PROPERTIES OF THE MAXIMUM LIKELIHOOD AND BAYESIAN ESTIMATORS OF AVAILABILITY

by

WAY KUO

B.S., Nuclear Engineering, National Tsing-Hua University, 1972

M.S., Nuclear Engineering, University of Cincinnati, 1975

M.S., Industrial Engineering, Kansas State University, 1977

Ph.D., Engineering, Kansas State University, 1979

A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1981

Approved by:

R. Nassar

Major Professor

TABLE OF CONTENTS

1. INTRODUCTION

Increasing complexity of modern equipment, both in the military and commercial areas, has brought with it new engineering problems involving high performance, reliability, and maintainability. Reliability has long been considered, during system design, as a measure of system effectiveness. However, it has proved to be an incompleted measure because it does not consider maintainability, another important aspect of system performance. With increasing complexity and the resulting high operational and maintenance costs, greater emphasis has been placed on reducing system maintenance while improving reliability. In this regard availability, which is a combined measure of maintainability and reliability, has received wide usage as a measure of maintained systems effectiveness.

Availability is defined as the probability that the system is operating satisfactorily at any point in time under stated conditions. Lie et al. [8] surveyed and classified the literature related to the availability of various systems. Depending on the time interval considered, availability can be classified into: (1) instantaneous availability, (2) average uptime availability, and (3) steady-state availability.

Instantaneous availability, $g(t)$ is defined as the probability that the system is operational at any time t, where $0 \le t < \infty$. Average uptime availability, $A(T)$, is the proportion of time in a specified time $(0,T)$ that the system is available for use and is expressed as

$$g(T) = \frac{1}{T} \int_0^T g(t)dt$$

Steady-state availability is then the instantaneous availability at time ∞, which is the limiting case of the instantaneous availability. Since both steady-state availability and average uptime availability are special cases of instantaneous availability, the derivation and evaluation of instantaneous availability is fundamental and of interest.

Availability estimation is no more than a typical statistical estimation problem. Two typical procedures can be used, namely, non-Bayesian inference such as the maximum likelihood estimate technique and Bayesian inference. Kuo [7] recently reported on the maximum likelihood estimator of availability and the Bayesian estimator of availability for gamma distributed system cycle time and on time. Properties of these estimators have not been studied.

It is controversial to use Bayesian approach in statistical estimation. However, Bayesian apprach has its merits in reliability/availability problems especially when (1) the sample size is small due to the expensive or time consuming testing procedure, and (2) prior information is available in practical engineering problem from past experience.

This study is an extension of a previous study [7, 11]. Some properties of the maximum likelihood estimator of availability and the Bayesian estimator of availability for negative exponential distributed system on time and off time are investigated through simulation. It can be concluded that: (1) Both the maximum likelihood estimator of availability and the Bayesian estimator of availability are biased, (2) the maximum likelihood estimator of availability has a larger variance, a wider range, and a wider 90% C.I. than those of the Bayesian estimator of availability, and (3) Bayesian estimator of availability is insensitive to the prior information within at least a certain range.

Future study in the availability estimation problems is directed toward the use of the nonparametric Bayesian estimation techniques. Some preliminary study of the nonparametric Bayesian estimation of life distributions, which can be applied to system on time and system off time, has been investigated.

2. THE SYSTEM AND THE ASSUMPTIONS

Statement of the System:

Consider a system which can be in one of two states, "on" or "off", when in the "on" state, the system is operating and in the "off" state, the system is failing and under repair. We assume that at time 0, it is "on". The system is then in service until it fails at a random time T_{on} with the distribution function $F_{on}(t)$. When it fails, it is then in the "off" state and under repair for a random time T_{off} with the distribution function $F_{off}(t)$. Then the cycle repeats by being operative for a random time and then being inoperative for another random time. Successive times to breakdown and to repair are assumed to be independent.

It is assumed that the events of either operative or inoperative are independent of time. A complete cycle time T is also a random variable which is equal to the addition of random variables T_{on} and T_{off}. T then is a random variable of the time from 0 to the time at which the system failed, was repaired and just restored to the operative state. (See Fig. 1.)

The Assumptions:

Assume that a system cycle time T, and on time, T_{on}, are gamma distributed with pdf's:

$$f_T(z) = \frac{\lambda}{(k-1)!}(\lambda z)^{k-1}e^{-\lambda z} \tag{1}$$

and

$$f_{T_{on}}(y) = \frac{\beta}{(\alpha-1)!}(\beta y)^{\alpha-1}e^{-\beta y} \tag{2}$$

where α, $\beta>0$, x, $y>0$, k and α are positive integers.

With the parameters λ, k, β, and α known and time, t, given, the time-dependent availability function evaluated in [7] is

$$g(t;\lambda,k,\beta,\alpha) = \sum_{=0}^{-1} p_0(\ell;\beta t)$$

$$+ \lambda \int_0^t \sum_{\ell=0}^{\alpha-1} p_0[\ell;\beta(t-s)] \sum_{\substack{q=k-1\\2k-1\\\vdots\\\text{etc.}}}^{\infty} p_0(q,\lambda s) \quad ds \qquad (3)$$

when T and T_{on} are independent.

To estimate the availability function, our main object is to present a Bayesian availability estimator and to compare it with the maximum likelihood estimator of availability.

3. AVAILABILITY ESTIMATORS

Maximum Likelihood Estimator of Availability:

Suppose that z samples of z and y are drawn from T and T_{on} respectively and the observations are denoted by (z_i, y_i), i=1,2,...,n. The maximum likelihood estimates of λ and k are given in Kuo[7]. They are given by the simultaneous solutions of

$$\lambda = \frac{nk}{\sum_{i=1}^{n} x_i} \tag{4}$$

and

$$e^{\ell n\lambda + \frac{1}{n}\ell n(\sum_{i=1}^{n} x_i)} \leq k \leq 1 + e^{\ell n\lambda + \frac{1}{n}\ell n(\sum_{i=1}^{n} x_i)} \tag{5}$$

Similarly the maximum likelihood estimates of β and α are the simultaneous solutions of

$$\beta = \frac{n\alpha}{\sum_{i=1}^{n} y_i} \tag{6}$$

and

$$e^{\ell n\beta + \frac{1}{n}\ell(\sum_{i=1}^{n} y_i)} \leq \alpha \leq 1 + e^{\ell n\beta + \frac{1}{n}\ell n(\sum_{i=1}^{n} y_i)} \tag{7}$$

Finally the maximum likelihood estimator of availability is given by eq. (3) after substituting $\hat{\lambda}_{ML}$, $\hat{k}_{ML}$, $\hat{\beta}_{ML}$, and $\hat{\alpha}_{ML}$ obtained from eqs. (4)-(7) into eq. (3).

Bayes Theorem

The primary mathematical tool for Bayesian analysis is called Bayes theorem in honor of Thomas Bayes who studied the topic in the mid-1700's. Crellin [3] discusses the theorem as well as its uses and misuses. The philosophy behind Bayes theorem is that two sources of information exist regarding the parameters of the data model. First, in assuming a prior model for the parameter or parameters of interest, we suppose that the assumed prior model summarizes and represents the totality of knowledge available concerning the parameters prior to the observation of data. Bayes theorem is a technique for combining the information about the parameters from both the prior model and the data information into a single model.

The combined model provided by Bayes theorem is called a posterior model because it represents the state of knowledge about the parameters after sample data information is combined with the prior data information. To state the Bayes theorem, let $f(t_i; \tilde{\theta})$ denote the data model for an observation t_i on a variable T. If $p(\tilde{\theta})$ is the prior model for the parameter vector $\tilde{\theta}$, and if a sample $(t_1, t_2, \ldots, t_n)$ of n independent observations on T is observed, the posterior model for $\tilde{\theta}$ is

$$h(\tilde{\theta}|t_1,t_2,\ldots,t_n) = \frac{p(\tilde{\theta}) \sum_{i=1}^{n} f(t_i|\tilde{\theta})}{\int_{\Omega} p(\tilde{q}) \sum_{i=1}^{n} f(t_i|q)\, d\tilde{q}} \tag{8}$$

where Ω is the parameter space of $\tilde{q}$.

Just as $p(\tilde{\theta})$ portrays the experimenter's feelings (prior model) regarding the possible values of $\tilde{\theta}$ before observing sample data, $h(\tilde{\theta}|t_1,t_2,\ldots,t_n)$ expresses the probability model for $\tilde{\theta}$ after adjusting

$p(\tilde{\theta})$ for the influence of sample data -- hence the name posterior model. Consequently, decisions and inferences made by using the posterior model are influenced by both the sample data information and the prior model information about $\tilde{\theta}$.

Bayesian Estimator of Availability:

To implement the Bayesian approach in availability, which is a measure of system effectiveness, the joint distribution functions of λ, k, β, and α should be assigned. This assignment is too complicated to work out analytically. It is usually possible to fix one of two parameters in a gamma distribution and allow the other parameter to have a certain distribution. Therefore, to approach this problem, we allow k and α to be fixed constant positive integers and λ and β to have the variations of negative exponential distributions

$$f_\lambda(\lambda) = \mu e^{-\mu\lambda} \tag{9}$$

and

$$f_\beta(\beta) = \nu e^{-\nu\beta} \tag{10}$$

where μ and ν are undetermined positive constants and λ and β are positive and mutually independent random numbers.

$f_\lambda(\lambda)$ and $f_\beta(\beta)$ are the so-called prior information or prior distributions of λ and β, respectively. Using the Bayes theorem, we combine eqs. (1) and (9) to obtain the posterior distribution of λ, and combine eqs. (2) and (10) to obtain the posterior distribution of β. Let $f_\lambda(\lambda; x_1,x_2,\ldots,x_n)$ be the posterior distribution of λ given the sample $x_1,x_2,\ldots,x_n$,

$$f_\lambda(\lambda;\, x_1,x_2,\ldots,x_n) = \frac{f_\lambda(\lambda) \cdot L(x_1,x_2,\ldots,x_n;\, \lambda,\, k)}{\int_0^\infty f_\lambda(\lambda)\, L(x_1,x_2,\ldots,x_n;\, \lambda,k)\, d\lambda}$$

$$= \frac{(\mu + \sum_{i=1}^{n} x_i)}{\Gamma(kn+1)} [\lambda(\mu + \sum_{i=1}^{n} x_i)]^{kn}\, e^{-\lambda(\mu + \sum_{i=1}^{n} x_i)} \tag{11}$$

Similarly we can obtain

$$f_\beta(\beta;\, y_1,y_2,\ldots,y_n)$$

$$= \frac{(\nu + \sum_{i=1}^{n} y_i)}{\Gamma(\alpha n+1)} [\beta(\nu + \sum_{i=1}^{n} y_i)]^{\alpha n}\, e^{-\beta(\nu + \sum_{i=1}^{n} y_i)} \tag{12}$$

The mean or expected value of the availability function $g(t;\, \lambda,\, k,\, \beta,\, \alpha)$ is a Bayes estimator of the availability if the squared error loss function is presumed. A Bayes estimator, $\hat{g}_B(t;\, \lambda,\, k,\, \beta,\, \alpha)$ for the $g(t;\, \lambda,\, k,\, \beta,\, \alpha)$, is the function that minimizes the expected value of the loss function with respect to the posterior model (or prior model when no data are available) of λ and β. That is

$$\hat{g}_B(t|k,\alpha) = \int_0^\infty \int_0^\infty g(t|\lambda,\, k,\, \beta,\, \alpha)\, f_\lambda(\lambda)\, f_\beta(\beta)\, d\lambda d\beta \tag{13}$$

when no data are observed, and

$$\hat{g}_B(t|\mu,\, k,\, \nu,\, \alpha) = \int_0^\infty \int_0^\infty g(t|\lambda,\, k,\, \beta,\, \alpha)\, f_\lambda(\lambda;\, x_1,x_2,\ldots,x_n) \cdot f_\beta(\beta;\, y_1,y_2,\ldots,y_n)\, d\lambda d\beta \tag{14}$$

when the samples (x_i, y_i), $i=1,2,\ldots,n$ are available.

4. AVAILABILITY ESTIMATORS OF EXPONENTIALLY DISTRIBUTED T_{off} AND T_{on}

If a system off time, T_{off}, and on time, T_{on}, are negative exponentially distributed, k in eq. (1) and α in eq. (2) are both equal to one. Replacing λ by η in eq. (1) we obtain

$$f_{T_{off}}(x) = \eta e^{-\eta x} , \quad x>0 \tag{15}$$

$$f_{T_{on}}(y) = \beta e^{-\beta y} , \quad y>0 \tag{16}$$

It can be shown that the availability function with the above pdf's has the form

$$g(t; \lambda, \beta) = \frac{\eta}{\beta+\eta} + (1 - \frac{\eta}{\beta+\eta}) e^{-(\beta+\eta)t} \tag{17}$$

Let two samples (x_i, y_i), i=1,2,...,n be drawn from T_{off} and T_{on}, respectively. The maximum likelihood estimators of η and β are then given by

$$\hat{\eta}_{MLE} = \frac{n}{\sum_{i=1}^{n} x_i} \tag{18}$$

$$\hat{\beta}_{MLE} = \frac{n}{\sum_{i=1}^{n} y_i} \tag{19}$$

and the maximum likelihood estimator of the availability function is:

$$\hat{g}_{MLE}(t; \lambda, \beta) = \frac{\hat{\eta}_{MLE}}{\hat{\beta}_{MLE} + \hat{\eta}_{MLE}} + (1 - \frac{\hat{\eta}_{MLE}}{\hat{\beta}_{MLE} + \hat{\eta}_{MLE}}) e^{-(\hat{\beta}_{MLE}+\hat{\eta}_{MLE})t} \tag{20}$$

To implement the Bayesian approach, let η and β be random variables with the pdf's given by eqs. (9) and (10). The posterior distributions of η and β are special cases obtained from eqs. (11) and (12):

$$f_\eta(\eta;\, x_1,x_2,\ldots,x_n) = \frac{(\mu + \sum_{i=1}^{n} x_i)}{\Gamma(n+1)} [\eta(\mu + \sum_{i=1}^{n} x_i)]^n \, e^{-\eta(\mu + \sum_{i=1}^{n} x_i)} \tag{21}$$

$$f_\beta(\beta;\, y_1,y_2,\ldots,y_n) = \frac{(\nu + \sum_{i=1}^{n} y_i)}{\Gamma(n+1)} [\beta(\nu + \sum_{i=1}^{n} y_i)]^n \, e^{-\beta(\nu + \sum_{i=1}^{n} y_i)} \tag{22}$$

when $k = \alpha = 1$.

