AD-A278 647

AFIT/GOR/ENS/ENC/94M-19

POINT AND INTERVAL ESTIMATION OF SERIES SYSTEM RELIABILITY USING SMALL DATA SETS

THESIS

Craig Joslyn Willits
Captain, USAF

AFIT/GOR/ENS/ENC/94M-19

DTIC
ELECTE
APR 22 1994
B
D

DTIC QUALITY INSPECTED 3

Approved for public release; distribution unlimited

94-12273

94 4 21 054

Best
Available
Copy

REPORT DOCUMENTATION PAGE

March 1994 | Master's Thesis

POINT AND INTERVAL ESTIMATION OF SERIES SYSTEM RELIABILITY USING SMALL DATA SETS

Craig Joslyn Willits, Captain, USAF

Air Force Institute of Technology
Wright-Patterson AFB OH 45433-6583

AFIT/GOR/ENC/ENS/94M-19

Lt Col J. R. Hays
HQ ACC/XP-JSG/SB
204 Dodd Boulevard, Suite 202
Langley AFB VA 23665-2778

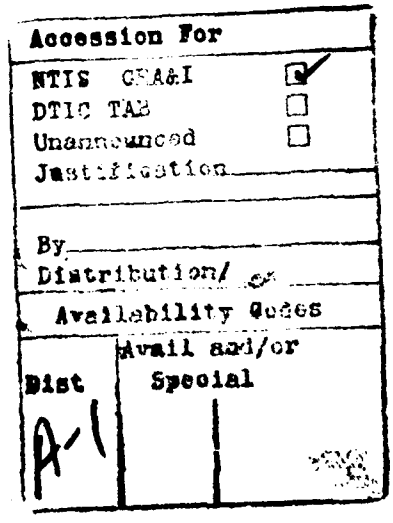
Approved for public release; distribution unlimited

This investigation explored the relative performance of several small-sample point and interval estimators for series system reliability. Among point estimators, the maximum likelihood estimator (MLE) was compared to the corresponding Bayes estimator. In addition, four interval estimators were compared: Easterling's modified maximum likelihood integer estimator, the Lindstrom-Madden estimator, and Bayesian probability interval estimators constructed using approximate beta and Bayes Monte Carlo empirical posterior densities. The relative performance of the point estimators was assessed by comparing their mean square errors. For the four interval estimators, the interval coverage probability and the average interval lower bound were examined. The values of these performance measures were generated for 32 representative series systems using Monte Carlo simulation. The results of the Monte Carlo study showed that the accuracy of the available prior information determines whether a Bayesian or a classical approach should be used to estimate series system reliability. If there is high confidence that the mean of each component prior distribution is within 20 percent of the true component reliability, Bayesian point and interval estimators constructed from an approximate beta posterior should be used. Otherwise, the MLE point estimator and the Lindstrom-Madden interval estimator should be chosen.

Reliability, Statistical Inference, Bayes' Theorem,
Maximum Likelihood Estimation, Weapon System Effectiveness | 95

Unclassified | Unclassified | Unclassified | UL

The views expressed in this thesis are those of the author and do not reflect the official policy of the Department of Defense or the U. S. Government.



AFIT/GOR/ENS/ENC/94M-19

POINT AND INTERVAL ESTIMATION OF SERIES SYSTEM RELIABILITY USING SMALL DATA SETS

THESIS

Presented to the Faculty of the Graduate School of Engineering
of the Air Force Institute of Technology
Air University
In Partial Fulfillment of the
Requirements for the Degree of
Master of Science in Operations Research

Craig Joslyn Willits, B.A.
Captain, USAF

March 1994

Approved for public release; distribution unlimited

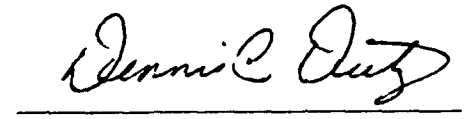
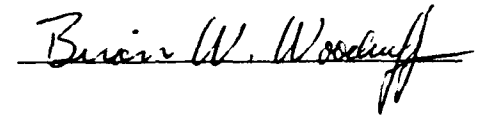
THESIS APPROVAL

STUDENT: Craig J. Willits, Captain, USAF CLASS: GOR-94M

THESIS TITLE: POINT AND INTERVAL ESTIMATION OF SERIES SYSTEM RELIABILITY USING SMALL DATA SETS

DEFENSE DATE: 17 February 1994

COMMITTEE:

Name/Title/Department	Signature
Co-Advisor: ALBERT H. MOORE Professor of Mathematics (Emeritus) Department of Mathematics and Statistics	
Co-Advisor/ENS Representative: DENNIS C. DIETZ, Lt Col, USAF Assistant Professor of Operations Research Department of Operational Sciences	
Reader: BRIAN W. WOODRUFF, Major, USAF Associate Professor of Statistics Department of Mathematics and Statistics	

Acknowledgements

Although the title page of this thesis lists me as the sole author, the project was a team effort. I am grateful for the contributions of several people; without their support, I would have quickly come to grief.

I am deeply indebted to my thesis committee for their guidance. Because I was exposed to their collective expertise, I grew in my ability as a researcher. My co-advisors, Dr Albert Moore and Lt Col Dennis Dietz, were with me step by step from the earliest days of the research; they challenged and encouraged me as the project developed. My reader, Maj Brian Woodruff, provided me with valuable insights.

Besides my committee, several others contributed to this effort. My sponsor, Lt Col Johnny Hays of Headquarters Air Combat Command, suggested this problem, helped me define it and gave me moral support. Col Gregory Parnell, head of the Department of Operational Sciences, made valuable input to my thesis proposal. Professor Daniel Reynolds of the Department of Mathematics and Statistics served as my sounding board during the validation of the catgorical data analysis. Capt Mark Abramson, my friend and classmate, read the draft of this report at several stages; his suggestions greatly improved the clarity of the final product.

Without the sacrifices my family made, this thesis would still be an idea. My wife, Sara, provided me with much prayer and steadfast support. She also read (and reread) the manuscript as it evolved, and made many helpful comments. The unconditional love of my sons, Steven, David and Joseph, helped keep me going through the long hours of hard work.

Finally, I give all praise to God, and dedicate this thesis to His glory. "A man's heart deviseth his way: but the Lord directeth his steps" (Proverbs 16:9).

Craig Joslyn Willits

Table of Contents

List of Figures

List of Tables

List of Symbols

γ	the level of significance/probability ($0 < \gamma < 1$)
$\chi^2_\alpha(\nu)$	the 100α percentile of a chi-squared distribution with ν degrees of freedom
$g(R)$	the prior density of system reliability
$g(R\|\vec{s})$	the posterior density of system reliability
k	the total number of components in the system
n_i	the number of tests of component i
n_i^o	number of prior pseudo-trials of component i
p_i	the failure probability of component i
$\hat{p}_i$	the maximum likelihood estimator (MLE) of p_i
r_i	the number of observed failures of component i
R_i	the reliability of component i ($R_i = 1 - p_i$)
$\hat{R}_i$	the MLE of R_i
R	the system reliability
R_L	the lower confidence/probability limit on system reliability
$\hat{R}$	the MLE of R
s_i	the number of observed successes of component i ($s_i = n_i - r_i$)
s_i^o	number of prior pseudo-successes of component i
z_α	the 100α percentile of a standard normal distribution

Abstract

This investigation explored the relative performance of several small-sample point and interval estimators for series system reliability. Among point estimators, the maximum likelihood estimator (MLE) was compared to the corresponding Bayes estimator. In addition, four interval estimators were compared: Easterling's modified maximum likelihood integer estimator, the Lindstrom-Madden estimator, and Bayesian probability interval estimators constructed using approximate beta and Bayes Monte Carlo empirical posterior densities.

The relative performance of the point estimators was assessed by comparing their mean square errors. For the four interval estimators, the interval coverage probability and the average interval lower bound were examined. The values of these performance measures were generated for 32 representative series systems using Monte Carlo simulation.

The results of the Monte Carlo study showed that the accuracy of the available prior information determines whether a Bayesian or a classical approach should be used to estimate series system reliability. If there is high confidence that the mean of each component prior distribution is within 20 percent of the true component reliability, Bayesian point and interval estimators constructed from an approximate beta posterior should be used. Otherwise, the MLE point estimator and the Lindstrom-Madden interval estimator should be chosen.

POINT AND INTERVAL ESTIMATION OF SERIES SYSTEM RELIABILITY USING SMALL DATA SETS

I. Introduction

Background

The allocation of firepower to a list of key enemy targets is an important part of strategic war planning. When assigning scarce strategic resources to a particular target, operations planners need an accurate assessment of the likelihood that an assigned weapon (the warhead and delivery system) will neutralize that target. This likelihood is called *damage expectancy,* and can be evaluated using the equation

$$D = P_a \times P_k \tag{1}$$

where P_a is the probability that the warhead will detonate on target and P_k is the probability that the detonation will inflict the desired damage level. The quantity P_a is a function of several other probabilities. It is calculated using the equation

$$P_a = A \times P_s \times R \times P_p \tag{2}$$

where A is weapon system availability, P_s is the probability that the weapon system survives launch, R is the weapon system reliability (WSR) and P_p is the probability that the weapon system penetrates enemy defenses [2].

While all aspects of damage expectancy are important, this research focuses on estimating WSR. According to Cardaronella et al., WSR is the probability that a

strategic weapon will detonate near the intended target, with no consideration of the effect of enemy defenses [7]. This means that, for a strategic mission to be successful, the delivery system must get within an acceptable radius of the target center and the warhead must detonate at the proper yield.

Modeling WSR. Headquarters Air Combat Command (ACC) is responsible to the Joint Staff for estimating WSR for weapon systems that include manned bomber aircraft. ACC's WSR model is based on statistical reliability theory. Since damage expectancy is a single-weapon probability, the model assumes the weapon delivery system is composed of three critical elements: the bomber aircraft, the air-launched missile, and a single warhead. The ACC model is comprehensively described in [7]; the remainder of this section is based on material from that source and from [13].

To construct the WSR model, the weapon system is decomposed into a series structure with three independent components. These components correspond to the three critical elements listed above. Each of these components can be viewed as a "subsystem," although the calculation of hardware reliability is not the goal of the model. Rather, each subsystem in the reliability model can be viewed as the contribution to mission success of each of the critical elements of the weapon system. The series model is used because if any of these elements fails to complete its portion of the mission, the mission will not be a success. The summary WSR model is shown in Figure 1.

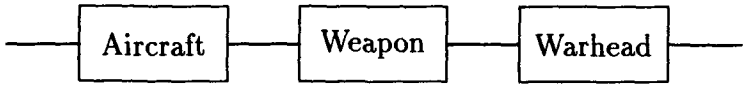


Figure 1. Summary WSR Model.

To meet the Joint Staff's reporting requirements, the aircraft's contribution to mission success is broken down one further level to a series of three components: the avionics system, the weapon release system, and the aircraft vehicle. The mission contribution of a component is declared successful if that component functions

properly until the weapon is deployed. At that point, the contribution of the aircraft to mission success ends, and whatever happens to it after missile launch is not considered in the model.

In contrast to the aircraft's contribution, the weapon's contribution is broken down into two mission phases: captive carry and free flight. (The warhead contribution is considered separately.) The weapon must function properly during both phases for the mission to be successful. The captive-carry phase starts when aircraft power is applied to the aircraft avionics during the operations preflight inspection and lasts until the aircraft launches the missile. The free-flight phase begins immediately upon missile launch and lasts until the planned warhead detonation time. During the free-flight phase, the weapon contribution is further decomposed into a series of components, each of which must function for the entire phase to guarantee mission success. These components are the engine, the fuel system, the navigation/guidance system, the air vehicle, and the fuze mechanism.

When the contributions of all the delivery systems are decomposed to the lowest level, the system devolves to a series structure with ten independent components. This expanded WSR model is shown in Figure 2.

Data Source. Since it is obviously not feasible to collect failure data under combat conditions, such data must come from other sources. ACC's Follow-On Operational Test and Evaluation Phase II (FOT&E II) programs are a primary source of system failure data. Since an FOT&E II program uses fully operational versions of the system being tested, ACC conducts FOT&E II in an attempt to assess the performance of its strategic systems in the most realistic environment possible. Currently, this program provides the only failure data that can be used to estimate WSR (although ACC is currently investigating using ground test data to supplement data from FOT&E II missions).

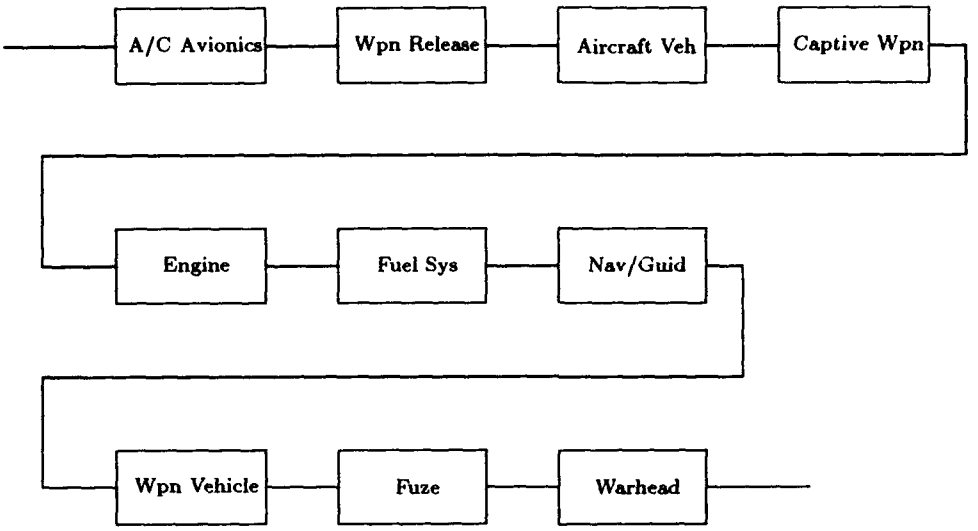


Figure 2. Expanded WSR Model.

Estimating WSR. ACC currently uses classical statistical theory to estimate WSR. A point estimate is obtained using the maximum likelihood estimator (MLE), while Easterling's Modified Maximum Likelihood Integer (MMLI) estimator is used to calculate a 90 percent lower confidence limit [7] (these techniques are fully explained in Chapter II).

Statement of the Problem

Funding for ACC's FOT&E II program has been substantially reduced in recent years. These reductions have severely limited ACC's ability to gather data for use in estimating WSR. This means that, for some newly fielded weapon systems, the failure database is likely to remain very small (that is, less than ten data points) for some time.

In the cases where failure data is scarce, the WSR estimates made using ACC's current approach may produce spurious results. It is well known that classical point estimators decrease in accuracy as sample size decreases [16:2]. Furthermore, because

the MMLI estimator has its roots in asymptotic theory, it is not intended to be used with small data sets [12]; it is often unusable in such a sampling situation because of mechanical difficulties with its implementation (see Chapter II for further discussion).

The need for strategic readiness demands meaningful estimates of WSR. Since the data sets available through the FOT&E II program are small, such estimates cannot be made using the MLE/MMLI approach. Therefore, the problem consists of finding one or more acceptable alternate estimation techniques that can be applied to small samples.

Assumptions and Limitations

The following assumptions and limitations apply to this research effort:

1. The context of this inquiry is limited to weapon systems composed of a combination of bomber aircraft and air-launched missiles (although the results of this study can easily be extended to delivery scenarios involving gravity weapons or ballistic missiles, since they can also be modeled as series structures).
2. In this investigation, all the system components are assumed to have binomial failure distributions. Although time dependence is an important part of the WSR model, the effect of estimating WSR using a mixture of pass-fail and time-dependent failure distributions is not explored. The only system reliability estimators that have been developed for the mixed-distribution case require the time-dependent failure distributions to be exponential; this would require gross assumptions about the mission process. Furthermore, these estimators are either derived using asymptotic approximations or are impractical to implement.
3. This effort focuses on one-sided intervals since they are the limits of interest in most practical reliability estimation situations. However, it is possible to derive either one- or two-sided interval estimates for system reliability.

