AN ABSTRACT OF THE THESIS OF

JERALD ROBERT WILLE for the degree DOCTOR OF PHILOSOPHY
(Name) (Degree)

in STATISTICS presented on August 12, 1975
(Major Department) (Date)

Title: BAYES-FIDUCIAL INFERENCE FROM CENSORED NORMAL DATA Redacted for Privacy

Abstract approved: ______________________
Dr. Donald A. Pierce

A study is made of a Bayes-Fiducial procedure for normally distributed data which is censored by a Type II censoring scheme. The Bayes procedure uses a prior density of the form $\pi(\mu, \sigma) \propto 1/\sigma$, and Fiducial procedure utilizes the approach of conditioning on ancillary statistics.

Comparisons are made between the exact fiducial limits and approximate confidence limits derived both from linear estimates based on order statistics and from maximum likelihood estimates. The comparisons are based on the posterior content of the approximate intervals. It is also shown that the ancillary information is for practical purposes negligible.

The computational procedure used in determination of posterior intervals is examined. Numerical integration using the Guassian-Laguerre-quadrature technique is explored. The number of sampling points necessary for the numerical integration and a method for increasing accuracy of the procedure are both examined experimentally.

Bayes-Fiducial Inference from Censored Normal Data

by

Jerald Robert Wille

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1976

APPROVED:

Redacted for Privacy

Associate Professor of Statistics

in charge of major

Redacted for Privacy

Chairman of Department of Statistics

Redacted for Privacy

Dean of Graduate School

Date thesis is presented Aug. 12 1975

Typed by Clover Redfern for Jerald Robert Wille

Dedication

This work is dedicated to my wife, Carole, my family and particularly my grandparents

ACKNOWLEDGMENT

The author greatly appreciates the comments, suggestions and vast amount of time spent by his major professor, Dr. Donald A. Pierce. Through-out the years of graduate work Dr. Pierce has been a constant help and encouraging friend.

The computer work would have been impossible without the generous grants provided through the Computer Center of Oregon State University. Many others have offered encouragement and helpful comments, the author is grateful to each.

TABLE OF CONTENTS

Chapter		Page
I.	INTRODUCTION	1
	Statement of the Problem	1
	Methods of Estimation	3
II.	UNDERLYING THEORY AND CHOICE OF PRIOR DISTRIBUTION	11
	The Fiducial Method	12
	Relation of the Fiducial Distribution to the Bayes Posterior Distribution and Their Frequency Interpretation	17
III.	COMPARISONS WITH APPROXIMATE METHODS	25
	Solutions to the Likelihood Equations	25
	Best Linear Estimates	27
	Importance of Ancillary Information	31
	Visual Comparisons of Approximate Distributions of the Maximum Likelihood Estimate, Best Linear Estimate and Posterior Distribution	33
	Numerical Comparisons of Confidence Intervals	46
IV.	NUMERICAL METHODS	53
	Posterior Distributions	53
	The Numerical Integration	55
	Size of the Gaussian Quadrature	61
	The Computer Program	62
V.	SUMMARY AND CONCLUSIONS	66
	Other Possible Topics	67
	BIBLIOGRAPHY	70
	APPENDIX: Computer Program	73

LIST OF FIGURES

Figure		Page
1.	Triangular distribution.	14
2.	Plot of censor scheme (10, 7, 1) using Gupta's data.	30
3.	Plot of censor scheme (10, 8, 3) sample 1.	34
4.	Plot of censor scheme (10, 8, 3) sample 2.	35
5.	Plot of censor scheme (10, 7, 1) sample 1.	36
6.	Plot of censor scheme (10, 7, 1) sample 2.	37
7.	Plot of censor scheme (10, 5, 1) sample 1.	38
8.	Plot of censor scheme (10, 5, 1) sample 2.	39
9.	Plot of censor scheme (20, 15, 6) sample 1.	40
10.	Plot of censor scheme (20, 15, 6) sample 2.	41
11.	Plot of censor scheme (20, 20, 11) sample 1.	42
12.	Plot of censor scheme (20, 20, 11) sample 2.	43
13.	Plot of censor scheme (20, 20, 6) sample 1.	44
14.	Plot of censor scheme (20, 20, 6) sample 2.	45
15.	Plot of $e^{-\theta}f(\theta)$ and $(1/S)e^{-\theta/S}f(\theta/S)$.	59
16.	Quadratures of order 3, 5, 10, 15, 20 with censor scheme (20, 20, 6).	63
17.	Quadratures of order 3, 5, 10, 15, 20 with censor scheme (20, 15, 6).	64
18.	Quadratures of order 3, 5, 10, 15, 20 with censor scheme (20, 15, 1).	65

LIST OF TABLES

Table | Page

1. Days after innoculation. 29

2. Confidence limits for Gupta's data. 30

3. Between sample differences of posterior content. 32

4. Values of K for M.L.E. and B.L.E. 47

5. Approximate posterior content of confidence intervals calculated from Maximum Likelihood Estimates. 51

6. Approximate posterior content of confidence intervals calculated from Best Linear Estimates. 52

7. Differences in posterior content due to scale factor. 60

BAYES-FIDUCIAL INFERENCE FROM CENSORED NORMAL DATA

I. INTRODUCTION

Statement of the Problem

The subject of this thesis is a Bayesian approach to the analysis of censored data where parent populations are assumed to be Gaussian. Censoring of data may occur naturally or may be done intentionally, and is found in many experimental situations.

