NUREG/CR-1110
KSU-2662

Bayesian Analysis of Component Failure Data

Prepared by J. K. Shultis, F. A. Tillman, N. D. Eckhoff, D. Grosh

College of Engineering
Kansas State University

Prepared for
U. S. Nuclear Regulatory
Commission

Available from

GPO Sales Program
Division of Technical Information and Document Control
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

and

National Technical Information Service
Springfield, Virginia 22161

NUREG/CR-1110
KSU-2662
RG, RD

Bayesian Analysis of Component Failure Data

Manuscript Completed: September 1979
Date Published: November 1979

Prepared by
J. K. Shultis, F. A. Tillman, N. D. Eckhoff, D. Grosh

College of Engineering
Kansas State University
Manhatten, Kansas 66506

Prepared for
Probabilistic Analysis Staff
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
NRC FIN No. B5973

1426 201

ABSTRACT

This report summarizes the many investigations made on the empricial Bayesian analysis of component failure data. In this study the analysis of attribute data of the failure-on-demand type was considered for components with low failure probabilities. Major areas emphasized in the study include (i) the development of computer techniques to obtain estimates of the prior distribution from observed failure data, (ii) the use of simulation studies to investigate the inherent properties of different prior parameter estimation techniques, (iii) the computation and comparison of probability and confidence intervals for the failure probability of individual components, and (iv) the use of non-beta prior distributions such as a mixture of beta distributions or a gamma distribution.

Four methods were examined for estimating parameters of the assumed prior beta distribution from failure data: (i) matching the moments of the prior distribution to those of the data, (ii) matching the moments of the marginal distribution to those of the data, (iii) the maximum likelihood method based on the prior distribution, and (iv) the maximum likelihood method based on the marginal distribution. From the analysis of actual failure data for diesel engines and the analysis of failure data randomly generated from a known beta distribution, it was found that method (i) is computationally the simplest, almost always yields parameter estimates, gives the smallest bias and mean square error in the parameter estimates for small sample sizes, and yields estimated prior distributions which are more conservative from a safety viewpoint than those estimated by the other estimation methods. These findings are very significant for application purposes particularly since methods (ii), (iii) and (iv) are generally used for estimation. Moreover the last three methods occasionally failed to give parameter estimates or occasionally produced totally unrealistic parameter estimates for low probability failure data of small sample size (≤ 10). Method (iii) almost always failed for samples of size greater than 20, and hence is judged unsuitable for the analysis of failure data from components with low failure probabilities.

Computer programs are presented for calculation of (i) beta parameter estimates by the three viable estimation techniques, (ii) variance and covariance estimates associated with the prior parameter estimates, (iii) plots of the estimated prior distributions, (iv) plots of the posterior distributions, and (v) confidence and probability intervals for component failure probabilities.

FOREWORD

The overall purpose of this project was to apply computer techniques to investigate properties of parameter estimation methods for use with Bayesian statistical analysis of component failure data. In this final report, the results obtained from the many investigations begun under this contract are summarized. During the course of this project several major statistical analysis programs were developed, and many important discoveries were made about the characteristics of several statistical analysis procedures. The success of this project depended upon the cooperative efforts of many people. In particular the authors would like to thank W. Buranapan, R. Lakshminarayan, Way Kuo, T. Applegate, and Yang Pan who helped the authors during various phases of this work. Also special appreciation is extended to W. E. Vesely who reviewed much of the work and suggested many avenues of fruitful investigation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (CONTINUED)

LIST OF FIGURES (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF TALBES (CONTINUED)

1426 212

1. REPORT SUMMARIES

1.1 Executive Summary

In this project, statistical methods were developed to estimate the uncertainty distributions for component failure probabilities ("per demand"). In WASH-1400, a log normal distribution was used to describe the uncertainties on the component failure probabilities. The log normal was chosen because it seemed to fit adequately the sparse data. The particular log normal distribution selected for a component was based on examination of general industrial data and on judgment.

As more failure data are collected the log normal distribution may not be adequate to describe the uncertainties and variations associated with the data. Also, instead of subjectively estimating the parameters of the distribution (e.g., the spread and median for the log normal), the parameters of the distribution should be estimated using formal statistical techniques. Such formal estimation of the parameters is based soley on the data themselves and not on any subjective judgment.

In this project, a beta distribution was used to describe the uncertainties in the component failure probabilities. The beta distribution is the distribution most often used to describe the variation of a quantity which ranges from 0 to 1 (here the component failure probability). The beta distribution is flexible in that it can accommodate a great many shapes over the interval 0 to 1, some of which are roughly similar to the log normal in shape and some of which are very different.

For the beta distribution, techniques are developed to estimate the parameters of the distribution soley from the observed data of failures and successes for a set of components treated as coming from the same population. For the components in the population, it is not assumed that their failure probabilities are necessarily equal but rather that their variation is describable by the beta distribution. Because of the different distribution shapes accommodated by the beta, this assumption for the population is much less restrictive than assuming equal probabilities. (If indeed the probabilities are very nearly equal, then the beta distribution which best describes the components will be very peaked about the representative value with small spread.)

1426 214

A particular estimation technique called "method i" in the subsequent technical discussions was found to be the best technique for estimating the beta parameters. There were several evaluation criteria used for "bestness" and method i was the best in all of these criteria. This finding is significant since method i is not the usual method employed in statistical analyses to estimate the beta parameters.

Comprehensive analyses and sensitivity studies were performed to evaluate the properties of four different parameter estimation techniques and the adequacy of using the resulting beta distribution (with the estimated parameters) to describe failure probability variations. Diesel data obtained from nuclear plant Licensee Evaluation Reports (LERs) were analyzed as an example of actual collected data. Monte Carlo calculations were also performed to generate simulated data representing other possible data behaviors. All these analyses are described in detail in this report.

Finally, computer codes were produced to allow the analyst or engineer to fit his own data with the best fitting beta distribution. These distributions can then be used in the same manner as the log normal distributions were used in WASH-1400--to determine the uncertainties in the system and accident probabilities from the uncertainties in component failure probabilities. The computer programs are documented in the Appendices to this report.

1.2 Technical Summary

This report is a summary of investigations into methods for the Bayesian analysis of failure-on-demand attribute data. Of particular interest was the analysis of components with low failure probabilities, and to illustrate the various analysis techniques, both actual failure data for emergency diesel engines at U.S. nuclear power plants and simulated failure data have been used. From this study many features of Bayesian analysis of low probability events have been determined and viable computational techniques to apply this analysis to low probability failure data have been developed.

1.2.1 Estimation Techniques for the Prior Distribution

In Section 3, four methods for estimating values of the parameters of the assumed beta prior distribution from observed failure data are reviewed. These methods are (i) matching the moments of the prior distribution to those of the failure data, (ii) matching moments of the marginal distribution to those of the data, (iii) the maximum likelihood method based on the prior distribution, and (iv) the maximum likelihood method based on the marginal distribution. In this phase of the study the following results were obtained:

- Computer codes were developed to estimate the beta prior parameters by each of four estimation techniques.
- Estimation of the variance of the parameter estimators were performed for methods (i) and (iv). For method (i) a first order Taylor's series expansion technique was used to obtain variance estimates of the beta parameters from the variances of the data moments. In method (iv) both an exact and an approximate method for values of the lower bound of the variances and covariance were used (based on the Cramer-Rao-Frechet inequality for the covariance matrix). The approximate method was found to give nearly identical results compared to those of the exact method.
- The prior matching moments technique (method (i)) was the only method which yields closed-form results for the parameter estimates. Further, the estimators were shown to be positive for very mild restrictions on the failure data.
- The prior maximum likelihood method (method (iii)) was shown to be infeasible for any failure data sample for which zero failures were observed for any component.
- For certain groupings of the diesel engine failure data, both marginal-based estimation methods (methods (ii) and (iv)) were observed to yield no numerical solutions.

- The observed diesel engine failure data were grouped by manufacturer and by number of starts and beta prior estimators were obtained for each grouping. For the results obtained with the prior matching moments method, only a few significant differences at the 0.05 level were found.
- Methods were developed for placing error bands on both the estimated prior density and prior cummulative distributions. These methods, which require variance and covariance estimates of the beta parameter estimators, were applied to estimated prior distributions for the diesel engine data.

Based on the diesel data analyzed, the prior matching moments technique (method i) appeared to be the best of the four methods for estimating the beta parameters from the data. The techniques for estimating variances and error spreads also seemed to be suitable for practical applications. The diesel data themselves did not show any strong clustering into distinct groups when analyzed by the various Bayesian approaches.

1.2.2 Characteristics of the Estimated Beta Prior

To determine how well the four estimation techniques for the prior parameters are able to predict the beta prior distribution, all four methods were used to analyze many samples of simulated failure data which were generated from a known beta-binomial (marginal) distribution. In this way, properties of the sampling distribution of the estimators and distributions of other related statistics were obtained. Important results from this phase of the study include:

- Only the prior matching moments estimation technique (method i) always yielded realistic prior parameter estimators for all 6500 simulated data samples of various sizes.
- Both marginal-based estimation techniques (methods ii and iv) would occasionally fail to yield parameter estimates or yield outlier estimates which were much too large in size. This deficiency was more severe for data generated from a beta prior skewed towards low failure probabilities than for data generated from a symmetric beta.
- The distributions of the prior parameters estimators for all four estimation techniques were found to have positive bias for small sample sizes ($N \leq 20$) which decreased in magnitude as the sample size increased. The prior matching moments estimators had smaller bias for all sample sizes, while the estimators from the two marginal-based techniques had the largest bias.
- The mean squared error and variance of the estimators for all four methods decrease as the sample size increases. The estimators obtained

from the prior matching moment methods have the smallest variance while the marginal-based methods produce estimators with the largest variances for samples of sizes N≤50.

- For small sample sizes (N≤10) the median of the prior parameter estimators from the matching moments method is nearest to the true values. However for larger sample sizes (N≥50) the median appears to underestimate the true values while the medians from both marginal-based methods approach the correct parameter values.

- There is a large correlation between the beta parameter estimates.

- The distribution of the estimated prior mean and variance was obtained from the parameter estimators. The distribution of the prior mean estimators was found to be nearly identical for the three estimation techniques considered (prior matching moments and the two marginal-based methods). No outliers were observed in the distribution of means since even the outlier estimates of the beta parameters yielded good values of the mean. However the large outlier parameter estimates (obtained only with the marginal-based methods) yielded prior variance estimates which were far too small.

- From the estimated prior distributions, the distribution of the estimated 95-th percentiles (i.e., the failure probability for which 95% of the area of the failure distribution falls below) was examined. The prior matching moments method appears to be slightly more conservative from a safety viewpoint since slightly higher values of the 95-th percentiles are obtained with this method than with the marginal-based techniques. Further, the marginal-based methods yielded several 95-th percentile estimates which were much too small, a result of the outliers obtained for the prior parameter estimators.

- The distribution of the fraction of the estimated prior distribution greater than the true 95-th percentile was also investigated. Again the prior matching moments method gave slightly more conservative results since the mean of these distributions were always slightly greater than the true value of 0.05, while the mean of the distributions produced by the marginal-based techniques were observed to oscillate around the true value. The variances of these distributions generated by the different estimation technique were nearly equal and they decreased as the sample size increased.

- The variance and covariance lower bounds for the parameter estimates determined with the marginal maximum likelihood method were compared to the variances of the parameter estimator distributions. The prior matching moments method (which produced no outliers and hence had the smallest variances) came closest to these lower bounds and for large sample sizes (N≥50) actually were smaller. The estimator variances from the marginal-based methods were more than 50 to 100% higher than the lower bounds even for sample sizes as large as 50.

- Bias removal schemes for the beta parameter estimators were briefly examined for the prior matching moments method. The bias was seen to decrease inversely to the sample size; however, no completely satisfactory empirical bias removing formula was found.

- The distribution of the beta parameter estimators as determined by the prior matching moments method was found to be described well by a shifted log normal distribution.

Thus based on these additional simulation studies, the prior matching moments technique (method i) was again the best method for estimating the beta parameters from the failure data. The parameters estimated by this method generally had the smallest bias and the smallest mean square error. Moreover, this simple prior matching moments technique always yielded realistic parameter estimates (unlike the other three estimation techniques examined) and consequently is well-suited for practical applications.

1.2.3 Probability Intervals for the Estimated Failure Probability

The calculation of both the classical confidence interval and the Bayesian probability interval for the estimated failure probability of an individual component with a given failure history was described by the equation involving the incomplete beta function. It was shown that the solution for the intervals could be expressed in terms of the Snedecor F-distribution. Also an approximate solution in terms of the χ^2 distribution was derived. For the special case of no failures observed for the component, explicit closed form results were obtained for the interval. Finally an algorithm to obtain a numerical solution for the probability limits was developed. Several numerical examples for low failure probability components are presented.

With these techniques, the analyst or engineer can thus calculate the uncertainty interval on the component failure probability by either Bayesian or classical techniques.

1.2.4 Extended Beta Priors

Two methods were briefly examined for describing the Bayesian prior distribution when this distribution was not a member of the beta family.

- For the case in which data are generated from a mixture of different beta distributions, the resulting overall prior (a weighted sum of betas) is itself not in the beta family. Methods are described whereby this overall prior may be approximated by a single beta. Numerical examples are given, and a method for constructing the weighting fractions is developed.

- It was shown that for low failure probability components, the binomial conditional distribution could be approximated by a Poisson distribution, further, the beta prior distribution was shown to be described approximately by a gamma distribution.
- For the diesel engine failure data, both the approximate gamma model and the beta distribution gave nearly identical results for the prior distribution.
- Both the binomial-beta model and the gamma-Poisson approximate model were found to give very similar results for the mean and variance of the posterior distribution for each diesel engine.

Based on these findings, the analyst confronted with a reliable component can thus treat its failure occurrences as being Poisson with the Poisson parameter having a gamma distribution to describe the uncertainty and parameter variations. This treatment, which is often simpler to apply, will give results which are essentially the same as the exact binomial-beta approach.

1.2.5 Computer Code Development

A major aspect of this study was the development of computer codes to perform many of the analyses described above. Although many programs were written in the course of this study, two were thought to be of general interest and are included in the Appendices of this report.

- BETA III calculates estimates of the beta prior parameters by all four estimation techniques as well as variance estimates of the parameters for methods (i) and (iv). Options are available to give plots of the estimated beta prior density and cummulative distributions.
- TAILS calculates both the classical confidence interval and the Bayesian probability interval for the failure probability of a component with a given failure history.

These codes give the analyst or engineer the capability to analyze data of failures and successes of a set of components which are assessed to be similar but not necessarily having exactly the same failure probabilities. The codes will estimate the parameters of the beta distribution describing the variation of the component failure probabilities. This distribution can then be used in subsequent reliability and risk analyses.

2. INTRODUCTION

Of considerable importance in the reliability analysis of nuclear power plants is a description of the distribution of failure probabilities for plant components, e.g., standby diesel generators. The performance data for a particular component, e.g., k failures in n start-ups, may be so sparse or may vary so much among "similar" components that classical estimates of the failure probability (k/n) may be deemed of little use. The classical estimates k/n are particularily noninformative when the component has never been observed to fail (k=0). In an effort to obtain a more meaningful description of the failure probability of such a component, additional external information is often inserted into a probability model for the component. For example, use of failure data from similar components and/or an engineer's judgemental estimates of the component's reliability can be incorporated with the actual performance data of a particular component to yield a better probability model for that component. The components which are judged to be similar do not all have to have exactly the same failure probabilities; it is only assumed that they are described by the same distribution. The insertion of extraneous information is the cornerstone of the Bayesian method [1] which over the past few years has been increasingly used in the description of components with low failure probabilities.

2.1 Bayesian Statistical Description of Failure-on-Demand Data

For any particular component in a power plant, e.g., a standby diesel generator, the probability of failure, p, is often assumed to be constant and not to vary among similar components. Under the assumption that p is constant, the probability of obtaining k failures in n tests, e.g., k nonstarts in n tries to start the standby diesel generator, is given by the binomial distribution,*

$$f(k|n,p) = \binom{n}{k} p^k (1-p)^{n-k} \quad . \qquad (2.1)$$

---

* In this report, a bar is used to separate the random variables from the constants, i.e., f(k|n,p) denotes k is a random variable and n and p are constants.

For a power plant component, the failure probability, p, is sometimes better modeled as being a random variable which will vary both with experience, e.g., learning to operate the generator better, and with the plant, e.g., different plant conditions may cause variation in the failure probability. In these cases when sampling similar components from different plants, a distribution of failure probabilities is more realistic a model than assuming all failure probabilities to be equal. The distribution for the failure probability between similar components is termed the prior distribution. Because of its ability to model a variety of different distributional shapes and because of the ease with which it is incorporated into the mathematical description, the beta distribution is usually used as the prior distribution to describe the variation in the failure probability [3]. The beta distribution (density function) for p, $g(p|a,b)$, is given by

$$g(p|a,b) = \frac{p^{a-1}\ (1-p)^{b-1}}{B(a,b)} \quad , \quad (a,b > 0) \ , \tag{2.2}$$

where

$$B(a,b) \equiv \int_0^1 x^{a-1}\ (1-x)^{b-1}\ dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \tag{2.3}$$

and Γ is the gamma function. The mean, μ, and variance, σ^2, are given by [2]

$$\mu = \frac{a}{a+b} \ , \tag{2.4}$$

and

$$\sigma^2 = \frac{ab}{[(a+b)^2(a+b+1)]} \ . \tag{2.5}$$

As previously stated, the beta distribution of Eq. (2.2) is often used because (i) the range of a and b describe a wide variety of distribution shapes with support on (0,1), and (ii) it is combined analytically with the binomial distribution with ease. The values of a and b which determine the explicit distribution of p must be subjectively assumed or can be estimated from experimental data, i.e., from records of failures and successes. Methods for the estimation of a and b are presented in the next section.

When p is treated as a random variable, the probability of exactly k failures in n tries, $h(k|n,a,b)$, is obtained by integrating the binomial distribution Eq. (2.1) over all p weighted with the beta distribution,

$$h(k|n,a,b) = \int_0^1 f(k|n,p)\ g(p|a,b)\ dp$$

$$= \binom{n}{k} \frac{B(a+k, b+n-k)}{B(a,b)} . \qquad (2.6)$$

The distribution $h(k|n,a,b)$ is termed the marginal distribution since all possible values of p are considered. This particular marginal distribution is called the "beta-binomial" or "hyperbinomial" and is encountered frequently in Bayesian statistics [3]. The expectation and variance of k described by the above marginal distribution are found to be

$$E(k|n,a,b) = \frac{a}{a+b}\, n, \qquad (2.7)$$

$$Var(k|n,a,b) = \frac{nab(a+b+n)}{(a+b)^2(a+b+1)} . \qquad (2.8)$$

The prior distribution, which in this study is assumed to belong to the beta family, describes the distribution of the failure probability among all components judged to be similar. The prior distribution is based on past experience and information. If a particular component is observed to fail k times in n demands, this additional (new) information can be used to revise the distribution for the possible values of p for the component. This updated distribution is called the *posterior* distribution and depends upon the original assessment of the distribution for p (the prior distribution) and the observed k failures in n demands (the new information). From Bayes' theorem one can calculate this posterior distribution, $\xi(p|k,n,a,b)$, for a component which has experienced k failures in n tries and which is assumed to belong to a class of components whose failure probabilities are distributed according to the prior distribution. Explicitly Bayes' theorem can be stated as

$$\xi(p|k,n,a,b) = \frac{f(k|n,p)g(p|a,b)}{h(k|n,a,b)} ,$$

which upon substitution of Eqs. (2.1), (2.2), and (2.6) yields the posterior distribution

$$\xi(p|k,n,a,b) = \frac{p^{a+k-1}\,(1-p)^{b+n-k-1}}{B(a+k, b+n-k)} . \qquad (2.9)$$

This posterior distribution of p for a particular component is also a beta distribution but with larger parameters, a+k and b+n-k. The larger parameters generally produce a smaller variance (see Eq. (2.5)) which corresponds to more knowledge or less uncertainty about p. This result is intuitively reasonable since the description of p is based on both prior intuition (Eq. (2.2)) as well as actual experimental knowledge. Consequently, one would expect a higher degree of certainty (about p) for this case than a case in which only prior intuition or actual experimental knowledge is used.

The posterior distribution can be used to obtain representative values for the failure probability of a particular component. For example, the posterior mean value for p, $\hat{p}_B$, is from Eq. (2.4)

$$E(p|k,n,a,b) \equiv \hat{p}_B = \frac{a+k}{a+b+n} \quad . \tag{2.10}$$

By contrast, the classical estimator of the failure probability for a particular component is

$$\hat{p}_c = \frac{k}{n} \; . \tag{2.11}$$

For many components the failure probability is intentionally designed to be very small, and in a relatively small number of tests, e.g., attempts to start a standby diesel generator, often zero failures will be observed. From these data, classical statistics would yield an estimate of the failure probability of the component to be zero, which is unrealistic. Bayesian statistics, however, which uses prior information based upon experience or information from similar components will give a nonzero value for the expected failure probability. Furthermore, the Bayesian approach gives a complete distribution $\xi(p|k,n,a,b)$ for the possible values of the failure probability for a particular component and not just one "best" estimate. In the Bayesian framework, the posterior distribution represents the complete knowledge of the uncertainty of the failure probability for a component.

2.2 Scope of Study

In this report the results of a study are reported on various techniques and applications of the preceding Bayesian analysis to describe the failure of components with expected low failure probabilities. A major portion of this study deals with methods to estimate values of the parameters of the beta prior distribution. Sometimes the particular prior distribution for a particular application is deduced from expert judgment; however in this study four techniques for estimating the prior parameter based upon only observed failure data are investigated. Such techniques which use only observed historical data are commonly referred to a "empirical" Bayes methods since the prior parameters are empirically deduced from the data. These techniques were then used to analyze failure data obtained from standby diesel engines at many U.S. nuclear power plants. Methods were also investigated to obtain estimates of the variance and covariance associated with the beta prior parameters. With these variance estimates, techniques were developed for obtaining confidence bands around the prior distributions to account for the fact that the beta parameters were estimated from data.

Also considered in this study was an evaluation of which of the parameter estimation procedures is "best" for use with low failure event situations. Through a simulation study, the biasedness and mean error of each estimation technique are evaluated. Further the effect of sample size is examined - an effect of considerable importance for situations characterized by a paucity of historical failure data.

Methods are also presented whereby both the classical confidence intervals and Bayesian probability intervals for the failure probability of a particular component can be evaluated. Of considerable importance in this stage of the study were the development of accurate numerical techniques to evaluate these intervals as well as the development of approximate methods.

In Section 6, brief investigations are presented of the effect of mixing two distributions and using a single prior distribution to model the mixed distribution. An alternative description of the failure-on-demand problem is also presented by using a Poisson conditional and its natural conjugate, the gamma distribution, as the prior distribution.

In the appendices of this report, two of the major computer programs developed in this study are described. These programs can be used to evaluate the beta prior parameters from historical failure data, plot estimated prior cumulative and probability distribution functions, and calculate probability and confidence intervals for the failure probability.

3. EMPIRICAL METHODS FOR ESTIMATING THE PRIOR DISTRIBUTION

To use the Bayesian approach, the prior distribution, $g(p)$, of Eq. (2.2) must first be obtained. Generally this is done by (i) subjective assessment, (ii) past experience, or (iii) from a fit to experimental data from similar components. For any particular component, given only its number of failures-on-demand and total number of demands, there are insufficient data to estimate a and b. However, if several independent sets of data, i.e., failure records for several components, are assumed belong to the same population and consequently to be described by same prior probability distribution, this observed data can be used to obtain estimates of the parameters of the prior distribution.* In this chapter four methods for obtaining estimates of the beta prior distribution from failure data are discussed and applied to the analysis of diesel engine data.

3.1 Method of Matching Moments of Prior to Data

Although there is no unique method to estimate the parameters of the prior distribution from the failure records, one method of estimation is to equate the mean (the first moment) and the variance (the second moment minus the square of the first moment) of the failure probability estimates to the corresponding expressions for the prior model involving the distribution parameters. In effect, these parameters are estimated by "matching moments" of the data to those of the prior model. If there are k_i failures out of n_i tries for the i-th component of a random sample of size N, an estimate of the failure probability, $\hat{p}_i$, for each sample is k_i/n_i, and thus the observed mean and variance of the $\hat{p}_i$ estimates are

$$\hat{\mu}_{ob} = \frac{1}{N} \sum_{i=1}^{N} \frac{k_i}{n_i} \tag{3.1}$$

and

$$\hat{\sigma}^2_{ob} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{k_i}{n_i} - \hat{\mu}_{ob}\right)^2 . \tag{3.2}$$

*It is interesting to compare this assumption with the usual classical analysis. In the classical analysis, the failure probabilities of similar components are assumed to be equal. Here, we allow the probabilities to vary and only assume the variation is describable by a general beta distribution whose parameters are to be determined.

where N is the total number of components in the same population for which failure data are available. By matching these sample moments, which use only the observed data, to the expressions of the mean and variance of the assumed beta prior distribution (Eqs. 2.4) and (2.5)), a relationship between the parameters of the distributions, a and b, and the observed data can be obtained, namely

$$\hat{p}_{ob} = \mu \equiv \frac{a}{a+b} \tag{3.3}$$

and

$$\hat{\sigma}^2_{ob} \simeq \sigma^2 \equiv \frac{ab}{(a+b)^2\ (a+b+1)} \quad . \tag{3.4}$$

These equations can be solved for a and b in terms of $\hat{p}_{ob}$ and $\hat{\sigma}^2_{ob}$ to give

$$a = \frac{\hat{p}^2_{ob}}{\hat{\sigma}^2_{ob}}\ (1-\hat{p}_{ob}) - \hat{p}_{ob} \tag{3.5}$$

and

$$b = \frac{\hat{p}_{ob}}{\hat{\sigma}^2_{ob}}\ (1-\hat{p}_{ob})^2 + \hat{p}_{ob} -1 \ . \tag{3.6}$$

One of the major advantages of this method is its simplicity and the existence of a closed-form solution for the parameter estimates (Eqs. (3.5) and (3.6)). However, these solutions for the parameter estimates do not necessarily yield positive values as is required for the beta parameters. For example the use of failure data $\{k_i, n_i\}$ = (1,100), (1,50), (99,100), (49,50) in Eq. (3.5) yields a negative value for a. Nevertheless, for low failure probability data, this estimation method generally gives positive and hence realistic values for the parameter estimates. To see this, rewrite Eq. (3.5) for a as

$$a = \frac{\hat{p}_{ob}}{\hat{\sigma}^2_{ob}}\ \{\hat{p}_{ob} - \hat{p}^2_{ob} - \hat{\sigma}^2_{ob}\},$$

which upon substitution for $\hat{p}_{ob}$ and $\hat{\sigma}^2_{ob}$ (which are always non-negative) from Eqs. (3.1) and (3.2) yields

$$a = \frac{\hat{p}_{ob}}{\hat{\sigma}^2_{ob}} \left\{\frac{1}{N} \sum_i \hat{p}_i - \frac{1}{N-1} \sum_i \hat{p}_i^2 + \frac{1}{(N-1)N^2} (\sum_i \hat{p}_i)^2\right\}$$

$$\geq \frac{1}{N} \frac{\hat{p}_{ob}}{\hat{\sigma}^2_{ob}} \left\{\sum_i \hat{p}_i \left(1 - \frac{N}{N-1} \hat{p}_i\right)\right\} \quad .$$

If the expression for the sample variance (Eq. (3.2)) had been divided by N rather than (N-1), the N-1 factor in the above inequality would have been replaced by N, and since $0 \leq \hat{p}_i \leq 1$, the right hand side of this inequality would then be ≥ 0. However if we require the $\hat{p}_i$ to be limited to a slightly more restrictive range, $0 \leq \hat{p}_i \leq (N-1)/N$, the above expression yields,

$$a \geq \frac{1}{N} \frac{\hat{p}_{ob}}{\hat{\sigma}^2_{ob}} \{\sum \hat{p}_i(1-\hat{p}_i)\} > 0 \quad . \tag{3.7}$$

For sufficiently large N or for small to moderate $\hat{p}_i$ values, this additional restriction on the $\hat{p}_i$ values is inconsequential. Even for the most restrictive case (N=2), positive estimates of a are always obtained if $0 \leq \hat{p}_i \leq \frac{1}{2}$ which is satisfied for low probability failure data. Finally, if the estimate for a is positive, then so must be the estimate of b since from Eq. (3.3)

$$b = a(1-\hat{p}_{ob})/\hat{p}_{ob} \;>\; 0 \;\text{ if } a > 0 \; . \tag{3.8}$$

Thus this simple prior matching moments method yields parameter estimates which are positive for the type of low probability failure data considered in this study. Although the estimation of p_i by k_i/n_i may appear to introduce a questionable approximation especially for low probability events (i.e., small p_i), it will been seen in Section 4 that this method has several additional advantages over the more complex estimation techniques also investigated in this study.

3.2 Maximum Likelihood Method Based on the Prior Distribution

The method of maximum likelihood can be used to obtain estimates of the prior parameters by constructing a likelihood function based on the prior beta distribution. Define the *likelihood function*

$$L(a,b|p_1,p_2,\ldots p_N) \equiv \prod_{i=1}^{N} g(p_i|a,b) \tag{3.9}$$

where g is the prior beta defined by Eq. (2.2). Explicitly, this likelihood function is the probability of observing $p_1,p_2,\ldots,p_N$ as values for the failure probabilities from components 1,2,...,N respectively. The values of a and b which maximize the likelihood function are called the maximum likelihood estimators, $\hat{a}$ and $\hat{b}$, i.e., the probability of obtaining the observed values is maximized. Intuitively, this choice is very appealing. The maximum likelihood approach has been shown to have many general properties and is widely used in statistical analysis [3].

For the actual failure-on-demand problem considered in this study, failure probabilities, p_i, are not observed directly, but rather must be approximated by the estimates $\hat{p}_i = k_i/n_i$. The maximum likelihood estimators of a and b are then the solutions to

$$\frac{\partial}{\partial a} \, ln L(a,b) = 0 \tag{3.10}$$

and

$$\frac{\partial}{\partial b} \, ln L(a,b) = 0 \tag{3.11}$$

Upon substitution of the explicit form of the beta function, g(p), these likelihood equations become

$$\psi(a) - \psi(a+b) - N^{-1} \sum_{i=1}^{N} ln p_i = 0 \tag{3.12}$$

$$\psi(b) - \psi(a+b) - N^{-1} \sum_{i=1}^{N} ln(1-p_i) = 0 \tag{3.13}$$

where $\psi(z) \equiv d[ln\Gamma(z)]/dz$, the digamma function. The solution to these simultaneous transcendental equations cannot be obtained analytically; however, if $\hat{a}$ and $\hat{b}$ are not too small the following approximate result may be used [3]:

$$\hat{a} \approx 1/2 \left(1 - \prod_{i=1}^{N} (1-p_i)^{1/n}\right) \left(1 - \prod_{i=1}^{N} p_i^{1/n} - \prod_{i=1}^{N} (1-p_i)^{1/n}\right)^{-1} \tag{3.14}$$

$$\hat{b} \simeq 1/2 \left[1 - \prod_{i=1}^{N} p_i^{1/n}\right]\left[1 - \prod_{i=1}^{N} p_i^{1/n} - \prod_{i=1}^{N} (1 - p_i)^{1/n}\right]^{-1} \tag{3.15}$$

This approximate solution may also be used as starting values for an iterative numerical solution of the likelihood equations.

This maximum likelihood method, while suitable for some problems, is not applicable to those situations in which some of the observed k_i are zero. In these cases the estimated failure probability p_i is also zero and the likelihood function becomes unbounded or zero depending upon the value of a. Consequently, little use was made of this estimation technique in this study which was concerned with small failure probabilities and with data for which k_i=0 is not unusual. A variation of this maximum likelihood technique based on the marginal distribution and which does not suffer from this deficiency in a zero failure case is discussed in Section 3.4.

3.3 Method of Matching Moments of the Marginal Distribution to Data Moments

An alternative to the technique of Section 3.1 is to substitute the moments of the marginal (or mixture) distribution of Eq. (2.6) for the moments of the prior distribution. Conceptually this technique is more attractive since only the proportion of failures k_i/n_i (which are observed data) are involved, whereas in matching the data to the prior moments, the failure probabilities, p_i, (which were not actually observed) had to be estimated as k_i/n_i.

For the present case, the sample sizes are of unequal sizes, i.e., different n_i, and thus a weighting scheme should be used in the estimation procedure. Define the following statistics:

$$\hat{p} = \frac{1}{w} \sum_{i=1}^{N} w_i \frac{k_i}{n_i} \tag{3.16}$$

$$S = \sum_{i=1}^{N} w_i \left(\hat{p} - \frac{k_i}{n_i}\right)^2 , \tag{3.17}$$

where

$$w = \sum_{i=1}^{N} w_i ,$$

and w_i is the weight assigned to the i-th sample. By setting the above statistics equal to their expected values (of the marginal distribution), estimates for the prior mean and variance are obtained [4]:

$$\hat{\mu} = \hat{p} \tag{3.18}$$

and

$$\hat{\sigma}^2 = \hat{\mu}(1-\hat{\mu}) \frac{S - \hat{p}\hat{q}[\sum_{i=1}^{N} \frac{w_i}{n_i}(1 - \frac{w_i}{w})]}{\hat{p}\hat{q}[\sum_{i=1}^{N} w_i(1 - \frac{w_i}{w}) - \sum_{i=1}^{N} \frac{w_i}{n_i}(1 - \frac{w_i}{w})]}, \tag{3.19}$$

where $\hat{q} \equiv 1-\hat{p}$. Kleinman [4] further suggests that better estimates are obtained if S, in Eq. (3.19), is replaced by (N-1)S/N. The choice of weights is made such that the estimate of μ is the linear unbiased estimate with minimum variance, i.e., weight each k_i/n_i with the inverse of its variance, namely

$$w_i = \frac{n_i}{1 + r(n_i - 1)} \tag{3-20}$$

where

$$r \equiv \sigma^2/(\mu(1-\mu)) . \tag{3-21}$$

Once $\hat{\mu}$ and $\hat{\sigma}^2$ of the prior distribution are calculated from Eqs. (3.18) and (3.19), the parameters a and b are found by solving Eqs. (2.4) and (2.5) for a and b. However, to calculate $\hat{\mu}$ and $\hat{\sigma}^2$, the weights, w_i, must be known, which from Eq. (3.20) implies that r (or $\hat{\sigma}^2$) must be known. Thus Eqs. (3.18)-(3.20) can be viewed as three equations for the quantities w_i, μ, and $\hat{\sigma}^2$ which can be solved by the following iteration scheme. Choose r = 0 so that $w_i = n_i$ ("binomial weighting") and solve for the resulting $\hat{\mu}$ and $\hat{\sigma}^2$. With these values of $\hat{\sigma}^2$ and $\hat{\mu}$ calculate r and new values of w_i from Eqs. (3.20) and (3.21) ("empirical weighting"). Continue iterating until $\hat{\mu}$, $\hat{\sigma}^2$, and w_i no longer change ("converged weighting"). Finally it should be noted that $\hat{\sigma}^2$ may be negative from Eq. (3.19). For this case r is set to zero, i.e., only binomial weighting is used. One major disadvantage of this method is that the iterative scheme just outlined occasionally does not converge or converges extremely slowly. Even the first iteration ("binomial weighting") occasionally produces infeasible solutions.

3.4 Maximum Likelihood Method Based on Marginal Distributions

A fourth technique for obtaining estimates of beta parameters a and b from the observed data is based on the marginal or mixture distribution of Eq. (2.6). The likelihood function

$$L(a,b|k_1,k_2\ldots k_N,n_1,n_2\ldots n_N) \equiv \prod_{i=1}^{N} h(a,b|k_i,n_i) \tag{3.22}$$

is the probability of obtaining $k_1,k_2,\ldots,k_N$ failures in $n_1,n_2,\ldots,n_N$ tries of components 1,2,...,N, respectively, for components whose probability distribution for failure is given by the prior distribution of Eq. (2.2) with parameters a and b. The values of a and b which maximize the likelihood function are called the maximum likelihood estimates, $\hat{a}$ and $\hat{b}$. If k_i and n_i are the observed data, then the maximum likelihood estimates maximize the probability of obtaining the observed values over all possible parameter values a and b.

Unfortunately the maximum likelihood estimators cannot be determined analytically when the marginal distribution, h, in Eq. (3.22) is a beta-binomial distribution. Thus numerical methods must be used. Substitution of Eq. (2.6) into Eq. (3.22) yields

$$L(a,b) \equiv L(a,b|k_1\ldots k_N,n_1\ldots n_N) =$$

$$\left\{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right\}^N \prod_{i=1}^{N} C_i \frac{\Gamma(a+k_i)\Gamma(b+n_i-k_i)}{\Gamma(a+b+n_i)} \tag{3.23}$$

where

$$C_i \equiv \binom{n_i}{k_i} = \frac{\Gamma(n_i+1)}{\Gamma(k_i+1)\Gamma(n_i-k_i+1)} \ . \tag{3.24}$$

The problem is to find the values of a and b (constrained such that a>o and b>o) which maximize L, or equivalently, which maximize $ln[L]$. This latter form is preferrable for numerical purposes since the $ln\Gamma$ function varies more slowly than does the Γ function. An example of a typical likelihood function is shown in Fig. 3.1. The extrema of

1426 232

1425 231

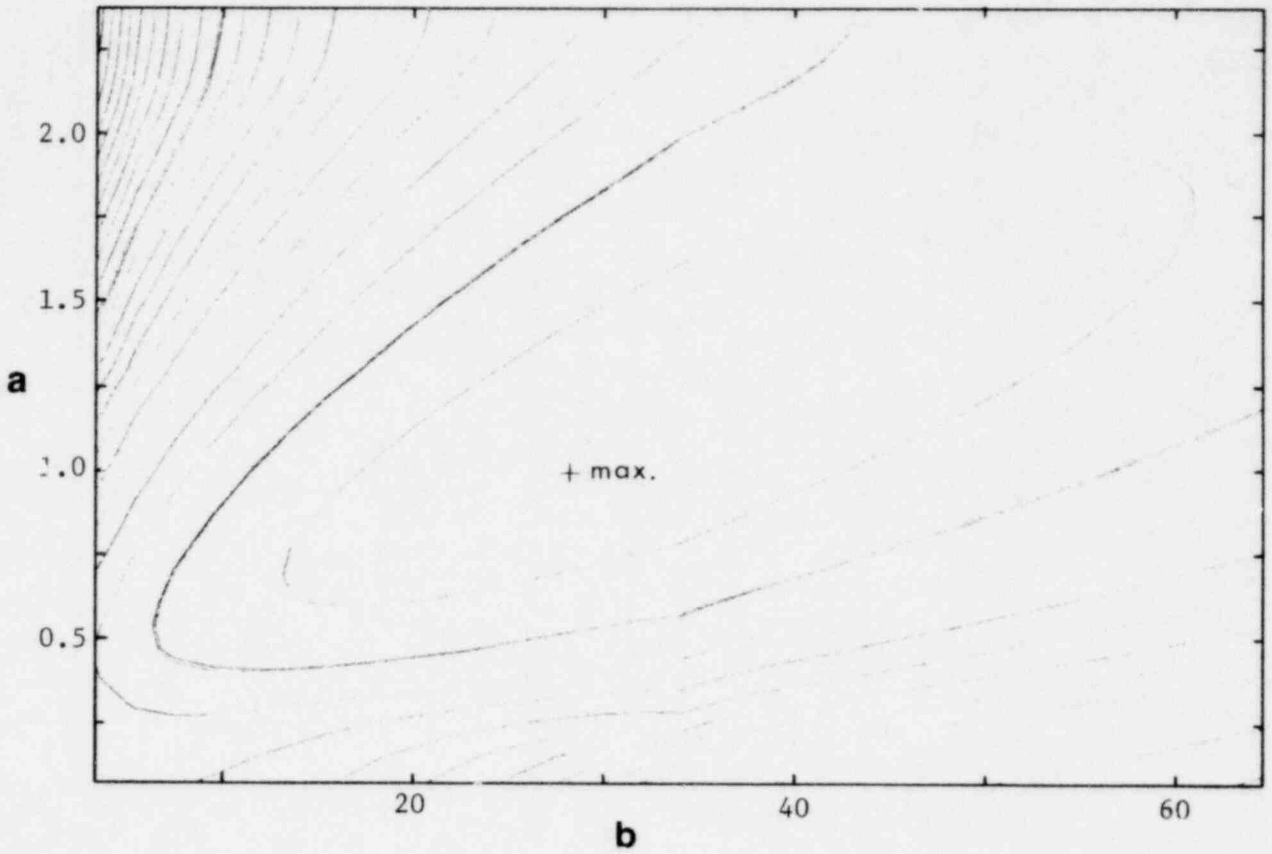


Fig. 3.1 A contour plot of the logarithm of the likelihood function for a three component case (n_1=100, n_2=392, n_3=230, k_1=6, k_2=1, k_3=11).

$\ell n L(a,b)$ are obtained from solutions to

$$\frac{\partial}{\partial a} \ell n L(a,b) = 0$$

$$\frac{\partial}{\partial b} \ell n L(a,b) = 0$$

or explicitly

$$N\left\{\psi(a+b) - \psi(a)\right\} + \sum_{i=1}^{N} \left\{\psi(a+k_i) - \psi(a+b+n_i)\right\} = 0 \qquad (3.25)$$

and

$$N\left\{\psi(a+b) - \psi(b)\right\} + \sum_{i=1}^{N} \left\{\psi(b+n_i-k_i) - \psi(a+b+n_i)\right\} = 0 \ . \qquad (3.26)$$

where $\psi(z) \equiv \frac{d}{dz}[\ell n\Gamma(z)]$, the digamma function. The numerical solution of these two simultaneous equations is obtained by standard numerical techniques (such as the Newton-Raphson method [5], with the matching moments solution as the starting points). Care must be taken since $(a,b)\to\infty$ is also a solution of Eqs. (3.25) and (3.26). If the sample data consist solely of one component (N=1), the only solution of the equation is for $a=b=\infty$ although a/b is finite such that from Eqs. (2.4) and (2.5) the mean of the prior is $\mu=k/n$ and the variance is $\sigma^2=0$ -- an expected result when only one sample is used (see Fig. 3.2). However, it has been found that for some data with N>1, Eqs. (3.25) and (3.26) may also have no finite positive solution.

3.5 Results for Diesel Engine Data

The beta prior distribution parameters (mean, variance, a and b) were estimated for standby diesel engine data (see Table 3.1) for various engine groupings by the three feasible methods described in the previous sections. The prior based maximum likelihood method (see Section 3.2) was not used as a result of inherent difficulties for zero failure cases. A listing of the computer code is given in Appendix I, and the results are summarized in Table 3.2.

From these results, several interesting features are apparent. First the maximum likelihood method (Method III) and the matching moments to the marginal distribution (Method II) did not always produce estimates of the prior variance, i.e, only values of b/a (or the mean) resulted. For the marginal-based maximum likelihood method, the solution, was for $a,b\to\infty$

1426 234

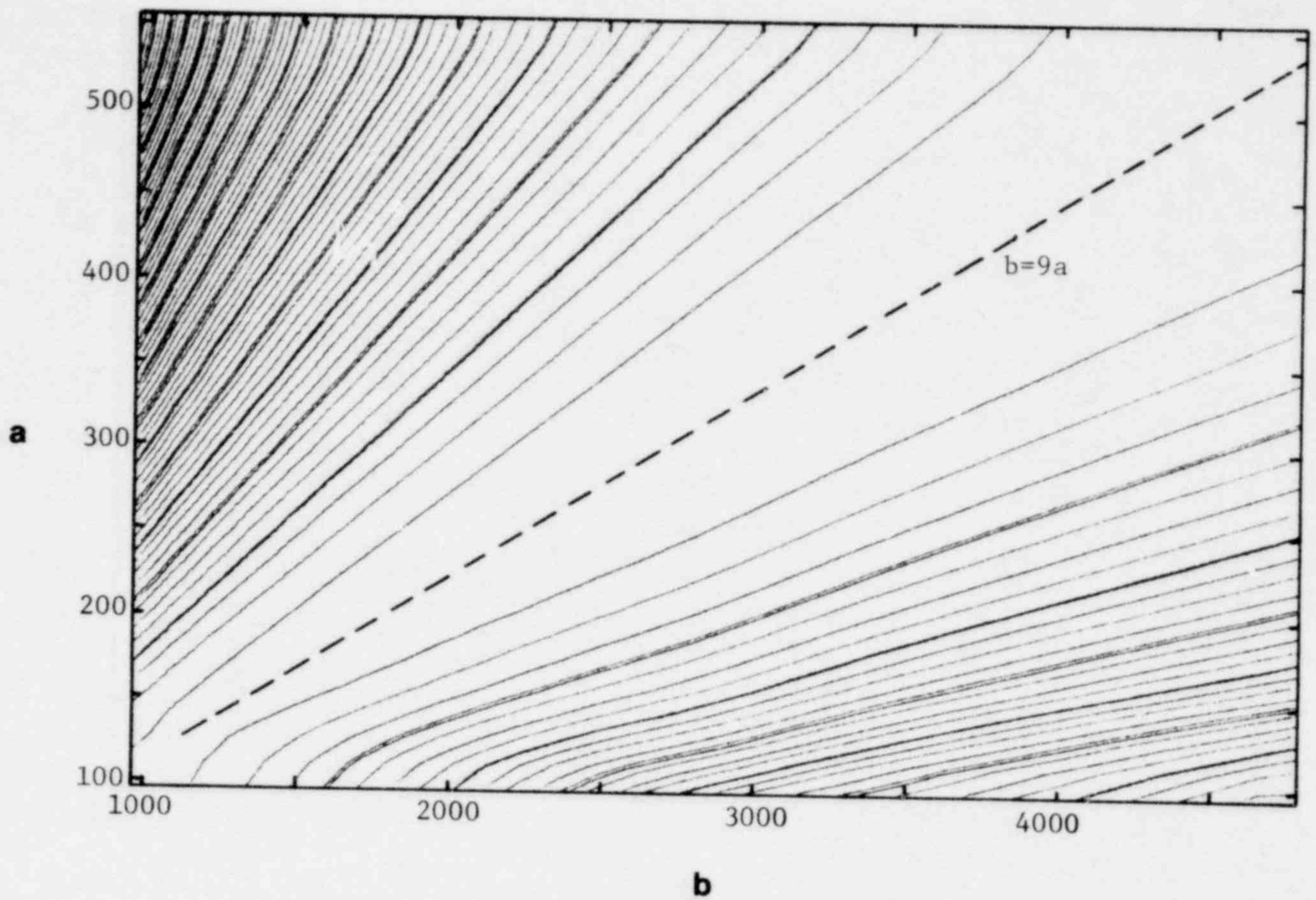


Fig. 3.2 Contour plot of the logarithm of the likelihood function for a one component case (n_1=100, k_1=10). The maximum occurs at $a,b \to \infty$ along the line $b/a = (n_1/k_1)-1 = 9$.

1426 233

Table 3.1 Diesel Engine Failure Probability Data [6].

Plant	No. of Units	Manufacturer	No. of Starts (n_i)	Failures (k_i)
Yankee	3	GM	100	6
Peach Bottom I	1	"	392	1
Oyster Creek	2	"	230	11
Monticello	2	"	68	5
Turkey Point 3	2	"	23	4
Maine Yankee	2	"	23	0
Fort Calhoun	2	"	12	2
Nine Mile Pt.	2	"	99	0
Surry 1, 2	3	"	33	3
Dresden 2, 3	3	"	126	9
Quad Cities 1, 2	3	"	47	2
Haddam Neck	2	"	87	1
Point Beach 1, 2	2	"	71	2
San Onofre	2	FAIRBANKS	656	3
HB-Robinson	2	"	73	5
Millstone 1	1	"	35	1
Vermont Yankee	2	"	37	1
Indian Pt.-2	3	ALCO	13	0
Ginna	2	"	95	2
Palisades	2	"	51	2
Pilgrim	2	"	35	2
Zion 1	3	COOPER	17	7
Dresden 1	1	GE	335	4
Big Rock Pt.	1	CATEPILLAR	206	9
LaCrosse	1	ALLIS-CHALMERS	76	1

but with a finite ratio and hence well-defined mean (see Fig. 3.3 for a contour plot of the maximum likelihood function for the four ALCO engine case). For the marginal distribution matching moments method, estimates of r of the prior variance were negative. Interestingly, these two methods failed for the same cases.

Second, while the method of matching moments to the assumed beta prior distribution (Method I) always yields finite positive results, the estimated means and standard deviations are always greater than the estimates obtained by the other methods.

Third, the iteration scheme used to calculate the weighting values, w_i, in Method II (marginal distribution matching) did not always converge evenly or quickly. For example, the iterated results for the four FAIRBANKS diesel engines are shown in Table 3.3. On the other hand, the thirteen GM diesel engines gave results which converged smoothly to five significant figures in only four iterations.

Finally, when they are obtainable the marginal-based maximum likelihood results and the converged results of matching marginal distribution moments are usually nearly equal, with the former usually yielding slightly larger estimates of the prior standard deviation. An assessment as to the ability of these three methods to estimate accurately the prior parameters from data generated from a pure beta distribution was undertaken in the second phase of this study. The results of this simulation study are presented in Section 4.

In Figs. 3.4 and 3.5 the estimated beta prior distributions obtained by the prior matching moments method (Method I) are shown for the diesel engine data grouped by manufacturer and by the number of starts, respectively. Notice that the Fairbanks and ALCO groupings appear to be very similar in shape, while the GM and Others, although of the same shape, have prior distributions which appear to be quite different from those of the Fairbanks and ALCO groupings. The estimated prior distributions for data grouped by number of starts reveal an apparent aging phenomenon. For the group 0-25 starts the prior distribution has no mode and is highly skewed towards zero failure probability. The three other groupings all are unimodal with the failure probability at the mode (most probable failure probability) decreasing as the engines age (or more experience is obtained). In Section 3.7 a more critical comparison is presented of these results for the diesel engine failure data.

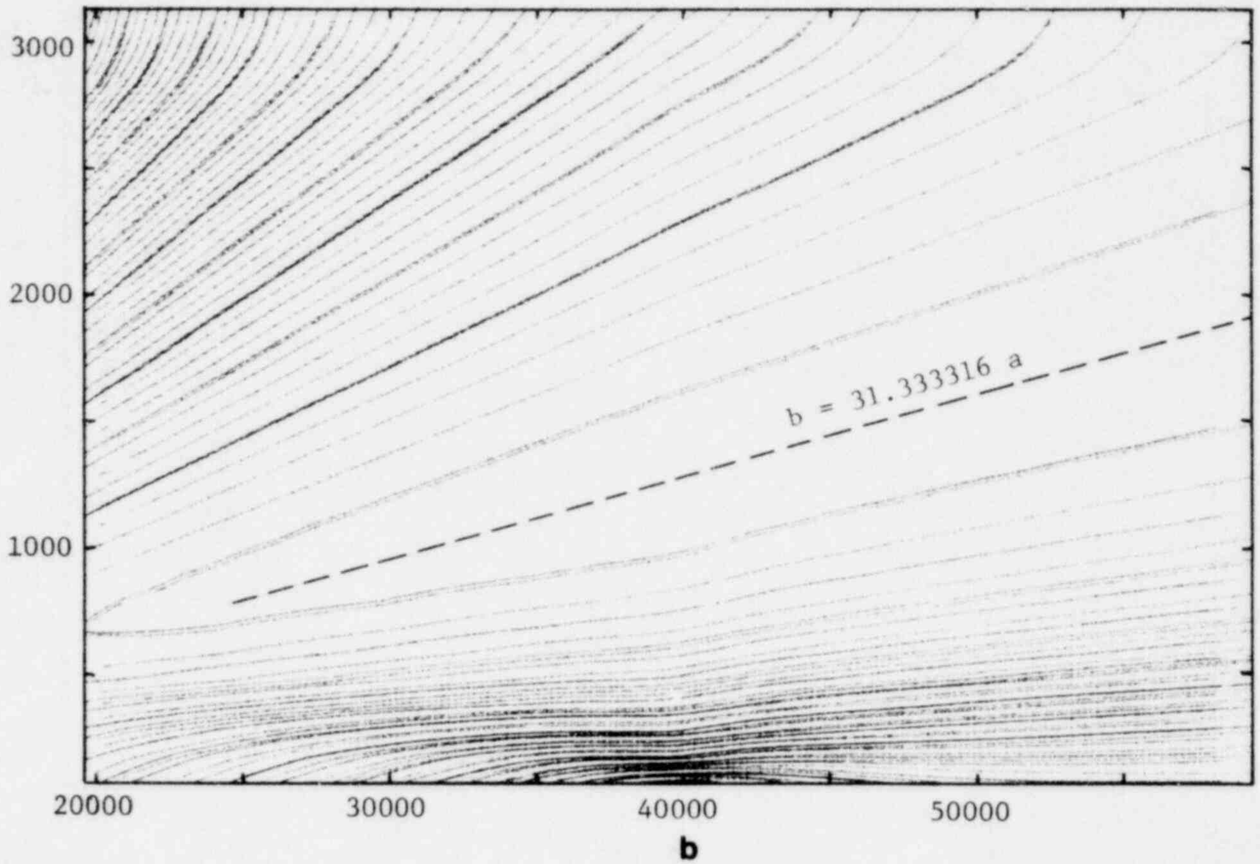


Fig. 3.3 Contour plot of the logarithm of the likelihood function for the four ALCO diesel engines of Table 3.1. The maximum occurs at $a, b \to \infty$ with the ratio $b/a = 31.333316$.

Tables 3.2. Comparison of calculated prior distribution parameters by three different techniques: (I) matching data to prior moments, (II) matching data to marginal moments, (III) marginal maximum likelihood method.

Problem	Method	Mean, μ	Stand. Dev., σ	a	b
13 GM	I	0.0592	0.0577	0.9303	14.80
Diesel	II	0.0491	0.0373	1.595	30.88
Engines	III	0.0502	0.0437	1.204	22.79
Fairbanks	I	0.0322	0.0266	1.385	41.66
Diesel	II	0.0270	0.0177	2.236	80.58
Engines	III	0.0291	0.0245	1.342	44.81
Four	I	0.0294	0.0245	1.364	45.12
ALCO	II	0.0309	negative	b/a = 31.333333	
Engines	III	0.0309	not obtained	b/a = 31.333316	
Other	I	0.120	0.195	0.2139	1.567
Four	II	0.110	0.159	0.3209	2.584
Engines	III	0.108	0.126	0.5550	4.570
Engines	I	0.150	0.169	0.5222	2.949
With 0-25	II	0.151	0.128	1.029	5.808
Starts	III	0.145	0.152	0.6318	3.728
Engines	I	0.0492	0.0263	3.287	63.46
With 25-50	II	0.0481	negative	b/a = 19.77778	
Starts	III	0.0481	not obtained	b/a = 19.77775	
Engines	I	0.0350	0.0268	1.612	44.44
With 50-100	II	0.0339	0.0154	4.626	131.7
Starts	III	0.0341	0.0186	3.192	90.55
Engines	I	0.0303	0.0281	1.100	35.16
With more	II	0.0283	0.0230	1.447	44.67
Than 100 starts	III	0.0287	0.0271	1.062	35.97

Table 3.3. Results of Matching Data to Marginal Distribution Moments (Method II) for the Fairbanks Engines.

Iteration	Mean	Stand. Dev.	a	b
1 (binomial)	0.012484	0.026654	0.2042	16.149
2 (empirical)	0.031138	0.0092698	10.9001	339.183
3	0.019762	0.023373	0.68098	33.778
4	0.029899	0.013094	5.0284	163.151
5	0.023544	0.020929	1.2123	50.279
6	0.028791	0.015238	3.4382	115.98
7	0.025300	0.019462	1.6220	62.486
8	0.028030	0.016395	2.8131	97.547
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
28	0.027004	0.017704	2.2368	80.596
29	0.027000	0.017708	2.2351	80.549
30	0.027003	0.017705	2.2363	80.583

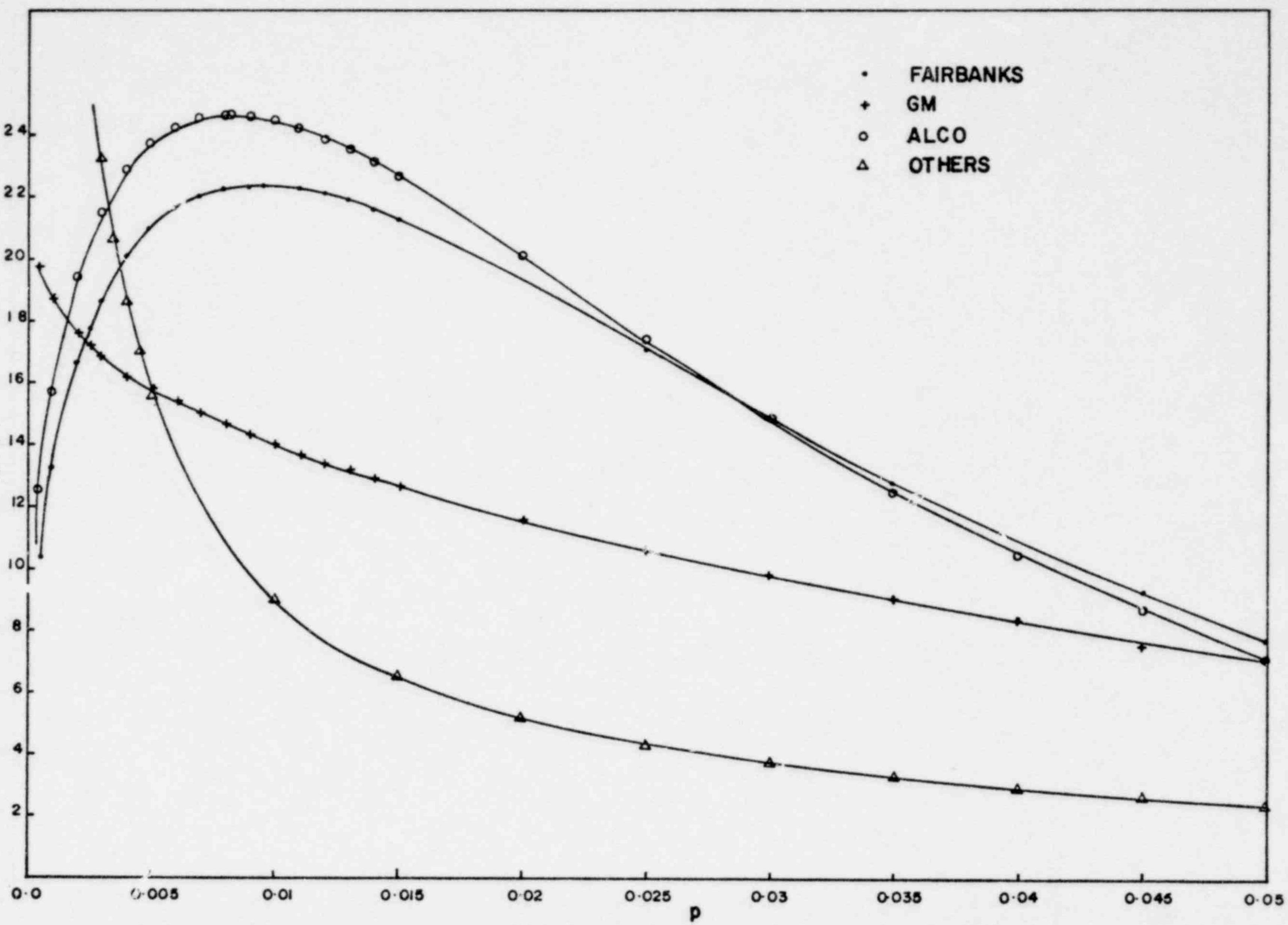


Fig. 3.4 The estimated beta prior distributions for the diesel engine data of Table 3.1 grouped by manufacturer. The beta parameters for each distribution were obtained by the prior matching moments technique.

