INTERNATIONAL SERIES IN OPERATIONS RESEARCH AND MANAGEMENT SCIENCE

Operations Research Management Science

System Signatures and their Applications in Engineering Reliability

Francisco J. Samaniego

Springer

System Signatures and their Applications in Engineering Reliability

Recent titles in the
INTERNATIONAL SERIES IN
OPERATIONS RESEARCH & MANAGEMENT SCIENCE
Frederick S. Hillier, Series Editor, *Stanford University*

Sethi, Yan & Zhang/ *INVENTORY AND SUPPLY CHAIN MANAGEMENT WITH FORECAST UPDATES*

Cox/ *QUANTITATIVE HEALTH RISK ANALYSIS METHODS: Modeling the Human Health Impacts of Antibiotics Used in Food Animals*

Ching & Ng/*MARKOV CHAINS: Models, Algorithms and Applications*

Li & Sun/ *NONLINEAR INTEGER PROGRAMMING*

Kaliszewski/*SOFT COMPUTING FOR COMPLEX MULTIPLE CRITERIA DECISION MAKING*

Bouyssou et al/ *EVALUATION AND DECISION MODELS WITH MULTIPLE CRITERIA: Stepping stones for the analyst*

Blecker & Friedrich/ *MASS CUSTOMIZATION: Challenges and Solutions*

Appa, Pitsoulis & Williams/ *HANDBOOK ON MODELLING FOR DISCRETE OPTIMIZATION*

Herrmann/ *HANDBOOK OF PRODUCTION SCHEDULING*

Axsäter/ *INVENTORY CONTROL, 2^{nd} Ed.*

Hall/ *PATIENT FLOW: Reducing Delay in Healthcare Delivery*

Józefowska & Węglarz/ *PERSPECTIVES IN MODERN PROJECT SCHEDULING*

Tian & Zhang/ *VACATION QUEUEING MODELS: Theory and Applications*

Yan, Yin & Zhang/ *STOCHASTIC PROCESSES, OPTIMIZATION, AND CONTROL THEORY APPLICATIONS IN FINANCIAL ENGINEERING, QUEUEING NETWORKS, AND MANUFACTURING SYSTEMS*

Saaty & Vargas/ *DECISION MAKING WITH THE ANALYTIC NETWORK PROCESS: Economic, Political, Social & Technological Applications w. Benefits, Opportunities, Costs & Risks*

Yu/ *TECHNOLOGY PORTFOLIO PLANNING AND MANAGEMENT: Practical Concepts and Tools*

Kandiller/ *PRINCIPLES OF MATHEMATICS IN OPERATIONS RESEARCH*

Lee & Lee/ *BUILDING SUPPLY CHAIN EXCELLENCE IN EMERGING ECONOMIES*

Weintraub/ *MANAGEMENT OF NATURAL RESOURCES: A Handbook of Operations Research Models, Algorithms, and Implementations*

Hooker/ *INTEGRATED METHODS FOR OPTIMIZATION*

Dawande et al/ *THROUGHPUT OPTIMIZATION IN ROBOTIC CELLS*

Friesz/ *NETWORK SCIENCE, NONLINEAR SCIENCE and INFRASTRUCTURE SYSTEMS*

Cai, Sha & Wong/ *TIME-VARYING NETWORK OPTIMIZATION*

Mamon & Elliott/ *HIDDEN MARKOV MODELS IN FINANCE*

del Castillo/ *PROCESS OPTIMIZATION: A Statistical Approach*

Józefowska/*JUST-IN-TIME SCHEDULING: Models & Algorithms for Computer & Manufacturing Systems*

Yu, Wang & Lai/ *FOREIGN-EXCHANGE-RATE FORECASTING WITH ARTIFICIAL NEURAL NETWORKS*

Beyer et al/ *MARKOVIAN DEMAND INVENTORY MODELS*

Shi & Olafsson/ *NESTED PARTITIONS OPTIMIZATION: Methodology And Applications*

** A list of the early publications in the series is at the end of the book **

Francisco J. Samaniego

System Signatures and their Applications in Engineering Reliability

Springer

Francisco J. Samaniego
University of California at Davis
Davis, CA, USA

Series Editor:
Fred Hillier
Stanford University
Stanford, CA, USA

ISBN: 978-0-387-71796-8 (HB) e-ISBN: 978-0-387-71797-5 (e-book)

Library of Congress Control Number: 2007931822

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

springer.com

Dedication

I wish to dedicate this monograph to two giants in the field of Reliability Theory, Frank Proschan and Dick Barlow. I learned a great deal from studying their jointly authored work, but I am particularly grateful to each of them for personally nurturing my interest in the subject through numerous discussions and stimulating correspondence. I was always impressed by their brilliance and creativity, but was even more impressed by their kindness.

I spent the 1971-72 academic year as a postdoctoral fellow in the Statistics Department at Florida State University. The highlight of that year was the privilege of learning about Reliability Theory at the feet of the master. Through Frank Proschan's Reliability Seminar, I slowly graduated from my early view of reliability as an acronym-laden subfield of applied probability to a more mature view (Frank's view) of the field as an area abundantly endowed with intriguing models, mathematical subtlety, interesting and challenging open problems and a wide range of important applications. Frank's legendary sense of humor contributed to a tone that made attendance at his seminar virtually mandatory – you didn't want to miss anything (be it substantive or comedic)! An example of Frank at his impertinent best was a comment he made when speaking about his 1983 "imperfect repair" paper with Mark Brown. Having proven that one could obtain the proportional hazards property for the "time until the first perfect repair" as a derived result from their modeling assumptions, Frank stated that he considered this as evidence that, among all the ways that one might model imperfect repair, "God favors our model." That remark is especially on my mind at this moment, having just completed a monograph on system signatures. I can't help but wonder whether Frank Proschan, were he still with us, would have favored, on the basis of the evidence presented in Chapters 3 - 7, this particular approach to classifying and characterizing system designs over the alternatives one might consider. It would have pleased me greatly if he had.

Dick Barlow had a different but similarly strong influence on the research directions I've taken over the years. This includes work on nonparametric modeling and inference in reliability, Bayesian thinking in reliability and in general, and a host of other special topics. Barlow introduced me to network reliability at a small workshop he hosted in Berkeley in the mid 1980s. Besides Dick himself, I was the only statistician in attendance. The participants were mostly engineers who were experts in subjects like graph theory and combinatorial mathematics. I recall jokingly referring to them as the counters (and they were, indeed, some of the most sophisticated counters one could have brought together). The group went to lunch together every day, usually at a Chinese restaurant on University Avenue. At the end of one of those lunches, I got a fortune cookie that still ranks as my favorite of all time. It actually said something seemingly relevant. It read "You will soon be recognized by people who count.” It made my day!

In a poem on the portraits in Dublin’s Municipal Gallery, W. B. Yeats wrote: “You that would judge me, do not judge alone this book or that . . . .Think of where man’s glory most begins and ends, and say my glory was I had such friends.” Yeats’ sentiments resonate strongly with me today as I dedicate this work to two great mentors and friends.

FJS

Preface

One of the most common problems in the practice of engineering reliability is that of selecting a particular system design among the several options available for achieving some particular performance goal. Often, the goal is a long-lived system. In the case of repairable systems, the goal might be identifying a system design that has as little downtime as possible. The best performance per unit cost is another worthy potential goal. The purpose of this monograph is to provide some guidance on how problems of this type might be formulated and solved. Our approach relies on the relatively new notion of "system signatures." We will introduce the concept here in the context of the well established theory of coherent systems and will seek to provide convincing evidence that the recommended approach to the problems mentioned above is efficacious.

I must admit that I stumbled on the notion of system signatures quite by accident. While visiting the University of Washington in 1982-3, on sabbatical leave from the University of California, Davis, I was, among other things, working on a number of research problems in Reliability Theory. One problem involved trying to understand more deeply the concept of "closure" under the formation of coherent systems. My being in Seattle, where much of the seminal work in this area had been done in the 60s by Birnbaum, Esary, Marshall, Proschan and Saunders, may have subconsciously driven me toward this topic. What was well known at the time was that a (coherent) system in components with increasing failure rates (IFR) is not "closed," that is, is not necessarily IFR, but that the larger class of systems in components with the IFRA (increasing failure rate average) property is closed, that is, these systems are themselves IFRA.

The question that interested me at the time was more or less halfway between these two results – what could one say about the class of systems that did enjoy the IFR closure property? The literature contained a partial answer: k-out-of-n systems in i.i.d. IFR components were known to be IFR.

Expanding upon this result seemed to require a new tool. My immediate goal was to find ways of identifying a direct connection between the failure rate of a system and the common failure rate of its components. The representation theorem in Samaniego [61] expressing the system's failure rate as a multiple of the component failure rate turned out to be the tool that permitted the complete characterization of systems that will be IFR when their components are i.i.d. IFR. That representation had the form $r_T(t) = h(\mathbf{s}, F)r(t)$, where $\mathbf{s}$ is the signature of the system, a vector essentially capturing the influence of the system design on the system's failure rate, and F and r are the underlying distribution and failure rate of the components. So the birth of the signature idea dates back nearly 25 years. In the preceding paragraph, I referred to this notion as "relatively new." In the grand scheme of things, 25 years goes by in a flash, so from that point of view, one could say that signatures are relatively new. My intent, however, was to acknowledge that the notion was not recognized as broadly useful until its properties were carefully studied in their own right. In Kochar, Mukerjee and Samaniego [51], some new preservation theorems were proven, and the comparison of system lifetimes via the properties of their signatures was shown to be feasible and fruitful. These results revealed the potential power and breadth of the concept. Both of the themes referred to above will be presented in detail in Chapter 4 of the present work.

This monograph consists of six substantive chapters on signatures and their applications, together with an opening chapter introducing the topic and a closing chapter summarizing the state of research on signatures and sharing some of my thoughts on future theoretical developments and potential applications of interest. Most of the existing theory on structural reliability is based on Birnbaum, Esary and Saunders' [9] seminal work on multicomponent, two-state systems. As in Birnbaum, et al. [9], our emphasis here will be on binary systems (which are either in a functioning (1) or failed (0) state). My work with the notion of signatures began with the publication of Samaniego [61], a paper entitled "On the Closure of the IFR Class under the Formation of Coherent Systems." In that paper, the signature vector was defined for a coherent system with components having i.i.d. lifetimes with common distribution F. In brief, the signature of such a system in n components is an n-dimensional probability vector $\mathbf{s}$ that is the distribution of the index of the ordered component failure time that corresponds to system failure.

Especially in the last decade, the broad applicability of system signatures has become apparent, and their utility in the comparison of coherent systems and communication networks has been more firmly established. Most recently, we have found that the tool can facilitate the reformulation of heretofore analytically intractable discrete optimization problems in the area of Reliability Economics, providing a mechanism which can make the analytical treatment of these problems feasible. My purpose here is to present a useful overview of work to date on the properties and applications of system signatures with a view toward opening up new potential applications. Some new results will be combined with work available in the literature. The present work is intended to be both comprehensive and unifying. If I succeed in accomplishing these goals, I am convinced that both the scope and depth of application of system signatures in reliability will be substantially enhanced in future work.

I wish to thank my students and coworkers who have participated in many of my studies in this general problem area. These include, in alphabetical order, Debasisis Bhattacharya, Henry Block, Philip Boland, Michael Dugas, Subhash Kochar, Myles Hollander, Michael McAssey, Hari Mukerjee, Moshe Shaked and Eric Vestrup. Many of the ideas presented in this monograph came to life in the course of my conversations with these collaborators, and I express to each of them my deep appreciation for the many stimulating discussions we have had and for each of their contributions to the theory and application of the signature idea. I hope that this attempt to present a coherent and unified version of our collective results does justice to both the work and to these key contributors.

I express my special appreciation to Dr. Robert Launer and to Dr. Harry Chang of the Army Research Office for their sustained support of this project. The present work combines new results developed under ARO support with a reworking and new presentation of years of work on signature-related ideas. My recent research in this area, as well as my work on the present project, was supported by ARO grants ARO19-02-1-0377 and WN11NF05-1-0118. It is also a pleasure to acknowledge other agencies that have supported this work over the years, including the Air Force Office of Scientific Research, The Ford Foundation, The National Security Agency and The National Science Foundation. This monograph was written during a sabbatical leave funded by the University of California, Davis, and I gratefully acknowledge that support as being a critical element in the completion of this project. I would like to thank Brad Efron for inviting me to spend a portion of that leave in the Statistics Department at Stanford University. This provided me with a stimulating yet quiet place to hide out while getting this project off the ground. Finally, I thank the students in my graduate reliability course in Winter Quarter, 2007, for reading the penultimate draft of this monograph and suggesting many improvements. My thanks to Ying Chen, Tammy Greasby, Yolanda Hagar, Jung Won Hyun, Michelle Norris, Clayton Schupp, and Li Zhu.

Francisco J. Samaniego
University of California, Davis
March, 2007

Contents

1

Introduction

The theory of Reliability and Life Testing has its roots in the research into the performance of engineered systems that was spawned by the applications arising in the second World War. An example of early studies of reliability issues is the work of Abraham Wald who, as a member of the heralded Statistical Research Group at Columbia University, treated the problem of estimating the vulnerability of aircraft used in WWII from data on "hits" taken by the planes that returned from various missions. Wald's work led to the addition of reinforcement of particularly vulnerable sections of the fuselage which ultimately led to a higher rate of returning aircraft. Wald's research on these problems was declassified in the late 1970s and is described in detail by Mangel and Samaniego [55].

Important advances in Reliability Theory were made in the early 1950s. Particularly notable is the paper by Epstein and Sobel [34], where optimal estimates are obtained for the mean lifetime of systems based on "type II" (or "order-statistic") censored data assumed to be drawn from an exponential distribution. The significance of that work was that it clearly demonstrated that characteristics of the population could be efficiently estimated from early failures, that is, from the first r systems to fail among the n systems placed on test. Grenander's [40] paper on nonparametric inference in reliability was highly influential. Zelen's [75] edited proceedings of a 1962 conference on the statistical theory of Reliability drew attention to the field and highlighted early research in the area.

A quantum leap in the development of a comprehensive theory of Reliability occurred concurrently with the formation of a statistical research team at Boeing Aircraft Company. That team, which achieved critical mass in the early 1960s, had as its core members James Esary, Albert Marshall, Frank Proschan and Sam Saunders. During their decade together at Boeing, they developed many of the key concepts, models and methods of modern Reliability Theory. This core group, in collaboration with Richard Barlow, Z. W.

Birnbaum, Ingram Olkin and others, published seminal work in each of the three primary subfields in the area: *structural reliability*, which concentrates on the way systems are designed and how these designs influence system performance, *stochastic reliability*, which concentrates on modeling the lifetime characteristics of systems and their components, and *statistical reliability*, which concentrates on the process of drawing inferences about general characteristics of systems from experimental data on their performance. Among the best known products of the Boeing group and its affiliates are papers by Birnbaum, Esary and Saunders [9] on the theory of coherent systems, by Proschan [59] on the occurrence of apparent improved performance over time (decreasing failure rate) in data sets consisting of combined failure data from several systems, the classic text on Mathematical Reliability by Barlow and Proschan [5], the important paper by Birnbaum, Esary and Marshall [8] on nonparametric modeling in reliability and the introduction of a multivariate exponential distribution as a shock model by Marshall and Olkin [56].

The basic theory and tools of structural reliability were pioneered by Birnbaum, Esary and Saunders in a seminal Technometrics paper published in 1961. In that work, the authors created a framework for studying the basic connection between the performance of a system and the performance of the components of which it is composed. Their study established the "structure function" as the predominant tool for distinguishing among systems and for determining whether one system will outperform another. In this sense, this class of functions can be used, though not with great ease, as an index on all systems of interest, and one might select one system over others on the basis of the characteristics of its structure function.

The aim of this monograph is to present a systematic examination of an alternative tool in structural reliability – system signatures. Both theory about, and applications of, system signatures are presented with a view toward demonstrating that this tool constitutes a powerful and versatile device for resolving a variety of problems in Reliability Theory, particularly those involving comparative analysis. I will begin, in Chapter 2, with a review of the traditional ideas and tools of structural reliability as found, for example, in Barlow and Proschan [6]. We formally define the notion of a coherent system and utilize structure functions and their properties as a vehicle for studying system behavior and for comparing one system with another. Central to this discussion is the important role of path sets and cut sets in studying the performance properties of coherent systems. The well-known representations of the structure function in terms of minimal path sets or minimal cut sets are developed. A constructive description of the class of all coherent systems of a given size is presented, and the intriguing open problem of counting the number of coherent systems of order n, for arbitrary fixed n, is discussed. The connection between the structure function and the reliability of a coherent system is presented, and the "reliability polynomial" is introduced for treatment of the i.i.d. case.

In Chapter 3, we introduce the notion of “signatures” of coherent systems in components with i.i.d. lifetimes and provide some guidance on computing and interpreting them. The problem of comparing two complex systems has typically been complicated by the fact that the traditional tools for characterizing system designs have proven to be rather awkward as indices in optimization problems. As we shall see, the existence of an easily interpretable summary of fixed dimension for the essential characteristics of the systems whose components have i.i.d. lifetimes has made the analytical investigation of many of these problems possible. Multiple examples of signature calculations are given, and mention is made of the elements of combinatorial mathematics that are relevant to such calculations. Under the assumption that the components of the systems to be considered have i.i.d. lifetimes, the distribution (and density and failure rate, if they exist) of the system's lifetime T will be represented explicitly as a function of the system's signature and the underlying distribution F of the component lifetimes. These representations are used with some regularity throughout the remainder of the monograph. The notion of signature is extended beyond the class of coherent systems to the family of all stochastic mixtures of coherent systems of a given size (to be referred to as “mixed systems”), and the motivation for doing so is discussed in detail.

In Chapters 4 and 5, the utility of signatures is demonstrated in various reliability contexts. Chapter 4 is dedicated to applications of system signatures to closure and preservation theorems in reliability and to the role that signatures can play in the comparison of coherent systems or mixtures of them. First, we present a description of the “IFR closure problem” and provide a characterization, in terms of system signatures, of systems whose lifetime distributions have an increasing failure rate whenever its components have i.i.d. lifetimes with an increasing failure rate. We then present a collection of preservation theorems showing that certain types of orderings of system signatures imply like orderings of the corresponding system lifetime distributions. Since the calculation of the lifetime distributions of complex systems is a challenging (and often unsolved) problem which makes the direct comparison of system lifetimes a tenuous matter, the utility of comparing some relatively simple summaries for two system designs and knowing immediately that one system has a longer lifetime (in some stochastic sense) is clearly useful. In Section 4.3, an example involving stochastic comparisons of different types of redundancy in coherent systems illustrates the utility of the preservation results developed in the preceding section.

Since the ordering conditions on signatures in the preservation theorems presented in Section 4.2 prove to be sufficient but not necessary for the ordering of system lifetimes, we turn, in the final section of Chapter 4, to the investigation of possible necessary and sufficient conditions (NASCs) for specific types of orderings to hold for the lifetimes T_1 and T_2 of two systems in i.i.d. components. For each of the contexts in which preservation theorems are established, NASCs are obtained for the ordering of system lifetimes. Interestingly, the precise crossing properties of the survival functions or failure rates of two systems of interest can be determined by the behavior of certain functions that depend on the systems' designs only through the respective system signatures. Results of this latter type lead to insights that extend beyond the partial ordering of systems via properties of their signatures. In situations in which systems are not comparable in the usual stochastic senses, it is possible to characterize the crossing behavior of pairs of survival functions or failure rates, as well as the alternating monotonicity of the likelihood ratio, through the precise behavior of the functions used in establishing NASCs for stochastic domination. This latter extension provides a vehicle for fully understanding the relative real-time behavior of the lifetime distributions, failure rates and density functions of competing systems.

In Chapter 5, several comparisons between pairs of special-purpose systems are pursued. In particular, direct and indirect majority systems are contrasted. Signatures are also employed in establishing monotonicity properties of consecutive k-out-of-n systems and in studying the limiting behavior of survival functions and failure rates of arbitrary mixed systems. In section 5.4, we present an important augmentation to the preceding theory on the comparison of two systems. The usual forms of stochastic comparisons (stochastic, hazard-rate and likelihood-ratio ordering) are powerful when they are applicable, but they have the limitation of inducing only a partial ordering on the class of coherent (or mixed) systems. Simply put, it is easy to find pairs of systems that are not comparable under any of these orderings. In Section 5.4, we consider the use of "stochastic precedence," introduced in Arcones, Kvam and Samaniego [3], which classifies system 2 as better than system 1 if $P(T_1 \leq T_2) \geq 1/2$. This criterion leads to definitive comparisons between any two systems of arbitrary size. An explicit formula for computing $P(T_1 \leq T_2)$ is displayed and a signature-based NASC is given for the inequality $P(T_1 \leq T_2) \geq 1/2$ to hold.

Chapter 6 is dedicated to the study of signatures in the context of network reliability. The chapter begins with a brief introduction to basic ideas and vocabulary of communication networks. The signature is a well-defined concept in each of several types of network problems, including two-terminal, k-terminal and all-terminal reliability (focusing, respectively, on whether two terminals, k terminals or all terminals in a network can communicate with each other). The treatment of network reliability includes a review the theory of "dominations" as introduced and developed by Satyanarayana and his co-workers. This is followed by a derivation of a closed-form functional relationship between dominations and the signature vector. The utility of this connection is illustrated by a comparison between two distinct networks with 9 terminals and 27 edges (i.e., 27 paths between pairs of terminals).

Problems in the field of “Reliability Economics” have been, until recently, largely resistant to analytical treatment. In Chapter 7, we will consider the problem of searching for the optimal system of order n relative to a specific criterion function which depends on both a system’s performance and its cost. The solution to this problem assumes i.i.d. component lifetimes and makes essential use of the $(n-1)$-dimensional simplex of signatures of all stochastic mixtures of coherent systems of order n. The latter strategy turns what has heretofore been treated as a large discrete optimization problem (i.e., finding the best coherent system) into a continuous problem (i.e., finding the best mixed system) to which the methods of differential calculus can be applied. Given such a framework, optimal systems are identified through the signatures that maximize the chosen criterion function. Examples are given in which the optimal system is a non-degenerate mixture of coherent systems and every coherent system is inferior to it. Since the solutions obtained depend on a known underlying lifetime distribution F of the competing systems’ components, a complete solution, usable in practice, would entail the estimation of this distribution or its relevant features. Chapter 7 closes with a treatment of the statistical problem that must be solved in order for the optimality results of Chapter 7 to be applicable in practice. The final chapter of this monograph is dedicated to a brief discussion of extensions of, and results related to, the theory and applications treated in Chapters 3 - 7, some further references to related work, and a description of several open problems of interest.

As indicated in the outline presented in this chapter, there’s a good deal of work to be done. Let us now proceed with the program described above.

2

Background on Coherent Systems

2.1 Basic Ideas

We will use the term “system” quite freely and regularly, even though it will remain an undefined term throughout this monograph. As we all have some experience with engineered “systems,” our use of the term should cause no confusion. Informally, we can think of a system as consisting of a collection of “components,” basic constituents which are connected in some fashion to create the whole. We might consider a radio, an automobile, a computer or a cell phone as concrete examples of systems in common use. The main characteristic of our use of the term is that a system works or fails to work as a function of the working or failure of its components. While there are various ways to formalize the notion of a system being partially functioning (for example, a car could technically be driven for a few miles with a flat tire), we will follow the convention established by Birnbaum et al. [9] and consider a system to be either working or failed at any given point in time. To quantify this fact, we assign a 1 to the event that the system works and a 0 to the event that the system fails. The same can be said of each component.

For a system with n components, this idea gives rise to the notion of a state vector, that is, a vector $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \{0,\ 1\}^n$, where for each i, $x_i = 1$ if the ith component is working and $x_i = 0$ if it is not working. We will be interested in whether or not the system is working when the components are in a specific state. A mapping called the structure function provides the desired link.

Definition 2.1. *Consider the space $\{0,\ 1\}^n$ of all possible state vectors for an n-component system. The* structure function $\varphi : \{0,\ 1\}^n \to \{0,\ 1\}$ *is a mapping that associates those state vectors $\mathbf{x}$ for which the system works with the value 1 and those state vectors $\mathbf{x}$ for which the system fails with the value 0.*

Some examples will help make this concept clear. Most people are familiar with two particular systems that arise frequently in reliability: the series system and the parallel system. The first works only if every component is working, while the second works as long as at least one component is working. The structure function for an n-component series system is given by

$$\varphi(\mathbf{x}) = \prod_{i=1}^{n} x_i \ , \tag{2.1}$$

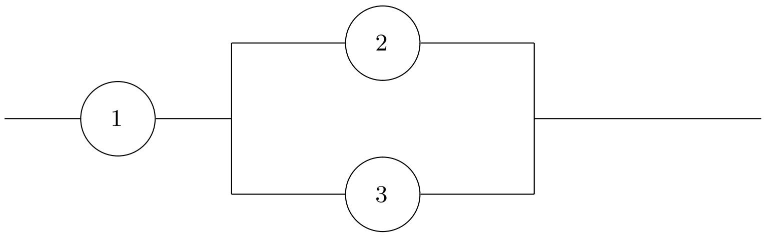
while for a parallel system, we have

$$\varphi(\mathbf{x}) = 1 - \prod_{i=1}^{n} (1 - x_i) \ . \tag{2.2}$$

The two systems above are extreme examples of an important class of systems called "k-out-of-n systems." This label has a bit of ambiguity to it, as it could represent systems that fail upon the kth component failure, but it might also represent systems that work as long as at least k components are working. For that reason, the former are often called k-out-of-n:F systems and the latter are called k-out-of-n:G systems, the "F" standing for the k failed components that ensure system failure and the "G" standing for the k good components that ensure that the system functions. Throughout this monograph, I will make reference to k-out-of-n systems without using the qualifiers F or G. In all instances, I will be referring to what I have called above the k-out-of-n:F systems. Thus, a series system is a 1-out-of-n system and a parallel system is an n-out-of-n system. For hybrid systems such as the system pictured in Figure 2.1, the structure functions will have elements of both of the functions above, that is,

$$\varphi(\mathbf{x}) = x_1[1 - (1 - x_2)(1 - x_3)] \ . \tag{2.3}$$

Fig. 2.1. A series-parallel system in three components

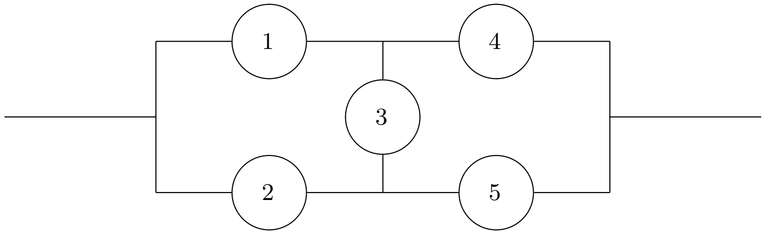


The structure function for a k-out-of-n system is most easily represented as:

$$\varphi(\mathbf{x}) = 0 \text{ if } \sum_{i=1}^{n} x_i \leq n-k \text{ and } \varphi(\mathbf{x}) = 1 \text{ if } \sum_{i=1}^{n} x_i \geq n-k+1 \,. \qquad (2.4)$$

A fourth example is the five-component bridge system displayed in Figure 2.2 below. Its structure function $\varphi(\mathbf{x})$, which follows the figure, is substantially more complex than those that precede it.

Fig. 2.2. A bridge system in 5 components



The structure function of the bridge system shown above is given by

$$\begin{aligned}\varphi(\mathbf{x}) = {} & x_1x_4 + x_2x_5 + x_1x_3x_5 + x_2x_3x_4 - x_1x_2x_3x_4 - x_1x_2x_3x_5 \\ & - x_1x_3x_4x_5 - x_1x_2x_4x_5 - x_2x_3x_4x_5 + 2x_1x_2x_3x_4x_5 \qquad (2.5)\end{aligned}$$

The reader might rightly wonder how the latter structure function was obtained. I placed it in the text at this point because it provides me with the opportunity to make some useful comments. Regarding how the function was obtained, one could use a rational, orderly process (resembling the inclusion-exclusion principle which is discussed later in this chapter) to account first for the ways in which the system will necessarily work – e.g., if components 1 and 4 are working or if components 1, 3 and 5 are working, etc. – and then compensating for their two-way, three-way and four-way intersections. We will introduce shortly an approach that results in (2.5) in a conceptually and practically simpler way. The fact remains that the structure function of a five-component bridge system is a complex object, no matter how you get it. You can imagine struggling with the structure function of a system of order 20. Dealing with the structure functions of systems of order 100 seems almost imponderable. In situations where systems are being compared or where an optimal system is sought relative to some fixed criterion, it is clear that the indexing of the class of systems by their structure functions complicates rather than simplifies the problem. A further complication is that any relabeling of a system's components gives rise to a new function that is equivalent to the original $\varphi(\mathbf{x})$ but looks different. One would have to inspect the two functions side by side to verify that the two representations have components that were in one-to-one correspondence. This, then, constitutes the first bit of motivation one might have for considering other ways of characterizing system designs. But structure functions are useful in their own way, and we will study them further before moving on. In particular, we will utilize them as a vehicle for focusing on the class of systems we will refer to as "coherent."

If an engineer is designing a system for performing a particular function, there are two basic requirements he or she would impose upon any design considered for use. First, the system would not contain any component whose functioning has absolutely no influence on whether or not the system works. If the vector $(x_1, \ldots, x_{i-1}, a, x_{i+1}, \ldots, x_n)$ represents a state vector for the n components of an arbitrary system for which $x_i = a \in \{0,\ 1\}$, then component i is said to be *irrelevant* if the system's structure function φ has the property that $\varphi(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) = \varphi(x_1, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n)$ for all possible values of $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \{0,\ 1\}^{n-1}$. If an n-component system contained a component that was irrelevant, that component would be removed from the design since there is a simpler system (of order no larger than $n-1$) that can provide identical performance.

A system that actually changed from a working state to a failed state upon the replacement of a failed component with a working one would be baffling indeed. One innate feature of systems with which we are familiar is that their failure coincides with the failure of some component. If we were monitoring a system over time, we would note that, as components begin to fail, the system may continue to work for a while but, eventually, one of the components that had remained working will prove to be critical to the system's functioning, and the system will fail upon the failure of that component. Fixing a failed component might get the system working again, but in no instance would we see a system, while working on the basis of k functioning components, fail as a result of the act of replacing a failed component with a working component. We call a system *monotone* if fixing a failed component cannot make the system worse. Symbolically, a monotone system has a structure function for which $\varphi(\mathbf{x}) \leq \varphi(\mathbf{y})$ whenever $\mathbf{x} \leq \mathbf{y}$, where the latter vector inequality is understood to be applied component-wise. These two natural properties of engineered systems form the basis for the following:

Definition 2.2. *A system is said to be* coherent *if each of its components is relevant and if its structure function is monotone.*

The definition above severely restricts the number of possible functions mapping $\{0,\ 1\}^n$ into $\{0,\ 1\}$ that could play the role of a structure function for a coherent system. For example, of the 256 possible functions mapping $\{0,\ 1\}^3$ into $\{0,\ 1\}$, there are only 5 that correspond to coherent systems. Definition 2.2 requires, for example, that the structure function of every coherent system satisfy the conditions $\varphi(\mathbf{0}) = 0$ and $\varphi(\mathbf{1}) = 1$. If either of these conditions were violated, the monotonicity property would imply that every component of the system is irrelevant. In spite of the fact that our restricting attention to coherent systems substantially reduces the number of possible structure functions, it remains true that the number $Z(n)$ of coherent systems of order n grows rapidly with n. There are 2 coherent systems of order 2, 5 of order 3 and 20 of order 4. That the number grows exponentially in n is clear from the fact that, to any system of order n, one may attach a new component either in series or in parallel, immediately doubling the number of coherent systems. This implies that there are more than a billion coherent systems of order 30. Counting the exact number of coherent systems of order n is a difficult open problem, though there exist published bounds for the rate of its growth. A notable upper bound was obtained by Kleitman and Markowsky [50], who addressed the equivalent calculation known as "Dedekind's problem" in the area of enumerative combinatorics:

$$\log_2 Z(n) \leq \{1 + O[(\log_2 n)/n]\} \binom{n}{[n/2]}, \tag{2.6}$$

where $\binom{n}{k}$ is the number of combinations of n things taken k at a time and $[x]$ represents the greatest integer less than or equal to x. The size of $Z(n)$ makes the identification of optimal designs in various settings somewhat imposing. In Chapter 7, we will illustrate a methodology that circumvents this difficulty.

Before proceeding to our discussion of an alternative summary of a coherent system's design, we will introduce an additional element of the Birnbaum, Esary and Saunders framework which will play an important role in the sequel. Let us focus for a moment on coherent systems of order n. A set of components P is said to be a *path set* if the system works whenever all the components in the set P work. It is clear that the set of all n components is a path set. If A is a path set, then any set B that has A as a proper subset will be a path set as well. The path sets of special interest are those that contain no proper subsets that are also path sets. Such a set is called a *minimal path set*. We will denote the minimal path sets of a coherent system as $P_1, P_2, \ldots, P_r$. If we examine the bridge system shown in Figure 2.2, we see that the minimal path sets consist of the collection $\{\{1,4\},\{1,3,5\},\{2,5\},\{2,3,4\}\}$. This collection has two interesting properties:

(i) No minimal path set is a proper subset of any other, and
(ii) The algebraic union of all minimal path sets is the set of all the system's components.

It is possible to characterize all coherent systems of a given order n by these two properties of its minimal path sets. Since, by (ii), every component from 1 to n is a member of at least one minimal path set, the relevance of every component is guaranteed. The monotonicity of the system corresponding to a fixed collection of minimal path sets can be argued as follows. If component k is not working and the system is also not working, then the structure function φ will either remain equal to 0 or will increase to 1 when component k is replaced by a working component. On the other hand, if the system is working, there is a minimal path set P whose components are all working. Since any set of components which contains P will also be a path set, it follows that the set $\{P \cup \{k\}\}$ is a path set and that the system's structure function will remain equal to 1 when component k is replaced by a working component.

There is a natural relationship between path sets and sets of components whose failure will guarantee that the system fails. A set of components C is said to be a *cut set* if the system fails whenever all the components in the set C fail. A cut set is *minimal* if it has no proper subset that is also a cut set. The relationship between cut sets and path sets is evident from the following facts: if P is a minimal path set and A is a proper subset of P, then A^c is a cut set, and if C is a minimal cut set and B is a proper subset of C, then B^c is a path set. Neither A^c nor B^c need be minimal. The family of minimal cut sets has properties analogous to (i) and (ii) above, viz., properties (iii) and (iv) below.

(iii) No minimal cut set is a proper subset of any other, and
(iv) The algebraic union of all minimal cut sets is the set of all the system's components.

In Table 2.1, coherent systems of order 4 are identified by their minimal cut sets.

Table 2.1. Coherent Systems of Order 4

System	Minimal cut sets
1	{1}, {2}, {3}, {4}
2	{1}, {2}, {3, 4}
3	{1}, {2, 3}, {2, 4}
4	{1}, {2, 3}, {2, 4}, {3, 4}
5	{1}, {2, 3, 4}
6	{1, 2}, {1, 3}, {1, 4}
7	{1, 2}, {1, 3}, {1, 4}, {2, 3}
8	{1, 2}, {1, 3}, {1, 4}, {2, 3}, {2, 4}
9	{1, 2}, {1, 3}, {1, 4}, {2, 3}, {2, 4}, {3, 4}
10	{1, 2}, {1, 3}, {2, 4}, {3, 4}
11	{1, 2}, {2, 4}, {3, 4}
12	{1, 2}, {3, 4}
13	{1, 2}, {1, 3}, {1, 4}, {2, 3, 4}
14	{1, 2}, {1, 3}, {2, 3, 4}
15	{1, 2}, {1, 3, 4}, {2, 3, 4}
16	{1, 2}, {1, 3, 4}
17	{1, 2, 3}, {1, 2, 4}
18	{1, 2, 3}, {1, 2, 4}, {1, 3, 4}
19	{1, 2, 3}, {1, 2, 4}, {1, 3, 4}, {2, 3, 4}
20	{1, 2, 3, 4}

Some of the 20 systems in Table 2.1 are easily recognizable. Clearly, system 1 is the series system in four components, system 9 is the 2-out-of-4 system, system 19 is the 3-out-of-4 system and system 20 is the parallel system. System 10 is the system formed by connecting a pair of two-component series systems in parallel, while system 12 is formed by connecting a pair of two-component parallel systems in series. Systems 1 - 5 can be viewed as the systems one obtains by adding a new component (1) in series to the five coherent systems of order 3, while systems 6, 16, 17, 18 and 20 are the systems that can be formed by adding a new component (1) in parallel to the five coherent systems of order 3.

Since all possible coherent systems of order 4 are listed in Table 2.1 according to their minimal cut sets, one might wonder what a list of the corresponding minimal path sets would look like. First, it should be mentioned that the 20 collections above can be thought of as the list of all possible minimal path sets of a particular coherent system, as they are an exhaustive list of collections of sets satisfying (i) and (ii) above. With that interpretation, system 1 would be the parallel system in four components, since if each component serves as a minimal path set, the system works if at least one component is working. Note, then, that the minimal cut sets of the series system are precisely the minimal path sets of the parallel system. This is an example of the "duality" of two coherent systems. Coherent system A is the *dual* of coherent system B if the minimal path sets of A are the minimal cut sets of B (and, similarly, the minimal cut sets of A are the minimal path sets of B). The relationship between dual coherent systems can be made precise through their respective structure functions.

Definition 2.3. *If A and B are dual systems, then their structure functions are related by the equation*

$$\varphi^A(\mathbf{x}) = 1 - \varphi^B(\mathbf{1} - \mathbf{x}) \; . \tag{2.7}$$

A state vector $\mathbf{x}$ is called a cut vector if $\varphi(\mathbf{x}) = 0$ and is called a path vector if $\varphi(\mathbf{x}) = 1$. Now if P is a minimal path set of system B, and $\mathbf{x}^*$ is the corresponding path vector, that is, $x_i^* = 1$ for all $i \in P$ and $x_i^* = 0$ for all $i \in P^c$, then $1 - \mathbf{x}^*$ is clearly a cut vector of B's dual system A since, by (2.7), $\varphi^A(\mathbf{1} - \mathbf{x}^*) = 0$. The set P corresponds to the zeros in the vector $\mathbf{1} - \mathbf{x}^*$ and is thus a cut set of system A. In addition, P is in fact a minimal cut set of A since if any zeros in that vector were turned into ones, the number of ones in the vector $\mathbf{x}^*$ would be diminished. Because P is assumed to be a minimal path set, φ^B would be zero for the diminished vector, implying that φ^A would be one for the augmented version of $\mathbf{1} - \mathbf{x}^*$. Thus, any set smaller than P cannot be a cut set for the system A. In short, the equation (2.7) guarantees that any minimal path set of a system is a minimal cut set of the system's dual and vice versa. This reasoning justifies the claim that for any collection of minimal cut sets defining a coherent system, the same collection can be regarded as the minimal path sets which define coherent systems of that order. Since the minimal path sets of a given system are the same as the minimal cut sets of its dual, one can infer the minimal path sets of a given system of order 4 from Table 2.2. For example, the minimal path sets of system 4 in Table 2.1 are $\{1, 2, 3\}$, $\{1, 2, 4\}$, $\{1, 3, 4\}$, the cut sets of its dual, system 18.

Table 2.2. Duality among systems of order 4

System	Dual
1	20
2	17
3	16
4	18
5	6
7	14
8	15
9	19
10	12
11	11
13	13

The closer look at the list of dual systems in Table 2.2 is instructive. An example of the reasoning which leads to the identification of a system's dual is as follows. Note that system 4 in Table 2.1 has minimal cut sets equal to $\{1\}$, $\{2, 3\}$, $\{2, 4\}$ and $\{3, 4\}$. Now consider the system that has these four sets as its minimal path sets, that is, the dual of system 4. Recognizing that each minimal cut set of this dual system must render all its minimal path sets inoperable, we may identify the sets $\{1, 2, 3\}$, $\{1, 2, 4\}$ and $\{1, 3, 4\}$ as the minimal cut sets of the dual system. These cut sets satisfy conditions (iii) and (iv) above, and, in fact, correspond to the minimal cut sets of system 18 in Table 2.1. Thus, systems 4 and 18 are duals of each other. A further insight one can glean from Table 2.2 is the fact that a system can be its own dual; among systems of order 4, systems 11 and 13 in Table 2.2 have this property.

Carrying out the process exemplified above to identify the dual of system 11 in Table 2.1 brings out an interesting idiosyncrasy of any attempt to classifying coherent systems by their minimal path or cut sets (an idiosyncrasy which carries over to their classification by their structure functions). The system with minimal path sets $\{1, 2\}$, $\{2, 4\}$ and $\{3, 4\}$ (which were the cut sets of system 11) has minimal cut sets $\{1, 4\}$, $\{2, 3\}$ and $\{2, 4\}$. The reader might find it disturbing that this collection of cut sets is not among the list of all possible collections of cut sets in Table 2.1 for systems of order 4. The reason for their apparent absence is the way that the components are labeled.

In the latter collection, if we were to relabel component 3 as component 1 and component 1 as component 3, we see that the resulting sets would be precisely the cut sets listed for system 11 in Table 2.1. Thus, system 11 is its own dual. A more important insight gained from this little exercise is that the labeling of components matters, so that two seemingly different systems might in fact be the same after its components are relabeled. On the other hand, while the minimal cut sets of systems 6 and 11 in Table 2.1 have a similar appearance, they are qualitatively different (since, for example, all three minimal cut sets of system 6 have component 1 in common), so that no relabeling can convert them into the minimal cut sets of system 11. The fact that relabeling of components can disguise a coherent system, making it look like a potentially different system, is a special difficulty that renders path sets, cut sets and structure functions (which share this property) quite imperfect as indexes for the class of coherent systems. The signature of a coherent system introduced in the next chapter does not suffer from this imperfection.

As has been noted earlier, the direct computation of the structure function for an arbitrary coherent system can be algebraically cumbersome. Fortunately, there is a simple connection between the structure function of a coherent system and its minimal path sets and minimal cut sets. In order for a system to work, it must be the case that all the components of at least one minimal path set are working. Similarly, the system will work if and only if at least one of the components in every minimal cut set is working. These observations provide the following tools for the computation of the structure function of a coherent system. Let the minimal path sets of the system be $P_1, \ldots, P_r$, and for each such set, define path structure function $p_j(\mathbf{x})$ as

$$p_j(\mathbf{x}) = \prod_{i \in P_j} x_i \ . \tag{2.8}$$

Now $p_j(\mathbf{x}) = 1$ precisely when every component in P_j is working. It follows that the structure function of the system may be represented as

$$\varphi(\mathbf{x}) = 1 - \prod_{j=1}^{r} (1 - p_j(\mathbf{x})) \ . \tag{2.9}$$

Equation (2.9) can also be written in the following equivalent and revealing form:

$$\varphi(\mathbf{x}) = \max_{\{1 \le j \le r\}} p_j(x) = \max_{\{1 \le j \le r\}} \min_{\{i \in P_j\}} \{x_i\} \tag{2.10}$$

Inspection of equations (2.9) or (2.10) confirms that the structure function is 1 if and only if there is at least one minimal path set for which all components are working. As an example of the approach above, let's revisit the calculation of the structure function for the bridge system in Figure 2.2. As we have noted, the minimal path sets for this system are $\{1, 4\}$, $\{2, 5\}$, $\{1, 3, 5\}$ and $\{2, 3, 4\}$. From (2.9), we thus obtain

$$\varphi(\mathbf{x}) = 1 - (1 - x_1 x_4)(1 - x_2 x_5)(1 - x_1 x_3 x_5)(1 - x_2 x_3 x_4) . \tag{2.11}$$

Noting that the constants cancel out of this expression and using the idempotency of the x variables (that is, $x_i^2 = x_i$), one can quite readily reduce the expression in (2.11) to that in (2.5).

Similar considerations apply to the development of the structure function using the properties of minimal cut sets. In this case, we denote the minimal cut sets of the system of interest to be $C_1, \ldots, C_k$, and we define cut structure functions as

$$c_j(\mathbf{x}) = 1 - \prod_{i \in C_j} (1 - x_i) . \tag{2.12}$$

Now $c_j(\mathbf{x}) = 1$ precisely when the minimal cut set C_j contains at least one working component. Since the system works if and only if every minimal cut set contains at least one working component, it follows that the system's structure function has the alternative representation

$$\varphi(\mathbf{x}) = \prod_{j=1}^{k} c_j(\mathbf{x}) . \tag{2.13}$$

Equation (2.13) can be written in the following equivalent form:

$$\varphi(\mathbf{x}) = \min_{\{1 \le j \le k\}} c_j(\mathbf{x}) = \min_{\{1 \le j \le k\}} \max_{\{i \in C_j\}} \{x_i\} \tag{2.14}$$

The discussion above, and in particular, the representations of the structure function in (2.9) and (2.13) make it apparent that a coherent system may be thought of as a parallel system in which each element is a series system in the components of a minimal path set. Similarly, the system can be thought of as a series system in which each element is a parallel system in the components of a minimal cut set.