Censoring may be defined as the situation where measurements on certain experimental units are known only to be greater or less than a certain value. This is not to be confused with truncation, in which case the range of the random variable is restricted. The most common types of censoring are known as Type I censoring, Type II censoring, and progressive censoring.

Type I censoring occurs when the variate-values occur outside of some fixed range, i.e., censoring takes place at certain fixed points. Type II censoring occurs when a fixed proportion of the sample is censored at the lower and/or upper ends of a range determined by the order statistics. The distinction then between Type I and Type II censoring is that in the former case the number of observations is a random variable, while in the latter case the number of observations is fixed in advance. Combinations of the two are of course possible. Progressive censoring takes place when experimental units are added or taken from the experiment on an individual basis prior to the recording of the observation. Progressive censoring may be either Type I or Type II, or both. In order to further clarify the distinction between the types of censoring three examples will now be given.

Example I: Type I Censoring. Rats are fed a diet containing varying amounts of selenium containing compounds. Time until death of the rats is recorded.

The test is terminated arbitrarily at 730 days. By the nature of the test termination, the number of observations is a random variable and censoring scheme is Type I.

Example II: Type II Censoring. One thousand light bulbs are tested to determine their lifetime and to determine confidence limits on the population mean. Prior to the experiment it is decided to stop the test when 750 bulbs have failed. Observations are recorded until 750 fail, and the test is curtailed. The decision to stop after a fixed number of bulbs have failed determines this to be a Type II censoring pattern.

Example III: Progressive Censoring. Mice are examined for affects of eating food which has been given massive doses of radiation as a preservation measure. At specified times mice are removed from the test to be pathologically examined for tumors, cancerous tissue, etc. Additional mice may be added to the experiment at subsequent times. The test is then curtailed at a fixed point in time. Additions to and removals from the test determines this to be progressive Type I censoring.

An important feature of the Bayes method studied here is its applicability to all forms of censoring. Although the focus is on Type II censoring in this thesis, the same Bayesian procedures are directly applicable to Type I and progressive censoring. The likelihood function for a given set of data is the same whether the data arose from Type I or Type II censoring.

In order to better understand the problem and also to put the Bayesian method in perspective, the following section presents the Bayesian method and the two most common alternate methods.

Methods of Estimation

The two most common methods of estimation for this problem are the method of maximum likelihood (M.L.E.) and the method of best linear estimation based on the order statistics (B.L.E.). When data are censored the distributions of maximum likelihood estimates and of the linear estimates are intractable and must be approximated. The normal approximations provided by the standard asymptotic theory are used to set confidence limits or test hypotheses.

The Bayesian analysis provides a method for using all available information and providing limits with exact frequency properties, while not being bound by the distributional difficulties associated with the estimates. This is accomplished through the joint posterior distribution and the marginal posterior distributions of the parameters. Bayes posterior confidence limits may be found directly from the marginal posterior distribution. Certain posterior limits will be shown in Chapter II to have exact frequency interpretation in the case of Type II censoring.

The notation used here for Type II censored data is as follows:

1. The triple (N, K_1, K_2) denotes a given Type II censoring scheme. N is the sample size, K_2 is the index of the maximum observed order statistic, K_1 is the index of the minimum observed order statistic.
2. The subset of the order statistics which are observed (not censored) is written as $x_{K_1}, x_{K_1+1}, \ldots, x_{K_2}$, where the order is from smallest to largest. $\underline{x}$ is the vector notation of the sample data.

The method of maximum likelihood can be applied as follows. Assume that a sample is to be taken (or tested) from a parent population which is $N(\mu, \sigma^2)$, and that the censoring scheme is (N, K_2, K_1). The likelihood function may be written as

$$(1.1) \qquad L(\underline{x}/\mu,\sigma) \propto F(x_{K_1}/\mu,\sigma)^{K_1-1}\,(1-F(x_{K_2}/\mu,\sigma))^{N-K_2}\prod_{i=K_1}^{K_2} f(x_i/\mu,\sigma)$$

where $f(x/\mu,\sigma)$ is the normal density with mean μ and standard deviation σ and

$$F(x/\mu,\sigma) = \int_{-\infty}^{x} f(t/\mu,\sigma)dt$$

The maximum likelihood estimates may then be found by a Newton Raphson iteration technique described by Harter and Moore (1966). Approximate confidence limits can be computed from the usual asymptotic distribution theory. This is described in more detail in Chapter III.

Lloyd (1952) gives a method of estimating parameters of any location and scale parameter distribution through linear functions of the order statistics which may be applied directly to the problem of Type II censored data. These estimators are unbiased and have minimum variance in the class of linear (in the order statistics) unbiased estimates.

Let 0_i, $i = 1, 2, \ldots, n$, be the expected value of order statistics of a random sample of size n from a $N(0,1)$ distribution. We may then write

$$E(x_i) = \mu + 0_i\sigma$$

where x_i is the i-th order statistic from a $N(\mu, \sigma^2)$ distribution. For $\underline{x} = (x_{K_1}, \ldots, x_{K_2})$, and $\theta = (\mu, \sigma)$ this may be written in matrix form as;

$$E(\underline{x}) = A\theta$$

If the covariance matrix of the $\underline{x}$ is $V(\underline{x}) = V$, then by standard weighted least squares theorem

$$\underline{\theta}* = (A'V^{-1}A)^{-1}A'V^{-1}\underline{x} , \tag{1.2}$$

Where $\theta* = (\mu*, \sigma*)$ is the estimator of θ. The covariance matrix of $\theta*$ is then $(A'V^{-1}A)^{-1}$ and the approximate normal distribution of $\mu*$ can be used to specify approximate confidence limits. This is discussed in greater detail in Chapter III.