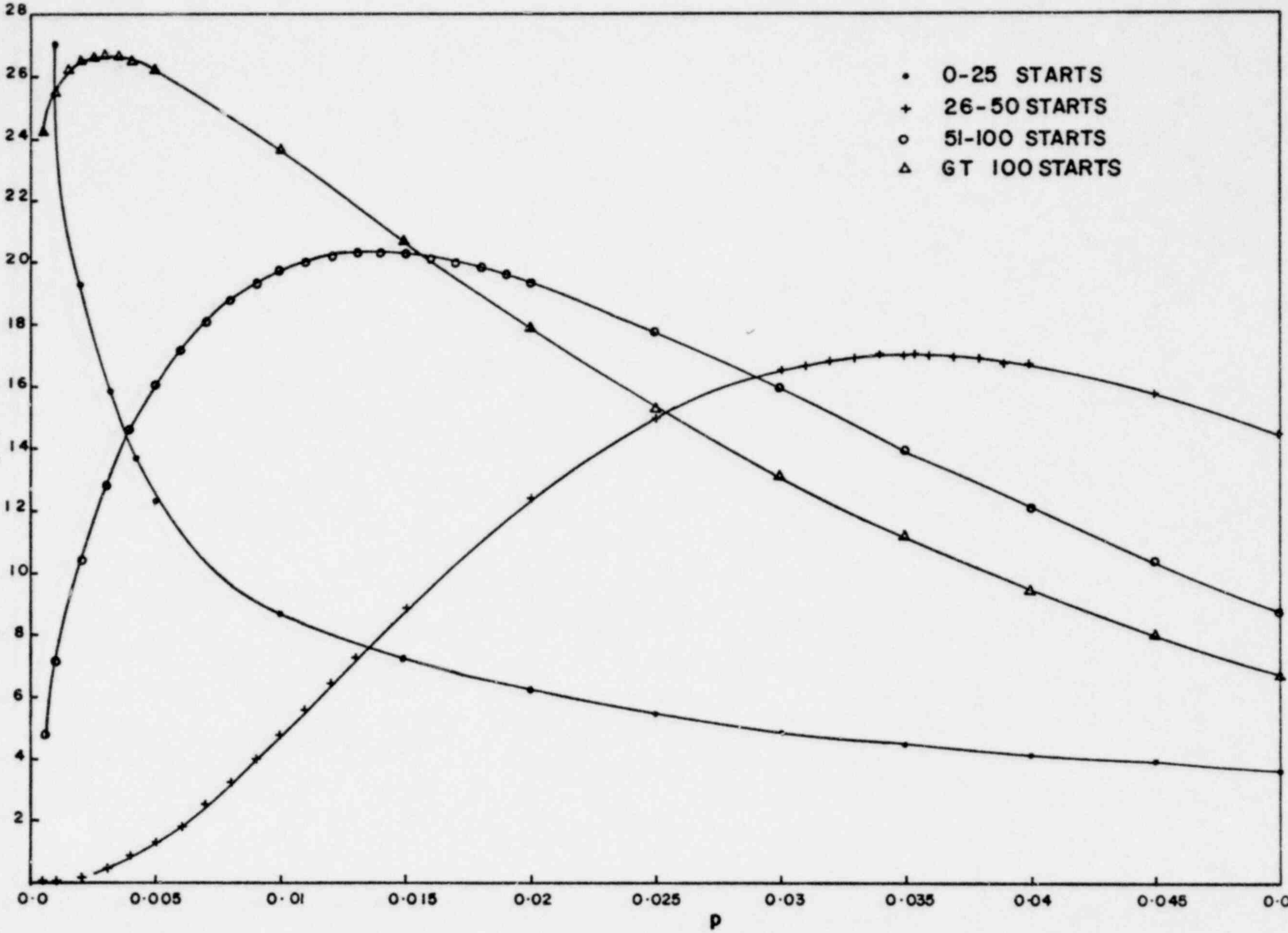


Fig. 3.5 The estimated beta prior distributions for the diesel engine data of Table 3.1 grouped by numer of starts. The beta parameters for each distribution were obtained by the prior matching moments technique.

3.6 Maximum Likelihood Bounds on the Variances of Prior Parameter Estimates

One of the most attractive features of the maximum likelihood method is that, besides yielding estimates of the parameters, this method can also yield lower bounds on the variances and the covariance of the parameters. These lower bounds can often be used as useful approximations to the variances and covariance. In this section a brief review of the pertinent aspects of this method is presented, and the method is applied to the problem of estimating variances and the covariance of the prior beta parameter estimates.

For N independent observation, $x_1, x_2, \ldots, x_N$, where the i-th observation is from a distribution $h_i(x|\underline{\theta})$, i.e., the marginal distribution for the i-th component, the *likelihood function* is defined by

$$L(\underline{\theta}|x_1, x_2 \ldots x_N) \equiv \prod_{i=1}^{N} h_i(\underline{\theta}|x_i) \tag{3.27}$$

where x and $\underline{\theta}$ represent the sample random variable and parameter vector, respectively. The maximum likelihood estimators of $\underline{\theta}$ are denoted by $\underline{\hat{\theta}}$, and are those values of the parameters which maximize L, i.e.,

$$\frac{\partial}{\partial \theta_i} L(\underline{\theta}|x_1, x_2 \ldots x_N)\bigg|_{\underline{\theta} = \underline{\hat{\theta}}} = 0 \ , \quad i = 1, 2, \ldots, N \tag{3.28}$$

or equivalently maximize $ln L$, i.e.,

$$\frac{\partial}{\partial \theta_i} L(\underline{\theta}|x_1 \ldots x_N)\bigg|_{\underline{\theta}=\underline{\hat{\theta}}} = 0 \ , \quad i = 1, 2, \ldots, N \ .$$

The elements of the *information matrix* $\underline{I}(\underline{\theta})$, are defined as

$$I_{ij}(\underline{\theta}) \equiv E\left(-\frac{\partial^2 lnL}{\partial \theta_i \partial \theta_j}\right) = -\int dx_1 \int dx_2 \ldots \int dx_N \frac{\partial^2 \, lnL}{\partial \theta_i \partial \theta_j} L(\underline{\theta}|x_1 \ldots x_N) \ ,$$

$$i,j = 1, 2, \ldots, N \tag{3.29}$$

where the integration (or summation in the case of a discrete distribution) is over all possible values of variables $x_1 \ldots x_N$. If the distribution of the likelihood function with respect to each parameter is symmetrical in the neighborhood of $\hat{\underline{\theta}}$, then

$$E\left(\frac{\partial^2 L}{\partial\theta_i \partial\theta_j}\right) \simeq \left(\frac{\partial^2 L}{\partial\theta_i \partial\theta_j}\right)\Bigg|_{\underline{\theta} = \hat{\underline{\theta}}} . \tag{3.30}$$

Asymptotic properties of the likelihood function guarantees that the above approximation is valid provided N is sufficiently large regardless of the symmetry of the likelihood function.

One of the most important theorems about the maximum likelihood method is known as the Cramer-Rao-Frechet inequality [3] which states

$$\sigma_{ii}(\underline{\theta}) \leq \text{variance}(\hat{\theta}_i) \tag{3.31}$$

and

$$|\sigma_{ij}(\underline{\theta})| \leq |\text{covariance}(\hat{\theta}_i, \hat{\theta}_j)| \tag{3.32}$$

where $\underline{\sigma}$ is the inverse of the information matrix $\underline{I}$. In effect this theorem provides lower bound estimates of the variance and covariance of the parameters. In fact under rather weak restrictions [3]

$$\lim_{N\to\infty} E[\hat{\underline{\theta}}] = \underline{\theta} \quad , \tag{3.33}$$

$$\lim_{N\to\infty} N[\text{var}(\hat{\theta}_i)] = \sigma_{ii} \tag{3.34}$$

and

$$\lim_{N\to\infty} N[\text{cov}(\hat{\theta}_i, \hat{\theta}_j)] = \sigma_{ij} . \tag{3.35}$$

With finite sample sizes, the information matrix is thus often used to give approximate values of the variances and covariance which asymptotically converge to the true values as the sample sizes become increasingly large [3].

To apply the above results to the problem of estimating the variances and covariances of the two parameters of the prior beta distribution, begin by constructing the information matrix for Eq. (3.27),

1426 243

$$\underline{I}(a,b) \equiv - \begin{pmatrix} E\left(\frac{\partial^2 \ln L}{\partial a^2}\right) & E\left(\frac{\partial^2 \ln L}{\partial a \partial b}\right) \\ & \\ E\left(\frac{\partial^2 \ln L}{\partial a \partial b}\right) & E\left(\frac{\partial^2 \ln L}{\partial b^2}\right) \end{pmatrix}. \tag{3.36}$$

The derivatives of the logarithm of the likelihood function, i.e., Eq. (3.23), are given by,

$$\frac{\partial^2 \ln L}{\partial a^2}(a,b) = N\{\psi'(a+b) - \psi'(a)\} + \sum_{i=1}^{N} \{\psi'(a+k_i) - \psi'(a+b+n_i)\} \tag{3.37}$$

$$\frac{\partial^2 \ln L}{\partial b^2}(a,b) = N\{\psi'(a+b) - \psi'(b)\} + \sum_{i=1}^{N} \{\psi'(b+n_i-k_i) - \psi'(a+b+n_i)\} \tag{3.38}$$

and

$$\frac{\partial^2 \ln L}{\partial a \partial b}(a,b) = N\psi'(a+b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \tag{3.39}$$

where $\psi'(x) \equiv d^2[\ln\Gamma(z)]/dz^2$ is the trigamma function [8] (see Appendix I for computational aspects of this function). The expectation values for the matrix elements in Eq. (3.36) are calculated from Eq. (3.22), by*

$$E[\cdot] \equiv \sum_{k_1=0}^{n_i} \sum_{k_2=0}^{n_2} \cdots \sum_{k_N=0}^{n_N} [\cdot]\, L(a,b|k_1 \ldots k_N, n_1 \ldots n_N) \; . \tag{3.40}$$

Since $\sum_{k_i=0}^{n_i} h(k_i|n_i,a,b) = 1$, the substitution of the explicit form of the likelihood function from Eq. (3.27) and simplification gives the following results for the matrix elements of the information matrix:

*The dot in the square brackets represents the various derivatives given in Eq. (3.36).

$$E\left(\frac{\partial^2 \ln L}{\partial a^2}\right) = N\{\psi'(a+b)-\psi'(a)\} + \sum_{i=1}^{N} \sum_{k_i=0}^{n_i} \psi'(a+k_i)h(k_i|n_i,a,b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \qquad (3.41)$$

$$E\left(\frac{\partial^2 \ln L}{\partial b^2}\right) = N\{\psi'(a+b)-\psi'(b)\} + \sum_{i=1}^{N} \sum_{k_i=0}^{n_i} \psi'(b+n_i-k_i)h(k_i|n_i,a,b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \qquad (3.42)$$

$$E\left(\frac{\partial^2 \ln L}{\partial a \partial b}\right) = N\psi'(a+b) - \sum_{i=1}^{N} \psi'(a+b+n_i) \ . \qquad (3.43)$$

Finally from Eqs. (3.31) and (3.32) we have the following approximations for the variance and covariance of the maximum likelihood estimators:

$$\mathrm{Var}(\hat{a}) \simeq [\underline{I}^{-1}(\hat{a},\hat{b})]_{11} \qquad (3.44)$$

$$\mathrm{Var}(\hat{b}) \simeq [\underline{I}^{-1}(\hat{a},\hat{b})]_{22} \qquad (3.45)$$

$$\mathrm{Cov}(\hat{a},\hat{b}) \simeq [\underline{I}^{-1}(\hat{a},\hat{b})]_{12} \qquad (3.46)$$

where the maximum likelihood estimates $\hat{a}$ and $\hat{b}$ are substituted for the true parameter values.

The numerical evaluation of the expected values of the matrix elements of the information matrix can be quite time consuming especially if the n_i are large and the number of components N grouped into the class is also large. Application of Eq. (3.30) allows a much more expedient, but approximate, evaluation of these matrix elements. Specifically one has

$$E\left(\frac{\partial^2 \ln L}{\partial a^2}\right)\Bigg|_{\substack{a=\hat{a}\\ b=\hat{b}}} \simeq \left(\frac{\partial^2 \ln L}{\partial a^2}\right)\Bigg|_{\substack{a=\hat{a}\\ b=\hat{b}}} = N\psi'(\hat{a}+\hat{b})-N\psi'(\hat{a}) + \sum_{i=1}^{N}\{\psi'(\hat{a}+k_i) - \psi'(\hat{a}+\hat{b}+n_i)\} \ , \qquad (3.47)$$

1426 245

$$E\left(\frac{\partial^2 \ln L}{\partial b^2}\right)\Bigg|_{\substack{a=\hat{a}\\ b=\hat{b}}} \simeq \left(\frac{\partial^2 \ln L}{\partial b^2}\right)_{\substack{a=\hat{a}\\ b=\hat{b}}} = N\psi'(\hat{a}+\hat{b})-N\psi'(\hat{b})$$

$$+ \sum_{i=1}^{N} \{\psi'(b+n_i-k_i) - \psi'(\hat{a}+\hat{b}+n_i)\} , \qquad (3.48)$$

$$E\left(\frac{\partial^2 \ln L}{\partial a \partial b}\right)_{\substack{a=\hat{a}\\ b=\hat{b}}} \simeq \left(\frac{\partial^2 \ln L}{\partial a \partial b}\right)_{\substack{a=\hat{a}\\ b=\hat{b}}} = N\psi'(\hat{a}+\hat{b}) - \sum_{i=1}^{N} \psi'(\hat{a}+\hat{b}+n_i). \qquad (3.49)$$

In practice, it has been found that the information matrix constructed from these approximations (Eqs. 3.49-3.51) gives very similar results for large sample size, N, as the more complicated, but exact, method of Eqs. (3.41)-(3.43). As an application of the covariance-variance calculations, the 25 diesel engines of Table 3.1 were fit to a single beta prior by the maximum likelihood method based upon the marginal distribution (Section 3.4). The results of the calculations of the variance and covariance bounds are presented in Table 3.4.

Table. 3.4 Estimates of Beta Prior Parameters and Variance Bounds for the 25 Diesel Engines of Table 3.1. The Maximum Likelihood Method Based on the Marginal Distribution (Eq. 3.27) was used.

Estimated Parameters		Exact Eqs. (3.41)-(3.43)	Aprox. Eqs. (3.47)-(3.49)
$\hat{a}$ = 1.0522	Var($\hat{a}$) =	0.1763	0.1545
	Var(b) =	81.67	93.73
$\hat{b}$ = 19.902	Cov($\hat{a}$,$\hat{b}$) =	3.273	3.283

The calculation of the variance bounds by both the exact and approximate information matrix is provided as an option in the computer program BETA III, listed and discussed in Appendix I.

1426 245

1426 246

3.7 Variance Estimates from the Method of Matching Moments to the Prior Moments

A simple, but approximate method to estimate variances for the beta parameters a and b can be obtained from the closed-form solution for the beta parameter estimates derived in Section 3.1. From the matching of data moments to those of the beta prior, the following results were previously obtained for the beta parameters (namely, Eqs. (3.5) and (3.6)):

$$a = \frac{\hat{\mu}_{ob}^2}{\hat{\sigma}_{ob}^2}(1-\hat{\mu}_{ob}) - \hat{\mu}_{ob} \tag{3.50}$$

and

$$b = \frac{\hat{\mu}_{ob}}{\hat{\sigma}_{ob}^2}(1-\hat{\mu}_{ob})^2 + \hat{\mu}_{ob} - 1 = a(1-\hat{\mu}_{ob})/\hat{\mu}_{ob} \ . \tag{3.51}$$

Equations (3.50) and (3.51) can be used to find expressions for estimates of the variances of a and b from the following first order Taylor series approximation [9]:

$$s^2(a) = \left(\frac{\partial a}{\partial \hat{\mu}_{ob}}\right)^2 s^2(\hat{\mu}_{ob}) + \left(\frac{\partial a}{\partial \hat{\sigma}_{ob}^2}\right)^2 s^2(\hat{\sigma}_{ob}^2) \ , \tag{3.52}$$

$$s^2(b) = \left(\frac{\partial b}{\partial \hat{\mu}_{ob}}\right)^2 s^2(\hat{\mu}_{ob}) + \left(\frac{\partial b}{\partial \hat{\sigma}_{ob}^2}\right)^2 s^2(\hat{\sigma}_{ob}^2) \ , \tag{3.53}$$

where $s^2(\hat{\mu}_{ob})$ and $s^2(\hat{\sigma}_{ob}^2)$ are estimates for the variances of $\hat{\mu}_{ob}$ and $\hat{\sigma}_{ob}^2$. In these first order approximations, the covariances are assumed to be negligible. Other approximations (discussed later) can incorporate the covariance between $\hat{\mu}_{ob}$ and $\hat{\sigma}_{ob}$. Estimates for $s^2(\hat{\mu}_{ob})$ and $s^2(\hat{\sigma}_{ob}^2)$ are [10]:

$$s^2(\hat{\mu}_{ob}) = \frac{\hat{\sigma}_{ob}^2}{N} \ , \tag{3.54}$$

and

$$s^2(\hat{\sigma}_{ob}^2) = \frac{2(\hat{\sigma}_{ob}^2)^2}{N-1} \ . \tag{3.55}$$

To obtain this last result it has been assumed that $s^2(\hat{\sigma}^2_{ob})$ is normally distributed. Wilks [11] presents a distribution independent formula:

$$s^2(\hat{\sigma}^2_{ob}) = \frac{1}{N}\left(\mu_4 - \frac{N-3}{N-1}\sigma^4\right) \tag{3.56}$$

where μ_4 is the fourth central moment, σ^4 is the square of the sample variance. Equations (3.52) and (3.53) become, upon substitution for $s^2(\hat{\mu}_{ob})$ and $s^2(\hat{\sigma}^2_{ob})$ from the normal based Eqs. (3.54) and (3.55)

$$s^2(a) = \left\{\left[\frac{1}{\hat{\sigma}^2_{ob}}(2\hat{\mu}_{ob}-3\hat{\mu}^2_{ob})\right] - 1\right\}^2 \frac{\hat{\sigma}^2_{ob}}{N} + 2\left(\frac{\hat{\mu}^2_{ob}(1-\hat{\mu}_{ob})}{(\hat{\sigma}^2_{ob})^2}\right)^2 \frac{(\hat{\sigma}^2_{ob})^2}{N-1} \tag{3.57}$$

and

$$s^2(b) = \frac{1}{N\hat{\sigma}^2_{ob}}\left[\hat{\sigma}^2_{ob} + 1 - 4\hat{\mu}_{ob} + 3\mu^2_{ob}\right]^2 + \frac{2}{N-1}\left(\frac{\hat{\mu}_{ob}}{\hat{\sigma}^2_{ob}}(1-\mu_{ob})^2\right)^2 \tag{3.58}$$

It should be emphasized that the above result is only approximate since the covariance between the mean and the variance of the beta prior have been assumed to be zero. Nevertheless, order of magnitude values for the variances can be obtained with this approximation. For example, the above method (based on Eqs. (3.54) and (3.56)) gives for the 25 diesel engines of Table 3.1 var(a) = 0.1393 and var(b) = 24.03. These values compare with the maximum likelihood results of var(a) ≈ 0.1763 and var(b) ≈ 81.66.

Once estimates have been obtained for the prior beta parameters and for their variances, various statistical tests can be used to search for significant differences between the estimates for various groupings of the diesel engine data considered in Section 3.5. One of the simplest tests is based on the statistic

$$z = (\xi_1-\xi_2)/[s^2(\xi_1) + s^2[(\xi_2)]^{\frac{1}{2}} \tag{3.59}$$

where ξ_i and $s^2(\xi_i)$ are respectively the estimated prior paramber ($\hat{a}$ or $\hat{b}$) and its estimated variance for the i-th data grouping. Under very general conditions, the z statistic will be asymptotically distributed as a unit normal deviate [16]. Thus the cumulative unit normal distribution can be used as a test criterion, if it is assumed that the sample sizes used to obtain the estimates of the prior parameters are sufficiently large for the asymptotic normality of z to be valid.

In Table 3.5 the estimates are presented for the prior beta parameters obtained by the prior matching moment technique, together with two estimates of their variances. The first variance estimates for $s^2(a)$ and $s^2(b)$ are based upon an assumption of normality for the distribution of $s^2(\hat{\sigma}^2_{ob})$ (Eq. (3.55)) and are computed directly from Eqs. (3.59) and (3.58). The second variance estimate is based on a distribution-independent result (Eq. (3.56)) for $s^2(\hat{\sigma}^2_{ob})$. Both variance estimation techniques are seen to give comparable results with the distribution-independent estimates always being slightly smaller than the normal-based estimates.

With these variance estimates, the z statistic may be computed from Eq. (3.59) for pairs of groupings of the diesel failure data. In Table 3.6 the z values are given for the case of the normal-based estimate of $s^2(\hat{\sigma}^2_{ob})$ while Table 3.7 presents the results based of the distribution-independent estimate of $s^2(\hat{\sigma}^2_{ob})$. From the values of the cumulative normal in these two tables it is apparent that one cannot conclude the estimated prior parameters for various diesel groupings are significantly different at the 5% level (i.e., $\Phi(z) < 0.025$ or $\Phi(z) > 0.975 > 0.975$). Thus while the estimated diesel prior distributions shown in Figs. 3.4 and 3.5 appear to have noticeable differences for the different diesel engine groupings, these differences may arise more from the paucity of the data used to estimate the prior parameters than from any real physical differences.

3.8 Error Bands for Estimated Prior Distributions

In this section a method is presented to estimate the confidence bounds on the estimated prior distribution, both for the estimated probability distribution function (pdf) and for the estimated cumulative distribution function (cdf). The pdf estimate for failure probability p is given as

$$g(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{(a-1)} (1-p)^{(b-1)} . \qquad (3.60)$$

1426 249

1426 250

Table 3.5. The estimated prior beta distribution parameters and their standard deviations as calculated by the prior matching moment technique for various groupings of the diesel engine data (see Table 3.1). The quantity N equals the number of plants in each grouping.

Grouping	N	a	$\sigma_1(a)^*$	$\sigma_2(a)^{**}$	b	$\sigma_1(b)^*$	$\sigma_2(b)^{**}$
<u>Manufacturers</u>							
GM	13	0.930	0.645	0.610	14.795	7.432	6.654
Fairbanks	4	1.385	1.623	1.308	41.662	38.502	25.444
Alco	4	1.364	1.606	1.256	45.120	41.751	25.472
Others	4	0.214	0.490	0.440	1.567	2.523	1.955
<u>Number of Starts</u>							
0-25	5	0.522	0.720	0.611	2.948	2.987	2.070
26-50	5	3.287	2.817	2.115	63.462	47.591	31.200
51-100	9	1.612	1.160	0.908	44.437	25.118	15.369
>100	6	1.100	1.095	0.892	35.162	26.202	16.553

*Based on normality of $s^2(\theta^2_{ob})$, Eq. (3.55)

**Distribution-independent estimate, Eq. (3.56)

Table 3.6. The z statistic and cumulative unit normal, $\Phi(z)$, used to compare the differences between pairs of the estimated prior parameters of Table 3.5. Variance estimates for $s^2(a)$ and $s^2(b)$ are based on the normality result of Eq. (3.55).

Grouping Comparison i=2 - i=1	a		b	
	z	$\Phi(z)$	z	$\Phi(z)$
By Manufacturer				
Fairbanks-GM	-0.261	0.397	-0.681	0.248
ALCO-GM	-0.251	0.401	-0.715	0.237
ALCO-Fairbanks	0.009	0.504	-0.061	0.476
Others-GM	0.884	0.812	1.685	0.954
Others-Fairbanks	0.641	0.755	1.039	0.851
Others-ALCO	0.685	0.753	1.041	0.851
By Number of Starts				
(26-50)-(0-25)	-0.951	0.171	-1.269	0.102
(51-100)-(0-25)	-0.798	0.212	-1.640	0.050
(51-100)-(26-50)	0.550	0.709	0.354	0.638
(>100)-(0-25)	-0.441	0.330	-1.222	0.111
(>100)-(26-50)	0.724	0.765	0.521	0.699
(>100)-(51-100)	0.321	0.626	0.256	0.601

Table 3.7. The z statistic and cumulative unit normal, $\Phi(z)$, used to compare the differences between pairs of the estimated prior parameters of Table 3.5. Variance estimates for $s^2(a)$ and $s^2(b)$ are based on the distribution-independent result for $s^2(\hat{\theta}_{ob})$, i.e., Eq. (3.56).

Grouping Comparison i=2 - i=1	a		b	
	z	$\Phi(z)$	z	$\Phi(z)$
By Manufacturer				
Fairbanks-GM	-0.315	0.376	-1.022	0.153
ALCO-GM	-0.311	0.378	-1.152	0.125
ALCO-Fairbanks	0.012	0.505	-0.096	0.462
Others-GM	0.952	0.829	1.907	0.972
Others-Fairbanks	0.849	0.802	1.571	0.942
Others-ALCO	0.864	0.806	1.705	0.956
By Number of Starts				
(26-50)-(0-25)	-1.256	0.105	-1.935	0.026
(51-100)-(0-25)	-0.996	0.160	-2.675	0.004
(51-100)-(26-50)	0.728	0.767	0.547	0.708
(>100)-(0-25)	-0.535	0.296	-1.931	0.027
(>100)-(26-50)	0.953	0.830	0.801	0.789
(>100)-(51-100)	0.402	0.656	0.411	0.659

If the estimators a and b are assumed to be uncorrelated, an estimate for the variance of g(p) can be obtained by the following propagation of error formula [9]*:

$$s^2[g(p)] = \left(\frac{\partial g}{\partial a}\right)^2 s^2(a) + \left(\frac{\partial g}{\partial b}\right)^2 s^2(b) \ . \qquad (3.61)$$

The first partial derivative of the prior distribution is given by

$$\frac{\partial g}{\partial a} = \left\{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} (1-p)^{(b-1)}\right\} p^{(a-1)} \ln p + \left\{\frac{p^{(a-1)}(1-p)^{(b-1)}}{\Gamma(a)\Gamma(b)}\right\} \frac{\partial \Gamma(a+b)}{\partial a}$$

$$+ \left\{\frac{\Gamma(a+b)\ p^{(a-1)}(1-p)^{(b-1)}}{\Gamma(b)}\right\} \frac{\partial[1/\Gamma(a)]}{\partial a} \ , \qquad (3.62)$$

with

$$\frac{\partial \Gamma(a+b)}{\partial a} = \psi(a+b)\Gamma(a+b) \ , \qquad (3.63)$$

and

$$\frac{\partial[1/\Gamma(a)]}{\partial a} = -\frac{\psi(a)}{\Gamma(a)} \ , \qquad (3.64)$$

where $\psi(a+b)$ and $\psi(a)$ are the digamma functions that can be calculated from a subroutine given in the BETA III computer code (given in Appendix I). Thus, this partial derivative can be simplified to

$$\frac{\partial g}{\partial a} = g(p)\ [\ln p + \psi(a+b) - \psi(a)] \ . \qquad (3.65)$$

The partial derivative with respect to b is given by

$$\frac{\partial g}{\partial b} = \left\{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} p^{(a-1)}\right\} [(1-p)^{(b-1)} \ln(1-p)] + \left\{\frac{p^{(a-1)}(1-p)^{(b-1)}}{\Gamma(a)\Gamma(b)}\right\} \frac{\partial \Gamma(a+b)}{\partial b}$$

$$+ \left\{\frac{\Gamma(a+b)\ p^{(a-1)}(1-p)^{(b-1)}}{\Gamma(a)}\right\} \frac{\partial[1/\Gamma(b)]}{\partial b} \ , \qquad (3.66)$$

---

*Equation (3.61) is based on a Taylor's series expansion. The second order and higher derivatives of g(p) with respect to a and b have been assumed to be small compared to the first order derivatives. Likewise the parameters a and b have been assumed to be uncorrelated. The inclusion of covariance is considered later.

1426 253

with

$$\frac{\partial\Gamma(a+b)}{\partial b} = \psi(a+b)\Gamma(a+b) \quad , \tag{3.67}$$

$$\frac{\partial[1/\Gamma(b)]}{\partial b} = - \frac{\psi(b)}{\Gamma(b)} \ . \tag{3.68}$$

Thus, this partial derivative becomes

$$\frac{\partial g}{\partial b} = g(p) \ [\mathit{ln}(1-p) + \psi(a+b) - \psi(b)] \ . \tag{3.69}$$

Hence, the estimate of the variance on g(p) is given by

$$s^2[g(p)] = [g(p)]^2 \ \{[\mathit{ln}p + \psi(a+b) - \psi(a)]^2 s^2(a)$$

$$+ [\mathit{ln}(1-p) + \psi(a+b) - \psi(b)]^2 s^2(a)\}. \tag{3.70}$$

A variance estimate can also be constructed in a similar manner for the cumulative distribution function (cdf) which utilizes the estimators for a and b. The cdf is given by

$$G(p) = \int_0^p g(t) \ dt, \tag{3.71}$$

or

$$G(p) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^p t^{(a-1)} \ (1-t)^{(b-1)} \ dt. \tag{3.72}$$

which is simply the *incomplete beta function* [8]. If the estimators, a and b, are again assumed, as a first approximation, to be uncorrelated random variables, the estimate for the variance of G(p) can be obtained in a similar fashion as was used in Eq. (3.61) for the pdf, i.e.,

$$s^2[G(p)] = \left(\frac{\partial G(p)}{\partial a}\right)^2 s^2(a) + \left(\frac{\partial G(p)}{\partial b}\right)^2 s^2(b) \ . \tag{3.73}$$

The partial derivative with respect to a is

$$\frac{\partial G}{\partial a} = \int_0^p \frac{\partial g(t)}{\partial a} dt \quad , \tag{3.74}$$

or

$$\frac{\partial G}{\partial a} = \int_0^p g(t) \ \mathit{ln}t \ dt + \int_0^p \psi(a+b) \ g(t) \ dt$$

$$- \int_0^p \psi(a) \ g(t) \ dt, \tag{3.75}$$

1426 255

or upon substitution for g

$$\frac{\partial G}{\partial a} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left[\int_0^p t^{(a-1)}(1-t)^{(b-1)} \ln t \; dt \right.$$

$$\left. + [\psi(a+b) - \psi(a)] \int_0^p t^{(a-1)}(1-t)^{(b-1)} \; dt \right] . \quad (3.76)$$

Similarly, the partial derivative with respect to b is

$$\frac{\partial G}{\partial b} = \int_0^p \frac{\partial g(t)}{\partial b} \; dt \quad (3.77)$$

or

$$\frac{\partial G}{\partial b} = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \left[\int_0^p t^{(a-1)}(1-t)^{(b-1)} \ln(1-t) \; dt \right.$$

$$\left. + [\psi(a+b) - \psi(b)] \int_0^p t^{(a-1)}(1-t)^{(b-1)} \; dt \right] \quad (3.78)$$

The integrals in Eqs. (3.76) and (3.78) must be calculated by numerical means although the second integral in both of these equations can be expressed in terms of the incomplete beta function (see Eq. (3.72)).

The above derivation for the variances of the prior density and cumulative distributions is based on a first order Taylor series expansion and on the assumption that the beta parameters a and b are uncorrelated. In the next chapter it is demonstrated that the estimated a and b parameters have a large positive covariance. If the covariance term is included in the derivation of Eqs. (3.61) and (3.73), these equations become

$$s^2[g(p)] = \left(\frac{\partial g}{\partial a}\right)^2 s^2(a) + \left(\frac{\partial g}{\partial b}\right)^2 s^2(b) + \frac{\partial g}{\partial a}\frac{\partial g}{\partial b} \operatorname{cov}(a,b) \quad 3.79)$$

and

$$s^2[G(p)] = \left(\frac{\partial G(p)}{\partial a}\right)^2 s^2(a) + \left(\frac{\partial G(p)}{\partial b}\right)^2 s^2(b) + \frac{\partial G(p)}{\partial a}\frac{\partial G(p)}{\partial b} \operatorname{cov}(a,b). \quad (3.80)$$

The expressions just obtained for the evaluation of the derivatives in the above expressions remain unchanged and hence to obtain approximate variances for the prior distribution, one needs only to have estimates of the variances and covariances of the beta prior parameters. With the matching moments technique, only estimates for $s^2(a)$ and $s^2(b)$ were obtained. However, with the maximum likelihood method, estimates for lower bounds of the covariance of a and b can be obtained from Eq. (3.32).

Often this bound is taken as an estimate of the actual covariance, and for the diesel engine data such an estimate was always found to be positive. With this estimate an additional term appears to be added to the variance estimates for the pdf and cdf if the first partials with respect to a and b are both positive or both negative (see Eq. (3.79) and (3.80) above); thus, the error bands around the estimated prior distribution would become even larger or further apart. However, it was found for the various diesel engine groupings that the covariance contribution generally decreased the variance estimates $s^2[g(p)]$ and $s^2[G(p)]$, although this decrease (compared to the results obtained without the covariance contribution) was usually quite small.

As an example, the beta prior density and cumulative distributions for all 25 diesel plants of Table 3.1 as estimated by the marginal maximum likelihood method are shown in Figs. 3.6 and 3.7 respectively. For this grouping of all the diesel data, the maximum likelihood estimates for the beta prior parameters are $\hat{a} = 1.0522$ and $\hat{b} = 19.902$ with variance estimates of $s^2(a) \approx 0.1763$, $s^2(b) \approx 81.67$ and $cov(a,b) \approx 3.273$. For both the density and cumulative distributions, the one sigma error bounds ($\pm s^2[g]$ or $\pm s^2[G]$) are also shown as calculated with and without the covariance contribution. It is seen from this example that the inclusion of the covariance contribution decreases the spread between the upper and lower error bound.

The error bounds for other subgroupings of the diesel engine data give similar results as for the 25 engines example, namely, the spread between the upper and lower error bounds are sufficiently large that the various estimated prior distributions tend to lie within the error bounds of each other. Such large uncertainty in the estimated prior distributions for the various groupings indicate there may be no significant differences between these estimated priors in the region where the bounds overlap.

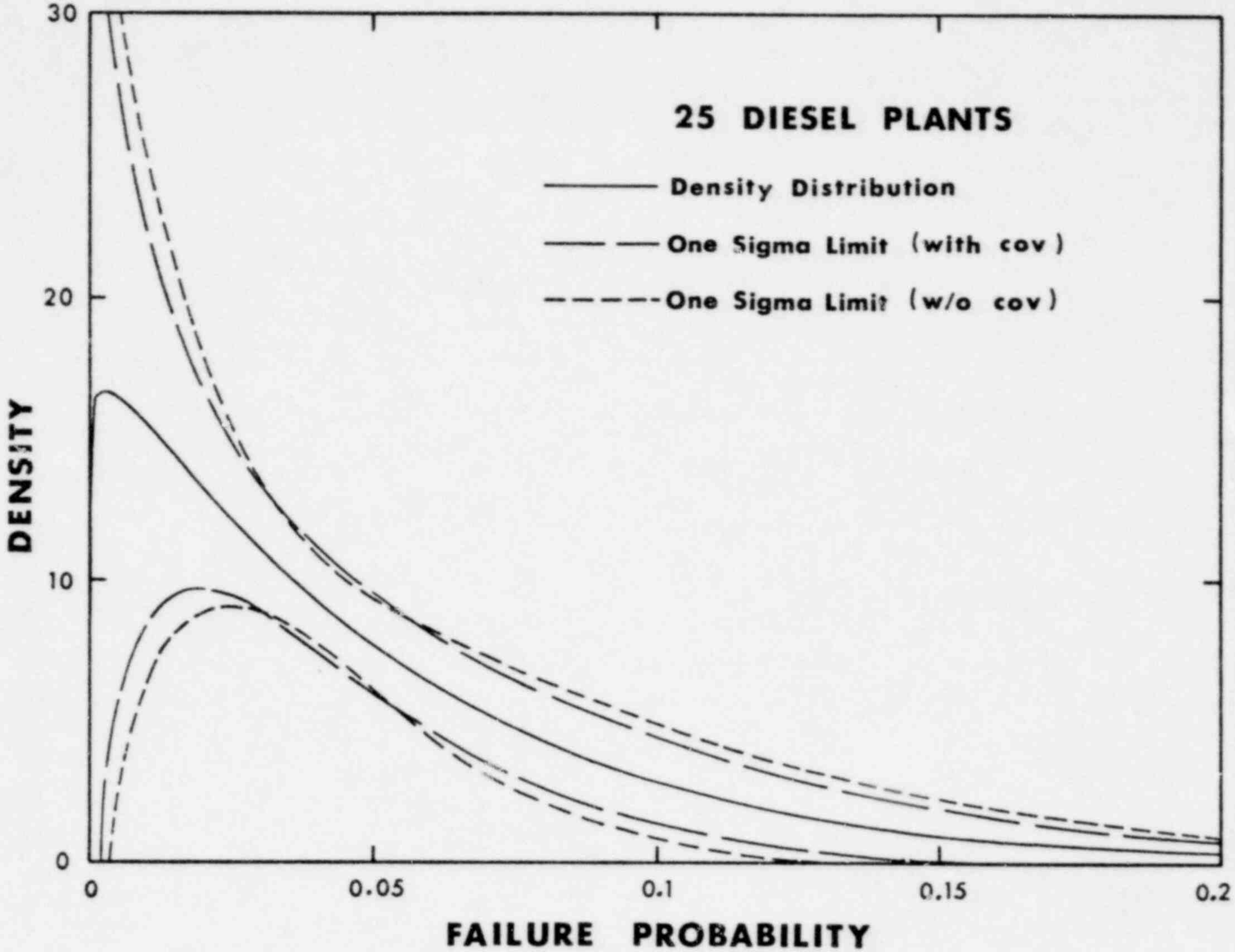


Fig. 3.6 The estimated prior density distribution with the estimated one sigma error bounds for all 25 diesel plants of Table 3.1. The prior parameters and their variances were estimated by the marginal-based maximum likelihood method which yielded $\hat{a}$=1.052, $\hat{b}$=19.90, var($\hat{a}$)=0.176, var($\hat{b}$)= 81.7, and cov($\hat{a}$,$\hat{b}$)=3.27 .

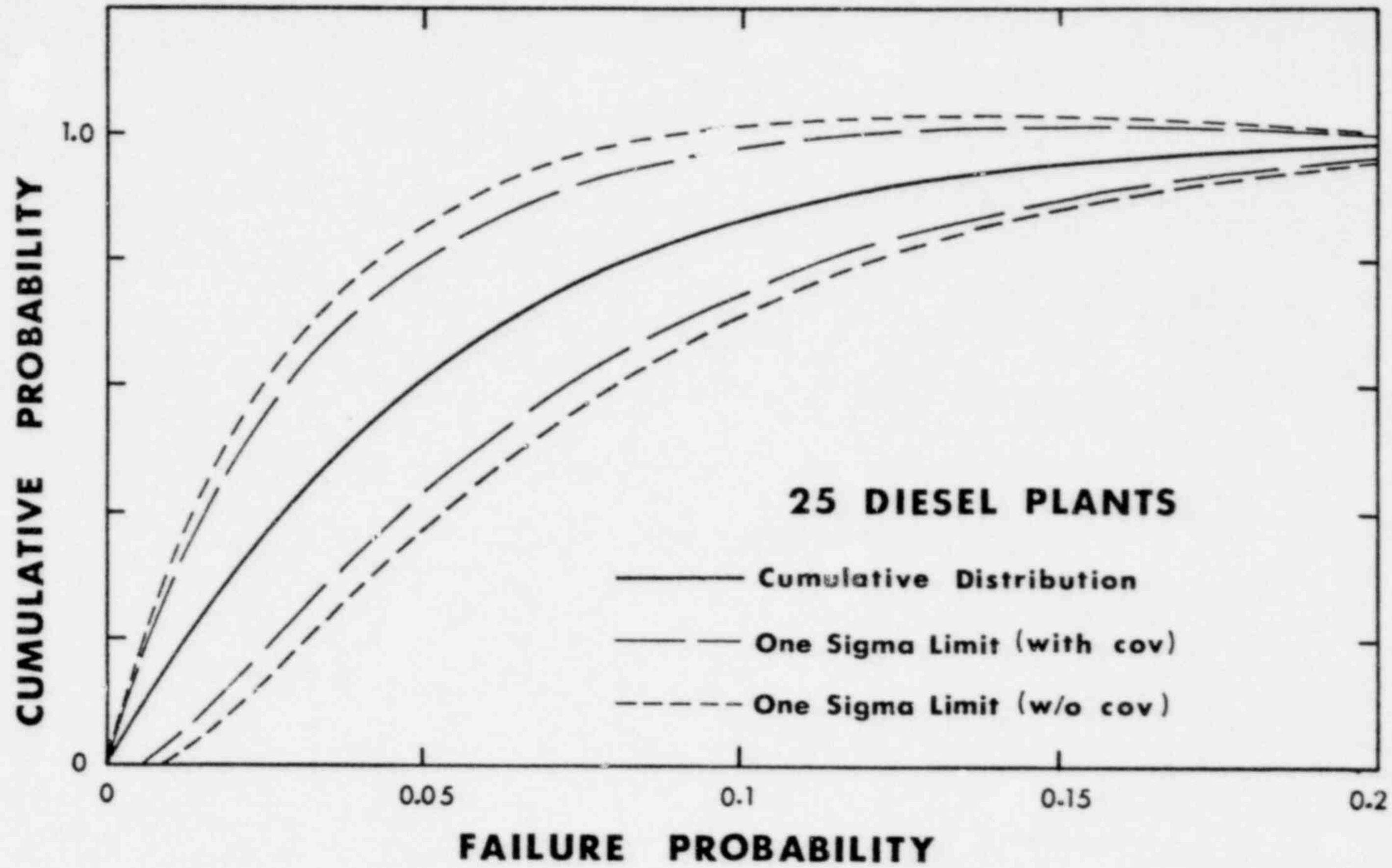


Fig. 3.7 The estimated prior cumulative distribution with the estimated one sigma error bounds for all 25 diesel plants of Table 3.1. The beta prior parameters and their variances were estimated by the marginal maximum likelihood method which yielded $\hat{a}$=1.052, $\hat{b}$=19.90, var($\hat{a}$)=0.176, var($\hat{b}$)= 81.7, and cov($\hat{a}$,$\hat{b}$)=3.27 .

4. SIMULATION STUDY OF PRIOR ESTIMATION TECHNIQUES

From the Bayesian analysis of the diesel engine failure data, the beta prior distributions, whose parameters were estimated from observed data, have modes in the region of small failure probabilities and are highly skewed away from high failure probabilities. Such mode behavior and skewness is expected for components which are designed to have low failure probabilities. However, the diesel data with which the early phase of this study was concerned have typically small sample sizes. Thus the question arises of biasedness and variance in the parameter estimates used for the beta priors and of the effects on the subsequent prediction of failure probability. To determine which of the four parameter estimation techniques discussed in the previous chapter is the most "conservative" or yields parameters closest to the true values, it is necessary to determine the distribution of the parameter estimates for each method. Consequently the objective of the study described here was to determine the properties of each of the four parameter estimation techniques. For such an investigation multiple sets of failure data in small sample sizes were generated randomly from known beta prior or marginal distributions. With these simulated failure data the distributions of the prior parameter estimates could be determined numerically for each estimation technique and from these distributions many properties of the four estimation techniques can be investigated.

4.1 Generation of Simulated Failure Data

To determine the distributional properties of each parameter estimation technique by numerical simulation, it is first necessary to generate a large number of failure data pairs (k failures in n tries) in which the number of failures k are distributed according to a known beta-binominal distribution with parameters a and b, i.e., according to the marginal distribution

$$h(k|n,a,b) = \binom{n}{k} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+k)\Gamma(b+n-k)}{\Gamma(a+b+n)} . \qquad (4.1)$$

Thus to generate the simulated failure data, the number of demands, n, is first selected randomly from a uniform distribution between n_1 and n_2. The number of demands n was allowed to vary in this manner to simulate better the type of failure data encountered in actual practice (see

Table 3.1). Then with n determined, and the beta parameters a and b fixed, the number of failures, k, is chosen from the above beta-binomial distribution. This two step process is repeated until a sufficient number of data pairs have been generated. Explicit details for each step are as follows:

For each step a random number, u, from a distribution, which was uniformly distributed between 0 and 1, was generated from the routine RANDU [12] and which subsequently was used to generate an n or k value. To select n, which for this study was assumed to be uniformly distributed between two positive integers n_1 and n_2, the following algorithm was used:

$$n = \begin{cases} n_1 + \text{integer}[u/p] \ , & u \neq p \\ n_1 + \text{integer}[u/p] - 1, & u = p \end{cases} \tag{4.2}$$

where $p \equiv (n_2 - n_1 - 1)^{-1}$ which is simply the probability of obtaining any integer between n_1 and n_2 inclusively, i.e., $n_1 \leq n \leq n_2$. The above algorithm is equivalent to

$$n = \begin{cases} n_1 & 0 \leq u \leq p \\ n_1+1 & p < u \leq 2p \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ n_1+i & ip < u \leq (i+1)p \ . \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ n_2 & 1-p < u \leq 1 \end{cases} \tag{4.3}$$

Once the number of failures, n, had been selected a new random number, u, was generated and used with the *inverse transformation* technique to obtain a value for k from the cumulative distribution of h(k), i.e., from

$$F(k) \equiv \sum_{m=0}^{k} h(m|n,a,b) \ , \quad k = 0,1 \ldots n \ . \tag{4.4}$$

The value of k selected is the minimum integer for which $u \leq F(k)$, or equivalently,

1426 260

$$k = \begin{cases} 0 & 0 \leq u \leq F(0) \\ 1 & F(0) < u \leq F(1) \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ i & F(i-1) < u \leq F(i) \; . \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ n & F(n-1) < u \leq F(n) = 1 \end{cases} \qquad (4.5)$$

In essence this method for changing a random variable, u, with a uniform distribution on (0,1) to a random variable, k, distributed according to a beta-binomial on (0,n) requires the sequential evaluation of the cumulative distribution, F(k). The use of Eq. (4.4) for each evaluation would be very time consuming if large amounts of simulated failure data were to be generated. However, considerable computational effort may be saved in the sequential evaluation of F by using the following recursion relation

$$F(k+1) = F(k) + h(k+1|n,a,b) \qquad (4.6)$$

with

$$h(k+1|n,a,b) = h(k|n,a,b)\,\frac{(a+k)\;(n-k)}{(b+n-k-1)(k+1)} \; . \qquad (4.7)$$

For situations involving beta parameters which yield a prior distribution with a low failure probability, (i.e., for which the above inverse technique would be expected to yield small values of k), the sequential search is best begun at k=0. Similarly if a prior corresponding to large expected values of k is used, then the sequential search is best begun at k=n. More generally, to minimize the length of the sequential search, the search should be begun near the mean of the beta-binomial distribution of interest. However, this optimal search method requires that the integer nearest to the mean and the cumulative distribution at that integer be initially evaluated and stored for all possible values of n. This search algorithm is outlined in Table 4.1.

Table 4.1. Algorithm for Optimal Calculation of Number of Failures, k, by the Inverse Transformation Technique.

Part I: Selection of Starting Values for Sequential Search

1. Calculate means, μ_i, of beta-binomials for all possible n_i (i.e., for $n_i = n_1, n_1+1, \ldots, n_2$).
2. Round off means to nearest integer, M_i
3. Calculate $F(M_i)$ and $h(M_i | n_i, a, b)$
4. Store values of M_i, $F(M_i)$ and $h(M_i)$ in a vector to be used as starting points in sequential search.

Part II: Sequential Search to Calculate k for Given n_i

1. Generate u from a uniform distribution on (0,1) by RANDU
2. If $u = F(M_i)$, then $k = M_i$
3. Otherwise, set $K = M_i$, $h(K) = h(M_i)$ and $F(K)=F(M_i)$
4. If $u<F(M_i)$ go to step 6; otherwise go to step 5
5. Compute:

$$h(K+1) = h(K)\,\frac{(a+K)(n_i-K)}{(b+n_i-K-1)(K+1)}$$

$$F(K+1) = F(K) + h(K+1)$$

If $u \leq F(K+1)$, then $k = K+1$ and exit; otherwise set $K=K+1$ and go back to beginning of step 5.

6. Compute

$$F(K-1) = F(K) - h(K)$$

If $u>F(K-1)$, then $k=K-1$ and exit; otherwise calculate,

$$h(K-1) = h(K)\,\frac{K\,(n_i-K+b)}{(K-1+a)(n_i-K+1)}$$

set $K=K-1$, and go back to beginning of step 6.

1426 262

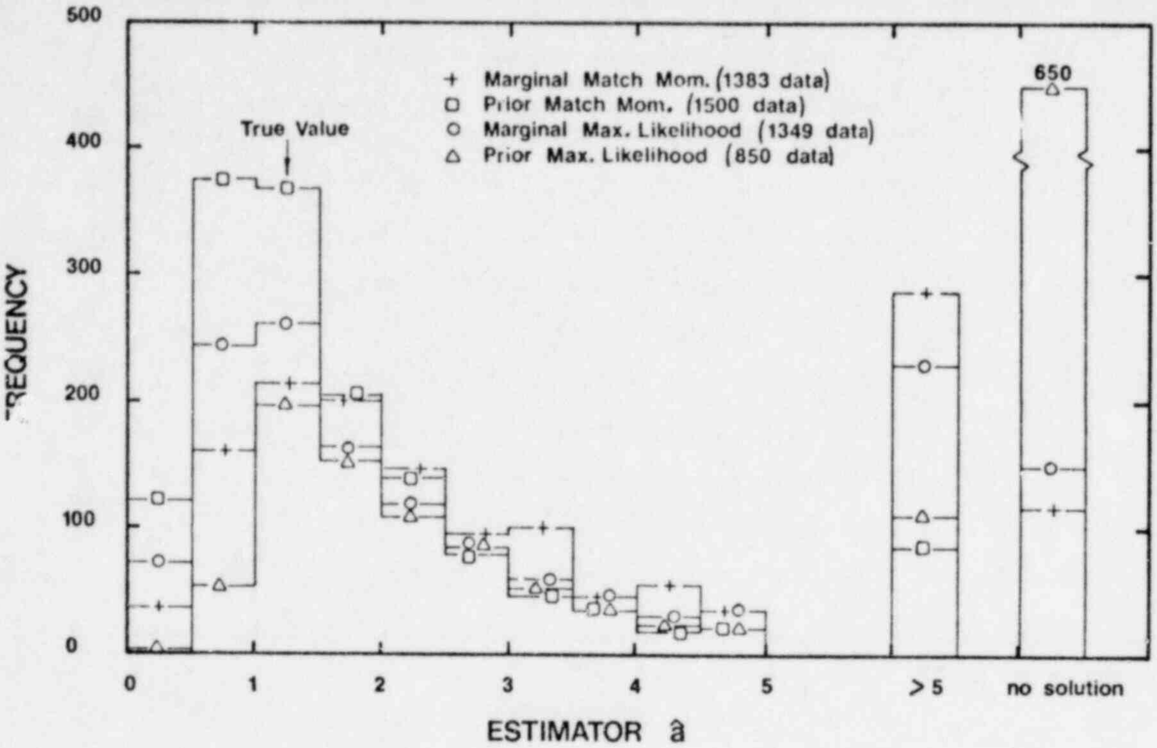
4.2 Distribution of Prior Parameter Estimates

To investigate how the estimates of the beta prior parameters are distributed, simulated failure data were analyzed by the four empirical estimation techniques described in Chapter 3. Since this study was concerned primarily with low failure probability events, a beta prior with parameters of a=1.2 and b=23 was used as the basis for generating the simulation failure data*. The number of starts, n_i, was randomly selected from a uniform distribution between 30 (n_1) and 300 (n_2), inclusively. For a given n_i, the number of failures, k_i, was selected randomly from a beta-binomial (marginal) distribution using the technique described in Section 4.1. In all, 1500 samples of size 5 (i.e., five k_i and n_i pairs), 10, and 20 were generated. Additionally 500 samples of size 50 were computed.

With these simulated failure data, estimates of the parameters a and b were calculated and compared to the true values of a=1.2 and b=23. The frequency distribution of the estimates $\hat{a}$ and $\hat{b}$ as calculated by the four estimation techniques for the four sample sizes are shown in Figs. 4.1 through 4.4. All these frequency distributions exhibit several common features. In particular all estimation methods exhibit a slowing decaying tail at high values. The mean of the distribution is always on the high side of the true value. For small sample sizes (N≤10) there were obtained an appreciable number of inordinately large estimators, or *outliers*, especially by the two most complicated estimation techniques--the marginal maximum likelihood method and the marginal matching moments method. Furthermore, only the simplest estimation method, the prior matching moments method, always yielded results for all samples regardless of size. For small sample sizes (N≤5) the marginal matching moments and marginal maximum likelihood methods often yielded no parameter estimates, while for large sample sizes the prior maximum likelihood method was unable to give an estimate as a result of at least one k_i=0 in the sample (a likely occurrence for the low failure probability case studied). In Table 4.2 the observed success history for each of the four methods is given.

---

*These particular values of a and b are the marginal maximum likelihood estimates for the failure data of the 13 GM diesel engines in Table 3.1.



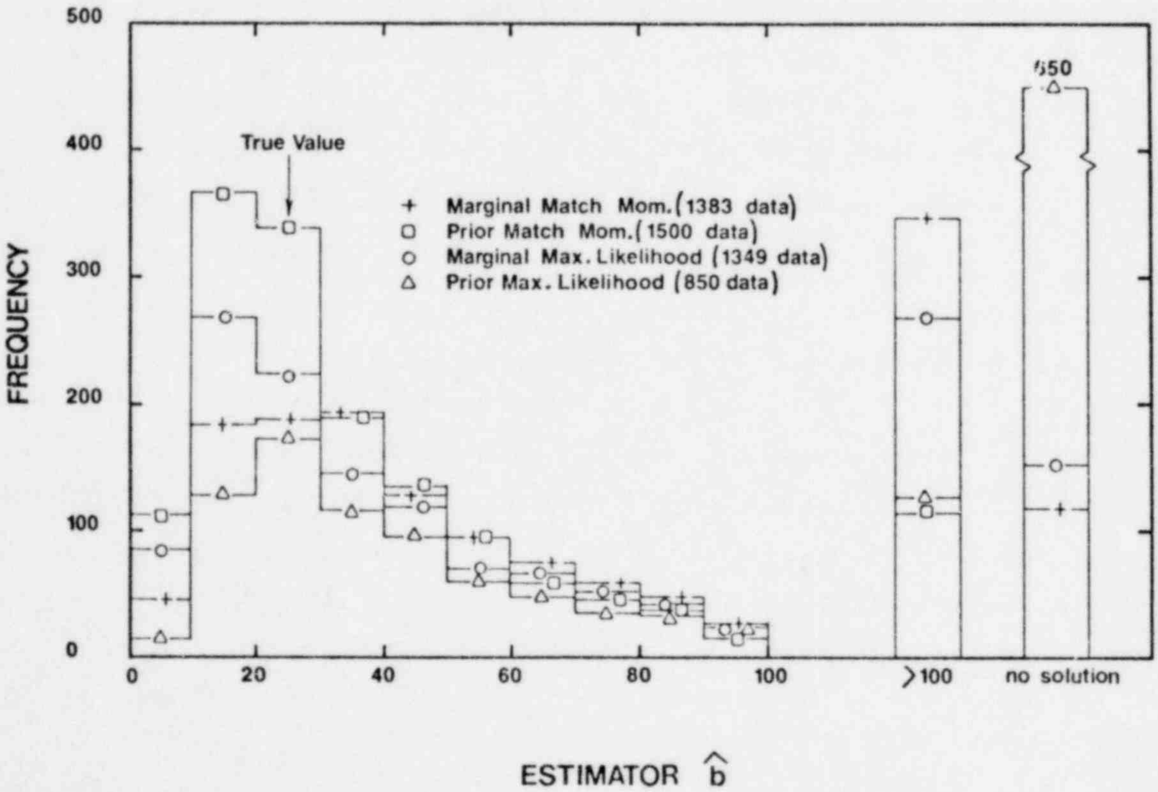


Fig. 4.1 Distribution of beta parameter estimators for samples of size N=5.

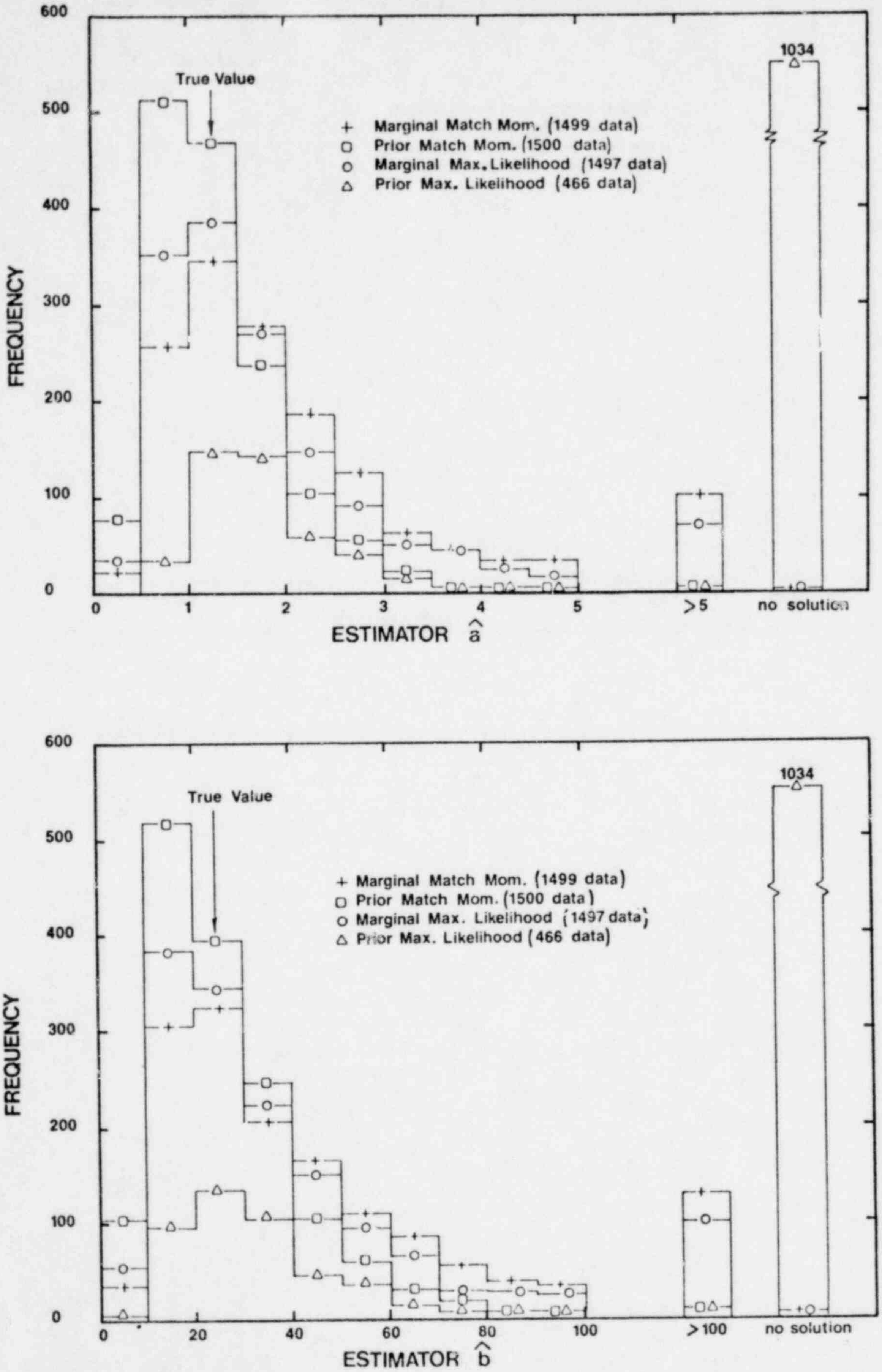
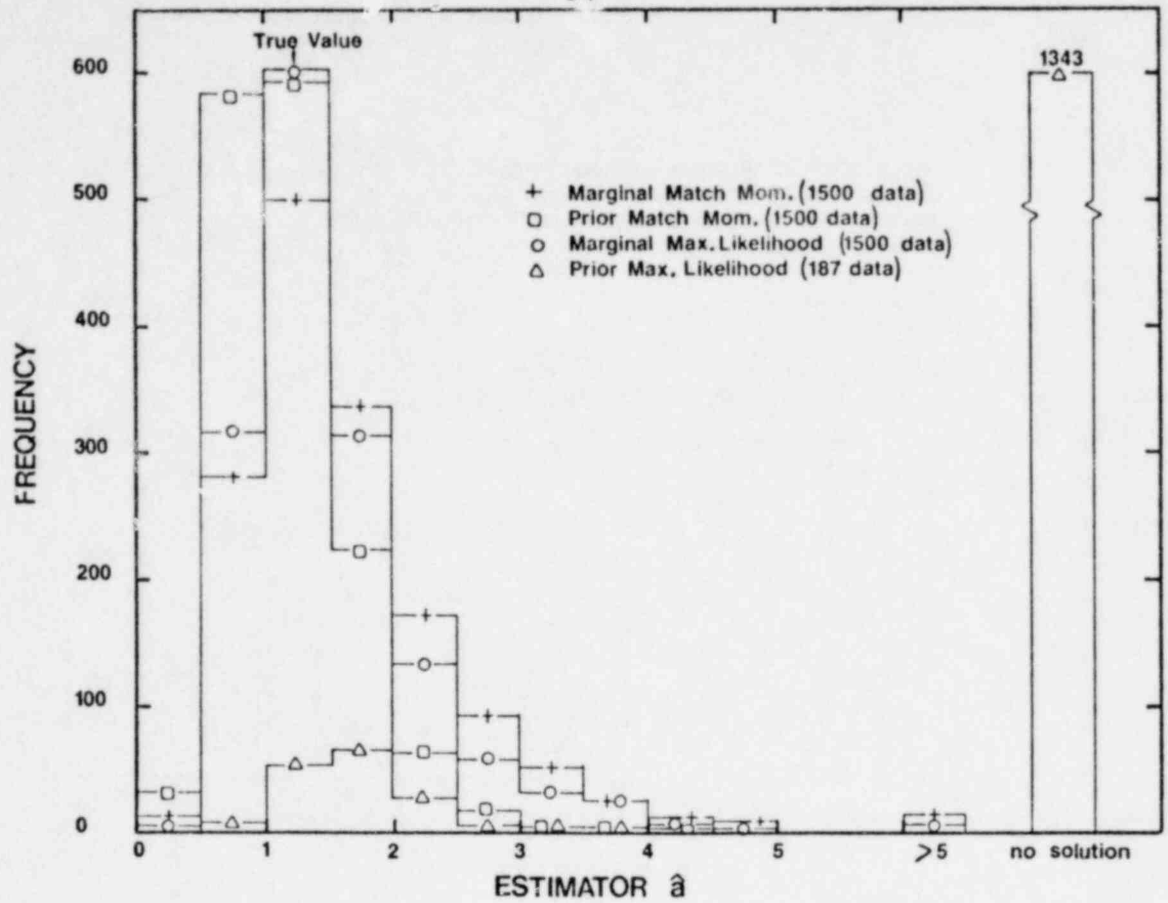
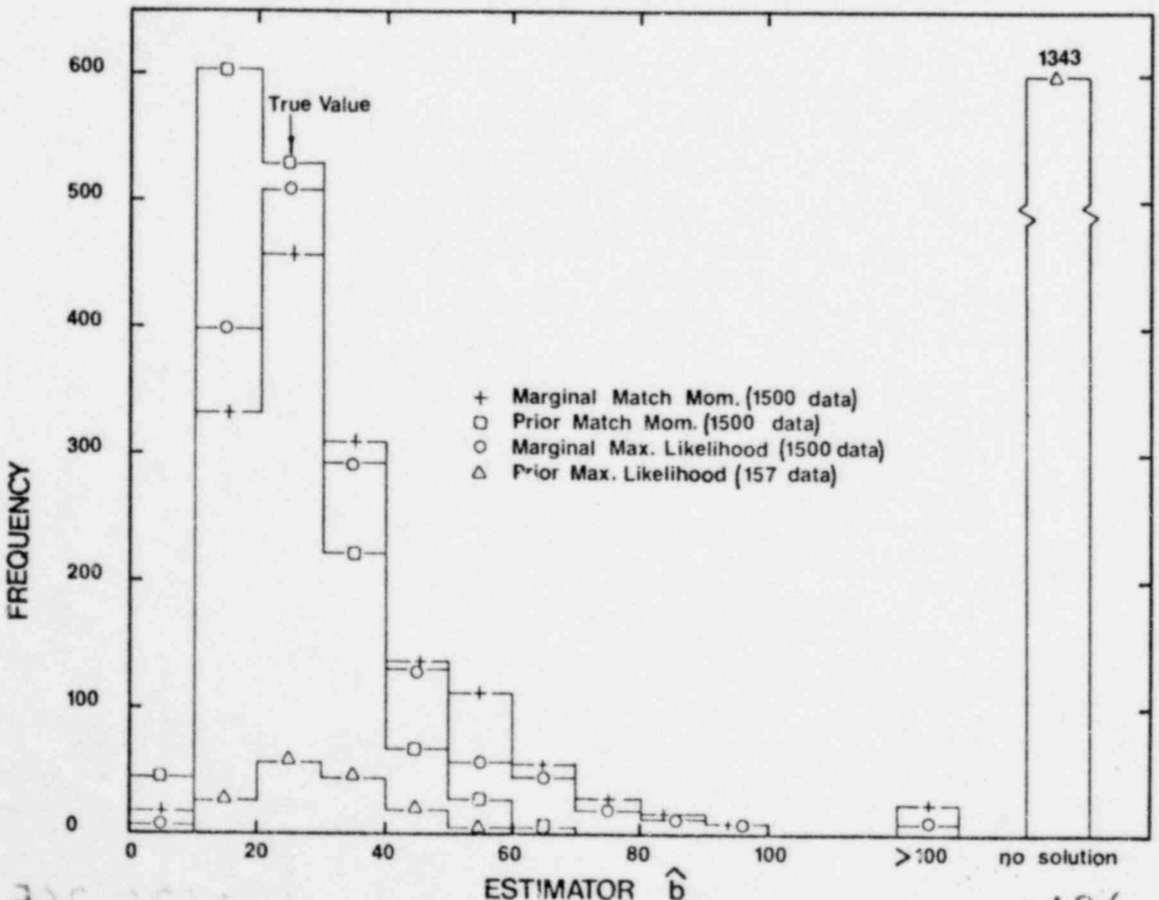


Fig. 4.2 Distribution of beta parameter estimators for samples of size N=10.

