}}{\lambda} \tag{6.35}$$

where t is the total time on test, and N is the cumulative number of failures.

Taking the log of both sides of (6.35), one gets:

$$\ln[\theta_c(t)] = (1-\beta)\ln(t) - \ln(\lambda), \qquad t>0$$

which does indeed present an equation of a straight line. Parameters λ and β are referred to as the growth parameters and can be estimated using the linear regression technique discussed in Section 2.8. The inverse of parameter λ is sometimes referred to as the *initial MTBF*. The latter term becomes self-explanatory, if one sets t to one in (6.35).

The instantaneous MTBF can be derived by differentiating (6.35) with respect to t.

$$\Theta_i(t) = \left(\frac{dN}{dt}\right)^{-1} = \left(\frac{d\left(\frac{t}{\theta_c}\right)}{dt}\right)^{-1} = \frac{t^{1-\beta}}{\lambda\beta} \tag{6.36}$$

It can be seen that under the Duane model, the cumulative and instantaneous MTBFs are related to each other through parameter β:

$$\theta_c(t) = \beta\,\theta_i(t)$$

Keeping in mind that ROCOF, $\lambda(t)$, is the inverse of the instantaneous MTBF $\Theta_i(t)$, equation (6.36) can be represented in the following form

$$\lambda(t) = \beta\lambda t^{\beta-1} \tag{6.37}$$

Note that this is the exact algebraic form of the NHPP ROCOF model discussed in Section 5.1.4.

Besides, expression (6.37) formally coincides with the Weibull hazard rate function. As such, $\beta < 1$ represents reliability growth and $\beta > 1$ represents reliability degradation. O'Connor (1991) has suggested the following engineering interpretation of the growth parameter β:

β	Interpretation
0.4–0.6	The program's top priority is the elimination of failure modes. The program uses accelerated tests and suggests immediate analysis and effective corrective action for all failures.
0.3–0.4	The program gives priority to reliability improvement. The program uses normal environmental tests and well-managed analysis. Corrective action is taken for important failure modes.
0.2–0.3	The program gives routine attention to reliability improvement. The program does not use applied environmental tests. Corrective action is taken for important failure modes.
0.0–0.2	The program gives no priority to reliability improvement. Failure data are not analyzed. Corrective action is taken for important failure modes, but with low priority.

Once the parameters of the Duane model are estimated, it becomes possible to determine the TAAF test time required to attain a given target instantaneous MTBF under a given rate of reliability growth β:

$$t = \left(\theta_i \lambda \beta\right)^{\frac{1}{1-\beta}} \tag{6.38}$$

Example 6.8

The following are the miles-between-consecutive-failures of a new automobile subsystem obtained through the TAAF program: 5940, 12,331, 21,010, 27,192, 19,910, 24,211, 26,422, 27,731, 26,862, 29,271. Estimate the parameters of the Duane model and find the total test mileage required to attain the

target MMBF (mean miles between failures) of 50,000 miles under the estimated rate of reliability growth.

Solution:

Cumulative failures, N	Failure interarrival mileage, Δt	Failure arrival mileage, t	CMMBF, $\Theta_c = t/N$	Log (t)	Log (Θ_c)
1	5940	5940	5940	3.77	3.77
2	12,331	18,271	9136	4.26	3.96
3	21,010	39,281	13,094	4.59	4.12
4	27,192	66,473	16,618	4.82	4.22
5	19,910	86,383	17,277	4.94	4.24
6	24,211	110,594	18,432	5.04	4.27
7	26,422	137,016	19,574	5.14	4.29
8	27,731	164,747	20,593	5.22	4.31
9	26,862	191,609	21,290	5.28	4.33
10	29,271	220,880	22,088	5.34	4.34

The plot of the two rightmost columns of the above table is shown in Figure 6.8.

Using (2.101), the slope and intercept of the regression line in Figure 6.8 are estimated as 0.37 and 2.41, respectively. Then, the β parameter of the Duane model is

$$\beta = 1 - 0.37 = 0.63$$

which corresponds to a program dedicated to the reduction of failures. The λ parameter is found as

$$\lambda = 10^{-2.41} = 0.0039 \text{ mile}^{-1}$$

Using (6.38), the test mileage required to attain the instantaneous MMBF of 50,000 miles is

$$t = (50{,}000 \times 0.0039 \times 0.63)^{\frac{1}{1-0.63}} = 443{,}568 \text{ miles}$$

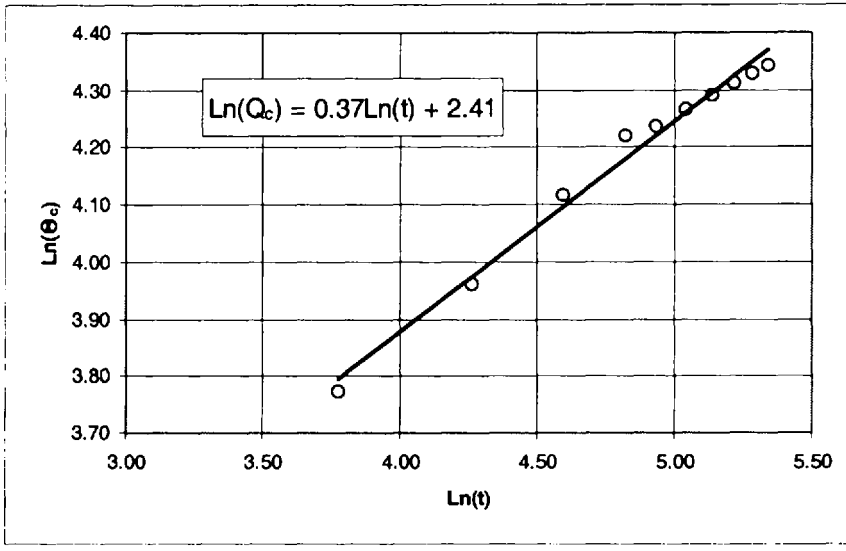


Figure 6.8 A Duane model plot in Example 6.8.

6.6.3 Army Material Systems Analysis Activity (AMSAA) Method

It is important to note that under Duane's assumption, (6.36) and (6.37) are deterministic models. Crow (1974) suggested that (6.37) could be treated probabilistically as the ROCOF of a nonhomogeneous Poisson process (see section 5.1.1). Such probabilistic interpretation of (6.37) is known as the AMSAA model and offers two major advantages.

First, the model parameters can be estimated through the maximum likelihood method using (5.30–5.31) and the confidence limits on these parameters can also be developed (Crow (1974)). Second, the distribution of the number of failures $f\{N(t)\}$ can be obtained based on:

$$\Pr(N(t) = n) = \frac{(\lambda t^{\beta})^n \exp(-\lambda t^{\beta})}{n!}, \quad n = 0, 1, 2, \ldots$$

Example 6.9

For the data in Example 6.8,

a. find the maximum likelihood estimates of the AMSAA model parameters,

b. determine the expected number of failures at 150,000 accumulated miles,

c. find the probability that the actual number of failures at 150,000 miles will be greater than the expected value determined in b.

Solution:

a. Using (5.30) and (5.31), the estimates of the AMSAA model parameters are

$$\hat{\beta} = \frac{10}{\ln\frac{220{,}880}{5940} + \ln\frac{220{,}880}{12{,}331} + \cdots + \ln\frac{220{,}880}{26{,}862}} = 0.86$$

$$\hat{\lambda} = 10 \times 220{,}880^{-0.86} = 0.00024 \text{ mile}^{-1}$$

b. The expected number of failures at 150,000 miles is

$$N(150{,}000) = \lambda t^{\beta} = 0.00024 \times 150000^{0.86} \approx 7$$

c. The probabilities of the actual number of failures taking a value of less than or equal to the expected number are provided in the table below.

n	$\Pr(N(150{,}000) = n) = \frac{7^n \exp(-7)}{n!}$
0	0.000912
1	0.006383
2	0.022341
3	0.052129
4	0.091226
5	0.127717
6	0.149003
7	0.149003
Total	0.598714

Thus, the probability of the actual number of failures being greater than 7 is

$$\Pr(N(150{,}000) > 7) = 1 - 0.5987 = 0.4013$$

 See the software supplement for the automated reliability growth analysis.

The concepts of reliability growth are discussed by a number of authors. Balaban (1978) presents the mathematical models of reliability growth. O'Connor (1991) discusses general methods for sequential testing, reliability demonstration, and growth monitoring. Fries and Sen (1996) present a comprehensive survey of discrete reliability growth models.

EXERCISES

6.1 An engine crankshaft is a good example of a high reliability part of a car. Although it is pounded by each cylinder with every piston stroke, that single bar remains intact for a long time. Assume the strength of the shaft is normal with the mean S and standard deviation s, while the load per stroke is L with standard deviation ℓ. Realize that a C cylinder engine hits the shaft at C different places along it, so these events can be considered independent. The problem will be to determine the reliability of the crankshaft.

a. Express the safety margin (SM) in terms of S, s, L, ℓ, and C.
b. Estimate the reliability. Assume the motor turns at $X(t)$ revolutions.
c. Express the total number of reversals. $N(t)$ seen by each piston as a function of time.
d. If the shaft is subject to fatigue, express the reliability as a function of time. Metals fatigue, generally, following the Manson–Coffin model: $S(N) = SN^{-1/q}$. Assume, also that the standard deviation, s does not change with N. Also, q is a constant. Determine the expected life (50% reliability) of the crankshaft turning at a constant rate, R (RPM).

6.2 Repeat Exercise 4.6 and calculate the Birnbaum and Fussell-Vesely importance measures for all events modeled in the fault tree.

6.3 The following data are given for a prototype of a system which undergoes design changes. A total of 10 failures have been observed since the beginning of the design. Estimate the Duane reliability growth model parameters. Discuss the results.

Failure number	1	2	3	4	5	6	7	8	9	10
Cumulative time on test (hrs)	12	75	102	141	315	330	342	589	890	1007

6.4 In response to an RFQ, two vendors have provided a system proposal consisting of subsystem modules *A*, *B*, and *C*. Each vendor has provided failure rates and average corrective maintenance time for each module. Determine which vendor system has the best MTTR and which one you would recommend for purchase.

		Vendor 1		Vendor 2	
Module	No. in system	Failure rate (per 10^4 hrs)	M_{ct} (min)	Failure rate (per 10^4 hrs)	M_{ct} (min)
A	2	45	15	45	20
B	1	90	20	30	15
C	2	30	10	90	10

a) Describe the advantages of a preventive maintenance program.
b) Is it worth doing preventive maintenance if the failure rate is constant?

6.5 You are a project engineer for the development of a new airborne radar system with a design goal of 7000 hours MTBF. The system has been undergoing development testing for the past six months, during which time eight failures have occurred in approximately 9000 test hours as follows.

Failure	Test hours to failure
1	1296
2	1582
3	1855
4	2310
5	3517
6	5188
7	6792
8	8902

How much time would you schedule for the balance of the test program, in order to have some confidence that your contractor had met its goal? (Note: each failure represented a different failure mode and corrective actions are being taken for each.)

REFERENCES

Amendola, A.U., Bersini, P.C., Cacciabue, C., and Mancini, G. "*Modeling Operators in Accident Conditions: Advances and Perspectives on Cognitive Model*," Int. J. Man-Machine Studies, 27: 599, 1987.

Balaban, H.S., "*Reliability Analysis for Complex Repairable Systems*," Reliability and Biometry, SIAM, 1978.

Birnbaum, Z.W., "*On the Importance of Different Components in a Multicomponent System*," in Multivariate Analysis-II (P.R. Krishnaiah, ed.), Academic Press, New York, NY, 1969.

Boehm W.B., "*A spiral model of software development and enhancement*," IEEE Computer, 61–72, 1988.

Crow, L. H., "*Reliability Analysis for Complex Repairable Systems, Reliability and Biometry*," F. Proschan and R. J. Serfling, eds., SIAM, Philadelphia, 1974.

Dougherty E.M. and Fragola, J.R. "*Human Reliability Analysis: A System Engineering Approach with Nuclear Power Plant Applications*," John Wiley and Sons, New York, NY, 1988.

Embry, D.E., Humphreys, P.C., Rosa, E.A., Kirwan, B., and Rea, K., "*SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Stuctured Expert Judgment*," U.S. Nuclear Regulatory Commission, NUREG/CR-3518, Washington DC, 1984.

Fries, A. and Sen, A., "*Survey of Descrete Reliability Growth Models*," IEEE Trans. Rel., Vol. 45, No.4, 1996.

Fussell, J., "*How to Hand Calculate System Reliability and Safety Characteristics*," IEEE Trans. Rel., Vol. R-24, No. 3, 1975.

Goel, A. L. and Okumoto, K., "*A Markovian Model for Reliability and Other Performance Measures of Software Systems*." In Proceedings of the National Computing Conference (New York), vol.48, 1979.

Goel, A.L. and Okumoto, K., "*A Time Dependent Error Detection Rate Model for Software Reliability and Other Performance Measures*." IEEE Trans. Rel., R28; 206, 1979.

Hall, R.E., Wreathall, J., and Fragola, J.R., "*Post Event Human Decision Errors: Operator Action/Time Reliability Correlation*," U.S. Nuclear Regulatory Commission, NUREG/CR-3010, Washington, DC, 1982.

Hannaman, G.W. and Spurgin, A.J., "*Systematic Human Action Reliability Procedure6(SHARP), Electric Power Research Institute*," NP-3583, Palo Alto, CA, 1984.

Hannaman, G.W., Spurgin, A.J., Lukic, Y.D., "*Human Cognitive Reliability Model for PRA Analysis*," NUS Corporation, NUS-4531, San Diego, CA, 1984.

Hoyland, A., and Rausand, M., "*System Reliability Theory: Models and Statistical Methods,*" John Wiley and Sons, New York, NY, 1994.

Hunns, D.M. and Daniels, B.K., "*The Method of Paired Comparisons,*" 3rd European Reliability Data Bank Seminar, University of Bradford, National Center of System Reliability, United Kingdom, 1980.

Jelinski, Z. and Moranda, P., "*Software Reliability Research.* In *Statistical Computer Performance Evaluation*", W. Freiberger, ed. Academic Press, New York, NY 1972.

Kapur, K.C. and Lamberson, L.R., "*Reliability in Engineering Design,*" John Wiley and Sons, New York, NY, 1977.

Lambert, H.E., "*Measures of Importance of Events and Cut Sets in Fault Trees in Reliability and Fault Tree Analysis,*" (R. Barlow, J. Fussell, and N. Singpurwalla eds.), SIAM, Philadelphia, PA, 1975.

Lawrence, J. D., "*Software Reliability and Safety in Nuclear Reactor Protection Systems,*" NUREG/CR-6101. Lawrence Livermore National Laboratory, 1993.

Littlewood, B., "*Software Reliability Model for Modular Program Structure,*" IEEE Trans. Rel., R-28(3), 1979.

Littlewood, B. and Verrall, J.K., "*A Bayesian reliability growth model for computer software,*" Appl. Statist. 22:332, 1973.

Littlewood, B., "*Software reliability model for modular program structure,*" IEEE Trans. Rel., R-28(3), 1979.

McDermid, J.A., "*Issues in Developing Software for Safety Critical Systems,*" Reliability Engineering and System Safety, Vol. 32, pp. 1–24, 1991.

Mills, H.D., "*On the Statistical Validation of ComputerPrograms,*" IBM Federal Systems Division, Rept. 72-6015, Gaithersburg, MD, 1972.

Musa, J.D., Iannino, A., and Okumoto, K., " *Software Reliability,*" McGraw-Hill, New York, NY, 1987.

Nelson, E., "*Estimating Software Reliability from TestDdate,*" Microelectronics Reliability 17:67, 1978.

Nelson, W., "*Applied Life Data Analysis,*" John Wiley and Sons, New York, NY, 1982.

O'Connor, "*Practical Reliability Engineering,*" 3rd ed., John Wiley and Son, New York, NY, 1991.

Petrella, S., et al., "*Random Testing of Reactor Shutdown System Software,*" Proceedings of the International Conference on Probabilistic Safety Assessment and Management, (G. Apostolakis, ed.), Elsevier, New York, NY 1991.

Potash, L.,Stewart, M., Diets, P.E., Lewis, C.M., and Dougherty, E.M., "*Experience in Integrating the Operator Contribution in the PRA of Actual Operating Plants,*" Proceedings of American Nuclear Society, Topical Meeting on Probabilistic Risk Assessment, Port Chester, New York, NY, 1981.

Poucet, A., "*Survey of Methods Used to Assess Human Reliability in the Human Factors Reliability Benchmark Exercise,*" Reliability Engineering and System Safety, 22, pp. 257–268, 1988.

Poucet, A., "*State of the Art in PSA Reliability Modeling as Resulting from the International Benchmark Exercise Project,*" NUCSAFE 88 Conference, Avignon, France, 1988.

Pressman, R.S., "*Software Engineering: A Practitioners Approach,*" 2nd ed., McGraw-Hill, 1987.

Ramamoorthy, C.V. and Bastani, F.B., "*Software Reliability: Status and Perspectives*," IEEE Trans. Soft. Eng. SE-8:359, 1982.

Rasmussen, J., "*Cognitive Control and Human Error Mechanisms*," Chapter 6 in (J. Rasmussen, K. Duncan, and J. LePlate ed.), New Technology and Human Error, John Wiley and Son, New York, NY, 1987.

Rasmussen, J., "*Skills, Rules and Knowledge: Signals, Signs and Symbols and Their Distinctions in Human Performance Models*," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-3, (3), pp. 257–268, 1982.

Reactor Safety Study: "*An Assessment of Accidents in U.S. Commercial Nuclear Power Plants*," U.S. Regulatory Commission, WASH-1400, Washington, DC, 1975.

Saaty, T.L., "*The Analytic Hierarchy Process*," McGraw-Hill, New York, NY, 1980.

Schick, G.J. and Wolverton, R.W., "*Assessment of Software Reliability*," 11th Annual Meeting German Oper. Res. Soc., DGOR, Hamburg, Germany; also in Proc. Oper. Res., Physica-Verlag, Wirzberg-Wien, 1973.

Severe Accident Risk: "*An Assessment for Five U.S. Nuclear Power Plants*," U.S. Nuclear Regulatory Commission, NUREG-1150, Washington, DC, 1990.

Sharirli, "*Methodology for System Analysis Using Fault Trees, Success Trees and Importance Evaluations*," Ph.D. dissertation, University of Maryland, Department of Chemical and Nuclear Engineering, College Park, MD, 1985.

Shooman, M.L., "*Software reliability measurements models*", In Proceedings of the Annual Reliability and Maintainability Symposium, Washington, DC, 1975.

Siegel, A.I., Bartter, N.D., Wolf, J., Knee, H.E., and Haas, P.M., "*Maintenance Personnel Performance Simulation (MAPPS) Model*," U.S. Nuclear Regulatory Commission, NUREG/CR-3626, Vol. I and II, Washington, DC, 1984.

Smidts, C., "*Software Reliability*," The Electronics Handbook IEEE Press, 1996.

Stillwell, W., Seaver, D.A., and Schwartz, J.P., "*Expert Estimation of Human Error Problems in Nuclear Power Plant Operations: A Review of Probability Assessment and Scaling*," U.S. Nuclear Regulatory Commission, NUREG/CR-2255, Washington, DC, 1982.

Swain, A.D., and Guttman, H.E., "*Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Applications*," U.S. Nuclear Regulatory Commission, NUREG/CR-1278, Washington, DC, 1983.

Woods, D.D., Roth, E.M., and Pole, H., "*Modeling Human Intention Formation for Human Reliability Assessment*," Reliability Engineering and System Safety, 22:169–200, 1988.

Wreathall, J., "*Operator Action Tree Method*," IEEE Standards Workshops on Human Factors and Nuclear Safety, Myrtle Beach, SC, 1981.

This page intentionally left blank

7

Selected Topics in Reliability Data Analysis

7.1 ACCELERATED LIFE TESTING

The reliability models considered in the previous chapters are expressed in terms of time-to-failure distribution or in terms of probability of failure on demand. Such models are not appropriate in the cases when one is interested in reliability dependence on such stress factors as ambient temperature, humidity, voltage applied to a unit, and operator's skill. This dependence is considered in the frame of the reliability models with explanatory variables or covariates (Leemis (1995)). Such models are traditionally referred to as *Accelerated Life Models* (ALM), which may be confusing because applications of these models are not necessarily limited to accelerated life testing as will be demonstrated in this section.

7.1.1 Basic Accelerated Life Notions

A *reliability model* (*Accelerated Life* (AL) *reliability model*) is defined as the relationship between the time-to-failure distribution of a device and stress factors, such as load, cycling rate, temperature, humidity, and voltage. The AL reliability models are based on physics of failure considerations.

Stress severity in terms of reliability (or time-to-failure distribution) is expressed as follows. Let $R_1(t; z_1)$ and $R_2(t; z_2)$ be the reliability functions of the item under constant stress conditions z_1 and z_2, respectively. It should be mentioned that stress condition, z, in general, is a vector of the stress factors. The stress condition z_2 is called more severe than z_1, if for all values of t the reliability of the item under stress condition z_2 is less than the reliability under stress condition z_1, i.e.,

$$R_2(t; z_2) < R_1(t; z_1) \tag{7.1}$$

Time-Transformation Function for the Case of Constant Stress

For the monotonic cdfs $F_1(t; z_1)$ and $F_2(t; z_2)$, if constant stress condition z_1 is less severe than z_2 and t_1 and t_2 are the times at which $F_1(t_1; z_1) = F_2(t_2; z_2)$, there exists a function g (for all t_1 and t_2) such that $t_1 = g(t_2)$, therefore

$$F_2(t_2; z_2) = F_1[g(t_2), z_1] \tag{7.2}$$

Because $F_1(t; z_1) < F_2(t; z_2)$, $g(t)$ must be an increasing function with $g(0) = 0$. The function $g(t)$ is called the *acceleration* or the *time transformation function*.

The AL reliability model is a deterministic transformation of time-to-failure. Two main time transformations are considered in reliability data analysis. These transformations are known as the *Accelerated Life* (AL) *Model* and the *Proportional Hazard* (PH) *Model.*

Accelerated Life Model

Accelerated Life model is the most popular type of reliability models with explanatory variables. For example, AT&T Reliability Model (AT&T Reliability Manual (1990)) is based on the AL model.

It may be assumed that $z = 0$ for the normal (use) stress condition. Denote a time-to-failure cdf under normal stress condition by $F_0(\cdot)$. The AL time transformation is expressed in terms of $F(t; z)$ and $F_0(\cdot)$, and it is given by the following relationship (Cox and Oaks (1984))

$$F(t; z) = F_0[t \cdot \psi(z, A)] \tag{7.3a}$$

where $\psi(z, A)$ is a positive function connecting time-to-failure with a vector of stress factors z; and A is a vector of unknown parameters; for $z = 0$, $\psi(z, A)$ is assumed to be 1.

The corresponding relationship for the pdf can be obtained from (7.3a) as

$$f(t; z) = f_0[t\,\psi(z, A)]\psi(z) \tag{7.3b}$$

where $f_0(\cdot)$ is the time-to-failure pdf under the normal stress condition. Relationship (7.3a) is the scale transformation. It means that a change in stress does not result in a change of the shape of the distribution function, but changes

its scale only. Relationship (7.3b) can be written in terms of the acceleration function as follows

$$g(t) = \psi(z, A)\, t \tag{7.4}$$

Relationship (7.3a) is equivalent to the linear with time acceleration function (7.4). The time-to-failure distributions of a device under the normal stress condition ($z = 0$) and the stress condition $z \neq 0$, are geometrically *similar* to each other. Such distributions are called *belonging to the class of time-to-failure distribution functions which is closed with respect to scale* (Leemis (1995)).

The similarity property is widely used in physics and engineering. Because it is difficult to imagine that any change of failure modes or mechanisms would not result in a change in the shape of the failure time distribution, relationship (7.3a) can be also considered as a principle of failure mechanism conservation or a *similarity* principal, which states that the failure modes and mechanisms remain the same over the stress domain of interest. The analysis of some sets of real life data often show that the similarity of time-to-failure distributions really exists, so that a violation of the similarity can identify a change in a failure mechanism.

The relationship for the $100p$th percentile of time-to-failure, $t_p(z)$, can be obtained from (7.3a) as

$$t_p(z) = \frac{t_p^0}{\psi(z, A)} \tag{7.5}$$

where t_p^0 is the $100p$th percentile for the normal stress condition $z = 0$.

The relationship (7.5) is the *percentile AL reliability model* and it is usually written in the form

$$t_p(z, B) = \eta(z, B) \tag{7.6}$$

where B is a vector of unknown parameters. Reliability models are briefly considered in Section 7.1.2.

The AL reliability model is related to the relationship for percentiles, (7.5), as

$$\eta(z, B) = \frac{t_p^0}{\psi(z, A)} \tag{7.7}$$

The corresponding relationship for failure rate can also be obtained from (7.3a) as

$$\lambda(t; z) = \psi(z, A)\, \lambda^0[t \cdot \psi(z, A)] \tag{7.8}$$

It is easy to see that the relationship for percentiles (7.5) is the simplest one.

Cumulative Damage Models and Accelerated Life Model

Some known cumulative damage models result in the similarity of time-to-failure distributions under quite reasonable restrictions. As an example, consider the Barlow and Proschan model (Barlow and Proschan (1981)) resulting in an aging (IFRA) time-to-failure distributions, introduced in Section 3.1.2.

An item subjected to shocks occurring randomly in time, is considered. Let these shocks arrive according to the Poisson process with constant intensity λ. Each shock causes a random amount x_i of damage, where $x_1, x_2, \ldots, x_i$ are random variables distributed with a common cdf, $F(x)$, called a *damage distribution function*. The item fails when accumulated damage exceeds a threshold x. It has been shown by Barlow and Proschan that for *any* damage distribution function $F(x)$, the time-to-failure distribution function is IFRA.

Now consider an item under the stress conditions characterized by different shock intensities λ_i and different damage distribution functions $F_i(x)$. It can be also shown that the similarity of the corresponding time-to-failure distribution functions will hold for all these stress conditions, z_i (λ_i, $F_i(x)$), if they have the same damage cdf, i.e., $F_i(x) = F(x)$. A similar example from fracture mechanics is considered in (Crowder et al. (1991)).

Proportional Hazard Model

For the proportional hazard (PH) model the basic relationship analogous to (7.3a) is given by

$$F(t; z) = 1 - \left[1 - F_0(t)\right]^{\psi(z, A)} \tag{7.9a}$$

or, in terms of reliability function, $R(t)$, as

$$R(t; z) = R_0(t)^{\psi(z, A)} \tag{7.9b}$$

The proper *PH (Cox) model* is known as the relationship for hazard rate (Cox and Oaks (1984)), which can be obtained from (7.9a or 7.9b) as

$$\lambda(t; z) = \psi(z, A)\, \lambda^0(t) \tag{7.10}$$

where $\psi(z, A)$ is usually chosen as a log-linear function.

Note that the PH model time transformation does not normally retain the shape of the cdf, and the function $\psi(z)$ no longer has a simple relationship to the acceleration function, nor has a clear physical meaning. That is why the PH model is not as popular in reliability applications as the AL model.

It should be mentioned that for the Weibull distribution (and only for the Weibull distribution) the PH model coincides with the AL model (Cox and Oaks (1984)).

7.1.2 Some Popular AL (Reliability) Models

The most commonly used AL models for the percentiles (including median) of time-to-failure distributions are log-linear models. Two of such models are the *Power Rule Model* and the *Arrhenius Reaction Model* (Nelson (1990)). The Power Rule model is given as:

$$t_p(x) = \frac{a}{x^c}, \quad a > 0, \quad c > 0, \quad x > 0 \tag{7.11a}$$

where x is a mechanical or electrical stress, c is a unitless constant, the unit of constant a being the product of *time* and the measure of x^c. In reliability of electrical insulation and capacitors, x is usually applied voltage. In estimating fatigue life the model is used as the analytical representation of the, so-called, *S-N* or *Wöhler curve*, where S is stress amplitude and N is life in cycles to failure, such that:

$$N = kS^{-b} \tag{7.11b}$$

where b and k are material parameters estimated from test data. Because of the probabilistic nature of fatigue life at any given stress level one has to deal with not one *S-N* curve, but with a family of *S-N* curves, so that each curve is related to a probability of failure as the parameter of the model. These curves are called *S-N-P* curves, or curves of constant probability of failure on a stress-versus life plot. It should be noted that relationship (7.11b) is an empirical model (Sobczyk and Spencer (1992)).

Another popular model is the Arrhenius Reaction Rate Model:

$$t_p(T) = a \exp\left(\frac{E_a}{T}\right) \tag{7.12}$$

where T is the absolute temperature, under which the unit is functioning, and E_a is the activation energy. This model is the most widely used expressing the effect of temperature on reliability. The application of the Arrhenius for electronic component reliability estimation was briefly discussed in Section 3.7. Originally the model was introduced as a chemical reaction rate model.

Another model is a combination of models (7.11) and (7.12):

$$t_p(x, T) = ax^{-c}\exp\left(\frac{E_a}{T}\right) \tag{7.13}$$

where x (as defined by (7.11)) is a mechanical or electrical stress. This model is used in fracture mechanics of polymers, as well as a model for the electromigration failures in aluminum thin films of integrated circuits. In the last case stress factor x is current density.

Jurkov's model (Nelson (1990)) is another popular AL reliability model:

$$t_p(x, T) = t_0 \exp\left(\frac{E_a - \gamma x}{T} \right) \tag{7.14}$$

This model is considered as an empirical relationship reflecting the thermal fluctuation character of long-term strength, i.e, durability under constant stress, (Goldman (1994)). For mechanical long-term strength, parameter t_0 is a constant, which is numerically close to the period of thermal atomic oscillations (10^{-11} - 10^{-13} s); E_a is the effective activation energy, which is numerically close to vaporization energy for metals and to chemical bond energies for polymers, and γ is a structural coefficient. The model is widely used for reliability prediction problems of mechanical and electrical (insulation, capacitors) long-term strength.

The *a priori* choice of a model (or some competing models) is made based on physical considerations. Meanwhile, statistical data analysis of accelerated life test results or collected field data, combined with failure mode and effects analysis (FMEA) can be used to check the adequacy of the chosen model, or to discriminate the most appropriate model among the competing ones.

7.1.3 Accelerated Life Data Analysis

Exploratory Data Analysis (Criteria of Linearity of Time Transformation Function for Constant Stress)

The experimental verification of the basic ALM assumption (7.3a) is not only important in failure mechanism study, but also has a great practical importance, because almost all statistical procedures for AL test planning and data analysis are based on this assumption. Several techniques can be used for verification of the linearity of the time transformation function. Some of them are briefly discussed below.

Two-Sample Criterion

Let's start with the first criterion which can make clear the physical meaning of the idea of similarity of time-to-failure distributions. This criterion requires two special tests. During the first test, a sample is tested at constant stress level z_1 over time period t_1 , at which z_1 is changed to a constant stress z_2 for time period t_2.

Such loading pattern (load as function of time) can be called *stress profile* S_1. During the second test, another sample is first tested under z_2 during t_2 and then it is tested under the stress level z_1 during the time t_1 (stress profile S_2). The time transformation function will be a linear function of time, if the reliability functions (or the corresponding failure probabilities) of the items after the first and the second tests are equal (i.e., a change of loading order does not change the cumulative damage).

The corresponding statistical procedure can be based on the analysis of the, so-called, 2x2 contingency tables (Nelson (1982)). This analysis was initially developed for comparing binomial proportions (probabilities).

The null hypothesis tested, H_0, is

$$H_0: \; p_1(S_1) = p_2(S_2) = p$$

where p is the failure probability during the test with stress profile S_1 (S_2), or, in terms of reliability functions, the null hypothesis is expressed as

$$H_0: R_1(S_1) = 1 - p_1(S_1) = R_2(S_2) = 1 - p_2(S_2) = R = 1 - p$$

The alternative hypothesis, H_1, is

$$H_1: \; p_1(S_1) \neq p_2(S_2)$$

Let n_1 and n_2 be the sample sizes tested under stress profiles S_1 and S_2, respectively. Further, let n_{1f} and n_{2f} be the number of items failed during these tests. Denote the corresponding numbers of nonfailed items by n_{1s} and n_{2s}. Obviously $n_1 = n_{1f} + n_{1s}$ and $n_2 = n_{2f} + n_{2s}$. Finally denote $N = n_1 + n_2$. These test data can be arranged in the following contingency table.

Stress profile 1	Stress profile 2
n_{1f}	n_{2f}
n_{1s}	n_{2s}

If H_0 is true:

1. the probability $p_1(S_1) = p_2(S_2) = p$ can be estimated as

$$\hat{p} = \frac{n_{1f} + n_{2f}}{n_1 + n_2} = \frac{n_{1f} + n_{2f}}{N}$$

2. the reliability functions $R_1(S_1) = R_2(S_2) = R$ can be estimated as

$$\hat{R} = \frac{n_{1s} + n_{2s}}{n_1 + n_2} = \frac{n_{1s} + n_{2s}}{N}$$

3. based on these estimates, the expected frequencies n_{1f}, n_{2f}, n_1s, and n_{2s} can be estimated as

$$\hat{n}_{1f} = \hat{p}\, n_1, \quad \hat{n}_{2f} = \hat{p}\, n_2$$
$$\hat{n}_{1s} = \hat{R}\, n_1, \quad \hat{n}_{2s} = \hat{R}\, n_2$$

The following measure of discrepancy between the observed and expected frequencies for the contingency table can be introduced as

$$W = \frac{(n_{1f} - \hat{n}_{1f})^2}{\hat{n}_{1f}} + \frac{(n_{1s} - \hat{n}_{1s})^2}{\hat{n}_{1s}} + \frac{(n_{2f} - \hat{n}_{2f})^2}{\hat{n}_{2f}} + \frac{(n_{2s} - \hat{n}_{2s})^2}{\hat{n}_{2s}}$$

which under the null hypothesis follows an approximate χ^2 distribution with (4 2 - 1) = 1 degree of freedom. Thus, for a significance level, α, the null hypothesis is rejected if the above sum W is greater than the critical value of $\chi^2_{1,\alpha}(1)$.

Example 7.1

Two samples of identical thin polymer film units were tested. The first sample of 48 units was tested under stress profile (S_1): during one hour the units were under voltage of 50 V, then the voltage was instantaneously increased to 70 V, under which the sample was tested for another hour. The second sample of 52 units was tested under the backward stress profile (S_2): it was put under 70 V during the first hour, then the voltage was decreased to 50 V and the test was continued during the next hour. The data obtained from the two sample tests are given in the table below. Test the null hypothesis: $p_1(S_1) = p_2(S_2) = p$

Stress profile 1	Stress profile 2
$n_{1f} = 19$	$n_{2f} = 32$
$n_{1s} = 29$	$n_{2s} = 20$

Solution:

Find: $n_1 = n_{1f} + n_{1s} = 48, \quad n_2 = n_{2f} + n_{2s} = 52, \quad n_1 + n_2 = N = 100.$

The probability $p_1(S_1) = p_2(S_2) = p$ is estimated as

$$\hat{p} = \frac{n_{1f} + n_{2f}}{n_1 + n_2} = \frac{n_{1f} + n_{2f}}{N}$$

$$\hat{p} = \frac{19 + 32}{100} = \frac{51}{100}$$

Similarly, the probability $R_1(S_1) = R_2(S_2) = R$ is estimated as

$$\hat{R} = \frac{n_{1s} + n_{2s}}{n_1 + n_2} = \frac{n_{1s} + n_{2s}}{N}$$

$$\hat{R} = \frac{29 + 20}{100} = \frac{49}{100}$$

The corresponding expected frequencies are calculated as

$$\hat{n}_{1f} = \hat{p}n_1 = \frac{(48)(51)}{100} = 24.48$$

$$\hat{n}_{2f} = \hat{p}n_2 = \frac{(52)(51)}{100} = 26.52$$

$$\hat{n}_{1s} = \hat{R}n_1 = \frac{(49)(48)}{100} = 23.52$$

$$\hat{n}_{2s} = \hat{R}n_2 = \frac{(49)(52)}{100} = 25.48$$

Finally, find the value of Chi-squared statistic as

$$W = \frac{(19 - 24.48)^2}{24.48} + \frac{(32 - 26.52)^2}{26.52} + \frac{(29 - 23.52)^2}{23.52} + \frac{(20 - 25.48)^2}{25.48} \approx 4.81$$

If α is chosen as 5%, $\chi^2_{0.95}(1) = 3.82$, therefore, our null hypothesis is rejected, which means that AL model (7.3a) is not applicable for the polymer film specimens, when the applied voltage is changed from 50 V up to 70 V. This conclusion can indicate a change in failure mechanisms due to a voltage increase.

Checking the Coefficient of Variation

The second criterion is associated with the coefficient of variation (i.e., standard deviation to mean ratio, σ/m). It is possible to show that if the time transformation function is linear with respect to time for some constant stress levels $z_1, z_2, \ldots, z_k$, the coefficient of variation of time-to-failure will be the same for all these stress levels.

Logarithm of Time-to-Failure Variance

It can also be shown that under the same assumption the variance of the logarithm of times to failure will be the same for the stress levels at which the AL model holds. For the lognormal time-to-failure distribution the Bartlett's and Cochran's tests can be used for checking if the variances are constant (Nelson (1990)).