4. Only series systems are considered in this research effort, since the WSR model is a series structure.

General Approach

This investigation explores the performance of several classical and Bayesian alternatives to ACC's current combination of estimators. (Appendix A contains an overview of Bayesian statistics for the benefit of the reader unfamiliar with the subject.) The performance of the MLE for system reliability is compared to that of the corresponding Bayes estimator. Among interval estimators, the MMLI estimator is compared to three alternatives: the Lindstrom-Madden estimator and Bayesian probability interval estimators constructed using approximate beta and Bayes Monte Carlo empirical posterior distributions.

The relative performance of the point estimators is assessed by comparing their mean square errors. For the four interval estimators, the interval coverage probability and the average interval lower bound are compared. The baseline values of these performance measures are generated using Monte Carlo simulation.

For the Monte Carlo comparison study, 32 representative systems are constructed. The number of components, the underlying component failure probabilities, the component test sample sizes, and the strength and accuracy of the prior information are varied across the systems. Sample sizes are limited to ten or less. Using sets of simulated failure data, the reliability of each system is estimated repeatedly, and performance measures generated for each estimator. The sensitivity of the results of the study to variations in system characteristics is analyzed.

Based on the analysis of estimator performance, the best point and interval estimators are recommended as replacements for the current WSR estimators. These best estimators are selected based on their performance across systems.

Sequence of Presentation

Chapters II and III present a survey of system reliability estimation methods. In Chapter II, classical point and interval estimators are considered, while Chapter III describes Bayesian estimators. Each major system reliability estimation approach is described, and its strengths and weaknesses are discussed with respect to small-sample analysis.

Chapter IV outlines the methodology used in the Monte Carlo comparison study. The simulation design is constructed and the methodology of the comparison is described in detail. Technical issues associated with the simulation are also discussed.

Chapter V presents the findings of the Monte Carlo study. The simulation results are reported and analyzed, conclusions are drawn about method performance, and the preferred estimation methods are identified. Recommendations for further research are also included.

II. Classical System Reliability Estimation

Overview

This chapter surveys the key results in classical system reliability estimation theory. In keeping with the scope of the investigation, the treatment is limited to the case of a series system of independent components, each with a binomial failure distribution.

The survey begins with point estimation. The MLE of system reliability is derived, and its asymptotic variance is developed. In addition, the advantages and disadvantages of using the MLE with small samples are explored.

The discussion of point estimation is followed by three sections on interval estimators. Buehler's exact lower confidence limit (LCL) is presented first, followed by several analytical and bootstrap approximations to the exact limit. In each section, the estimators are described, and their performance with respect to small samples is considered.

Point Estimation

The MLE for System Reliability. In general, the *reliability* of a component is the probability that it will function for a specified period of time. If the component is repeatedly tested over a fixed operating time, its reliability can be modeled with a binomial failure distribution. In mathematical terms, if the ith component in a system has a binomial failure distribution, its reliability is defined as

$$R_i = 1 - p_i \tag{3}$$

where p_i is the probability that the component will fail. Since a series system will not function unless all of its components function, it is easy to see that the reliability

of that system must be the product of the component reliabilities. That is,

$$R = \prod_{i=1}^{k} R_i = \prod_{i=1}^{k} (1 - p_i) \tag{4}$$

where k is the number of components in the system. Since the p_i are unknown parameters, they must be estimated (which, in turn, requires that the R_i and R be estimated). MLEs are usually used to obtain the point estimates, because they have convenient mathematical properties like asymptotic normality and invariance [9:53].

The invariance property of MLEs is useful for developing a point estimator for R. This property states that an arbitrary transformation of an MLE space is the MLE of that tranformation (Mood et al. formally state and prove the invariance theorem [33:285]). Suppose that, for component i, r_i failures are observed in n_i component tests. Since it can be shown that the MLE for p_i is $\hat{p}_i = r_i/n_i$ [32:420], the invariance property can be applied to Equation (3) to get the MLE of the component reliability:

$$\hat{R}_i = 1 - \hat{p}_i \tag{5}$$

A second application of the invariance property to Equations (4) and (5) yields the MLE of system reliability:

$$\hat{R} = \prod_{i=1}^{k} \hat{R}_i \tag{6}$$

The Reliability Estimator's Asymptotic Variance. The variance of $\hat{R}$ is important to the development of certain interval estimation methods. The delta method, which is based on the asymptotic normality property of MLEs, can be used to derive an asymptotic approximation for this variance [9]. Let $\hat{\theta} = \{\hat{\theta}_1, \ldots, \hat{\theta}_k\}$ be the MLE of $\theta = \{\theta_1, \ldots, \theta_k\}$. Also, define the arbitrary function $\xi \equiv g(\theta)$. By the invariance property, $\hat{\xi} \equiv g(\hat{\theta})$ is the MLE of ξ. It can be shown that the variance of $\hat{\xi}$ can be

asymptotically approximated by

$$\hat{\sigma}^2_{\hat{\xi}} = \sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{\partial g}{\partial \theta_i}\right)\Big|_{\hat{\theta}}\left(\frac{\partial g}{\partial \theta_j}\right)\Big|_{\hat{\theta}} \mathrm{Cov}[\hat{\theta}_i, \hat{\theta}_j] \tag{7}$$

[9:52]. Equation (7) can be applied to a reliability context by letting $\hat{\xi} = \hat{R}$, $\hat{\theta} = \{\hat{R}_i\}$ and $\hat{\theta}_i = \hat{R}_i$. This yields

$$\hat{\sigma}^2_{\hat{R}} = \sum_{i=1}^{k}\sum_{j=1}^{k}\left(\frac{\partial R}{\partial R_i}\right)\Big|_{\{\hat{R}_i\}}\left(\frac{\partial R}{\partial R_j}\right)\Big|_{\{\hat{R}_i\}} \mathrm{Cov}[\hat{R}_i, \hat{R}_j] \tag{8}$$

Since the system's components are assumed to be independent, Equation (8) becomes

$$\hat{\sigma}^2_{\hat{R}} = \sum_{i=1}^{k}\left[\left(\frac{\partial \hat{R}}{\partial \hat{R}_i}\right)^2\right]\Big|_{\{\hat{R}_i\}} \cdot \sigma^2_{\hat{R}_i} \tag{9}$$

where $\sigma^2_{\hat{R}_i}$ is the variance of $\hat{R}_i$. This variance can be approximated in the limit by

$$\hat{\sigma}^2_{\hat{R}_i} = \frac{\hat{R}_i(1-\hat{R}_i)}{n_i} \tag{10}$$

[12], so Equation (9) reduces to

$$\hat{\sigma}^2_{\hat{R}} = \hat{R}^2 \sum_{i=1}^{k}\left[\frac{r_i}{n_i(n_i - r_i)}\right] \tag{11}$$

Small-Sample Reliability Point Estimation. The advantages of using MLEs instead of other estimators are treated at length in many advanced statistics texts (such as [16]). These advantages are such that MLEs are the preferred point estimators in a variety of situations. Like any other classical estimator, however, the accuracy of the estimates produced by an MLE decreases as the size of the data set used in the estimate decreases [16:2]. Further, the approximate normality of MLEs may not hold for small samples, so the asymptotic variance of the MLE is likely to

be a poor approximation of its true variance [9:64]. Therefore, the MLE of system reliability should be used with care when data is scarce.

Interval Estimation: Buehler's Exact LCL

Because the point estimator of series system reliability is the product of the estimators of the component reliabilities, it seems intuitive that a confidence interval for R could be constructed by taking the product of the component interval bounds. While this approach will indeed produce a confidence interval, it will not yield the shortest possible interval estimate for R [6]. The construction of this shortest interval is not a simple task.

In 1957, R. J. Buehler proposed a method for obtaining the exact limit of the shortest one-sided confidence interval for the product of two binomial parameters [6]. Buehler's work was generalized and applied to system reliability by Lipow [21] and Steck [42]. The general form of Buehler's method is examined in detail in several sources; the discussion in Crowder et al. [9:189-191] is followed here.

Let $\vec{r} = \{r_i, i = 1, \ldots, k\}$ be a vector of component failure observations from some system test. Further, let $\vec{\rho} = \{\rho_i, i = 1, \ldots, k\}$ be any possible combination of observed failure levels for the system other than that defined by $\vec{r}$. Also, define an arbitrary ordering function $\omega(\cdot)$ such that lower values of $\omega(\cdot)$ are more desirable. Given these definitions, Buehler's exact $100(1-\gamma)$ percent LCL for the reliability of a series system is the solution to the following nonlinear program:

$$\text{minimize} \quad R_L = \prod_{i=1}^{k} \pi_i$$

subject to

$$\sum_{\Omega} \left[\prod_{i=1}^{k} \binom{n_i}{r_i} (1-\pi_i)^{r_i} \pi_i^{(n_i - r_i)} \right] = \gamma$$

$$0 \leq \pi_i \leq 1, \quad i = 1, \ldots, k \tag{12}$$

where $\Omega = \{\vec{\rho} : \omega(\vec{\rho}) \leq \omega(\vec{r})\}$. Crowder et al. suggest that $\omega(\cdot) = \prod_i \rho_i/n_i$ (the MLE of the system failure probability) is an intuitive ordering function [9:191]. This form of $\omega(\cdot)$ does not always produce intuitive results. For example, Reiser and Jaegar find that the resulting interval estimate does not uniformly shorten as the component sample sizes increase; nonetheless, the failure probability estimator is the ordering function most commonly used in the solution of the Buehler minimization problem [37].

Lipow and Riley provide tables of solutions to Equation (12) when the component sample sizes are equal and $k \leq 3$ [22], [23]. For $k > 3$ or large n_i values, solving Equation (12) may not be possible, or may require computational resources beyond those available to most analysts [9:191]. Willits finds that, for series systems, the nonlinear optimization program MINOS can easily be used to find the solution to Equation (12), provided the number of components is less than three and the test samples are not too large; however, a general solution algorithm cannot be implemented for systems with four or more components because of the computational complexity of the problem [48].

Interval Estimation: Approximation Methods

Because Buehler's exact LCL is difficult to obtain in most practical experimental situations, several analytical methods for approximating it have been proposed. Some of the more significant of these approaches are summarized below.

This section does not contain an exhaustive catalog of analytical approximations. Many of the proposed strategies for approximating Buehler's limit have been shown to be ineffective. Only those estimators that have been considered in a wide variety of sources are included. The reader interested in a comprehensive treatment of analytical interval estimation methods is referred to Section 10.3 of [27].

The Lindstrom-Madden Estimator. D. L. Lindstrom and J. H. Madden propose a conceptually simple method for approximating the exact LCL of R. In the Lindstrom-Madden method, the system test is construed as a single binomial experiment where $n^* = \min_i\{n_i\}$ tests produce $r^* = (1 - \hat{R})n^*$ failures. The $100(1-\gamma)$ percent LCL of R can then be calculated, as it is for a single binomial component, by solving for R_L in the following expression [24]:

$$\sum_{i=1}^{r^*} \binom{n^*}{i} (1 - R_L)^i (R_L)^{(n^*-i)} = \gamma \tag{13}$$

It is clear that r^* will not always be an integer; in that case, Equation (13) cannot be used to evaluate R_L. This problem can be avoided by using the identity

$$\begin{aligned} \sum_{i=1}^{r} \binom{n}{i} (1-p)^i p^{(n-i)} &= \frac{\Gamma(n-r)\Gamma(r+1)}{\Gamma(n+1)} \int_0^p \tau^{n-r-1}(1-\tau)^r d\tau \\ &= I(p; n-r, r+1) \end{aligned} \tag{14}$$

where $I(\cdot)$ is the incomplete beta function [17:214-215]. The approximate $100(1-\gamma)$ percent LCL can be found by solving the following equation for R_L [24]:

$$I(R_L; n^* - r^*, r^* + 1) = \gamma \tag{15}$$

Equation (15) can only be used when $r^* \neq n^*$; fortunately, this condition holds in nearly all practical circumstances, since $r^* = n^*$ if and only if $\hat{R} = 0$.

The Lindstrom-Madden estimator has both weaknesses and strengths. According to Winterbottom, Lindstrom-Madden performance deteriorates as the minimum component sample size decreases [50]. Also, Martz and Duran find that the Lindstrom-Madden confidence limits can be far too conservative in certain circumstances [29]. However, Sudakov finds that if r^* is an integer, the Lindstrom-Madden estimator produces Buehler's exact LCL. Sudakov also shows that if r^* is not an

integer, Buehler's limit is bounded by R_L (the solution to Equation (15)) and R'_L, which is the solution to

$$I(R_L; n^* - \lfloor r^*, \lfloor r^* + 1) = \gamma \tag{16}$$

where $\lfloor r^*$ is the greatest integer less than r^*. Thus, Buehler's exact LCL is always bounded by two Lindstrom-Madden limits [43]. (Soms claims that Sudakov's proof is flawed, but that in most practical situations the Lindstrom-Madden solutions still bound the exact limit [39].)

Interval Estimation Based on Asymptotic Normality. Since $\hat{R}$ is a MLE, a straightforward way to obtain a LCL on system reliability is to take advantage of the fact that MLEs are asymptotically normal [16:43-44]. It is clear that if $\hat{R}$ is normal with mean $\hat{R}$ and variance defined by Equation (11), then $(\hat{R} - R)/\sigma_{\hat{R}}$ has an asymptotic standard normal distribution. This means that an approximate $100(1 - \gamma)$ percent LCL can be estimated by solving the following equation for R_L:

$$\Pr\left[\frac{(\hat{R} - R)}{\sigma_{\hat{R}}} \leq R_L\right] = \gamma \tag{17}$$

In the literature, this approach is often referred to as the maximum likelihood (ML) method (although occasionally it is called "linearization" [25]).

An obvious weakness of the ML method is its use of a symmetric distribution with infinite support (the normal distribution) to approximate a random variable with an unsymmetric, bounded distribution (R_L). In other words, it is possible in some situations to produce a LCL that lies outside the unit interval. The next five estimators presented attempt to improve this situation.

The MMLI Estimator. Easterling's MMLI estimator is a modification of the ML approach described in the previous section. As with the Lindstrom-Madden estimator, Easterling conceives of the system as undergoing a single binomial experiment rather than a set of independent component-level experiments. The difference between the MMLI and the Lindstrom-Madden estimators is that Easterling derives r^* and n^* by equating the variance of $\hat{R}$ under the single experiment with the asymptotic variance of the estimator of a binomial probability.

If R is construed as a parameter of a Binomial(n^*, R) distribution, and $\hat{R}$ is the MLE of R, Equation (10) can be applied to get the variance of $\hat{R}$:

$$\sigma^2_{\hat{R}} = \frac{\hat{R}(1-\hat{R})}{n^*} \tag{18}$$

Setting the variances in Equations (11) and (18) equal yields

$$\hat{R}^2 \sum_{i=1}^{k} \left[\frac{r_i}{n_i(n_i - r_i)} \right] = \frac{\hat{R}(1-\hat{R})}{n^*} \tag{19}$$

which, when solved for n^*, produces the expression

$$n^* = \frac{(1-\hat{R})}{\hat{R} \sum_{i=1}^{k} \left[\frac{r_i}{n_i(n_i - r_i)} \right]} \tag{20}$$

It follows immediately that $r^* = (1-\hat{R})n^*$.

The values of r^* and n^* are often not integers. Based on empirical evidence, Easterling concludes that rounding both r^* and n^* up to the next integer improves the approximation of Buehler's limit. Applying the identity in Equation (14), the $100(1-\gamma)$ percent LCL can be found by solving

$$I(R_L; (\lceil n^* - \lceil r^*), (\lceil r^* + 1)) = \gamma \tag{21}$$

for R_L, where $\lceil r^*$ and $\lceil n^*$ are the smallest integers greater than r^* and n^*, respectively [12]. It should be noted that when $\hat{R} = 1$, $n^* = 0$ and r^* is indeterminate, so the MMLI estimator cannot be used. With respect to small-sample analysis, this is a major weakness, since $\hat{R}$ often equals one when n_i values are small and the R_i values are reasonably high.