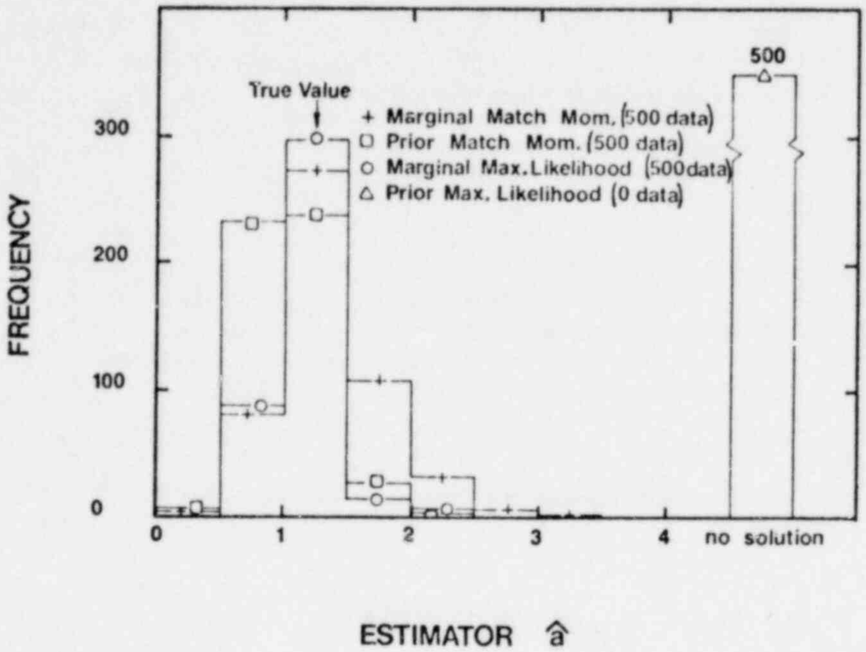




1426 265

1426 266

Fig. 4.3 Distribution of beta parameter estimators for samples of size N=20.



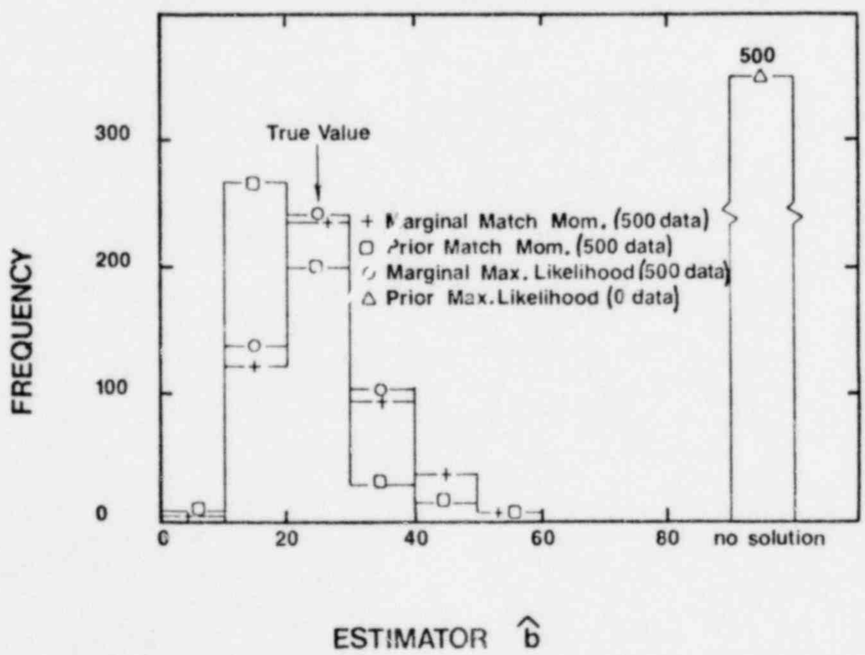


Fig. 4.4 Distribution of beta parameter estimators for samples of size N=50.

1426 267

Table 4.2 Number of successful solutions and failures for prior parameter estimates from the simulation failure data for the four estimation techniques.

Sample Size	Marginal Matching Mom.			Prior Matching Mom.		
	Sol.	No-Sol.	% Success	Sol.	No-Sol.	% Success
5	1383	117	92.20	1500	0	100.0
10	1499	1	99.93	1500	0	100.0
20	1500	0	100.0	1500	0	100.0
50	500	0	100.0	500	0	100.0

Sample Size	Marginal Max. Likelihood			Prior Max. Likelihood		
	Sol.	No-Sol.	% Success	Sol.	No-Sol.	% Success
5	1349	151	89.93	850	650	56.67
10	1497	3	99.80	466	1034	31.07
20	1500	0	100.0	157	1343	10.47
50	500	0	100.0	0	500	0.00

Table 4.3 displays some simulation data samples for which no parameter estimates could be obtained by three of the estimation techniques. No noticeable features about these particular data seem to distinguish them from other data samples for which the estimation methods yielded solutions. A test to screen small data samples to determine whether a particular sample permits a solution by each method has not been found.

4.2.1 Bias and Variance of Prior Parameter Estimates

The degree of bias inherent in any parameter estimation technique is often of concern. The *bias* of an estimator, $\hat{\theta}$, is defined as

$$\text{Bias} \equiv E[\hat{\theta}-\theta] = \bar{\theta}-\theta \qquad (4.8)$$

where θ is the true value of the parameter (e.g., a or b) and $\bar{\theta}$ is the mean of of the estimators. All of the estimation techniques investigated in this study were found to yield biased estimates of the prior parameters, especially for small sample sizes.

In the estimation of the mean or bias of the estimators from the empirically derived distributions of Figs. 4.1-4.4, the treatment of outliers present some difficulties. For the estimation techniques based on the marginal distribution, estimates of a and b would occasionally be obtained which were orders of magnitude greater than the true values. In this section those outlier estimates which were greater than one hundred times the true value were classified together with those samples which yielded no solution and hence were not used in the computation of statistics from the distribution of estimates. If those outlier values were included, values of bias and variance of the estimator distributions would be determined principally by the outlier values. For example, the distribution for N=5 of Fig. 4.1 for a estimated by the marginal maximum likelihood method yields a mean $\bar{a}$=7.23 and a variance var($\hat{a}$)=2581 if all data are used, while if the outliers ($\hat{a}$>100a) are suppressed, a mean $\bar{a}$=3.79 and a variance var($\hat{a}$)=59.5 results (the true value of a is 1.2). Unless explicitly specified to the contrary, all outliers are suppressed in the subsequent analyses of the distributions of $\hat{a}$ and $\hat{b}$.

In Table 4.4 the results are presented of the bias of the beta parameter estimators for each estimation method considered. The variation of

Table 4.3 Simulated failure data $\binom{n_i}{k_i}$ from a beta-binomial (a=1.2, b=23) for which the marginal-based estimation methods yielded no solution.

Sample Size N=5:

1. Data for which marginal maximum likelihood and marginal matching moments give no solution:

$$\begin{pmatrix}129 & 235 & 290 & 30 & 97\\ 8 & 8 & 14 & 1 & 5\end{pmatrix} \quad \begin{pmatrix}38 & 207 & 87 & 114 & 108\\ 1 & 8 & 3 & 3 & 5\end{pmatrix}$$

$$\begin{pmatrix}110 & 218 & 123 & 282 & 226\\ 3 & 10 & 7 & 11 & 13\end{pmatrix} \quad \begin{pmatrix}237 & 74 & 287 & 245 & 147\\ 3 & 2 & 1 & 4 & 2\end{pmatrix}$$

$$\begin{pmatrix}113 & 64 & 81 & 56 & 145\\ 4 & 2 & 4 & 6 & 7\end{pmatrix} \quad \begin{pmatrix}49 & 154 & 155 & 48 & 264\\ 3 & 7 & 4 & 0 & 9\end{pmatrix}$$

$$\begin{pmatrix}64 & 65 & 62 & 197 & 166\\ 0 & 2 & 1 & 2 & 5\end{pmatrix} \quad \begin{pmatrix}274 & 60 & 250 & 197 & 60\\ 11 & 5 & 14 & 15 & 4\end{pmatrix}$$

$$\begin{pmatrix}84 & 33 & 266 & 242 & 133\\ 7 & 3 & 20 & 17 & 7\end{pmatrix} \quad \begin{pmatrix}215 & 221 & 76 & 32 & 70\\ 4 & 7 & 1 & 2 & 3\end{pmatrix}$$

2. Data for which only the marginal matching moments method failed:

$$\begin{pmatrix}92 & 263 & 225 & 71 & 146\\ 3 & 18 & 11 & 2 & 4\end{pmatrix} \quad \begin{pmatrix}193 & 192 & 292 & 277 & 264\\ 11 & 8 & 22 & 11 & 12\end{pmatrix}$$

$$\begin{pmatrix}85 & 87 & 123 & 269 & 63\\ 3 & 6 & 7 & 6 & 3\end{pmatrix} \quad \begin{pmatrix}253 & 32 & 39 & 150 & 97\\ 2 & 2 & 2 & 10 & 4\end{pmatrix}$$

$$\begin{pmatrix}38 & 128 & 46 & 175 & 223\\ 0 & 2 & 1 & 1 & 7\end{pmatrix} \quad \begin{pmatrix}246 & 249 & 227 & 167 & 255\\ 12 & 13 & 4 & 8 & 14\end{pmatrix}$$

$$\begin{pmatrix}166 & 59 & 61 & 104 & 150\\ 5 & 4 & 3 & 6 & 13\end{pmatrix} \quad \begin{pmatrix}208 & 60 & 33 & 253 & 151\\ 6 & 2 & 1 & 7 & 10\end{pmatrix}$$

$$\begin{pmatrix}237 & 67 & 77 & 227 & 47\\ 2 & 0 & 1 & 7 & 1\end{pmatrix} \quad \begin{pmatrix}213 & 89 & 209 & 248 & 122\\ 9 & 3 & 5 & 3 & 2\end{pmatrix}$$

3. Data for which only the marginal maximum likelihood method failed:

$$\begin{pmatrix}100 & 87 & 253 & 181 & 97\\ 7 & 3 & 22 & 19 & 5\end{pmatrix} \quad \begin{pmatrix}187 & 151 & 50 & 45 & 272\\ 3 & 2 & 4 & 0 & 5\end{pmatrix}$$

$$\begin{pmatrix}271 & 43 & 253 & 273 & 169\\ 10 & 3 & 10 & 7 & 1\end{pmatrix} \quad \begin{pmatrix}98 & 101 & 60 & 229 & 81\\ 7 & 7 & 9 & 18 & 11\end{pmatrix}$$

$$\begin{pmatrix}279 & 206 & 59 & 64 & 122\\ 8 & 8 & 0 & 0 & 3\end{pmatrix} \quad \begin{pmatrix}137 & 80 & 123 & 88 & 45\\ 11 & 0 & 0 & 8 & 5\end{pmatrix}$$

$$\begin{pmatrix}144 & 284 & 220 & 207 & 277\\ 1 & 11 & 8 & 7 & 5\end{pmatrix} \quad \begin{pmatrix}31 & 205 & 68 & 48 & 255\\ 0 & 18 & 2 & 6 & 22\end{pmatrix}$$

$$\begin{pmatrix}238 & 237 & 35 & 39 & 261\\ 8 & 8 & 0 & 4 & 6\end{pmatrix} \quad \begin{pmatrix}289 & 37 & 280 & 91 & 204\\ 5 & 2 & 5 & 1 & 8\end{pmatrix}$$

Sample Size N=10

1. Sample for which marginal matching moments method found no solution:

$$\begin{pmatrix}225 & 85 & 73 & 71 & 238 & 167 & 245 & 91 & 187 & 67\\ 7 & 1 & 2 & 0 & 7 & 4 & 4 & 1 & 0 & 1\end{pmatrix}$$

2. Samples for which marginal maximum likelihood method failed:

$$\begin{pmatrix}40 & 111 & 108 & 273 & 217 & 207 & 254 & 31 & 284 & 108\\ 0 & 6 & 8 & 14 & 8 & 10 & 7 & 0 & 14 & 1\end{pmatrix}$$

$$\begin{pmatrix}36 & 206 & 254 & 97 & 95 & 99 & 276 & 233 & 253 & 281\\ 1 & 10 & 8 & 3 & 7 & 10 & 15 & 11 & 11 & 12\end{pmatrix}$$

$$\begin{pmatrix}152 & 86 & 85 & 206 & 75 & 88 & 267 & 279 & 111 & 229\\ 10 & 8 & 4 & 12 & 11 & 4 & 18 & 11 & 3 & 16\end{pmatrix}$$

the bias in $\hat{a}$ and $\hat{b}$ with sample size is shown in Fig. 4.5. Notice that as the sample size increases, the bias of the estimators decreases towards zero as would be expected. However, from Fig. 4.5 all of the methods except the simplest method - the prior matching moments - always yield a positive bias. The prior matching moments method has the smallest bias of all four methods and actually changes sign for sample sizes of about 20 or larger.

The bias results for the prior-based maximum likelihood method, however, are relatively poor for the large sample sizes since, for the assumed prior beta, many of the simulated samples contain at least one $k_i=0$ which makes this estimation method fail (see Table 4.1). Since all the samples which preclude estimation of the prior parameters with this method have at least one $k_i=0$, it can be expected that the estimators may inherently contain a bias. In fact, from Fig. 4.5 it is seen that the bias appears to level off at some small positive value as the sample size increases.

The mean values of $\hat{a}$ and $\hat{b}$ for the various sample sizes and estimation techniques are readily obtained from Table 4.4 by adding to the tabulated values of bias the true value of the parameter, $a=1.2$, or $b=23$. The variance and covariance of the distribution of the estimates are presented in Table 4.5. As would be expected, the variances and covariance for all estimation techniques decrease as the sample size increases. The minimum variance for a given sample size was always obtained with the simplest estimation technique, i.e., with the prior matching moment method. Those estimation methods based on the marginal distribution always yielded the largest variances, a result of the slowly decaying tail of the distributions for $\hat{a}$ and $\hat{b}$ and of the presence of unsuppressed outliers which were more prevalent with these methods.

The covariance of $\hat{a}$ and $\hat{b}$ were always observed to be positive which indicates that large values of $\hat{a}$ are associated with large values of $\hat{b}$. In fact, the outliers were observed to have just this property, namely that a sample which produced a large value for $\hat{a}$ also generated a large value for $\hat{b}$.

Table 4.4 The bias or deviation of mean of estimators from true parameters [a=1.2, b=23.0]. Each data set consists of 500 simulation samples.

Sample Size (N)	Data Set No.	Marg. Match. Mom.		Prior Match. Mom.		Marg. Max. Like.		Prior Max. Like.	
		$\bar{a}-a$	$\bar{b}-b$	$\bar{a}-a$	$\bar{b}-b$	$\bar{a}-a$	$\bar{b}-b$	$\bar{a}-a$	$\bar{b}-b$
5	1	3.24	76.2	0.566	16.5	2.76	63.1	1.49	2.0
	2	2.68	61.8	0.739	21.9	2.08	30.6	2.05	47.01
	3	3.30	72.8	0.835	2.41	2.91	68.4	1.95	45.6
10	1	1.20	26.3	0.124	3.72	0.887	21.2	0.673	10.6
	2	1.12	26.5	0.104	3.72	0.772	19.2	0.691	11.8
	3	1.38	33.1	0.125	4.82	0.872	23.5	0.660	13.0
20	1	0.471	10.2	-0.0238	0.0602	0.325	7.37	0.479	6.07
	2	0.412	9.50	-0.0574	0.299	0.268	6.71	0.439	6.48
	3	0.568	13.4	0.0118	1.44	0.373	9.34	0.491	7.38
50	1	0.164	3.40	-0.142	-2.58	0.100	2.22	*	*

*Method always failed for sample size N=50 since each sample contained at least one k_i=0.

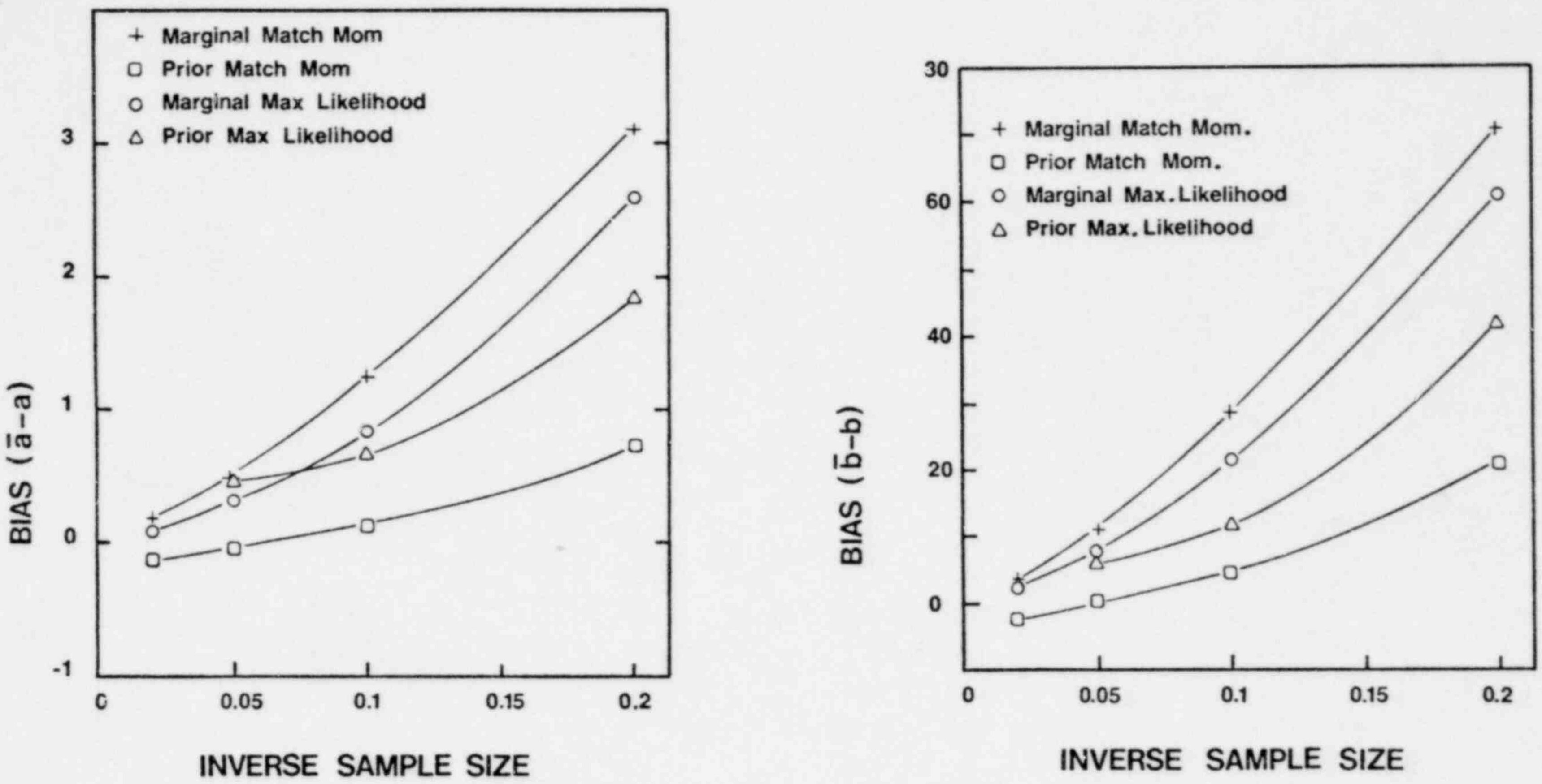


Fig. 4.5 Variation of the bias of the beta parameter estimators with sample size for the different estimation techniques. True values of the beta parameters are a=1.2 and b=23.

Table 4.5 Variances and covariance of parameter estimators for different sample sizes and estimation techniques. True beta parameter values are a=1.2 and b=23.0. Results for marginal-based methods are presented with and without outliers ($\hat{a}$>100a or $\hat{b}$>100b) included.

Sample Size	Prior Matching Moments var ($\hat{a}$)	Prior Matching Moments var ($\hat{b}$)	Prior Matching Moments cov ($\hat{a}$,$\hat{b}$)	Prior Maximum Likelihood var ($\hat{a}$)	Prior Maximum Likelihood var ($\hat{b}$)	Prior Maximum Likelihood cov ($\hat{a}$,$\hat{b}$)
5	4.42	3.79 (3)	1.03 (2)	9.29	8.44 (3)	2.23 (2)
10	5.50 (-1)*	2.86 (2)	1.01 (1)	8.40 (-1)	4.39 (2)	1.63 (1)
20	2.11 (-1)	9.97 (1)	3.79	2.02 (-1)	1.08 (2)	3.81
50	6.72 (-2)	3.05 (1)	1.23	-	-	-

*read as 5.50×10^{-1}

Sample Size	Marg. Match Mom. w/o Outliers var ($\hat{a}$)	Marg. Match Mom. w/o Outliers var ($\hat{b}$)	Marg. Match Mom. w/o Outliers cov ($\hat{a}$,$\hat{b}$)	Marg. Match. Mom. with Outliers var ($\hat{a}$)	Marg. Match. Mom. with Outliers var ($\hat{b}$)	Marg. Match. Mom. with Outliers cov ($\hat{a}$,$\hat{b}$)
5	5.20 (1)	2.50 (4)	9.90 (2)	8.15 (4)	3.41 (7)	1.64 (6)
10	1.23 (1)	5.76 (3)	2.51 (2)	2.69 (1)	1.15 (4)	5.40 (2)
20	8.01 (-1)	4.49 (2)	1.69 (1)	8.01 (-1)	4.49 (2)	1.69 (1)
50	1.75 (-1)	8.13 (1)	3.44	1.75 (-1)	8.13 (1)	3.44

Sample Size	Marg. Match. Like. w/o Outliers var ($\hat{a}$)	Marg. Match. Like. w/o Outliers var ($\hat{b}$)	Marg. Match. Like. w/o Outliers cov ($\hat{a}$,$\hat{b}$)	Marg. Max. Like. with Outliers var ($\hat{a}$)	Marg. Max. Like. with Outliers var ($\hat{b}$)	Marg. Max. Like. with Outliers cov ($\hat{a}$,$\hat{b}$)
5	5.94 (1)	2.74 (4)	1.15 (3)	2.58 (3)	6.39 (6)	1.18 (5)
10	5.60	4.09 (3)	1.37 (2)	2.89 (3)	1.08 (6)	5.59 (4)
20	5.70 (-1)	3.39 (2)	1.22 (1)	5.70 (-1)	3.39 (2)	1.22 (1)
50	1.14 (-1)	5.83 (1)	2.27	1.14 (-1)	5.83 (1)	2.27

4.2.2 Mean Squared Error of Estimators

For safety analyses the mean square error of an estimator is generally of concern. Although a particular method may have a small bias, the variance of the estimates may be quite large and hence the analysis of an individual sample could lead to parameter estimates which are significantly different from the true values. For safety considerations in which only a few samples are to be analyzed it is important that the mean square error of the estimates be small even if the estimates are slightly biased.

For the simulated data the mean squared error (MSE) is estimated as

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (\hat{\theta}_i - \theta)^2 \qquad (4.9)$$

where $\hat{\theta}_i$ represents the estimate $\hat{a}$ or $\hat{b}$ and θ represents the true value. From this equation, it is seen that outliers (i.e., estimates which are far removed from the true value) will change the value of the mean squared error greatly, and that estimates close to the true value have little influence. From the distributions of $\hat{a}$ and $\hat{b}$ shown in Figs. 4.1-4.4, it is seen that there are typically several outliers produced by the marginal-based estimation methods, especially for small sample sizes. To compare the mean squared error for the different estimation methods, these outliers were suppressed by ignoring those values of $\hat{a}$ or $\hat{b}$ which were more than one hundred times the true values of a and b. The results of the mean squared error analysis for the simulated failure data are presented in Table 4.5 and in Fig. 4.6.

From these results it is seen that for small or moderate sample sizes ($N \le 50$) the prior matching moment estimation techniques yields the lowest mean squared error. The two estimation methods based on the marginal distribution produce the poorest results, i.e., the largest mean squared errors. These large errors are a direct result of the occasional high estimates of a and b obtained with these methods.

4.2.3 Median of Estimators

To suppress naturally the effect of outliers without actually ignoring them, the median of the empirical distributions for $\hat{a}$ and $\hat{b}$ were calculated. The results for the median of the distributions are given in Table 4.7 and the variation of the median with sample size is shown in Fig. 4.7. In the calculation of the median values, the outlier estimators were included.

Table 4.6 Mean squared error about the true beta parameters (a=1.2, b=23) for the simulated failure data. Each data set contained 500 samples.

Sample Size (N)	Data Set No.	Marginal Match. Mom. MSE($\hat{a}$)	Marginal Match. Mom. MSE($\hat{b}$)	Prior Match. Mom. MSE($\hat{a}$)	Prior Match. Mom. MSE($\hat{b}$)	Marginal Max. Likelihood MSE($\hat{a}$)	Marginal Max. Likelihood MSE($\hat{b}$)	Prior Max. Likelihood MSE($\hat{a}$)	Prior Max. Likelihood Var($\hat{b}$)
	L	55.9	33,200	2.57	1,740	77.6	35,000	6.76	3,780
5	2	58.0	27,500	6.43	4,860	43.4	21,000	17.1	11,500
	3	70.8	28,900	5.80	6,050	77.1	37,100	14.9	15,100
	1	11.7	4,880	0.629	308	7.12	4,670	1.52	524
10	2	7.61	4,480	0.526	290	4.68	3,090	1.20	566
	3	22.0	10,400	0.535	310	7.12	5,860	1.14	639
	1	0.971	472	0.215	95.3	0.618	314	0.422	125
20	2	0.806	455	0.185	89.1	0.680	382	0.288	128
	3	1.33	782	0.235	115	0.781	503	0.520	193
50	1	0.201	92.7	0.0874	37.1	0.123	63.1	-	-

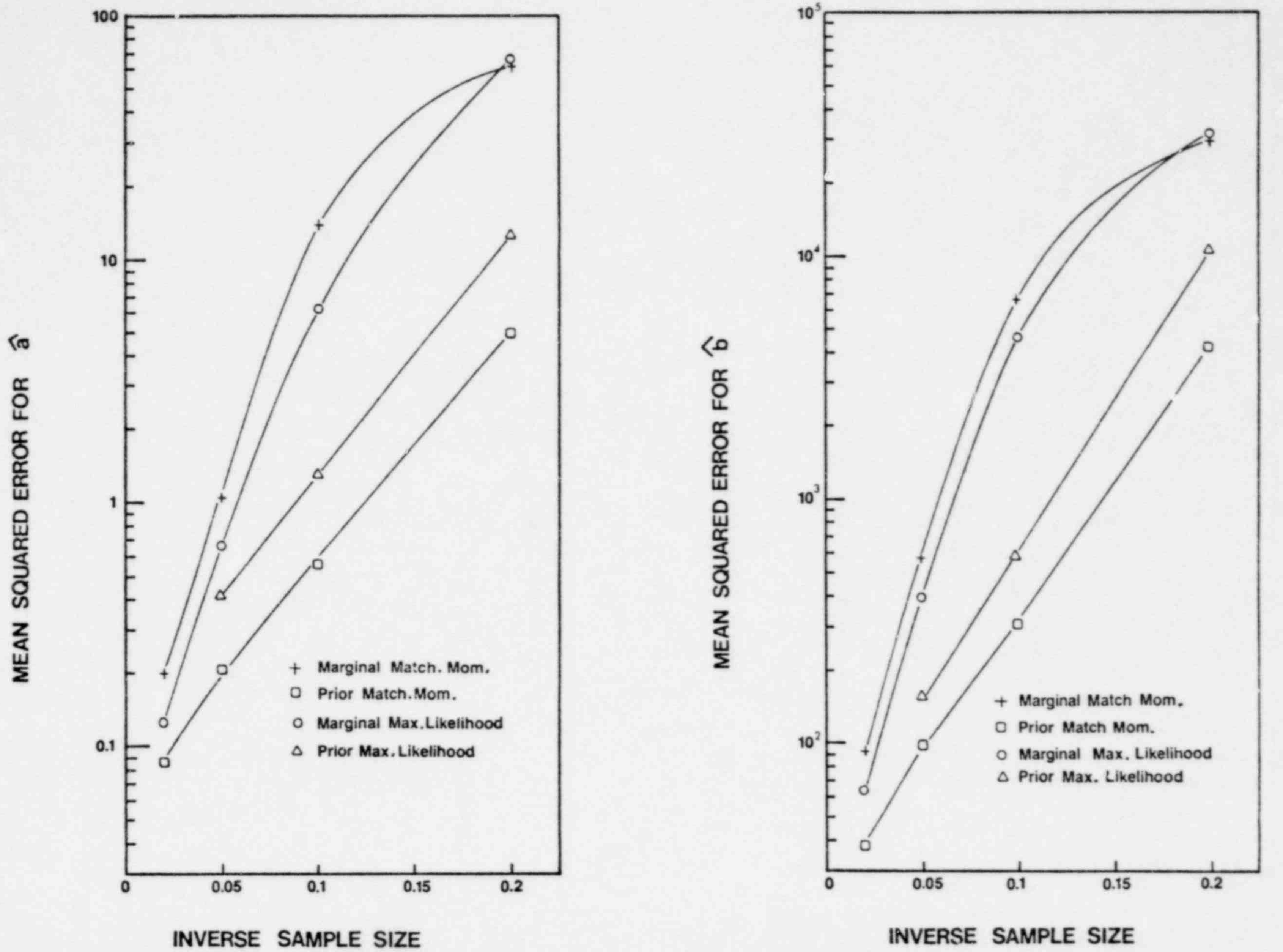


Fig. 4.6 Variation of the mean squared error of the beta parameter estimators with sample size for the different estimation techniques. True values of the beta parameters are a=1.2 and b=23.

Table 4.7 Median values for the estimates $\hat{a}$ and $\hat{b}$ for different sample sizes and estimation techniques. For sample sizes of 5, 10 and 20, 1500 simulated failure data were used, and for sample size 50, 500 simulated data were used. The true value of the parameters are a=1.2 and b=23.0.

Sample Size (n)	Marginal Match. Mom. $\hat{a}$	$\hat{b}$	Prior Match. Mom. $\hat{a}$	$\hat{b}$
5	2.22	46.3	1.31	27.8
10	1.72	33.5	1.76	23.0
20	1.47	28.4	1.10	21.4
50	1.28	24.4	1.02	19.6

Sample Size (N)	Marg. Max. Like. $\hat{a}$	$\hat{b}$	Prior Max. Like. $\hat{a}$	$\hat{b}$
5	1.77	36.9	2.09	39.2
1C	1.47	28.9	1.65	29.9
20	1.33	25.6	1.67	29.2
50	1.23	23.3	-	-

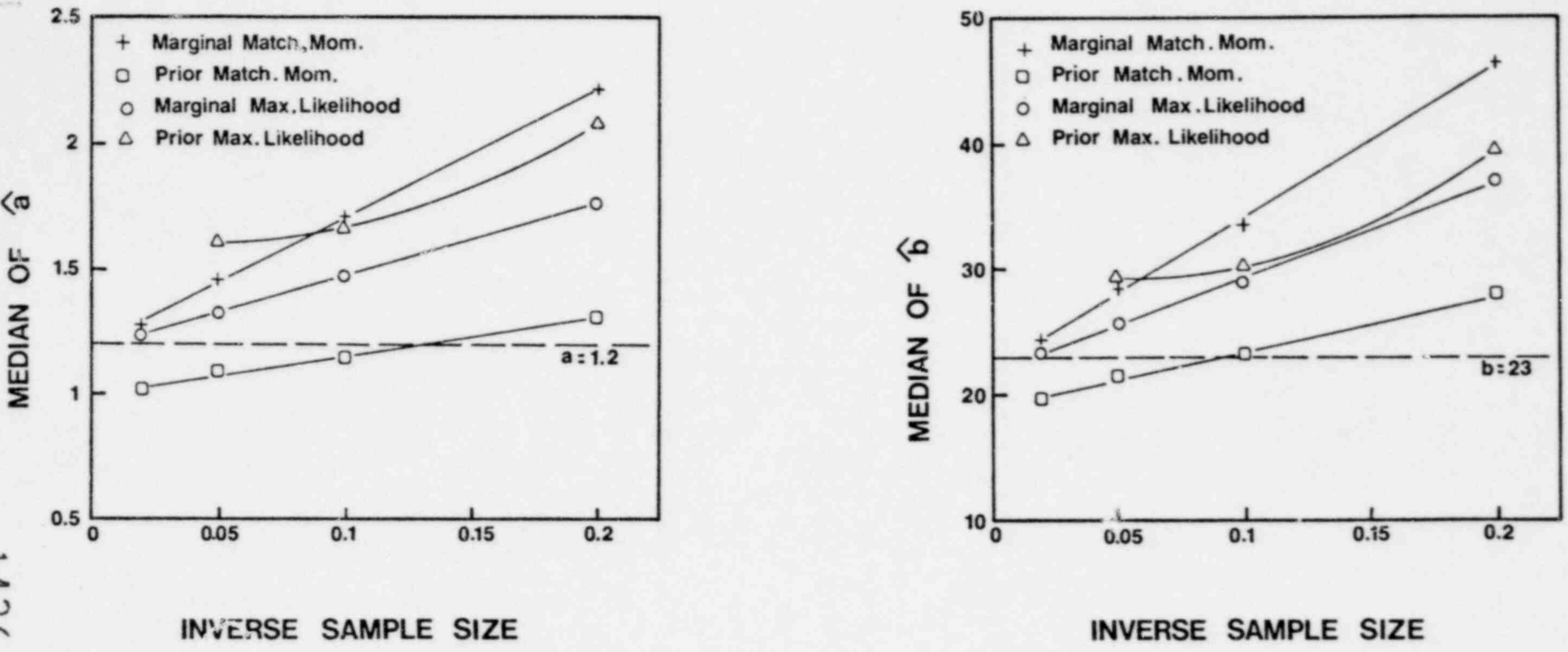


Fig. 4.7 Variation of the median of the beta parameter estimators with sample size for the different estimation methods. True parameter values are a=1.2 and b=23.

For small sample sizes ($N \leq 10$) the simple prior matching moments method yields median values which are closest to the true values of the parameters. However, for larger sample sizes the prior matching moment methods gives a median which is smaller than the true value. Only the estimation methods based on the marginal distribution appear to yield medians which approach the true value as the sample size becomes very large.

4.2.4 Comparison to Results from a Symmetric Beta Prior

The results in the previous section were estimated from simulation failure data based on a specific beta prior distribution which was highly skewed towards low failure probabilities (the mean of the beta prior = $a/(a+b)$ = $1.2/(1/2+23) = 0.043$). To determine whether the results obtained for the estimators of this particular beta prior are applicable only to similarly skewed beta priors or to more generally distributed beta priors, failure data were simulated for a symmetrically distributed beta prior with parameters $a=b=5$ and consequently with a mean of 0.5 Simulated failure data sets of 500 samples of size 5, 10 and 20 were generated from this symmetric beta distribution. The four estimation techniques were used to analyze these data.

From this analysis of failure data generated from a symmetric beta prior, it was found that both marginal-based estimation techniques yielded numerical solutions for a larger fraction of the samples than they did for the nonsymmetric case. For example, 98.8% of the size 5 samples yielded results with the marginal matching moments method and 98.0% of the same samples were successfully analyzed by the marginal maximum likelihood method. For the nonsymmetric case these success rate percentages were (see Table 4.2) 92.2% and 89.9%,respectively. Unlike the nonsymmetric case, all data samples of size greater than 5 yielded solutions by all four methods. Moreover, the estimator outliers obtained with the symmetric samples were far less objectionable (i.e., fewer in number and closer in value to the main distribution) than were the outliers for the corresponding nonsymmetric cases. For the case of a symmetric beta prior, none of the simulated failure samples contained a $k_i=0$ (or $k_i=n_i$), and hence, unlike the skewed beta prior case, the prior maximum likelihood estimation method produced parameter estimates for all samples.

1426 280

The results for the bias and the mean squared error of the estimators are given in Table 4.8 for various sample sizes. Figures 4.8 and 4.9 show the variation with sample size of the bias and mean square error, respectively. Because the true beta parameters are equal ($a=b=5$), one would expect the plots of the bias for $\hat{a}$ to be the same as for $\hat{b}$. Indeed the small observed differences in Fig. 4.8 or in Table 4.8 are a result of statistical uncertainties arising from the relatively small number of samples (500) used to construct the distributions of $\hat{a}$ and $\hat{b}$.

From Fig. 4.8 all four methods appear to give zero or very small bias if the sample size becomes sufficiently large. As with the skewed case, all four methods tend to overestimate the prior parameters for small sample size, and only the simplest method, the prior matching moments technique gives a slight negative bias for samples of size greater than about N=15. Also, as was seen with the skewed case, the two estimation techniques based on the marginal distribution give essentially identical results which are considerably poorer than those obtained with the prior based methods. Thus the prior matching moments techniques had a performance which was as good or better than the other techniques in this symmetric case also.

4.3 Distribution of Estimators for the Mean and Variance of the Prior Distribution

For small sample sizes ($N \leq 20$) all four parmeter estimation techniques investigated in this study tended to overestimate values of the parameters a and b for the beta prior distribution. In fact, for very small sample sizes ($N \simeq 5$) and for data generated from the beta prior distribution skewed towards low probability values ($a=1.2$, $b=23$), occasional estimates of a and b were obtained from the marginal-based techniques which were several orders of magnitude too large.

As previously stated, it was observed that whenever an inordinately large value of one beta parameter was obtained, the estimate for the other parameter was also very large. For these overestimation cases, it was observed that a reasonable estimate of the mean of the beta prior was obtained even with these large parameter estimates, since the mean depends only on the ratio a/b, i.e., from Eq. (2.4)

$$\mu = (1 + b/a)^{-1} \quad . \qquad (4.10)$$

1426 281

Table 4.8 The bias and mean squared error of the estimators of the parameters for a symmetric beta prior distribution (a=b=5) as calculated by different estimation techniques from simulated failure data of various sample sizes. Each data set consisted of 500 samples.

Sample Size (N)	Marginal Matching Moments				Prior Matching Moments			
	$\hat{a}-a$	$\hat{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)	$\bar{a}-a$	$\bar{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)
5	10.98	10.8	1076.	1092.	3.68	3.38	164.0	124.1
10	2.50	2.56	69.1	94.9	0.535	0.533	12.3	13.2
20	0.79	0.764	6.36	5.91	0.110	-0.13	3.47	3.19

Sample Size (N)	Marginal Maximum Likelihood				Prior Maximum Likelihood			
	$\bar{a}-a$	$\bar{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)	$\bar{a}-a$	$\bar{b}-b$	MSE($\hat{a}$)	MSE($\hat{b}$)
5	10.3	9.99	936.	862	6.16	5.80	272.	210.
10	2.65	2.70	75.3	102.	1.3	1.30	16.7	12.8
20	0.827	0.805	6.19	5.74	0.208	0.186	3.89	3.51

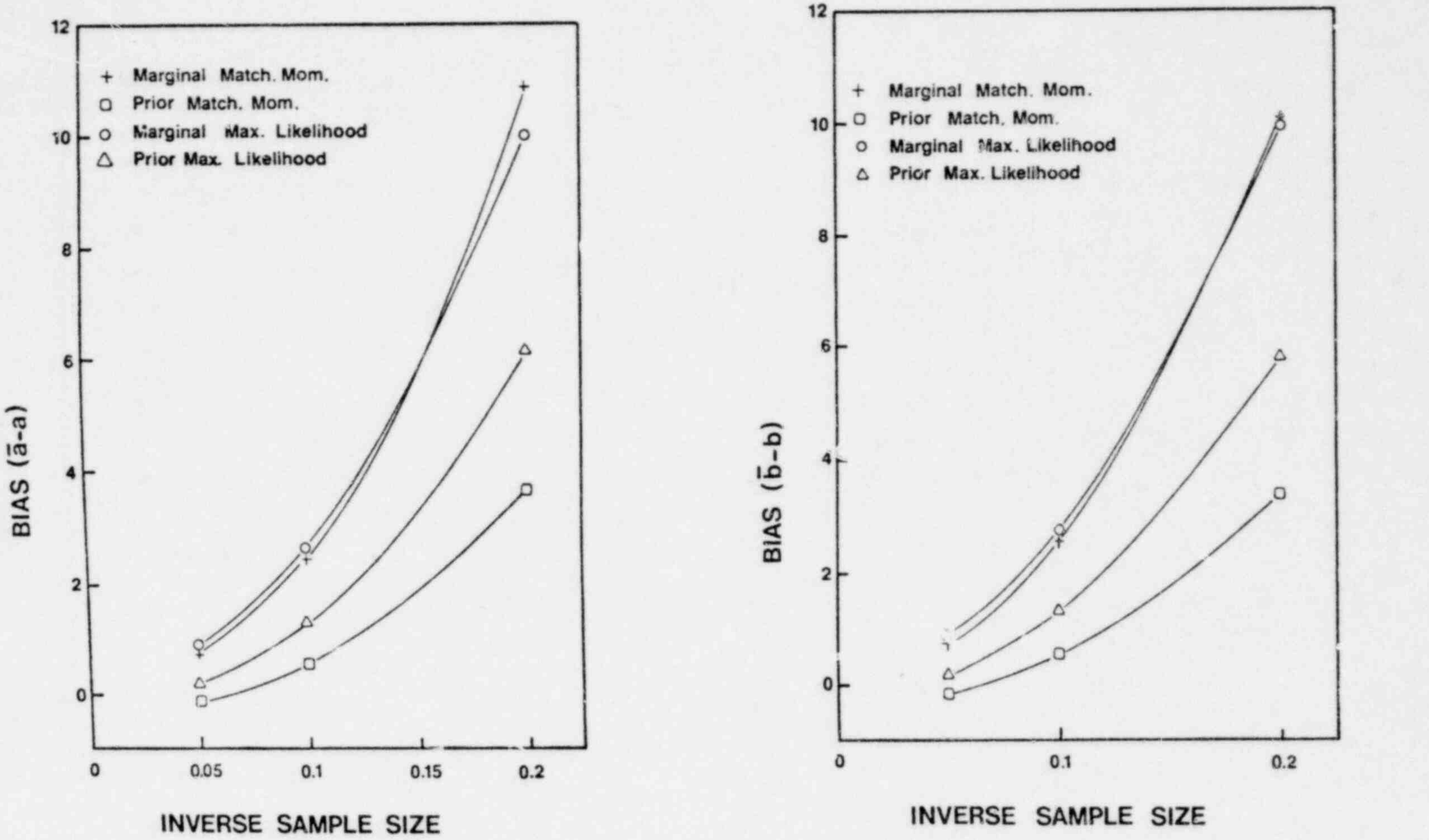


Fig. 4.8 Variation of the bias of the beta parameter estimators with sample size for the symmetric beta distribution (a=b=5).

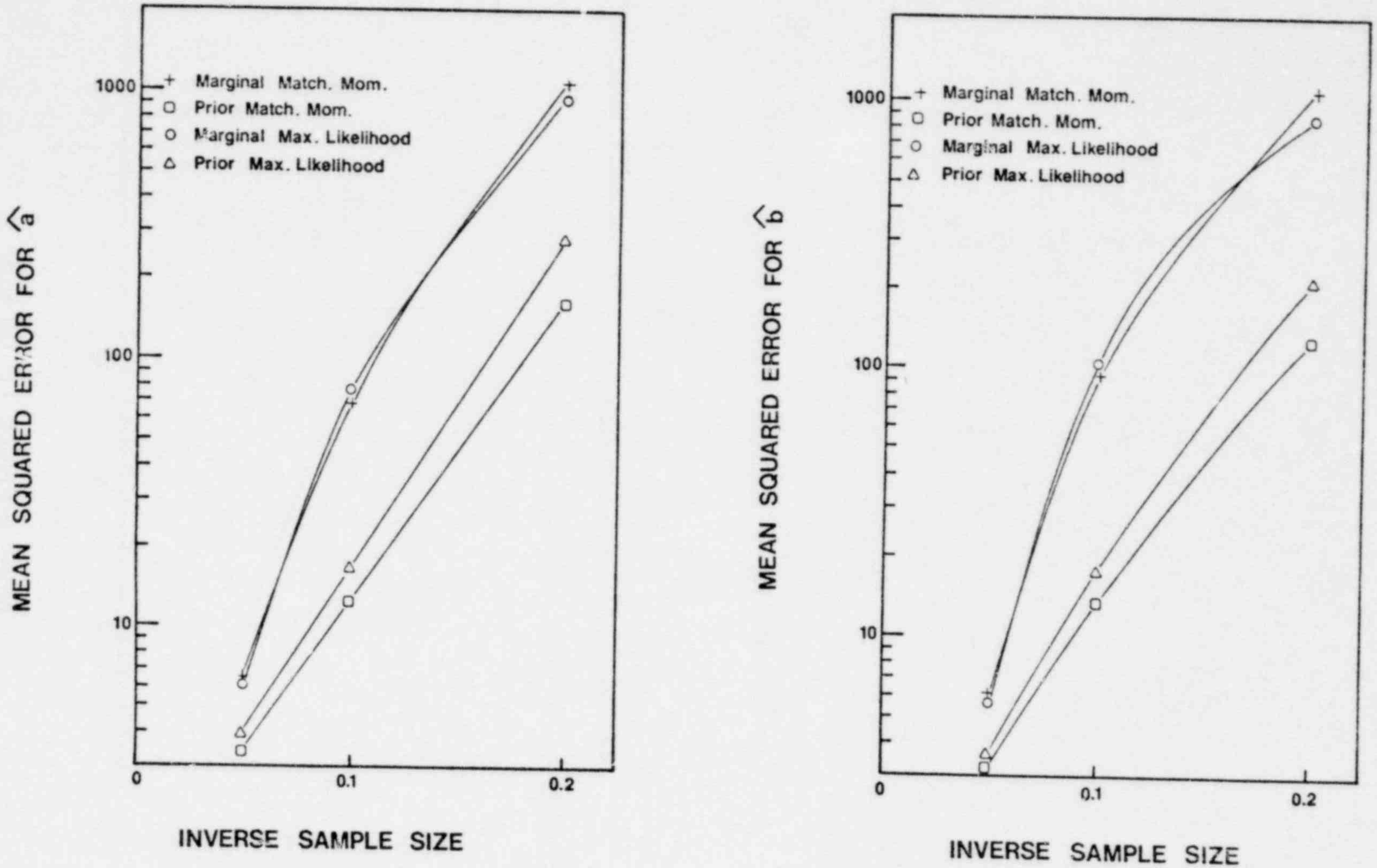


Fig. 4.9 Variation with sample size of the mean squared error of the beta parameter estimators for the symmetric beta distribution (a=b=5).

The empirical distributions of the estimate of the prior mean was calculated for different sample sizes, by using the estimators $\hat{a}$ and $\hat{b}$ in Eq. (4.10) previously obtained with the simulated failure data for the skewed prior case (true mean = $(1 + 23/1.2)^{-1} = 0.0496$). These distributions are shown in Figs. 4.10-4.13 and the mean and variance of these distributions are given in Table 4.9. Because of the inability of the prior maximum likelihood method to treat low failure probability cases, this method was not included in the analysis.

From these distributions of mean estimators it is seen that no apparent outliers are present. Further the mean of the distributions are all within a small percentage of the true value, although a very slight bias to over-estimate the mean is noted. As would be expected, the variances of the distributions decrease as the sample size increases. The most important feature, however, of these distributions of $\hat{\mu}$ is that all three estimation techniques appear to give nearly the same distribution for a given sample size.

Although the presence of outlier estimators for a and b does not affect the distribution of the mean estimators, the high a and b estimates will have a profound effect on the estimation of the variance of the beta prior distribution. The variance of the beta prior is given by (Eq. (2.5))

$$\sigma^2 = [(1 + b/a)(1 + a/b)(a + b + 1)]^{-1} \qquad (4.11)$$

which becomes very small as a and b both become large. Thus the use of outlier estimators $\hat{a}$ and $\hat{b}$ to produce an estimate of the variance for the beta prior will give unrealistically small values. In Figs. 4.14-4.17, the distributions of the variance estimators for the prior beta are shown for different sample sizes.

Notice that for small sample sizes (e.g., Fig. 4.14) for which outlier values are expected for the marginal-based estimation methods, the empirical frequency distributions of the variance estimators (Eq. 4.11) are peaked towards the low end. However as the sample size increases, outlier values for a and b are no longer obtained, and the variance estimator distribution becomes increasing centered around the true variance of $\sigma^2 = 0.00187$. Finally it should be noted from these variance distributions, that the distribution produced by the prior matching moments results is always slightly more skewed towards the high values as compared to the distributions for the two marginal-based methods.

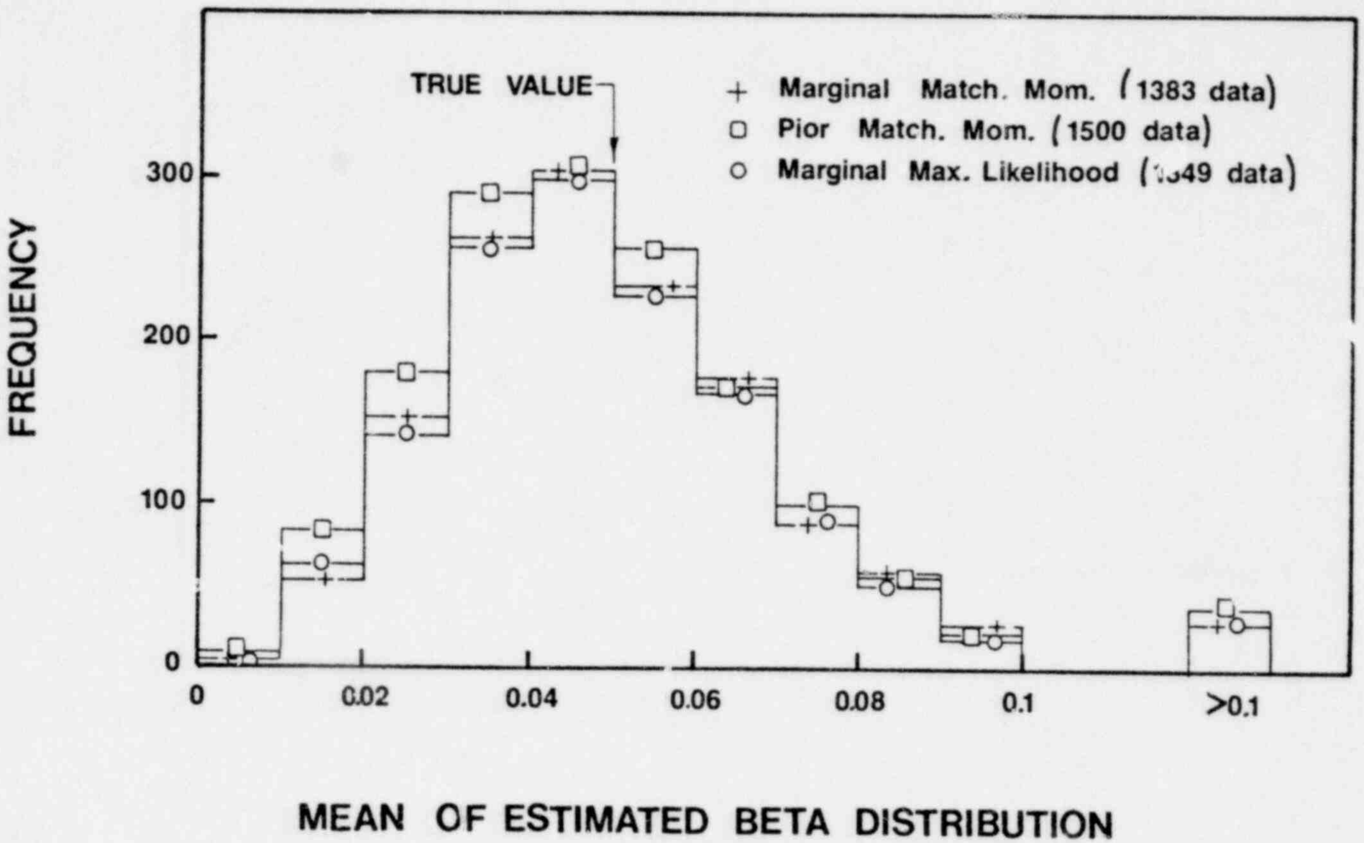


Fig. 4.10 Distribution of the means of the estimated beta prior distributions from samples of size N=5. Samples were generated from a beta-binomial distribution with parameters a=1.2 and b=23 which yield a true prior mean of 0.0496.