Bayesian inference regarding μ and σ can be carried out as follows. Let $\pi(\mu, \sigma)$ represent the density of a prior distribution representing the state of our knowledge of the parameters μ and σ. If we consider taking a random sample from a distribution with density $f(x/\mu, \sigma)$, where μ and σ are fixed but unknown and the function f is known; the likelihood function is then given by Equation (1.1) and is denoted by $L(\underline{x}/\mu, \sigma)$. The density of the joint posterior distribution is given by

$$\pi(\mu, \sigma/\underline{x}) \propto \pi(\mu, \sigma)L(\underline{x}/\mu, \sigma). \tag{1.3}$$

This posterior distribution represents how one's subjective beliefs should be modified from the prior beliefs $\pi(\mu, \sigma)$ by the sample data $\underline{x}$.

Inferences may be made and Bayes confidence regions found from marginal posterior distributions of the parameters of interest. That is, inferences about μ would be generated from

$$g(\mu/\underline{x}) \propto \int_0^\infty \pi(\mu, \sigma/\underline{x})d\sigma$$

and inferences about σ from

$$h(\sigma/\underline{x}) \propto \int_{-\infty}^\infty \pi(\mu, \sigma/\underline{x})d\mu.$$

The methods used in this thesis are for applications where the prior information is negligible in comparison to the sample information, or for the case where one wishes to make such inferences even when the prior information might be substantial. For Type II censoring, the methods developed here have exact relative frequency, as well as Bayesian interpretation.

Bayesian confidence intervals will have the standard relative frequency interpretation and will be optimal in a certain sense. Optimality will be developed and expanded further in Chapter II where it will be shown that the Bayesian methods are unique in that (i) they use all the sample information and (ii) are invariant under location and scale parameter transformations.

One of the first papers which described methods for estimating the parameters of a normal distribution when the data are censored was by Gupta (1952). In this paper Gupta develops a method for calculating the maximum likelihood estimates using successive substitution, i.e., for an initial estimate for μ, say μ_0, find an estimate for σ, say σ_1. Use the value σ_1 to then calculate μ_1 and so on until the differences between the successive estimates and suitably small. Gupta's method is quite awkward in that a separate table of constants is needed in order to solve the equations.

Because of the bias he noted in the maximum likelihood estimates, Gupta developed what he called an alternative linear estimate for use with samples larger than size ten. Gupta obtains coefficients by solving the Equation (1.2), but with $V = I$.

$$\theta^* = (A'A)^{-1}A'\underline{x}$$

He shows that this estimate is very good when compared to the best linear estimate of (1.2).

Cohen (1959) gives tables for calculating the maximum likelihood estimates of Type II censored data, for one-sided censoring only. Values of λ are given so that estimates for the mean μ^* and variance σ^{*2} may be calculated as; $\mu^* = \bar{x} - \lambda(\bar{x}-x_K)$ and $\sigma^{*2} = s^2 + \lambda(\bar{x}-x_K)^2$, where

$$s^2 = \sum_{i=1}^{K} (x_i - \bar{x})^2 / K, \quad \bar{x} = \sum_{i=1}^{K} (x_i)/K$$

and x_K is the maximum order statistic. The advantage of this estimate over Gupta's is that only one statistic, λ, is used in calculating μ^* and σ^*.

J.G. Saw (1959) proposed for the Type II singly censored problem, estimators which he calls "symmetrical" in the order statistics. These are $\tilde{\mu} = \epsilon \bar{x}_{K-1} + (1-\epsilon)x_K$ and

$$\tilde{\sigma}^2 = \alpha \sum_{i=1}^{K} (x_i - x_K)^2 + \beta \left(\sum_{i=1}^{K} (x_i - x_K)\right)^2$$

where

$$\bar{x}_{K-1} = \sum_{i=1}^{K} (1/(K-1))x_i \ .$$

ϵ is chosen to make μ unbiased and α and β to make σ^2 unbiased and of minimum variance. These estimates are good in that they allow one to estimate parameters from large samples $(N > 20)$ but they require extensive tables.

Harter and Moore (1966) solved the maximum likelihood equations by iteration and developed asymptotic variances of $\hat{\mu}$ and $\hat{\sigma}$. Approximate means, variances and covariances were also obtained for $N \geq 20$.

Other authors studying similar problems were Dixon (1961) who looked at symmetrical censoring; Rao, Savage and Sobel (1961) who studied the two-sample problem from a rank order statistic viewpoint; and Tiku (1967) who used the relation $f(x)/F(x) \doteq \alpha + \beta x$ over an interval $c \leq x \leq d$ to approximate the M. L. E.

Except for the Bayesian method each of these share a major drawback; that is that an approximate distribution for $\hat{\mu}$ must be used in order to set confidence intervals. In each of these papers methods for approximating maximum likelihood estimates or linear estimates based on the order statistics were given. With an electronic computer, the M. L. E. and B. L. E. are readily calculable, thus it will be adequate to use only these estimates for comparative purposes in this thesis.

In Chapter II theoretical properties of certain methods for certain Bayesian methods for censored data will be explored. The results of Fraser (1961) and Hora and Buehler (1965) will be applied to establish the frequency properties of the Bayesian results and the use of a "flat" prior will be discussed.