Asymptotic Expansion. Winterbottom proposes another refinement to the ML method based on Cornish-Fisher asymptotic expansions. This approach improves the approximation of the LCL by using the expansions to correct for bias and skewness. The result is an approximate LCL that, for certain large-sample situations, performs very well. Winterbottom's approximate $100(1-\gamma)$ percent LCL is

$$R_L = \hat{R} \cdot K \tag{22}$$

where

$$K = 1 - \frac{z_{1-\gamma}\sigma}{\sqrt{n^*}} - \frac{(z_{1-\gamma}^2 + 1)\left[\sum_{i=1}^{k} \frac{(1-\hat{R}_i^2)(n^*)^2}{(n_i\hat{R}_i)^2} - 3\sigma^4\right] + \left[\sum_{i=1}^{k} \frac{(1-\hat{R}_i)(1-2\hat{R}_i)(n^*)^2}{(n_i\hat{R}_i)^2}\right]}{6n^*\sigma^2},$$

$\sigma^2 = \sigma_{\hat{R}}^2/\hat{R}^2$, $n^* = \min_i\{n_i\}$ and $z_{1-\gamma}$ is the $100(1-\gamma)$th percentile of the standard normal distribution. Winterbottom notes that the approximation requires $r_i > 0 \ \forall i$. The derivation of Equation (22) is omitted here; for details, see [49] and [50].

The Likelihood Ratio (LR) Lower Bound. Madansky proposes an asymptotic approximation method based on the LR test statistic. For a series system, $\Lambda = R/\hat{R}$ is the LR test statistic for system reliability. Wilks shows that, for general Λ, $-2\ln\Lambda$ has a limiting χ^2 distribution with one degree of freedom [47]. Using Wilks' result, Madansky derives the following $100(1-\gamma)$ percent LCL for R [25]:

$$R_L = \inf\left\{R : -2\ln\Lambda \leq \chi_\gamma^2(1)\right\} \tag{23}$$

Madansky's result, while of historical importance, does not improve the gross approximation of the ML method enough to be practically useful. In addition, Equation (23) is very difficult to solve on a computer [49],[50].

The Approximately Optimal (AO) Nonrandomized LCL. Based on asymptotic theory, Mann develops an approximate LCL that compensates for the case where $r_i = 0$ for some i. (For a series system, most other approximation methods ignore the impact of this case.) Mann's derivation, which is beyond the scope of this discussion, centers around the development of asymptotic approximations for first two moments of the random variable $-\ln R$.

If Ω is the set of components with at least one observed failure in a sample (i.e., $\Omega \equiv \{i : 1 \leq i \leq k, r_i \neq 0\}$), then the approximations of the first two moments of $-\ln R$ are given by

$$\mu_1 = \frac{0.5[1 + (1/\nu)]}{N} + \frac{1 - \hat{R}}{1 - 0.5(1 - \hat{R})} \tag{24}$$

and

$$\mu_2 = \frac{0.5[1 + (1/\nu)]\mu_1}{N} \tag{25}$$

where

$$N = \left[\inf_{i \in \Omega}\{n_i\}\right]\left[1 - 0.5(1 - \hat{R})^2\right]\left[1 - 0.5(1 - \hat{R})\right] \tag{26}$$

and

$$\nu = \left[\inf_{i \in \Omega}\{n_i\}\right]\sum_{i \in \Omega}\frac{1}{n_i} \tag{27}$$

Mann applies Equations (24) and (25) to get the following expression for the AO nonrandomized $100(1 - \gamma)$ percent LCL [26],[27]:

$$R_L = \exp\left\{-\mu_1 \cdot \left[1 - \frac{\mu_2}{9{\mu_1}^2} + \frac{z_{1-\gamma}\sqrt{\mu_2}}{3\mu_1}\right]^3\right\} \tag{28}$$

Poisson Approximation. When the component reliabilities are known to be high, Buehler's exact LCL can be approximated by assuming each r_i has a Poisson distribution. If the sample sizes are equal and large, it can be shown that

$$R_L = 1 - \frac{\chi^2_\gamma(\nu)}{2n^*} \tag{29}$$

where $\nu = 2[(\sum_{i=1}^{k} r_i) + 1]$ and $n^* = \min_i\{n_i\}$; if the sample sizes differ, Equation (29) produces a conservative LCL [50]. The Poisson assumption breaks down when samples are small; for a detailed treatment of Poisson-related techniques, see Section 10.3 of [27].

Using Approximation Methods with Small Samples. All the interval estimators presented in this section, except the Lindstrom-Madden estimator, are based on asymptotic theory. Asymptotically-derived estimators are meant to be used with sufficiently large data sets; when sample sizes are small, the approximation is often inaccurate [9:64]. In some cases, as with the Poisson approximation, the inaccuracy is severe; in other instances, as with ML and LR, the approximations are not the best available even when they are used with a large sample [50].

Some approximate estimators suffer because they fail to account for the case where no failures are observed for one or more components [26]. When $r_i = 0$, ML, MMLI and LR all ignore the contribution of component i to R_L; when $r_i = 0$ for every component, these estimators cannot be used at all. Also, asymptotic expansion is useless if $r_i = 0$ for one or more components. Mann's AO LCL is an attempt to rectify this situation, but its derivation is still based on asymptotic theory.

The only approximate estimator that can be legitimately used for small data sets is the Lindstrom-Madden estimator. While this method has its drawbacks, it is nonetheless a viable alternative to the MMLI method for small test samples.

Interval Estimation: Bootstrap Methods

According to Crowder et al., the use of Monte Carlo simulation to approximate system reliability confidence limits dates to the 1963 work of Rosenblatt. Efron refined and extended the method, and is generally credited with originating the term "bootstrap" [9:192]. In bootstrap interval estimation, Monte Carlo random variates are generated and used to construct an empirical cumulative distribution function (CDF). The appropriate quantile of this CDF is used to approximate the exact confidence limit. The steps in the general procedure, as implemented by Levy and Moore, are as follows:

1. Determine the underlying component failure distributions from the observed data set using the methods of ML.
2. Generate a set of simulation component reliabilities using the component distributions identified in Step 1.
3. Calculate the simulated system reliability as a function of the component reliabilities.
4. Repeat Steps 2 and 3 $N-1$ times, where N is some large number (say, 10000).
5. Rank order the simulated system reliabilities, with the lowest magnitude first.
6. Use the $(N\gamma)$th simulated system reliability as the approximate $100(1-\gamma)$ percent LCL.

While Levy and Moore are concerned primarily with continuous component failure distributions, the above procedure is applicable to systems with binomially-distributed components [20].

Asymptotic Normal Monte Carlo (ANMC). Rice and Moore proposed a variant of the bootstrap method for the all-binomial case. This refinement recognizes that since the parameters of the underlying failure distributions are MLEs, their

values can vary under repeated sampling. The asymptotic normality of these MLEs is used to account for the estimation uncertainty. (The refined method also compensates for cases when $r_i = 0$.) The steps in ANMC are as follows:

1. From the observed test data, determine r_i for $i = 1, \ldots, k$. If $r_i = 0$ for some i, calculate r_i^* by solving

$$(r_i^*)^2 + (2n_i - 3)r_i^* + (n_i - 1)\ln[\chi^2_{(1-\gamma)}(2)] = 0. \tag{30}$$

2. Calculate

$$\hat{R}_i = \begin{cases} 1 - (r_i/n_i) & \text{if } r_i \geq 0 \\ 1 - (r_i^*/n_i) & \text{if } r_i = 0 \end{cases} \tag{31}$$

and

$$\hat{\sigma}_i^2 = \frac{\hat{R}_i(1 - \hat{R}_i)}{n_i}. \tag{32}$$

3. Set N large (say, 10000). For $j = 1, \ldots, N$, generate a set of k standard normal deviates $z_{i,j}$ $(i = 1, \ldots, k)$.
4. Calculate $\tilde{R}_{i,j} = z_{i,j}\sigma_{i,j} + \hat{R}_i$, $\forall(i,j)$.
5. Calculate $\tilde{R}_j = \prod_{i=1}^{k} \tilde{R}_{i,j}$, $\forall(i,j)$.
6. Rank order the values of $\tilde{R}_j$ to get the order statistics $\tilde{R}_{(j)}$.
7. Approximate the lower limit as $R_L = \tilde{R}_{(N\gamma)}$.

Rice and Moore observe a slight upward bias in the confidence intervals ANMC produces [38].

The Modified ANMC Method. Chao and Huwang propose a modification to ANMC to correct its inherent bias. In modified ANMC, Bayes estimates are used for the component reliabilities. (Although this makes the modified method pseudo-Bayesian, it is considered here instead of in Chapter III because it is intended to be

used in a classical context.) The steps for the modified method are the same as those for the original method, except that the following terms are defined differently:

$$\hat{R}_i = 1 - \frac{r_i + \alpha}{n_i + \alpha + \beta} \tag{33}$$

and

$$\tilde{R}_{i,j} = 1 - \frac{\rho_i + \alpha}{n_i + \alpha + \beta} \tag{34}$$

where $\rho_i \sim \text{Bi}(n_i, \hat{R}_i)$, and α and β are fixed prior parameters. Based on empirical evidence, Chao and Huwang find that the modified ANMC method produces the best interval coverage if $0.1 \leq \alpha \leq 0.2$ and $\beta = 0$ [8].

Using the Bootstrap with Small Samples. Martz and Duran find that the performance of the standard bootstrap varies widely, depending on the situation in which it is used; in many cases, it produces quite optimistic intervals [29]. Also, the standard bootstrap does not accomodate the case where one or more components has no failures observed; if $r_i = 0$ for some i, the component is treated as if it had perfect reliability. Rice and Moore rectified this situation with the ANMC method [38]. Unfortunately, the device they use to incorporate the variability of the reliability estimates is based on the asymptotic normality of the estimators, and their method produces slightly optimistic LCLs. Chao and Huwang avoid the use of asymptotic theory when correcting the bias; for component sample sizes of ten, the method appears to yield reasonable results [8]. However, no computational experience has been reported for smaller samples or for unequal sample sizes.

III. Bayesian System Reliability Estimation

Overview

This chapter summarizes Bayesian system reliability estimation theory. As with Chapter II, the methods considered are limited to the case of a series system of independent binomial components. After introductory remarks concerning the selection of the component prior densities and the nature of reliability estimation, methods for determining the posterior density of R are surveyed and evaluated. The chapter concludes with a brief justification for investigating Bayesian alternatives to classical reliability techniques.

Selecting the Prior Density of R

When a component has an underlying binomial failure density, one possible prior probability density function (pdf) is the Beta(α,β) density

$$g(p) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \tag{35}$$

(other forms, such as the uniform density, are used as well). Since a beta prior is a natural conjugate prior for a binomial proportion, it has desirable mathematical properties (see Appendix A). Further, Weiler showed that assumption of a beta prior when the true form is other than beta has negligible effect on the posterior of R [46]. Because of these advantanges, the beta family is most often used in a binomial sampling situation.

When the Beta(α,β) density is used as the prior pdf of a component, its hyperparameters are usually chosen as $\alpha = s_i^o + 1$ and $\beta = n_i^o - s_i^o + 1$. The values of s_i^o and n_i^o, which represent the number of successes and trials in a prior pseudo-test of component i, attempt to adequately translate the subjective prior information about the reliability of component i into numerical values [30:236]. The values of

s_i^o and n_i^o can be determined somewhat objectively (using, for example, test data from a similar system or component), or the estimation process may be completely subjective.

If the subjective approach must be used, the prior hyperparameters should not be assigned values directly. Mosleh and Apostolakis recommend that the source of the prior assessment be the perceived quantiles of the prior distribution. This is because the resulting prior variance tends to be too small otherwise [34]. Building on research by Weiler [46], Waterman et al. (reproduced in [30:236-237]) propose an assessment method for beta hyperparameters which uses the mean and one other quantile. The procedure requires that the assessor assign perceived values of the mean and the fifth or 95th percentile of the prior density. Using tables provided by Waterman et al., these assignments can be translated into equivalent values of s_i^o and n_i^o.

Using Waterman's procedure, Duran and Booker examine the effect on system reliability estimation of faulty subjective assessments of beta prior hyperparameters. They find, not surprisingly, that as the accuracy of the prior assessment worsens, the posterior variance of R increases [11]. Since this adversely affects the quality of the system reliability estimates, extreme care should be taken when subjectively assessing the hyperparameters of the component prior densities.

Bayes Estimators and Probability Intervals

When the posterior pdf of R is determined, point and interval estimates for R can easily be derived. As was shown in Appendix A, the Bayes estimator is the mean of the posterior pdf if a squared-error loss function is assumed. Similarly, the $100(1 - \gamma)$ percent Bayesian lower probability limit is the (100γ)th percentile of the posterior CDF, which is found by solving Equation (49) (see Appendix A). Since these calculations are usually straightforward when the form of the posterior density is known, the primary task in Bayesian system reliability estimation is to derive the

posterior for R. Three methods for generating this posterior can be applied to the all-binomial case. These methods are described in the next section.

Methods for Deriving the Posterior Density of R

The Exact Posterior. Springer and Thompson propose an analytical method for deriving the posterior distribution of series system reliability in the case where all components have binomial failure distributions. Since the posterior derived using this method is exact, the Bayesian point and interval estimates it yields are also exact.

Mellin transform theory can be readily applied to the problem of obtaining the posterior density for R (see Appendix B for a discussion of the Mellin transform). Let $M\{f(x);u\}$ be the Mellin transform of $f(x)$ with respect to the complex parameter u. Springer and Thompson show that

$$M\{g(R|\vec{s});u\} = \prod_{i=1}^{k} M\{g_i(R_i|s_i);u\} \tag{36}$$

where $\vec{s} = \{s_1, \ldots, s_k\}$. If R_i has a Beta$(s_i^o+1, n_i^o-s_i^o+1)$ prior, the Mellin transform of the posterior pdf of R is

$$M\{g(R|\vec{s});u\} = \left[\prod_{i=1}^{k} \frac{(n_i^*+1)!}{s_i^*!}\right] \cdot \frac{1}{\prod_{j=0}^{n-s}(u+s+j)^{c_j}} \tag{37}$$

where $s_i^* = s_i^o + s_i$, $n_i^* = n_i^o + n_i$, $s = \max_i s_i^*$, $n = \max_i n_i^*$, $\mathrm{Re}(u) > -s$, and c_j is the number of times $(u+s+j)$ occurs over the component Mellin transforms. The posterior pdf of R, which is not stated here, can be obtained by evaluating the inversion integral of the Mellin transform in Equation (37).

The mean of the posterior pdf is

$$E[R|\vec{s}] = M\{g(R|\vec{s};u=2\} = \prod_{i=1}^{k}\left(\frac{s_i^*+1}{n_i^*+2}\right) \tag{38}$$

This expression is also the Bayes estimator of R for a squared-error loss function [41].

The closed form of the posterior CDF of R is obtained by integrating the pdf. Springer and Thompson show that this CDF is

$$G(R|\vec{s}) = R^{s+1} \sum_{j=1}^{n-s} \left[\sum_{i=1}^{c_j} A_{j,i} (-ln(R))^{i-1} \right] \tag{39}$$

where

$$A_{j,i} = \frac{1}{(i-1)!} \sum_{m=0}^{c_j - i} \frac{K_{j,i+m}}{(s+j+i)^{m+1}} \tag{40}$$

and

$$K_{j,i+m} = \begin{cases} \frac{1}{(c_j - i - m)!} \frac{d^{c_j - i - m}}{du^{c_j - i - m}} [(u+s+j)^{c_j} M\{g(R|\vec{s}); u\}]|_{u=-s+j} & c_j \neq 0 \\ 0 & c_j = 0 \end{cases} \tag{41}$$

Springer and Thompson develop computer code to evaluate quantiles of this CDF. However, their experience (and that of others, notably Martz et al. [31]) shows that such implementations are numerically unstable, especially for systems with high reliability.

In addition to the application of Mellin transforms, other methods exist for finding the exact posterior CDF of R. While these approaches are of theoretical interest, they are not without their difficulties. Tang and Gupta provide a set of recurrence relations that can be used to find the exact pdf of a product of independent beta random variables [44]; the problem with this approach is that it requires a cumbersome term-by-term integration of the resulting pdf to find the CDF. As another alternative, Wolfe uses a probabilistic argument to develop a multiple integral expression for the exact posterior CDF [51]; however, when k is large, the dimensionality of the integral makes it difficult to solve. A detailed discussion of these methods is omitted. The interested reader is directed to the appropriate references.

The Approximate Beta Posterior. Computationally, it is very difficult to find the exact posterior of R. Springer and others suggest overcoming this difficulty by approximating the exact posterior density of R with a beta density sharing the same first two moments [40:268-269],[45],[50]. According to Springer, such approximations can often be quite close to the exact form [40:268]. Thompson and Haynes report that, at least for moderate sample sizes ($n_i \geq 20$), this beta approximation yields excellent results [45].