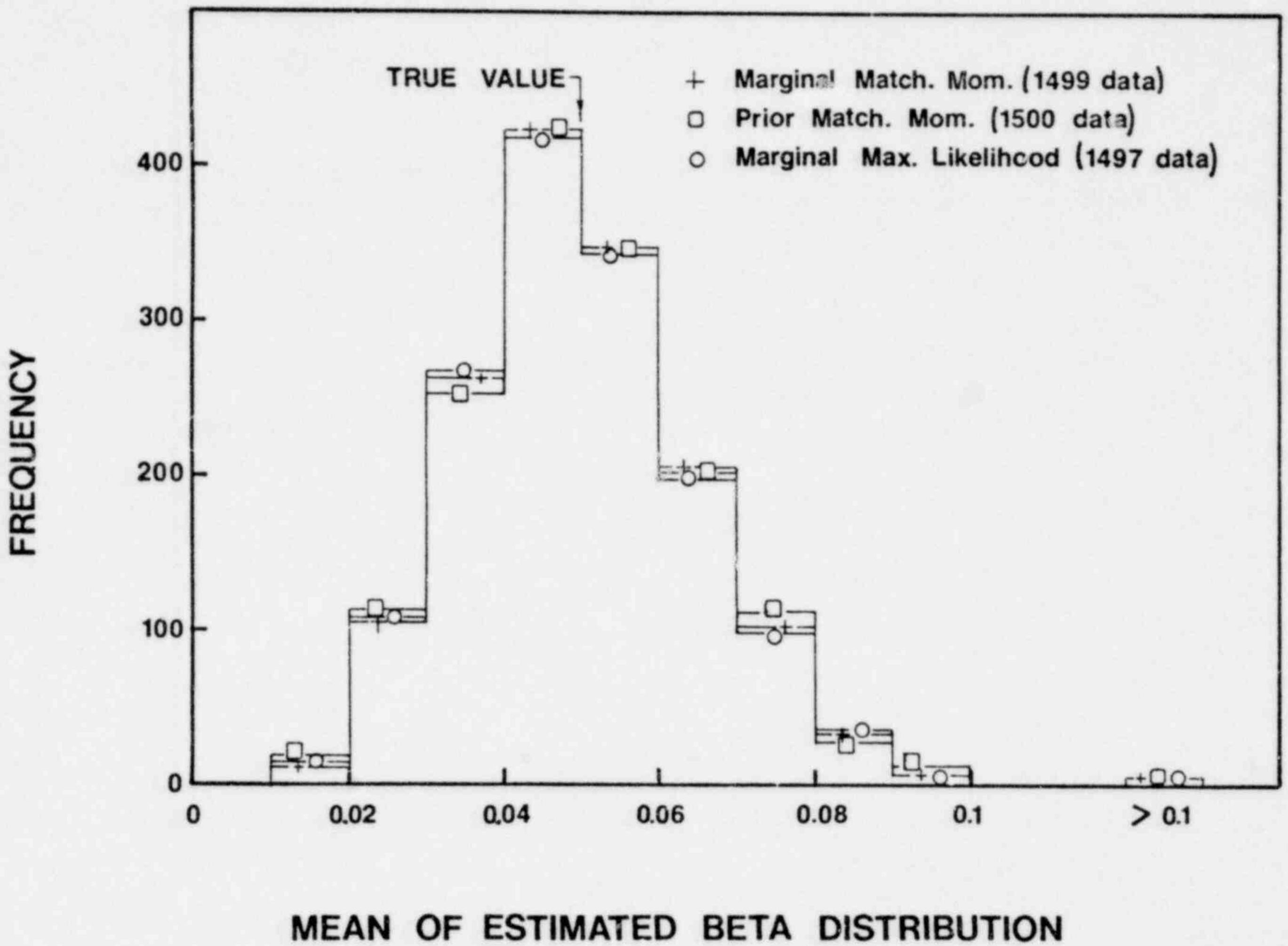


Fig. 4.11 Distribution of the means of the estimated beta prior distributions from samples of size N=10. Samples were generated from a beta-binomial distribution with parameters a=1.2 and b=23 which yield a true prior mean of 0.0496

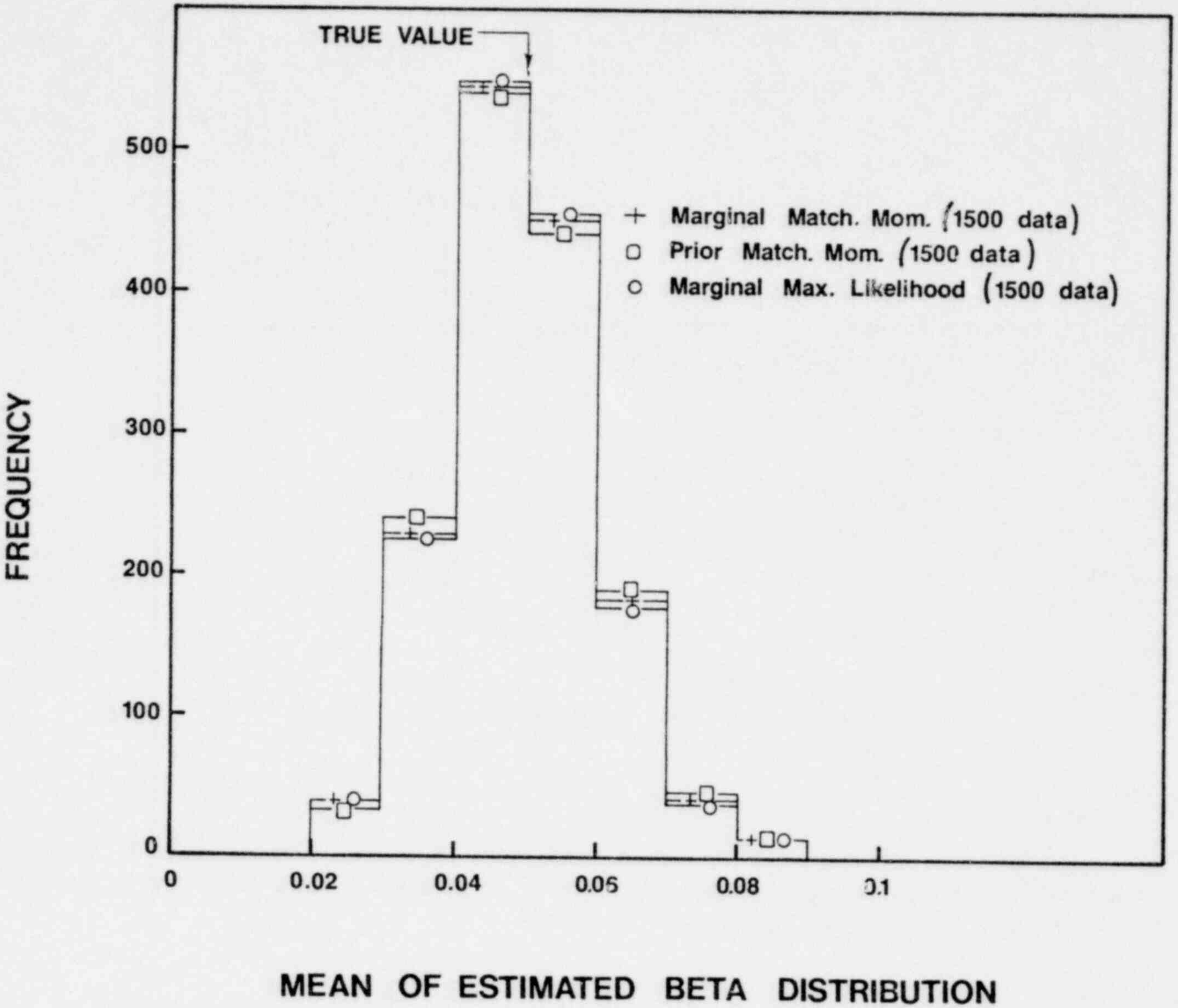


Fig. 4.12 Distribution of the means of the estimated beta prior distributions from samples of size N=20. Samples were generated from a beta-binomial distribution with parameters a=1.2, b=23 which yield a true prior mean of 0.0496.

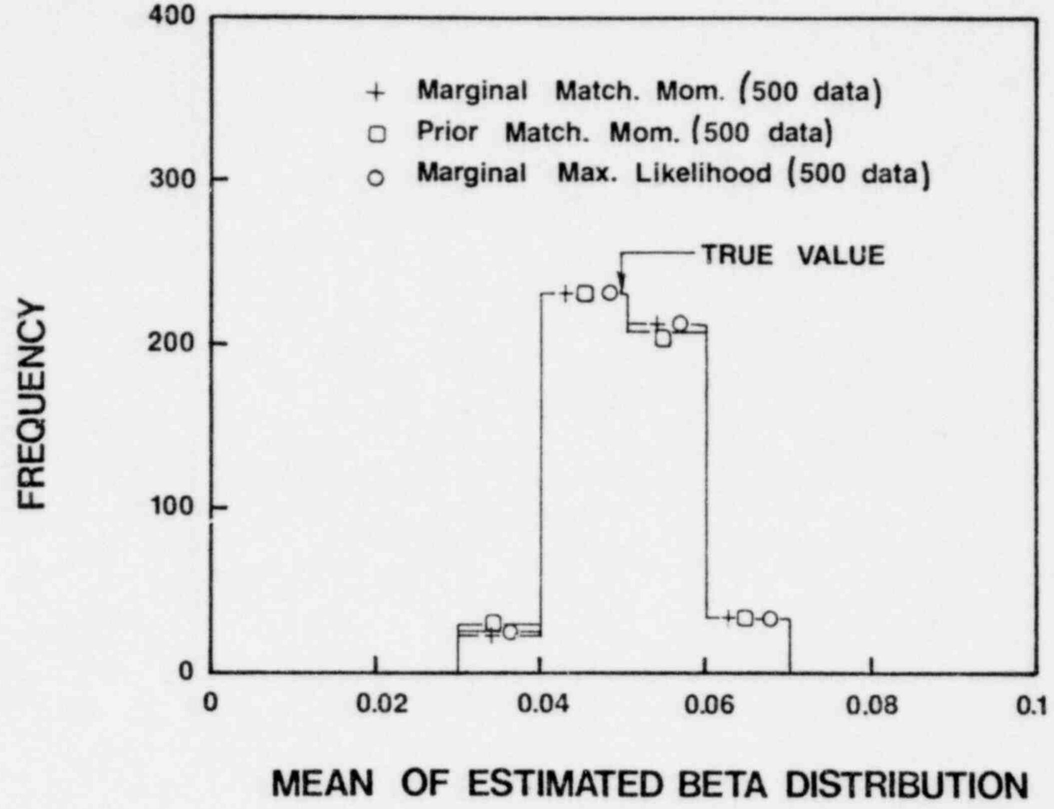


Fig. 4.13 Distribution of the means of the estimated beta prior distributions from samples of size N=50. Samples were generated from a beta-binomial distribution with parameters a=1.2, b=23 which yield a true prior mean of 0.0496.

Table 4.9. Mean and variance of the estimators for the mean of the beta prior (a=1.2, b=23) for different sample sizes. True prior mean is 0.0496.

Sample Size*	Marg. Match. Mom.		Prior Match. Mom.		Marg. Max. Likelihood	
	Mean	Variance	Mean	Variance	Mean	Variance
5	0.0500	0.0000422	0.0488	0.000423	0.0497	0.000415
10	0.0500	0.000218	0.0500	0.000221	0.0498	0.000218
20	0.0496	0.000113	0.04963	0.000114	0.0495	0.000112
50	0.0500	0.0000422	0.049928	0.0000419	0.0499	0.0000419

*1500 samples were used for size 5-20 results; 500 samples were used for size 50 results.

Table 4.10. Mean and variance of the estimators for the variance of the beta prior (a=1.2, b=23) for different sample sizes. True prior variance is 0.00187.

Sample Size*	Marg. Match. Mom.		Prior Match. Mom.		Marg. Max. Likelihood	
	Mean	Var. [x10^5]	Mean	Var. [x10^5]	Mean	Var. [x10^5]
5	0.00141	0.298	0.00207	0.507	0.00171	0.393
10	0.00167	0.215	0.00227	0.295	0.00185	0.225
20	0.00172	0.116	0.00222	0.145	0.00181	0.102
50	0.00184	0.0468	0.00227	0.0558	0.00188	0.0406

*1500 samples were used for sizes 5-20 results; 500 samples were used for size 50 results.

1426 290

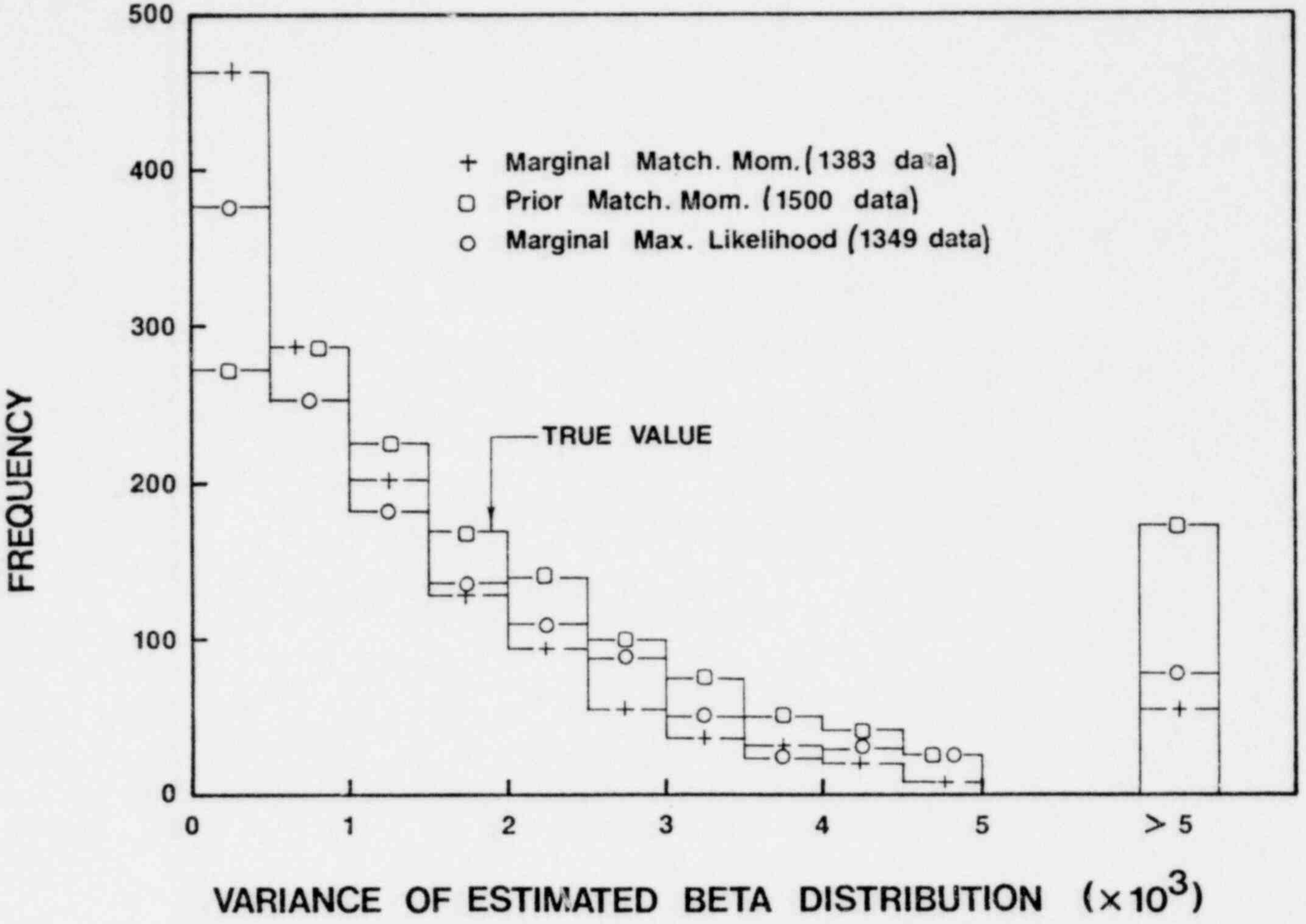


Fig. 4.14 Distribution of the variances of the estimated beta prior distributions from samples of size N=5. Samples were generated from a beta-binomial distribution with parameters a=1.2 and b=23 which gives a variance of 0.00187 for the beta prior distribution.

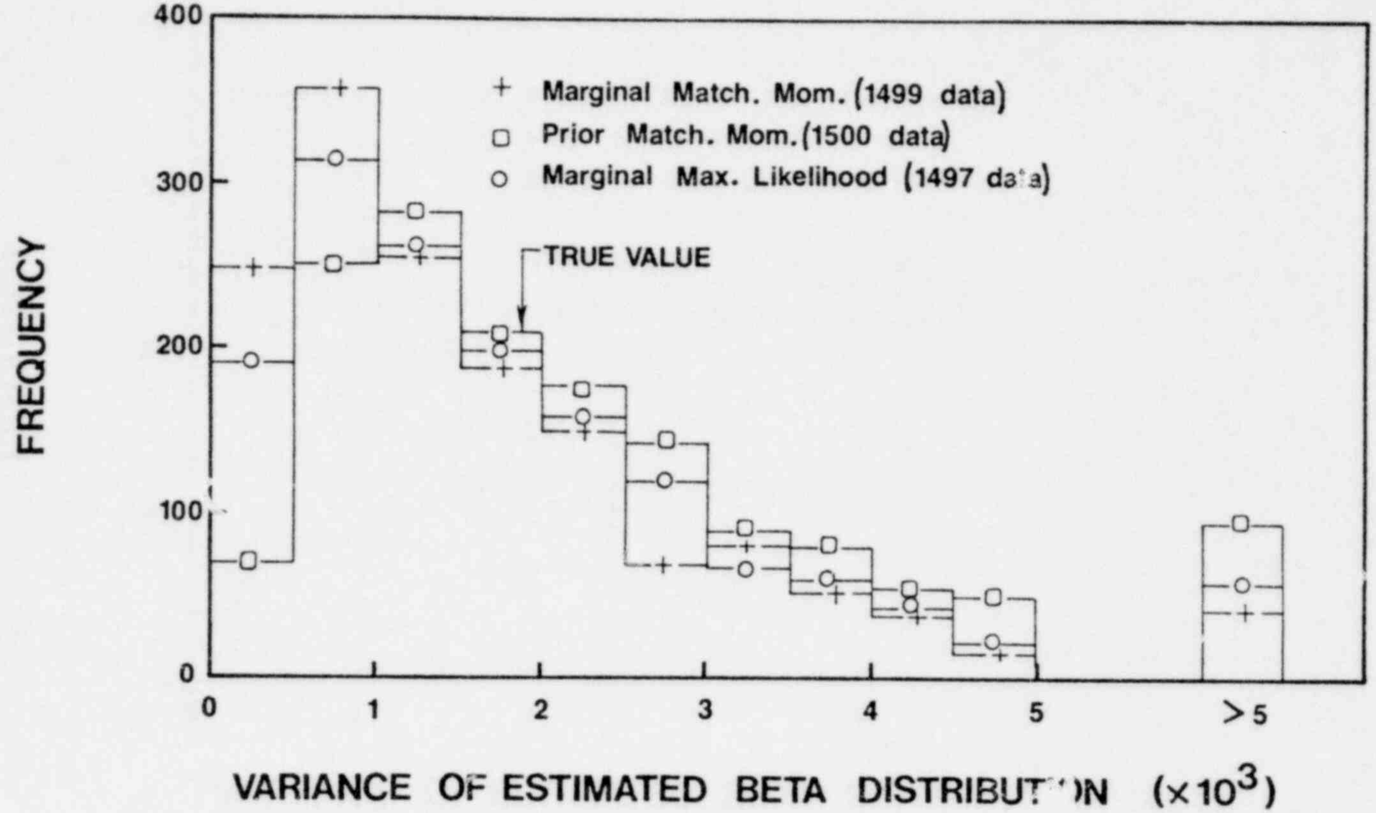


Fig. 4.15 Distribution of the variances of the estimated beta prior distributions from samples of size N=10. Samples were generated from a beta-binomial distribution with parameters of a=1.2 and b=23 which gives a variance of 0.00187 for the true beta prior distribution.

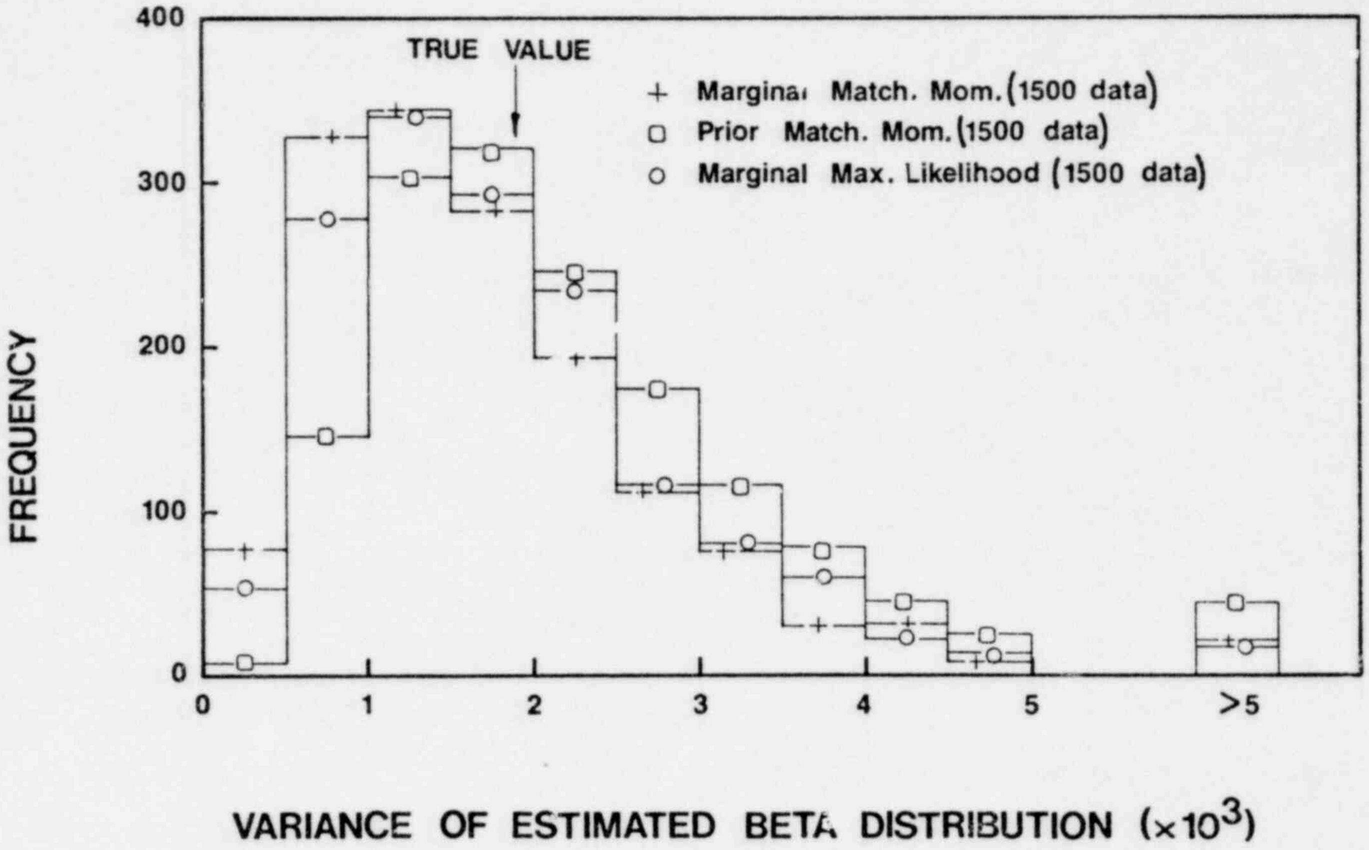


Fig. 4.16 Distribution of the variances of the estimated beta prior distributions from samples of size N=20. Samples were generated from a beta-binomial distribution with parameters of a=1.2 and b=23 which gives a variance of 0.00187 for the true beta prior distribution.

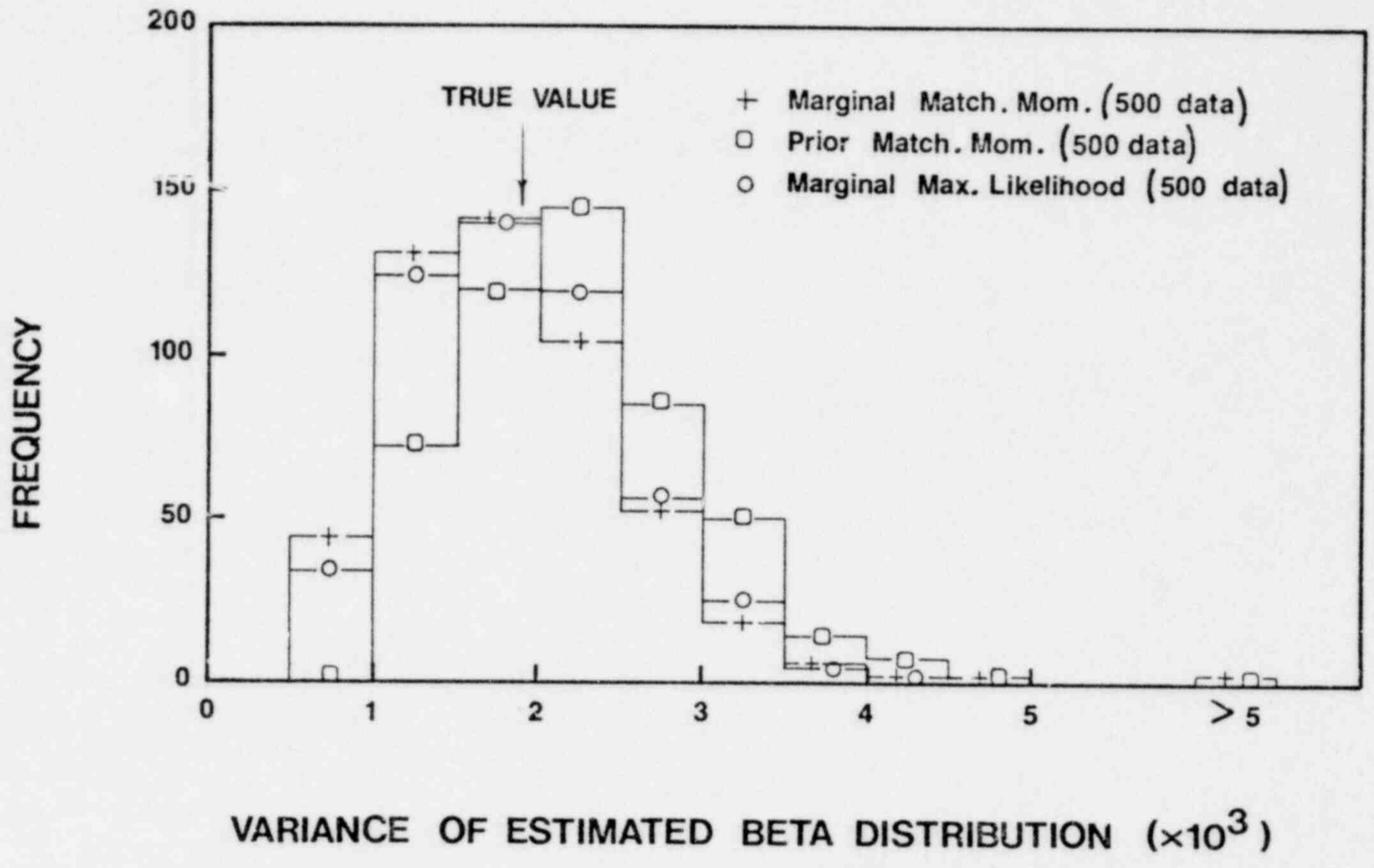


Fig. 4.17 Distribution of the variances of the estimated beta prior distributions from samples of size N=50. Samples were generated from a beta-binomial distribution with parameters of a= 1.2 and b=23 which gives a variance of 0.00187 for the true beta prior distribution.

In Table 4.10 the mean and variance of these variance estimator distributions are given. It is noted that the mean of the distribution is always slightly less than the true prior variance ($\sigma^2 = 0.00187$) but approaches the true value as the sample size increases. The means of the prior matching moments distributions, however, always overestimate the true mean. More importantly, these overestimates do not appear to approach the true value even as the sample size increases, but rather appear to remain about 20% higher than the true value.

4.4 Distribution of 95-th Percentile Estimators

Of considerable interest in safety analysis is the estimation of the prior distribution at high failure probabilities. One widely used measure of the high probability tail is the 95-th percentile, i.e., the failure probability, p_{95}, above which there is only a 5% chance that the true failure probability lies for a component described by the prior distribution, g(p). For the beta prior distribution used in this study, the 95-th percentile, p_{95}, is the solution of the following equation:

$$0.5 = \int_0^{p_{95}} g(p)dp = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^{p_{95}} p^{a-1} (1-p)^{b-1} dp. \qquad (4.12)$$

The numerical solution of this equation for p_{95} is discussed in detail in Chapter 5, and a program for performing this calculation is included in Appendix II.

For each simulated failure data set generated for the beta prior which was skewed towards the low probability end (a=1.2, b=23), an estimator of the 95-th percentile was obtained by using the estimators $\hat{a}$ and $\hat{b}$ for each set in Eq. (4.12) and solving numerically for the 95-th percentile. The distribution of the 95-th percentile estimators so obtained are shown in Figs. 4.18-4.21 for the three estimation techniques suitable for analyzing low probability failure data. The mean, variance and median of these distributions are presented in Table 4.11.

From a safety viewpoint, one would like to use an estimation technique which has a low inherent probability of yielding 95-th percentile estimates which are very much less than the true value. In other words, if the estimator is biased, then it would be better if it were biased so as to yield overestimates of p_{95} (with hopefully small minimum mean square error). Further, there should be little if any chance of

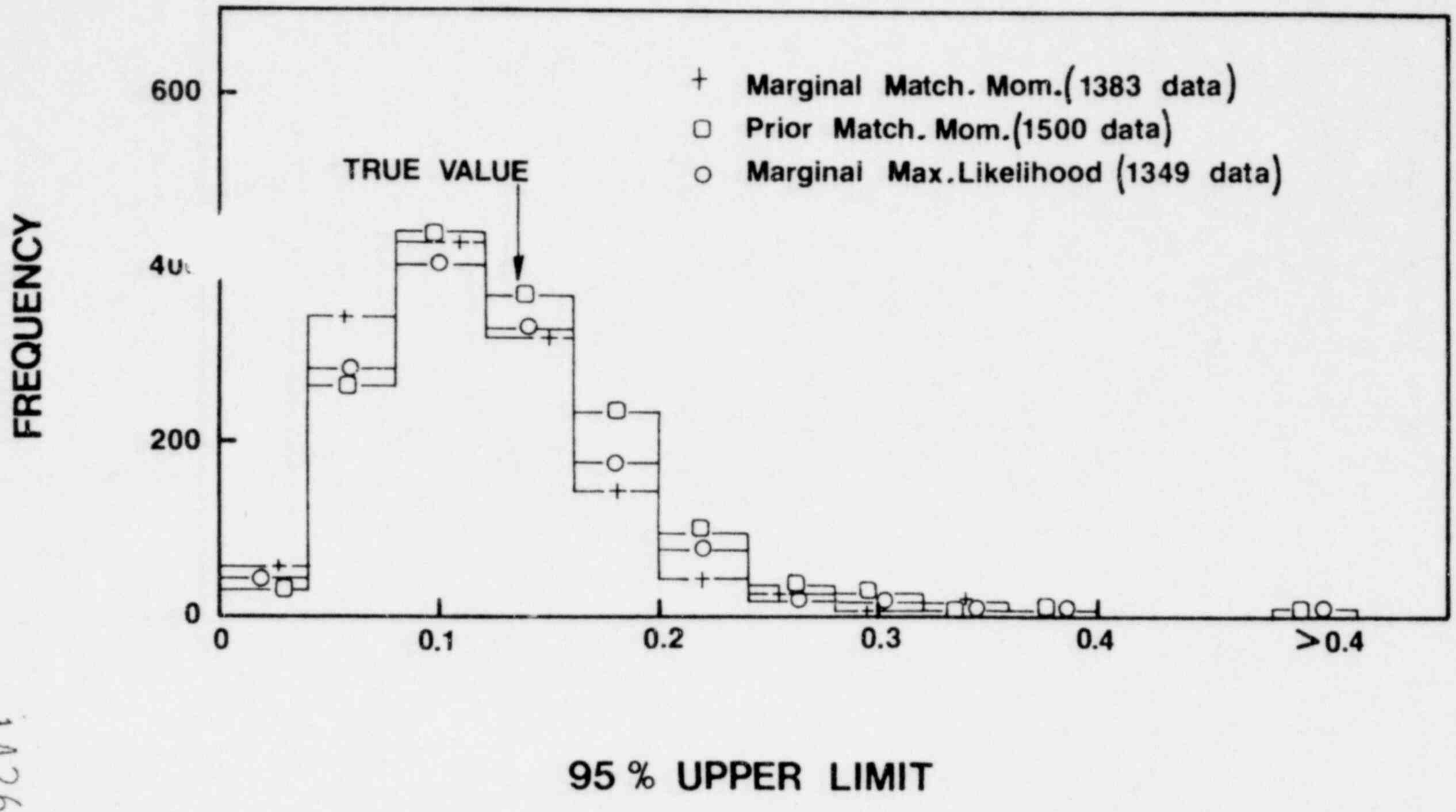


Fig. 4.18 Distribution of the 95-th percentiles oi the estimated beta prior distributions for samples of size N=5. The 95-th percentile of the true beta distribution (a=1.2, b=23) used to generate the simulated failure data is 0.136.

1426 296

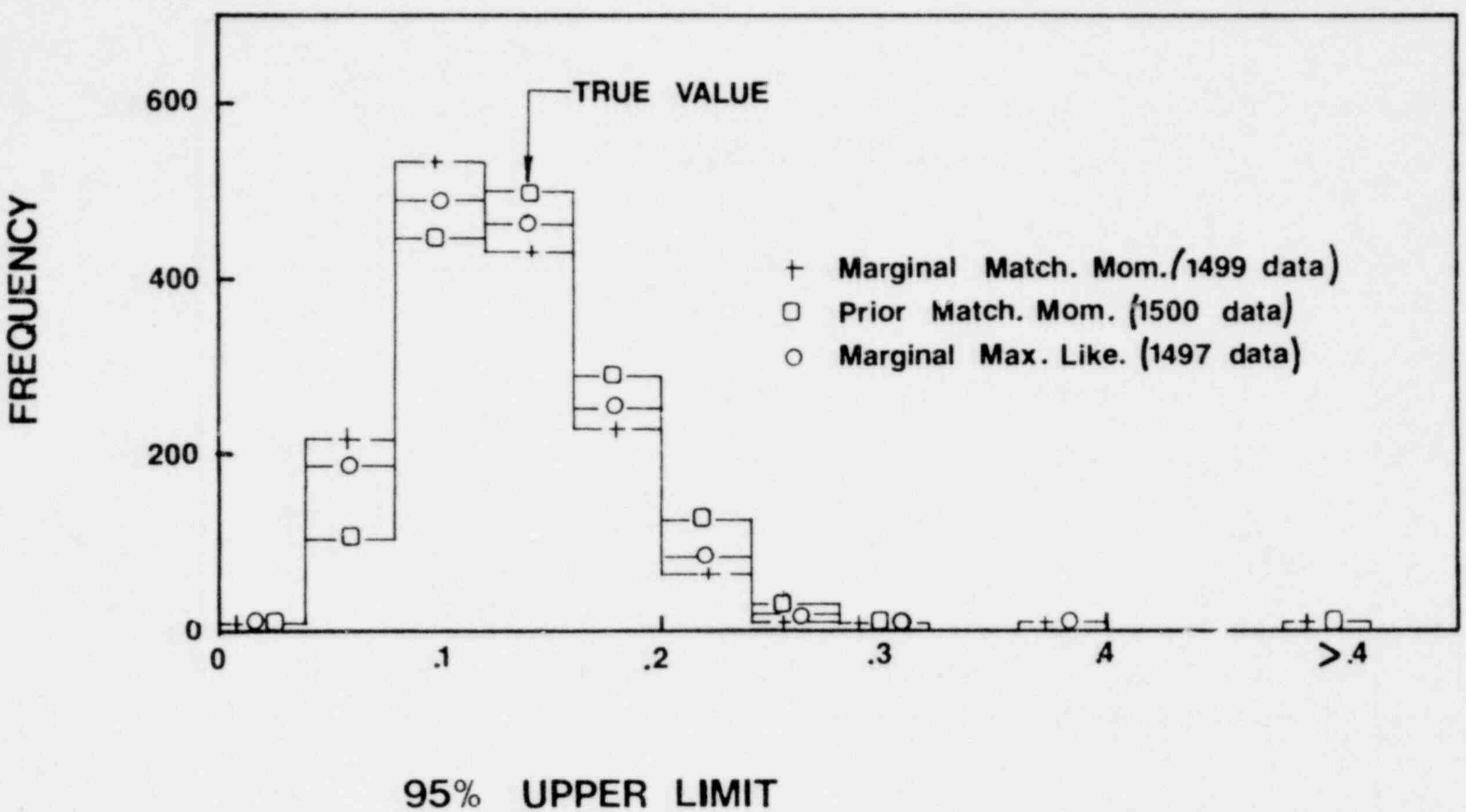


Fig. 4.19 Distribution of the 95-th percentiles of the estimated beta prior distributions for samples of size N=10. The 95-th percentile of the true beta distribution (a=1.2, b=23) used to generate the simulated failure data is 0.136.

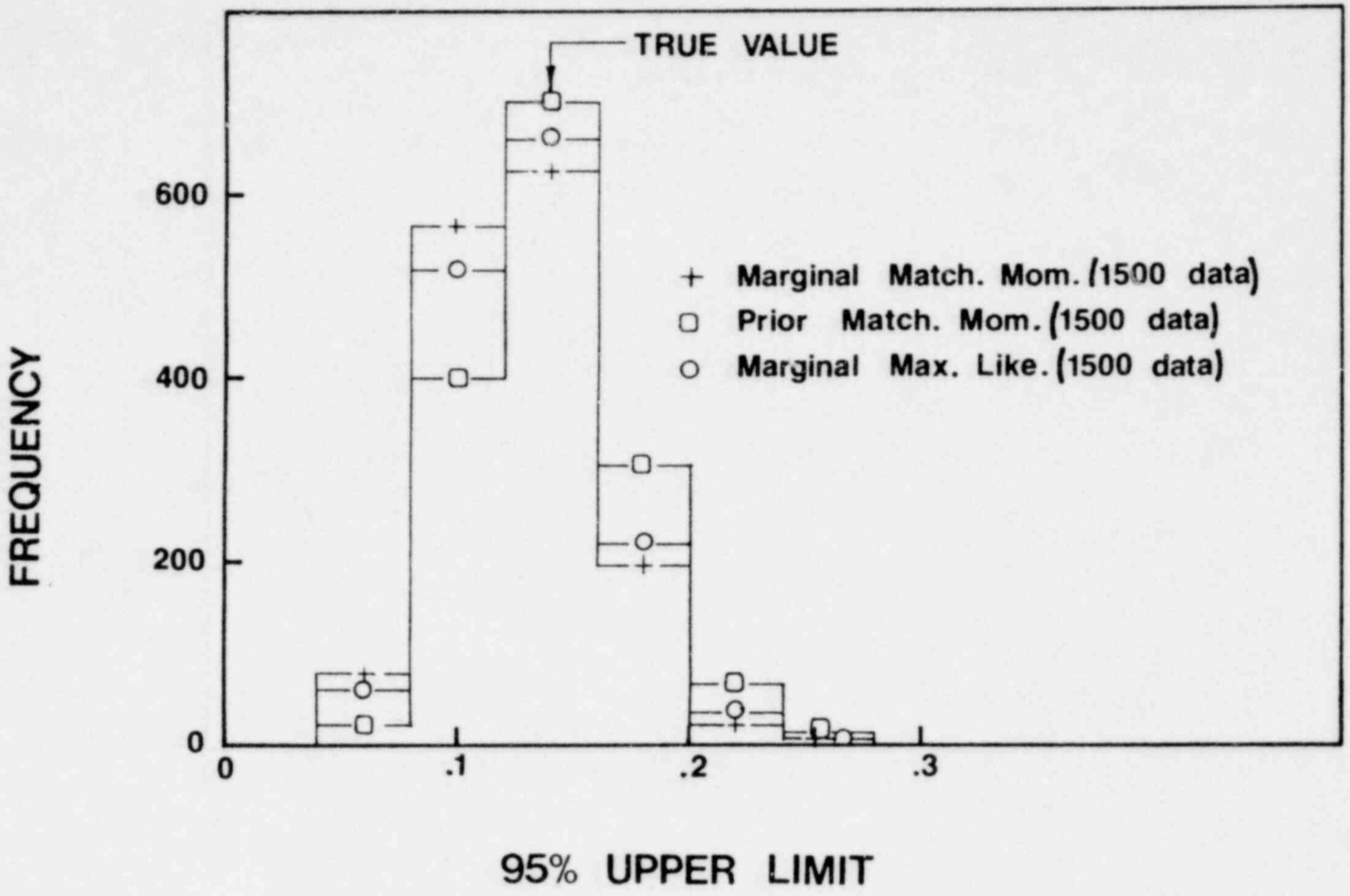


Fig. 4.20 Distribution of the 95-th percentiles of the estimated beta prior distributions for samples of size N=20. The 95-th percentile of the true beta distribution (a=1.2, b=23) used to generated the simulated failure data is 0.136.

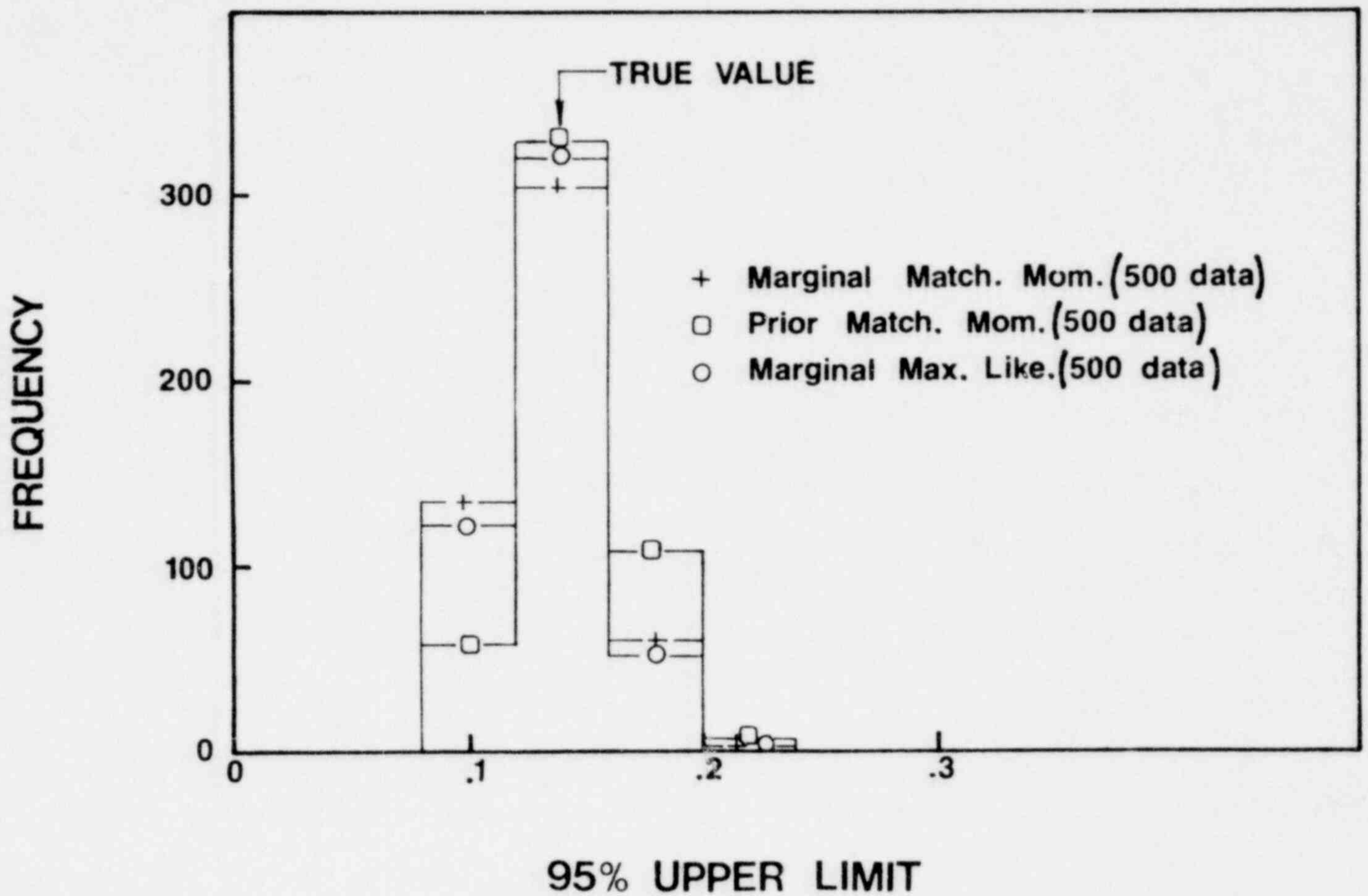


Fig. 4.21 Distribution of the 95-th percentiles of the estimated beta prior distributions for samples of size N=50. The 95-th percentile of the true beta distribution (a=1.2, b=23) used to generate the simulated failure data is 0.136.

Table 4.11 Median, mean and variance of the distributions of the 95-th percentile estimators. True 95-th percentile = 0.13586.

Sample Size*	Marginal Matching Moments			Prior Matching Moments			Marginal Max. Likelihood		
	Median	Mean	Var.	Median	Mean	Var.	Median	Mean	Var.
5	0.106	0.114	0.0029	0.121	0.130	0.0035	0.113	0.121	0.0032
10	0.119	0.124	0.0020	0.136	0.140	0.0021	0.125	0.129	0.0020
20	0.125	0.128	0.0011	0.138	0.141	0.0011	0.129	0.131	0.0010
50	0.133	0.134	0.00045	0.144	0.145	0.00044	0.134	0.135	0.00042

*1500 samples were used for size 5-20; 500 samples for size 50.

yielding outliers or values of $\hat{p}_{95}$ which are orders of magnitude less than the true value. For the present case the true value of the 95-th percentile for a=1.2 and b=23 is p_{95}=0.13586. In Table 4.11, the number of simulated data samples which yielded estimators greater than or less than the true p_{95} are given. Notice that for small samples all three estimation methods are non-conservative ($\text{Prob}\{\hat{p}_{95}<p_{95}\}>0.5$), while as the sample size increases, the prior matching moments becomes increasingly conservative while the medians for the other two methods approach the true p_{95} value.

From Table 4.10, all three methods are seen to yield distributions for $\hat{p}_{95}$ with almost equal variance. However, the two marginal-based estimation techniques yield distributions with means and medians smaller than the true value for all sample sizes although as the sample size increases the medians and means increase and approach the true value of p_{95}. The simple prior matching moments technique also yields distributions of $\hat{p}_{95}$ whose mean and median also increase with increasing sample size, but unlike the other techniques, for sample sizes greater than about seven, the means and medians become greater than the true values, i.e., the distribution becomes conservative. Further for very large sample sizes this positive bias does not disappear, although the bias may not be significantly large.

For small sample sizes (N=5) (see Fig. 4.18) all three methods yield some estimators $\hat{p}_{95}$ in the lowest value bin (0-0.04). These values are, of course, not conservative. Of considerable concern is how these low estimates are distributed in this low end bin. Since the marginal-based estimation techniques occassionally yield very large estimators for a and b, i.e., outliers, the resulting estimated prior distribution will have a very small variance and hence the 95-th percentile will be only slightly greater than the mean. If the mean should turn out to be very small, the $\hat{p}_{95}$ values for these outliers could be very much smaller than the true value. Clearly such a feature of these estimation techniques would preclude their use in safety analyses. In Table 4.13, the lowest 5 values of $\hat{p}_{95}$ found in the present simulation study are listed. It is seen that only one estimate is smaller than 10% of the true value, and hence the possibility of obtaining in the $\hat{p}_{95}$ distribution severe outliers which are orders of magnitude smaller than the true value does not appear to be very likely.

Table 4.12 Number and percent of simulated failure data samples which yielded estimated 95-th percentiles greater than (GT) or less than (LT) the true value of 0.13586).

Sample Size	Marg. Match. Mom.				Prior Match. Mom.				Marg. Max. Likelihood			
	LT		GT		LT		GT		LT		GT	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
5	978	70.7	405	29.3	890	59.3	610	40.7	873	74.7	476	35.3
10	953	63.6	546	36.4	755	50.3	745	49.7	883	59.0	614	41.0
20	873	58.2	627	41.8	701	46.7	799	53.3	820	54.7	680	45.3
50	277	55.4	223	44.6	176	35.2	324	64.8	261	52.2	239	47.8

Table 4.13 Smallest 95-th percentile estimators observed for simulated failure data samples of size N. True value of the 95-th percentile is 0.13586.

Marg. Matching Moments			
N=5	N=10	N=20	N-50
0.0193	0.0362	0.0428	0.0863
0.0206	0.0364	0.0446	0.0871
0.0221	0.0371	0.0503	0.0881
0.0223	0.0387	0.0533	0.0845
0.0234	0.0395	0.0554	0.0922

Prior Matching Moments			
N=5	N=10	N=20	N=50
0.0115	0.0385	0.0592	0.0974
0.0196	0.0451	0.0622	0.101
0.0242	0.0491	0.0658	0.101
0.0243	0.0500	0.0673	0.101
0.0256	0.0509	0.0695	0.102

Marginal Maximum Likelihood			
N=5	N=10	N=20	N=50
0.0152	0.0269	0.0426	0.0848
0.0154	0.0306	0.0461	0.0870
0.0170	0.0360	0.0503	0.0892
0.0209	0.0369	0.0505	0.0922
0.0239	0.0400	0.0572	0.0924

4.5 Fraction of the Estimated Prior Distribution Above the True 95-th Percentile

The extent of the high probability tail of the estimated beta prior distribution is of considerable concern in safety analysis. In the previous section the distribution of the 95-th percentiles of the estimated prior distributions was discussed. An alternative perspective is to consider the fraction of the estimated prior that is supported above the true 95-th percentile, i.e., the probability that the estimated failure probability is greater than the true 95-th percentile. This quantity is given by

$$\text{Prob}\left\{\text{estimated } p \geq p_{95}^{true}\right\} = \int_{p_{95}^{true}}^{1} g_{est}(p)\, dp \ , \qquad (4.13)$$

where p_{95}^{true} is the 95-th percentile of the beta distribution used to generate the simulated failure data (a=1.2, b=23), and $g_{est}(p)$ is the estimated prior distribution for a particular failure data sample (i.e., a beta distribution with $a=\hat{a}$ and $b=\hat{b}$).

If the estimation technique used to analyze the failure data should yield estimators $\hat{a}$ and $\hat{b}$ equal to the true values of the beta prior, then the probability given by Eq. (4.13) would equal 0.05. Of course, the estimation techniques will not in general yield exact values for the beta parameters, and those methods which tend to yield estimated priors skewed more towards higher probability values than the true prior are preferred for safety analysis since the resulting estimated failure probabilities will be overestimated and hence conservative.

The distribution of the probability estimates given by Eq. (4.13), for the three parameter estimation techniques suitable for analyzing low failure probability data, are shown in Figs. 4.22-4.25. It is seen that all three estimation methods yield a considerable portion of values of $\text{Prob}\{p \geq p_{95}^{true}\}$ below the ideal value of 0.05. As the sample size increases, these distributions become increasingly centered about 0.05. However, the distribution for $N \leq 20$ are all highly skewed towards small probabilities with a long slowly decaying behavior at high values. The prior matching moments method in all cases appears to be slightly more "conservative" by giving a distribution which is not as concentrated at the low probability values as compared to the distributions obtained with the other two estimation techniques.

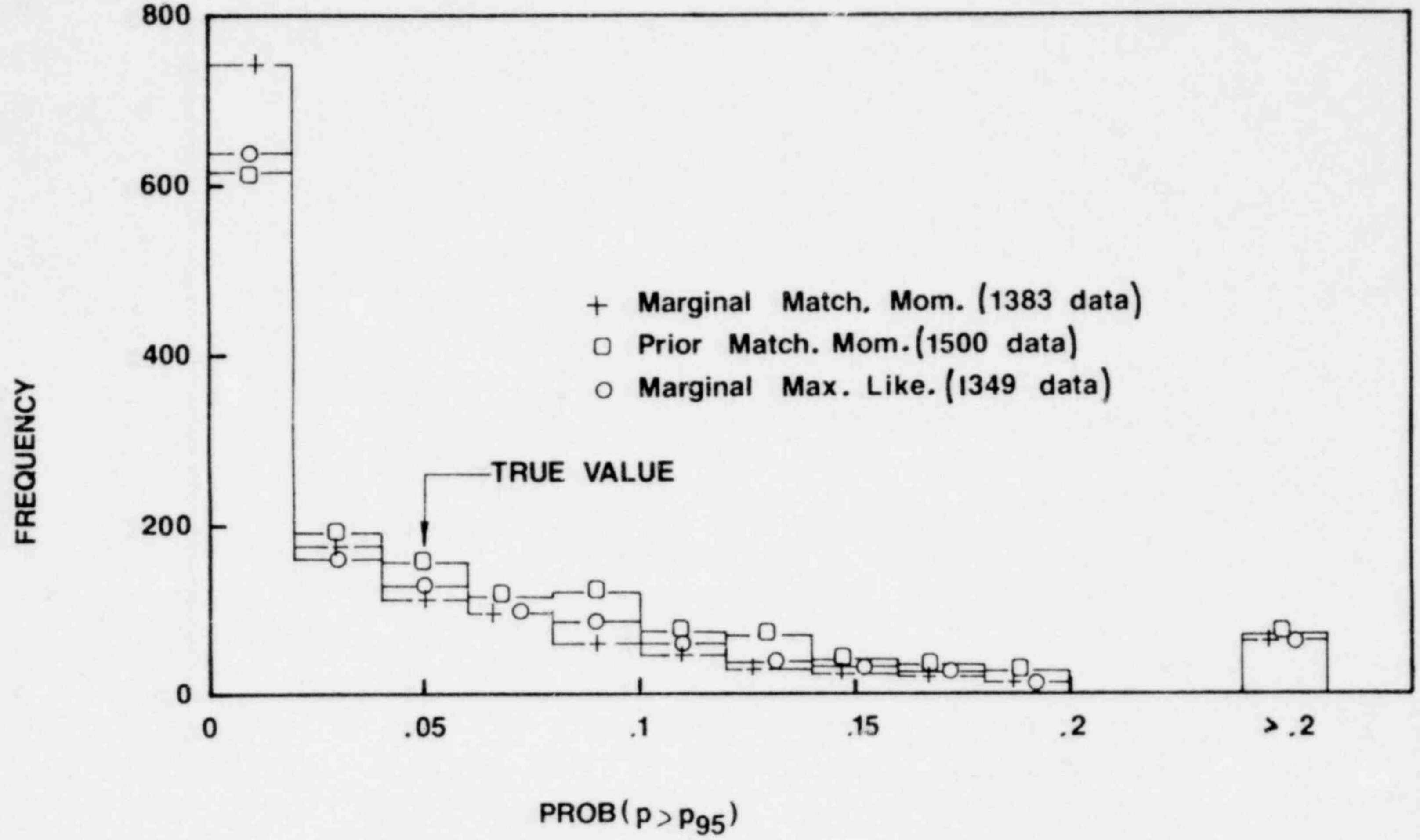


Fig. 4.22 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2, b=23). Size of samples used to obtain estimates was N=5.

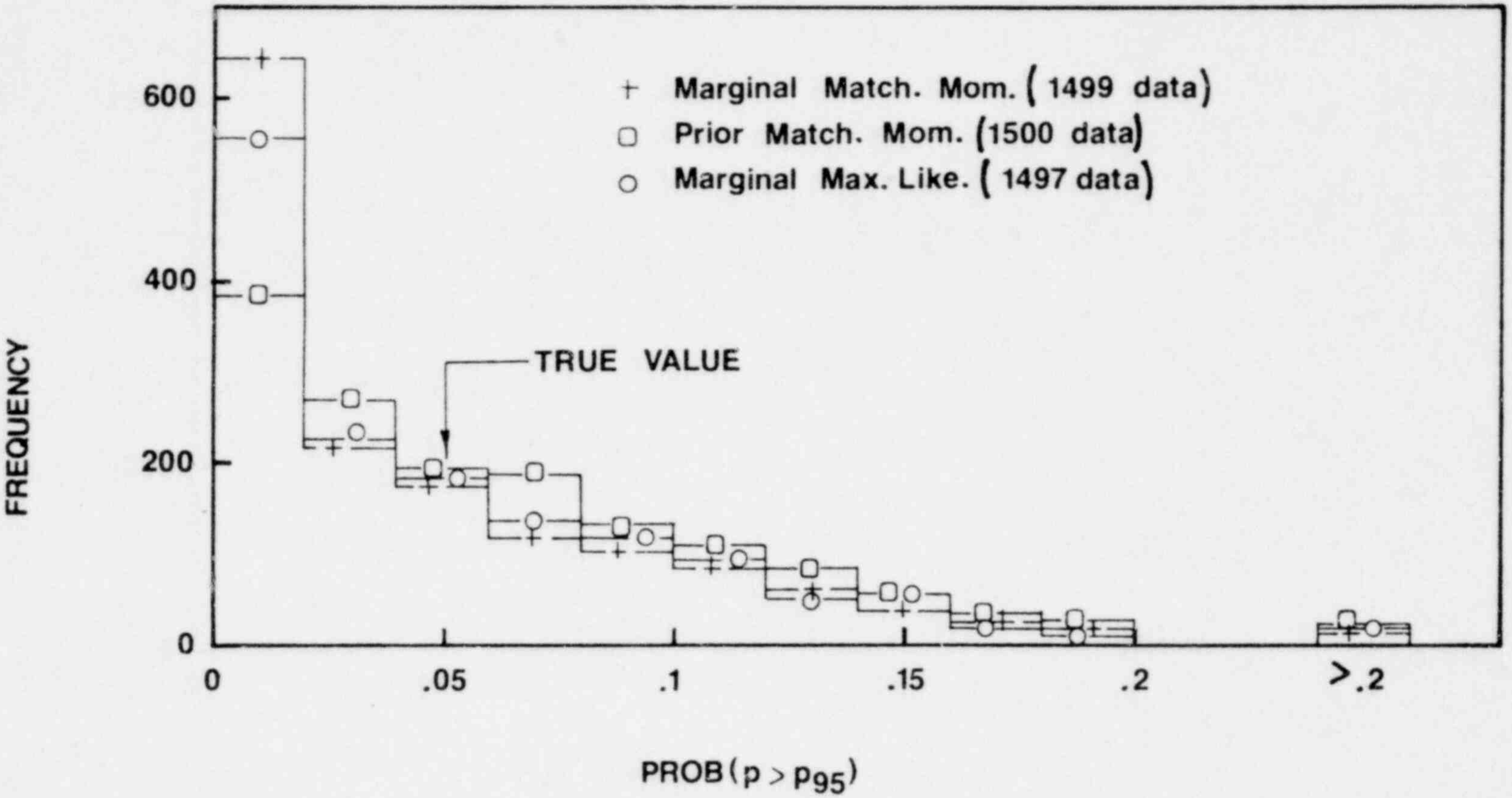


Fig. 4.23 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2, b=23). Size of samples used to obtain estimates was N=10.

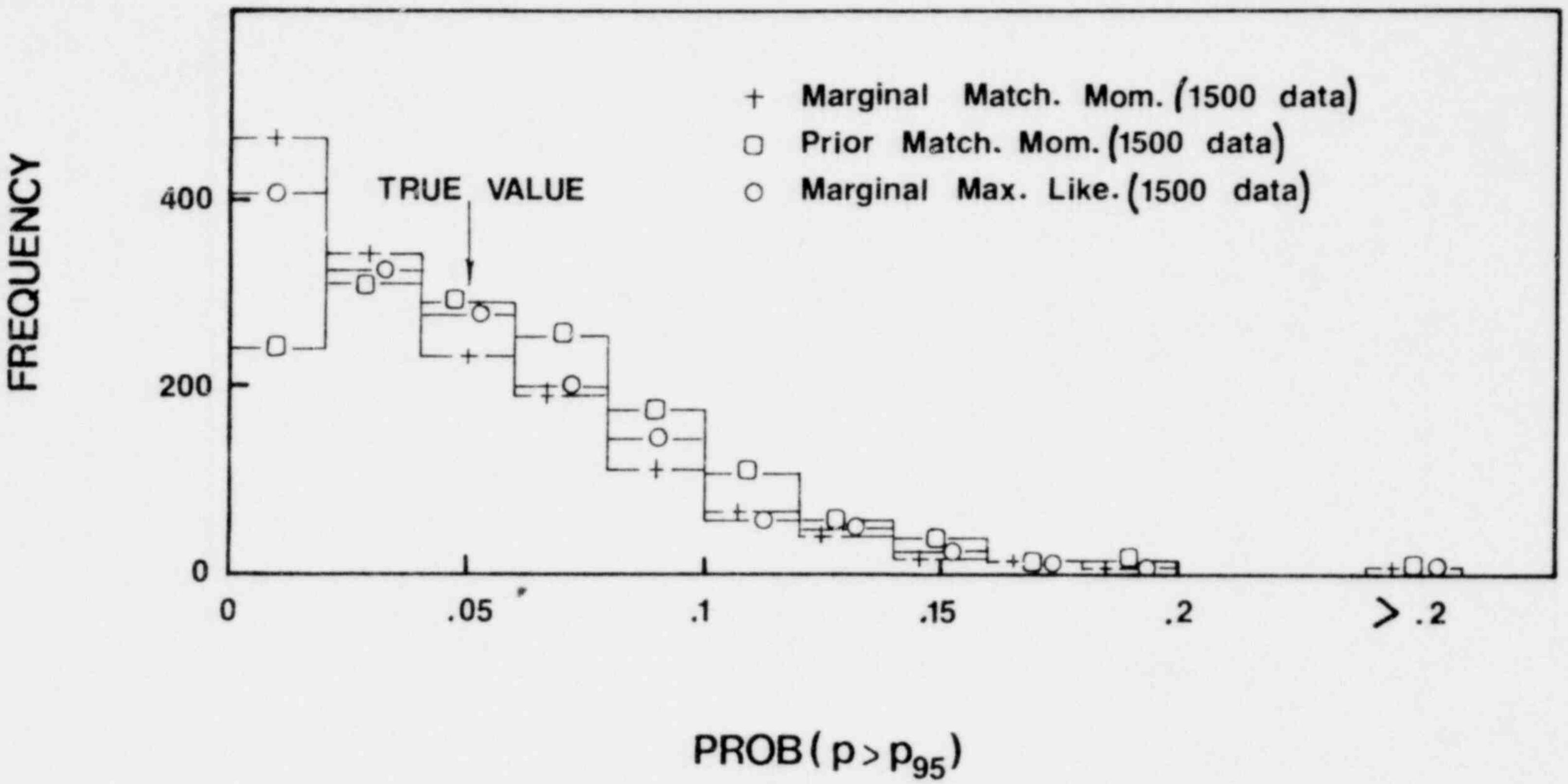


Fig. 4.24 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2, b=23). Size of samples used to obtain estimates was N=20.