Chapter III involves comparisons of the Bayesian method with the approximate M. L. E. and B. L. E. methods.

Chapter IV is devoted to methods used in computing the Bayesian results. Certain key results such as numerical integration procedures and schemes which insure the accuracy of these procedures will be investigated.

II. UNDERLYING THEORY AND CHOICE OF PRIOR DISTRIBUTION

For a random sample from any location-scale parameter probability function, Bayes intervals on μ, σ, and certain functions thereof are confidence intervals in the relative frequency sense when the prior density of μ and $\log \sigma$ is taken to be Lebesque measure on the plane. When sufficient estimators exist these are the usual results of pivotal arguments; when sufficient estimators do not exist, these are what Fisher (1934) introduced as fiducial intervals. By sufficient estimators is meant a pair of estimators $\hat{\mu}$ and $\hat{\sigma}$ which are jointly sufficient for μ and σ. Discussion is given below of the precise frequency behavior of such procedures, their Bayesian interpretation, and their relationship to fiducial and structural inferences.

Various authors have considered the frequency interpretation of the Bayes-Fiducial method of analysis. Welch and Peers (1963), for example, generalize an earlier result of Lindley (1958) and show that for a single parameter problem the Bayesian confidence intervals will have the standard frequency interpretation if and only if it is possible to write the density $f(\underline{x}, \theta)$ in the form $f(y-\tau)$, where y and τ are monotonic functions of $\underline{x}$ and θ respectively, $-\infty < \tau, y < \infty$, and the prior density of τ is uniform over the real line.

In the two parameter problem, it is well known that when a non-censored sample of size N is drawn from a $N(\mu, \sigma^2)$ distribution, and the prior density is taken as the improper prior $\pi(\mu, \sigma) \propto 1/\sigma$ (i.e., μ and $\log \sigma$ have a uniform density) that the posterior marginal density of μ is proportional to a t distribution with $N-1$ degrees of freedom (see Lindley, 1965, p. 36). Further, Hora and Buehler (1966) show for a generalized class of problems, which include location-scale problems, that certain Bayesian α level confidence intervals, for certain scalar functions of the parameters have probability α of coverage. This concept is explained in more detail below. For the two parameter negative exponential distribution Pierce (1973) has discussed Bayes- fiducial intervals calculated for invariantly estimable functions. Bogdanoff and Pierce (1972) also show that for the two parameter Weibull distribution with Type II censoring Bayesian inferences have exact frequency interpretation and coincide with fiducial methods when a similar "flat" prior distribution is used.

The Fiducial Method

R.A. Fisher (1930) first described the fiducial method as a tool of statistical analysis which could be used in the absence of prior information. When sufficient estimates μ and σ exist one can use pivotal arguments. One pivotal based on the sufficient statistics of a location-scale parameter distribution is $(\hat{\mu}-\mu/\hat{\sigma}, \hat{\sigma}/\sigma)$. This pivotal quantity gives rise to not only confidence regions but a unique fiducial distribution for μ and σ in which $\hat{\mu}$ and $\hat{\sigma}$ are regarded as fixed. The development of the fiducial distribution and fiducial intervals is shown below.

It may happen that an estimator $(\hat{\mu}, \hat{\sigma})$ which is a one-to-one function of the minimal sufficient statistic, does not exist. In this case the dimension of $\underline{S}$, the sufficient statistic is greater than the dimension of the parameter $\underline{\theta}$. For location-scale problems we shall see that we can then write $S = (T, \underline{A})$, where $\underline{A}$ has a marginal distribution not depending on θ. $\underline{A}$ is said to be an ancillary statistic. One could think of T as being sufficient conditional on $\underline{A} = \underline{a}$. The approach used by Fisher in such situations is to condition on $\underline{A}$, meaning that if we have data $\underline{X}$ with $\underline{A} = \underline{a}$, rather than basing frequency type inferences on the sequence of all repetitions of the experiment, base it on the sequence of trials for which $\underline{A} = \underline{a}$. That is to say, use the distribution $\{(f_\theta(x|\underline{A})\ \theta \in \Omega)\}$ rather than the original family $(f_\theta(x)\ \theta \in \Omega)$.

Below are several examples which demonstrate the principle of conditioning on ancillary statistics.

<u>Example 2.1</u>. Draw a sample of size $N = 2$ from a symmetric triangular distribution where:

$$
\begin{aligned}
f_\theta(x) &= x - \theta + 1 & \theta-1 \le x < \theta \\
&= \theta + 1 - x & \theta \le x \le \theta+1 \\
&= \theta & \text{otherwise}
\end{aligned}
$$

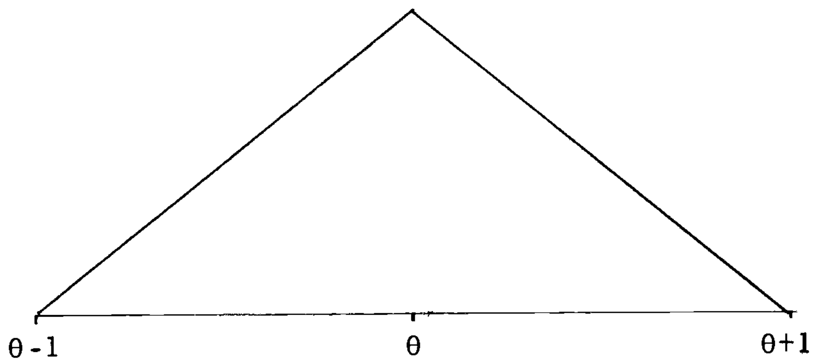

Figure 1. Triangular distribution.