If sample data are not available, the Bayesian estimate of availability is simply

$$\hat{g}_B(t) = \int_0^\infty\int_0^\infty [\frac{\eta}{\beta+\eta} + (1 - \frac{\eta}{\beta+\eta})\, e^{-(\beta+\eta)t}]\, \mu\nu e^{-\mu\eta}\, e^{-\nu\beta}\, d\beta d\eta \tag{23}$$

If, instead, the sample information (x_i, y_i), $i=1,2,\ldots,n$ are available, the Bayesian availability estimate is

$$\hat{g}_B(t;\, \mu,\, \nu) = (\frac{d}{t+d})^{n+1}\, (\frac{h}{t+h})^{n+1}$$

$$+ \frac{(dh)^{n+1}}{f^2} \int_0^\infty \int_\eta^\infty \frac{\eta}{\beta'} [1-e^{-\beta' t}]\, \eta^n(\beta'-\eta)^n\, e^{-\eta d}\, e^{-(\beta'-\eta)h}\, d\beta' d\eta \tag{24}$$

where $d = \mu + \sum_{i=1}^{n} x_i$, $h = \nu + \sum_{i=1}^{n} y_i$, and $f = \Gamma(n+1)$.

Notice that μ is the mean repairing time and ν is the mean failure time. Hence d and h are the weighted mean repairing time and mean failure time. In other words, the maximum likelihood estimator of availability considers sample information only, Bayesian estimator of availability without data considers prior mean information only, while Bayesian estimator of availability with data considers both.

Difficulty of the Proposed Bayesian Approach

When the sample data are not available, the proposed Bayesian approach is evaluated by

$$\hat{g}_B(t) = \frac{\mu}{t+\mu} \times \frac{\nu}{t+\nu} + \mu\nu\int_0^\infty\int_\eta^\infty \eta e^{-(\mu-\nu)} \quad [\frac{1}{\beta'}(1 - e^{-\beta' k})\ e^{-\nu\beta'}]\ d\beta' d\eta \qquad (25)$$

When the sample data are available, the proposed Bayesian availability is evaluated by

$$\hat{g}_B(t) = (\frac{t}{t+d})^{n+1} \ (\frac{h}{t+h})^{n+1}$$

$$+ \frac{(dh)^{n+1}}{f^2} \int_0^\infty\int_\eta^\infty \frac{\eta}{\beta'} [1 - e^{-\beta' t}]\ \eta^n(\beta' - \eta)^n\ e^{-\eta d}\ e^{-(\beta'-\eta)n}\ d\beta' d\eta \qquad (26)$$

To evaluate eqs. (25) and (26), numerical integration is necessary. One should be careful in selecting the upper and lower limits in applying the numerical integration method. The results presented for this numerical example are obtained by Simpson's Rule. Any advanced numerical integration techniques may improve the results of the proposed Bayesian approach.

5. A NUMERICAL EXAMPLE AND SIMULATION RESULTS

Two samples of size 49 each in Epstein's work [4] have been tested and found to conform to one parameter negative exponential distributions. Let the first set of data, denoted by $y_1^{'}$, $y_2^{'}$ ..., $y_{49}^{'}$, be the observed failure times (on times) for a system in a life testing. Let another set of data denoted by $x_1^{'}$, $x_2^{'}$, ..., $x_{49}^{'}$ be the observed repair times (off times) for the same system in the down state. It is assumed that $y_i^{'}$ and $x_i^{'}$ are independent events for i = 1, 2, ..., 49. The data for $y_i^{'}$ and $x_i^{'}$ are given in Table 1 where the mean values of $y_i^{'}$ and $x_i^{'}$ are 1042.00 and 104.20 respectively. It is shown in the Table that the failure times are 10 times of the repair times.

It is usually expensive and time consuming to obtain all 49 pairs of data as given by Table 1. It is common that only a small part of the 49 pairs is obtainable. Let us assume that 5 pairs of data points are available from the life testing experiment. Five data points are randomly drawn from the 49 failure times and another 5 points from the 49 repair times as given in Table 2. Let y_i and x_i be the failure time and repair time, respectively, for i = 1, 2, ..., 5. The mean values of y_i and x_i for the sample of size 5 are 1083.80 and 95.28, respectively. It is clear that 1083.80 and 95.28 are very close to 1042.00 and 104.20 which are the mean values of system on time and off time in the sample of size 49. Let the sample of size 5 with means 1083.80 and 95.28 be designated by Set 1.

By the same sampling procedure, another set of 5 data points is drawn from the 49 failure times and a set of 5 data points is drawn from the 49 repair times. These 5 pairs of data points are given in Table 3 and designated by Set 2. The mean values of the failure time and repair time in Set 2 sample are 350.8 and 132.42, respectively.

Since the failure times and repair times given in Set 1 and Set 2 are drawn from the sample of size 49 which comes from 2 negative exponential distributions, it would be valid to assume that y_i and x_i, $i = 1, 2, \ldots, 5$ in Set 1 and Set 2 samples both have negative exponential underlying distributions.

With equal probability that each failure time and repair time can be drawn as in the above case, we need to simulate other possible situations to determine the properties of the proposed Bayesian estimator. Specifically, we wish to determine, the mean, the variance, the 90% C.I. and the range of the estimator given a set of prior information as compared with those of the maximum likelihood estimator.

The simulation procedures are given as follows:

Step 1. Through a random procedure, 100 samples, each of 5 failure times and 5 repair times, are selected from negative exponential distributions with parameters 1042 and 104.2 for the failure time and the repair time.

Step 2.

2.1 A set of prior information on the failure times and the repair times are specified.

2.2 Eq. (23) is used to calculate the Bayesian availability when no data are available.

2.3 For each of the 100 samples, eq. (24) is used to calculate Bayesian availabilities. The mean, the variance, and the 90% C.I. for these estimators are calculated.

Step 3. Eq. (20) is used to find the maximum likelihood estimates of availabilities for 100 samples. The mean, the variance, and the 90% C.I. are obtained for these estimators.

Step 4. Steps 2 and 3 are repeated for both the unsteady-state condition (T = 200), and the steady-state condition (solution converges by time T = 400).

A Fortran program based on the proposed Bayesian approach with various prior information is written to evaluate Step 2. The program is listed in Appendix 1. A Fortran program to evaluate Step 3 is listed in Appendix 2.

Four different prior information are applied for the Bayesian approach (Table 4). Prior 1 overestimates the failure time and underestimates the repair time which results in overestimating the availability. Prior 2 underestimates the failure time and gives the right information on the repair time which results in underestimating the availability. Prior 3 gives the right information on the failure time but underestimates the repair time which results in overestimating the availability. Prior 4 gives the right information on the failure time and the repair time which results in the true information on the availability.

In the following discussions, prior 1 refers to Bayesian estimation with μ = 90 and ν = 1200, prior 2 with μ = 104.2 and ν = 800, prior 3 with μ = 80.0 and ν = 1042 and prior 4 with μ = 104.2, and ν = 1042.

For T = 200, the maximum likelihood estimates of availabilities are shown in Table 5. Bayesian estimates of availabilities with prior 1, prior 2, prior 3, and prior 4 are shown in Table 6, Table 7, Table 8, and Table 9, respectively. Based on Tables 5 - 9, frequency distributions of the instantaneous availability estimators are drawn in Fig. 2. Similarly, for the steady-state situation, the maximum likelihood estimates of availabilities are shown in Table 10. Bayesian estimates of availabilities with prior 1, prior 2, prior 3, and prior 4 are shown in Table 11, Table 12, Table 13, and Table 14, respectively. Based on Tables 10 - 14, frequency distributions of the steady-state availability estimators are drawn in Fig. 3.

The means, the variances, the 90% C.I.'s, and the ranges of various estimators are outlined in Table 15 for T = 200 and Table 16 for the steady-state situation. Some conclusions can be drawn from these tables. Referring to Table 15 for the instantaneous case, it is seen that (1) Availability esimators with a small amount of data are biased whether by the maximum likelihood estimation technique or the Bayesian approach. (2) Without sample information available, Bayesian availability with prior 1 has the highest value (0.928), while with prior 2 it has the lowest value (0.886). These results reflect the fact that prior 1 overestimates the failure time but underestimates the repair time, and prior 2 underestimates the failure time. These results, however, do not exist when the sample information is available. (3) With sample information available, there is no significant difference among the means of the various estimators (ranged from 0.8067 of Bayesian with prior 2 to 0.8383 of Bayesian with prior 4). However, the variance of the maximum likelihood estimator (0.0154) is the highest among the estimators (0.0041 for prior 1, 0.0074 for prior 2, 0.0059 for prior 3, and 0.0048 for prior 4). This confirms the small variance property of Bayesian inference. This is at least true for the prior information of a ratio of failure time to repair time between 8.00 to 13.33, when the true ratio is 10.00. (4) With the sample information available, availability estimator obtained by the maximum likelihood estimation technique has the widest 90% simulation confidence interval (0.612, 0.924) and the widest range (0.115, 0.950) among the 5 different estimators.

The above statements are also true for the steady-state availabilities as shown in Table 16.

6. CONCLUSIONS

Availability is an important measure of the system effectiveness. In this study, negative exponential distributions have been imposed on a system's on time and off time. Some statistical properties of the maximum likelihood and Bayesian estimators of availability have been investigated through computer simulation. It has been shown that the maximum likelihood estimator has larger variance and wider range than those of the Bayesian estimators, while both of the estimators are biased. Therefore, for a small amount of data, the Bayesian approach seems superior.

The Bayesian approach in this study also shows its insensitivity to a prior chosen within a certain range. This may not be true when the prior chosen is far from the true value.. However, prior information about a system of interest is always available, hence Bayesian inference is valuable in dealing with engineering reliability problems.

The proposed simulation procedures should be extended to a wider range of parameters and to a system of T_{on} and T_{off} other than the negative exponential distributions.

7. NONPARAMETRIC BAYESIAN ESTIMATION OF AVAILABILITIES: PRELIMINARY RESULTS AND FUTURE INVESTIGATIONS

In the availability estimation problems, both the distribution functions of the system on time and off time, i.e. $F_{T_{on}}$ and $F_{T_{off}}$, should be estimated. Parametric estimation of the distribution in the Bayesian sense has been widely studied, whereas nonparametric Bayesian approach has not been used. There is a strong impetus to use nonparametric approach in solving engineering reliability problems when only a small amount of data is available, and Bayesian inference when one wants to use one's past experience or subjective judgment.

Several recent studies in nonparametric Bayesian estimation of life distribution functions have been reviewed in [5]. The feature of these nonparametric estimations of life distributions is using a weak set of assumptions, as compared to the more restrictive parametric models, to get the estimation of the distribution. Once the estimate of the distribution is obtained, one can predict the probability of failure at any given time. Besides, nonparametric estimation techniques have the advantage of being relatively insensitive to outlines in the data.

Type of data. Life testing has the following common sampling forms. (1) Accelerated sample: Samples of certain devices are subject to conditions of greater stress than that encountered under normal operation, and from the results for those high-stress environments (may or may not include normal stress), an estimate of performance of the device under normal operation is obtained. This sampling method is used when lifetime tends to be long and the time consumed in testing a sample of a certain device may be excessive. (2) Nonaccelerated sample: Samples are tested under conditions of normal operation only.

The above sampling schemes are distinguished by the following types of data.

(1) Type I censored data: A test is conducted on n items, as each failure occurs, the time is recorded. $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$ are the observed ordered failure times of the r items, $r \leq n$. The test terminates at a preassigned time.

(2) Type II censored data: A test is conducted on n items and as each failure occurs, the time is recorded. $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$ are the observed ordered lifetimes of the r items, $r \leq n$. The test teminates when a preassigned number of failures, r, has occurred.

(3) Mixed censored data: A test is conducted on n items and as each failure occurs, the time is recorded. $X_{(1)}, X_{(2)}, \ldots, X_{(r)}$ are observed lifetimes of the r items, $r \leq n$. The test terminates when a preassigned number of failures, r, has occurred or a preassigned time has been reached, whichever comes first.

In either type of data, we have two methods of sampling. (1) With replacement: Items that fail are immediately replaced by new items having the same expected life distribution. (2) Without replacement: Items that fail are not replaced.

Moreover, in each operating method of Type I censored data there are three types of observations.

(1) Real observation: $X_i = x_i$

(2) Right censored data: $X_i > x_i$ (exclusive censoring) or

$X_i \geq x_i$ (inclusive censoring)

This is usually encountered when one preassigns a different time (t_i) for each different sample, X_i.