Quantile-Quantile Plots

The quantile-quantile plot is a curve, such that the coordinates of every point are the time-to-failure quantiles (percentiles) for a pair of stress conditions of interest. If the time transformation function is linear in time (i.e., relationship (7.3a) holds), the quantile-quantile plot will be a straight line going through the origin. A sample quantile, $\hat{t}_p$, of level p (i.e., an estimate of the respective true quantile, t_p) for a sample of size n is defined as :

$$\hat{t}_p = \begin{cases} t_{(np)}, & \text{if } np \text{ is not integer, and} \\ \text{any value from the interval } \left[t_{(np)}, t_{(np+1)}\right], & \text{if } np \text{ is integer} \end{cases}$$

where $t_{(i)}$ is the failure time (order statistic), and $[x]$ means *the greatest integer which does not exceed x*.

The corresponding data analysis procedure is realized in the following way. All the sample quantiles of a given constant stress condition are plotted on one axis and the sample quantiles of another stress condition are plotted on the other axis. If the sample sizes for two stress conditions are equal, the corresponding order statistics can be used as the sample quantiles. Using the points obtained (a

pair of quantiles of the same level gives a point), a straight regression line can be fitted. The AL model will be applicable, provided one gets linear dependence between the sample quantilies, and if the hypothesis that the intercept of the fitted line is equal to zero, is not rejected (for more details see (Crowder et al. (1991))

Example 7.2

For the data given in Example 2.33, verify the applicability of AL model (7.3) assumptions.

Solution:

The values of sample coefficients of variation (i.e., sample standard deviation to sample mean ratio) for the time-to failure data obtained under the temperatures 50, 60, and 70°C as well as the corresponding logarithms of the time-to-failure variances are given in the following table. It is easy to see that the values of sample coefficients of variation and the values of logarithms of the time-to-failure variances are very close to each other for the respective temperatures. Thus, the ALT model assumptions look realistic for the data given.

Temperature °C	Sample coefficients of variation	Logarithm time to failure variances
50	0.678	0.632
60	0.573	0.302
70	0.626	0.521

The same conclusion could be drawn using the quantile-quantile plots for these data. They show strong linear dependence between the sample quantilies (all the correlation coefficients are greater than 0.95) and the respective intercepts of the fitted lines are reasonably insignificant. Figure 7.1 provides an example of the quantile-quantile plots for the temperatures 50 and 70°C .

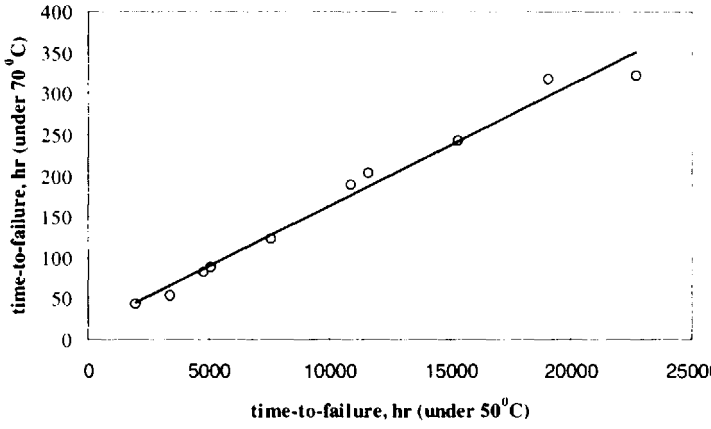


Figure 7.1 Quantile-quantile plot in Example 7.2.

Reliability Models Fitting: Constant Stress Case

Statistical methods for reliability model fitting on the basis of AL tests or field data collected can be divided into groups—parametric and nonparametric. For the former, the time-to-failure distribution is assumed to be a specific parametric distribution—normal, exponential, Weibull; while for the latter the only assumption is that the time-to-failure distribution belongs to a particular class of time-to-failure distribution, i.e., continuous, IFR, IFRA.

The most commonly used parametric methods are the parametric regression (normal and lognormal, exponential, Weibull and extreme value), least squares method, and maximum likelihood method. The following discusses the least squares method for uncensored data (Cox and Oaks (1984); Nelson (1990)).

The relationship for quantiles (7.5) can be written in terms of random variables as

$$t = \frac{t_0}{\psi(z)}$$

where the time-to-failure, t_0, under normal stress has cdf, $F_0(\cdot)$.

Designate the expectation of $\log t_0$ by μ_0, i.e.,

$$E(\log t_0) = \mu_0$$

Using the equation above one can write

$$\log t = \mu_0 - \log\psi(z) + \epsilon \tag{7.15}$$

where ϵ is a random variable of zero mean with a distribution not depending on z. To make (7.15) clear, note that any random variable x, having an expectation, $E(x)$, and a finite variance, var(x), can be represented as

$$x = E(x) + \epsilon$$

here $E(\epsilon) = 0$, and var(ϵ) = var(x)

If log $\psi(z)$ is a linear function with respect to parameters B function (the case of loglinear reliability model), i.e.,

$$\log\psi(Z, B) = ZB$$

equation (7.15) can be written as

$$\log t = \mu_0 - ZB + \epsilon$$

which is a linear, with respect to parameters B, regression model.

When time-to-failure samples are uncensored, the regression equation for observations t_i, Z_i ($i = 1, 2, \ldots, n$) is

$$\log t_i = \mu_0 - Z_i B + \epsilon_i$$

where for any time-to-failure distribution, ϵ_i ($i = 1, 2, \ldots, n$) are independent and identically distributed random variables with an unknown variance and known distribution form (if the time-to-failure distribution is known). Thus, on the one hand, the least squares technique (briefly considered in Section 2.8) for AL data analysis can be used as a nonparametric model, on the other hand, if time-to-failure distribution is known, one can use a parametric approach. The lognormal time-to-failure distribution is an example of the last case, which is reduced to standard normal regression. This is why the lognormal distribution is popular in AL practice. The respective example of a model parameter estimation problem for the Arrhenius model has already been considered in Chapter 2 (Example 2.33). The problems of optimal Design of Experiments (DoE) for ALT are considered in (Nelson (1990)).

7.1.4 Accelerated Life Model for Time-Dependent Stress

The models considered in the previous sections are related to constant stress. The case of time-dependent stress is not only more general, but also of more practical importance because its applications in reliability are not limited by accelerated life testing problems. As an example, consider the time-dependent stress analog of the power rule model (7.11b).

The stress amplitude, S, experienced by a structural element often varies during its service life, so that the straightforward use of Equation (7.11b) is not possible. In such situations the, so-called, *Palmgren–Miner rule* is widely used to estimate the fatigue life. The rule treats fatigue fracture as a result of a *linear accumulation* of partial fatigue damage fractions. According to the rule, the damage fraction, Δ_i, at any stress level S_i is linearly proportional to the ratio n_i / N_i, where n_i is the number of cycles of operation under stress level S_i, and N_i is the total number of cycles to failure (life) under the *constant* stress level S_i, i.e.,

$$\Delta_i(S_i) = \frac{n_i(S_i)}{N_i(S_i)}, \quad n_i \le N_i$$

The total accumulated damage, D, under different stress levels S_i ($i = 1, 2, \ldots, n$) is defined as

$$D = \sum_i \Delta_i = \sum_i \frac{n_i}{N_i}$$

It is assumed that failure occurs if the total accumulated damage $D \ge 1$.

Accelerated life tests with time dependent stress such as step-stress and ramp-tests are also of great importance. For example, one of the most common reliability tests of thin silicon dioxide films in metal-oxide-semiconductor integrated circuits is the so-called ramp-voltage test. In this test the oxide film is stressed to breakdown by a voltage which increases linearly with time (Chan, (1990)).

Let $z(t)$ be a time dependent stress such that $z(t)$ is integrable. In this case the basic relationship (7.3a) can be written in the form given by Cox and Oaks (1984):

$$F[t;\ z(t)] = F_0[\Psi(t)] \tag{7.16}$$

where

$$\Psi(t^{(z)}) = \int_0^{t^{(z)}} \psi[z(s), A]\, ds$$

and $t^{(z)}$ is the time related to an item under the stress condition $z(t)$.

Based on (7.16), the analogous relationships for the pdf and failure rate function can be obtained. The corresponding relationship for the $100p$th percentile of time-to-failure $t_p[z(t)]$ for the time-dependent stress, $z(t)$, can be obtained from (7.16) as

$$t_p^0 = \int_0^{t_p[z(t)]} \psi\,[z(s), A]\,ds \tag{7.17}$$

Using (7.6) and (7.7), (7.17) can be rewritten as

$$1 = \int_0^{t_p[z(t)]} \frac{1}{t_p^0\{\psi[z(s), A]\}^{-1}}\,ds$$

or, using (7.7), in terms of the percentile reliability models, as

$$1 = \int_0^{t_p[z(t)]} \frac{1}{\eta[z(s), B]}\,ds \tag{7.18}$$

AL reliability model for time-dependent stress and Palmgren–Miner's Rule

It should be noted that relationship (7.18) is an exact nonparametric probabilistic continuous form of the Palmgren–Miner rule. So, the problem of using AL tests with time-dependent stress is identical to the problem of cumulative damage addressed by the Palmgren–Miner rule. Moreover, there exists a useful analogy between mechanical damage accumulation and electrical breakdown. For example, the power rule and Jurkov's models are used as the relationship for mechanical as well as for electrical long-term strength. There are two main applications of Equation (7.18):

1. fitting an AL reliability model (estimating the vector of parameters, B, of percentile reliability model, $\eta(z, B)$, on the basis of AL tests with time-dependent stress), and
2. reliability (percentiles of time-to-failure) estimation (when reliability model is known) for the given time-dependent, in the stress domain, where conservation of failure mechanisms holds.

Example 7.3

The constant stress reliability model for a component is based on the Arrhenius model for the 5th percentile of time-to-failure given by the following equation

$$t_{0.05} = 2.590 \exp\left(\frac{0.400}{0.862 \times 10^{-4}\,(273 + T)}\right)$$

where $t_{0.05}$ is 5th percentile in hours, and T is temperature in °C. Find the 5th percentile of time-to-failure for the following cycling temperature profile, $T(t)$:

$$T(t) = 25°C \text{ for } 0 < t \leq 24 \text{ h}$$
$$T(t) = 35°C \text{ for } 24 < t \leq 48 \text{ h}$$
$$T(t) = 25°C \text{ for } 48 < t \leq 72 \text{ h}$$
$$T(t) = 35°C \text{ for } 72 < t \leq 96 \text{ h}$$

Solution:

An exact solution can be found as a solution for the following equation (based on relationship (7.18)):

$$\int_0^{t_p[T(s)]} \frac{1}{A \exp\left(\dfrac{E_a}{b(T(s) + 273)} \right)} ds = 1$$

Replacing the integral by the following sum, one gets:

$$\sum_{i=1}^{k(t)} \delta_i + \delta(t^*) = 1$$

where $\delta_i = \delta$ is damage accumulated in a complete cycle (48 hour period), $\delta(t^*)$ is damage accumulated during the last incomplete cycle, having duration t^*, k is the largest integer, for which $k\delta < 1$ and

$$\delta = \Delta_1 + \Delta_2$$

where Δ_1 is the damage associated with the first 24 hours of the cycle (under 25°C), and Δ_2 is the damage associated with the second part of the cycle (under 35°C). These damages can be calculated as:

$$\Delta_1 = \frac{24}{A \exp\left(\dfrac{E_a}{b(T_1 + 273)} \right)}$$

$$\Delta_2 = \frac{24}{A \exp\left(\dfrac{E_a}{b(T_2 + 273)} \right)}$$

where $T_1 = 25°C$ and $T_2 = 35°C$. The numerical calculations result in $\Delta_1 = 1.6003E-6$ and $\Delta_2 = 2.6533E-6$ Thus,

$$\delta = \Delta_1 + \Delta_2 = 4.2532E-6$$

The integer k is calculated as $k = [1/\delta] = 235100$, where $[x]$ means the greatest integer which does not exceed x.

Estimate the damage accumulated during the last incomplete cycle, $\delta(t^*)$, as

$$\delta(t^*) = 1 - k\delta = 1 - 2.35E-5 \times 4.25E-6 = 2.1510E-6 > 1.6003E-6$$

which means that the last temperature in the profile is 35°C. Find t^* as a solution of the following equation

$$\int_0^{t^*-24} \frac{1}{A \exp\left(\frac{E_a}{b(35+273)}\right)} ds = (2.15E-6) - (1.60E-6)$$

which gives $t^* - 24 = 4.97$ (hrs). Finally, the exact solution is

$$t_p = 48k + 24 + 4.97 \approx 1.13 \times 10^7 \text{ (hr)}$$

It is clear that the correction obtained is negligible, but in the case when the cycle period is comparable with the anticipated life, the correction can be significant.

7.1.5 Exploratory Data Analysis for Time-Dependent Stress

Basically, the two sample criterion considered earlier, is the criterion for the particular time-dependent stress. Generally speaking, the value of the integral in (7.18) does not change when a stress history $z(t)$, is changed to $z(t_p - t)$, $t_p \geq t > 0$; which means that time is reversible under the AL model. Based on this property, it is not very difficult to verify if the AL model assumptions are applicable to a given problem. For example, each sample which is going to be tested under time-dependent stress can be divided in two equal parts, so that the first sub-sample could be tested under the forward stress history, while the second sub-sample is tested under the backward stress.

Statistical Estimation of AL Reliability Models on the Basis of AL Tests with Time-Dependent Stress

Using Equation (7.18) the time-dependent percentile regression model can be obtained in the following form

$$t_p^0 = \int_0^{\hat{t}_p[z(t)]} \psi\,[z(s), A]\,ds \tag{7.19}$$

where $\hat{t}_p[z(t)]$ is the sample percentile for an item under the stress condition (loading history) $z(t)$. The problem of estimating the vector A and t_p^0 in this case cannot be reduced to parameter estimation for a standard regression model as in the case of constant stress.

Consider k different time-dependent stress conditions (loading histories) $z_i(t)$, $i = 1,2,\ldots k$, $[k > (\dim A) + 1]$, where the test results are complete or Type II censored samples and the number of uncensored failure times and the sample sizes are large enough to estimate the t_p as the sample percentile $\hat{t}_p$. In this situation the parameter estimates for the AL reliability model (A and t_p^0) can be obtained using a least squares method solution of the following system of integral equations:

$$t_p^0 = \int_0^{\hat{t}_p[z_i(t)]} \psi\,[z_i(s), A]\,ds, \quad i = 1, 2, \cdots, k \tag{7.20}$$

Example 7.4 (Kaminskiy et al., (1995))

Assume a model (7.13) for the 10th percentile of time-to-failure $t_{0.1}$ of a ceramic capacitor in the form

$$t_{0.1}(U, T) = a\,U^{c} \exp\left(\frac{E_a}{T}\right)$$

where U is applied voltage and T is absolute temperature. Consider a time-step-stress AL test plan using step-stress voltage in conjunction with constant temperature as accelerating stress factors. A test sample starts at a specified low voltage U_0 and it is tested for a specified time Δt. Then the voltage is increased by ΔU, and the sample is tested at $U_0 + \Delta U$ during Δt, i.e.,

$$U(t) = U_0 + \Delta U \times En\left(\frac{t}{\Delta t}\right)$$

where $En(x)$ means "nearest integer not greater than x." The test will be terminated after the portion $p \geq 0.1$ of items fails. So the test results are sample percentiles at each voltage-temperature combination. The test plan and simulated results with

$\Delta U = 10$ V, $\Delta t = 24$ h are given in Table 7.1 Estimate the model parameters a, c, and E_a.

Solution:

For the example considered the system of integral equations (7.20) takes the form:

$$a = \int_0^{t_{0,i}} \exp\left(-\frac{E_a}{T_i}\right)\left[U(s_i)\right]^c ds, \quad i = 1, 2, 3, 4$$

or

$$a = \int_0^{347.9} \exp\left(-\frac{E_a}{398}\right)\left[U(s_1)\right]^c ds$$

$$a = \int_0^{1688.5} \exp\left(-\frac{E_a}{358}\right)\left[U(s_2)\right]^c ds$$

$$a = \int_0^{989.6} \exp\left(-\frac{E_a}{373}\right)\left[U(s_3)\right]^c ds$$

$$a = \int_0^{1078.6} \exp\left(-\frac{E_a}{373}\right)\left[U(s_4)\right]^c ds$$

where $U(s_1) = U(s_2)$. Solving this system for the data above yields the following estimates for the model (7.13): $a = 2.23E-8\ \text{hV}^{1.88}$, $E_a = 1.32E4\,^\circ\text{K}$, $c = 1.88$, which are close to the following values of the parameters used for simulating the data: $a = 2.43E-8\ \text{H/V}^{1.87}$, $E_a = 1.32E4\,^\circ\text{K}$, $c = 1.87$. The values of the percentiles predicted using the model, are given in the last column of Table 7.1

Table 7.1 Ceramic Capacitors Test Results

Temperature °K	Voltage U_0, V	Sample time-to-failure percentile hr	Time-to-failure percentile (predicted) hr
398	100	347.9	361.5
358	150	1688.5	1747.8
373	100	989.6	1022.8
373	63	1078.6	1108.6

7.2 ANALYSIS OF DEPENDENT FAILURES

Dependent failures are extremely important in reliability analysis and must be given adequate treatment so as to minimize gross overestimation of reliability. In general, dependent failures are defined as events in which the probability of each failure is dependent on the occurrence of other failures. According to (2.14), if a set of dependent events $\{E_1, E_2, \ldots, E_n\}$ exists, then the probability of each failure in the set depends on the occurrence of other failures in the set.

The probabilities of dependent events in the left-hand side of (2.14) are usually, but not always, greater than the corresponding independent probabilities. Determining the conditional probabilities in (2.14) is generally difficult. However, there are parametric methods that can take into account the conditionality and generate the probabilities directly. These methods are discussed later in this section.

Generally, dependence among various events, e.g., failure events of two items, is either due to the internal environment of these systems or external environment (or events). The internal aspects can be divided into three categories: internal challenges, intersystem dependencies, and intercomponent dependencies. The external aspects are natural or human-made environmental events that make failures dependent. For example, the failure rates for items exposed to extreme heat, earthquakes, moisture, and flood will increase. The intersystem and intercomponent dependencies can be categorized into four broad categories: functional, shared equipment, physical, and human caused dependencies. These are described in Table 7.2.

The major causes of dependence among a set of systems or components as described in Table 7.2 can be explicitly described and modeled, e.g., by system reliability analysis models, such as fault trees. However, the rest of the causes can be collectively modeled using the concept of common cause failures (CCFs). Common cause failures are considered as the collection of all sources of dependencies described in Table 7.2 (especially between components) that are not known, or are difficult to explicitly model in the system or component reliability analysis. For example, functional and shared equipment dependencies are often handled by explicitly modeling them in the system analysis, but other dependencies are considered collectively using CCF.

CCFs have been shown by many reliability studies to contribute significantly to the overall unavailability or unreliability of complex systems. There is no unique and universal definition for CCFs. However, a fairly general definition of CCF is given by Mosleh et al. (1988) as " . . a subset of dependent events in which two or more component fault states exist at the same time, or in a short time interval, and are direct results of a shared cause."

Table 7.2 Types of Dependent Events

Dependent event type	Dependent event category	Subcategory	Example
Internal	1. Challenge	—	Internal transients or deviations from the normal operating envelope introduce a challenge to a number of items.
	2. Intersystem (Failure between two or more systems)	1. Functional	Power to several independent systems is from the same source.
		2. Shared equipment	The same equipment, e.g., a valve, is shared between otherwise independent systems.
		3. Physical	The extreme environment, (e.g., high-temperature, causes dependencies between independent systems.
		4. Human	Operator error causes failure of two or more independent systems.
	3. Intercomponent	1. Functional	A component in a system provides multiple functions.
		2. Shared equipment	Two independent trains in a hydraulic system share the same common header.
		3. Physical	Same as system interdependency above.
		4. Human	Design errors in redundant pump controls introduces a dependency in the system
External	—	—	Earthquake or fire fails a number of independent systems or components.

To better understand CCFs, consider a system with three redundant components A, B, and C. The total failure probability of A can be expressed in terms of its independent failure A_I and dependent failures as follows.

C_{AB} is the failure of components A and B (and not C) from common causes;
C_{AC} is the failure of components A and C (and not B) from common causes;
C_{ABC} is the failure of components A, B, and C from common causes.

Component A fails if any of the above events occur. The equivalent Boolean representation of total failure of component A is $A_T = A_I + C_{AB} + C_{AC} + C_{ABC}$. Similar expressions can be developed for components B and C.

Now, suppose that the success criteria for the system is 2-out-of-3 for components A, B, and C. Accordingly, the failure of the system can be represented by the following events (cut sets):
$(A_I \cdot B_I)$, $(A_I \cdot C_I)$, $(B_I \cdot C_I)$, C_{AB}, C_{AC}, C_{BC}, C_{ABC}. Thus, the Boolean representation of the system failure will be

$$S = (A_I \cdot B_I) + (A_I \cdot C_I) + (B_I \cdot C_I) + C_{AB} + C_{AC} + C_{BC} + C_{ABC}$$

It is evident that if only independence is assumed, the first three terms of the above Boolean expression are used, and the remaining terms are neglected. Applying the rare event approximation, the system failure probability Q_S is given by

$$Q_S \approx \Pr(A_I) \cdot \Pr(B_I) + \Pr(A_I) \cdot \Pr(C_I) + \Pr(B_I) \cdot \Pr(C_I) + \Pr(C_{AB}) + \Pr(C_{AC}) + \Pr(C_{BC}) + \Pr(C_{ABC})$$

If components A, B, and C are similar (which is often the case since common causes among different components have a much lower probability), then

$$\Pr(A_I) = \Pr(B_I) = \Pr(C_I) = Q_1,$$
$$\Pr(C_{AB}) = \Pr(C_{AC}) = \Pr(C_{BC}) = Q_2, \text{ and}$$
$$\Pr(C_{ABC}) = Q_3$$

Therefore,

$$Q_S = 3(Q_1)^2 + 3Q_2 + Q_3$$

In general, one can introduce the probability Q_k representing the probability of CCF among k specific components in a component group of size m, such that

$1 \le k \le m$. The CCF models for calculating Q_k are summarized in Table 7.3. In this table, Q_t is the total probability of failure accounting both for common cause and independent failures, and α, β, γ, δ, μ, ρ, and ω are the parameters, estimated from the failure data on these components.

Table 7.3 Key Characteristics of the CCF Parametric Models, Mosleh (1991)

Estimation approach	Model	Model parameters	General form for multiple component failure probabilities
Nonshock models single	Beta factor	β	$Q_k = \begin{cases} (1-\beta)\,Q_t & k = 1 \\ 0 & 1 < k < m \\ \beta\,Q_t & k = m \end{cases}$
Nonshock models multi-parameter	Multiple Greek letters	β, γ, δ	$Q_k = \dfrac{1}{\binom{m-1}{k-1}} \left(1 - \rho_{k-1}\right) \left(\prod_{i=1}^{k} \rho_i\right) Q_t$ $k = 1, \cdots, m$ $\rho_1 = 1,\ \rho_2 = \beta, \cdots, \rho_{k-1} = 0$
	Alpha factor	$\alpha_1, \alpha_2, \ldots, \alpha_m$	$Q_k = \dfrac{k}{\binom{m-1}{k-1}} \dfrac{\alpha_k}{\alpha_t} Q_t, \quad k = 1, \cdots, m$ $\alpha_t = \sum_{k=1}^{m} k\,\alpha_k$
Shock models	Binomial failure rate	μ, ρ, ω	$Q_k = \begin{cases} \mu\,\rho^k (1-\rho)^{m-k} & k < m \\ \mu\,\rho^m + \omega & k = m \end{cases}$

CCF parametric models can be divided into two categories: single parameter models and multiple parameter models. The remainder of this section discusses these two categories in more detail as well as elaborates on the parameter estimation of the CCF models.

7.2.1 Single Parameter Models

Single parameter models are those that use one parameter in addition to the total component failure probability to calculate the CCF probabilities. One of the most commonly used single parameter models defined by Fleming (1975) is called the β-factor model. It is the first parametric model applied to CCF events in risk and reliability analysis. The sol10e parameter of the model, β, can be associated with that fraction of the component failure rate that is due to the common cause failures experienced by the other components in the system. That is,

$$\beta = \frac{\lambda_c}{\lambda_c + \lambda_I} = \frac{\lambda_c}{\lambda_t} \tag{7.21}$$

where λ_c is a failure rate due to common cause failures, λ_I is a failure rate due to independent failures, and $\lambda_t = \lambda_c + \lambda_I$.

An important assumption of this model is that whenever a common cause event occurs, all components of a redundant component system fail. In other words, if a CCF shock strikes a redundant system, all components are assumed to fail instantaneously.

Based on the β-factor model, for a system of m components, the probabilities of basic events involving k specific components (Q_k), where $1 \le k \le m$, are equal to zero, except Q_1 and Q_m. These quantities are given as

$$\begin{aligned} Q_1 &= (1 - \beta) Q_t \\ Q_2 &= 0 \\ &\vdots \\ Q_{m-1} &= 0 \\ Q_m &= \beta Q_t \end{aligned}$$

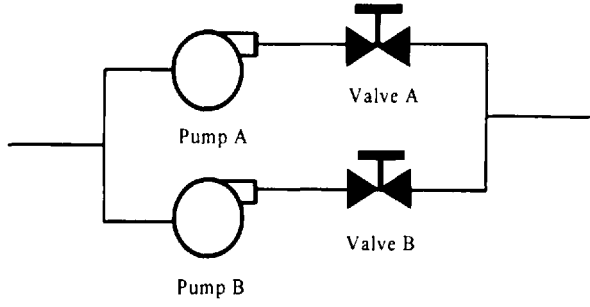
with $m = 1, 2, \ldots$

In general, the estimate for the total component failure rate is generated from generic sources of failure data, while the estimators of the corresponding β-factor do not explicitly depend on generic failure data, but rather rely on specific assumptions concerning data interpretation. The point estimator of β is discussed in Section 7.2.3. Besides, some recommended values of β are given in (Mosleh et al. (1988)). It should be noted that although this model can be used with a certain degree of accuracy for two component redundancy, the results tend to be

conservative for a higher level of redundancy. However, due to its simplicity, this model has been widely used in risk and reliability studies. To get more reasonable results for a higher level of redundancy, more generic parametric models should be used.

Example 7.5

Consider the following system with two redundant trains. Suppose each train is composed of a valve and a pump (each driven by a motor). The pump failure modes are "failure to start" (PS) and "failure to run following a successful star" (PR). The valve failure mode is "failure to open" (VO). Develop an expression for the probability of system failure.



Solution:

Develop a system fault tree to include both independent and common cause failures of the components.

where

P_A is the independent failure of pump A,

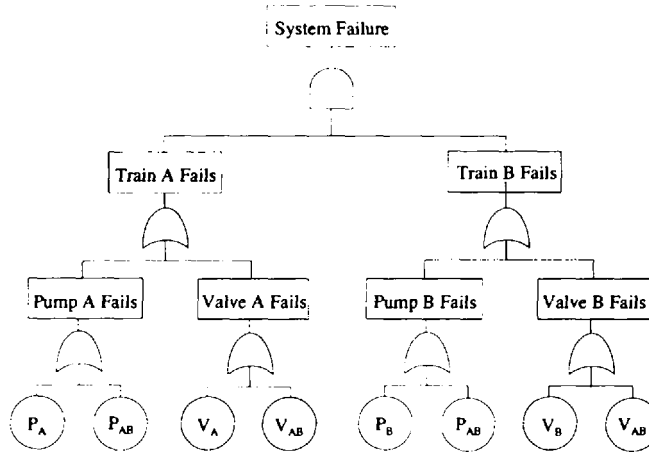
P_B is the independent failure of pump B,

P_{AB} is the dependent failure of pumps A and B,

V_A is the independent failure of valve A,

V_B is the independent failure of valve B,

V_{AB} is the dependent failure of valves A and B.



By solving the fault tree, the following cut sets can be identified:

$$C_1 = (P_A, P_B), \quad C_2 = (P_{AB})$$
$$C_3 = (V_A, V_B), \quad C_4 = (V_{AB})$$
$$C_5 = (P_A, V_B), \quad C_6 = (P_B, V_A)$$

Use the β-factor method to calculate the probability of each cut set.

$$\Pr(C_1) = (1 - \beta_{PS})^2 (q_{PS})^2 + (1 - \beta_{PR})^2 (\lambda_{PR} t)^2$$
$$\Pr(C_2) = \beta_{PS}(q_{PS}) + \beta_{PR}(\lambda_{PR} t)$$
$$\Pr(C_3) = (1 - \beta_{VO})^2 (q_{VO})^2$$
$$\Pr(C_4) = \beta_{VO}(q_{VO})$$
$$\Pr(C_5) = \Pr(C_6) = (q_{PS} + \lambda_{PR} t)(q_{VO})$$

where q is the probability of failure rate on demand, λ is the failure rate to run, and t is mission time.

System failure probability is calculated using rare event approximation, as follows:

$$Q_S \cong \sum_{i=1}^{6} \Pr(C_i)$$

7.2.2 Multiple Parameter Models

Multiple parameter models are used to get a more accurate assessment of CCF probabilities in systems with a higher level of redundancy. These models have several parameters that are usually associated with different event characteristics. This category of models can be further divided into two subcategories, namely, shock and nonshock models. Multiple Greek Letter models and Alpha-Factor models are nonshock models, whereas a Binomial Failure Rate model is a shock model. These models are further discussed below.

Multiple Greek Letter Model

The Multiple Greek Letter (MGL) model introduced by Fleming et al. (1986) is a generalization of the β-factor model. New parameters such as γ, δ, etc., are used in addition to β to distinguish among common cause events affecting different numbers of components in a higher level of redundancy. For a system of m redundant components, $m - 1$ different parameters are defined. For example, for $m = 4$ the model includes the following 3 parameters (see Table 7.3):

Conditional probability that the common cause of failure of an item will be shared by one or more additional items, β ;

Conditional probability that the common cause of an item failure that is shared by one or more items will be shared by two or more items in addition to the first, γ ;

Conditional probability that the common cause of an item failure shared by two or more items will be shared by three or more items in addition to the first, δ.

It should be noted that the β-factor model is a special case of the MGL model in which all other parameters excluding β are equal to 1.

The following estimates of the MGL model parameters are used as generic values:

Number of components (m)	MGL parameters		
	β	γ	δ
2	0.1	X	X
3	0.1	0.27	X
4	0.11	0.42	0.4

Consider the 2-out-of-3 success model described before. If we were to use the MGL model, then equivalent equations for (7.22) for $m = 3$ (see Table 7.3) take the form:

$$Q_1 = \frac{1}{\binom{3-1}{1-1}}\left(1-\rho_{1-1}\right)\left(\prod_{i=1}^{1}\rho_i\right)Q_t = \frac{1}{2}\left(1-\rho_2\right)\left(\rho_1\right)Q_t$$

since $\rho_1 = 1$ and $\rho_2 = \beta$ then

$$Q_1 = \frac{1}{2}\left(1-\beta\right)Q_t$$

Similarly,

$$Q_2 = \frac{1}{2}\left(1-\rho_3\right)\left(\rho_1\rho_2\right)Q_t$$

with $\rho_1 = 1$, $\rho_2 = \beta$ and $\rho_3 = \gamma$,

$$Q_2 = \frac{1}{2}\beta\left(1-\gamma\right)Q_t$$

Also,

$$Q_3 = \frac{1}{2}\left(1-\rho_4\right)\left(\rho_1\rho_2\rho_3\right)Q_t$$

with $\rho_1 = 1$, $\rho_2 = \beta$, $\rho_3 = \gamma$, and $\rho_4 = 0$,

$$Q_3 = \beta\gamma Q_t$$

To compare the result of the β-factor and MGL, consider a case where the total failure probability of each component (accounting for both dependent and independent failures) is 8×10^{-3}. According to the β-factor model, failure probability of the system including common cause failures, if $\beta = 0.1$, would be

$$\begin{aligned} Q_3 &= 3\left(1-\beta\right)^2 Q_t + \beta Q_t \\ &= 3\left(1-0.1\right)^2\left(8E-3\right)^2 + (0.1)(8E-3) \\ &= 9.6E-4 \end{aligned}$$

However, MGL model with $\beta = 0.1$ and $\gamma = 0.27$ will predict the system failure probability as

$$Q_s = \frac{3}{4}(1-\beta)^2 Q_t^2 + \frac{3}{2}\beta(1-\gamma)^2 Q_t + \beta\gamma Q_t$$

$$= \frac{3}{4}(1-0.1)^2(8E-3)^2$$

$$+ \frac{3}{2}\,0.1(1-0.27)(8E-3) + (0.1)(0.27)(8E-3)$$

$$= 1.1E-3$$

The difference is obviously small, but the MGL model is more accurate than the β-factor model.

Alpha Factor Model

The α-factor model discussed by Mosleh and Siu (1987) develops CCF failure probabilities from a set of failure ratios and the total component failure rate. The parameters of the model are the fractions of the total probability of failure in the system that involves the failure of k components due to a common cause, a_k .

The probability of a common cause basic event involving failure of k components in a system of m components is calculated according to the equation given in Table 7.3. For example, the probabilities of the basic events of the three-component system described earlier will be

$$Q_1 = (\alpha_1 / \alpha_t)\, Q_t$$
$$Q_2 = (\alpha_2 / \alpha_t)\, Q_t$$
$$Q_3 = (3\alpha_3 / \alpha_t)\, Q_t$$

where $\alpha_t = \alpha_1 + 2\alpha_2 + 3\alpha_3$. The table below (Mosleh (1991)) provides generic values of α-factors.

Number of items (m)	α-Factor			
	α_1	α_2	α_3	α_4
2	0.95	0.05	—	—
3	0.95	0.04	0.01	—
4	0.95	0.035	0.01	0.005

Therefore, the system failure probability for the three redundant components discussed earlier can now be written as

$$Q_s = 3\left(\frac{\alpha_1}{\alpha_t} Q_t\right)^2 + 3\left(\frac{\alpha_2}{\alpha_t} Q_t\right) + 3\left(\frac{\alpha_3}{\alpha_t} Q_t\right)$$

Accordingly, using the generic α values for the 2-out-of-3 success $\alpha_t = 0.95 + 0.08 + 0.03 = 1.06$. Thus,

$$Qs = 3\left[\frac{0.95}{1.06}(8\times10^{-3})\right]^2 + 3\left[\frac{0.04}{1.06}(8\times10^{-3})\right]^2 + 3\left[\frac{0.01}{1.06}(8\times10^{-3})\right]^2$$

$$= 1.3\times10^{-3}$$

which is closely consistent with the MGL model results.

Binomial Failure Rate Model

The binomial failure rate (BFR) model discussed by Atwood (1983), unlike the α-factor model and MGL model, is a shock dependent model. It estimates the failure frequency of two or more components in a redundant system as the product of the CCF shock arrival rate and the conditional failure probability of components given the shock has occurred. This model considers two types of shock: lethal and nonlethal. The assumption is that, given a nonlethal shock, components fail independently, each with a probability of ρ, whereas in the case of a lethal shock, all components fail with a probability of 1. The expansion of this model is called the *Multinomial Failure Rate* (MFR) model. In this model, the conditional probability of failure of k components is calculated directly from component failure data without any further assumptions. Therefore, the MFR model becomes essentially the same as the nonshock models, because the separation of the CCF frequency into the shock arrival rate and conditional probability of failure given shock has occurred is, in general, a statistical rather than a physical modeling step. The parameters of the BFR model generally include:

Nonlethal shock arrival rate, μ ;
Conditional probability of failure of each component given the occurrence of a nonlethal shock, ρ ;
Lethal shock arrival rate, ω .

It should be noted that due to the BFR model complexity and the lack of data to estimate its parameters, it is not widely used in practice.

7.2.3 Data Analysis for Common Cause Failures

Despite the difference among the models described in Section 7.2.1, they all have similar data requirements in terms of parameter estimation. One should not expect

Table 7.4 Simple Point Estimators for Various Parametric Models

Model	Point estimator
Beta-factor	$\hat{Q}_t = \frac{1}{mN_D}\sum_{i=1}^{m} kn_k$ $\hat{\beta} = \left(\sum_{i=2}^{m} kn_k\right) \Big/ \left(\sum_{i=1}^{m} kn_k\right)$
Multiple Greek letters	$\hat{Q}_t = \frac{1}{mN_D}\sum_{i=1}^{m} kn_k$ $\hat{\beta} = \left(\sum_{i=2}^{m} kn_k\right) \Big/ \left(\sum_{i=1}^{m} kn_k\right)$ $\hat{\gamma} = \left(\sum_{i=3}^{m} kn_k\right) \Big/ \left(\sum_{i=2}^{m} kn_k\right)$ $\hat{\delta} = \left(\sum_{i=4}^{m} kn_k\right) \Big/ \left(\sum_{i=3}^{m} kn_k\right)$
Alpha-factor	$\hat{Q}_t = \frac{1}{mN_D}\sum_{i=1}^{m} kn_k$ $\hat{\alpha}_k = n_k \Big/ \left(\sum_{i=1}^{m} kn_k\right) \qquad (k = 1, \ldots, m)$

any significant difference among the numerical results provided by these models. The relative difference in the results may be attributed to the statistical aspects of the parameter estimation, which has to do with the assumptions made in developing a parameter estimator and the dependencies assumed in CCF probability quantification.