The minimal sufficient statistic in this case is the sample (x_1, x_2). The maximum likelihood estimate $\theta = \bar{x}$, is not a sufficient estimate, but if we condition on the ancillary statistic $\underline{A} = |x_2 - x_1|$, then θ is conditionally sufficient. $\underline{A}$ by itself contains no information about θ, it does however tell something about how much $\hat{\theta}$ is saying about θ. The closer A is to two, the more precise the information in (x_1, x_2). In particular if $A = 2 - e$, one could determine θ to be within an interval of width e.

<u>Example 2.2.</u> Suppose that a random variable Y is equally likely to be $N(\mu, \sigma_1^2)$ or $N(\mu, \sigma_2^2)$, where σ_1 and σ_2 are different and known. An indicator variable A is observed, taking the value 1 or 2 according to whether y has the first or

second distribution. Thus it is known from which distribution y comes. Then the likelihood of the data (y, A) is

$$f_{c,y}(y, A) = (2\pi\sigma_A^2)^{-1/2} \exp\{(-(y-\mu)^2/(2\sigma_A^2)\}$$

so that $s = (y, A)$ is sufficient for μ with σ_1^2 and σ_2^2 known. Clearly one would want to use the conditional distribution of y given A for purpose of inference from observed data.

Example 2.3. In a linear regression problem, suppose that the values of the explanatory variable have a known joint p.d.f. $f_x(x)$ and that conditional on $X = x$ the $Y_1, Y_2, \ldots, Y_N$ are independent, Y_j having the distribution $N(\gamma+\beta x, \sigma^2)$. The full likelihood of the data is

$$f_{xy}(xy) = f_x(x)(2\pi\sigma^2)^{-(1/2)n} \exp - \frac{1}{2\sigma^2} \sum_{i=1}^{N} (Y_j - \gamma - \beta x_j)^2$$

the sufficient statistic for $(\gamma, \beta, \sigma^2)$ is $S = (\gamma, \beta, SS_{res}, \sum x_j \sum x_j)^2$, the last two components of which form an ancillary statistic. It is quite conventional to make inferences in this situation conditional on x. That is even if the explanatory variable is random, conditioning on the ancillary statistic would lead to treating the explanatory variable as fixed.

Fisher (1958) describes a sufficient statistic as "an exhaustive estimate without ancillary statistics," which we call here a sufficient estimator. His recommendation was that in the absence of a sufficient estimator, estimates should be conditional on ancillary statistics. In the case where data are normally distributed but are censored by a Type II scheme, a sufficient estimator does not exist. Estimates such as the maximum likelihood or best linear are not sufficient statistics and are in a sense incomplete. They should be conditioned on ancillary statistics to give the correct relative frequency interpretation.

In general the choice of ancillary statistics is a very difficult problem. In problems of location and scale, however, it is quite clear that one should take A to be some version of the maximal invariant statistic; that is take T as any statistic such that $(T, \underline{A})$ is a 1-1 function of the sample. The maximal invariant A for uncensored location-scale data can be taken as

$$\underline{A} = \left(\frac{x_3 - \hat{\mu}}{\hat{\sigma}}, \ldots, \frac{x_N - \hat{\mu}}{\hat{\sigma}} \right)$$

where μ, σ are any equivariant estimators. For the case of censored data this may be taken as

$$\left(\frac{x_{K_1} - \hat{\mu}}{\hat{\sigma}}, \ldots, \frac{x_{K_2} - \hat{\mu}}{\hat{\sigma}} \right)$$

which is our choice for $\underline{A}$.

Fraser (1961) shows that any "pivotal quantity" based on sufficient statistics (or sufficient conditioned on the ancillary statistic), which yields regions invariant under change of location and scale of data will lead to the same confidence regions. This invariance or equivariance simply means that relocation and rescaling the data will relocate and rescale the confidence regions in the appropriate manner.

In summary, then, if one adopts the principle of conditioning on ancillary statistics as a device to cope with non-existence of sufficient estimators, the scheme of inference given here is the only conclusion to be reached. The question of which ancillary statistics to use is not a problem in the location-scale situation--use the maximal invariant. The choice of T and the form of the pivotal quantity is not a difficulty because Fraser has shown that any choice which yields equivariant confidence regions must lead to the same procedure.

Relation of the Fiducial Distribution to the Bayes Posterior Distribution and Their Frequency Interpretation

The relation of these fiducial methods to Bayesian methods and the frequency interpretation that the results share can be studied through the results of Hora and Buehler (1966). The concepts developed there can be seen more directly in the simpler setting of the location-scale parameter problem.