(3) Left censored data: $X_i < x_i$ (exclusive censoring) or

$$X_i \leq x_i \quad \text{(inclusive censoring)}$$

For nonaccelerated type I data without replacement, the following 3 nonparametric techniques have been investigated:

(1) Kaplan and Meier's PL method [6]. Let $T_1, \ldots, T_N$ be a random sample of values of the random variable T (called the lifetime), and $L_1, \ldots, L_N$ be a sample of the random variable L (called limits of observation) where T and L are assumed independent. We observe $t_i = \min(T_i, L_i)$ $i = 1, 2, \ldots, N$. For each item it is known whether one has

$$T_i \leq L_i \quad t_i = T_i \quad \text{(a death)}$$

or

$$L_i < T_i \quad t_i = L_i \quad \text{(a loss)}$$

Let N be the total sample size. If one lists and labels the N observed lifetimes (whether to death or loss) in order of increading magnitude $0 \leq T_1' \leq t_2' \leq \ldots \leq t_N'$, then the estimator of survival function is

$$P(t) = \prod_r [(N-r)/(N-r+1)]$$

where r assumes those values for which $t_r' \leq t$, and t_r' measures the time to death.

(2) Susarla and Van Ryzin's method [10]. Let $X_1, \ldots, X_n$ be the true survival times of n individuals which are censored on the right by n

follow-up times, $Y_1, \ldots, Y_n$. It is assumed that the X_i are independent identically distribution function $F(u)$, where F is distributed as a Dirichlet process on $R^+ = (0,\infty)$, and that the parameter $\alpha(\cdot)$ is known (see Ferguson [5], P. 116 for the definition of a Dirichlet process). The observable data are:

$$Z_i = \min\{X_i, Y_i\}$$

$$\delta_i = \begin{cases} 1 \text{ if } X_i \leq Y_i \\ 0 \text{ if } X_i < Y_i \end{cases} \quad i = 1, \ldots, n$$

Assume that $Y_1, \ldots, Y_n$ are mutually independent random variables which are also independent of $X_1, \ldots, X_n$ where Y_i is distributed as H_i, $H_i(u) = P_r(Y_i \leq u)$, $i = 1, \ldots, n$. Note that if $\delta_i = 1$, the Z_i in the pair (Z_i, δ_i) which is observed is a true lifetime; and if $\delta_i = 0$, then Z_i is an exclusive right censored data. Let $Z_1, \ldots, Z_k$ be the real observations and $Z_{k+1}, \ldots, Z_n$ be the exclusive right censored observations. Also, let $Z_{(k+1)}, \ldots, Z_{(m)}$ denote the distinct observations among the exclusive right censored observations $Z_{k+1}, \ldots, Z_n$. Let λ_j denote the number of exclusive right censored observations that are equal to $Z_{(j)}$, for $j = k+1, \ldots, m$, and let $N(u)$ and $N^+(u)$ denote the number of observations greater than or equal to u and the number greater than u, respectively. Then the nonparametric estimator $\hat{S}(u)$ of survival function $S(u)$ under the squared errors loss

$$L(\hat{S},S) = \int_0^\infty (\hat{S}(u)-S(u))^2 \, dw(u)$$

with w being a weight function, is

$$\hat{S}(u) = \frac{\alpha(u,\infty)+N^+(u)}{\alpha(R^+)+n} \prod_{j=k+1}^{\ell} \left\{ \frac{\alpha[Z_{(j)},\infty)+N(Z_{(j)})}{\alpha[Z_{(j)},\infty)+N(Z_{(j)})-\lambda_j} \right\}$$

(3) Ferguson and Phadia's Method [5].

This method is an extension of the Susarla and Van Ryzin's Method to a more general class of prior distribution for F(u), namely the process neutral to the right introduced by Doksum [12].

A process F(t) is said to be a random distribution function neutral to the right if it can be written in the form

$$F(t) = 1 - e^{-Y_t}$$

where Y_t is a process with independent increments such that (a) Y_t is non-decreasing a.s., (b) Y_t is right continuous a.s., (c) $\lim_{1\to-\infty} Y_t = 0$ a.s., and (d) $\lim_{1\to+\infty} Y_t = \infty$ a.s.

Let $F = 1-e^{-Y_t}$ be a random distribution function neutral to the right, and let $X_1, \ldots, X_n$ be a sample of size n from F. Assume that the observational data has three forms, m_1 real observations $X_1 = x_1, \ldots, X_{m_1} = x_{m_1}$, m_2 exclusive censorings $X_{m_1+1} > X_{m_1+1} \ldots, X_{m_1+m_2} > X_{m_1+m_2}$, and m_3 inclusive censorings $X_{m_1+m_2+1} \geq x_{m_1+m_2+1}, \ldots, X_{m_1+m_3} \geq x_{m_1+m_2+m_3}$ where $m_1+m_2+m_3 = n$. Let $u_1, \ldots, u_k$ be the distinct values among $X_1, \ldots, x_n$, ordered so that $u_1 < \ldots < u_k$. Let $\delta_1, \ldots, \delta_k$ denote the number of real observations at $u_1, \ldots, u_k$ respectively, let $\lambda_1, \ldots, \lambda_k$ denote the number of exclusive censorings at $u_1, \ldots, u_k$ respectively, and let $\mu_1, \ldots, \mu_k$ denote the number of inclusive censorings at $u_1, \ldots, u_k$ respectively so that $\Sigma_1^k \delta_i = m_1$, $\Sigma_1^k \lambda_i = m_2$, and $\Sigma_1^k \mu_i = m_3$. Let $h_j = \Sigma_{i=j+1}^k(\delta_i+\lambda_i+\mu_i)$ denote the number of the x_i greater than u_j, and j(t) denote the number of u_i less than or equal to t.

Assume that the independent increments of a process Y_t has gamma distribution with shape parameter $\nu(t)$ and scale parameter τ independent of t, and that $\nu(t)$ is continuous. Then,

$$\hat{S}(t) = \left(\frac{h_{j(t)}+\tau}{h_{j(t)}+\tau+1}\right)^{\nu(t)}$$

$$\Pi_{i=1}^{j(t)} \left[\left(\frac{(h_{i-1}+\tau)(h_i+\tau+1)}{(h_{i-1}+\tau+1)(h_i+\tau)}\right)^{\nu(u_i)} \frac{\zeta_G(h_i+\lambda_i+\tau+1,\delta_i)}{\zeta_G(h_i+\lambda_i+\lambda,\delta_i)} \right]$$

where

$$\begin{aligned}\zeta_G(\alpha,\beta) &= \Sigma_{i=0}^{\beta-1} \binom{\beta-1}{i} (-1)^i \log\left(\frac{\alpha+i+1}{\alpha+1}\right) & \beta \geq 1 \\ &= 1 & \beta = 0\end{aligned}$$

If our prior guess at the shape of S(t) is given by $S_0(t)$, then for fixed τ, $\nu(t)$ is

$$\nu(t) = \log(S_0(t))/\log(\tau/(\tau+1))$$

The nonparametric Bayesian estimation procedures, as the ones described, applied to the failure time distribution can also be used to estimate the repair time distribution.

We generate, from a gamma distribution (α=3, β=0.5 in eq. (2)), 5 true lifetime observations and 5 exclusive right censored data. The probabilities of survival at 10 different time indices by 3 different estimation techniques are listed in Table 17 and drawn in Fig. 4. From the same gamma distribution, another 19 set of observations are drawn. The probability of survivals for the 19 sets of data are obtained by using 3 different estimation techniques. The means, the variances, and the ranges for ten different time index and by different estimation techniques are summarized in Table 18.

From Table 18, the variances of the estimators of both the Susarla and Van Ryzin's Method and the Ferguson and Phadia's Method are smaller than those of the Kaplan and Meier's Method. The estimates of the Kaplan and Meier's Method drop fast with regard to time. The estimates of the Susarla and Van Ryzin's Method are greater than the estimates of the Ferguson and Phadia's Method by time 8. After time 8, the Susarla and Van Ryzin's Method underestimates the survival probability, while Ferguson and Phadia's Method overestimates the survival probability.

The computer program to calculate Table 17 is given in Appendix 3. For details of the computation, see [5, 6, 10].

These nonparametric (Bayesian) estimations of distributions serve as the distributions of T_{on} as well as T_{off}. These estimators of T_{on} and T_{off} can be used to estimate the availability according to the methodology presented in [7]. This form of a nonparametric Bayesian estimate of availability is in variance with the classical Bayesian estimators presented in the previous chapters. It would be of interest to compare these two kinds of availability estimates as to efficiency and robustness. It is also of interest to follow up on this preliminary investigation and study the properties of these nonparametric (Bayesian) estimations of life distribution as well as the properties of the availability estimators resulting from them.

REFERENCES

1. D. M. Brender, "The prediction and measurement of system availability: a Bayesian treatment," IEEE Trans Reliability, vol. R-17, no. 3, pp. 127-138, (Sept., 1968).

2. D. N. Brender, "The Bayesian assessment of system availability: advanced applications and techniques," IEEE Trans Reliability, vol. R-17, no. 3, pp. 138-147 (1968).

3. G. L. Crellin, "The philosophy and mathematics of Bayes equation," IEEE Trans Reliability, vol. R-21, no. 3, pp. 131-135, (Aug., 1972).

4. B. Epstein, "Tests for the validity of the assumption that the underlying distribution of life is exponential: Part II," Technometrics, vol. 2, no. 2, pp. 167-183, (May, 1960).

5. T. S. Ferguson, E. G. Phadia, "Bayesian nonparametric estimation based on censored data," Ann. Statist., Vol. 7, pp. 163-186 (1979).

6. E. L. Kaplan, P. Meier, "Nonparametric estimation from incomplete observations," J. Amer. Statist. Assoc., Vol. 53, pp. 457-581, (1958).

7. Way Kuo, "Systems Effectiveness Models via Renewal Theory and Bayesian Inference," Ph.D. Dissertation, Kansas State University, Manhattan, KS. 66621, 1980.

8. C. H. Lie, C.L. Hwang, F. A. Tillman, "Availability of maintained systems: a state-of-the-art survey," AIEE Trans., vol. 9, no. 3, pp. 247-259, (1977).

9. D. V. Lindley, "Bayesian statistics, a review," Society for Industrial and Applied Mathematics, Philadelphia, PA. (1972).

10. V. Susarla, J. Van Ryzin, "Nonparametric Bayesian estimation of survival curves from incomplete observations," J. Amer. Statist. Assoc., Vol. 71, pp. 897-902, (1976).

11. F. A. Tillman, C. L. Hwang, Way Kuo, Optimization of Systems Reliability, Marcel Dekker, N. Y., 1980.

12. K. Doksum, "Tailfree and neutral random probabilities and their posterior distributions," Ann. Prob., vol. 2, pp. 183-201, (1974).

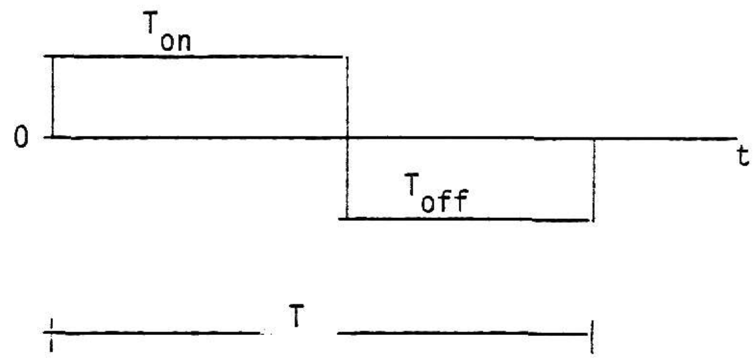

Fig. 1. A complete cycle of "on" and "off" for describing the operative characteristics of a system. T_{on}, T_{off}, and T are all random variables.

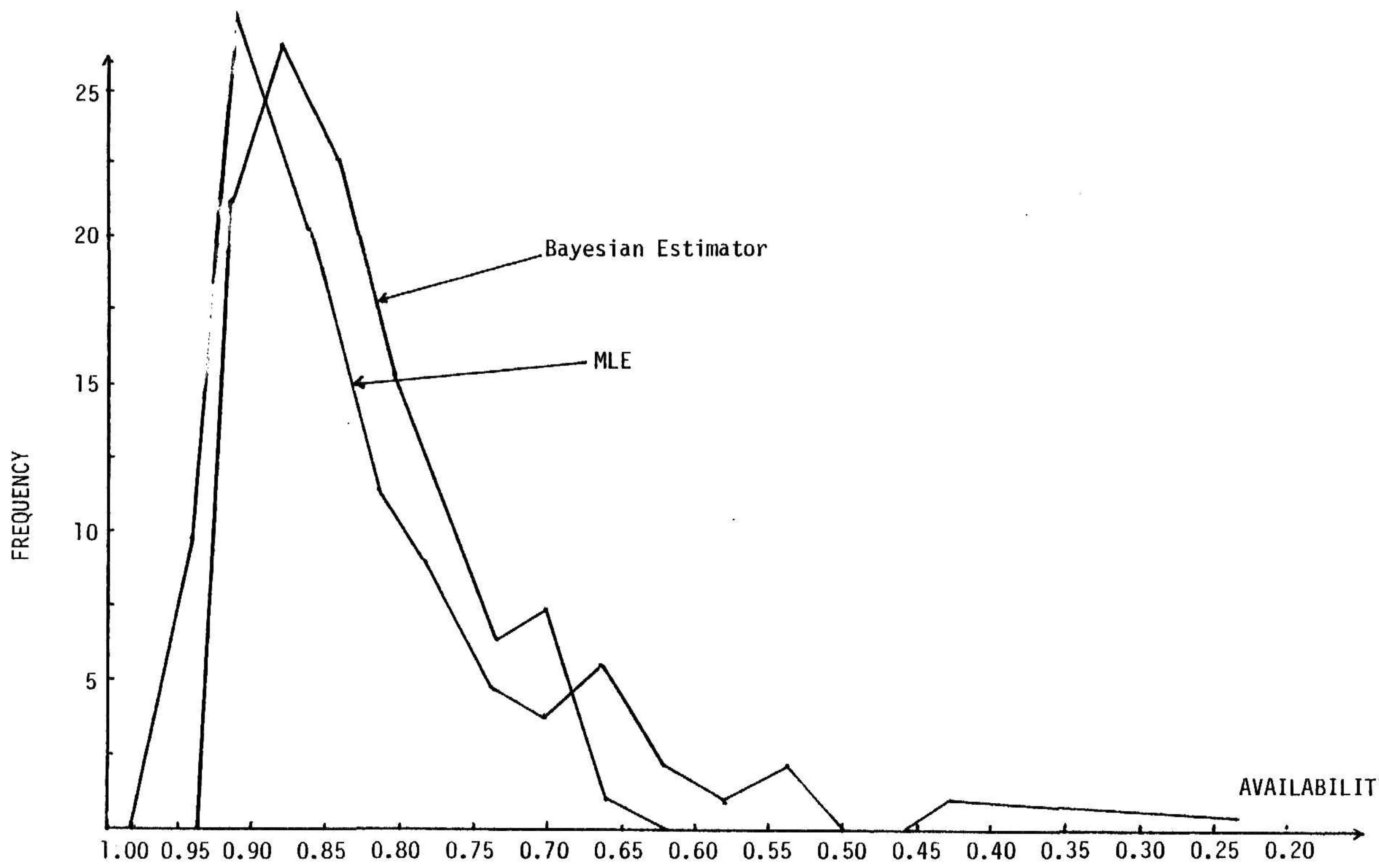

Fig. 2. Frequency distributions of the maximum likelihood and Bayesian Estimators of availability at T = 200.