1426 307

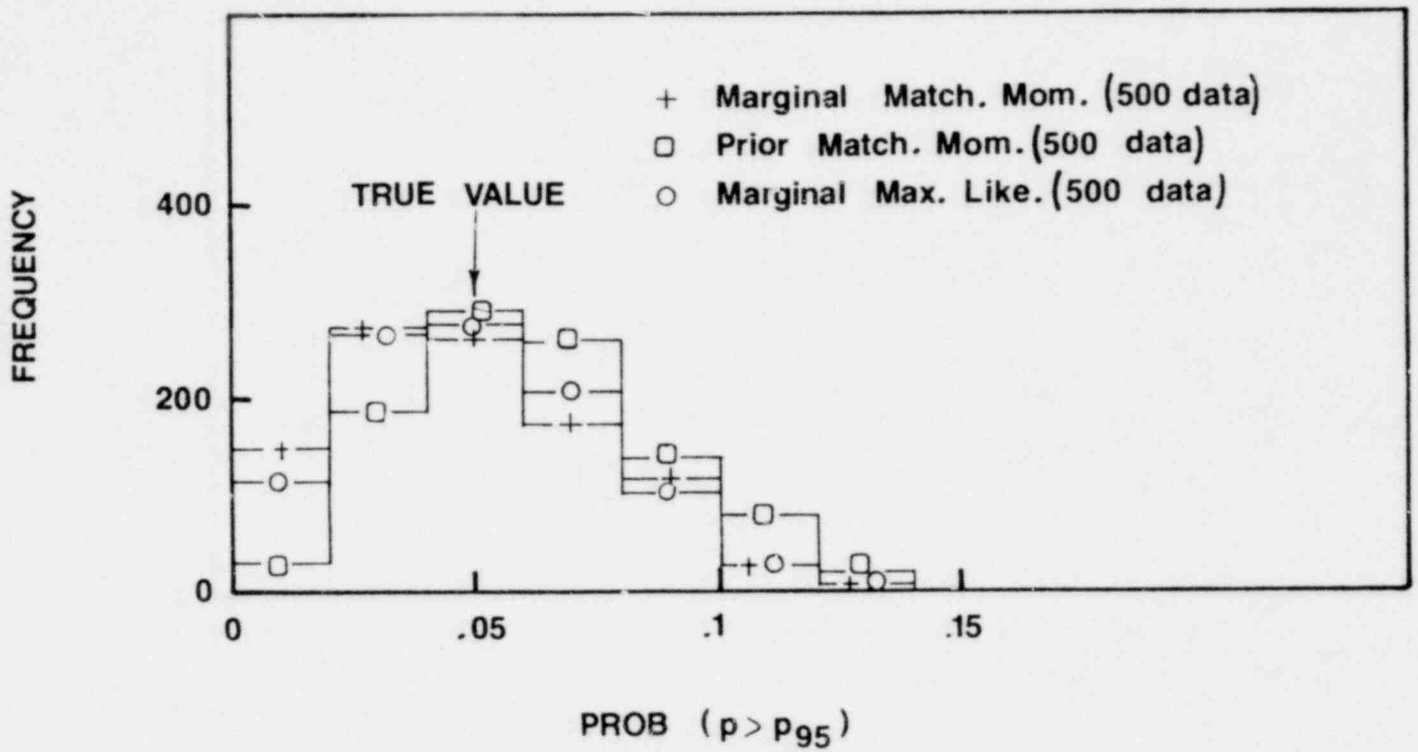


Fig. 4.24 Distribution of the fraction of the estimated beta prior distribution that lies above the 95-th percentile of the beta function used to generate the simulated failure data (a=1.2, b=23). Size of samples used to obtain estimates was N=50.

The median, mean and variance of these distributions are presented in Tal 4.13. From these results the variances for all three methods are within a few percent of each other although the mean for the prior matching moment distribution is considerably higher than that for the distributions produced by the marginal-based methods. Moreover, even for large sample sizes the mean of the distribution for the prior matching moments method is about 20% greater than the ideal value of 0.05. The marginal-based methods, in contrast, appear to approach the ideal value as the sample size becomes sufficiently large.

4.6 Comparison of Maximum Likelihood Variance Bounds to Measured Variances

In Section 3.6 expressions for the variance and covariance of the parameter estimators were derived for the marginal maximum likelihood method. Although these expressions are strictly asymptotic values, the expressions are often used as actual estimators of the variance or covariance of the parameter estimates for finite size data samples. Since the values of the variances and covariances of the parameter estimates are important for error propagation (see Section 3.8), one would like to know how close these maximum likelihood estimated values are to the true values of the variances and covariance.

Such a determination was started during this project and some preliminary results are presented in this section. The actual variances and covariance for the parameter estimators found in the simulation study are listed in Table 4.4. Because of the presence of estimator outliers for small sample sizes ($N \leq 10$) obtained with both marginal-based estimation techniques, the experimental values of variances and covariance depends greatly on how these outliers are treated. In this study estimators greater than 100 times the true beta parameter values ($a=1.2$, $b=23$) were ignored.

To evaluate the effectiveness of using the maximum likelihood expressions as estimators, simulated failure data samples were selected which produced either excellent or very poor parameter estimates. With these data samples the marginal maximum likelihood variance bounds were calculated from Eqs. (3.43)-(3.48). The results for the "good" and

Table 4.14 Median, mean and variance of the distribution for the Prob$\{p \geq p_{95}^{true}\}$. For samples of size 5, 10 and 20, 1500 simulated failure data sets were used, while for the size 50 sample, 500 sets were used. Beta prior parameters are a=1.2 and b=23.

Sample Size	Marginal Matching Moments			Prior Matching Moments			Marginal Maximum Likelihood		
	Median	Mean	Var.	Median	Mean	Var.	Median	Mean	Var.
5	0.0142	0.0425	0.0041	0.0321	0.0570	0.0045	0.0230	0.0493	0.0043
10	0.0287	0.0462	0.0025	0.0498	0.0616	0.0027	0.0363	0.0511	0.0026
20	0.0367	0.0456	0.0015	0.0532	0.0595	0.0015	0.0415	0.0489	0.0015
50	0.0467	0.0491	0.00068	0.0596	0.0618	0.00065	0.0478	0.0508	0.00065

"bad" data samples are shown in Table 4.15 and the data samples themselves are given in Table 4.16. From these results it is seen that the "bad" data samples which yield inordinantly large values for $\hat{a}$ and $\hat{b}$, also produce extremely large estimates for the variances and covariance and are much larger than the empirical estimates in Table 4.5.

Table 4.15 Variance bounds [bnd($\hat{a}$) and bnd($\hat{b}$)], and the covariance bounds [bnd($\hat{a}$,$\hat{b}$)] for parameter estimators [$\hat{a}$ and $\hat{b}$], as calculated by the marginal maximum likelihood method for selected simulated failure data samples. True values of the beta parameters is a=1.2 and b=23. The selected data samples are given in Table 4.16.

Sample Size	ID No.	$\hat{a}$	$\hat{b}$	bnd($\hat{a}$)	bnd($\hat{b}$)	bnd($\hat{a}$,$\hat{b}$)
5	1	1.2444	22.823	0.89839	393.129	16.179
	2	528.92	11338.	3.0843×10^{8}	1.417×10^{11}	6.6111×10^{9}
10	3	1.2673	23.541	0.42806	193.50	7.8072
	4	2080.8	40183.	3.9119×10^{10}	1.4589×10^{13}	7.5545×10^{11}
20	5	1.2248	22.720	0.20962	94.534	3.8150
	6	7.1495	137.61	19.074	7309.5	366.41
50	7	1.1728	23.094	0.076788	39.481	1.4846
	8	2.8889	58.522	0.67451	308.08	13.580

The maximum likelihood estimates for the "good" data samples appear much more reasonable and are generally smaller than the empirically observed variances listed in Table 4.5. To compare these maximum likelihood estimates to the variances and covariance measured from the distributions of the parameter estimators, the ratio of the measured value to the likelihood bound was calculated.

Table 4.16 Selected simulated failure data samples used to estimate variance bounds in Table 4.15. Data were simulated from a beta binomial with parameters a=1.2 and b=23. Data are read from left to right with the number of failures, k_i, following the number of tries, n_i.

Sample Size	ID No.	(n_i, k_i)									
5	1	45	4	216	5	213	25	92	0	260	9
	2	246	12	249	13	227	4	167	8	255	14
10	3	100	3	109	9	83	11	242	5	287	19
		247	4	116	6	248	5	195	21	256	0
	4	45	3	265	14	43	1	164	7	288	14
		44	4	180	15	247	13	163	4	247	8
20	5	46	4	43	1	276	35	139	0	168	16
		160	3	84	9	175	2	169	0	219	13
		264	37	271	22	247	12	111	4	106	1
		243	16	111	1	191	9	105	1	228	9
	6	227	4	91	5	287	17	184	3	121	10
		264	26	137	6	286	8	255	9	118	8
		175	7	128	3	31	2	225	12	150	11
		166	3	34	3	150	11	188	10	173	7
50	7	261	20	33	0	281	11	237	29	203	8
		157	35	227	7	44	1	245	6	59	1
		155	8	176	10	48	2	192	14	82	1
		241	7	150	25	255	4	265	3	131	4
		119	14	148	6	102	8	103	5	87	7
		266	0	137	0	178	1	261	34	280	2
		144	4	227	11	284	7	244	6	56	1
		184	3	101	4	196	2	213	3	125	16
		137	0	172	0	122	19	218	8	261	9
		80	7	60	2	254	16	241	5	263	6
	8	209	2	77	2	158	13	168	18	213	1
		209	19	63	0	196	9	30	2	104	1
		224	8	173	11	155	5	143	7	266	20
		250	27	42	0	290	17	153	7	101	4
		286	18	213	15	132	6	56	1	62	4
		68	3	273	14	199	2	116	4	80	3
		142	14	140	9	208	7	243	13	235	19
		287	12	204	0	167	8	300	16	262	8
		226	13	142	6	227	2	169	6	124	6
		165	7	267	3	97	8	163	15	193	1

1426 312

These ratios are presented in Table 4.17 for each of the three estimation techniques suitable for the low failure probability case studied. From these results it is seen that the empirical variances of the parameter estimator as determined by the prior matching moment technique are much closer to the likelihood estimates than are the variances for the estimators as determined by either of the marginal based techniques. The marginal-based estimators, $\hat{a}$ and $\hat{b}$, have empirical variances which are many times larger than the likelihood expressions for samples less than 20 in size, although the variances still appear to approach the bounds as the sample size becomes very large.

It should be emphasized that the above conclusions hold for particular examples of "good" failure data. Whether they hold true on the average for all data samples is the subject of further investigation. However, it is seen by the "bad" data samples used here, that the likelihood bounds are capable of yielding completely unrealistic values, and hence for the analysis of a single failure data sample, care must be used in using the likelihood bounds as estimates for the variances of the prior parameter estimators.

4.7 Bias Removal for the Prior Matching Moments Method

In Section 4.2 it was seen that all of the prior parameter estimation techniques produced a bias in the distribution of the estimators, $\hat{a}$ and $\hat{b}$, especially for small sample sizes. Ideally, one would like an expression for the amount of bias inherent in each estimator. Thus a cursory examination of the relation between parameter estimator bias and the sample size was undertaken. Since the prior matching moment estimation technique was found from several considerations, to be the best of the four techniques studied for analysis of low probability failure data, e.g., no outliers, smallest bias, simplest computationally, and most conservative in describing the high probability tail of the estimated prior, only this estimation technique was examined in the bias removal study.

To simplify the generation of failure data, random samples of the failure probability, p_i, were made directly from a known beta prior distribution, rather than to simulate failure-on-demand data, n_i and k_i, by sampling from a beta-binomial distribution as was done in all the previous

Table 4.17 Ratio of measured variances and covariances of the parameter estimators (listed in Table 4.5) to the marginal maximum likelihood bounds (bnd) (listed in Table 4.15) for the "good" data samples

	Prior Matching Moments			Marg. Max. Likelihood			Marg. Match. Moments		
Sample Size	$\frac{var(\hat{a})}{bnd(\hat{a})}$	$\frac{var(\hat{b})}{bnd(\hat{b})}$	$\frac{cov(\hat{a},\hat{b})}{bnd(\hat{a},\hat{b})}$	$\frac{var(\hat{a})}{bnd(\hat{a})}$	$\frac{var(\hat{b})}{bnd(\hat{b})}$	$\frac{cov(\hat{a},\hat{b})}{bnd(\hat{a},\hat{b})}$	$\frac{var(\hat{a})}{bnd(\hat{a})}$	$\frac{var(\hat{b})}{bnd(\hat{b})}$	$\frac{cov(\hat{a},\hat{b})}{bnd(\hat{a},\hat{b})}$
5	4.92	9.64	63.7	61.1	69.7	71.8	57.8	63.6	61.2
10	1.28	1.48	1.29	13.1	21.1	17.6	28.7	29.8	32.1
20	1.01	1.06	0.993	2.72	3.59	3.20	3.82	4.57	4.43
50	0.875	0.773	0.829	1.48	1.48	1.53	2.28	2.06	2.32

sections. The failure probability samples, p_i, were generated by the inverse transformation technique (described in Section 4.1) where a random number u was transformed to a failure probability p through the cumulative distribution of a beta distribution, i.e.,

$$u = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^p x^{a-1} (1-x)^{b-1} dx \; . \tag{4.14}$$

For a given value of u, the failure probability p can readily be obtained by solving the above equation numerically using techniques described in Chapter 5.

For this bias removal investigation, 500 failure probability samples of various sizes, N, were generated from two beta distributions:

Population 1: $a = 0.39$ $b = 6.14$

Population 2: $a = 3$ $b = 7$

Population 1 was selected because this beta was found to describe the prior distribution for a particular grouping of the diesel engine data of Table .1, while population 2 represents a more centered distribution. For each data sample, the sample mean and variance were calculated, and beta parameter estimators were obtained by the method of prior matching moments using Eqs. (3.5) and (3.6).

As would be expected from the earlier study on the estimators $\hat{a}$ and $\hat{b}$, these estimators were again highly biased towards the high values and $\hat{a}$ and $\hat{b}$ were highly correlated. The results are summarized in Table 4.18 where the average of the estimators (denoted by $\bar{a}$ and $\bar{b}$), their ranges, variances, and the coefficient of linear correlation (r) between $\hat{a}$ and $\hat{b}$ are tabulated.

There is one surprising difference between these results and those obtained in Section 4.2 from data simulated from the beta-binomial, i.e., using k_i and n_i data . The data simulated directly from the beta distribution always yielded estimators with positive bias whereas the earlier results indicated the bias becomes slightly negative for a sample size over 20. This difference is thought to arise because of the inability of the simulated data taken from the beta-binomial distribution to yield failure probabilities between k/n and (k+1)/n. The data

Table 4.18 Results of the beta parameter estimators as calculated by the prior matching moments technique from simulated failure probability data.

Population 1 (a = 0.39, b = 6.14)

N	$\bar{a}$	$\bar{b}$	min $\hat{a}$	max $\hat{a}$	min $\hat{b}$	max $\hat{b}$	var $\hat{a}$	var $\hat{b}$	r
10	0.633	13.4	0.0641	3.12	0.642	143.	0.147	201.	0.620
20	0.507	9.03	0.111	1.76	1.50	43.7	0.0481	32.1	0.708
40	0.449	7.45	0.130	1.02	2.31	2.60	0.0204	9.59	0.757
50	0.444	7.33	0.178	1.05	2.75	21.8	0.0165	8.36	0.740
60	0.432	7.05	0.182	0.845	2.67	16.8	0.0119	5.53	0.741
70	0.429	6.95	0.173	0.770	3.00	17.0	0.0108	4.67	0.743
80	0.425	6.84	0.167	0.792	2.66	15.4	0.0099	4.04	0.780
90	0.423	6.79	0.212	0.805	2.89	14.7	0.0084	3.41	0.754
100	0.418	6.67	0.199	6.836	3.56	12.9	0.0075	2.84	0.765

Population 2 (a = 3, b = 7)

n	$\bar{a}$	$\bar{b}$	min $\hat{a}$	max $\hat{a}$	min $\hat{b}$	max $\hat{b}$	var $\hat{a}$	var $\hat{b}$	r
10	4.06	9.51	0.933	30.6	2.18	49.5	7.60	43.7	0.923
20	3.42	8.04	1.10	13.2	2.76	28.4	1.75	10.6	0.921
30	3.29	7.71	1.32	7.63	2.71	16.9	1.00	5.98	0.922
40	3.19	7.47	1.52	7.04	3.47	17.8	0.644	3.86	0.917
50	3.18	7.44	1.84	5.58	3.86	14.5	0.506	3.11	0.906
60	3.13	7.32	1.79	5.07	3.83	12.7	0.369	2.34	0.907
70	3.12	7.29	1.99	5.36	4.25	13.0	0.341	2.09	0.911
80	3.09	7.22	1.87	4.96	3.85	12.1	0.2681	1.52	0.901
90	3.09	7.21	1.93	5.16	4.12	12.7	0.223	1.31	0.889
100	3.06	7.15	1.844	4.64	4.53	11.0	0.194	1.14	0.893

simulated from the beta distribution, on the other hand, may assume non-fractional values and be more smoothly distributed.

From the results in Table 4.19, it is seen that the bias on the parameter estimates (i.e., $\bar{a}$-a or $\bar{b}$-b) decreases with increasing sample size, N. In an attempt to find an empirical expression for the bias of the estimators the following two models were used:

$$\text{Exponential:} \quad \text{bias} = \alpha\, n^{\beta}$$
$$\text{Linear:} \quad \text{bias} = \gamma + \delta n^{-1} .$$

The coefficients for each model were computed by fitting each model to the bias given in Table 4.18 by the methods of least squares. (For the exponential model the logarithm was taken before performing the least squares analysis.) The values of the coefficients so obtained and the coefficient of determination, R^2, for each fit are given in Table 4.19.

The high values of R^2 for both models implies that either model may be considered satisfactory for estimating bias. Furthermore, the fact that β is close to the value -1 in all cases implies that there is not much practical difference between the two models. What is distressing is that the values of α, γ, and δ are so disparate. It had been hoped that these coefficients would be sufficiently close in the four cases that the same bias-removing formula could be used for all beta parameters a and b. Clearly these coefficients are functions of these parameters. Further work to find a bias-removing factor (or term) that is independent of the true values of a and b is needed. No use has been made so far of the high correlation between $\hat{a}$ and $\hat{b}$, and this should also be incorporated into future studies.

4.8 Fit of Empirical Distribution for $\hat{a}$ and $\hat{b}$ to the Gamma and Log Normal Distributions

In the study of the distribution of the beta parameter estimators, a preliminary investigation was undertaken to see if these empirically derived distributions could be described adequately by a simple model. For this investigation the estimator distributions obtained in the previous Section 4.7 by the prior matching moments technique for the

Table 4.19 Least squares coefficients for the bias predicting formulas.

	R^2	α	β	γ	δ
POP'N 1: $\bar{a}$					
Exponential	.9931	1.8704	-.9125		
Linear	.9997			.00415	2.3619
POP'N 1: $\bar{b}$					
Exponential	.9513	82.0148	-1.0915		
Linear	.9906			-.30437	73.0225
POP'N 2: $\bar{a}$					
Exponential	.9952	13.7106	-1.1398		
Linear	.9915			-.04964	10.5452
POP'N 2: $\bar{b}$					
Exponential	.9971	34.2282	-1.1439		
Linear	.9934			-.1132	25.5715

simulated failure data generated directly from the beta distribution were used. Both the shifted and unshifted gamma and log normal distributions were fit to the empirical distributions. The results of this modelling of the estimator distributions are summarized in Section 4.8.1 and 4.8.2.

4.8.1 The Gamma Model

The first model fit to the observed estimator distributions was the gamma distribution

$$f(v|\alpha,\beta) = \frac{v^{\alpha-1} e^{-v/\beta}}{\Gamma(\alpha)\ \beta^{\alpha}}\ , \quad 0 \leq v < \infty\ , \qquad (4.15)$$

where v represents either estimator $\hat{a}$ or $\hat{b}$. Values for the gamma parameters α and β were obtained by equating, respectively, the variance, s^2, and mean, $\bar{v}$, of the empirical estimator distribution to the mean, $\alpha\beta$, and variance, $\alpha\beta^2$, of the gamma distribution. The resulting estimates for the gamma parameters are thus

$$\hat{\alpha} = \bar{v}^2/s^2 \qquad (4.16)$$

and

$$\hat{\beta} = s^2/\bar{v}\ . \qquad (4.17)$$

The results of these fits to several distributions were not encouraging as can be seen from Table 4.20 in which are presented the results of a χ^2 goodness-of-fit test using 20 equi-probability intervals in v (and thus 17 degrees of freedom).

Table 4.20 χ^2 Goodness-of-fit results for the gamma model. The critical values of χ^2 for the test are: $\chi^2_{.05}(17)=27.59$, $\chi^2_{.01}(17)=33.41$, $\chi^2_{.005}(17)=35.72$

Sample Size	Beta Population	χ^2 $\hat{a}$	χ^2 $\hat{b}$
10	1	31.68*	141.36*
50	1	14.56	34.64*
100	1	16.96	31.44*
10	2	97.76*	103.36*
50	2	19.04	24.64
100	2	23.76	19.60

*Indicates a significant difference at the 0.05 level or lower.

Upon examination of the estimator distribution within the 20 equal probability cells, it was found that for cases which yielded large χ^2 values there were disproportionately fewer estimates in the cells for small values of v. This underpopulation in the initial cells results in the large χ^2 values. In other words, the fitted gamma model predicted far more small v values than were observed in the simulation results.

This emphasis of the gamma distribution for small v values suggests that instead of the usual two parameter gamma function, a three parameter shifted gamma function might be a useful model to fit to the empirical distributions. The shifted gamma function is given by

$$f(w|\alpha,\beta,\theta) = \frac{(w-\theta)^{\alpha-1} e^{-(w-\theta)/\beta}}{\Gamma(\alpha)\ \beta^{\alpha}} \ , \quad \theta \leq w < \infty \ , \qquad (4.18)$$

where $w=v+\theta$. For a given θ, the estimates for the parameters α and β can be obtained, as before, by matching the mean and variance of the gamma model to those of the empirical distribution. The result is given by Eqs. (4.16) and (4.17) or equivalently by

$$\hat{\alpha} = (\bar{w}-\theta)^2/S^2 \ , \qquad (4.19)$$

and

$$\hat{\beta} = S^2/(\bar{w}-\theta) \quad . \qquad (4.20)$$

The choice of a value for θ, however is not so straightforward Clearly θ must be constrained between zero and the minimum observed value for v. Ideally θ should be chosen so as to minimize the χ^2 statistic. Such a technique would require computer analysis; but for this preliminary investigation on modeling the estimator distributions, a more cursory treatment was indicated. The shift parameter θ was given several values between zero and the minimum v observed.

While this increase in θ generally lowered the χ^2 statistic, it was found that the best χ^2 values were still too large for the fit by a shifted gamma model to be acceptable. For example, the case for $\hat{b}$ from sample size 10 generated from the population 1 beta, the χ^2 statistic decreased from 141.36 for $\theta=0$ to 118.88 for $\theta=0.32$ to 115.92 for $\theta=0.6395$. From these and other examples it is concluded that neither a gamma or a shifted gamma distribution is a reasonable model for the empirical $\hat{a}$ or $\hat{b}$ distributions.

4.8.2 The Log Normal Model

As an alternative to the gamma distribution, the log normal distribution was also investigated as a possible model for the $\hat{a}$ and $\hat{b}$ distributions. In this model it was assumed that $\ln v$ is distributed normally, i.e.,

$$f(v|\alpha,\beta) = \frac{1}{\sqrt{2\pi}\,\beta} \exp[-\frac{(\ln v-\alpha)^2}{2\beta^2}] , \quad 0 < v < \infty. \tag{4.21}$$

Again estimates of the parameters α and β are obtained by matching the mean, $\bar{v}$, and variance, s^2, of the empirical distribution to the mean and variance of the log normal distribution, respectively. The mean and variance of the log normal distribution are

$$\mu = \exp[\alpha + \beta^2/2] \tag{4.22}$$

and

$$\sigma^2 = \mu^2 (e^{\beta^2}-1) . \tag{4.23}$$

The inverse relations are

$$\beta^2 = \ln[1 + \sigma^2/\mu^2] \tag{4.24}$$

and

$$\alpha = \ln\mu - \beta^2/2 . \tag{4.25}$$

Thus the estimates $\hat{\alpha}$ and $\hat{\beta}$ are obtained by replacing μ and σ in the above equations by $\bar{v}$ and s respectively.

With Eqs. (4.24) and (4.25), log normal distributions were fit to the same example $\hat{a}$ and $\hat{b}$ distributions as were used in the preceding gamma analysis. Again a χ^2 goodness-of-fit test using 20 equi-probability intervals was used to compare the fit to the empirical distribution. The results, which are much more encouraging, are shown in Table 4.21.

Table 4.21 χ^2 goodness-of-fit results for the log normal model. The critical values of χ^2 for the test are: $\chi^2_{.05} = 27.59$, $\chi^2_{.01} = 33.41$, $\chi^2_{.005} = 35.72$.

Sample Size	Beta Population	χ^2 $\hat{a}$	χ^2 $\hat{b}$
10	1	14.64	27.28
50	1	22.32	14.08
100	1	12.40	15.52
10	2	35.76*	32.16*
50	2	20.40	13.36
100	2	19.84	24.24

*Indicates a significant difference at the 0.05 level.

Most of the computed χ^2 values indicate adequate fits to the log normal model and those which show poor fits are, as might be expected, for the small sample sizes. Thus the log normal appears to fit the data much better than the gamma models (see Tables 4.20 and 4.21).

However, there is an indication that a better model could be found. Inspection of the frequency of observed data ($\hat{a}$ or $\hat{b}$ values) in the lower probability intervals used for the χ^2 analysis again showed that these intervals were populated with fewer than expected observations, and hence made the largest contribution to the calculated χ^2 values. To increase the population in the lower probability cells, a shifted log normal distribution,

$$f(w|\alpha,\beta,\theta) = \frac{1}{\sqrt{2\pi}\,\beta} \exp\left[-\frac{(\ln(w-\theta)-\alpha)^2}{2\beta^2}\right], \qquad (4.26)$$

could be used. The shift parameter must be constrained between 0 and the smallest observed $\hat{a}$ or $\hat{b}$. For a fixed θ, the parameters α and β can be found by matching moments to those of the empirical distribution. In this way one finds

$$\hat{\beta}^2 = \ln[1 + s^2/(\bar{w}-\theta)^2] \quad (4.27)$$

and

$$\hat{\alpha}' = \ln(\bar{w}-\theta) - \beta^2/2 \; . \quad (4.28)$$

To fit Eq. (4.26) to the empirical distributions, the shift parameter was varied to find the value which yielded the lowest χ^2 statistic. It was found that the use of a non-zero value for θ decreased the goodness-of-fit statistic, χ^2. (However, it must be remembered that use of a non-zero θ reduces the degrees of freedom from 17 to 16). Some results are shown in Table 4.22 where it is seen that the fits for small sample sizes have been greatly improved over the non-shifted log normal and gamma models. In fact all of the example distributions have an acceptable χ^2 value.

Table 4.22 χ^2 goodness-of-fit results for the shifted log normal model. For θ=0 critical value $\chi^2_{.05}(17) = 27.59$, while for θ>0 the critical value is $\chi^2_{.05}(16)$ = [illegible].30.

Sample Size	Beta Population	χ^2 a	χ^2 $\hat{b}$
10	1	14.64 (θ=0)	27.28 (θ=0)
			18.88 (θ=0.3)
50	1	22.32 (θ=0)	14.08 (θ=0)
			18.24 (θ=1)
100	1	12.4 (θ=0)	15.52 (θ=0)
			14.00 (θ=1)
10	2	35.76 (θ=0)*	32.16 (θ=0)*
		14.56 (θ=0.6)	21.36 (θ=1.1)
50	2	20.40 (θ=0)	13.36 (θ=0)
		22.24 (θ=1)	13.44 (θ=1)
100	2	19.83 (θ=0)	24.24 (θ=0)
		17.44 (θ=1)	

*Indicates a significant difference at the 0.05 level or lower.

5. CALCULATION OF CONFIDENCE AND PROBABILITY INTERVALS FOR COMPONENT FAILURE PROBABILITIES

In the previous chapter, techniques were developed to estimate the mean failure probability of plant components from the observed number of failures and the sample size. Both the classical and Bayesian estimation techniques were analyzed and applied to diesel engine failure data.

This chapter represents an extension of this estimation work. In particular, the question of the confidence of the failure probability estimates is examined. Of prime concern is the determination of a "confidence interval" for the classical description (or a "probability interval" for the Bayesian description) into which the true failure probability of a particular component falls with an associated degree of certainty (or "confidence level"). The question of such interval determination is reviewed for both the classical and Bayesian descriptions.

5.1 Classical Estimation of Confidence Levels

The classical description of the failure probability distribution for obtaining k failures in n tries is given by the binomial distribution

$$f(k|n,p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}, \tag{5.1}$$

where p is the failure probability. For an observed k failures in n attempts the failure probability can be estimated by $\hat{p} = k/n$. With what degree of precision is this estimate made? Equivalently, what is the maximum (or minimum) reasonable value of p for which we would expect to obtain the observed k failures in n tries at some confidence level α?

The probability of observing k or fewer failures in n tries is

$$F(k|n,p) = \sum_{\ell=0}^{k} \frac{n!}{(\ell)!(n-\ell)!} p^\ell (1-p)^{n-\ell} \tag{5.2}$$

i.e., the cumulative distribution of the binomial. For a fixed n and k (observed values), F will decrease (increase) continuously as p increases (decreases). Thus the maximum reasonable value of p at the *a-level*, is that value of the failure probability, p_1, for which one would observe, with a probability of $\alpha/2$, k <u>or fewer</u> failures in n tries, i.e.,

$$F(k|n,p_1) = \alpha/2 \quad . \tag{5.3}$$

Similarly the minimum reasonable value of the failure probability at the α-level, is that value, p_o, for which the probability of observing k <u>or more</u> failures in n tries is α/2, i.e.,

$$1 - F(k-1|n,p_o) = \alpha/2 \ . \tag{5.4}$$

To find the upper and lower bounds of the component failure probability, Eqs. (5.3) and (5.4) must be solved for p_1 and p_o. However such solutions require numerical evaluation, and it is easier to convert these equations into a form more amenable to numerical analysis. In particular, the cumulative binomial distribution, Eq. (5.2), can be written in terms of the incomplete beta function. To find this relation, differentiate Eq. (4.2) with respect to p and simplify the result to obtain

$$\frac{\partial F(k|n,p)}{\partial p} = - \frac{p^k(1-p)^{n-k-1}}{B(k+1,n-k)} \ , \tag{5.5}$$

where

$$B(x,y) \equiv \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \ . \tag{5.6}$$

Integration of Eq. (5.5) over p from 0 to p yields

$$F(k|n,p) - F(k|n,o) = - \int_0^p \frac{z^k(1-z)^{n-k-1}}{B(k+1,n-k)} dz, \tag{5.7}$$

or equivalently

$$F(k|n,p) = 1 - I_p(k+1,n-k) \ , \tag{5.8}$$

where the incomplete beta function I_p is defined by

$$I_p(a,b) \equiv \frac{1}{B(a,b)} \int_0^p z^{a-1} (1-z)^{b-1} dz. \tag{5.9}$$

With this relation between F and I_p, the equations which determine the upper and lower bounds on p may be written as

$$I_{p_o}(k,n-k+1) = \alpha/2 \tag{5.10}$$

and

$$I_{p_1}(k+1,n-k) = 1 - \alpha/2. \tag{5.11}$$

The advantage of this form, which still must be solved numerically for p_o and p_1, is that the corresponding probability limits for the Bayesian analogue are given by equations of the same functional form, and the same numerical algorithm used to solve the above equation can be used in the Bayesian analysis.

It is easily shown that $p_o \leq \hat{p} \equiv k/n \leq p_1$, with the equality defined only if k=0 ($p_o=\hat{p}=0$) or with k=n ($p_1=\hat{p}=1$).* Of special interest are situations involving events with low probabilities of failure, for which one often encounters observed values of k=0 for relatively large values of n. For this case, the upper bound, p_1, can be obtained analytically. From Eq. (5.11) one obtains

$$\frac{\alpha}{2} = 1 - n \int_0^{p_1} (1-z)^{n-1}\, dz = (1-p_1)^n ,$$

or upon solving for p_1

$$p_1 = 1 - [\tfrac{\alpha}{2}]^{1/n} , \quad k=0. \tag{5.12}$$

Similarly for high probability events for which one often observes k=n (and for which $\hat{p}=p_1=1$), Eq. (5.10) yields

$$\frac{\alpha}{2} = n \int_{p_o}^{1} z^{n-1}\, dz = 1 - p_o^{\,n},$$

or solving for p_o,

$$p_o = (1 - \tfrac{\alpha}{2})^{1/n} . \tag{5.13}$$

5.2 Bayesian Estimation of Probability Intervals

In the Bayesian description of the failure probability for a component, it is assumed that the failure probability comes from a particular *prior distribution* which is known from previous experience or which is

*For k=0, the integrand on the left hand side of Eq. (5.10) becomes singular and the equation has no solution. In this case the entire confidence level is often associated with the "upper tail" of the distribution. However to be consistent with the more general case (k≠0,n), we will always associate only half of the total confidence level with each end of the tail. A similar convention is used with the k=n case.

1426 326

assumed. For the present study, we have assumed that the prior distribution is given by a beta distribution

$$g(p) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} \quad , \quad (a,b>0) \; . \tag{5.14}$$

If we assume, as in the classical case, the failure distribution is given by a binomial distribution (Eq. (5.1)), then the use of Bayes' theorem gives for the *posterior* distribution

$$\xi(p|k,n,a,b) = \frac{p^{a+k-1}(1-p)^{b+n-k-1}}{B(a+k,b+n-k)} \; . \tag{5.15}$$

This quanti y (also a beta distribution), is the Bayesian estimate of the distribution of the failure probability, p, for a particular component which has previously experienced k failures in n tries and which is assumed to belong to a class of components whose failure probabilities are distributed according to Eq. (5.14).

With the posterior distribution, the probability intervals about the mean of the posterior distribution,

$$\hat{p} = \frac{a+k}{(a+k) + (b+n-k)} \; , \tag{5.16}$$

are readily formulated for a component which has experienced k failures in n tries. Explicitly the probability that the true failure probability is greater than some upper bound, p_1, at the $\alpha/2$ level is given by

$$\text{Prob}\{p>p_1\} = \frac{\alpha}{2} = \int_{p_1}^{1} \xi(p|k,n,a,b)\, dp \; . \tag{5.17}$$

Similarly the probability that p is less than some lower bound, p_o, at the $\alpha/2$ level is

$$\text{Prob}\{p<p_o\} = \frac{\alpha}{2} = \int_0^{p_o} \xi(p|k,n,a,b)\, dp. \tag{5.18}$$

Upon substitution for ξ, the confidence limits are readily expressed in terms of the incomplete beta function as

$$I_{p_o}(a+k,n+b-k) = \alpha/2 \tag{5.19}$$

and

$$I_{p_1}(a+k,n+b-k) = 1 - \alpha/2 \ . \qquad (5.20)$$

Again these equations have the same form as those for the classical case (Eqs. (5.10) and (5.11)), although with different arguments for the incomplete beta function.

5.3 Solution for Interval Limits in Terms of the Snedecor F-Distribution

For other than the extreme cases when one of the arguments of the incomplete beta function equals zero, Eqs. (5.10) and (5.11) or Eqs. (5.19) and (5.20) cannot be solved analytically for p_o and p_1. However, the solutions can be expressed in terms of the inverse values of the Snedecor F-distribution [13] (also known as the variance-ratio distribution [8]). Consider the general form of Eqs. (5.9), (5.10), (5.19), or (5.20), namely

$$\int_{p_i}^{1} \frac{z^{x-1}(1-z)^{y-1}}{B(x,y)} \, dz = \beta \ . \qquad (5.21)$$

With the change of variable $z=w/(1+w)$, one obtains

$$\frac{1}{B(x,y)} \int_{w_i}^{\infty} w^{x-1}(1+w)^{-(x+y)} \, dw = \beta \qquad (5.22)$$

where $w_i = p_i/(1-p_i)$. To solve for w_i, let $w = \nu_1 F/\nu_2$ with $\nu_1 = 2x$ and $\nu_2=2y$. With this substitution, Eq. (5.22) becomes

$$\frac{\nu_1}{\nu_2 B[\frac{\nu_1}{2},\frac{\nu_2}{2}]} \int_{F_i}^{\infty} [\frac{\nu_1}{\nu_2} F]^{(\nu_1-2)/2} [1 + \frac{\nu_1}{\nu_2}F]^{-(\nu_1+\nu_2)/2} \, dF = \beta \quad (5.23)$$

where $F_i = \nu_2 w_i/\nu_i$. The quantity on the left hand side of the Eq. (5.23) is the cumulative distribution of the Snedecon F-distribution between F_i and ∞. The solution of Eq. (5.23) is often denoted by

$$F_i = F_\beta(\nu_1,\nu_2) \qquad (5.24)$$

1426 328

where values of F_β are tabulated for integral values of ν_1 and ν_2 [13].

$$p_i = \frac{w_i}{1+w_i} = [1 + \frac{\nu_2}{\nu_1} \frac{1}{F_\beta(\nu_1,\nu_2)}]^{-1}$$

or

$$p_i = [1 + \frac{y}{x} \frac{1}{F_\beta(2x,2y)}]^{-1} = [1 + \frac{y}{x} F_{1-\beta}(2y,2x)]^{-1} \qquad (5.25)$$

Only for the classical results of Eqs. (5.10) and (5.11) do the parameters x and y (and hence ν_1 and ν_2) always assume integer values and therefore standard tables of F_β can be used. Even most computer programs written to calculate F_β require that the "degrees of freedom" parameters ν_1 and ν_2 be integer values. Consequently the above reduction is of little practical consequence for the calculation of the Bayesian estimates of the confidence limits.

5.4 Approximate Solution for the Interval Limits

As an alternative to the above procedure, the exact interval bound equations (Eqs. (5.10) and (5.11) or Eqs. (5.19) and (5.20)) can be expressed approximately in terms of the Chi-squared distribution [8], i.e.,

$$P(\chi^2|\nu) = [2^{\nu/2}\,\Gamma(\nu/2)]^{-1} \int_0^{\chi^2} t^{\frac{\nu}{2}-1}\, e^{-t/2}\, dt\ , \quad 0 \le \chi^2 < \infty\ , \qquad (5.26)$$

where ν is the *degrees of freedom*. Consider the general form of the exact interval equation, Eq. (5.21), which can be written as

$$I_p(x,y) \equiv \int_0^p \frac{z^{x-1}(1-z)^{y-1}}{B(x,y+1)} = \beta\ . \qquad (5.27)$$

Upon change of variables u=yz and the use of Eq. (5.6), this equation can be written as

$$\beta = \frac{\Gamma(x+y)}{\Gamma(y)} \frac{1}{y^x} \frac{1}{\Gamma(x)} \int_0^{yp} u^{x-1}\,(1 - \frac{u}{y})^{y-1}\, du\ . \qquad (5.28)$$

For large y, $(1 + \frac{a}{y})^y \approx e^a$, and with Stirling's approximation for $\Gamma(x+y)$ and $\Gamma(y)$ one has for large y

$$\frac{\Gamma(x+y)}{\Gamma(y)} \frac{1}{y^x} \approx 1\ . \qquad (5.29)$$

Thus Eq. (4.28) may be approximated for large y by

$$\beta \simeq \frac{1}{\Gamma(x)} \int_0^{up} u^{x-1} e^{-u} du \equiv P(2yp|\nu) \qquad (5.30)$$

with $\nu=2x$. If the solution χ^2_β is defined by $P(\chi^2_\beta|\nu) = \beta$, the solution of Eq. (5.30) for p (and the approximate solution of Eq. (5.27)) can be written as

$$p \simeq \chi^2_\beta/(2y) \ . \qquad (5.31)$$

As an example, consider the solution of Eq. (5.11) for p_1 when k=0. For this case x=1, y=n, and $\beta=1-\alpha/2$. Equation (5.30) can be solved directly when x=1, namely

$$\beta \simeq \frac{1}{\Gamma(1)} \int_0^{np_1} e^{-u} du = 1 - e^{-np_1} \ . \qquad (5.32)$$

Solving for p_1, one obtains

$$p_1 \simeq -\frac{1}{n}\, ln(1-\beta) = -\frac{1}{n}\, ln(\tfrac{\alpha}{2}) \ . \qquad (5.33)$$

Use of a series expansion for the logarithm reduces this result, in the limit of large n, to the exact result of Eq. (5.12). For n=69 and $\alpha=0.50$, Eq. (5.33) yields $p_1 \simeq 0.02009$ which is only 1% higher than the exact value of $p_1=0.01989$.

The approximate interval equation, Eq. (5.30) or (5.31), cannot be solved analytically except for the case x=1 (k=0). However the use of the approximately χ^2-distribution is often preferable to the exact expression in terms of the Snedecor F-values (Eq. (5.25)) because the values χ^2_β are extensively tabulated (albeit for integral degrees of freedom, ν). However, even for the Bayesian description, for which non-integral values of ν results, interpolation of χ^2_β tables is readily effected and approximate solution for the interval limits, p_i, (via Eq. (5.31)) can be obtained. In Fig. 5.1, a comparison between the approximate and exact values of p_1 of the classical description is presented. The agreement is excellent except for very small values of n.

5.5 Numerical Evaluation of Interval Bounds

A computer program TAILS was developed to solve the general form of the confidence interval equation (Eqs. (5.10) and (5.11) or Eqs. (5.19) and (5.20)), i.e.,

$$I_p(x,y) = \beta \ , \tag{5.34}$$

for the value of p (given x,y, and β). The complete program is listed in Appendix III.

The incomplete beta function $I_{p_i}(x,y)$ is calculated from the following expression [14]

$$I_p(x,y) = \frac{\text{INFSUM}\ p^x\ \Gamma(PS+x)}{\Gamma(PS)\ \Gamma(x+1)} + \frac{p^x\ (1-p^y\ \Gamma(x+y)\ \text{FINSUM}}{\Gamma(x)\ \Gamma(y+1)} \tag{5.35}$$

where INFSUM and FINSUM represent two series summations defined as follows:

$$\text{INFSUM} = \sum_{j=1}^{\infty} \frac{x(1-PS)_j}{x+j} \frac{p^j}{j!} \ , \tag{5.36}$$

where

$$(1-PS)_j = \begin{cases} 1 \ , & j = 0 \\ \Gamma(1+y-PS)/\Gamma(1-PS) \ , & j > 0 \end{cases} \tag{5.37}$$

and

$$\text{FINSUM} = \sum_{j=1}^{[y]} \frac{y(y-1)\ldots(y-j+1)}{(x+y-1)(x+y-2)\ldots(x+y-j)} \frac{1}{(1-p)^j} \tag{5.38}$$

where [y] is equal to the largest integer less than y. If [y]=0, the FINSUM=0. The quantity PS is defined as

$$PS = \begin{cases} 1 & \text{if } y \text{ is integer} \\ y - [y] \ , & \text{otherwise} \ . \end{cases} \tag{5.39}$$

The above algorithm (combined with scaling to avoid numerical inaccuracies encountered when using the gamma function with large arguments) was incorporated into a FORTRAN program MDBETA by Bosten and Battiste [14]. This program (modified in accordance to remarks made by Pike and Soo Hoo [14]) was used in the present analysis. The program MDBETA is significantly more accurate than the widely used program BDTR [13], especially for large arguments. For example in the case p=0.5, x=y=2000, MDBETA gives the correct value, 0.5, while BDTR gives 0.497026.

Once the incomplete beta function can be evaluated numerically, Eq. (5.34) is readily solved by standard numerical root finding techniques. The solution of Eq. (5.34) for p is limited to the left and the right by $0 \leq p \leq 1$, and consequently a "bracketing" technique, i.e., one which successively approaches the solution from opposite sides, is well suited to this problem. The procedure RIMI [13], which solves non-linear equations by means of Mueller's iteration scheme of successive bisection and inverse parabolic interpolation, was found to be effective.

5.6 Numerical Results

With the program listed in Appendix III, sample calculations of confidence intervals were obtained for the low failure probability events characteristic of the diesel generators in nuclear power plants. Of special concern are those records in which zero failures are observed in n startups. Classically the upper failure probability for the classical description is given by Eq. (5.12); however, the Bayesian description requires the numerical solution of Eq. (5.20). Results are shown in Figs. 5.1 and 5.2.

For most of the diesel engine failure data studied in this project, Bayesian estimates of the prior beta distribution parameters of Eq. (5.14) were approximately given by a=1, b=20. For this case it is found that the Bayesian estimate of the upper limit of the failure probability, p_1, was always less than the classical estimate (see Fig. 5.1). For example, for k=0 and n=69, the upper limit on the classical failure probability at the $\alpha/2$ = 25% confidence level is 0.02, while to achieve the same upper limit with the Bayesian estimates one has only to observe zero failures in 49 startups. In fact for the case a=1, b=20 the Bayesian description requires about 20 fewer startups to achieve the same upper confidence limit when k=0 for all confidence levels! This reduction in the number of startups required to estimate a given upper limit on the failure probability with the Bayesian description, makes this particular description quite attractive for establishment of initial acceptance criteria or maintenance criteria for diesel generators.

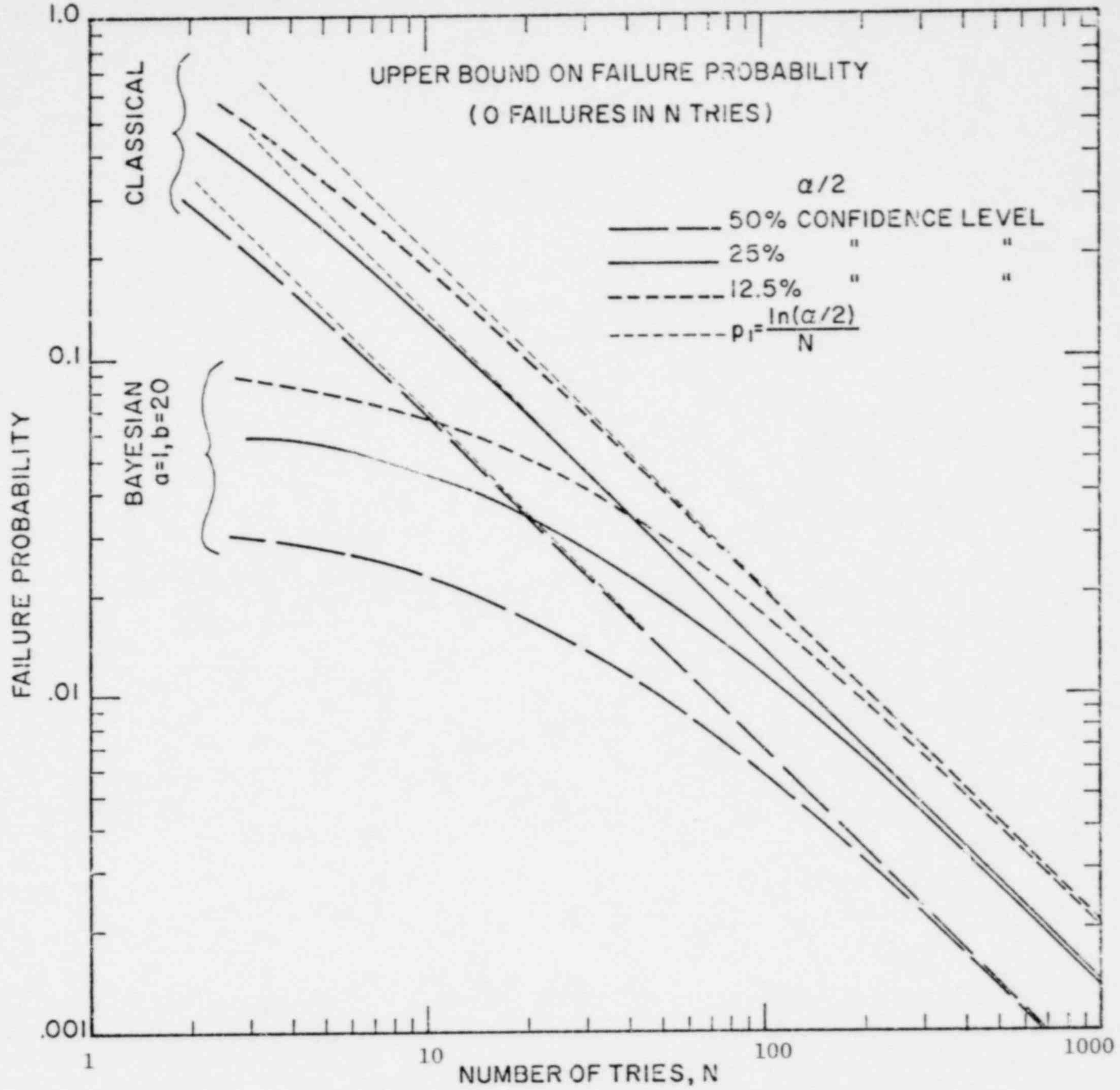


Fig. 5.1 Comparison of classical and Bayesian estimates of the upper limit on the failure probability at various confidence levels.

1426 333

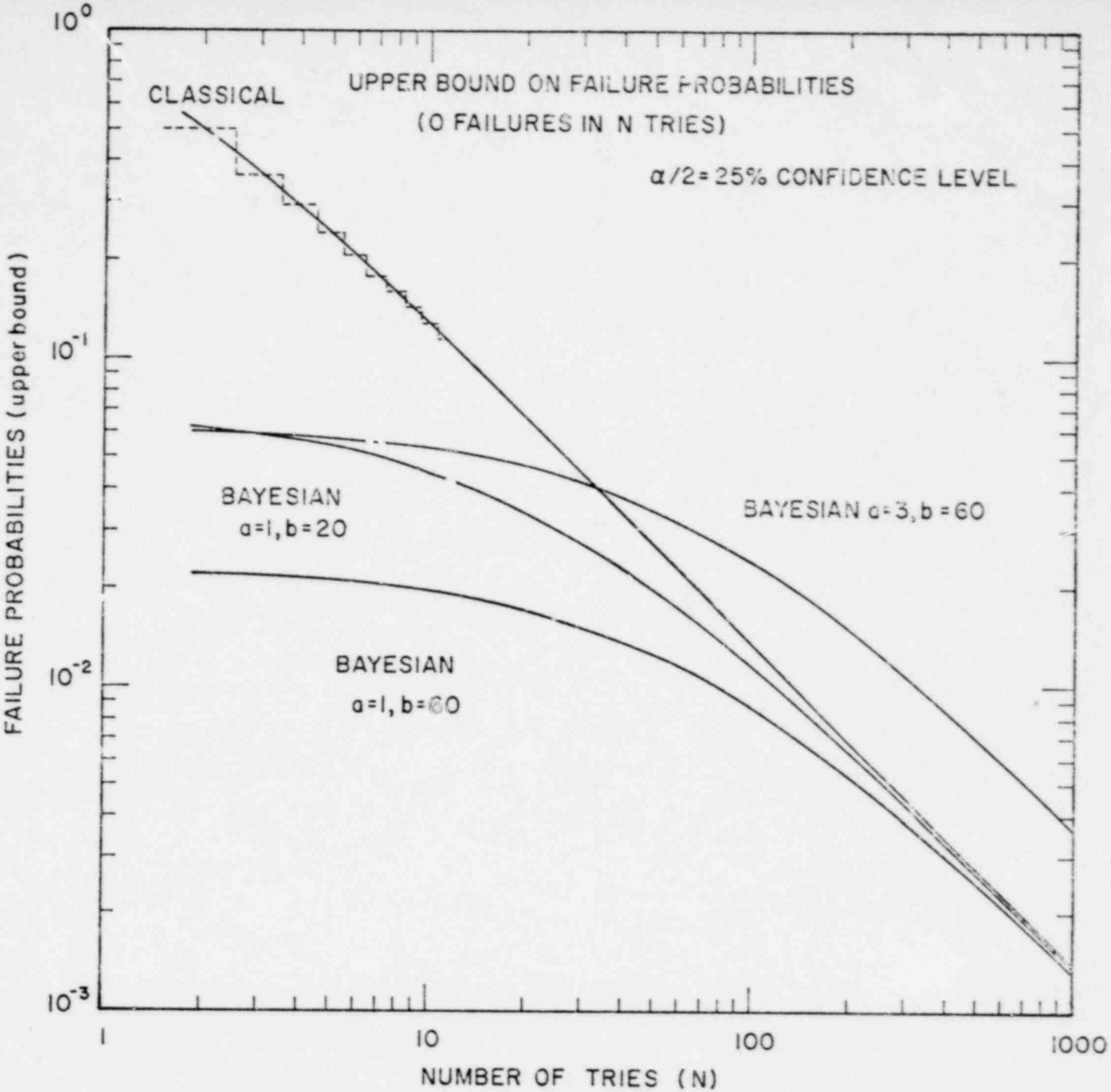


Fig. 5.2 Comparison of the Classical and Bayesian estimates for the upper confidence limit at the α=0.5 level.

However, the Bayesian estimate does not always require fewer startups than the classical description to achieve a given confidence level estimate of a failure probablity. For example, with $a=3$, $b=60$, (the same prior mean as $a=1$, $b=20$), the Bayesian estimate of p_1 is less than the classical estimate for $1 < n < 33$ with $k=0$, while for $n > 33$ the Bayesian estimate is greater (see Fig. 5.2). This result is not surprising, since for $a=3$, $b=60$ the prior distribution is highly peaked around the mean = $a/(a+b) = 0.048$ (i.e., it has a very small variance) and consequently a great deal of subsequent experimental observation is required to reduce the estimate of p_1 below this preconceived or biased value. Thus, not only is the mean (or the a/b ratio) of the prior distribution significant in establishing p_1, but the variance is also of major concern.

6. NON-BETA PRIOR DISTRIBUTIONS

A brief investigation was initiated to examine the effect of using non-beta prior distributions in the analysis of failure-on-demand attribute data. While this phase of the study is incomplete, some progress was made in two areas. First, a mixture of several beta distributions to form a contagious distribution [15] was examined. Then it was shown that a gamma prior distribution could be used for the Bayesian analysis of failure-on-demand data if the failure probability for the components is small. The results of these two investigations are summarized in this section.

6.1 Mixture Distributions

Contrasted to the familiar case in which two or more random variables are combined in a linear fashion is the case in which two or more probability distribution functions are combined in a linear fashion This is called a mixture (or *contagious*) distribution [15].

In the first case two variables are added to form a new variable , e.g.,

$$z = c_1x_1 + c_2x_2 \; . \tag{6.1}$$

In this case the x_1 and x_2 values are assumed to be from the same probability distribution function (pdf). The expected value, E[z], and variance, V[z], of z are given by

$$E[z] = c_1E[x_1] + c_2E[x_2] \tag{6.2}$$

and

$$V[z] = c_1^2V[x_1] + c_2^2V[x_2] + 2c_1c_2 \, Cov[x_1,x_2] \; . \tag{6.3}$$

In the second case, the mixture (or contagious) distribution is formed as a linear combination of the pdfs, i.e.,

$$f(x) = \alpha_1 \, f_1(x) + \alpha_2 \, f_2(x), \tag{6.4}$$

where α_1, α_2 are the relative weights ($0 \leq \alpha_1 \leq 1$, $0 \leq \alpha_2 \leq 1$) and

$$\alpha_1 + \alpha_2 = 1 \; .$$

1426 336

The pdf, $f(x)$, of Eq. (6.4) can be viewed as the pdf which contains variables from two distinctly different pdfs, $f_1(x)$ and $f_2(x)$. It is convenient to establish formulas for the mean (μ) and variance (σ^2) of the mixture population in terms of the means (μ_i) and variances (σ_i^2) of the component pdfs. Since

$$E[x] = \int x f(x) dx \ . \tag{6.5}$$

substitution for $f(x)$ in Eq. (6.5) from Eq. (6.4) yields

$$E[x] = \alpha_1 \int x f_1(x) dx + \alpha_2 \int x f_2(x) dx$$

or

$$E[x] = \alpha_1 \mu_1 + \alpha_2 \mu_2 \ . \tag{6.6}$$

For the variance, one obtains

$$Var[x] = E[x^2] - \{E[x]\}^2$$

$$= \alpha_1 \int x^2 f_1(x) dx + \alpha_2 \int x^2 f_2(x) dx - [\alpha_1 \mu_1 + \alpha_2 \mu_2]^2 \tag{6.7}$$

However

$$\sigma_i^2 = \int x^2 f_i(x) dx - \mu_i^2 \ ,$$

and Eq. (6.7) can be simplified to give

$$Var[x] = \alpha_1 \sigma_1{}^2 + \alpha_2 \sigma_2{}^2 + \alpha_1 \alpha_2 (\mu_1 - \mu_2)^2 \tag{6.8}$$

Thus the mean (or expected) value (Eq. 6.6) of the random variable governed by the mixture distribution has the same form as that for the case when two or more random variables are combined in a linear fashion, Eq. (6.2). However, the variance is substantially different for these two cases (compare Eq. (6.3) to Eq. (6.8)).

The above results can be generalized to a mixture of N probability density functions, i.e.,

$$f(x) = \sum_{i=1}^{N} \alpha_i \ f_i(x)$$

where the weights are subject to $\sum_i \alpha_i = 1$. For this case the mean and variance of the mixture distribution can be expressed in terms of the means and variances of the component distributions. The mean is given by

$$\mu = \sum_{i=1}^{N} \alpha_i \mu_i \; , \tag{6.9}$$

and the variance is given by

$$\sigma^2 = \sum_{i=1}^{N} \alpha_i \sigma_i^{\,2} + \sum_{i=1}^{N} \alpha_i (1-\alpha_i) \mu_i^{\,2} - \sum_{i=1}^{N} \alpha_i \mu_i \sum_{j \neq i}^{N} \alpha_j \mu_j \; . \tag{6.10}$$

6.1.1 Mixture of Two Beta Distributions

If two beta distributions are mixed according to Eq. (6.4), the shape of f(x) can vary quite widely, e.g., from bimodal to unimodal to exponential shaped. Thus f(x) may or may not be adequately expressed as a beta distribution. The object of this section is to investigate the problems of estimating the weights (α_1 and α_2) and the parameters of the beta used to approximate the mixture distribution. Thus one can write

$$\frac{p^{(a-1)}(1-p)^{(b-1)}}{B(a,b)} \simeq \alpha_1 \frac{p^{(a_1-1)}(1-p)^{(b_1-1)}}{B(a_1,b_1)} + \alpha_2 \frac{p^{(a_2-1)}(1-p)^{(b_2-1)}}{B(a_2,b_2)} \tag{6.11}$$

or

$$\mathrm{be}(a,b) \simeq \alpha_1 \, \mathrm{be}(a_1,b_1) + \alpha_2 \, \mathrm{be}(a_2,b_2) \; . \tag{6.12}$$

If a_1, b_1, a_2, b_2, α_1, and α_2 are known, one can use Eqs. (6.6) and (6.8) together with the relationships for a and b as functions of μ and σ^2, the mean and variance of be(a,b), to obtain estimates for a and b in terms of known quantities. Thus, by matching moments one obtains

$$\mu \equiv \frac{a}{a+b} = \alpha_1 \mu_1 + \alpha_2 \mu_2 \tag{6.13}$$

$$\sigma^2 \equiv \frac{ab}{(a+b)^2 (a+b+1)} = \alpha_1 \sigma_1^{\,2} + \alpha_2 \sigma_2^{\,2} + \alpha_1 \alpha_2 (\mu_1 - \mu_2)^2 \; . \tag{6.14}$$

1426 338

In Table 6.1 the values of a and b which result from mixing two beta distributions are listed. The two mixed beta distributions are of the exponential-type ($a_1, a_2 < 1.0$) and the resulting mixture beta distributions are also of the exponential-type ($a < 1.0$). The values for the a parameter increase monotonically with increasing α_1, and the values for the b parameter decrease, although not monotonically.

Table 6.1. The mean, variance, and beta parameters of mixed beta distributions of the exponential type[a].

α_1	μ	σ	a	b
0	0.010	0.00160	0.0519	5.1356
0.1	0.019	0.00467	0.0568	2.9352
0.2	0.028	0.00758	0.0726	2.5198
0.3	0.037	0.01032	0.0907	2.3615
0.4	0.046	0.01290	0.1104	2.2904
0.5	0.055	0.01532	0.1315	2.2600
0.6	0.064	0.01758	0.1540	2.2527
0.7	0.073	0.01968	0.1780	2.2604
0.8	0.082	0.02162	0.2036	2.2789
0.9	0.091	0.02339	0.2308	2.3058
1.0	0.100	0.02500	0.2600	2.3400

[a]The two beta distributions used for mixing have the following means and variances:

$$\mu_1 = 0.1 \quad , \quad \mu_2 = 0.01$$

$$\sigma_1^{\,2} = 0.025, \; \sigma_2^{\,2} = 0.0016$$

Thus the relationships for the mean and variance of the mixed beta distribution in terms of weighting value α_1 ($\alpha_2 = 1-\alpha_1$) are given as

$$\mu = 0.1\alpha_1 + 0.01\,(1-\alpha_1)$$

$$\sigma^2 = 0.025\alpha_1 + 0.0016(1-\alpha_1) + 0.0081\alpha_1(1-\alpha_1)$$

As further examples of mixing two beta distributions, several pairs of beta distributions used to describe diesel engine failure data were mixed in varying proportions. The mean and variance (calculated by the prior matching moments method) of several diesel engine grouping were reported in Section 3.5. The results of the mixture of

1426 339

two groupings for two different manufacturers are shown in Table 6.2 for "13 GM diesel engines" with "Four ALCO engines". Table 6.3 shows the results from grouping "13 GM diesel engines" with "Four Fairbanks diesel engines". Similarly, Table 6.4 shows the results of mixing "0-25 starts" with "more than 100 starts" and, Table 6.5 "0-25 starts with "26-50 starts".