They point out that for any location-scale problem, fiducial distributions as defined by Fraser and developed below, are exactly equivalent to Bayes posterior distributions using the prior distribution,

$$\pi(\mu, \sigma) \propto \frac{1}{\sigma} . \tag{2.1}$$

In the general location-scale setting define $\alpha(C_1, C_2, \underline{A})$ such that

$$P\left(\frac{\hat{\mu}-\mu}{\hat{\sigma}} > C_1, \frac{\hat{\sigma}}{\sigma} > C_2 \,\middle|\, \underline{A}\right) = \alpha(C_1, C_2, \underline{A}) \tag{2.2}$$

and where A is the maximal invariant statistic. Now define the fiducial distribution of $(\mu, \sigma \mid \text{data})$ as:

$$P_{fid}\left(\mu < \hat{\mu} - C_1\sigma, \sigma < \frac{\hat{\sigma}}{C_2} \,\middle|\, \text{data}\right) = \alpha(C_1, C_2, \underline{A}) \tag{2.3}$$

that is,

$$P_{fid}(\mu < \mu_0, \sigma < \sigma_0 \mid \text{data}) = \alpha\left(\frac{\hat{\mu}-\mu}{\hat{\sigma}}, \frac{\hat{\sigma}}{\sigma}, \underline{A}\right) \tag{2.4}$$

This defines for any μ_0 and σ_0 a joint probability distribution for μ and σ given the sample data. The density of this distribution will be denoted by $\pi_{fid}(\mu, \sigma \mid \underline{A})$. The fiducial distribution is then simply a representation of all confidence regions of the form (2.4). These are the usual type of confidence regions based on pivotal quantities. The fiducial density will then be:

$$(2.5) \qquad \pi_{fid}(\mu, \sigma | \underline{A}) = \frac{\partial^2 P_{fid}}{\partial\mu\partial\sigma}[\mu < \mu_0, \ \sigma < \sigma_0]$$

$$= \frac{\partial^2 a}{\partial\mu\partial\sigma}\left[\frac{\hat{\mu}-\mu}{\hat{\sigma}}, \ \frac{\hat{\sigma}}{\sigma} | \underline{A}\right]$$

$$= \frac{\partial^2 a}{\partial C_1 \partial C_2}(C_1, C_2, \underline{A}) \begin{vmatrix} \frac{dC_1}{d\mu} & \frac{dC_2}{d\sigma} \end{vmatrix}$$

Thus

$$= \frac{\partial^2 a}{\partial C_1 \partial C_2}(C_1, C_2 | \underline{A})(-\frac{1}{\sigma})(-\frac{\sigma}{\sigma^2})$$

But $\frac{\partial^2 a}{\partial C_1 \partial C_2}(C_1, C_2 | \underline{A})$ is the density of $(z_1 = \frac{\mu-\mu}{\sigma}, \ z_2 = \frac{\sigma}{\sigma})$ by definition. Thus

$$(2.6) \qquad \pi_{fid}(\mu, \sigma | \text{data}) \underset{\mu, \sigma}{\propto} \frac{g(z_1, z_2 | A)}{\sigma^2}$$

Now the likelihood $f(\underline{x} | \mu, \sigma)$ is proportional to the density $f^*(\hat{\mu}, \hat{\sigma}, A)$. But the joint density of $z_1 z_2$ and $\underline{A}$ is

$$g(z_1, z_2, A) = f^*_{\mu, \sigma}(\hat{\mu}, \hat{\sigma}, \underline{A}) \left| \frac{d(\mu, \sigma)}{d(z_1, z_2)} \right|$$

$$\propto f^*_{\mu, \sigma}(\hat{\mu}, \hat{\sigma}, \underline{A})\sigma$$

$$\propto f_{\mu, \sigma}(\underline{x})\sigma$$

which incidently is

$$(\frac{1}{\sigma})^{N-1} \prod_{i=1}^{N} f(\frac{x_i - \mu}{\sigma})$$

in the uncensored case. Thus

$$\frac{g(z_1 z_2 | A)}{\sigma^2} \propto \frac{1}{\sigma} f(\underline{x} | \mu, \sigma) \tag{2.7}$$

This then shows that the fiducial distribution is exactly equivalent to the Bayes posterior with prior distribution given by (2.1).

More specifically, let $\hat{\mu}$ and $\hat{\sigma}$ be the maximum likelihood estimates calculated from data which are censored via a Type II censoring scheme and are drawn from a normal population. Let the pivotal quantity be defined as before, i.e., $(\hat{\mu}-\mu/\hat{\sigma}, \hat{\sigma}/\sigma)$ and define

$$\underline{A} = \left(\frac{x_{K_1} - \hat{\mu}}{\hat{\sigma}}, \ldots, \frac{x_{K_2} - \hat{\mu}}{\hat{\sigma}}\right)$$

to be an ancillary statistic. Inferences about the parameters and certain functions of the parameters may be made from the distribution of the pivotal quantity conditioned on the ancillary statistic, that is

$$g_{fid}\left(\frac{\hat{\mu}-\mu}{\hat{\sigma}}, \frac{\hat{\sigma}}{\sigma} \middle| \underline{A}\right) .$$

Thus for example, a means of determining an upper confidence limit for μ from censored normal data, which will be the same as one based on the sampling distribution of $\hat{\mu}-\mu/\hat{\sigma}$ given $\underline{A}$ is as follows. Compute the posterior distribution

$$\pi(\mu, \sigma | \text{data}) \propto \left(\frac{1}{\sigma}\right) f(\underline{x} | \mu, \sigma) ,$$

then compute

$$\pi(\mu | \text{data}) \propto \int \pi(\mu, \sigma | \text{data}) d\sigma \ .$$

Take the upper confidence limit for μ to be a quantile, corresponding to the level of confidence of this posterior distribution.

In many cases one may wish to make inferences about some other function $\phi(\mu, \sigma)$ of the parameters. The fiducial density (2.6) or the equivalent posterior distribution (1.3) suggest that a marginal distribution of the form

$$(2.8) \qquad g_1(\phi | \hat{\mu}, \hat{\sigma}, \underline{\mathbf{A}}) = \iint_{\phi} g(\mu, \sigma | \hat{\mu}, \hat{\sigma}, \underline{\mathbf{A}}) d(\mu, \sigma)$$

be used for setting confidence intervals and making inferences.