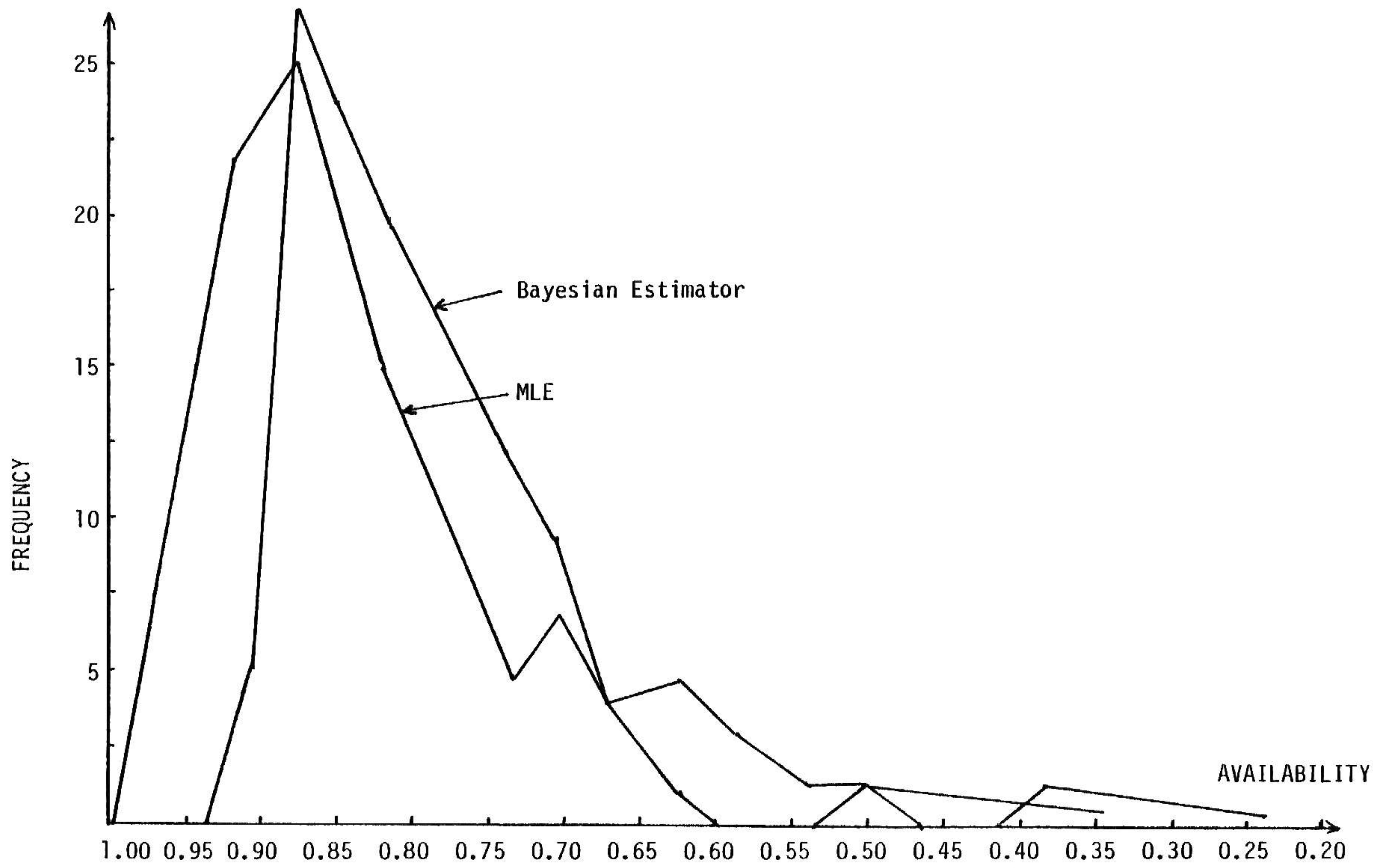

Fig. 3. Frequency distributions of the maximum likelihood and Bayesian estimators of availability at steady-state.

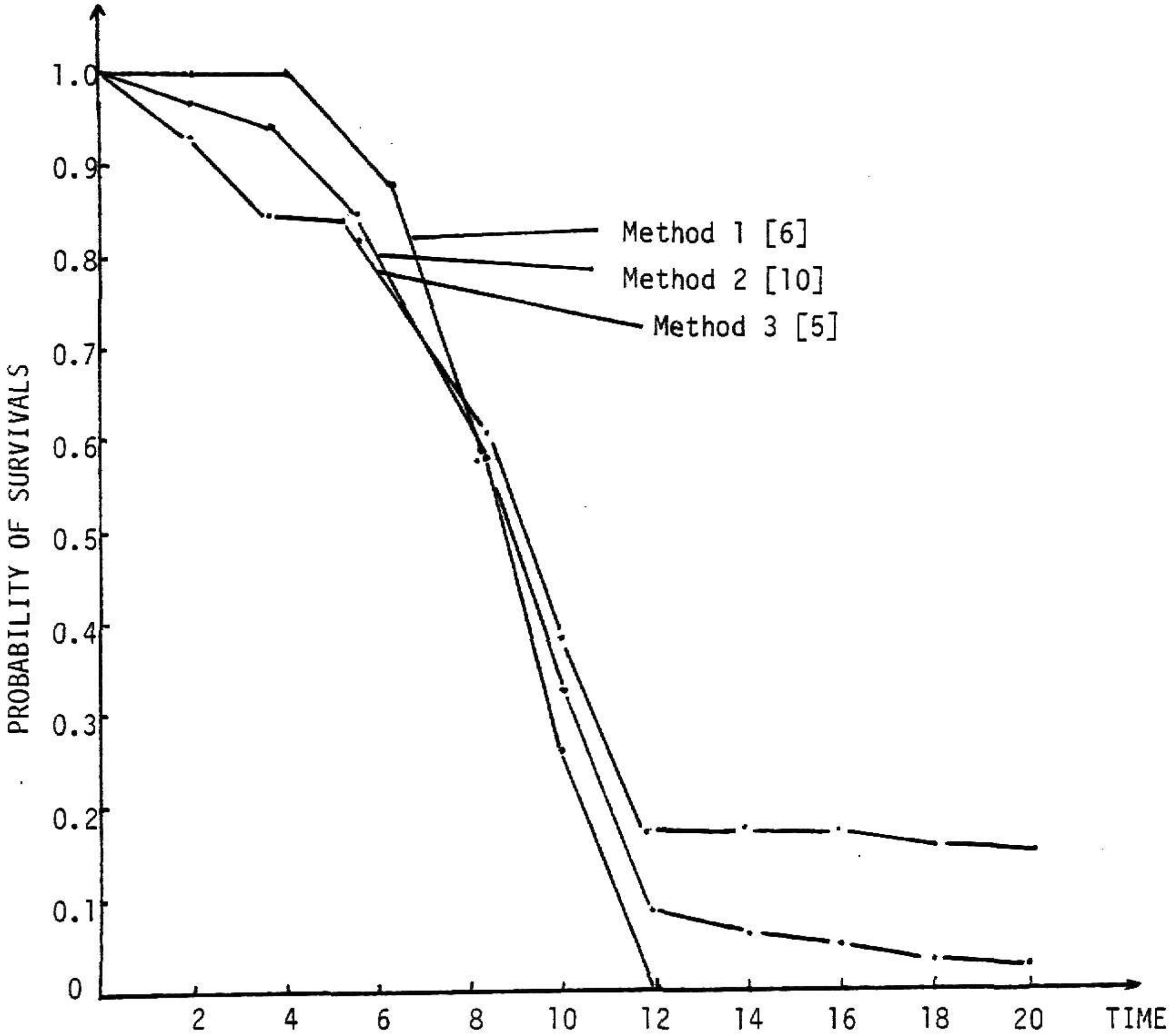

Fig. 4. Probabilities of survivals by 3 estimation techniques.

TABLE 1: 49 Pairs of Exponentially Distributed System off time and on time [4]

off time	on time	off time	on time
1.2	12	95.1	951
2.2	22	97.9	979
4.9	49	99.6	996
5.0	50	102.8	1028
6.8	68	108.5	1055
7.0	70	128.7	1227
12.1	121	133.6	1256
13.7	137	144.1	1351
15.1	151	147.6	1426
15.2	152	150.6	1491
23.9	239	151.6	1516
24.3	243	152.6	1526
25.1	251	164.2	1592
35.8	358	166.8	1668
38.9	389	178.6	1746
47.9	479	185.2	1852
48.4	484	187.1	1871
49.3	493	203.0	2031
53.2	532	204.3	2043
55.6	556	229.5	2295
62.7	627	253.1	2591
72.4	734	304.1	3041
73.6	736	341.7	3427
76.8	768	341.7	3427
83.8	858	354.4	3544

TABLE 2: A Sample of Size 5 (Set 1)

Observation	Off Time	On Time
1	144.10	2043.00
2	76.80	1746.00
3	97.90	1227.00
4	150.60	251.00
5	7.00	152.00
Mean	95.28	1083.80

TABLE 3: A Sample of Size 5 (Set 2)

Observation	Off Time	On Time
1	354.40	251.00
2	76.80	358.00
3	53.20	68.00
4	25.10	49.00
5	152.60	1028.00
Mean	132.42	350.80

TABLE 4. Various Prior Information Applied in The Simulation

prior	μ	ν	$\frac{1}{\mu}/\frac{1}{\nu}$	Remarks
1	90	1200	13.33	overestimate the failure time but underestimate the repair time
2	104.2	800	8.00	underestimate the failure time
3	80.0	1042	12.50	underestimate the repair time
4	104.2	1042	10.00	true information on the population

TABLE 5: Maximum Likelihood Estimates of Availability at T = 200

ESTIMATES ARE

0.950	0.928	0.928	0.927	0.927	0.924	0.922
0.922	0.922	0.916	0.916	0.915	0.914	0.913
0.912	0.911	0.910	0.910	0.907	0.905	0.902
0.902	0.899	0.898	0.896	0.896	0.896	0.895
0.894	0.893	0.889	0.889	0.888	0.887	0.886
0.885	0.884	0.884	0.879	0.879	0.878	0.878
0.876	0.874	0.874	0.869	0.869	0.862	0.861
0.854	0.854	0.853	0.851	0.850	0.849	0.846
0.843	0.841	0.840	0.840	0.838	0.838	0.836
0.832	0.831	0.831	0.819	0.819	0.809	0.799
0.798	0.796	0.791	0.788	0.778	0.773	0.767
0.767	0.744	0.738	0.737	0.726	0.724	0.717
0.704	0.698	0.686	0.670	0.668	0.663	0.663
0.663	0.662	0.629	0.612	0.589	0.545	0.536
0.432	0.115					

TABLE 6: Bayesian Estimates of Availability with Prior 1 at T = 200

ESTIMATES ARE

0.911	0.908	0.906	0.901	0.901	0.900	0.900
0.899	0.896	0.894	0.893	0.893	0.892	0.889
0.886	0.886	0.886	0.884	0.883	0.881	0.881
0.880	0.880	0.879	0.877	0.877	0.876	0.873
0.872	0.871	0.867	0.867	0.867	0.867	0.864
0.864	0.861	0.860	0.858	0.857	0.857	0.850
0.850	0.849	0.847	0.842	0.841	0.840	0.838
0.837	0.834	0.834	0.833	0.830	0.827	0.820
0.820	0.820	0.820	0.820	0.818	0.818	0.818
0.817	0.817	0.815	0.811	0.807	0.807	0.807
0.793	0.791	0.785	0.780	0.780	0.777	0.775
0.774	0.774	0.774	0.771	0.770	0.769	0.762
0.762	0.751	0.751	0.751	0.745	0.730	0.729
0.720	0.720	0.715	0.705	0.701	0.693	0.683
0.666	0.581					