Given two systems with structure functions φ_1 and φ_2 respectively, it is clear that the second system performs better than the first if $\varphi_1(\mathbf{x}) \le \varphi_2(\mathbf{x})$ for all $x \in \{0, 1\}^n$, since this inequality implies that the first system will always fail when the second system does. Two useful results concerning coherent systems are easily established by comparing appropriate structure functions. The first is that no system can perform better than the parallel system nor worse than the series system. This follows from the self-evident inequalities

$$\prod_{i=1}^{n} x_i \le \varphi(\mathbf{x}) \le 1 - \prod_{i=1}^{n} (1 - x_i) . \tag{2.15}$$

The second result is both more interesting and more useful. It concerns the effect of redundancy in system designs. If one has an n-component system and had the opportunity to enhance its performance by incorporating redundancy through the addition of n more components, one might reasonably ask which of two options would be preferable: (1) backing up every component, essentially replacing every component in the original design by a parallel system in two components, or (2) backing up the entire system by an identical system of order n (placed in parallel with the original system). The following theorem establishes that option (1) is better, documenting the well known engineering principle that componentwise redundancy is always better than systemwise redundancy.

Theorem 2.1. *Let φ be the structure function of a coherent system of order n. Then for any $\mathbf{x}$ and $\mathbf{y} \in \{0,1\}^n$,*

$$\varphi(1-(1-x_1)(1-y_1),\ldots,1-(1-x_n)(1-y_n)) \geq 1-(1-\varphi(\mathbf{x}))(1-\varphi(\mathbf{y})) \quad (2.16)$$

Proof. The structure function of the left-hand side of (2.16) is obtained by adopting the view that the system with componentwise redundancy can be considered to have the same structure as the original system, but with each component replaced by a parallel subsystem in two components. The ith subsystem works when $[1-(1-x_i)(1-y_i)] = 1$ and fails when $[1-(1-x_i)(1-y_i)] = 0$. Since the inequalities $1-(1-x_i)(1-y_i) \geq x_i$ and $1-(1-x_i)(1-y_i) \geq y_i$ hold for all i, it follows from the mononicity of φ that

$$\varphi(1-(1-x_1)(1-y_1),\ldots,1-(1-x_n)(1-y_n)) \geq \varphi(\mathbf{x})$$

and

$$\varphi(1-(1-x_1)(1-y_1),\ldots,1-(1-x_n)(1-y_n)) \geq \varphi(\mathbf{y})\ .$$

The latter inequality implies that

$$\varphi(1-(1-x_1)(1-y_1),\ldots,1-(1-x_n)(1-y_n)) \geq \max\{\varphi(\mathbf{x}),\varphi(\mathbf{y})\},$$

an inequality that is equivalent to (2.16). ∎

2.2 The Reliability of a Coherent System

Consider a coherent system in n independent components. If we fix a time t at which the system is examined, we may treat the ith component as working with probability p_i, that is, we may let $p_i = P(X_i = 1)$, where X_i is a Bernoulli variable representing the random state of the ith component at time t. We will define the *reliability* of a system at time t as the probability that it is working at that time. This probability will be denoted by $h(\mathbf{p})$ and can be computed from the structure function as

$$h(\mathbf{p}) = P(\varphi(\mathbf{X}) = 1) = E\varphi(\mathbf{X})\ . \quad (2.17)$$

The function $h(\mathbf{p})$ is multilinear, that is, it is linear in every p_i. For the bridge system pictured in Figure 2.2, the system reliability is given by

$$\begin{aligned}h(\mathbf{p}) &= E(\varphi(\mathbf{x}))\\ &= E(X_1X_4 + X_2X_5 + X_1X_3X_5 + X_2X_3X_4 - X_1X_2X_3X_4 - X_1X_2X_3X_5\\ &\quad - X_1X_3X_4X_5 - X_1X_2X_4X_5 - X_2X_3X_4X_5 + 2X_1X_2X_3X_4X_5)\\ &= p_1p_4 + p_2p_5 + p_1p_3p_5 + p_2p_3p_4 - p_1p_2p_3p_4 - p_1p_2p_3p_5\\ &\quad - p_1p_3p_4p_5 - p_1p_2p_4p_5 - p_2p_3p_4p_5 + 2p_1p_2p_3p_4p_5\ .\end{aligned}$$

When components are identically distributed, we have $p_i \equiv p$, and the reliability function h simplifies. For the bridge system in i.i.d. components, the reliability function can be written in terms of this common p, and reduces to

$$h(p) = 2p^2 + 2p^3 - 5p^4 + 2p^5\ . \tag{2.18}$$

In the i.i.d. case, we refer to h as the *reliability polynomial*. For the n-component series, parallel and k-out-of-n systems, the respective reliability polynomials are given by

$$h_1(p) = p^n, \quad h_2(p) = 1 - (1-p)^n \quad \text{and} \quad h_3(p) = \sum_{i=0}^{k-1} \binom{n}{i} (1-p)^i p^{n-i}\ . \tag{2.19}$$

For complex systems, computing the system's reliability can be cumbersome, even under the i.i.d. assumption. A useful tool for making this computation (and as we shall see, for bounding $h(p)$ above and below by partial sums) is the so-called inclusion-exclusion formula. It is simply the general formula for calculating the union of a collection of (possibly overlapping) events. For two events, it is usually called "the addition rule." The general formula is given below. For a proof, see Feller [35].

Theorem 2.2. *Let $A_1, A_2, \ldots, A_n$ be n events, that is, subsets of the sample space of a random experiment. Then the probability that at least one of the events occurs is given by*

$$\begin{aligned}P\left(\cup_{i=1}^n A_i\right) &= \sum_{i=1}^n P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)\\ &\quad - \cdots \pm P\left(\cap_{i=1}^n A_i\right)\ .\end{aligned} \tag{2.20}$$

If we denote the ith summation of the inclusion-exclusion formula by S_i, we may write

$$P\left(\cup_{i=1}^n A_i\right) = \sum_{i=1}^n (-1)^{i+1} S_i\ . \tag{2.21}$$

Further, partial sums of the formula in (2.21) can be shown to provide increasingly precise upper and lower bounds for the probability of interest. For instance,

$$P\left(\cup_{i=1}^n A_i\right) \le S_1, \quad P\left(\cup_{i=1}^n A_i\right) \ge S_1 - S_2, \quad P\left(\cup_{i=1}^n A_i\right) \le S_1 - S_2 + S_3, \quad \text{etc.}$$

Let's apply this tool in (2.20) to compute the reliability function for the coherent system of order 4 with minimal path sets $\{1\}, \{2,3\}, \{2,4\}, \{3,4\}$, that is, for system 4 in Table 2.1. Note that the inclusion-exclusion formula is precisely tailored for this calculation since a coherent system will function if and only if all the components in at least one of its minimal path sets are working. Let A_1, A_2, A_3 and A_4 be the events that all the components in each of the four minimal path sets above are working. Then the reliability function $h(\mathbf{p})$ of the system is equal to $P\left(\cup_{i=1}^{4} A_i\right)$. For simplicity, let's assume that each component has the same probability p of working. Then

$$\begin{aligned} h(p) &= S_1 - S_2 + S_3 - S_4 \\ &= (p + 3p^2) - (6p^3) + (3p^4 + p^3) - (p^4) \\ &= p + 3p^2 - 5p^3 + 2p^4 \ . \end{aligned} \tag{2.22}$$

More generally, consider a system in n i.i.d. components, and let p be the common probability that any given component is working at a fixed point in time. We will say that the reliability polynomial is in "standard form" when it is written as

$$h(p) = \sum_{i=1}^{n} d_i p^i \ . \tag{2.23}$$

The polynomial in (2.22) is in standard form. The other form we will be interested in is the so-called "pq form," where h is written as

$$h(p) = \sum_{i=1}^{n} c_i p^i q^{n-i} \ . \tag{2.24}$$

In Chapters 6 and 8, we discuss when and why one might prefer one form over the other.

3

System Signatures

The basic notion of a coherent system has been defined and illustrated in the preceding chapter, and various properties and tools have been established to assist in the study of such systems. Structure functions are admittedly quite revealing. They are in one-to-one correspondence with the coherent systems themselves and provide a way of indexing systems and also of comparing them. Also, they are unambiguous summaries of a system's design, and are more useful summaries than schematic diagrams or flow charts, which often look different but may correspond to one and the same system. We must, however, acknowledge that the artillery we have discussed thus far has some limitations. Since the number of coherent systems of order n grows exponentially with n, the indexing of systems through their structure functions tends to be of limited use in problems involving comparisons or optimization among systems. Structure functions are complex algebraic expressions that, in general, admit to multiple equivalent representations. For example, the expressions in (2.5) and (2.11) look quite different but are, in fact, equivalent forms of the structure function of the bridge system in Figure 2.2. In this chapter, we introduce an alternative index which, although less general than a structure function, has the virtues of being both quite manageable and easily interpreted. Most importantly, for systems of order n, this index is of fixed dimension; in fact, it resides in a bounded simplex in n-dimensional Euclidean space. We call this index the system's signature. Its precise meaning is specified below.

Definition 3.1. *Let τ represent a coherent system of order n. Assume that the lifetimes of the system's n components are independent and identically distributed (i.i.d.) according to the (continuous) distribution F. The* signature *of the system τ, denoted by $\mathbf{s}_\tau$, or simply by $\mathbf{s}$ when the corresponding system is clear from the context, is an n-dimensional probability vector whose ith element s_i is equal to the probability that the ith component failure causes the system to fail. In brief, $s_i = P(T = X_{i:n})$, where T is the failure time of the* *system and $X_{i:n}$ is the ith order statistic of the n component failure times, that is, the time of the ith component failure.*

Before illustrating the concept of system signatures, it seems advisable to scrutinize the definition above. In particular, it seems reasonable to question the wisdom of the i.i.d. assumption on component lifetimes. The notion of signature, as a certain probability vector, is well defined without this assumption, but the assumption is nonetheless made, and is made for good reason. Signatures will be used primarily in the comparison of system designs. It should be noted that a comparison between two systems with quite different component characteristics may well be either misleading or inconclusive. It is clear, for example, that a series system with four highly reliable components will outperform a four-component parallel system with relatively poor components. If the probability that the components of the series system last beyond a fixed mission time is 0.9, its reliability at that mission time is 0.6561, while that of a parallel system having four components with reliability 0.1 is 0.3439. It is clear, however, that parallel systems are preferable, in a general sense, to series systems. Indeed, the former's structure function uniformly dominates the latter's. Once the i.i.d. assumption is made, any remaining differences in system performance must be attributable to the system's design. In that sense, the assumption levels the playing field so that one has a basis for comparing the designs themselves. From an analytical point of view, signatures, as defined above in the i.i.d. setting, provide three major advantages. They allow one to utilize the tools of combinatorial mathematics for the calculation of system characteristics. Also, the well-known distribution theory for the order statistics of an i.i.d. sample from a continuous distribution F is available for studying the performance of a system with a given signature. Finally, signatures depend only on the permutation distribution of the n observed failure times and do not depend on the underlying distribution F. The signature vector can therefore be viewed as a pure measure of a system's design.

More can be said about the comparisons we will indulge in as we proceed. We will, for example, be primarily interested in comparing systems of the same order. While one could, in some instances, be interested in comparing systems of different sizes, it is far more common to compare systems of the same size, essentially investigating questions such as "Which of several possible configurations of components would be preferable for certain specific purposes?" In the words of the great Eastern philosopher Confucius, comparing apples to oranges is a rather fruitless endeavor. A second issue that should be addressed before proceeding is the possibility that results depending on signature vectors as defined above might in fact be irrelevant in studying and comparing the performance of real systems whose components are neither independent nor identically distributed. In addressing this concern, it should be said that, in any application of signature-related results in which the foundations of their definition are in doubt, one should proceed with considerable caution. It is probably worth adding that signature-based results may be inexact in such applications but are not necessarily irrelevant. Any mathematical result can only give guidance in real applications, as the assumptions under which the result is developed can't be checked with certainty in a given practical situation. What signatures do is tell us something about the design of the associated system. Knowing that one design is better than another (everything else being equal) is useful information as one diverges from the basic assumption of i.i.d. component lifetimes. If, for example, the component lifetimes could be considered independent and, while not identically distributed, nonetheless roughly comparable, selecting the system with a better signature should lead to better performance. Although an exact analysis would of course be desirable, characterizing system performance in non i.i.d. settings is a formidable analytical task, quite unlike the i.i.d. setting to be studied here.

The computation of system signatures is, in essence, a combinatorial exercise. That doesn't mean that it's simple. It only means that there is a well-organized body of knowledge and tools that can be applied to such problems. To describe the counting problem of interest, let's suppose that the random variables $X_1, X_2, \dots, X_n$ represent the failure times of the components of the n-component system under study. Since the Xs are assumed to be i.i.d. from some continuous distribution on $(0, \infty)$, the $n!$ permutations of these n distinct failure times are equally likely. As noted above, the ith element of $\mathbf{s}$ can be obtained as the probability $s_i = P(T = X_{i:n})$, where T is the failure time of the system and $X_{i:n}$ is the ith order statistic (that is, the ith smallest value) among the i.i.d. failure times $X_1, X_2, \dots, X_n$. Equivalently, we may obtain s_i as the ratio of n_i, the number of orderings for which the ith component failure causes system failure, to $n!$, the total number of possible orderings of the n failure times. The essential feature of the calculation of signatures is the counting of the number of permutations of the n potential component failure times that correspond with system failure upon the ith failure among the n components. Since T resides in the set $\{X_{1:n}, X_{2:n}, \dots, X_{n:n}\}$ with probability one, it follows that the signature $\mathbf{s}$ is a probability vector, that is, $s_i \geq 0$ for all i and $\sum_{i=1}^{n} s_i = 1$.

We now turn to the computation of the signature vector for some simple coherent systems. As an example of this computation, consider the three-component system pictured in Figure 3.1 below. The failure times X_1, X_2 and X_3 of the three components of this system can be ordered in $3! = 6$ ways, and these six possible permutations are equally likely due to the i.i.d. assumption. The "order-statistic equivalent" for the system failure time T is shown below for each permutation of the component failure times.

Fig. 3.1. A 3-component system with structure function $\varphi^*(\mathbf{x}) = x_1(x_2+x_3-x_2x_3)$

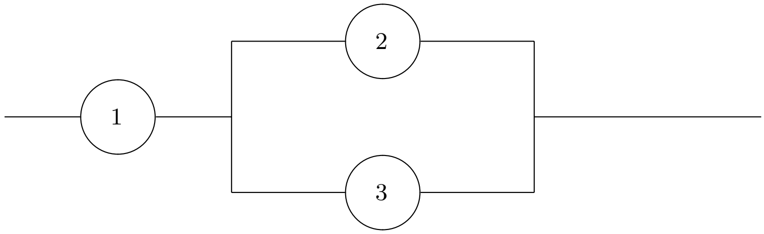


Table 3.1. The ordered component failure time which causes failure of system φ^* above

Ordered Component Failure Times	Order Statistic Equal to System Failure Time T
$X_1 < X_2 < X_3$	$X_{1:3}$
$X_1 < X_3 < X_2$	$X_{1:3}$
$X_2 < X_1 < X_3$	$X_{2:3}$
$X_2 < X_3 < X_1$	$X_{2:3}$
$X_3 < X_1 < X_2$	$X_{2:3}$
$X_3 < X_2 < X_1$	$X_{2:3}$

It follows that the system above has signature vector $\mathbf{s} = (1/3, 2/3, 0)$. It is easy to show that the five distinct coherent systems of order 3 have the signatures $(1, 0, 0), (0, 1, 0), (0, 0, 1), (1/3, 2/3, 0)$ and $(0, 2/3, 1/3)$. The first three of these signatures correspond to the i-out-of-3 systems for $i = 1, 2, 3$, and the fifth corresponds to the system in which one component is in parallel with a series system in the other two components. While the combinatorics involved in calculating the signatures of systems of a given order can be fairly complex, it is worth noting that, via the notion of "duality," the amount of calculation can be cut roughly in half, as the signature of a system's "dual" can be obtained from the system's signature via symmetry arguments. Of the five systems mentioned above, the first system is the dual of the third and the fourth is the dual of the fifth. Table 3.2 provides the signatures of the 20 distinct coherent systems of order 4.

Table 3.2. Coherent Systems of Order 4

System	Minimal cut sets	Signature
1	{1}, {2}, {3}, {4}	(1, 0, 0, 0)
2	{1}, {2}, {3, 4}	(1/2, 1/2, 0, 0)
3	{1}, {2, 3}, {2, 4}	(1/4, 7/12, 1/6, 0)
4	{1}, {2, 3}, {2, 4}, {3, 4}	(1/4, 3/4, 0, 0)
5	{1}, {2, 3, 4}	(1/4, 1/4, 1/2, 0)
6	{1, 2}, {1, 3}, {1, 4}	(0, 1/2, 1/4, 1/4)
7	{1, 2}, {1, 3}, {1,4}, {2, 3}	(0, 2/3, 1/3, 0)
8	{1, 2}, {1, 3}, {1,4}, {2, 3}, {2, 4}	(0, 5/6, 1/6, 0)
9	{1, 2}, {1, 3}, {1,4}, {2, 3}, {2, 4}, {3, 4}	(0, 1, 0, 0)
10	{1, 2}, {1, 3}, {2, 4}, {3, 4}	(0, 2/3, 1/3, 0)
11	{1, 2}, {2, 4}, {3, 4}	(0, 1/2, 1/2, 0)
12	{1, 2}, {3, 4}	(0, 1/3, 2/3, 0)
13	{1, 2}, {1, 3}, {1, 4}, {2, 3, 4}	(0, 1/2, 1/2, 0)
14	{1, 2}, {1, 3}, {2, 3, 4}	(0, 1/3, 2/3, 0)
15	{1, 2}, {1, 3, 4}, {2, 3, 4}	(0, 1/6, 5/6, 0)
16	{1, 2}, {1, 3, 4}	(0, 1/6, 7/12, 1/4)
17	{1, 2, 3}, {1, 2, 4}	(0, 0, 1/2, 1/2)
18	{1, 2, 3}, {1, 2, 4}, {1, 3, 4}	(0, 0, 3/4, 1/4)
19	{1, 2, 3}, {1, 2, 4}, {1, 3, 4}, {2, 3, 4}	(0, 0, 1, 0)
20	{1, 2, 3, 4}	(0, 0, 0, 1)

We now establish a fundamental property of a system's signature $\mathbf{s}$, namely, that the distribution of the system lifetime T, given i.i.d. components lifetimes with c.d.f. F, can be expressed as a function of $\mathbf{s}$ and F alone. The following representation is drawn from Samaniego [61].

Theorem 3.1. *Let $X_1, \ldots, X_n$ be the i.i.d. component lifetimes of an n-component coherent system with signature $\mathbf{s}$, and let T be the system's lifetime. Then*

$$\overline{F}_T(t) \equiv P(T > t) = \sum_{i=1}^{n} s_i \sum_{j=0}^{i-1} \binom{n}{j} (F(t))^j \left(\overline{F}(t)\right)^{n-j} . \tag{3.1}$$

Proof. We first note that the system fails concurrently with the failure of one of its components, so that T will necessarily take on the values of one of the order statistics $X_{i:n}$ of the sample $X_1, \ldots, X_n$, that is, $T \in \{X_{1:n}, X_{2:n}, \ldots, X_{n:n}\}$ with probability 1. Then, utilizing the Law of Total Probability and the i.i.d. assumption on component lifetimes, we may write

$$\begin{aligned} P(T > t) &= \sum_{i=1}^{n} P\left(T > t, T = X_{i:n}\right) \\ &= \sum_{i=1}^{n} P\left(T > t \mid T = X_{i:n}\right) P\left(T = X_{i:n}\right) \\ &= \sum_{i=1}^{n} s_i P\left(X_{i:n} > t\right) \\ &= \sum_{i=1}^{n} s_i \sum_{j=0}^{i-1} \binom{n}{j} (F(t))^j \left(\overline{F}(t)\right)^{n-j} . \qquad \blacksquare \end{aligned}$$

By interchanging the order of the summations in (3.1), the representation of the system survival function in (3.1) may be written in the alternative form

$$\overline{F}_T(t) = \sum_{j=0}^{n-1} \left(\sum_{i=j+1}^{n} s_i \right) \binom{n}{j} (F(t))^j \left(\overline{F}(t)\right)^{n-j} . \tag{3.2}$$

If one considers the chances that a system based on n i.i.d components is working at a fixed point in time t_0, then setting $p = \overline{F}(t_0)$ and $q = F(t_0)$, we derive the reliability polynomial $h(p)$ in pq form from (3.2). Two equivalent versions of h are

$$h(p) = \sum_{j=0}^{n-1} \left(\sum_{i=j+1}^{n} s_i \right) \binom{n}{j} q^j p^{n-j} \text{ and } h(p) = \sum_{j=1}^{n} \left(\sum_{i=n-j+1}^{n} s_i \right) \binom{n}{j} p^j q^{n-j} . \tag{3.3}$$

The representation in (3.2) can also be written as a function involving the odds $G(t) = F(t)/\overline{F}(t)$ of failure vs. survival. This form of the representation of $\overline{F}_T(t)$ in terms of signatures will prove to be quite useful in the sequel, and is recorded below:

$$\overline{F}_T(t) = \left(\overline{F}(t)\right)^n \sum_{j=0}^{n-1} \left(\sum_{i=j+1}^{n} s_i \right) \binom{n}{j} (G(t))^j . \tag{3.4}$$

A more detailed proof of Theorem 3.1 invoking Lemma 8.3.11 of Randles and Wolfe [60] on the independence of the order statistics $\{X_{i:n}\}$ and the ranks of the original observations $X_1, X_2, \ldots, X_n$ is given in Kochar, Mukerjee and Samaniego [51].

The proof of Theorem 3.1 contains an elementary fact that is of independent interest and will also prove quite useful in the sequel. Note that the survival function of the system lifetime T may be written in terms of the survival functions of the order statistics of the component failure times, that is,

$$P(T > t) = \sum_{i=1}^{n} s_i P(X_{i:n} > t) \, . \tag{3.5}$$

Utilizing the well-known identity for positive random variables Y, namely,

$$EY = \int_0^\infty \overline{F}(y) \, \mathrm{d}y \, , \tag{3.6}$$

another useful connection between the system lifetime and the ordered failure times follows:

$$ET = \sum_{i=1}^{n} s_i EX_{i:n} \, . \tag{3.7}$$

The representation in (3.1) can be applied to obtain useful representations of a system's density function and failure rate when F is absolutely continuous. For example, the density function $f(x)$ may be obtained from (3.1) as follows.

Corollary 3.1. *Let $X_1, \ldots, X_n \sim F$ be the i.i.d. component lifetimes of an n-component coherent system with signature* $\mathbf{s}$*, and let T be the system's lifetime. If F is absolutely continuous, then*

$$f_T(t) = -(\partial/\partial t)P(T > t) = \sum_{i=1}^{n} i \; s_i \binom{n}{i} \left(F(t)\right)^{i-1} \left(\overline{F}(t)\right)^{n-i} f(t) \, . \tag{3.8}$$

Proof. Differentiating $\overline{F}_T(t)$ in formula (3.1) yields an interior summation with alternating signs in which all elements but the one shown in (3.8) cancel out. ∎

It follows that the system failure rate $r_T(t)$, defined as the ratio

$$\frac{f_T(t)}{\overline{F}_T(t)} \, ,$$

can be written in terms of the signature vector $\mathbf{s}$ and the underlying component distribution F. The ratio of the density in (3.8) to the survival function in (3.1) may be algebraically simplified to obtain a useful representation of the system's failure rate.

Corollary 3.2. *Consider an n-component coherent system with signature* $\mathbf{s}$*, and assume that the component lifetimes $X_1, \ldots, X_n$ are i.i.d. with distribution F and density f. Let T be the system lifetime. Then*

$$r_T(t) = \frac{\sum_{i=1}^{n} i \; s_i \binom{n}{i} \left(F(t)\right)^{i-1} \left(\overline{F}(t)\right)^{n-i+1}}{\sum_{i=1}^{n} s_i \sum_{j=0}^{i-1} \binom{n}{j} \left(F(t)\right)^{j} \left(\overline{F}(t)\right)^{n-j}} \, r(t) \, , \tag{3.9}$$

where $r(t) = \left(f(t)/\overline{F}(t)\right)$, the common failure rate of the components.

An equivalent and occasionally more useful version of (3.9) is the following:

$$r_T(t) = \frac{\sum_{i=0}^{n-1}(n-i)s_{i+1}\binom{n}{i}(F(t))^i\left(\overline{F}(t)\right)^{n-i}}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_j\right)\binom{n}{i}(F(t))^i\left(\overline{F}(t)\right)^{n-i}}\, r(t) \ , \tag{3.10}$$

or, in terms of the odds function $G(t) = F(t)/\overline{F}(t)$,

$$r_T(t) = \frac{\sum_{i=0}^{n-1}(n-i)s_{i+1}\binom{n}{i}(G(t))^i}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_j\right)\binom{n}{i}(G(t))^i}\, r(t) \ . \tag{3.11}$$

We will be interested in the comparison of two systems with i.i.d. components. As is clear from equation (3.1), the lifetime of a coherent system with i.i.d. components depends on the structure of the system only through the signature $\mathbf{s}$. Indeed, if two systems in i.i.d. components have the same signature, the stochastic behavior of their lifetimes is identical. It is natural to ask if two different coherent systems can have the same signature. The answer is yes; one can see from Table 3.2 that the four-component systems labeled as systems 11 and 13 have the same signature. The twenty coherent systems of order four give rise to precisely 17 distinct signatures.

While the class of all coherent systems of a given size (or even the overall collection of coherent systems of arbitrary order) is arguably the collection of systems on which one would wish to concentrate in a particular application, the class does have some limitations which will lead us to broaden our perspective. I mentioned earlier that the number of coherent systems of order n is not precisely known for general n and is quite large, even for moderate size n. For any fixed n, the space of coherent systems of order n is, obviously, discrete. As we shall see in the sequel, this has some negative consequences, both mathematically and practically. The mathematical difficulty is that in problems in which an optimal coherent system is sought, one typically must focus on finding approximately optimal systems via some appropriate discrete search algorithm. In other words, problems aimed at finding optimal coherent systems tend to be analytically intractable. The practical problem with coherent systems is less apparent at this point, but will become quite clear in certain specific problems taken up in Chapter 7. Briefly, the fact is that it is possible to expand the class of coherent systems to a larger collection and that the solution to certain optimality problems lies outside of the subset of coherent systems. We will see that, in certain problems, one can actually do better, in a sense that will be made specific, by using a "system" that is not coherent. The expansion pursued below is based on the familiar notion of "randomization." Indeed, what will be advocated here, in selected circumstances, is the process of selecting a coherent system at random. This process leads to the concept of a mixed system, to which I now turn. Mixed systems were first treated in Boland and Samaniego [20].

Let us suppose that we have an (essentially) unlimited supply of components whose lifetimes are i.i.d. with common distribution F. Consider the collection of all coherent systems of order n. We could, in principle, have a warehouse in which all such systems are in stock, and we could conceive of the possibility of making available, upon demand, any particular coherent system with n i.i.d. components. While this collection may be quite large, it is finite. Limiting ourselves to this collection may have negative consequences, as alluded to above. Let us consider, instead, the process of selecting a coherent system at random according to a fixed and known probability distribution $\mathbf{p}$. Let's suppose that the probability vector $\mathbf{p}$ is m-dimensional and gives positive weight to m distinct coherent systems of order n with signature vectors $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_m$ (each assumed to have components whose lifetimes are i.i.d with distribution F). Then it is clear that

$$\begin{aligned}&P(\text{system fails upon the } i\text{th component failure})\\&= \sum_{k=1}^{m} P(k\text{th syst. chosen}) P(i\text{th comp. failure kills syst.} \mid k\text{th syst. chosen})\\&= \sum_{k=1}^{m} p_k s_{ki} \ . \qquad (3.12)\end{aligned}$$

It follows that the signature $\mathbf{s}^*$ associated with the process of selecting among these m systems according to the probability distribution $\mathbf{p}$ is the vector equal to the mixture of the signature vectors $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_m$, that is, $\mathbf{s}^* = \sum_{k=1}^{m} p_k \mathbf{s}_k$.

To what extent does the consideration of stochastic mixtures of coherent systems (hereafter referred to as "mixed systems") broaden the class of coherent systems? The broadening is, in fact, quite substantial. The signature of a k-out-of-n system is the n-dimensional unit vector $\mathbf{s}_{k:n} = (0, \ldots, 0, 1, 0, \ldots, 0)$, with 1 as the vector's kth element. It is thus clear that any probability vector $\mathbf{p}$ in the simplex

$$\left\{ \mathbf{p} \in [0,1]^n : \sum_{i=1}^{n} p_i = 1 \right\} \qquad (3.13)$$

is the signature of a mixed system, namely, the system which mixes the k-out-of-n systems with mixing distribution $\mathbf{p}$. This simple observation follows from (3.12) since the n-dimensional probability vector $\mathbf{p}$ can be written as

$$\mathbf{p} = \sum_{k=1}^{n} p_k \mathbf{s}_{k:n} \ . \qquad (3.14)$$

Expanding the space of coherent systems of order n to the space of all mixed systems of order n has a number of mathematical benefits. It makes the index of the systems under consideration continuous, taking values in the simplex in (3.13). Thus, instead of being restricted to discrete search methods and approximate solutions to optimization problems of interest, such problems will be amenable to analytical treatment using the tools of differential calculus. As mentioned earlier, the solutions to certain optimization problems turn out to be mixed rather than coherent systems; in such problems, the potential exists for improving over the "score" of any coherent system by using a particular mixed system. One further comment about the utility of mixed systems should be made: the representation results in (3.1) - (3.11), as well as all signature-related results to be presented in the sequel, apply equally to coherent or mixed systems. Indeed, since mixed systems include coherent systems as special cases, that is, as degenerate mixtures placing all their mass on a single coherent system, it suffices to say that all signature-related results in this monograph apply to the entire class of mixed systems.

But a critical question remains to be answered. Do mixed systems correspond to a physical reality that can be utilized in practice? If not, the potential mathematical advantages alluded to above are merely window-dressing. I will therefore briefly indulge in some apologetics for the concept of mixtures of coherent systems. To a decision theorist, or to a statistician working in the field of survey sampling or experimental design, the notion of randomization is both natural and essential. In the first instance, it is clear from the theory of games that the best strategies often involve randomization, and the decision theorist's goal is to select the "best" decision rule available, a goal that, in certain situations, leads to some form of randomization. In the latter fields, randomization is seen as a tool that protects the statistician from both known and unknown biases. Sir Ronald Fisher, the brilliant pioneer in the field of experimental design (and in mathematical statistics generally) was an early and strong advocate of randomization in designed experiments. What we have advocated above is nothing more than randomization within the class of coherent systems. The fact that a mixed system can be physically realized by a simple randomization process is the key observation in recognizing the concept's practical utility.

To utilize a mixed system, one simply selects a coherent system at random according to a particular mixing distribution, and one then uses the coherent system so chosen. The signature of the mixed system is interpreted as an expectation or long-run average over many applications of the mixed system. The signature's first element, for example, will be the limiting probability that the first component failure causes the system to fail. This interpretation suggests that the natural domain of application of mixed systems is a scenario in which such a system will be used many times. Thus, its failure history will, in the long run, be well represented by the mixed system's signature. Recalling (3.14), we note that, in the i.i.d. setting, any system, coherent or mixed, will have the same long run performance as a mixture of k-out-of-n systems. The mixing distribution would simply be the signature vector of the desired system. An important and useful implication of these considerations is that any coherent or mixed system of order n will have the same expected performance as a specific mixture of k-out-of-n systems. The "warehouse" mentioned on page 29 could therefore provide a client with a mixture of k-out-of-n systems whose performance matched the client's requirements, that is, had the desired signature.

I mentioned above that a non-degenerate mixture of coherent systems is not itself coherent. This may not be self evident, especially in view of the fact that each application of a mixed system amounts to picking a particular coherent system at random. Of course, successive applications of the mixed system will typically result in the use of different coherent systems from one application to the next. The latter fact notwithstanding, every application of a mixed system selects a coherent system with probability one. Since every coherent system is monotone, every mixed system also enjoys the monotonicity property. Fixing a non-working component cannot adversely affect the performance of a mixed system. So what is incoherent about a mixed system? The incoherence comes from the following technicality: only the components of the coherent system selected for use are relevant. The state of the components of all other (unused) coherent systems to which the mixed system gives weight makes no difference to the success or failure of the system actually chosen. One could argue that in the course of repeated applications, all these other coherent systems will come into play. But relevance is defined in terms of a single application of the system and, according to that definition, a mixed system will have irrelevant components. This technicality proves to be of minor importance, since when one focuses on relevance over repeated trials, the issue vanishes entirely.

One final comment about the utility of mixtures should be made. The intended domain for use of mixed systems is in applications in which a large number of systems will be purchased and used. As will be seen in Chapter 7, the best available system, when both performance and cost are taken into consideration, may well be a non-degenerate mixture of coherent systems. In some settings, the exclusive use of a particular coherent system might be justified by convenience considerations or by its being "almost optimal" relative to the criterion used to rank systems. The fact remains that one may be able to achieve better results by randomizing among two or more coherent systems. In using the optimal "mixed system," one must recognize that the system used in a given instance will be suboptimal, but that mixing these systems according to the prescribed recipe produces an optimal result when averaged over the many times in which the mixed system is applied.

In the next two chapters, we will consider a variety of problems in which our primary interest will be in comparing one coherent (or mixed) system with another. The most natural comparison one can make is between two systems of the same size. The preservation theorems of Section 4.2 and the characterization results in section 4.4 deal precisely with that type of comparison. But there are also occasions in which one is interested in comparing two systems in i.i.d. components whose orders n and m do not happen to match (say, with $n < m$). Our approach to such a problem is to attempt to "convert" the smaller system into an equivalent system of order m, (that is, a system in m i.i.d. components with exactly the same lifetime distribution as the smaller system). If this were possible, we would be in a position to use the results of Sections 4.2 and 4.4 on comparing two systems of the same size. It is not obvious, at first view, that the equivalence mentioned can in fact be realized. The fact that it can is a consequence of the following theorem. Assuming i.i.d. component lifetimes, this result ensures that for any system of size n (a positive integer assumed to be smaller than the integer m), there exists a system of size m with the same lifetime distribution. To get from the smaller system to the larger one requires $m - n$ consecutive applications of the result below.

Theorem 3.2. *Let* $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ *be the signature of a coherent or mixed system based on n i.i.d. components with common lifetime distribution F. Then the coherent or mixed system with $(n+1)$-components with i.i.d. lifetimes $\sim F$ and corresponding to the signature vector*

$$\mathbf{s}^* = \left(\frac{n}{n+1}s_1, \frac{1}{n+1}s_1 + \frac{n-1}{n+1}s_2, \frac{2}{n+1}s_2 + \frac{n-2}{n+1}s_3, \ldots, \right. \\ \left. \frac{n-1}{n+1}s_{n-1} + \frac{1}{n+1}s_n, \frac{n}{n+1}s_n\right) \tag{3.15}$$

has the same lifetime distribution as the n-component system with signature $\mathbf{s}$.

Proof. Assume that all components in the following discussion have i.i.d. lifetimes with common distribution F. We wish to show that, for a given n-component mixed system with signature $\mathbf{s}$, there is an $(n+1)$-component system (namely, the system with signature $\mathbf{s}^*$) with the same lifetime distribution, thereby being stochastically equivalent to the original system. It suffices to prove (3.15) when the original system is an arbitrary k-out-of-n system, that is, when $\mathbf{s}$ is the unit vector

$$\mathbf{s}_{k:n} = (0, \ldots, 0, 1, 0, \ldots, 0) \ , \tag{3.16}$$

where only the kth element is different from zero. We will show that this n-dimensional signature is equivalent to the $(n+1)$-dimensional signature

$$\begin{aligned} \mathbf{s}^* &= \frac{n-k+1}{n+1}\mathbf{s}_{k:n+1} + \frac{k}{n+1}\mathbf{s}_{k+1:n+1} \\ &= \left(0, \ldots, 0, \frac{n-k+1}{n+1}, \frac{k}{n+1}, 0, \ldots, 0\right) \ , \end{aligned} \tag{3.17}$$

where the two non-zero elements of $\mathbf{s}^*$ are in the kth and $(k+1)$st positions. To see that this implies the claimed equivalence in the theorem, assume that, for any k, the signature in (3.16) is equivalent to the signature in (3.17). Then, since an arbitrary n-dimensional signature $\mathbf{s}$ may be written as a mixture of the signatures of k-out-of-n systems, that is, as

$$\mathbf{s} = \sum_{i=1}^{n} s_i \mathbf{s}_{i:n} \ , \tag{3.18}$$

one may use the equivalence of (3.16) and (3.17) to convert $\mathbf{s}$ into the vector in (3.15).

A direct proof of the equivalence of the signatures in (3.16) and (3.17) could proceed by determining the validity of equating the corresponding survival functions in (3.1). We will examine such an equation at a fixed time point t and denote by p and q, respectively, the probability of the success or failure of any given component at time t; thus, we will let $q = F(t)$ and $p = \overline{F}(t)$. As we will see, the equation of interest holds for all values of p and thus for all choices of t. Writing the reliability polynomial as on the left-hand side of (3.3) for the signature vectors in (3.16) and (3.17), and setting these expressions equal to each other, yields

$$\sum_{j=0}^{k-1} \binom{n}{j} q^j p^{n-j} = \frac{n-k+1}{n+1} \sum_{j=0}^{k-1} \binom{n+1}{j} q^j p^{n+1-j} + \frac{k}{n+1} \sum_{j=0}^{k} \binom{n+1}{j} q^j p^{n+1-j} \ . \tag{3.19}$$

One could then replace each "q" in equation (3.19) by "$1-p$," write both sides of the equation as polynomials in p, and then verify that the coefficients of p^k on either side of (3.19) are equal for $k = 0, 1, \ldots, n+1$. This approach to the desired identity is clearly a cumbersome algebraic exercise. We shall, instead, demonstrate that the equality in (3.19) is valid by working with an alternative representation of the equation.

Suppose that X is a binomial random variable and that Y is a Bernoulli variable, with $X \sim B(n, q)$ and $Y \sim B(1, q)$, and assume that X and Y are independent. (Note that $q = F(t)$ is being used as the probability of "success" in the Bernoulli trials associated with these variables.) We can then identify equation (3.19) as the following equation involving X and Y:

$$P(X < k) = \frac{n-k+1}{n+1} P(X+Y < k) + \frac{k}{n+1} P(X+Y < k+1) \ . \tag{3.20}$$

It is evident that equation (3.20) can be rewritten as

$$P(X < k) = P(X + Y < k) + \frac{k}{n+1} P(X + Y = k)\,. \tag{3.21}$$

Upon dividing both sides of (3.21) by $P(X < k)$, we note that

$$\frac{P(X + Y < k)}{P(X < k)}$$

is equal to the conditional probability $P(X + Y < k \mid X < k)$. We thus see that (3.21) holds if and only if

$$P(X + Y \geq k \mid X < k) = \frac{k}{n+1} \; \frac{P(X + Y = k)}{P(X < k)}\,. \tag{3.22}$$

We will now demonstrate that (3.22) holds. We do so by displaying a sequence of equivalent statements, the last of which is transparently true. We begin by exploiting the fact that Y is a Bernoulli variable and is independent of X; this allows us to rewrite (3.22) as

$$q\,\frac{P(X = k-1)}{P(X < k)} = \frac{k}{n+1} \; \frac{qP(X = k-1) + pP(X = k)}{P(X < k)} \tag{3.23}$$

$$\begin{aligned}
\Longleftrightarrow\; & q\binom{n}{k-1} q^{k-1} p^{n-k+1} \\
& = \frac{k}{n+1}\left(q\binom{n}{k-1} q^{k-1} p^{n-k+1} + p\binom{n}{k} q^k p^{n-k}\right)
\end{aligned} \tag{3.24}$$

$$\Longleftrightarrow \binom{n}{k-1} q^k p^{n-k+1} = \frac{k}{n+1}\left(\binom{n}{k-1} + \binom{n}{k}\right) q^k p^{n-k+1} \tag{3.25}$$

$$\Longleftrightarrow \binom{n}{k-1} = \frac{k}{n+1}\binom{n+1}{k}\,. \tag{3.26}$$

We thus see that equation (3.19) is equivalent to the trivial relationship between binomial coefficients displayed in (3.26). Since (3.19) is equivalent to the fact that the systems with the signatures **s** and **s*** in the statement of the theorem have identical distributions, the proof is complete. ■

The form of the signature **s*** in (3.15) makes it evident that, when **s** is symmetric, the equivalent signature **s*** will inherit this property. This result is recorded as

Corollary 3.3. *Let* $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ *be the signature of a mixed system based on* n *i.i.d. components with common lifetime distribution* F*, and let* $\mathbf{s}^*$ *be the signature in* (3.15) *of the equivalent system with* $(n+1)$*-components based on i.i.d. lifetimes* $\sim F$*. If* $\mathbf{s}$ *is symmetric, that is, if* $s_i = s_{n-i+1}$ *for all* i*, then* $\mathbf{s}^*$ *is symmetric as well. Furthermore, if* (3.15) *is applied repeatedly to a signature vector* $\mathbf{s}_1$ *of length* n *to obtain an equivalent signature* $\mathbf{s}_2$ *of length* $m > n$*, then the symmetry of* $\mathbf{s}_1$ *implies the symmetry of* $\mathbf{s}_2$*.*

The concept of the duality of two systems was defined formally in Chapter 2. Since taking advantage of duality relationships reduces the calculation of the signatures of all coherent systems of a given size approximately in half, it is useful to record the relationship between the signature of a system and its dual. This is done in the following theorem. For a proof, see Kochar, Mukerjee and Samaniego [51].

Theorem 3.3. *Let* $\mathbf{s}$ *be the signature of a coherent system* φ *whose* n *components have i.i.d. lifetimes, and let* $\mathbf{s}^D$ *be the signature of its dual system* φ^D*. Then*

$$s_i = s_{n-i+1}^D \quad \textit{for } i = 1, 2, \ldots, n \,. \tag{3.27}$$

We note that Theorem 3.3 holds as well for the broader class of mixed systems of order n.

We close this chapter with a useful notion and a related expression to which we will return in Chapter 5. In (3.5), we displayed an explicit representation of the survival function of a system in terms of its signature vector and the survival functions of the order statistics corresponding to the n i.i.d. component failure times. A similar representation may be developed for the probability $P(T_1 \leq T_2)$, where T_1 and T_2 are the lifetimes of mixed systems of orders n and m based on two independent i.i.d. samples of sizes n and m from underlying distributions F_1 and F_2, respectively. The representation of interest is drawn from Hollander and Samaniego [43], whose proof differs from the one below in that it utilizes integral representations of probabilities of the form $P(X \leq Y)$. Here, we provide a simpler, direct proof.

Theorem 3.4. *Let* T_1 *and* T_2 *represent the lifetimes of mixed systems of orders* n *and* m *with respective signatures* $\mathbf{s}_1$ *and* $\mathbf{s}_2$*. Assume that the* n *components of system 1 have i.i.d. lifetimes governed by the continuous distribution* F_1*, and let* $\{X_{1:n}, X_{2:n}, \ldots, X_{n:n}\}$ *be the corresponding ordered component lifetimes. Similarly, assume that the* m *components of system 2 have i.i.d. lifetimes governed by the continuous distribution* F_2*, and let* $\{Y_{1:m}, Y_{2:m}, \ldots, Y_{m:m}\}$ *be the corresponding ordered component lifetimes. Finally, assume the two samples are independent. Then*

$$P(T_1 \leq T_2) = \sum_{i=1}^{n} \sum_{j=1}^{m} s_{1i} s_{2j} P\left(X_{i:n} \leq Y_{j:m}\right) \tag{3.28}$$

Proof. Using the Law of Total Probability and the i.i.d. assumption on component lifetimes, we may obtain the desired result by writing

$$
\begin{aligned}
&P(T_1 \leq T_2) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} P(T_1 = X_{i:n}) P(T_2 = Y_{j:m}) P(T_1 \leq T_2 \mid T_1 = X_{i:n}, T_2 = Y_{j:m}) \\
&= \sum_{i=1}^{n} \sum_{j=1}^{m} s_{1i} s_{2j} P(X_{i:n} \leq Y_{j:m}). \qquad \blacksquare
\end{aligned}
$$

4

Signature-Based Closure, Preservation and Characterization Theorems

4.1 An Application to the IFR Closure Problem

In Chapter 3, we referred to the failure rate function of a component or system and mentioned its standard definition. The function merits some further discussion. We will begin this section by reviewing the concept and its relevance in reliability studies. The existence of a "failure rate" requires that the distribution function F of the random variable of interest be absolutely continuous, that is, that F be differentiable. Suppose T is a random variable with distribution function F and density function f, and as usual, let $\overline{F} = 1 - F$. Then the failure rate of T is defined as the ratio $r(t) = f(t)/\overline{F}(t)$. Here, as in typical applications of the concept, we will assume that $T \geq 0$ with probability 1. This, of course, is an intuitive and necessary assumption when dealing with the lifetime or time until failure of an engineered system. Because of the approximation

$$f(t)\Delta t \;/\; \overline{F}(t) \approx P(t < T \leq t + \Delta t) \;/\; P(T > t) \;, \tag{4.1}$$

the function $r(t)$ is generally thought of as the instantaneous rate of failure of the system at time t given that it has survived until time t. The function thus tracks a system's vulnerability to failure as the system ages.