Using Springer's approach, the approximate beta posterior can be constructed easily. The first two moments of the exact posterior pdf of R can be found by setting $u = 2, 3$ in the density's Mellin transform. The resulting moments are

$$\mu = \prod_{i=1}^{k} \left(\frac{s_i^* + 1}{n_i^* + 2} \right) \tag{42}$$

and

$$\sigma^2 = \left\{ \prod_{i=1}^{k} \frac{(s_i^* + 1)(s_i^* + 2)}{(n_i^* + 2)(n_i^* + 3)} \right\} - \mu^2 \tag{43}$$

The Beta(α, β) density with these same first and second moments has the parameters

$$\alpha = \frac{\mu^2(1 - \mu) - \sigma^2 \mu}{\sigma^2} \tag{44}$$

and

$$\beta = \frac{\mu(1 - \mu)^2 - \sigma^2(1 - \mu)}{\sigma^2} \tag{45}$$

[40:269].

The Bayes Monte Carlo Empirical Posterior. Martz and Duran propose a Monte Carlo method for approximating a $100(1 - \gamma)$ *lower probability limit*. This approach, which is a variant of the bootstrap, has as its goal the generation of an empirical posterior CDF for R. While Martz and Duran use noninformative priors

in their implementation, their method can be easily adapted to handle informative priors.

The steps in Bayes Monte Carlo are as follows:

1. Set N large (say, 10000). For $j = 1, \ldots, N$, generate the random variate $\tilde{R}_{i,j}$ from a Beta$(n_i - r_i + \alpha_i, r_i + \beta_i)$ distribution, where α_i and β_i are the parameters of the ith prior density.
2. Calculate $\tilde{R}_j = \prod_{i=1}^{k} \tilde{R}_{i,j}$, $\forall j$.
3. Rank order the values of $\tilde{R}_j$ to obtain the order statistics $\tilde{R}_{(j)}$.
4. Approximate the $100(1-\gamma)$ lower probability limit as $R_L = \tilde{R}_{(N\gamma)}$ [29].

Martz and Duran find that this method generally outperforms the classical bootstrap [29].

Bayes Monte Carlo is only intended to produce approximate probability limits. If a Bayes estimate of R is desired, Equation (38) (the first moment of Springer and Thompson's exact posterior distribution) should be used.

The Bayesian Estimation of WSR

Martz and Waller argue convincingly for the use of Bayesian inference in reliability analysis. They claim that the Bayesian approach has the following advantages over classical reliability techniques:

1. If the prior information is valid, the Bayesian inferences are more accurate.
2. Bayesian analyses usually require smaller samples.
3. In a Bayesian test, unacceptable results are a consequence of bad assumptions (i.e., inaccurate prior information) rather than questionable methodology.
4. Bayesian probability intervals for R are easier to obtain and understand than classical confidence intervals [30:170-174].

On the face of it, Martz and Waller's arguments are persuasive, but the Bayesian approach is far from universally accepted. In fact, controversy has raged between Bayesian and classical statisticians for many years as to the validity of the list of "advantages" given above (see Appendix A).

The discussion in Chapter II showed, however, that (1) ML estimation performance suffers when sample sizes decrease, and (2) the Lindstrom-Madden estimator provides the only classical interval estimator that can be legitimately applied when component test samples are small. In the face of these difficulties, it seems reasonable to accept the Bayesian paradigm as a valid alternative to frequentist theory. This increases the list of potential replacement estimators, and avoids having to accept the MLE/Lindstrom-Madden combination by default.

The problem is that there is no small-sample comparison between Bayesian and classical system reliability estimation techniques published in the reliability literature. The remainder of this thesis is devoted to conducting such a study.

IV. The Monte Carlo Comparison Study

Overview

This chapter describes the Monte Carlo study undertaken in this research effort. First, the estimation methods that were compared are identified, as are the measures of merit upon which the comparison was based. The series systems used in the study are described next, followed by a discussion of technical considerations related to the Monte Carlo implementation. The chapter concludes with a description of the methodology used in the simulation.

Candidate Estimation Methods

The purpose of this study was to find some set (or sets) of point and upper-tail interval estimators that performs better for small test samples than those presently used by ACC. Accordingly, the performance of several classical and Bayesian estimators were considered as possible replacements for both the MLE ($\hat{R}$) and the MMLI interval estimator.

The Bayes estimator in Equation (38) was examined as an alternative to the MLE. This estimator was chosen for comparison because it is the first moment of all three posterior forms considered in Chapter III.

Two forms of the Bayesian interval estimator in Equation (49) (see Appendix A) were compared to the MMLI estimator. The first form was constructed from the approximate beta posterior CDF; the second used the BMC empirical posterior CDF (the exact posterior was not used because of the computational difficulties associated with its implementation). Since the Lindstrom-Madden estimator is the only classical interval estimator suitable for use with small test samples, it was also compared to the MMLI estimator. All interval estimators considered were constructed at two levels of significance (0.10 and 0.05) to aid sensitivity analysis.

Measures of Merit

The measure of merit used to compare the point estimators was *mean square error* (MSE) (see Appendix A). Since the Bayes estimator under consideration minimizes the squared-error loss function, MSE was the logical measure to use. The point estimator that produced the smallest estimated MSE for a given system was considered the superior performer for that system.

For the interval estimators, the measures of merit were the *coverage proportion* (the percentage of generated intervals containing the true system reliability) and the *average interval lower bound.* These measures are often used to compare interval estimators (see, for example, Martz and Duran [29]). An interval estimator that produced a coverage proportion less than $(1 - \gamma)$ was considered *optimistic.* For a given system, the interval estimator that produced the largest average interval bound at both significance levels without being optimistic was considered the superior estimator for that system.

Hypothetical System Structures

The basic concepts of experimental design were used to construct a comprehensive set of hypothetical series systems for study. Table 1 shows the design strategy that was used to develop the systems. This strategy was based on a completely randomized 2^5 factorial design. Although a more complex approach could have been taken, only two factor levels were considered so that the number of systems remained manageable. Table 2 presents the details of the randomized design.

The hypothetical systems were developed using the components described in Table 3. These components were constructed so that their true reliabilities were between 0.90 and 0.999, a reasonable range for many real-world applications. The strength and accuracy of the component prior information were determined by the values chosen for s_i^o and n_i^o. For the purposes of this study, it was assumed that setting $n_i^o \approx 25$ provided sufficiently strong prior information; for a weak prior, n_i^o

was fixed at around 10. The accuracy of the prior information was controlled by varying the magnitude of the difference between (s_i^o/n_i^o) and the true component reliability R_i (that is, $|(s_i^o/n_i^o) - R_i|$). To provide a component with accurate prior information, s_i^o and n_i^o were selected so that $|(s_i^o/n_i^o) - R_i| \leq 0.2$; an inaccurate prior was constructed by choosing the hyperparameters $|(s_i^o/n_i^o) - R_i| > 0.2$. Table 4 shows the component-by-component structure of each series system.

Table 1. Design Strategy.

Factor	Level 1	Level 2
k	5	10
Sample sizes	$n_i = 4 \; \forall \; i$	$n_i = 10 \; \forall \; i$
R_i	equal	unequal
Component Prior Info Strength	uniformly weak	uniformly strong
Component Prior Info Accuracy	uniformly inaccurate	uniformly accurate

Table 2. Design Strategy Implementation.

System	k	n_i	R_i	Component Prior Information	
				Strength	Accuracy
1	5	4	equal	weak	inaccurate
2	5	4	equal	weak	accurate
3	5	4	equal	strong	inaccurate
4	5	4	equal	strong	accurate
5	5	4	unequal	weak	inaccurate
6	5	4	unequal	weak	accurate
7	5	4	unequal	strong	inaccurate
8	5	4	unequal	strong	accurate
9	5	10	equal	weak	inaccurate
10	5	10	equal	weak	accurate
11	5	10	equal	strong	inaccurate
12	5	10	equal	strong	accurate
13	5	10	unequal	weak	inaccurate
14	5	10	unequal	weak	accurate
15	5	10	unequal	strong	inaccurate
16	5	10	unequal	strong	accurate
17	10	4	equal	weak	inaccurate
18	10	4	equal	weak	accurate
19	10	4	equal	strong	inaccurate
20	10	4	equal	strong	accurate
21	10	4	unequal	weak	inaccurate
22	10	4	unequal	weak	accurate
23	10	4	unequal	strong	inaccurate
24	10	4	unequal	strong	accurate
25	10	10	equal	weak	inaccurate
26	10	10	equal	weak	accurate
27	10	10	equal	strong	inaccurate
28	10	10	equal	strong	accurate
29	10	10	unequal	weak	inaccurate
30	10	10	unequal	weak	accurate
31	10	10	unequal	strong	inaccurate
32	10	10	unequal	strong	accurate

Table 3. Component Specifications.

Type	True R_i	Prior Info (s_i^o/n_i^o)
1	0.999	25/25
2	"	10/10
3	"	19/25
4	"	7/10
5	0.99	25/25
6	"	10/10
7	"	19/25
8	"	7/10
9	0.95	24/25
10	"	10/10
11	"	19/25
12	"	7/10
13	0.90	22/25
14	"	10/11
15	"	18/25
16	"	6/10

Table 4. System Structures.

No.	Number of Each Component Type																R
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1				5													0.9960
2						5											0.9510
3											5						0.7738
4													5				0.5905
5				2				2				1					0.9292
6						1				2				2			0.7237
7			2								2				1		0.8106
8	1				2								2				0.7931
9																5	0.5905
10										5							0.7738
11							5										0.9510
12	5																0.9960
13				2				3									0.9684
14						3				2							0.8757
15											2				3		0.6579
16	2				1				1				1				0.8448
17								10									0.9044
18		10															0.9900
19															10		0.5987
20									10								0.3487
21				3				3				2				2	0.7072
22						4								6			0.5105
23			6				4										0.9548
24	4												6				0.5293
25												10					0.5987
26														10			0.3487
27			10														0.9900
28					10												0.9044
29				3				3								4	0.6347
30			3							4				3			0.5920
31							4				3				3		0.6004
32	3				4				3								0.8211

Technical Considerations

The Monte Carlo software was written in FORTRAN 77; the accuracy of each module's output was verified using a set of hand-computed results. Most of the statistical calculations required by the simulation were performed using the International Mathematical and Statistical Libraries (IMSL) Stat/Library. IMSL routines called by the simulation software included the binomial deviate generator DRNBIN [14:967], the beta deviate generator DRNBET [14:993], and the inverse beta function DBETIN [14:915].

IMSL Stat/Library's uniform random number stream, which it uses to generate other random deviates, was evaluated for independence and uniformity. A stream of random numbers was generated on each computer used in the simulation; each stream was then evaluated using a runs-up test and a chi-square goodness-of-fit test. Each test was conducted using the guidelines in Section 7.4 of Law and Kelton [19]. The tests showed that the uniformity and independence assumptions were justified.

Methodology

The following series of steps was executed for each of the 32 hypothetical systems:

1. A set of system failure data was constructed by simulating a binomially-distributed failure count for each of the system's k components. Five thousand of these data sets were generated.
2. ML and Bayes estimates of R were calculated using each of the 5000 data sets; the MSE was then calculated for each estimate.
3. MMLI 90 and 95 percent lower bounds on R were computed using each of the 5000 data sets. Each data set with no observed failures for any component (that is, where $\hat{R} = 1.0$) was ignored. If more than 500 data sets were ignored, the coverage proportion and the average lower bound were not reported; other-

wise, these measures were calculated using only those lower bounds that were evaluated.

4. Lindstrom-Madden 90 and 95 percent lower bounds on R were computed using each of the 5000 data sets, and the coverage proportion and the average lower bound were calculated.

5. The performance measures for the approximate beta posterior 90 and 95 percent lower proportion limits on R were derived using the methods in Step 4.

6. For each of the 5000 data sets, a BMC empirical posterior CDF of R was constructed according to the method described in Chapter III. Each CDF was composed of 10000 data points. The appropriate order statistics were used as the 90 and 95 percent lower proportion bounds on R.

Using these six steps, a set of performance measures was compiled for each estimator under consideration. These performance measures are reported and analyzed in the next chapter.

V. Findings

Overview

This chapter presents the findings of the Monte Carlo comparison study. In the first section, the performance measures of the point estimators are reported and analyzed; the same is done in the second section for the interval estimators. After the analysis of the results, conclusions are drawn about estimator performance. The chapter ends with a list of recommendations for further research.

Results and Analysis: Point Estimators

Table 5 lists the MSE by system for the ML and Bayes estimators of R. The ratio of the losses (ML to Bayes) is also included; if the ratio is greater than one, the Bayes loss is smaller than that of the MLE (and therefore the Bayes estimator performed better). The MLE outperformed the Bayes estimator for 23 of the 32 systems. Appendix C contains a discussion of the Monte Carlo sampling statistics for the point estimators.

An analysis was conducted to determine if any of the structural aspects of the hypothetical systems had a statistically significant effect on the choice of the best estimator for a given system. Contingency tables were used because of the nominal nature of the explanatory variables (the structural factors in Table 1) and the response variable (the estimator judged best). The analysis, which followed the guidelines in Chapters 2 and 3 of Agresti [1], was conducted at significance level 0.05. Only main factor effects and pair-wise factor interactions were investigated; higher level interactions were assumed to be negligible. The results of the analysis are in Tables 6 and 7; the relevant contingency tables are in Appendix D.

Table 5. Point Estimator Performance Measures.

System	MSE($\hat{R}$)	MSE(Bayes)	Ratio of MSEs (ML to Bayes)
1	0.0012	0.5747	0.0021
2	0.0114	0.0562	0.2026
3	0.0414	0.2573	0.1608
4	0.0504	0.0162	3.1052
5	0.0158	0.4867	0.0324
6	0.0463	0.0197	2.3504
7	0.0374	0.3060	0.1224
8	0.0393	0.0167	2.3447
9	0.0203	0.1506	0.1347
10	0.0160	0.0079	2.0290
11	0.0046	0.3676	0.0125
12	0.0005	0.0155	0.0307
13	0.0030	0.3695	0.0081
14	0.0102	0.0190	0.5366
15	0.0197	0.1467	0.1345
16	0.0126	0.0116	1.0830
17	0.0211	0.7224	0.0292
18	0.0025	0.2181	0.0113
19	0.0375	0.0938	0.3995
20	0.0509	0.0149	3.4237
21	0.0471	0.4424	0.1065
22	0.0499	0.0427	1.1699
23	0.0105	0.7710	0.0136
24	0.0491	0.0326	1.5062
25	0.0194	0.2474	0.0785
26	0.0138	0.0117	1.1712
27	0.0009	0.7528	0.0012
28	0.0106	0.0191	0.5543
29	0.0201	0.3052	0.0658
30	0.0196	0.0282	0.6964
31	0.0184	0.3388	0.0544
32	0.0136	0.0262	0.5204

Table 6. Main Factor Effects: Point Estimators.

Factor	P-Value	Significant at $\alpha = 0.05$?	Goodman and Kruskal's τ
System Size	0.5000	No	0.005
Sample Size	0.2166	No	0.043
Reliability Levels	0.5000	No	0.005
Component Prior Info Strength	0.5000	No	0.005
Component Prior Info Accuracy	0.0004	Yes	0.391

Table 7. Pair-Wise Factor Interactions: Point Estimators.

Factor 1	Factor 2	P-Value	Significant at $\alpha = 0.05$?	Goodman and Kruskal's τ
System Size	Sample Size	0.6368	No	0.053
System Size	Reliability Levels	0.9268	No	0.014
System Size	Component Prior Info Strength	0.9268	No	0.014
System Size	Component Prior Info Accuracy	0.0050	Yes	0.401
Sample Size	Reliability Levels	0.4014	No	0.092
Sample Size	Component Prior Info Strength	0.4014	No	0.092
Sample Size	Component Prior Info Accuracy	0.0016	Yes	0.478
Reliability Levels	Component Prior Info Strength	0.9268	No	0.014
Reliability Levels	Component Prior Info Accuracy	0.0050	Yes	0.401
Component Prior Info Strength	Component Prior Info Accuracy	0.0050	Yes	0.401

Table 6 shows that prior information accuracy is the only factor with a significant main effect. However, all pair-wise interactions involving prior information accuracy were significant as well (see Table 7). Since the association between prior accuracy and estimator choice (as measured by Goodman and Kruskal's τ) is an order of magnitude larger than any of the other association levels, the significance of the pair-wise interactions is most likely an artifact of the dominance of the prior accuracy main effect. According to these tests, then, prior information accuracy is the only structural characteristic of the hypothetical systems that affects the choice of the best estimator.