TABLE 7: Bayesian Estimates of Availability with Prior 2 at T = 200

ESTIMATES ARE

0.909	0.909	0.906	0.903	0.903	0.897	0.897
0.897	0.896	0.896	0.894	0.894	0.890	0.889
0.886	0.886	0.886	0.886	0.886	0.886	0.886
0.886	0.882	0.879	0.879	0.879	0.876	0.875
0.872	0.871	0.867	0.863	0.861	0.858	0.855
0.854	0.854	0.850	0.849	0.848	0.848	0.845
0.845	0.845	0.842	0.838	0.835	0.831	0.829
0.827	0.827	0.827	0.824	0.821	0.821	0.818
0.815	0.811	0.809	0.807	0.806	0.806	0.800
0.797	0.796	0.795	0.795	0.794	0.792	0.790
0.788	0.761	0.761	0.754	0.745	0.745	0.743
0.743	0.736	0.736	0.733	0.725	0.723	0.721
0.721	0.718	0.680	0.680	0.679	0.678	0.664
0.664	0.658	0.657	0.653	0.651	0.651	0.639
0.631	0.452					

TABLE 8: Bayesian Estimates of Availability with Prior 3 at T = 200

ESTIMATES ARE

0.912	0.911	0.909	0.906	0.903	0.899	0.898
0.897	0.897	0.896	0.894	0.893	0.891	0.890
0.890	0.890	0.888	0.886	0.883	0.883	0.881
0.880	0.878	0.877	0.877	0.874	0.872	0.872
0.869	0.867	0.867	0.867	0.863	0.863	0.862
0.859	0.857	0.855	0.855	0.852	0.852	0.851
0.842	0.841	0.840	0.838	0.833	0.833	0.831
0.831	0.829	0.825	0.823	0.823	0.819	0.817
0.812	0.812	0.811	0.809	0.809	0.809	0.808
0.808	0.806	0.805	0.805	0.803	0.798	0.794
0.779	0.777	0.771	0.764	0.764	0.758	0.755
0.755	0.755	0.753	0.750	0.746	0.741	0.741
0.730	0.724	0.707	0.703	0.692	0.692	0.688
0.688	0.673	0.673	0.670	0.666	0.662	0.651
0.641	0.538					

TABLE 9: Bayesian Estimates of Availability with Prior 4 at T = 200

ESTIMATES ARE

0.950	0.928	0.928	0.927	0.927	0.924	0.922
0.922	0.922	0.916	0.916	0.915	0.914	0.913
0.912	0.911	0.910	0.910	0.907	0.905	0.902
0.902	0.899	0.898	0.896	0.896	0.896	0.895
0.894	0.893	0.889	0.889	0.888	0.887	0.886
0.885	0.884	0.884	0.879	0.879	0.878	0.876
0.876	0.874	0.874	0.869	0.869	0.862	0.861
0.854	0.854	0.853	0.851	0.850	0.849	0.846
0.843	0.841	0.840	0.840	0.838	0.838	0.836
0.832	0.831	0.831	0.819	0.819	0.809	0.799
0.798	0.796	0.791	0.788	0.778	0.773	0.767
0.767	0.744	0.738	0.737	0.726	0.724	0.717
0.704	0.698	0.686	0.670	0.668	0.663	0.663
0.663	0.662	0.629	0.612	0.589	0.545	0.536
0.432	0.115					

TABLE 10: Maximum Likelihood Estimates of Availability at Steady-state

ESTIMATES ARE

0.920	0.920	0.918	0.917	0.914	0.912	0.910
0.909	0.903	0.897	0.896	0.896	0.895	0.895
0.893	0.891	0.889	0.886	0.885	0.884	0.881
0.881	0.879	0.879	0.879	0.878	0.877	0.876
0.875	0.873	0.872	0.872	0.871	0.869	0.869
0.864	0.861	0.856	0.856	0.855	0.854	0.853
0.851	0.849	0.847	0.846	0.845	0.836	0.832
0.830	0.829	0.821	0.820	0.818	0.817	0.816
0.814	0.814	0.811	0.809	0.807	0.805	0.796
0.792	0.789	0.788	0.782	0.774	0.769	0.766
0.764	0.763	0.753	0.725	0.724	0.722	0.721
0.718	0.716	0.712	0.709	0.701	0.691	0.689
0.654	0.653	0.652	0.641	0.636	0.621	0.615
0.608	0.608	0.587	0.586	0.566	0.524	0.497
0.388	0.112					

TABLE 11: Bayesian Estimates of Availability with Prior 1 at Steady-state

ESTIMATES ARE

0.889	0.885	0.884	0.884	0.881	0.879	0.875
0.874	0.871	0.871	0.865	0.865	0.864	0.863
0.863	0.862	0.861	0.861	0.858	0.858	0.856
0.854	0.854	0.853	0.853	0.853	0.852	0.852
0.848	0.841	0.841	0.841	0.837	0.834	0.832
0.832	0.830	0.829	0.827	0.827	0.826	0.823
0.822	0.821	0.819	0.818	0.818	0.815	0.815
0.812	0.810	0.809	0.807	0.805	0.805	0.799
0.798	0.796	0.794	0.790	0.790	0.786	0.785
0.785	0.778	0.778	0.778	0.777	0.776	0.770
0.764	0.764	0.762	0.762	0.756	0.756	0.749
0.745	0.745	0.743	0.740	0.731	0.729	0.726
0.725	0.721	0.720	0.720	0.711	0.698	0.695
0.688	0.688	0.681	0.678	0.676	0.661	0.651
0.634	0.491					

TABLE 12: Bayesian Estimates of Availability with Prior 2 at Steady-state

ESTIMATES ARE

0.886	0.885	0.885	0.881	0.880	0.880	0.878
0.878	0.875	0.873	0.870	0.869	0.864	0.863
0.857	0.857	0.856	0.853	0.852	0.852	0.852
0.851	0.849	0.846	0.846	0.845	0.844	0.842
0.842	0.841	0.841	0.837	0.836	0.829	0.824
0.822	0.818	0.818	0.816	0.812	0.812	0.811
0.811	0.810	0.809	0.809	0.808	0.807	0.806
0.803	0.803	0.798	0.793	0.792	0.790	0.789
0.785	0.782	0.782	0.774	0.773	0.771	0.768
0.766	0.766	0.763	0.762	0.752	0.752	0.747
0.741	0.724	0.716	0.716	0.709	0.707	0.707
0.705	0.701	0.689	0.688	0.688	0.671	0.660
0.660	0.658	0.658	0.644	0.641	0.636	0.629
0.627	0.626	0.625	0.625	0.616	0.608	0.608
0.571	0.401					

TABLE 13: Bayesian Estimates of Availability with Prior 3 at Steady-state

ESTIMATES ARE

0.889	0.889	0.886	0.883	0.880	0.880	0.879
0.878	0.877	0.876	0.874	0.871	0.868	0.867
0.862	0.862	0.860	0.859	0.857	0.856	0.852
0.852	0.851	0.850	0.849	0.849	0.848	0.847
0.847	0.843	0.834	0.829	0.827	0.827	0.826
0.824	0.821	0.821	0.821	0.821	0.818	0.818
0.815	0.811	0.810	0.809	0.807	0.807	0.807
0.804	0.802	0.801	0.800	0.797	0.793	0.787
0.786	0.786	0.786	0.786	0.778	0.778	0.774
0.772	0.772	0.767	0.766	0.763	0.761	0.742
0.741	0.739	0.730	0.730	0.729	0.728	0.728
0.721	0.712	0.710	0.710	0.705	0.705	0.704
0.695	0.671	0.668	0.668	0.656	0.649	0.649
0.646	0.642	0.630	0.630	0.629	0.626	0.607
0.582	0.451					

TABLE 14: Bayesian Estimates of Availability with Prior 4 at Steady-state

ESTIMATES ARE

0.920	0.920	0.918	0.917	0.914	0.912	0.910
0.909	0.903	0.897	0.896	0.896	0.895	0.895
0.893	0.891	0.889	0.886	0.885	0.884	0.881
0.881	0.879	0.879	0.879	0.878	0.877	0.876
0.875	0.873	0.872	0.872	0.871	0.869	0.869
0.864	0.861	0.856	0.856	0.855	0.854	0.853
0.851	0.849	0.847	0.846	0.845	0.836	0.832
0.830	0.829	0.821	0.820	0.818	0.817	0.816
0.814	0.814	0.811	0.809	0.807	0.805	0.796
0.792	0.789	0.788	0.782	0.774	0.769	0.766
0.764	0.763	0.753	0.725	0.724	0.722	0.721
0.718	0.716	0.712	0.709	0.701	0.691	0.689
0.654	0.653	0.652	0.641	0.636	0.621	0.615
0.608	0.608	0.587	0.586	0.566	0.524	0.497
0.388	0.112					

TABLE 15. Means, Variances, and 90% C.I.'s for Various Availability Estimators at T = 200

	Population	Without Data				With Data				
	MLE	Prior 1	Prior 2	Prior 3	Prior 4	MLE	Prior 1	Prior 2	Prior 3	Prior 4
MEAN	0.920	0.928	0.886	0.913	0.911	0.8179	0.8241	0.8067	0.8129	0.8383
VARIANCE						0.0154	0.0041	0.0074	0.0059	0.0048
90% C.I.						0.612	0.705	0.653	0.670	0.713
						0.924	0.900	0.897	0.899	0.906
RANGE						0.115	0.581	0.452	0.538	0.546
						0.950	0.911	0.909	0.912	0.912

TABLE 16. Means, Variances, and 90% C.I.'s for Various Availability Estimators at Steady-state

	Population		Without Data			With Data				
	MLE	Prior 1	Prior 2	Prior 3	Prior 4	MLE	Prior 1	Prior 2	Prior 3	Prior 4
MEAN	0.910	0.904	0.852	0.888	0.881	0.7899	0.7957	0.7735	0.7809	0.8085
VARIANCE						0.0165	0.0048	0.0083	0.0071	0.0060
90% C.I.						0.586	0.678	0.625	0.630	0.665
						0.912	0.879	0.880	0.880	0.884
RANGE						0.112	0.491	0.401	0.451	0.458
						0.920	0.889	0.886	0.889	0.889

TABLE 17: The Probabilities of Survivals at 10 Different Time Indices by 3 Estimation Techniques

	Probabilities of Survivals		
Time	Method 1 [6]	Method 2 [10]	Method 3 [5]
2.0	1.0000	0.9835	0.9256
4.0	1.0000	0.9691	0.8482
6.0	0.8889	0.8545	0.8539
8.0	0.5556	0.5500	0.5745
10.0	0.2778	0.3071	0.3676
12.0	0.0000	0.0676	0.1574
14.0	0.0000	0.0554	0.1697
16.0	0.0000	0.0453	0.1798
18.0	0.0000	0.0371	0.1880
20.0	0.0000	0.0304	0.1947

TABLE 18: Means, Variances of the Probabilities of Survivals at 10 Different Indices by 3 Different Methods

	Probabilities of Survivals					
	Method 1 [6]		Method 2 [10]		Method 3 [5]	
Time	Mean	Variance	Mean	Variance	Mean	Variance
2	0.9944	0.0006	0.9784	0.0005	0.9155	0.0009
4	0.9203	0.0117	0.8984	0.0091	0.8462	0.0073
6	0.6641	0.0820	0.6613	0.0602	0.6563	0.0613
8	0.4908	0.0818	0.4825	0.0528	0.5094	0.0523
10	0.2152	0.0596	0.2527	0.0421	0.3074	0.0409
12	0.0961	0.0326	0.1676	0.0307	0.2297	0.0208
14	0.0525	0.0166	0.1158	0.0176	0.2019	0.0105
16	0.0337	0.0109	0.0715	0.0077	0.1988	0.0070
18	0.0187	0.0070	0.0513	0.0049	0.1949	0.0044
20	0.0000	0.0000	0.0298	0.0000	0.1891	0.0004