Among parametric lifetime distributions, the exponential distribution with density function

$$f(t) = \lambda \mathrm{e}^{-\lambda t}, \qquad t > 0 \tag{4.2}$$

is ubiquitous for a number of reasons. For example, it is the unique continuous probability model with the "memoryless property": if T is exponentially distributed, then $P(T > x + t \mid T > t) = P(T > x)$. In words, this equation implies that, if an item with an exponentially distributed lifetime has survived until time t, then it is as good as new, i.e., the distribution of its remaining (or residual) lifetime has the same exponential distribution as when the item was new. A further defining property of exponential variables is their constant failure rate – as long as the item is working, the chances of proximate failure is always the same. This follows from the fact that a variable T with density in (4.2) has survival function

$$\overline{F}(t) = \int_t^\infty \lambda e^{-\lambda x} dx = e^{-\lambda t}, \qquad t > 0 \tag{4.3}$$

so that its failure rate is given by

$$r(t) = f(t)/\overline{F}(t) = \lambda \quad \text{for all } t > 0 \, . \tag{4.4}$$

In view of the memoryless property of the exponential distribution, its constant failure rate is not unexpected; in fact, either property can be derived from the other. While the latter property has a certain aesthetic appeal, it is the very property that limits the applicability of the model. Most real systems have failure rates that change over time. For many real systems, the failure rate that one might expect would have a bathtub shape – initially decreasing, as the item grows stronger in early life, but eventually increasing as the system begins to deteriorate and has an increasing propensity for failure. While families of distributions with bathtub-shaped failure rates have received some attention in the literature, their general lack of tractability has been an impediment to their use. Perhaps the most widely studied models for failure time data are those with a monotone failure rate. Models for components or systems which deteriorate with age are of special interest in practical applications. It is often reasonable to suppose that an engineered system begins a slow but steady process of deterioration from its very inception. There are many parametric models which have such a property, including gamma and Weibull models with their "shape" parameters greater than 1. The entire class of distributions with this property has the special appeal of including many different shapes and parametric forms yet being, in total, a non-parametric class which makes no specific assumption about the form of the distribution or the density function. This class is formally defined below.

Definition 4.1. *Let X be a positive random variable with distribution F, density f and failure rate r. Then X is said to have an* increasing[1] failure rate *if $r(t)$ is increasing in t. The collection of all distributions with an increasing failure rate is referred to as the* IFR class*.*

The definition above notwithstanding, the IFR property can be defined solely in terms of the distribution function F and does not require its absolute continuity (i.e., the existence of the density function f). The idea that a system is deteriorating with age can be captured by the condition that the conditional probability $P(X > t+x \mid X > t)$ is decreasing in t, that is, that the probability

[1] Here, and in general in this monograph, the term "increasing" ("decreasing") is taken to mean "nondecreasing" ("nonincreasing").

of surviving x additional units of time is a decreasing function of age. This condition may be written as

$$\frac{\overline{F}(t+x)}{\overline{F}(t)} \text{ is decreasing for } t > 0 . \tag{4.5}$$

Another well known equivalent formulation of the IFR property involves the so-called hazard function R, which is defined as

$$R(t) = -\ln \overline{F}(t) . \tag{4.6}$$

A distribution has the IFR property if the corresponding hazard function R is a convex function of t. When F is absolutely continuous, all three definitions can be easily shown to be equivalent. For example, if F has density f, the hazard function $R(t)$ is convex if and only if $\partial/\partial t R(t)$ is increasing in t, but of course

$$\frac{\partial}{\partial t} R(t) = -\frac{\partial}{\partial t} \ln \overline{F}(t) = \frac{f(t)}{\overline{F}(t)} = r(t) , \tag{4.7}$$

implying that the convexity of R is equivalent to Definition 4.1. The relationship between R and F is often expressed as

$$\overline{F}(t) = \mathrm{e}^{-R(t)} , \tag{4.8}$$

and the relationship between R and r is often expressed as

$$R(t) = \int_0^t r(x)\mathrm{d}x . \tag{4.9}$$

Because of the latter representation, R is sometimes called the "cumulative hazard" or the "cumulative failure rate." The nomenclature varies among those who study and/or use these ideas. The phrase "hazard rate" is often used in referring to the function r.

In addition to being a reasonable way to model the behavior of systems that tend to deteriorate with time, the IFR class has a number of interesting properties. It is known, for example, that an IFR distribution is absolutely continuous on the set $\{t \mid F(t) < 1\}$. In other words, if an IFR distribution F has a jump point, it can occur only at the supremum of its support set $(0, z]$ and $F(z) = 1$. This includes the possibility that $z = \infty$, in which case F is absolutely continuous on the whole real line. The IFR class is known to be closed under convolutions (that is, if X and Y are independent IFR variables, then $X + Y$ is also IFR) but it is not closed under mixtures. Indeed, the mixture of two different exponential distributions, both boundary members of the IFR class, has a decreasing failure rate. Another form of closure that is of great interest in reliability is closure under the formation of coherent systems. Specifically, it is of interest to know whether or not a coherent system in n

independent components will have a given property when each component has that property. It is known that the IFR class does not enjoy such closure. For example, one can show that a parallel system in two components with independent exponentially distributed (and thus IFR) lifetimes will be IFR if and only if the two lifetime distributions are identical. In all other circumstances, the failure rate of the system will be initially strictly increasing but eventually strictly decreasing. This finding motivated Birnbaum, Esary and Marshall [8] to search for a nonparametric class of distributions containing the IFR class that was in fact closed under the formation of coherent systems. The outcome of their search was the important discovery of the IFRA (increasing failure rate average) class, that is, the class of distributions for which $R(t)/t$ is increasing in t. The IFRA class was shown to be the smallest class of distributions which contains the exponential distributions and is closed under the formation of coherent systems and taking limits in distribution. However, with additional restrictions, certain IFR closure properties do hold for coherent systems in IFR components. Following Samaniego [61], we discuss below what can be said when component lifetimes are i.i.d. from an IFR distribution.

In the i.i.d. case, the representations given in (3.9) - (3.11) for a system's failure rate in terms of the system's signature and the common component lifetime distribution F and failure rate r, are valid for any absolutely continuous distribution F. As mentioned above, an IFR distribution is absolutely continuous in the interior of its support set. Thus, the representation in (3.11) applies for all $t < \inf\{z \mid F(z) = 1\}$. For an n-component mixed system based on components with i.i.d. lifetimes $\sim F$, the system failure rate can be written as

$$r_T(t) = \frac{\sum_{i=0}^{n-1}(n-i)s_{i+1}\binom{n}{i}(G(t))^i}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_j\right)\binom{n}{i}(G(t))^i} r(t) , \tag{4.10}$$

where $G(t) = F(t)/\overline{F}(t)$ represents the odds of failure versus survival. Since $G(t)$ is an increasing function of t, it is apparent from (4.10) that, if the component lifetimes of a mixed system are i.i.d. from an IFR distribution F, then the system is IFR whenever the rational function h, given by

$$h(x) = \frac{\sum_{i=0}^{n-1}(n-i)s_{i+1}\binom{n}{i}x^i}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_j\right)\binom{n}{i}x^i} , \tag{4.11}$$

is an increasing function of x. This fact may be utilized to provide an elementary proof of the following result.

Theorem 4.1. *Consider a k-out-of-n system* $\tau_{k:n}$. *If the component lifetimes of* $\tau_{k:n}$ *are i.i.d. according to an IFR distribution* F*, then the distribution of the system lifetime is IFR.*

Proof. A proof of this result using properties of convex ordering may be found in Barlow and Proschan [6]. A simple, direct proof using the signatures of the systems involved proceeds as follows. The signature vector of the system $\tau_{k:n}$ is the n-dimensional probability vector $(0, \ldots, 0, 1, 0, \ldots, 0)$, with a "1" as its kth element. Thus, the function h in (4.11) takes the particularly simple form

$$h_0(x) = \frac{(n-k+1)\binom{n}{k-1}x^{k-1}}{\sum_{i=0}^{k-1}\binom{n}{i}x^i} \,. \tag{4.12}$$

The derivative of the function h_0 above is easily shown to be positive for all $x > 0$. The fact that h_0 is increasing on $(0, \infty)$ implies that the lifetime distribution of $\tau_{k:n}$ is IFR. ■

While the result above has been in the reliability literature for some time, the more general problem of characterizing the class of all systems that enjoy this type of closure is not easy to investigate using traditional methods. The fact that a broader class of systems in i.i.d. components are IFR when their components are, and also the fact that such closure does not hold for all coherent systems, are demonstrated in the following two examples.

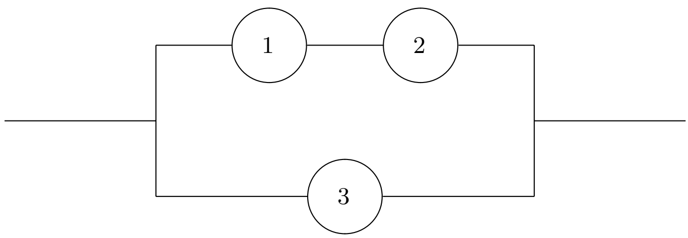
Example 4.1. Let us first consider the bridge system pictured in Figure 2.2 of Chapter 2. By elementary combinatorial calculations, the signature of this system may be shown to be $\mathbf{s} = (0, 1/5, 3/5, 1/5, 0)$. Using this $\mathbf{s}$ and $n = 5$ in (4.11), one obtains that the function h reduces to the following in this case:

$$h_1(x) = \frac{4x^3 + 18x^2 + 4x}{2x^3 + 8x^2 + 5x + 1} \,. \tag{4.13}$$

The derivative of h_1 above is positive for all $x > 0$, implying that the lifetime of the bridge structure will be IFR when its components have i.i.d. IFR lifetimes.

Let's now examine next the 3-component system pictured below.

Fig. 4.1. A parallel-series system in three components



Example 4.2. Let τ be the 3-component system pictured in Figure 4.1 above. Its signature is easily calculated to be $(0, 2/3, 1/3)$. We will assume here that the three components have lifetimes that are i.i.d. according to a particular IFR distribution, namely, the exponential distribution with common failure rate λ. From (4.10), we have that the failure rate of the system's lifetime T can be written as

$$r_T(t) = \lambda \, \frac{G^2(t) + 4G(t)}{G^2(t) + 3G(t) + 1} \;, \tag{4.14}$$

where $G(t) = \mathrm{e}^{\lambda t} - 1$. The corresponding function h of (4.11) is in this case

$$h_2(x) = \frac{x^2 + 4x}{x^2 + 3x + 1} \;. \tag{4.15}$$

Since the component failure rate is the constant λ, and the function h_2 is increasing in the interval $(0, 1 + \sqrt{5})$ and decreasing for $x > 1 + \sqrt{5}$, it follows that the system τ does not have an increasing failure rate.

The next result, proven by Samaniego [61] for coherent systems, characterizes the class of systems which are IFR when their components have i.i.d. IFR lifetimes. The characterization is based on the behavior of the function h of (4.11) and thus implicitly on the system signatures for which h has the requisite behavior.

Theorem 4.2. *Consider a mixed system of order n based on components having lifetimes that are i.i.d. IFR. The system's lifetime is IFR if and only if the rational function $h(x)$ in (4.11) is increasing for $x > 0$.*

Proof. As has been noted, the fact that the system lifetime is IFR when the function h is increasing is directly implied by the failure rate representation in (4.10). Now, suppose that the function h in (4.11) is not increasing over the entire positive real line. The continuity of h, as a ratio of polynomials with non-negative coefficients, ensures that there is an interval $(a, b) \subset (0, \infty)$ in which h is decreasing. If the n components of the system have lifetimes that are i.i.d. according to an exponential distribution with mean 1, the failure rate of the system's lifetime T may be written as

$$r_T(t) = h(\mathrm{e}^t - 1) \;. \tag{4.16}$$

From this, we conclude that $r_T(t)$ is decreasing in the interval

$$(\ln(1 + a),\ \ln(1 + b)) \;,$$

confirming that the distribution of T is not in the IFR class. ∎

As the preceding theorem indicates, the rational function h in (4.11) captures the precise information concerning the signature of a system that is required to characterize the class of systems in i.i.d. components that have IFR lifetime distribution when the components do. While checking that h is increasing may seem to be an imposing problem when the system involved is complex, it is worth noting that it is equivalent to the problem of verifying that a polynomial of degree no greater than $2n-3$ is non-negative on the positive real line. This is a problem which can be attacked numerically without great difficulty.

4.2 Preservation Theorems Based on Signature Properties

In this section, we consider the problem of comparing the performance of two mixed systems. Three different scenarios will be treated, each identifying conditions which yield increasingly stronger conclusions about the superiority of one system over another. Our treatment focuses on three common formulations of the notion that a random variable X_1 is smaller than a random variable X_2 in some stochastic sense. The versions of such orderings to be utilized below are defined as follows. These definitions apply equally when the random pairs (X_1, X_2) involved are discrete or continuous.

Definition 4.2. *The random variables X_1 and X_2 are* stochastically ordered *(denoted by $X_1 \leq_{st} X_2$) if and only if their respective survival functions are suitably ordered, that is, if and only if $\overline{F}_1(x) \leq \overline{F}_2(x)$ for all x.*

Definition 4.3. *The random variable X_1 is smaller than the random variable X_2 in the hazard rate (or uniform stochastic) ordering if and only if the ratio of survival functions $\overline{F}_2(x)/\overline{F}_1(x)$ is increasing in x. This ordering will be denoted by $X_1 \leq_{hr} X_2$.*

Definition 4.4. *The random variable X_1 is smaller than the random variable X_2 in the likelihood ratio ordering (denoted by $X_1 \leq_{lr} X_2$) if and only if the ratio $f_2(x)/f_1(x)$ is increasing in x, where f_i represents the density or probability mass function of X_i.*

For all three orderings above, we will use the statements "$X_1 \leq X_2$" and "$F_1 \leq F_2$" interchangeably. Stochastic ordering occurs when one survival function dominates another uniformly; in the boundary case in which $\overline{F}_1(x) \equiv \overline{F}_2(x)$, each distribution can be said to be "stochastically larger" than the other, and the latter condition can in fact be taken as a definition of equality of the two distributions. For discrete distributions (e.g., a pair of signatures $\mathbf{s}_1$ and $\mathbf{s}_2$), the stochastic ordering condition $\overline{F}_1(x) \leq \overline{F}_2(x)$ for all x reduces to $\sum_{i=j}^{n} s_{1i} \leq \sum_{i=j}^{n} s_{2i}$ for $j = 1, \ldots, n$. When the underlying distributions are absolutely continuous, hazard rate ordering is equivalent to the ordering of the failure rates, with $X_1 \leq_{\text{hr}} X_2$ if and only if $r_1(t) \geq r_2(t)$ for all t. It is well known that likelihood ratio ordering is the most stringent of these orderings; indeed, lr $\Rightarrow$ hr $\Rightarrow$ st. It is also well known that $X_1 \leq_{\text{st}} X_2 \Rightarrow EX_1 \leq EX_2$.

Preservation theorems are of considerable interest in reliability theory. In essence, such theorems indicate that a property of the basic elements of a system of interest (e.g., its components or its design) will be inherited by the system itself. Since properties of a complex system are generally more difficult to ascertain than properties of its basic elements, preservation theorems, which are vehicles for identifying system properties by inspection of more basic elements of the system, can be very useful. In section 4.1, we proved a theorem involving the IFR property, giving conditions under which the system lifetime has an IFR distribution when its components have IFR distributions. Results which show that a system inherits properties held by its components are generally referred to as "closure theorems." Here, we will be interested in whether ordering properties of the signature vectors of two systems are preserved by the lifetimes of the systems corresponding to these signatures. The three preservation results presented below were proven for coherent systems by Kochar, Mukerjee and Samaniego [51]. They are stated and proved here for general mixed systems. The first result examines the consequences of the stochastic ordering of two signatures.

Theorem 4.3. *Let* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *be the signatures of the two mixed systems of order* n*, both based on components with i.i.d. lifetimes with common distribution* F*. Let* T_1 *and* T_2 *be their respective lifetimes. If* $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$*, then* $T_1 \leq_{st} T_2$*.*

Proof. From the representation in (3.2) of a system's survival function, we have that, for all non-negative t,

$$\begin{aligned}\overline{F}_1(t) &= \sum_{j=0}^{n-1}\left(\sum_{i=j+1}^{n} s_{1i}\right)\binom{n}{j}(F(t))^j\left(\overline{F}(t)\right)^{n-j}\\ &\leq \sum_{j=0}^{n-1}\left(\sum_{i=j+1}^{n} s_{2i}\right)\binom{n}{j}(F(t))^j\left(\overline{F}(t)\right)^{n-j}\\ &= \overline{F}_2(t)\ ,\end{aligned}$$

the inequality above being directly implied by the assumption $\mathbf{s}_1 \leq_{\text{st}} \mathbf{s}_2$. ■

Note that the five coherent systems of order three are totally ordered in the sense of the theorem above. The twenty coherent systems of order four cannot be totally ordered in this way. For example, the signatures of systems 5 and 7 in Table 3.2 are not stochastically ordered, so that the comparison of these systems' lifetimes via Theorem 4.3 is not possible. However, of the 190 possible pair-wise comparisons of coherent systems of order 4, the stochastic ordering of system signatures obtains in 180 cases.

We next examine the implications of the hazard rate ordering of two signatures. Following the line of argument in Kochar, Mukerjee and Samaniego [51], we show below that such an ordering between the signatures of two mixed systems implies the corresponding system lifetimes obey the hazard rate ordering. Establishing the desired result will require use of the following lemma proven in Joag-dev, Kochar and Proschan [46].

Lemma 4.1. *Let $\alpha(\cdot)$ and $\beta(\cdot)$ be real valued functions such that $\beta(\cdot)$ is nonnegative and $\alpha(\cdot)/\beta(\cdot)$ and $\beta(\cdot)$ are nondecreasing. If $X_1 \leq_{hr} X_2$, where $X_1 \sim H_1$ and $X_2 \sim H_2$, then*

$$\frac{\int_{-\infty}^{\infty} \alpha(x)\,\mathrm{d}H_1(x)}{\int_{-\infty}^{\infty} \beta(x)\,\mathrm{d}H_1(x)} \leq \frac{\int_{-\infty}^{\infty} \alpha(x)\,\mathrm{d}H_2(x)}{\int_{-\infty}^{\infty} \beta(x)\,\mathrm{d}H_2(x)} \tag{4.17}$$

Theorem 4.4. *Let $\mathbf{s}_1$ and $\mathbf{s}_2$ be the signatures of the two mixed systems of order n, both based on components with i.i.d. lifetimes with common distribution F. Let T_1 and T_2 be their respective lifetimes. If $\mathbf{s}_1 \leq_{hr} \mathbf{s}_2$, then $T_1 \leq_{hr} T_2$.*

Proof. Recall from (3.5) that, for $j = 1$ and 2, the survival function $\overline{F}_j$ of T_j may be written as

$$\overline{F}_j(t) = \sum_{i=1}^{n} s_{ji} P(X_{i:n} > t)\ .$$

We assume that $\mathbf{s}_1 \leq_{\mathrm{hr}} \mathbf{s}_2$. We will prove that $T_1 \leq_{\mathrm{hr}} T_2$ by showing that $\overline{F}_2(t)\ /\ \overline{F}_1(t)$ is increasing in t. This is equivalent to showing that

$$\frac{\overline{F}_2(x)}{\overline{F}_1(x)} \leq \frac{\overline{F}_2(y)}{\overline{F}_1(y)} \text{ for any } x \leq y\ , \tag{4.18}$$

an inequality that may be rewritten as

$$\frac{\sum_{i=1}^{n} s_{1i} P(X_{i:n} > y)}{\sum_{i=1}^{n} s_{1i} P(X_{i:n} > x)} \leq \frac{\sum_{i=1}^{n} s_{2i} P(X_{i:n} > y)}{\sum_{i=1}^{n} s_{2i} P(X_{i:n} > x)} \text{ for all } x < y\ . \tag{4.19}$$

But (4.19) can be shown to follow directly from (4.17) in Lemma 4.1 by taking α and β to be the discrete functions $\alpha(i) = P(X_{i:n} > y)$ and $\beta(i) = P(X_{i:n} > x)$ and by taking H_1 and H_2 to be the discrete distributions $\mathbf{s}_1$ and $\mathbf{s}_2$, respectively. We need only to verify that the chosen functions α and β satisfy the hypotheses of Lemma 4.1. The required monotonicity of β follows from the fact that successive order statistics are stochastically ordered. The inequality

$$\frac{\alpha(i)}{\beta(i)} \leq \frac{\alpha(i+1)}{\beta(i+1)}$$

can be rewritten as

$$\frac{P(X_{i+1:n} > x)}{P(X_{i:n} > x)} \leq \frac{P(X_{i+1:n} > y)}{P(X_{i:n} > y)} \text{ for all } x < y ,$$

an inequality that is equivalent to $X_{i:n} \leq_{\text{hr}} X_{i+1:n}$, a well known property of independently drawn order statistics (see, for example, Boland, El-Newehi and Proschan [18]). ■

The following result establishes that likelihood ratio ordering between system signatures will imply the same ordering between system lifetimes.

Theorem 4.5. *Let* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *be the signatures of the two mixed systems of order* n*, both based on components with i.i.d. lifetimes with common distribution* F*. Let* T_1 *and* T_2 *be their respective lifetimes. If* $\mathbf{s}_1 \leq_{lr} \mathbf{s}_2$*, then* $T_1 \leq_{lr} T_2$*.*

Proof. This result is proven in Kochar, Mukerjee and Samaniego [51] using the variation diminishing properties of totally positive functions of order 2. Here, we will present an elementary proof. Let f_1 and f_2 be the density functions of T_1 and T_2 respectively. We must show that the ratio $f_2(t) \,/\, f_1(t)$ is increasing in t. Using the density representation in (3.8) and dividing the numerator and denominator of this ratio by $f(t)\left(\overline{F}(t)\right)^{n-1}$, we may write

$$\frac{f_2(t)}{f_1(t)} = \frac{\sum_{i=1}^{n} i\ s_{2i}\binom{n}{i}\left(G(t)\right)^{i-1}}{\sum_{i=1}^{n} i\ s_{1i}\binom{n}{i}\left(G(t)\right)^{i-1}} , \tag{4.20}$$

where $G(t) = F(t) \,/\, \overline{F}(t)$. A necessary and sufficient condition for the ratio in (4.20) to be increasing on $(0, \infty)$ is that, for any real number c, the difference $f_2(t) - cf_1(t)$ changes signs at most once, with this change, if it occurs, being from negative to positive, as t goes from 0 to ∞. (Clearly, this condition is satisfied if the ratio in (4.20) is increasing, while if the difference crosses 0 more than once, it would necessarily be decreasing in some interval.) Letting $x = G(t)$, we wish to study the crossing properties of the function

$$t(x) = f_2\left(G^{-1}(x)\right) - cf_1\left(G^{-1}(x)\right) \tag{4.21}$$

$$= \sum_{i=1}^{n} i\binom{n}{i}(s_{2,i} - cs_{1,i})x^{i-1} \tag{4.22}$$

which, of course, is simply a polynomial of degree at most $n - 1$. From our assumption that $\mathbf{s}_1 \leq_{\text{lr}} \mathbf{s}_2$, we have that the ratios $s_{2,i} \,/\, s_{1,i}$ are increasing as i goes from 1 to n. This implies that, for any real number c, the sequence $\{s_{2,i} - cs_{1,i}\}$ has at most one change of sign, from negative to positive, as i goes from 1 to n. We may thus surmise that the coefficients of the polynomial in (4.22) have at most one sign change. Now, Descartes' "Rule of Signs" states that the number of positive roots of a polynomial of arbitrary degree is no greater than the number of sign changes in its sequence of coefficients. We conclude that, for any real number c, the polynomial in (4.22) will cross zero at most once as x goes from 0 to ∞. The fact that the coefficients in (4.22) can only change signs from negative to positive implies that $t(x)$ itself, if it undergoes a sign change, will change from negative to positive. Together, these facts justify the conclusion that the ratio $f_2(t)$ / $f_1(t)$ is increasing for $t \in [0, \infty)$, that is, that $T_1 \leq_{\text{lr}} T_2$. ■

The results above are by no means an exhaustive list of the ways in which the results on signatures in Chapter 3 can be used to infer properties of systems from information available about system designs or the properties of a system's components. The preservation theorems above represent one natural question one could ask: when all components have i.i.d. lifetimes, does one system perform better than another when its design (that is, its signature) is better in some specific sense? The results above show that this tends to be the case. Another natural question to explore is: how do changes in the underlying distribution of the components affect the performance of a given system? This latter question deals with a single system signature but with two sets of independent components, each set having its own lifetime distribution. The following result provides an answer to this question in one particular setting and serves as an example of how the properties of signatures may be used in examining questions of this type.

Theorem 4.6. *Consider a mixed system of order n based on components with i.i.d. lifetimes. Let T_1 be the lifetime of the system if its components have lifetime distribution F_1, and let T_2 be the lifetime of the system if its components have lifetime distribution F_2. If $F_1 \leq_{st} F_2$, then $T_1 \leq_{st} T_2$.*

Proof. Let $X_1, X_2, \ldots, X_n$ be i.i.d. lifetimes with distribution F_1, and let $Y_1, Y_2, \ldots, Y_n$ be i.i.d. lifetimes with distribution F_2. From (3.5), we have that, for $i = 1$ or 2, the survival function of T_i may be written as

$$P(T_1 > t) = \sum_{i=1}^{n} s_i P(X_{i:n} > t) \tag{4.23}$$

and

$$P(T_2 > t) = \sum_{i=1}^{n} s_i P(Y_{i:n} > t) \tag{4.24}$$

respectively. Let X and Y be independent binomial variables, with $X \sim B\left(n, \overline{F}_1(t)\right)$ and $Y \sim B\left(n, \overline{F}_2(t)\right)$. We may write $P(X_{i:n} > t) = P(X > n-i)$ and $P(Y_{i:n} > t) = P(Y > n-i)$. Thus, the fact that $\overline{F}_2(t) \geq \overline{F}_1(t)$ for all $t > 0$ implies that

$$P(Y_{i:n} > t) = P(Y > n-i) \geq P(X > n-i) = P(X_{i:n} > t) \text{ for all } t\,. \tag{4.25}$$

Using (4.25), we may conclude from (4.23) and (4.24) that $P(T_2 > t) \geq P(T_1 > t)$ for all t, which is the desired conclusion. ■

Up to this point, the comparative results discussed in this section have been restricted to two systems of the same size. It has been shown that the ordering properties of the system signatures imply similar orderings between the corresponding system lifetimes. Now let us suppose that we are interested in comparing two systems that are not of the same order. Recall that in section 4.3, we showed that for any system of a given order, there exists an equivalent system of any fixed larger size. Indeed, for $n < m$, Theorem 3.2 provides an explicit representation which can be used $m - n$ times in succession to identify the signature of the m-component system that is equivalent to that of any fixed n-component system. The following example illustrates the use of such a result in comparing systems of different orders.

Example 4.3. Assume components with i.i.d. lifetimes ($\sim F$) for all systems under discussion here. Let us consider a 2-component mixed system (based on i.i.d. components) with signature vector (1/2, 1/2), a system that can be realized as a 50-50 mixture of the series and parallel systems in two components. A design engineer might legitimately ask whether a particular 4-component coherent system to which he or she has access performs better or worse than the 2-component system above. Using Theorem 3.2, we find that the 3- and 4-component systems with signature vectors (1/3, 1/3, 1/3) and (1/4, 1/4, 1/4, 1/4), respectively, are stochastically equivalent to the original system. Now, using results from Theorems 4.1-4.3, one can compare the four dimensional signature vectors directly and reach the appropriate conclusion. From Table 3.2, we see that there are precisely five 4-component coherent systems whose lifetimes are stochastically larger than the 2-component system above (namely, systems 6, 16, 17, 18 and 20), and also precisely 5 with stochastically smaller lifetimes (namely, 1, 2, 3, 4 and 5). The lifetimes of each of the remaining 10 coherent systems of order 4 have survival functions that cross that of the 2-component system, so that neither is uniformly superior to the other. Using this knowledge, the engineer can make a well-informed decision about the choice alluded to above. In section 4.4, we will take a closer look at the relative behavior of pairs of systems that are non-comparable via the stochastic ordering of lifetimes. This further discussion of system comparisons might well be relevant to the engineering decision alluded to above, as a system that is better than another up to the system's intended mission time would be preferable, whether or not it has uniformly superior reliability at any arbitrary time t.

4.3 An Application to Redundancy Comparisons

In Chapter 2, we presented a brief discussion of the engineering principle that redundancy at the system level (that is, backing up a system with a second, identical system) is less effective than redundancy at the component level (that is, backing up every component with an identical component). This principle can be rigorously established by comparing the structure functions φ_1 and φ_2 of the two augmented systems. The fact that $\varphi_1 \leq \varphi_2$ was established in Theorem 2.1. When the components of both systems have lifetimes that are i.i.d. according to the common distribution F, it follows from this theorem that the two system signatures are stochastically ordered, that is, $\mathbf{s}_1 \leq_{\mathrm{st}} \mathbf{s}_2$. This of course implies, by Theorem 4.3, that the system's lifetimes T_1 and T_2 are similarly ordered. Boland and El-Newehi [17] showed that stronger inequalities failed to hold for systemwise and componentwise redundancy in general but conjectured that they might hold for the class of k-out-of-n systems in i.i.d. components. Boland and El-Newehi's conjecture that $T_1 \leq_{\mathrm{lr}} T_2$ holds for such systems was established by Singh and Singh [68] using a technical lemma which developed some new, complex inequalties. Kochar, Mukerjee and Samaniego [51] provided an alternative proof based on signatures. Specifically, Theorem 4.5 will imply the desired ordering of the two system lifetimes. The results in the latter paper were proven for systems which survive as long as they have k working components. The results are rederived below for k-out-of-n systems as defined in this monograph, namely for systems which fail upon the kth component failure. First, the signatures of k-out-of-n systems with systemwise and componentwise redundancy are derived. I will take a different and perhaps more natural approach to these derivations, using arguments based on "combinations" rather than on "permutations" as utilized in the referenced paper. This renders the proofs more direct and more intuitive. The redundant systems based on a 2-out-of-3 system are pictured in Figures 4.2 and 4.3 below.

Fig. 4.2. Systemwise redundancy for a 2-out-of-3 system

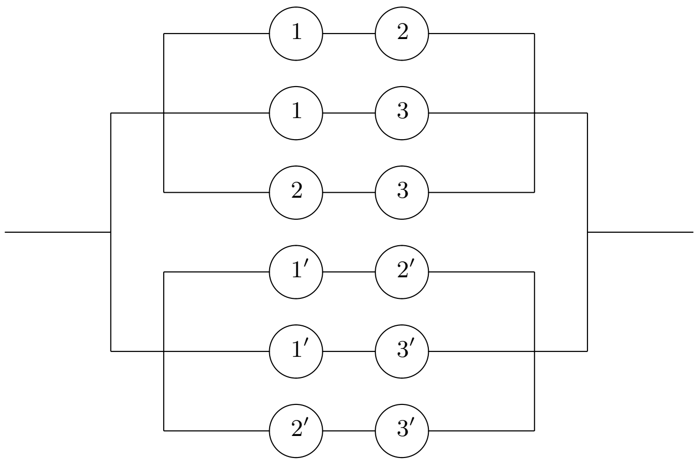
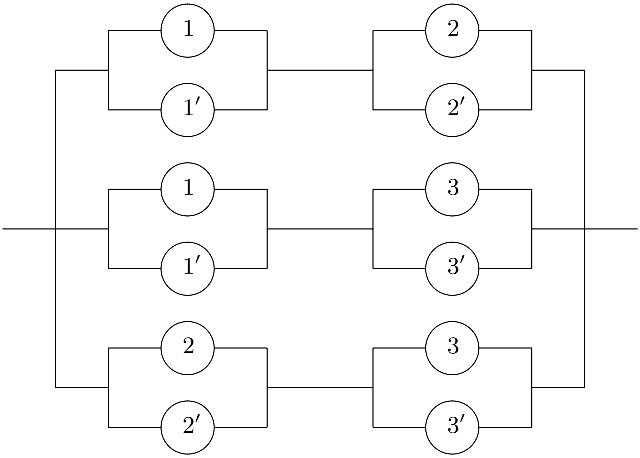


Fig. 4.3. Componentwise redundancy for a 2-out-of-3 system



Theorem 4.7. *The signature* $\mathbf{s}$ *of a* k*-out-of-*n *system in i.i.d. components with systemwise redundancy has the elements*

$$s_{2k+r} = \frac{\binom{n-1}{k-1}\binom{n}{k+r}}{\binom{2n-1}{2k+r-1}} \quad \text{for } r = 0, \ldots, n-k \tag{4.26}$$

with $s_i = 0$ *for* $1 \leq i < 2k$ *and for* $n + k < i \leq 2n$*.*

Proof. We first establish the fact that the elements of $\mathbf{s}$ are equal to zero for both sufficiently small and sufficiently large i, as stated in the theorem. Let i represent the number of failed components. If $i < 2k$, it is not possible for both the original k-out-of-n system and its backup system to have k failed components, and thus one of the two must be working, which implies that the overall system is working. If $i > n + k$, then both the original and the backup systems have at least $k + 1$ failed components (the two extreme cases being that all components in one of these systems have failed and $k + 1$ components of the other system have failed, and vice versa). It follows that the overall system necessarily failed upon the prior failure of a component. Thus, $s_i = 0$ if either $i < 2k$ or $i > n + k$.

To assist us in making the combinatorial argument leading to (4.26), let us think of the component lifetimes of the original system as the Xs and the component lifetimes of the backup system as the Ys. Let $r \in \{0, \ldots, n - k\}$. To obtain the likelihood of the system failing upon the $(2k + r)$th component failure, we will condition on the fact that a particular X or a particular Y is the time at which the $(2k+r)$th component fails. Since the conditional probability obtained turns out to be the same for any component lifetime X or Y, it is also the unconditional probability that the system fails upon the $(2k+r)$th component failure, that is, it is the probability in (4.26). Suppose that a particular component lifetime X_i corresponds to the $(2k+r)$th component failure. For the system to fail upon this component failure, one would need to have exactly $k-1$ other failures in the original system (i.e., X failures) before the failure time X_i. This can happen in $\binom{n-1}{k-1}$ ways. There would also need to be precisely $k+r$ backup component failures (i.e., Y failures) preceding the failure time X_i. This can happen in $\binom{n}{k+r}$ ways. The product of these two numbers is the total number of ways in which components could fail prior to the failure time X_i, the $(2k+r)$th component failure time and the time at which the system fails. On the other hand, the total number of ways in which $2k+r-1$ components could fail prior to X_i is $\binom{2n-1}{2k+r-1}$, a number which serves as the denominator of the desired probability. It follows that the probability of system failure at time X_i is simply the ratio $\binom{n-1}{k-1}\binom{n}{k+r} / \binom{2n-1}{2k+r-1}$. Since this probability is independent of X_i, and in fact, is the same for any X or Y which kills the system and occurs as the $(2k+r)$th component failure time, it is equal to the unconditional probability of system failure upon the $(2k+r)$th component failure. ■

Theorem 4.8. *The signature* **s** *of a k-out-of-n system in i.i.d. components with componentwise redundancy has the elements*

$$s_{2k+r} = \frac{\binom{n-1}{k-1}\binom{n-k}{r}}{\binom{2n-1}{2k+r-1}} \times 2^r \quad \text{for } r = 0, \ldots, n-k \tag{4.27}$$

with $s_i = 0$ *for* $1 \le i < 2k$ *and for* $n+k < i \le 2n$.

Proof. We will find it helpful to view the system of interest here as being composed of n modules, each being a parallel system containing two components – the original component j and its identical backup j'. Let us first deal with the elements of the signature of the system that are claimed to be equal to zero. Let i represent the number of failed components. If $i < 2k$, it is clearly impossible for each of the $2k$ components in a given collection of k modules to have failed, which means that every collection of k modules has at least one working component. This ensures that there is an operative path set and thus that the system is working. We conclude that $s_i = 0$ when $i < 2k$. Now suppose that $i > n+k$. In this case, there are, for some j in the set $\{0, 1, \ldots, n-k-1\}$, $n-j$ failed original components and at least $k+j+1$ failed backup components. Since only j original components are working, at least $(k+j+1) - j = k+1$ failed backup components served as backups for failed original components. This implies that there exists more than one collection of k modules in which all $2k$ components have failed. From this, we surmise that the system would have failed upon a prior component failure. It follows that $s_i = 0$ when $i > n+k$.

As in the proof of the preceding theorem, let us think of the component lifetimes of the original system as the Xs and those of the backup components as the Ys. Let $r \in \{0, \ldots, n-k\}$. To obtain the likelihood of the system failing upon the $(2k+r)$th component failure, we will condition on the fact that the failure time X of a particular original component or Y of a particular backup component causes the system to fail upon the $(2k+r)$th component failure. Since the conditional probability obtained turns out to be the same for any X or Y component, it is also the unconditional probability, that is, it is the probability we seek. Suppose that a particular component lifetime X_i corresponds to the $(2k+r)$th component failure. For the system to fail upon this component failure, one would need to have exactly $k-1$ failed modules preceding time X_i. This can happen in $\binom{n-1}{k-1}$ ways. We may think of this as the occurrence of $k-1$ "doubletons," that is, failed pairs of original and backup components, among the failures preceding time X_i. The remaining failures must be "singletons," since the occurrence of a kth doubleton in the course of the first $2k+r-1$ failures would cause the system to fail before X_i. One of these singletons must be the failure time Y_i, for otherwise, the failure time X_i will not correspond to the failure time of the system. The remaining r singletons must be chosen from the $n-k$ modules yet unaccounted for. These singletons can be chosen in $\binom{n-k}{r} \times 2^r$ ways, the factor 2^r accounting for that fact that each of these singletons can be either an X or a Y component. The failed components preceding the failure time X_i will total $2k+r-1$ as required, $(2k-2)$ from the failed modules, 1 from the failed singleton at time Y_i and r from the additional singletons. The product of the two counts above is precisely the total number of ways in which components could fail prior to the time X_i, the failure time of the $(2k+r)$th component and the time at which the system fails. On the other hand, the total number of ways in which $2k+r-1$ components could fail prior to time X_i as the $(2k+r)$th component failure is $\binom{2n-1}{2k+r-1}$. It follows that the probability of system failure at time X_i is simply the ratio $\binom{n-1}{k-1}\binom{n-k}{r} \times 2^r \;/\; \binom{2n-1}{2k+r-1}$. Since this probability is independent of X_i, and is in fact the same for any X or Y which kills the system and occurs as the $(2k+r)$th component failure time, it is equal to the unconditional probability of system failure upon the $(2k+r)$th component failure. ■

We are now in a position to establish a strong form of domination of componentwise redundancy over systemwise redundancy.

Theorem 4.9. *Let T_1 be the lifetime of a k-out-of-n system with i.i.d. components under systemwise redundancy and let T_2 be the lifetime of the corresponding system under componentwise redundancy. Then $T_1 \leq_{\mathrm{lr}} T_2$.*

Proof. Let $\mathbf{s}_1$ and $\mathbf{s}_2$ be the signatures of the systems above. From Theorems 4.7 and 4.8, we have that, for $r = 0, \ldots, n-k$, the likelihood ratio of the elements of these signatures is

$$V(r) \equiv \frac{s_{2,2k+r}}{s_{1,2k+r}} = \frac{\binom{n-k}{r} \times 2^r}{\binom{n}{k+r}} = \frac{(n-k)!(k+r)! \times 2^r}{n!r!} \tag{4.28}$$

We can verify that $V(r)$ is increasing in r by noting that for $r = 0, \ldots, n-k-1$,

$$\frac{V(r+1)}{V(r)} = \frac{2(k+r+1)}{r+1} > 1 \; . \tag{4.29}$$

This establishes that $\mathbf{s}_1 \leq_{\text{lr}} \mathbf{s}_2$ which, by Theorem 4.5, implies that $T_1 \leq_{\text{lr}} T_2$, as desired. ■

4.4 Signature-Based Characterizations of Relative System Performance

The various ordering conditions on signatures of Section 4.2 are sufficient to imply corresponding orderings of the system lifetimes. They are not, however, necessary. Block, Dugas and Samaniego [11] provide counterexamples for the necessity of these conditions, and they also provide new necessary and sufficient conditions (NASCs) which ensure ordering of two system lifetimes in each of the three senses considered in Section 4.2. These results are the subject of the present section. The results below are based on arguments found in Block, Dugas and Samaniego [10]. We first consider an NASC for the stochastic ordering of system lifetimes.

Theorem 4.10. *Let* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *be the signatures of two arbitrary mixed systems based on coherent systems in* n *i.i.d. component lifetimes, and let* T_1 *and* T_2 *denote the system lifetimes. Then* $T_1 \leq_{st} T_2$ *if and only if* $g(x) \geq 0$ *for all* $x \geq 0$*, where* $g(x)$ *is the polynomial given by*

$$g(x) = \sum_{j=0}^{n-1} \binom{n}{j} \left(\sum_{i=j+1}^{n} (s_{2i} - s_{1i}) \right) x^j \; . \tag{4.30}$$

Proof. Suppose that two mixed systems are based on components with lifetimes that are i.i.d. according to F, and suppose that $T_1 \leq_{\text{st}} T_2$. Writing this condition as

$$\overline{F}_1(t) \leq \overline{F}_2(t) \quad \text{for all } t > 0 \; , \tag{4.31}$$

where $\overline{F}_1$ and $\overline{F}_2$ are the respective survival functions for T_1 and T_2, we see that, in light of (3.2), the stochastic inequality $T_1 \leq_{\text{st}} T_2$ is equivalent to

$$\sum_{i=1}^{n} s_{1i} \sum_{j=0}^{i-1} \binom{n}{j} (G(t))^j \leq \sum_{i=1}^{n} s_{2i} \sum_{j=0}^{i-1} \binom{n}{j} (G(t))^j \, , \quad \text{for all } t > 0 \; , \tag{4.32}$$

where $G(t) = F(t) \,/\, \overline{F}(t)$ and $\mathbf{s}_1$ and $\mathbf{s}_2$ are the respective system signatures. The inequality (4.32) is algebraically equivalent to

$$\sum_{i=1}^{n} (s_{2i} - s_{1i}) \sum_{j=0}^{i-1} \binom{n}{j} (G(t))^j \geq 0 \quad \text{for all } t > 0 \,, \tag{4.33}$$

or, upon interchanging the order of summation, is equivalent to

$$\sum_{j=0}^{n-1} \binom{n}{j} \left(\sum_{i=j+1}^{n} (s_{2i} - s_{1i}) \right) (G(t))^j \geq 0 \quad \text{for all } t > 0 \,. \tag{4.34}$$

Now, define the function g to be the polynomial of degree $(n-1)$ displayed in (4.30). From the equivalence of the inequalities (4.31) - (4.34), and the fact that $G(t)$ is an increasing function of t mapping $(0, \infty)$ onto $(0, \infty)$, it follows that the stochastic ordering in (4.31) holds if and only if $g(x) \geq 0$ for all $x \geq 0$. ■

Remark 4.1. First, it should be noted that, if two system signatures are stochastically ordered, i.e. $\mathbf{s}_1 \leq_{\text{st}} \mathbf{s}_2$, then (4.30) clearly holds, since g is then a polynomial with nonnegative coefficients. Thus, Theorem 4.10 contains Theorem 4.3 as a special case. It also extends the earlier result, since condition (4.30) is not only sufficient but also necessary for the stochastic ordering of T_1 and T_2 to hold. Just as it is possible for a polynomial $p(x)$ having some negative coefficients to be positive for all $x > 0$, it is clear that the restrictive condition $\mathbf{s}_1 \leq_{\text{st}} \mathbf{s}_2$ is not required for the polynomial in (4.30) to be positive for all positive x. Secondly, it is worth noting that, although condition (4.30) is a complex statement concerning the relationship between the two system signatures involved, no essential simplification is possible, as the condition is both necessary and sufficient. In practice, condition (4.30) can be checked using standard numerical methods.

Let us now consider an NASC for the hazard-rate ordering between system lifetimes. For simplicity, we will state and prove the following result under the assumption that the underlying component distribution F is absolutely continuous.

Theorem 4.11. *Let $\mathbf{s}_1$ and $\mathbf{s}_2$ be the signatures of two arbitrary mixed systems based on coherent systems in n i.i.d. components and the same component distribution F with density f and failure rate r, and let T_1 and T_2 be the respective system lifetimes. Then $T_1 \leq_{hr} T_2$ if and only if*

$$h_1(x) - h_2(x) \geq 0 \quad \text{for all } x \geq 0 \,, \tag{4.35}$$

where h_j represents the rational function of x given by

$$h(x) = \frac{\sum_{i=0}^{n-1}(n-i)s_{i+1}\binom{n}{i}x^i}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_j\right)\binom{n}{i}x^i} \tag{4.36}$$

with $\mathbf{s} = \mathbf{s}_l$ *for* $l = 1$ *or* 2.