The most important steps in the quantification of CCFs are collecting information from the raw data and selecting a model that can use most of this information. Statistical estimation procedures discussed in Chapters 2 and 3 can be applied to estimate the CCF model parameters. If separate models rely on the same type of information in estimating the CCF probabilities, and similar assumptions regarding the mechanism of CCFs are used, comparable numerical results can be expected. Table 7.4 summarizes simple point estimators for parameters of various nonshock CCF models. In this table, n_k is the total number of observed failure events involving failure of k similar components due to a common cause, m is the total number of redundant items considered; and N_D is the total number of system demands. If the item is normally operating (not on a standby), then N_D can be replaced by the total test (operation) time T. The estimators in Table 7.4 are based on the assumption that in every system demand, all components and possible combination of components are challenged. Therefore, the estimators apply to systems whose tests are nonstaggered.

Example 7.6

For the system described in Example 7.5, estimate the β parameters, λ and q, for the valves and pumps based on the following failure data:

Failure mode	Event statistic		
	n_1	n_2	T (hr) or N_D
Pump fails to start (*PS*)	10	1	500 (demands)
Pump fails to run (*PR*)	50	2	10,000 (hours)
Valve fails to open (*VO*)	10	1	10,000 (demands)

In the above table, n_1 is the number of observed independent failures, n_2 is the number of observed events involving double CCF. Calculate the system unreliability for a mission of 10 hours.

Solution:

From Table 7.4,

$$\beta = \frac{2n_2}{n_1 + 2n_2}$$

Apply this formula to β_{PR}, β_{PS}, and β_{VO}, by using appropriate values for n_1 and n_2.

$$n_{PS} = n_1 + 2n_2 = 12$$
$$n_{PR} = n_1 + 2n_2 = 54$$
$$n_{VO} = n_1 + 2n_2 = 17$$

Accordingly, use (3.62) and (3.77) for estimating λ and q, respectively

$$q_{PS} = \frac{12}{500} = 2.4E\text{-}2\,D^{-1}, \quad \beta_{PS} = \frac{7}{12} = 0.17$$

$$\lambda_{PR} = \frac{54}{10{,}000} = 5.4E\text{-}3\ \text{hr}^{-1}, \quad \beta_{PR} = \frac{4}{54} = 0.07$$

$$q_{VO} = \frac{17}{10{,}000} = 1.7E\text{-}3\,D^{-1}, \quad \beta_{VO} = \frac{2}{17} = 0.12$$

Therefore, using the cut set probability equations developed in Example 7.5, the estimates of the failure probabilities at 10 hours of operation for each cut set are

$$\Pr(C_1) = (1 - 0.17)^2 (2.4E\text{-}2)^2$$
$$+ (1 - 0.07)^2 (5.4E\text{-}3 \times 10)^2 = 2.9E\text{-}3$$
$$\Pr(C_2) = (0.17)(2.4E\text{-}2) + (0.07)(5.4E\text{-}3)(10) = 7.9E\text{-}3$$
$$\Pr(C_3) = (1 - 0.12)^2 (1.7E\text{-}3)^2 = 2.2E\text{-}6$$
$$\Pr(C_4) = (0.12)(1.7E\text{-}3) = 2.0E\text{-}4$$
$$\Pr(C_5) = [2.4E\text{-}2 + (5.4E\text{-}3)(10)](1.7E\text{-}3) = 1.3E\text{-}4$$
$$\Pr(C_6) = \Pr(C_5) = 1.39E\text{-}2$$

Thus, the system failure probability is

$$Q_s \cong \sum_{i=1}^{6} \Pr(C_i) = 1.1E\text{-}2$$

7.3 UNCERTAINTY ANALYSIS

Uncertainty arises primarily due to lack of reliable information, e.g., lack of information about the ways a given system may fail. Uncertainty may also arise due to

linguistic imprecision, e.g., the expression "System A is highly reliable." Furthermore, uncertainty may be divided into two kinds: the *aleatory* models of the world and *epistemic* uncertainty. For example, the Poisson model for modeling the inherent randomness in the occurrence of an event (e.g., failure event) can be considered the "world model" of the occurrence of failure. The variability associated with the results obtained from this model represents the aleatory uncertainty. The epistemic uncertainty, on the other hand, describes our state of knowledge about this model. For example, the uncertainty associated with the choice of the Poisson model itself and its parameter λ is considered epistemic.

Consider a Weibull distribution used to represent the time to failure. The choice of the distribution model itself involves some modeling uncertainty (epistemic); however, the variability of time-to-failure is the aleatory uncertainty. We may even be uncertain about the way we construct the failure model. For example, our uncertainty about parameters α and β of the Weibull distribution representing time to failure distribution may be depicted by another distribution, e.g., a lognormal distribution. In this case, the lognormal distribution models represent the epistemic uncertainty about the Weibull distribution model.

The most common practice in measuring uncertainty is the use of the probability concept. In this book, we have only used this measure of uncertainty. As we discussed in Chapter 2, there are different interpretations of probability. This also affects the way uncertainty analysis is performed. In this section, we first briefly discuss uncertainty in choice of models and then present methods of measuring the uncertainty about the parameters of the model. Then we discuss methods of propagating uncertainty in a complex model. For example, in a fault tree model representing a complex system, the uncertainty assigned to each leaf of the tree can be propagated to obtain a distribution of the top event probability.

The simplest way to measure uncertainty is to use sample mean $\bar{x}$ and variance S^2, described by (2.81) and (2.83). We have discussed earlier in Chapter 2 that estimations of $\bar{x}$ and S^2 are themselves subject to some uncertainty, it is important to describe this uncertainty by confidence intervals of $\bar{x}$ and S^2, e.g., by using (2.90). This brings another level of uncertainty. The confidence intervals associated with different types of distributions were discussed in Chapter 3. For a binomial model, the confidence intervals can be obtained from (3.78) and (3.79). Similarly, if the data are insufficient, then the subjectivist definition of probability can be used and different Bayesian probability intervals can be obtained (see Section 3.6).

Generally, the problem of finding the distribution of a function of random variables is difficult, which is why for most of the reliability and risk assessment applications, the problem is reduced to estimation of mean and variance (or standard deviation) of function of random variables. Such techniques are considered in the following sections. It should be mentioned that the uses of these techniques

are, by no mean, limited to reliability and risk assessment problems. They are widely used in engineering.

7.3.1 Types of Uncertainty

Because different types of uncertainties are generally characterized and treated differently, it is useful to identify three types of uncertainty: *parameter uncertainty*, *model uncertainty*, and *completeness uncertainty*.

Parameter Uncertainties

Parameter uncertainties are those associated with the values of the fundamental parameters of the reliability or risk model, such as failure rates, event probabilities including human error probabilities etc. They are typically characterized by establishing probability distributions on the parameter values.

Parameter uncertainties can be explicitly represented and propagated through the reliability or risk model, and the probability distribution of the relevant metrics (e.g., reliability, unavailability, risk) can be generated. Various measures of central tendency, such as the mean, median and mode, can be evaluated. For example, the distribution can be used to assess the confidence with which reliability targets are met. The results are also useful to study the contributions from various elements of a model and to see whether it can be determined that the tails of the distributions are being determined by uncertainties on a few significant elements of the reliability or risk model. If so, these elements can be identified as candidates for compensatory measures and/or monitoring.

In Chapter 3, we discussed measures for quantifying uncertainties of parameter values of distribution models for both the frequentist and subjectivist (Bayesian) methods. Examples of these parameters are MTTF, μ, failure rate, λ, and probability of failure on demand, p, of a component. Uncertainty of the parameters is primarily governed by the amount of field data available about failures and repairs of the items. Because of these factors, a parameter does not take a fixed and known value, and has some random variability. In Section 7.3.3, we discuss how the parameter uncertainty is propagated in a system to obtain an overall uncertainty about the system failure.

Model Uncertainties

There are also uncertainties as to how to model specific elements of the reliability or risk. Model uncertainty may be analyzed in different ways. It is possible to include some model uncertainty by incorporating with the reliability/risk model a discrete probability distribution over a set of models for a particular issue (e.g., various models for reliability growths or human reliability). In principle,

uncertainty in choosing a model can be handled in the same way as parameter uncertainty. For example, if a set of candidate models are available, one could construct a discrete probability distribution (M_i, p_i), where p_i is the degree of belief (in subjectivist terms) in model M_i as being the most appropriate representation. This has been done for the modeling of a seismic hazard, for example, where the result is a discrete probability distribution on the frequencies of earthquakes. This uncertainty can then be propagated in the same way as the parameter uncertainties. Other methods are also available. For example, see Mosleh et al. (1995).

It is often instructive to understand the impact of a specific assumption on the prediction of the model. The impact of using alternate assumptions or models may be addressed by performing appropriate sensitivity studies, or they may be addressed using qualitative arguments. This may be a part of the model uncertainty evaluation.

There are two aspects of modeling uncertainty at the component level or system level. In estimating uncertainty associated with unreliability or unavailability of a basic component, a modeling error can occur as a result of using an incorrect distribution model. Generally, it is very difficult to estimate an uncertainty measure for these cases. However, in a classical (frequentist) approach, the confidence level associated with a goodness-of-fit test can be used as a measure of uncertainty. For the reliability analysis of a system, one can say that a model describes the behavior of a system as viewed by the analyst. However, the analyst can make mistakes due to a number of constraints, namely, his degree of knowledge and understanding of the system design and his assumptions about the system, as reflected in the reliability model (e.g., a fault tree).

Clearly one can minimize these sources of uncertainty, but one cannot eliminate them. For example, a fault tree based on the analyst's understanding of the success criteria of the system can be incorrect, if the success criteria used are in error. For this reason, a more accurate dynamic analysis of the system may be needed to obtain correct success criteria.

Definition and quantification of the uncertainty associated with a model are very complex and cannot easily be associated with a quantitative representation (e.g., probabilistic representation). The readers are referred to Morgan and Henrion (1990) for more discussion on this topic.

Completeness Uncertainty

Completeness is not in itself an uncertainty, but a reflection of scope of reliability and risk analysis limitations. The result is, however, an uncertainty about where the true reliability or risk lies. The problem with completeness uncertainty is that, because it reflects unanalyzed contributions (e.g., contribution due to exclusion of certain failure modes in a fault tree analysis), it is difficult (if not impossible) to estimate the uncertainty magnitude. Thus, for example, the impact

on actual reliability/risk from unanalyzed issues such as the influences of organization factor on equipment performance (e.g., reliability) quality assurance cannot be explicitly assessed.

7.3.2 Uncertainty Propagation Methods

Consider a general case of a system performance characteristic Y (e.g., system reliability or unavailability). Based on an aleatory model of the system, a general function of uncertain quantities x_i and uncertain parameters θ_i can describe this system performance characteristic as

$$Y = f(x_1, x_2, \ldots, x_n, \theta_1, \theta_2, \ldots, \theta_m) \tag{7.23}$$

A simple example is a system composed of elements having the exponential time-to-failure distributions. In this case, Y can be the MTTF of the system, x_i ($i = 1, 2, \ldots, n$) are the estimates of MTTFs of the system components, and θ_i ($i = 1, 2, \ldots, m$) are the standard deviations (errors) of these estimates. System performance characteristic, Y, can also be the probability of the top event of a fault tree, in which case x_i will be the failure probability (unavailability) of each component represented in the fault tree, and θ_is will be the parameters of the distribution models representing x_i.

The variability of Y as a result of the variability of the basic parameters x_i and θ_i is estimated by the methods of propagation. We will discuss these methods below.

Method of Moments

Write the function (7.23) in the following form:

$$Y = f(x_1, x_2, \ldots, x_n; S_1, S_2, \ldots, S_n) \tag{7.24}$$

where $x_i (i = 1, 2, \ldots, n)$ are the estimates of reliability parameters (e.g., MTTF, failure rate, probability of failure on demand, etc.) of system component, and $S_i (i = 1, 2, \ldots, n)$ are the respective standard deviations (errors).

Assume that:

$f(x_1, x_2, \ldots x_n; S_1, S_2, \ldots, S_n) \equiv f(X, S)$ satisfies the conditions of Taylor's theorem

the estimates $x_i (i = 1, 2, \ldots, n)$ are independent and unbiased with expectations (true values) $\mu_i (i = 1, 2, \ldots, n)$.

Using the Taylor's series expansion about μ_i, and denoting $(x_1, x_2, \ldots, x_n)$ by X and $(S_1, S_2, \ldots, S_n)$ by S, we can write:

$$
\begin{aligned}
Y &= f(X;S) \\
&= f(\mu_1:\mu_2,\ldots,\mu_n;S) + \sum_{i=1}^{n}\left[\frac{\partial f(X)}{\partial X}\right]_{x_i=\mu_i}(x_i-\mu_i) + \\
&\quad \frac{1}{2!}\sum_{j=1}^{n}\sum_{i=1}^{n}\left[\frac{\partial^2 f(X)}{\partial x_i \partial x_j}\right]_{x_i=\mu_i, x_j=\mu_j}(x_i-\mu_i)(x_i-\mu_j) + R
\end{aligned}
\tag{7.25}
$$

where R represents the residual terms.

Taking the expectation of (7.25) (using the algebra of expectations given in Table 2.2), one gets

$$
\begin{aligned}
E(Y) &= f(\mu_1,\mu_2,\ldots,\mu_n;S) + \sum_{i=1}^{n}\left[\frac{\partial f(X)}{\partial X}\right]_{x_i=\mu_i}E(x_i-\mu_i) + \\
&\quad \frac{1}{2!}\sum_{j=1}^{n}\sum_{i=1}^{n}\left[\frac{\partial^2 f(X)}{\partial f x_i \partial f x_j}\right]_{x_i=\mu_i, x_j=\mu_j}E\left[(x_i-\mu_i)(x_i-\mu_j)\right] + E(R)
\end{aligned}
\tag{7.26}
$$

Because the estimates $x_i (i = 1, 2, \ldots, n)$ are unbiased with expectations (true values) μ_i, the second term in the above equation is canceled. Dropping the residual term, $E(R)$, and assuming that the estimates x_i are independent, one gets the following approximation:

$$
E(Y) \approx f(\mu_1,\mu_2,\ldots\mu_n;S) + \frac{1}{2}\sum_{i=1}^{n}\left[\frac{\partial^2 f(X)}{\partial x_i^2}\right]_{x_i=\mu_i}S^2(x_i) \tag{7.27}
$$

For the more general and practical applications of the method of moments, we need to get the point estimate $\hat{Y}$ and its variance $\mathrm{var}(\hat{Y})$. Replacing μ_i by x_i $(i = 1, 2, \ldots, n)$, we get

$$
\hat{Y} \approx f(x_1, x_2, \ldots, x_n; S) + \frac{1}{2}\sum_{i=1}^{n}\left[\frac{\partial^2 f(X)}{\partial x_i^2}\right]_{x=x_i}S^2(x_i) \tag{7.28}
$$

If, for a given uncertainty analysis problem, the second term can be neglected the estimate (7.28) is reduced to the following simple form, which can be used as the point estimate:

$$
\hat{Y} \approx f(x_1, x_2, \ldots, x_n) \tag{7.29}
$$

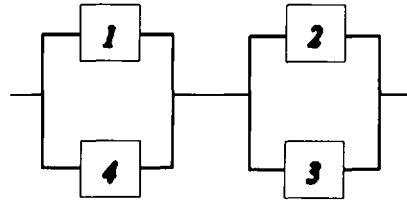
To get a simple approximation for the variance (as a measure of uncertainty) of the system performance characteristic Y, consider the first two-term

approximation for (7.25). Taking the variance and treating the first term as constant, one gets

$$\begin{aligned} \mathrm{var}(\hat{Y}) &= \mathrm{var}\left\{\sum_{i=1}^{n}\left[\frac{\partial f(X)}{\partial x_i}\right]_{x_i=\mu_i}\left(x_i-\mu_i\right)\right\} \\ &= \sum_{i=1}^{n}\left[\frac{\partial f(X)}{\partial x_i}\right]^2_{x_i=\mu_i} S^2(x_i) \end{aligned} \tag{7.30}$$

Example 7.7

For the system shown below, the constant failure rate of each component has a mean value of 5 10^{-3} hr^{-1}. If the failure rate can be represented by a r.v. which follows a lognormal distribution with a coefficient of variation of 2, calculate the mean and standard derivation of the system unreliability at t = 1, 10, and, 100 hours.



Solution:

System unreliability can be obtained from the following expression

$$Q = q_1 \times q_4 + q_3 \times q_2 - q_1 \times q_2 \times q_3 \times q_4$$

since $q_i = 1 - e^{-\lambda t}$, then

$$\begin{aligned} Q = &\left(1-e^{-\lambda_1 t}\right)\left(1-e^{-\lambda_2 t}\right)+\left(1-e^{-\lambda_3 t}\right)\left(1-e^{-\lambda_4 t}\right)- \\ &\left(1-e^{-\lambda_1 t}\right)\left(1-e^{-\lambda_2 t}\right)\left(1-e^{-\lambda_3 t}\right)\left(1-e^{-\lambda_4 t}\right) \end{aligned}$$

note that $\hat{\lambda}_1 = \hat{\lambda}_2 = \hat{\lambda}_3 = \hat{\lambda}_4 = \hat{\lambda}$ =5E - 3 hr^{-1}. Using (7.28) and neglecting the second term (due to its insignificance):

$$\hat{Q} = \left(1 - e^{-\hat{\lambda}_1 t}\right)\left(1 - e^{-\hat{\lambda}_2 t}\right) + \left(1 - e^{-\hat{\lambda}_3 t}\right)\left(1 - e^{-\hat{\lambda}_4 t}\right) - \left(1 - e^{-\hat{\lambda}_1 t}\right)\left(1 - e^{-\hat{\lambda}_2 t}\right)\left(1 - e^{-\hat{\lambda}_3 t}\right)\left(1 - e^{-\hat{\lambda}_4 t}\right)$$

$$\hat{Q} = 2\left(1 - e^{-\hat{\lambda} t}\right)^2 - \left(1 - e^{-\hat{\lambda} t}\right)^4 = 1 - 4e^{-2\hat{\lambda} t} + 4e^{-3\hat{\lambda} t} - e^{-4\hat{\lambda} t}$$

$$\hat{Q}_{1\text{ hour}} = 4.97 \times 10^{-5}$$

$$\hat{Q}_{10\text{ hours}} = 4.75 \times 10^{-3}$$

$$\hat{Q}_{100\text{ hours}} = 0.286$$

Calculate the derivatives.

For example,

$$\frac{\partial Q}{\partial \lambda_1} = \lambda_1 e^{-\lambda_1 t}\left(1 - e^{-\lambda_2 t}\right) - \lambda_1 e^{-\lambda_1 t}\left(1 - e^{-\lambda_2 t}\right)\left(1 - e^{-\lambda_3 t}\right)\left(1 - e^{-\lambda_4 t}\right)$$

Repeating for other derivatives of Q with respect to λ_2, λ_3, and λ_4, yields

$$\frac{\partial Q}{\partial \lambda_i} = \sum_{i=1}^{4} \lambda_i e^{-\lambda_i t}\left[\left(1 - e^{-\lambda t}\right) - \left(1 - e^{-\lambda t}\right)^3\right]$$

and by (7.30)

$$S^2(Q) = 4\sum_{i=1}^{4} \mathrm{var}(\lambda_i)\left\{\hat{\lambda}_i e^{-\hat{\lambda}_i t}\left[\left(1 - e^{-\hat{\lambda} t}\right) - \left(1 - e^{-\hat{\lambda} t}\right)^3\right]\right\}^2$$

$$S^2(Q) = 4\sum_{i=1}^{4} (2\hat{\lambda})^2\left\{\hat{\lambda}_i e^{-\hat{\lambda}_i t}\left[\left(1 - e^{-\hat{\lambda} t}\right) - \left(1 - e^{-\hat{\lambda} t}\right)^3\right]\right\}^2$$

$$S^2(Q)_{1\text{ hour}} = 2.51 \times 10^{-13}$$

$$S^2(Q)_{10\text{ hours}} = 2.14 \times 10^{-11}$$

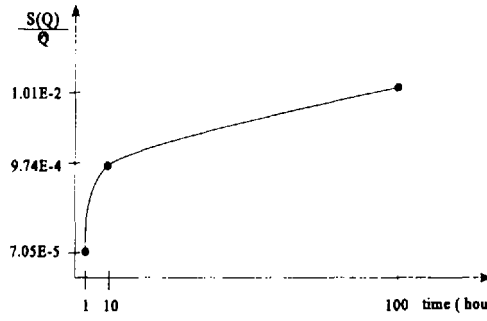
$$S^2(Q)_{100\text{ hours}} = 4.07 \times 10^{-10}$$

Using $S(\lambda_i) = 2 \times \hat{\lambda} = 2 \times 5E-3 = 0.01$. Therefore, $\mathrm{var}(\lambda_i) = S2(\lambda_i) = 10^{-4}$. It is now possible to calculate coefficient of variation for system unreliability as

$$\left.\frac{S(Q)}{\hat{Q}}\right|_{1\text{ hour}} = 1.01E-2$$

$$\left.\frac{S(Q)}{\hat{Q}}\right|_{10\text{ hours}} = 9.74E-4$$

$$\left.\frac{S(Q)}{\hat{Q}}\right|_{100\text{ hours}} = 7.05E-5$$



For more detailed consideration of the reliability applications of the method of moments, the reader is referred to (Morchland and Weber (1972)). Apostolakis and Lee (1977) propagate the uncertainty associated with parameters x_i by generating lower order moments, such as the mean and variance for Y, from the lower order moments of the distribution for x_i. A detailed treatment of this method is covered in a comparison study of the uncertainty analysis method by Martz (1983).

For a special case when $Y = \sum_{i=1}^{n} x_i$ (for example, a series composed of components having the exponential time-to-failure distributions with failure rates x_i), and dependent x_is, the variance of $\hat{Y}$ is given by

$$\text{var}[\hat{Y}] = \sum_{i=1}^{n} \text{var}[x_i] + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} \text{cov}[x_i, x_j] \tag{7.31}$$

In the case where $Y = \prod_{i=1}^{n} x_i$, and x_is are independent (a series system composed of components having reliability functions, x_i ($i = 1, 2, \ldots, n$))

$$E(Y) = \prod_{i=1}^{n} E(x_i)$$

and

$$\text{var}[Y] \approx \left[\sum_{i=1}^{n} \frac{\text{var}(x_i)}{E^2(x_i)}\right] \times E^2(Y)$$

Dezfuli and Modarres (1984) have expanded this approach to efficiently estimate a distribution fit for Y when x_is are highly dependent. The method of

moments provides a quick and accurate estimation of lower moments of Y based on the moments of x_i, and the process is simple. However, for highly nonlinear expressions of Y, the use of only low-order moments can lead to significant inaccuracies, and the use of higher moments is complex.

7.3.3 System Reliability Confidence Limits Based on Component Failure Data

Estimation of system reliability, usually, is associated with system component models uncertainties. In this section, we consider some practical approaches to eliminating this type of uncertainty for series systems.

Lloyd–Lipow Method

Consider a series system composed of m different components. Let $p_i (i = 1, 2, \ldots, m)$ be the respective component failure probabilities. They can be treated as $F_i(t)$, i.e., the time-to-failure cdfs at a given time t, for the time-dependent reliability models. Similarly, they can be the time-independent failure probabilities (the binomial model), for example, the probabilities of failure on demand.

The reliability of the system, R_s, is given by

$$R_s = \prod_{i=1}^{m} (1 - p_i)$$

The probabilities, p_i, are not known but can be estimated. The respective estimates are obtained based on component tests or field data. In the following we consider methods of system point and confidence reliability estimation, based on straightforward use of component tests' data, i.e., without estimating the components reliability characteristics.

We start with the *Lindstrom–Madden method* which is more frequently referred to as the *Lloyd–Lipow method,* due to the book by Lloyd and Lipow (1962), where the method was first described. Note that the Lindstrom–Madden method is a heuristic one.

To simplify our consideration, let's limit ourselves by the case of a two-component series system. Assume that the test results for the components are given in the following form:

N_1 is the number of the first components tested, and d_1 is the number of failures observed during the test

N_2 is the number of the second components tested, and d_2 is the number of respective failures observed during the test.

Without loss of generality, suppose that $N_2 > N_1$. These test results can be represented by the following two sets:

$$x_{1i} \ (i = 1, 2, \ldots, N_1)$$
$$x_{2i} \ (i = 1, 2, \ldots, N_2)$$

where x_{1i} and x_{2i} take on the value 1, if the respective component failed during the test and they take zero values if the respective component did not fail during the test. Let us have d_1 survived units among N_1 first components tested, and d_2 survived units among N_2 second components tested.

Select randomly N_1 elements from the set x_{2i}. Randomly combining each of these elements with elements from the set x_{2i} $(i = 1, 2, \ldots, N_2)$, obtain N_1 pairs (x_{1j}, x_{2j}) with $j = 1, 2, \ldots, N_1$. The idea of the Lindstrom–Madden method is to treat these pairs as fictitious test results of N_1 series systems composed of the first and the second components. Expected number of the fictitious series systems failed (i.e., having at least one component failed), D_s, is given by

$$D_s = N_1\left(1 - \hat{R}_S\right) \tag{7.32}$$

where

$$\hat{R}_s = \left(1 - \frac{d_1}{N_1}\right)\left(1 - \frac{d_2}{N_2}\right)$$

is the point estimate of the series system reliability function. The value of D_s is considered as "equivalent" number of failures for a sample of N_1 series systems of interest (Ushakov (1994)). Note that, similar to Bayes' approach, D_s is not necessarily an integer.

To get confidence limits for the system reliability, one needs to use the Clopper–Pearson procedure, considered in Chapter 3.

In general, the case of a system composed of k components, the expected number of the fictitious series systems failed, D_s, is given by:

$$D_s = N_{1s}\left(1 - \hat{R}_s\right) \tag{7.33}$$

where $N_{1s} = \min(N_1, N_2, \ldots, N_k)$ and

$$\hat{R}_s = \prod_{i=1}^{k}\left(1 - \frac{d_i}{N_i}\right) \tag{7.34}$$

Based on D_s and N_{1s} the respective confidence limits for the system reliability, are constructed in a similar way, using the Clopper–Pearson procedure.

Example 7.8

Two components were tested under the following time-terminated test plans. A sample of 110 units of the first component was tested during 2000 hours. The failures were observed at: 3, 7, 58, 145, 155, 273, 577, 1104, 1709, and 1999 hours. A sample of 100 units of the second component was tested during 1000 hours. The failures were observed at: 50, 70, 216, 235, 295, 349, 368, and 808 hours. Find the point estimate and 90% lower confidence limit for the reliability function at 1000 hours for the two-component series system composed of these components.

Solution:

Find the number of the series systems "tested" as

$$N_{1,} = \min(N_1, N_2) = \min(110, 100) = 100$$

For the 1000 hour interval we have $d_1 = 7$ and $d_2 = 8$.

Using Equation (7.33) find

$$\hat{R}_{,}(1000) = \left(1 - \frac{7}{110}\right)\left(1 - \frac{8}{100}\right) = 0.861$$

$$D_{,} = 100(1 - 0.861) = 13.9$$

Using the Clopper–Pearson procedure in the form (3.85) with $n = 100$ and $r = 13.9$ find the 90% lower confidence limit for the system reliability function at 1000 hours, $R_l(1000)$, as a solution of the following equation

$$I_{R_l}(100 - 13.9, 13.9 + 1) \leq 0.1$$

which gives $R_l(1000) = 0.806$.

Note that the solution of the problem does not depend on particular time-to-failure distributions of the components, which shows that Lindstrom–Madden is nonparametric.

7.3.4 Maximus Method

As mentioned, the Lindstrom–Madden method considered in the previous section can be applied to a series system only. The Maximus method, we briefly discuss

below, is a generalization of the Lindstrom–Madden method for series-parallel arrangement of subsystems of components (Martz and Duran (1985)).

Under this method, the basic steps for constructing the lower confidence limit for a system reliability, based on component failure data are:

1. Reduce each subsystem to an equivalent component. Treat the components of the reduced system as each having its equivalent failure data obtained from the reduction performed.
2. Obtain the maximum likelihood point estimate of system reliability, $\hat{R}_s$, based on the system configuration and component equivalent failure data.
3. Calculate the equivalent system sample size, N_S, according to the reduced system configuration and the respective equivalent component failure data from step 1.
4. Calculate the equivalent number of system failures, D_S, as

$$D_s = N_S\left(1 - \hat{R}_S\right) \tag{7.35}$$

Note that the above equation coincides with Equation (7.34).

5. Using the Clopper–Pearson procedure (Equation (3.85)) with N_s, D_s and a chosen confidence probability, calculate the lower confidence limit for the system reliability.

Classical Monte Carlo Simulation

There are three techniques for system reliability confidence estimation based on Monte Carlo simulation: classical Monte Carlo simulation, bootstrap method, and Bayes' Monte Carlo method.

The classical Monte Carlo method is based on classical component probabilistic models (failure distributions) which are obtained using failure data only. In other words, each component of the system analyzed is provided with a failure (time-to-failure or failure on demand) distribution, fitted using real failure data.

If we knew the *exact* values of the reliability characteristic of the system components, we would be able, in principal, to calculate the system reliability using the system reliability function, e.g., using equations (4.1) and (4.7). Instead of exact component reliability characteristics we deal with their estimates which are random variables. Thus, if there are no failure data for the system as a whole, we have to treat any system reliability characteristic as a random variables transformation result, obtained using the system reliability function. As mentioned in Section .3.1, generally, it is not easy to find the distribution of the transformed random

variables, which is why the Monte Carlo approach turns out to be a practical tool for solving many problems associated with complex system reliability estimation.

In the framework of the classical Monte Carlo approach, there could be different algorithms for system reliability estimation. The following example illustrates the general steps for constructing the lower confidence limit for system reliability using this method. These steps are:

1. For each component of the system given, obtain a classical estimate (e.g., the maximum likelihood estimate) of component reliability, R_i, ($i = 1, 2, \ldots, n$, where n is the number of component in the system) generating it from the respective estimate distribution.
2. Calculate the corresponding classical estimate of the system reliability

$$\hat{R}_S = f(R_1, R_2, \ldots, R_n) \tag{7.36}$$

where $f(\cdot)$ is the system reliability function.

3. Repeat steps 1–2 a sufficiently large number of times, n, (for example, 10,000) to get a large sample of $\hat{R}_S$.
4. Using the sample obtained, and a chosen confidence level $(1 - \alpha)$, construct the respective lower confidence limit for the system reliability of interest as a sample percentile of level α (discussed in Section 7.1):

$$\hat{R}_{S_p} = \begin{cases} \hat{R}_{S(np)}, & \text{if } np \text{ is not integer and} \\ \text{any value from the interval } \left[\hat{R}_{S(np)}, \hat{R}_{S(np+1)}\right], & \text{if } np \text{ is integer} \end{cases} \tag{7.37}$$

Bayes' Monte Carlo Simulation

The principal and the only difference between the classical Monte Carlo approach and the Bayesian, is related to component reliability estimation. Under the Bayes' approach, we need to provide prior information and respective prior distribution for each unique component in the given system. Then we need to get the corresponding posterior distributions. Having these distributions obtained, the same steps as under the classical Monte Carlo approach are performed.

In the absence of prior information about reliability of the system components and binomial data with moderate sample size, Martz and Duran (1985) recommend using the beta distribution having parameters 0.5 and 0.5, as an

appropriate prior distribution, which they call *noninformative* prior. Note that such noninformative prior has the mean 0.5 and the coefficient of variation which is very closed to the coefficient of variation of the standard uniform distribution (0, 1). Also recall that the standard uniform distribution is a particular case of the beta distribution with parameters 1 and 1 (see Section 2.3).

Bootstrap Method

The bootstrap method introduced by Efron in 1979 is a Monte Carlo simulation technique in which new samples are generated from the data of an original sample. The method's name, derived from the old saying about pulling yourself up by your own bootstraps, reflects the fact that one available sample gives rise to many others.

Unlike the classical and Bayes' Monte Carlo techniques, the bootstrap method is a universal nonparametric method. To illustrate the basic idea of this method, consider the following simple example, in which the standard error of a median is estimated (Efron and Tibshirani (1993)).

Consider an original sample, $x_1, x_2, \ldots, x_n$, from an unknown distribution. The respective *bootstrap sample*, $x_1^b, x_2^b, \ldots, x_n^b = X^b$, is obtained by randomly sampling n times with replacement from the original sample $x_1, x_2, \ldots, x_n$.

The bootstrap procedure consists of the following steps:

Generating a large number, N, of bootstrap samples X_i^b $(i = 1, 2, \ldots, N)$

For each bootstrap sample obtained, the sample median, $x_{0.5}(X_i^b)$ is evaluated and called the *bootstrap replication*

The *bootstrap estimate* of standard error of the median of interest is calculated as

$$\hat{S}(x_{0.5}) = \sqrt{\frac{\sum_{i=1}^{N} \left(x_{0.5}\left(X_i^b\right) - \hat{x}_{0.5}\right)^2}{N-1}}$$

where

$$\hat{x}_{0.5} = \sum_{i=1}^{N} \frac{x_{0.5}\left(X_i^b\right)}{N}$$

Note, that no assumption about the distribution of random variable x was introduced.

For some estimation problems, the results obtained using the bootstrap approach coincide with respective known classical ones. This can be illustrated by the following example related to binomial data (Martz and Duran (1985)).

Assume that for each component of the system of interest, we have the data

collected in the form $\{S_i, N_i\}$ ($i = 1, 2, \ldots, n$, where n is the number of component in the system), where N_i is the number of units of ith component tested (or observed) during a fixed time interval (the same for all n components of the system) and S_i is the respective number of units survived.

The basic steps of the corresponding bootstrap simulation procedure are as follows:

1. For each component of the system given, obtain the bootstrap estimate of component reliability, R_i, ($i = 1, 2, \ldots, n$, where n is the number of component in the system), generating it from the binomial distribution with parameters N_i and $p = S_i/N_i$. In the case when $S_i = N_i$, i.e., $p = 1$, one needs to smooth the bootstrap, replacing p by $(1 - \epsilon)$, where $\epsilon << 1$. This procedure is discussed in (Efron and Tibshirani (1979)).
2. Calculate the corresponding classical estimate of the system reliability, using (7.36) with R_i ($i = 1, 2, \ldots, n$) obtained from the results of step 1.
3. Repeat steps 1-2 a sufficiently large number of times, n, (for example, 10,000) to get a large sample of $\hat{R}_S$.
4. Based on the sample obtained, and a chosen confidence level $(1 - \alpha)$, construct the respective lower confidence limit for the system reliability of interest as a sample percentile of level α, using (7.37).

Example 7.9

Consider a fault tree, the top event, T, of which is described by the following expression:

$$T = C_1 + C_2C_3$$

where C_1, C_2, and C_3 are the cut-sets of the system modeled by the fault tree. If the following data are reported for the components representing the respective cut-sets, determine a point estimate and 95% confidence interval for the system reliability $\hat{R}_S = 1 - \Pr(T)$ using a) the system reduction methods, and b) the bootstrap method.

Component	Number of failures d	Number of trials N
C_1	1	1785
C_2	8	492
C_3	4	371

Solution:

a) The second cut-set can be considered a parallel subsystem containing components C_2 and C_3. Therefore, we shall apply the Maximus method to reduce this subsystem to an equivalent component C_{23}. The maximum likelihood point estimate of the equivalent component reliability can be obtained as

$$\begin{aligned}\hat{R}_{C_{23}} &= 1 - \Pr(C_{23}) \\ &= 1 - \left(\frac{8}{492}\right)\left(\frac{4}{371}\right) \\ &= 0.99982\end{aligned}$$

The equivalent number of trials for C_{23} is

$$\begin{aligned}N_{C_{23}} &= \min(N_{C_2}, N_{C_3}) \\ &= \min(492, 371) \\ &= 371\end{aligned}$$

and, using (7.35), the equivalent number of the failures is

$$\begin{aligned}D_{C_{23}} &= N_{C_{23}}(1 - \hat{R}_{C_{23}}) \\ &= (371)(1 - 0.99982) \\ &= 0.06678\end{aligned}$$

Now we can treat C_1 and C_{23} as a series system and apply the Lloyd–Lipow method to reduce it. Using (7.33), the estimate of system reliability:

$$\begin{aligned}\hat{R}_s &= \left(1 - \frac{1}{1785}\right)\left(1 - \frac{0.067}{371}\right) \\ &= 0.99926\end{aligned}$$

Then, keeping in mind the equivalent number of system trials of

$$N_S = \min\left(N_{C_1}, N_{C_{23}}\right)$$

$$= \min(1785, 371)$$

$$= 371$$

from (7.32) the fictitious number of system failures is:

$$D_S = N_S \times \left(1 - \hat{R}_S\right)$$

$$= (371)[1 - (0.99926)]$$

$$= 0.27454$$

Using (3.83–3.84) with $n = 371$ and $r = 0.27$, the 95% lower confidence limits for the system reliability estimate are found to be:

$$0.99758 \leq \hat{R}_S \leq 0.99988$$

b) The bootstrap estimation can be obtained as follows:

1. Using the failure data for each component, compute the estimate of the binomial probability of failure and treat it as a nonrandom parameter p.
2. Simulate N binomial trials of a component and count the observed number of failures.
3. Obtain a bootstrap replication of p dividing the observed number of failures by the number of trials. Once the bootstrap replications are computed for each component, find the estimate of system reliability using (7.36).
4. Repeat steps 2 and 3 sufficiently large number of times, and use (7.37) to obtain the interval estimates of system reliability.