According to Table 5, systems with inaccurate prior information (the odd-numbered systems) overwhelmingly favored the MLE, while those with accurate component priors were evenly split between the estimators. These results support the intuition that, if there is a large gap between the prior distribution mean and the true component reliability, the MLE should outperform the Bayes estimator. However, the results are *not* consistent with Martz and Waller's rule of thumb that the Bayes estimator outperforms the MLE if the prior information is accurate. The inconsistency may result from the fact that, in this study, a component's prior information was deemed "accurate" if the prior mean was within ± 20 percent of the truth. For some types of systems, this latitude in prior information accuracy may be too broad.

Results and Analysis: Interval Estimators

Tables 8 through 11 show the interval estimator performance measures by system. The estimator judged the best performer for a given system is designated with an asterisk (*). For most systems, either the beta approximation or the Lindstrom-Madden method provided the most desirable interval estimates, although for three systems (14, 26 and 32) there was no clearly superior estimator. Appendix C contains a discussion of the Monte Carlo sampling statistics for the interval estimators.

Table 8. Interval Estimator Performance Measures, Systems 1-8.

System	R	Method	90% Lower Bound		95% Lower Bound	
			Avg Limit	Coverage	Avg Limit	Coverage
1	0.9960	MMLI	**	**	**	**
		L-M*	0.5579	1.0000	0.4689	1.0000
		Beta Approx	0.1435	1.0000	0.1228	1.0000
		BMC	0.0952	1.0000	0.0784	1.0000
2	0.9510	MMLI	**	**	**	**
		L-M	0.5176	1.0000	0.4316	1.0000
		Beta Approx*	0.5769	1.0000	0.5333	1.0000
		BMC	0.5114	1.0000	0.4609	1.0000
3	0.7738	MMLI	**	**	**	**
		L-M*	0.3603	1.0000	0.2895	1.0000
		Beta Approx	0.1932	1.0000	0.1752	1.0000
		BMC	0.1367	1.0000	0.1208	1.0000
4	0.5905	MMLI	**	**	**	**
		L-M	0.2256	1.0000	0.1726	1.0000
		Beta Approx*	0.3686	1.0000	0.3424	1.0000
		BMC	0.2979	1.0000	0.2709	1.0000
5	0.9292	MMLI	**	**	**	**
		L-M*	0.4959	1.0000	0.4116	1.0000
		Beta Approx	0.1396	1.0000	0.1193	1.0000
		BMC	0.0927	1.0000	0.0763	1.0000
6	0.7237	MMLI	**	**	**	**
		L-M	0.3230	1.0000	0.2567	1.0000
		Beta Approx*	0.4507	1.0000	0.4108	1.0000
		BMC	0.4029	1.0000	0.3600	1.0000
7	0.8106	MMLI	**	**	**	**
		L-M*	0.3947	1.0000	0.3202	1.0000
		Beta Approx	0.1857	1.0000	0.1681	1.0000
		BMC	0.1329	1.0000	0.1175	1.0000
8	0.7931	MMLI	**	**	**	**
		L-M	0.3789	1.0000	0.3601	1.0000
		Beta Approx*	0.5634	1.0000	0.5326	1.0000
		BMC	0.5357	1.0000	0.5027	1.0000

* Best Estimator. ** Not reported.

Table 9. Interval Estimator Performance Measures, Systems 9-16.

System	R	Method	90% Lower Bound		95% Lower Bound	
			Avg Limit	Coverage	Avg Limit	Coverage
9	0.5905	MMLI	0.3265	1.0000	0.2832	1.0000
		L-M*	0.3528	0.9664	0.3044	0.9664
		Beta Approx	0.1306	1.0000	0.1142	1.0000
		BMC	0.0857	1.0000	0.0726	1.0000
10	0.7738	MMLI	0.4515	1.0000	0.4006	1.0000
		L-M	0.5312	0.9236	0.4762	1.0000
		Beta Approx*	0.5849	1.0000	0.5482	1.0000
		BMC	0.5217	1.0000	0.4795	1.0000
11	0.9510	MMLI	**	**	**	**
		L-M*	0.7328	1.0000	0.6780	1.0000
		Beta Approx	0.2669	1.0000	0.2469	1.0000
		BMC	0.2020	1.0000	0.1828	1.0000
12	0.9960	MMLI	**	**	**	**
		L-M	0.7879	1.0000	0.7346	1.0000
		Beta Approx*	0.7995	1.0000	0.7740	1.0000
		BMC	0.7611	1.0000	0.7292	1.0000
13	0.9684	MMLI	**	**	**	**
		L-M*	0.7563	1.0000	0.7021	1.0000
		Beta Approx	0.2581	1.0000	0.2324	1.0000
		BMC	0.1940	1.0000	0.1697	1.0000
14	0.8757	MMLI	**	**	**	**
		L-M	0.6403	1.0000	0.5844	1.0000
		Beta Approx	0.6293	1.0000	0.5924	1.0000
		BMC	0.5767	1.0000	0.5337	1.0000
15	0.6579	MMLI	0.3709	1.0000	0.3244	1.0000
		L-M*	0.4154	0.9156	0.3639	0.9832
		Beta Approx	0.2067	1.0000	0.1895	1.0000
		BMC	0.1547	1.0000	0.1390	1.0000
16	0.8448	MMLI	**	**	**	**
		L-M	0.6083	1.0000	0.5525	1.0000
		Beta Approx*	0.6512	1.0000	0.6231	1.0000
		BMC	0.6236	1.0000	0.5926	1.0000

* Best Estimator. ** Not reported.

Table 10. Interval Estimator Performance Measures, Systems 17-24.

System	R	Method	90% Lower Bound		95% Lower Bound	
			Avg Limit	Coverage	Avg Limit	Coverage
17	0.9044	MMLI	**	**	**	**
		L-M*	0.4721	1.0000	0.3899	1.0000
		Beta Approx	0.0252	1.0000	0.0200	1.0000
		BMC	0.0177	1.0000	0.0140	1.0000
18	0.9900	MMLI	**	**	**	**
		L-M*	0.5526	1.0000	0.4638	1.0000
		Beta Approx	0.3864	1.0000	0.3491	1.0000
		BMC	0.3479	1.0000	0.3092	1.0000
19	0.5987	MMLI	0.0654	1.0000	0.0448	1.0000
		L-M*	0.0856	0.9848	0.0593	0.9848
		Beta Approx	0.0248	1.0000	0.0212	1.0000
		BMC	0.0171	1.0000	0.0145	1.0000
20	0.3487	MMLI	**	**	**	**
		L-M	0.2315	1.0000	0.1778	1.0000
		Beta Approx*	0.3832	1.0000	0.3571	1.0000
		BMC	0.3456	1.0000	0.3184	1.0000
21	0.7072	MMLI	**	**	**	**
		L-M*	0.3129	1.0000	0.2478	1.0000
		Beta Approx	0.0184	1.0000	0.0144	1.0000
		BMC	0.0132	1.0000	0.0104	1.0000
22	0.5105	MMLI	0.1114	1.0000	0.0796	1.0000
		L-M	0.1728	0.9320	0.1288	1.0000
		Beta Approx*	0.2022	1.0000	0.1775	1.0000
		BMC	0.1835	1.0000	0.1597	1.0000
23	0.9548	MMLI	**	**	**	**
		L-M*	0.5203	1.0000	0.4340	1.0000
		Beta Approx	0.0481	1.0000	0.0420	1.0000
		BMC	0.0354	1.0000	0.0305	1.0000
24	0.5293	MMLI	**	**	**	**
		L-M	0.1833	0.9262	0.1373	1.0000
		Beta Approx*	0.2659	1.0000	0.2444	1.0000
		BMC	0.2546	1.0000	0.2333	1.0000

* Best Estimator. ** Not reported.

Table 11. Interval Estimator Performance Measures, Systems 25-32.

System	R	Method	90% Lower Bound		95% Lower Bound	
			Avg Limit	Coverage	Avg Limit	Coverage
25	0.5987	MMLI	0.3374	1.0000	0.2937	1.0000
		L-M*	0.3616	0.9624	0.3126	0.9624
		Beta Approx	0.0597	1.0000	0.0510	1.0000
		BMC	0.0449	1.0000	0.0379	1.0000
26	0.3487	MMLI	0.1704	0.9772	0.1423	0.9936
		L-M	0.1546	0.9770	0.1227	0.9842
		Beta Approx	0.1688	1.0000	0.1503	1.0000
		BMC	0.1403	1.0000	0.1232	1.0000
27	0.9900	MMLI	**	**	**	**
		L-M*	0.7820	1.0000	0.7284	1.0000
		Beta Approx	0.0845	1.0000	0.0759	1.0000
		BMC	0.0655	1.0000	0.0581	1.0000
28	0.9044	MMLI	**	**	**	**
		L-M	0.6355	1.0000	0.5797	1.0000
		Beta Approx*	0.6433	1.0000	0.6156	1.0000
		BMC	0.6138	1.0000	0.5831	1.0000
29	0.6347	MMLI	0.3555	1.0000	0.3101	1.0000
		L-M*	0.3939	0.9378	0.3433	0.9884
		Beta Approx	0.0468	1.0000	0.0396	1.0000
		BMC	0.0366	1.0000	0.0308	1.0000
30	0.5920	MMLI	0.3298	1.0000	0.2864	1.0000
		L-M*	0.3544	0.9660	0.3058	0.9660
		Beta Approx	0.3226	1.0000	0.2643	1.0000
		BMC	0.3012	1.0000	0.2725	1.0000
31	0.6004	MMLI	0.3856	1.0000	0.3382	1.0000
		L-M*	0.4343	0.9846	0.3819	0.9846
		Beta Approx	0.0656	1.0000	0.0585	1.0000
		BMC	0.0509	1.0000	0.0449	1.0000
32	0.8211	MMLI	**	**	**	**
		L-M	0.5820	1.0000	0.5263	1.0000
		Beta Approx	0.5696	1.0000	0.5421	1.0000
		BMC	0.5466	1.0000	0.5172	1.0000

* Best Estimator. ** Not reported.

For the interval estimators, a contingency table analysis was conducted using methodology similar to that of the analysis of point estimator performance (see the previous section); Appendix D contains the related contingency tables. As can be seen from Tables 12 and 13, this set of tests yielded the same conclusion as the point estimator tests: only the accuracy of the prior information significantly affected the choice of the best estimator for the hypothetical systems.

One obvious feature of the results is the conservatism of the interval estimates. To capture the degree of conservatism for each method, each system's 90 percent Lindstrom-Madden and approximate beta interval estimates were arbitrarily reduced by 10, 20 and 50 percent. Table 14 shows the effect of the reductions on the coverage proportions. The reduction data show that, for many systems, both estimators produced interval estimates that were so conservative they all would have covered R even if they were cut in half.

System 10 was chosen at random to illustrate the effect of the interval reduction on best estimator selection. For each interval estimation method, the length of each of System 10's 90 percent interval estimates was reduced until the coverage proportions were equal and as close to 0.9 as possible. A comparison of the average lower bounds of the reduced intervals (see Table 15) shows that the beta approximation still yields the shortest intervals; this outcome is unchanged from the original comparison.

Table 12. Main Factor Effects: Interval Estimators.

Factor	P-Value	Significant at $\alpha = 0.05$?	Goodman and Kruskal's τ
System Size	0.3612	No	0.016
Sample Size	0.6388	No	0.000
Reliability Levels	0.6388	No	0.000
Component Prior Info Strength	0.3612	No	0.016
Component Prior Info Accuracy	0.0000	Yes	0.778

Table 13. Pair-Wise Factor Interactions: Interval Estimators.

Factor 1	Factor 2	P-Value	Significant at $\alpha = 0.05$?	Goodman and Kruskal's τ
System Size	Sample Size	0.9171	No	0.016
System Size	Reliability Levels	0.9171	No	0.016
System Size	Component Prior Info Strength	0.6768	No	0.048
System Size	Component Prior Info Accuracy	0.0000	Yes	0.810
Sample Size	Reliability Levels	0.9171	No	0.016
Sample Size	Component Prior Info Strength	0.9171	No	0.016
Sample Size	Component Prior Info Accuracy	0.0000	Yes	0.778
Reliability Levels	Component Prior Info Strength	0.9171	No	0.016
Reliability Levels	Component Prior Info Accuracy	0.0000	Yes	0.778
Component Prior Info Strength	Component Prior Info Accuracy	0.0000	Yes	0.810

Table 14. Coverage Proportions for Reduced Intervals.

System	Method	Reduction Percentage			
		0	10	30	50
1	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	1.0000
2	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	1.0000
3	L-M	1.0000	1.0000	1.0000	0.6410
	Beta Approx	1.0000	1.0000	1.0000	1.0000
4	L-M	1.0000	0.8806	0.8806	0.3624
	Beta Approx	1.0000	1.0000	1.0000	0.0000
5	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	1.0000
6	L-M	1.0000	1.0000	1.0000	0.7218
	Beta Approx	1.0000	1.0000	1.0000	0.3528
7	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	1.0000
8	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	0.5986
9	L-M	0.9664	0.8848	0.7474	0.0776
	Beta Approx	1.0000	1.0000	1.0000	0.9948
10	L-M	0.9236	0.9236	0.9236	0.4562
	Beta Approx	1.0000	1.0000	0.9236	0.2404
11	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	1.0000
12	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	1.0000
13	L-M	1.0000	1.0000	1.0000	1.0000
	Beta Approx	1.0000	1.0000	1.0000	1.0000
14	L-M	1.0000	1.0000	1.0000	0.7474
	Beta Approx	1.0000	1.0000	1.0000	1.0000
15	L-M	0.9156	0.9156	0.7724	0.2126
	Beta Approx	1.0000	1.0000	1.0000	1.0000
16	L-M	1.0000	1.0000	0.8116	0.8116
	Beta Approx	1.0000	1.0000	1.0000	0.8116

Table 15. Interval Reduction Data for System 10.

Method	Original Limit	Reduction Multiplier	New Coverage	New Avg Limit
Beta Approx	0.5849	0.3706	0.9236	0.7387
L-M	0.5312	0.3284	0.9236	0.6852
BMC	0.5217	0.4395	0.9236	0.7320

Conclusions

Based on the assumptions made and the results of the Monte Carlo comparison study, four conclusions can be drawn:

1. If the accuracy of the component prior information is in doubt, then the MLE and the Lindstrom-Madden estimator are clearly the best estimators for R. However, if confidence in the prior information is high, the Bayes estimators constructed from the approximate beta posterior should be used instead.
2. Because of its extreme conservatism, the interval estimator constructed from the Bayes Monte Carlo empirical posterior distribution should not be used for component sample sizes under ten.
3. Easterling's MMLI method should not be used when the sample sizes are uniformly under ten, especially when the component reliabilities are known to be high.
4. Any small-sample interval estimate generated with any of the methods considered will be conservative.

In summary, the results of this investigation show that both the beta approximation and the Lindstrom-Madden method to be useful with small data sets. However, neither of these estimators is strongly preferable to the other. ACC could use either approach, depending on the quality of the prior information available.

Recommendations for Further Research

Since this study is only a first step in understanding small-sample reliability estimation, ample opportunity exists to broaden the scope of the results. Several areas with potential for additional research are suggested:

1. Additional hypothetical systems should be considered to incorporate the following extensions to the design:

   (a) More than two levels of k and n_i should be examined. For example, a possible extension could be $k = \{3,\ 5,\ 10,\ 15\}$ and $n_i = \{4,\ 10,\ 20\}$.

   (b) The accuracy and strength of the component prior information and the component sample size should be allowed to vary across components.

   (c) Parallel and complex systems should be considered as well as series systems.

2. As mentioned previously, the definition of "accuracy" applied to the component prior information in this study may have been both simplistic and overly liberal. The effect on the study results of variations in prior density accuracy should be analyzed more thoroughly. Possible avenues for incorporating error into the prior include adjustments to its mean or variance (or both). Such a study would yield valuable information about the impact of prior information on small-sample Bayes estimation.