Proof. From (3.11), we may infer that the lifetimes T_1 and T_2 of two mixed systems based on coherent systems in n i.i.d. components, both having the same component distribution F with density f and failure rate r, satisfy $T_1 \leq_{\rm hr} T_2$ if and only if for all $t > 0$,

$$\frac{\sum_{i=0}^{n-1}(n-i)s_{2,i+1}\binom{n}{i}\left(G(t)\right)^i}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_{2,j}\right)\binom{n}{i}\left(G(t)\right)^i} \leq \frac{\sum_{i=0}^{n-1}(n-i)s_{1,i+1}\binom{n}{i}\left(G(t)\right)^i}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_{1,j}\right)\binom{n}{i}\left(G(t)\right)^i} \tag{4.37}$$

where $G(t) = F(t) \,/\, \overline{F}(t)$ and $\mathbf{s}_1$ and $\mathbf{s}_2$ are the respective system signatures. Now let h be the rational function displayed in (4.36), where $\mathbf{s}$ is an arbitrary n-dimensional probability vector (or system signature in the setting of interest here). It is clear that, upon making the transformation $x = G(t)$ in (4.37), the inequality reduces to (4.35), showing that the latter inequality is necessary and sufficient for the ordering of the failure rates of two systems under consideration. ■

Remark 4.2. Note that the NASC in (4.35) for the hazard rate ordering of the lifetimes T_1 and T_2 does not depend functionally on the density or failure rate of the common component distribution F. It can be shown that the condition is necessary and sufficient for $T_1 \leq_{\rm hr} T_2$ provided only that the common component lifetime distribution is continuous. The proof of this fact can be obtained by verifying that condition (4.35) is an NASC for the ratio of survival functions, that is, $\overline{F}_2(t)/\overline{F}_1(t)$, to be increasing in t. The fact that the derivative of this ratio is non-negative for all $t > 0$ reduces to (4.35). Condition (4.35) is, like (4.30), a complex but both necessary and sufficient condition on two signature vectors for the hazard rate ordering of system lifetimes. While not immediately transparent, it can be shown that $\mathbf{s}_1 \leq_{\rm hr} \mathbf{s}_2$ implies condition (4.35). Although condition (4.35) is mathematically complex, we note that, after cross multiplying in the inequality $h_1(x) \geq h_2(x)$, the condition reduces to checking that a certain polynomial of degree $2n-3$ is nonnegative for all $x \geq 0$. In a given problem of interest, this can be determined using standard numerical methods.

Finally, let us consider the case of likelihood ratio ordering between two system lifetimes.

Theorem 4.12. *Let* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *be the signatures of two arbitrary mixed systems based on coherent systems in* n *i.i.d. components with the same component lifetime distribution* F, *and let* T_1 *and* T_2 *be the respective system* *lifetimes. Then* $T_1 \leq_{lr} T_2$ *if and only if the rational function* $m_2(x) \,/\, m_1(x)$ *is increasing for all* $x \geq 0$*, where* $m_l(x)$ *is given by*

$$m_l(x) = \sum_{i=0}^{n-1} (n-i) s_{l,i+1} \binom{n}{i} x^i \;, \tag{4.38}$$

for $l = 1$ *or* 2*.*

Proof. One may derive from (3.8) that $T_1 \leq_{\mathrm{lr}} T_2$ if and only if the rational function

$$\frac{\sum_{i=0}^{n-1} (n-i) s_{2,i+1} \binom{n}{i} \left(G(t)\right)^i}{\sum_{i=0}^{n-1} (n-i) s_{1,i+1} \binom{n}{i} \left(G(t)\right)^i} \tag{4.39}$$

is increasing in $t > 0$. If we define the polynomials $m_1(x)$ and $m_2(x)$ as in (4.38), it is clear from (4.39), upon setting $G(t) = x$, that the condition that the ratio $m_2(x) \,/\, m_1(x)$ increases for all $x \geq 0$ is necessary and sufficient for $T_1 \leq_{\mathrm{lr}} T_2$ to hold. ■

Remark 4.3. The NASC in Theorem 4.12 is complex but, again, immutable, as no nontrivial simplification is possible. Determining the intervals over which the ratio $m_2(x) \,/\, m_1(x)$ is increasing or decreasing is equivalent to finding the roots of a polynomial of degree no greater than $2n - 3$ and would, in a given application, be addressed using numerical methods.

The tools utilized above, namely, the inspection of the behavior of the functions $g(x)$, $h(x)$ and $m(x)$ in Theorems 4.10 - 4.12, can lead to some additional useful insights. We know that if the function $g(x)$ in (4.30) is positive for all $x > 0$, the survival functions corresponding to the two system survival functions will not cross, thus implying their stochastic ordering. A closer look at the proof of Theorem 4.10 reveals that the positive roots of the function g coincide precisely with points at which the two survival functions cross. Thus, the roots of g completely determine the crossing properties of the survival functions $\overline{F}_1$ and $\overline{F}_2$. Similarly, the crossing points of two failure rates $r_1(t)$ and $r_2(t)$ are completely determined and identified by the values at which the functions $h_1(x)$ and $h_2(x)$ of Theorem 4.11 are equal. Thus, one can characterize the crossing behavior of pairs of survival functions or failure rates in terms of the behavior of the functions g and h of the respective signature vectors. This permits the comparison of survival functions and failure rates in real time (in contrast to the asymptotic comparisons featured in recent papers such as Block and Joe [12] and Block, Li and Savits [13]).

Similar results can be obtained regarding intervals in which the likelihood ratio $f_2(t) \,/\, f_1(t)$ is increasing or decreasing, as these behaviors correspond to intervals in which the function $m_2(x) \,/\, m_1(x)$ of Theorem 4.12 is increasing or decreasing. The three types of order relations studied above are illustrated in the following three examples. These examples are based on two systems in i.i.d. components for which both the survival functions and the failure rates of the systems cross exactly once and whose likelihood ratio changes monotonicity exactly once. In each case, the changes can be identified to occur at a specific quantile of the common component lifetime distribution F. In all three examples, the two systems to be compared are the same, namely, the 3-component systems having the signatures $\mathbf{s}_1 = (1/2, 0, 1/2)$ and $\mathbf{s}_2 = (0, 1, 0)$, respectively. The first system is a mixed system which results from randomly selecting a series or a parallel system, each with probability 1/2, while the second is simply a 2-out-of-3 system.

Example 4.4. To execute a comparison of the survival functions of the two systems above, we note that the polynomial g of (4.30) can be identified as

$$g(x) = -1.5x^2 + 1.5x \tag{4.40}$$

in this problem. The polynomial g has a unique positive root at $x = 1$. From Theorem 3.1, it follows that the two survival functions will cross at the time t_0 for which $G(t_0) = 1$, that is, at time $t_0 = F^{-1}(1/2)$. This leads to the conclusion that the 2-out-of-3 system is as good as or better than the mixed system if and only if $t \le t_0$. This would be a particularly important finding if the mission time for the chosen system happens to be smaller that t_0, as it would then serve to identify a system that is uniformly superior to the other in the time internal of interest. If the system's mission time is substantially larger than t_0, then one might well prefer the mixed system since it has superior performance when $t > t_0$.

Example 4.5. To compare the failure rates of these same two systems, we proceed by computing the relevant functions h_1 and h_2 in (4.36). For the systems in question, we obtain

$$h_1(x) = \frac{3 + 3x^2}{2 + 3x + 3x^2} \qquad \text{and} \qquad h_2(x) = \frac{6x}{1 + 3x}\ . \tag{4.41}$$

The inequality $h_1(x) \ge h_2(x)$ is equivalent to

$$3x^3 + 5x^2 + x - 1 \ge 0\ . \tag{4.42}$$

For $x > 0$, the inequality in (4.42) holds if and only if x is sufficiently small, as the coefficients of the cubic involved have only one sign change. By Descartes' Rule of Signs, this cubic can have at most one positive root. Its unique root occurs at $x = 1/3$. We thus conclude that the functions h_1 and h_2 in (4.41) cross exactly once, as do the failure rates of the two systems involved. The two failure rates will cross at the time t_0 for which $G(t_0) = 1/3$ or, equivalently, at time $t_0 = F^{-1}(1/4)$. It follows that the 2-out-of-3 system has a smaller failure rate than the mixed system for $0 \le t < F^{-1}(1/4)$ and has a larger failure rate than the mixed system if $t > F^{-1}(1/4)$. This implies that, relative to the hazard rate ordering, the 2-out-of-3 system would be preferable to the mixed system if the mission time T_0 of the system satisfies the inequality $T_0 \leq F^{-1}(1/4)$.

Our final example explores the behavior of the likelihood ratio for the two systems above.

Example 4.6. From (4.38) we identify the polynomials m_1 and m_2 in this problem as

$$m_1(x) = 1.5x^2 + 1.5 \qquad \text{and} \qquad m_2(x) = 6x \; .$$

The derivative of the ratio $m_2(x) \; / \; m_1(x)$ has its lone positive root at $x = 1$. Since $F(t) = 0.5$ when $x = 1$, the relative likelihood of the 2-out-of-3 system to the mixed system is increasing for times t less than the median component lifetime and decreases thereafter.

5

Further Signature-Based Analysis of System Lifetimes

5.1 An Application to Direct and Indirect Majority Systems

A *direct majority system* of order n is an n-component system that works if and only if a majority of its components are working. When n is odd, a direct majority system is simply an $(n+1)/2$-out-of-n system. When n is even, randomization is generally employed in defining the system. Specifically, if $n = 2m$ for some positive integer m, then a direct majority system works either if more than m components are working or if exactly m are working and a Bernoulli random variable $X \sim B(1, 1/2)$ yields a success. (Examples of the case where n is even tend to be sociological rather than engineering-related; for instance, if four people go out to dinner and two favor one restaurant while the other two favor another, the matter is likely to be settled by a coin toss.) When n is odd (say, $n = 2m+1$), the reliability polynomial of the direct majority system $\tau(m+1 \mid 2m+1)$ in "pq form" is given by

$$h_{\tau(m+1 \mid 2m+1)}(p) = \sum_{i=m+1}^{2m+1} \binom{2m+1}{i} p^i q^{2m+1-i} \,. \qquad (5.1)$$

When $n = 2m$, it is easy to show that the system's reliability polynomial is equal to that of a direct majority system in $2m-1$ components. Hence we will, without loss of generality, restrict attention to the case where the number of components n is odd.

The frequent occurrence of the phrase "the majority rules" in common parlance suggests that direct majority systems have wide applicability. The Marquis of Condorcet (1743-94) was an early proponent of the use of majority systems in social and political situations. He advocated the use of juries that make decisions on the basis of the will of the majority (see Boland [15] for details). One form of the Condorcet Jury Theorem states that if n is an odd integer greater than or equal to 3 and $p > 1/2$, then $h_{\tau((n+1)/2 \mid n)}(p) \geq p$ and, moreover, $h_{\tau((n+1)/2 \mid n)}(p)$ increases to 1 as $n \to \infty$. This result indicates that if a jury operating on a "majority rules" basis consists of n individuals whose disposition toward conviction p is the same and exceeds 1/2, then the jury is more likely to convict the defendant than any individual member of the jury making the decision alone. Condorcet's theorem makes an assumption of homogeneity (that is, a common p) and independence of individuals – essentially the i.i.d. assumptions we have made in our discussion of coherent and mixed systems. It thus seems reasonable to expect that some additional insights can be gained on majority systems through the application of system signatures in this particular context. The subject matter is of some interest in the field of engineering, since many safety systems are constructed using some form of "majority logic" in their design.

Some engineering applications require a more intricate design than that of a direct majority system. For any odd integers R and S, a system in $R \times S$ components is said to be an *indirect majority system* if the components are divided into R subsystems of size S and the system works if and only if a majority of the subsystems work, with a subsystem working if and only if a majority of its components are working. The definition of indirect majority systems generalizes to subsystems of different sizes (the American presidential electoral system being a prime example), but this extension will not be pursued here, as our present purpose is simply to demonstrate the applicability of signatures to the general area. In this section, we will be interested in comparing the performance of direct and indirect majority systems of a given size n.

Let R and S be odd positive integers, and let $\tau_{R \times S}$ denote an indirect majority system of size $n = R \times S$. Using the theory of total positivity, Boland, Proschan and Tong [19] showed that for $n = R \times S$, $h_{\tau((n+1)/2 \mid n)}(p) \geq h_{\tau(R \times S)}(p)$ if and only if $p \geq 1/2$. In words, this indicates that when components are independent, homogeneous and reasonably reliable (i.e., $p \geq 1/2$), a direct majority system is more reliable than any indirect majority system of the same size. While this result would seem to imply that a direct majority system is preferable to an indirect majority system of the same size, we will see below that, using the concept of signature, one can prove that the mean life of an indirect majority system always exceeds that of the direct majority system of the same size when the i.i.d. components have a distribution with a decreasing failure rate (DFR). The result presented below is due to Boland and Samaniego [20]. One ingredient in its proof is a certain symmetry property of the signature vectors of indirect majority systems. Boland [16] showed that the signature vector of the $n = R \times S$ indirect majority system is symmetric (i.e., $s_i = s_{n-i+1}$ for $i = 1, \ldots, n$). Table 5.1 illustrates this property for the signatures of 5×3 and 3×5 indirect majority systems.

Table 5.1. Signatures for two indirect majority systems of size $n = 15$

	$n = 5 \times 3$	$n = 3 \times 5$
i	$\mathbf{s}_i$	$\mathbf{s}_i$
≤ 5	0	0
6	0.054	0.060
7	0.240	0.220
8	0.412	0.440
9	0.240	0.220
10	0.054	0.060
≥ 11	0	0

Now, let τ be a mixed system based on components with i.i.d. lifetimes, and let $\mathbf{s}$ be its signature vector. From Chapter 3 (equations (3.6) and (3.7)), we may represent the survival function of the system's lifetime T as

$$P(T > t) = \sum_{i=1}^{n} s_i P(X_{i:n} > t) \tag{5.2}$$

and the system's expected lifetime as

$$E(T) = \sum_{i=1}^{n} s_i EX_{i:n} \ . \tag{5.3}$$

Using these representations, we prove the claim made above about the comparative means of the lifetimes of direct and indirect majority systems.

Theorem 5.1. *Let n, R and S be odd integers, with $n = R \times S$. Let $T_{((n+1)/2 \mid n)}$ and $T_{R\times S}$ be the lifetimes of the direct majority system and the $R \times S$ indirect majority system, respectively, with i.i.d. DFR component lifetimes given by $X_1, \ldots, X_n$. Then*

$$ET_{((n+1)/2 \mid n)} \leq ET_{R\times S} \ . \tag{5.4}$$

Proof. Consider the function $g : \{1, 2, \ldots, n\} \to (0, \infty)$ defined by $g(i) = E(X_{i:n})$, where $\{X_{i:n}, i = 1, \ldots, n\}$ are the order statistics corresponding to the i.i.d. component lifetimes $X_1, \ldots, X_n$. Under the assumed condition that the underlying distribution of the Xs is DFR, Kirmani and Kochar [49] showed that g is a convex function in i. Now, let $\mathbf{s}$ be the signature of the $R\times S$ indirect majority system. As noted above, $\mathbf{s}$ is symmetric about the central position in this n-dimensional vector. This symmetry implies that $\sum_{i=1}^{n} is_i = (n+1)/2$, as this sum is simply the mean of a symmetric distribution on the integers $\{1, \ldots, n\}$. From (5.3) and the convexity of g, we have that

$$ET_{R\times S} = \sum_{i=1}^{n} g(i)s_i \geq g\left(\sum_{i=1}^{n} i\ s_i\right) = ET_{((n+1)/2 \mid n)} \ , \tag{5.5}$$

the last equality following from the symmetry of $\mathbf{s}$ about its mean $(n+1)/2$. This completes the proof. ■

Example 5.1. The exponential distribution with mean μ is a boundary member of the DFR class. Thus, Theorem 5.1 is applicable for a sample of n i.i.d. exponential random variables and implies, for $n = 15$ say, that $ET_{8|15} \leq ET_{5\times 3}$. An exact calculation results in values which illustrate inequality (5.4), with $ET_{8|15} = 0.7254\mu$ and $ET_{5\times 3} = 0.7336\mu$.

5.2 An Application to Consecutive k-out-of-n Systems

Chiang and Niu [27] introduced the concept of the consecutive k-out-of-n system. As the name suggests, such systems have their components arranged in order, and they fail if k components in a row fail. Interest in these systems is motivated, in part, by their occurrence in applications involving the performance of oil pipelines and of telecommunication systems. They have also been found to be useful as models in the design of integrated circuitry. A substantial literature on this type of system design has accumulated since the appearance of Chiang and Niu's paper, much of it directed at the challenging problem of the computation of such a system's reliability (see, for example, Derman et al. [30], and Du and Hwang [32]). Other issues receiving recent attention include investigations of consecutive k-out-of-n systems with dependent components (see Boland et. al. [19] and Gera [39]), investigations into component importance in such designs (e.g. Zuo [76], Chang et. al. [24]) and studies on optimal arrangement of components (see Tong [70] and Du et. al. [31]). For comprehensive reviews of the literature on this topic, see Chao, Fu and Koutras [26] and, especially, the book by Chang, Cui, and Hwang [25].

In this section, we will apply the notion of system signatures to obtain some further insights into consecutive k-out-of-n systems. The developments below are based on work by Boland and Samaniego [21]. There are two types of consecutive k-out-of-n systems that one finds in the reliability literature – linear and circular consecutive k-out-of-n systems. We will restrict our attention here to the former type. A system is said to be a *linear consecutive k-out-of-n system* if its n components are ordered linearly, from first to last, and the system fails as soon as a set of consecutive components $i, i+1, \ldots, n(i)$ have failed, where $1 \leq i \leq n(i) \leq n$ and $n(i) = i + k - 1$. Since only the linear version will be discussed here, any reference made below to a consecutive k-out-of-n system will tacitly assume linearity.

Among linear consecutive k-out-of-n systems, the simplest case is the consecutive 2-out-of-n system. We will utilize $\tau_{c:2|n}$ as shorthand notation for this system. Recall that our definition of the ith element s_i of the signature vector is the probability that the system fails upon the ith component failure. The concept of signature is fundamentally related to cut sets of the system. In what follows, it will be useful to use a complementary concept related to the system's path sets. Let a_j represent the proportion of path sets among the $\binom{n}{j}$ sets with exactly j working components. It then follows that the reliability polynomial $h(p)$ of (3.4) can be written as

$$h(p) = \sum_{j=1}^{n} a_j \binom{n}{j} p^j q^{n-j} \ . \tag{5.6}$$

Comparing (5.6) to (3.4), in which $h(p)$ is written as a function of the signature of the system, one may deduce that the vectors $\mathbf{a}$ and $\mathbf{s}$ are related as follows:

$$a_j = \sum_{i=n-j+1}^{n} s_i \qquad \text{for } j = 1, \dots, n \tag{5.7}$$

or equivalently

$$s_j = a_{n-j+1} - a_{n-j} \qquad \text{for } j = 1, \dots, n \tag{5.8}$$

We will now obtain the vector $\mathbf{a}$ explicitly for the system $\tau_{c:2|n}$. It can be shown that there are precisely $\binom{i+1}{n-i}$ ways one can have i successes and $n-i$ failures in a linear arrangement of n components for which there is at least one success between every two failures. (We interpret $\binom{A}{B}$ as 0 when $A < B$.) The key idea in establishing the claim above is the following: Consider a sequence of i successes (Ss) arranged linearly. These define $i+1$ "spaces," the first one preceding the first S, the second being between the first and second S, and so on, with the $(i+1)$st following the ith and final S. Then the event described above will occur if and only if there is at most one failure (F) in each space. What remains is to determine the number of ways in which this can happen. This is equivalent to choosing $n-i$ spaces among $i+1$ and placing an F in each slot selected. The answer is clearly $\binom{i+1}{n-i}$. This argument holds for any value of i for which $n-i \leq i+1$, that is, for $i \geq (n-1)/2$. Since such sequences of Ss and Fs are necessary and sufficient for the system $\tau_{c:2|n}$ to work with precisely i working components, we may surmise that

$$a_i \binom{n}{i} = \binom{i+1}{n-i} \qquad \text{for } \frac{n-1}{2} \leq i \leq n \tag{5.9}$$

It therefore follows that the reliability polynomial for a consecutive 2-out-of-n system is given by the two equivalent expressions below.

$$h_{c:2|n}(p) = \sum_{i=[n/2]}^{n} \binom{i+1}{n-i} p^i q^{n-i} = \sum_{i=0}^{[(n+1)/2]} \binom{n-i+1}{i} p^{n-i} q^i \ . \tag{5.10}$$

Using (5.8) and (5.9), one can obtain the signature of any consecutive 2-out-of-n system. In Table 5.2, the signature vectors of consecutive 2-out-of-n systems

Table 5.2. Signatures for consecutive 2-out-of-n systems with $2 \le n \le 8$

i	$s_{c:2\|2}$	$s_{c:2\|3}$	$s_{c:2\|4}$	$s_{c:2\|5}$	$s_{c:2\|6}$	$s_{c:2\|7}$	$s_{c:2\|8}$
1	0	0	0	0	0	0	0
2	1	2/3	1/2	4/10	5/15	10/35	7/28
3		1/3	1/2	5/10	7/15	15/35	11/28
4			0	1/10	3/15	9/35	8/28
5				0	0	1/35	2/28
6					0	0	0
7						0	0
8							0

are displayed for $2 \le n \le 8$.

In order to investigate the relationship between consecutive 2-out-of-n systems for different values of n, we will make use of the following result.

Lemma 5.1. *Let $h_n(p)$ and $h_{n+1}(p)$ be the reliability polynomials of mixed systems of order n and $n+1$ respectively. From* (5.6)*, these polynomials can be written as*

$$h_r(p) = \sum_{j=1}^{r} a_{j,r} \binom{r}{j} p^j q^{r-j} , \tag{5.11}$$

with $r = n$ or $n+1$, where the elements $a_{j,r}$, for $j = 1, 2, \ldots, r$, are defined as in (5.7)*, that is,*

$$a_{j,r} = \sum_{i=r-j+1}^{r} s_{i,r} \qquad \text{for } j = 1, \ldots, r ,$$

with the subscript r representing the order of the system and $(s_{1,r}, \ldots, s_{r,r})$ being its signature. If, for all $j \in \{1, 2, \ldots, n\}$,

$$a_{j,n} \binom{n}{j} + a_{j+1,n} \binom{n}{j+1} \ge a_{j+1,n+1} \binom{n+1}{j+1} , \tag{5.12}$$

then $h_n(p) \ge h_{n+1}(p)$ for all $p \in [0, 1]$.

Proof. Note that the left hand side of (5.12) is the coefficient of $p^{j+1}q^{n-j}$ in the polynomial $(p+q)h_n(p)$, while the right hand side of (5.12) is the coefficient of $p^{j+1}q^{n-j}$ in $h_{n+1}(p)$. Condition (5.12) states that, of these two polynomials of degree $n+1$, the coefficients of $h_n(p)$ are uniformly equal to or larger than those of $h_{n+1}(p)$. The condition is therefore clearly a sufficient condition for the claimed domination. ■

From this, we may obtain the following result.

Theorem 5.2. *Let $T_{c:2|n}$ be the lifetime of a consecutive 2-out-of-n system with i.i.d. component lifetimes. Then $T_{c:2|n} \geq_{st} T_{c:2|n+1}$.*

Proof. The vectors $\mathbf{a}_n$ and $\mathbf{a}_{n+1}$ corresponding to the systems $\tau_{c:2|n}$ and $\tau_{c:2|n+1}$ can be identified from (5.9). One can show that the former system is uniformly superior to the latter simply by verifying that these vectors satisfy condition (5.12). From (5.9), we see that for the consecutive 2-out-of-N systems compared in this theorem, condition (5.12) reduces to

$$\binom{j+1}{n-j} + \binom{j+2}{n-j-1} \geq \binom{j+2}{n-j} \qquad \text{for } \frac{n-3}{2} \leq j \leq n\,, \tag{5.13}$$

where $\binom{A}{B}$ is taken as 0 if $A < B$ or if $B < 0$. Letting $r = j+1$ and $k = n-j$ in (5.13), the inequality may be rewritten as

$$\binom{r}{k} + \binom{r+1}{k-1} \geq \binom{r+1}{k} \qquad \text{for } k = 0, 1, \ldots, r+2\,. \tag{5.14}$$

But since $\binom{r+1}{k-1} \geq \binom{r}{k-1}$, the inequality in (5.14) follows immediately from the well-known identity

$$\binom{r}{k} + \binom{r}{k-1} = \binom{r+1}{k} \qquad \text{for } k = 0, 1, \ldots, r+1\,. \tag{5.15}$$

We may thus conclude that condition (5.12) holds for the systems in question. It follows from Lemma 5.1 that $T_{c:2|n} \geq_{\text{st}} T_{c:2|n+1}$. ∎

Theorem 5.2 generalizes to consecutive k-out-of-n systems for $2 < k \leq n$. This result is established in Boland and Samaniego [21] using a different method of proof.

5.3 The Limiting Behavior of System Failure Rates and Survival Curves

In this section, the representations of $\overline{F}_T(t)$ and $r_T(t)$ in Chapter 3 are utilized to study the behavior of a mixed system as $t \to \infty$. The results presented here are due to Block, Dugas and Samaniego [10]. The first result identifies the asymptotic failure rate of the system as a multiple of the limiting failure rate of an individual component, that multiple depending on the largest index i of the signature vector for which $s_i > 0$. This result extends the recently established results by Block, Li and Savits [13] where conditions which determine the asymptotic behavior of the failure rate of the mixture of lifetime distributions are identified. One of their principal results is that if the failure rates of the component lifetimes have limits, the failure rate of the mixture converges to the limit of the failure rate of the strongest component. The approach taken here differs from earlier studies in that our focus and our results are based on signatures of coherent or mixed systems.

Theorem 5.3. *Let T be the lifetime of a mixed system based on a set of coherent systems in n i.i.d. components, each component having a common failure rate $r(t)$. Assume that $r(t)$ has limit r as $t \to \infty$, where $0 \le r \le \infty$. If the system has signature $\mathbf{s} = (s_1, s_2, \ldots, s_n)$ and failure rate $r_T(t)$, then*

$$r_T(t) \to (n - K + 1)r \ , \tag{5.16}$$

where $K = \max\{i \mid s_i > 0\}$.

Proof. Consider the representation of $r_t(t)$ given in (3.11), namely,

$$r_T(t) = \frac{\sum_{i=0}^{n-1}(n-i)s_{i+1}\binom{n}{i}(G(t))^i}{\sum_{i=0}^{n-1}\left(\sum_{j=i+1}^{n} s_i\right)\binom{n}{i}(G(t))^i}\, r(t) \ , \tag{5.17}$$

where $G(t) = F(t) \ / \ \overline{F}(t) \in (0, \infty)$. Note that, by the definition of K, both the numerator and the denominator of (5.17) are polynomials in the variable "$G(t)$" and have degree $K-1$. If we divide the numerator and the denominator of (5.17) by $(G(t))^{K-1}$, we see that $r_T(t)$ has the same limit, as $t \to \infty$, as

$$\frac{o(1) + (n-K+1)s_K\binom{n}{K-1}}{o(1) + s_K\binom{n}{K-1}}\, r(t) \ . \tag{5.18}$$

It follows from (5.18) that $r_T(t) \to (n - K + 1)r$ as claimed. ∎

Remark 5.1. We note that the failure rate representation in (5.17) can also be employed to identify the behavior of the system failure rate near zero. If T is the lifetime of a mixed system based on a set of coherent systems in n i.i.d. components, each component having a common failure rate $r(t)$, then letting $t \to 0$ in (5.17) shows that $r_T(t) \to ns_1 r(0^+)$, where

$$r(0^+) = \lim_{t \to 0^+} r(t).$$

We now turn to an examination of the relative rate of convergence to zero of the survival functions of two systems under comparison. For simple systems, information on this question is quite transparent. For example, the survival function of a parallel system in i.i.d. components tends to zero more slowly than the survival function of an individual component, and the opposite is true for series systems. The following results utilize the representation of $\overline{F}_T(t)$ in (3.2) to provide definitive descriptions of the rates of convergence of survival functions to zero for general mixed systems.

Theorem 5.4. *Let T be the lifetime of a mixed system with signature* $\mathbf{s}$ *based on a set of coherent systems in n i.i.d. components. Let F be the common lifetime distribution of the components. Then*

$$\frac{\overline{F}_T(t)}{\left[\overline{F}(t)\right]^{n-K+1}} \to \binom{n}{K-1} s_K \ , \tag{5.19}$$

where $K = \max\{i \mid s_i > 0\}$.

Proof. From (3.2), we have that the system has lifetime distribution

$$\overline{F}_T(t) = \sum_{j=0}^{K-1} \left(\sum_{i=j+1}^{n} s_i \right) \binom{n}{j} (F(t))^j \left(\overline{F}(t)\right)^{n-j} \ . \tag{5.20}$$

If one divides the above quantity by $\left[\overline{F}(t)\right]^{n-K+1}$ and lets t approach ∞, one obtains a summation of K terms, only the last of which has a non-zero limit, that limit being the product $\binom{n}{K-1} s_K$. ∎

Remark 5.2. Note that the rate of convergence of $\overline{F}_T(t)$ to zero is only affected by the largest index of a positive element of the signature vector. It follows that a mixed system in n components for which the element s_n of the signature vector is positive achieves the same basic rate of convergence to zero as that of the parallel system (which achieves the best (slowest) possible rate). On the other hand, systems with the same value of $K = \max\{i \mid s_i > 0\}$ have the same rate of convergence, but the survival functions, for large t, are ordered and are proportional to their corresponding values of s_K.

Theorem 5.5. *Let T_1 and T_2 be the lifetimes of two mixed systems, with signatures* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *and lifetime distributions F_1 and F_2, respectively, each based on a set of coherent systems in n i.i.d. components with common distribution F. Let* $K_1 = \max\{i \mid s_{1i} > 0\}$ *and* $K_2 = \max\{i \mid s_{2i} > 0\}$. *Then*

$$\frac{\overline{F}_2(t)}{\overline{F}_1(t)} \to \frac{s_{2K}}{s_{1K}} \qquad \text{if } K_1 = K_2 = K \tag{5.21}$$

and

$$\frac{\overline{F}_2(t)}{\overline{F}_1(t)} \to \infty(0) \qquad \text{if } K_1 < (>) K_2 \ . \tag{5.22}$$

Proof. Using the representation of $\overline{F}_T(t)$ in equation (3.4) and dividing both numerator and denominator of $\overline{F}_2(t)/\overline{F}_1(t)$ by $(G(t))^M$, where $M = \max\{K_1, K_2\}$, the desired result follows. ∎

5.4 Comparing Arbitrary Mixed Systems via Stochastic Precedence

As in previous sections, we will assume that all systems under discussion are based on components having independent lifetimes with common distribution F. It will often be the case that some form of stochastic ordering, and/or the necessary and sufficient conditions of section 4.4 which guarantee that system lifetimes will be ordered in some stochastic sense, apply to the comparison of two systems of interest. However, none of these conditions apply to all possible system comparisons. The relationships alluded to above each induce a partial ordering on the class of mixed systems. Some pairs of mixed systems are simply non-comparable using the signature conditions discussed thus far. An example of such non-comparability is the 2-component mixed system with signature $(1/2, 1/2)$ (the subject of comparisons examined in Example 4.3) and the 4-component mixed system with signature $\mathbf{s} = (0, 1/2, 1/2, 0)$. It has been noted that this 2-component system has the same lifetime distribution as the system with signature $\mathbf{s}^* = (1/4, 1/4, 1/4, 1/4)$. It is easy to verify that $\mathbf{s}$ and $\mathbf{s}^*$ fail to satisfy the usual (lr, hr or st) orderings and also fail to satisfy the less stringent NASCs of Theorems 4.7 - 4.9. Thus, none of the theorems of Chapter 4 apply to the lifetimes T_1 and T_2 of these two particular mixed systems. It is easy to find examples of pairs of coherent systems that are non-comparable in this same sense. For instance, assuming i.i.d. component lifetimes, the survival functions of coherent systems #5 and #7 in Table 3.2 cross each other and the corresponding system lifetimes are non-comparable using the tools of Chapter 4. This type of limitation of the comparative methods considered thus far leads naturally to the exploration of different metrics that might be applied to the comparative analysis of system performance.

"Stochastic precedence" was introduced by Arcones, Kvam and Samaniego [3] as an alternative approach to the notion that one random variable is smaller than another. They defined the "sp" relationship as follows.

Definition 5.1. *Let Y_1 and Y_2 be independent random variables with respective distributions F_1 and F_2. Then Y_1 is said to* stochastically precede Y_2 *(written $Y_1 \le_{sp} Y_2$) if and only if $P(Y_1 \le Y_2) \ge 1/2$. If both $Y_1 \le_{sp} Y_2$ and $Y_2 \le_{sp} Y_1$ hold, the variables are said to be* sp-equivalent. *Continuous variables Y_1 and Y_2 will be sp-equivalent if and only if $P(Y_1 \le Y_2) = 1/2$.*

It is worth noting that probabilities of the form $P(X \ge Y)$ have a long history in probability and statistics. The probability arises, most notably, in nonparametric tests for distributional equality and in "stress-strength" testing in reliability. In the former context, $P(X \ge Y)$ is the expected value of the famous Mann-Whitney statistic used to test the equality of two distributions F and H based on independent random samples (Xs and Ys) from these distributions. In that setting, $P(X \ge Y)$ is a well-accepted measure of the extent to which one distribution (or random variable) is larger than another and is thus clearly relevant in the problem of interest as well. In stress-strength testing, the variable X represents the strength of the material of interest and the variable Y represents the level of stress to which the material is subjected. With this understanding, $P(X \geq Y)$ represents the probability that randomly chosen material will survive the amount of stress that is randomly applied to it. If $P(X \geq Y)$ exceeds $1/2$, the material tends to be stronger than the stress to which it will be subjected. See, for example, Johnson [47] or Samaniego [62] for further details.

We now turn to the problem of interest here, namely, the problem of comparing the performance of two mixed systems based on i.i.d. components from a common distribution F. Let T_1 and T_2 be independent random variables representing the lifetimes of two such systems. In this section, we will treat the comparison of T_1 and T_2 via stochastic precedence. We derive below an explicit formula for computing $P(T_1 \leq T_2)$. Armed with the computed value of $P(T_1 \leq T_2)$, we will always be able to classify the second system as better than, equivalent to or worse than the first system according to whether this probability is greater than, equal to or less than $1/2$. It is therefore apparent that any pair of mixed systems of arbitrary order can be definitively compared via stochastic precedence. We note that the potential for comparison extends to the lifetimes of pairs of mixed systems in which the underlying component distributions are allowed to be different. We may, for example, posit that the components from which the first system is constructed have i.i.d. lifetimes with distribution F_1 and the components from which the second system is constructed have i.i.d. lifetimes with distribution F_2. Indeed, the expression for this probability in 3.4 is obtained under these broader assumptions. In what follows, however, we will restrict our attention to the case in which $F_1 = F_2$. The results presented here are drawn from Hollander and Samaniego [43]. See also Hollander and Samaniego [44].

Given two mixed systems τ_1 and τ_2 with corresponding lifetimes T_1 and T_2 satisfying the assumptions in the preceding paragraph, one can infer that system τ_2 will tend to last longer than system τ_1 if their lifetimes satisfy the condition $P(T_1 \leq T_2) > 1/2$. In Theorem 3.4, we established the general expression needed for stochastic-precedence calculations:

$$P(T_1 \leq T_2) = \sum_{i=1}^{n} \sum_{j=1}^{m} s_{1i} s_{2j} P(X_{i:n} \leq Y_{j:m}) \ . \tag{5.23}$$

We now develop an explicit formula from which $P(T_1 \leq T_2)$ can be computed. We first establish a lemma which treats stochastic precedence between two independently drawn order statistics. This result is a special case of Theorem 4 in Kvam and Samaniego [54], a result that applies to inequalities involving an arbitrary number $r \geq 2$ of independent order statistics. We give an elementary proof of the lemma.

Lemma 5.2. *Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a continuous distribution F, and let $X_{1:n}, X_{2:n}, \ldots, X_{n:n}$ be the corresponding order statistics. Let $Y_1, Y_2, \ldots, Y_m$ be a random sample of size m from F, and let $Y_{1:m}, Y_{2:m}, \ldots, Y_{m:m}$ be the corresponding order statistics. Assume that the two samples are independent. Then for any such distribution F,*

$$P(X_{i:n} \leq Y_{j:m}) = \sum_{k=i}^{n} \frac{\binom{n}{k}\binom{m}{j}}{\binom{n+m}{k+j}} \frac{j}{k+j} . \tag{5.24}$$

Proof. All possible orderings of the combined sample of Xs and Ys are equally likely, each occurring with probability $\frac{1}{(n+m)!}$. For the event $\{X_{i:n} \leq Y_{j:m}\}$ to occur, one must have, for some $k \in \{i, i+1, \ldots, n\}$, exactly k Xs and j Ys occurring as the first $k+j$ values in the combined sample, and given that, the last of the $k+j$ values must be a Y. The probability of the first event is $\binom{n}{k}\binom{m}{j} / \binom{n+m}{k+j}$ and the conditional probability of the second event is $j/(k+j)$. Applying the law of total probability, we obtain (5.24). ■

Lemma 5.2 provides an explicit formula which allows us to refine the expression for $P(T_1 \leq T_2)$ given in Theorem 3.4. An interesting and important by-product of the lemma is that it establishes the fact that the probability $P(T_1 \leq T_2)$ is a distribution-free measure which takes the same value for all continuous distributions F. This latter fact does not generalize to the case where the distributions of the Xs and the Ys differ. If the X sample is drawn from the distribution F_1 while the Y sample is independently drawn from the distribution F_2, the probability $P(T_1 \leq T_2)$ will depend on F_1 and F_2. This probability will also be substantially more difficult to calculate, though explicit formulas may be obtained in certain circumstances. For example, Hollander and Samaniego [43] derive an explicit formula for $P(T_1 \leq T_2)$ when F_2 is in the class of Lehmann alternatives to F_1, that is, when $F_2 = F_1^k$ for some positive k. For the case of interest here (that is, where $F_1 = F_2 = F$), the computational formula for $P(T_1 \leq T_2)$ is given below in its final form.

Theorem 5.6. *Let T_1 and T_2 represent the lifetimes of mixed systems of orders n and m with respective signatures $\mathbf{s}_1$ and $\mathbf{s}_2$ and ordered component lifetimes $\{X_{1:n}, X_{2:n}, \ldots, X_{n:n}\}$ and $\{Y_{1:m}, Y_{2:m}, \ldots, Y_{m:m}\}$ obtained from two independent i.i.d. samples of sizes n and m from a common distribution F. Then*

$$P(T_1 \leq T_2) = \sum_{i=1}^{n}\sum_{j=1}^{m} s_{1i}s_{2j} \sum_{k=i}^{n} \frac{\binom{n}{k}\binom{m}{j}}{\binom{n+m}{k+j}} \frac{j}{k+j} . \tag{5.25}$$

The following pair of examples illustrates the utility of stochastic precedence in making system comparisons. Systems that are non-comparable using the stronger conditions on pairs of signatures found in earlier sections of this monograph will now be compared via stochastic precedence.

Example 5.2. Consider the two mixed systems discussed in the opening paragraph of this section. To conform to the notation of Theorem 5.6, we relabel these two systems' signatures as $\mathbf{s}_1 = (1/4, 1/4, 1/4, 1/4)$ and $\mathbf{s}_2 = (0, 1/2, 1/2, 0)$. Theorem 5.6 allows us to compute $P(T_1 \leq T_2)$ explicitly. Adding the eight terms of (5.25) for which the product $s_{1i}s_{2j} > 0$, we find that $P(T_1 \leq T_2) = 1/2$. Thus, although the two systems above are not comparable on the basis of the criteria treated in Chapter 4, we see that they are comparable in the alternative metric of stochastic precedence and are found to be equivalent in that metric.

Example 5.3. We have noted that the coherent systems #5 and #7 in Table 3.2 are non-comparable in the "st" sense (which of course implies the same in the "hr" and "lr" senses). Here, we will compare these systems via stochastic precedence. System #5 has signature $\mathbf{s}_1 = (1/4, 1/4, 1/2, 0)$ and system #7 has signature $\mathbf{s}_2 = (0, 2/3, 1/3, 0)$. Using Theorem 5.6, the probability $P(T_1 \leq T_2)$ is found to be equal to 109/210 or 0.519. From this we conclude that system #7 will last longer than system #5 slightly more than half the time. Thus, under a stochastic precedence criterion, system #7 would be judged to be preferable to system #5.

As shown in Arcones, Kvam and Samaniego [3], stochastic precedence is weak ordering, being implied, for example, by the usual stochastic ordering. It is especially weak in the discrete case. For example, the 3-component systems with signatures $\mathbf{s}_1 = (0, 3/8, 5/8)$ and $\mathbf{s}_2 = (0, 0, 1)$ are sp-equivalent. If $X_i \sim \mathbf{s}_i$ for $i = 1, 2$, then $P(X_1 \leq X_2) = 5/8$ and $P(X_2 \leq X_1) = 1$, so that both $X_1 \leq_{\text{sp}} X_2$ and $X_2 \leq_{\text{sp}} X_1$ hold. Thus, even though $X_1 \leq_{\text{st}} X_2$, the two signatures are equivalent in the "sp" sense. It is apparent from the above that preservation theorems such as those featured in section 4.2 do not hold for stochastic precedence. The condition $\mathbf{s}_1 \leq_{\text{sp}} \mathbf{s}_2$ will not necessarily imply that $T_1 \leq_{\text{sp}} T_2$. It is not easy to identify simple sufficient conditions on two signature vectors for the inequality $T_1 \leq_{\text{sp}} T_2$ to hold. It is thus perhaps surprising that necessary and sufficient conditions on the signatures of two systems which guarantee the inequality $T_1 \leq_{\text{sp}} T_2$ can be obtained without any additional strain. This is due entirely to the interesting and important fact that the probability $P(T_1 \leq T_2)$ is a distribution-free measure of the comparative behavior of the two systems involved. This immediately leads to an NASC for stochastic precedence between the lifetimes of two mixed systems of arbitrary order based on i.i.d. component lifetimes. Indeed, if the RHS of (5.25) is denoted by $W(\mathbf{s}_1, \mathbf{s}_2)$, then

$$P(T_1 \leq T_2) > 1/2 \text{ (or } = 1/2 \text{ or } < 1/2)$$

if and only if

$$W(\mathbf{s}_1, \mathbf{s}_2) > 1/2 \text{ (or } = 1/2 \text{ or } < 1/2) . \tag{5.26}$$

Remark 5.3. From Theorem 5.6, we may identify $W(\mathbf{s}_1, \mathbf{s}_2)$ as precisely equal to the probability $P(T_1 \le T_2)$. It is, in fact, simply a formula for computing $P(T_1 \le T_2)$. Thus, the condition $W(\mathbf{s}_1, \mathbf{s}_2) \ge 1/2$ is necessary and sufficient to ensure that $T_1 \le_{\text{sp}} T_2$. The legitimacy of this as an NASC derives from the fact $W(\mathbf{s}_1, \mathbf{s}_2)$ is independent of the underlying distribution F and depends only on the signatures of the two systems of interest.

Remark 5.4. It would seem from (5.25) that computing $W(\mathbf{s}_1, \mathbf{s}_2)$ can be a cumbersome matter for large systems with complex signatures. However, it is clear that the formula in (5.25) is easily programmed. Thus, checking for stochastic precedence of one mixed system relative to another is a process that may be executed without much difficulty.

We began this section with a discussion of two mixed systems having signatures $(1/2, 1/2)$ and $(0, 1/2, 1/2, 0)$ respectively. Example 5.2 continued the discussion, carrying out the comparison via stochastic precedence of the second system with a 4-component system whose lifetime was stochastically identical to the first. The two systems were shown to be equivalent in the "sp" sense. This outcome is no accident, as it is typical of "sp" comparisons when the signature vectors are symmetric. A general result in this regard is established below. We will prove the next two results for systems of the same size (i.e., $n = m$), but as we shall see in the corollary that follows, the main result holds for $n \ne m$ as well. The following tool will be required.

Lemma 5.3. *Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be two independent random samples of size n from a continuous distribution F, and let $X_{1:n}, X_{2:n}, \ldots,$ $X_{n:n}$ and $Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$ be the corresponding order statistics. Then (i) $P(X_{i:n} \le Y_{i:n}) = 1/2$, (ii) $P(X_{i:n} \le Y_{j:n}) = P(X_{n-j+1:n} \le Y_{n-i+1:n})$ and (iii) $P(X_{i:n} \le Y_{j:n}) = 1 - P(X_{j:n} \le Y_{i:n})$.*

Proof. Because of the continuity of F, the equality in each of the inequalities above has probability zero. Consider the possible $(2n)!$ permutations of the combined sample of Xs and Ys. Each ordering is equally likely. Claim (i) follows from the fact that, for every ordering for which $X_{i:n} \le Y_{i:n}$, the ordering that relabels X_i as Y_i and vice versa results in $Y_{i:n} \le X_{i:n}$. Thus these events have the same probability, namely 1/2. Claim (ii) follows from the fact that if one reflects the Xs and Ys in each particular permutation so that the ith smallest item becomes the ith largest, the event $\{X_{i:n} \le Y_{j:n}\}$ in the original collection of permutations will correspond to the event $\{X_{n-i+1:n} \ge Y_{n-j+1:n}\}$ in the reflected set, implying that these events have the same probability. Finally, since $\{X_{i:n} \le Y_{j:n}\} = \{X_{i:n} \ge Y_{j:n}\}^c$ when F is continuous and $P(X_{j:n} \le Y_{i:n}) = P(X_{i:n} \ge Y_{j:n})$, claim (iii) follows. ■

Theorem 5.7. *Let $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ be two independent random samples of size n from a continuous distribution F, and let $X_{1:n}, X_{2:n}, \ldots,$*

$X_{n:n}$ and $Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$ be the corresponding order statistics. Let τ_1 and τ_2 be mixed systems of order n based on components with i.i.d. lifetimes $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ as above. Denote the signatures of τ_1 and τ_2 by $\mathbf{s}_1$ and $\mathbf{s}_2$ and their lifetimes T_1 and T_2. If both signature vectors are symmetric, that is, that $s_{1,i} = s_{1,n-i+1}$ for all i and $s_{2,i} = s_{2,n-i+1}$ for all i, then $P(T_1 \leq T_2) = 1/2$, that is, T_1 and T_2 are sp-equivalent.