APPENDIX 1

A FORTRAN PROGRAM TO EVALUATE STEP 2

OF THE PROPOSED SIMULATION PROCEDURES

```
C
C
C     TRAP IS A DOUBLE PRECISION FUNCTION USING THE
C     TRAPEZOIDAL RULE TO EVALUATE THE FINITE INTEGRAL
C     OF FX.
C
C
      DOUBLE PRECISION FUNCTION TRAP(A,B,M,FX)
      DOUBLE PRECISION A,B,SUM,H,T,V,DXJ,U,Q,DYJ
      DOUBLE PRECISION FX,D
      DOUBLE PRECISION FX1
      COMMON/BL1/DXJ
      COMMON/BL3/DYJ,D,H,N
      COMMON/BL2/T,V,U
      Q=(B-A)/M
      SUM=0.0D0
      K=M-2
      DO 6 I=2,K,2
      I2=I+1
    6 SUM=SUM+0.2D01*FX(A+I*Q)+0.4D01*FX(A+I2*Q)
      SUM=SUM+0.4D01*FX(A+Q)
      TRAP=(FX(A)+FX(B)+SUM)*(Q/(0.3D01))
      RETURN
      END
C
C
C     FX IS A DOUBLE PRECISION FUNCTION SERVING
C     AS AN EXTERNAL FUNCTION OF THE MAIN PROGRAM.
C     FX DEFINES THE INTEGRAND USED IN TRAP, WHEN
C     SAMPLE DATA INFORMATION ARE NOT AVAILABLE.
C
C

      DOUBLE PRECISION FUNCTION FX(Z)
      DOUBLE PRECISION Z,T,V,DXJ,U
      COMMON/BL1/DXJ
      COMMON/BL2/T,V,U
      FX=(0.1D01)/Z*(0.1D01-DEXP(-Z*T))*DEXP(-V*Z)
      FX=FX*DXJ*DEXP(-DXJ*(U-V))
      RETURN
      END
C
C
C     FX1 IS A DOUBLE PRECISION FUNCTION SERVING
C     AS AN EXTERNAL FUNCTION OF THE MAIN PROGRAM.
C     FX1 DEFINES THE INTEGRAND USED IN TRAP, WHEN
C     SAMPLE DATA INFORMATION ARE AVAILABLE.
C
C
```

```
      DOUBLE PRECISION FUNCTION FX1(Z)
      DOUBLE PRECISION H,D
      DOUBLE PRECISION Z,T,V,DYJ,U
      COMMON/BL3/DYJ,D,H,N
      COMMON/BL2/T,V,U
      FX1=DYJ**(N+1)*(1.0D00-DEXP(-Z*T))
      FX1=FX1/Z*(Z-DYJ)**N
      FX1=FX1*DEXP(-DYJ*(D-H))*D**(N+1)
      FX1=FX1*DEXP(-Z*H/3.0D00)*H**(N+1)
      FX1=FX1*DEXP(-Z*H/3.0D00)
      FX1=FX1*DEXP(-Z*H/3.0D00)
      RETURN
      END
C
C
C     THIS PROGRAM CALCULATES THE SYSTEM AVAILABILITIES BY
C
C     1) CLASSICAL APPROACH OF MAXIMUM LIKELIHOOD ESTIMATE
C        TECHINQUE WHEN THE SYSTEM CYCLE TIME AND ON TIME
C        HAVE VALID NEGATIVE EXPONENTIAL DISTRIBUTIONS,
C     2) BAYESIAN APPROACH WHEN THE SYSTEM CYCLE TIME AND ON
C        TIME HAVE VALID NEGATIVE EXPONENTIAL DISTRIBUTIONS,
C        THE PARAMETERS OF THE SYSTEM CYCLE TIME AND ON TIME
C        ARE EXPONENTIALLY DISTRIBUTED, AND
C     3) BAYESIAN APPROACH OF CASE 2) WHEN THE SAMPLE
C        DATA INFORMATIONS ARE AVAILABLE.
C
C
C     BAYESIAN AVAILABILITY EVALUATED IN THIS PROGRAM IS
C     BASED UPON THE MEAN SQUARED ERROS LOSS FUNCTION.
C
C
C
C     NOTATIONS:
C
C     N: NUMBER OF OBSERVATION PAIRS.
C     AN: REAL VALUE OF N.
C     X(I): RANDOM VARIABLE OF SYSTEM OFF TIME,TOFF,
C           I=1,2,....,N
C     Y(I): RANDOM VARIABLE OF SYSTEM ON TIME,TON,
C           I=1,2,....,N
C     ALAMDA: PARAMETER OF DISTRIBUTION FUNCTION OF T.
C     ABETA: PARAMETER OF DISTRIBUTION FUNCTION OF TON.
C     G: AVAILABILITY FUNCTION, EXPRESSED BY
C        ALAMDA/ABETA+(1.-ALAMDA/ABETA)*EXP(-ABETA*T).
C     U: PARAMETER OF NEGATIVE EXPONENTIAL DISTRIBUTION
C        FUNCTION OF ALAMDA.
C     V: PARAMETER OF NEGATIVE EXPONENTIAL DISTRIBUTION
C        FUNCTION OF ABETA.
C     D: U+SUM OF X(I) FOR I=1,2,...,N.
```

```
C       H: V+SUM CF Y(I) FCR I=1,2,...,N.
C       ST: LOWER LIMIT CF INTEGRATICN
C           ST=1.0D-09 FOR WITH AND WITHOUT DATA INFCRMATICN
C              ANC FOR ETA.
C       ED: UPPER LIMIT CF INTEGRATICN
C           ED=1.0D-01 FOR NO DATA AND ETA
C           ED=7.0D-02 FOR NO DATA AND BETA
C           ED=2.5D-02 FOR SAMPLE SET 1,2 ANC ETA
C           IN TERMS OF BETA
C           ED=1.9D-02 FOR SAMPLE SET 1 AND PARAMETER SET 1
C           IN TERMS OF BETA
C           ED=2.0D-02 FOR SAMPLE SET 1,2 AND PARAMETER SET 1,2
C
C
C

      DOUBLE PRECISION B1,B2,B3,T,V,U,DXJ
      DOUBLE PRECISION DX,C1,C2,C12
      DOUBLE PRECISICN F1,DY,DYJ,BA,BB,BC,ED
      DOUBLE PRECISION D1,D2,D12,E12
      DOUBLE PRECISION H,D
      DOUBLE PRECISICN CC,C34,DD,D34,G12
      DOUBLE PRECISICN TRAP
      DOUBLE PRECISION Q1,FA
      DOUBLE PRECISION ST,ED
      EXTERNAL FX
      EXTERNAL FX1
    8 FORMAT (D6.2)
    9 FORMAT (I2)
   10 FORMAT (2F7.2)
   11 FORMAT (D8.2,C8.2)
   19 FORMAT ('1',///)
   21 FORMAT (//,5X,'MAXIMUM LIKELIHOOD ESTIMATE CF ETA IS',
     11X,F8.6)
   22 FORMAT (5X,'MAXIMUM LIKELIHOOD ESTIMATE CF BETA IS',
     11X,F8.6)
   23 FORMAT (//,5X,'MAXIMUM LIKELIHOCD ESTIMATE OF',1X,
     1'AVAILABILITY AT',/,5X,'TIME T OF',1X,
     2D8.2,1X,'IS',1X,F5.3)
   24 FORMAT (//)
   29 FORMAT (//,5X,'CATA SET 2')
   31 FORMAT (///,5X,'BAYESIAN AVAILABILITY ESTIMATE AT ',
     1'TIME T OF',1X,C8.2,/,5X,'WHEN THE SAMPLE ',
     2'DATA INFORMATICN ARE AVAILABLE, IS'
     3/,5X,F5.3)
   39 FORMAT (//,5X,'PARAMETER SET',I2)
   41 FORMAT (///,5X,'BAYESIAN AVAILABILITY ESTIMATE AT',1X,
     1'TIME T OF',1X,D8.2,/5X,',WHEN THE SAMPLE DATA'
     2,' INFORMATICN ARE NOT AVAILABLE,'
     3,' IS',/5X,F5.3)
```

```
   51 FORMAT (15X,I2,12X,F7.2,10X,F7.2)
   52 FORMAT (/,5X,'U= ',D8.2)
   53 FORMAT (/,5X,'V= ',D8.2)
   98 FORMAT(///,10X,'OBSERVATION',10X,'TOFF',12X,'TON',
     1///)
   99 FORMAT('1',////,10X,'TABLE     THE DATA OF OFF AND ON TIMES')
  100 FORMAT('1')
      DIMENSION X(50),Y(50),W(50)
      COMMON/BL1/DXJ
      COMMON/BL3/DYJ,D,H,N
      COMMON/BL2/T,V,U
      READ(5,8) T
      READ(5,9) N
      READ(5,10) (Y(I),X(I),I=1,N)
      WRITE(6,99)
      WRITE(6,98)
      WRITE(6,51) (I,X(I),Y(I),I=1,N)
      SX=0.0
      SY=0.0
      DO 81 I=1,N
      SX=SX+X(I)
      SY=SY+Y(I)
   81 CONTINUE
C
C
C     USE THE MAXIMUM LIKELIHOOD ESTIMATE TECHNIQUE TO
C     OBTAIN LAMDA,BETA,AND AVAILABILITY.
C
C
      AN=N
      HLAMDA=AN/SX
      HBETA=AN/SY
      H1=HLAMDA/(HBETA+HLAMDA)
      WRITE (6,19)
      WRITE (6,29)
      WRITE(6,21) HLAMDA
      WRITE(6,22) HBETA
C
C
      DO 502 KUV=1,3
      READ(5,11) U,V
      WRITE(6,52) U
      WRITE(6,53) V
C
C
      D=U+SX
      H=V+SY
      DO 501 KKY=1,2
      H2=(1.0-H1)*DEXP(-T*(HBETA+HLAMDA))
      HG=H1+H2
```

```
      WRITE(6,23) T,HG
C
      GO TO 555
C
C
C
C     USE BAYESIAN APPROACH TO EVALUATE THE AVAILABILITY
C     FUNCTION WHEN THE SAMPLE DATA INFORMATION ARE NOT
C     AVAILABLE.
C
C
      ST=1.0D-09
      ED=2.0D-01
      KK=100
      DX=(ED-ST)/KK
      KKK=KK-1
      KQ=KK-2
      C12=0.0D00
      D12=0.0D00
      DXJ=ST
      F1=TRAP(DXJ,1.0D-01,60,FX)
      DO 71 J=1,KKK,2
      DXJ=DX*J+ST
      C1=0.0D00
      IF (DXJ.GE.7.0D-02) GO TO 73
      C2=TRAP(DXJ,7.0D-02,60,FX)
   71 C12=4.0D00*(C1+C2)+C12
   73 DO 72 J=2,KQ,2
      DXJ=DX*J+ST
      D1=0.0D00
      IF (DXJ.GE.7.0D-02) GO TO 74
      D2=TRAP(DXJ,7.0D-02,60,FX)
   72 D12=2.0D00*(D1+D2)+D12
   74 E12=DX*(C12+D12+F1)/(3.0D00)
      BGN=U*V/((U+T)*(V+T))+U*V*E12
      WRITE(6,41) T,BGN
C
C
C     USE BAYESIAN APPROACH TO EVALUATE THE AVAILABILITY
C     FUNCTION WHEN THE SAMPLE DATA INFORMATION ARE AVAILABLE.
C
  555 CONTINUE
C
      FA=1.0D00
      DO 60 I=1,N
      FA=FA*I
   60 CONTINUE
      ST=1.0D-09
      ED=2.5D-02
      KM=100
```

```
      DY=(ED-ST)/KM
      KKM=KM-1
      KP=KM-2
      C34=0.0D00
      D34=0.0D00
      DYJ=ST
      Q1=TRAP(DYJ,1.8D-02,80,FX1)
      DO 61 J=1,KKM,2
      DYJ=DY*J+ST
      IF (DYJ.GE.1.8D-02) GO TO 63
      CC=TRAP(DYJ,1.8D-02,80,FX1)
   61 C34=4.0D00*CC+C34
   63 DO 62 J=2,KP,2
      DYJ=DY*J+ST
      IF (DYJ.GE.1.8D-02) GO TO 64
      DD=TRAP(DYJ,1.8D-02,80,FX1)
   62 D34=2.0D00*DD+D34
   64 G12=DY*(C34+D34+Q1)/(3.0D00)
      BA=1.0D00/FA
      BB=1.0D00/FA
      BC=(D/(T+D))**(N+1)
      BD=(H/(T+H))**(N+1)
      BG=BC*BC+BA*BB*G12
      WRITE(6,31) T,BG
  501 T=T*2
      T=T/4.0
  502 CONTINUE
      WRITE (6,19)
      WRITE(6,100)
      STOP
      END
```

APPENDIX 2

A FORTRAN PROGRAM TO EVALUATE STEP 3

OF THE PROPOSED SIMUATION PROCEDURES

N (I3)   NP (I2)

P (F 5.3)

10 F 5.3

MLE

10 F 5.3

.
.
.

10 F 5.3

Bayesian

Prior 1

10 F 5.3

.
.
.

10 F 5.3

Bayesian

Prior 2

10 F 5.3

.
.
.

Etc.