Quantiles of this distribution might then be used for setting confidence limits of ϕ.

$$(2.9) \qquad P_{fid}(\overline{\phi}(\mu, \sigma) \leq \phi_\gamma(\mu, \sigma) | \hat{\mu}, \hat{\sigma}, \underline{\mathbf{A}}) = \gamma$$

Such confidence limits, although uniquely defined for invariant pivotal quantities need not have the usual frequency interpretation.

Hora and Buehler (1966) however, define a wide class of functions for which these procedures will have the usual frequency interpretation. If $\phi(\mu_1, \sigma_1) = \phi(\mu_2, \sigma_2)$ implies that

$\phi(a\mu_1+b, a\sigma_1) = \phi(a\mu_2+b, a\sigma_2)$ for all $a > 0$ and b in $(-\infty, \infty)$, then ϕ is said to be invariantly estimable. Hora and Buehler (1966) prove a theorem which states that:

i) If $\phi(\mu, \sigma)$ is invariantly estimable,

ii) For any $a > 0$ and b, $\phi(a\mu+b, b\sigma)$ increases and

iii) (2.9) has a unique solution for $\overline{\phi}(\mu, \sigma)$, then $\phi(\mu, \sigma)$ is a fiducial limit and $P[(\overline{\phi}(\hat{\mu}, \hat{\sigma}) > \phi(\mu, \sigma) | \mu, \sigma] = \gamma$, i.e., it has the confidence interval property.

Some common functions satisfying this criteria are μ, σ, $a\mu+b\sigma$, μ^k. One of importance which does not satisfy this is $\mu + \sigma^2/2$, the mean of the lognormal variate $\exp(x)$.

The confidence interval property means that $P(\overline{\phi}(\mu, \sigma) > \phi(\mu, \sigma) | \mu, \sigma) = \gamma$ has the usual frequency interpretation associated with confidence regions.

The methods developed in this chapter as fiducial methods are shown in a previous section to be exactly Bayesian methods for the prior distribution $\pi(\mu, \sigma) \propto 1/\sigma$ $\sigma > 0$, $-\infty < \mu < \infty$. It is interesting at this point to note the Bayesian interpretation of the use of this type of prior distribution.

Suppose that we have data such that the likelihood function is essentially non-zero only in a region of the parameter space within which our subjective prior belief is quite uniform. That is our prior information is negligible in relation to the sample information, or we wish to suppose that is so. In such a case the actual subjective prior will be approximately the same as if we had taken a prior nearly uniform over all the parameter space. As the parameters of interest (μ, σ) have infinite range, the uniform distribution cannot be defined in the usual way; that is there is no $\pi(\mu, \sigma) = c$ such that

$$\int_{-\infty}^{\infty} c\,d(\mu, \sigma) = 1.$$

The prior must be defined instead as a conditional density, such that if F is some set of (μ, σ) of finite length or measure, then the distribution, conditional on (μ, σ) belonging to F, has density $\pi(\theta | F) = m(F)^{-1}$ where $m(F)$ is the Lebesque measure of F so that

$$\int_F \pi(\mu, \sigma | F) d(\mu, \sigma) = m(F)^{-1} \int d(\mu, \sigma) = 1 .$$

This principle of using the uniform distribution as subjective prior knowledge has been called by Savage (1962), "The principle of precise measurement." The important concept here is that we are supposing that the density function $\pi(\mu, \sigma) = c$ represents our prior knowledge only over the range of appreciable likelihood and that it suitably tails off to zero outside that range. We are thus relieved of a theoretical difficulty by assuming that the prior actually used is proper.

To summarize, one may obtain the Bayes posterior distribution of the parameters of interest. This is equivalent to the fiducial distribution of the pivotal quantity using sufficient statistics or statistics conditioned on ancillary statistics. This fiducial distribution is simply a representation of all confidence regions of the form $(\mu < \mu_0,\ \sigma \leq \sigma_0 | A)$. These fiducial distributions are unique when based on invariant pivotal quantities. Inferences about functions of the parameters $\phi(\mu, \sigma)$ may be made in the relative frequency sense if $\phi(\mu, \sigma)$ is invariantly estimable. Confidence statements about invariantly estimable functions likewise have the standard frequency interpretation.

If one adopts this principle of conditioning on ancillary statistics as a device to cope with non-existence of sufficient estimators, the scheme of inference given here is the only conclusion to be reached. The question of which ancillary statistics to use is not a problem in the location-scale situation; use the maximal invariant. The choice of T and the form of the pivotal quantity is not a difficulty because Fraser has shown that any choice which yields equivariant confidence regions must lead to the same procedure.

III. COMPARISONS WITH APPROXIMATE METHODS

The most frequently used estimators for the parameters of the normal distribution when data are censored are the maximum likelihood estimators (M.L.E.) and the linear estimators based on the order statistics (B.L.E.). In this chapter comparisons will be made between the approximate confidence intervals derived from these estimators and the Bayesian confidence intervals which were previously shown to have exact frequency interpretation.

Solutions to the Likelihood Equations

Solutions to the likelihood equations may be found by the Newton-Raphson iteration technique as shown by Harter and Moore (1966).