Proof. For simplicity, we will denote $P(X_{i:n} \leq Y_{j:n})$ by $p_{i,j}$. From Lemma 5.3, we have that, for all i and j, (i) $p_{i,i} = 1/2$, (ii) $p_{i,j} = p_{n-j+1,n-i+1}$ and (iii) $p_{i,j} = 1 - p_{j,i}$. Using property (i), we may rewrite the expression for $P(T_1 \leq T_2)$ in (5.23) as

$$\begin{aligned} P(T_1 \leq T_2) &= \frac{1}{2}\sum_{i=1}^{n} s_{1,i}s_{2,i} + \sum_{i<j} s_{1,i}s_{2,j}p_{i,j} + \sum_{i>j} s_{1,i}s_{2,j}p_{i,j} \\ &= \frac{1}{2}\sum_{i=1}^{n} s_{1,i}s_{2,i} + \sum_{i<j} s_{1,i}s_{2,j}p_{i,j} + \sum_{i>j} s_{1,i}s_{2,j}(1 - p_{j,i}) \end{aligned}$$

(by virtue of property (iii))

$$= \frac{1}{2}\sum_{i=1}^{n} s_{1,i}s_{2,i} + \sum_{i<j} s_{1,i}s_{2,j}p_{i,j} + \sum_{i>j} s_{1,n-i+1}s_{2,n-j+1}(1 - p_{n-i+1,n-j+1})$$

(by virtue of the symmetry of the signatures and property (ii))

$$\begin{aligned} &= \frac{1}{2}\sum_{i=1}^{n} s_{1,i}s_{2,i} + \sum_{i<j} s_{1,i}s_{2,j}p_{i,j} + \sum_{i<j} s_{1,i}s_{2,j}(1 - p_{i,j}) \\ &= \frac{1}{2}\sum_{i=1}^{n} s_{1,i}s_{2,i} + \sum_{i<j} s_{1,i}s_{2,j} \ . \end{aligned} \tag{5.27}$$

By the symmetry of $\mathbf{s}_1$ and $\mathbf{s}_2$, we have

$$\sum_{i<j} s_{1,i}s_{2,j} = \sum_{i<j} s_{1,n-i+1}s_{2,n-j+1} = \sum_{i>j} s_{1,i}s_{2,j} \ ,$$

which implies that we may rewrite the expression for $P(T_1 \leq T_2)$ in (5.27) as

$$\begin{aligned} P(T_1 \leq T_2) &= \frac{1}{2}\left\{\sum_{i=1}^{n} s_{1,i}s_{2,i} + \sum_{i<j} s_{1,i}s_{2,j} + \sum_{i>j} s_{1,i}s_{2,j}\right\} \\ &= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} s_{1i}s_{2j} = 1/2. \qquad \blacksquare \end{aligned}$$

Corollary 5.1. *Let τ_1 and τ_2 be mixed systems of orders n and m, respectively, based on components with i.i.d. lifetimes $\sim F$, and denote their signatures by $\mathbf{s}_1$ and $\mathbf{s}_2$ and their lifetimes by T_1 and T_2. If both signature vectors are symmetric, that is, if $s_{1,i} = s_{1,n-i+1}$ for all i and $s_{2,i} = s_{2,m-i+1}$ for all i, then $P(T_1 \leq T_2) = 1/2$. Thus T_1 and T_2 are equivalent in the sense of stochastic precedence.*

Proof. If $n = m$, the result is proven in Theorem 5.7. Assume, without loss of generality, that $n < m$. Recall from Theorem 3.2 that, assuming i.i.d. component lifetimes $\sim F$, the system with signature $\mathbf{s}_1$ has the identical lifetime distribution as the $(n+1)$-component system with signature

$$\mathbf{s}^* = \left(\frac{n}{n+1}s_1,\ \frac{1}{n+1}s_1 + \frac{n-1}{n+1}s_2,\ \frac{2}{n+1}s_2 + \frac{n-2}{n+1}s_3, \ldots, \frac{n-1}{n+1}s_{n-1} + \frac{1}{n+1}s_n,\ \frac{n}{n+1}s_n\right) .$$

As stated in Corollary 3.3, the signature $\mathbf{s}^*$ will be symmetric if $\mathbf{s}_1$ is. Repeated applications of Theorem 3.2 lead to a signature $\mathbf{s}_1^*$ of an m-component system that is symmetric and has the same lifetime distribution as T_1. The fact that $P(T_1 \leq T_2) = 1/2$ will then follow from Theorem 5.7. ■

Example 5.4. The symmetry conditions on the signatures $\mathbf{s}_1$ and $\mathbf{s}_2$ in Theorem 5.7 and Corollary 5.1 are sufficient conditions for the sp-equivalence of two system lifetimes, but they are not necessary. Let T_1 and T_2 be the lifetimes of two mixed systems based on components with i.i.d. lifetimes governed by the continuous distribution F, and let the corresponding system signatures be $\mathbf{s}_1 = (1/4, 1/4, 1/2, 0)$ and $\mathbf{s}_2 = (0, 3/4, 1/4, 0)$. It is easy to confirm that T_1 and T_2 are sp-equivalent.

6

Applications of Signatures to Network Reliability

6.1 An Introduction to Communication Networks

Communication networks have been used and studied for several decades, but their application and broad utilization has expanded dramatically in the last 15 years. The original DARPA nationwide electronic network was established largely for military communications and served as the first truly large-scale application in the field. It is still in the memory banks of many present-day researchers. The World Wide Web, a network which we take completely for granted today, as if it has always been there, dates back only to the early nineties and owes much to the DARPA model for its design. These examples of modern electronic communication systems have important predecessors in the area of telephone and telegraph networks, but the specific problems that will interest us here are of more modern origin. We are interested in the mathematical and probabilistic analysis involved in addressing "connectivity" problems. Specifically, we will want to consider questions such as: "What are the chances that all 'terminals' in a given network can communicate with each other?" and "Which of two competing network designs has a higher probability of being connected?" One can, of course, ask more fundamental questions regarding one's ability to compute answers to these two questions. A good deal of research in network theory is directed at classifying computational algorithms in network reliability into problem types such as NP-complete, NP-hard, problems solvable in polynomial time, etc., classes which indicate whether or not a problem is hopelessly complex or is, at the other extreme, computationally feasible. Chapter 5 of Colbourn [28] has an excellent summary of problem types relative to this classification system. We will not deal with this latter problem here. Instead, we will focus exclusively on ways of computing the probability that a network is operational.

Following Colbourn [28], we will think of a "network" as an undirected graph with a fixed number of vertices and a fixed number of edges connecting pairs of vertices. We will assume that every vertex in such a network is connected to any other through at least one sequence of edges. A network with v vertices and n edges is typically denoted by the symbol $G(v, n)$. Further, we take a viewpoint that is quite standard in the field, namely, that vertices cannot fail, but that edges can be in either a functioning or a failed state. It can be argued that such an assumption does not in fact limit the scope of the problems we will study, as relaxing that assumption can be handled by embedding a given network problem into a larger network problem satisfying this assumption (see Colbourn [28], Chapter 4, for details). In communication networks, "connectivity" is the "quality" characteristic of primary interest, and most of the optimization problems that arise are directed at maximizing the probability that the network is "connected" in some sense. Two-terminal connectivity means that there is at least one set of functioning edges providing a path from one distinguished terminal to another. One can also consider k-terminal connectivity for $k > 2$ and its fullest extension, v-terminal connectivity, the latter term meaning that every vertex is connected to every other vertex by at least one functioning path.

Whichever version of the connectivity problem is of interest, we have assumed that the network is potentially connected, but that the connectivity can be lost through the failure of one or more edges. These edges have specified probabilities of working or failing. We will focus on the problem of determining the probability that a given network is working (that is, determining the reliability of the network) and on the problem of comparing the reliability of two competing networks, assuming, in both circumstances, that the probabilities that the edges work are given. When all edges work independently of each other and have a common probability p of working, the reliability of a network with n edges can be written as an nth degree polynomial. The goal of this chapter is to clarify the connection between two natural ways of computing a network's reliability polynomial and, of course, to exploit that connection in extracting information about the network as efficiently as possible. In the sequel, we derive a general formula for linking the reliability polynomials in "standard" and "pq" forms (as defined in equations (2.23) and (2.24)) for an arbitrary network reliability problem. We limit our examples to the "two-terminal" connectivity problem in section 6.2 and to an "all-terminal" connectivity problem in section 6.3. We nonetheless aver that the treatment given here is easily extended to the other versions of the connectivity problems mentioned above.

In its most general form, the problem of computing the reliability of a communication network is a highly challenging one, as it involves handling the potentially complex multivariate distribution describing edge behavior, including any and all dependencies among edges, and it requires a calculation which takes account of this dependence structure. Not only do computational issues pose difficulties, the comparison of competing networks magnifies them. We will restrict our attention to a simpler but still quite important problem, namely, that of addressing reliability calculations and comparisons when the edges' lifetimes or their states (working or failed) at a given point in time are independent and identically distributed. As we shall see, even this simpler problem poses formidable analytical challenges. Solving such problems is clearly the first step toward the solution of the more general problem of computing network reliability when edges are independent but not identically distributed. The most general form of the problem, which permits dependencies among the lifetimes of individual edges, is likely to resist analytical solution in the foreseeable future, although certain special cases may prove to be manageable. The developments below are largely based upon the work in Boland, Samaniego and Vestrup [22].

An additional and quite typical assumption we shall make is that the networks on which we focus attention are "coherent," that is, every edge of the network is relevant and any set of edges containing a path set is also a path set (a condition equivalent to monotonicity). Proceeding under the assumptions of coherency and i.i.d. edge lifetimes with common distribution F, let us consider the calculation of the distribution of T, the failure time of the network (i.e., the time at which it becomes disconnected in the sense relevant to the context under study). We will focus, in particular, on the calculation of the probability that the network is connected at a given time t_0. In the latter instance, we'll treat the states of edges (i.e. working or failed states) as independent Bernoulli variables. In what follows, we will be interested in the special case in which these variables are also identically distributed, that is, they are i.i.d. Bernoulli variables with common "success" probability $p = \overline{F}(t_0)$ at time t_0.

Let us suppose that the n edges of the network $G(v, n)$ have i.i.d. lifetimes and thus that, for each, the probability of its functioning at a fixed time t_0 is $p \in (0, 1)$. The reliability polynomial of the network can be expressed, in standard form, as

$$h(p) = \sum_{r=1}^{n} d_r p^r \ , \tag{6.1}$$

Satyanarayana and Prabhakar [63] provided an alternative (via "signed dominations") to the use of the traditional "inclusion-exclusion" formula for determining the coefficients $\{d_r\}$ in (6.1). Agrawal and Barlow [1] referred to "domination theory" as a major breakthrough in the area of computational network reliability, and today's continued use of this computational tool in network reliability confirms their assessment. For the reader's convenience, we will review the concept of dominations below. While domination theory has proven to be a major boon in simplifying the computation of the reliability of a network, it has not been found particularly useful in comparing one network design with another. The use of "dominations" in the important problem of identifying uniformly optimal networks of a given size has been limited to the role of computing the reliability functions to be compared. See Boesch, Li and

Suffel [14] for a description of the latter problem.

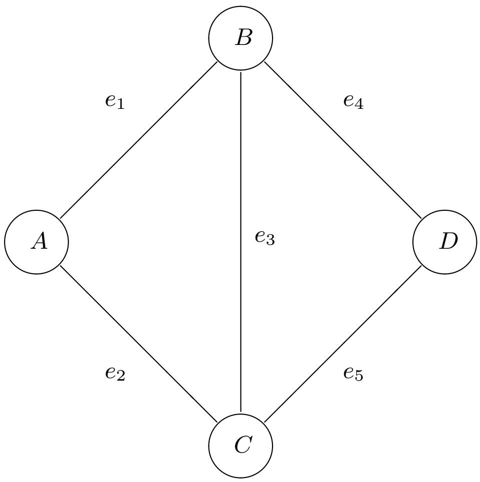
In the sequel, the notion of the "signature" of a network, obtained as the natural extension to network theory of system signatures introduced in Chapter 3, will be shown to be quite useful in problems involving network reliability. The signature of a network, like the signature of a coherent or mixed system, is a probability vector **s** whose components are the probabilities that the 1st, 2nd, ..., and nth ordered edge failures cause the network to fail. Assuming, again, i.i.d. edge states at a fixed time t_0, the reliability polynomial of a network can be expressed in terms of the network's signature vector. Unlike the domination vector, the properties of the signature vector are readily interpretable and have a close relationship to the failure time T of the network itself. Before discussing domination theory in some detail, we will briefly examine the notion of the "signature" of a network.

The signature of a network of order n (that is, having n edges) is defined as the probability distribution **s** on the integers $\{1, 2, \ldots, n\}$ for which

$$s_i = P(T = X_{i:n}), \qquad i = 1, 2, \ldots, n \; , \tag{6.2}$$

where $X_{1:n} < X_{2:n} < \cdots < X_{n:n}$ are the order statistics from a random (i.i.d.) sample of component lifetimes drawn from a continuous lifetime distribution F and T is the lifetime of the network. As an illustration, we compute below the signature of the "Wheatstone bridge" network.

Fig. 6.1. The Wheatstone bridge



The bridge network above connects the four vertices A, B, C and D via the five edges labeled e_1, e_2, e_3, e_4 and e_5. For the two-terminal connectivity problem where we are interested in whether or not terminals A and D can communicate, the network's minimal cut sets are $\{1, 2\}$, $\{4, 5\}$, $\{1, 3, 5\}$ and $\{2, 3, 4\}$. The network cannot fail upon the first edge failure, so we have that $s_1 = 0$. The second edge failure can cause network failure only if the first two failures both fall into one of the two minimal cut sets $\{1, 2\}$ or $\{4, 5\}$. There are four possible arrangements of the first two failures that will be fatal to the network. These failures could then be followed by 3! possible orderings of the three edges yet unaccounted for. Thus, 4! of the 5! permutations of the 5 failure times result in network failure upon the second edge failure. This implies that $s_2 = 1/5$. While calculating s_4 and s_5 is simpler than calculating s_3, we will do the latter as a means of giving an additional example of the kind of combinatorial thinking that is often required in the calculation of a signature. Note that the edges e_1, e_2, e_4 and e_5 have equal status in Figure 6.1. Thus, if we identify the number of permutations in which the failure of edge e_1 is the third to occur and is fatal to the network, the same number will be applicable to these other three edges. Now edge e_1 will be the third failure and will cause the network to fail if and only if edge failures occur in one of the 16 permutations

3, 5, 1, _, _; 5, 3, 1, _,_; 2, 4, 1, _, _; 4, 2, 1, _, _;
2, 3, 1, _, _; 3, 2, 1, _, _; 2, 5, 1, _, _; 5, 2, 1, _, _,

where the blanks indicate that the remaining edge failures can occur in either order. Extending this reasoning to edges e_2, e_4 and e_5, we have identified 64 permutations in which the failure of the network occurs upon the third edge failure. Finally, we note that there are precisely 8 permutations in which edge e_3 is the third to fail and causes the network to fail. These permutations are

1, 5, 3, _, _; 5, 1, 3, _, _; 2, 4, 3, _, _; 4, 2, 3, _, _.

It follows that there are a total of 72 permutations corresponding to the failure of the network upon the third failure. Since there are $5! = 120$ permutations in all, we obtain that $s_3 = 72/120 = 3/5$. Since it is clear that the bridge network cannot function with just one working edge, we know that $s_5 = 0$ and thus deduce that $s_4 = 1/5$. In summary, the signature of the 5-component bridge network above is $\mathbf{s} = (0, 1/5, 3/5, 1/5, 0)$.

As shown in Samaniego [61], the survival function of a system's lifetime T can be written as a simple function of $\mathbf{s}$ and F. This representation is shown in equation (3.1) in Chapter 3. The same representation is transparently applicable to a communication network with signature $\mathbf{s}$. When focusing on the reliability of the network at a fixed time t_0, where $P(X_j > t_0) = p$ for all j, this representation reduces to the reliability polynomial in "pq-form," that is, in the form

$$h(p) = \sum_{j=1}^{n} \left(\sum_{i=n-j+1}^{n} s_i \right) \binom{n}{j} p^j q^{n-j} . \tag{6.3}$$

Boland [16] noted that the "tail probabilities" of the signature vector **s** may be viewed as the proportion of path sets of order j. It is precisely those sets, among the collection of $\binom{n}{j}$ sets with exactly j working components, that each contribute the positive probability $p^j q^{n-j}$ to the reliability polynomial in (6.3). Letting a_j stand for the proportion of path sets among the $\binom{n}{j}$ distinct sets of j working components (with the complementary components non-working), we may re-express the reliability polynomial in pq-form as

$$h(p) = \sum_{j=1}^{n} a_j \binom{n}{j} p^j q^{n-j} . \tag{6.4}$$

The relationship between the vectors **a** and **s** was made explicit in section 5.2. For the reader's convenience, we record this relationship anew:

$$a_j = \sum_{i=n-j+1}^{n} s_i \qquad \text{for } j = 1, \ldots, n , \tag{6.5}$$

or equivalently,

$$s_j = a_{n-j+1} - a_{n-j} \qquad \text{for } j = 1, \ldots, n , \tag{6.6}$$

where $a_0 \equiv 0$. We can put the relationship between **a** and **s** in a particularly useful form by writing $\mathbf{a} = \mathbf{P}\mathbf{s}$ and $\mathbf{s} = \mathbf{P}^{-1}\mathbf{a}$, where P and P^{-1} are the $n \times n$ matrices

$$\mathbf{P} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & \cdots & 1 & 1 & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & 1 & \cdots & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 1 & 1 & 1 \end{pmatrix} \tag{6.7}$$

and

$$\mathbf{P}^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 \end{pmatrix} . \tag{6.8}$$

As has hopefully become apparent from the preceding chapters, signature vectors are rich in interpretation; they stand to be particularly useful in the comparison of competing networks. The preservation and characterization results of sections 4.2 and 4.4, which hold equally for the signatures of networks and of coherent or mixed systems, are examples of the utility of signatures in the comparisons of systems or of networks. We now proceed to the main agenda of this chapter, the identification of the exact relationship between the vector of dominations $\mathbf{d}$ and the signature vector $\mathbf{s}$. This is accomplished in Section 6.3. Because domination theory provides essential tools for computing the reliability of a network, and signatures have high interpretive value regarding network performance, the exact functional relationship $\mathbf{s} = f(\mathbf{d})$ developed below enables one to exploit the benefits of both. Our closing example, which compares the reliability of two complex networks, illustrates the utility of this linkage quite effectively. Before establishing the linkage between signature and domination vectors, we briefly digress to present the fundamental ideas and jargon of domination theory.

6.2 A Brief Look at Domination Theory

We now discuss a very useful shortcut for dealing with calculations calling for the inclusion-exclusion formula. The approach is due to Satyanarayana and his coworkers, and is generally known under the rubric "domination theory." The idea first appeared in Satyanarayana and Prabhakar [63]. The notion of dominations was discovered in the process of seeking a reduction in the complexity of the inclusion-exclusion formula for calculating the probability that all components are functioning in at least one of a given network's minimal path sets. The inclusion-exclusion rule (see Chapter 2 for a statement, with discussion) applies to the union of any n sets. When applied to a reliability computation, each of the intersections appearing in the inclusion-exclusion formula represents the collection of components appearing in one or more events of the intersection in question, and each term in the expansion of $P\left(\cup_{i=1}^{n} A_i\right)$ contributes elements of the form p^k or $-p^k$ to the reliability polynomial. (The example below will serve to explain and clarify this seemingly counterintuitive description.) Summing over all r-fold intersections for $r = 1, \ldots, n$, and grouping elements of the same order, one obtains the expression in (6.1). While the inclusion-exclusion formula provides an explicit expression for the reliability polynomial, it can entail substantial computational complexity. The generation of the (say m) minimal path sets of a given system involves an algorithm that is exponential in m, and the number of different intersections of m sets is also exponential in m. What results is a doubly exponential algorithm for computing system reliability. From this "inconvenient truth," and the need for something simpler, domination theory was born.

Let $A_1, A_2, \ldots, A_m$ be a list of all minimal path sets of a given network of order n (that is, with n edges). Let us view a given path set as a set of working components (with the property that their working ensures the working of the network and no proper subset has that characteristic). Note that the "intersection" of any k sets among the m minimal path sets of the network may then be thought of as the event that every component that is a member of at least one of these k minimal path sets is working. It follows that the probability of an intersection of events representing functioning minimal path sets can be seen to be equivalent to the probability that every one of these components work. The union of the components in a fixed collection of minimal path sets is called a *formation*. Further, an *i-formation* is defined as a union of the components in a collection of i minimal path sets. For example, the union $\{1, 2, 3\}$ of the minimal path sets $\{1, 2\}$ and $\{2, 3\}$ is an example of a 2-formation. We will refer to a particular formation as *even* (*odd*) if it is the union of an even (odd) number of minimal path sets. A particular formation can be both odd and even. For instance, in the network used in the example below, we'll see that $\{1, 2, 3, 4, 5\}$ is simultaneously a 2-, 3- and 4-formation.

One may view domination theory as an accounting mechanism that helps one to keep track of the basic elements of the inclusion-exclusion calculation. An even formation occurs with each k-fold intersection in (2.20) for which k is even, while an odd formation results when k is odd. The signs associated with formations are drawn directly from (2.20), since the intersection of k events contributes a negative term to the sum precisely when k is even. A network having m minimal path sets has a total of $2^m - 1$ formations. We illustrate these ideas for the Wheatstone bridge network shown in Figure 6.1. We will calculate the reliability of the two-terminal Wheatstone bridge (where A and D are the terminals of interest) given that each edge functions independently with probability p. The minimal path sets of this two-terminal network are the sets of edges $\{1, 4\}$, $\{2, 5\}$, $\{1, 3, 5\}$, and $\{2, 3, 4\}$. The *signed domination* of a given union of minimal path sets is simply the difference between the number of even and odd formations for that union. As an example of this calculation, note that, if we denote the four minimal path sets above as A_1, A_2, A_3 and A_4, then the formations associated with the set $\{1, 2, 3, 4, 5\}$ are the 4 odd formations (namely, the four unions of the sets $\{A_1, A_2, A_3\}$, $\{A_1, A_2, A_4\}$, $\{A_1, A_3, A_4\}$ and $\{A_2, A_3, A_4\}$) and the 2 even formations (namely, the two unions of the sets $\{A_3, A_4\}$ and $\{A_1, A_2, A_3, A_4\}$). The relevant accounting is tabulated below.

Table 6.1. Signed dominations for the Wheatstone bridge network

Unions of Min Path Sets	Number of Odd Formations	Number of Even Formations	Signed Domination
{1, 4}	1	0	1
{2, 5}	1	0	1
{1, 3, 5}	1	0	1
{2, 3, 4}	1	0	1
{1, 2, 3, 4}	0	1	-1
{1, 2, 3, 5}	0	1	-1
{1, 2, 4, 5}	0	1	-1
{1, 3, 4, 5}	0	1	-1
{2, 3, 4, 5}	0	1	-1
{1, 2, 3, 4, 5}	4	2	2

Suppose that a particular union A of minimal path sets consists of exactly r edges. Then the marginal probability that this set "works" is p^r and the sum of the coefficients (1's and -1's) in the terms of the inclusion-exclusion formula corresponding to the occurrences of the probability $P(A)$ in the expansion (2.20) is precisely the signed domination of A. This term contributes the product of the signed domination of A and the probability element p^r to the reliability polynomial. Given this type of accounting, the signed dominations of all sets of size r can be summed and then multiplied by p^r. Summing over $r = 1, \ldots, n$ yields the reliability polynomial. For the bridge network above, the reliability polynomial is thus computed as

$$h(p) = 2p^2 + 2p^3 - 5p^4 + 2p^5 \; . \tag{6.9}$$

In the notation of (2.23), we have identified the coefficient vector $\mathbf{d}$ of the reliability polynomial as $(0, 2, 2, -5, 2)$ for the Wheatstone bridge network in Figure 6.1. We will refer to $\mathbf{d}$ as the vector of dominations. The general process described above applies equally well to the computation of the reliability function for networks with independent but non-identical components, though the final expression, being a multilinear function of $p_1, \ldots, p_n$, will be both more complex and more cumbersome. We will not deal with this generalization further in the developments in this chapter.

6.3 The Linkage Between Dominations and Signatures

The goal of this section is to identify the functional relationship between the signature vector $\mathbf{s}$ and the vector $\mathbf{d}$ of dominations. As we shall see, it is quite easy to write $\mathbf{d}$ as a function of $\mathbf{s}$, say $\mathbf{d} = f(\mathbf{s})$, but since the domination vector is generally easier to compute than the signature vector, one would ordinarily expect to have $\mathbf{d}$ in hand first. Since signatures have greater interpretive value, one would then typically seek to obtain $\mathbf{s}$ from $\mathbf{d}$. Thus, it will be the functional relationship $\mathbf{s} = f^{-1}(\mathbf{d})$ that is our real goal. We begin by identifying the function f above. Recall that we have two different but equivalent representations of the reliability polynomial, namely,

$$h(p) = \sum_{r=1}^{n} d_r p^r \tag{6.10}$$

from (6.1) and

$$h(p) = \sum_{j=1}^{n} a_j \binom{n}{j} p^j q^{n-j} \ . \tag{6.11}$$

from (6.4), where $a_j = \sum_{i=n-j+1}^{n} s_i$ for $j = 1, \ldots, n$. Writing q^{n-j} in (6.11) as $(1-p)^{n-j}$ and expanding the term via the binomial theorem, one may rewrite the polynomial in (6.11) as

$$\begin{aligned} h(p) &= \sum_{j=1}^{n} a_j \binom{n}{j} p^j \left(\sum_{i=0}^{n-j} \binom{n-j}{i} (-1)^i p^i \right) \\ &= \sum_{r=1}^{n} \left(\sum_{j=1}^{r} a_j \binom{n}{j} \binom{n-j}{r-j} (-1)^{r-j} \right) p^r \ , \end{aligned} \tag{6.12}$$

the latter sum resulting from the change of indexes from (j, i) to (r, j), where $r = i + j$. From (6.10) and (6.12), it is apparent that

$$d_r = \sum_{j=1}^{r} a_j \binom{n}{j} \binom{n-j}{r-j} (-1)^{r-j} \quad \text{for } r = 1, \ldots, n \ . \tag{6.13}$$

It follows from (6.5) that the functional relationship $\mathbf{d} = f(\mathbf{s})$ is given by

$$d_r = \sum_{j=1}^{r} \left(\sum_{i=n-j+1}^{n} s_i \right) \binom{n}{j} \binom{n-j}{r-j} (-1)^{r-j} \quad \text{for } r = 1, \ldots, n \ . \tag{6.14}$$

Since $\mathbf{d}$ and $\mathbf{a}$ are linearly related via (6.13), their relationship may be expressed as $\mathbf{d} = \mathbf{M}\mathbf{a}$, where $\mathbf{M}$ is the matrix specified in (6.15) below.

$$\mathbf{M} = \begin{pmatrix} \binom{n}{1}\binom{n-1}{0} & 0 & 0 & \cdots & 0 \\ -\binom{n}{1}\binom{n-1}{1} & \binom{n}{2}\binom{n-2}{0} & 0 & \cdots & 0 \\ \binom{n}{1}\binom{n-1}{2} & -\binom{n}{2}\binom{n-2}{1} & \binom{n}{3}\binom{n-3}{0} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \pm\binom{n}{1}\binom{n-1}{n-1} & \mp\binom{n}{2}\binom{n-2}{n-2} & \pm\binom{n}{3}\binom{n-3}{n-3} & \cdots & \binom{n}{n}\binom{n-n}{n-n} \end{pmatrix} \tag{6.15}$$

Moreover, we have that $\mathbf{a}$ and $\mathbf{s}$ are linearly related, that is, $\mathbf{a} = \mathbf{Ps}$, where $\mathbf{P}$ is the matrix displayed in (6.7). Since $\mathbf{d} = \mathbf{MPs}$, the relationship of interest to us is $\mathbf{s} = \mathbf{P}^{-1}\mathbf{M}^{-1}\mathbf{d}$, where the matrix $\mathbf{M}$ is given in (6.15). The inverse of $\mathbf{P}$ was identified in (6.8). What remains is to invert the matrix $\mathbf{M}$ above and make the relationship $\mathbf{s} = f^{-1}(\mathbf{d})$ explicit. Our expression for $\mathbf{M}^{-1}$ will require the use of notation for the number of permutations of k items taken j at a time. We will use the symbol $(k)_j$ for this number, where it is understood that $0 \le j \le k$. Rather than compute $\mathbf{M}^{-1}$ directly using the usual inversion methodology (co-factors, parity, determinants, etc.), we will simply claim that $\mathbf{M}^{-1}$ is the matrix $\mathbf{M}^*$ whose ith row is given by

$$(m^*_{i1}, \cdots, m^*_{ii}, 0, \cdots, 0) = \left(\underbrace{\frac{(i)_1}{(n)_1}, \frac{(i)_2}{(n)_2}, \cdots, \frac{(i)_i}{(n)_i}}_{i \text{ slots}}, \underbrace{0, \cdots, 0}_{n-i \text{ slots}} \right) . \tag{6.16}$$

This claim will be proven by showing that

$$(\mathbf{M}^*)(\mathbf{M}) = I , \tag{6.17}$$

where I is the $n \times n$ identity matrix. Students (and even seasoned researchers) reading this demonstration might make special note of this unusual method of proof and make a point of adding it to their toolbox. For clarity's sake, I will make the method of proof especially explicit: what you do is *guess* the answer to your problem, and then verify that your answer is correct. Cool, huh? While not an elegant method, when it works, you really can't beat it. The good news is that it works here. We now show that the inner product π_{ij} of the ith row of $\mathbf{M}^*$ and the jth column of $\mathbf{M}$ (that is, $\pi_{ij} = \sum_{k=1}^{n} \mathbf{M}^*_{ik}\mathbf{M}_{kj}$) is 1 if $i = j$ and is 0 otherwise. We will consider separately the following three cases (a) $i < j$, (b) $i > j$ and (c) $i = j$. Clearly, (6.17) holds in case (a), as every element of the inner product π_{ij} is a product of two numbers, one of which is 0. For case (b), let $m = i - j$; we then have

$$\begin{aligned}
\pi_{ij} &= \binom{n}{i-m} \sum_{k=0}^{m} (-1)^k \frac{(i)_{i-m+k}}{(n)_{i-m+k}} \binom{n-i+m}{k} \\
&= \binom{n}{i-m} \sum_{k=0}^{m} (-1)^k \frac{i \times (i-1) \times \cdots \times (m-k+1)}{n \times (n-1) \times \cdots \times (n-i+m-k+1)} \\
&\qquad \times \frac{(n-i+m)!}{k!(n-i+m-k)!} \\
&= \binom{n}{i-m} \frac{(n-i+m)!i!}{n!m!} \sum_{k=0}^{m} (-1)^k \frac{m!}{k!(m-k)!} \\
&= \binom{i}{m} \sum_{k=0}^{m} (-1)^k \binom{m}{k} \\
&= 0 ,
\end{aligned}$$

since the sum in the penultimate term above is simply the binomial expansion of $(1-1)^m$. The arguments above show that all of the off-diagonal elements of $(\mathbf{M}^*)(\mathbf{M})$ are 0. We now turn to the case in which $i = j$. In this case, we have that π_{ii} is given by

$$\frac{(i)_i}{(n)_i} \times \binom{n}{i}\binom{n-i}{0} = \frac{i!}{n \times \cdots \times (n-i+1)} \frac{n!}{i!(n-i)!} = 1 \,.$$

This completes our proof of (6.17). ■

Given the explicit expressions for $\mathbf{P}^{-1}$ in (6.8) and $\mathbf{M}^{-1}$ in (6.16), one may obtain the ith row of the matrix product $\mathbf{M}^{-1}\mathbf{P}^{-1}$, for $i = 1, \ldots, n$, as

$$\Big(\underbrace{-m^*_{n-i,1}+m^*_{n-i+1,1},\cdots,-m^*_{n-i,n-i}+m^*_{n-i+1,n-i},m^*_{n-i+1,n-i+1}}_{n-i+1 \text{ slots}},\ \underbrace{0,\cdots,0}_{i-1 \text{ slots}}\Big)$$

where $m^*_{ij} = (i)_j/(n)_j$. Since $\mathbf{s} = \mathbf{P}^{-1}\mathbf{M}^{-1}\mathbf{d}$, we may now identify the relationship $\mathbf{s} = f^{-1}(\mathbf{d})$ in the following explicit form.

Theorem 6.1. *Let* $\mathbf{d}$ *and* $\mathbf{s}$ *denote the domination and signature vectors for a given network of order* n*. Then for* $i = 1, \ldots, n$*, we have*

$$s_i = \sum_{j=1}^{n-i}(-m^*_{n-i,j} + m^*_{n-i+1,j})d_j + m^*_{n-i+1,n-i+1}d_{n-i+1} \,. \tag{6.18}$$

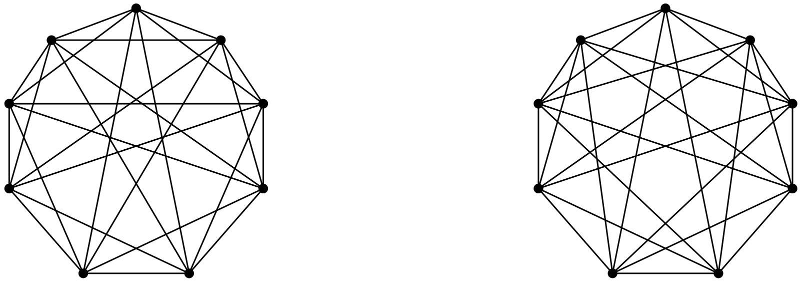
That is, for $i = 1, \ldots, n$*, we have*

$$s_i = \sum_{j=1}^{n-i} \frac{-(n-i)_j + (n-i+1)_j}{(n)_j} d_j + \frac{(n-i+1)_{n-i+1}}{(n)_{n-i+1}} d_{n-i+1} \,. \tag{6.19}$$

In Section 6.1, we mentioned that the comparison of networks via their domination vectors was unintuitive and, for complex networks, likely to be unproductive. The reason for this is that for two complex networks, the difference of the reliability polynomials, that is, $\sum(d_{2r} - d_{1r})p^r$, will typically be of quite high degree. Because of the requirement $\sum d_r = 1$ on the domination vector of an arbitrary network, the coefficients of the difference polynomial will have some terms with opposite signs. Thus, determining whether one reliability polynomial is uniformly larger than another for all $0 < p < 1$ is a task equivalent to finding the roots of a high degree polynomial. But Galois Theory shows that the problem of finding roots of polynomials of degree greater than 4 is not "solvable by radicals." One cannot, in general, obtain closed-form expressions for the solutions to such problems.

The ability to transform the problem from one involving the domination vector to one involving the signature vector changes things substantially. To see this more graphically, let us consider the comparison between the two $G(9, 27)$ networks pictured below.

Fig. 6.2. Networks G_1 and G_2



The reliability polynomials for the connectivity of the 9 vertexes of these two networks are displayed below in standard form.

$$\begin{aligned} h_{G_1}(p) = {} & 419904p^{27} - 6021144p^{26} + 41705280p^{25} - 18489826p^{24} \\ & +586821717p^{23} - 1413876060p^{25} + 2677774329p^{21} \\ & -4074363810p^{20} + 5048856414p^{19} - 5135792742p^{18} \\ & +4303029693p^{17} - 2967712776p^{16} + 1676975886p^{15} \\ & -769265910p^{14} - 282176568p^{13} + 80853282p^{12} \\ & +17445456p^{11} - 2667060p^{10} + 257634p^{9} - 11828p^{8} \end{aligned}$$

$$\begin{aligned} h_{G_2}(p) = {} & 414720p^{27} - 5934288p^{26} + 41015964p^{25} - 181453380p^{24} \\ & +574666025p^{23} - 1381692972p^{22} + 2611463517p^{21} \\ & -3965536554p^{20} - 4904464002p^{19} + 4979513718p^{18} \\ & +4164454729p^{17} - 2867022480p^{16} + 1617256842p^{15} \\ & -740601350p^{14} - 271201476p^{13} + 77576922p^{12} \\ & +16709916p^{11} - 2550156p^{10} + 245898p^{9} - 11268p^{8} \end{aligned}$$

To the naked eye, the two networks pictured above are difficult to distinguish, and little intuition can be brought to bear on the question of which might offer better performance. Further, it is clear from an inspection of the two polynomials above that a comparison of them is not an inviting proposition. It is by no means obvious that one polynomial is uniformly larger than the other for all $p \in (0, 1)$. The comparison could be attacked by brute (numerical and computational) force, and one would, in the end, conclude that

$h_{G_1}(p) \geq h_{G_2}(p)$ for $0 < p < 1$. However, the question of "why?" would remain a mystery. Let us examine the same comparison using the signatures of the two networks. In Table 6.2, the tail probabilities $S_j(x) = \sum_{i=x}^{27} s_{ji}$ have been calculated for each of the networks G_1 and G_2. From the second and third columns of Table 6.2, we see that $\mathbf{s}_{G_1} \geq_{\text{st}} \mathbf{s}_{G_2}$, an inequality that immediately implies that $h_{G_1}(p) \geq h_{G_2}(p)$ for $p \in (0, 1)$. Further, we see that the superiority of network G_1 over G_2 is due to the superiority of one design over the other as reflected by the comparisons of their signatures. This seems sufficiently rewarding to justify, by itself, the treatment of this comparison through the use of signatures. But as Sesame Street's Count Von Count likes to say while counting bats in the belfry: "But wait, there's more!" The ratios of the two signatures' survival functions displayed in the last column of Table 6.2 shows that $\mathbf{s}_{G_1} \geq_{\text{hr}} \mathbf{s}_{G_2}$, a conclusion that is not possible to obtain from an analysis of the polynomials $h_{G_1}(p)$ and $h_{G_2}(p)$ alone. The latter inequality implies that, when viewed as a continuous process, the lifetime of the network G_1 dominates that of the network G_2 in the hazard rate ordering. This additional fact establishes that G_1 is not only better than the network G_2, it's actually better in quite a strong sense. While far from obvious by inspection, G_1 rules!

The main motivation for dealing with signatures in the context of network reliability is the fact that domination theory, while useful in making network reliability calculations, is of limited value in comparing the performance characteristics of competing networks. On the other hand, the examination of the signatures of competing networks can be quite revealing. Theorem 6.1 makes the relationship between dominations and signatures explicit and does so in a particularly useful form. Clearly, both tools have a role to play. Domination theory remains the go-to approach for computing the reliability of a network in i.i.d. edges. With dominations in hand, one may readily identify the network's signatures via Theorem 6.1. This explicit connection allows one to exploit in tandem the computational advantages of domination theory and the interpretive advantages of signatures when making comparisons among networks. The literature on the existence, uniqueness and identification of uniformly optimal networks of a given size (see, for example, Ath and Sobel [4], Boesch, Li and Suffel [14], Myrvold, Cheung, Page and Perry [57] and Wang [74]) is still quite incomplete, with a good many open problems remaining to be solved. The coordinated application of dominations and signatures in such problems should facilitate progress in this area.

Table 6.2. Signature Tail Probabilities $S(x) = \sum_{i=x}^{27} s_i$ and Their Ratios

x	$S_{G_1}(x)$	$S_{G_2}(x)$	$S_{G_1}(x)/S_{G_2}(x)$
1	1.0	1.0	1.0
2	1.0	1.0	1.0
3	1.0	1.0	1.0
4	1.0	1.0	1.0
5	1.0	1.0	1.0
6	1.0	1.0	1.0
7	0.999970	0.999970	1.0
8	0.999787	0.999787	1.0
9	0.999149	0.999149	1.0
10	0.997367	0.997367	1.0
11	0.993612	0.993612	1.0
12	0.985922	0.985922	1.0
13	0.971744	0.971743	1.0000005
14	0.947220	0.947214	1.0000063
15	0.906907	0.906867	1.0000442
16	0.843421	0.843240	1.0002148
17	0.747317	0.746717	1.0008024
18	0.607883	0.606416	1.0024183
19	0.417560	0.415077	1.0059834
20	0.189140	0.186804	1.0125000
21	0.0	0.0	—
22	0.0	0.0	—
23	0.0	0.0	—
24	0.0	0.0	—
25	0.0	0.0	—
26	0.0	0.0	—
27	0.0	0.0	—

7

Applications of Signatures in Reliability Economics

7.1 Prototypical Problems in Reliability Economics

It is difficult to give a succinct description of the field of Reliability Economics. At present, the field might be thought of as a somewhat scattered collection of results in which a decision is made or an action is taken with a view toward balancing the natural tension that exists between the performance of a system and its cost. In assessing the performance characteristics of a particular system design, for instance, one's goal would be to identify a system whose performance is good but whose cost is modest. Three imposing challenges immediately arise in seeking to address a problem in Reliability Economics analytically. One must begin by identifying a reasonable way of quantifying the performance of each system under consideration (perhaps all possible systems of a given size). One must also quantify the cost of each of these systems. Recall that the number of system designs under consideration may be huge. The third challenge is that of determining some reasonable criterion for comparing the systems of interest. These challenges are formidable because of the need to accurately assess each system's potential performance and actual cost, a process that can require a substantial amount of expertise. Often, they are imposing because of the sheer size of the problem. Successfully meeting the challenges above will invariably involve the development of a problem formulation that is both conceptually sound and analytically tractable. In what follows, we will devote considerable attention to the presentation and justification of a particular formulation of the problem of finding the optimal system design relative to a specific family of criterion functions that take performance and cost into account.

It is apparent from the literature on modeling and inference in reliability that economic concerns have not garnered their due attention. Most published studies in the field concentrate on performance issues. On the several occasions that I've had the opportunity to lecture on the subject of Reliability Economics, I've motivated the interest we should have in the economic questions that arise in our field by retelling a story my father told me when I was young. It seems that there was a boy in my father's home town who was roundly regarded as mentally challenged. The town's elders would gather in the town square and laugh about the boy's quirks. One prominent one was this. Whenever the boy was offered his choice between a nickel and a quarter, he would immediately take the nickel. This behavior was seen as very amusing. Anytime there was a guest in town, the boy would be summoned to the town square and the guest would be prompted to try this experiment first hand. Sure enough, the boy would take the nickel and be on his way. Unbeknownst to the cackling folks in the square, the boy would saunter home, go to his room and toss the nickel into an enormous jar full of nickels. The boy knew that the choice he had made, long before the current episode, was the choice between one quarter and a jar full of nickels!

Economic concerns are so natural in the area of reliability that it may seem surprising that the analytical treatment of problems in Reliability Economics is, at present, in a relatively primitive state. General developments in Reliability Economics are, as of this writing, quite sparse. (One notable exception is the area of "warranty analysis"). This state of affairs is all the more curious given the fact that circumstances in which one would be willing to ignore the economics associated with a reliability problem are extremely rare. Except for those cases in which either "cost is no object" or one can only afford the cheapest available system and must accept whatever performance comes with it, one is naturally inclined to try to optimize relative to one's investment. Informally, one can think of the aim of Reliability Economics as that of "getting the most bang for the buck." As a result, problems in the area of Reliability Economics arise frequently and broadly. Many of us make use of the publication **Consumer Reports** to assist us in making rational decisions in balancing performance and cost in the things we buy. The field extends as well to the U.S. military acquisitions program where the goal is to develop a system designed to meet certain performance and suitability goals while respecting the applicable budgetary constraints. In this chapter, we will focus on the problem of optimal system design. The goal is to identify a system that strikes an appropriate balance between one's positive expectations regarding its reliability and one's concerns or constraints regarding cost. The results presented here are largely drawn from Dugas and Samaniego [33].

It is useful to examine why problems in Reliability Economics have heretofore tended to resist clean analytical solutions. Let us for a moment restrict our attention to the search for an "optimal" coherent system of order n. While the entire collection of such systems is easy to enumerate when n is small (as is done, for example, for $n = 3$ and $n = 4$ in Chapters 2 and 3), we know that the number of distinct coherent systems of order n grows exponentially with n and is quite overwhelming even for relatively modest values of n. It is therefore apparent that the problem of finding the best coherent system of a given order is, typically, a discrete optimization problem in which the space to be searched is huge. A second obstacle to the analytical treatment of this problem is the fact that there has been no obvious, manageable index with respect to which one might optimize. Without such a tool, one must attack the optimization problem by obtaining the cost and expected performance of each available system and then comparing the value of the criterion function across such systems. Not only is the exercise a complex one system by system (since, for example, estimating a given system's cost can by itself be a challenging problem), the exercise must be repeated for every system under consideration. When all coherent systems of a given size are being compared, an exact cost-benefit analysis of all these systems will almost always be infeasible. We should mention one seemingly appropriate index for coherent systems in Reliability Economics problems, namely the system structure functions discussed in Chapter 2. Some of the limitations of this potential index have already been mentioned. Structure functions tend to be awkward to compute and each has many equivalent forms. Further, structure functions have the same cardinality as the class of coherent systems themselves (so their exhaustive computation is, at best, exhausting) and they have no clear connection to a system's cost. Thus, the assessment of system cost would still need to be dealt with system by system.