3. It was shown earlier in this chapter that large decreases in small-sample interval estimates can sometimes be made at a relatively low cost in significance. An analytical understanding of this phenomenon could result in the development of interval reduction methods, which would in turn lead to small-sample interval estimates of higher quality.

Appendix A. *Bayesian Statistical Inference*

Introduction

This appendix provides a brief introduction to Bayesian statistical inference. With the aid of this presentation, the reader with little or no exposure to the Bayesian paradigm can understand the Bayesian reliability methods discussed in this thesis.

The Nature of Probability: Two Views

Before undertaking a meaningful discussion of Bayesian statistical methods, it is important to understand how and why those methods differ from the classical statistical techniques taught in most universities. The two approaches are dissimilar because they are rooted in different understandings of the fundamental nature of probability.

The *classical* school of probability (also called the *frequentist* or *objectivist school*) holds that probability is a measure of the relative frequency of an event during repeated experimentation. Consider an experiment for which Ω is the set of all possible outcomes. Define as one of those outcomes the event E such that $E \in \Omega$. A frequentist statistician would view the probability of E (that is, the likelihood of event E occurring) as $P[E] = \lim_{n\to\infty}(e/n)$, where e is the number of times E occurs during n experimental trials. Thus, $P[E]$ is the long-run frequency with which E occurs. According to the classical viewpoint, the value of this frequency is an unknown constant that can be reasonably approximated if enough experiments are conducted.

In contrast, the Bayesian school (which takes its name from the British clergyman and amateur mathematician Thomas Bayes) adopts a *subjectivistic* (or *personalistic*) view of probability. The Bayesian views $P[E]$ not as an objective estimate of the unknown constant likelihood of E, but as a rational expression of his or her

degree of belief that E will occur. Since probability is subjective, different people can assign different values to $P[E]$ based on varying perceptions of prior experience. Thus, from the Bayesian perspective, this prior experience replaces experimentation as the source of the probability estimate. However, the Bayesian does not abandon experimentation, because it is an important means of refining his or her assessment of $P[E]$.

To illustrate these divergent views of the nature of probability, consider a simple experiment. A certain coin is flipped, with the outcome being either "heads" or "tails"; that is, $\Omega = \{\text{heads, tails}\}$. Suppose the likelihood of each outcome is unknown. Further, define the event $E = \{\text{heads}\}$. In order to make a probability statement about E, the frequentist would perform n experiments (where n is large, say 1000), observe m occurrences of E, and declare that, for this particular coin, $P[E] \approx m/n$. On the other hand, the Bayesian might conclude, based on experience with similar coins, that E is just as likely to occur as not. Therefore, he or she would declare that $P[E] = 1/2$, then conduct a series of experimental trials to validate and refine that estimate.

There is a great deal of controversy over the legitimacy of the Bayesian view of probability. Many theoretical statisticians reject the Bayesian approach because they believe that statistics, as a science, should be as objective as possible. However, Press and others have pointed out that the methods of classical statistical inference incorporate a great deal of subjectivity, although they purport to be objective. The subjective nature of the Bayesian statistics is not a weakness but a strength, because valuable prior experience can be accounted for in the inference process [36:1-14].

Bayes' Theorem

Bayes' Theorem, which is presented below, is both stated and proven using the definition of conditional probability. This definition, which can be found in any

introductory probability text, provides a convenient mechanism for incorporating subjective probability information into a statistical analysis.

Theorem A.1 (Bayes) *Let Ω be a sample space. Define the mutually exclusive events F_i such that (1) $F_i \in \Omega \ \forall i$ and (2) $\bigcup_{i=1}^{\infty} F_i = \Omega$. Further, let E be an event such that (1) the conditional probabilities $P[E|F_i]$ are defined for all i, and (2) $P[E] > 0$. Then*

$$P[F_i|E] = \frac{P[E|F_i]P[F_i]}{\sum_{i=1}^{\infty} P[E|F_i]P[F_i]}$$

The proof of this theorem, which is a straightforward application of the definition of conditional probability, is given by Biswas [5:28].

Bayes' Theorem had its origin in Bayes' 1763 paper *An Essay Towards Solving a Problem in the Doctrine of Chances*, where it was originally limited to the binomial sampling problem [3]. Laplace, working independently, published the general form of this theorem in 1774 [18]; he developed it more fully in several subsequent papers. The modern fusion of Bayes' Theorem with subjective probability is due mostly to the 1939 work of Jeffreys; De Groot, Savage, de Finetti, Ramsey and Rényi also made significant contributions to the development of Bayesian theory [36:8,16].

Foundations of Bayesian Inference

Consider a failure density with some parameter θ. Based on earlier discussion, the frequentist view of θ is that its true value is a fixed, unknown constant that can be discovered through infinite sampling. From the Bayesian perspective, however, the true value of θ could be any of a range of values, some of which are more likely to be the true value of θ than others. It is not surprising, then, that the perceived value of θ is viewed as a random variable in Bayesian inference. (In the discussion that follows, the symbol Θ will be used to refer to the random variable; θ will refer to a particular value of Θ.)

Since Θ is viewed as a random variable, subjective information about the likelihood of its possible values may be encoded in a probability density or mass function $g(\theta)$. Once $g(\theta)$ is defined, experiments can be conducted to validate and refine the assessment of Θ. Bayes' Theorem is the mechanism for combining the subjective information with the experimental results.

The subjective assessment of Θ is refined in a straightforward manner. First, a random sample of size n of independent, identically-distributed random variables X_i is taken ($\{X_i = x_i, i = 1 \ldots n\}$). The probability mass/density function of X_i is $f(x_i|\theta)$ (that is, the form of the density/mass function of X_i depends on the value of Θ). After the random sample is taken, the information from the sample is combined with the *à priori* information assumed about Θ (i. e., $g(\theta)$) to get an *à posteriori* refined probability statement about Θ. If Θ is assumed to be a discrete random variable, this procedure is captured in the following extension to Bayes' Theorem.

Theorem A.2 *Let Θ be a discrete random variable with probability mass function $g(\theta)$. Further, let $\vec{X} \equiv \{X_1, \ldots, X_n\}$ be a set of independent, identically-distributed random variables, each with probability mass/density function $f(x|\theta)$, and let $\vec{x} \equiv \{x_1, \ldots, x_n\}$ be a set of random samples from those random variables. Then*

$$h(\theta|\vec{x}) = \frac{g(\theta) \prod_{i=1}^{n} f(x_i|\theta)}{\sum_{\theta} \{[\prod_{i=1}^{n} f(x_i|\theta)]\, g(\theta)\}}$$

The proof of this theorem is omitted because it is trivial [36:24]. A similar extension to Bayes' Theorem can be derived for continuous Θ. (Note that X_i can be either discrete or continuous, despite of the form of $g(\theta)$.)

Since $g(\theta)$ incorporates *à priori* information about Θ, it is called the *prior density* of Θ; its parameters are called *hyperparameters* to differentiate them from the parameters of, say, $f(x|\theta)$. Similarly, $h(\theta|\vec{x})$ is called the *posterior density* of Θ. These respective terms are usually shortened to "prior" and "posterior" for simplicity.

The Prior Distribution

It should be obvious that a Bayesian estimate is only as good as the accuracy of the prior assessment of Θ. This means, then, that the selection of the prior density $g(\theta)$ must be made with great care.

The prior density of Θ can take any form that seems reasonable. Usually, a *proper* prior (a probability density or mass function) is used. However, the prior need not integrate to unity (in which case it is called an *improper prior*). The prior can even be an empirical function. In most practical applications, however, priors are selected from two classes: *natural conjugate priors* (sometimes called *convenience priors*), and *noninformative priors* (also called *indifference* or *vague priors*).

When a natural conjugate prior is used, the resulting posterior is of the same family. The form of this prior depends on the form of $f(x|\theta)$. For example, a natural conjugate prior of a binomial probability is the beta density. This prior returns a posterior that is also a beta density. Press discusses the advantages of natural conjugate priors at length, and catalogs the common conjugate prior families [36].

A noninformative prior is used when there is little or no reason to prefer any one value of Θ over another. There are various types of noninformative priors discussed in the literature; the most common form used is the uniform distribution. The use of noninformative priors has been the subject of much debate; many statisticians see no justification for the Bayesian approach if no meaningful prior information exists. See Press [36:46-51] for a detailed discussion of the noninformative prior.

Once the family of the prior is chosen, its hyperparameters can be assigned values based on any reasonable criterion. In some cases it is perfectly acceptable to choose hyperparameter values based on "gut feel." For example, a group of design engineers might feel strongly (perhaps because of prior design or manufacturing experience) that a certain switch has a failure probability of 0.01. In other situations, it may be advisable to estimate the hyperparameters from a data set derived from

a different (but conceptually related) experimental situation. In any case, the hyperparameters should be estimated with care. If the choice is made blindly, it can render the entire analysis meaningless. For example, Kapur and Lamberson show that the parameters of a beta prior can be selected so that the experimental results will not substantially affect the form of the posterior, even if those results completely contradict the the prior information [15:380-382].

Determining the form of $g(\theta)$ is a difficult process, and the results are often controversial. The key point to remember when selecting the prior is that, although the choice is ultimately based on subjective factors, that choice should be both reasonable and justifiable.

Bayesian Point Estimation

A point estimate is essentially a decision about the value of the parameter being estimated. It is possible, if not rigorous, to avoid decision theory when introducing the basics of classical point estimation. However, Bayesian point estimation must be introduced in a decision-theoretic framework.

Before Bayesian point estimation is discussed, several concepts from decision theory must be defined. These definitions, which deal with a single parameter only, are adapted from the general forms given by Mood et al. [33:297-299, 344].

Definition A.1 *Let* $\hat{\Theta} = \hat{\Theta}(X_1, \ldots, X_n)$ *be an estimator of a parameter* θ*, and let* $\hat{\theta}$ *be a particular value of* $\hat{\Theta}$*. The* loss function *measures the error sustained by estimating* θ *by* $\hat{\theta}$*. This function is denoted by* $\lambda(\hat{\theta};\theta)$*, and has the following properties: (1)* $\lambda(\hat{\theta};\theta)$ *is a real function, (2)* $\lambda(\hat{\theta};\theta) \geq 0 \ \forall\theta$*, and (3)* $\lambda(\hat{\theta};\theta) = 0$ *when* $\hat{\theta} = \theta$*.*

There are several possible loss functions. The one most often seen in practical applications is the *squared-error loss function:*

$$\lambda(\hat{\theta};\theta) = (\hat{\theta} - \theta)^2 \tag{46}$$

Definition A.2 *For some loss function* $\lambda(\hat{\theta};\theta)$*, the* risk function *of the estimator* $\hat{\Theta}$ *is*

$$\rho_{\hat{\Theta}}(\theta) = E[\lambda(\hat{\Theta};\theta)]$$

The risk associated with a squared-error loss function is $\mathrm{E}[(\hat{\theta} - \theta)^2]$, which is easily recognized as the mean square error of $\hat{\Theta}$.

Definition A.3 *The* Bayes risk *of the estimator* $\hat{\Theta}$ *with respect to the loss function* $\lambda(\hat{\theta};\theta)$ *and the prior density* $g(\theta)$ *is*

$$\mathcal{R}_{\lambda,g}(\hat{\Theta}) = \int_{\Theta} \rho_{\hat{\Theta}}(\theta) g(\theta) d\theta$$

With the above terms defined, Bayes estimators can now be introduced.

Definition A.4 *The* Bayes estimator $\hat{\Theta}^*$ *of* Θ *(with respect to the loss function* $\lambda(\hat{\theta};\theta)$ *and the prior density* $g(\theta)$*) is the estimator with the smallest Bayes risk; that is,* $\hat{\Theta}^*$ *is the estimator for which*

$$\mathcal{R}_{\lambda,g}(\hat{\Theta}^*) = \inf_{\hat{\Theta}} \{\mathcal{R}_{\lambda,g}(\hat{\Theta})\}$$

It is straightforward to find the Bayes estimator for a given loss function and prior density. For a squared-error loss function, Mood et al. show that the Bayes estimator is the mean of the posterior density (i.e., $\hat{\Theta}^* = \mathrm{E}[\Theta|\vec{x}]$) [33:345].

One other property of Bayes estimators is useful in certain situations. Biswas shows that Bayes estimates and maximum likelihood estimates converge as sample

size approaches infinity. Further, this property holds regardless of the form of the prior [5:338].

Bayesian Interval Estimation

The interpretation of a Bayesian interval estimate is different from that of its classical counterpart, the confidence interval. Recall that a frequentist assumes the parameter θ estimated is an unknown constant. In a classical interval estimation method, the estimators of the end points of the interval are random variables. Therefore, a classical confidence coefficient is not an explicit probability statement about θ. Instead, it is the probability that the interval estimator will generate an interval that will contain the true value of θ (and is therefore a declaration of confidence in the estimate) [32:353].

In the Bayesian view, however, Θ is a random variable, so a Bayesian probability (or credibility) interval is a direct probability statement about the value of Θ. Thus, the two-sided $100(1-\gamma)$ percent probability interval $[\theta_L, \theta_U]$ has an intuitive interpretation: Θ will take on a value within the interval with probability $1-\gamma$ (that is, $P\{\Theta \in [\theta_L, \theta_U]\} = 1-\gamma$). The bounds of the interval $[\theta_L, \theta_U]$ can be evaluated directly by solving

$$\int_{-\infty}^{\theta_L} h(\theta|\vec{x})d\theta = \frac{\gamma}{2} \tag{47}$$

and

$$\int_{\theta_U}^{\infty} h(\theta|\vec{x})d\theta = \frac{\gamma}{2} \tag{48}$$

for θ_L and θ_U. This leads to a symmetric probability interval for Θ. There are other types of two-sided Bayesian probability intervals, but the symmetric interval is most often used because it is simple to calculate [30:208].

A one-sided Bayesian probability interval has an interpretation similar to that of the two-sided interval. The lower probability bound, which is considered exten-

sively in this thesis, is found by solving the following equation for θ_L:

$$\int_{-\infty}^{\theta_L} h(\theta|\vec{x})d\theta = \gamma \tag{49}$$

This is an obvious adaptation of Equation (47). An expression for the upper bound can be formulated by similarly manipulating Equation (48).

Empirical Bayesian Inference

A school of thought that tries to mitigate the subjectivity of Bayesian inference is the *empirical Bayes* approach. Empirical Bayes procedures compensate for the uncertainty over the form of the prior density by estimating its hyperparameters from the current data set. Because they violate Bayes Theorem by making the hyperparameters dependent on the test data, empirical Bayes procedures are controversial [36:43]. Furthermore, empirical Bayes estimates require a data set of adequate size; when the data set is small, it makes no sense to use these methods (philosophical issues notwithstanding). Since empirical procedures are not used in this thesis, they are not developed here. The interested reader is referred to [28] and Chapter 13 of [30].

Further Reading

Press, who has been cited frequently in this appendix, provides a general introduction to Bayesian statistics that is well-written and easy to understand [36]. Also, Martz and Waller's reliability text [30] and Berger's book on decision theory [4] contain brief, but comprehensible, introductions to Bayesian inference as a foundation for their primary subject matter (with Berger's treatment being considerably more rigorous). For an advanced treatment of Bayesian probability theory, the reader should refer to De Groot [10].

Most standard reliability textbooks contain at least a chapter introducing Bayesian inference and describing its applications to reliability. Some of the texts containing particularly enlightening expositions of Bayesian theory include Crowder et al. [9], Kapur and Lamberson [15], and Mann et al. [27].

For a reliability text written exclusively from the Bayesian viewpoint, the reader is referred to Martz and Waller [30]. This book is the standard reference work on Bayesian reliability methods.

Appendix B. *The Mellin Transform*

The Mellin transform of $f(x)$ with respect to the complex parameter u is

$$M\{f(x);u\} = \int_0^\infty x^{u-1} f(x)dx \tag{50}$$

If $f(x)$ is the density function of the continuous random variable X and u is an integer greater than 2, it immediately follows that

$$M\{f(x);u\} = E[X^{u-1}] \tag{51}$$

which is the $(u-1)$th central moment of X.