The procedure and results of the bootstrap solution are summarized in Table 7.5

As seen from Table 7.5, the point estimate of system reliability closely coincides with the one obtained in part a).

From the distribution of the system reliability estimates, the 95% confidence bounds can be obtained as the 2.5% and 97.5% sample percentiles—see (7.37):

$$0.99779 \leq \hat{R}_S \leq 0.99999$$

Table 7.5 The Bootstrap Solution in Example 7.9

Monte Carlo Run No.	Component	C_1	C_2	C_3	Estimate of system reliability R_s
	Number of failures d	1	8	4	
	Number of t rials N	1785	492	371	
	Binomial probability of failure $p = d/N$	0.00056	0.01626	0.01078	
1	Observed number of failures in N binomial trials with parameter p, d_i	0	7	4	0.99985
	Bootstrap replication, $p_i^b = d_i / N$	0.00000	0.01423	0.01078	
2	Observed number of failures in N binomial trials with parameter p, d_i	0	7	6	0.99977
	Bootstrap replication, $p_i^b = d_i / N$	0.00000	0.01423	0.01617	
3	Observed number of failures in N binomial trials with parameter p, d_i	1	9	4	0.99924
	Bootstrap replication, $p_i^b = d_i / N$	0.00056	0.01829	0.01078	
...	...	...	...	...	...
	...	...	...	...	...
10,000	Observed number of failures in N binomial trials with parameter p, d_i	2	10	5	0.99861
	Bootstrap replication, $p_i^b = d_i / N$	0.00112	0.02033	0.01348	
				$E(R_s)$	9.9926×10^{-1}
				$\text{var}(R_s)$	3.4500×10^{-7}

The system reliability confidence bounds can also be estimated through the use of the Clopper-Pearson procedure. From (3.81), the fictitious number of system trials is:

$$N_S = \frac{\hat{R}_s \times (1 - \hat{R}_s)}{\text{var}(\hat{R}_s)}$$

$$= \frac{(9.9926E-7)[1-(9.9926E-7)]}{3.45E-7}$$

$$= 2155.5$$

Then, the fictitious number of system failures is

$$D_S = N_S \times (1 - \hat{R}_S)$$

$$= (2155.5)[1-(0.99926)]$$

$$= 1.6$$

Using (3.83–3.84) with $n = 2155.5$ and $r = 1.6$, the 95% lower confidence limits for the system reliability estimate are found to be:

$$0.99899 \leq \hat{R}_s \leq 0.99944$$

which is quite consistent with the results obtained from the other two methods.

Martz and Duran (1985) performed some numerical comparisons of the Maximus, bootstrap and Bayes' Monte Carlo methods applied to 20 simple and moderately complex system configurations and simulated binomial data for the system components. Martz and Duran made the following conclusions about the regions of superior performance of the methods:

1. The Maximus method is, generally, superior for: a) moderate to large series systems with small quantities of test data per component, and b) small series systems composed of repeated components.
2. The Bootstrap method is recommended for highly reliable and redundant systems.
3. The Bayes' Monte Carlo method is, generally, superior for: a) moderate to large series systems of reliable components with moderate to large samples of test data, and b) small series systems, composed of reliable nonrepeated components.

7.3.5 Graphic Representation of Uncertainty

The results of a probabilistic uncertainty analysis should be presented in a clear manner that aids analysts in developing appropriate qualitative insights. Generally, we will discuss three different ways of presenting probability distributions (so, their use is not limited by uncertainty analysis): plotting the pdf or the cdf, or

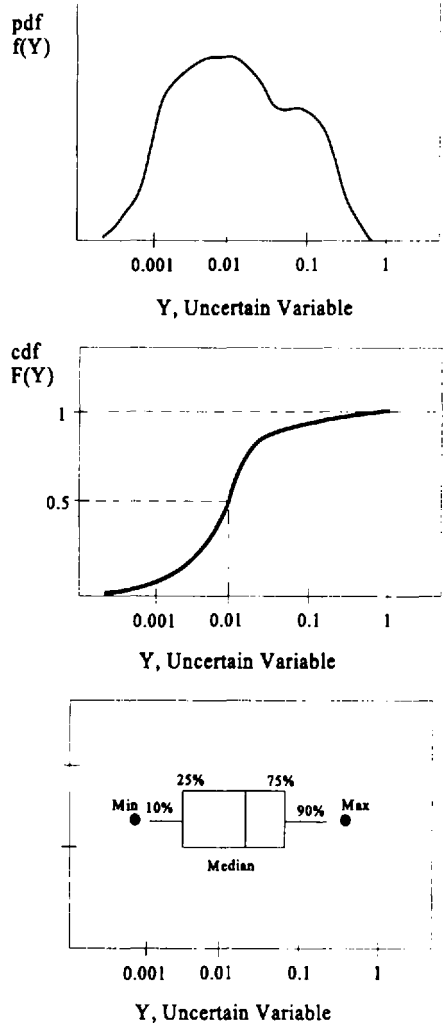


Figure 7.2 Three conventional methods of displaying distribution.

displaying selected percentiles, as in a Tukey (1979) box plot. Figure 7.2 shows examples. The probability density function shows the relative probabilities of different values of the parameters. One can easily see the areas or ranges where high densities (occurrences) of the r.v. occur (e.g., the modes). One can easily judge symmetry and skewness and the general shape of the distribution (e.g., bell-shaped vs. J-shaped). The cdf is best for displaying percentiles (e.g., median) and the respective confidence intervals. It is easily used for both continuous and discrete distributions.

The standard Tukey box shows a horizontal line from the 10th to 90th percentiles, a box between the lower percentiles (e.g., from the 25th to 75th percentiles), and a vertical line at the median, and points at the minimum and maximum observed values. This method clearly shows the important quantities of the r.v.

In cases where statistical uncertainty limits are estimated, the Tukey box can be used to describe the confidence intervals. Consider a case where the distribution of a variable Y is estimated and described by a pdf. For example, a pdf of time-to-failure (the aleatory model) can be represented by an exponential distribution and the value of λ for this exponential distribution is represented, under the Bayes' approach, by a lognormal distribution. Then, $f(Y|\lambda)$ for various values of λ can be plotted and families of curves can be developed to show an aggregate effect of both kinds of uncertainty. This (epistemic uncertainty) is shown in Figure 7.3.

In general, a third method can be shown by actually displaying the probability densities of λ in a multidimensional form. For example, Figure 7.4 presents such a case for a two-dimensional distribution. In this figure $f(Y|\lambda)$ is shown for various values of λ.

7.4 USE OF EXPERT OPINION FOR ESTIMATING RELIABILITY PARAMETERS

The use of expert opinions is often desired in reliability analysis, and in many cases is unavoidable. One reason for using experts is the lack of a statistically significant amount of empirical data necessary to estimate new parameters. Another reason for using experts is to assess the likelihood of a one-time event, such as the chance of rain tomorrow.

However, the need for expert judgement that requires extensive knowledge and experience in the subject field is not limited to one-time events. For example, suppose we are interested in using a new and highly capable microcircuit device currently under development by a manufacturer and expected to be available for use soon. The situation requires an immediate decision on whether or not to design an electronic box around this new microcircuit device. Reliability is a critical decision criterion for the use of this device. Although reliability data on the new

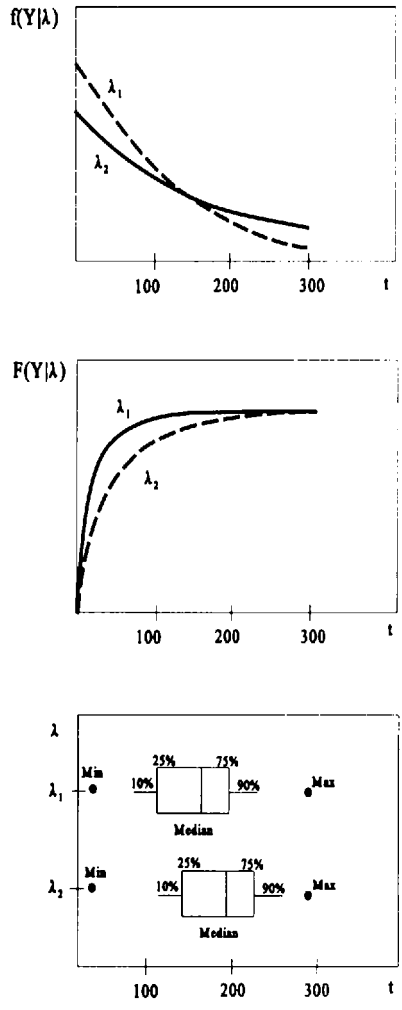


Figure 7.3 Representation of uncertainties.

device are not available, reliability data on other types of devices employing similar technology are accessible. Therefore, reliability assessment of the new device requires both knowledge of and expertise in similar technology, and can be achieved through the use of expert opinion.

Some specific examples of expert use are the Reactor Safety Study (1975); IEEE-Standard 500 (1984); and *Severe Accident Risk: An Assessment for Five U.S.*

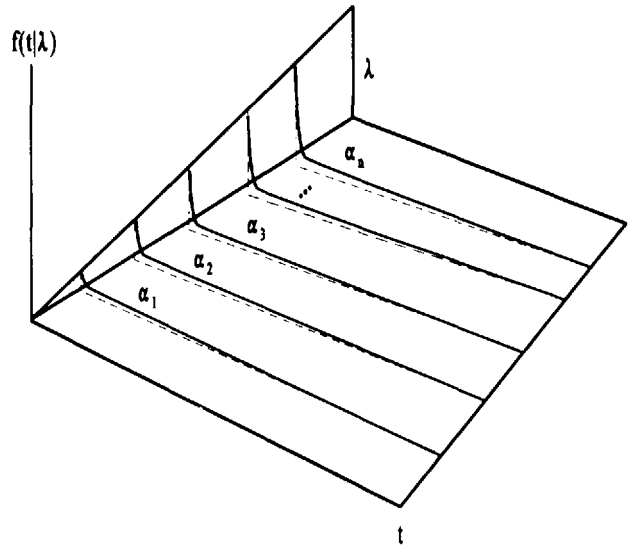


Figure 7.4 Two-dimensional uncertainty representation.

Nuclear Power Plants (1990), where expert opinion was used to estimate the probability of components failure and other rare events. The power industry's Electric Power Research Institute has relied on expert opinion to assess seismic hazard rates. Other applications include weather forecasting. For example, Clemens et al. (1990) discusses the use of expert opinion by meteorologists. Another example is the use of expert opinion in assessing human error rates discussed by Swain and Guttman (1983).

The use of expert opinion in decision making is a two-step process: elicitation and analysis of expert opinion. The method of elicitation may take the form of individual interviews, interactive group sessions, or the Delphi approach discussed by Dalkey and Helmer (1963). The relative effectiveness of different elicitation methods has been addressed extensively in the literature. Techniques for improving the accuracy of expert estimates include calibration, improvement in questionnaire design, motivation techniques, and other methods, although clearly no technique can be applied to all situations. The analysis portion of expert use involves combining expert opinions to produce an aggregate estimate that can be used by reliability analysts. Again, various aggregation techniques for pooling

expert opinions exists, but of particular interest are those adopting the form of mathematical models. The usefulness of each model depends on both the reasonableness of the assumptions (implicit and explicit) carried by the model as it mimics the real world situation, and the ease of implementation from the user's perspective. The term "expert" generally refers to any source of information that provides an estimate and includes human experts, measuring instruments, and models.

Once the need for expert opinion is determined and the opinion is elicited, the next step is to establish the method of opinion analysis and application. This is a decision task for the analysts, who may simply decide that the single best estimate of the value of interest is the estimate provided by the arithmetic average of all estimates, or an aggregate from a nonlinear pooling method, or some other opinions. Two methods of aggregating expert opinion are discussed in more detail, the geometric averaging technique and the Bayesian technique.

7.4.1 Geometric Averaging Technique

Suppose n experts are asked to make an estimate of the failure rate of an item. The estimates can be pooled using the geometric averaging technique. For example, if λ_i is the estimate of the ith expert, then an estimate of the failure rate is obtained from

$$\hat{\lambda} = \sqrt[n]{\prod_{i=1}^{n} \lambda_i} \tag{7.38}$$

This was the primary method of estimating failure rates in IEEE-Standard 500 (1984). The IEEE-Standard 500 contains rate data for electronic, electrical, and sensing components. The reported values were synthesized primarily from the opinions of some 200 experts (using a form of the Delphi procedure). Each expert reported "low," "recommended," and "high" values for each failure rate under normal conditions, and a "maximum" value that would be applicable under all conditions (including abnormal conditions). The estimates were pooled using (7.38). For example, for maximum values,

$$\hat{\lambda}_{\max} = \sqrt[n]{\prod_{i=1}^{n} \lambda_{\max,\, i}}$$

As discussed by Mosleh and Apostolakis (1983), the use of geometric averaging implies that 1) all the experts are equally competent, 2) the experts do not have any systematic biases, 3) experts are independent, and 4) the preceding three

assumptions are valid regardless of which value the experts are estimating, e.g., high, low, or recommended.

The estimates can be represented in the form of a distribution. Apostolakis et al. (1980) suggests the use of a lognormal distribution for this purpose. In this approach, the "recommended" value is taken as the median of the distribution, and the error factor (EF) is defined as

$$EF = \sqrt{\frac{\hat{\lambda}_{0.95}}{\hat{\lambda}_{0.05}}} \tag{7.39}$$

7.4.2 Bayesian Approach

As discussed by Mosleh and Apostolakis (1983), the challenge of basing estimates on expert opinion is to maintain coherence throughout the process of formulating a single best estimate based on the experts' actual estimates and credibilities. Coherence is a notion of internal consistency within a person's state of belief. In the subjectivist school of thought, a probability is defined as a measure of personal uncertainty. This definition assumes that a coherent person will provide his or her probabilistic judgements in compliance with the axioms of probability theory.

An analyst often desires a modeling tool that can aid him or her in formulating a single best estimate from expert opinion(s) in a coherent manner. Informal methods such as simple averaging will not guarantee this coherence. Bayes' theorem, however, provides a framework to model expert belief, and ensures coherence of the analysts in arriving at a new degree of belief in light of expert opinion. According to the general form of the model given by Mosleh and Apostolakis, the state-of-knowledge distribution of a failure rate λ, after receiving an expert estimate $\hat{\lambda}$, can be obtained by using Bayes' theorem in the following form:

$$\Pi(\lambda \mid \hat{\lambda}) = \frac{1}{k} L(\hat{\lambda} \mid \lambda)\, \pi_0(\lambda) \tag{7.40}$$

where $\pi_0(\lambda)$ is the prior distribution of λ; $\pi(\lambda \mid \hat{\lambda})$ is the posterior distribution of λ; $L(\hat{\lambda} \mid \lambda)$ is the likelihood of receiving the estimate $\hat{\lambda}$, given the true failure rate λ; k is a normalizing factor.

One of the models suggested for the likelihood of observing $\hat{\lambda}$ given λ, is based on the lognormal distribution in the following form:

$$L(\hat{\lambda} \mid \lambda) = \frac{1}{\sqrt{2\pi}\sigma\hat{\lambda}} \exp\left[-\frac{1}{2}\left(\frac{\ln(\hat{\lambda}) - \ln(\lambda) - \ln(b)}{\sigma}\right)^2\right] \qquad (7.41)$$

where b is a bias factor ($b = 1$ when no bias is assumed) and σ is the standard deviation of the logarithm of $\hat{\lambda}$, given λ. When the analyst believes no bias exists among the experts, he or she can set $b = 1$. The quantity σ, therefore, represents the degree of accuracy of the experts' estimate as viewed by the analyst. The work by Kim (1991), which includes a Bayesian model for a relative ranking of experts, is an extension of the works by Mosleh and Apostolakis.

7.4.3 Statistical Evidence on the Accuracy of Expert Estimates

Among the attempts to verify the accuracy of expert estimates, two types of expert estimates are studied—assessment of single values and assessment of distributions.

Notable among the studies on the accuracy of expert assessments of a single estimate is Snaith's study (1981). In this study, observed and predicted reliability parameters for some 130 pieces of different equipment and systems used in nuclear power plants were evaluated. The predicted values included both direct assessments by experts and the results of analysis. The objective was to determine correlations between the predicted and observed values. Figure 7.5 shows the ratio ($R = \lambda/\hat{\lambda}$) of observed to predicted values plotted against their cumulative frequency. As shown, the majority of the points lie within the dashed boundary lines. Predicted values are within a factor of 2 from the observed values, and 93% are within a factor of 4. The figure also shows that $R = 1$ is the median value, indicating that there is no systematic bias in either direction. Finally, the linear nature of the curve shows that R tends to be lognormally distributed, at least within the central region. This study clearly supports the use and accuracy of expert estimation.

Among the studies of expert estimation are the works by cognitive psychologists. For example, Lichtenstein et al. (1977) described the results of testing the adequacy of probability assessments and concluded that "the overwhelming evidence from research on uncertain quantities is that people's probability distributions tend to be biased." Commenting on judgemental biases in risk perception, Slovic et al. (1980) stated: "A typical task in estimating uncertain quantities like failure rates is to set upper and lower bounds such that there is a 98% chance that the true value lies between them. Experiments with diverse groups of people making different kinds of judgements have shown that, rather than 2% of true values falling outside the 98% confidence bounds, 20% to 50% do so. Thus, people think that they can estimate such values with much greater precision than is actually the case."

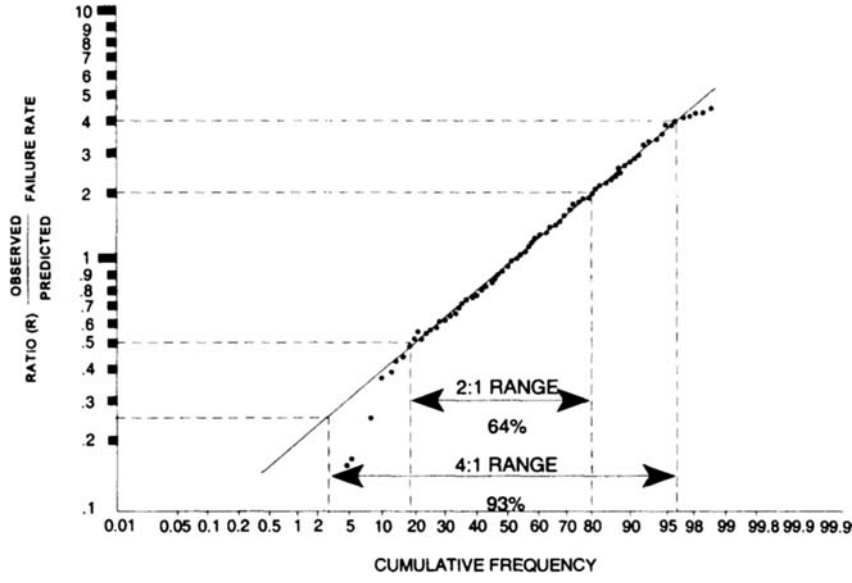


Figure 7.5 Frequency distribution of the failure rate ratio, Snaith (1981).

Based on the above conclusion Apostolakis (1982) has suggested the use of the 20th and 80th percentiles of lognormal distributions instead of the 5th and 95th when using (7.40), to avoid a bias toward low values, overconfidence of experts, or both. When using the Bayesian estimation method based on (7.41), the bias can be accounted for by using larger values of σ and b in (7.41).

7.5 PROBABILISTIC FAILURE ANALYSIS

Statistical, probabilistic, or deterministic methods are used to analyze failures. While all three methods or combinations of them can be used, in this section we rely primarily on the statistical methods for the analysis of failures. However, for evaluating the results of the analysis, mainly deterministic techniques are used. Probabilistic (Bayesian) and deterministic techniques are equally applicable. However, since Bayesian techniques may require expert or prior knowledge about equipment failures, they should be used only when observed failure data are sparse.

The statistical methods described in this section are based on the classical inference methods discussed in Chapters 3 and 5. That is, the history of failure or event occurrences is first studied to determine whether or not a statistically significant trend can be detected. If not, the traditional maximum likelihood parameter

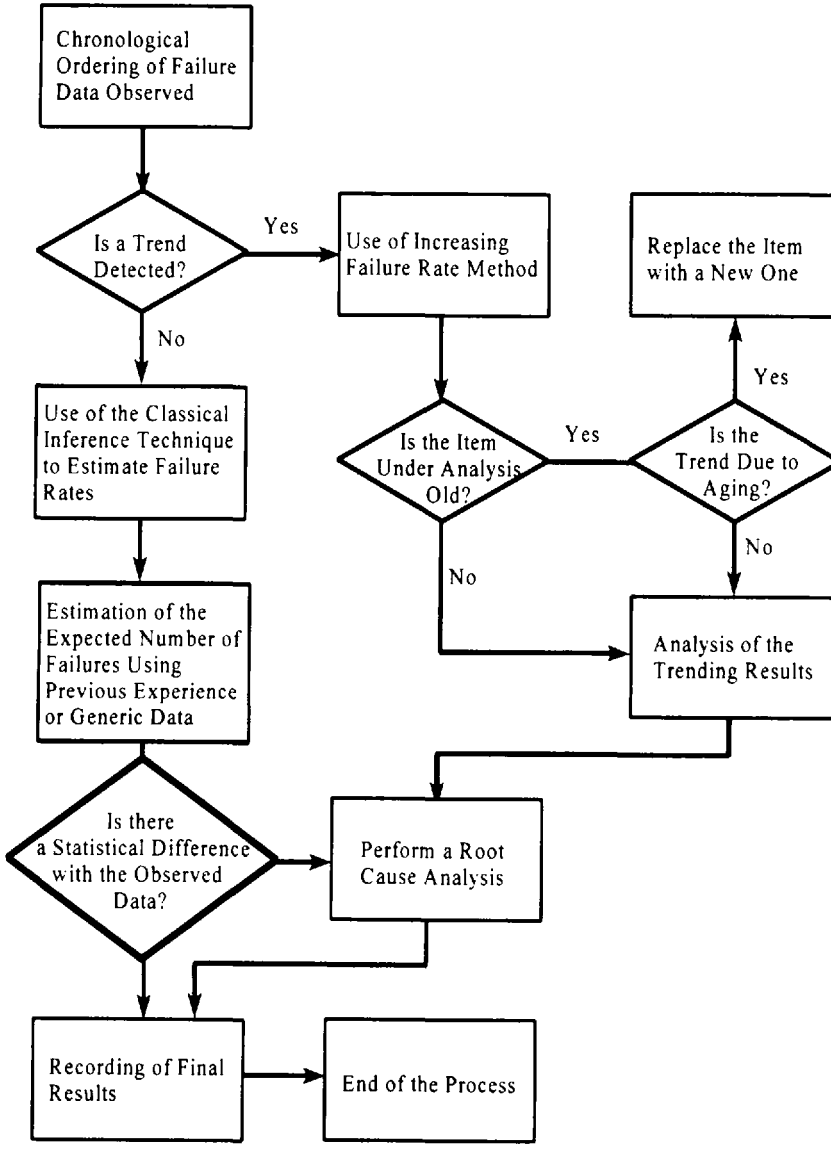


Figure 7.6 Failure analysis process.

estimating method is used to determine the failure characteristic of the item, e.g., to determine the failure rate or demand failure probability of an item.

If the trend analysis method discussed in Chapter 5 shows a significant trend in the data, it is important to determine the nature and degree to which the failure characteristic of the item is changed. Classical statistics methods are used to determine the failure characteristics of equipment if no trends are exhibited.

When the failure characteristics of an item with or without trend are determined, one needs to evaluate them to determine whether or not they show any change in the capability of the item. Both statistical and nonstatistical techniques can be used to detect changes.

When significant changes are detected, it is very important to search for possible reasons for such changes. This may require an analysis to determine the root causes of the detected changes or observed failure events. No standard practice exists for determining root-cause failures. Engineers often use ad hoc techniques for this purpose.

Figure 7.6 shows the overall approach employed in this section. The basis for the statistical methods used in this document are explained in the remainder of this section.

7.5.1 Detecting Trends in Observed Failure Events

The use of statistical estimators for equipment failure characteristics should be done only after it has been determined whether the failure occurrence is reasonably constant, i.e., there is no evidence of an increasing or decreasing trend. In Chapter 5, we described the Centroid method to test for the possibility of a trend. In (5.26), since statistic U is a sensitive measure, one could use the following practical criteria to ensure detection of trends, especially when the amount of data are limited:

1. When $U > 0.5$ or $U < -0.5$, assume a reasonable trend exists.
2. Otherwise, depending on the age and the item's recent failure history, assume a mostly constant failure rate (or failure probability) or a mild trend exists.

7.5.2 Failure Rate and Failure Probability Estimation for Data with No Trend

Sections 3.4–3.5 dealt with statistical methods for estimating failure rate and failure probability parameters of components when there is no trend in failures. The objective is to find a point estimate and a confidence interval for the parameters of interest.

Parameter Estimation When Failures Occur by Time

When failures of equipment occur by time (r failures in T hours), the exponential distribution is most commonly used. Therefore, when the failure events are believed to occur at a constant rate (i.e., with no trend), the exponential model is reasonable and the parameter estimation should proceed. In this case, λ parameter must be estimated. The point estimator is for the failure rate parameter (λ) of the exponential distribution obtained from $\hat{\lambda} = r/T$. Depending on the method of observing data, the confidence interval of λ can be obtained from one of the expressions in Table 3.2.

Parameter Estimation When Failures Occur on Demand (Binomial Model)

When the data are in the form of X failures in n trials (or demands), no time relationship exists and the binomial distribution best represents the data. This situation often occurs for equipment in the standby mode, e.g., a redundant pump that is demanded for operation n times in a fixed period of time. In a binomial distribution, the only parameter of interest is p. An estimate of p and its confidence interval can be obtained from (3.78–3.79).

7.5.3 Failure Rate and Failure Probability Estimation for Data with Trend

The existence of a trend in the data indicates that the interarrivals of failures are not statistically similar, and thus (5.6) should be used. Chapter 5 describes the methods of estimating the rate of failure occurrence $\lambda(t)$.

7.5.4 Evaluation of Statistical Data

After the data are analyzed, it is important to determine whether or not any significant changes between the past data and more recent data can be detected. If such changes are detected, it is important to formulate a procedure for dealing with them.

Evaluation of Data with No Trend

Two methods of evaluation are considered, statistical and nonstatistical. One effective statistical technique is the Chi-square method. The nonstatistical technique only considers degrees of change in the failure characteristics of an item (e.g.,

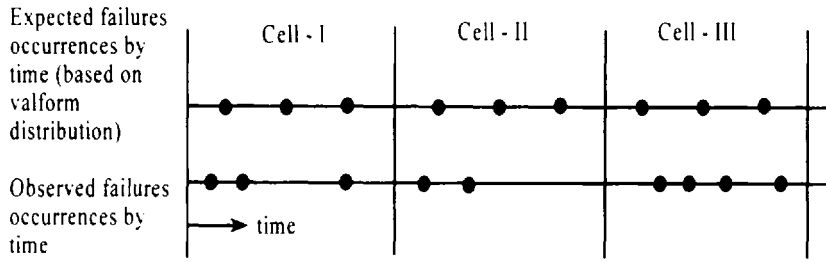


Figure 7.7 Comparison of expected and observed failure occurrences by item.

in the form of a percent difference from a generic value or prior experience). Proper action is suggested based on a predefined criterion.

As mentioned earlier, the Chi-square method can be adapted to the type of problems considered here. The Chi-square method was described in Chapter 2. In failure analysis, the Chi-square test can be used to determine whether or not the observed failure data are statistically different from generic data, or from past history of the same or a similar item. For example, consider Figure 7.7.

If the expected number of failures, based on generic failure data or previously calculated values (e.g., using statistical analysis), are determined and compared with the
observed failures, one can statistically measure the difference.

It is easy to divide the time line (or in a demand type item, the number of demands) into equal time demand intervals (e.g., three intervals as in Figure 7.7) and compare them to see whether or not the observed and expected failures in each interval are statistically different.

For example, for data in Figure 7.7, the following Chi-square statistic can be calculated:

$$W = (3 - 3)^{2/3} + (3 - 2)^{2/3} + (3 - 4)^{2/3} = 2/3$$

This shows that there is a slight difference between the observed and expected data, but depending on the desired level of confidence, this may or may not be acceptable.

The nonstatistical technique uses only a percent difference between the estimated failure rate $\hat{\lambda}$ and the generic failure rate λ_g. For example, by using

$$e = \left| \frac{\hat{\lambda} - \lambda_g}{\lambda_g} \right| \times 100 \tag{7.42}$$

if the difference is large (more than 100), one can assume the data are different and further root-cause analysis is required.

Evaluation of Data with Trend

Generally, there is no set rule for this purpose. One approach is to use the doubling failure concept. If two consecutive intervals of (t_1, t_2) and (t_2, t_3) are such that $t_2 - t_1 = t_3 - t_2$, and the expected number of failures in each interval (N_1 and N_2 respectively) are such that $N_2/N_1 = 2$, then it is easy to prove, using (5.2) and (5.3), that $\beta = 1.58$. Accordingly, for $N_2/N_1 = 5$, $\beta = 2.58$. These can be used as guidelines for determining the severity of the trend. For example, one can assume the following:

If $1 \leq \beta \leq 1.58$, the trend is mildly increasing. Suggest a root-cause analysis and implement a careful monitoring system.
If $1.58 < \beta \leq 2.58$, the trend is major. Suggest replacement or root-cause analysis.
If $\beta > 2.58$, the trend is significant. Cease operation of the item and determine the root cause of the trend.

7.5.5 Root-Cause Analysis

Root causes are the most basic causes that can be reasonably identified by experts and can be corrected so as to minimize their recurrence. The process of identifying root causes is generally performed by a group of experts (investigators). Modarres et al. (1989) explains the application of expert systems in root-cause analysis. The goal of the experts is to identify the basic causes. The more specific they can be about the reasons an incident occurred, the easier it is to arrive at a recommendation that will prevent recurrence of the failure events. However, investigation of root causes should not be carried to the extreme. The analysis should yield the most out of the time spent, and only identify root causes for which a reasonable corrective action exists. Therefore, very complex and specific mechanisms of failure do not need to be identified, especially when corrective actions can be determined at a higher level of abstraction. The recommended corrective actions should be specific and should directly address the root causes identified during the analysis.

Root-cause analysis involves three steps:

1. Determining events and causal factors.
2. Coding and documenting root causes.
3. Generating recommendations.

Charting the event and causal factors provides a road map for experts to organize and analyze the information that they gather, identify their findings, and highlight gaps in knowledge as the investigation progresses. For example, a sequence diagram similar to that in Figure 7.8 is developed, showing the events leading up to and following an occurrence as well as the conditions and their causes surrounding the failure event. The process is performed inductively and in progressively more detail.

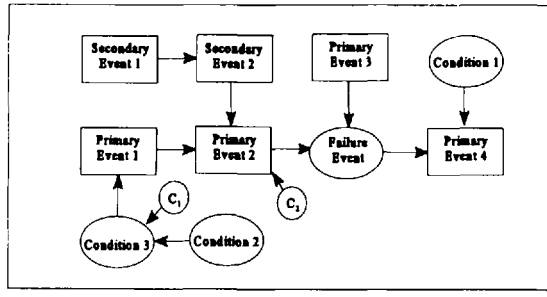
Figure 7.8a shows the causal relations leading to a "failure event," including the conditions, events, and causal factors.

Following this step, the causal factors and events should be documented. One method suggested by the *Root-Cause Analysis Handbook* (1991) uses a root cause tree involving six levels. From the event and causal factors chart, these levels are described and documented. Figure 7.8b shows an example of the levels used and Figure 7.8c shows an example of a report made based on this classification.

The final and most important step in this process is to generate of recommendations. This process is based on the experience of the experts. However, as a general guideline, the following items should be considered when recommending corrective actions:

1. At least one corrective action should be identified for each root cause.
2. The corrective action should directly and unambiguously address the root cause.
3. The corrective action should not have secondary degrading effects.
4. The consequences of the recommended (or not recommended) corrective actions should be identifiable
5. The cost associated with implementation of the corrective action should be estimated.
6. The need for special resources and training for implementation of the action should be identified.
7. The effect on the frequency of item failure should be estimated.
8. The impact the corrective action is expected to have on other items or on workers should be addressed.
9. The effect of the corrective action should be easily measurable.
10. Other possible corrective actions that are more resource intensive but more effective should be listed.

The root-cause analysis is a major field of study. For further reading in this subject, see Chu (1989), Ferry (1988), Kendrick (1987, 1990).



Levels of the Root Cause Tree

Level	Shape	Description	Examples
A		Primary Difficulty Source	· Equipment Difficulty · Operations Difficulty · Technical Difficulty
B		Area of Responsibility	· Equipment Reliability/ Design · Production Organization · Technical Support Organization
C		Equipment Problem Category	· Design · Installation/Corrective/ Preventive Maintenance Difficulty · Fabrication Difficulty
D		Major Root Cause Category	· Design Review/ Verification · Training · Management Systems
E		Near Root Cause	· Procedures Followed Incorrectly · Workplace Layout · Supervision During Work
F		Root Cause	· More Than One Action Per Step · Conflicting Layouts · No Supervision

Causal Factor	Path Through Root Cause Tree	Recommendations
Operator had not previously been to Motor Control Center BACKGROUND: The Operator who went to verify the position of the Motor-Generator (M-G) switchgear had not been required to use this particular switchgear in the past. If he had been shown the switchgear as part of the training, it is unlikely that he would have forgotten its location.	· Operations Difficulty · Production Organization · Training The Operator had never been trained on location of equipment (including switchgear) in Motor Control Center	· Include a tour of the Motor Control Centers as part of on-the-job training. Provide specific instructions on how to use drawings for verifying the positions of important places of equipment. (Production Training Department.

Figure 7.8 Events and causal factors chart.

EXERCISES

7.1 Consider two resistors in parallel configuration. The mean and standard deviation for the resistance of each are as follows:

$$\mu_{R1} = 25\ \Omega \quad \sigma_{R1} = 0.1\ \mu_{R1}$$
$$\mu_{R2} = 50\ \Omega \quad \sigma_{R2} = 0.1\ \mu_{R2}$$

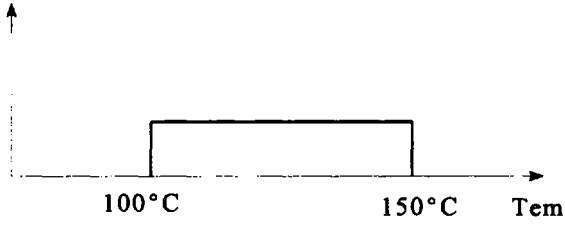
Using one of the statistical uncertainty techniques, obtain:

a) mean and standard deviation of the equivalent resistor,
b) in what ways the uncertainty associated with the equivalent resistance is different from the individual resistor? Discuss the results.

7.2 The results of a bootstrap evaluation gives: $\mu = 1 \times 10^{-4}$, and $\sigma = 1 \times 10^{-3}$. Evaluate the number of pseudo failures F, in N trials for an equivalent binomial distribution. Estimate the 95% confidence limits of μ.

7.3 Repeat Exercise 4.6 and assume that a common cause failure between the valves and the pumps exist. Using the generic data in Table C.1, calculate the probability that the top event occurs. Use a β-factor method with $\beta = 0.1$ for valves and pumps. Discuss if the selection of $\beta = 0.1$ is sensitive to the end result.

7.4 A class of components is temperature sensitive in that they will fail if temp erature is raised too high. Uncertainty associated with a component's failure temperature is characterized by a continuous uniform distribution such as shown below:



If the temperature for a particular component is uncertain but can be characterized by an exponential distribution with $\lambda = 0.05$ per degree Celsius, calculate the reliability of this component.

7.5 Consider the cut-sets below describing the failure of a simple system: $F = AB + BC$. The following data have been found for components A, B, and C

Components	A	B	C
Number of failure	5	12	1
Total test time (hr)	1250	4315	2012

a) Use the system reduction methods to calculate equivalent number of failures and total test time for failure of the system.
b) Given the results of (a), calculate the 90% confidence limits for the unreliability of this system.

REFERENCES

Apostolakis, G., "*Data Analysis in Risk Assessment*," Nuclear Engineering and Design, 71:375–381, 1982.

Apostolakis, G. and Lee, V.T., "*Methods for the Estimation of Confidence Bounds for the Top Event Unavailability of Fault Trees*," Nuclear Engineering and Design, Vol. 41, pp. 411–419, 1977.

AT&T Reliability Manual, edited by Klinger, D.J., Nakada, Y., and Menendez, M., Van Nostrand Reinhold, New York, 1990.

Atwood, C.L., "*Common Cause Failure Rates for Pumps*," NUREG/CR-2098, U.S. Nuclear Regulatory Commission, Washington, DC, 1983.