These two difficulties – the large discrete space over which the optimization is to take place and the absence of an easily managed index for the optimization – have led to the reliance on searching techniques for seeking good (that is, nearly optimal) solutions as efficiently as possible. The literature on general search methods for optimization purposes is vast. In the Reliability Economics framework discussed above, where discrete search methods have been required, genetic algorithms appear to be the favored approach in the recent literature. See the monograph by Kuo, Prasad, Tillman and Hwang [53] and their references for a discussion of the algorithmic approach to constrained optimization problems in reliability. Concrete examples of the use of genetic algorithms in reliability economics problems include papers by Deeter and Smith [29] and by Usher, Kamal and Sayed [71].

Both of the difficulties mentioned above are directly addressed in our treatment of the problems described in this chapter, one, perhaps unexpectedly, by making the space of systems to be considered even larger and the other by indexing that space in a new way. In Chapter 3, we defined the signature vector $\mathbf{s}$ of a system in n i.i.d. components. In this chapter, we will take the signature vector as an index of the system of interest. Implicit in this selection is the fact that we are, again, restricting attention to systems whose components have i.i.d. lifetimes with common distribution F. While signatures were initially defined for coherent systems of a given order, limiting their use as an index for that specific collection merely addresses the same indexing problem alluded to above. One still has to deal with an optimization problem over a potentially huge discrete space. A further difficulty is that, even though the dimension of the index is bounded (in fact, is equal to the order of the systems being compared), there are still a great many of them, and their individual computation can be combinatorially quite complex.

For both practical and analytical reasons, the expansion of the space of signatures beyond those of coherent systems will pay some positive dividends. We will, in fact, consider the problem of finding the optimal system in the Reliability Economics context studied in the sequel by maximizing a chosen criterion function over the class of all mixed systems of order n. As discussed in Chapter 3, a mixed system with signature $\mathbf{s}$ can be physically realized through a randomization process that selects a k-out-of-n system with probability s_k, with k ranging from 1 to n. Since the class of mixed systems includes all coherent systems as special (degenerate) cases, nothing is lost by considering this larger space. And while the space of signatures of mixed systems based on coherent systems of order n in i.i.d. components is uncountably infinite (leading us to replace a discrete search by a search over a much larger space), the larger problem is much more amenable to analytical treatment, as the tools of differential calculus become available in the maximization problem of interest. The optimization problem to which we now turn involves the maximization of a chosen criterion function over the $(n-1)$-dimensional simplex $\{\mathbf{s} \in [0,1]^n \mid \sum_{i=1}^n s_i = 1\}$. As will be seen in the sequel, we will encounter problems in which the optimal signature is in fact a coherent system (i.e., a degenerate mixed system), so that the artifact of randomization serves there only as a tool for finding the best coherent system analytically. Interestingly, we will also encounter problems in which a non-degenerate mixed system proves to be optimal, showing that the strategy of randomizing among a collection of coherent systems is more than an analytical device. It can in fact lead to the identification of a mixed system that improves upon any individual coherent system.

7.2 Optimality Criteria

As mentioned above, the formulation of any Reliability Economics problem will invariably include the specification of a criterion function that quantifies the precise manner in which the performance and cost of the system (or policy) of interest will be weighed relative to each other. There is no question that the choice of criterion function is, at least in part, a subjective matter and that its choice over some alternative function will have some influence upon the outcome of the optimization process. Whatever criterion function is chosen, it is reasonable to require it to possess two basic properties that seem essential in Reliability Economics applications: the criterion function should vary proportionately with measures of system performance and inversely with measures of system cost. These properties are equivalent, for example, to the position that, given two systems with the same performance characteristics, we would prefer the one which costs less, and given two systems which cost the same, we would prefer the better-performing system.

The class of criterion functions on which we will focus is general enough to admit a wide variety of interpretations. As we shall see, however, the subclass we study in detail, and propose for practical use, does indeed have the two "essential" properties alluded to above. In treating system performance, two particular measures seem the most natural: the expected lifetime ET of the system and the system's reliability function $R_T(t) = P(T > t)$. If the system has been designed to survive beyond a predetermined mission time t_0, then the reliability $R_T(t_0)$ at this mission time might be taken as the appropriate measure of performance. Let us now recall that, for a mixed system of order n based on components with i.i.d. lifetimes, both $R_T(t)$ and ET can be expressed in terms of the system's signature and the particular properties of the component lifetime distribution F. To be more specific, we have from (3.5) and (3.7) that

$$R_T(t) = P(T > t) = \sum_{i=1}^{n} s_i P(X_{i:n} > t) \tag{7.1}$$

and

$$ET = \sum_{i=1}^{n} s_i EX_{i:n} \; . \tag{7.2}$$

Both of these measures have convenient forms, i.e., are linear combinations of the elements of the signature vector, and both have natural interpretations in terms of performance.

Similarly, one could model the expected cost of a system as a different linear combination of the elements of $\mathbf{s}$, that is, as

$$EC = \sum_{i=1}^{n} c_i s_i \; . \tag{7.3}$$

One example where such a representation of expected cost arises is the "salvage model" in which there are three fundamental components of cost: C_I, the initial, fixed cost of manufacturing the systems of interest, the cost A of an individual component and the salvage value B of a used but working component removed after system failure. Assuming that these elements determine a system's cost, then the expected cost of the system may be written as

$$EC = \sum_{i=1}^{n} (C_I + n(A - B) + iB) s_i \; . \tag{7.4}$$

The criterion function which we use in the sequel is somewhat more general than simply the ratio of the measures of performance in (7.1) or (7.2) to the measures of cost in (7.3) or (7.4). Such ratios serve as notable special cases of the class of criterion functions on which we will focus. Specifically, the measure to be used of the relative value of performance and cost has the general form

$$m_r(\mathbf{s},\mathbf{a},\mathbf{c}) = \frac{\sum_{i=1}^n a_i s_i}{\left(\sum_{i=1}^n c_i s_i\right)^r} , \tag{7.5}$$

where $r > 0$ and the vectors $\mathbf{a}$ and $\mathbf{c}$ can be chosen arbitrarily within the context of two natural constraints: $0 < a_1 < a_2 < \cdots < a_n$ and $0 < c_1 < c_2 < \cdots < c_n$. These monotonicity constraints are motivated by the intuition that the constants a_i and c_i are measures of performance and cost, respectively, and since they are the coefficients of the signature element $s_i = P(T = X_{i:n})$, both a_i and c_i should increase as the index i of the order statistic fatal to the system increases. Without such a requirement, it would be possible, for example, for a series system to have better performance, on average, and also cost more, than a parallel system (the respective consequences of the inequalities $a_1 > a_n$ and $c_1 > c_n$, were those inequalities permitted to hold). Consequences such as these are counterintuitive. Since we expect that systems that tend to survive through more component failures have better overall performance and also generally cost more, the constraints on $\mathbf{a}$ and $\mathbf{c}$ above are, arguably, both natural and necessary. We note that the monotonicity condition on $\mathbf{a}$ is satisfied for the choices of $\mathbf{a}$ of primary interest, namely $a_i = EX_{i:n}$ for $i = 1, \ldots, n$ or $a_i = P(X_{i:n} > t)$ for $i = 1, \ldots, n$, and the monotonicity conditions on $\mathbf{c}$ seem quite reasonable in general and are satisfied, in particular, for the salvage model. That a better-designed system will tend to perform better than a system with a weaker design, but will also be more expensive, are facts that are both implied by the following result when the monotonicity constraints on the vectors $\mathbf{a}$ and $\mathbf{c}$ hold.

Proposition 7.1. *Let* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *be the signatures of two mixed systems based on* n *components with i.i.d. lifetimes, and suppose that* $\mathbf{s}_1 \neq \mathbf{s}_2$ *and that* $\mathbf{s}_1 \leq_{\text{st}} \mathbf{s}_2$. *If* $\mathbf{b}$ *is an* n*-dimensional vector satisfying* $0 < b_1 < b_2 < \cdots < b_n$, *then it follows that* $\sum_{i=1}^n b_i s_{1i} < \sum_{i=1}^n b_i s_{2i}$.

Proof. Since $\mathbf{s}_1 \leq_{\text{st}} \mathbf{s}_2$, we may write

$$\sum_{i=1}^n b_i s_{1i} = \sum_{i=1}^n (b_i - b_{i-1}) \sum_{j=i}^n s_{1j} < \sum_{i=1}^n (b_i - b_{i-1}) \sum_{j=i}^n s_{2j} = \sum_{i=1}^n b_i s_{2i} ,$$

where we take $b_0 = 0$ by definition. ■

The general form of the criterion function in (7.5) ensures that it will have the desirable property of being an increasing function of performance and a decreasing function of cost. It is clear that if $\mathbf{a}^{(1)} \leq \mathbf{a}^{(2)}$, where this vector inequality is interpreted componentwise, then it follows trivially that $\sum_{i=1}^{n} a_i^{(1)} s_i \leq \sum_{i=1}^{n} a_i^{(2)} s_i$; similarly, if $\mathbf{c}^{(1)} \leq \mathbf{c}^{(2)}$, it follows that $\sum_{i=1}^{n} c_i^{(1)} s_i \leq \sum_{i=1}^{n} c_i^{(2)} s_i$. Thus, for example, if the elements of the "performance vector" **a** undergo an increase resulting from an upgrade in component quality, this will result in an overall increase in our measure $\sum_{i=1}^{n} a_i s_i$ of system performance, and if the elements of the "cost vector" **c** undergo an increase due to inflation, this will result in an overall increase in our measure of the cost $\sum_{i=1}^{n} c_i s_i$ of the system.

The denominator of (7.5) requires further comment. Of course the salvage model serves as motivation for the form of our measure of cost, but a quite different justification is possible and is both more general and more compelling. Let us assume that it is possible, through appropriate consultation with experts in engineering design and budgetometry (to coin a term), to obtain an expert assessment of the cost of constructing a k-out-of-n system. Let c_k be equal to that cost. Now suppose that one wishes to assign a cost to a particular (coherent or mixed) system in n components, with i.i.d. lifetimes $\sim F$, having signature vector **s**. As noted earlier, one can find a mixture of k-out-of-n systems in n components with i.i.d. lifetimes $\sim F$ having precisely the same signature and thus the same performance characteristics (that is, exactly the same lifetime distribution). Thus, the cost of producing a system with signature **s** can reasonably be considered to be the cost of producing the equivalent mixture of k-out-of-n systems. This latter system is produced using the probability vector **s** as the mixing distribution for k-out-of-n systems, and its expected cost would be exactly $\sum_{i=1}^{n} c_i s_i$. Since the performance of any system in i.i.d. components can be replicated by a mixed system with a cost of the form $\sum_{i=1}^{n} c_i s_i$, the use of this linear function of the elements of the signature vector in the denominator of our criterion function m is quite appropriate. Over repeated applications of the mixed system above, its average cost would converge to $\sum_{i=1}^{n} c_i s_i$.

Finally, something should be said about the positive parameter r in (7.5). The exponent r in the denominator of the criterion function serves as a calibration parameter. In some applications, one will wish to weigh performance and cost differently, and the value of r can be adjusted to accommodate the application at hand. While setting r equal to 1 seems like the most natural choice, rendering the criterion roughly interpretable as "performance per unit cost," other values of r may be required in particular applications. If controlling costs is of special importance, one might choose a value of r greater than 1, as this would accentuate the impact of the system's cost. On the other hand, if high performance is seen as essential, a value of r less than one might be called for. The choice of r will vary with the application, and should be thought of as a parameter to be selected, upon careful consideration, by the client who would be using the chosen system.

The class of criterion functions to which we restrict attention here is by no means uniquely appropriate to the problem of interest. This choice might be thought of as carrying the same import as the selection of, say, squared error loss in a statistical estimation problem. In both scenarios, it is reasonable to investigate alternative criteria. In the present problem, one might wish to maximize $E(T/C)$ instead of ET/EC. Further, one might consider different functional forms for the criterion m or at least introduce some nonlinearity in the relationships between performance and/or cost and system signatures. Criterion functions having the form in (7.5) are chosen here for two good reasons, their desirable monotonicity properties and their amenability to analytical treatment. The exploration of alternative criteria with respect to which one might seek to resolve problems in Reliability Economics seems, nonetheless, to be both warranted and desirable.

7.3 Characterizing Optimal Systems

We reiterate the tacit assumption made in signature-related analyses that the systems under consideration have components whose lifetimes are i.i.d. according to a common distribution F. Since the optimality results of this section rely crucially on the choice of criterion function, and the criterion function in (7.5) is an explicit function of the system signature, applications should be guided by careful judgments regarding the relevance of these assumptions. Under the criterion function in (7.5), we will now characterize optimal systems for arbitrary but fixed r. We treat the cases $r = 1$ and $r \neq 1$ separately.

Consider first the case in which $r = 1$. Our search for an optimal system will begin with an examination of the impact on the criterion function of a small shift in mass from one element to another in a given signature vector. Let $\mathbf{s}$ be the signature of a mixed system based on n i.i.d. components, and suppose that the ith and jth elements of $\mathbf{s}$ are positive. Assume, without loss of generality, that $1 \leq i < j \leq n$. Let $\mathbf{s}^*$ be a signature whose elements are identical to those of $\mathbf{s}$ except for its ith and jth elements, with $s_i^* = s_i + \varepsilon$ and $s_j^* = s_j - \varepsilon$, where ε obeys the constraint $-s_i \leq \varepsilon \leq s_j$. If $\varepsilon > 0$, this change represents a transfer of mass from s_j to s_i, while if $\varepsilon < 0$, the transfer of mass is in the other direction. It is convenient, at this point, to subsume the dependence of the criterion function m on all parameters save the signature, treating the remaining parameters (that is, $\mathbf{a}$ and $\mathbf{c}$) as fixed. The parameter r is, for now, set equal to 1. We may then write the inequality of interest as

$$m(\mathbf{s}^*) \geq m(\mathbf{s}) . \tag{7.6}$$

It is easy to verify that for $\varepsilon > 0$ ($\varepsilon < 0$), (7.6) is equivalent to the inequality

$$\frac{a_j - a_i}{c_j - c_i} \geq (\leq) \frac{\sum_{i=1}^n a_i s_i}{\sum_{i=1}^n c_i s_i} . \tag{7.7}$$

Note that the inequality in (7.7) does not depend on the value of ε. This fact has quite heavy implications! Indeed, it provides sufficient information to allow us to identify the form of the optimal system. We note first that, for any fixed vectors **a** and **c**, the criterion function is monotone in ε. Suppose a small positive ε (representing a shift of the amount ε from the jth to the ith elements of **s**) causes a positive increase in the criterion function m. Because the inequality in (7.7) doesn't depend on ε, one can show that further shifts of probability from s_j to s_i will cause further increases in m. (Indeed, the inequality which specifies that consecutive shifts of positive mass ε_1 and ε_2 from the jth to the ith elements of **s** further increases the value of m over its value following the single shift of mass ε_1 is easily shown to be equivalent to (7.7).) We conclude that, in such a circumstance, the maximum possible increase in m is obtained by shifting the entire amount s_j to the element s_i, that is, setting the ith element of **s** equal to $s_i + s_j$ and the jth element of **s** equal to zero. A similar argument, leading to a shift of mass in the opposite direction, obtains when ε is negative. In each case, m is monotone in ε, and the criterion function will necessarily increase by either shifting all the mass from s_j to s_i or from s_i to s_j. The main consequence of this argument is that, if **s** is a signature vector with at least two non-zero elements, it is always possible to increase the value of the criterion function by reducing the number of non-zero elements, that is, by shifting all the mass from one element of **s** to another. Repeated application of this procedure leads us to the following result.

Theorem 7.1. *Consider the class of all mixed systems based on* n *components with i.i.d. lifetimes. For* $r = 1$*, the criterion function* m *in* (7.5) *is maximized, for some* $k = 1, \ldots, n$*, by the signature* $s_{k:n}$ *of a* k*-out-of-*n *system.*

While identifying a specific k-out-of-n system that maximizes the criterion function m in (7.5) is a straightforward matter, it should be recognized that the maximizer need not be unique, and if there are two or more such systems which maximize m, any mixture of these systems will also be optimal. It remains true that, when $r = 1$, there is always at least one k-out-of-n system that is as good or better than any other system one might utilize. The collection of equivalent optimal systems is identified in the result below.

Corollary 7.1. *Assume that* $r = 1$ *in the criterion function* m *in* (7.5)*, with* **a** *and* **c** *fixed. Let*

$$K^* = \left\{ k \mid k = \operatorname*{argmax}_i \left\{ \frac{a_i}{c_i}, i = 1, \ldots, n \right\} \right\} .$$

Then, for any $k \in K^*$*,* m *is maximized by the* k*-out-of-*n *system with signature* $\mathbf{s} = \mathbf{s}_{k:n}$ *as well as by any mixture of* i*-out-of-*n *systems with* $i \in K^*$*.*

Turning our attention to the case in which $r \neq 1$ in the criterion function in (7.5), we begin by proving the following key lemma. From this result, we will be able to infer that, given any signature vector with at least three positive elements, there is a signature having one less positive element that will attain a criterion value that is at least as large.

Lemma 7.1. *Let κ be a fixed positive constant no larger than 1, and let S_κ be the simplex of vectors $\mathbf{s} \in [0,\kappa]^3$ for which $s_1 + s_2 + s_3 = \kappa$. Define the function $f : S_\kappa \to \mathbb{R}^+$ as*

$$f(\mathbf{s}) = \frac{b_1 + \sum_{i=1}^{3} a_i s_i}{\left(b_2 + \sum_{i=1}^{3} c_i s_i\right)^r} \tag{7.8}$$

with $0 < a_1 < a_2 < a_3$, $0 < c_1 < c_2 < c_3$, and $b_1, b_2 \geq 0$. Then $f(\mathbf{s})$ is maximized over $\mathbf{s} \in S_\kappa$ at a point on the boundary of S_κ; i.e. if $\mathbf{s}^* \in \{\operatorname{argmax} f(\mathbf{s}), \mathbf{s} \in S_\kappa\}$, then at least one of its elements is equal to 0.

Proof. The set S_κ is the convex hull of the set $\{(\kappa,0,0),(0,\kappa,0),(0,0,\kappa)\}$. Since $s_1 = \kappa - s_2 - s_3$, we may write the function f as

$$f(\mathbf{s}) = g(s_2, s_3) = \frac{x_1 + s_2 x_2 + s_3 x_3}{(y_1 + s_2 y_2 + s_3 y_3)^r} \tag{7.9}$$

where $x_1 = b_1 + \kappa a_1$, $y_1 = b_2 + \kappa c_1$ and $x_i = a_i - a_1$, $y_i = c_i - c_1$ for $i = 2,3$. We note that the function g in (7.9) is a continuous and differentiable function of (s_2, s_3) in the interior of S_κ, and can have an extreme point (that is, a local maximum or minimum) in the interior of S_κ only if that point is a critical point of g, that is, simultaneously satisfies the equations $\partial g/\partial s_2 = 0$ and $\partial g/\partial s_3 = 0$. Setting the first partial derivatives of g with respect to s_2 and s_3 equal to 0 results in the equations

$$x_2(y_1 + s_2 y_2 + s_3 y_3) - r y_2 (x_1 + s_2 x_2 + s_3 x_3) = 0 \tag{7.10}$$

and

$$x_3(y_1 + s_2 y_2 + s_3 y_3) - r y_3 (x_1 + s_2 x_2 + s_3 x_3) = 0\,. \tag{7.11}$$

If $x_2 y_3 = x_3 y_2$, equations (7.10) and (7.11) have no solutions in the plane, which implies that the function g has no local extrema, and the desired conclusion holds. If $x_2 y_3 \neq x_3 y_2$, equations (7.10) and (7.11) have the unique solution (s_2^*, s_3^*) given by

$$s_2^* = \frac{x_3 y_1 - x_1 y_3}{x_2 y_3 - x_3 y_2} \tag{7.12}$$

and

$$s_3^* = \frac{x_1 y_2 - x_2 y_1}{x_2 y_3 - x_3 y_2}\,. \tag{7.13}$$

We now demonstrate that s_2^* and s_3^* must be of opposite signs; this will imply that the point $(\kappa - s_2^* - s_3^*, s_2^*, s_3^*)$ lies outside of the simplex S_κ. The cases in which $(x_2y_3 - x_3y_2)$ is positive or negative are similar. Let us suppose that $(x_2y_3 - x_3y_2) > 0$. If s_2^* and s_3^* were both positive, we would have, simultaneously,

$$\text{a) } x_2y_3 > x_3y_2, \quad \text{b) } x_3y_1 > x_1y_3 \quad \text{and} \quad \text{c) } x_1y_2 > x_2y_1 \; . \tag{7.14}$$

Since y_1, y_2 and y_3 are positive by assumption, we obtain, after division by the appropriate y's, the equivalent inequalities

$$\text{d) } \frac{x_2}{y_2} > \frac{x_3}{y_3}, \quad \text{e) } \frac{x_3}{y_3} > \frac{x_1}{y_1} \text{ and f) } \frac{x_1}{y_1} > \frac{x_2}{y_2} \; . \tag{7.15}$$

Note that d) and f) in (7.15) imply that $x_1/y_1 > x_3/y_3$; they are thus incompatible with e), showing that s_2^* and s_3^* cannot both be positive. Analyzing the case in which $(x_2y_3 - x_3y_2) < 0$ yields the same conclusion. Together, these cases show that the unique critical point of the function f, namely $(\kappa - s_2^* - s_3^*, s_2^*, s_3^*)$, lies outside of the first quadrant and thus outside of the set S_κ. This in turn implies that the maximizer of the function f in the closed simplex S_κ must lie on the boundary. ■

Theorem 7.2. *Consider the class of all mixed systems of order n, and let $0 < r \neq 1$ in the criterion function in (7.5), with $0 < a_1 < a_2 < \cdots < a_n$ and $0 < c_1 < c_2 < \cdots < c_n$. Any system signature $\mathbf{s}^*$ maximizing the function $m_r(\mathbf{s}, \mathbf{a}, \mathbf{c})$ has at most two non-zero elements.*

Proof. Suppose that the signature vector $\mathbf{s}$ has at least three non-zero elements. Let's assume that $s_{j_i} > 0$ for $i = 1, 2, 3$. Note that the criterion function m can be written as

$$m(s) = \frac{b_1 + \sum_{i=1}^{3} a_i s_{j_i}}{\left(b_2 + \sum_{i=1}^{3} c_i s_{j_i}\right)^r} \; , \tag{7.16}$$

where

$$b_1 = \sum_{i \neq j_1, j_2, j_3} a_i s_i \quad \text{and} \quad b_2 = \sum_{i \neq j_1, j_2, j_3} c_i s_i \; .$$

Lemma 7.1 implies that there exists another system with fewer non-zero elements that has a criterion function value greater than or equal to $m(\mathbf{s})$. This process can be repeated as long as the resulting signature has at least three positive elements, eventually arriving at a signature vector $\mathbf{s}^*$ with at most two non-zero elements for which $m(\mathbf{s}^*) \geq m(\mathbf{s})$. ■

Theorem 7.2 implies that an optimal system may be found among the $\binom{n}{2}$ possible mixtures of two k-out-of-n systems. It follows that, when $r \neq 1$, one can find either a k-out-of-n system or a mixture of at most two k-out-of-n systems that maximizes the criterion function among all mixed systems of order n. While the lemma and theorem above leave open the possibility that an optimal system can always be found among degenerate mixtures, that is, among the class of the k-out-of-n systems, the following example shows that this is not the case. When $r \neq 1$, a non-degenerate mixture of coherent systems may indeed be superior to any individual coherent system relative to the criterion function in (7.5).

Example 7.1. Let us examine the class of all mixed systems of order two, including, of course, the two possible coherent systems, the parallel system and the series system. Let the performance criterion be expected system lifetime and the cost criterion be the salvage model. Assume that the component lifetimes are distributed according to the Uniform $(0, 1)$ distribution, so that $(a_1, a_2) = (\mu_{1:2}, \mu_{2:2}) = (1/3, 2/3)$. Letting $A = 2$, $B = 1$ and $C_I = 0$ in the salvage model (so that $c_1 = 3$ and $c_2 = 4$) and letting $r = 2.5$, the criterion m in (7.5) is given by

$$m_{2.5}(\mathbf{s}, \mathbf{a}, \mathbf{c}) = \frac{\frac{2}{3} - \frac{s_1}{3}}{(4 - s_1)^{2.5}}\ , \tag{7.17}$$

where s_2 has been replaced by $1 - s_1$. Upon differentiating m in (7.17) with respect to s_1, it is easy to verify that the value of s_1 that maximizes the function m above is $s_1 = 2/3$. Indeed, the mixed system with signature $\mathbf{s}^* = (2/3, 1/3)$ is the uniquely optimal system in this problem and outperforms each of the two coherent systems of order two. ■

The example above is sufficient to demonstrate that there are circumstances in which the class of coherent systems is suboptimal and that there exists a mixed system that does better. On the other hand, it is possible to identify conditions under which only certain types of two-fold mixtures of k-out-of-n systems can be optimal. Under such conditions, an optimal system must reside in a restricted subspace of the simplex of n-dimensional probability vectors. In exploring this possibility, we shall see that the optimality of non-degenerate mixed systems is not a rare peculiarity, but actually occurs in problems which might well arise in practice. In the developments that follow, we identify a particular, quite ordinary, setting in which only series and parallel systems and their mixtures can be optimal. We first establish the following tool.

Lemma 7.2. *Let κ be a fixed positive constant no larger than 1, and let S_κ be the simplex of vectors $\mathbf{s} \in [0, \kappa]^3$ for which $s_1 + s_2 + s_3 = \kappa$. Define the function $f : S_\kappa \to \mathbb{R}^+$ as*

$$f(\mathbf{s}) = \frac{b_1 + \sum_{i=1}^{3} a_i s_i}{\left(b_2 + \sum_{i=1}^{3} c_i s_i\right)^r} \tag{7.18}$$

for $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ *satisfying* $0 < a_1 < a_2 < a_3$, $b_1, b_2 \geq 0$ *and* $0 < c_1 < c_2 < c_3$. *For* $i = 2, 3$, *let* $a_i^* = a_i - a_{i-1}$ *and* $c_i^* = c_i - c_{i-1}$, *and assume that*

$$\frac{a_2^*}{c_2^*} < \frac{a_3^*}{c_3^*} \; . \tag{7.19}$$

If $f(\mathbf{s})$ *is maximized over* S_κ *by* $\mathbf{s}^*$, *then* $s_2^* = 0$.

Proof. Set $s_1 = \kappa - s_2 - s_3$ in (7.18), and let $g(s_2, s_3) = f(\kappa - s_2 - s_3, s_2, s_3)$. After a little algebra, we have that

$$\frac{\partial}{\partial s_2} g(s_2, s_3) = \frac{A(r)}{B(r)} \; , \tag{7.20}$$

where

$$A(r) = a_2^*[b_2 + \kappa c_1 + s_2 c_2^* + s_3(c_2^* + c_3^*)] - r c_2^*[b_1 + \kappa a_1 + s_2 a_2^* + s_3(a_2^* + a_3^*)]$$

and

$$B(r) = [b_2 + \kappa c_1 + s_2 c_2^* + s_3(c_2^* + c_3^*)]^{r+1} \; .$$

Now, set $s_3 = \kappa - s_1 - s_2$ in (7.18), and let $h(s_1, s_2) = f(s_1, s_2, \kappa - s_1 - s_2)$. We then have that

$$\frac{\partial}{\partial s_2} h(s_1, s_2) = \frac{C(r)}{D(r)} \; , \tag{7.21}$$

where

$$C(r) = -a_3^*[b_2 + \kappa c_3 - s_1(c_2^* + c_3^*) - s_2 c_3^*] + r c_3^*[b_1 + \kappa a_3 - s_1(a_2^* + a_3^*) - s_2 a_3^*]$$

and

$$D(r) = [b_2 + \kappa c_3 - s_1(c_2^* + c_3^*) - s_2 c_3^*]^{r+1} \; .$$

If the value of either (7.20) or (7.21) is negative, then for each fixed s_1 or s_3, shifting mass away from element s_2 will cause the value of the criterion function m to increase. Since the denominators in (7.20) and (7.21) are simply transformed versions of the denominator in (7.18), they are both positive. Thus, the signs of both (7.20) and (7.20) are the same as the signs of the numerators. We now show that, under the assumptions of the theorem, either one or both of the numerators of (7.20) and (7.21) are negative at any $r > 0$ and for any value of $s_2 < \kappa$. Now, the numerators of both (7.20) and (7.21) are linear functions of r. The numerator in (7.20) is decreasing in r while the numerator in (7.21) increasing in r. (The latter claim follows from the fact that the coefficients of r in (7.20) and (7.21) have the appropriate signs). Setting (7.20) equal to 0 and solving for r, we obtain:

$$r^* = \frac{a_2^*}{c_2^*} \left[\frac{b_2 + \kappa c_1 + s_2 c_2^* + s_3(c_2^* + c_3^*)}{b_1 + \kappa a_1 + s_2 a_2^* + s_3(a_2^* + a_3^*)} \right] , \tag{7.22}$$

while setting (7.21) equal to 0 and solving for r, we obtain:

$$r^{**} = \frac{a_3^*}{c_3^*} \left[\frac{b_2 + \kappa c_3 - s_1(c_2^* + c_3^*) - s_2 c_3^*}{b_1 + \kappa a_3 - s_1(a_2^* + a_3^*) - s_2 a_3^*} \right] , \tag{7.23}$$

Now, since $s_1 + s_2 + s_3 = \kappa$, we may make the substitution $s_1 = \kappa - s_2 - s_3$ in (7.23) to obtain

$$r^{**} = \frac{a_3^*}{c_3^*} \left[\frac{b_2 + \kappa c_1 + s_2 c_2^* + s_3(c_2^* + c_3^*)}{b_1 + \kappa a_1 + s_2 a_2^* + s_3(a_2^* + a_3^*)} \right] . \tag{7.24}$$

It is clear from (7.22) and (7.24) that $r^* < r^{**}$ if and only if $a_2^*/c_2^* < a_3^*/c_3^*$, which is precisely the assumed condition in (7.19). But $r^* < r^{**}$ implies that, at any fixed value of $r > 0$, either one or both of the derivatives (7.20) and (7.21) are negative. This in turn implies that the function f with positive s_2 can be increased by shifting mass away from s_2 to either s_1 or s_3 or both. Since this is true as long as $s_2 > 0$, it follows that any maxima of f in the closed simplex S_κ must have $s_2 = 0$. ■

The proof of Lemma 7.2 is roughly depicted in the figure below.

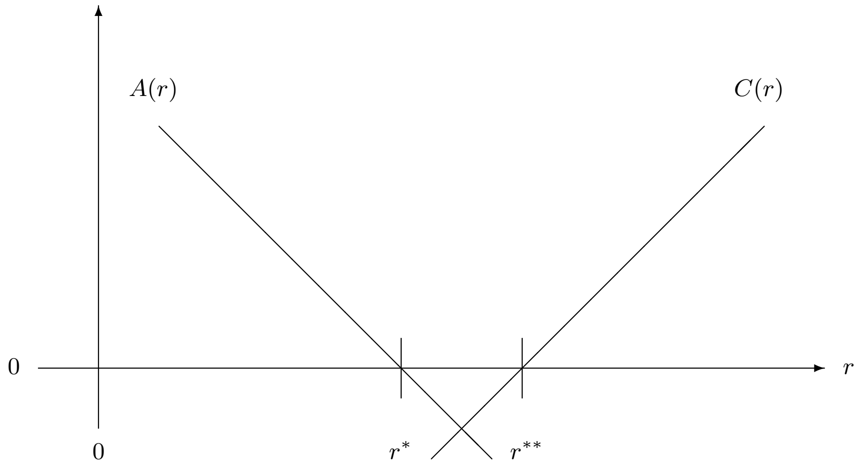


Fig. 7.1. The graph depicts the value of the numerators $A(r)$ and $C(r)$ of (7.20) and (7.21) as a function of $r > 0$

The lemma above provides the necessary tool to investigate the possible existence of special subclasses of mixed systems to which the search for an optimal system may be restricted. The result below shows that, under particular conditions on the vectors **a** and **c** of the criterion function of (7.5), the signature of the optimal system will put all of its mass on series and parallel systems.

Theorem 7.3. *Consider the class of all mixed systems based on n components with i.i.d. lifetimes, and let* $m_r(\mathbf{s}, \mathbf{a}, \mathbf{c})$ *be the criterion function in* (7.5), *with* $0 < a_1 < a_2 < \cdots < a_n$ *and* $0 < c_1 < c_2 < \cdots < c_n$. *Denote the successive differences of the elements of* **a** *and* **c** *as* $a_i^* = a_i - a_{i-1}$ *and* $c_i^* = c_i - c_{i-1}$ *for* $i = 2, \ldots, n$. *Suppose that*

$$\frac{a_2^*}{c_2^*} < \frac{a_3^*}{c_3^*} < \cdots < \frac{a_n^*}{c_n^*} . \tag{7.25}$$

Then for $r > 0$, *the optimal signature is a mixture of a series and a parallel system.*

Proof. Let d be a positive integer no greater that $n - 2$, and let $\mathbf{s}_1$ be a signature vector with d nonzero elements among the 2nd through $(n-1)$st elements of the signature; label these indices $i_1 < i_2 < \cdots < i_d$. We can write

$$m(\mathbf{s}_1) = \frac{b_1 + a_1 s_{11} + a_{i_1} s_{1i_1} + a_3 s_{1n}}{(b_2 + c_1 s_{11} + c_{i_1} s_{1i_1} + c_3 s_{1n})^r} ,$$

where $b_1 = \sum a_i s_i$ and $b_2 = \sum c_i s_i$, with the sum taken over all values of $i \neq 1, i_1, n$. Under the assumption (7.25), it follows from Lemma 7.2 that there exists another signature, $\mathbf{s}_2$ with $s_{2i_1} = 0$ and with $s_{2i} = s_{1i}$ for $i \neq 1, i_1, n$, for which $m(\mathbf{s}_2) > m(\mathbf{s}_1)$ obtains. Repeating this process, that is, successively shifting all the mass from each of the d "interior" nonzero elements of $\mathbf{s}_1$ to the elements s_{11} or s_{1n} or both, one achieves successive increases in the criterion function m. The resulting signature, after d iterations of this routine, is a mixture of the signatures of a series system and a parallel system. ■

Example 7.2. Consider the class of mixed systems based on n components with i.i.d. exponentially distributed lifetimes (with common mean μ, say). Let us take expected lifetime as the performance measure and assume the salvage model for costs. Since the values of c_i are linear and increasing in i, while the values of $a_i = EX_{i:n} = \mu \sum_{j=1}^{i} 1/(n - j + 1)$ are strictly convex in i, the vectors **a** and **c** obey condition (7.25) of Theorem 7.3. Table 7.3 shows the optimal signature for various values of n, r and A. These examples demonstrate a rather surprising lack of continuity in the nature of optimal systems, as the best system design "leaps" over all intermediate k-out-of-n systems, transitioning, as r increases, from a parallel system to mixtures of parallel and series systems to a series system.

Table 7.1. The optimal mixtures of 4 and 8 component systems with independent, exponentially distributed components. The salvage model is assumed with varying values of A and r, with the constant B set equal to 1 without loss of generality, as the optimal system depends on A and B only through the ratio A/B.

$n = 4$ components				$n = 8$ components			
A	r	s_1	s_4	A	r	s_1	s_8
1.5	≤ 1.6	0	1	1.5	≤ 1.6	0	1
1.5	1.7	0	1	1.5	1.7	.0967	.9033
1.5	1.8	.0568	.9432	1.5	1.8	.2156	.7844
1.5	1.9	.1768	.8232	1.5	1.9	.3081	.6919
1.5	2	.2727	.7273	1.5	2	.3821	.6179
1.5	2.5	.5606	.4394	1.5	2.5	.6042	.3958
1.5	3	.7045	.2955	1.5	3	.7152	.2848
1.5	4	.8485	.1515	1.5	4	.8262	.1738
1.5	5	.9205	.0795	1.5	5	.8817	.1183
1.5	10	1	0	1.5	10	.9742	.0258
1.5	≥ 10	1	0	1.5	≥ 15	1	0
2	≤ 1.6	0	1	2	≤ 1.6	0	1
2	1.7	0	1	2	1.7	0	1
2	1.8	0	1	2	1.8	0	1
2	1.9	0	1	2	1.9	0	1
2	2	0	1	2	2	0	1
2	2.5	.1162	.8838	2	2.5	.2232	.7768
2	3	.3712	.6288	2	3	.4295	.5705
2	4	.6263	.3737	2	4	.6357	.3643
2	5	.7538	.2462	2	5	.7388	.2612
2	10	.9663	.0337	2	10	.9107	.0893
2	≥ 12	1	0	2	≥ 27	1	0
3	≤ 1.6	0	1	3	≤ 1.6	0	1
3	1.7	0	1	3	1.7	0	1
3	1.8	0	1	3	1.8	0	1
3	1.9	0	1	3	1.9	0	1
3	2	0	1	3	2	0	1
3	2.5	0	1	3	2.5	0	1
3	3	0	1	3	3	0	1
3	4	.1818	.8181	3	4	.2548	.7452
3	5	.4205	.5795	3	5	.4531	.5469
3	10	.8181	.1818	3	10	.7837	.2163
3	≥ 21	1	0	3	≥ 51	1	0

7.4 Estimating the Relevant Characteristics of the Component Distribution

Thus far, we have assumed that the underlying component distribution F is a known quantity and that, consequently, the F-dependent values such as the expected order statistics $\{\mu_{i:n}, i = 1, \ldots, n\}$ of the n i.i.d. component lifetimes of the systems of interest are known. This assumption will not be satisfied in practical applications. Thus, utilizing the optimality results above requires the development of estimators for F or for the functions of that distribution on which the criterion function m in (7.5) depends. The developments above serve to identify the mixed system that maximizes the criterion function $m_r(\mathbf{s}, \mathbf{a}, \mathbf{c})$ in (7.5) for any fixed values of the vectors $\mathbf{a}$ and $\mathbf{c}$ and the constant r. The cost vector $\mathbf{c}$ and the calibration parameter r involve assessments on the part of the experimenter (the producer, the consumer or both), and it seems reasonable to assume that the value of $\mathbf{c}$ can be determined, in most applications of interest, with the assistance of engineering judgment and other expert advice. Further, the value of r can be chosen to fit a given application, perhaps with the help of a suitable sensitivity analysis. When no good reason appears to exist for setting r different from 1, the value $r = 1$ can serve as a default setting that can provide useful guidance.

A statistical problem remains to be addressed. The vector $\mathbf{a}$, being a function of the unknown distribution F, must be estimated from data. For the vector $\mathbf{a} = (\mu_{1:n}, \ldots, \mu_{n:n})$ of the expected order statistics of the components' lifetimes (the most natural choice for these performance-related constants), applying the optimality results above calls for reliable estimates of these μ's. We will hereafter assume, for concreteness, that this specification of the vector $\mathbf{a}$ has been chosen. The developments below are directed exclusively at the estimation of $\{\mu_{i:n}, i = 1, \ldots, n\}$. Specifically, consistent, asymptotically normal estimators of these parameters are developed below. It is then demonstrated that the maximized criterion function utilizing these estimates of the expected order statistics can be made arbitrarily close to the true maximum value, that is, that the optimal design relative to an estimated criterion function will eventually be ε-optimal, for arbitrary $\varepsilon > 0$, as the data set upon which our estimates are based grows.

The approach taken below to estimating $\mu_{i:n}$ requires access to data from an auxiliary experiment yielding a random sample of N component lifetimes having the common distribution F. In an environment in which the systems under consideration have components with i.i.d. lifetimes, it is reasonable to expect that a sample of such components would be available for independent life testing. The failure times of these N components can be used to estimate the expected order statistics $\{\mu_{i:n}, i = 1, \ldots, n\}$ of the lifetimes of the n components upon which the systems of interest rely. One might estimate $\mu_{i:n}$ by taking repeated bootstrap samples of size n from $X_1, \ldots, X_N$; the average of the ith largest values of each bootstrap sample stands to yield a good approximation of $\mu_{i:n}$. The estimator proposed below is closely related to this empirical procedure. It is the expected value of the bootstrap estimate given the observed values $x_1, \ldots, x_N$. Hutson and Ernst [45] proposed such an estimator for $\mu_{i:n}$ based only on samples of size n from F. The estimator proposed for study here is based on random samples of arbitrary size N from

F.

For a continuous distribution F, the expected value of the ith order statistic from a sample of size n from F can be represented by the familiar expression

$$\mu_{i:n} = \int_{-\infty}^{\infty} \mathrm{Be}(i, n-i+1) x (F(x))^{i-1} (1-F(x))^{n-i} \mathrm{d}F(x) , \tag{7.26}$$

where $\mathrm{Be}(\alpha, \beta) = \Gamma(\alpha+\beta) \,/\, \Gamma(\alpha)\Gamma(\beta)$ for $\alpha > 0$ and $\beta > 0$. Since we will be interested in estimating $\mu_{i:n}$ from an independent sample of size N, we will find the following alternative expression more useful:

$$\begin{aligned} \mu_{i:n} &= \int_0^1 \mathrm{Be}(i, n-i+1) u^{i-1} (1-u)^{n-i} F^{-1}(u) \, \mathrm{d}u \\ &= \sum_{j=1}^{N} \int_{(j-1)/N}^{j/N} \mathrm{Be}(i, n-i+1) u^{i-1} (1-u)^{n-i} F^{-1}(u) \, \mathrm{d}u \, . \end{aligned} \tag{7.27}$$

The "analog estimator" of $\mu_{i:n}$ is obtained by replacing $F^{-1}(u)$ in (7.27) by the sample quantile function, that is, by replacing $F^{-1}(u)$ for $u \in ((j-1)/N, j/N)$ by $\hat{F}^{-1}(u) = X_{[Nu]+1:N}$, where $[y]$ is the greatest integer less than or equal to y. This substitution leads to the estimator

$$\hat{\mu}_{i:n} = \sum_{j=1}^{N} \left\{ \int_{(j-1)/N}^{j/N} J_{i:n}(u) \, \mathrm{d}u \right\} X_{j:N} , \tag{7.28}$$

where $J_{i:n}(u) = \mathrm{Be}(i, n-i+1) u^{i-1} (1-u)^{n-1}$. The estimator $\hat{\mu}_{i:n}$ is the expected value of $X_{i:n}$ under assumed sampling from the empirical distribution F_N of the observed component failure times $X_1, \ldots, X_N$. It is an explicit formula for the mean bootstrap estimate of $\mu_{i:n}$ based on repeated samples of size n from $x_1, \ldots, x_N$. For our purposes, a more important feature of the estimator $\hat{\mu}_{i:n}$ is the fact that it is an L-estimator, that is, a linear combination of order statistics. This fact makes available for application in the present problem the extensive literature and theoretical developments that exist for such estimators. We take a brief digression here to make note of some relevant background on L-estimation.

The study of L-estimators over the years has largely been motivated by the interest in robust methods of estimation. When estimating a location parameter, for example, it has long been known that the extreme order statistics in one's sample tend to have a strong and often undesirable influence on an estimator's precision. Of course the mean $\overline{X}$ of a sample of size N is itself an L-estimator (since it can be expressed as the linear combination $\sum(1/N)X_{i:N}$), but it is not typical of the L-estimators employed to estimate the population mean. Because of the strong influence that extreme order statistics can have on $\overline{X}$, alternative L-estimators like the sample median, trimmed means (which assign weight zero to a certain percentage of the sample among the largest and the smallest observations) or more refined estimators of the form $\sum w_{j:N} X_{j:N}$, which assign higher weights to observations close to the middle of the data and lower weights to outlying observations, are typically recommended for practical use. A detailed treatment of the early history of L-estimators can be found in Stigler [69]. Prominent among the early contributors to the theory of L-estimation include Bickel, Govidarajulu, Stigler, Shorack and van Zwet. Reviews of the L-estimator literature can be found in Chapter 1 of Helmers [41], in Shorack and Wellner [67], and in Chapter 22 of van der Vaart [72].

All existing results on the consistency or asymptotic normality of estimators of the form $\sum w_{j:N} X_{j:N}$ require restrictions on both the weights $w_{j:N}$ and on the underlying distribution F. Published results are often non-comparable in that some make strong assumptions about the weights and weak assumptions about F while others do the opposite. The goal of balancing these two ways of restricting the sum to obtain asymptotic stability can be achieved in a variety of ways. From equation (7.28), we note that $\hat{\mu}_{i:n}$ is an L-estimator with weights,

$$w_{j:N} = \left\{ \int_{(j-1)/N}^{j/N} J_{i:n}(u)\, \mathrm{d}u \right\} . \tag{7.29}$$

These weights are generated from the integral of a bounded continuous function with bounded derivative existing almost everywhere. This particular choice of weights will thus tend to satisfy the most stringent of regularity conditions typically placed on the weights. Because of this, we shall be able to obtain a satisfactory asymptotic theory for the estimators $\{\hat{\mu}_{i:n}\}$ while placing minimal restrictions on F.