INPUT DATA FORMAT FOR APPENDIX 2

```
C
C
C PURPOSES: THIS PROGRAM FINDS THE MEANS AND THE
C           VARIANCES OF GROUPS OF BAYESIAN ESTIMATES
C           WHICH ARE OBTAINED FROM SIMULATTED
C           RESULTS.  FOR EACH PRIOR INFORMATION,
C           THE ESTIMATES ARE ALSO PUT IN DESCENDING
C           ORDER.
C DISTRIBUTION HISTOGRAMS WILL BE FOLLOWING THE
C EDITED RESULTS OF THIS PROGRAM.
C
C
C NOTATIONS:
C N: NUMBER OF OBSERVATIONS (ESTIMATES)
C NP: NUMBER OF PRIOR INFORMATION
C P: CLASS INTERVAL
C I: INDEX TO DENOTE PRIOR INFORMATION
C J: INDEX TO DENOTE ESTIMATIONS
C EM(I,J): ESTIMATE OBTAINED FROM ITH PRIOR
C          AND JTH OBSERVATION
C EMX(J): JTH ESTIMATE FOR A PRIOR
C AVGEM(I): MEAN OF ESTIMATORS FOR ITH PRIOR
C VAREM(I): VARIANCE OF ESTIMATORS FOR ITH PRIOR
C EB(K): KTH CLASS MARK
C EL(K): KTH LOWER CLASS LIMIT
C EU(K): KTH UPPER CLASS LIMIT
C C(K): NUMBER OF OBSERVATIONS IN THE KTH CLASS
C
C
C FORMAT
C
    1 FORMAT (I3,I2)
   11 FORMAT (10F5.3)
   21 FORMAT ('1',10X,'PRIOR NUMBER',I2,/,11X,
     2'MEAN IS',F7.4,/,11X,'VARIANCE IS',F7.4)
   31 FORMAT (///,8X,'ESTIMATES ARE',//,(5X,7F8.3/))
   41 FORMAT (//,9X,'FROM',10X,'TO',9X,
     2'CLASS MARK',3X,'NO OF OBSERVATIONS')
   51 FORMAT (/,3(8X,F6.3),10X,F6.2)
   61 FORMAT ('1',////)
   71 FORMAT (F5.3)
C
C MAIN PROGRAM
C
      DIMENSION EM(10,120),EMX(120),AVGEM(10),VAREM(10)
      DIMENSION C(50),EU(50),EL(50),EB(50)
      READ(5,1) N,NP
      READ (5,71) P
      DO 3 I=1,NP
```

```
      READ(5,11) (EM(I,J),J=1,N)
      DO 5 J=1,N
    5 EMX(J)=EM(I,J)
      CALL AVGV(N,EMX,AVG,VAR)
      AVGEM(I)=AVG
    3 VAREM(I)=VAR
      DO 4 I=1,NP
      WRITE (6,21) I,AVGEM(I),VAREM(I)
      DO 6 J=1,N
    6 EMX(J)=EM(I,J)
      CALL SORT(EMX,N)
      WRITE(6,31) (EMX(J),J=1,N)
      WRITE (6,61)
      WRITE (6,41)
      DO 99 KJ=1,50
      EL(KJ)=1.00-P*KJ
      EU(KJ)=1.00-P*(KJ-1)
   99 EB(KJ)=0.50*(EU(KJ)+EL(KJ))
C
C FIND THE FREQUANCY COUNTS
C
      K=1
      C(K)=0.0
      DO 7 J=1,N
      IF (EMX(J).GT.EL(K)) GO TO 8
      WRITE (6,51) EL(K),EU(K),EB(K),C(K)
      IF (EU(K).LE. 0.00) GO TO 4
   13 K=K+1
      IF (EMX(J).GT.EL(K)) GO TO 12
      C(K)=0.0
      WRITE(6,51) EL(K),EU(K),EB(K),C(K)
      GO TO 13
   12 C(K)=1.0
      GO TO 9
    8 C(K)=C(K)+1.0
    9 IF (J.EQ.N) GO TO 14
      GO TO 7
   14 WRITE (6,51) EL(K),EU(K),EB(K),C(K)
    7 CONTINUE
    4 WRITE (6,61)
      STOP
      END
C
C THIS IS THE SUBROUTINE WHICH EVALUATES THE
C MEAN AND THE VARIANCE OF A GROUP OF DATA
C

      SUBROUTINE AVGV(N,E,AVG,VAR)
      DIMENSION E(120)
      S=0.0
```

```
      DO 1 I=1,N
    1 S=S+E(I)
      AVG=S/N
      SV=0.0
      DO 2 I=1,N
    2 SV=SV+(AVG-E(I))**2
      VAR=SV/(N-1)
      RETURN
      END
C
C SUBROUTINES SORT,SORTL1, AND MAXL2 ARE THE ONES
C WHICH SORT A GROUP OF ESTIMATES IN THE DESCENDING
C ORDER
C
      SUBROUTINE SORT(X,N)
      DIMENSION X(120)
      CALL SORTL1(X,N,1)
      RETURN
      END

      SUBROUTINE SORTL1(B,M,II)
      DIMENSION B(120)
      I=II
  280 IF (I.GT.(M-1)) GO TO 200
      J=I
      K=I+1
      CALL MAXL2(B,M,J,K)
      IQ=K
      IF (IQ.GT.M) GO TO 200
      T=B(I)
      B(I)=B(J)
      B(J)=T
      I=I+1
      GO TO 280
  200 RETURN
      END

      SUBROUTINE MAXL2(C,MM,JJ,K)
      DIMENSION C(120)
      KK=K
  290 IF (KK.GT.MM) GO TO 210
      IF (C(KK).GT.C(JJ)) GO TO 130
      KK=KK+1
      GO TO 290
  130 JJ=KK
      KK=KK+1
      GO TO 290
  210 RETURN
      END
```