Bier, V.M., "*A Measure of Uncertainty Importance for Components in Fault Trees*," Transactions of the 1983 Winter Meeting of the Am. Nucl. Soc., San Fransisco, CA, 1983.

Barlow, R.E. and Proschan, F., "*Statistical Theory of Reliability and Life Testing: Probability Models*," To Begin With, Silver Spring, MD, 1981.

Chan, C. K., "*A Proportional Hazard Approach to SiO_2 Breakdown Voltage*," IEEE Trans. on Reliability, R-39, 147–150, 1990.

Chu, C., "*Root Cause Guidebook: Investigation and Resolution of Power Plant Problems*," Failure Prevention, Inc., San Clemente, CA, 1989.

Clemens, R.J. and Winkler, R.L., "*Unanimity and Compromise Among Probability Forecasters*," Mgmt. Science, 36:767–779, 1990.

Cox, D.R., and Oaks, D., "The Analysis of Survival Data," Chapman & Hall, London, New York, NY, 1984.

Crowder, M.J., Kimber A.C., Smith, R.L., and Sweeting, T.J., "*Statistical Analysis of Reliability Data*," Chapman & Hall, London, New York, NY, 1991.

Dalkey, N. and Helmer, O., *"An Experimental Application of the Delphi Method to the Use of Experts,"* Mgmt. Science, 9:458–467, 1963.

Dezfuli, H. and Modarres, M., *"Uncertainty Analysis of Reactor Safety Systems with Statistically Correlated Failure Data,"* Reliability Engineering Journal, Vol. 11, 1, pp. 47–64, 1984.

Efron, B.A. and Tibshirani, R.J., *"An Introduction to the Bootstrap,"* Chapman and Hall, London, New York, NY, 1979.

Ferry, T. S., "Modern Accident Investigation Analysis," 2nd Ed., Wiley, New York, 1988.

Fleming, K.N., *"A Reliability Model for Common Mode Failures in Redundant Safety Systems,"* Proceeding of the Sixth Annual Pittsburgh Conference on Modeling and Simulations, Instrument Society of America, Pittsburgh, PA, 1975.

Fleming, K.N., Mosleh, A., and Deremer, R.K., *"A Systematic Procedure for the Incorporation of Common Cause Event, Into Risk and Reliability Models,"* Nuclear Engineering and Design, 58, 415–424, 1986.

Goldman, A.Ya., *"Prediction of the Deformation Properties of Polymeric and Composite Materials,"*American Chemical Society, Washington, DC, 1994.

Hahn, G.J. and Shapiro, S.S., *"Statistical Models in Engineering,"* John Wiley & Sons, New York, 1967.

IEEE Standard-500, *"IEEE Guide to the Collection and Presentation of Electrical, Electronic and Sensing Component Reliability Data for Nuclear Powered Generation Stations,"* Institute of Electrical and Electronic Engineers, Piscataway, NJ, 1984.

Iman, R.L., Davenport, J.M., and Zeigler, D.K., *"Latin Hypercube Sampling (Program User's Guide),"* SAND79-1473, Sandia National Laboratories, Albuquerque, NM, 1980.

Kaminskiy, M., *"Accelerated Life Testing,* In *Statistical Reliability Engineering,* (to be published), Gnedenko, B.V. Ushakov, I., eds., John Wiley & Sons, New York, 1998.

Kaminskiy, M., Ushakov, I., and Hu, J., *"Statistical Inference Concepts,* In Product Reliability, Maintainability, and Supportability Handbook," Pecht, M., ed., CRC Press, 1995.

Kaplan, S., *"On the Method of Discrete Probability Distributions in Risk and Reliability Calculation—Application to Seismic Risk Assessment," Risk Analysis Journal,* 1, pp. 189–196, 1981.

Kendrick, "Investigating Accidents with STEP," Marcel Dekker, New York, NY, 1987.

Kendrick, "Systematic Safety Training," Marcel Dekker, New York, NY, 1990.

Kim, J.H., *"A Bayesian Model for Aggregating Expert Opinions,"* Ph.D. Dissertation, University of Maryland, Department of Materials and Nuclear Engineering, College Park, MD, 1991.

Leemis, L.M., "Reliability: Probabilistic Models and Statistical Methods," Prentice-Hall, Englewood Cliffs, NJ, 1995.

Lichtenstein, S.B., Fischoff, B., and Phillips, L.D., *"Calibration of Probabilities: The State of the Art,"* Decision Making and Change in Human Affairs, Jungerman, J. and deZeeuw, G., ed., D. Reidel, Dordrecht, Holland, 1977.

Lloyd, D.K. and Lipow, M., *"Reliability: Management, Methods and Mathematics,"* Prentice Hall, Englewood Cliff, NJ, 1962.

Martz, H.F., *"A Comparison of Methods for Uncertainty Analysis of Nuclear Plant Safety System Fault Tree Models,"* U.S. Nuclear Regulatory Commission and Los Alamos National Laboratory, NUREG/CR-3263, Los Alamos, NM, 1983.

Martz, H.F. and Duran B.S., "*A Comparison of Three Methods for Calculating Lower Confidence Limits on System Reliability Using Binomial Component Data*," IEEE Transactions on Reliability, Vol R-34, N 2, pp. 113–121, 1985.

Modarres, M., Chen, L., and Danner, M., "*A Knowledge-Based Approach to Root-Cause Failure Analysis*," Proceeding of the Expert Systems Applications for the Electric Power industry Conference, Orlando, FL, 1989.

Morchland, J.D. and Weber, G.G., "*A Moments Method for the Calculation of Confidence Interval for the Failure Probability of a System*," Proceeding of the 1972 Annual Reliability and Maintainability Symposium, pp. 505–572, 1972.

Morgan, M.G. and Henrion, M., "*Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*," Cambridge Press, Cambridge, UK, 1990.

Mosleh, A. and Siu, N., Smidts, C., and Lui, C., "*Model Uncertainty: Its Characterization and Quantification*," International Workshop Series on Advanced Topics in Reliability and Risk Analysis, Center for Reliability Engineering, University of Maryland, College Park, MD, 1995.

Mosleh, A. and Apostolakis, G., "*Combining Various Types of Data in Estimating Failure Rates*," Transaction of the 1983 Winter Meeting of the American Nuclear Society, San Fransisco, CA, 1983.

Mosleh, A. et al., "*Procedure for Treating Common Cause Failures in Safety and Reliability Studies*," U.S. Nuclear Regulatory Commission, NUREG/CR-4780, Vol. I and II, Washington, DC, 1988.

Mosleh, A., "*Common Cause Failures: An Analysis Methodology and Examples*," Reliability Engineering and System Safety, 34, 249–292, 1991.

Mosleh, A. and Siu, N.O., "*A Multi-parameter, Event-based Common-cause Failure Model*," Proc. of the Ninth International Conference on Structural Mechanics in Reactor Technology, Lausanne, Switzerland, 1987.

Nelson, W., "*Applied Life Data Analysis*," Wiley, New York, 1982.

Nelson, W, "*Accelerated Testing: Statistical Models, Test Plans and Data Analysis*," Wiley, New York, 1990.

Reactor Safety Study: An Assessment of Accidents in U.S. Commercial Nuclear Power Plants, U.S. Regulatory Commission, WASH-1400, Washington, DC, 1975.

Root Cause Analysis Handbook, Westinghouse Savannah River Company, Savannah River Site, WSRC-IM-91-3, 1991.

Severe Accident Risk: An Assessment for Five U.S. Nuclear Power Plants, U.S. Nuclear Regulatory Commission, NUREG-1150, Washington, DC, 1990.

Slovic, P., Fischhoff, B., and Lichtenstein, S., "*Facts Versus Fears: Understanding Perceived Risk*," *Societal Risk Assessment*, Schwing, R.C. and Albers, W.A., Jr., eds., Plenum, New York, 1980.

Snaith, E. R., "*The Correlation Between the Predicted and Observed Reliabilities of Components, Equipment and Systems*," National Center of Systems Reliability, UK Atomic Energy Authority, NCSR-R18, 1981.

Sobczyk, K. and Spencer, B.F., Jr., "*Random Fatigue: From Data to Theory*," Academic Press, New York, 1992.

Swain, A.D., and Guttman, H.E., "*Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Applications*," U.S. Nuclear Regulatory Commission, NUREG/CR-1278, Washington, DC, 1983.

Tukey, J., "*Protection Against Depletion of Stratospheric Ozone by Chlorofluorocarbons*," Report by the Committee on Impacts of Stratospheric Change and the Committee on Alternative for the Reduction of Chlorofluorocarbon Emission, National Research Council, Washington, DC, 1979.

Ushakov, I.A., ed., "*Handbook of Reliability Engineering*," John Wiley & Sons, New York, NY, 1994.

Wheeler, T.A., and Spulak, R.G., "*The Importance of Data and Related Uncertainties in Probabilistic Risk Assessments*," Amer. Nucl. Soc. PSA Topical Meeting, San Fransisco, CA, 1985.

8
Risk Analysis

Risk analysis is a technique for identifying, characterizing, quantifying, and evaluating hazards. It is widely used by private and government agencies to support regulatory and resource allocation decisions. Risk analysis consists of two distinct phases: a qualitative step of identifying, characterizing, and ranking hazards; and a quantitative step of risk evaluation, which includes estimating the likelihood (e.g., frequencies) and consequences of hazard occurrence. After risk has been quantified, appropriate risk-management options can be devised and considered; risk-benefit or cost-benefit analysis may be performed; and risk-management policies may be formulated and implemented. The main goals of risk management are to minimize the occurrence of accidents by reducing the likelihood of their occurrence (e.g., minimize hazard occurrence); reduce the impacts of uncontrollable accidents (e.g., prepare and adopt emergency responses); and transfer risk (e.g., via insurance coverage). The estimation of likelihood or frequency of hazard occurrence depends greatly on the reliability of the system's components, the system as a whole, and human-system interactions. These topics have been extensively addressed in previous chapters of this book. In this chapter we discuss how the reliability evaluation methods addressed in the preceding chapters are used, collectively, in a risk-analysis process. We will discuss some relevant topics which are not discussed in the previous chapters (e.g., risk perception).

8.1 RISK PERCEPTION AND ACCEPTABILITY

8.1.1 Risk Perception

Perceptions of risk often differ from objective measures and may distort or politicize risk-management decisions. Subjective judgement, beliefs, and societal bias

against events with low probability but high consequences may influence the understanding of the results of a risk analysis. Public polls indicate that societal perception of risk, associated with certain unfamiliar or incorrectly publicized activities, is far out of proportion to the actual damage or risk measure. For example, according to Litai (1980), the risk of motor and aviation accidents is perceived to be less than its actual value by a factor of 10 to 100 by the public, but the risk of nuclear power and food coloring is overestimated by a factor of greater than 10,000. Risk conversion and compensating factors must often be applied to determine risk tolerance thresholds accurately, to account for public bias against risks that are unfamiliar (by a factor of 10), catastrophic (by a factor of 30), involuntary (by a factor of 100), or uncontrollable (by a factor of 5 to 10), or have immediate consequences (by a factor of 30). For example, people perceive a voluntary action to be less risky by a factor of 100 than an identical involuntary action. Although the exact values of the above conversion factors are debatable, they generally show the direction and the degree of bias in people's perception.

Different risk standards often apply in the workplace, where risk exposure is voluntary and exposed workers are indemnified. Stricter standards apply to public risk exposure, which is involuntary. The general guide to risk standards is that occupational risk should be small compared with natural sources of risk. Some industrial and voluntary risks may be further decreased by strict enforcement or adequate implementation of known risk-avoidance measures (e.g., wearing seat belts, not drinking alcohol, or not smoking). Therefore, some of these risks are controllable by the individual (who can choose whether to fly, to work, to drive, or to smoke), while others are not (e.g., chemical dumps, severe floods, and earthquakes).

8.1.2 Risk Acceptability

Risk acceptance is a complex subject and is often the subject of controversial debate. However, using the results of risk assessment in a relative manner is a common method of ranking risk-exposure levels. For example, consider Table 8.1. In this table societal risks of individual death due to the leading causes are ranked. An assessed risk from any controllable activity should be required to be lower than the risks of these causes, so as to be defined acceptable. These de facto levels of socially tolerated (acceptable) levels of risk exposure can define acceptable risk thresholds of risk. Although regulators often strive to assess absolute levels of risk, the relative ranking of risks is a better risk-management strategy for allocating resources toward regulatory controls. Cost-benefit analysis is often required as an adjunct to formulating risk-control strategies to socially acceptable levels.

Table 8.1 Major Causes of Death in the United States in 1996

No.	Cause	Number
1	Cardiovascular diseases	948,000
2	Malignancies	522,000
3	Accidents	92,000
	(Motor vehicle)	(42,000)
4	Pneumonia	83,000
5	Pulmonary diseases, Chronic	75,000
6	Diabetes	54,000
7	H.I.V. Infection (AIDS)	37,000
8	Suicide	31,000
9	Liver diseases	27,000
10	Homicide (including police)	26,000
11	Other	434,000
	Total	2,269,000

Another form of risk ranking is to use odds or probability of hazard ex posure per unit of time. For example, Table 8.2 is a typical ranking for some societal causes. It should be noted that for an objective ranking the risk exposure should be the same group. For example, risk of breast causes is different for different age group, and largely applies to women.

Table 8.2 Risk of Dying from Selected Causes

Cause	Odds
Breast Cancer (at age 60)	1 in 500
Breast Cancer (at age 40)	1 in 1000
Car crash	1 in 5300
Drowning	1 in 20,000
Choking	1 in 68,000
Bicycle crash	1 in 75,000

Source: Paulos (1991).

As the third and perhaps a more objective method of risk comparison, sometimes risk exposure is normalized both to the population exposed and to the duration of the exposure and is used for comparison purpose. To compare the risk associated with each cause, consistent units are used (such as number of fatalities or dollar loss per year, per 100,000 population, per event, per person-year of exposure). Table 8.3 shows a risk comparison based on the amount of exposure that yields the same risk value.

The typical guideline for establishing risk-acceptance criteria for involuntary risks to the public has been that fatality rates from the activity of interest should never exceed average individual fatality rates from natural causes (about 0.07 per 100,000 population, from all natural causes) and should be further reduced by risk-control measures to the extent feasible and practical. For example, the U.S. Nuclear Regulatory Commission (1986) has recently suggested quantitative safety goals which implicitly define acceptable risk in nuclear power plants. These safety goals state that the risk from nuclear power plants should not exceed 0.1% of the sum of prompt fatality or cancer fatality risk to which all other risks that individual U.S. residents and the public as a whole are generally exposed. Also it requires that reactors be designed such that the overall mean frequency of a large radioactive release to the environment from a reactor accident be less than $1E-6$ per year of reactor operation.

The societal benefits and the cost trade-offs for risk reduction are widely used guides to set and justify risk acceptability limits. By comparing the risks and benefits associated with certain activities, fair, balanced and consistent limits for

Table 8.3 Risk Exposures That Increase Chance of Death by 1 in 1,000,000 per Year

Nature of risk exposure	Cause of death
Smoking 1.4 Cigarettes	Cancer, heart disease
Spending 1 hour in a coal mine	Black lung disease
Spending 3 hours in a coal mine	Accident
Living 2 days in New York or Boston	Air pollution
Traveling 10 miles by bicycle	Accident
Traveling 300 miles by car	Accident
Traveling 10,000 miles by jet	Accident
Having chest X-ray taken in a good hospital	Cancer caused by radiation
Living 50 years within 5 miles of a nuclear plant	Cancer caused by plant

Source: Wilson (1979).

risk acceptability can be set and institutional controls on risk can be established. Rowe (1977) describes methods of risk-benefit and cost trade-off for risk analysis.

8.2 DETERMINATION OF RISK VALUES

There are two major parts in risk analysis:

Determination of the likelihood, (e.g., prob. P_i or frequency of occurrence, F_i), of an undesirable event, E_i. Sometimes the likelihood estimates are generated from a detailed analysis of past experience and available historical data; sometimes they are judgemental estimates based on an expert's view of the situation, or simply a best guess. This assessment of event likelihood can be useful, but the confidence in such estimates depends on the quality and quantity of the data and the methods used to determine event likelihood.

Evaluation of the consequence, C_i, of this hazardous event. The choice of the type of consequence may affect the acceptability threshold and the tolerance level for the risk.

Risk analysis, generally, consists of the following three steps, sometimes called the "Risk Triplet" which is represented by expression (1.4).

1. Selection of a specific hazardous reference event E_i or scenario S_i (sequence or chain of events) for quantitative analysis (hazard identification)
2. Estimation of the likelihood or frequencies of events, P_i (or F_i)
3. Estimation of the consequences of these events, C_i

In most risk assessments the likelihood of event E_i is expressed in terms of the probability of that event. Alternatively, a frequency per year or per event (in units of time) may be used. Consequence C_i, is a measure of the impacts of event E_i. This can be in the form of mission loss, payload damage, damage to property, number of injuries, number of fatalities, dollar loss, etc.

The results of the risk estimation are then used to interpret the various contributors to risk, which are compared, ranked, and placed in perspective. This process consists of:

1. Calculating and graphically displaying a risk profile based on individual failure event risks, similar to the process presented in Figure 8.1. This method will be discussed in more details in this section.

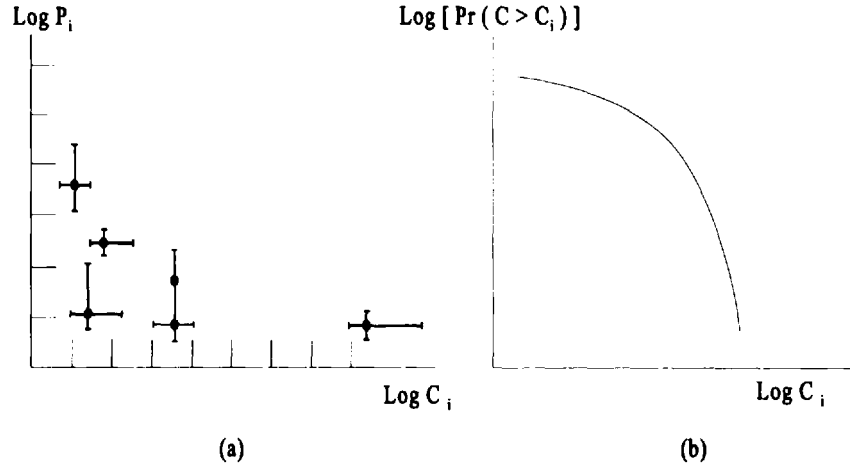


Figure 8.1 Construction of a risk profile.

2. Calculating a total expected risk value R from

$$R = \sum_i P_i \times C_i \tag{8.1}$$

Naturally, all the calculations described involve some uncertainties, approximations, and assumptions. Therefore, uncertainties must be considered explicitly, as discussed in Section 7.3. Using expected losses and the risk profile, one can evaluate the amount of investment that is reasonable to control risks, alternative risk-management decisions to avoid risk (i.e., decrease the risk probability), and alternative actions to mitigate consequences. Therefore, the following two additional planning steps are usually included in risk analysis:

1. Identification of cost-effective risk management alternatives
2. Adoption and implementation of risk-management methods

The risk estimation results are often shown in a general form similar to (8.1). There are two useful ways to interpret such results: determining expected risk values, R_i, and constructing risk profiles. Both methods are used in quantitative risk analysis.

Expected values are most useful when the consequences C_i are measured in financial terms or other directly measurable units. The expected risk value R_i (or expected loss) associated with event E_i is the product of its probability P_i and consequence values, as described by (8.1). Thus, if the event occurs with a frequency of 0.01 per year, and if the associated loss is $1 million, then the expected loss (or

risk value) is: $R_i = 0.01 \times \$1{,}000{,}000 = \$10{,}000$. Conversely, if the frequency of event occurrence is 1 per year, but the loss is \$10,000, the risk value is still $R_i = 1 \times \$10{,}000 = \$10{,}000$. Thus, the risk value for these two situations is the same, i.e., both events are equally risky.

Table 8.4 General Form of Output from the Analytic Phase of Risk Analysis

Undesirable Event	Likelihood	Consequences	Risk Level
E_1	P_1	C_1	$R_1 = P_1C_1$
E_2	P_2	C_2	$R_2 = P_2C_2$
E_3	P_3	C_3	$R_3 = P_3C_3$
.	.	.	.
.	.	.	.
.	.	.	.
E_n	P_n	C_n	$R_n = P_nC_n$

Since this is the expected annual loss, the total expected loss over 20 years (assuming a constant dollar value) would be \$200,000. This assumes the parameters do not vary significantly with time, and ignores the low probability of multiple losses over the period. Expression (8.1) can be used to obtain the total expected loss per year for a whole set of possible events. This expected loss value assumes that all events (E_i) contributing to risk exposure have equal weight. Occasionally, for risk decisions, value factors (weighting factors) are assigned to each event contributing to risk. The relative values of the terms associated with the different hazardous events give a useful measure of their relative importance, and the total risk value can be interpreted as the average or "expected" level of loss over a period of time.

As discussed earlier another method for interpreting the results is construction of a risk profile. With this method, the probability values are plotted against the consequence values. Figure 8.1 illustrates these methods. Figure 8.1a shows the use of logarithmic scales, which are usually used because one can cover a wide range of values. The error brackets denote uncertainties in the probability estimate (vertical) and the consequences (horizontal). This approach provides a means of easily illustrating events with high probability, high consequence, or high uncertainty. It is useful when discrete probabilities and consequences are known. Figure 8.1b shows the construction of the complementary cumulative probability risk profile (sometimes known as a Farmer's curves (1960)). In this case, the logarithm of the probability that the total consequence C exceeds C_i is plotted against the logarithm of C_i. The most notable application of this method was in the landmark

Reactor Safety Study (1975). With this method, the low probability/high consequence risk values and high probability/low consequence risk values can be easily seen. That is, the extreme values of the estimated risk can be easily displayed.

8.3 FORMALIZATION OF RISK ASSESSMENT

The hazardous events E_i discussed in the previous section can occur as a result of a chain of basic events. In combination, these events are called a "scenario." The risk-assessment process is therefore primarily one of scenario development, with the risk contribution from each possible scenario that leads to the outcome or event of interest. This concept is described in terms of the triplet represented by (1.4). Because the risk-assessment process focuses on scenarios that lead to hazardous events, the general methodology becomes one that allows the identification of all possible scenarios, calculation of their individual probabilities, and a consistent description of the consequences that result from each. Scenario development requires a set of descriptions of how a barrier confining a hazard is threatened, how the barrier fails, and the effects on the subject when it is exposed to the uncontained hazard. This means that one needs to formally address the items described below.

Identification of Hazards

A survey of the process under analysis should be performed to identify the hazards of concern. These hazards can be categorized as follows:

Chemical hazard (e.g., toxic chemicals released from a chemical process)
Thermal hazard (e.g., high-energy explosion from a chemical reactor)
Mechanical hazard (e.g., kinetic energy from a moving object)
Electrical hazard (e.g., potential difference, electrical and magnetic fields, electrical shock)
Ionizing radiation (e.g., radiation released from a nuclear plant)
Nonionizing radiation (e.g., radiation from a microwave oven)
Biological hazard (e.g., spread of certain bacteria)

Presumably, each of these hazards will be part of the process and normal process boundaries will be used as their containment. This means that, provided there is no disturbance in the process, the barrier that contains the hazard will be unchallenged. However, in a risk scenario one postulates the challenges to such barriers and tries to estimate the probability of these challenges.

Identification of Barriers

Each of the identified hazards must be examined to determine all the physical barriers that contain it or can intervene to prevent or minimize exposure to the hazard. These barriers may physically surround the hazard (e.g., walls, pipes, valves, fuel clad, structures); they can be based on a specified distance from a hazard source to minimize exposure to the hazard (e.g., minimize exposure, to radioactive materials); or they may provide direct shielding of the subject from the hazard (e.g., protective clothing, bunkers).

Identification of Challenges to Barriers

Identification of each of the individual barriers is followed by a concise definition of the requirements for maintaining each one. This can be done by developing an analytical model that has a hierarchical character. One can also simply identify what is needed to maintain the integrity of each barrier. These are due to the degradation of strength of the barrier and high stress in the barrier.

Barrier strength degrades because of:

reduced thickness (due to deformation, erosion, corrosion etc.),
change in material properties (e.g., toughness, yield strength). This may be affected by the local environment, e.g., temperature).

Stress on the barrier increases by:

internal forces or pressure,
penetration or distortion by external objects or forces.

The above causes of degradation are often the result of one or more of the following conditions:

Malfunction of process equipment (e.g., the emergency cooling system in a nuclear plant)
Problems with man-machine interface
Poor design or maintenance
Adverse natural phenomena
Adverse human-made environment.

Estimation of Hazard Exposure

The next step in the risk-assessment procedure is to define those scenarios in which the barriers may be breached, and then make the best possible estimate

of the probability or frequency for each sequence. Those scenarios that pose similar levels of hazard under similar conditions of hazard dispersal are grouped together, and the probabilities or frequencies of the respective event sequences associated with these groups are determined.

Consequences Evaluation

The range of effects produced by exposure to the hazard may encompass harm to people, damage to equipment, and contamination of land or facilities. These effects are evaluated from knowledge of the toxic behavior of the particular material(s) and the specific outcomes of the scenarios considered. In the case of the dispersal of toxic materials, the size of the release is combined with the potential dispersion mechanisms to calculate the outcome.

From the generic nature of risk analysis, there appears to be a common approach to understanding the ways in which hazard exposure occurs. This understanding is key in the development of logical scenario models that can then be solved. Quantitative and qualitative solutions can provide estimates of barrier adequacy and methods of effective enhancement. This formalization provides a basis from which we can describe a commonly used practice in risk analysis called probabilistic risk assessment (PRA). This technique, pioneered by the nuclear industry, is the basis of a large number of formal risk assessments today. We describe this approach in Section 8.4 and provide an example in Section 8.5.

8.4 STEPS IN CONDUCTING A PROBABILISTIC RISK ASSESSMENT

The following subsections provide a discussion of the basic elements of PRA as we walk our way through the steps that must be performed. We also describe the methods that are useful for this analysis as described in previous chapters of the book. Figure 8.2 illustrates the general PRA process.

8.4.1 Methodology Definition

Preparing for a PRA begins with a review of the objectives of the risk analysis. Aninventory of possible techniques for the desired analysis should be developed. The available techniques range from required computer codes to facility experts and analytical experts. This, in essence, provides a road map for the analysis. The methods described in the preceding chapters of this book discussed most of the techniques currently used for PRA.

The resources required for each analytical option should be evaluated, and the most cost-effective option selected. The basis for the selection should be

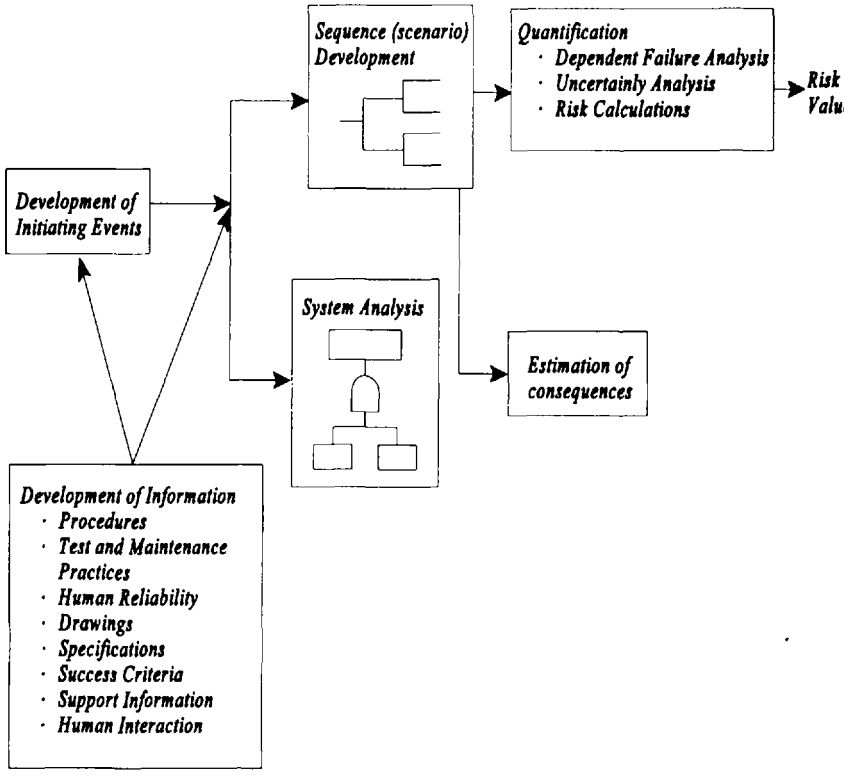


Figure 8.2 The process of probabilities risk analysis.

documented briefly, and the selection process reviewed to ensure that the objectives of the analysis will be adequately met.

8.4.2 Familiarization and Information Assembly

A general knowledge of the physical layout of the system or process (e.g., facility, plant, design), administrative controls, maintenance and test procedures, as well as protective systems whose functions maintain safety, is necessary to begin the PRA. All systems, locations, and activities expected to play a roll in the initiation, propagation, or arrest of an upset or hazardous condition must be understood in sufficient detail to construct the models necessary to capture all possible scenarios. A detailed inspection of the process must be performed in the areas expected to be of interest and importance to the analysis.

The following items should be considered in this step:

1. Major safety and emergency systems (or methods) should be identified.
2. Physical interactions among all major systems should be identified and explicitly described. The result should be summarized in a dependency matrix.
3. Past major failures and abnormal events that have been observed in the facility should be noted and studied. Such information would help ensure inclusion of important applicable scenarios.
4. Consistent documentation is key to ensuring the quality of the PRA. Therefore, a good filing system must be created at the outset, and main tained throughout the study.

With the help of designers, operators, or owners, one should determine the ground rules for the analysis, the scope of the analysis, and the configuration to be analyzed. One should also determine the faults and conditions to be included or excluded, the operating modes of concern, the freeze date design, and the hardware configuration on the design freeze date. The freeze date is an arbitrary date after which no additional changes in the facility design and configuration will be modeled. Therefore, the results of the PRA are only applicable to the facility at the freeze date.

8.4.3 Identification of Initiating Events

This task involves identifying those events (abnormal events) that could, if not correctly responded to, result in hazard exposure. The first step involves identifying sources of hazard and barriers around these hazards. The next step involves identifying events that can lead to a direct threat to the integrity of the barriers.

A system or process may have one or more operational modes which produce its output. In each operational mode, specific functions are performed that result in the output. Each function is directly related to one or more systems that perform the necessary functional actions. These systems, in turn, are composed of more basic units (e.g., components) that accomplish the objective of the system. As long as a system is operating within its design parameter tolerances, there is little chance of challenging the system boundaries in such a way that hazards will escape those boundaries. These operational modes are called normal operation modes.

During normal operation mode loss of certain functions or systems will cause the process to enter an off-normal condition. Once in this condition, there are two possibilities. First, the state of the process could be such that no other function is required to maintain the process in a safe condition. (*safe* refers to a

mode where the chance of exposing hazards beyond the facility boundaries is negligible.) The second possibility is a state wherein other functions or systems are required to prevent exposing hazards beyond the system boundaries. For this second possibility, the loss of a functional or loss of a system is an initiating event. Since such an event is related to the operating process equipment, it is called. an operational initiating event.

Operational initiating events can also apply to shutdown and start-up modes of the process. The terminology remains the same since, for a shutdown or start-up procedure, certain equipment must be functioning. For example, an operational initiating event found during the PRA of a test nuclear reactor was Low Primary Coolant System Flow. Flow is required to transfer heat produced in the reactor to heat exchanges and ultimately to the cooling towers and the air. If this coolant flow function is reduced to the point where an insufficient amount of heat is transferred, core damage could result. Therefore, another protective system must operate to remove the heat produced by the reactor. By definition, then, Low Primary Coolant System Flow is an operational initiating event.

One method for determining the operational initiating events begins with first drawing a functional diagram of the facility (similar to the MLD method described in Chapter 4). From the functional diagram, a hierarchical relationship is produced, with the process objective being successful completion of the desired process. Each function can then be decomposed into its systems, and components can be combined in a logical manner to represent success of that function. (Figure 8.3 illustrates this hierarchical decomposition). Potential initiating events are the failures of particular functions, systems, or components, the occurrence of which causes the process to fail. These potential initiating events are grouped such that members of a group require similar process system and safety system responses to cope with the initiators. These groupings are the operational initiator categories.

An alternative to the use of functional hierarchy for identifying initiating events is the use of FMEA, discussed in Chapter 4. The difference between these two methods is noticeable, namely, the functional hierarchy method is deductive and systematic, whereas FMEA is inductive. The use of FMEA for identifying initiating events consists of identifying failure events (modes of failure) whose effect is a threat to hazard barriers. In both of the above methods, one can always supplement the set of initiating events with generic initiating events (if known). For example, see NUREG/CR-4550 (1990) for these initiating events for nuclear reactors.

To simplify the process, it is necessary, after identifying all initiating events, to combine those initiating events that pose the same threat to hazard barriers and require the same mitigating functions of the process to prevent hazard exposure. The following inductive procedures should be followed when grouping initiating events:

1. Combine the initiating events that directly break all hazard barriers.
2. Combine the initiating events that break the same hazard barriers (not necessarily all the barriers).
3. Combine the initiating events that require the same group of mitigating personnel or automatic actions following their occurrence.
4. Combine the initiating events that simultaneously disable the normal process as well as some of the available mitigating human or automatic actions.

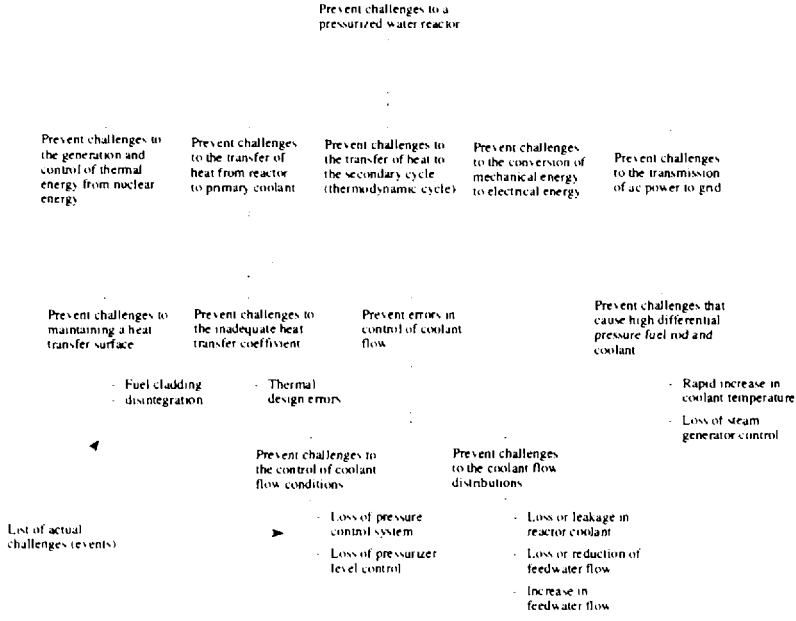


Figure 8.3 Partial goal tree to determine challenges to a pressurized water reactor.

Events that cause off-normal operation of the facility and require other systems to operate to maintain process materials within their desired boundaries, but are not directly related to a process system or component, are nonoperational initiating events. Nonoperational initiating events are identified with the same methods used to identify operating events. However, the events of interest are those that are primarily external to the facility. These are discussed in more detail in Sections 8.4.6 and 8.4.7.

The following procedures should be followed in this step of the PRA:

1. Select a method for identifying specific operational and nonopertional initiating events. Two representative methods are functional hierarchy and FMEA. If a generic list of initiating events is available, it can be used as a supplement.
2. Using the method selected, identify a set of initiating events.
3. Group the initiating events such that those having the same effect on the process and requiring the same mitigating functions to prevent hazard exposure are grouped together.

8.4.4 Sequence or Scenario Development

The goal of scenario development is to derive a complete set of scenarios that encompasses all of the potential propagation paths that can lead to loss of confinement of the hazard following the occurrence of an initiating event. To describe the cause and effect relationship between initiators and the event progression, it is necessary to identify those functions (e.g., safety functions) that must be maintained to prevent loss of hazard barriers. The scenarios that describe the functional response of the process to the initiating events are frequently displayed by event-trees.

As discussed in Chapter 4, event trees order and depict (in approximately chronological manner) the success or failure of key mitigating actions (e.g., human actions or mitigative hardware that automatically responds) that are required to respond following an initiating event. In PRA, two types of event trees can be developed: functional and systemic. The functional event tree uses mitigating functions as its heading. The main purpose of the functional tree is to better understand the scenario of events at a high level following the occurrence of an initiating event. The functional tree also guides the PRA analyst in the development of a more detailed systemic event tree. The systemic event tree reflects the mitigative scenarios of specific events (specific human actions or mitigative system operations or failures) that lead to a hazardous outcome. That is, the functional event tree can be further decomposed to show specific hardware or human actions that perform the functions described in the functional event tree. Therefore, a systemic event tree fully delineates the process or system response to an initiating event and serves as the main tool for further analysis in the PRA.

The following procedures should be followed in this step of the PRA:

1. Identify the mitigating functions for each initiating event (or group of events).

2. Identify the corresponding human actions, systems or hardware operations associated with each function, along with their necessary conditions for success.
3. Develop a functional event tree for each initiating event (or group of events).
4. Develop a systemic event tree for each initiating event, delineating the success conditions, initiating event progression phenomena, and end effect of each scenario.