The Mellin transform is useful for finding density functions for products and quotients of random variables. Consider the independent, nonnegative random variables $X_1, \ldots, X_k$ with density functions $f_1(x_1), \ldots, f_k(x_k)$. If the random variable $Y = \prod_{i=1}^k X_i$ has a density function $h(y)$, it can be shown that

$$M\{h(y);u\} = \prod_{i=1}^k M\{f_i(x_i);u\} \tag{52}$$

The function $f(x)$ associated with the Mellin transform $M\{f(x);u\}$ can be derived using the Mellin inversion integral:

$$f(x) = \frac{1}{2\pi i}\left[\lim_{\kappa\to\infty}\int_{\rho-\kappa i}^{\rho+\kappa i} x^{-u} M\{f(x);u\}du\right] \tag{53}$$

Since tables of Mellin transform pairs are available for common functional forms, it is often possible to avoid evaluating the inversion integral.

Springer, from whose book this discussion is digested, provides a comprehensive treatment of the Mellin transform [40:91-97]. Oberhettinger provides an extensive catalog of Mellin transform pairs [35].

Appendix C. *Assessment of the Monte Carlo Process*

Overview

When a Monte Carlo study is conducted, it is important to understand the variability inherent in the simulation process. If a simulation's variance is too high, its results are suspect; this usually happens when the sample size of the simulation is too small. Two measures of merit are often used to judge the quality of a Monte Carlo simulation: the Monte Carlo standard error (MCSE) and the Monte Carlo confidence interval (MCCI). This appendix presents the theory behind these statistics, then uses them to assess the simulation conducted during this investigation.

Theory

Since the variability of the simulation process is under consideration, the MCSE is the standard error constructed from the expected value of the parameter being simulated. Consider a simulation parameter $\hat{\theta}$ with mean $\mu_{\hat{\theta}}$ and variance $\sigma^2_{\hat{\theta}}$. Further, define N as the Monte Carlo sample size. From Theorem 7.1 in Mendenhall et al. [32:304], it follows that the Monte Carlo variance is

$$\sigma^2_{MC} = \text{Var}[\mu_{\hat{\theta}}] = \frac{\sigma^2_{\hat{\theta}}}{N} \tag{54}$$

so the MCSE is obviously

$$\sigma_{MC} = \frac{\sigma_{\hat{\theta}}}{\sqrt{N}} \tag{55}$$

The unbiased estimator for σ_{MC} is

$$S_{MC} = \frac{S_{\hat{\theta}}}{\sqrt{N}} = \frac{1}{\sqrt{N}(N-1)} \sum_{i=1}^{N} (\hat{\theta}_i - \hat{\mu}_{\hat{\theta}})^2 \tag{56}$$

where $\{\hat{\theta}_i,\ i = 1, \ldots, N\}$ is the set of simulated estimates and

$$\hat{\mu}_{\hat{\theta}} = \frac{(\sum_{i=1}^{N} \hat{\theta}_i)}{N} \tag{57}$$

is the estimator of $\mu_{\hat{\theta}}$.

If the MCSE is known, the general form of the MCCI can easily be derived using the pivot method. Since $\mu_{\hat{\theta}}$ is normally distributed with standard error σ_{MC}, it immediately follows that the $100(1-\gamma)$ percent MCCI for $\mu_{\hat{\theta}}$ is

$$[\hat{\mu}_{\hat{\theta}} \pm z_{\gamma/2} \cdot \sigma_{MC}] \tag{58}$$

Results and Analysis

Tables 16 lists the estimated MCSE and a 95 percent MCCI by system for each type of point estimate. Table 17 contains the same information for lower confidence and probability limits generated by the Lindstrom-Madden and approximate beta interval estimators (due to technical difficulties, no statistics are presented for the other two interval estimators). The sampling statistics show that, for $N = 5000$, the variability in the simulation process is reasonably low.

The Monte Carlo statistics for the Bayesian estimators are much lower than those derived for the classical estimators. However, this is not surprising since the prior information dominates the small sets of simulated test data. A comparison of the Bayesian results in Table 16 with the MSE values for the Bayes estimator (see Table 5) suggests that there is a significant bias component in the Bayes estimator's MSE; the size of the bias seems to be a function of the accuracy of the prior information.

Table 16. Monte Carlo Statistics: Point Estimates.

System	MLE		Bayes Estimator	
	S_{MC}	95% MCCI	S_{MC}	95% MCCI
1	0.0005	[0.9945, 0.9964]	0.0000	[0.2369, 0.2370]
2	0.0015	[0.9498, 0.9557]	0.0003	[0.7143, 0.7155]
3	0.0029	[0.7663, 0.7776]	0.0002	[0.2663, 0.2669]
4	0.0032	[0.5850, 0.5975]	0.0003	[0.4646, 0.4659]
5	0.0018	[0.9261, 0.9331]	0.0001	[0.2314, 0.2320]
6	0.0030	[0.7185, 0.7305]	0.0006	[0.5890, 0.5914]
7	0.0027	[0.8078, 0.8185]	0.0001	[0.2573, 0.2579]
8	0.0028	[0.7890, 0.8000]	0.0003	[0.6650, 0.6662]
9	0.0020	[0.5860, 0.5939]	0.0004	[0.2026, 0.2041]
10	0.0018	[0.7705, 0.7775]	0.0008	[0.7011, 0.7040]
11	0.0010	[0.9498, 0.9536]	0.0001	[0.3445, 0.3449]
12	0.0003	[0.9945, 0.9957]	0.0001	[0.8706, 0.8709]
13	0.0008	[0.9689, 0.9719]	0.0002	[0.3603, 0.3609]
14	0.0014	[0.8714, 0.8770]	0.0006	[0.7428, 0.7450]
15	0.0020	[0.6536, 0.6614]	0.0003	[0.2749, 0.2759]
16	0.0016	[0.8424, 0.8486]	0.0004	[0.7397, 0.7413]
17	0.0021	[0.8994, 0.9074]	0.0000	[0.0544, 0.0545]
18	0.0007	[0.9885, 0.9912]	0.0001	[0.5228, 0.5232]
19	0.0027	[0.3436, 0.3544]	0.0000	[0.0423, 0.0425]
20	0.0032	[0.5934, 0.6059]	0.0003	[0.4784, 0.4797]
21	0.0031	[0.7055, 0.7176]	0.0001	[0.0420, 0.0422]
22	0.0032	[0.5026, 0.5150]	0.0005	[0.3056, 0.3073]
23	0.0014	[0.9529, 0.9586]	0.0000	[0.0767, 0.0768]
24	0.0031	[0.5208, 0.5330]	0.0003	[0.3492, 0.3503]
25	0.0020	[0.5962, 0.6039]	0.0002	[0.1012, 0.1019]
26	0.0017	[0.3437, 0.3502]	0.0005	[0.2457, 0.2477]
27	0.0004	[0.9897, 0.9914]	0.0000	[0.1224, 0.1224]
28	0.0015	[0.8670, 0.8727]	0.0003	[0.7311, 0.7324]
29	0.0020	[0.6308, 0.6387]	0.0001	[0.0821, 0.0827]
30	0.0020	[0.5881, 0.5959]	0.0007	[0.4293, 0.4318]
31	0.0019	[0.6740, 0.6815]	0.0001	[0.0973, 0.0976]
32	0.0017	[0.8184, 0.8249]	0.0004	[0.6607, 0.6621]

Table 17. Monte Carlo Statistics: Interval Lower Bounds (Selected Estimators).

System	Lindstrom-Madden Limit		Approximate Beta Limit	
	S_{MC}	95% MCCI	S_{MC}	95% MCCI
1	0.0005	[0.5570, 0.5589]	0.0000	[0.1434, 0.1435]
2	0.0014	[0.5148, 0.5203]	0.0003	[0.5763, 0.5776]
3	0.0024	[0.3556, 0.3649]	0.0001	[0.1930, 0.1935]
4	0.0023	[0.2211, 0.2300]	0.0003	[0.3680, 0.3692]
5	0.0016	[0.4926, 0.4991]	0.0001	[0.1394, 0.1398]
6	0.0024	[0.3183, 0.3277]	0.0006	[0.4495, 0.4519]
7	0.0023	[0.3901, 0.3992]	0.0001	[0.1855, 0.1859]
8	0.0023	[0.3744, 0.3835]	0.0003	[0.5628, 0.5641]
9	0.0018	[0.3492, 0.3565]	0.0003	[0.1301, 0.1312]
10	0.0019	[0.5275, 0.5349]	0.0008	[0.5833, 0.5865]
11	0.0012	[0.7305, 0.7351]	0.0001	[0.2667, 0.2671]
12	0.0004	[0.7872, 0.7887]	0.0001	[0.7993, 0.7996]
13	0.0010	[0.7544, 0.7582]	0.0001	[0.2578, 0.2583]
14	0.0016	[0.6371, 0.6436]	0.0006	[0.6281, 0.6305]
15	0.0019	[0.4116, 0.4191]	0.0002	[0.2063, 0.2072]
16	0.0018	[0.6049, 0.6118]	0.0004	[0.6504, 0.6521]
17	0.0019	[0.4684, 0.4758]	0.0000	[0.0251, 0.0252]
18	0.0007	[0.5512, 0.5539]	0.0001	[0.3862, 0.3866]
19	0.0015	[0.0827, 0.0884]	0.0000	[0.0247, 0.0248]
20	0.0023	[0.2270, 0.2360]	0.0003	[0.3826, 0.3838]
21	0.0024	[0.3082, 0.3177]	0.0000	[0.0184, 0.0185]
22	0.0021	[0.1687, 0.1769]	0.0004	[0.2015, 0.2029]
23	0.0014	[0.5176, 0.5229]	0.0000	[0.0481, 0.0482]
24	0.0021	[0.1792, 0.1874]	0.0002	[0.2654, 0.2664]
25	0.0018	[0.3581, 0.3652]	0.0001	[0.0595, 0.0600]
26	0.0012	[0.1522, 0.1570]	0.0004	[0.1680, 0.1696]
27	0.0006	[0.7809, 0.7831]	0.0000	[0.0846, 0.0846]
28	0.0017	[0.6323, 0.6388]	0.0004	[0.6426, 0.6439]
29	0.0019	[0.3902, 0.3976]	0.0001	[0.0466, 0.0470]
30	0.0018	[0.3509, 0.3580]	0.0006	[0.3215, 0.3237]
31	0.0019	[0.4306, 0.4380]	0.0001	[0.0655, 0.0658]
32	0.0018	[0.5785, 0.5856]	0.0004	[0.5689, 0.5704]

Appendix D. *Contingency Tables*

Point Estimator Tables

This section contains the contingency tables that were used to test for the significance of factor effects on the choice of the best point estimator. The single-factor tables (numbers 18 to 22) are listed first, followed by the pair-wise tables (numbers 23 to 32). For two-by-two tables, Fisher's Exact Test was used to derive the p-values; for larger tables, the p-value was found using Pearson's chi-squared statistic. The values for Goodman and Kruskal's τ, a measure of nominal association, are also included.

Table 18. Main Effect Test: System Size.

	MLE	Bayes	Row Totals
$k = 5$	11	5	16
$k = 10$	12	4	16
Column Totals	23	9	32

P-Value = 0.5000 $\quad \tau = 0.005$

Table 19. Main Effect Test: Sample Size.

	MLE	Bayes	Row Totals
$n_i = 4$	10	6	16
$n_i = 10$	13	3	16
Column Totals	23	9	32

P-Value = 0.2166 $\quad \tau = 0.043$

Table 20. Main Effect Test: R_i Levels.

	MLE	Bayes	Row Totals
equal	12	4	16
unequal	11	5	16
Column Totals	23	9	32

P-Value = 0.5000 $\tau = 0.005$

Table 21. Main Effect Test: Prior Information Strength.

	MLE	Bayes	Row Totals
uniformly weak	12	4	16
uniformly strong	11	5	16
Column Totals	23	9	32

P-Value = 0.5000 $\tau = 0.005$

Table 22. Main Effect Test: Prior Information Accuracy.

	MLE	Bayes	Row Totals
uniformly inaccurate	16	0	16
uniformly accurate	7	9	16
Column Totals	23	9	32

P-Value = 0.0004 $\tau = 0.391$

Table 23. Interaction Test: System Size and Sample Size.

	MLE	Bayes	Row Totals
$k = 5$ and $n_i = 4$	5	3	8
$k = 5$ and $n_i = 10$	6	2	8
$k = 10$ and $n_i = 4$	5	3	8
$k = 10$ and $n_i = 10$	7	1	8
Column Totals	23	9	32

$X^2 = 1.700$ P-Value $= 0.6368$ $\tau = 0.053$

Table 24. Interaction Test: System Size and R_i Levels.

	MLE	Bayes	Row Totals
$k = 5$ and R_i equal	6	2	8
$k = 5$ and R_i unequal	5	3	8
$k = 10$ and R_i equal	6	2	8
$k = 10$ and R_i unequal	6	2	8
Column Totals	23	9	32

$X^2 = 0.464$ P-Value $= 0.9268$ $\tau = 0.014$

Table 25. Interaction Test: System Size and Prior Information Strength.

	MLE	Bayes	Row Totals
$k = 5$ and weak priors	6	2	8
$k = 5$ and strong priors	5	3	8
$k = 10$ and weak priors	6	2	8
$k = 10$ and strong priors	6	2	8
Column Totals	23	9	32

$X^2 = 0.464$ P-Value $= 0.9268$ $\tau = 0.014$

Table 26. Interaction Test: System Size and Prior Information Accuracy.

	MLE	Bayes	Row Totals
$k = 5$ and inaccurate priors	8	0	8
$k = 5$ and accurate priors	3	5	8
$k = 10$ and inaccurate priors	8	0	8
$k = 10$ and accurate priors	4	4	8
Column Totals	23	9	32

$X^2 = 12.831$ P-Value $= 0.0050$ $\tau = 0.401$

Table 27. Interaction Test: Sample Size and R_i Levels.

	MLE	Bayes	Row Totals
$n_i = 4$ and R_i equal	6	2	8
$n_i = 4$ and R_i unequal	4	4	8
$n_i = 10$ and R_i equal	6	2	8
$n_i = 10$ and R_i unequal	7	1	8
Column Totals	23	9	32

$X^2 = 2.937$ P-Value $= 0.4014$ $\tau = 0.092$

Table 28. Interaction Test: Sample Size and Prior Information Strength.

	MLE	Bayes	Row Totals
$n_i = 4$ and weak priors	6	2	8
$n_i = 4$ and strong priors	4	4	8
$n_i = 10$ and weak priors	6	2	8
$n_i = 10$ and strong priors	7	1	8
Column Totals	23	9	32

$X^2 = 2.937$ P-Value $= 0.4014$ $\tau = 0.092$

Table 29. Interaction Test: Sample Size and Prior Information Accuracy.

	MLE	Bayes	Row Totals
$n_i = 4$ and inaccurate priors	8	0	8
$n_i = 4$ and accurate priors	2	6	8
$n_i = 10$ and inaccurate priors	8	0	8
$n_i = 10$ and accurate priors	5	3	8
Column Totals	23	9	32

$X^2 = 15.304$ P-Value $= 0.0016$ $\tau = 0.478$

Table 30. Interaction Test: R_i Levels and Prior Information Strength.

	MLE	Bayes	Row Totals
R_i equal and weak priors	6	2	8
R_i equal and strong priors	6	2	8
R_i unequal and weak priors	6	2	8
R_i unequal and strong priors	5	3	8
Column Totals	23	9	32

$X^2 = 0.464$ P-Value $= 0.9268$ $\tau = 0.014$

Table 31. Interaction Test: R_i Levels and Prior Information Accuracy.

	MLE	Bayes	Row Totals
R_i equal and inaccurate priors	8	0	8
R_i equal and accurate priors	4	4	8
R_i unequal and inaccurate priors	8	0	8
R_i unequal and accurate priors	3	5	8
Column Totals	23	9	32

$X^2 = 12.831$ P-Value $= 0.0050$ $\tau = 0.401$

Table 32. Interaction Test: Prior Information Strength and Accuracy.

	MLE	Bayes	Row Totals
weak, inaccurate priors	8	0	8
weak, accurate priors	4	4	8
strong, inaccurate priors	8	0	8
strong, accurate priors	3	5	8
Column Totals	23	9	32

$X^2 = 12.831$ P-Value = 0.0050 $\tau = 0.401$

Interval Estimator Tables

This section contains the contingency tables that were used to test for the significance of factor effects on the choice of the best interval estimator. Frequency counts for the tied cases (Systems 14, 26 and 32) were grouped with the frequency counts of those systems preferring the approximate beta estimator; this was necessary to keep the expected cell frequencies sufficiently large.