6.1.2 Estimates of the Weights from Test Samples

To form the contagious distribution, one must first determine values for the weights, α_i, for each subgroup or component distribution. Since α_i can be interpreted as the probability of a failure data sample being chosen from subgroup i, the probability of obtaining s_i samples from the i-th subgroup is

$$f(s_i,\alpha_i) = \alpha_i^{s_i} , \quad i = 1,2,\ldots,N . \tag{6.15}$$

The *likelihood function*, L, which is the probability of obtaining $s_1,s_2,\ldots,s_n$ samples from subgroup 1,2,...,N is thus given by

$$L = C \prod_{i=1}^{N} \alpha_i^{s_i} , \tag{6.16}$$

where C is simply the number of permutations of $s_1,s_2,\ldots,s_N$ in $S = \sum_i s_i$ samples, i.e.,

$$C = S!/(\prod_{i=1}^{N} s_i!) . \tag{6.17}$$

The choice of the mixture weights to describe the mixture distributions is taken as those values of α_i which maximize the likelihood function, or equivalently minimize lnL. Since the sum of the weights must be unity, the logarithm of Eq. (6.16) may be written as

$$lnL = lnC + \sum_{i=1}^{N-1} s_i \; ln\alpha_i + s_N \; ln\left(1-\sum_{i=1}^{N-1} \alpha_i\right) \tag{6.18}$$

To find the values of α_i which minimize this result, differentiate with respect to α_i, i=1,...,N-1, set the result to zero, and solve for α_i to obtain

Table 6.2. The mean and variance and beta parameters of the mixture distribution of 13 GM diesel engines with 4 ALCO engines.

α_1	μ	σ^2	a	b
0.0000000E 00	0.2940000E-01	0.5999999E-03	0.1368845E 01	0.4519054E 02
0.1000000E 00	0.3238000E-01	0.9509996E-03	0.1034408E 01	0.3091148E 02
0.2000000E 00	0.3536000E-01	0.1284000E-02	0.9039839E 00	0.2466116E 02
0.3000001E 00	0.3833999E-01	0.1599000E-02	0.8457108E 00	0.2121246E 02
0.4000001E 00	0.4132000E-01	0.1896000E-02	0.8219684E 00	0.1907077E 02
0.5000001E 00	0.4430000E-01	0.2175000E-02	0.8180226E 00	0.1764749E 02
0.6000001E 00	0.4728000E-01	0.2436000E-02	0.8269845E 00	0.1666422E 02
0.7000002E 00	0.5026000E-01	0.2679000E-02	0.8452634E 00	0.1597253E 02
0.8000002E 00	0.5324000E-01	0.2904000E-02	0.8708609E 00	0.1548639E 02
0.9000002E 00	0.5622000E-01	0.3111000E-02	0.9026338E 00	0.1515274E 02
0.1000000E 01	0.5920000E-01	0.3300000E-02	0.9399409E 00	0.1493744E 02

Table 6.3 The mean and variance and beta parameters of the mixture distribution of 13 GM diesel engines with 4 ALCO engines.

α_1	μ	σ^2	a	b
0.0000000E 00	0.3220000E-01	0.7000000E-03	0.1401304E 01	0.4211749E 02
0.1000000E 00	0.3490000E-01	0.1023000E-02	0.1114172E 01	0.3081055E 02
0.2000000E 00	0.3760000E-01	0.1443000E-02	0.9849840E 00	0.2518297E 02
0.3000001E 00	0.4030000E-01	0.1627000E-02	0.9176834E 00	0.2195361E 02
0.4000001E 00	0.4300000E-01	0.1908000E-02	0.8844073E 00	0.1968320E 02
0.5000001E 00	0.4570000E-01	0.2175000E-02	0.8706429E 00	0.1818060E 02
0.6000001E 00	0.4840000E-01	0.2428000E-02	0.8697135E 00	0.1709955E 02
0.7000002E 00	0.5110000E-01	0.2667000E-02	0.8779503E 00	0.1630305E 02
0.8000002E 00	0.5380000E-01	0.2892000E-02	0.8931983E 00	0.1570898E 02
0.9000002E 00	0.5650000E-01	0.3103000E-02	0.9141374E 00	0.1526526E 02
0.1000000E 01	0.5920000E-01	0.3300000E-02	0.9399409E 00	0.1493744E 02

Table 6.4 The mean and variance and beta parameters of the mixture distribution of "0-25 starts" with "more than 100 starts".

α_1	μ	σ^2	a	b
0.0000000E 00	0.3030000E-01	0.8000000E-03	0.1082539E 01	0.3464482E 02
0.1000000E 00	0.4227000E-01	0.4866000E-02	0.3093279E 00	0.7008579E 01
0.2000000E 00	0.5424000E-01	0.8647997E-02	0.2674996E 00	0.4664279E 01
0.3000001E 00	0.6620997E-01	0.1214299E-01	0.2708988E 00	0.3820613E 01
0.4000001E 00	0.7817996E-01	0.1535200E-01	0.2888247E 00	0.3405534E 01
0.5000001E 00	0.9014994E-01	0.1827500E-01	0.3144662E 00	0.3173791E 01
0.6000001E 00	0.1021199E 00	0.2091200E-01	0.3456383E 00	0.3038993E 01
0.7000002E 00	0.1140899E 00	0.2326300E-01	0.3816093E 00	0.2963202E 01
0.8000002E 00	0.1260599E 00	0.2532800E-01	0.4222611E 00	0.2927423E 01
0.9000002E 00	0.1380299E 00	0.2710700E-01	0.4678091E 00	0.2921375E 01
0.1000000E 01	0.1500000E 00	0.2860000E-01	0.5187059E 00	0.2939333E 01

Table 6.5 The mean and variance and beta parameters of the mixture distribution of "0-25 starts" with 26-50 starts".

α_1	μ	σ^2	a	b
0.0000000E 00	0.4920000E-01	0.7000000E-03	0.3238720E 01	0.6259901E 02
0.1000000E 00	0.5928000E-01	0.4407998E-02	0.6906751E 00	0.1096039E 02
0.2000000E 00	0.6935996E-01	0.7911995E-02	0.4965054E 00	0.6661884E 01
0.3000001E 00	0.7943994E-01	0.1121200E-01	0.4386995E 00	0.5083706E 01
0.4000001E 00	0.8951998E-01	0.1430800E-01	0.4204345E 00	0.4276107E 01
0.5000001E 00	0.9959996E-01	0.1720000E-01	0.4197084E 00	0.3794231E 01
0.6000001E 00	0.1096799E 00	0.1988799E-01	0.4288494E 00	0.3481156E 01
0.7000002E 00	0.1197599E 00	0.2237200E-01	0.4445519E 00	0.3267472E 01
0.8000002E 00	0.1298400E 00	0.2465200E-01	0.4652240E 00	0.3117834E 01
0.9000002E 00	0.1399199E 00	0.2672800E-01	0.4900669E 00	0.3012412E 01
0.1000000E 01	0.1500000E 00	0.2860000E-01	0.5187059E 00	0.2930333E 01

$$\alpha_i = \alpha_N \, s_i/s_N \ , \quad i=1,\ldots,N-1 \ . \tag{6.19}$$

Summation of this result over i from 1 to N-1 and use of the relation $\alpha_N = 1 - \sum_{i=1}^{N-1} \alpha_i$ yields

$$\alpha_N = s_N/S \ . \tag{6.20}$$

Substitution for α_N into Eq. (6.19) then gives the maximum likelihood estimate for the i-th subgroup weighting factor as

$$\hat{\alpha}_i = s_i/S \ , \quad i=1,\ldots,N \ , \tag{6.21}$$

i.e., the weight factor for the i-th subgroup is simply the observed fraction of the total samples which are taken from the i-th subgroup.

6.2 Gamma Prior Distribution with the Conjugate Poisson Conditional Distribution

The beta family is usually chosen to represent the prior distribution in the Bayesian analysis of failure-on-demand data because of the mathematical convenience of using the conjugate distribution to the binomial conditional distribution. As an alternative to a beta prior distribution, one could also use a "truncated" gamma distribution as the prior distribution, namely

$$g(p) = \frac{\delta^{\alpha} \, p^{\alpha-1} \, e^{-\delta p}}{\Gamma(\alpha)} \left[1 - \int_1^{\infty} \frac{\delta^{\alpha} \, x^{\alpha-1} \, e^{-\delta x}}{\Gamma(\alpha)} \, dx \right]^{-1} \tag{6.22}$$

where p is restricted to $0 \le p \le 1$. If the parameters, α and δ, of this truncated gamma distribution are such that the normalization factor in brackets in the above equation is very close to unity, then this truncated gamma distribution may be approximated by the usual gamma distribution,

$$g(p) \simeq \frac{\delta^{\alpha} \, p^{\alpha-1} \, e^{-\delta p}}{\Gamma(\alpha)} \ . \tag{6.23}$$

This approximation will be valid whenever the function is highly skewed towards small failure probabilities. Such skewness of the prior distribution can be expected for components whose failure probabilities are much less than unity.

The use of either the truncated or regular gamma distribution as a prior distribution with a binomial conditional distribution does not lead to closed form results for the marginal and posterior distributions since the gamma and binomial distributions are not natural conjugates. However, for the type of failure-on-demand data considered in this study (i e., failure data from components with low failure probabilities), the bionomial conditional distribution may be approximated by a Poisson distribution, which is the natural conjugate of the gamma distribution. If the number of demands, n, is large and the number of failures, k, is much smaller, then [8]

$$\frac{n!}{(n-k)!} \simeq n^k \; . \qquad (6.24)$$

Further, if the failure probability, p, for each component is very small, ($p<<1$) then

$$(1-p)^{n-k} \simeq (e^{-p})^{n-k} \simeq e^{-np} \; . \qquad (6.25)$$

With these two approximations, the binomial conditional distribution of Eq. (2.1) can be approximated by a Poisson distribution, i.e.,

$$f(k|n,p) = \frac{n!}{(n-k)!k!} \; p^k \; (1-p)^{n-k} \simeq \frac{(np)^k \; e^{-np}}{k!} \quad . \qquad (6.26)$$

The marginal distribution can now be evaluated readily using the above approximations. Recall

$$h(k|n,\alpha,\delta) = \int_0^1 f(k|n,p) \; g(p) \; dp \; , \qquad (6.27)$$

which, if g(p) is highly skewed towards the lower limit, can be approximated mathematically by

$$h(k|n,\alpha,\delta) \simeq \int_0^\infty f(k|n,p) \; g(p) \; dp \; . \qquad (6.28)$$

Substituting for $f(k|n,p)$ and $g(p)$ from Eqs. (6.26) and (6.23), respectively, gives

$$h(k|n,\alpha,\delta) = \frac{\delta^{\alpha} n^{k}}{k!\Gamma(\alpha)} \; \frac{\Gamma(k+\alpha)}{(n+\delta)^{k+\alpha}} \; . \qquad (6.29)$$

The posterior distribution, $\delta(p|k,n,\alpha,\delta)$, is

$$\xi(p\; k,n,\alpha,\delta) = \frac{f(k|n,p)g(p)}{h(k|n,\alpha,\delta)} \; , \qquad (6.30)$$

and upon substitution of Eqs. (6.15), (6.16), and (6.20) yields

$$\xi(p|k,n,\sigma,\delta) = \frac{(n+\delta)^{k+\alpha} \; e^{-p(\delta+n)} \; p^{\alpha+k-1}}{\Gamma(k+\alpha)} \; , \qquad (6.31)$$

which is also a gamma distribution. The mean of this posterior distribution is

$$E(p|k,n\alpha,\delta) = \frac{k+\alpha}{n+\delta} \equiv \hat{p}_B \; , \qquad (6.32)$$

while the classical estimate of the mean of p is

$$\hat{p}_c = \frac{k}{n} \; . \qquad (6.33)$$

6.2.1 Estimation of Gamma Parameters

To estimate values for the gamma prior parameters from failure data, any of the four estimation methods previously discussed for the beta-binomial model could also be used. The simplest method is to match the prior moments to those of the data. The mean and variance of the gamma prior of Eq. (6.23) are

$$\mu = \alpha/\delta \; , \qquad (6.34)$$

and

$$\sigma^2 = \alpha/\delta^2 \; . \qquad (6.35)$$

The data mean and variance are

$$\hat{\mu}_{ob} = \frac{1}{N} \sum_{i=1}^{N} \frac{k_i}{n_i} \; , \qquad (6.36)$$

and

$$\hat{\sigma}^2_{ob} = \frac{1}{N-1} \sum_{i=1}^{N} \left(\frac{k_i}{n_i} - \hat{\mu}_{ob}\right)^2 \; . \qquad (6.37)$$

By matching these calculated values to the mean and variance of the prior distribution, a relation between α and δ in terms of the observed data can be obtained, namely

$$\alpha = \mu\delta = \hat{\mu}^2_{ob}/\hat{\sigma}^2_{ob} \qquad (6.38)$$

and

$$\delta = \mu/\sigma^2 = \hat{\mu}_{ob}/\hat{\sigma}^2_{ob} \qquad (6.39)$$

Equations (6.38) and (6.39) can be used to find expressions for estimates of the variances of α and δ from the following relations:

$$s^2(\alpha) = \left(\frac{\partial\alpha}{\partial\hat{\mu}_{ob}}\right)^2 s^2(\hat{\mu}_{ob}) + \left(\frac{\partial\alpha}{\partial\hat{\sigma}^2_{ob}}\right)^2 s^2\left(\hat{\sigma}^2_{ob}\right), \qquad (6.40)$$

and

$$s^2(\delta) = \left(\frac{\partial\delta}{\partial\hat{\mu}_{ob}}\right)^2 s^2(\hat{\mu}_{ob}) + \left(\frac{\partial\delta}{\partial\hat{\sigma}^2_{ob}}\right) s^2(\sigma^2_{ob}) \ , \qquad (6.41)$$

where $s^2(\hat{\mu}_{ob})$ and $s^2(\hat{\sigma}^2_{ob})$ are estimates for the variances of $\hat{\mu}_{ob}$ and $\hat{\sigma}^2_{ob}$. Expressions for $s^2(\hat{\mu}_{ob})$ and $s^2(\hat{\sigma}^2_{ob})$ are (of Section 3.7)

$$s^2(\hat{\mu}_{ob}) = \frac{\hat{\sigma}^2_{ob}}{N} \qquad (6.42)$$

$$s^2(\hat{\sigma}^2_{ob}) = \frac{2(\hat{\sigma}^2_{ob})^2}{N-1} \quad . \qquad (6.43)$$

The maximum likelihood method can be used to estimate the parameters of the prior distribution by using the likelihood function

$$L(k_1,k_2,\ldots,k_N|n_1,n_2,\ldots,n_N,\alpha,\delta) = \prod_{i=1}^{N} h(k_i|n_i,\alpha,\delta) \ , \qquad (6.44)$$

which is the probability of obtaining, simultaneously, $k_1,k_2,\ldots,k_N$ failures in $n_1,n_2,\ldots,n_N$ tries for components $1,2,\ldots,N$, respectively, for components whose probability distribution for failure is given by the prior distribution of Eq. (6.23) with parameters α and δ. Substitution of Eq. (6.29) into Eq. (6.44) yields

$$L(\alpha,\delta) = L(k_1,\ldots,k_N|n_1,\ldots,n_N,\alpha,\delta)$$

$$= \left(\frac{\delta^\alpha}{\Gamma(\alpha)}\right)^N \prod_{i=1}^{N} \frac{n_i^{k_i}}{k_1!} \frac{\Gamma(k_i + \alpha)}{(n_i+\delta)^{k_i+\alpha}} \quad . \qquad (6.45)$$

To find the values of α and δ which maximizes L, or equivalently, minimizes lnL. The extrema of $ln\text{L}(\alpha,\delta)$ are obtained by solving

$$\frac{\partial ln\text{L}(\alpha,\delta)}{\partial\alpha} = 0 \qquad (6.46)$$

$$\frac{\partial ln\text{L}(\alpha,\delta)}{\partial\delta} = 0 \qquad (6.47)$$

The numerical solution of these two simultaneous equations can be obtained by several standard numerical techniques.

6.2.2 Comparison of Beta and Gamma Priors for Diesel Engine Data

To test the ability of the gamma function to serve as a prior distribution for low probability failure data, the diesel engine failure data of Table 3.1 were analyzed by both the approximate gamma-Poisson description and the beta binomial description. As before the diesel failure data were grouped by manufacturer and by number of starts, and each group was then separately examined.

The method of matching the prior moments to those of the failure data were used to obtain values for the prior parameters of each data group (i.e., Eqs. (6.38) and (6.39) for the gamma distribution, and Eqs. (3.5) and (3.6) for the beta distribution). The resulting beta and gamma parameters for the various data groupings are given in Table 6.6.

One immediate result to be seen from these parameter results, is that both prior models generally yield unimodal priors ($a,b>1$ or $\alpha>1$) for most groupings. However, the estimated beta priors for the "GM engines", and "other engines" and "0-25 starts" groupings and the gamma priors for the "Other engines" and "0-25 starts" groupings are all monotonically decreasing functions which become unbounded as $p\to0$. Moreover, for the "GM engines" grouping, the estimated beta prior is monotonically decreasing while the estimated gamma prior is unimodal and everywhere bounded.

Table 6.6 Parameter Values for the beta and gamma prior models obtained by the prior matching moments method for various groupings of the diesel engine failure data of Table 3.1.

Grouping	No. Engines	Mean	Variance	Beta Prior a	Beta Prior b	Gamma Prior α	Gamma Prior δ
GM	13	0.05916	0.003328	0.9303	14.795	1.0516	17.777
Fairbanks	4	0.03217	0.000707	1.3846	41.662	1.4639	45.511
ALCO	4	0.02935	0.000600	1.3644	45.120	1.4359	48.920
Other	4	0.12014	0.038014	0.2139	1.567	0.3797	3.160
0-25 Starts	5	0.15047	0.028592	0.5222	2.949	0.7919	5.263
25-50 Starts	5	0.04924	0.000691	3.2868	63.462	3.5088	71.258
50-100 Starts	9	0.03501	0.000718	1.6123	44.437	1.7071	48.757
>100 Starts	6	0.03033	0.000789	1.1000	35.162	1.1657	38.428

As p→0 the difference between these two distributions diverges! Nevertheless both of these estimated prior distributions give approximately the same values for all but very small values of p (see Fig. 6.1 in which some of the beta and gamma distributions are shown). From Fig. 6.1 it is seen that the difference between the beta and prior models for the same data group is typically very small. This excellent agreement was found for all the data groupings.

As an additional comparison between the approximate gamma-Poisson model and the beta-binomial model, the posterior distribution for each diesel engine in each grouping was calculated. Again the corresponding beta and gamma posteriors distributions were very similar. In Tables 6.7 and 6.8 the mean and variance of these posterior distributions are shown together with the classical estimate of the failure probability for each engine (k_i/n_i). Notice how closely the means and variance of the gamma posterior distributions are to those of the corresponding beta posterior distributions.

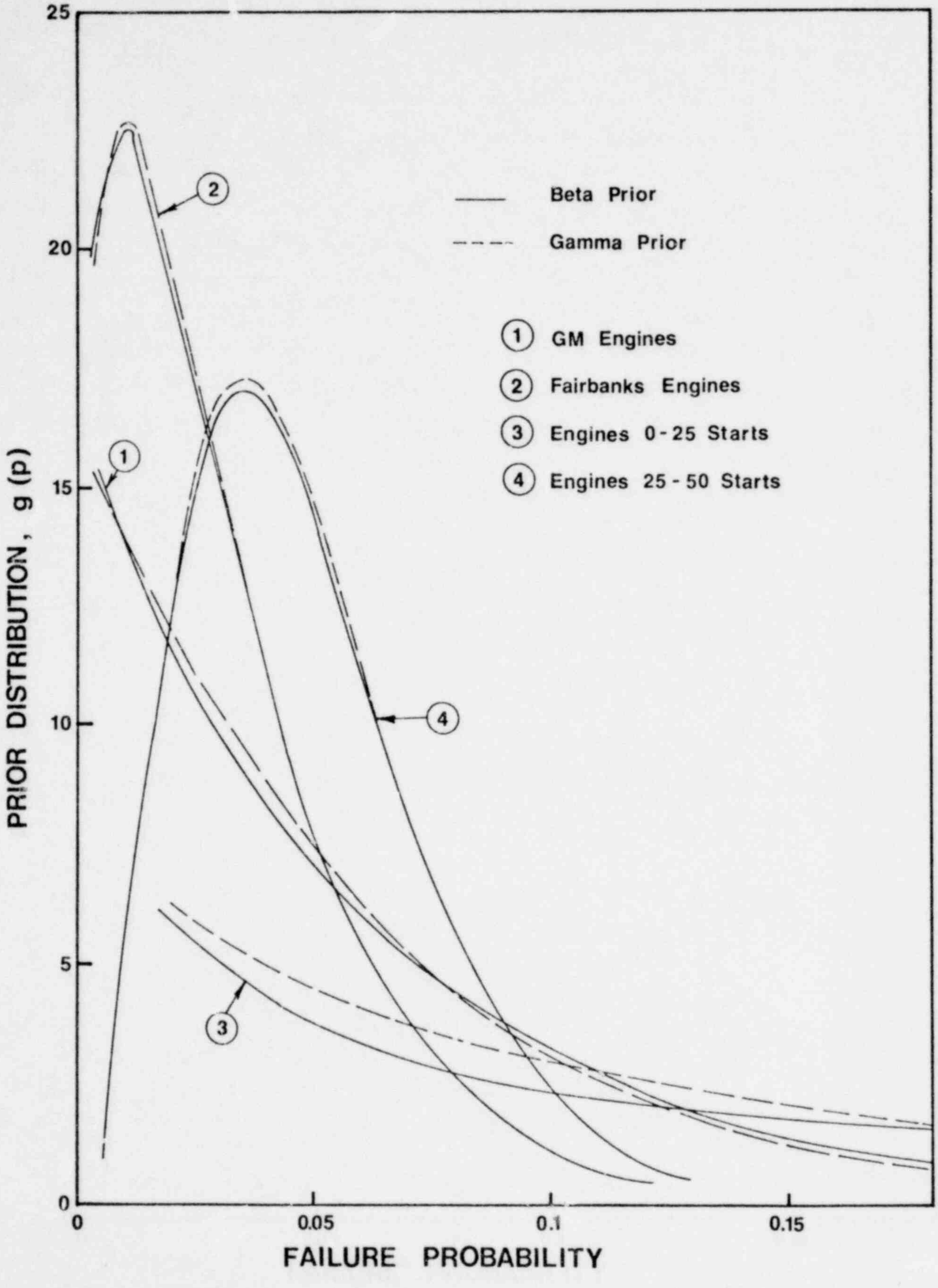


Fig. 6.1 Estimated gamma and beta prior distributions obtained by the prior matching moments method for several groupings of the diesel engine failure data.

1426 350

Table 6.7 Mean and variance of component posterior distributions for both the beta and gamma models of the prior distribution for the diesel engine failure data of Table 3.1 grouped by manufacturer.

Component k_i	n_i	Beta Posterior Mean	Beta Posterior Variance	Gamma Posterior Mean	Gamma Posterior Variance	Classical Mean
GM Diesel Engines						
6	100	0.599E-01	0.482E-03	0.599E-01	0.508E-03	0.600E-01
1	392	0.473E-02	0.115E-04	0.501E-02	0.122E-04	0.255E-02
11	230	0.486E-01	0.187E-03	0.486E-01	0.196E-03	0.478E-01
5	68	0.708E-01	0.777E-03	0.706E-01	0.822E-03	0.735E-01
4	23	0.127	0.280E-02	0.124	0.304E-02	0.174
0	23	0.240E-01	0.590E-03	0.258E-01	0.632E-03	0.000
2	12	0.106	0.329E-02	0.102	0.344E-02	0.167
0	99	0.811E-02	0.695E-04	0.901E-02	0.771E-04	0.000
3	33	0.807E-01	0.149E-02	0.798E-01	0.157E-02	0.909E-01
9	126	0.701E-01	0.457E-03	0.699E-01	0.486E-03	0.714E-01
2	47	0.467E-01	0.699E-03	0.471E-01	0.727E-03	0.426E-01
1	87	0.188E-01	0.178E-03	0.196E-01	0.187E-03	0.115E-01
2	71	0.338E-01	0.372E-03	0.344E-01	0.387E-03	0.282E-01
Fairbanks Diesel Engine						
3	656	0.627E-02	0.890E-05	0.636E-02	0.907E-05	0.457E-02
5	73	0.550E-01	0.444E-03	0.545E-01	0.460E-03	0.685E-01
1	35	0.306E-01	0.375E-03	0.306E-01	0.380E-01	0.286E-01
1	37	0.298E-01	0.357E-03	0.299E-01	0.362E-03	0.270E-01
ALCO Diesel Engines						
0	13	0.229E-01	0.371E-03	0.232E-01	0.375E-03	0.000
2	95	0.238E-01	0.163E-03	0.239E-01	0.166E-03	0.211E-01
2	51	0.345E-01	0.338E-03	0.344E-01	0.344E-03	0.392E-01
2	35	0.413E-01	0.480E-03	0.409E-01	0.488E-03	0.571E-01
Diesels by Other Manufacturers						
7	17	0.384	0.120E-01	0.366	0.182E-01	0.412
4	335	0.125E-01	0.366E-04	0.130E-01	0.383E-04	0.119E-01
9	206	0.443E-01	0.203E-03	0.448E-01	0.214E-03	0.437E-01
1	76	0.156E-01	0.195E-03	0.174E-01	0.220E-03	0.132E-01

1426 351

Table 6.8 Mean and variance of component posterior distributions for both the beta and gamma models of the prior distribution for the diesel engine failure data of Table 3.1 grouped by number of starts.

Component k_i	n_i	Beta Posterior Mean	Beta Posterior Variance	Gamma Posterior Mean	Gamma Posterior Variance	Classical Mean
0-25 Starts						
4	23	0.171	0.516E-02	0.170	0.600E-02	0.174
0	23	0.197E-01	0.704E-03	0.280E-01	0.991E-03	0.000
2	12	0.163	0.828E-02	0.162	0.937E-02	0.167
0	13	0.317E-01	0.176E-02	0.434E-01	0.237E-02	0.000
7	17	0.367	0.108E-01	0.350	0.157E-01	0.412
25-50 Starts						
3	33	0.630E-01	0.586E-03	0.624E-01	0.599E-03	0.909E-01
2	47	0.465E-01	0.386E-03	0.466E-01	0.394E-03	0.426E-01
1	35	0.421E-01	0.393-03	0.424E-01	0.399E-03	0.286E-01
1	37	0.413E-01	0.378E-03	0.416E-01	0.385E-03	0.270E-01
2	35	0.520E-01	0.479E-03	0.518E-01	0.488E-03	0.571E-01
50-100 Starts						
6	100	0.521E-01	0.336E-03	0.518E-01	0.348E-03	0.600E-01
5	68	0.580E-01	0.475E-03	0.574E-01	0.492E-03	0.735E-01
0	99	0.111E-01	0.753E-04	0.116E-01	0.782E-04	0.000
1	87	0.196E-01	0.144E-03	0.199E-01	0.147E-03	0.115E-01
2	71	0.309E-01	0.253E-03	0.310E-01	0.258E-03	0.282E-01
5	73	0.555E-01	0.437E-03	0.551E-01	0.452E-03	0.685E-01
2	95	0.256E-01	0.176E-03	0.258E-01	0.179E-03	0.211E-01
2	51	0.372E-01	0.365E-03	0.372E-01	0.373E-03	0.392E-01
1	76	0.214E-01	0.170E-03	0.217E-01	0.174E-03	0.132E-01
>100 Starts						
1	392	0.490E-02	0.114E-04	0.503E-02	0.117E-04	0.255E-02
11	230	0.454E-01	0.162E-03	0.453E-01	0.169E-03	0.478E-01
9	126	0.622E-01	0.358E-03	0.618E-01	0.376E-03	0.714E-01
3	656	0.592E-02	0.849E-05	0.600E-02	0.864E-05	0.457E-02
4	335	0.137E-01	0.364E-04	0.138E-01	0.370E-04	0.119E-01
9	206	0.417E-01	0.164E-03	0.416E-01	0.170E-03	0.4[illegible]E-01

7. REFERENCES

1. Bayes formulated his work in 1763 and his classic paper can be found in *Biometrika*, 45, 293 (1958).

2. A. M. Mood, and F. A. Graybill, *Introduction to the Theory of Statistics*, 3rd Ed., p. 115, McGraw-Hill, New York, (1963).

3. N. L. Johnson and S. Kotz, *Discrete Distributions*, Houghton Mifflin Co., (1969).

4. J. C. Kleinman, "Proportions with Extraneous Variance: Single and Independent Samples," J. Am. Stat. Assoc., 68, 46 (1973).

5. A. C. Bajpai, L. R. Mastoe, and D. Walker, *Advanced Engineering Mathematics*, Wiley, New York (1977).

6. W. E. Vesely, U.S. Nuclear Regulatory Commission, private communication.

7. J. B. Kennedy, and A. M. Neville, *Basic Statistical Methods for Engineers and Scientists*, 2nd Ed., p. 205, A Dun-Donnelley Publisher, New York, (1976).

8. M. Ambramovitz and L. A. Stegun, "Handbook of Mathematical Functions," National Bureau of Standards, Applied Mathematics Series - 55, (1966).

9. W. E. Deming, *Statistical Adjustment of Data*, p. 37, Dover, New York, (1964).

10. K. A. Brownlee, *Statistical Theory and Methodology in Science and Engineering*, 2nd Ed., p. 300, John Wiley and Sons, New York, (1965).

11. S. S. Wilks, *Mathematical Statistics*, Wiley, New York (1962).

12. Scientific Subroutine Package (360-A-CM-03X), Version III Programmer's Manual, H20-0205, IBM (1963).

13. J. A. Greenwood and H. O. Hartley, "Guide to Tables in Mathematical Statistics," Princeton University Press (1961).

14. O. G. Ludwig, "Incomplete Beta Ratio," Comm. ACM, 6 (1963) 314; also see "Collected Algorithms from CACM," Algorithm 179 and modifications by N. E. Bosten and E. L. Battiste (1972), and M. C. Pike and J. Soo Hoo (1975).

15. A. M. Mood, F. A. Graybill, and D. C. Boes, *Introduction to the Theory of Statistics*, McGraw-Hill, New York (1974).

16. C. R. Rao, *Statistical Inference and Its Application*, 2nd Ed., Wiley, New York (1976).

APPENDIX I

A User's Guide to the Program

BETA III

KSU-2662-9
CES-64
Nov. 30, 1978

A User's Guide to the Program

BETA III

by

J. K. Shultis and W. Buranapan
Dept. of Nuclear Engineering
Kansas State University
Manhattan, Kansas 66506

ABSTRACT

Beta III is a FORTRAN program which evaluates from observed component failure data the two parameters of a beta distribution which is assumed to describe the prior distribution of the failure probability among the components considered. Four methods are used to evaluate these prior parameters: (1) matching the mean and variance of the component data to those of the marginal distribution, (2) matching the mean and variance of the observed failure probabilities to those of the prior distribution, (3) the maximum likelihood method based on the marginal distribution, and (4) the maximum likelihood method based on the prior distribution. Beta III also calculates and plots both the probability density function and the cumulative distribution function of the beta prior distribution as calculated by each method.

CONTENTS

1. THEORY

1.1 Summary of Pertinent Probability Functions [1]

The probability of failure, p, is often assumed constant for a particular component. Thus, the probability of obtaining k failures in n tests is given by the binomial distribution, a *conditional* probability with respect to parameters n and p,

$$f(k|n,p) = \binom{n}{k} p^k (1-p)^{n-k}. \tag{1}$$

In sampling many similar components, it is often assumed that the distribution of failure probabilities among the components, called the *prior* distribution, can be described by a beta distribution,

$$g(p) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)} \qquad a,b>0 \tag{2}$$

where

$$B(a,b) \equiv \int_0^1 x^{a-1}(1-x)^{b-1}\,dx = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \tag{3}$$

and Γ is the gamma function. The program described in this report estimates values of the parameters a and b from observed component failure data.

The probability of k failures in n tries, $h(k|n,a,b)$, independent of the particular component, i.e. averaged over all component failure probabilities, is obtained by integrating Eq. (1) over all p weighted with the probability function g(p). This result is known as the *marginal* distribution, and is given by

$$h(k|n,a,b) = \int_0^1 f(k|n,p)g(p)\,dp = \binom{n}{k}\frac{B(a+k,\ b+n-k)}{B(a,b)}. \tag{4}$$

From Bayes' theorem one can determine the *posterior* distribution, $\xi(p|k,n,a,b)$, which is the distribution of the failure probability, p, for a particular component which previously has experienced k failures in n tries and which belongs to a class of components whose failure probabilities are distributed according to the prior distribution of Eq. (2) with parameters a and b. Explicity Bayes' theorem says

$$\xi(p|k,n,a,b) = \frac{f(k|n,p)g(p|a,b)}{h(k|n,a,b)}$$

which upon substitution of Eqs. (1), (2), and (4) yields

$$\xi(p|k,n,a,b) = \frac{p^{a+k-1}\,(1-p)^{b+n-k-1}}{B(a+k,b+n-k)}\,. \tag{5}$$

1.2 Summary of Techniques For Calculation of Prior Distribution

In this section a summary of the methods used to estimate the parameters of the prior beta distribution from observed component failure data is presented.

Matching Data to Moments of the Prior Distribution [1]

If there are k_i failures out of n_i tries for the i-th component, an *estimate* of the failure probability, p_i, is k_i/n_i, and thus the observed mean and variance are

$$\hat{p}_{ob} = \frac{1}{N}\sum_{i=1}^{N}\frac{k_i}{n_i} \tag{6}$$

and

$$\hat{\sigma}^2_{ob} = \frac{1}{N-1}\sum_{i=1}^{N}\left(\frac{k_i}{n_i} - \hat{p}_{ob}\right)^2. \tag{7}$$

where N is the total number of components in the same class and for which failure data are available. By matching these calculated values (which use only the observed data), to the mean and variance of the assumed prior distribution, the parameters a and b of the beta prior distribution are obtained as

$$a = \frac{\hat{p}^2_{ob}}{\hat{\sigma}^2_{ob}}\,(1-\hat{p}_{ob}) - \hat{p}_{ob} \tag{8}$$

and

$$b = \frac{\hat{p}_{ob}}{\hat{\sigma}^2_{ob}}\,(1-\hat{p}_{ob})^2 + \hat{p}_{ob} - 1. \tag{9}$$

Matching Data to Moments of the Marginal Distribution [2]

An alternative to the preceeding technique is to match the experimental data to the moments of the marginal or mixture distribution of Eq. (4). In general, the sample sizes will be unequal (i.e. different n_i), and thus, a weighting scheme should be used to calculate the mean and variance of the observed failure proportions, i.e.

$$\hat{p} = \frac{1}{w} \sum_{i=1}^{N} w_i \frac{k_i}{n_i}, \quad \text{where } w = \sum_{i=1}^{N} w_i$$

$$S = \frac{N-1}{N} \sum_{i=1}^{N} w_i \left(\left(\frac{k}{n}\right) - \frac{k_i}{n_i} \right)^2$$

By setting the above statistics equal to their expected values (of the marginal distribution) and solving the resulting equations for the prior mean and variance one obtains the following estimates:

$$\hat{\mu} = \hat{p} \tag{10}$$

and

$$\hat{\sigma}^2 = \hat{\mu}(1-\hat{\mu}) \frac{S - \hat{p}\hat{q}[\sum_{i=1}^{N} \frac{w_i}{n_i} (1 - \frac{w_i}{w})]}{\hat{p}\hat{q}[\sum_{i=1}^{N} w_i(1 - \frac{w_i}{w}) - \sum_{i=1}^{N} \frac{w_i}{n_i} (1 - \frac{w_i}{w})]} \tag{11}$$

where $\hat{q} = 1-\hat{p}$. The parameters a and b of the beta prior are then given by

$$a = \frac{\hat{\mu}^2}{\hat{\sigma}^2} (1-\hat{\mu}) - \hat{\mu}, \tag{12}$$

and

$$b = \frac{\hat{\mu}}{\hat{\sigma}^2} (1-\hat{\mu})^2 + \hat{\mu} - 1. \tag{13}$$

The choice of weights is made such that the estimate of μ is the linear unbiased estimate with minimum variance, i.e. weight each k_i/n_i with the inverse of its variance, namely

$$w_i = \frac{n_i}{1 + r(n_i - 1)} \tag{14}$$

where

$$r \equiv \sigma^2/(\mu(1-\mu)). \tag{15}$$

Equations (10),(11) and (14) can be viewed as three equations for the quantitites w_i, μ and σ^2 which can be solved by the following iteration scheme. Choose $r = 0$ so that $w_i = n_i$ (*binomial weighting*) and solve for the resulting $\hat{\mu}$ and $\hat{\sigma}^2$. With this value of $\hat{\sigma}^2$ and $\hat{\mu}$, calculate r and new values of w_i from Eqs. (14) and (15) (*empirical weighting*). Continue iterating until $\hat{\mu}, \hat{\sigma}^2$ and w_i no longer change (*converged weighting*). Finally it should be noted that $\hat{\sigma}^2$ may be negative from Eq. (11). For this case r is set to zero (i.e. only binomial weighting is used). For each estimate of $\hat{\sigma}$ and $\hat{\mu}$, the corresponding values of a and b of the beta prior are calculated from Eqs. (12) and (13).

The Maximum Likelihood Method Based on the Marginal Distribution [1]

The maximum likelihood method chooses the parameters a and b as those values which maximize the *likelihood function*

$$L(a,b) \equiv L(k_1 \ldots k_N | n_1 \ldots n_N, a, b) = \left\{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\right\}^N \prod_{i=1}^{N} C_i \frac{\Gamma(a+k_i)\Gamma(b+n_i-k_i)}{\Gamma(a+b+n_i)} \quad (16)$$

where

$$C_i \equiv \begin{pmatrix} n_i \\ k_i \end{pmatrix} = \frac{\Gamma(n_i+1)}{\Gamma(k_i+1)\Gamma(n_i-k_i+1)} \quad . \quad (17)$$

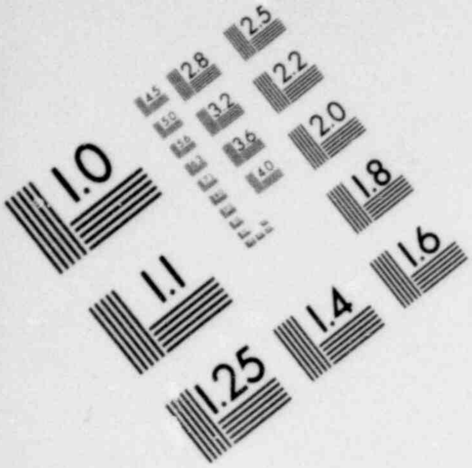
Equivalently, one seeks values of a and b which minimize the logarithm of L, $ln[L]$, since L is always less than unit. This latter form is preferrable for numerical purposes since the $ln\Gamma$ function varies more slowly than does the Γ function. The extrema of $lnL(a,b)$ are obtained from solutions to

$$\frac{\partial lnL}{\partial a}(a,b) = 0$$

$$\frac{\partial lnL}{\partial b}(a,b) = 0$$

or explicitly

$$N\{\psi(a+b) - \psi(a)\} + \sum_{i=1}^{N} \{\psi(a+k_i) - \psi(a+b+n_i)\} = 0 \quad (18)$$



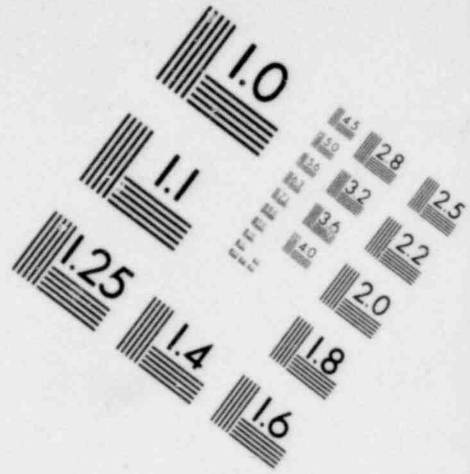
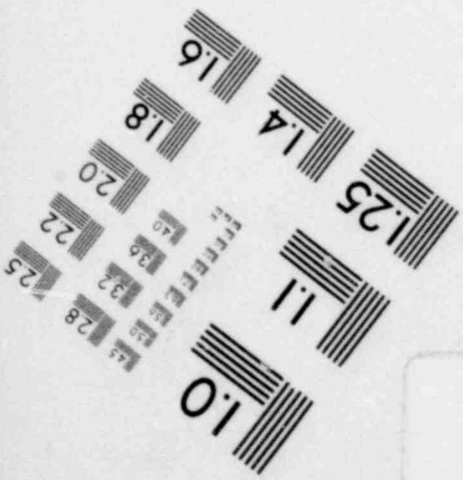


IMAGE EVALUATION TEST TARGET (MT-3)



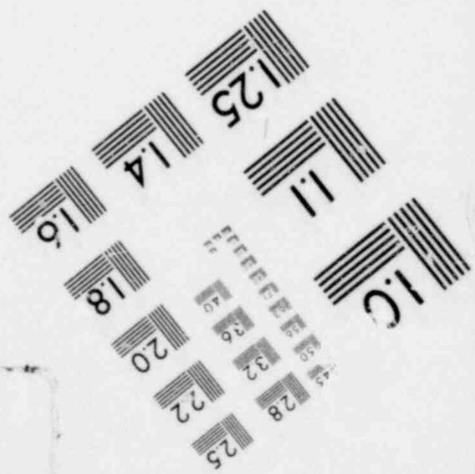


**IMAGE EVALUATION
TEST TARGET (MT-3)**

6"

IMAGE EVALUATION TEST TARGET (MT-3)

and

$$N\{\psi(a+b) - \psi(b)\} + \sum_{i=1}^{N} \{\psi(b+n_i-k_i) - \psi(a+b+n_i)\} = 0, \tag{19}$$

where $\psi(z) = \frac{d}{dz}[\ln\Gamma(z)]$, the digamma function. The numerical solution of these two simultaneous equations can be obtained by pattern search techniques [3]. However, since the first and second derivatives of $\ln L$ are readily evaluated with the polygamma functions*, BETA III uses a Newton-Raphson numerical solution. Care must be taken since $a,b\to\infty$ is also a solution of Eqs. (18) and (19). Also, if the sample data consist soley of one component (N=1), the only solution of the equation is for $a,b=\infty$ but with a/b finite. Also for some data for n>1, it has been found that Eqs. (18) and (19) may have no finite positive solutions.

The Maximum Likelihood Method Based on the Prior Distribution [6]

The maximum likelihood method can also be applied to the prior distribution [Eq. (2)] by defining the likelihood function as

$$L(a,b) \equiv L(p_1,p_2\ldots,p_N|a,b) = \prod_{i=1}^{N} \frac{p_i^{a-1}\,(1-p_i)^{b-1}}{B(a,b)} \tag{20}$$

where $p_i = \frac{k_i}{n_i}$.

The estimates of parameters a and b are chosen to be the values which minimize the logarithm of R, and consequently are solutions to

$$\left.\begin{aligned} &\sum_{i=1}^{N} \ln p_i + N[\psi(a+b)-\psi(a)] = 0 \\ &\sum_{i=1}^{N} \ln(1-p_i) + N[\psi(a+b) -\psi(b)] = 0 \end{aligned}\right\} \tag{21}$$

where $\psi(z)$ is the digamma function. The program BETA III uses a Newton-Raphson method to evaluate numerically the solutions of Eqs. (21).

1.3 Classical and Bayesian Estimates of Mean Failure Probability. [1]

For a given component the classical estimate of the failure probability is simply

$$\hat{P}_c = \frac{k}{n}\,. \tag{22}$$

A Bayesian estimate is obtained by using the expected value of p from the posterior distribution of Eq. (5), namely

$$\hat{P}_B = \frac{a+k}{(a+k) + (b+n-k)}\,.$$

---

*The procedure for evaluation of the polygamma functions is outlined in Addendum A.

1427 001

1.4 Variance of Estimators from the Maximum Likelihood Method Based on the Marginal Distribution. [7]

The information matrix [A] is defined as

$$[A] = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \tag{24}$$

where

$$a_{11} = -E\left(\frac{\partial^2 \ln L}{\partial a^2}\right)$$

$$a_{22} = -E\left(\frac{\partial^2 \ln L}{\partial b^2}\right)$$

$$a_{12} = a_{21} = -E\left(\frac{\partial^2 \ln L}{\partial a \partial b}\right)$$

In the limit of a large number of failure data, the covariance matrix $[\sigma]$ can be obtained from the inversion of the information matrix [6], i.e.,

$$[\sigma] \equiv \begin{pmatrix} \mathrm{var}(\hat{a}) & \mathrm{cov}(\hat{a},\hat{b}) \\ \mathrm{cov}(\hat{a},\hat{b}) & \mathrm{var}(\hat{b}) \end{pmatrix} = [A]^{-1} \tag{25}$$

The elements of the information matrix may be evaluated directly from their definitions as

$$E\left(\frac{\partial^2 \ln L}{\partial a^2}\right) = N[\psi'(\hat{a}+\hat{b})-\psi'(\hat{a})] + \sum_{i=1}^{N} \sum_{k_i=0}^{n_i} \psi'(\hat{a}+k_i)h(k_i|n_i,\hat{a},\hat{b}) - \sum_{i=1}^{N} \psi'(\hat{a}+\hat{b}+n_i)$$

$$E\left(\frac{\partial^2 \ln L}{\partial b^2}\right) = N[\psi'(\hat{a}+\hat{b})-\psi'(\hat{b})] + \sum_{i=1}^{N} \sum_{k_i=0}^{n_i} \psi'(\hat{b}+n_i-k_i)h(k_i|n_i,\hat{a},\hat{b}) - \sum_{i=1}^{N} \psi'(\hat{a}+\hat{b}+n_i) \tag{26}$$

$$E\left(\frac{\partial^2 \ln L}{\partial a \partial b}\right) = N\psi'(\hat{a}+\hat{b}) - \sum_{i=1}^{N} \psi'(\hat{a}+\hat{b}+n_i)$$

where $\psi'(z) = \dfrac{d^2 \ln\Gamma(z)}{dz^2}$ = trigamma function.

In some cases, if there is some evidence showing that the distribution of the likelihood function L is symmetric about the maximum, the expectations in Eq. (26) may be approximated by the relations

$$E\left(\frac{\partial^2 \ln L}{\partial a^2}\right) \simeq \left.\frac{\partial^2 \ln L}{\partial a^2}\right|_{\substack{\hat{a}=\hat{a}\\ \hat{b}=\hat{b}}} = N[\psi'(\hat{a}+\hat{b})-\psi'(\hat{a})]+ \sum_{i=1}^{N} [\psi'(\hat{a}+k_i)-\psi'(\hat{a}+\hat{b}+n_i)]$$

$$E\left(\frac{\partial^2 \ln L}{\partial b^2}\right) \simeq \left.\frac{\partial^2 \ln L}{\partial b^2}\right|_{\substack{a=\hat{a}\\ b=\hat{b}}} = N[\psi'(\hat{a}+\hat{b})-\psi'(\hat{b})] + \sum_{i=1}^{N} [\psi'(\hat{b}+n_i-k_i)-\psi'(\hat{a}+\hat{b}+n_i)] \tag{27}$$

$$E\left(\frac{\partial^2 \ln L}{\partial a \partial b}\right) \simeq \left.\frac{\partial^2 \ln L}{\partial a \partial b}\right|_{\substack{a=\hat{a}\\ b=\hat{b}}} = N\psi'(\hat{a}+\hat{b}) - \sum_{i=1}^{N} \psi'(\hat{a}+\hat{b}+n_i) .$$

Asymptotic properties of the likelihood function guarantees that Eqs. (27) is true when N is sufficiently large.

1.5 Evaluation of the Cumulative Prior Distribution Function

The cumulative distribution function of the beta prior distribution is computed numerically from

$$G(p) = \frac{1}{B(a,b)} \int_0^b z^{a-1} (1-z)^{b-1} dz \tag{28}$$

which is the incomplete beta function. In Addendum B the numerical evaluation of this function is discussed.

2. DESCRIPTION OF BETA III

The FORTRAN program BETA III computes estimates of the a and b parameters of the beta prior distribution by each of the four methods outlined in Section 1.2. As an option, the classical and Bayesian estimates of the failure probability of each component are calculated (see Section 1.3) and plots of the prior distributions as calculated by each method may be specified. A complete listing of BETA III and all its subroutines is given in Addendum C.

2.1 Input Data

The data required by BETA III consists of the observed failure data (k_i and n_i) for each component in the class to be analyzed as well as several program and option parameters. Sequential analyses may be performed for multiple classes (sets of components) by simply adding a set of input data cards for each class to be analyzed.

For each set of components, a complete data set is required. Each data set consists of the four card types described below.

CARD 1 (20A4)

TITLE = any 80 character title to identify the component set.

CARD 2 (3I5, 5G10.3) (NITER, IOUT, IPROB, Y1, Y2, EPS, Z1, Z2)

NITER = maximum number of iterations to be used in the numerical solutions of the maximum likelihood result and in the iterative solution of the marginal matching moments method (default = 30). If this parameter is set to zero, only the two matching moments methods are used.

IOUT = intermediate calculation output parameter. If IOUT = 0, only the final results of all four estimation methods are printed. If IOUT = 1, the results of each iteration in the marginal matching moments method are printed, as well as the results of each Newton-Raphson iteration in the maximum likelihood method.

IPROB = component probability calculation parameter. If IPROB = 1, the classical and Bayesian estimates of the failure probability are computed for each component. Bayesian estimates are given for the prior distribution as determined by the maximum likelihood method and by the prior matching moments method. If IPROB = 0, none of the component failure probabilities is calculated.

Y1,Y2 = initial guess for parameters a and b which are used as the starting values in the Newton-Raphson procedure used in the maximum likelihood method based on the marginal distribution. If Y1=Y2=0.0, the results of the prior matching moments method are used.

EPS = convergence parameter used to terminate the maximum likelihood method iterative solution, and the marginal matching moments iterative solution. For the maximum likelihood solution, iterations end when differences between successive estimates of a and b are less than EPS. For the marginal matching moments method, iterations end when the difference between successive estimates of the prior mean [=a/(a+b)] is less than EPS.

Z_1, Z_2 = initial guess for parameters a and b which are used as the starting values in the Newton-Raphson procedure used in the maximum likelihood method based on the prior distribution. If Z1,Z2=0.0, the results of the prior matching moments method are used.

Card 3 (4G10.3,6I5) (PI,PJ,PK,PF1,NI,NJ,NL,IXOUT,IVAL,IPL)

The parameters on this card control the line printer plots of the density and cummulative distributions of the estimated prior beta functions. The distributions as a function of the failure probability, p, are in general performed for two ranges of the failure probability: First Range PI $\leq$ p $\leq$ PJ, and Second Range PJ $\leq$ p $\leq$ PK. This flexibility allows the use of a fine grid for a range of the independent variable p over which the distributions vary rapidly, and a coarser grid for a range over which the distributions are more slowly varying. The parameter IPL determines whether only one or both ranges of p are to be used.

PI = the lower limit of the failure probability for the First Range over which the estimated prior density function is to be plotted.

PJ = the upper limit of the First Range and the lower limit of the Second Range over which the estimated prior density and cumulative distributions are to be plotted.

PK = the upper limit of the Second Range over which the estimated prior density and cumulative distributions are to be plotted.

PF1 = the lower limit of the First Range over which the estimated prior cumulative distribution is to be plotted. Often PF1 = PI, although when the density distribution becomes unbounded (typically at p = 0), the lower limits of the First Range should be different for the density and cummulative distributions.

NI = the number of points or values to be plotted in the First Range (between PI and PJ for the estimated density functions or between PF1 and PJ for the estimated cumulative distributions). If NI = 0 then NI is set to 51.

NJ = the number of points or values to be plotted in the Second Range (between PJ and PK) for both the density and cumulative distributions. If NJ = 0 then the program sets NJ = 2.

NL = the number of lines used for printing the independent variable axis. If NL = 0 then 51 lines are used.

IXOUT = controls printing of tic marks and values of the independent variable on the independent variable axis (failure probability axis) every IXOUT lines; if IXOUT = 0, tic marks and values are printed every five lines.

IVAL = parameter to control which distributions are tabulated and plotted. If IVAL = -1 the prior density and cumulative distributions are tabulated for each of the four estimation results. If IVAL = 0 results from the four estimation methods are plotted on the same figure (comparison plot). If IVAL = 1 gives both separate and comparison plots as well as tabulations of the density and cumulative distributions.

IPL = parameter to control over which ranges the prior distributions are to be plotted. If IPL = 0, First Range only is plotted. If IPL = 1, Second Range only is plotted. If IPL = 2, plots for both ranges are produced.

Card 4 (use multiple cards if necessary) (16I5) (NN, N(1), K(1), N(2), K(2),...)

NN = number components in class (maximum number 50)

N(I),K(I) = n_i (number of demands) and k_i (number of failures-on-demand) for the i-th component in class being considered. NN pairs of data are required.

2.2 Sample Input

```
FAIRBANKS DIESEL ENGINE DATA -- FOUR PLANTS

   30    0    1   0.0D00   0.0D00   1.0D-12   0.0D00   0.0D00

 0.0D00   2.0D-01   1.0D00   0.0D00    0    0    0    0    1    0

    4  656    3   73    5   35    1   37    1
```

2.3 Sample Output

The output from BETA III can be quite voluminous if all the program options are elected by the user. On the next few pages, portions of example output are given.

3. ACKNOWLEDGMENT

The development of this code was supported by the U. S. Nuclear Regulatory Commission.