APPENDIX 3

A FORTRAN PROGRAM TO CALCULATE THE PROBABILITIES OF SURVIVALS BY THREE NONPARAMETRIC (BAYESIAN) LIFE ESTIMATION METHODS

```
C
C       PURPOSE:
C         THIS PRCGRAM COMPUTES ESTIMATORS CF
C         DISTRIBUTICN FUNCTICNS FCR ANY GIVEN TIME BY
C         3 NCNPARAMETRIC BAYESIAN ESTIMATICN TECHNIQUES
C
C       SUBROUTINES REQUIRED IN MAIN PRCGRAM:
C         1. METH1(NI,XI,DA,AR,DC,DC,DCT)
C         2. METH2(N,M,X,DFN,DP,H,MN,AR,DCT)
C         3. METH3(NI,XI,Y,DCT)
C         4. METH4(NI,XI,Y,CA,AR,DC,DC,DCT,K1,LN)
C         5. METH5(NI,XI,Y,ID,DHN,AR,DCH,DCH,DCT)
C         6. METH6(NI,XI,Y,Z,DR,TOW,CA,AR,DK1,DK2,DC,DC,DCT,K)
C         7. METH7(NI,XI,Y,IC,AR,CT,DCT)
C         8. SORT1(XI,NI,IST)
C         9. SDIF1(XI,Y,NI,LN,IST)
C        10. SORT2(XI,Y,NN)
C        11. SDIF2(XI,Y,Z,NN,LN)
C
C
C       NOTATIONS USED ALL THROUGH THE MAIN PROGRAM ANC SUBRCUTINES:
C
C       N: NUMBER OF DISTRIBUTION FUNCTICNS
C       M(I): SAMPLE SIZE OF THE ITH RANDCM SAMPLE,I IS A VARIABLE
C       X(I,J): THE JTH OBSERVATICN FROM THE KTH RANDCM
C       SAMPLE I=1,N, J=1,M(I)
C       NI: SAMPLE SIZE CFT8&E9C8 RA-DOM SAMPLE
C       XI(J): THE JTH CBSERVATICN FROM THE ITH RANDCM SAMPLE J=1,NI
C       Y(J): AN INTEGER VALUE INDICATES XI(J) IS A REAL
C       CBSERVATICN OR EXCLUSIVE CENSCRED DATA CR INCLUSIVE
C       CENSOREC CATA
C       AR: PRICR SAMPLE SIZE
C       C: TIME SPAN
C       LM: AN INTEGER. LM*C=UPPER BCUNC CF TIME INTERVAL
C       CA: AN EXTERNAL FUNCTION,PARAMETER CF THE DIRICHLET PROCESS
C       MN: MAX( M(I),I=1,N )
C       ID: THE ITH INDIVIDUAL CNE WANTS TO ESTIMATE
C       ITS DISTRIBUTION FUNCTION, I=1,2,...NI.
C       DR: TIME DEPENDENT SHAPE PARAMETER CF THE
C       GAMMA PROCESS AND THE SIMPLE HOMCGENEROUS PRCCESS
C       DHN: AN EXTERNAL FUNCTICN
C       TOW: SCALE PARAMETER OF THE GAMMA PRCCESS
C       AND THE SIMPLE HOMOGENERCUS PROCESS
C       DC: A GIVEN CCNSTANT USED IN CA1
C       DO: A GIVEN CCNSTANT USED IN CA1
C       DCH: A GIVEN CCNSTANT USED IN DHN1
C       DOH: A GIVEN CONSTANT USED IN DHN1
C       Z: A 2-DIMENSICNAL MATRIX
C       Z(1,I) DENOTES THE NUMBER OF REAL CBSERVATICNS AT XI(I)
```

```
C      Z(2,I) DENOTES THE NUMBER OF EXCLUSIVE
C      CENSORINGS AT XI(I)
C      Z(3,I) DENOTES THE NUMBER OF INCLUSIVE
C      OBSERVATIONS AT XI(I), WHERE XI(I) IS
C      AN ARRAY AFTER CALLING SDIF2 FOR I=1,K
C      CT: A GIVEN CONSTSNT USED IN METH7
C      DCT: C*IT, A TIME INDEX AT WHICH THE PROBABILITY
C      OF SURVIVAL IS CALCULATED
C
       IMPLICIT REAL*8(D)
       EXTERNAL CA1,CR1,CHN1
       REAL*8 XI1(10),C,AR1,AR2,TOW,XI2(10),XI3(10)
       INTEGER Y1(10),Z(3,10),Y2(10),Y3(10)
       COMMON/FIRST/DEFN
       COMMON/THIRD/DEP
       COMMON/FORTH/DES
       COMMON/SIXTH/DSG,DSH,DSD
       NI=10
       LM=20
       C=2.D0
       AR2=1.D0
       III=1
       TOW=1.D0
       DK1=0.1443D0
       DK2=0.1D0
       DC6=1.D0
       DD6=-0.1D0
       AR1=1.D0
       DC4=1.D0
       DD4=-.1D0
       DO 1 IGE=1,4
       READ(5,100) (XI1(I),I=1,5),(Y1(I),I=1,5)
   100 FORMAT(5(1X,F7.4),T46,5I1)
       READ(5,100) (XI1(I),I=6,10),(Y1(I),I=6,10)
       IK=0
       DO 210 I=1,NI
       IK=IK+Y1(I)
       XI2(I)=XI1(I)
       XI3(I)=XI1(I)
       Y2(I)=Y1(I)
       Y3(I)=Y1(I)
       IF(Y1(I) .EQ. 0) Y3(I)=2
   210 CONTINUE
       K1=IK+1
       CALL SORT2(XI1,Y1,NI)
       CALL SORT1(XI2,NI,K1)
       CALL SDIF1(XI2,Y2,NI,LN,K1)
       CALL SORT2(XI3,Y3,NI)
       CALL SDIF2(XI3,Y3,Z,NI,K)
       WRITE (6,300)
```

```
  300 FORMAT('1',////////)
      DO 10 IT=1,10
      DCT=C*IT
      CALL METH1(NI,XI1,DA1,AR2,DC6,DO6,DCT)
      CALL METH3(NI,XI1,Y1,DCT)
      CALL METH4(NI,XI2,Y2,DA1,AR1,DC4,DO4,DCT,K1,LN)
      CALL METH6(NI,XI3,Y3,Z,DR1,TCW,DA1,AR2,DK1,DK2,DC6,DO6,DCT,K)
      WRITE (6,301) DCT,DEP,DES,DSD
  301 FORMAT(16X,F4.1,3F12.4,/)
   10 CONTINUE
    1 CONTINUE
      WRITE(6,400)
  400 FORMAT('1')
      STOP
      END
C
C     FUNCTION DA1
C
C        DA1 IS A DOUBLE PRECISION FUNCTION  BEING SERVED
C     AS AN EXTERNAL FUNCTION OF THE MAIN PROGRAM.
C     DA1 IS A PARAMETER OF THE DIRICHLET PROCESS.
C

      FUNCTION DA1(DU,DC,DC)
      IMPLICIT REAL*8(D)
      DA1=DC*DEXP(DC*DU)
      RETURN
      END
C
C     FUNCTION DR1
C
C        DR1 IS A DOUBLE PRECISION FUNCTION BEING SERVED
C     AS AN EXTERNAL FUNCTION OF THE MAIN PROGRAM.
C     DR1 IS A TIME DEPENDENT AHAPE PARAMETER.
C

      FUNCTION DR1(DT,DK)
      IMPLICIT REAL*8(D)
      DR1=DK*DT
      RETURN
      END
C
C     FUNCTION DHN1
C
C        DHN1 IS A DOUBLE PRECISION FUNCTION BEING
C     SERVED AS AN EXTERNAL FUNCTION OF THE MAIN PROGRAM.
C

      FUNCTION DHN1(I,DX,DCH,DCH)
      IMPLICIT REAL*8(D)
```

```
      DHN1=DCH*DEXP(DCH*DX)
      RETURN
      END
C
C     SUBROUTINE SORT1
C       SORT THE IST'TH ELEMENT TO THE NI'TH ELEMINT OF
C     THE ARRAY XI IN THE ASENCDING ORDER.
C

      SUBROUTINE SORT1(XI,NI,IST)
      REAL*8 XI(NI),TEMP
      I=IST
  100 IF(I .GT. NI-1) GO TO 101
      IF(XI(I) .GT. XI(I+1)) GO TO 201
      GO TO 202
  201 CONTINUE
      TEMP=XI(I+1)
      J=I
  300 IF(J .LE. 0) GO TO 301
      IF(XI(J) .LE. TEMP) GO TO 301
      XI(J+1)=XI(J)
      J=J-1
      GO TO 300
  301 CONTINUE
      XI(J+1)=TEMP
  202 CONTINUE
      I=I+1
      GO TO 100
  101 CONTINUE
      RETURN
      END
C
C
C     SUBROUTINE SDIF1
C
C        SORT THE IST'TH ELEMENT TO THE NI'TH ELEMENT OF THE
C     ASCENDING ORDER ARRAY XI INTO DISTINCT OBSERVATIONS.
C     THE TOTAL NUMBER OF DISTINCT OBSERVATIONS IS LN.
C

      SUBROUTINE SDIF1(XI,Y,NI,LN,IST)
      REAL*8 XI(NI)
      INTEGER Y(NI)
      I=IST
      MM=IST
      K=IST
  100 Y(K)=1
      XI(K)=XI(MM)
  200 IF(I .GT. NI-1) GO TO 101
      I=I+1
```

```
      IF(XI(I) .NE. XI(MM)) GO TO 201
      Y(K)=Y(K)+1
      GO TO 200
  201 CONTINUE
      MM=I
      K=K+1
      GO TO 100
  101 CONTINUE
      LN=K
      RETURN
      END
C
C     SUBROUTINE SORT2
C
C       SORT THE ARRAY XI OF NN ELEMENTS IN ASCENDING ORDER
C     WITH EACH NEW XI(I), I=1,NN HAVING THE ORIGINAL
C     MATCHED Y VALUE.
C

      SUBROUTINE SORT2(XI,Y,NN)
      REAL*8 XI(NN),TEMP1
      INTEGER Y(NN),TEMP2
      I=1
  100 IF(I .GT. NN-1) GO TO 101
      IF(XI(I) .GT. XI(I+1)) GO TO 201
      GO TO 202
  201 CONTINUE
      TEMP1=XI(I+1)
      TEMP2=Y(I+1)
      J=I
  300 IF(J .LE. 0) GO TO 301
      IF(XI(J) .LE. TEMP1) GO TO 301
      XI(J+1)=XI(J)
      Y(J+1)=Y(J)
      J=J-1
      GO TO 300
  301 CONTINUE
      XI(J+1)=TEMP1
      Y(J+1)=TEMP2
  202 CONTINUE
      I=I+1
      GO TO 100
  101 CONTINUE
      RETURN
      END
C
C     SUBROUTINE DSIF2
C
C       SORT THE ASCENDING ORDER ARRAY XI INTO DISTINCT
C     OBSERVATIONS AND FIND THE NUMBER OF REAL OBSERVATIONS,
```

```
C       Z(1,I), EXCLUSIVE RIGHT CENSCREC CATA, Z(2,I),
C       AND INCLUSIVE RIGHT CENSCRED DATA, Z(3,I) CF EACH
C       DISTINCT CBSERVATICN. THE TCTAL NUMBER OF OBSERVATICNS IS LN.
C

        SUBRCUTINE SDIF2(XI,Y,Z,NN,LN)
        REAL*8 XI(NN)
        INTEGER Y(NN),Z(3,NN)
        DO 1 I1=1,3
        DO 1 J1=1,NN
      1 Z(I1,J1)=0
        K=1
        I=1
        MM=1
    100 Z(Y(MM),K)=1
        XI(K)=XI(MM)
    200 I=I+1
        IF(I .GT. NN) GC TO 101
        IF(XI(I) .NE. XI(MM)) GO TO 201
        Z(Y(I),K)=Z(Y(I),K)+1
        GO TC 200
    201 CONTINUE
        MM=I
        K=K+1
        GC TC 100
    101 CCNTINUE
        LN=K
        RETLRN
        END
C
C       SUBRCUTINE METH1
C
C       PURPCSE:
C          USE FERGUSCN'S METHODE TO COMPUTE DEFN= PRCBABILITY
C       THAT THE RANDCM VARIABLE F, DISTRIBUTED  ACCCRDING
C       TO THE DIRICHLET PROCESS WITH PARAMETER DR,
C       IS LESS THAN CR EQUAL TO DCT. THIS ESTIMATOR IS
C       EVALUATED UPCN THE MEAN SQUARED ERRCR LCSS FLNCTICN.
C         CK: NUMBER CF CBSERVATICNS LESS THAN OR EQUAL TC DCT
C         DFO: DA(DCT,DC,DC)/AR
C         DPN: AR/(AR+NI)
C         DEFN: OUTPUT PROBABILITY
C

        SUBRCUTINE METH1(NI,XI,DA,AR,DC,DC,DCT)
        IMPLICIT REAL*8(D)
        REAL*8 XI(NI),CK,AR
        COMMCN/FIRST/DEFN
        DNI=NI
        DFO=DA(DCT,DC,DC)/AR
```

```
      DPN=AR/(AR+DNI)
      CK=0.D0
      DO 100 I=1,NI
      DD=DCT-XI(I)
  100 IF(DD .GE. -1.D-10) CK=CK+1.D0
      DFN=CK/DNI
      DEFN=DPN*DF0+(1.D0-DPN)*DFN
      DEFN=1.D0-DEFN
      RETURN
      END
C
C     SUBROUTINE METH3
C
C     PURPOSE:
C        USE KAPLAN AND MEIER'S METHOD TO COMPUTE DEP=
C     PROBABILITY OF SURVIVAL UNTIL THE TIME DCT.
C     NOTATIONS:
C       DEP: OUTPUT PROBABILITY
C

      SUBROUTINE METH3(NI,XI,Y,DCT)
      IMPLICIT REAL*8(D)
      REAL*8 XI(NI)
      INTEGER Y(NI)
      COMMON/THIRD/DEP
      I=1
      DEP=1.D0
      DD=XI(I)-DCT
  100 IF(DD .GT. 1.D-10) GO TO 101
      DNII=NI-I
      IF(XI(I) .LE. DCT .AND. Y(I) .EQ. 1) DEP=DEP*DNII/(DNII+1.D0)
      I=I+1
      IF(I .GT. NI) GO TO 101
      GO TO 100
  101 CONTINUE
      RETURN
      END
C
C     SUBROUTINE METH4
C
C     PURPOSE:
C        USE SUSARLA AND VAN RYZIN'S METHOD TO COMPUTE
C     DES= PROBABILITY THAT THE RANDOM VARIABLE F, DISTRIBUTED
C     ACCORDING TO THE DIRICHLET PROCESS WITH PARAMETER DA,
C     IS GREATER THAN OR EQUAL TO DCT.
C     THIS ESTIMATOR IS EVALUATED UPON THE MEAN SQUARED
C     ERROR LOSS FUNCTION.
C     NOTATIONS:
C       K: NUMBER OF REAL OBSERVATIONS
C       NI: NUMBER OF OBSERVATIONS GREATER THAN DCT
```

```
C         NTE: NUMBER CF CBSERVATIONS GREATER THAN CR EQUAL TC DCT
C          K2: THE LARGEST DISTINCT NUMBER WITH INCREASING
C       CRDER CBSERVATICNS AMONG THE EXCLUSIVE RIGHT CENSCRED DATA
C       WHICH IS LESS THAN CR EQUAL TC DCT.
C         K1: K+1
C         NXZ: NUMBER CF CBSERVATICNS GREATER THAN OR
C       EQUAL TO X(J) J=K1,K2
C         DES: CUTPUT PROBABILITY
C

        SUBROUTINE METH4(NI,XI,Y,DA,AR,DC,DC,DCT,K1,LN)
        IMPLICIT REAL*8(D)
        REAL*8 XI(NI),AR,NT
        INTEGER Y(NI)
        CCMMCN/FORTH/DES
        K=K1-1
        NT=0.D0
        DO 100 I=1,NI
        DD1=XI(I)-DCT
        IF(DD1 .GT. 1.D-10) NT=NT+1.D0
  100   CONTINUE
        K2=K
  200   DD2=XI(K2+1)-DCT
        IF(DD2 .GT. 1.D-10) GO TC 201
        K2=K2+1
        IF(K2 .EQ. LN) GC TO 201
        GO TO 200
  201   CONTINUE
        DES=1.DC
        IF(K2 .LT. K1) GC TO 300
        DO 400 J=K1,K2
        NXJ=0
        DO 401 L=1,LN
  401   IF(XI(L) .GE. XI(J)) NXJ=NXJ+Y(L)
        DES=DES*((DA(XI(J),DC,DC)+NXJ)/(DA(XI(J),DC,DC)+NXJ-Y(J)))
  400   CONTINUE
  300   CONTINUE
        DES=DES*((DA(DCT,DC,DC)+NT)/(AR+NI))
        RETURN
        END
C
C       SUBRCUTINE METH6
C
C       PURPOSE:
C         USE FERGUSCN AND PHADIA'S METHCD TC COMPUTE
C       (1).DSG = PRCBABILITY THAT THE RANDCM VARIABLE F,
C       DISTRIBUTED ACCCRDING TO THE SIMPLE HOMCGENERCUS PROCESS
C       WITH PARAMETER DR AND TOW, IS GREATER THAN OR EQUAL TC DCT.
C       (2).DSH = PRCBABILITY THAT THE RANDCM VARIABLE F,
C       DISTRIBUTED ACCCRDING TO THE SIMPLE HOMOGENERCUS PROCESS
```

```
C      WITH PARAMETER DR AND TOW, IS GREATER THAN OR EQUAL TO DCT.
C      (3).DSD = PROBABILITY THAT THE RANDOM VARIABLE F,
C      DISTRIBUTED ACCORDING TO THE DIRICHLET PROCESS WITH
C      PARAMETER DA, IS GREATER THAN OR EQUAL TO DCT.
C      NOTATIONS:
C         JCT: THE LARGEST DISTINCT NUMBER WITH INCREASING
C      ORDER OBSERVATIONS AMONG THE ARRAY XI WHICH IS
C      LESS THAN OR EQUAL TO DCT.
C        K: NUMBER OF DISTINCT OBSERVATIONS
C        DHJ: DH(JCT,X,K,NI)
C        DHJT: DHJ+TOW
C        DHI1: DH(I-1,X,K,NI)
C        DHI: DH(I,X,K,NI)
C        DHI1T: DHI1+TOW
C        DHIT: DHI+TOW
C        DSG: OUTPUT PROBABILITY
C        DSH: OUTPUT PROBABILITY
C        DSD: OUTPUT PROBABILITY
C
C      SUBROUTINE AND FUNCTION SUBPROGRAMS REQUIRED
C        FUNCTION DH
C        FUNCTION DFG
C

       SUBROUTINE METH6(NI,XI,Y,Z,DR,TOW,DA,AR,DK1,DK2,DC,DD,DCT,K)
       IMPLICIT REAL*8(D)
       EXTERNAL DA,DR
       REAL*8 XI(NI),AR,TOW
       INTEGER Y(NI),Z(3,NI),Z1
       COMMON/SIXTH/DSG,DSH,DSD
       J=1
   100 DD=XI(J)-DCT
       IF(DD .GT. 1.D-10) GO TO 101
       J=J+1
       IF(J .GT. NI) GO TO 101
       GO TO 100
   101 JCT=J-1
       DHJ=DH(JCT,Z,K,NI)
       DHJT=DHJ+TOW
       DSG=(DHJT/(DHJT+1.D0))**DR(DCT,DK1)
       DSH=DEXP(-DR(DCT,DK2)/DHJT)
       DSD=(AR-DA(DCT,DC,DD)+DHJ)/(AR+NI)
       IF(JCT .LE.1) GO TO 201
       DO 200 I=1,JCT
       DHI1=DH(I-1,Z,K,NI)
       DHI1T=DHI1+TOW
       DHI=DH(I,Z,K,NI)
       DHIT=DHI+TOW
       DHIZ2=DHIT+Z(2,I)
       Z1=Z(1,I)
```

```
      DHIZ21=DHIZ2+1.D0
      DSG=DSG*(DHI1T/(DHI1T+1.D0)*(DHIT+1.D0)/DHIT)**DR(XI(I),DK1)
     $*(DFG(DHIZ21,Z1)/DFG(DHIZ2,Z1))
      DEI=DR(XI(I),DK2)*(DHI1-DHI)/(DHI1T*DHIT)
      DSH=DSH*DEXP(DEI)*DHIZ2/(DHIZ2+Z(1,I))
      DAN=AR-DA(XI(I),DC,DC)
      DAM=AR-DA(XI(I)-1.0D-20,DC,DC)
      DSD=DSD*(DAM+DHI1)/(DAN+DHI)*(DAN+DHI+Z(2,I))/
     $(DAM+DHI+Z(1,I)+Z(2,I))
  200 CONTINUE
  201 CONTINUE
      RETURN
      END
C
C     FUNCTION DH
C
C        COMPUTE THE NUMBER OF OBSERVATIONS GREATER THAN
C     XI(II) WHERE XI IS AN ARRAY AFTER CALLING SDIF2.

      FUNCTION DH(II,Z,K,NI)
      IMPLICIT REAL*8(D)
      INTEGER Z(3,NI)
      IF(II .GE. K) GO TO 20
      H=0
      J=II+1
      DO 10 L=J,K
   10 H=H+Z(1,L)+Z(2,L)+Z(3,L)
      DH=H
      GO TO 30
   20 DH=0.D0
   30 CONTINUE
      RETURN
      END
C
C     FUNCTION DFG
C
C        EVALUATE THE EQUATION (3.12) WITH PARAMETERS A AND IB.
C

      FUNCTION DFG(A,IB)
      IMPLICIT REAL*8(D)
      REAL*8 A
      IF(IB .EQ. 0) GO TO 20
      DB=IB
      DO 10 I=1,IB
      DI=I
      DAFI=A+DI
      DFG=(-1.D0)**(DI-1.D0)*DGAMMA(DB)/DGAMMA(DI)/DGAMMA(DB-DI+1.D0)
     $*DLOG(DAFI/(DAFI-1.D0))
   10 CONTINUE
```

```
      GO TO 30
   20 CONTINUE
      DFG=1.D0
   30 CONTINUE
      RETURN
      END
```

ACKNOWLEDGMENTS

The author appreciates his major professor, Dr. Nassar, for his valuable guidance in the preparation of this work. Appreciation is also extended to Dr. Perng and Dr. Kemp for their comments on this study.

PROPERTIES OF THE MAXIMUM LIKELIHOOD AND BAYESIAN

ESTIMATORS OF AVAILABILITY

by

WAY KUO

B.S., Nuclear Engineering, National Tsing-Hua University, 1972

M.S., Nuclear Engineering, University of Cincinnati, 1975

M.S., Industrial Engineering, Kansas State University, 1977

Ph.D., Engineering, Kansas State University, 1979

AN ABSTRACT OF A MASTER'S REPORT

submitted in partial fulfillment of the

requirements for the degree

MASTER OF SCIENCE

Department of Statistics

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1981

ABSTRACT

Availability, a combined measure of reliability and maintainability, plays an important role in the system effectiveness measure. While not much has been written on availability from the statistical estimation viewpoint, it appears that availability estimation is a natural trend toward the use of comprehensive probabilistic methods for dealing with the uncertainties associated with modern engineering problems.

In this study, availability is estimated by the maximum likelihood estimate technique and Bayesian inference. Statistical properties of the maximum likelihood and Bayesian estimators such as the mean, the variance, the range, and the 90% confidence interval (C.I.) are obtained through simulation for a negative exponentially distributed system on time and off time. We conclude that 1) both the maximum likelihood estimator and Bayesian estimator of availabilities are biased, 2) the maximum likelihood estimator of availability has larger variance, wider range, and wider 90% C.I. than those of Bayesian estimators of availability, and 3) Bayesian estimators of availability is insensitive to the prior information.

Applying Bayesian inference in availability estimator has merits when a small amount of data is available and past experience is important.