The single-factor tables (numbers 33 to 37) are listed first, followed by the pair-wise tables (numbers 38 to 47). For two-by-two tables, Fisher's Exact Test was used to derive the p-values; for larger tables, the p-value was found using Pearson's chi-squared statistic. The values for Goodman and Kruskal's τ, a measure of nominal association, are also included.

Table 33. Main Effect Test: System Size.

	L-M	Not L-M	Row Totals
$k = 5$	8	8	16
$k = 10$	10	6	16
Column Totals	18	14	32

P-Value = 0.3612 $\tau = 0.016$

Table 34. Main Effect Test: Sample Size.

	L-M	Not L-M	Row Totals
$n_i = 4$	9	7	16
$n_i = 10$	9	7	16
Column Totals	18	14	32

P-Value = 0.6388 $\quad \tau = 0.000$

Table 35. Main Effect Test: R_i Levels.

	L-M	Not L-M	Row Totals
equal	9	7	16
unequal	9	7	16
Column Totals	18	14	32

P-Value = 0.6388 $\quad \tau = 0.000$

Table 36. Main Effect Test: Prior Information Strength.

	L-M	Not L-M	Row Totals
uniformly weak	10	6	16
uniformly strong	8	8	16
Column Totals	18	14	32

P-Value = 0.3612 $\quad \tau = 0.016$

Table 37. Main Effect Test: Prior Information Accuracy.

	L-M	Not L-M	Row Totals
uniformly inaccurate	16	0	16
uniformly accurate	2	14	16
Column Totals	18	14	32

P-Value = 0.0000 $\quad \tau = 0.778$

Table 38. Interaction Test: System Size and Sample Size.

	L-M	Not L-M	Row Totals
$k = 5$ and $n_i = 4$	4	4	8
$k = 5$ and $n_i = 10$	4	4	8
$k = 10$ and $n_i = 4$	5	3	8
$k = 10$ and $n_i = 10$	5	3	8
Column Totals	18	14	32

$X^2 = 0.508$ $\quad$ P-Value = 0.9171 $\quad \tau = 0.016$

Table 39. Interaction Test: System Size and R_i Levels.

	L-M	Not L-M	Row Totals
$k = 5$ and R_i equal	4	4	8
$k = 5$ and R_i unequal	4	4	8
$k = 10$ and R_i equal	5	3	8
$k = 10$ and R_i unequal	5	3	8
Column Totals	18	14	32

$X^2 = 0.508$ $\quad$ P-Value = 0.9171 $\quad \tau = 0.016$

Table 40. Interaction Test: System Size and Prior Information Strength.

	L-M	Not L-M	Row Totals
$k = 5$ and weak priors	4	4	8
$k = 5$ and strong priors	4	4	8
$k = 10$ and weak priors	6	2	8
$k = 10$ and strong priors	4	4	8
Column Totals	18	14	32

$X^2 = 1.524$ P-Value $= 0.6768$ $\tau = 0.048$

Table 41. Interaction Test: System Size and Prior Information Accuracy.

	L-M	Not L-M	Row Totals
$k = 5$ and inaccurate priors	8	0	8
$k = 5$ and accurate priors	0	8	8
$k = 10$ and inaccurate priors	8	0	8
$k = 10$ and accurate priors	2	6	8
Column Totals	18	14	32

$X^2 = 25.905$ P-Value $= 0.0000$ $\tau = 0.810$

Table 42. Interaction Test: Sample Size and R_i Levels.

	L-M	Not L-M	Row Totals
$n_i = 4$ and R_i equal	5	3	8
$n_i = 4$ and R_i unequal	4	4	8
$n_i = 10$ and R_i equal	4	4	8
$n_i = 10$ and R_i unequal	5	3	8
Column Totals	18	14	32

$X^2 = 0.508$ P-Value $= 0.9171$ $\tau = 0.016$

Table 43. Interaction Test: Sample Size and Prior Information Strength.

	L-M	Not L-M	Row Totals
$n_i = 4$ and weak priors	5	3	8
$n_i = 4$ and strong priors	4	4	8
$n_i = 10$ and weak priors	5	3	8
$n_i = 10$ and strong priors	4	4	8
Column Totals	18	14	32

$X^2 = 0.508$ P-Value $= 0.9171$ $\tau = 0.016$

Table 44. Interaction Test: Sample Size and Prior Information Accuracy.

	L-M	Not L-M	Row Totals
$n_i = 4$ and inaccurate priors	8	0	8
$n_i = 4$ and accurate priors	1	7	8
$n_i = 10$ and inaccurate priors	8	0	8
$n_i = 10$ and accurate priors	1	7	8
Column Totals	18	14	32

$X^2 = 24.889$ P-Value $= 0.0000$ $\tau = 0.778$

Table 45. Interaction Test: R_i Levels and Prior Information Strength.

	L-M	Not L-M	Row Totals
R_i equal and weak priors	5	3	8
R_i equal and strong priors	4	4	8
R_i unequal and weak priors	5	3	8
R_i unequal and strong priors	4	4	8
Column Totals	18	14	32

$X^2 = 0.508$ P-Value $= 0.9171$ $\tau = 0.016$

Table 46. Interaction Test: R_i Levels and Prior Information Accuracy.

	L-M	Not L-M	Row Totals
R_i equal and inaccurate priors	8	0	8
R_i equal and accurate priors	1	7	8
R_i unequal and inaccurate priors	8	0	8
R_i unequal and accurate priors	1	7	8
Column Totals	18	14	32

$X^2 = 24.889$ P-Value = 0.0000 $\tau = 0.778$

Table 47. Interaction Test: Prior Information Strength and Accuracy.

	L-M	Not L-M	Row Totals
weak, inaccurate priors	8	0	8
weak, accurate priors	2	6	8
strong, inaccurate priors	8	0	8
strong, accurate priors	0	8	8
Column Totals	18	14	32

$X^2 = 25.905$ P-Value = 0.0000 $\tau = 0.810$

Bibliography

[1] Agresti, A. *Categorical Data Analysis*. New York: John Wiley and Sons, 1990.

[2] Auclair, P. F. Lecture materials distributed in OPER 742, Analysis for Defense Decisions. Graduate School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB OH, 1993.

[3] Bayes, T. "An Essay Towards Solving a Problem in the Doctrine of Chances," *Philosophical Transactions of the Royal Society of London*, *53*:370–418 (1763).

[4] Berger, J. O. *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag, 1985.

[5] Biswas, S. *Topics in Statistical Methodology*. New York: John Wiley and Sons, 1991.

[6] Buehler, R. J. "Confidence Limits for the Product of Two Binomial Parameters," *Journal of the American Statistical Association*, *52*:482–493 (1957).

[7] Cardaronella, R. I., and others. *Evaluation of Weapon System Capabilities: Analysts'* [sic] *Handbook*. Headquarters Strategic Air Command, Office of Science and Research, Offutt Air Force Base NE, 1989.

[8] Chao, A. and L.-C. Huwang. "A Modified Monte Carlo Technique for Confidence Limits of System Reliability Using Pass-Fail Data," *IEEE Transactions on Reliability*, *R-36*(1):109–112 (1987).

[9] Crowder, M. J., and others. *Statistical Analysis of Reliability Data*. London: Chapman and Hall, 1991.

[10] De Groot, M. *Optimal Statistical Decisions*. New York: McGraw Hill, Inc., 1970.

[11] Duran, B. S. and J. M. Booker. "A Bayes Sensitivity Analysis When Using the Beta Distribution as a Prior," *IEEE Transactions on Reliability*, *R-37*:239–247 (1988).

[12] Easterling, R. G. "Approximate Confidence Intervals for System Reliability," *Journal of the American Statistical Association*, *67*:220–222 (1972).

[13] Hays, LtCol J. R. Telephone interview. Headquarters Air Combat Command, 1993.

[14] International Mathematical and Statistical Libraries (IMSL), Inc., Houston TX. *IMSL Stat/Library User's Manual, Version 1.1*, 1989.

[15] Kapur, K. C. and L. R. Lamberson. *Reliability in Engineering Design*. New York: John Wiley and Sons, 1977.

[16] Kendall, M. G. and A. Stuart. *The Advanced Theory of Statistics, Volume 2: Inference and Relationship* (Second Edition). New York: Hafner Publishing Company, 1967.

[17] Kendall, M. G., and others. *Kendall's Advanced Theory of Statistics, Volume 1: Distribution Theory* (Fifth Edition). New York: Oxford University Press, 1987.

[18] Laplace, Pierre Simon, Marquis de. "Mémoir sur la probabilité des causes par les évenements," *Mémoirs de Mathématique et de Physique, Presentés a l'Académie Royale des Sciences par divers Savans et lûs dans ses Assemblés*, *6*:621–656 (1774).

[19] Law, A. M. and W. D. Kelton. *Simulation Modeling and Analysis* (Second Edition). New York: McGraw-Hill, Inc., 1991.

[20] Levy, L. J. and A. H. Moore. "A Monte Carlo Technique for Obtaining System Reliability Confidence Limits from Component Test Data," *IEEE Transactions on Reliability*, *R-16*(2):69–72 (1967).

[21] Lipow, M. *Measurement of Over-all* [sic] *Reliability Utilizing Results of Independent Sybsystem Tests*. Technical Report GM-TR-0165-00506, Space Technology Laboratories, Inc., 1958.

[22] Lipow, M. and J. Riley. *Tables of Upper Confidence Limits on Failure Probability of 1, 2, and 3 Component Serial Systems, Volume 1*. Technical Report AD-609-100, Washington DC: U. S. Department of Commerce, 1960.

[23] Lipow, M. and J. Riley. *Tables of Upper Confidence Limits on Failure Probability of 1, 2, and 3 Component Serial Systems, Volume 2*. Technical Report AD-636-718, Washington DC: U. S. Department of Commerce, 1960.

[24] Lloyd, D. K. and M. Lipow. *Reliability: Management, Methods and Mathematics*. Englewood Cliffs NJ: Prentice Hall, 1962.

[25] Madansky, A. "Approximate Confidence Limits for the Reliability of Series and Parallel Systems," *Technometrics*, *7*:495–503 (1965).

[26] Mann, N. L. and F. E. Grubbs. "Approximate Optimum Confidence Bounds for System Reliability Based on Component Test Data," *Technometrics*, *16*(3):335–347 (1974).

[27] Mann, N. R., and others. *Methods for Statistical Analysis of Reliability and Life Data*. New York: John Wiley and Sons, 1974.

[28] Maritz, J. S. *Empirical Bayes Methods*. London: Methuen and Co., 1970.

[29] Martz, H. F. and B. S. Duran. "A Comparison of Three Methods for Calculating Lower Confidence Limits on System Reliability Using Binomial Component Data," *IEEE Transactions on Reliability*, *R-34*(2):113–120 (1985).

[30] Martz, H. F. and R. A. Waller. *Bayesian Reliability Analysis* (Reprint Edition). Malabar FL: Krieger Publishing Company, 1991. Corrected reprint of 1982 John Wiley and Sons edition.

[31] Martz, H. F., and others. "Bayesian Reliability Analysis of Series Systems of Binomial Subsystems and Components," *Technometrics*, *30*(2):143–154 (1988).

[32] Mendenhall, W., and others. *Mathematical Statistics with Applications* (Fourth Edition). Boston: PWS-Kent Publishing Company, 1990.

[33] Mood, A. M., and others. *Introduction to the Theory of Statistics* (Third Edition). New York: McGraw-Hill Book Company, 1974.

[34] Mosleh, A. and G. Apostolakis. "Some Properties Useful in the Study of Rare Events," *IEEE Transactions on Reliability*, *R-31*:87–94 (1982).

[35] Oberhettinger, F. *Tables of Mellin Transforms*. New York: Springer-Verlag, 1974.

[36] Press, S. J. *Bayesian Statistics: Principles, Models and Applications*. New York: John Wiley and Sons, 1989.

[37] Reiser, B. and M. Jaegar. "A Comment on Buehler Optimal Confidence Bounds for Series Systems Reliability," *Statistics and Probability Letters*, *11*:65–67 (1991).

[38] Rice, R. E. and A. H. Moore. "A Monte Carlo Technique for Estimating Lower Confidence Limits on System Reliability Under Pass-Fail Data," *IEEE Transactions on Reliability*, *R-32*(4):366–369 (1983).

[39] Soms, A. P. "Lindstrom-Madden Method." *Encyclopedia of Statistical Sciences* edited by S. Kotz and N. L. Johnson, New York: John Wiley and Sons, 1986.

[40] Springer, M. D. *The Algebra of Random Variables*. New York: John Wiley and Sons, 1979.

[41] Springer, M. D. and W. E. Thompson. "Bayesian Confidence Limits for the Product of N Binomial Parameters," *Biometrika*, *53*:611–613 (1966).

[42] Steck, G. P. *Upper Confidence Limits for the Failure Probability of Complex Networks*. Technical Report SC-4133(TR), Albuquerque NM: Sandia Corporation, 1957.

[43] Sudakov, R. S. "On the Question of Interval Estimation of the Index of Reliability of a Sequential System," *Engineering Cybernetics*, *12*:55–63 (1974).

[44] Tang, J. and A. K. Gupta. "On the Distribution of the Product of Independent Beta Random Variables," *Statistics and Probability Letters*, *2*:165–168 (1984).

[45] Thompson, W. E. and R. D. Haynes. "On the Reliability, Availability and Bayes Confidence Intervals for Multicomponent Systems," *Naval Research Logistics Quarterly*, *27*:345–358 (1980).

[46] Weiler, H. "The Use of Incomplete Beta Functions for Prior Distributions in Binomial Sampling," *Technometrics*, 7:335–347 (1965).

[47] Wilks, S. S. "The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses," *Annals of Mathematical Statistics*, *9*:60–62 (1938).

[48] Willits, C. J., "The Calculation of Buehler's Optimal Confidence Limit for Series Systems with Small Test Samples." Unpublished report in partial fulfillment of the requirements for OPER 768, Nonlinear Programming. Graduate School of Engineering, Air Force Institute of Technology, Wright-Patterson AFB OH, 1993.

[49] Winterbottom, A. "Asymptotic Expansions to Improve Large Sample Confidence Intervals for System Reliability," *Biometrika*, *67*:351–357 (1980).

[50] Winterbottom, A. "The Interval Estimation of System Reliability from Component Test Data," *Operations Research*, *32*(3):628–640 (1984).

[51] Wolfe, J. E. "Bayesian Reliability Assessment from Test Data." *Proceedings of the 1976 Annual Reliability and Maintainability Symposium*. 411–419. New York: IEEE Press, 1976.

Vita

Craig Joslyn Willits was born on 13 October 1961 in Newark, New Jersey. After graduating from high school in Rumson, New Jersey in 1979, he attended Rutgers, The State University of New Jersey on a four-year Air Force Reserve Officers Training Corps (AFROTC) scholarship. In May 1983, he received a Bachelor of Arts degree with a major in mathematics and a commission in the Air Force Reserve.

Upon entering active duty in October 1983, Captain Willits was assigned to the Space Transportation System Program Office at Headquarters Space Division, Los Angeles AFS, California, where he performed duties in resource management and requirements analysis. After attending the Aircraft Maintenance Officer Course (AMOC) at Chanute AFB, Illinois, from December 1986 to May 1987, he was transferred to the 22nd Air Refueling Wing, March AFB, California, to serve a career-broadening tour as an aircraft maintenance officer. In December 1989, he was assigned to the 49th Test Squadron, Barksdale AFB, Louisiana, as Chief of the Analysis Section. He entered the Graduate Operations Research program at the Air Force Institute of Technology in August 1992.

Captain Willits, who holds a regular Air Force commission, is a Distinguished Graduate of both AFROTC and AMOC. He is a member of the Omega Rho honor society, the Operations Research Society of America, the Military Operations Research Society and the Association for Computing Machinery.

Captain Willits and his wife, the former Sara Elizabeth Brown of Hackettstown, New Jersey, have three sons: Steven, David and Joseph.

Permanent address: 25 Lakeside Avenue
Rumson, New Jersey 07760

END