8.4.5 System Analysis

Event trees commonly involve branch points at which a given system (or event) either works (or happens) or does not work (or does not happen). Sometimes, failure of these systems (or events) is rare and there may not be an adequate record of observed failure events to provide a dependable database of failure rates. In such cases, other system analysis methods described in Chapter 4 may be used, depending on the accuracy desired. The most common method used in PRA to calculate the probability of system failure is fault tree analysis. This analysis involves developing a system model in which the system is broken down into basic components or modules for which adequate data exist. In Chapter 4, we discussed how a fault tree can represent the event headings of an event tree.

Different event-tree modeling approaches imply variations in the complexity of the system models that may be required. If only main functions or systems are included as event-tree headings, the fault trees become more complex and must accommodate all dependencies among front-line and support functions (or systems) within the fault tree. If support functions (or systems) are explicitly included as event-tree headings, more complex event trees and less complex fault trees will result.

The following procedures should be followed as a part of developing the fault tree:

1. Develop a fault tree for each event in the event tree heading.
2. Explicitly model dependencies of a system on other systems and inter component dependencies (e.g., common cause failure as described in Section 7.2).
3. Include all potential causes of failure, such as hardware, software, test and maintenance, and human error, in the fault tree.

8.4.6 Internal Events External to the Facility

Events that originate within a complex system are called internal events. Events that adversely affect the process and occur outside of the facility boundaries, but

within the facility, are defined as internal events external to the facility. Typical internal events external to the process are internal fires, internal floods, and high-energy events within the complex system. The effects of these events should be modeled with event trees to show all possible scenarios.

8.4.7 External Events

The clear counterpoint to the type of initiating event discussed in Section 8.4.6 is an initiating event that originates outside of the complex system, called an external event. Examples of external events are fires and floods that originate outside of the system, seismic events, transportation events, volcanic events, and high-wind events. Again, this classification can be used in grouping the event-tree scenarios.

8.4.8 Dependent Failure Considerations

To attain the very low levels of risk, the systems and hardware that comprise the barriers to hazard exposure must have very high levels of reliability. This high reliability is typically achieved through the use of redundant and/or diverse hardware, which provides multiple success paths. The problem then becomes one of ensuring the independence of the paths, since there is always some degree of coupling between their failure mechanisms, either through the operating environment (events external to the hardware) or through functional and spatial dependencies. In Section 7.2, we elaborated on the nature and mathematics of these dependencies. Treatment of dependencies should be carefully included in both event-tree and fault-tree development and analysis in PRA. As the reliability of individual systems and subsystems increases due to redundancy, the contribution from dependent failures becomes more important; at some point, dependent failures may dominate the overall reliability. Including the effects of dependent failures in the reliability models is difficult and requires some sophisticated, fully integrated models be developed and used to find those failure combinations that lead to mission failure. The treatment of dependent failures is not just a single step performed during the PRA; it must be considered throughout the analysis (e.g., in event trees, fault trees, and human actions).

The following procedures should be followed in the dependent failure analysis:

1. Identify the items that are similar and could cause dependent or common cause failures. For example, similar pumps, motor-operated valves, air-operated valves, diesel generators, and batteries are major components in process plants, and are considered important sources of common cause failures.

2. Items that are potentially susceptible to common cause failure should be explicitly incorporated into the fault trees and event trees where appli cable.
3. Functional dependencies should be identified and explicitly modeled in the fault trees and event trees.

8.4.9 Failure Data Analysis

A critical building block in assessing the reliability and availability of items in complex systems is the failure data on the performance of items. In particular, the best resources for predicting future availability of equipment are past experiences or tests. Component reliability data are inputs to system reliability studies, and the validity of the results depends highly on the quality of the input information. It must be recognized, however, that historical data have predictive value only to the extent that the conditions under which the data were generated remain applicable. Collection of the various component failure data consists essentially of the following steps: collecting generic data, assessing generic data, statistically evaluating facility-specific data, and specializing the failure probability distributions using facility-specific data. Three types of events identified during the accident-sequence definition and system modeling must be quantified for the event trees and fault trees to estimate the frequency of occurrence of sequences: initiating events, component failures and human errors.

The quantification of initiating events and components failure probabilities involves two separate activities. First, the probabilistic model for each event must be established; then the parameters of the model must be estimated. The necessary data include component failure rates, repair times, test frequencies, test downtimes, common-cause probabilities, and uncertainty characterizations. In Chapter 3 we discussed available methods for analyzing data to obtain the probability of failure or the probability of occurrence of equipment failure. In Chapter 5 we discussed analysis of data relevant to repairable systems. Finally, in Chapter 6 we discussed analysis of data for dependent failures and human reliability. The establishment of the database to be used will generally involve the collection of some equipment or facility-specific data or the use of generic reliability databases.

The following procedures should be followed as part of the data analysis task:

1. Determine generic values of failure rate and failure on demand proba bilities for each component identified in the fault-tree analysis. This can be obtained either from facility-specific experiences or from generic sources of data (see Chapter 3.)
2. Determine test, repair, and maintenance outages primarily from experi ence, if available. Otherwise use generic sources.

3. Determine the frequency of initiating events and other component failure events from experience, expert judgement, or generic sources. (see Chapters 3 and 7.)
4. Determine the common cause failure probability for similar items, primarily from generic values. However, when significant specific data are available, they can be used (see Chapter 7.)

8.4.10 Quantification

Fault-tree/event-tree sequences are quantified to determine the frequencies of scenarios and associated uncertainties in the calculation. The approach depends somewhat on the manner in which system dependencies have been handled. We will describe the more complex situation in which the fault trees are not independent, i.e., there are dependencies (e.g., through support systems).

Normally, the quantification will use a Boolean reduction process to arrive at a Boolean representation for each sequence. Starting with fault-tree models for the various systems or event headings in the event trees, and using probability estimates for each of the events in the fault trees, the probability of each event-tree heading is obtained (if the heading is independent of other headings). The fault trees for support systems (e.g., cooling, power) are merged where needed with the front-line systems (i.e., systems that utilize main factions of the facility) and converted into Boolean equation representations. The equations are solved for the minimal cut-sets for each of the front-line systems (those identified as headings on the event trees). The minimal cut-sets for the front-line systems are then appropriately combined to determine the cut-sets for the event-tree sequences. The process is described in Chapter 4.

If all possible cut-sets are retained during this process, an unmanageably large collection of terms will almost certainly result. Therefore, the collection of cut-sets is truncated (i.e., insignificant members are discarded based on the number of terms in a cut-set or on the probability of the cut-set.) This is usually a practical necessity because of the overwhelming number of cut-sets that can result from the combination of a large number of failures, even though the probability of any of these combinations may be vanishingly small. The truncation process does not disturb the effort to determine the dominant scenarios since we are discarding scenarios that are very often unlikely.

A valid concern is sometimes voiced that even though the individual discarded cut-sets may be at least several orders of magnitude less probable than the average of those retained, the large number of them might represent a significant part of the risk. The actual risk might thus be considerably larger than the PRA results indicate. Detailed examination of a few PRA studies of nuclear power plants show that truncation did not have a significant effect on the total risk assess-

ment results in those particular cases. The process of quantification is generally straightforward, and the methods used are described in Chapter 4. More objective truncation methods are discussed by Dezfuli and Modarres (1985).

The following procedures should be followed as part of the quantification process:

1. Merge corresponding fault trees associated with each failure or success event in the event tree sequences (i.e., combine them in a Boolean form). Develop a reduced Boolean function for each sequence.
2. Calculate the total frequency of each sequence, using the frequency of initiating events, the probability of hardware failure, test and mainten ance frequency (outage), common cause failure probability, and human error probability.
3. Use the minimal cut-sets of each sequence for the quantification process. If needed, simplify the process by truncating based on the cut-sets or probability.
4. Calculate the total frequency of each sequence.

8.5 A SIMPLE EXAMPLE OF RISK ANALYSIS

Consider the fire protection system shown in Figure 8.4. This system is designed to extinguish all possible fires in a plant with toxic chemicals. Two physically independent water extinguishing nozzles are designed such that each is capable of controlling all types of fires in the plant. Extinguishing nozzle 1 is the primary method of injection. Upon receiving a signal from the detector/alarm/actuator device, pump-1 starts automatically, drawing water from the reservoir tank and injecting it into the fire area in the plant. If this pump injection path is not actuated, plant operators can start a second injection path manually. If the second path is not available, the operators will call for help from the local fire department, although the detector also sends a signal directly to the fire department. However, due to the delay in the arrival of the local fire department, the magnitude of damage would be higher than it would be if the local fire extinguishing nozzles were available to extinguish the fire. Under all conditions, if the normal off-site power is not available due to the fire or other reasons, a local generator would provide electric power to the pumps. The power to the detector/alarm/actuator system is provided through the batteries, which are constantly charged by the off-site power. Even if the ac power is not available, the dc power provided through the battery is expected to be available at all times. The manual valves on the two sides of pump-1 and pump-2 are normally open, and only remain closed when they are being repaired. The entire fire system and generator are located outside of the

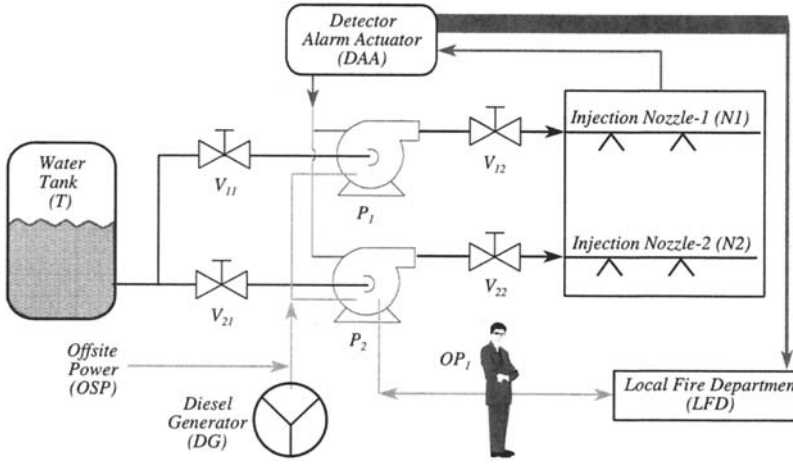


Figure 8.4 A fire protection system.

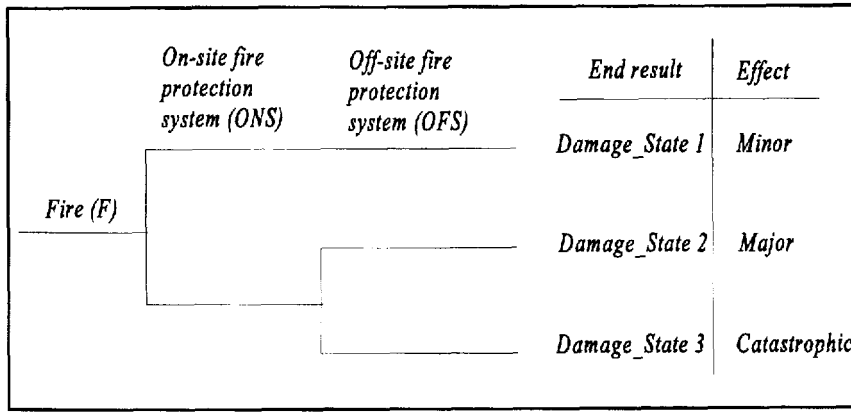


Figure 8.5 Scenario of events following a fire using the event-tree methods.

reactor compartment, and are therefore not affected by an internal fire. The risk-analysis process for this situation consists of the steps explained below.

1. Identification of Initiating Events

In this step, all events that lead to or promote a fire in the reactor compartment must be identified. These should include equipment malfunctions, human errors, and facility conditions. The frequency of each event should be estimated. Assuming that all events would lead to the same magnitude of fire, the ultimate initiating event is a fire, the frequency of which is the sum of the frequencies of the individual fire-causing events. Assume for this example that the frequency of fire is estimated at 1×10^{-6} per year. Since fire is the only challenge to the plant in this example, we end up with only one initiating event. However, in more complex situations, a large set of initiating events can be identified, each posing a different challenge to the plant.

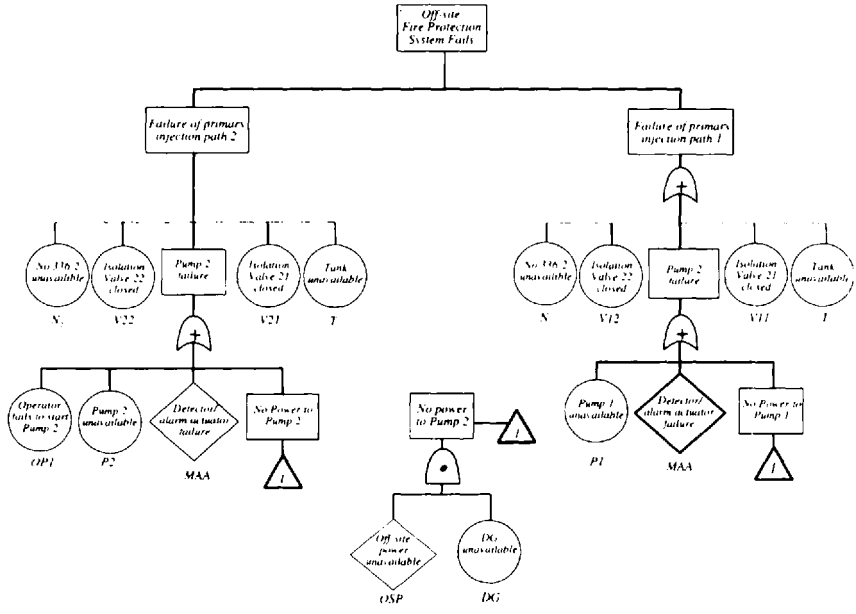


Figure 8.6 Fault tree for on-site fire protection system failure.

2. Scenario Development

In this step, we should explain the cause and effect relationship between the fire and the progression of events following the fire. We will use the event-tree method to depict this relationship. Generally, this is done inductively, and the level of detail considered in the event tree is somewhat dependent on the analyst. Two

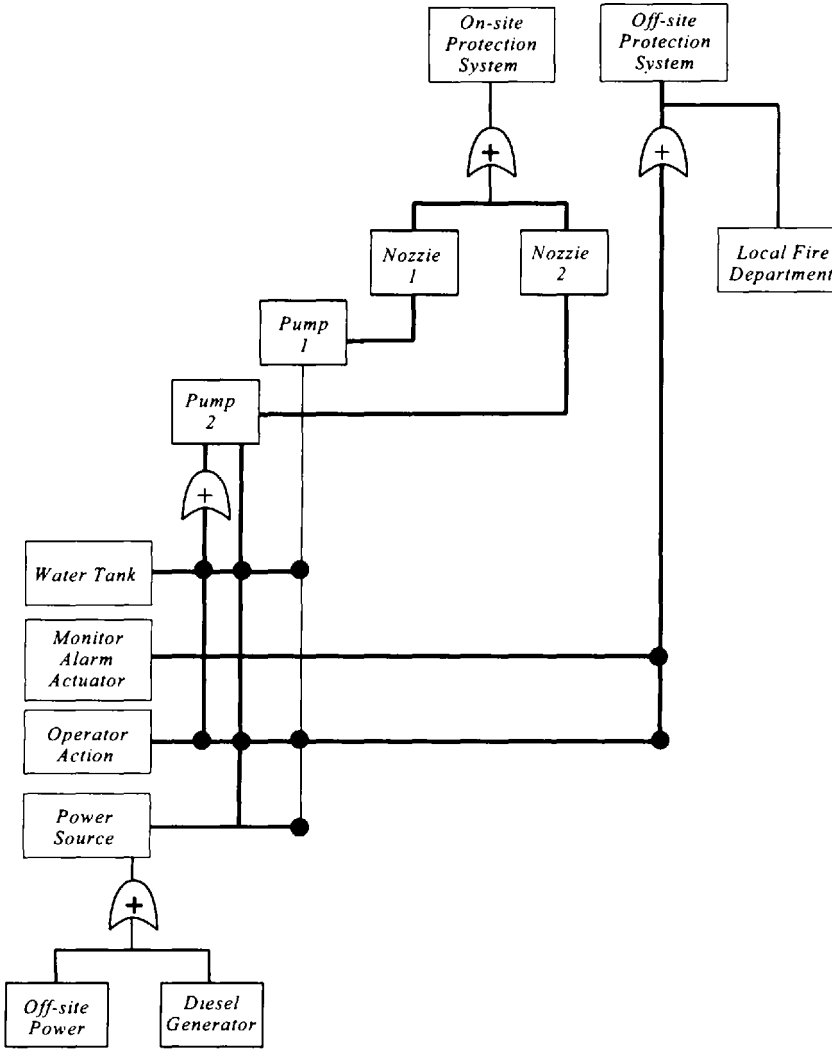


Figure 8.7 Fault tree for off-site fire protection system failure.

protective measures have been considered in the event tree shown in Figure 8.5: on-site protective measures (on-site pumps, tanks, etc.), and off-site protective fire department measures. The selection of these measures is based on the fact that availability or unavailability of the on-site or off-site protective measures would lead to different states of plant damage.

3. System Analysis

In this step, we should identify all failures (equipment or human) that lead to failure of the event-tree headings (on-site or off-site protective measures).

For example, Figure 8.6 shows the fault tree developed for the on-site fire protection system. In this fault tree, all basic events that lead to the failure of the two independent paths are described. Note that MAA, electric power to the pumps, and the water tank are shared by the two paths. Clearly these are considered as physical dependencies. This is taken into account in the quantification step of the risk analysis. In this tree, all external event failures and passive failures are neglected.

Figure 8.7 shows the fault tree for the off-site fire protection system. This tree is simple since it only includes all failures that does not lead to an on-time response from the local fire department.

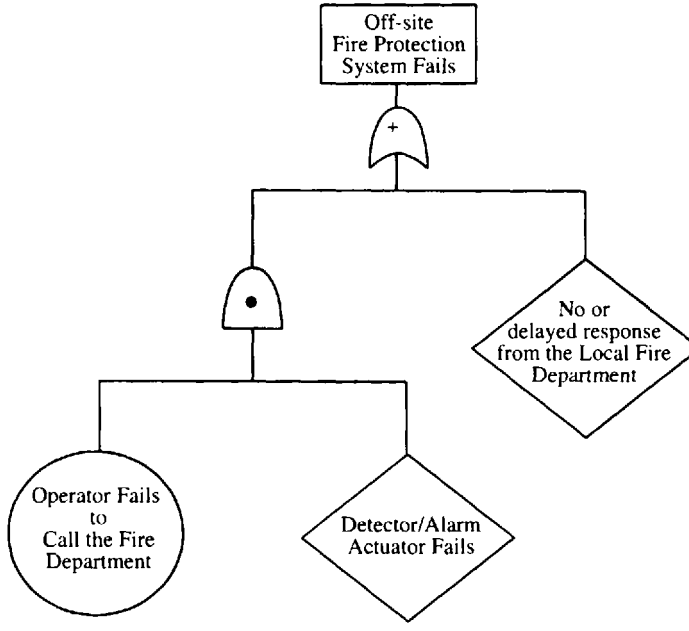


Figure 8.8 MLD for the fire protection system.

It is also possible to use the master logic diagram (MLD) for system analysis. An example of the MLD for this problem is shown in Figure 8.8. However, here only the fault trees are used for risk analysis, although MLD can also be used.

Table 8.5 Sources of Data and Failure Probabilities

Failure event	Plant-specific experience	Generic data	Probability used	Comments
Fire initiation frequency	No such experience in 10 years of operation.	5 fires in similar plants. There are 70,000 plant-years of experience.	$F = 5/70{,}000 = 7.1E-4$/yr.	Use generic data.
Pump 1 and Pump 2 failure	4 failure of two pumps to start. Monthly tests are performed which takes negligible time. Repair time takes about 10 hours at a frequency of 1 per year. No experience of failure to run.	Failure to run = $1E-5$/hr.	$\frac{4}{2 \times 12 \times 10} = 1.7E-2$/demand. Unavailability $= 1.7E-2 + \frac{10}{8760}$ $= 1.8E-2$/demand. Failure to run $= 1E-5$/hr. $P_t = P_2 = 1.7E-2 + 1E-5 \times 10 \approx 1.7E-2$	For failure to start, use plant-specific data. For failure to run, use generic data. If possible, use Bayesian updating technique described in Section 3.6. Assume 10 years of experience and 8760 hours in one year.
Common cause failure between Pump 1 and Pump 2	No such experience	Using the β-factor method, $\beta = 0.1$ for failure of pumps to start.	Unavailability due to common cause failure: $CCF = 0.1 \times 1.8E-2 = 1.8E-3$/demand.	Assume no significant common cause failure exists between valves and nozzles. See Section 7.2 for more detail.

Table 8.5 Continued

Failure event	Plant-specific xperience	Generic data	Probability used	Comments
Failure of isolation valves	1 failure to leave the valve in open position following a pump test	Not used.	$V_{11} = V_{12} = V_{21} = V_{22}$ $= \frac{1}{(10)(12)(2)}$ $= 4.2E-3$/demand.	Plan-specific data used.
Failure of nozzles	No-such experience	1×10^{-5}/demand	$N_1 = N_2 = 1.0E-5$/demand.	Generic data used.
Diesel generator failure	3 failures in monthly tests. 40 hours of repair per year.	$3E-2$/demand $3E-3$/hr 40 run	failure on demand $= 3/[(12)(10)]$ $= 2.5E-2$/demand. failure to run $= 3E-3$/hr Total Failure of DG $= 2.5E-2+3E-3\times 10 = 5.5E-2$.	Plant-specific data used for demand failure. Assume 10 years of experience.
Loss of off-site power	No experience.	0.1/yr.	$OSP = 0.1 \times \frac{10}{8760} = 1.1E-4$/demand.	Assume 104 hours of operation for fire extinguisher and use generic data.

Failure of MAA	No experience.	No data available.	MAA = 1*E* - 4/demand.	This estimate is based on expert judgement. See Section 6.3 for the methods.
Failure of operator to start Pump 2	No such experience	Using the THERP method for tasks of this kind, 1*E* - 2 is suggested.	OP_1 = 1*E* - 2/demand.	Use the THERP Handbook data discussed in Section 6.3.
Failure of operator to call the fire department	No such experience	1*E* - 3	OP_2 = 1*E* - 3/demand.	This is based on experience from no response to similar situations. Generic probability is used.
No or delayed response from fire department	No such experience	1*E* - 4	LFD = 1*E* - 4/demand.	This is based on response to similar cases from the fire department. Delayed/no arrival is due to accidents, traffic, communication problems, etc.
Tank failure	No such experience	1*E* - 5	T = 1*E* - 5/demand.	This is based on date obtained from rupture of the tank or insufficient water content.

4. Failure Data Analysis

It is important at this point to calculate the probabilities of the basic failure events described in the event trees and fault trees. As indicated earlier, this can be done by using either plant-specific data, generic data, or expert judgement. Table 8.5 describes the data used and their sources. It is assumed that at least 10 hours of operation is needed for the fire to be completely extinguished.

5. Quantification

To calculate the frequency of each scenario defined in Figure 8.5, we must first determine the cut-sets of the two fault trees shown in Figures 8.6 and 8.7. From this, the cut-sets of each scenario are determined, followed by calculation of the probabilities of each scenario based on the occurrence of one of its cut sets. These steps are described below.

1. The cut-sets of the on-site fire protection system failure are obtained using the technique described in Section 4.2. These cut-sets are listed in Table 8.6. Only cut-set number 22, which is failure of both pumps is subject to a common cause failure. This is shown by adding a new cut-set (cut-set number 24), which represents this common cause failure.
2. The cut-sets of the off-site fire protection system failure are similarly obtained and listed in Table 8.7.
3. The cut-sets of the three scenarios are obtained using the following Boo lean equations representing each scenario:

$$\text{Scenario–1} = F \cdot \overline{ONS}$$
$$\text{Scenario–2} = F \cdot ONS \cdot \overline{OFS}$$
$$\text{Scenario–3} = F \cdot ONS \cdot OFS\,.$$

   The process is described in Section 4.3.2.
4. The frequency of each scenario is obtained using data listed in Table 8.5. These frequencies are shown in Table 8.8.
5. The total frequency of each scenario is calculated using the rare event approximation. These are also shown in Table 8.8.

6. Consequences

In the scenario development and quantification tasks, we identified three distinct scenarios of interest, each with different outcomes and frequencies. The consequences associated with each scenario should be specified in terms of both economic and/or human losses. This part of the analysis is one of the most difficult for several reasons.

Table 8.6 Cut-Sets of the On-Site Fire Protection System Failure

Cut set no.	Cut set	Probability/ (% of total)
1	T	$1.0E-5$ (0.35)
2	MAA	$1.0E-4$ (3.5)
3	OSP · DG	$6.0E-6$ (0.21)
4	$N_2 \cdot N_1$	$1.0E-10$ (~0)
5	$N_2 \cdot V_{12}$	$4.2E-8$ (~0)
6	$N_2 \cdot P_1$	$1.7E-7$ (~0)
7	$N_2 \cdot V_{11}$	$4.2E-8$ (~0)
8	$V_{22} \cdot N_1$	$4.2E-8$ (~0)
9	$V_{22} \cdot V_{12}$	$1.8E-5$ (0.64)
10	$V_{22} \cdot P_1$	$7.1E-5$ (2.5)
11	$V_{22} \cdot V_{11}$	$1.8E-5$ (0.64)
12	$V_{21} \cdot N_1$	$4.2E-8$ (~0)
13	$V_{21} \cdot V_{12}$	$1.8E-5$ (0.35)
14	$V_{22} \cdot P_1$	$7.1E-5$ (2.5)
15	$V_{21} \cdot V_{11}$	$1.8E-5$ (0.64)
16	$OP_1 \cdot N_1$	$1.0E-7$ (~0)
17	$OP_1 \cdot V_{12}$	$4.2E-5$ (1.5)
18	$OP_1 \cdot P_1$	$1.7E-4$ (6.0)
19	$OP_1 \cdot V_{11}$	$4.2E-5$ (1.5)
20	$P_2 \cdot N_1$	$1.7E-7$ (~0)
21	$P_2 \cdot V_{12}$	$7.1E-5$ (2.5)
22	$P_2 \cdot P_1$	$2.9E-4$ (0.3)
23	$P_2 \cdot V_{11}$	$7.1E-5$ (2.5)
24	CCF	$1.8E-3$ (63.8)
		$Pr(ON) = \Sigma_i C_i$ $= 2.8E-3$

Table 8.7 Cut-Sets of the Off-Site Fire Protection System

Cut set no.	Cut set	Probability
1	LFD	$1E-4$
2	$OP_2 \cdot MAA$	$1E-7$
		Total $Pr^{(OFF)} \approx 1E-4$

1. Each scenario poses different hazards and methods of hazard exposure, and requires careful monitoring. In this case, the model should include the ways how the fire can spread through the plant, how people can be exposed, evacuation procedures, the availability of protective clothing, etc.

Table 8.8 Cut-Sets of the Scenarios

Scenario no.	Cut sets	Frequency	Comment
1	$F \cdot \overline{ON}$	$7.1E-4 \cdot (1-2.0E-2) = 7.1E-4$	Since the probability can be directly evaluated for $\overline{ON}$ without evaluating the need to generate cut sets, only the probability is calculated.
2	$F \cdot MAA \cdot LFD \cdot OP_2$	$7.0E-8$	1. Only cut sets from Table 7.6 that have a contribution greater than 1% are shown.
	$F \cdot V_{22} \cdot P_1 \cdot \overline{LFD} \cdot \overline{OP_2}$	$5.0E-8$	2. Cut set $F \cdot MAA \cdot \overline{LFD} \cdot \overline{MAA}$ is eliminated since $MAA \cdot \overline{MAA} = \phi$.
	$F \cdot V_{22} \cdot P_1 \cdot \overline{LFD} \cdot \overline{MAA}$	$5.0E-8$	
	$F \cdot V_{21} \cdot P_1 \cdot \overline{LFD} \cdot \overline{OP_2}$	$5.0E-8$	
	$F \cdot V_{22} \cdot P_1 \cdot \overline{LFD} \cdot \overline{MAA}$	$5.0E-8$	
	$F \cdot OP_1 \cdot V_{12} \cdot \overline{LFD} \cdot \overline{OP_2}$	$2.9E-9$	
	$F \cdot OP_1 \cdot V_{12} \cdot \overline{LFD} \cdot \overline{MAA}$	$2.9E-9$	
	$F \cdot OP_1 \cdot P_1 \cdot \overline{LFD} \cdot \overline{OP_2}$	$1.1E-7$	
	$F \cdot OP_1 \cdot P_1 \cdot \overline{LFD} \cdot \overline{MAA}$	$1.1E-7$	

2		2.9*E* - 9
		2.9*E* - 9
	$F \cdot OP_1 \cdot V_{11} \cdot \overline{LFD} \cdot \overline{OP_2}$	5.0*E* - 8
	$F \cdot OP_1 \cdot V_{11} \cdot \overline{LFD} \cdot \overline{MAA}$	5.0*E* - 8
	$F \cdot P_2 \cdot V_{12} \cdot \overline{LFD} \cdot \overline{OP_2}$	5.0*E* - 8
	$F \cdot P_2 \cdot V_{12} \cdot \overline{LFD} \cdot \overline{MAA}$	2.0*E* - 7
	$F \cdot P_2 \cdot P_1 \cdot \overline{LFD} \cdot \overline{OP_2}$	2.0*E* - 7
	$F \cdot P_2 \cdot P_1 \cdot \overline{LFD} \cdot \overline{MAA}$	5.0*E* - 8
	$F \cdot P_2 \cdot V_{11} \cdot \overline{LFD} \cdot \overline{OP_2}$	5.0*E* - 8
	$F \cdot P_2 \cdot V_{11} \cdot \overline{LFD} \cdot \overline{MAA}$	1.3*E* - 6
	$F \cdot CCP \cdot \overline{LFD} \cdot \overline{OP_2}$	1.3*E* - 6
	$F \cdot CCP \cdot \overline{LFD} \cdot \overline{MAA}$	

$$\sum_i = 1.3E - 6$$

1. Only cut sets from Table 7.6 that have a contribution greater than 1% are shown.
2. Cut set $F \cdot MAA \cdot \overline{LFD} \cdot \overline{MAA}$ is eliminated since $MAA \cdot \overline{MAA} = \phi$.

Table 8.8 Continued

Scenario no.	Cut-sets	Frequency	Comment
3	$F \cdot MAA \cdot LFD$	7.1×10^{-12}	1. Only cut sets from Tables 7.6 and 7.7 that have the highest contribution to the scenario are shown.
	$F \cdot V_{22} \cdot P_1 \cdot LFD$	5.0×10^{-12}	
	$F \cdot V_{21} \cdot P_1 \cdot LFD$	5.0×10^{-12}	
	$F \cdot OP_1 \cdot V_{12} \cdot LFD$	2.9×10^{-12}	
	$F \cdot OP_1 \cdot P_1 \cdot LFD$	2.8×10^{-12}	
	$F \cdot OP_1 \cdot V_{11} \cdot LFD$	2.9×10^{-12}	
	$F \cdot P_2 \cdot P_{12} \cdot LFD$	5.0×10^{-12}	
	$F \cdot P_2 \cdot P_1 \cdot LFD$	2.0×10^{-11}	
	$F \cdot P_2 \cdot V_{11} \cdot LFD$	5.0×10^{-12}	
	$F \cdot CCP \cdot LFD$	3.0×10^{-11}	
		$\sum_i = 8.4E-11$	

2. The outcome of the scenario can be measured in terms of human losses. It can also be measured in terms of financial losses, i.e., the total cost associated with the scenario. This involves assigning a dollar value to human life or casualties, which is a source of controversy.

Suppose a careful analysis of the spread of fire and fire exposure is performed, with consideration of the above issues, and ultimately results in damages measured only in terms of economic losses. These results are shown in Table 8.9.

The low value (in dollars) at risk indicates that fire risk is not important for this plant. However, scenarios 1 and 2 are significantly more important than scenario 3. Therefore, if the risk were high, one should improve those components that are major contributors to scenario 1 and 2. Scenario 1 is primarily due to common cause failure between pumps P_1 and P_2 so reducing this failure is a potential source of improvement.

Table 8.9 Economic Consequences of Fire Scenarios

Scenario number	Economic consequence
1	$1,000,000
2	$92,000,000
3	$210,000,000

7. Risk Calculation and Evaluation

Using values from Table 8.9, we can calculate the risk associated with each scenario. These risks are shown in Table 8.10.

Since this analysis shows that risk due to fire is rather low, uncertainty analysis is not very important. However, one of the methods described in Section 7.3 could be used to estimate the uncertainty associated with each component and the fire-initiating event if necessary. The uncertainties should be propagated through the cut sets of each scenario to obtain the uncertainty associated with the frequency estimation of each scenario. The uncertainty associated with the consequence estimates can also be obtained. When uncertainty associated with the consequence values are combined with the scenario frequencies and their uncertainty, the uncertainty associated with the estimated risk can be calculated. Although this is not a necessary step in risk analysis, it is reasonable to make an estimate of the uncertainties when risk values are high.

Figure 8.9 shows the risk profile based on the values in Table 8.10.

Table 8.10 Risk Associated With Each Scenario

Scenario number	Economic consequence
1	(7.1E - 4) ($1,000,000) = $710.000
2	(3.7) ($92,000,000) = $340.000
3	(8.4E - 11) ($210,000,000) = $0.017

8.6 PRECURSOR ANALYSIS

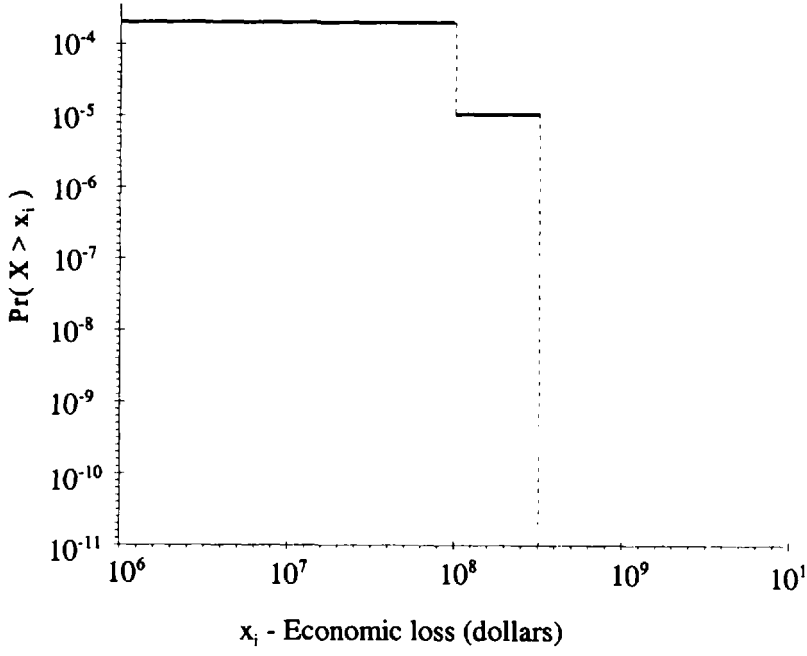


Figure 8.9 Risk profile.

8.6.1 Introduction

Risk analysis may be carried out by completely hypothesizing scenarios of events, which can lead to exposure of hazard, or may be based on actuarial scenarios of events. Sometimes, however, certain actuarial scenarios of events may have occurred without leading to an exposure of hazard, but involve a substantial erosion of barriers that prevent or mitigate hazard exposure. These scenarios are considered as precursors to accidents (exposure of hazard).

Accident *precursor events* or simply *precursor events* (PEs), in the reliability context given, can be defined as those operational events that constitute important elements of accident sequences leading to accidents (or hazard exposure) in complex systems, such as a severe core damage in a nuclear power plant, severe aviation or marine accidents, chemical plant accidents etc. The significance of a PE is measured through the conditional probability that the actual event or scenarios of events would result in exposure of hazard. In other words, PEs are those events that substantially reduce the margin of safety available for prevention of accidents.

Accident precursor analysis (APA) can be used as a convenient tool for complex system safety and performance monitoring and analysis. The APA methodology considered in this section is mainly based on the methodology developed for nuclear power plants (Modarres et al. (1996)), nevertheless, its application to other complex systems seems to be straightforward.

8.6.2 Basic Methodology

Considering a sequence of accidents in a system given as one following the homogeneous poisson process (HPP), the maximum likelihood estimate (MLE) for the rate of occurrence of accidents, λ, can be written as

$$\hat{\lambda} = \frac{n}{t} \tag{8.2}$$

where n is the total number of accidents observed in nonrandom exposure (or cumulative exposure) time t. The total exposure time can be measured in such units as *reactor-years* (for nuclear power plants), *aircraft hours flown*, *aircraft miles flown*, etc.

Because a severe accident is a rare event (i.e., n is quite small), estimator (8.2) cannot be applied, so one must resort to postulated events, whose occurrence would lead to the severe accident. The marginal contribution from each precursor event in the numerator of (8.2) can be counted as a positive number less than 1. For nuclear power plants Apostolakis and Mosleh (1979) have suggested using *conditional core damage probability* given a precursor event in the numerator of equation (8.2). Obviously this approach can be similarly used for other complex systems.