Sample Output for 25 Diesel Plants

```
ALL DIESEL ENGINES -- 25 PLANTS
    TRIES:       100  392  230   68   23   23   12   99   33  126   47   87   71  656   73   35   37   13   95   51   3[illegible]   17  335
                 206   76
    FAILURES:      6    1   11    5    4    0    2    0    3    9    2    1    2    3    5    1    1    0    2    2    2    7    4
                   9    1

MATCHING MOMENTS OF DATA TO THOSE OF MARGINAL DISTRIBUTION:
NO WEIGHTING        :  MEAN= 0.59826741D-01     SIGMA= 0.79387034D-01;      PRIOR PARAMETERS:  A= 0.47412184       B=  7.4507932
BINOMIAL WEIGHTING  :  MEAN= 0.28231293D-01     SIGMA= 0.39744376D-01;      PRIOR PARAMETERS:  A= 0.46208133       B= 15.905619
EMPIRICAL WEIGHTING:   MEAN= 0.51166396D-01     SIGMA= 0.65718644D-01;      PRIOR PARAMETERS:  A= 0.5[illegible]8545       B=  9.7168269
CONVERGED RESULT    :  MEAN= 0.53729114D-01     SIGMA= 0.69383896D-01;      PRIOR PARAMETERS:  A= [illegible] 1370766    B=  9.0473594


MATCHING MOMENTS OF THE DATA TO THOSE OF THE PRIOR DISTRIBUTION:
PRIOR MOMENTS:         MEAN= 0.59826741D-01     SIGMA= 0.86416118D-01;      PRIOR PARAMETERS:  A= 0.39079183       B=  6.1412676
VARIANCE AND STANDARD DEVIATION ESTIMATES (ASSUMING NORMAL DISTRIBUTION) :          VAR(A)= 0.720476D-01VAR(B)=  7.42852
                                                                                    SIG(A)= 0.268417     SIG(B)=  2.72553
                                                                                    VAR(A)= 0.139273     VAR(B)= 24.0305
VARIANCE AND STANDARD DEVIATION ESTIMATES (DISTRIBUTION INDEPENDENT) :              SIG(A)= 0.373194     SIG(B)=  4.90210


MAXIMUM LIKELIHOOD METHOD WITH BETA-BINOMIAL DISTRIBUTION:
    INITIAL STARTING POINTS CALCULATED BY MATCHING MOMENTS TO PRIOR 0.39079183      6.1412676
    ACCURACY PARAMETER=  0.1000D-11
    MAXIMUM NUMBER OF ITERATIONS=  30
    SOLUTION CONVERGED TO:   A=  1.0521510       AND   B=  19.901503       AFTER  8 ITERATIONS.
PRIOR MOMENTS:          MEAN= 0.50213245D-01     SIGMA= 0.46608893D-01;      PRIOR PARAMETERS:  A=  1.0521510       B= 19.901503
        EXACT SOLUTION
        INFORMATION MATRIX :          22.1501       -0.887711
                                     -0.887711       0.478217D-01
                                                                                    VAR(A)= 0.176316     VAR(B)=  81.6664
                                                                                COVAR(A,B)=  3.27295

        APPROXIMATE SOLUTION
        INFORMATION MATRIX :          25.3432       -0.887711
                                     -0.887711       0.417636D-01
                                                                                    VAR(A)= 0.154454     VAR(B)=  93.7267
                                                                                COVAR(A,B)=  3.28301


MAXIMUM LIKELIHOOD METHOD WITH BETA DISTRIBUTION:
    INITIAL STARTING POINTS CALCULATED BY MATCHING MOMENTS TO PRIOR 0.39079183      6.1412676
THIS DATA SET IS REJECTED BECAUSE OF 0 NO.OF FAILURE
```

MATCHING DATA MOMENTS TO PRIOR DISTRIBUTION MOMENTS

PROBABILITY DENSITY FUNCTION
OF BETA DISTRIBUTION

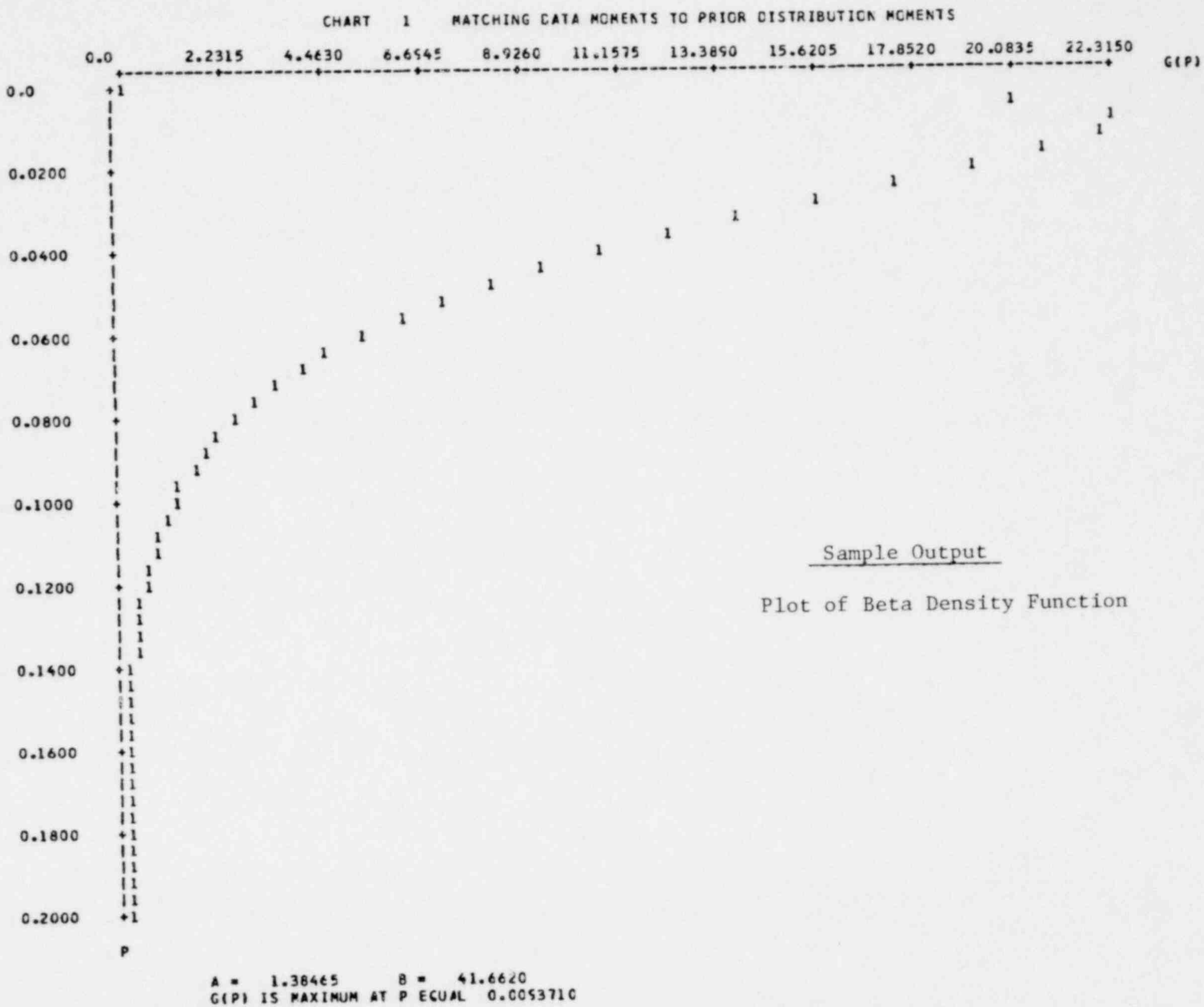
WITH PARAMETERS : A = 1.384647
B = 41.66201
B(A,B) = 0.5046495718D-02
G(P) IS MAXIMUM AT P EQUAL 0.0093710

P	G(P)
0.0	0.0
0.4000000D-02	20.13125
0.8000000D-02	22.31503
0.1200000D-01	22.12951
0.1600000D-01	20.56027
0.2000000D-01	19.35267
0.2400000D-01	17.57809
0.2800000D-01	15.78337
0.3200000D-01	14.05025
0.3600000D-01	12.42328
0.4000000D-01	10.92470
0.4400000D-01	9.563060
0.4800000D-01	8.338519
0.5200000D-01	7.246112
0.5600000D-01	6.277915
0.6000000D-01	5.424431
0.6400000D-01	4.675512
0.6800000D-01	4.020941
0.7200000D-01	3.450818
0.7600000D-01	2.955778
0.8000000D-01	2.527128
0.8400000D-01	2.156904
0.8800000D-01	1.837885
0.9200000D-01	1.563582
0.9600000D-01	1.328202
0.1000000D 00	1.126605
0.1040000	0.9542515
0.1080000	0.8071510
0.1120000	0.6818083
0.1160000	0.5751727
0.1200000	0.4845854
0.1240000	0.4077547
0.1280000	0.3426746
0.1320000	0.2876273
0.1360000	0.2411252
0.1400000	0.2019047
0.1440000	0.1688551
0.1480000	0.1410549
0.1520000	0.1176902
0.1560000	0.9808079D-01
0.1600000	0.8164339D-01
0.1640000	0.6786178D-01
0.1680000	0.5637438D-01
0.1720000	0.4676354D-01
0.1760000	0.3874632D-01
0.1800000	0.3206647D-01
0.1840000	0.2650752D-01
0.1880000	0.2188687D-01
0.1920000	0.1805070D-01
0.1960000	0.1486960D-01
0.2000000	0.1223482D-01

0.2000000	0.1223482D-01
1.000000	0.0

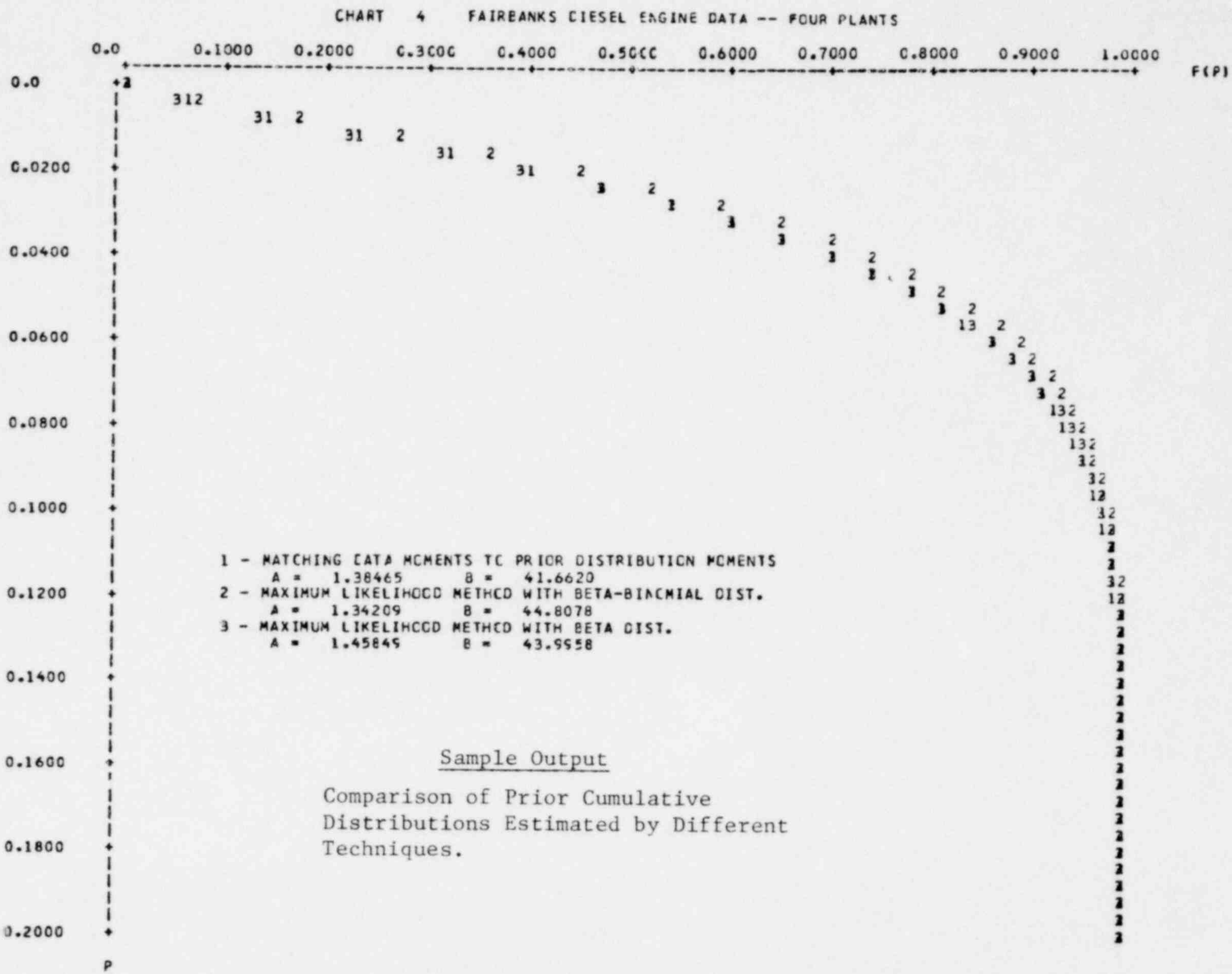
Sample Output

Tabulation of Estimated
Prior Density Function



Sample Output

Plot of Beta Density Function

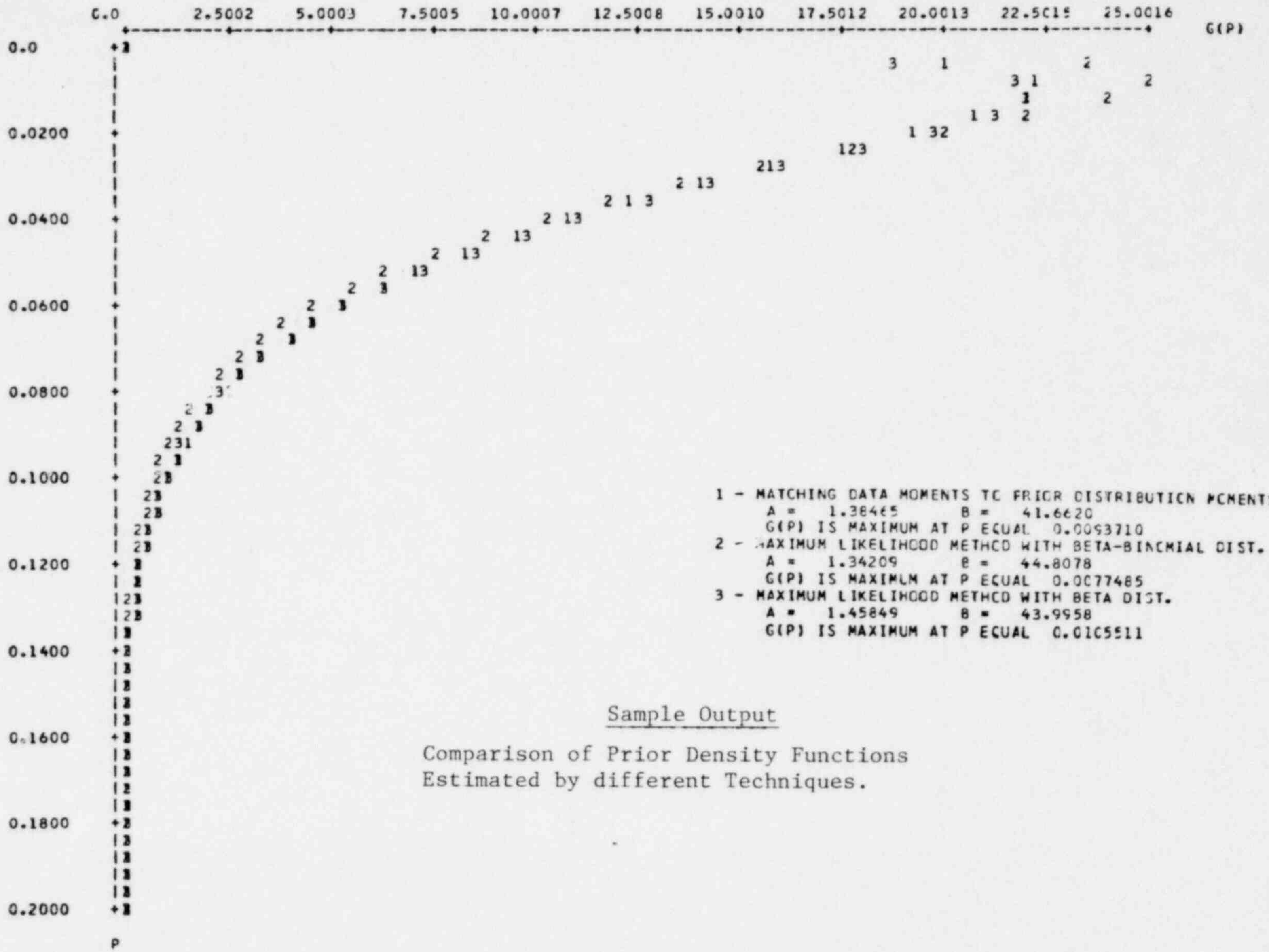


Sample Output

Comparison of Prior Cumulative Distributions Estimated by Different Techniques.

1427 010

CHART 4 FAIRBANKS DIESEL ENGINE DATA -- FOUR PLANTS



Sample Output

Comparison of Prior Density Functions
Estimated by different Techniques.

1427 011

4. REFERENCES

1. J. K. Shultis and N. D. Eckhoff, "Selection of Beta Prior Distribution Parameters From Component Failure Data," published IEEE Transactions, July, 1978.

2. J. C. Kleinman, "Proportions with Extraneous Variance: Single and Independent Samples," J. Am. Stat. Assoc., 68, 46 (1973).

3. J. K. Shultis, "Solution of the Prior Distribution Parameters from the Maximum Likelihood Method by Pattern Search Techniques," Center for Energy Studies Report CES-26, Kansas State University, Manhattan, Kansas, Dec., 1975.

4. K.S. Kölbig, "Programs for Computing the Logarithm of the Gamma Function, and the Digamma Function, For Complex Arguments," Comp. Phys. Comm., 4, 221 (1972).

5. A. T. de Medeiros and G. Schwachleim, "Polygamma Functions with Arbitrary Precision," Comm. ACM, 12, 213 (1969).

6. N. L. Johnson and S. Kotz, "Continous Univariate Distributions-2," John Wiley & Sons (1970).

7. S. L. Meyer, "Data Analysis for Scientists and Engineers, Wiley, New York (1975).

8. O. G. Ludwig, "Incomplete Beta Ratio," Comm. ACM, 6 (1963) 314; also see "Collected Algorithms from CACM," Algorithm 179 and modifications by N.E. Boston and E. L. Battiste (1972), and by M. C. Pike and J. Soo Hoo (1975).

9. Scientific Subroutine Package (360-A-CM-03X) Version III Programmer's Manual, 420-0205, IBM(1963).

Addendum A

Evaluation of Polygamma Functions

In the Newton-Raphson evaluation of the numerical solution of the maximum likelihood estimates by Eqs. (18), (19) or (21) and in evaluation of the variance bounds (Eqs. (26)), both the digamma function and its derivative, the trigamma function, must be evaluated over a wide range of arguments. The procedure used in BETA III is based on a power series expansion of these functions for large arguments, and a recursion relation for small agruments [4,5].

The polygamma function $\psi^m(z)$ is defined as

$$\psi^m(z) = \frac{d^m\psi(z)}{dz^m} = \frac{d^{m+1}}{dz^{m+1}}\,[\ln\Gamma(z)].$$

The digamma function and trigamma functions are special cases of the polygamma function (m=0 and 1 respectively). These functions may be evaluated accurately by the formulae below:

1. Digamma (m=0):

$$z \geq 8 \qquad \psi(z) \approx \ln z - \frac{1}{2z} - \sum_{k=1}^{10} \frac{B_{2k}}{2k}\, z^{-2k}$$

$$z < 8 \qquad \psi(z) = \psi(n+z) - \sum_{k=1}^{n} (z+k-1)^{-1}$$

where B_{2k} are the Bernoulli numbers.

2. Trigamma (m=1):

$$z \geq 8 \qquad \psi^1(z) \approx \frac{1}{2} + \frac{1}{2z^2} + \sum_{k=1}^{10} B_{2k} z^{-(2k+1)}$$

$$z < 8 \qquad \psi^1(z) = \psi^1(n+z) + \sum_{k=1}^{n} (z+k-1)^{-2}$$

3. Polygamma (m>1):

$$z \geq 8 \qquad \psi^m(z) = (-1)^{m-1}\Big[\frac{(m-1)!}{z^m} + \frac{m!}{2z^{m+1}} +$$

$$\sum_{k=1}^{10} B_{2k} \frac{(2k+m-1)!}{(2k)!}\, z^{-(2k+m)}\Big]$$

$$z < 8 \qquad \psi^m(z) = \psi^m(2+n) - (-1)^m\, m! \sum_{k=1}^{n} (z+k-1)^{-m-1}$$

1427 013

Addendum B

Evaluation of Incomplete Beta Function

The incomplete beta function $I_p(x,y)$ is calculated from the following expression [8]

$$I_p(x,y) = \frac{\text{INFSUM}\ p^x\ \Gamma(PS+x)}{\Gamma(PS)\ \Gamma(x+1)} + \frac{p^x\ (1-p)^y\ \Gamma(x+y)\ \text{FINSUM}}{\Gamma(x)\ \Gamma(y+1)} \tag{35}$$

where INFSUM and FINSUM represent two series summations defined as follows:

$$\text{INFSUM} = \sum_{j=1}^{\infty} \frac{x(1-PS)_j}{x+j} \frac{p^j}{j!}, \quad \text{where} \tag{36}$$

$$(1-PS)_j = \begin{cases} 1 & , \ j = 0 \\ \Gamma(1+y-PS)/\Gamma(1-PS) & , \ j > 0 \end{cases} \tag{37}$$

and

$$\text{FINSUM} = \sum_{j=1}^{[y]} \frac{y(y-1)\ldots(y-j+1)}{(x+y-1)(x+y-2)\ldots(x+y-j)} \frac{1}{(1-p)^j} \tag{38}$$

where [y] is equal to the largest integer less than y. If [y]=0, the FINSUM=0. The quantity PS is defined as

$$PS = \begin{cases} 1 & \text{if } y \text{ is integer} \\ y - [y] & , \text{ otherwise} \end{cases} \tag{39}$$

The above algorithm (combined with scaling to avoid numerical inaccuracies encountered when using the gamma function with large arguments) was incorporated into a FORTRAN program MDBETA by Bosten and Battiste [8]. This program (modified in accordance to remarks made by Pike and Soo Hoo [8] was used in the present analysis. The program MDBETA is significantly more accurate than the widely used program BDTR [9], especially at large arguments. For example, in the case p=0.5, x=y=2000, MDBETA gives the correct value, 0.5, while BDTR gives 0.497026.

Addendum C

Listing of Program Beta III

FORTRAN IV G LEVEL 21 MAIN DATE = 79045 13/21/20

```
C********************************* BETA III ************************************
C*                                                                             *
C* THIS PROGRAM                                                                *
C*    - CALCULATES THE PARAMETERS A AND B OF AN ASSUMED BETA MIXING            *
C* DISTRIBUTION BY  FOUR TECHNIQUES:  (1) MATCHING MOMENTS OF THE EXPERIMENTAL *
C* DATA TO THOSE OF THE MARGINAL DISTRIBUTION,   (2) MATCHING MOMENTS OF THE   *
C* DATA TO THOSE OF THE PRIOR DISTRIBUTION,      (3) THE MAXIMUM LIKELIHOOD    *
C* METHOD WITH BETA-BINOMIAL DISTRIBUTION,AND (4) THE MAXIMUM LIKELIHOOD       *
C* METHOD WITH BETA DISTRIBUTION                                               *
C*    - ALSO CALCULATES AND PLOTS BETA DISTRIBUTION (BOTH PROBABILITY DENSITY  *
C* FUNCTION AND CUMULATIVE DISTRIBUTION FUNCTION)                              *
C* FOR EACH METHOD AND COMPARISON OF FOUR METHODS.                             *
C*                                                                             *
C* INPUT DATA:                                                                 *
C*                                                                             *
C*    CARD 1     (20A4)                                                        *
C*       TITLE  = THE TITLE OF THE PROBLEM (80 COLUMNS)                        *
C*                                                                             *
C*    CARD 2     (3I5,5G10.3)                                                  *
C*       NITER  = MAXIMUM NUMBER OF ITERATIONS FOR METHOD 1 AND FOR NUMERICAL  *
C*                SOLUTION IN METHOD 3&4.IF =0 ONLY MOMENTS METHODS            *
C*                CALCULATIONS ARE PERFORMED.                                  *
C*       IOUT   = 1 IF INTERMEDIATE OUTPUT IS DESIRED FOR THE ITERATIONS IN    *
C*                METHOD 1 AND FOR THE NUMERICAL SOLUTION IN METHOD 3&4;IF =0  *
C*                ONLY FINAL RESULTS FOR ALL FOUR  METHODS ARE PRINTED OUT.    *
C*       IPROB  = 1 IF A COMPARISON OF THE CLASSICAL AND BAYESIAN FAILURE
C*                PROBABILITIES FOR EACH COMPONENT IS DESIRED;  IF =0 THIS     *
C*                OPTION IS BYPASSED.                                          *
C*       Y1     = INITIAL GUESS FOR A IN METHOD 3;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*       Y2     = INITIAL GUESS FOR B IN METHOD 3;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*       EPS    = CONVERGENCE PARAMETER FOR METHODS 1,3 & 4.    IN METHOD 1    *
C*                ITERATIONS CONTINUE UNTIL PRIOR MEAN CHANGES BY LESS THAN    *
C*                EPS.IN METHOD 3&4 NEWTON-RAPHSON ITERATIONS CONTINUE UNTIL   *
C*                DERIVATIVES ARE < EPS.                                       *
C*       Z1     = INITIAL GUESS FOR A IN METHOD 4;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*       Z2     = INITIAL GUESS FOR B IN METHOD 4;  IF =0 RESULT FROM METHOD 2 *
C*                WILL BE USED FOR INITIAL GUESS.                              *
C*                                                                             *
C*    CARD 3      (4G10.3,7I5)                                                 *
C*       PI,PJ,PK,PF1,NI,NJ,NL,IXCUT,IVAL,IPL,IBETA                            *
C*       IBETA  = 0; COMPUTED VALUES & PLOTS OF BETA DISTRIBUTIONS ARE         *
C*                    IGNORED.                                                 *
C*       IBETA  = 1; COMPUTED VALUES & PLOTS OF BETA DISTRIBUTIONS ARE         *
C*                    DISPLAYED ( SEE IVAL & IPL FOR MORE DETAILS).            *
C*       SEE MORE EXPLANATION IN SUBROUTINE BETDIS.                            *
C*                                                                             *
C*    CARD 4...   (16I5)                                                       *
C*       NN     = NUMBER OF PAIRS OF DATA POINTS TO BE READ                    *
C*       N(I),K(I) = NUMBER OF TRIES, NUMBER OF FAILURES FOR I-TH PLANT        *
C*                   NN PAIRS OF N(I) AND K(I) ARE TO BE ENTERED.               *
C*                                                                             *
C* SUBROUTINES REQUIRED:                                                       *
C*    NEWRAL - NEWTON-RAPHSON SOLUTION OF TWO SIMULTANEOUS EQUATIONS           *
C*    FNDATA - READS IN STARTS AND FAILURES, N(I) AND K(I). ALSO CALCULATES    *
C*             THE LIKELIHOOD FUNCTION AND ITS DERIVATIVES                     *
```

```
C*               (BETA-BINOMIAL DISTRIBUTION)                                  *
C*     FBT    - CALCULATES THE LIKELIHOOD FUNCTION AND ITS DERIVATIVES         *
C*               (BETA DISTRIBUTIIN)                                           *
C*     POLGAM - CALCULATES THE POLYGAMMA FUNCTION                              *
C*     VARMLE - CALCULATES VARIANCES AND COVARIANCE OF MAXIMUM LIKELIHOOD      *
C*               ESTIMATORS (EXACT EXPECTATION VALUES;BETA-BINOMIAL DIST.)     *
C*     APPMLE - CALCULATES VARIANCES AND COVARIANCE OF MAXIMUM LIKELIHOOD      *
C*               ESTIMATORS (APPROX. EXPECTATION VALUES;BETA-BINOMIAL DIST.)   *
C*     BETDIS - CALCULATE AND PLOT BETA DISTRIBUTION (PROBABILITY DENSITY      *
C*               FUNCTION AND CUMULATIVE DISTRIBUTION FUNCTION)                *
C*     GPA    - USED IN SUBROUTINE BETDIS                                      *
C*     PLOT   - USED IN SUBROUTINE BETDIS                                      *
C*     MOBETA - USED IN SUBROUTINE BETDIS                                      *
C*                                                                             *
C* REMARKS:                                                                    *
C*     DIMENSION OF P,PB,W,N,K ARE NN                                          *
C*                                                                             *
C*******************************************************************************
      REAL*8 Y1,Y2,AA,BB,EPS,F,G,MEAN,SIG,P,PB(50),DFLOAT
      REAL*8 SIGA,SIGB,DSQRT,VARP,VARSIG,A(4),TITLE(20),DABS
      REAL*8 RBAR,W(50) ,WW,PBAR,S,QBAR,SUM1,SUM2,SSS,BA,PPBAR
      REAL*8 HEMT1(20),HEMT2(20),HEMT3(20),HEADT(4,20),CA(4),DB(4)
      REAL*8 PI,PJ,PK,V11,V22,V12,W11,W22,W12,PF1
      REAL*8 Z1,Z2,HEMT4(20)
      REAL*8 VARA,VARB,VARAND,VARBND,SIGAND,SIGBND
      REAL*8 XPBAR,XQBAP,XS,XPQ,XSUM,XSIG,XRBAR,XAA,XBB
      COMMON/DATA/NN,N(50) ,K(50)
      COMMON /Z/ P(50)
      EXTERNAL FNDER,FBT
      DATA HEMT1/'MATC','HING',' CAT','A MO','MENT','S TO',' MAR','GINA'
     *,'L DI','STRI','BUTI','ON M','OMEN','TS  ',6*'    '/
      DATA HEMT2/'MATC','HING',' CAT','A MO','MENT','S TO',' PRI','OR D'
     *,'ISTR','IBUT','ION ','MOME','NTS ',7*'    '/
      DATA HEMT3/'MAXI','MUM ','LIKE','LIHO','OD M','ETHO','D WI','TH B'
     *,'ETA-','BINO','MIAL',' DIS','T.  ',7*'    '/
      DATA HEMT4/'MAXI','MUM ','LIKE','LIHO','OD M','ETHO','D WI','TH B'
     *,'ETA ','DIST','.   ',9*'    '/
C
C**** READ IN THE PROBLEM TITLE AND DATA
   99 READ (5,12,END=98) (TITLE(I),I=1,20)
   12 FORMAT(20A4)
      PRINT 13,(TITLE(I),I=1,20)
   13 FORMAT('1',20A4)
      READ 10,NITER,IOUT,IPROB,Y1,Y2,EPS,Z1,Z2
   10 FORMAT(3I5,5G10.3)
      READ 150,PI,PJ,PK,PF1,NI,NJ,NL,IXOUT,IVAL,IPL,IBETA
  150 FORMAT(4G10.3,7I5)
      CALL FNDATA(Y1,Y2,F,G,A)
      PRINT 14, (N(I),I=1,NN)
      PRINT 17, (K(I),I=1,NN)
   14 FORMAT(5X,'TRIES:    ',23I5,/(15X,23I5))
   17 FORMAT(5X,'FAILURES: ',23I5,/(15X,23I5))
      NITE=NITER
      NOM=0
C
C*** CALCULATE THE PRIOR PARAMETERS BY MATCHING DATA MOMENTS TO MARGINAL DISTRI-
C*** BUTIONS MOMENTS.
      PRINT 610
```

```
0032      610   FORMAT('OMATCHING MOMENTS OF DATA TO THOSE OF MARGINAL DISTRIBUTIO
               *N:')
0033            NMAX=NITER
0034            IF(NMAX.EQ.0) NMAX=20
0035            ITER=0
0036            MCONV=0
0037            DO 51 I=1,NN
0038         51 P(I)=DFLOAT(K(I))/N(I)
0039            XPBAR=0.0D0
0040            DO 300 I=1,NN
0041      300   XPBAR=XPBAR+P(I)
0042            XPBAR=XPBAR/NN
0043            XQBAR=1.0D0-XPBAR
0044            XS=0.0D0
0045            DO 305 I=1,NN
0046      305   XS=XS+(P(I)-XPBAR)**2
0047            XPQ=XPBAR*XQBAR*(NN-1)
0048            XSUM=0.0D0
0049            DO 310 I=1,NN
0050      310   XSUM=XSUM+1.0D0/N(I)
0051            XSUM=XSUM*XPQ/NN
0052            XRBAR=(XS-XSUM)/(XPQ-XSUM)
0053            IF(XPBAR.LE.0.0D0) GO TO 315
0054            XSUM=XRBAR*XPBAR*XQBAR
0055            XSIG=DSQRT(XSUM)
0056            IF(XPBAR*XQBAR.LE.XSUM) GO TO 316
0057            XPQ=1.0D0/XRBAR-1.0D0
0058            XAA=XPBAR*XPQ
0059            XBB=XQBAR*XPQ
0060            PRINT 612,XPBAR,XSIG,XAA,XBB
0061      612   FORMAT(' NO WEIGHTING       :  MEAN=',G15.8,'     SIGMA=',G15.8,
               *';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
0062            GO TO 320
0063      315   PRINT 613,XPBAR
0064      613   FORMAT(' NO WEIGHTING       :  MEAN=',G15.8,'     R IS NEGATIVE')
0065            GO TO 320
0066      316   PRINT 614,XPBAR,XSIG
0067      614   FORMAT(' NO WEIGHTING       :  MEAN=',G15.8,'     SIGMA=',G15.8,
               *';',7X,'PRIOR PARAMETERS: A&B ARE NEGATIVE')
0068      320   CONTINUE
0069            PPBAR=10.0D0
0070            RBAR=0.0D0
0071         50 ITER=ITER+1
          C*** CALCULATE THE WEIGHTS
0072            WW=0.0D0
0073            DO 52 I=1,NN
0074            W(I)=N(I)/(1.0D0+RBAR*(N(I)-1))
0075         52 WW=WW+W(I)
          C*** CALCULATE PBAR AND S
0076            PBAR=0.0D0
0077            DO 53 I=1,NN
0078         53 PBAR=PBAR+W(I)*P(I)
0079            PBAR=PBAR/WW
0080            QBAR=1.0D0-PBAR
0081            S=0.0D0
0082            DO 54 I=1,NN
0083         54 S=S + W(I)*(P(I)-PBAR)**2
0084            S=(NN-1)*S/NN
```

```
      C*** CALCULATE MEAN OF PRIOR AND RBAR
0085        SUM1=0.0D0
0086        SUM2=0.0D0
0087        DO 55 I=1,NN
0088        SSS=W(I)*(1.0D0-W(I)/WW)
0089        SUM1=SUM1+SSS/N(I)
0090     55 SUM2=SUM2+SSS
0091        RBAR=(S-PBAR*QBAR*SUM1)/(PBAR*QBAR*(SUM2-SUM1))
0092        IF (RBAR.LE.0.0D0) RBAR=0.0D0
      C*** CHECK FOR CONVERGENCE
0093        SSS=DABS((PBAR-PPBAR)/PBAR)
0094        IF(SSS.LE.EPS) MCONV=1
0095        PPBAR=PBAR
      C*** CALCULATE THE A AND B PARAMETERS OF THE PRIOR DISTRIBUTION
0096        IF(RBAR) 56,56,57
0097     57 AA=PBAR*(1.0D0/RBAR-1.0D0)
0098        BB=QBAR*(1.0D0/RBAR-1.0D0)
0099        SIG=DSQRT(RBAR*PBAR*QBAR)
0100        IF (ITER.GT.2) GO TO 59
0101        IF (ITER.EQ.1) PRINT 65,PBAR,SIG,AA,BB
0102        IF (ITER.EQ.2) PRINT 66,PBAR,SIG,AA,BB
0103        GO TO 81
0104     59 IF(IOUT.EQ.1) PRINT 69,PBAR,SIG,AA,BB
0105     81 IF (MCONV.EQ.1) PRINT 67,PBAR,SIG,AA,BB
0106        IF((ITER.EQ.NMAX).AND.(MCONV.EQ.0)) PRINT 68,PBAR,SIG,AA,BB
0107        IF((MCONV.EQ.1).OR.(ITER.EQ.NMAX)) GO TO 85
0108        GO TO 50
0109     56 BA=1.0D0/PBAR - 1.0D0
0110        IF(ITER.GT.2) GO TO 61
0111        IF (ITER.EQ.1) PRINT 75,PBAR,BA
0112        IF (ITER.EQ.2) PRINT 76,PBAR,BA
0113        GO TO 82
0114     61 IF(IOUT.EQ.1) PRINT 79,PBAR,BA
0115     82 IF (MCONV.EQ.1) PRINT 77,PBAR,BA
0116        IF((ITER.EQ.NMAX).AND.(MCONV.EQ.0)) PRINT 78,PBAR,BA
0117     65 FORMAT(' BINOMIAL WEIGHTING :  MEAN=',G15.8,'      SIGMA=',
           2G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
0118     66 FORMAT(' EMPIRICAL WEIGHTING:  MEAN=',G15.8,'      SIGMA=',
           1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
0119     67 FORMAT(' CONVERGED RESULT   :  MEAN=',G15.8,'      SIGMA=',
           1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
0120     68 FORMAT(' NO CONVERGENCE     :  MEAN=',G15.8,'      SIGMA=',
           1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
0121     69 FORMAT(23X,'MEAN=',G15.8,'      SIGMA=',
           1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
0122     75 FORMAT(' BINOMIAL WEIGHTING :  MEAN=',G15.8,'      SIGMA=',
           23X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
0123     76 FORMAT(' EMPIRICAL WEIGHTING:  MEAN=',G15.8,'      SIGMA=',
           13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
0124     77 FORMAT(' CONVERGED RESULT   :  MEAN=',G15.8,'      SIGMA=',
           13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
0125     78 FORMAT(' NO CONVERGENCE     :  MEAN=',G15.8,'      SIGMA=',
           13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
0126     79 FORMAT(20X,'   MEAN=',G15.8,'      SIGMA=',
           13X,'NEGATIVE',8X,'     PRIOR PARAMETER B/A=',G15.8)
0127     85   IF(MCONV.NE.1.OR.RBAR.LE.0.0D00) GO TO 86
0128          NCM=NCM+1
0129          DA(NCM)=AA
```

```
            DB(NOM)=BB
            DO 110 I=1,20
      110   HEADT(NOM,I)=HEMT1(I)
      C
      C**** CALCULATE A AND B BY MATCHING THE DATA MOMENTS TO THOSE OF THE PRIOR.
      86    MEAN=0.0D0
            SIG4=0.0D00
            SIG=0.0D0
            DO 34 I=1,NN
            P(I)=DFLOAT(K(I))/N(I)
         34 MEAN=MEAN+P(I)
            MEAN= MEAN/NN
            DO 35 I=1,NN
            SIG4=SIG4+(P(I)-MEAN)**4
         35 SIG=SIG + (P(I)-MEAN)**2
            SIG=SIG/(NN-1)
            SIG4=SIG4/NN
            AA=(MEAN**2/SIG)*(1.0D0-MEAN) - MEAN
            BB=(1.0D0-MEAN)*AA/MEAN
            SSS=DSQRT(SIG)
            IF(NN.LE.2) GO TO 40
            VARP=SIG/NN
            VARSIG=2.0D0*SIG**2/(NN-1)
            VAR=(SIG4-(NN-3)*SIG**2/(NN-1))/NN
            VARA=        ((((2.0D0-3.0D0*MEAN)*MEAN/SIG)-1.0D0)**2)*VARP
           1  + VARSIG*(MEAN**2*(1.0D0-MEAN)/SIG**2)**2
            SIGA=DSQRT(VARA)
            VARB=VARP*(((1.0D0-4.0D0*MEAN+3.0D0*MEAN**2)/SIG)+1.0D0)**2 +
           2 VARSIG*((MEAN*(1.0D0-MEAN)**2)/(SIG**2))**2
            SIGB=DSQRT(VARB)
            VARAND=        ((((2.0D00-3.0D00*MEAN)*MEAN/SIG)-1.0D00)**2)*VARP
           @  +VAR*(MEAN**2*(1.0D00-MEAN)/SIG**2)**2
            SIGAND=DSQRT(VARAND)
            VARBND=VARP*(((1.0D0-4.0D0*MEAN+3.0D0*MEAN**2)/SIG)+1.0D0)**2 +
           2     VAR*((MEAN*(1.0D0-MEAN)**2)/(SIG**2))**2
            SIGBND=DSQRT(VARBND)
            PRINT 37,MEAN,SSS,AA,BB
            PRINT 38,VARA,VARB,SIGA,SIGB,VARAND,VARBND,SIGAND,SIGBND
      38    FORMAT(' VARIANCE AND STANDARD DEVIATION ESTIMATES (ASSUMING NORMA
           *L DISTRIBUTION) :',T92,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
           *T92,'SIG(A)=',G13.6,'SIG(B)=',G13.6/
           *' VARIANCE AND STANDARD DEVIATION ESTIMATES (DISTRIBUTION INDEPEND
           *ENT) :',T92,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
           *T92,'SIG(A)=',G13.6,'SIG(B)=',G13.6)
            GO TO 39
         40 PRINT 37, MEAN,SSS,AA,BB
         37 FORMAT('0',//'0MATCHING MOMENTS OF THE DATA TO THOSE OF THE PRIOR
           1DISTRIBUTION:',/' PRIOR MOMENTS:',8X,'MEAN=',G15.8,5X,'SIGMA=',
           2G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
      39    NOM=NOM+1
            DA(NOM)=AA
            DB(NOM)=BB
            DO 120 I=1,20
      120   HEADT(NOM,I)=HEMT2(I)
      C
      C*** CALCULATE A AND B BY MAX.LIKELIHOOD METHOD WITH BETA-BINOMIAL DISTRIBUTION
            IF(NITER.EQ.0) GO TO 41
            IF(Y1.EQ.0.0D0) GO TO 32
```

```
FORTRAN IV G LEVEL  21                MAIN              DATE = 79045         13/21/20

0173              PRINT 11, Y1,Y2,EPS,NITER
0174           11 FORMAT('0',/'OMAXIMUM LIKELIHOOD METHOD WITH BETA-BINOMIAL DISTRIB
                 *UTION:',
                 1/5X,'INITIAL STARTING POINTS',2G15.8,/5X,'ACCURACY PARAMETER=',
                 2G12.4,/5X,'MAXIMUM NUMBER OF ITERATIONS=',I4)
0175              GO TO 33
0176           32 Y1=AA
0177              Y2=BB
0178              PRINT 36,Y1,Y2,EPS,NITER
0179           36 FORMAT('0',/'OMAXIMUM LIKELIHOOD METHOD WITH BETA-BINOMIAL DISTRIB
                 *UTION:',
                 1/5X,'INITIAL STARTING POINTS CALCULATED BY MATCHING MOMENTS TO PRI
                 2OR',2G15.8,/5X,'ACCURACY PARAMETER=',G12.4,/5X,'MAXIMUM NUMBER OF
                 3ITERATIONS=',I4)
          C*   SOLVE FOR A AND B BY THE NEWTON-RAPHSON METHOD
0180           33 IOT=IOUT
0181              CALL NEWRAL(Y1,Y2,F,G,FNDER,EPS,NITER,IOT)
0182              MEAN=Y1/(Y1+Y2)
0183              SIG=DSQRT(Y1*Y2/((Y1+Y2+1)*(Y1+Y2)**2))
0184              IF (IOT)  15,20,15
0185           15 PRINT 16,Y1,Y2,IOT
0186           16 FORMAT(5X,'SOLUTION CONVERGED TO:   A=',G15.8,'   AND   B=',G15.8,
                 1'     AFTER',I3,' ITERATIONS.')
0187              PRINT 24, MEAN,SIG,Y1,Y2
0188           24 FORMAT(' PRIOR MOMENTS:',8X,'MEAN=',G15.8,'     SIGMA=',
                 1G15.8,';',7X,'PRIOR PARAMETERS:  A=',G15.8,'   B=',G15.8)
          C *** CALCULATE VARIANCES & COVARIANCE OF MAX. LIKELIHOOD ESTIMATORS
0189              CALL VARMLE(Y1,Y2,NN,N,V11,V22,V12)
0190              CALL APPMLE(Y1,Y2,NN,N,K,W11,W22,W12)
0191              NOM=NOM+1
0192              DA(NOM)=Y1
0193              DB(NOM)=Y2
0194              DO 130 I=1,20
0195          130 HEADT(NOM,I)=HEMT3(I)
0196              GO TO 241
0197           20 PRINT 21, Y1,Y2
0198           21 FORMAT(5X,'SOLUTION DID NOT CONVERGE -- LAST VALUES OF A AND B ARE
                 1',2G15.8)
0199              BA=Y2/Y1
0200              PRINT 25, MEAN, \
0201           25 FORMAT(' PRIOR MOMENTS:',8X,'MEAN=',G15.8,'     SIGMA=   ',
                 1'NOT DEFINED          PRIOR PARAMETER B/A=',G15.8)
0202              NITER=0
          C
          C***CALCULATE A AND B BY MAX.LIKELIHOOD METHOD WITH BETA DISTRIBUTION
0203      241    IF(Z1.NE.0.0D0) GO TO 232
0204              Z'=AA
0205              Z2=BB
0206              PRINT 231,Z1,Z2
0207      231    FORMAT('0',/'OMAXIMUM LIKELIHOOD METHOD WITH BETA DISTRIBUTION:',
                 1/5X,'INITIAL STARTING POINTS CALCULATED BY MATCHING MOMENTS TO PRI
                 *OR',2G15.8)
0208              GO TO 233
0209      232    PRINT 211,Z1,Z2
0210      211    FORMAT('0',/'OMAXIMUM LIKELIHOOD METHOD WITH BETA DISTRIBUTION:',
                 */5X,'INITIAL STARTING POINTS',2G15.8)
          C*  REJECT THE DATA SET CONTAINING 0 NO.OF FAILURE
0211      233    DO 210 I=1,NN
```

```
               IF(K(I).GT.0) GO TO 210
               PRINT 615
     615       FORMAT(T2,'THIS DATA SET IS REJECTED BECAUSE OF 0 NO.OF FAILURE')
               GO TO 41
     210       CONTINUE
     C*   SOLVE FOR A AND B BY THE NEWTON-RAPHSON METHOD
               IOT=IOUT
               CALL NEWRAL(Z1,Z2,F,G,FBT,EPS,NITE,IOT)
               IF(IOT)215,220,215
     215       PRINT 16,Z1,Z2,IOT
               MEAN=Z1/(Z1+Z2)
               SIG=DSQRT(Z1*Z2/((Z1+Z2+1)*(Z1+Z2)**2))
               PRINT 24,MEAN,SIG,Z1,Z2
               NOM=NOM+1
               CA(NOM)=Z1
               CB(NOM)=Z2
               DO 230 I=1,20
     230       HEADT(NOM,I)=HEMT4(I)
               GO TO 41
     220       PRINT 21,Z1,Z2
               BA=Z2/Z1
               PRINT 25,MEAN,BA
     C
     C*** CALCULATION OF CLASSICAL AND BAYESIAN FAILURE PROBABILITIES FOR EACH
     C      COMPONENT USING RESULTS OF METHODS 2 AND 3
     41        IF(IPROB.EQ.0) GO TO 140
               PRINT 31
               PRINT 42
          42   FORMAT('0',///'OESTIMATED FAILURE PROBABILITY FOR EACH COMPONENT.
              1 BAYESIAN ESTIMATE BASED ON RESULTS OF MATCHING MOMENTS TO PRIOR')
               PRINT 46
          46   FORMAT(47X,'TRIES    FAILURES    PMEAN-CLASS.    PMEAN-BAYS.')
               DO 45 I=1,NN
          45   PB(I)=(AA+K(I))/(AA+BB+N(I))
               PRINT 47,(N(I),K(I),P(I),PB(I),I=1,NN)
          47   FORMAT(48X,I3,7X,I3,G16.3,G14.3)
               IF (NITER.EQ.0) GO TO 140
     C*** CALCULATION FROM THE A AND B OF THE MAX. LIKELIHOOD FUNCTION SOLUTION
               PRINT 43
          43   FORMAT('OESTIMATED FAILURE PROBABILITY FOR EACH COMPONENT.  BAYESI
              1AN ESTIMATE BASED ON RESULTS FROM MAXIMUM LIKELIHOOD CALCULATIONS.
              2')
               PRINT 46
               DO 48 I=1,NN
          48   PB(I)=(Y1+K(I))/(Y1+Y2+N(I))
               PRINT 47,(N(I),K(I),P(I),PB(I),I=1,NN)
     C*** CALCULATE AND PLOT BETA DISTRIBUTION
     140       IF(IBETA.EQ.0) GO TO 99
               CALL BETDIS(NOM,HEADT,CA,CB,NI,NJ,NL,IXOUT,IVAL,PI,PJ,PK,IPL,
              *TITLE,PF1)
               GO TO 99
     C*
          98   PRINT 31
          31   FORMAT('1')
               STOP
               END
```

1427 021

1427 022

FORTRAN IV G LEVEL 21 DATE = 78309 16/03/17

```
C*****************************************************************************
C*    SUBROUTINE BETDIS(NOC,HEADT,A,B,NI,NJ,NL,IXOUT,IVAL,PI,PJ,PK,IPL,CBT,PF1)*
C*    PURPOSES :                                                              *
C*       - COMPUTE BETA DISTRIBUTION                                          *
C*       - PLOT BETA DISTRIBUTION                                             *
C*       - COMPARE BETA DISTRIBUTION OF DIFFERENT PARAMETERS                  *
C*         (BOTH PROBABILITY DENSITY AND CUMULATIVE DISTRIBUTION FUNCTIONS)   *
C*    DESCRIPTION OF PARAMETERS :                                             *
C*       NOC  - NO. OF BETA DISTRIBUTIONS TO BE COMPARED IN ONE FIGURE        *
C*       HEADT- DESCRIPTION FOR EACH DISTRIBUTION                             *
C*       CBT  - COMPARISON CHART HEADING                                      *
C*       A,B  - BETA DISTRIBUTION PARAMETERS                                  *
C*       IVAL - CONTROL PARAMETER FOR DISPLAYING RESULTS                      *
C*             IVAL=-1 PRINT COMPUTED VALUES ONLY                             *
C*             IVAL= 0 PLOT COMPARISON FIGURE ONLY;IF NOC=1,PLOT 1 CURVE      *
C*             IVAL=1 PRINT COMPUTED VALUES,PLOT INDIVIDUAL CURVE             *
C*                    AND COMPARISON CHART                                    *
C*       IPL  - CONTROL PARAMETER FOR PLOTTING                                *
C*             IPL=0 PLOT NI DATA POINTS FROM PI TO PJ (IF NI=0,NI=51 IS USED) *
C*             IPL=1 PLOT NJ DATA POINTS FROM PJ TO PK (IF NJ=0,NJ=2 IS USED) *
C*             IPL=2 PLOT NI+NJ-1 DATA POINTS FROM PI TO PK                   *
C*       PI   - INDEPENDENT VARIABLE (FIRST DATA POINT)                       *
C*       PJ   - INDEPENDENT VARIABLE (INTERMEDIAT DATA POINT)                 *
C*       PK   - INDEPENDENT VARIABLE (LAST DATA POINT) POINT)                 *
C*             IPL,PI,PJ,PK - USED IN COMPUTING & PLOTTING DENSITY FUNCTION   *
C*       IXOUT- PRINT MARK ON BASE-VARIABLE AXIS EVERY IXOUT DATA POINT       *
C*             IXOUT=0,PRINT EVERY 5 DATA POINTS.                             *
C*       NL   - NO. OF LINES USED FOR PRINTING BASE-VARIABLE AXIS             *
C*             IF NL=0, 51 LINES WILL BE USED                                 *
C*       PF1  - FIRST DATA POINT(=0 USUALLY) USED IN COMPUTING & PLOTTING     *
C*                 DISTRIBUTION FUNCTION.                                     *
C*    SUBROUTINE REQUIRED : GPA,PLOT & MDBETA                                 *
C*    REMARKS :                                                               *
C*       NI AND NJ MUST BE ODD INTEGERS                                       *
C*       DIMENSION OF G,P,GX,PX,F,PF ARE NI+NJ-1                              *
C*       DIMENSIONS OF AA,AAA,FF SHOULD BE 5 TIMES OF G,P,GX,PX,F,PF          *
C*****************************************************************************
      SUBROUTINE BETDIS(NOC,HEADT,A,B,NI,NJ,NL,IXOUT,IVAL,PI,PJ,PK,IPL,
     *CBT,PF1)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION HEADT(4,20),A(4),B(4)
      DIMENSION G(53) ,P(53) ,HEAD(20),GX(53) ,PX(53)
      DIMENSION CBT(20),AA(265),AAA(265)
      DIMENSION F(53),PF(53),FF(265)
      DATA FAX/'F(P)'/
      DATA NS/0/,M/2/
      DATA XAX/'P'/,YAX/'G(P)'/
      DATA GMIN,GMAX,FMIN,FMAX/3*0.0D0,1.0D0/
C
      IF(IXOUT.EQ.0) IXOUT=5
      IF(NI.EQ.0) NI=51
      IF(NJ.EQ.0) NJ=2
      DO 900 NO=1,NOC
      DO 100 I=1,20
  100 HEAD(I)=HEADT(NO,I)
      BAB=DEXP(DLGAMA(A(NO))+DLGAMA(B(NO))-DLGAMA(A(NO)+B(NO)))
      IF(IVAL.EQ.0) GO TO 200
      PRINT 600
```

1427 023

```
FORTRAN IV G LEVEL  21                  BETDIS              DATE = 78309           16/03/17

      600   FORMAT('1')
            PRINT 602,HEAD
      602   FORMAT(//T15,20A4)
            PRINT 605,A(NO),B(NO)
      605   FORMAT(//T15,'PROBABILITY DENSITY FUNCTION',
           */T15,'OF BETA DISTRIBUTION'//
           *T15,'WITH PARAMETERS : A      = ',G15.7/T33,'B       = ',G15.7)
            PRINT 610,BAB
      610   FORMAT(T33,'B(A,B) = ',G18.10)
            CALL GPAI(A(NO),B(NO))
            PRINT 612
      612   FORMAT(/T15,40('-')/T22,'P',T45,'G(P)'/T15,40('-')/)
      C *** CALCULATE DENSITY FUNCTION
      200   CALL GPA(PI,PJ,NI,A(NO),B(NO),AREA1,P,G,IVAL,BAB)
            PRINT 622
      622   FORMAT(' ')
            CALL GPA(PJ,PK,NJ,A(NO),B(NO),AREA2,PX,GX,IVAL,BAB)
            DO 150 I=2,NJ
            NC=NI+I-1
            P(NC)=PX(I)
      150   G(NC)=GX(I)
            IF(IVAL.EQ.0) GO TO 250
            AREA=AREA1+AREA2
            PRINT 444
      C *** PLOT INDIVIDUAL CURVE OF DENSITY FUNCTION
      250   IF(IPL-1) 253,252,251
      251   NT=NI+NJ-1
            IP=0
            GO TO 255
      252   NT=NJ
            GO TO 254
      253   NT=NI
      254   IP=IPL
      255   IF(IVAL.EQ.-1) GO TO 399
            DO 300 I=1,NT
            ID=NI*IP-IP+I
            AA(I)=P(ID)
            AA(NT+I)=G(ID)
            AAA(I)=P(ID)
      300   AAA(NT*NO+I)=G(ID)
            IF(IVAL.EQ.0.AND.NOC.GT.1) GO TO 399
            CALL PLOT(NO,AA,NT,M,NL,NS,HEAD,XAX,YAX,IXOUT,GMAX,GMIN)
            PRINT 660,A(NO),B(NO)
            CALL GPAI(A(NO),B(NO))
      C *** CALCULATE DISTRIBUTION FUNCTION
      399   NF=NT
            NI1=NI-1
            DPF=(PJ-PF1)/NI1
            PF(1)=PF1
            PF(NI)=PJ
            DO 400 I=2,NI1
      400   PF(I)=PF(I-1)+DPF
            DO 401 I=2,NJ
            NC=NI+I-1
      401   PF(NC)=P(NC)
            NI=NI+NJ-1
            DO 410 I=1,NI
      410   CALL MDBETA(PF(I),A(NO),B(NO),F(I),IER)
```

```
              IF(IVAL.EQ.0) GO TO 450
              PRINT 600
              PRINT 602,HEAD
              PRINT 606,A(NC),B(NC)
        606   FORMAT(//T15,'CUMULATIVE DISTRIBUTION FUNCTION',
             */T15,'OF BETA DISTRIBUTION'//
             *T15,'WITH PARAMETERS : A      = ',G15.7/T33,'B      = ',G15.7)
              PRINT 670
        670   FORMAT(/T15,40('-')/T22,'P',T45,'F(P)'/T15,40('-')/)
              DO 420 I=1,NI
        420   PRINT 415,PF(I),F(I)
        415   FORMAT(T14,G15.7,T39,G15.7)
              PRINT 622
              DO 421 I=NI,N1
        421   PRINT 415,PF(I),F(I)
              PRINT 444
        444   FORMAT(/T15,40('-'))
        C *** PLOT INDIVIDUAL CURVE OF DISTRIBUTION FUNCTION
        450   CONTINUE
              IF(IVAL.EQ.-1) GO TO 900
              DO 455 I=1,NF
              ID=NI*IP-IP+I
              AA(I)=PF(ID)
              AA(NF+I)=F(ID)
              FF(I)=PF(ID)
        455   FF(NF*NC+I)=F(ID)
              IF(IVAL.EQ.0.AND.NOC.GT.1) GO TO 900
              CALL PLOT(NO,AA,NF,M,NL,NS,HEAD,XAX,FAX,IXOUT,FMAX,FMIN)
              PRINT 660,A(NC),B(NC)
        900   CONTINUE
        C
        C *** PLOT COMPARISON CURVES
              IF(IVAL.EQ.-1) RETURN
              IF(NOC.EQ.1) RETURN
              NC=NOC+1
              CALL PLOT(NO,AAA,NT,NO,NL,NS,CBT,XAX,YAX,IXCUT,GMAX,GMIN)
              DO 350 I=1,NOC
              PRINT 650,I,(HEADT(I,J),J=1,20)
        650   FORMAT(T20,I2,' - ',20A4)
              PRINT 660,A(I),B(I)
        660   FORMAT(T26,'A = ',G13.6,2X,'B = ',G13.6)
              CALL GPAI(A(I),B(I))
        350   CONTINUE
              CALL PLOT(NO,FF,NF,NO,NL,NS,CBT,XAX,FAX,IXOUT,FMAX,FMIN)
              DO 360 I=1,NOC
              PRINT 650,I,(HEADT(I,J),J=1,20)
              PRINT 660,A(I),B(I)
        360   CONTINUE
              RETURN
              END
```

```
      SUBROUTINE GPA(P1,P2,N,A,B,AREA,P,G,IVAL,BAB)
C*********************************************************************************************
C*    PURPOSE:THIS PROGRAM IS USED IN CONJUNCTION WITH BETDIS ONLY                          *
C*********************************************************************************************
      IMPLICIT REAL*8(A-H,C-Z)
      DIMENSION G(53) ,P(53)
      DP=(P2-P1)/(N-1)
C *** CALCULATE DENSITY AND DISTRIBUTION FUNCTIONS
      P(1)=P1
      P(N)=P2
      N1=N-1
      DO 105 I=2,N1
105   P(I)=P(I-1)+DP
      DO 110 I=1,N
      IF(P(I).LT.1.0D00.AND.P(I).GT.0.0D00) GO TO 106
      GO TO 107
106   GI=(A-1.D0)*DLOG(P(I))+(B-1.D0)*DLOG(1.D0-P(I))-DLOG(BAB)
      IF(GI.GT.-168.D00) GO TO 107
      G(I)=0.0D00
      GO TO 110
107   G(I)=P(I)**(A-1.D0)*(1.D0-P(I))**(B-1.D0)/BAB
110   CONTINUE
      AREA=0.0D00
      IF(IVAL.EQ.0) RETURN
      DO 120 I=1,N
120   PRINT 620,P(I),G(I)
620   FORMAT(T14,G15.7,T39,G15.7)
C *** CHECK VALUES OF DENSITY FUNCTION BY COMPUTING AREA UNDER CURVE
C *** (USING SIMPSON'S RULE)
      GE=0.D0
      GO=0.D0
      DO 125 I=2,N1,2
      GE=GE+G(I)
125   GO=GO+G(I+1)
      AREA=DP/3.D0*(G(1)+4.D0*GE+2.D0*GO-G(N))
      RETURN
C
      ENTRY GPAI(A,B)
C
C *** PRINT REMARK ON EACH BETA DISTRIBUTION
      IF(A-1.0D00) 400,410,420
400   IF(B-1.0D00) 401,402,402
401   PRINT 501
501   FORMAT(T26,'G(P) GOES TO INFINITY AT P EQUAL 0 AND 1')
      RETURN
402   PRINT 502
502   FORMAT(T26,'G(P) GOES TO INFINITY AT P EQUAL 0')
      RETURN
410   IF(B-1.0D00) 411,412,413
411   PRINT 511
511   FORMAT(T26,'G(P) GOES TO INFINITY AT P EQUAL 1')
      RETURN
412   PRINT 512
512   FORMAT(T26,'G(P) IS UNIFORMLY DISTRIBUTED')
      RETURN
413   PRINT 513
513   FORMAT(T26,'G(P) IS MAXIMUM AT P EQUAL 0')
      RETURN
```

1427 026

```
0050      420   IF(B-1.0D00) 411,422,423
0051      422   PRINT 522
0052      522   FORMAT(T26,'G(P) IS MAXIMUM AT P EQUAL 1')
0053            RETURN
0054      423   PMAX=(A-1.0D00)/(A+B-2.0D00)
0055            PRINT 523,PMAX
0056      523   FORMAT(T26,'G(P) IS MAXIMUM AT P EQUAL ',F10.7)
0057            RETURN
0058            END
```

FORTRAN IV G LEVEL 21 NEWRAL DATE = 78305 16/03/17

```
      SUBROUTINE NEWRAL(Y1,Y2,F,G,FN,EPS,NITER,ICONV)
C*******************************************************************************
C*                                                                             *
C* THIS SUBROUTINE SOLVES TWO SIMULTANEOUS EQUATIONS OF THE FORM F(Y1,Y2)=0    *
C* AND G(Y1,Y2)=0 BY THE NEWTON-RAPHSON METHOD.                                *
C* WRITTEN BY J.K. SHULTIS, SEPTEMBER, 1976.                                   *
C*                                                                             *
C* INPUT PARAMETERS:                                                           *
C*       Y1     = STARTING ESTIMATE OF Y1.                                     *
C*       Y2     = STARTING ESTIMATE OF Y2.                                     *
C*       F      = FINAL VALUE OF THE FUNCTION F(Y1,Y2).                        *
C*       G      = FINAL VALUE OF THE FUNCTION G(Y1,Y2).                        *
C*       FN     = NAME OF THE FUNCTION SUBROUTINE WHICH CALCULATES VALUES      *
C*                OF F AND G AND ITS DERIVATIVES.                              *
C*       EPS    = CONVERGENCE CRITERION -- ACCURACY OF SOLUTION                *
C*       NITER  = MAXIMUM NUMBER OF ITERATIONS DESIRED.                        *
C*       ICONV  = 1 IF OUTPUT FOR EACH ITERATION IS DESIRED, =0 OTHERWISE.     *
C*                THIS PARAMETER IS SET TO 0 IF CONVERGENCE IS NOT ACHIEVED    *
C*                OR TO THE ITERATION NUMBER FOR WHICH CONVERGENCE OCCURRED.   *
C*                                                                             *
C*******************************************************************************
      REAL*8 Y1,Y2,F,G,EPS,A(4),X1,X2,DET,DABS,CONVA,CONVB
      IPRINT=ICONV
      IF (IPRINT.EQ.1) PRINT 40
   40 FORMAT('0ITERATION ',7X,'Y1',13X,'Y2',9X,'F(Y1,Y2)',6X,'G(Y1,Y2)')
      ICONV=0
      DO 30 I=1,NITER
      ICONV=ICONV+1
C  THE NEXT TWO CARDS ARE TO BE INCLUDED ONLY IF Y1 AND Y2 MUST BOTH BE >0
      IF (Y1.LT.0.0D0) Y1=DABS(Y1)
      IF (Y2.LT.0.0D0) Y2=DABS(Y2)
      CALL FN(Y1,Y2,F,G,A)
      DET=A(1)*A(4) - A(2)*A(3)
      IF (DET) 10,20,10
   10 X1=(F*A(4) - G*A(3))/DET
      X2=(G*A(1) - F*A(2))/DET
      IF (IPRINT.EQ.1) PRINT 41, ICONV,Y1,Y2,F,G
   41 FORMAT(I5,5X,4G15.8)
      CONVA=DABS(X1/Y1)
      CONVB=DABS(X2/Y2)
      IF (CONVA.LT.EPS) GO TO 1
      GO TO 2
    1 IF (CONVB.LT.EPS) GO TO 3
    2 Y1=Y1-X1
   30 Y2=Y2-X2
      ICONV=0
    3 RETURN
   20 PRINT 11
      ICONV=0
   11 FORMAT(' DETERMINANT IS ZERO -- NO SOLUTION')
      RETURN
      END
```

1427 028

```
FORTRAN IV G LEVEL  21                                  DATE = 78309        16/03/17

      C
            SUBROUTINE FNDATA(Y1,Y2,F,G,A)
            REAL*8 Y1,Y2,X1,X2,X3,F,G,A(4),SUM1,SUM2,SUM3,PCLGAM
            COMMON/DATA/NN,N(50) ,K(50)
      C
      C** READ IN THE PLANT FAILURE DATA
            READ 10,NN,(N(I),K(I),I=1,NN)
         10 FORMAT(16I5)
            RETURN
      C
      C***  BEGIN THE CALCULATION OF THE DERIVATIVES
            ENTRY FNDER(Y1,Y2,F,G,A)
            SUM1=0.0D0
            SUM2=0.0D0
            SUM3=0.0D0
            DO 20 I=1,NN
            X1=Y1+K(I)
            X2=Y2+N(I)-K(I)
            X3=Y1+Y2+N(I)
            SUM1=SUM1 + POLGAM(X1,1)
            SUM2=SUM2 + POLGAM(X2,1)
         20 SUM3=SUM3 + PCLGAM(X3,1)
            X1=PCLGAM(Y1+Y2,1)
            A(1)=NN*(X1-POLGAM(Y1,1)) + SUM1 - SUM3
            A(4)=NN*(X1-POLGAM(Y2,1)) + SUM2 - SUM3
            A(2)=NN*X1 - SUM3
            A(3)=A(2)
      C
      C***  CALCULATE ONLY THE VALUE OF THE F AND G FUNCTIONS
            ENTRY FNONLY(Y1,Y2,F,G)
            SUM1=0.0D0
            SUM2=0.0D0
            SUM3=0.0D0
            DO 30 I=1,NN
            X1=Y1+K(I)
            X2=Y2+N(I)-K(I)
            X3=Y1+Y2+N(I)
            SUM1=SUM1 + PCLGAM(X1,0)
            SUM2=SUM2 + PCLGAM(X2,0)
         30 SUM3=SUM3 + POLGAM(X3,0)
            X1=PCLGAM(Y1+Y2,0)
            F=NN*(X1 - POLGAM(Y1,0)) + SUM1 - SUM3
            G=NN*(X1 - PCLGAM(Y2,0)) + SUM2 - SUM3
            RETURN
            END
```