Van Zwet [73] established the strong consistency of L-estimators under the relatively mild assumptions specified in the theorem below. His proof applies Hölder's inequality to the difference $|\hat{\mu}_{i:n} - \mu_{i:n}|$ (see Hewitt and Stromberg [42], pp. 188 - 190, for relevant definitions). We state van Zwet's [73] theorem without proof. It provides for strong rather than weak consistency of L-estimators, in contrast with most consistency results in the literature on L statistics.

Theorem 7.4. *Consider functions F and J used in defining $\mu_{i:n}$ and $\hat{\mu}_{i:n}$ in equations* (7.27) *and* (7.28), *and assume that $J \in L_p$ and $F^{-1} \in L_q$ with $(1/p + 1/q) = 1$. Let $X_1, \ldots, X_N$ be i.i.d $\sim F$, and assume that $E(|X_1|^{1+\alpha}) < \infty$, where $\alpha > 0$. If $\hat{\mu}_{i:n}$ is the L-estimator given in* (7.28), *then, for every fixed i and n, $\hat{\mu}_{i:n} \to \mu_{i:n}$ with probability 1.*

The asymptotic distribution theory for $\hat{\mu}_{i:n}$ will provide the rate of its convergence to $\mu_{i:n}$ and will fully characterize the estimator's asymptotic behavior. To achieve the desired results, we employ the theory of statistical functionals (see, for example, Fernholz [38]). We proceed by expressing the target parameter as a statistical functional. We then show that the Taylor series expansion of an empirical counterpart of the functional has a linear term to which the central limit theorem applies and a remainder term converging to zero at an appropriate rate. The mild condition on F that will be required in these developments is the existence of a finite second moment.

Let us define the statistical functional $T(F)$ as follows:

$$T(F) = \int_{-\infty}^{\infty} x J_{i:n}(F(x))\, \mathrm{d}F(x)\ , \tag{7.30}$$

an integral that may alternatively be written as

$$T(F) = \int_{0}^{1} F^{-1}(u) J_{i:n}(u)\, \mathrm{d}u\ . \tag{7.31}$$

We proceed by taking the Taylor series expansion of $T(F + t(F_N - F))$ about $t = 0$:

$$T(F_N) - T(F) = T^{(1)}(F + t(F_N - F))\Big|_{t=0} + \sum_{k=2}^{\infty} \frac{1}{k!}\, T^{(k)}(F + t(F_N - F))\Big|_{t=0}\ . \tag{7.32}$$

Expressions for the derivatives in (7.32) are given in the following result.

Lemma 7.3. *For* $k = 1, 2, \ldots,$

$$\begin{aligned}
&T^{(k)}(F + t(F_N - F)) = \frac{\mathrm{d}^k}{\mathrm{d}t^k}[T(F + t(F_N - F)]\\
&= \int_{-\infty}^{\infty} x(F_N(x) - F(x))^k J_{i:n}^{(k)}(F(x) + t(F_N(x) - F(x)))\, \mathrm{d}F(x)\\
&+ k \int_{-\infty}^{\infty} x(F_N(x) - F(x))^{k-1}\\
&\qquad\qquad \times J_{i:n}^{(k-1)}(F(x) + t(F_N(x) - F(x)))\, \mathrm{d}[F_N(x) - F(x)]\\
&+ t \int_{-\infty}^{\infty} x(F_N(x) - F(x))^{k}\\
&\qquad\qquad \times J_{i:n}^{(k)}(F(x) + t(F_N(x) - F(x))\, \mathrm{d}[F_N(x) - F(x)]\ .
\end{aligned} \tag{7.33}$$

Proof. By Fubini's theorem, one may pass the first derivative under the integral sign and verify that the theorem holds for $k = 1$. We proceed by induction. Assuming that the theorem holds for a fixed positive integer k, we consider the $(k + 1)$st derivative. We have

$$\frac{\mathrm{d}^{k+1}}{\mathrm{d}t^{k+1}}[T(F+t(F_N-F)] = \frac{\mathrm{d}}{\mathrm{d}t}\left\{\frac{\mathrm{d}^k}{\mathrm{d}t^k}[T(F+t(F_N-F)]\right\}$$
$$= \frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} x(F_N(x)-F(x))^k J_{i:n}^{(k)}(F(x)+t(F_N(x)-F(x)))\,\mathrm{d}F(x)$$
$$+ k\frac{\mathrm{d}}{\mathrm{d}t}\int_{-\infty}^{\infty} x(F_N(x)-F(x))^{k-1}$$
$$\times J_{i:n}^{(k-1)}(F(x)+t(F_N(x)-F(x)))\,\mathrm{d}[F_N(x)-F(x)]$$
$$+ t\frac{\mathrm{d}}{\mathrm{d}t}\left\{\int_{-\infty}^{\infty} x(F_N(x)-F(x))^k\right.$$
$$\left.\times\, J_{i:n}^{(k)}(F(x)+t(F_N(x)-F(x)))\,\mathrm{d}[F_N(x)-F(x)]\right\} \qquad (7.34)$$
$$= \int_{-\infty}^{\infty} x(F_N(x)-F(x))^{k+1} J_{i:n}^{(k+1)}(F(x)+t(F_N(x)-F(x)))\,\mathrm{d}F(x)$$
$$+ k\int_{-\infty}^{\infty} x(F_N(x)-F(x))^k J_{i:n}^{(k)}(F(x)+t(F_N(x)-F(x)))\,\mathrm{d}[F_N(x)-F(x)]$$
$$+ \int_{-\infty}^{\infty} x(F_N(x)-F(x))^k J_{i:n}^{(k)}(F(x)+t(F_N(x)-F(x)))\,\mathrm{d}[F_N(x)-F(x)]$$
$$+ t\int_{-\infty}^{\infty} x(F_N(x)-F(x))^{k+1}$$
$$\times J_{i:n}^{(k+1)}(F(x)+t(F_N(x)-F(x)))\,\mathrm{d}[F_N(x)-F(x)]\,. \qquad (7.35)$$

Combining the second and third integrals in (7.35), we see that the $(k+1)$st derivative obeys the prescription in (7.33). ■

Next, we rewrite the first term on the right-hand side of (7.32) using integration by parts:

$$T^{(1)}(F+t(F_N-F))\Big|_{t=0} = \int_{-\infty}^{\infty} x(F_N(x)-F(x))J_{i:n}^{(1)}(F(x))\,\mathrm{d}F(x)$$
$$+\int_{-\infty}^{\infty} xJ_{i:n}(F(x))\,\mathrm{d}[F_N(x)-F(x)]$$
$$= \int_{-\infty}^{\infty} x(F_N(x)-F(x))J_{i:n}^{(1)}(F(x))\,\mathrm{d}F(x)$$
$$-\int_{-\infty}^{\infty} (F_N(x)-F(x))\,\mathrm{d}[xJ_{i:n}(F(x))]$$
$$= -\int_{-\infty}^{\infty} (F_N(x)-F(x))J_{i:n}(F(x))\,\mathrm{d}x$$
$$= -\frac{1}{N}\sum_{j=1}^{N}\int_{-\infty}^{\infty} (I_{(-\infty,x)}(X_j)-F(x))J_{i:n}(F(x))\mathrm{d}x\,. \qquad (7.36)$$

Invoking Fubini's Theorem, we see that the i.i.d. random variables in (7.36) have expectation 0. Since $E(X^2) < \infty$, we obtain, again using Fubini's Theorem, the following variance expression for an arbitrary term in the sum in (7.36):

$$\begin{aligned}\sigma_{i:F}^2 &= E\left[\int_{-\infty}^{\infty} (I_{(-\infty,x)}(X_j) - F(x))J_{i:n}(F(x))\,\mathrm{d}x\right]^2 \\ &= E\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (I_{(-\infty,x)}(X_j) - F(x))(I_{(-\infty,y)}(X_j) - F(y)) \\ &\qquad\qquad \times J_{i:n}(F(x))J_{i:n}(F(y))\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} J_{i:n}(F(x))J_{i:n}(F(y))[\min(x,y) - F(x)F(y)]\,\mathrm{d}x\mathrm{d}y\,. \end{aligned} \tag{7.37}$$

Applying the central limit theorem to the sum in (7.36), we may identify the asymptotic distribution of $\hat{\mu}_{i:n}$. This follows from the fact that

$$\sqrt{N}\sum_{k=2}^{\infty}\frac{1}{k!}\,T^{(k)}(F + t(F_N - F))\Big|_{t=0} \xrightarrow{p} 0\,,$$

as is shown below.

Theorem 7.5. *Let $X_1, X_2, \ldots, X_N$ be i.i.d. with common distribution F, and assume that $0 < E(X_1^2) < \infty$. For every fixed $i = 1, 2, \ldots, n$,*

$$\sqrt{N}(\hat{\mu}_{i:n} - \mu_{i:n}) \xrightarrow{D} Y \sim N(0, \sigma_{i:F}^2)\,,$$

where $\sigma_{i:F}^2$ is given in (7.37).

Proof. Note that, since $J_{i:n}(u)$ is a polynomial of degree $n-1$,

$$J_{i:n}^{(n+j)}(u) = 0 \text{ for } j = 0, 1, 2, \ldots\,;$$

it follows that

$$\sum_{k=n+1}^{\infty}\frac{1}{k!}\,T^{(k)}(F + t(F_N - F))\Big|_{t=0} = 0\,.$$

Moreover, we may write

$$J_{i:n}^{(k)}(u) = \frac{n!}{(n-k)!}\sum_{j=0}^{k}(-1)^j\binom{k}{j}J_{i-j:n-k}(u), \text{ for } k = 1, 2, \ldots, n-2\,,$$

where $J_{i-j:n-k}(u) = 0$ whenever $i - j \le 0$ or $i - j \ge n - k$. Since $J_{a:b}(u)$ is bounded for all $0 \le a \le b \le n$, this shows that $\left|J_{i:n}^{(k)}(u)\right|$ is bounded.

We now show the convergence of the remaining terms in (7.32) to zero, that is,

$$\sqrt{N}\sum_{k=2}^{n}\frac{1}{k!}\,T^{(k)}(F+t(F_N-F))\Big|_{t=0}\xrightarrow{p}0\,. \tag{7.38}$$

It suffices to show that the quadratic term (i.e., the term in which $k=2$) converges to 0. The higher order terms (with $k\geq 3$) converge at least as fast to 0 as the quadratic term since $|F_N(x)-F(x)|\leq 1$ and for all k, $|J_{i:n}^{(k)}|<M_k$ for some $M_k<\infty$. Now

$$\begin{aligned}
&\sqrt{N}\frac{1}{2!}\,T^{(2)}(F+t(F_N-F))\Big|_{t=0}\\
&=\frac{1}{\sqrt{N}}\int_{-\infty}^{\infty}x\left[\sqrt{N}(F_N(x)-F(x))\right]^2J_{i:n}^{(2)}(F(x))\,\mathrm{d}F(x) &(7.39)\\
&+\frac{2}{\sqrt{N}}\int_{-\infty}^{\infty}x\sqrt{N}(F_N(x)-F(x))J_{i:n}^{(1)}(F(x))\mathrm{d}[F_N(x)-F(x)]\,. &(7.40)
\end{aligned}$$

To show that the quantities in (7.39) and (7.40) each converge to 0 in probability, we apply Donsker's Theorem (see, for example, van der Vaart [72]) to deduce that $\sqrt{N}\sup_x(F_N(x)-F(x))=O_p(1)$. Let M_2 be such that $|J_{i:n}^{(2)}(F(x))|<M_2$. Then the convergence of (7.39) to 0 follows from

$$\begin{aligned}
&\left|\frac{1}{\sqrt{N}}\int_{-\infty}^{\infty}x\left[\sqrt{N}(F_N(x)-F(x))\right]^2J_{i:n}^{(2)}(F(x))\,\mathrm{d}F(x)\right|\\
&\leq\frac{1}{\sqrt{N}}\int_{-\infty}^{\infty}\left[\sup_x\sqrt{N}(F_N(x)-F(x))\right]^2\left|xJ_{i:n}^{(2)}(F(x))\right|\,\mathrm{d}F(x)\\
&\leq\frac{1}{\sqrt{N}}O_p(1)M_2\int_{-\infty}^{\infty}|x|\,\mathrm{d}F(x)\\
&=O_p\left(\frac{1}{\sqrt{N}}\right)\,.
\end{aligned}$$

The convergence of (7.40) to zero follows by a similar argument:

$$\begin{aligned}
&\left|\frac{2}{\sqrt{N}}\int_{-\infty}^{\infty}x\sqrt{N}(F_N(x)-F(x))J_{i:n}^{(1)}(F(x))\,\mathrm{d}[F_N(x)-F(x)]\right|\\
&\leq\left|\frac{2}{\sqrt{N}}\int_{-\infty}^{\infty}\sqrt{N}\sup_x|F_N(x)-F(x)|\times\left|xJ_{i:n}^{(1)}(F(x))\right|\mathrm{d}[F_N(x)+F(x)]\right|\\
&\leq\frac{2}{\sqrt{N}}O_p(1)M_1\frac{1}{N}\sum_{j=1}^{N}|X_j|+\frac{2}{\sqrt{N}}O_p(1)M_1\int_{-\infty}^{\infty}|x|\,\mathrm{d}F(x)\\
&=O_p\left(\frac{1}{\sqrt{N}}\right)\,.
\end{aligned}$$

The two bounds above serve to establish that

$$\sqrt{N}\frac{1}{2!}\,T^{(2)}(F+t(F_N-F))\Big|_{t=0} \xrightarrow{p} 0\,,$$

as claimed. ■

7.5 Approximately Optimal System Designs

In this section, it will be shown that the consistency of each $\hat{a}_{Ni} = \hat{\mu}_{i:n}$ demonstrated above implies that $m_r(s_N^*;\hat{\mathbf{a}}_N;\mathbf{c})$ converges to $m_r(\mathbf{s}^*;\mathbf{a};\mathbf{c})$ in some appropriate sense, where $\mathbf{s}^*$ is the true optimal signature and

$$\mathbf{s}_N^* = \underset{\mathbf{s}}{\operatorname{argmax}}\left\{\frac{\sum_{i=1}^n \hat{a}_{Ni}s_i}{\left(\sum_{i=1}^n c_i s_i\right)^r}\right\}\,. \tag{7.41}$$

In the following theorem, conditions are specified under which the optimized approximate criterion function converges to the true maximum of $m_r(\mathbf{s};\mathbf{a};\mathbf{c})$ over all signature vectors $\mathbf{s}\in[0,1]^n$.

Theorem 7.6. *Let* $m_r(\mathbf{s};\mathbf{a};\mathbf{c})$ *be the criterion function in* (7.5)*, where* $\mathbf{a}$ *and* $\mathbf{c}$ *satisfy the constraints* $0<a_1<a_2<\cdots<a_n$ *and* $0<c_1<c_2<\cdots<c_n$*. Let* $\mathbf{a}^*$ *be the true value of* $\mathbf{a}$*. If* $\hat{\mathbf{a}}_N$ *is the estimated value of* $\mathbf{a}$ *based on a random sample of size* N*, and* $\hat{\mathbf{a}}_N \xrightarrow{p} \mathbf{a}^*$ *as* $N\to\infty$*, then*

$$m_r(\mathbf{s}_N^*;\hat{\mathbf{a}}_N;\mathbf{c}) \xrightarrow{p} m_r(\mathbf{s}^*;\mathbf{a}^*;\mathbf{c})\,. \tag{7.42}$$

Proof. We will develop an appropriate upper bound for the distance between the quantities of interest. Note that

$$\begin{aligned}
|m_r(\mathbf{s}_N^*;\hat{\mathbf{a}}_N;\mathbf{c})-m_r(\mathbf{s}^*;\mathbf{a};\mathbf{c})| &= \left|\sup_s m_r(\mathbf{s};\hat{\mathbf{a}}_N;\mathbf{c})-\sup_s m_r(\mathbf{s};\mathbf{a}^*;\mathbf{c})\right|\\
&\le \sup_s |m_r(\mathbf{s};\hat{\mathbf{a}}_N;\mathbf{c})-m_r(\mathbf{s};\mathbf{a}^*;\mathbf{c})|\\
&\le \sup_s\left|\frac{\sum_{i=1}^n \hat{a}_{Ni}s_i}{\left(\sum_{i=1}^n c_i s_i\right)^r}-\frac{\sum_{i=1}^n a_i^* s_i}{\left(\sum_{i=1}^n c_i s_i\right)^r}\right|\\
&\le (c_1)^{-r}\sup_s\left|\sum_{i=1}^n(\hat{a}_{Ni}-a_i^*)s_i\right|\\
&\le (c_1)^{-r}\max_i|\hat{a}_{Ni}-a_i^*| = o_p(1).
\end{aligned}$$

■

The underlying component lifetime distribution F is unknown in virtually all practical settings in which the optimality results of section 7.3 might be of interest. In such settings, the vector of expected order statistics $\mathbf{a}$ can be consistently estimated from auxiliary experiments. Theorem 7.6 shows that, if F is assumed to have a finite second moment, the optimality results of section

7.3 are applicable with the vector of expected order statistics $\mathbf{a}$ replaced by the consistent estimator $\hat{\mathbf{a}}_N$. When N is sufficiently large, the optimal value $m_r(\mathbf{s}^*_N; \hat{\mathbf{a}}_N; \mathbf{c})$ of the approximate criterion function can be made arbitrarily close to the true optimal value $m_r(\mathbf{s}^*; \mathbf{a}^*; \mathbf{c})$ with high probability. One thus has the mechanisms needed to identify ε-optimal systems, for arbitrary $\varepsilon > 0$, based on the sampling experiments used to obtain our estimates of the expected order statistics of the component lifetimes.

There is always an element of uncertainty in applying asymptotic theory to problems that occur in practice. A perennial question that arises is how large a sample size is needed for the asymptotic approximation of interest to hold. In the present context, one might ask how large the auxiliary sample size N would need to be for the "approximately optimal" system to be close to the true optimal system (in the sense of Theorem 7.6). To gain some insight into this question, Dugas and Samaniego [33] carried out a simulation to compare approximate and true optimal systems in a particular problem involving systems of order 4. Component lifetimes were assumed to be independent exponential variables with mean 1. Dugas and Samaniego took the cost vector in the criterion function m of (7.5) to be $\mathbf{c} = (2, 3, 4, 10)$ and the calibration parameter to be $r = 1$. The true values of the expected order statistics from Exp(1) are (1/4, 7/12, 13/12, 25/12). From Corollary 7.1, the optimal system is found to be the 3-out-of-4 system with signature $\mathbf{s} = (0, 0, 1, 0)$. Samples of size $N = 100$ were repeatedly generated from the distribution $F = \text{Exp}(1)$. The data below were displayed as an example of a typical (ordered) outcome among their 200 replications.

Table 7.2. One hundred simulated Exp(1) variables.

0.025	0.064	0.065	0.069	0.075	0.081	0.091	0.094	0.094	0.100
0.133	0.145	0.174	0.175	0.181	0.189	0.214	0.219	0.228	0.235
0.262	0.273	0.276	0.278	0.283	0.321	0.324	0.333	0.346	0.393
0.407	0.407	0.442	0.444	0.485	0.507	0.524	0.528	0.536	0.558
0.567	0.579	0.599	0.609	0.609	0.625	0.651	0.701	0.708	0.709
0.721	0.730	0.741	0.827	0.870	0.889	0.902	0.913	0.926	0.944
0.953	0.973	0.997	1.052	1.150	1.150	1.156	1.185	1.198	1.291
1.313	1.330	1.351	1.486	1.507	1.529	1.588	1.624	1.636	1.731
1.747	1.768	1.831	1.880	1.908	1.915	1.934	1.966	2.403	2.415
2.453	2.642	2.726	2.907	3.160	3.840	3.958	4.796	4.926	6.094

From the data above, the estimate of $\mathbf{a}$ can be calculated to be $\hat{\mathbf{a}} = (0.2801, 0.6261, 1.1571, 2.2913)$, and the ratios of $\hat{a}_i$ to c_i are given by 0.1401, 0.2087, 0.2893 and 0.2291. Thus, for these data, the process of approximating the expected order statistics of F by the L-estimators in (7.28) leads to an "approximately optimal" system that is in fact optimal. The most impressive aspect of the simulation is the fact that the "approximately optimal" system coincided with the true optimal system in each of the 200 simulations based on 100 Exp(1) variables. This suggests that the estimation approach developed above has the potential to work with substantial dependability even when the size of the auxiliary sample is small to moderate. The success of the approach is, of course, dependent on the underlying distribution F, the applicable cost vector $\mathbf{c}$ and the calibration parameter r.

7.6 Discussion

Many problems in the area of Reliability Economics have heretofore proven to be quite resistant to analytical treatment. In contrast, the formulation of the problem treated above makes an analytical treatment completely feasible. The key technical ingredients have been (i) the use of system signatures as an index for the systems under consideration, (ii) the identification of a particular class of reasonable criterion functions that depend on the system design solely through the signature and (iii) the broadening of the class of coherent systems to the class of mixed systems, making the criterion function m a continuous function of the family of n-dimensional probability vectors. In this framework, it is possible to characterize optimal system designs precisely instead of simply approximating them using searching techniques. In the problems to which the present formulation applies, this represents a useful conceptual advance. To apply these optimization results in practice, that is, when the underlying component distribution F is unknown, one must estimate the features of the criterion function that depend on F. For the important case where F influences the criterion function through its expected order statistics $\{\mu_{i:n}\}$, specific estimators for these parameters have been proposed (assuming the availability of an auxiliary sample of component lifetimes), and their consistency and asymptotic normality have been confirmed. Finally, it has been shown that the system which optimizes the approximate criterion function will be ε-optimal relative to the true criterion function when the size of the auxiliary sample is sufficiently large. We conclude this section with some discussion regarding the envisioned domain of application of these results.

Throughout this monograph, we have made the assumption that the components of the systems of interest have i.i.d. lifetimes. Since the notion of system signature extends beyond this setting, generalizations of these results are possible. There are, however, some technical obstacles to such generalizations. These are discussed and briefly explored in the concluding chapter. The results established here serve several purposes. First, they represent a concrete entrée into a challenging area of application with an approach that permits an analytical treatment of optimization problems. Secondly, the theory developed is directly applicable to systems in i.i.d. components, be they the chips or wafers in a computer, optical cables in a communication network or batteries in a flashlight. Finally, since the i.i.d. assumption "evens the playing field" when comparing two competing systems, these results provide a useful benchmark when comparing system designs. The optimal system in an i.i.d. scenario can reasonably be expected to perform well, if not optimally, in a "neighborhood" of that scenario, and perhaps even more broadly.

Finally, let us briefly revisit the notion of mixed systems, as they have been a key feature of the formulation and solution of the problem considered in this chapter. The use of mixed strategies in game theory and the use of randomized decision rules in decision theory are essential in the search for and identification of optimal procedures. From a mathematical point of view, the use of randomization in statistical work is beyond reproach. One simply can't have a comprehensive optimization theory without them. In the problems studied here, we've seen that a mixed system may well dominate all other feasible solutions (including all non-randomized solutions – the coherent systems themselves). Table 7.3 demonstrates quite vividly that mixed systems do in fact, in a setting as common as exponential life testing, provide "the best bang for the buck." While acknowledging this mathematical reality, one might still ask whether mixed systems have true functionality. Can they be employed in practice? Will an engineer accept the fact that, in a given situation, he/she should use a mixed system which is specified by using a randomization device? A consumer may be uncomfortable with the supplier's randomization, providing him/her, for example, with a series system with probability 1/2 and a parallel system with probability 1/2. The fact remains that if the consumer is going to make repeated use of these types of systems, the sequence of systems provided according to the optimal mixture produces the best results when performance and cost are considered simultaneously. The consequences of using a suboptimal system can, of course, be measured and compared by their corresponding values of $m_r(\mathbf{s};\mathbf{a};\mathbf{c})$.

The real domain for application of the results in this chapter is in problems in which a large number of systems will be purchased and used. The developments above suggest that it may be best to randomize in the use of the collection of systems used. To make the argument for the use of mixtures more concrete, consider a car rental company in the process of purchasing a new fleet of cars. Suppose that a mixed strategy that dictates purchasing new cars according to a fixed mixing distribution rather than to buying the same model repeatedly is found to be optimal with respect to a "performance per adjusted unit cost" criterion. Using such a strategy, that is, mixing these purchases according to the prescribed recipe, would produce an optimal overall result.

An interesting feature of these results is that optimality can be achieved by restricting attention to k-out-of-n systems and mixtures thereof. It is also worth noting that particular k-out-of-n systems are often optimal in the class of all mixed systems (always if $r = 1$). In fact, in the general framework we have studied, the system one should employ in any given application will never involve a mixture of more than two such systems. Thus, the implementation of the optimal mixing strategy is, in the end, a fairly painless exercise.

8

Summary and Discussion

8.1 Introduction

There are many problems of probability and statistics in which characterizing a large and awkward space of objects by a simpler index of the space facilitates analysis and makes the identification of optimal or at least rational solutions possible. The notion of a sufficient statistic, one that can reduce the data to a simple summary measure without loss of information about the unknown features of the model involved, is perhaps the quintessential example of this phenomenon. In linear model theory, results on dimension reduction have the same aim, though the possibility of such reduction without some (at least minor) loss of information is rarely possible. In this latter case, the compromise is generally deemed to be worth making. The theory and applications of system signatures can be thought of in the same way. The signature of a system is a characteristic of the system's design which captures an essential feature of that design. Specifically, it provides a measure of how component failures influence system failures when the components are independent and have the same lifetime distributions. As mentioned earlier, this leveling of the playing field among the components' theoretical performance allows one to focus exclusively on system design. Signatures are deterministic measures that are properly classified as tools within the field of Structural Reliability, providing information solely about the design of the corresponding system.

Regarding the information lost in using a signature vector as a proxy for a particular system design, there are two sources of lost information that require mention. The first is that there is not a one-to-one correspondence between systems and signatures. There are, for example, only 17 distinct signatures among the 20 different coherent systems of order 4. There are three pairs of systems of order 4 that have the same lifetime distributions when the component lifetimes are i.i.d. with a common distribution F. Of course, all 20 system lifetime distributions would differ if one relaxes the i.i.d. assumption. Secondly, the lifetime distributions of the components of most real systems cannot reasonably be considered identical. The independence of the component lifetimes is a less restrictive assumption, but dependencies can certainly occur due to stress, wear-out or early failures, and a theory for comparing two systems which relaxes both the independence and identically distributed assumptions could well be viewed as the ultimate goal of the type of analyses developed in this monograph.

This final chapter has several purposes. I will present, in Section 8.2, an overview of the theory and applications of system signatures, summarizing what I consider to be the highlights of the present monograph. This will include some brief commentary on the definition and interpretation of system signatures, related representations of system lifetime distributions, preservation and characterization results based on traditional stochastic orders, alternative signature-based metrics for comparing systems, the relationship between dominations and signatures in the context of communication networks and the search for optimal systems in a Reliability Economics setting. Possible extensions of the developments mentioned above are discussed in Section 8.3, where I attempt to provide some indication of the extent of generalization that appears to be feasible. In Section 8.4, I will review some signature-related literature that has not been mentioned in this monograph but gives further evidence of the broad applicability of the concept. Finally, in Section 8.5, I will mention a number of open problems for which solutions would be most welcome. In the spirit of the great mathematician Paul Erdös, I will offer financial rewards for published solutions to these problems. Being of comparatively modest means, however, I cannot match the tantalizing offers that Erdös enjoyed sprinkling throughout his lectures. I will pay 50 cents for solutions to easy problems and 1 dollar for solutions to hard ones. While these miserly offers won't serve as much of an incentive for anyone to work on these problems, I will count on old-fashioned self satisfaction, plus the right to add something like "Winner of the Samaniego Prize for Contributions to Signature Theory - 2043" to one's resume, as sufficient inducement for readers to spend at least a few minutes considering the problems I will mention. (Don't worry about the year of the prize; a generous endowment has been added to my will which will sustain this prize indefinitely. Indeed, because of this endowment, I am able to extend the range of the prize to any contribution to signature theory that I or my descendents deem to be "not bad.")

8.2 A Retrospective Overview

The signature of a system of order n whose components have i.i.d. lifetimes with common distribution F has been defined as an n-dimensional probability vector $\mathbf{s}$ whose ith element is $s_i = P(T = X_{i:n})$, where T is the system's lifetime and $X_{i:n}$ is the ith ordered component lifetime. As the order statistics of a random sample are stochastically ordered, it is clear from the definition of signatures that probability vectors **s** which place most of their weight on the larger integers in the set $\{1, 2, \ldots, n\}$ will correspond to the better performing systems simply because these systems will tend to fail later, that is, they will fail upon the failure of one of the larger order statistics. It has been shown that these particular proxies for system designs give rise to representations of systems' survival functions, and also of the systems' density and failure rate functions when the underlying component distribution F is absolutely continuous. These representations are used as essential tools in studying the performance of individual systems in i.i.d. components and in comparing such systems with each other. In the latter context, it is shown in Chapter 4 that the existence of certain ordering relationships between the signatures of two systems ensures that a similar relationship holds between the systems' lifetimes. Such preservation results are established for stochastic, hazard-rate and likelihood-ratio ordering between signature vectors of the same size. The sufficient conditions of these preservation theorems are extended in Section 4.4 to necessary and sufficient conditions on two signature vectors for the aforementioned relationships between system lifetimes to hold. Theorem 3.2, a result that establishes a recursive relationship between a given system's signature and that of a system of an arbitrary larger size having the same lifetime distribution, renders comparisons between systems of different sizes feasible.

Both of the developments mentioned above – the signature-based representations of system behavior and the relationships between signatures that imply or characterize similar relationships between system lifetimes – hold for all coherent systems and hold as well for all stochastic mixtures of coherent systems. It has been argued that the notion of mixed systems is more than a mathematical artifact which extends the reach of some theoretical results of interest and serves as a useful tool in certain optimization problems. Indeed, in applications in which a given system is to be used repeatedly, a mixed system represents a potential selection among systems. Its implementation can be physically realized through a simple process of randomization. In the i.i.d. setting studied here, employing a mixed system involves the selection, in each particular instance of its use, of a coherent system chosen according to a fixed probability distribution. Further, since for any (coherent or mixed) system in i.i.d. components, there exists a mixture of k-out-of-n systems with the same lifetime distribution, one can restrict attention to the class of k-out-of-n systems in carrying out the randomized selection of a coherent system at each stage of the application. In Chapter 7, examples of problems are given in which the optimal system relative to a chosen criterion function is not a coherent system but rather a nondegenerate mixture of k-out-of n systems. Thus, in selected circumstances, a particular mixed system can exhibit better expected behavior than any competing coherent system and can thus be reasonably recommended for practical use. The natural domain of applicability of mixed systems is in settings in which the opportunity to select a system for a particular purpose occurs repeatedly.

The various forms of stochastic ordering considered in Chapter 4 have the common characteristic of generating only a partial order among signature vectors or system lifetimes. Even the most liberal of these orderings, the "st" order, will not apply to all possible pairs of coherent systems, and there are uncountably many pairs of mixed systems that are not comparable via stochastic ordering. In Section 5.4, this limitation is addressed through the consideration of an alternative metric between system lifetimes. If T_1 and T_2 are the lifetimes of two mixed systems in i.i.d. components and the orders of these systems are potentially different, then T_1 is said to stochastically precede T_2 if and only if $P(T_1 \leq T_2) \geq 1/2$. Three characteristics of the "sp" metric that make it especially useful in comparing systems are that (i) any two mixed systems of arbitrary size are necessarily comparable, that is, the first is better than, equivalent to or worse than the second, and (ii) the relevant probability $P(T_1 \leq T_2)$ is independent of the underlying common component lifetime distribution F, that is, it is distribution-free (provided only that F is continuous) and (iii) a closed form expression for computing $P(T_1 \leq T_2)$ is available (and is given in Lemma 5.2). This type of comparison offers a potential refinement of comparisons via the traditional stochastic orderings when the latter yield inconclusive results.

The comparative reliability of communication networks is a research area that abounds with problems having both theoretical interest and practical importance. As is typical in the field, a given network is pictured as an undirected graph with a certain number of vertices and with a set of edges connecting different pairs of vertices. The primary problem of interest is the determination of the probability that a given set of vertices can communicate with another set. Two scenarios of special interest are the "two-terminal" problem, where interest is restricted to the question of whether two particular vertices are connected, and the "all terminal" problem, where the probability that each vertex can communicate with every other vertex is of primary interest. In these and other communication network problems, much attention has been given to the problem of computing the reliability of the network of interest. In Chapter 6, our focus is directed at one particularly efficient mode of computation of network reliability – Satyanarayana's theory of dominations. The main goal of that chapter is to identify an explicit relationship between the domination vector $\mathbf{d}$ of a given network and its signature vector $\mathbf{s}$ in the form $\mathbf{s} = g(\mathbf{d})$. Such a formula allows one to exploit simultaneously the computational efficiency of dominations and the utility and interpretive power of signatures in the comparative analysis of networks. In Theorem 6.1, the relationship between dominations and signatures is clarified through an explicit expression of the form $\mathbf{s} = \mathbf{P}^{-1}\mathbf{M}^{-1}\mathbf{d}$, where the matrices $\mathbf{P}$ and $\mathbf{M}$ are specifically identified. The comparison of the networks displayed in Figure 7.3 provides a striking illustration of the potential benefits of the joint use of these two tools.

The generality of the main result in Chapter 6 should be noted. In that chapter, we first call attention to the two different forms generally used to represent the reliability polynomial of a mixed system based on components with i.i.d. lifetimes in Chapter 2. The standard form and pq-form of these polynomials were displayed in equations (2.23) and (2.24). Then the relationship was derived between the domination vector $\mathbf{d}$ and the signature vector $\mathbf{s}$ which, respectively, define the coefficients of the reliability polynomials of a given communication network in standard and pq forms. This relationship, in the form $\mathbf{s} = g(\mathbf{d})$, is displayed explicitly in Theorem 6.1. The fact that this theorem applies equally to the respective coefficients of the reliability polynomials of mixed systems was not stated explicitly in Chapter 6, but is readily apparent from the algebraic developments in that chapter. The general problem solved in Chapter 6 is that of obtaining the exact relationship between the vectors $\mathbf{d}$ and $\mathbf{s}$ in the two polynomials

$$\sum_{j=1}^{n} d_j p^j \quad \text{and} \quad \sum_{j=1}^{n} \left(\sum_{i=n-j+1}^{n} s_i \right) \binom{n}{j} p^j q^{n-j}.$$

This is, of course, precisely the same problem one faces when transforming the reliability polynomial of a mixed system from standard to pq form. It is thus apparent that the relationship between $\mathbf{d}$ and $\mathbf{s}$ in Theorem 6.1 provides the required link in the "systems" setting as well. This connection makes it possible to exploit in tandem the computational advantages of dominations and the broad utility of signatures in the comparative analysis of mixed systems.

Chapter 7 is dedicated to a particular problem in the area of Reliability Economics. Specifically, we are interested in the problem of finding optimal system designs relative to a class of criterion functions depending on both a system's performance and its cost. The criterion functions employed (see (7.5)) depend on a system's design solely through its signature, and they have the desirable property that they are increasing functions of system performance and decreasing functions of system cost. In a case of special interest, the criterion function can be viewed as a system's "performance per unit cost" (PPUC), but the class of criteria considered includes functions that admit to a variety of other interpretations. The optimization problem considered in Chapter 7 is divided into two mutually exclusive cases, and the precise nature of the optimal design is obtained for each. In the first case (corresponding to $r = 1$, i.e., the PPUC case alluded to above), it is shown that the criterion function can be maximized by a given coherent system (indeed, by a particular k-out-of-n system), while in the complementary case, where $r \neq 1$, an optimal system may be represented as a stochastic mixture of at most two k-out-of-n systems. Several examples are given in which the class of coherent systems are dominated by a particular mixed system and are thus suboptimal. Finally, the problem of estimating the features of the underlying component lifetime distribution F upon which the criterion function depends is treated in Section 7.5. The availability of an auxiliary sample of N i.i.d. component lifetimes is assumed. For the particular case in which the measure of performance used in the criterion function is the expected lifetime of the system (and thus depends on F only through the expected order statistics $\{\mu_{i:n}, i = 1, \ldots, n\}$ of component lifetimes), consistent, asymptotically normal estimates of the parameters $\mu_{i:n}$ are obtained. This leads to the important practical conclusion that the system that maximizes the estimated criterion function will be, for arbitrary $\varepsilon > 0$, an ε-optimal system relative to the true criterion function if N is sufficiently large.

8.3 Desiderata

The theory and applications of system signatures treated in this monograph have been developed under the assumption that the systems on which we have focused are based on components with i.i.d. lifetimes. Since this is an overarching assumption and since signature vectors are well defined with or without this assumption, it is natural to explore possible generalizations of signatures in which the i.i.d. assumption is relaxed. We will begin such an exploration in this section with a view toward making the case that certain generalizations are in fact both feasible and useful. Before tackling this issue, however, it seems worth expanding upon the defense of signatures as defined herein. In Chapter 3, we argued that signatures based on an i.i.d. assumption on component lifetimes have the conceptual benefit of "evening the playing field" among system designs we might wish to compare. Further, because the signature provides information about a system that is a function of the system design alone, it is a valuable measure of system characteristics that can be useful quite apart from the consideration of the behavior of the components one might deal with in practice. For example, if one system has a signature vector that stochastically dominates that of a second system, then the fact that the second system performs better in a particular application constitutes an indication that the lifetimes of the components are either exhibiting some form of dependence or are quite differently distributed or both. Since information on the behavior of a system's components is not always available or easy to obtain, the insight about components gained from observed system behavior can be helpful. Finally, when the i.i.d. assumption is a reasonable approximation to the true behavior of a system's components, signature-based calculations of the system's theoretical behavior can be expected to provide useful approximations.

One generalization that can be developed involves replacing the i.i.d. assumption by the assumption that component lifetimes are exchangeable. As noted by Kochar, Mukerjee and Samaniego [51], the representation in (3.5) of the system survival function holds under this less stringent assumption, that is, the identity

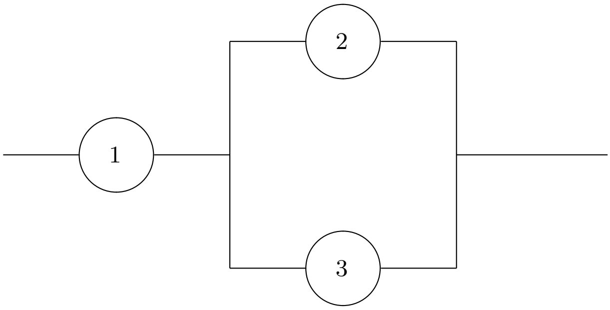
$$P(T > t) = \sum_{i=1}^{n} s_i P(X_{i:n} > t) \tag{8.1}$$

holds when component lifetimes are exchangeable. Thus, one possible direction of further research is to seek to establish the results presented in this monograph under the weaker assumption of exchangeability. However, since exchangeability is but another way to quantify the notion that components behave in a similar manner, such generalizations are not likely to make an appreciable difference in practical applications. Let us, therefore, consider for a moment a more useful generalization, that is, the relaxation of the i.i.d. assumption to the case in which components have independent lifetimes with possibly different lifetime distributions. The signature of an n-component system whose components have independent lifetimes is defined as before, that is, the signature vector $\mathbf{s}$ is an n-dimensional probability vector whose ith element is given by $s_i = P(T = X_{i:n})$, where T is the failure time of the system and $X_{i:n}$ is the lifetime of the ith component to fail. However, representation results such as those in Chapter 3, while not entirely lost, will emerge in a somewhat more cumbersome form. For example, the representation in (8.1) becomes

$$P(T > t) = \sum_{i=1}^{n} s_i P(T > t \mid T = X_{i:n}) , \tag{8.2}$$

and the conditional probability in (8.2) can be computationally complex. The signature vector itself may be computed as the sum of the probabilities of all permutations of the component failure times that correspond to system failure upon the ith component failure. Since the signature vector provides an indication of how long a system will tend to last, it is useful to have it in hand. Although there are no existing results under the sole assumption of independent component lifetimes that state that the domination of one signature over another in some stochastic sense implies some form of domination of the respective system lifetimes, it still makes some intuitive sense to utilize the system with the dominating signature. The potential for theoretical results in this setting will not be pursued further here. Instead, we turn to an example of this latter setting in which the signature vector and the survival function are obtained for the system displayed in the figure below.

Fig. 8.1. A 3-component series-parallel system



Example 8.1. Consider the 3-component coherent system in Figure 8.1 above. Let us assume that the system's components have independent exponential lifetimes X_1, X_2, X_3 with

$$X_i \sim \text{Exp}(\lambda_i) \quad \text{for } i = 1, 2, 3 \, .$$

We first compute the signature vector of the system. Note that the system will fail upon the first component failure if and only if $T = X_1$. Thus, the permutations of the component failure times that result in system failure upon the first component failure are $\{X_1 < X_2 < X_3\}$ and $\{X_1 < X_3 < X_2\}$. We thus obtain s_1 as

$$\begin{aligned} s_1 &= \int_0^\infty \int_{x_1}^\infty \int_{x_2}^\infty \lambda_1 \lambda_2 \lambda_3 \exp\{-(\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3)\} \; \mathrm{d}x_3 \mathrm{d}x_2 \mathrm{d}x_1 \\ &\quad + \int_0^\infty \int_{x_1}^\infty \int_{x_3}^\infty \lambda_1 \lambda_2 \lambda_3 \exp\{-(\lambda_1 x_1 + \lambda_2 x_2 + \lambda_3 x_3)\} \, \mathrm{d}x_2 \mathrm{d}x_3 \mathrm{d}x_1 \\ &= \frac{\lambda_1 \lambda_2}{(\lambda_1 + \lambda_2 + \lambda_3)(\lambda_2 + \lambda_3)} + \frac{\lambda_1 \lambda_3}{(\lambda_1 + \lambda_2 + \lambda_3)(\lambda_2 + \lambda_3)} \\ &= \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \, . \end{aligned}$$

It follows that

$$s_2 = \frac{\lambda_2 + \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \, .$$

The survival function of the system may be computed directly as

$$\begin{aligned} P(T > t) &= P(X_1 > t, X_1 < X_2 < X_3) + P(X_1 > t, X_1 < X_3 < X_2) \\ &\quad + P(X_1 > t, X_2 < X_1 < X_3) + P(X_3 > t, X_2 < X_3 < X_1) \\ &\quad + P(X_1 > t, X_3 < X_1 < X_2) + P(X_2 > t, X_3 < X_2 < X_1). \end{aligned} \tag{8.3}$$

The first two probabilities in (8.3) correspond to system failure upon the first component failure (since $T = X_{1:3}$ only if $X_{1:3} = X_1$). A typical calculation of such probabilities would proceed as follows:

$$
\begin{aligned}
&P(X_1 > t, X_1 < X_2 < X_3) \\
&\quad = \int_t^\infty \int_{x_1}^\infty \int_{x_2}^\infty \lambda_1\lambda_2\lambda_3 \exp\{-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_3 x_3\}\ \mathrm{d}x_3 \mathrm{d}x_2 \mathrm{d}x_1 \\
&\quad = \int_t^\infty \int_{x_1}^\infty \lambda_1\lambda_2 \exp\{-\lambda_1 x_1 - (\lambda_2 + \lambda_3)x_2\}\ \mathrm{d}x_2 \mathrm{d}x_1 \\
&\quad = \frac{\lambda_2}{\lambda_2 + \lambda_3} \int_t^\infty \lambda_1 \exp\{-(\lambda_1 + \lambda_2 + \lambda_3)x_1\}\ \mathrm{d}x_1 \\
&\quad = \frac{\lambda_1\lambda_2}{(\lambda_1 + \lambda_2 + \lambda_3)(\lambda_2 + \lambda_3)} \exp\{-(\lambda_1 + \lambda_2 + \lambda_3)t\}\ .
\end{aligned}
$$

From this we may infer that the first two terms on the RHS of (8.3) add to

$$
P(T > t, T = X_{1:3}) = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \exp\{-(\lambda_1 + \lambda_2 + \lambda_3)t\}\ . \tag{8.4}
$$

The expression in (8.4) is, of course, equal to

$$
s_1 \times P(T > t \mid T = X_{1:3})\ .
$$

Proceeding similarly, one may obtain an expression for $P(T > t, T = X_{2:3})$ which is equivalent to the sum of the last four terms in (8.3). Evaluating the integrals associated with these four terms yields

$$
\begin{aligned}
P(T > t, T = X_{2:3}) = {} & \exp\{-(\lambda_1 + \lambda_3)t\}(1 - \exp\{-\lambda_2 t\}) \\
& + \exp\{-(\lambda_1 + \lambda_2)t\}(1 - \exp\{-\lambda_3 t\}) \\
& + \frac{\lambda_2 + \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3} \exp\{-(\lambda_1 + \lambda_2 + \lambda_3)t\}\ .
\end{aligned} \tag{8.5}
$$

Adding (8.4) and (8.5), we obtain the final expression

$$
\begin{aligned}
P(T > t) = {} & \exp\{-(\lambda_1 + \lambda_3)t\}(1 - \exp\{-\lambda_2 t\}) \\
& + \exp\{-(\lambda_1 + \lambda_2)t\}(1 - \exp\{-\lambda_3 t\}) \\
& + \exp\{-(\lambda_1 + \lambda_2 + \lambda_3)t\}\ .
\end{aligned} \tag{8.6}
$$

The reader will notice that the expression in (8.6) can also be obtained utilizing three independent Bernoulli variables associated with the events $\{X_i > t\}$ for $i = 1, 2$ and 3.