Considering all such precursor events that have occurred in exposure time t, the estimator (8.2) is replaced by

$$\hat{\lambda} = \frac{\sum_i p_i}{t} \tag{8.3}$$

where p_i is the conditional probability of a severe accident given precursor event i.

The methodology of precursor analysis has two major components—screening, i.e., identification of events with anticipated high p_i values, and quantification, i.e., estimation of p_i and λ, and developing corresponding trend analysis, as an indicator of the overall system(s) safety, which are discussed below.

8.6.3 Categorization and Selection of Precursor Events

The conditional probabilities of hazard exposure events given precursor events i ($i = 1, 2, \ldots$). p_i, are estimated based on the data collected on the observed operational events in order to identify those events that are above a threshold level. These events are known as *significant precursor events.* The process of estimating the p_is is rather straightforward. Events are mapped onto an event tree, and other failures, which eliminate remaining barriers, are postulated so as to complete a severe accident scenario. The event trees are developed the same way as in regular PRA methods. The probabilities that such postulated events occur are multiplied to estimate the conditional probability of a severe accident of interest.

The process of mapping an event i onto event trees and subsequently calculating the conditional probability p_i turns out to be time consuming. However, because the majority of the events are rather minor, only a small proportion of events—those which are expected to yield high p_i values (meet some qualitative screening criteria)—need to be analyzed. On the other hand to estimate the rate of occurrence of hazard exposure events, $\hat{\lambda}$, using equation (8.3), it would be advisable to include the risk significance of all precursor events because the more frequent but less significant events are not considered. For example, in a system having no events that meet some precursor selection criteria, (8.3) yields a zero estimate for $\hat{\lambda}$. However, provided the system may have had some other incidents with potentially small p_i values which do not meet the selection criteria chosen, the zero value of $\hat{\lambda}$ underestimates the system true rate of occurrence of hazard exposure events, λ. Therefore, a *background risk correction factor* that collectively accounts for these less serious incidents is sometimes introduced (Modarres et al. (1996)).

Additionally, when a system is shut down or not in use, some potentially risk-significant states and corresponding precursor events might be identified to avoid risk underestimation. Another potentially major underestimation of the rate of occurrence of severe accidents is associated with such very low-frequency high-consequence external events as earthquakes, floods, etc. Bier and Mosleh (1991) have discussed this problem using a Bayesian framework.

Ideally, the following expression for total annual (or another appropriatereference interval) frequency of occurrence of hazard exposure events, F(HE), can be used:

$$\begin{aligned} F(HE) = {} & F(\text{HE due to significant precursors}) + \\ & F(\text{HE during shutdown or not in use}) + \\ & F(\text{HE due to background events}) + \\ & F(\text{HE due to low-frequency high-consequence events}) \end{aligned}$$

8.6.4 Properties of Precursor Estimator for the Occurrence Rate of Hazard Exposure Events and Its Interpretation

Because the set of hazard exposure event (e.g., accidents) sequences corresponding to the observed precursor events usually overlap, it was shown (see Rubenstein (1985), Cooke et al. (1987), Bier (1993), Abramson (1994), Modarres et al. (1996)) that there is over counting in the numerator of (8.3), i.e., (8.3) is a positively biased estimator of λ, in contrast with MLE (8.2) which is generally unbiased. It is interesting to note that in the case when no failures are observed during time t (which is a typical situation for rare events), the estimate based on (8.2) takes on zero value, which in a sense, means a negative bias.

Bayesian interpretation of estimator λ is discussed by Modarres et al. (1996). Suppose we partition the total exposure time t into two distinct parts: (a) the exposure time t_1 associated with those systems in which all the precursor events (excluding actual severe accident events) have been observed, and (b) exposure time t_2 associated with the remaining systems in which no precursors (but including zero or more actual severe accident events) have been observed. Thus, $t = t_1 + t_2$.

Because we are interested in estimating the rate of occurrence of severe accidents λ, we consider the conjugate gamma prior distribution of λ (see Section 3.6) with shape parameter Σp_i and scale parameter t_1. Because the word *precursor* quite naturally means *prior* (to an actual event), we can interpret Σp_i as a prior pseudo number of prior (or precursor) events in prior (or precursor) exposure time t_1. Due to the over counting inherent in Σp_i, the positive bias mentioned before is likewise inherent in this prior distribution. In other words, the gamma prior is likely to be centered over values that are larger than λ.

Using Bayes' theorem to combine this gamma prior with the HPP data consisting of zero or more actual severe accident events in exposure time t_2 yields a

gamma posterior distribution of λ with shape parameter Σp_i and scale parameter t. The aforementioned partition of t also avoids overlap (or over counting) in Bayes' theorem. The mean of this gamma posterior (the Bayesian estimator of λ under square-error loss function) is given by (8.3). Depending on the magnitude of Σp_i and t, this posterior gamma distribution may be excessively positively skewed such that the posterior mean lies in the extreme right-hand tail. In such cases, the use of the posterior mean as a Bayesian point estimator may be undesirable and other more appropriate point estimators should be considered (such as the median). Using this gamma posterior, one can also calculate a corresponding Bayesian one- or two-sided probability interval estimate of λ.

To assess the appropriateness of using (8.3) as an estimator of the rate of occurrence of hazard exposure events λ, it is essential to evaluate the statistical properties of this estimator. To do this, one needs a probabilistic model for the number of precursor events and a model for the magnitude of the p_i values. Usually it is assumed that the number of precursors observed in exposure time t follows the HPP with a rate (intensity) μ, and p_i is assumed to be an independently distributed continuous random variable having a truncated (due to the threshold mentioned above) pdf $h(p)$. For the U.S. nuclear power plants examples considered below, the lower truncation value p_0, as a rule, is 10^{-6}.

Under these assumptions the estimator (8.3) can be written as

$$\hat{\lambda} = \frac{\sum_{i=1}^{N(t)} p_i}{t} \tag{8.4}$$

where the number of items in the numerator $N(t)$ has the Poisson distribution with mean μt, and the conditional probabilities p_i are all independent identically distributed according to pdf $h(p)$. Suppose now that $N(t) = n$ precursors have occurred in exposure time t, thus, $\hat{\mu} = n/t$. As it was mentioned, the exposure time t may be cumulative exposure time. For example, for the U.S. nuclear power plants, for the period 1984 through 1993, $n = 275$ precursors were observed in $t = 732$ reactor-year of operation (Modarres et al. (1996)); thus, $\hat{\mu} = 0.38$ precursors/reactor-year.

There exist numerous parametric and nonparametric methods that can be used to fit $h(p)$, based on the available values of p_i. Some parametric and nonparametric approaches are considered in (Modarres et al. (1996)).

For an appropriately chosen (or fitted) distribution $h(p)$, one is interested in determining the corresponding distribution of the estimate (8.4), from which one can then get any moments or quantiles of interest, such as the mean or 0.95th quantile. In general, it is difficult analytically to determine the distribution of $\hat{\lambda}$, therefore, Monte Carlo simulation is recommended as a universal practical approach.

The HPP model considered can be generalized by using the nonhomogeneous Poisson process (NHPP) model (introduced in Section 5.1) with intensity $\mu(t)$ for $N(t)$, which allows one to get an analytical trend for λ.

Another approach is based on the use of a truncated nonparametric pdf estimator of $h(p)$ (Scott, 1992, Chapter 6) and Monte Carlo simulation to estimate the distribution of $\hat{\lambda}$. This approach is known as the smooth bootstrap method.

An alternative but similar model can be obtained through the use of the extreme value theory. An analogous example for earthquakes is considered in (Castillo (1988)) in which the occurrence of earthquakes is treated as the homogeneous Poisson process, and severity (or *intensity* in geophysical terms) of each earthquake is assumed to be a positively defined random variable. It is clear that the conditional probability of hazard exposure p_i given a precursor considered, is analogous to earthquake severity given the occurrence of an earthquake.

To further illustrate the application of extreme value theory, suppose that we are interested in the distribution of the maximum value of conditional probability of severe accidents, which we denote by P_{max}, for exposure time t, based on random sample of size n precursors that occur in t. Let $H(p)$ denote the cumulative distribution function corresponding to $h(p)$. The distribution function of P_{max} for a nonrandom sample of size n is given by $H^n(p)$ (see Section 3.2.6). Because for the case considered n has the Poisson distribution with parameter μt, the cumulative distribution function of P_{max} becomes

$$H_{max}(p, t) = \sum_{n=0}^{\infty} \frac{e^{-\mu t}(\mu t)^n H^n(p)}{n!}$$

Using the MacLaurin expansion for an exponent, this relationship can be written as

$$H_{max}(p, t) = \exp\{-\mu t[1 - H(p)]\} \tag{8.5}$$

Correspondingly the probability that the maximum value is greater than p (probability of exceedance) is simply $1 - H_{max}(p, t)$. Equation (8.5) can be generalized for the case of the NHPP with the rate $\mu(t)$ as:

$$H_{max}(p, t_1, t_2) = \exp\left\{-\int_{t_1}^{t_2} \lambda(s)\, ds\, [1 - H(p)]\right\}$$

Using the corresponding sample (empirical) cumulative distribution function for the precursor events to estimate $H(p)$, it is possible to estimate the probability of exceeding any value p in any desired exposure time t. The corresponding example associated with nuclear power plant safety problems is given in the following section.

8.6.5 Applications of Precursor Analysis

From the discussion above it is obvious that the precursor analysis (PA) results can be used as follows:

To select and compare safety significance of operational events, which are then considered as major precursors

To show trends in the number and significance of the precursor events selected

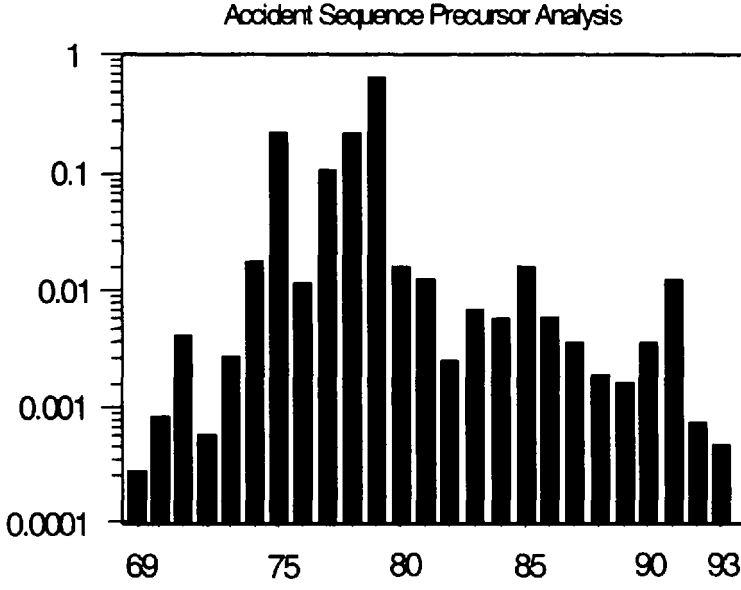


Figure 8.10 Annual sum of ASP conditional core damage probabilities.

Some examples of PA for the nuclear power plant data for the 1984 through 1993 period (Modarres et al. (1996)) are considered below. In the framework of nuclear power plant terminology "severe accident" is referred to as *core damage*, correspondingly the term *conditional probability of core damage* is used as a substitute of conditional probability of severe accidents.

The results of analysis of precursor data for the 1984 through 1993 period are given in Table 8.11. The table gives a breakdown of important precursors but it does not show trends in the occurrence of precursors as an indicator of overall plant safety. Figure 8.10 represents one such indicator. In this figure, the

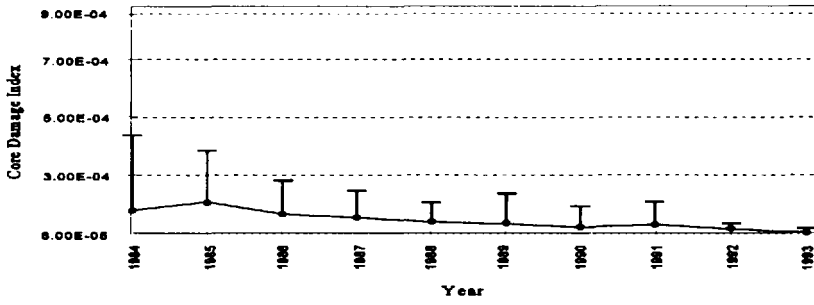


Figure 8.11 Trunicated lognormal distribution for $h(p)$.

Table 8.11 Analysis of Nuclear Power Plant Precursor Data for the 1984 through 1993 Period

Year	Cumulative reactor-years	Cumulative number of precursors, n_i	Cumulative Σp_i	Rate of occurrence of core damage/year [Equation (8.3)]
1984	52.5	32	0.00579	1.1E-4
1985	114.2	71	0.02275	2.0E-4
1986	178.1	89	0.02857	1.6E-4
1987	248.6	122	0.03268	1.3E-4
1988	324.7	154	0.03509	1.1E-4
1989	400.7	184	0.03741	9.3E-5
1990	481.4	212	0.04124	8.6E-5
1991	565.4	238	0.05124	9.0E-5
1992	649.1	262	0.05358	8.1E-5
1993	732.0	275	0.05440	7.2E-5

conditional core damage probabilities p_i of the precursors for each year are summed to calculate a value which is then used as an indicator of the overall safety of plants.

Provided the bias in (8.4) is constant or approximately constant, one can use the estimator to analyze an overall trend in the safety performance. The accumulated precursor data for 1984 through 1993 are used at the end of each subsequent year to sequentially estimate the intensity of the HPP μ for the occurrence of the precursors.

Figure 8.11 illustrates the trend obtained from an approach based on the truncated lognormal distribution of the conditional core damage probabilities p_i which was fitted using the method of moments (see Section 2.5) and the sample of 275 values of p_i.

Having this distribution estimated, the distribution of $\hat{\lambda}$ in (8.4) was estimated using Monte Carlo simulation from which the mean and upper 95% quantile were calculated. The maximum for 1985 is associated with the outlying precursor observed in the year for which $p_i = 0.011$.

Finally, Figure 8.12 shows the trend based on the extreme value approach (Equation (8.5)). The probabilities that p_{max} exceeds the two indicated values (0.01 and 0.001) are plotted based on the same precursor data. Note that the results in Figure 8.12 indicate the same general trend as in Figure 8.11.

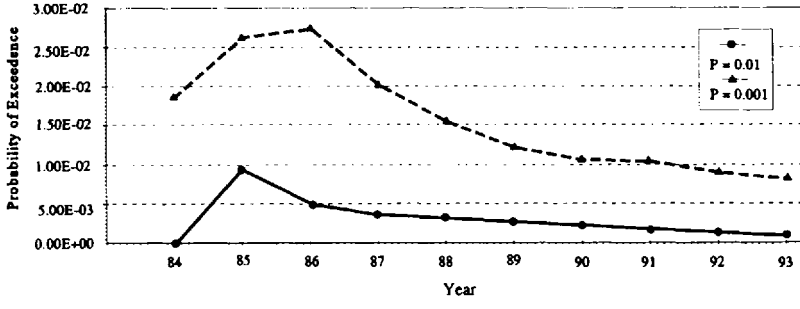


Figure 8.12 Safety trends b ased on Equation (8.5). Probability that P_{max} exceeds P.

8.6.6 Differences Between Precursor Analysis and Probabilistic Risk Assessments

The precursor analysis (PA) originated from the problems associated with nuclear power plant safety problems. Originally, its objective was to validate the probabi-

listic risk assessment (PRA) results, so that PA was traditionally viewed as a different approach from PRA. However, the two approaches are fundamentally the same but with different emphasis. For example, both approaches rely on event trees to postulate accident sequences and both use plant-specific data to obtain failure probability of severe accidents (core damage in the case of nuclear power plants). The only thing that differentiates the two approaches is the process of identifying significant events. Readers are referred to Cooke and Goossens (1990), which conclude that PRA and PA are only different in the way the analysis is performed; however, both approaches use the same models and data for the analysis. Therefore, PA and PRA results cannot be viewed as totally independent, and one cannot validate the other.

Another small difference between the two approaches is the way dependent failures are treated. Dependent failures such as common-cause failures, are considered in PA because a precursor event may include dependent failures. This is a favorable feature of PA calculations. One can also estimate the contribution that common-cause or other events make to the overall rate of occurrence of severe accidents. Common-cause failures are explicitly modeled in PRA the same way as discussed in Section 7.2.

The last difference to be mentioned is that PRAs limit themselves to a finite number of postulated events. However, some events that are not customarily included in PRA may occur as precursor events, and these may be important contributions to risk. This is certainly an important strength of PA methodology.

REFERENCES

Apostolakis, G.A. and Mosleh, A., "*Expert Opinion and Statistical Evidence: An Application to Reactor Core Melt Frequency*," Nucl. Sci. Eng., 70, 135, 1979.

Bier, V. M. and Mosleh, A., "An Approach to the Analysis of Accident Precursors: The Analysis, Communication, and Perception of Risk," B. J. Garrick and W. C. Gekler, Eds., Plenum Press, New York (1991).

Bier, V. M., "Statistical Methods for the Use of Accident Precursor Data in Estimating the Frequency of Rare Events," Reliability, Engineering & System Safety, 39, 267, 1993.

Castillo, E., "*Extreme Value Theory in Engineering*," Academic Press, New York, 1988.

Cooke, R. M., Goossens, H. J., Hale, A. R., and Von der Horst, J., "Accident Sequence Precursor Methodology: A Feasibility Study for the Chemical Process Industries," Technical University of Delft Report, 1987.

Cooke, R. and Goosens, L., "*The Accident Sequence Precursor Methodology for the European Post-Seveso Era*," Reliab. Eng. System Safety, 27, 117, 1990.

Dezfuli, H. and Modarres, M., "*A Truncation Methodology for Evaluation of Large Fault Trees*," IEEE Transactions on Reliability, Vol. R-33, 4, pp. 325–328, 1984.

Farmer, F., "*Containment and Siting of Nuclear Power Plants*," Proc. of a Symp. on Con tain. and Siting of Nucl. Power Plants, Int. Atomic Energy Org., Vienna, Austria, 1967.

Litai, D., "*A Risk Comparison Methodology for the Assessment of Acceptable Risk*," Ph.D. Thesis, Dept. of Nucl. Eng., Mass. Inst. Tech., Cambridge, MA, 1980.

Modarres, M., Martz, H., and Kaminskiy, M., "*The Accident Sequence Precursor Analysis: Review of the Methods and New Insights*," Nucl. Sci. Eng., 123, 238–258, 1996.

NUREG/CR-4550, "*Analysis of Core Damage Frequency from Internal Events*," Vol.1, U.S. Nuclear Regulatory Commission, Washington, DC, 1990.

Paulos, J.A., "*Temple University Report*," Philadelphia, 1991.

Reactor Safety Study, "*Reactor Safety Study—An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants*," WASH-1400, U.S. Nuclear Regulatory Commission, Washington, DC, 1975.

Rowe, W.D., "*An Anatomy of Risk*," Wiley, New York, 1977.

Rubenstein, D., "Core Damage Overestimation," U.S. Nuclear Regulatory Commission, NUREG/CR-3591, 1985.

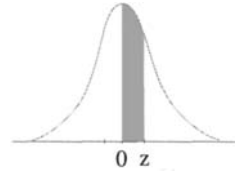
Scott, D.W., "*Multivariate Density Estimation*," John Wiley & Sons, New York, 1992.

U.S. Nuclear Regulatory Commission, "*Safety Goals for the Operation of Nuclear Power Plants: Policy Statement*," Fed. Regist., 51 (149), Washington, DC, 1986.

Wilson, R., "*Analyzing the Daily Risks of Life*," *Technology Review*, Vol. 81, No. 4, pp. 41–46, Cambridge, MA, 1979.

Appendix A: Statistical Tables

Table A.1 Standard Normal Distribution Table*



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.464
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.424
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.385
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.348
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.312
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.277
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.245
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.214
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.186
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.161
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.137
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.117
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.098
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.082
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.068
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.055
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.045
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.036
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.029
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.023
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.018
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.014
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.011
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.008
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.006
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.004
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.003
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.002
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.001
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.001
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.001
3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.000
3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.000
3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.000
3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.000
3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.000

*Adapted from Table 1 of Pearson, E.S., and Hartley, H.O., Eds.: *Biometrika Tables for Statisticians,* Vol. 1, 3rd ed. Cambridge Univ. Press,Cambridge, U.K., 1966. Used by permission.

Table A.2 Percentiles of the t Distribution*

df	$t_{.60}$	$t_{.70}$	$t_{.80}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	.325	.727	1.376	3.078	6.314	12.706	31.821	63.657
2	.289	.617	1.061	1.886	2.920	4.303	6.965	9.925
3	.277	.584	.978	1.638	2.353	3.182	4.541	5.841
4	.271	.569	.941	1.533	2.132	2.776	3.747	4.604
5	.267	.559	.920	1.476	2.015	2.571	3.365	4.032
6	.265	.553	.906	1.440	1.943	2.447	3.143	3.707
7	.263	.549	.896	1.415	1.895	2.365	2.998	3.499
8	.262	.546	.889	1.397	1.860	2.306	2.896	3.355
9	.261	.543	.883	1.383	1.833	2.262	2.821	3.250
10	.260	.542	.879	1.372	1.812	2.228	2.764	3.169
11	.260	.540	.876	1.363	1.796	2.201	2.718	3.106
12	.259	.539	.873	1.356	1.782	2.179	2.681	3.055
13	.259	.538	.870	1.350	1.771	2.160	2.650	3.012
14	.258	.537	.868	1.345	1.761	2.145	2.624	2.977
15	.258	.536	.866	1.341	1.753	2.131	2.602	2.947
16	.258	.535	.865	1.337	1.746	2.120	2.583	2.921
17	.257	.534	.863	1.333	1.740	2.110	2.567	2.898
18	.257	.534	.862	1.330	1.734	2.101	2.552	2.878
19	.257	.533	.861	1.328	1.729	2.093	2.539	2.861
20	.257	.533	.860	1.325	1.725	2.086	2.528	2.845
21	.257	.532	.859	1.323	1.721	2.080	2.518	2.831
22	.256	.532	.858	1.321	1.717	2.074	2.508	2.819
23	.256	.532	.858	1.319	1.714	2.069	2.500	2.807
24	.256	.531	.857	1.318	1.711	2.064	2.492	2.797
25	.256	.531	.856	1.316	1.708	2.060	2.485	2.787
26	.256	.531	.856	1.315	1.706	2.056	2.479	2.779
27	.256	.531	.855	1.314	1.703	2.052	2.473	2.771
28	.256	.530	.855	1.313	1.701	2.048	2.467	2.763
29	.256	.530	.854	1.311	1.699	2.045	2.462	2.756
30	.256	.530	.854	1.310	1.697	2.042	2.457	2.750
40	.255	.529	.851	1.303	1.684	2.021	2.423	2.704
60	.254	.527	.848	1.296	1.671	2.000	2.390	2.660
120	.254	.526	.845	1.289	1.658	1.980	2.358	2.617
∞	.253	.524	.842	1.282	1.645	1.960	2.326	2.576
df	$-t_{.40}$	$-t_{.30}$	$-t_{.20}$	$-t_{.10}$	$-t_{.05}$	$-t_{.025}$	$-t_{.01}$	$-t_{.005}$

When the table is read from the foot, the tabled values are to be prefixed with a negative sign. Interpolation should be performed using the reciprocals of the degrees of freedom.

* The data of this table are taken from Table III of Fischer and Yates: *Statistical Tables for Biological, Agricultural and Medical Research*, published by Longman Group U.K., Ltd., London (previously published by Oliver & Boyd, Ltd., Edinburgh and by permission of the author and publishers. From *Introduction to Statistical Analysis*, 2nd ed., by W. J. Dixon and F. J. Massey, Jr. Copyright, 1957. McGraw-Hill Book Company.). Used by permission.

Table A.3 Percentiles of the χ^2 Distribution*

df	Per Cent									
	.5	1	2.5	5	10	90	95	97.5	99	99.5
1	.000039	.00016	.00098	.0039	.0158	2.71	3.84	5.02	5.63	7.88
2	.0100	.0201	.0506	.1026	.2107	4.61	5.99	7.38	9.21	10.60
3	.0717	.115	.216	.352	.584	6.25	7.81	9.35	11.34	12.84
4	.207	.297	.484	.711	1.064	7.78	9.49	11.14	13.28	14.86
5	.412	.554	.831	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	.676	.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40,00
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
120	83.85	86.92	91.58	95.70	100.62	140.23	146.57	152.21	158.95	163.64

For large values of degrees of freedom the approximate formula

$$\chi_\alpha^2 = n\left(1 - \frac{2}{9n} + z_\alpha\sqrt{\frac{2}{9n}}\right)^3$$

where z_α is the normal deviate and n is the number of degrees of freedom, may be used. For example:

$\chi_{.99^2} = 60\,[1 - .00370 + 2.326(.06086)]^3 = 60(1.1379)^3 = 88.4$ for the 99th percentile for 60 degrees of freedom.

* From *Introduction to Statistical Analysis*, 2d ed., by W. J. Dixon and F. J. Massey, Jr., Copyright, 1957. McGraw-Hill Book Company. Used by permission.

Table A.4 Critical Values $D_n^{(\gamma)}$ for the Kolmogorov Goodness-of-Fit Test*

n	γ				
	0.20	0.15	0.10	0.05	0.01
1	0.900	0.925	0.950	0.975	0.995
2	0.684	0.726	0.776	0.842	0.929
3	0.565	0.597	0.642	0.708	0.828
4	0.494	0.525	0.564	0.624	0.733
5	0.446	0.474	0.510	0.565	0.669
6	0.410	0.436	0.470	0.521	0.618
7	0.381	0.405	0.438	0.486	0.577
8	0.358	0.381	0.411	0.457	0.543
9	0.339	0.360	0.388	0.432	0.514
10	0.322	0.342	0.368	0.410	0.490
11	0.307	0.326	0.352	0.391	0.468
12	0.295	0.313	0.338	0.375	0.450
13	0.284	0.302	0.325	0.361	0.433
14	0.274	0.292	0.314	0.349	0.418
15	0.266	0.283	0.304	0.338	0.404
16	0.258	0.274	0.295	0.328	0.392
17	0.250	0.266	0.286	0.318	0.381
18	0.244	0.259	0.278	0.309	0.371
19	0.237	0.252	0.272	0.301	0.363
20	0.231	0.246	0.264	0.294	0.356
25	0.210	0.220	0.240	0.270	0.320
30	0.190	0.200	0.220	0.240	0.290
35	0.180	0.190	0.210	0.230	0.270
>35	$\frac{1.07}{\sqrt{n}}$	$\frac{1.14}{\sqrt{n}}$	$\frac{1.22}{\sqrt{n}}$	$\frac{1.36}{\sqrt{n}}$	$\frac{1.63}{\sqrt{n}}$

* With permission from F. J. Massey (1951). The Kolmogorov-Smirnov Test for Goodness of Fit, *Journal of the American Statistical Association*, Vol. 46, p. 70.

Table A.5a Percentage Points of the F-Distribution (90th Percentile Values of the *F*-Distribution)

f1 / f2	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	39.86	49.50	53.59	55.83	57.24	58.20	58.91	59.44	59.86	60.19	60.71	61.22	61.74	62.00	62.26	62.53	62.79	63.06	63.33
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39	9.41	9.42	9.44	9.45	9.46	9.47	9.47	9.48	9.49
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23	5.22	5.20	5.18	5.18	5.17	5.16	5.15	5.14	5.13
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92	3.90	3.87	3.84	3.83	3.82	3.80	3.79	3.78	3.76
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30	3.27	3.24	3.21	3.19	3.17	3.16	3.14	3.12	3.10
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94	2.90	2.87	2.84	2.82	2.80	2.78	2.76	2.74	2.72
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70	2.67	2.63	2.59	2.58	2.56	2.54	2.51	2.49	2.47
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54	2.50	2.46	2.42	2.40	2.38	2.36	2.34	2.32	2.29
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42	2.38	2.34	2.30	2.28	2.25	2.23	2.21	2.18	2.16
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32	2.28	2.24	2.20	2.18	2.16	2.13	2.11	2.08	2.06
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25	2.21	2.17	2.12	2.10	2.08	2.05	2.03	2.00	1.97
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19	2.15	2.10	2.06	2.04	2.01	1.99	1.96	1.93	1.90
13	3.14	2.76	2.56	2.43	2.35	2.28	2.23	2.20	2.16	2.14	2.10	2.05	2.01	1.98	1.96	1.93	1.90	1.88	1.85
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10	2.05	2.01	1.96	1.94	1.91	1.89	1.86	1.83	1.80
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06	2.02	1.97	1.92	1.90	1.87	1.85	1.82	1.79	1.76
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03	1.99	1.94	1.89	1.87	1.84	1.81	1.78	1.75	1.72
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00	1.96	1.91	1.86	1.84	1.81	1.78	1.75	1.72	1.69
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98	1.93	1.89	1.84	1.81	1.78	1.75	1.72	1.69	1.66
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96	1.91	1.86	1.81	1.79	1.76	1.73	1.70	1.67	1.63
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94	1.89	1.84	1.79	1.77	1.74	1.71	1.68	1.64	1.61
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92	1.87	1.83	1.78	1.75	1.72	1.69	1.66	1.62	1.59
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.86	1.81	1.76	1.73	1.70	1.67	1.64	1.60	1.57
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.74	1.72	1.69	1.66	1.62	1.59	1.55
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.83	1.78	1.73	1.70	1.67	1.64	1.61	1.57	1.53
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87	1.82	1.77	1.72	1.69	1.66	1.63	1.59	1.56	1.52
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.81	1.76	1.71	1.68	1.65	1.61	1.58	1.54	1.50
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85	1.80	1.75	1.70	1.67	1.64	1.60	1.57	1.53	1.49
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.79	1.74	1.69	1.66	1.63	1.59	1.56	1.52	1.48
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83	1.78	1.73	1.68	1.65	1.62	1.58	1.55	1.51	1.47
30	2.88	2.49	2.28	2.14	2.03	1.98	1.93	1.88	1.85	1.82	1.77	1.72	1.67	1.64	1.61	1.57	1.54	1.50	1.46
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.71	1.66	1.61	1.57	1.54	1.51	1.47	1.42	1.38
60	2.79	2.39	2.18	2.14	1.95	1.87	1.82	1.77	1.74	1.71	1.66	1.60	1.54	1.51	1.48	1.44	1.40	1.35	1.29
120	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.60	1.55	1.48	1.45	1.41	1.37	1.32	1.26	1.19
∞	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.55	1.49	1.42	1.38	1.34	1.30	1.24	1.17	1.00

Table A.5b Percentage Points of the F-Distribution (95th Percentile Values of the *F*-Distribution)

f2 \ f1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.88	5 80	5 77	5 75	5 72	5 69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3 84	3 81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.84	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3 15	3 12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.88	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2 62	258	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2:20	2.13	2.05	2.01	1.98	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

f_1 = degrees of freedom in numerator

f_2 = degrees of freedom in denominator

*E.S. Pearson and H.O. Hartley, *Biometrika Tables for Statisticians*, Vol. 2 (1972), Table 5, p. 178. Used by permission.

Table A.5.c Percentage Points of the F-distribution (99th Percentile Values of F-distribution)

f2 \ f1	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	4052	4999.5	5403	5625	5764	5859	5928	5982	6022	6056	6106	6157	6209	6235	6261	6287	6313	6339	6366
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40	99.42	99.43	99.45	99.46	99.47	99.47	99.48	99.49	99.50
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23	27.05	26.87	26.69	26.60	26.50	26.41	26.32	26.22	26.13
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55	14.37	14.20	14.02	13.93	13.84	13.75	13.65	13.56	13.46
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05	9.89	9.72	9.55	9.47	9.38	9.29	9.20	9.11	9.02
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87	7.72	7.56	7.40	7.31	7.23	7.14	7.06	6.97	6.88
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62	6.47	6.31	6.16	6.07	5.99	5.91	5.82	5.74	5.65
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81	5.67	5.52	5.36	5.28	5.20	5.12	5.03	4.95	4.46
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26	5.11	4.96	4.81	4.73	4.65	4.57	4.48	4.40	4.31
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85	4.71	4.56	4.41	4.33	4.25	4.17	4.08	4.00	3.91
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54	4.40	4.25	4.10	4.02	3.94	3.86	3.78	3.69	3.60
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30	4.16	4.01	3.86	3.78	3.70	3.62	3.54	3.45	3.36
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10	3.96	3.82	3.66	3.59	3.51	3.43	3.34	3.25	3.17
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94	3.80	3.66	3.51	3.43	3.35	3.27	3.18	3.09	3.00
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80	3.67	3.52	3.37	3.29	3.21	3.13	3.05	2.96	2.87
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69	3.55	3.41	3.26	3.18	3.10	3.02	2.93	2.84	2.75
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59	3.46	3.31	3.16	3.08	3.00	2.92	2.83	2.75	2.65
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51	3.37	3.23	3.08	3.00	2.92	2.84	2.75	2.66	2.57
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43	3.30	3.15	3.00	2.92	2.84	2.76	2.67	2.58	2.59
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37	3.23	3.09	2.94	2.86	2.78	2.69	2.61	2.52	2.42
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31	3.17	3.03	2.88	2.80	2.72	2.64	2.55	2.46	2.36
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.12	2.98	2.83	2.75	2.67	2.58	2.50	2.40	2.31
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21	3.07	2.93	2.78	2.70	2.62	2.54	2.45	2.35	2.26
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.03	2.89	2.74	2.66	2.58	2.49	2A0	2.31	2.21
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13	2.99	2.85	2.70	2.62	2.54	2.45	2.36	2.27	2.17
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	2.96	2.81	2.66	2.58	2.50	2.42	2.33	2.23	2.13
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06	2.93	2.78	2.63	2.55	2.47	2.38	2.29	2.20	2.10
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.90	2.75	2.60	2.52	2.44	2.35	2.26	2.17	2.06
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00	2.87	2.73	2.57	2.49	2.41	2.33	2.23	2.14	2.03
30	7.56	5.39	4.51	4.02	3.70	3.47	3.3(1	3.17	3.07	2.98	2.84	2.70	2.55	2.47	2.39	2.30	2.21	2.11	2.01
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.66	2.52	2.37	2.29	2.20	2.11	2.02	1.92	1.80
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.50	2.35	2.20	2.12	2.03	1.94	1.84	1.73	1.60
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.34	2.19	2.03	1.95	1.86	1.76	1.66	1.53	1.38
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.18	2.04	1.88	1.79	1.70	1.59	1.47	1.32	1.00

Appendix B: Generic Failure Data

Table B.1 Generic Failure Data for Mechanical Items

Component Failure Mode	Range from other source	Suggested mean value	Lognormal error factor*
Air operated values			
Failure to operate	3E - 4/D to 2E - 2/D	2E - 3/D	3
Failure due to plugging	2E - 5/D to 1E - 4/D 1E - 7/yr	1E 7/hr	3
Unavailability due to test and maintenance	6E - 5/D to 6E - 3/D	8E - 4/D	10
Spurious closure	—	1E 7/hr	3
Spurious open	—	5E 7/hr	10
Pressure regulator valve			
Failure to open	—	2E - 3/D	3
Motor operated valves			
Failure to operate	1E - 3/D to 9E - 3/D	3E 3/D	10
Failure due to plugging	2E - 5/D to 1E - 4/D	1E - 7/hr	3
Unavailability due to test and maintenance	6E 5/D to 6E - 3/D	8E - 4/D	10
Failure to remain closed	—	5E 7/hr	10
Failure to remain open	—	1E 7/hr	3
Solenoid operated valves			
Failure to operate	1E - 3/D to 2E - 2/D	2E -3/D	3
Failure due to plugging	2E - 5/D to 1E - 4/D 1E - 7/yr	1E - 7/hr	3
Unavailability due to test and maintenance	6E 5/D to 6E 3/D	8E 4/D	10
Hydraulic operated valves			
Failure to operate	3E - 4/D to 2E - 2/D	2E - 3/D	3
Failure due to plugging	2E - 5/D to 1E - 4/D 1E - 7/yr	1E 7/hr	3
Unavailability due to test and maintenance	6E - 5/D to 6E - 3/D	8E - 4/D	10
Explosive operated valves			
Failure to operate	1E - 3/D to 9E - 3/D	3E - 3/D	3
Failure due to plugging	2E 5/D to 1E - 4/D, 1E - 7/yr	1E - 7/hr	3
Unavailability due to test and maintenance	6E - 5/D to 6E - 3/D	8E - 4/D	10

Component Failure Mode	Range from other source	Suggested mean value	Lognormal error factor*
Manual valve			
Failure due to plugging	2E - 5/D to 1E - 4/D, 1E - 7/yr	1E - 7/hr	3
Unavailability due to test and maintenance	6E - 5/D to 6E - 3/D	8E - 4/D	10
Failure to open	—	1E - 4/D	3
Failure to remain closed	—	1E - 4/D	3
Check value			
Failure to open	6E - 5/D to 1.2E - 4/D,	1E - 4/D	3
Failure to close	—	1E - 3/hr	3
Safety relief valves (SRVs)— BWR			
Failure to open for pressure relief	—	1E - 5/D	3
Failure to open on actuation	—	1E - 2/D	3
Failure to reclose on pressure relief	—	3.9E - 6/hr	10
Relief valve (not SRV or PORV)			
Spurious open	—	3.9E - 6/hr	10
Power operated relief valves (PORVs)—PWR			
Failure to open on actuation	—	2E - -3/D	3
Failure to open for pressure relief	—	3E - 4/D	10
Failure to reclose	—	2E - -3/D	3
Motor driven pump			
Failure to start	5E - 4/D to 1E - 4/D	3E - 3/D	10
Failure to run	1E - 6/hr to 1E - 3/hr	3E - 5/hr	10
Unavailability due to test and maintenance	1E - 4/D to1E - 2/D	2E - 3/D	10
Turbine driven pump			
Failure to start	5E - 3/D to 9E - 2/D	3E - 2/D	10
Failure to fun	8E - 6/hr to 1E - 3/hr	5E - 3/hr	10
Unavailability due to test and maintenance	3E - 3/D to 4E - 2/D	1E - 2/D	10

Table B.1 Continued

Component Failure Mode	Range from other source	Suggested mean value	Lognormal error factor*
Diesel driven pump			
Failure to start	1E-3/D to 1E-2/D	3E-2/D	3
Failure to run	2E-5/hr to 1E-3/hr	8E-4/hr	10
Unavailability due to test and maintenance	—	1E-2/D	10
Heat exchanger			
Failure due to blockage	—	5.7E-6/hr	10
Failure due to rupture (leakage)	—	3E-6/hr	10
Unavailability due to test and maintenance	—	3E-5/hr	110
AC electric power diesel generator (DG) hardware failure			
Failure to start	8E-3/D to 1E-3/D	3E-2/D	3
Failure to run	2E-4/hr to 3E-3/hr	2E-2/hr	10
DG test and maintenance unavailability	1 to 4E-2/D	6E-3/D	10
Loss of offsite power other than initiator		2E-4/hr	3
AC bus hardware failure	1E-8/hr to 4E-6/hr	1E-7/hr	5
Circuit breaker			
Spurious open	—	1E-6/hr	3
Fail to transfer	—	3E-3/D	10
Time delay relay			
Fail to transfer	—	3E-4/hr	10
Transformer			
Short or open	—	2E-6/hr	10
DC electric power hardware failure	6E-10/hr to 1E-4/hr		
Battery	—	1E-6/hr	3
Bus	—	1E-7/hr	5
Charger	—	1E-6/hr	3
Inverter	—	1E-4/hr	3

Component Failure Mode	Range from other source	Suggested mean value	Lognormal error factor*
Test and maintenance unavailability			
Battery	—	1*E* - 3/D	10
Bus	—	8*E* - 6/hr	10
Charger	—	3*E* 4/D	10
Inverter	—	1*E* 3/D	10
Orifice			
Failure due to plugging	—	3*E* 4/D	3
Strainer			
Failure due to plugging	—	3*E* - 5/hr	10
			10
Sump			
Failure due to plugging	—	5*E* - 5/D	100
Cooling coil			
Failure to operate	—	1*E* - 6/hr	3
Transmitter			
Failure to operate	—	1*E* 6/hr	3
Fan (HVAC)			
Failure to start	—	3*E* - 4/D	3
Failure to run	—	1E 5/hr	3
Unavailability due to test and maintenance	—	2*E* - 3/D	10
Instrumentation (includes sensor, transmitter and process switch)			
Failure to operate	—	3*E* - 6/hr	10
Temperature switch			
Failure to transfer	—	1*E* 4/D	3

Table B.1 Continued

Component Failure Mode	Range from other source	Suggested mean value	Lognormal error factor*
Transfer switch			
Failure to transfer	—	1*E* - 3/D	3
Instrument air compressor			
Failure to start	—	8*E* - 2/D	3
Failure to run	—	2*E* - 4/hr	10
Unavailability due to test and maintenance	—	2*E* - 3/D	10
Flow controller			
Failure to operate	—	1*E* - 4/D	3
Cooling tower fan			
Failure to start	—	4*E* - 3/D	3
Failure to run	—	7*E* - 6/hr	10
Unavailability due to test and maintenance	—	2*E* - 3/D	10
Damper			
Failure to open	—	3*E* - 3/D	10

* Defined as $EF = PU/m = m/P_L$, where P_U and P_L are upper and lower 95th percentile of lognormal distribution and m is its median.