1427 029

```
C*** FBT ***
      SUBROUTINE FBT(XA,XB,F,G,A)
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION A(4)
      COMMON /DATA/ NN,N(50),K(50)
      COMMON /Z/ P(50)
C***CALCULATE DERIVATIVES
      DUM=POLGAM(XA+XB,1)
      A(1)=NN*(DUM-POLGAM(XA,1))
      A(2)=NN*DUM
      A(3)=A(2)
      A(4)=NN*(DUM-POLGAM(XB,1))
C***CALCULATE VALUES OF THE FUNCTIONS
      SUM1=0.0D00
      SUM2=0.0D00
      DO 100 I=1,NN
      SUM2=SUM2+DLOG(1.0D00-P(I))
  100 SUM1=SUM1+DLOG(P(I))
      DUM=POLGAM(XA+XB,0)
      F=SUM1+NN*(DUM-POLGAM(XA,0))
      G=SUM2+NN*(DUM-POLGAM(XB,0))
      RETURN
      END
```

```
FORTRAN IV G LEVEL  21                  POLGAM             DATE = 78305          16/03/17

0001            REAL FUNCTION POLGAM*8(Z,M)
          C*************************************************************************************
          C*                                                                                   *
          C* THIS FUNCTION CALCULATES THE POLYGAMMA FUNCTION FOR REAL POSITIVE ARGUMENTS *
          C* USING AN ASYMPTOTIC SERIES EXPANSION FOR LARGE ARGUMENTS, AND THEN A RECUR- *
          C* SION RELATION FOR SMALL ARGUMENTS.  THIS METHOD IS DESCRIBED BY A. TADEU DE *
          C* MEDEIROS AND G. SCHWACHHEIM, COMM. ACM, 12 (1969) 213.      CODE PREPARED BY *
          C* J.K. SHULTIS, JULY 1976.                                                          *
          C*                                                                                   *
          C* INPUT PARAMETERS:                                                                 *
          C*        Z = REAL POSITIVE ARGUMENT FOR POLYGAMMA FUNCTION                          *
          C*        M = INDEX OR DERIVATIVE ORDER OF THE POLYGAMMA FUNCTION                    *
          C*                                                                                   *
          C*************************************************************************************
0002            REAL*8 B(10),Z,X,DLOG,DGAMMA,PSI,TRI,NFAC,ARG1,ARG2,AA
          C
          C*** INITIALIZE THE VECTOR B TO THE EVEN BERNOULLI NUMBERS
0003            NBERN=10
0004            IF(Z.GT.100.0D0) NBERN=IDINT(10.0D0/DLOG(Z)) + 1
0005            B(1)=0.1666666666666667D0
0006            B(2)=-0.3333333333333333D-01
0007            B(3)=0.2380952380952380-01
0008            B(4)=-0.3333333333333333D-01
0009            B(5)=0.7575757575757576
0010            B(6)=-0.2531135531135531
0011            B(7)=1.1666666666666667
0012            B(8)=-7.0921568627451
0013            B(9)= 54.9711779448621
0014            B(10)=-529.124242424242
0015            IF (M-1) 12,13,20
          C
          C*** CALCULATE THE DIGAMMA OR PSI FUNCTION (M=0)
          C*** CALCULATE WHETHER Z > 8
0016         12 NN=Z
0017            N=8-NN
0018            N=MAX0(0,N)
0019            X=Z+N
          C*** CALCULATE PSI FOR X > 8
0020            PSI=0.0D0
0021            DO 10 K=1,NBERN
0022            I=2*K
0023         10 PSI=PSI + B(K)/(K*X**I)
0024            PSI=DLOG(X) - 0.5D0*(1.D0/X + PSI)
          C*** CALCUATE FOR Z< 8 IF NECESSARY
0025            IF (N) 15,15,14
0026         14 DO 16 NN=1,N
0027         16 PSI=PSI - 1.D0/(Z+NN-1)
0028         15 POLGAM=PSI
0029            RETURN
          C
          C*** CALCULATION OF THE TRIGAMMA FUNCTION (M=1)
0030         13 NN=Z
0031            N=8-NN
0032            N=MAX0(0,N)
0033            X=Z+N
          C*** CALCULATE FOR Z >8
0034            TRI=0.0D0
0035            DO 17 K=1,NBERN
```

```
      I=2*K+1
   17 TRI=TRI+ B(K)/X**I
      TRI=1.D0/X + 0.5D0/X**2 + TRI
C*** CALCULATE FOR Z < 8
      IF (N) 18,18,19
   19 DO 11 NN=1,N
   11 TRI=TRI + 1.0D0/(Z+NN-1)**2
   18 POLGAM=TRI
      RETURN
C
C*** CALCULATION OF THE GENERAL POLYGAMMA FUNCTION (M > 1)
   20 NN=Z
      N=8-NN
      N=MAX0(0,N)
      X=Z+N
      POLGAM=0.0D0
      MM=M+1
      ARG1=MM
      NFAC=DGAMMA(ARG1)
      ISIGN=4*(M/2) - 2*M + 1
C*** CALCULATE FOR Z > 8
      DO 27 K=1,NBERN
      I=2*K+M
      ARG1=I
      ARG2=2*K+1
   27 POLGAM=POLGAM + B(K)*DGAMMA(ARG1)/(DGAMMA(ARG2)*X**I)
      POLGAM=-ISIGN*(NFAC/(M*X**M) + 0.5D0*NFAC/X**MM + POLGAM)
C*** CALCULATE FOR Z < 8
      IF (N) 28,28,29
   29 AA=0.0D0
      DO 21 NN=1,N
   21 AA=AA + 1.0D0/(Z+NN-1)**MM
      POLGAM=POLGAM - ISIGN*NFAC*AA
   28 RETURN
      END
```

FORTRAN IV G LEVEL 21 APPMLE DATE = 78305 16/03/17

```
      SUBROUTINE APPMLE(A,B,NN,N,K,U11,U22,U12)
C**************************************************************************
C*                                                                        *
C*    PURPOSE : CALCULATE VARIANCES AND COVARIANCES                       *
C*              OF MAXIMUM LIKELIHOOD ESTIMATORS                          *
C*              OF PARAMETERS A AND B                                     *
C*              OF BETA PRIOR DISTRIBUTION                                *
C*    PARAMETER DESCRIPTION :                                             *
C*       A     ESTIMATOR OF A                                             *
C*       B     ESTIMATOR OF B                                             *
C*       NN    NUMBER OF OBSERVED DATA                                    *
C*       N(I)  NUMBER OF TRIES                                            *
C*       U11   VARIANCE(A)                                                *
C*       U22   VARIANCE(B)                                                *
C*       U12   COVARIANCE(A,B)                                            *
C*    SUBROUTINE REQUIRED :                                               *
C*       POLGAM  CALCULATE POLYGAMMA FUNCTIONS                            *
C*    REMARK :                                                            *
C*       APPROX. EXPECTATION VALUES BY 2-ND DERIVATIVES OF                *
C*       LIKELIHOOD FUNCTION                                              *
C*                                                                        *
C**************************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION N(50),K(50)
C *** CALCULATE INFORMATION MATRIX
      W11=NN*(POLGAM(A+B,1)-POLGAM(A,1))
      W22=NN*(POLGAM(A+B,1)-POLGAM(B,1))
      W12=NN*POLGAM(A+B,[illegible]
      DO 100 I=1,NN
      AG1=A+K(I)
      AG2=A+B+N(I)
      AG3=B+N(I)-K(I)
      W11=W11+POLGAM(AG1,1)-POLGAM(AG2,1)
      W22=W22+POLGAM(AG3,1)-POLGAM(AG2,1)
      W12=W12-POLGAM(AG2,1)
  100 CONTINUE
      W11=-W11
      W22=-W22
      W12=-W12
      PRINT 605
  605 FORMAT(T10,'APPROXIMATE SOLUTION')
      PRINT 620,W11,W12,W12,W22
  620 FORMAT(T10,'INFORMATION MATRIX : ',(T35,2(2X,G13.6)))
C *** CALCULATE VARIANCES AND COVARIANCE
      DET=W11*W22-W12*W12
      U11= W22/DET
      U22= W11/DET
      U12=-W12/DET
      PRINT 630,U11,U22,U12
  630 FORMAT(91X,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
     *87X,'COVAR(A,B)=',G13.6)
      RETURN
      END
```

```
      SUBROUTINE VARMLE(Y1,Y2,NN,N,V11,V22,V12)
C***********************************************************************
C*                                                                     *
C*    PURPOSE : CALCULATE VARIANCES AND COVARIANCES                    *
C*              OF MAXIMUM LIKELIHOOD ESTIMATORS                       *
C*              OF PARAMETERS A AND B                                  *
C*              OF BETA PRIOR DISTRIBUTION                             *
C*    PARAMETER DESCRIPTION :                                          *
C*       Y1   ESTIMATOR OF A                                           *
C*       Y2   ESTIMATOR OF B                                           *
C*       NN   NUMBER OF OBSERVED DATA                                  *
C*       N(I) NUMBER OF TRIES                                          *
C*       V11  VARIANCE(A)                                              *
C*       V22  VARIANCE(B)                                              *
C*       V12  COVARIANCE(A,B)                                          *
C*    SUBROUTINE REQUIRED :                                            *
C*       POLGAM  CALCULATE POLYGAMMA FUNCTIONS                         *
C*    REMARK :                                                         *
C*       USING EXACT EXPECTATION VALUES                                *
C*                                                                     *
C***********************************************************************
      IMPLICIT REAL*8(A-H,O-Z)
      DIMENSION N(50)
C *** CALCULATE INFORMATION MATRIX
      HL1=DLGAMA(Y1+Y2)-DLGAMA(Y1)-DLGAMA(Y2)
      PG1=POLGAM(Y1+Y2,1)
      E11=NN*(PG1-POLGAM(Y1,1))
      E22=NN*(PG1-POLGAM(Y2,1))
      E12=NN*PG1
      DO 200 I=1,NN
      AG1=N(I)+1
      AG2=Y1+Y2+N(I)
      HL2=DLGAMA(AG1)-DLGAMA(AG2)
      PG2=POLGAM(AG2,1)
      E11=E11-PG2
      E22=E22-PG2
      E12=E12-PG2
      NI=N(I)+1
      DO 200 KK=1,NI
      KI=KK-1
      AG3=Y1+KI
      AG4=Y2+N(I)-KI
      AG5=KI+1
      AG6=N(I)-KI+1
      HL3=DLGAMA(AG3)+DLGAMA(AG4)-DLGAMA(AG5)-DLGAMA(AG6)
      H=DEXP(HL1+HL2+HL3)
      E11=E11+POLGAM(AG3,1)*H
      E22=E22+POLGAM(AG4,1)*H
  200 CONTINUE
      E11=-E11
      E22=-E22
      E12=-E12
      PRINT 606
  606 FORMAT(T10,'EXACT SOLUTION')
      PRINT 620,E11,E12,E12,E22
  620 FORMAT(T10,'INFORMATION MATRIX : ',(T35,2(2X,G13.6)))
C *** CALCULATE VARIANCES AND COVARIANCE
      DET=E11*E22-E12*E12
```

```
FORTRAN IV G LEVEL  21                VARMLE              DATE = 78309          16/03/17

 0037              V11= E22/DET
 0038              V22= E11/DET
 0039              V12=-E12/DET
 0040              PRINT 630,V11,V22,V12
 0041        630   FORMAT(91X,'VAR(A)=',G13.6,'VAR(B)=',G13.6/
                  *87X,'COVAR(A,B)=',G13.6)
 0042              RETURN
 0043              END
```

1427 035

1427 036

```
      SUBROUTINE MCBETA(X, P, Q, PROB, IER)
C*****************************************************************************
C*                                                                           *
C*  FUNCTION:       EVALUATE THE INCOMPLETE BETA DISTRIBUTION FUNCTION       *
C*                                                                           *
C*  PARAMETERS:                                                              *
C*  X    - VALUE TO WHICH FUNCTION IS TO BE INTERGRATED. X MUST BE IN THE    *
C*          RANGE (0,1) INCLUSIVE.                                           *
C*  P    - INPUT (1ST) PARAMETER (MUST BE GREATER THAN 0)                    *
C*  Q    - INPUT (2ND) PARAMETER (MUST BE GREATER THAN 0)                    *
C*  PROB - OUTPUT PROBABILITY THAT A RANDOM VARIABLE FROM A BETA DISTRIBUTION *
C*          HAVING PARAMETERS P AND Q WILL BE LESS THAN OR EQUAL TO X.       *
C*  IER  - ERROR PARAMETER.                                                  *
C*            IER = 0 INDICATES A NORMAL EXIT                                *
C*            IER = 1 INDICATES THAT X IS NOT IN THE RANGE (0,1) INCLUSIVE   *
C*            IER = 2 INDICATES THAT P AND/OR Q IS LESS THAN OR EQUAL TO 0.  *
C*                                                                           *
C* CODE BASED ON SIMILAR CODE BY N. BOSTEN AND E.BATTISTE AS MODIFIED BY     *
C* M. PIKE AND J. HCC.                                                       *
C*                                                                           *
C*****************************************************************************
      DOUBLE PRECISION PS,PX,Y,P1,DP,INFSUM,CNT,WH,XB,DQ,C,EPS,EPS1
      DOUBLE PRECISION ALEPS,FINSUM,PQ,DA,DLGAMA
      DOUBLE PRECISION X,P,Q,PROB
C DOUBLE PRECISION FUNCTION DLGAMA
C MACHINE PRECISION
      DATA EPS/1.D-6/
C SMALLEST POSITIVE NUMBER REPRESENTABLE
      DATA EPS1/1.D-78/
C NATURAL LOG OF EPS1
      DATA ALEPS/-179.6016D0/
C CHECK RANGES OF THE ARGUMENTS
      Y = X
      IF ((X.LE.1.0) .AND. (X.GE.0.0)) GO TO 10
      IER = 1
      GO TO 140
   10 IF ((P.GT.0.0) .AND. (Q.GT.0.0)) GO TO 20
      IER = 2
      GO TO 140
   20 IER =0
      IF (X.GT.0.5) GO TO 30
      INT = 0
      GO TO 40
C SWITCH ARGUMENTS FOR MORE EFFICIENT USE OF THE POWER
C SERIES
   30 INT = 1
      TEMP = ?
      P = Q
      Q = TEMP
      Y = 1.D0 - Y
   40 IF (X.NE.0. .AND. X.NE.1.) GO TO 60
C SPECIAL CASE - X IS 0. OR 1.
   50 PROB = 0.0D00
      GO TO 130
   60 IB = Q
      TEMP = IB
      PS = Q -DFLOAT(IB)
      IF (Q.EQ.TEMP) PS = 1.D0
```

```
FORTRAN IV G LEVEL  21                 MOBETA              DATE = 78309          16/03/17

            DP = P
            DQ = Q
            PX = DP*DLOG(Y)
            PQ = DLGAMA(DP+DQ)
            P1 = DLGAMA(DP)
            C = DLGAMA(DQ)
            C4 = DLOG(DP)
      C DLGAMA IS A FUNCTION WHICH CALCULATES THE DOUBLE
      C PRECISION LOG GAMMA FUNCTION
            XB = PX + DLGAMA(PS+DP) - DLGAMA(PS) - C4 - P1
      C SCALING
            IB = XB/ALEPS
            INFSUM = 0.D0
      C FIRST TERM OF A DECREASING SERIES WILL UNDERFLOW
            IF (IB.NE.0) GO TO 90
            INFSUM = DEXP(XB)
            CNT = INFSUM*DP
      C CNT WILL EQUAL DEXP(TEMP)*(1.D0-PS)I*P*Y**I/FACTORIAL(I)
            WH = 0.0D0
         80 WH = WH + 1.D0
            CNT = CNT*(WH-PS)*Y/WH
            XB=DP+WH
            IF(CNT.LE.XB*EPS1) GO TO 90
            XB=CNT/XB
            INFSUM = INFSUM + XB
            IF (XB/EPS.GT.INFSUM) GO TO 80
      C DLGAMA IS A FUNCTION WHICH CALCULATES THE DOUBLE
      C PRECISION LOG GAMMA FUNCTION
         90 FINSUM = 0.D0
            IF (DQ.LE.1.D0) GO TO 120
            XB = PX+ DQ*DLOG(1.D0-Y) + PQ - P1 - DLOG(DQ) - C
      C SCALING
            IB = XB/ALEPS
            IF (IB.LT.0) IB = 0
            C = 1.D0/(1.D0-Y)
            CNT = DEXP(XB-DFLOAT(IB)*ALEPS)
            PS = DQ
            WH = DQ
        100 WH =WH -1.D0
            IF (WH.LE.0.0D0) GO TO 120
            PX = (PS*C)/ (DP+WH)
            IF (PX.GT.1.0D0) GO TO 105
            IF (CNT/EPS.LE.FINSUM.OR.CNT.LE.EPS1/PX) GO TO 120
        105 CNT =CNT*PX
            IF (CNT.LE.1.D0) GO TO 110
      C RESCALE
            IB = IB - 1
            CNT = CNT*EPS1
        110 PS =WH
            IF (IB.EQ.0) FINSUM = FINSUM + CNT
            GO TO 100
        120 PROB =FINSUM + INFSUM
        130 IF (INT.EQ.0) GO TO 140
            PROB = 1.0 - PROB
            TEMP = P
            P = Q
            Q = TEMP
        140 RETURN
            END
```

```
FORTRAN IV G LEVEL  21                                   CATE = 78309          16/03/17
              C
              C     .................................. ......................................
              C
              C        SUBROUTINE PLOT
              C
              C        PURPOSE
              C           PLOT SEV   CROSS-VARIABLES VERSUS A BASE VARIABLE
              C
              C        USAGE
              C           CALL PLOT (NO,A,N,M,NL,NS,CDT,XAX,YAX,IXOUT,AXMX,AXMN)
              C
              C        DESCRIPTICN OF PARAMETERS
              C           NO - CHART NUMBER (3 DIGITS MAXIMUM)
              C           A  - MATRIX OF DATA TO BE PLOTTED. FIRST COLUMN REPRESENTS
              C                BASE VARIABLE AND SUCCESSIVE COLUMNS ARE THE CROSS-
              C                VARIABLES (MAXIMUM IS 9).
              C           N  - NUMBER OF ROWS IN MATRIX A
              C           M  - NUMBER OF COLUMNS IN MATRIX A (EQUAL TO THE TOTAL
              C                NUMBER OF VARIABLES). MAXIMUM IS 10.
              C           NL - NUMBER OF LINES IN THE PLOT. IF 0 IS SPECIFIED, 51
              C                LINES ARE USED.
              C           NS - CODE FOR SORTING THE BASE VARIABLE DATA IN ASCENDING
              C                ORDER
              C                  0  SORTING IS NOT NECESSARY (ALREADY IN ASCENDING
              C                     ORDER).
              C                  1  SORTING IS NECESSARY.
              C           CDT- CHART DESCRIPTION (60 CHARACTERS,DIMENSICN 20)
              C           XAX- BASE VARIABLE-AXIS DESCRIPTION (6 CHARACTERS)
              C           YAX- CROSS VARIABLE-AXIS DESCRIPTION (6 CHARACTERS)
              C           IXOUT - MARKS ON BASE VARIABLE-AXIS WILL BE PRINTED
              C                   EVERY IXOUT DATA POINTS
              C                   IXOUT=0 PRINT MARK ON EVERY DATA POINT
              C           AXMX  - MAXIMUM VALUE ON THE CROSS VARIABLE AXIS
              C           AXMN  - MINIMUM VALUE ON THE CROSS VARIABLE AXIS
              C                   IF AXMX & AXMN = 0.0000,MAX.& MIN. VALUES
              C                   IN THE MATRIX A WILL BE USED
              C
              C        REMARKS
              C           NONE
              C
              C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
              C           NONE
              C
              C     ..........................................................................
              C
                    SUBROUTINE PLOT(NO,A,N,M,NL,NS,CDT,XAX,YAX,IXOUT,AXMX,AXMN)
                    IMPLICIT REAL*8(A-H,O-Z)
                    DIMENSICN OUT(101),YPR(11),ANG(9),A(265)
                    DIMENSICN CDT(20)
                    DATA BLANK/' '/,ANG/'1','2','3','4','5','6','7','8','9'/
              C
                  1 FORMAT(1H1,37X,'CHART ',I3,4X,20A4/)
                  7 FORMAT(1H ,16X,'+',10('---------+'),5X,A6)
                  8 FORMAT(1H ,9X,11F10.4)
                  9 FORMAT(1H0,15X,A6/)
              C
              C     ..........................................................................
              C
```

FORTRAN IV G LEVEL 21 PLOT DATE = 78305 16/03/17

```
            IF(IXOUT.EQ.0) IXOUT=1
            NLL=NL
      C
            IF(NS) 16, 16, 10
      C
      C        SORT BASE VARIABLE DATA IN ASCENDING ORDER
      C
         10 DO 15 I=1,N
            DO 14 J=I,N
            IF(A(I)-A(J)) 14, 14, 11
         11 L=I-N
            LL=J-N
            DO 12 K=1,M
            L=L+N
            LL=LL+N
            F=A(L)
            A(L)=A(LL)
         12 A(LL)=F
         14 CONTINUE
         15 CONTINUE
      C
      C        TEST NLL
      C
         16 IF(NLL) 20, 18, 20
         18 NLL=51
      C
      C        PRINT TITLE
      C
         20 WRITE(6,1)NO,CDT
      C
      C        FIND SCALE FOR BASE VARIABLE
      C
            XSCAL=(A(N)-A(1))/(NLL-1)
      C
      C        FIND SCALE FOR CROSS-VARIABLES
      C
            IF(AXMX.LE.AXMN) GO TO 22
            YMIN=AXMN
            YMAX=AXMX
            GO TO 41
       22   M1=N+1
            YMIN=A(M1)
            YMAX=YMIN
            M2=M*N
            DO 40 J=M1,M2
            IF(A(J)-YMIN) 28,26,26
         26 IF(A(J)-YMAX) 40,40,30
         28 YMIN=A(J)
            GO TO 40
         30 YMAX=A(J)
         40 CONTINUE
       41   YSCAL=(YMAX-YMIN)/100.0D0
      C
      C        PRINT CROSS-VARIABLES NUMBERS
      C
            YPR(1)=YMIN
            DO 90 KN=1,9
         90 YPR(KN+1)=YPR(KN)+YSCAL*10.0D0
```

1427 039

```
          YPR(11)=YMAX
          WRITE(6,8)(YPR(IP),IP=1,11)
          WRITE(6,7)YAX
C
C         FIND BASE VARIABLE PRINT POSITION
C
          XB=A(1)
          L=1
          MY=M-1
          I=1
       45 F=I-1
          XPR=XB+F*XSCAL
          IF(A(L)-XPR) 50,50,46
46        IF(DABS(A(L)-XPR)-XSCAL*0.5D00) 50,70,70
C
C         FIND CROSS-VARIABLES,PRINT LINE AND CLEAR,OR SKIP
C
       50 WRITE(6,100)
      100 FORMAT(1H )
          DO 60 J=1,MY
          DO 55 IX=1,101
       55 OUT(IX)=BLANK
          LL=L+J*N
          JP=((A(LL)-YMIN)/YSCAL)+1.0D00
          OUT(JP)=ANG(J)
          IF((L-1)-(L-1)/IXOUT*IXOUT)56,57,56
56        IF(J.GT.1) GO TO 58
          WRITE(6,110) (OUT(IZ),IZ=1,101)
110       FORMAT(1H+,15X,'|',101A1)
          GO TO 60
58        WRITE(6,111) (OUT(IZ),IZ=1,101)
111       FORMAT(1H+,16X,101A1)
          GO TO 60
57        IF(J.GT.1) GO TO 58
          WRITE(6,2) XPR,(OUT(IZ),IZ=1,101)
        2 FORMAT(1H+,F11.4,4X,'+',101A1)
       60 CONTINUE
          L=L+1
          GO TO 80
       70 WRITE(6,3)
        3 FORMAT(1H ,15X,'|')
       80 I=I+1
          IF(I-NLL) 45, 84, 86
       84 XPR=A(N)
          GO TO 50
       86 CONTINUE
          WRITE(6,9) XAX
          RETURN
          END
```

APPENDIX II

A User's Guide to the Program

TAILS

KSU-2662-7
CES-58
April 24, 1978

A User's Guide to the Program

TAILS

By

J. Kenneth Shultis
Dept. of Nuclear Engineering
Kansas State University
Manhattan, Kansas 66506

ABSTRACT

The FORTRAN program TAILS calculates confidence limits and probability intervals for the failure probability of a component. In particular, confidence limits at arbitrary confidence levels, are calculated by a classical description of the failure probability for a component which has experienced a given number of failures in a specified number of operations. A Bayesian analysis of the same component (whose failure probability is assumed to come from a specified beta prior distribution) is performed to obtain from the posterior distribution the probability interval for the component failure probability.

CONTENTS

1. THEORY

In the reliability analysis of a system, the probability of failure of a particular component is often of great concern. Estimates of the component failure probability can be obtained by both classical and Bayesian analyses [1]. In this document, the theory of obtaining confidence intervals or probability intervals for such estimates is reviewed, and a code to compute these intervals is described. A more complete description is given in Ref. [2].

1.1 Review of the Classical Analysis

For a component which has experienced k failures in n operations, classical analysis estimates the component failure probability to be $\hat{p} = k/n$. Further if p is the true failure probability, then the probability of obtaining k failures in n operations is given by the binomial distribution

$$f(k|n,p) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}. \quad (1)$$

The probability of observing k or fewer failures in n tries is then

$$F(k|n,p) = \sum_{\ell=0}^{k} \frac{n!}{(\ell)!(n-\ell)!} p^{\ell} (1-p)^{n-\ell}. \quad (2)$$

To obtain *confidence intervals* for p, one seeks a lower value, p_o, and an upper value, p_1, such that the probability of obtaining at most and at least k failures in n operations is $\alpha/2$ (i.e., half the confidence level).* Thus, to obtain k or fewer failures in n operations with a probability $\alpha/2$ p_1 is chosen such that

$$F(k|n,p_1) = \alpha/2 \ . \quad (3)$$

Similarly the minimum reasonable value of the failure probability at the α-level, is that value p_o, for which the probability of observing k or more failures in n tries is $\alpha/2$, i.e.,

$$1 - F(k-1|n,p_o) = \alpha/2 \ . \quad (4)$$

Although the confidence limits, p_o and p_1, could be obtained by numerical solution of Eqs. (3) and (4), the potentially large summations in these equations can be avoided by recognizing

*The confidence level refers to the total probability in both upper or lower tails. Half of the total confidence level is associated with each tail region.

$$F(k|n,p) = 1 - I_p(k+1,n-k), \tag{5}$$

where the incomplete beta function I_p is defined by

$$I_p(a,b) \equiv \frac{1}{B(a,b)} \int_0^p z^{a-1} (1-z)^{b-1} dz \tag{6}$$

with $B(a,b) \equiv \Gamma(a)\Gamma(b)/\Gamma(a+b)$ and Γ is the gamma function. With this relation between F and I_p, the equations which determine the upper and lower confidence limits on p may be written as

$$I_{p_o}(k,n-k+1) = \alpha/2 \tag{7}$$

and

$$I_{p_1}(k+1,n-k) = 1 - \frac{\alpha}{2} . \tag{8}$$

The advantage of this form, which still must be solved numerically for p_o and p_1, is that the corresponding probability limits for the Bayesian analogue are given by equations of the same functional form, and the same numerical algorithm used to solve the above equation can be used in the Bayesian analysis.

It is easily shown that $p_o \leq \hat{p} \equiv k/n \leq p_1$, with the equality defined* only if k=0 ($p_o=\hat{p}=0$) or k=n ($p_1=\hat{p}=1$). Of special interest are situations involving events with low probabilities of failure, for which one often encounters observed values of k=0 for relatively large values of n. For this case, the upper bound, p_1, can be obtained analytically. From Eq. (8) one obtains upon solving for p_1

$$p_1 = 1 - [\tfrac{\alpha}{2}]^{1/n}, \quad \text{for } k=0. \tag{9}$$

Similarly for high probability events for which one often observes k=n (and for which $\hat{p}=p_1=1$), Eq. (7) yields

$$p_o = (1 - \tfrac{\alpha}{2})^{1/n}, \quad \text{for } k=n. \tag{10}$$

---

*For k=0, the integrand on the left hand side of Eq. (7) becomes singular and the equation has no solution. In this case the entire confidence level is often associated with the "upper tail" of the distribution. However, to be consistent with the more general case (k≠0,n), we will always associate only half of the total confidence level with each end of the tail. A similar convention is used with the k=n case.

1.2 Review of the Bayesian Analysis

In the Bayesian description of the failure probability for a component, it is assumed that the failure probability comes from a particular *prior distribution* which is known either from previous experience or from the analysis of similar components [1]. In this document, it is assumed that the prior distribution is given by a beta distribution

$$g(p) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}, \quad (a,b>0). \tag{11}$$

If it is assumed, as was done in the classical case, the failure distribution is given by a bionomial distribution, then the use of Bayes' theorem gives for the *posterior* distribution [1]

$$\xi(p|k,n,a,b) = \frac{p^{a+k-1}(1-p)^{b+n-k-1}}{B(a+k,b+n-k)}. \tag{12}$$

This quantity (also a beta distribution), is the Bayesian estimate of the distribution of the failure probability, p, for a particular component which has previously experienced k failures in n tries and which is assumed to belong to a class of components whose failure probabilities are distributed according to Eq. (11).

With the posterior distribution, the *probability limits* are readily formulated for a component which has experienced k failures in n tries. Explicity the probability that the component failure probability is greater than some upper bound p_1 at the $\alpha/2$ level

$$\text{Prob}\ \{p>p_1\} = \frac{\alpha}{2} = \int_{p_1}^{1} \xi(p|k,n,a,b)dp. \tag{13}$$

Similarly the probability that the component failure probability, p, is less than some lower bound, p_o, at the $\alpha/2$ level is

$$\text{Prob}\ \{p<p_o\} = \frac{\alpha}{2} \int_{0}^{p_o} \xi(p|k,n,a,b)dp. \tag{14}$$

Upon substitution for ξ, the probability limits are readily expressed in terms of the incomplete beta function as

$$I_{p_o}(a+k,n+b-k) = \alpha/2 \tag{15}$$

and

$$I_{p_1}(a+k,n+b-k) = 1 - \alpha/2. \tag{16}$$

Again these equations have the same form as those defining the confidence interval in the classical case (Eqs. (7) and (8)), although with different arguments for the incomplete beta function.

1.3 Estimates of Component Failure Probability

Classical analysis estimates the probability of failure for component with k failures in n tries as

$$\hat{p} = \frac{k}{n} \tag{17}$$

The Bayesian approach uses as its estimate of the component failure probability the mean of the posterior distribution (Eq. (12)), namely

$$\hat{p} = \frac{a+k}{(a+k) + (b+n-k)} \ . \tag{18}$$

2. DESCRIPTION OF PROGRAM 'TAILS'

For a given α-level and component history (i.e., values for n and k), the code TAILS calculates (i) the upper and lower confidence limits on the component failures probability from Eqs. (7) and (8), and (ii) the upper and lower probability limits of the Bayesian posterior distribution for the component from Eqs. (15) and (16) with any specific beta prior distribution (defined by parameters a and b). The four equations to be solved, Eqs. (7), (8), (15) and (16), all are of the same form, and are readily solved for p_o or p_1 by numerical techniques involving methods of successive bisection and interpolation in the interval (0,1) [2]. To evaluate the incomplete beta function, a very accurate subroutine by N. Bosten and E. Battiste is used [3], and is briefly described in Appendix A.

A complete listing of the program is given in Appendix B.

2.1 Input Data

For each component to be analyzed, input data consists of the component performance history (n and k), the desired statistical level α, and, if the Bayesian probability limits are sought, the parameters of the assumed beta prior distribution (a and b). One input card is required for each component to be analyzed, and analysis continues until all data cards are processed.

For each component the data card contains the following information:

COMPONENT DATA CARD: Format (2I5,4G10.4,I5)

K = number of observed failures for component ($\equiv$k)
N = total number of operations in which K failures were observed ($\equiv$n)
AALPHA = confidence level or fraction of distribution in both the upper and lower tails ($\equiv\alpha$)
AA = parameter "a" of the assume beta prior distribution for the component. If no Bayesian analysis is desired then AA is set to 0.0.
BB = parameter "b" of the assumed beta prior distribution for the component. If no Bayesian analysis is desired then BB is set to 0.0.
EPS = accuracy parameter for iterative solution. Iterations stop when the magnitude of the difference between two successive values of p_1 or p_o is less than EPS
IPRINT = option variable for intermediate output. If IPRINT = 0 only final values for the confidence interval and probability limits are printed. If IPRINT = 1, results of the interative solution at each step are also printed.

2.2 Sample Output

In the numerical solution of the probability or confidence limits, iterative procedure is used. The output indicates an "error code" for each limit which indicates whether a successive result was obtained in the iterative solution procedure. Explicitly,

ERROR CODE = 0 successive solution
= 1 no solution found in 20 iterations
= 2 solution not in interval (0.1) - should never occur.

In Fig. 1 a sample output is shown for a component which has experienced five failures in 100 operations. The output is self-explanatory.

ACKNOWLEDGMENT

This code was developed with support from the U.S. Nuclear Regulatory Commission under Contract AT(49-24)-0339.

```
CALCULATION OF CONFIDENCE INTERVALS FOR THE TRUE FAILURE PROBABILITY P AT THE 0.500     LEVEL
PLANT DATA:      5 FAILURES IN  100 TRIES
ESTIMATED PARAMETERS OF THE PRIOR DISTRIBUTION: A= 1.000      B= 20.00
REQUESTED ACCURACY FOR P0 AND P1= 0.100E-04


CLASSICAL RESULT:      ESTIMATE OF FAILURE PROBABILITY P=   0.500000E-01

UPPER LIMIT P1=   0.733268E-01   (ERROR CODE= 0)

LOWER LIMIT P0=   0.337948E-01   (ERROR CODE= 0)


BAYESIAN RESULT:      ESTIMATE OF FAILURE PROBABILITY P=   0.495868E-01

UPPER LIMIT P1=   0.612309E-01   (ERROR CODE= 0)

LOWER LIMIT P0=   0.352770E-01   (ERROR CODE= 0)
```

Fig. 1. Sample Output from TAILS for component with k=5 and n=100 which is assumed to come from a class described by a beta prior with a=1 and b=20.

4. REFERENCES

1. J. K. Shultis and N. D. Eckhoff, "Selection of Beta Prior Distribtuion Parameters from Component Failure Data," to be published IEEE Transactions, July, 1978.

2. J. K. Shultis, D. Grosh and Y. Pan, "Calculation of Confidence Intervals for Component Failure Probabilities," Center for Energy Studies Report CES-42, Kansas State University, Manhattan, Kansas, March, 1977.

3. Scientific Subroutine Package (360-A-CM-03X) Version III Programmer's Manual, H20-0205, IBM (1963).

4. O. G. Ludwig, "Incomplete Beta Ratio," Comm. ACM, 6 (1963) 314; also see "Collected Algorithms from CACM," Algorithm 179 and modifications by N. E. Bosten and E. L. Battiste (1972), and by M. C. Pike and J. Soo Hoo (1975).

ADDENDUM A

Evaluation of the Incomplete Beta Functions

The incomplete beta function $I_p(x,y)$ is calculated from the following expression: [3]

$$I_p(x,y) = \frac{\text{INFSUM}\ p^x\ \Gamma(PS+x)}{\Gamma(PS)\ \Gamma(x+1)} + \frac{p^x\ (1-p)^y\ \Gamma(x+y)\ \text{FINSUM}}{\Gamma(x)\ \Gamma(y+1)}$$

where INFSUM and FINSUM represent two series summations defined as follows:

$$\text{INFSUM} = \sum_{j=1}^{\infty} \frac{x(1-PS)_j}{x+j} \frac{p^j}{j!}, \quad \text{where}$$

$$(1-PS) = \begin{cases} 1\ ,\ j = 0 \\ \Gamma(1+y-PS)/\Gamma(1-PS)\ ,\ j > 0 \end{cases}$$

and

$$\text{FINSUM} = \sum_{j=1}^{[y]} \frac{y(y-1)\ldots(y-j+1)}{(x+y-1)(x+y-2)\ldots(x+y-j)} \frac{1}{(1-p)^j}$$

where [y] is equal to the largest integer less than y. If [y]=0, the FINSUM=0. The quantity PS is defined as

$$PS = \begin{cases} 1 \quad \text{if } y \text{ is integer} \\ y - [y],\ \text{otherwise}\ . \end{cases}$$

The above algorithm (combined with scaling to avoid numerical inaccuracies encountered when using the gamma function with large arguments) was incorporated into a FORTRAN program MDBETA by Bosten and Battiste [5]. This program (modified in accordance to remarks made by Pike and Soo Hoo [5] was used in the present analysis. The program MDBETA is significantly more accurate than the widely used program BDTR [3], especially at large arguments. For example, in the case p=0.5, x=y=2000, MDBETA gives the correct value, 0.5, while BDTR gives 0.497026.

ADDENDUM B

Listing of the Program TAILS

```
FORTRAN IV G LEVEL  21                  MAIN              DATE = 78111          22/27/18

      C
      C**********************************  TAILS  ***************************************
      C*                                                                               *
      C* THIS PROGRAM CALCULATES (1) THE CONFIDENCE LIMITS ON THE CLASSICAL ESTIMATE *
      C* OF A COMPONENT FAILURE PROBABILITY,  AND (2) THE PROBABILITY INTERVALS OF    *
      C* THE BAYESIAN POSTERIOR FAILURE PROBABILITY DISTRIBUTION FOR THE SAME         *
      C* COMPONENT. ARBITRARY CONFIDENCE LEVELS (OR TAIL AREAS) MAY BE SPECIFIED.      *
      C*                                                                               *
      C*                                                                               *
      C*  INPUT DATA:   (ONE CARD FOR EACH COMPONENT ANALYSIS)   (2I5,4G10.4,I5)      *
      C*             K       = NUMBER OF OBSERVED FAILURES FOR COMPONENT              *
      C*             N       = TOTAL NUMBER OF TRIES IN WHICH K FAILURES WERE OBSERVED *
      C*             AALPHA  = CONFIDENCE LEVEL (DISTRIBUTION FRACTION IN BOTH TAILS) *
      C*             AA      = 'A' PARAMETER OF THE ASSUMED BETA PRIOR DISTRIBUTION   *
      C*                       ( = 0 IF NO BAYESIAN ANALYSIS IS DESIRED)              *
      C*             BB      = 'B' PARAMETER OF THE ASSUMED BETA PRIOR DISTRIBUTION   *
      C*                       ( = 0 IF NO BAYESIAN ANALYSIS IS DESIRED)              *
      C*             EPS     = REQUESTED ACCURACY FOR THE CONFIDENCE LIMITS           *
      C*             IPRINT  = 1 IF INTERMEDIATE OUTPUT IS DESIRED; = 0 IF ONLY FINAL *
      C*                       RESULT IS TO BE PRINTED                                *
      C*                                                                               *
      C*                                                                               *
      C*  WRITTEN BY J. K. SHULTIS, KANSAS STATE UNIVERSITY, MARCH 1977                *
      C*                                                                               *
      C*********************************************************************************
      C
      C
            COMMON IPRINT,A,B,ALPHA
            EXTERNAL FCT
      C
      C*** READ IN THE INPUT DATA
         99 READ(5,10,END=100) K,N,AALPHA,AA,BB,EPS,IPRINT
         10 FORMAT(2I5,4G10.4,I5)
            PRINT 11,AALPHA,K,N,AA,BB,EPS
         11 FORMAT('1CALCULATION OF CONFIDENCE INTERVALS FOR THE TRUE FAILURE
           1PROBABILITY P AT THE',G10.3,' LEVEL',
           2/' PLANT DATA:   ',I3,' FAILURES IN ',I4,' TRIES',
           3/' ESTIMATED PARAMETERS OF THE PRIOR DISTRIBUTION: A=',G10.4,
           4'  B=',G10.4,/' REQUESTED ACCURACY FOR P0 AND P1=',G10.3)
      C
      C*** CLASSICAL CALCULATIONS
            P=K/FLOAT(N)
            PRINT 12,P
         12 FORMAT('0',///'0CLASSICAL RESULT:      ESTIMATE OF FAILURE PROBABIL
           1ITY P=',G15.6,/)
            A=K+1.
            B=N-K
            ALPHA=1.0 - 0.5*AALPHA
            CALL RTMI(P1,F,FCT,0.0,1.0,0.0001,20,IER)
         13 FORMAT(' UPPER LIMIT P1=',G15.6,'   (ERROR CODE=',I2,')',/)
            PRINT 13,P1,IER
            IER=0
            P0=0.0
            IF (K.EQ.0) GO TO 15
            A=K
            B=N-K+1
            ALPHA=0.5*AALPHA
            CALL RTMI(P0,F,FCT,0.0,1.0,0.0001,20,IER)
```

1427 053

FORTRAN IV G LEVEL 21 MAIN DATE = 78111 22/27/18

```
          15 PRINT 16,PO,IER
          16 FORMAT(' LOWER LIMIT PO=',G15.6,'   (ERROR CODE=',I2,')',/)
      C
      C*** BAYESIAN ESTIMATES
             IF((AA+BB).EQ.0.0) GO TO 99
             P=(AA+K)/(AA+BB+N)
             PRINT 20,P
          20 FORMAT('0',///'OBAYESIAN RESULT:     ESTIMATE OF FAILURE PROBABILI
            1TY P=',G15.6,/)
             A=AA+K
             B=BB+N-K
             ALPHA=1.0-0.5*AALPHA
             CALL RTMI(P1,F,FCT,0.0,1.0,0.0001,20,IER)
             PRINT 13, P1,IER
             ALPHA=0.5*AALPHA
             CALL RTMI(PO,F,FCT,0.0,1.0,0.0001,20,IER)
             PRINT 16,PO,IER
             GO TO 99
         100 PRINT 30
          30 FORMAT('1')
             STOP
             END
```

FORTRAN IV G LEVEL 21 FCT DATE = 78111 22/27/18

```
             FUNCTION FCT(X)
      C*  THIS FUNCTION EVALUATES CONFIDENCE LIMIT EQUATION
             COMMON IPRINT,A,B,ALPHA
             IF((X.EQ.1.0).OR.(X.EQ.0.0)) GO TO 20
             CALL MDBETA(X,A,B,P,IER)
             FCT=P-ALPHA
             IF(IPRINT.EQ.1) PRINT 10,X,FCT,IER
          10 FORMAT(' X=',G12.5,'  I(X|A,B)-ALPHA=',G13.5,'    IER=',I3)
             RETURN
          20 FCT=X-ALPHA
             RETURN
             END
```

```
FORTRAN IV G LEVEL  21                  MOBETA              DATE = 78111          22/27/18

0001                 SUBROUTINE MOBETA(X, P, Q, PROB, IER)
               C
               C*************************************************************************
               C*                                                                       *
               C*  FUNCTION:      EVALUATE THE INCOMPLETE BETA DISTRIBUTION FUNCTION    *
               C*                                                                       *
               C*  PARAMETERS:                                                          *
               C*  X     - VALUE TO WHICH FUNCTION IS TO BE INTERGRATED. X MUST BE IN THE *
               C*          RANGE (0,1) INCLUSIVE.                                       *
               C*  P     - INPUT (1ST) PARAMETER (MUST BE GREATER THAN 0)               *
               C*  Q     - INPUT (2ND) PARAMETER (MUST BE GREATER THAN 0)               *
               C*  PROB - OUTPUT PROBABILITY THAT A RANDOM VAPIABLE FROM A BETA DISTRIBUTION *
               C*             HAVING PARAMETERS P AND Q WILL BE LESS THAN OR EQUAL TO O. *
               C*  IER  - ERROR PARAMETER.                                              *
               C*             IER = 0 INDICATES A NORMAL EXIT                           *
               C*             IER = 1 INCICATES THAT X IS NOT IN THE RANGE (0,1) INCLUSIVE *
               C*             IER = 2 INDICATES THAT P ANC/OR Q IS LESS THAN OR EQUAL TO 0. *
               C*                                                                       *
               C* CODE BASED ON SIMILAR CODE BY N. BOSTEN AND E.BATTISTE AS MODIFIED BY *
               C* M. PIKE AND J. HOO.                                                   *
               C*                                                                       *
               C*************************************************************************
0002                 DOUBLE PRECISION PS, PX, Y, P1, DP, INFSUM, CNT, WH, XB,
                    * DQ, C, EPS, EPS1, ALEPS, FINSUM, PQ, DA, DLGAMA
               C DOUBLE PRECISION FUNCTION DLGAMA
               C MACHINE PRECISION
0003                 DATA EPS/1.D-6/
               C SMALLEST POSITIVE NUMBER REPRESENTABLE
0004                 DATA EPS1/1.D-78/
               C NATURAL LOG CF EPS1
0005                 DATA ALEPS/-179.601600/
               C CHECK RANGES OF THE ARGUMENTS
0006                 Y = X
0007                 IF ((X.LE.1.0) .AND. (X.GE.0.0)) GO TO 10
0008                 IER = 1
0009                 GO TO 140
0010              10 IF ((P.GT.0.0) .AND. (Q.GT.0.0)) GO TO 20
0011                 IER = 2
0012                 GO TU 140
0013              20 IER =0
0014                 IF (X.GT.0.5) GO TO 30
0015                 INT = 0
0016                 GO TO 40
               C SWITCH ARGUMENTS FOR MORE EFFICIENT USE OF THE POWER
               C SERIES
0017              30 INT = 1
0018                 TEMP = P
0019                 P = Q
0020                 Q = TEMP
0021                 Y = 1.D0 - Y
0022              40 IF (X.NE.0. .AND. X.NE.1.) GO TO 60
               C SPECIAL CASE - X IS 0. OR 1.
0023              50 PROB = 0.
0024                 GO TO 130
0025              60 IB = Q
0026                 TEMP = IB
0027                 PS = Q - FLOAT(IB)
0028                 IF (Q.EQ.TEMP) PS = 1.D0
```

1427 055

```
          DP = P
          DQ = Q
          PX = DP*DLOG(Y)
          PQ = DLGAMA(DP+DQ)
          P1 = DLGAMA(DP)
          C = DLGAMA(DQ)
          D4 = DLOG(DP)
C DLGAMA IS A FUNCTION WHICH CALCULATES THE DOUBLE
C PRECISION LOG GAMMA FUNCTION
          XB = PX + DLGAMA(PS+DP) - DLGAMA(PS) - D4 - P1
C SCALING
          IB = XB/ALEPS
          INFSUM = 0.D0
C FIRST TERM OF A DECREASING SERIES WILL UNDERFLOW
          IF (IB.NE.0) GO TO 90
          INFSUM = DEXP(XB)
          CNT = INFSUM*DP
C CNT WILL EQUAL DEXP(TEMP)*(1.D0-PS)I*P*Y**I/FACTORIAL(I)
          WH = 0.0D0
     80 WH = WH + 1.D0
          CNT = CNT*(WH-PS)*Y/WH
          XB = CNT/(DP+WH)
          INFSUM = INFSUM + XB
          IF (XB/EPS.GT.INFSUM) GO TO 80
C DLGAMA IS A FUNCTION WHICH CALCULATES THE DOUBLE
C PRECISION LOG GAMMA FUNCTION
     90 FINSUM = 0.D0
          IF (DQ.LE.1.D0) GO TO 120
          XB = PX+ DQ*DLOG(1.D0-Y) + PQ - P1 - DLOG(DQ) - C
C SCALING
          IB = XB/ALEPS
          IF (IB.LT.0) IB = 0
          C = 1.D0/(1.D0-Y)
          CNT = DEXP(XB-FLOAT(IB)*ALEPS)
          PS = DQ
          WH = DQ
    100 WH =WH -1.D0
          IF (WH.LE.0.0D0) GO TO 120
          PX = (PS*C)/ (DP+WH)
          IF (PX.GT.1.0D0) GO TO 105
          IF (CNT/EPS.LE.FINSUM.OR.CNT.LE.EPS1/PX) GO TO 120
    105 CNT =CNT*PX
          IF (CNT.LE.1.D0) GO TO 110
C RESCALE
          IB = IB - 1
          CNT = CNT*EPS1
    110 PS =WH
          IF (IB.EQ.0) FINSUM = FINSUM + CNT
          GO TO 100
    120 PROB =FINSUM + INFSUM
    130 IF (INT.EQ.0) GO TO 140
          PROB = 1.0 - PROB
          TEMP = P
          P = Q
          Q = TEMP
    140 RETURN
          END
```

1427 056

```
FORTRAN IV G LEVEL  21                                  DATE = 78111          22/27/18
            C        SUBROUTINE RTMI
            C
            C        PURPOSE
            C           TO SOLVE GENERAL NONLINEAR EQUATIONS OF THE FORM FCT(X)=0
            C           BY MEANS OF MUELLER-S ITERATION METHOD.
            C
            C        USAGE
            C           CALL RTMI (X,F,FCT,XLI,XRI,EPS,IEND,IER)
            C           PARAMETER FCT REQUIRES AN EXTERNAL STATEMENT.
            C
            C        DESCRIPTION OF PARAMETERS
            C           X      - RESULTANT ROOT OF EQUATION FCT(X)=0.
            C           F      - RESULTANT FUNCTION VALUE AT ROOT X.
            C           FCT    - NAME OF THE EXTERNAL FUNCTION SUBPROGRAM USED.
            C           XLI    - INPUT VALUE WHICH SPECIFIES THE INITIAL LEFT BOUND
            C                    OF THE ROOT X.
            C           XRI    - INPUT VALUE WHICH SPECIFIES THE INITIAL RIGHT BOUND
            C                    OF THE ROOT X.
            C           EPS    - INPUT VALUE WHICH SPECIFIES THE UPPER BOUND OF THE
            C                    ERROR OF RESULT X.
            C           IEND   - MAXIMUM NUMBER OF ITERATION STEPS SPECIFIED.
            C           IER    - RESULTANT ERROR PARAMETER CODED AS FOLLOWS
            C                     IER=0 - NO ERROR,
            C                     IER=1 - NO CONVERGENCE AFTER IEND ITERATION STEPS
            C                             FOLLOWED BY IEND SUCCESSIVE STEPS OF
            C                             BISECTION,
            C                     IER=2 - BASIC ASSUMPTION FCT(XLI)*FCT(XRI) LESS
            C                             THAN OR EQUAL TO ZERO IS NOT SATISFIED.
            C
            C        REMARKS
            C           THE PROCEDURE ASSUMES THAT FUNCTION VALUES AT INITIAL
            C           BOUNDS XLI AND XRI HAVE NOT THE SAME SIGN. IF THIS BASIC
            C           ASSUMPTION IS NOT SATISFIED BY INPUT VALUES XLI AND XRI, THE
            C           PROCEDURE IS BYPASSED AND GIVES THE ERROR MESSAGE IER=2.
            C
            C        SUBROUTINES AND FUNCTION SUBPROGRAMS REQUIRED
            C           THE EXTERNAL FUNCTION SUBPROGRAM FCT(X) MUST BE FURNISHED
            C           BY THE USER.
            C
            C        METHOD
            C           SOLUTION OF EQUATION FCT(X)=0 IS DONE BY MEANS OF MUELLER-S
            C           ITERATION METHOD OF SUCCESSIVE BISECTIONS AND INVERSE
            C           PARABOLIC INTERPOLATION, WHICH STARTS AT THE INITIAL BOUNDS
            C           XLI AND XRI. CONVERGENCE IS QUADRATIC IF THE DERIVATIVE OF
            C           FCT(X) AT ROOT X IS NOT EQUAL TO ZERO. ONE ITERATION STEP
            C           REQUIRES TWO EVALUATIONS OF FCT(X). FOR TEST ON SATISFACTORY
            C           ACCURACY SEE FORMULAE (3,4) OF MATHEMATICAL DESCRIPTION.
            C           FOR REFERENCE, SEE G. K. KRISTIANSEN, ZERO OF ARBITRARY
            C           FUNCTION, BIT, VOL. 3 (1963), PP.205-206.
            C
            C     ..................................................................
            C
0001                 SUBROUTINE RTMI(X,F,FCT,XLI,XRI,EPS,IEND,IER)
            C
            C
            C        PREPARE ITERATION
0002                 IER=0
0003                 XL=XLI
```

```
0004              XR=XRI
0005              X=XL
0006              TOL=X
0007              F=FCT(TOL)
0008              IF(F)1,16,1
0009            1 FL=F
0010              X=XR
0011              TOL=X
0012              F=FCT(TOL)
0013              IF(F)2,16,2
0014            2 FR=F
0015              IF(SIGN(1.,FL)+SIGN(1.,FR))25,3,25
          C
          C       BASIC ASSUMPTION FL*FR LESS THAN 0 IS SATISFIED.
          C       GENERATE TOLERANCE FOR FUNCTION VALUES.
0016            3 I=0
0017              TOLF=100.*EPS
          C
          C
          C       START ITERATION LOOP
0018            4 I=I+1
          C
          C       START BISECTION LOOP
0019              DO 13 K=1,IEND
0020              X=.5*(XL+XR)
0021              TOL=X
0022              F=FCT(TOL)
0023              IF(F)5,16,5
0024            5 IF(SIGN(1.,F)+SIGN(1.,FR))7,6,7
          C
          C       INTERCHANGE XL AND XR IN ORDER TO GET THE SAME SIGN IN F AND FR
0025            6 TOL=XL
0026              XL=XR
0027              XR=TOL
0028              TOL=FL
0029              FL=FR
0030              FR=TOL
0031            7 TOL=F-FL
0032              A=F*TOL
0033              A=A+A
0034              IF(A-FR*(FR-FL))8,9,9
0035            8 IF(I-IEND)17,17,9
0036            9 XR=X
0037              FR=F
          C
          C       TEST ON SATISFACTORY ACCURACY IN BISECTION LOOP
0038              TOL=EPS
0039              A=ABS(XR)
0040              IF(A-1.)11,11,10
0041           10 TOL=TOL*A
0042           11 IF(ABS(XR-XL)-TOL)12,12,13
0043           12 IF(ABS(FR-FL)-TOLF)14,14,13
0044           13 CONTINUE
          C       END OF BISECTION LOOP
          C
          C       NO CONVERGENCE AFTER IEND ITERATION STEPS FOLLOWED BY IEND
          C       SUCCESSIVE STEPS OF BISECTION OR STEADILY INCREASING FUNCTION
          C       VALUES AT RIGHT BOUNDS. ERROR RETURN.
```

1427 058

```
FORTRAN IV G LEVEL  21                 RTMI                DATE = 78111          22/27/18

 0045              IER=1                                                           RTMI1200
 0046           14 IF(ABS(FR)-ABS(FL))16,16,15                                     RTMI1210
 0047           15 X=XL                                                            RTMI1220
 0048              F=FL                                                            RTMI1230
 0049           16 RETURN                                                          RTMI1240
             C                                                                     RTMI1250
             C     COMPUTATION OF ITERATED X-VALUE BY INVERSE PARABOLIC INTERPOLATIONRTMI1260
 0050           17 A=FR-F                                                          RTMI1270
 0051              DX=(X-XL)*FL*(1.+F*(A-TOL)/(A*(FR-FL)))/TOL                     RTMI1280
 0052              XM=X                                                            RTMI1290
 0053              FM=F                                                            RTMI1300
 0054              X=XL-DX                                                         RTMI1310
 0055              TOL=X                                                           RTMI1320
 0056              F=FCT(TOL)                                                      RTMI1330
 0057              IF(F)18,16,18                                                   RTMI1340
             C                                                                     RTMI1350
             C     TEST ON SATISFACTORY ACCURACY IN ITERATION LOOP                 RTMI1360
 0058           18 TOL=EPS                                                         RTMI1370
 0059              A=ABS(X)                                                        RTMI1380
 0060              IF(A-1.)20,20,19                                                RTMI1390
 0061           19 TOL=TOL*A                                                       RTMI1400
 0062           20 IF(ABS(DX)-TOL)21,21,22                                         RTMI1410
 0063           21 IF(ABS(F)-TOLF)16,16,22                                         RTMI1420
             C                                                                     RTMI1430
             C     PREPARATION OF NEXT BISECTION LOOP                              RTMI1440
 0064           22 IF(SIGN(1.,F)+SIGN(1.,FL))24,23,24                              RTMI1450
 0065           23 XR=X                                                            RTMI1460
 0066              FR=F                                                            RTMI1470
 0067              GO TO 4                                                         RTMI1480
 0068           24 XL=X                                                            RTMI1490
 0069              FL=F                                                            RTMI1500
 0070              XR=XM                                                           RTMI1510
 0071              FR=FM                                                           RTMI1520
 0072              GO TO 4                                                         RTMI1530
             C     END OF ITERATION LOOP                                           RTMI1540
             C                                                                     RTMI1550
             C                                                                     RTMI1560
             C     ERROR RETURN IN CASE OF WRONG INPUT DATA                        RTMI1570
 0073           25 IER=2                                                           RTMI1580
 0074              RETURN                                                          RTMI1590
 0075              END                                                             RTMI1600
```

NRC FORM 335 (7-77)

U.S. NUCLEAR REGULATORY COMMISSION

BIBLIOGRAPHIC DATA SHEET

1. REPORT NUMBER (Assigned by DDC): NUREG/CR-1110

2. (Leave blank)

3. RECIPIENT'S ACCESSION NO.

4. TITLE AND SUBTITLE (Add Volume No., if appropriate)

Bayesian Analysis of Component Failure Data

5. DATE REPORT COMPLETED

MONTH	YEAR
September	1979

7. AUTHOR(S)

J. K. Shultis and F. A. Tillman

DATE REPORT ISSUED

MONTH	YEAR
November	1979

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

Kansas State University
Department of Nuclear Engineering
Manhattan, Kansas 66506

6. (Leave blank)

8. (Leave blank)

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code)

U.S. Nuclear Regulatory Commission
Office of Nuclear Regulatory Research
Probabilistic Analysis Staff
Washington, DC 20555

10. PROJECT/TASK/WORK UNIT NO.

11. CONTRACT NO.

NRC-04-76-339

13. TYPE OF REPORT

Final - Technical

PERIOD COVERED (Inclusive dates)

15. SUPPLEMENTARY NOTES

14. (Leave blank)

16. ABSTRACT (200 words or less)

This report summarizes the many investigations made on the empirical Bayesian analysis of component failure data. In this study the analysis of attribute data of the failure-on-demand type was considered for components with low failure probabilities. Major areas emphasized in the study include (i) the development of computer techniques to obtain estimates of the prior distribution from observed failure data, (ii) the use of simulation studies to investigate the inherent properties of different prior parameter estimation techniques, (iii) the computation and comparison of probability and confidence intervals for the failure probability of individual components, and (iv) the use of non-beta prior distributions such as a mixture of beta distributions or a gamma distribution.

Four methods were examined for estimating parameters of the assumed prior beta distribution from failure data: (i) matching the moments of the prior distribution to those of the data, (ii) matching the moments of the marginal distribution to those of the data, (iii) the maximum likelihood method based on the prior distribution, and (iv) the maximum likelihood method based on the marginal distribution.

17. KEY WORDS AND DOCUMENT ANALYSIS

17a. DESCRIPTORS

Bayesian
Component Failure Probabilities
Prior Distribution
Beta or Gamma Distributions

1427 060

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT

19. SECURITY CLASS (This report)	21. NO. OF PAGES
Unclassified	212
20. SECURITY CLASS (This page)	**22. PRICE**
Unclassified	$

NRC FORM 335 (7-77)

UNITED STATES
NUCLEAR REGULATORY COMMISSION
WASHINGTON, D. C. 20555

OFFICIAL BUSINESS
PENALTY FOR PRIVATE USE, $300

POSTAGE AND FEES PAID
U.S. NUCLEAR REGULATORY
COMMISSION

U.S. MAIL

Document Control
O/C

POOR ORIGINAL

120555031837 2 ANRGRD
US NRC
SECY PUBLIC DOCUMENT ROOM
BRANCH CHIEF
HST LOBBY
WASHINGTON DC 20555

1427 061