Problems at the next level of generalization, where the i.i.d. assumption is relaxed in its entirety, are likely to resist solution for some time. There are a variety of reasons for this. First, the modeling of dependence in lifetime distributions is itself a challenging problem, with only a few models available that are both tractable and easily interpreted. Among such models, the Marshall-Olkin [56] multivariate exponential (MVE) model is the best known. But even this model, which is well-motivated as a shock model and highly tractable, has found rather limited applicability in practice. One of the reasons for this is the complexity of the model in higher dimensions, where, in the most general case, $2^n - 1$ unknown parameters are required to describe the distribution of an n-dimensional vector of component lifetimes. In computing signature vectors under an MVE assumption, there is a further difficulty. Since the MVE is not absolutely continuous with respect to Lebesgue measure of the appropriate dimension, and indeed gives positive probability to the events that two or more component lifetimes are equal (via the action of a shock which causes several components to fail simultaneously), the term $P(T = X_{i:n})$ has some ambiguity. One could impose some convention on the interpretation of the term (such as that the event occurs if $i - 1$ failures preceded the failure of the system, and one or more components then fail simultaneously, causing the system to fail). But whatever convention is adopted, it is clear that the calculation of signature vectors will be substantially more complex than when component lifetimes are i.i.d. according to some continuous distribution F. Although some tractability is lost in using continuous multivariate lifetime models, the fact that all the component lifetimes are different with probability 1 at least removes the ambiguity discussed above. Time will tell whether the extension of the notion of signatures to the general multivariate domain proves feasible and useful.

8.4 Some Additional Related Literature

There are a variety of further applications of system signatures which merit mention. We present a brief summary.

Boland [16] derives the signatures of indirect majority systems and executes a comparison of such systems with direct majority systems of the same size. He proves that the signature vector of an indirect majority system of odd order n is symmetric about $(n + 1)/2$, and uses this fact to show that, for $n = R \times S$, the expected lifetime of an n-component indirect majority system exceeds the expected lifetime of a direct majority system of size n when the components have i.i.d. lifetimes with a common exponential distribution. Some of this work is presented without proof in Section 5.1 and is applied to the problem studied there.

Shaked and Suarez-Llorens [66] compare the information content of reliability experiments when components are assumed to have i.i.d. lifetimes distributed according to a two-parameter exponential distribution. They introduce the “convolution ordering” and provide sufficient conditions in terms of this ordering for one experiment to be more informative than another. They also use conditions on the signature of a system to obtain certain information inequalities. Specifically, when a system has a signature vector of the form $(0, \ldots, 0, s_k, \ldots, s_n)$ and its components have i.i.d. exponential lifetimes, its lifetime is dominated in the information ordering by a k-out-of-n system with similarly distributed components. An analogous result is shown to hold for the dispersive ordering.

Belzunce and Shaked [7] define the new and useful concept of "failure profiles" for studying the behavior of systems with independent but not necessarily identically distributed component lifetimes. In describing their main results, I will define terms slightly differently and more simply than is done in the referenced paper (eliminating, for example, their use of the term "admissible"), but the results to be described are isomorphic to theirs. A *failure profile* of a coherent system is a pair (I, i), where I is a set of components and $i \notin I$, such that I is a path set of τ and $I \cup \{i\}$ is a cut set of τ. Belzunce and Shaked demonstrate the relevance of failure profiles in two standard formulations of component importance. In their Theorem 2.5, they obtain a useful representation of the density of the lifetime of a system based on components with independent lifetimes in terms of its collection of failure profiles and the individual densities and distribution functions of the system's components. This result generalizes the representation (3.8) of the density of the system lifetime in the i.i.d. case. In their Theorem 3.5, they prove the likelihood ratio ordering between two competing systems (assuming only independent component lifetimes from distributions that are allowed to vary) under specific conditions on the underlying component distributions and on the failure profiles of the two systems. They utilize this latter result to establish a likelihood ratio ordering result for two systems with i.i.d. component lifetimes whose respective signature vectors have a particular form.

Khaledi and Shaked [48]) study the behavior of the conditional residual system lifetime given that a certain number of components are known to be working. The motivation for this study is the fact that, for some systems, it is possible to design a warning mechanism which alerts the user, before the system fails, that at least a certain number of components are still functioning. The authors' main interest is the comparison of two systems conditional on such information. They provide conditions on the signature vectors of the two n-component systems, and on the component distributions, which ensure that the conditional system lifetimes of two competing systems, given that at least $n - i + 1$ components are functioning, are stochastically ordered. For example, they prove the following result.

Theorem 8.1. *Let F_1 and F_2 be two continuous distributions on $(0, \infty)$. Let τ_1 and τ_2 be coherent systems of order n based on components with i.i.d.*

lifetimes

$$X_1, X_2, \ldots, X_n \sim F_1 \quad \textit{and} \quad Y_1, Y_2, \ldots, Y_n \sim F_2 \ ,$$

and let

$$X_{1:n}, X_{2:n}, \ldots, X_{n:n} \quad \textit{and} \quad Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$$

be the corresponding order statistics. Denote the signatures of τ_1 *and* τ_2 *by* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *and their lifetimes* T_1 *and* T_2*. Suppose* $\mathbf{s}_1$ *is of the form* $(0, \ldots, 0, s_{1j}, \ldots, s_{1n})$ *and* $\mathbf{s}_2$ *is of the form* $(0, \ldots, 0, s_{2j}, \ldots, s_{2n})$*. If* $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$ *and* $F_1 \leq_{hr} F_2$*, then for* $i \leq j$*,*

$$F_{T_1 - y \mid X_{i:n} > y} \leq_{st} F_{T_2 - y \mid Y_{i:n} > y} \ .$$

They also obtain the following complementary result for component distributions F_1 and F_2 that are "reverse hazard rate ordered" (denoted by $\leq_{\mathrm{rh}}$), that is, for which $F_2(t) \ / \ F_1(t)$ is increasing in t.

Theorem 8.2. *Let* F_1 *and* F_2 *be two continuous distributions on* $(0, \infty)$*. Let* τ_1 *and* τ_2 *be coherent systems of order* n *based on components with i.i.d. lifetimes*

$$X_1, X_2, \ldots, X_n \sim F_1 \quad \textit{and} \quad Y_1, Y_2, \ldots, Y_n \sim F_2 \ ,$$

and let

$$X_{1:n}, X_{2:n}, \ldots, X_{n:n} \quad \textit{and} \quad Y_{1:n}, Y_{2:n}, \ldots, Y_{n:n}$$

be the corresponding order statistics. Denote the signatures of τ_1 *and* τ_2 *by* $\mathbf{s}_1$ *and* $\mathbf{s}_2$ *and their lifetimes* T_1 *and* T_2*. Suppose* $\mathbf{s}_1$ *is of the form* $(s_{11}, \ldots, s_{1i}, 0, \ldots, 0)$ *and* $\mathbf{s}_2$ *is of the form* $(s_{21}, \ldots, s_{2i}, 0, \ldots, 0)$*. If* $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$ *and* $F_1 \leq_{rh} F_2$*, then for* $i \leq j$*,*

$$F_{T_1 - y \mid X_{j:n} > y} \leq_{st} F_{T_2 - y \mid Y_{j:n} > y} \ .$$

Further, under specific conditions on the signature vector of the system, Khaledi and Shaked [48] obtain upper and lower bounds for

$$E[T - y \mid X_{i:n} > y] \ ,$$

where T and $X_{i:n}$ are the lifetime and ith ordered component failure time of a given system.

Navarro and Shaked [58] utilize system signatures in showing hazard-rate ordering among independently drawn minima $\{X_{1:1}, X_{1:2}, \ldots, X_{1:n}, \ldots\}$ and in studying the limiting behavior of failure rates of selected systems. They develop a representation of the system survival function as a linear combination of the survival functions of minima such as those above, calling the vector of coefficients in the expression the "minimal signature" of the system. For signatures as defined in this monograph, conditions are given under which the ratio of the failure rates of an n-component system with a given signature and that of a k-out-of-n system is asymptotically equal to 1.

8.5 Some Open Problems of Interest

8.5.1 The Ordering of Expected System Lifetimes

In section 5.4, we treated the comparison of systems via stochastic precedence. This metric has the appealing characteristic of rendering all pairs of mixed systems comparable, making it always possible to judge one system as better or worse than (or equivalent to) the other. It would, of course, be useful to have definitive results using the even simpler and most commonly used metric for system performance, the expected system lifetime. It should be noted that comparisons of the expected lifetimes of two systems may be "too rough" a comparison in some problems, as it completely ignores the variability in system lifetime. Further, it is possible that a system whose expected lifetime exceeds that of a second system will be less reliable than the second system at the systems' planned mission time. It is nonetheless of interest to know when one could expect that, on average, one system will last longer than another. While the condition $ET_1 \leq ET_2$, where T_1 and T_2 are the lifetimes of the two systems involved, is a fairly weak stochastic relationship (implied, for example, by $T_1 \leq_{\mathrm{st}} T_2$), questions about this ordering are likely to arise more often in applications than questions about the more stringent relationships discussed in Chapters 4 and 5. Thus, the goal of finding conditions which guarantee that the above ordering of expectations holds seems worthy of attention. Boland and Samaniego [20] discuss this problem and note that for two given systems having components with i.i.d. lifetimes $\sim F$, it is possible for $ET_1 \leq ET_2$ when $F = F_1$ and for $ET_2 \leq ET_1$ when $F = F_2$. However, they prove the following result for a particular group of small systems.

Theorem 8.3. *Consider two mixed systems of order $n = 3$ based on coherent systems in i.i.d. components with lifetime distribution F. Denote their respective signature vectors as $\mathbf{s}_1$ and $\mathbf{s}_2$. Then $ET_1 \leq ET_2$ for all distribution functions F if and only if $\mathbf{s}_1 \leq_{st} \mathbf{s}_2$.*

The extension of this result to systems of arbitrary order n has not been shown, nor have counterexamples been identified which demonstrate that the result fails to hold for other values of n. The former possibility depends on the special properties of the spacings between order statistics and appears to be quite challenging. Counterexamples are inherently quirky, so it is difficult to assess the level of difficulty in showing that the theorem above does not hold for general n if, in fact, that is the case.

8.5.2 Other Preservation Results

In Section 4.2, it was shown that three specific versions of the stochastic ordering of system signatures carry over to the lifetimes of these systems. There are, of course, numerous other orderings for which preservation results may hold (or fail to hold). Shaked and Shantikumar [65] discuss a host of alternative orderings of univariate distributions. Among these are mean residual life ordering, reverse hazard rate ordering, dispersive ordering, the Laplace transform and the moment generating function orderings, convex ordering and star-shaped ordering. The latter two orderings are also discussed in Chapter 4 of Barlow and Proschan [6]. These and a number of other formulations of univariate ordering have been found useful in reliability. If ORD represents any given ordering among them, then it would be of interest to know whether or not the implication $\mathbf{s}_1 \leq_{\mathrm{ORD}} \mathbf{s}_2 \Rightarrow T_1 \leq_{\mathrm{ORD}} T_2$ holds for systems in i.i.d components. I am inclined to classify problems consisting of proofs or counterexamples for these implications as dollar-valued. For all those disposed to think about such problems, it seems appropriate to say at this time: ready, set, go!

8.5.3 The limiting monotonicity of $r_T(t)$

In Section 5.3, the asymptotic behavior of the failure rate of an arbitrary mixed system was examined, and its limiting value was explicitly identified. This result was established by Block, Dugas and Samaniego [11] using the failure rate representation in (3.11). A natural question that arises in this same context concerns the potential relationship between the monotonicity of the component failure rate $r(t)$ and that of the system's failure rate $r_T(t)$. More specifically, it would be of interest to identify conditions that imply that, for sufficiently large t, the system failure rate $r_T(t)$ is strictly increasing (decreasing) if and only if the common component failure rate $r(t)$ is eventually strictly increasing (decreasing).

8.5.4 Further Results on Stochastic Precedence

In section 5.4, it was shown that, using the metric of stochastic precedence, any pair of mixed systems based on components with i.i.d. lifetimes $\sim F$ are comparable, with either one being superior to the other or the two systems being sp-equivalent. Stochastic precedence is well defined when both systems are based on components with i.i.d. lifetimes with differing component distributions F_1 and F_2. Indeed, the representation in (3.28) was shown to hold under these more general conditions. However, the computation of $P(T_1 \leq T_2)$ is considerably more challenging in this latter scenario. Results facilitating the comparison of two system lifetimes when the systems are based on components with independent lifetimes but different distributions would be a worthwhile extension of the results in section 5.4. Hollander and Samaniego [43] demonstrate the feasibility of such generalizations, providing a formula for the exact calculation of the probability $P(T_1 \leq T_2)$ when F_2 is in the class of Lehmann alternatives to F_1, that is, when $F_2(t) \equiv [F_1(t)]^k$ for some $k > 0$. The comparative analysis of systems in the general setting in which $F_1 \neq F_2$ will require further investigation.

8.5.5 Uniformly Optimal Networks

Let us, for concreteness, limit our discussion to the "all-terminal" problem for communication networks, that is, to the problem of determining whether or not all the vertices of a given network can communicate with each other. Among all networks of a given size, that is, within the class of $G(v, n)$ networks with v vertices and n edges, a network is said to be *uniformly optimal* if the probability that all vertices can communicate is maximal. Even under the simplifying assumption that that edges are independent and have a common reliability p, the problem of identifying uniformly optimal networks (that is, networks that are optimal for all $p \in (0, 1)$) in the class $G(v, n)$ is an open problem and appears to be a quite challenging one. In certain special cases, the problem has been solved, but results to date are quite limited. Boesch et al. [14], for example, identified a uniformly optimal network (UON) in the class $G(v, v+1)$. The general problem is complicated by the fact that, for some values of v and n, no uniformly optimal network exists, as demonstrated by Myrvold et al. [57]. Thus, the open problems that remain include the problem of characterizing those classes of networks for which a uniformly optimal network exists and, given such a class, identifying the UON explicitly. Because of the challenging nature of these problems, certain intermediate problems are also of interest. For example, direct comparisons among two or more networks of special interest or between two subclasses of networks of the same size, can be of use in particular applications.

In Chapter 6, it was shown that the signature vectors of competing networks can be a useful tool in comparing their reliability. To my knowledge, the tool has not yet been applied in the search for UONs. To illustrate the utility of signatures in this context, we provide a brief illustration in a problem alluded to above. Let us consider the class of networks in the class $G(5, 6)$. Three particular networks in this class are displayed below.

Fig. 8.2. Network $G_1(5,6)$

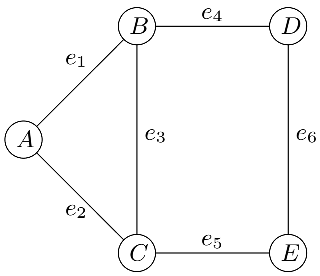


Fig. 8.3. Network $G_2(5,6)$

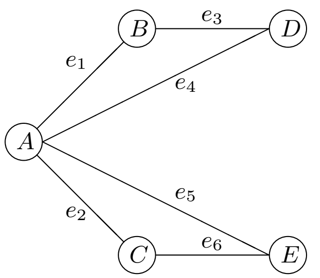
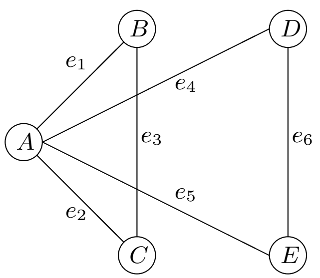


Fig. 8.4. Network $G_3(5,6)$



The three networks above can be compared using the associated reliability polynomials computed via domination theory. Alternatively, one can identify the signatures of the three systems as $\mathbf{s}_1 = (0, 4/15, 11/15, 0, 0, 0)$, $\mathbf{s}_2 = (0, 2/5, 3/5, 0, 0, 0)$ and $\mathbf{s}_3 = (0, 2/5, 3/5, 0, 0, 0)$, from which we see that G_2 and G_3 are equivalent and that both are inferior to G_1 (with $\mathbf{s}_i \geq_{\text{lr}} \mathbf{s}_1$ for $i = 2$ and 3). In larger problems, where existence and uniqueness questions remain unresolved and available methods of finding UONs, if they exist, amount to numerical searches, the ability to establish the superiority of one network over another by comparing their signatures, and the possibility of optimizing network reliability as a function of the signature vector, should offer some hope of successfully attacking these challenging problems.

8.5.6 Other Problems in Reliability Economics

I'll begin by making brief mention of two problems in Reliability Economics that are related to the problem treated in Chapter 7, both involving the search for an optimal system design relative to a criterion such as that in (7.5) under specified constraints. One obvious class of open problems involves maximizing the criterion function (7.5) among systems that are mixtures of a fixed sub-collection of coherent systems. Another class of open problems would involve maximizing the criterion function under a budgetary constraint such as $\sum_{i=1}^{n} c_i \leq K$. Both of these problems are of practical interest, as the selection of a system will often be restricted to the choice among certain available systems and mixtures thereof, and there are often budgetary limits that restrict the selection of the system one might purchase. In either of these constrained scenarios, the optimal system is likely to differ from the optimal systems identified in Chapter 7. Problems involving the characterization of optimal solutions in constrained Reliability Economics contexts constitute a set of interesting open problems of some practical importance.

Example 8.2. As an illustration of the first of these problems, consider the problem of selecting among stochastic mixtures of the two coherent systems of order $n = 3$ having signature vectors $\mathbf{s}_1 = (1/3, 2/3, 0)$ and $\mathbf{s}_2 = (0, 2/3, 1/3)$ respectively. Let us take $r = 1$ in the criterion function in (7.5). Without loss of generality, we set $c_1 = 1$ and allow c_2 and c_3 to be arbitrary values satisfying $1 < c_2 < c_3$. We will take the vector $\mathbf{a}$ to be equal to $(1/4, 1/2, 3/4)$, the expected order statistics of the Uniform distribution $U[0, 1]$. Then the criterion function of the two systems above will be

$$m_1 = \frac{5/12}{1/3 + c_2(2/3)} \quad \text{and} \quad m_2 = \frac{7/12}{c_2(2/3) + c_3(1/3)} .$$

A mixed system giving weights p and $(1 - p)$ to these two systems will have criterion function equal to

$$m_3 = \frac{p(5/12) + (1 - p)(7/12)}{p(1/3) + c_2(2/3) + c_3(1/3)(1 - p)} .$$

It's easy to verify that

$$m_1 < m_3 \quad \text{if and only if} \quad c_3 < \frac{7}{5} + c_2 \, \frac{4}{5} ,$$

while

$$m_2 < m_3 \quad \text{if and only if} \quad c_3 > \frac{7}{5} + c_2\,\frac{4}{5}\,.$$

It follows that if $1 < c_2 < c_3$ and $c_3 = 7/5 + c_2(4/5)$, both coherent systems above, as well as all stochastic mixtures of them, yield the same value of the criterion function, while for any other choice of $c_3 > c_2 > 1$, the criterion function is uniquely maximized by one of the coherent systems or the other. In all cases, there exists a coherent system (among the two available) that is optimal relative to the chosen criterion function.

The example above suggests the conjecture that, when $r = 1$ in the criterion function in (7.5), there exists a system in any collection of coherent systems which will be optimal within the class of all mixtures of these systems. This conjecture agrees with the result in Theorem 7.1 in the case that one seeks an optimal system among the stochastic mixtures of all coherent systems. While it is not readily apparent that Theorem 7.2 generalizes in the same way, it seems reasonable to conjecture that, when $r \neq 1$ in (7.5), an optimal system within any collection of coherent systems and their mixtures can be found within the class of mixtures of two systems in the collection. The method of proof used in establishing Theorem 7.2 (showing, essentially, that any mixture of three systems in the collection can be improved upon by an appropriate mixture of two) may well be successful in showing this.

We will also comment briefly on the problem of searching for an optimal system under cost constraints. In the criterion function in (7.5), the cost of a given n-component system design is quantified in terms of the positive constants $c_1 < c_2 < \cdots < c_n$. Note that if the constraint $\sum_{i=1}^{n} c_i \leq K$ will place no restriction on the choice of system if in fact $c_n \leq K$. This is the case because the optimal system in the unconstrained problem is either a k-out-of-n system costing $c_k \leq K$ or a mixture of two k-out-of-n systems costing $pc_i + (1-p)c_j \leq K$. Thus, for any value of r in (7.5), the system that optimizes the criterion function overall will satisfy the constraint and is thus obviously optimal in the constrained problem. At the other extreme, if $K < c_1$, then there is no system that will satisfy the constraint. Intermediate problems in which $c_1 < K < c_n$ will require individual optimal solutions that may well differ from the optimal solution in the unconstrained problem. We conjecture that, in such constrained problems, the optimal solutions will be of a form similar to those given in Theorems 7.1 and 7.2, and that the methods of proof utilized in those theorems can be adapted to obtain the new results. Without repeating the commentary above, we mention that a similar set of considerations arise when the budgetary constraint of interest is on the total cost of a mixed system, that is, has the form $\sum_{i=1}^{n} c_i s_i \leq K$.

The problem of finding an optimal design while accounting for both performance and cost is but one of many optimization problems of interest in the general area of Reliability Economics. Among topics in reliability in which performance and cost enter in a central way are the areas of maintenance and repair. The varied policies employed in these areas (including, for example, block replacement policies and maintenance through the use of spare parts) have both reliability and economic implications. While there is a literature on optimization on these topics, it remains to be seen whether the problems of interest can benefit from formulations based on system signatures.

8.5.7 Wholly New Stuff

Hey, don't be greedy. I've got to leave something for me to do!

References

[1] Agrawal, A., Barlow, R.E.: A Survey of Network Reliability and Domination Theory. Operations Research, **32**, 478–492 (1984)

[2] Aki, S., Hirano, K.: Lifetime Distributions of Consecutive-k-out-of-n : F Systems. Nonlinear Analysis - Theory, **30**, 555–562 (1997)

[3] Arcones, M., Kvam, P., Samaniego, F. J.: Nonparametric Estimation of a Distribution Subject to a Stochastic Precedence Constraint. Journal of the American Statistical Association, **97**, 170–182 (2002)

[4] Ath, Y., Sobel, M.: Some Conjectured Uniformly Optimal Reliable Networks. Probability in the Engineering and Informational Sciences, **14**, 375–383 (2000)

[5] Barlow, R.E., Proschan, F.: Mathematical Theory of Reliability. John Wiley and Sons, New York (1965)

[6] Barlow, R. E., Proshan, F.: Statistical Theory of Reliability and Life Testing. To Begin With Press, Silver Springs, MD (1981)

[7] Belzunce, F., Shaked, M.: Failure Profiles of Coherent Systems. Naval Research Logistics, **51**, 477–490 (2003)

[8] Birnbaum, Z.W., Esary, J,. Marshall, A.: Stochastic Characterization of Wearout for Components and Sysytems. Annals of Mathematical Statistics, **37**, 816–825 (1966)

[9] Birnbaum, Z. W., Esary, J. D., Saunders, S. C.: Multicomponent Systems and Structures, and their Reliability. Technometrics, **3**, 55–77 (1961)

[10] Block, H., Dugas, M., Samaniego, F. J.: Characterizations of the Relative Behavior of two Systems via Properties of their Signature Vectors. In: Balakrishnan, N., Castillo, E., Sarabia, J. M. (ed) Advances in Distribution Theory, Order Statistics and Inference. Birkhauser, Boston, 279–289 (2006)

[11] Block, H., Dugas, M., Samaniego F. J.: Signature-related Results on System Failure Rates and Lifetimes. In: Nair, V. N. (ed) Advances in Statistical Modeling and Inference: Essays in Honor of Kjell A. Doksum. World Scientific, Singapore, 115-130 (2007)

[12] Block, H., Joe, H.: Tail behavior of the failure rate functions of mixtures. Lifetime Data Analysis, **3**, 269–288 (1997)

[13] Block, H., Li, Y., Savits. T.: Initial and final behavior of the failure rate functions for mixtures and systems. Journal of Applied Probability, **40**, 721–740 (2003)

[14] Boesch, F. T., Li, X., Suffel, C.: On the Existence of Uniformly Optimally Reliable Networks. Networks, **21**, 181–191 (1991)

[15] Boland, P. J.: Majority Systems and the Condorcet Jury Theorem. The Statistician, **26**, 81–88 (1989)

[16] Boland, P. J.: Signatures of Indirect Majority Systems. Journal of Applied Probability, **38**, 597–603 (2001)

[17] Boland, P. J., El-Newehi, E.: Component Redundancy vs. System Redundancy in the Hazard Rate Ordering. IEEE – Transactions on Reliability, **44**, 614–619 (1995)

[18] Boland, P. J., El-Newehi, E., Proschan, F.: Application of the Hazard Rate Ordering in Reliability and Order Statistics. Journal of Applied Probability, **31**, 180–192 (1994)

[19] Boland, P. J., Proschan, F., Tong, Y. L.: Linear Dependence in Consecutive k-out-of-n Systems. Probability in the Engineering and Informational Sciences, **4**, 391–397 (1990)

[20] Boland, P. J., Samaniego, F. J.: The Signature of a Coherent System and its Applications in Reliability. In: Soyer, R., Mazzuchi, T., Singpurwalla, N. (ed) Mathematical Reliability: An Expository Perspective. Kluwer Academic Publishers, Boston 1–29 (2004)

[21] Boland, P. J., Samaniego, F. J.: Stochastic Ordering Results for Consecutive k-out-of-n Systems. IEEE Transactions on Reliability, **TR-53**, 1–4 (2004)

[22] Boland, P. J., Samaniego, F. J., Vestrup, E. M.: Linking Dominations and Signatures in Network Reliabity Theory. In: Lindqvist, B., Doksum, K. (ed) Mathematical and Statistical Methods in Reliability. World Scientific, Singapore, 89–103 (2003)

[23] Boland, P. J., Shaked, M., Shanthikumar, J. G.: Stochastic Ordering of Order Statistics. In: Handbook of Statistics, Vol. 16. Academic Press, San Diego, 89–103 (1998)

[24] Chang, C. J., Cui, L. R., Hwang, F. K.: New Comparisons in Birnbaum Importance for Consecutive k-out-of-n Systems. Probabability Engineering and Information Science, **13:(2)**, 187–192 (1999)

[25] Chang, C. J., Cui, L. R., Hwang, F. K.: Reliabilities of Consecutive k-out-of-n Systems. Kluwer Academic Publishers, Dorrecht (2000)

[26] Chao, M., Fu, J., Koutras, M.: Survey of Reliability Studies of Consecutive k-out-of-$n : F$ and related systems. IEEE Transactions on Reliability, **TR-34**, 120–127 (1995)

[27] Chiang, D., Niu, S. C.: Reliability of Consecutive k-out-of-n Systems. IEEE Transactions on Reliability, **TR-30**, 87–89 (1981)

[28] Colbourn, C. J.: The Combinatorics of Network Reliability. Oxford University Press, Oxford (1987)

[29] Deeter, D. L., Smith, A. E.: Economic Design of Reliable Networks. IIE Transactions on Reliability and Quality Management, **30**, 1161–1174 (1998)

[30] Derman, C., Lieberman, G., Ross, S.: On the Consecutive k-out-of-$n : F$ System. IEEE Transactions on Reliability, **TR-31**, 57–63 (1982)

[31] Du, D. Z., Hwang, F. K., Jung, Y.: Optimal Consecutive-k-out-of-$(2k + 1) : G$ Cycles. Journal of Global Optimization, **19:(1)**, 51–60 (2001)

[32] Du, D. Z., Hwang, F. K.: A Direct Algorithm for Computing Reliabilities of Consecutive k-out-of-n Cycles. IEEE Transactions on Reliability, **TR-37**, 70–72 (1988)

[33] Dugas, M., Samaniego F. J.: On Optimal System Design in Reliability-Economics Frameworks. Naval Research Logistics **54(5)**, 568–582 (2007)

[34] Epstein, S., Sobel, M.: Life testing. Journal of the American Statistical Association, **48**, 486–502 (1953)

[35] Feller, W.: An Introduction to Probability Theory and its Applications. Wiley and Sons, New York (1950)

[36] Feller, W.: Introduction to Probability Theory and its Applications, Third Edition. Wiley and Sons, New York (1968)

[37] Ferguson, T. S.: A Course in Large Sample Theory. Chapman & Hall, London (1996)

[38] Fernholz, L. T.: Von Mises Calculus for Statistical Functionals. Springer-Verlag, New York (1983)

[39] Gera, A. E.: A Consecutive k-out-of-n : G System with Dependent Elements – a Matrix Formulation and Solution. Reliability Engineering System Safety, **68:(1)**, 61–67 (2000)

[40] Grenander, U.: On the Theory of Mortality Measurement, Part II. Scandinavian Journal of Actuarial Science, **39**, 125–153 (1956)

[41] Helmers, R.: Edgeworth Expansions for Linear Combinations of Order Statistics. In: Mathematical Centre Tract 105. Mathematisch Centrum, Amsterdam (1982)

[42] Hewitt, E., Stromberg, K.: Real and Abstract Analysis. Springer-Verlag, New York (1965)

[43] Hollander, M., Samaniego, F. J.: A Nonparametric, Signature-Based Method for Comparing the Reliability of Arbitrary Systems via Stochastic Precedence. Technical Report, Department of Statistics, University of California, Davis (2006)

[44] Hollander, M., Samaniego, F. J.: The Use of Stochastic Precedence in the Comparison of Engineered Systems. Proceedings of the International Conference on Mathematical Methods in Reliability, Glasgow, Scotland (2007)

[45] Hutson, A. D., Ernst, M. D.: The Exact Bootstrap Mean and Variance of an L-Estimator. Journal of the Royal Statistical Society, Ser. B, **62**, 89–94 (2000)

[46] Joag-dev, K., Kochar, S., Proschan, F.: A General Composition Theorem and its Applications to Certain Partial Orderings of Distributions. Statistics and Probability Letters, **22**, 111–119 (1995)

[47] Johnson, R.: Stress-Strength Models for Reliability. In: Krishniah, P. R. (ed) Handbook of Statistics, Vol. 7: Quality Control and Reliability. Elsevier, New York, 27–54 (1988)

[48] Khaledi, B., Shaked, M.: Ordering Conditional Lifetimes of Coherent Systems. Journal of Statistical Planning and Inference (in press) (2006)

[49] Kirmani, S., Kochar, S.: Some New Results on Spacings from Restricted Families of Distributions. Journal of Statistical Planning and Inference, **46**, 47–57 (1995)

[50] Kleitman, D., Markowsky, G.: On Dedekind's problem: the number of isotone Boolean functions. II. Transactions of the American Mathematical Society, A, **213**, 373–390 (1975)

[51] Kochar, K., Mukerjee, H., Samaniego, F.J.: The 'signature' of a coherent system and its application to comparisons among systems. Naval Research Logistics, **46**, 507–523 (1999)

[52] Koutras, M., Papastavridis, S. G.: Consecutive k-out-of-n : F Systems and their Generalizations. In: Misra, K.B. (ed) New Trends in System Reliability Evaluation, Elsevier, Amsterdam, 228–248 (1993)

[53] Kuo, W., Prasad, V. R., Tillman, F. A., Hwang, C.: Optimal Reliability Design, Cambridge University Press, Cambridge, MA (2001)

[54] Kvam, P., Samaniego, F. J.: On the Inadmissibility of Empirical Averages as Estimators in Ranked Set Sampling, Journal of Statistical Planning and Inference, **36**, 39–55 (1993)

[55] Mangel, M., Samaniego, F. J.: Abraham Wald's work on Aircraft Survivability, Journal of the American Statistical Association, **79**, 259–271 (with discussion) (1984)

[56] Marshall, A., Olkin, I.: A Multivariate Exponential Distribution. Journal of the American Statistical Association, **62**, 30–44 (1967)

[57] Myrvold, W., Cheung, K.H., Page, L. B., Perry, J. E.: Uniformly most reliable networks do not always exist. Networks, **21**, 417–419 (1991)

[58] Navarro, J., Shaked, M.: Hazard Rate Ordering of Order Statistics and Systems. Journal of Applied Probability, **43**, 391–408 (2006)

[59] Proschan, F.: Theoretical Explanation of Observed Decreasing Failure Rate. Technometrics, **5**, 375–383 (1963)

[60] Randles, R., Wolfe, D.: Introduction to the Theory of Nonparametric Statistics. Wiley and Sons, New York (1991)

[61] Samaniego, F. J.: On Closure of the IFR Class Under Formation of Coherent Systems. IEEE Transactions on Reliability Theory, **TR-34**, 69–72 (1985)

[62] Samaniego, F. J.: Estimation Based on Autopsy Data from Stress-Strength Experiments. In: Quality Technology and Quantitative Management, 4(1), 1-15 (2007)

[63] Satyarananaya, A., Prabhakar, A.: A New Topological Formula and Rapid Algorithm for Reliability Analysis of Complex Networks. IEEE Transactions on Reliability, **TR-30**, 82–100 (1978)

[64] Satyanarayana, A., Chang, M. K.: Network Reliability and the Factoring Theorem. Networks, **13**, 107–120 (1983)

[65] Shaked, M., Shanthikumar, J.G.: Stochastic Orders and Their Applications. Academic Press, San Diego (1994)

[66] Shaked, M., Suarez-Llorens, J.: On the Comparison of Reliability Experiments Based on the Convolution Order. Journal of the American Statistical Association, **98**, 693–702 (2003)

[67] Shorack, G. R., Wellner, J. A.: Empirical Processes with Applications to Statistics. Wiley and Sons, New York (1986)

[68] Singh, H., Singh, R. S.: On Allocation of Spares at Component Level versus System Level. Journal of Applied Probability, **34**, 283–287 (1997)

[69] Stigler, S. M.: Simon Newcomb, Percy Daniell, and the History of Robust Estimation 1885-1920. Journal of the American Statistical Association, **68**, 872–879 (1973)

[70] Tong, Y. L.: A Rearrangement Inequality for the Longest Run, with an Application to Network Reliability. Journal of Applied Probability, **22**, 386-393 (1985)

[71] Usher, J., Kamal, A., Sayed, H.: Cost Optimal Preventive Maintenance and Replacement Scheduling. IIE Transactions on Reliability and Quality Management, **30**, 1121-1128 (1998)

[72] van der Vaart, A. W.: Asymptotic Statistics. Cambridge University Press, Cambridge (1998)

[73] van Zwet, W. R.: A Strong Law for Linear Functions of Order Statistics, Annals of Probability, **8**, 986–990 (1980)

[74] Wang, G.: A Proof of Boesch's Conjecture. Networks, **24**, 277–284 (1994)

[75] Zelen, M. (ed): Statistical Theory of Reliability. The University of Wisconsin University Press, Madison (1963)

[76] Zuo, M. J.: Reliability and Component Importance of a Consecutive-k-out-n system. Microelectronics Reliability, **33:(2)**, 243–258 (1993)

Index

Early Titles in the

INTERNATIONAL SERIES IN OPERATIONS RESEARCH & MANAGEMENT SCIENCE

Frederick S. Hillier, Series Editor, *Stanford University*

Saigal/ *A MODERN APPROACH TO LINEAR PROGRAMMING*

Nagurney/ *PROJECTED DYNAMICAL SYSTEMS & VARIATIONAL INEQUALITIES WITH APPLICATIONS*

Padberg & Rijal/ *LOCATION, SCHEDULING, DESIGN AND INTEGER PROGRAMMING*

Vanderbei/ *LINEAR PROGRAMMING*

Jaiswal/ *MILITARY OPERATIONS RESEARCH*

Gal & Greenberg/ *ADVANCES IN SENSITIVITY ANALYSIS & PARAMETRIC PROGRAMMING*

Prabhu/ *FOUNDATIONS OF QUEUEING THEORY*

Fang, Rajasekera & Tsao/ *ENTROPY OPTIMIZATION & MATHEMATICAL PROGRAMMING*

Yu/ *OR IN THE AIRLINE INDUSTRY*

Ho & Tang/ *PRODUCT VARIETY MANAGEMENT*

El-Taha & Stidham/ *SAMPLE-PATH ANALYSIS OF QUEUEING SYSTEMS*

Miettinen/ *NONLINEAR MULTIOBJECTIVE OPTIMIZATION*

Chao & Huntington/ *DESIGNING COMPETITIVE ELECTRICITY MARKETS*

Weglarz/ *PROJECT SCHEDULING: RECENT TRENDS & RESULTS*

Sahin & Polatoglu/ *QUALITY, WARRANTY AND PREVENTIVE MAINTENANCE*

Tavares/ *ADVANCES MODELS FOR PROJECT MANAGEMENT*

Tayur, Ganeshan & Magazine/ *QUANTITATIVE MODELS FOR SUPPLY CHAIN MANAGEMENT*

Weyant, J./ *ENERGY AND ENVIRONMENTAL POLICY MODELING*

Shanthikumar, J.G. & Sumita, U./ *APPLIED PROBABILITY AND STOCHASTIC PROCESSES*

Liu, B. & Esogbue, A.O./ *DECISION CRITERIA AND OPTIMAL INVENTORY PROCESSES*

Gal, T., Stewart, T.J., Hanne, T. / *MULTICRITERIA DECISION MAKING: Advances in MCDM Models, Algorithms, Theory, and Applications*

Fox, B.L. / *STRATEGIES FOR QUASI-MONTE CARLO*

Hall, R.W. / *HANDBOOK OF TRANSPORTATION SCIENCE*

Grassman, W.K. / *COMPUTATIONAL PROBABILITY*

Pomerol, J-C. & Barba-Romero, S./ *MULTICRITERION DECISION IN MANAGEMENT*

Axsäter, S./ *INVENTORY CONTROL*

Wolkowicz, H., Saigal, R., & Vandenberghe, L. / *HANDBOOK OF SEMI-DEFINITE PROGRAMMING: Theory, Algorithms, and Applications*

Hobbs, B.F. & Meier, P. / *ENERGY DECISIONS AND THE ENVIRONMENT: A Guide to the Use of Multicriteria Methods*

Dar-El, E. / *HUMAN LEARNING: From Learning Curves to Learning Organizations*

Armstrong, J.S. / *PRINCIPLES OF FORECASTING: A Handbook for Researchers and Practitioners*

Balsamo, S., Personé, V., & Onvural, R./ *ANALYSIS OF QUEUEING NETWORKS WITH BLOCKING*

Early Titles in the
INTERNATIONAL SERIES IN
OPERATIONS RESEARCH & MANAGEMENT SCIENCE
(Continued)

Bouyssou, D. et al. / *EVALUATION AND DECISION MODELS: A Critical Perspective*

Hanne, T. / *INTELLIGENT STRATEGIES FOR META MULTIPLE CRITERIA DECISION MAKING*

Saaty, T. & Vargas, L. / *MODELS, METHODS, CONCEPTS and APPLICATIONS OF THE ANALYTIC HIERARCHY PROCESS*

Chatterjee, K. & Samuelson, W. / *GAME THEORY AND BUSINESS APPLICATIONS*

Hobbs, B. et al. / *THE NEXT GENERATION OF ELECTRIC POWER UNIT COMMITMENT MODELS*

Vanderbei, R.J. / *LINEAR PROGRAMMING: Foundations and Extensions, 2nd Ed.*

Kimms, A. / *MATHEMATICAL PROGRAMMING AND FINANCIAL OBJECTIVES FOR SCHEDULING PROJECTS*

Baptiste, P., Le Pape, C. & Nuijten, W. / *CONSTRAINT-BASED SCHEDULING*

Feinberg, E. & Shwartz, A. / *HANDBOOK OF MARKOV DECISION PROCESSES*: *Methods and Applications*

Ramík, J. & Vlach, M. / *GENERALIZED CONCAVITY IN FUZZY OPTIMIZATION AND DECISION ANALYSIS*

Song, J. & Yao, D. / *SUPPLY CHAIN STRUCTURES: Coordination, Information and Optimization*

Kozan, E. & Ohuchi, A. / *OPERATIONS RESEARCH/ MANAGEMENT SCIENCE AT WORK*

Bouyssou et al. / *AIDING DECISIONS WITH MULTIPLE CRITERIA: Essays in Honor of Bernard Roy*

Cox, Louis Anthony, Jr. / *RISK ANALYSIS: Foundations, Models and Methods*

Dror, M., L'Ecuyer, P. & Szidarovszky, F. / *MODELING UNCERTAINTY: An Examination of Stochastic Theory, Methods, and Applications*

Dokuchaev, N. / *DYNAMIC PORTFOLIO STRATEGIES: Quantitative Methods and Empirical Rules for Incomplete Information*

Sarker, R., Mohammadian, M. & Yao, X. / *EVOLUTIONARY OPTIMIZATION*

Demeulemeester, R. & Herroelen, W. /*PROJECT SCHEDULING: A Research Handbook*

Gazis, D.C. / *TRAFFIC THEORY*

Zhu/*QUANTITATIVE MODELS FOR PERFORMANCE EVALUATION AND BENCHMARKING*

Ehrgott & Gandibleux/ *MULTIPLE CRITERIA OPTIMIZATION: State of the Art Annotated Bibliographical Surveys*

Bienstock/ *Potential Function Methods for Approx. Solving Linear Programming Problems*

Matsatsinis & Siskos/ *INTELLIGENT SUPPORT SYSTEMS FOR MARKETING DECISIONS*

Alpern & Gal/ *THE THEORY OF SEARCH GAMES AND RENDEZVOUS*

Hall/*HANDBOOK OF TRANSPORTATION SCIENCE - 2^{nd} Ed.*

Glover & Kochenberger/ *HANDBOOK OF METAHEURISTICS*

Graves & Ringuest/ *MODELS AND METHODS FOR PROJECT SELECTION: Concepts from Management Science, Finance and Information Technology*

Hassin & Haviv/ *TO QUEUE OR NOT TO QUEUE: Equilibrium Behavior in Queueing Systems*

Gershwin et al/ *ANALYSIS & MODELING OF MANUFACTURING SYSTEMS*

Maros/ *COMPUTATIONAL TECHNIQUES OF THE SIMPLEX METHOD*

Early Titles in the
INTERNATIONAL SERIES IN
OPERATIONS RESEARCH & MANAGEMENT SCIENCE
(Continued)

Harrison, Lee & Neale/ *THE PRACTICE OF SUPPLY CHAIN MANAGEMENT: Where Theory and Application Converge*

Shanthikumar, Yao & Zijm/ *STOCHASTIC MODELING AND OPTIMIZATION OF MANUFACTURING SYSTEMS AND SUPPLY CHAINS*

Nabrzyski, Schopf & Wglarz/*GRID RESOURCE MANAGEMENT: State of the Art and Future Trends*

Thissen & Herder/ *CRITICAL INFRASTRUCTURES: State of the Art in Research and Application*

Carlsson, Fedrizzi, & Fullér/ *FUZZY LOGIC IN MANAGEMENT*

Soyer, Mazzuchi & Singpurwalla/ *MATHEMATICAL RELIABILITY: An Expository Perspective*

Chakravarty & Eliashberg/ *MANAGING BUSINESS INTERFACES: Marketing, Engineering, and Manufacturing Perspectives*

Talluri & van Ryzin/ *THE THEORY AND PRACTICE OF REVENUE MANAGEMENT*

Kavadias & Loch/*PROJECT SELECTION UNDER UNCERTAINTY: Dynamically Allocating Resources to Maximize Value*

Brandeau, Sainfort & Pierskalla/ *OPERATIONS RESEARCH AND HEALTH CARE: A Handbook of Methods and Applications*

Cooper, Seiford & Zhu/ *HANDBOOK OF DATA ENVELOPMENT ANALYSIS: Models and Methods*

Luenberger/ *LINEAR AND NONLINEAR PROGRAMMING, 2^{nd} Ed.*

Sherbrooke/*OPTIMAL INVENTORY MODELING OF SYSTEMS: Multi-Echelon Techniques, Second Edition*

Chu, Leung, Hui & Cheung/ *4th PARTY CYBER LOGISTICS FOR AIR CARGO*

Simchi-Levi, Wu & Shen/*HANDBOOK OF QUANTITATIVE SUPPLY CHAIN ANALYSIS: Modeling in the E-Business Era*

Gass & Assad/ *AN ANNOTATED TIMELINE OF OPERATIONS RESEARCH: An Informal History*

Greenberg/ *TUTORIALS ON EMERGING METHODOLOGIES AND APPLICATIONS IN OPERATIONS RESEARCH*

Weber/ *UNCERTAINTY IN THE ELECTRIC POWER INDUSTRY: Methods and Models for Decision Support*

Figueira, Greco & Ehrgott/ *MULTIPLE CRITERIA DECISION ANALYSIS: State of the Art Surveys*

Reveliotis/ *REAL-TIME MANAGEMENT OF RESOURCE ALLOCATIONS SYSTEMS: A Discrete Event Systems Approach*

Kall & Mayer/ *STOCHASTIC LINEAR PROGRAMMING: Models, Theory, and Computation*

** A list of the more recent publications in the series is at the front of the book **