Obtained from NUREG /CR-4550, 1990. Analysis of Core Damage Frequency from Internal Events, U.S. NRC, Washington, D.C., Vol. 1.

Appendix C: Software for Reliability and Risk Analysis

Table C.1 Selected PC-Based Software for Logic (Boolean-Based) Analysis

Software	Primary functions	For more information
CAFTA Windows	Full screen fault tree editor Top-event cut set generator Cut-set screening editor Event tree editor Integrates fault trees and event trees Cut-set generator Cut-set screening editor Database for failure data Cut-set quantification	Science Applications International Corp. 5150 El Camino Real, Suite C-31 Los Altos, CA 94022 http://www.saic.com
EOOS	Reliability and risk-based analyses System status or "alarm" panel Simplified Gantt chart to fill reliability, maintenance related schedules Color-coded system diagram display	Science Applications International Corp. 5150 El Camino Real, Suite C-31 Los Altos, CA 94022 http://www.saic.com
ORAM	Evaluates safety functions Provides guidelines for managing risk Displays quantitative risk profiles	ERIN Engineering 2033 N. Main Street, Suite 1000 Walnut Creek, CA 94596 http://www.erineng.com
R&R-Workstation	Integrate other software (e.g., CAFTA, EOOS, RISKMAN Application of risk and performance tools Integrates software tools to application environments	Science Applications International Corp. 5150 El Camino Real, Suite C-31 Los Altos, CA 94022 http://www.saic.com

Software	Primary functions	For more information
REVEAL_W	Graphically constructs MLD and success trees Propagates effects of failure in the MLD Reliability and risk-based ranking Common cause failures Connects with MS ACCESS™ database Connects with MS EXCEL™ for report generation	Scientech, Inc. 11140 Rockville Pike Rockville, MD 20852 http://www.scientech.com
RISKMAN	Event tree editor Database editor Fault tree editor Cut-set generator Handles and combine large event trees Calculate event tree sequence probabilities Bayesian analysis	PLG, Inc. 2260 University Drive Newport Beach, CA 92660 http://www.plg.com
SAPHIRE	Full screen fault tree editor Top-event cut set generator Cut-set generator Event tree editor Cut-set and event tree sequence quantification Database for failure date Integrates fault trees and event trees Uncertainty analysis	U.S. Nuclear Regulatory Commission Office of Nuclear Regulatory Research Washington, DC 20555 http://www.nrc.gov
SENTINEL	Evaluates maintenance and testing Maintenance effectiveness analysis Safety function assessment Performs integrated safety assessment Performance criteria assessment	ERIN Engineering 2033 N. Main Street, Suite 1000 Walnut Creek, CA 94596 http://www.erineng.com

Software	Primary functions	For more information
SETS	Boolean equation reduction Handles complex Boolean equations or fault tree Logically combines (merges) fault trees Quantifies fault trees or Boolean equations	Logic Analysts, Inc. 1717 Louisiana Ave. Suite 102A Albuquerque, NM 87110
SAFETY MONITOR	Calculates On-line assessment of performance and risk of system and plants reliability and risk Uses fault tree and event trees for assessments Uses a "gauge" display of safety significance of actions or system operating configurations Provide a database for storing past performance data	Scientech, Inc. 11140 Rockville Pike Rockville, MD 20852 http://www.scientech.com

Table C.2 Capabilities of Other PC-Based Software

Software name	Time-dependent availability	Reliability block diagram	Fault tree	Markov analysis	Data trend analysis	Dependent failure analysis	FMEA FMECA	Uncertainty analysis	Importance analysis	Human reliability analysis	Address
FRANTIC ABC	X										Applied Biomathematics 100 North County Rd., Bld. B Setauket, NY 1173
CAFTA								X	X		See Table C.1
RAMS			X	X	X				X		Item Software Inc. 2030 Main Street, Suite 1130 Irvine, CA 92614
QUALITY ASSURANCE							X		X		
SAPHIRE			X			X		X	X		See Table C.1
RISKMAN						X		X	X		See Table C.1
TEMAC								X	X		Sandia National Laboratories Albuquerque, NM 87185
MARKOV1				X							Decision System Associate 746 Crompton Redwood City, CA 94061
RAMS-CALS							X				Management Sciences Inc. 6022 Constitution Ave., NE Albuquerque, NM 87110
RELCON			X					X	X		Relcon Teknik AB Box 1288 S-172 25 Sundbyberg, Sweden
RBDA		X									Science Application Int. Corp. 5150 El Camino Real Suite C-31 Los Altos, CA 94022

BRAT								X			The Craig Mark Company P O Box 192 Del Mar, CA 92014
SETS						X			X		See Table C 1
SLIM-MAUD										X	Scientech, Inc 11140 Rockville Pike Rockville, MD 20852
NOPRA			X			X		X	X		Scientech, Inc 11140 Rockville Pike Rockville, MD 20852

Appendix D: Reliability Analysis and Risk Evaluator (RARE) Quick User's Manual

D.1 INTRODUCTION

The objective of the reliability analysis and risk evaluator (RARE) is to help a reader better understand major concepts in reliability engineering and risk analysis. It is intended to illustrate numerical examples provided in the book as well as assist the reader in working out the homework problems. Apart from that, it presents a finalized software tool, which can be used to analyze a variety of the real world reliability data. Written in Visual Basic, RARE has a friendly user interface, and is compatible with a popular MS Excel spreadsheet. The RARE user is expected (but not required) to have a general proficiency in MS Excel. Table D.1 presents a summary of RARE programs.

D.2 RARE INSTALLATION

D.2.1 Hardware and Software Requirements

1. IBM or compatible PC
2. ~ 1.5 MB of hard drive space
3. Windows 3.1 or higher
4. MS Excel 5.0 or higher. MS Excel is essential for running RARE programs. MS Excel must be a fully installed registered copy, which includes the following "Add-In" modules:
   Analysis ToolPak
   Analysis ToolPak—VBA
   Solver Add-In

Table D.1 Summary of RARE Programs

RARE program	Program description	Program concept covered in
Goodness of fit test	The program demonstrates the Chi-square and Kolmogorov–Smirnov tests to perform goodness-of-fit testing for the following distributions: exponential, normal, lognormal, Weibull, and Poisson distributions.	Section 2.7
Nonparameteric estimation	The program demonstrates nonparameteric graphical estimation procedures.	Sections 3.3.1 and 3.3.3
Sample size estimation	The program demonstrates a sample size estimation procedure used in nonparameteric reliability analysis.	Section 3.5.1
Distribution estimation	The program demonstrates the maximum likelihood and probability paper methods of parameter estimation for some popular distributions including exponential, normal, lognormal and Weibull.	Sections 2.5.1 and 3.3.2
Exponential distribution estimation	The program demonstrates a classical estimation of the exponential distribution based on type I and II life test data with and without replacement.	Sections 3.4.1 and 3.4.2
Interval esimation	The program demonstrates interval estimation for the binomial distribution parameter, unknown CDF, as well as normal and lognormal distribution parameters.	Sections 2.5.2, 3.4.5, and 3.5.1
Bayesian estimation	The program demonstrates the Bayesian estimation procedures for binomial and Poisson distributions using conjugate and nonconjugate prior distributions including beta, gamma, uniform, normal, and lognormal.	Section 3.6
Repairable system analysis	The program demonstrates the Laplace test as well as the estimation procedures used in data analysis of homogeneous and nonhomogeneous Poisson processes and reliability growth modeling.	Sections 5.1.3, 5.1.4, and 6.6

If the above-mentioned modules are not available in your current version of MS Excel, they can be added by selecting the "Add-Ins . . . " option from the MS Excel "Tools" menu and by checking the names of the above modules in the "Add-Ins" window.

D.2.2 Installation Procedure

Please follow the instructions on the diskette label for the installation procedure.

D.3 DISCLAIMER

The authors disclaim all warranties as to the RARE software, whether expressed or implied, including without limitation any implied warranties of merchant ability, fitness for a particular purpose, functionality or data integrity or protection.

The RARE software is protected from unintentional modifications by a user. Nevertheless, it is strongly recommended to use only the program control buttons and not alter Excel files. All modifications to RARE programs can be done at the user's risk.

D.4 RUNNING RARE PROGRAMS

All RARE programs have a similar set of control buttons. Every program has a Help and a Quit button, the functions of which are self-explanatory.

Some programs have the Import Data function. The imported data should be an ASCII file containing a column of numbers—for ungrouped data, three tab delimited columns (interval beginning, end, frequency)—for grouped data, and two tab delimited columns (r.v. realization and frequency)—for Poisson data. The file extension for grouped data is *.txg

RARE is supplied with a library of examples from the book, which can be imported to the respective programs using the Import Data function. To print the output of the RARE programs, use the MS Excel print function.

D.4.1 Main Controls Program

Figure D.1 shows the main controls program of RARE.

To run a RARE program:

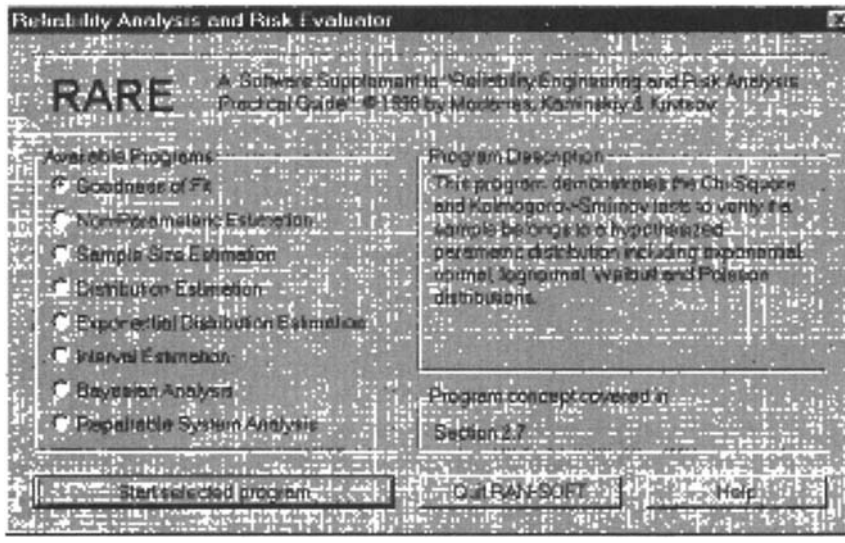


Figure D.1 Main controls window of RARE.

1. Select the program of interest in the Available Programs section of the Main Controls window.
2. Click the Start Selected Program button.

D.4.2 Goodness of Fit Program

The program demonstrates the Chi-square and Kolmogorov–Smirnov tests for exponential, normal, lognormal, Weibull, and Poisson distributions.

To run the program:

1. Click the New Data or the Import Data button to get the initialization window.
2. In the initialization window (see Figure D.2):
   a) select the test type (Chi-square or Kolmogorov–Smirnov);
   b) select the hypothesized distribution type;
   c) select whether the distribution parameters will be estimated from data, or they are known a priori; if the parameters are known a priori, provide their estimates;

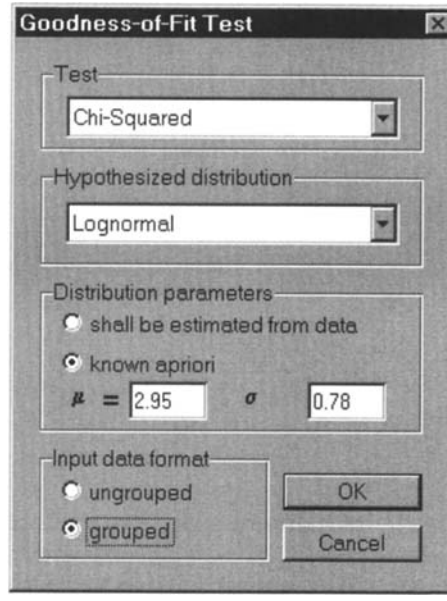


Figure D.2 Initialization window of the goodness of fit program.

d) select whether the data are in an ungrouped or grouped form;
e) click OK to import data from a file, or to type in the respective cells of the main window.

3. Select the desirable significance level, at which the null hypothesis will be checked (see Figure D.3).
4. Click the Compute button to process the data.

☞ Note:

1. Once the initial computation for a given data set have been completed, the hypothesized distribution can be changed (see Figure D.3) to dynamically analyze the results.

D.4.3 Nonparametric Estimation Program

The program demonstrates nonparameteric methods of failure data estimation including procedures for small and large samples on the total-time-on-test plot.

Test	Chi-Squared	Distribution	Lognormal	μ	2.95
Significance	0.05	Parameters	estimated from data	σ	7.80E-01
Critical Value	1.26E+01	New Data	Import Data	Compute	Help	Quit

	Grouped Data Class Interval		Observed Frequency	Expected Frequency	χ^2 Statistic
	Beginning	End	O_i	E_i	$(O_i-E_i)^2/E_i$
	1.00E-08	2.00E+01	7.90E+01	7.54E+01	1.75E-01
	2.00E+01	4.00E+01	3.70E+01	4.39E+01	1.09E+00
	4.00E+01	6.00E+01	1.50E+01	1.45E+01	1.85E-02
	6.00E+01	8.00E+01	6.00E+00	5.47E+00	5.13E-02
	8.00E+01	1.00E+02	2.00E+00	2.34E+00	5.00E-02
	1.00E+02	1.20E+02	1.00E+00	1.11E+00	1.00E-02
	1.20E+02	> 120	4.00E+00	1.33E+00	5.35E+00
			1.44E+02	1.44E+02	6.74E+00

Figure D.3 Main window of the goodness of fit program.

To run the program:

1. Click the New Data or the Import Data button to get the initialization window.
2. In the initialization window: (a) select whether the data are in an ungrouped or grouped form; (b) click OK to import data from a file, or to type in the respective cells of the main window.
3. Click the Compute button.
4. In the Graph window, select which estimated function is to be displayed in the chart.

D.4.4 Sample Size Estimation Program

The program demonstrates a sample size estimation procedure used in nonparameteric reliability analysis.

To run the program (see Figure D.4):

1. Select the lower bound of reliability to be demonstrated in the test.
2. Select the confidence level.
3. Select the number of failures, at which the test will be terminated.
4. Click the Compute button.

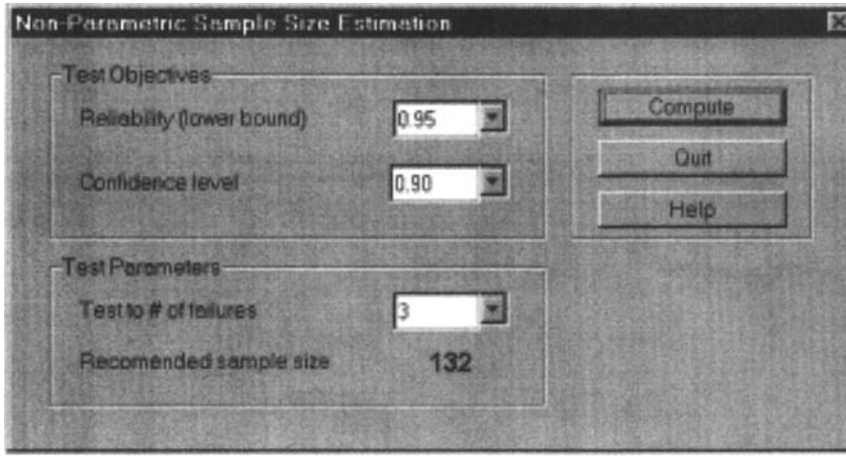


Figure D.4 Main window of the sample size estimation program.

D.4.5 Distribution Estimation Program

The program demonstrates the maximum likelihood and probability paper methods of parameter estimation for some popular distributions including exponential, normal, lognormal and Weibull.

To run the program:

1. Click the New Data or the Import Data button to input the ungrouped failure data into the first column.
2. Choose the type of distribution, parameters of which need to be estimated.
3. Choose the method of plotting position computation for the rank regression analysis.
4. Click the Compute button.
5. In the Graph window (see Figure D.5), select which estimated function is to be displayed in the chart.

☞ Note:

1. Censored data points should be marked with a negative sign.
2. Once the initial computation for a given data set have been completed,

both the plotting position method and the distribution type can be changed to dynamically analyze the results.

3. All functions displayed in the graph window employ probability paper estimates of distribution parameters.

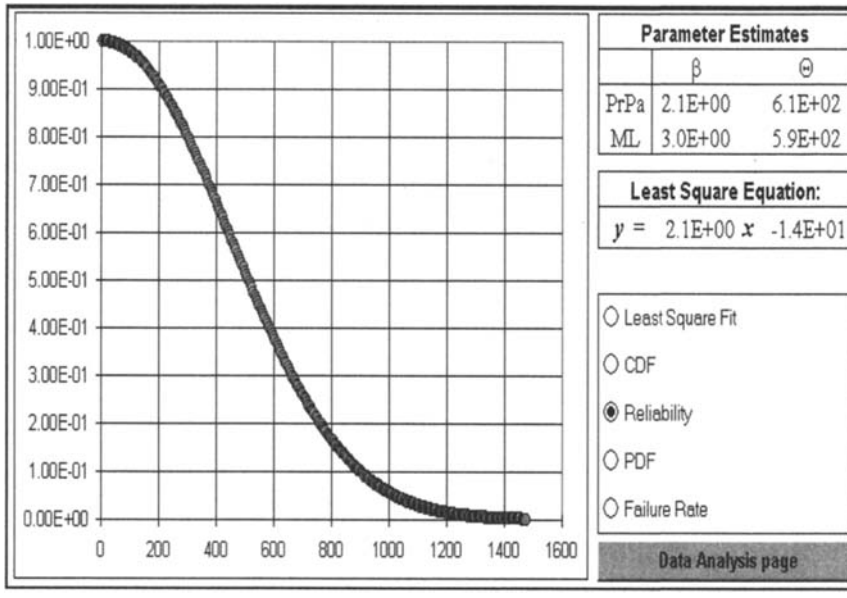


Figure D.5 Graph window of the distribution estimation program.

D.4.6 Exponential Distribution Estimation Program

The program demonstrates a classical estimation of the exponential distribution based on type I and II life test data with and without replacement.

To run the program (see Figure D.6):

1. Click the New Data button.
2. Select whether the test is of type I (time terminated) or type II (failure terminated).
3. Select whether the testing was conducted on with replacement or without replacement scheme.
4. Fill out the missing information in the data input table.
5. Click the Compute button.
6. In the analysis window, select the confidence level of interest for interval estimation.

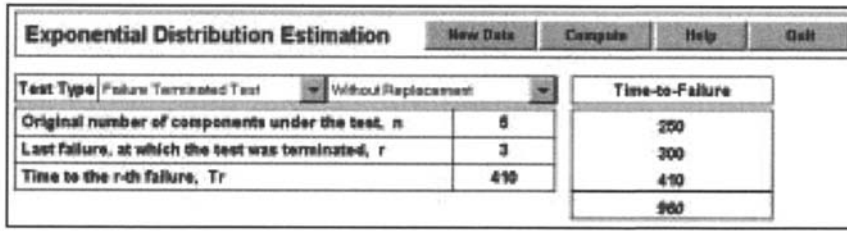


Figure D.6 Main window of the exponential distribution estimation program.

☞ Note:

1. By default, the reliability function is estimated at the time of the last failure or the test termination time. This time can be changed to the operating time of interest by adjusting the value of the blue cell in the analysis window.

D.4.7 Interval Estimation Program

The program demonstrates interval estimation for the binomial distribution parameter, unknown CDF, as well as normal and lognormal distribution parameters.

To run the program (see Figure D.7):

1. Click the New Data button.
2. Select the distribution, parameter(s) of which need(s) to be estimated.
3. Fill out the input data cells.
4. Select the confidence level of interest.
5. Click the Compute button.

☞ Note:

1. Once the initial computation for a given data set have been completed, the confidence level can be changed to dynamically analyze the results.
2. For interval estimation of the exponential and Poisson distribution parameters, please use the **Exponential Distribution Estimation** program.

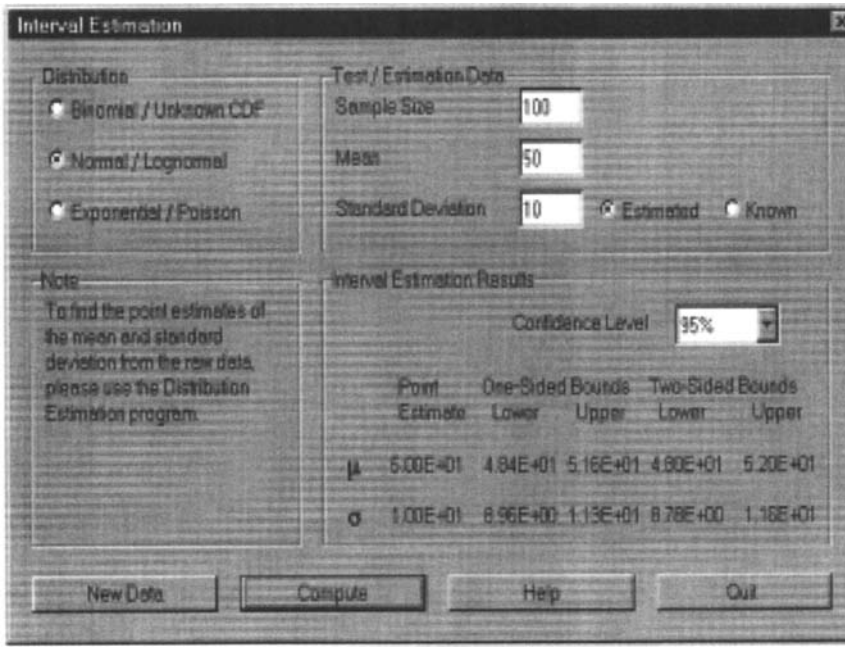


Figure D.7 Main window of the interval estimation program.

D.4.8 Bayesian Analysis Program

The program demonstrates the Bayesian estimation procedures for binomial and Poisson distributions using conjugate and nonconjugate prior distributions including beta, gamma, uniform, normal, and lognormal.

To run the program:

1. Click the New Data button to get the initialization window.
2. In the initialization window (see Figure D.8):
   a) choose the type of prior distribution;
   b) choose the prior distribution evaluation method;
   c) for the method of moments, provide the mean and either the standard deviation or the coefficient of variation; for the method of quantiles, provide the available quantiles and their levels;
   d) choose the likelihood function;

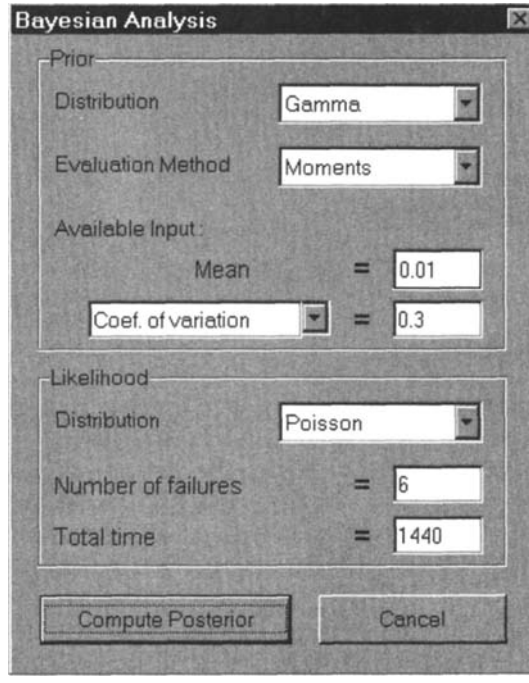


Figure D.8 Initialization window of the Bayesian analysis program.

e) provide the test data corresponding to the chosen likelihood function;

f) click the compute posterior button.

3. Click the Zoom on Graph button to get the enlarged graphical output (see Figure D.9).

☞ Note:

1. Confidence level corresponds to the two-sided confidence bounds.
2. Once the initial computation for a given data set have been completed, the input data (values in the blue cells) can be changed to dynamically analyze the results.

D.4.9 Repairable System Analysis Program

The program demonstrates estimation procedures used in data analysis of homogeneous and nonhomogeneous Poisson processes as well as reliability growth modeling.

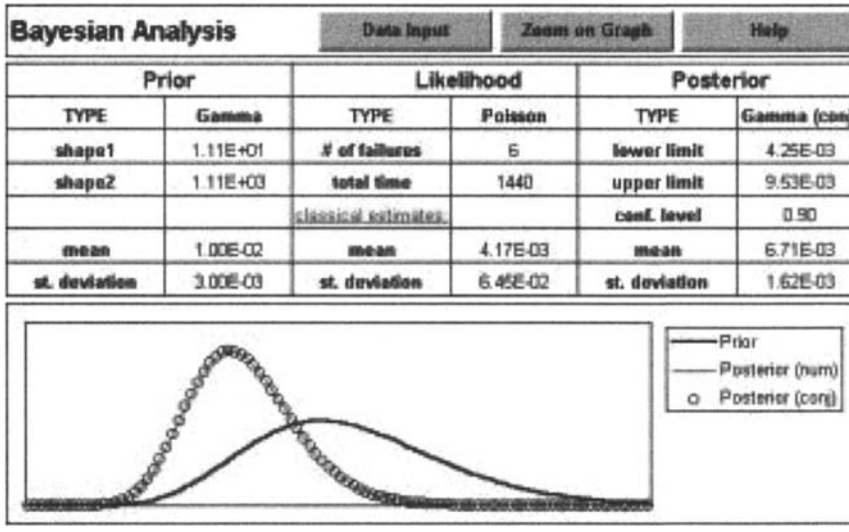


Figure D.9 Main window of the Bayesian analysis program.

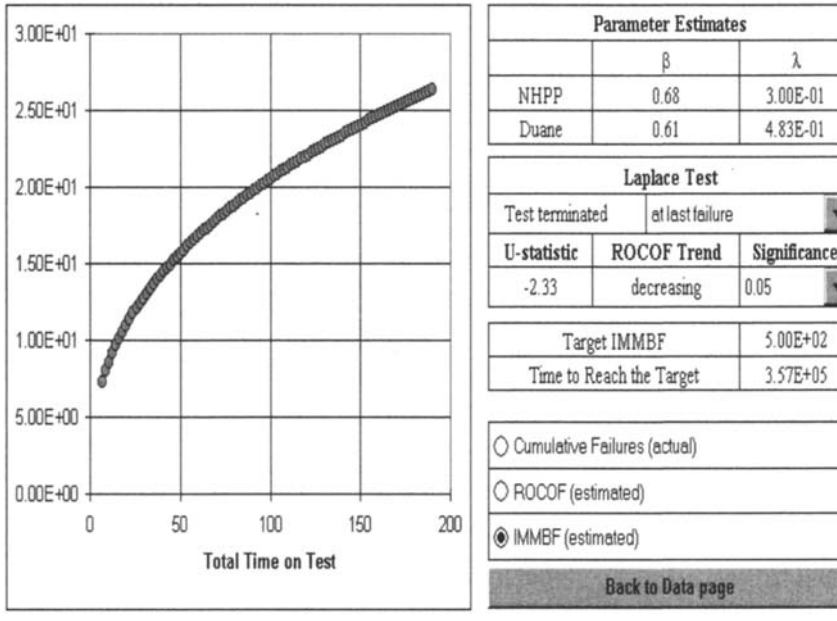


Figure D.10 Graph window of the repairable system analysis program.

To run the program:

1. Click the New Data or the Import Data button to input the failure arrival times into the first column.
2. Click the Compute button.
3. In the graph window (see Figure D.10), select whether the test was terminated at or after the last failure.
4. Select the significance level for the trend hypothesis test.
5. Provide the target IMMBF to have the time to reach the target computed.

☞ Note:

1. The input data should be the failure arrival as opposed to failure inter-arrival times.
2. All functions displayed in the graph window employ maximum likelihood estimates of the NHPP parameters.

Index

Accelerated life data analysis, 394
Accelerated life model, 390
Accident precursor analysis, 495, 500
Alpha factor model, 417
Arrhenius reaction model, 393
Availability, 281
  average, 307, 311
  definition of, 17
  instantaneous, 307
  limiting average, 307
  limiting point, 310

Bathtub curve, 109
Bayes' theorem, 33
Boolean algebra, 25

Censoring, 145
  left, 145
  random, 146
  right, 145
  type I, 145
  type II, 146
Central limit theorem, 120
Challenge response model, 3
Common cause failures, 408
Confidence interval, 78
Confidence level, 78
Correlation coefficient, 70
Counting function, 282
Covariance, 70
Cumulative distribution function, 48
Cumulative hazard function, 108
Cut set, 212
  minimal, 212

Damage endurance model, 2
Distribution
  beta prior, 178
  conjugate prior, 168
  empirical, 81
  lognormal prior, 182
  posterior, 165
  prior, 165
  uniform prior, 171

Empirical distribution function, 158
Estimation
  based on expert opinion, 442
  Bayes', 164
  of binomial distribution, 154, 173
  classical nonparametric, 158
  classical parametric, 144
  of exponential distribution, 147, 150, 166
  graphical nonparametric, 127
  of lognormal distribution, 154
  of Weibull distribution, 155

Estimator, 74
  efficient, 74
  minimum variance, 74
  unbiased, 74
Event, 26
  accident precursor, 495
  desirable top, 232
  external, 477
  internal, 476
  primary, 216
  rare approximation of, 31, 225
  top, 213
Expectation, 65
  algebra of, 68
Expected value, 65

Failure mechanism
  electrical, 5
  extrinsic, 9
  intrinsic, 9
  mechanical, 4
Failure rate, 107
  decreasing, 109
  decreasing average, 110
  generic, 185
  increasing, 110
  increasing average, 110
Fault tree method, 213
FMEA, 248, 267, 473
  design, 249
  process, 249
FMECA, 249, 262, 267

Gamma function, 58
Goodness-of-fit test, 83
  Chi-square, 83
  Kolmogorov, 87
Greenwood's formula, 163

Hazard rate, 107
Homogeneous Poisson process (HPP), 284, 290
Human reliability, 346
  analysis, 346
  models, 352

Hypothesis,
  alternative, 79
  null, 79
  testing, 73

Kaplan-Meier estimation, 162

Laplace test, 301
Lloyd-Lipow method, 430
Logic tree, 219

Maintenance
  optimal preventive, 374
  reliability-centered, 370
Master logic diagram, 238
Maximus method, 432
Mean, 65
Mean-time-between-failures (MTBF), 107
Mean-time-to-failure (MMTF), 106
Measure of importance, 360
  Birnbaum, 360
  Fussell-Vesely, 363
  risk achievement worth, 365
  risk reduction worth, 364
Median, 106
Method of maximum likelihood, 76
Method of moments, 75
Multiple Greek letter model, 415

Nonhomogeneous Poisson Process (NHPP), 282, 295

Palgren-Minor rule, 402
Path
  minimal, 234
  success, 234
Power rule model, 393
Probability
  calculus of, 27
  classical interpretation of, 27
  frequency interpretation of, 27
  posterior, 34
  prior, 34
  subjective interpretation of, 27
Probability density function, 48

Probability distributions
beta, 60
binomial, 40
conditional, 63
continuous, 47
discrete, 39
exponential, 55, 115
extreme value, 121
Frechet, 124
gamma, 58, 118
geometric, 47
Gumbel, 124
hypergeometric, 42
joint, 61
lognormal, 53, 120
marginal, 62
normal, 50, 120
Poisson, 44
uniform, 39
Weibull, 56, 116
Probability plotting of
exponential distribution, 133
lognormal distribution, 138
normal distribution, 138
Weibull distribution, 135
Proportional hazard model, 392

Random variable, 39
Rate of occurrence of failures (ROCOF), 282
Regression analysis, 82
Reliability
component model, 127
definition of, 14
function, 106
human, 346
software, 339
system, 197
Reliability block diagram, 198
Reliability growth, 376
AMSAA method of estimation, 381
Duane method of estimation, 377
Renewal
elementary theorem, 287
equation, 286
Renewal process, 285
overdispersed, 286
underdispersed, 286
Risk
acceptability, 461
analysis, 461, 465
definition of, 18
evaluation, 493
perception, 461
probabilistic assessment, 470, 475
Root-cause analysis, 453

Safety margin, 334
Set
compliment of, 22
disjoint, 24
empty, 23
exclusive, 228
null, 23
universal, 21
Software life cycle model, 345
Software reliability
count model of, 342
analysis of, 338
model of, 339, 341
Nelson's model of, 342
Standard deviation, 68
Stress-strength analysis, 333
Symbol
event, 216
gate, 216
transfer, 216
System
complex, 209
decomposition method, 210
event space method, 210
inspection method, 210
path-trace method, 210
K-out-of-*N*, 202
load-sharing, 207
parallel, 200
series, 198
standby, redundant, 203

Time dependent stress, 401, 405
Tolerance requirements model, 3
Total-time-on-test plot, 141
Type I error, 80
Type II error, 80

Uncertainty, 421
  completeness, 424
  graphical representation of, 441
  model, 423
  parameter, 423
  propagation, 425

Variance, 68
  residual, 94
  sample, 75