That is, for the likelihood function (1.1);

$$L(\underline{x}/\mu,\sigma) \propto \prod_{i=K_1}^{K_2} f(x_i)(F(x_{K_1}))^{K_1}(1-F(x_{K_1}))^{N-K_2})$$

where

$$f(x) = 1/(\sqrt{2\pi})\exp - \frac{1}{2}(x-\mu)^2/\sigma^2$$

and

$$F(x) = \int_{-\infty}^{x} f(t)dt ,$$

and where the sample $\underline{x}$ is censored by a Type II censoring scheme

(N, K_2, K_1). The estimates of the parameters may be found iteratively using the relation;

$$\theta_i = \theta_{i-1} - [\partial^2 L/\partial\theta^2]^{-1}[\partial L/\partial\theta]_{\theta=\theta_{i-1}}$$

where

$$\theta = \begin{bmatrix} \mu \\ \sigma \end{bmatrix}$$

The initial estimates for μ and σ, i.e., μ_0 and σ_0, need only be of the same order of magnitude as the final estimates to keep the number of iterations to a reasonable number. The initial estimates used here are; the medium for μ; and 1/3 the sample range for σ. If censoring is extreme, say 50 percent or more, the initial estimates might well be modified. The preliminary estimate for σ, for example, may not be the most appropriate estimate if censoring is extreme. In some cases it is possible for σ_i to become negative during iteration. This may be avoided by insertion of a conditioning step in the computer program to set $\hat{\sigma}_i = \hat{\sigma}_{i-1}/2$ if $\hat{\sigma}_i$ becomes negative.

The approximate covariance matrix of the maximum likelihood estimates is found using the asymptotic approximation

$$V(\theta) = \begin{bmatrix} -\partial^2 \ln L/\partial\mu^2 & -\partial^2 \ln L/\partial\mu\partial\sigma \\ -\partial^2 \ln L/\partial\sigma\partial\mu & -\partial^2 \ln L/\partial\sigma^2 \end{bmatrix}^{-1}$$

Best Linear Estimates

Linear estimates based on the order statistics may be found using the method given by Lloyd (1952) and Sarhan and Greenberg (1956). Although this method of estimation was developed originally for full samples, it is immediately applicable to Type II censored observations. All that needs to be done is to delete from the matrix of expected values and from the covariance matrix those rows and columns relating to the missing observations.

Let 0_i be the expected value of the i-th order statistic of a sample from a $N(0,1)$ distribution. Then for a sample from $N(\mu, \sigma^2)$ the expected value of the i-th order statistic is $E(x_i) = \mu + \sigma 0_i$. This may be written in matrix form as $E(\underline{x}) = T\theta$ where

$$T = \begin{bmatrix} 1 & 0_{K_1} \\ 1 & \cdot \\ \vdots & \vdots \\ 1 & 0_{K_2} \end{bmatrix} \quad \text{and} \quad \theta = \begin{bmatrix} \mu \\ \sigma \end{bmatrix}$$

Denoting by V the covariance matrix of order statistics from $N(0,1)$ then the covariance matrix of the x_i is $\sigma^2 V$. Then the standard weighted least squares solution $\hat{\theta} = (T'V^{-1}V)^{-1}T'V^{-1}\underline{x}$ is the result of minimizing $(\underline{x}-T\theta)'V^{-1}(x-T\theta)$ with respect to θ. The covariance matrix of $\hat{\theta}$ is given by $V(\hat{\theta}) = \sigma^2(T'V^{-1}T)^{-1}$.

The exact distributional properties of maximum likelihood and of best linear estimates from censored normal data must be found by

Monte-Carlo simulation techniques. Halperin (1952) proved that under mild regularity conditions the maximum likelihood estimator of a single parameter from singly censored samples is consistent, asymptotically normal and of minimum variance for large samples. Plackett (1958) showed that maximum likelihood estimators are asymptotically linear and that linear estimates based on the order statistics are asymptotically normal, efficient and unbiased for all censoring schemes.

Harter and Moore (1966) in a discussion of the bias of the maximum likelihood estimates state that the estimate for μ is negatively biased when $K_1 < K_2$, is positively biased when $K_2 > K_1$ and is unbiased when $K_1 = K_2$; and further that the estimate $\hat{\sigma}$ of σ is negatively biased regardless of the magnitude of K_1 and K_2. Saw (1961) shows that the magnitude of the bias where bias is $(E(\mu)-\hat{\mu})$ may be as much as 13 percent for a Type II censor scheme with 70 percent censored observations (in this case a (19, 7, 1) situation) or as low as .9 percent for a (19, 17, 1) censoring pattern.

The difficulties in both methods of estimation lies not in the estimation procedures, for on a modern electronic computer this is relatively simple, but in determining confidence intervals. It is common practice to set confidence intervals by using asymptotic normality in both M.L.E and B.L.E. cases. A more reliable method of setting confidence intervals might be to use a t-distribution with (K_2-K_1)

degrees of freedom. This is reasonable since with no censoring we use a t-distribution with $N-1$ degrees of freedom.

Gupta (1952) gives the following example which we repeat here to illustrate the methods under discussion.

Ten mice were innoculated with a uniform culture of human tuberculosis. Table 1 below shows the day on which the first seven of the ten mice died. Since the reaction times are more likely to be lognormally than normally distributed the estimates are based on $\log_{10}(x)$.

Table 1. Days after innoculation.

x	41	44	46	54	55	58	60
log x	1.613	1.643	1.663	1.732	1.740	1.763	1.778

Using the methods shown earlier, the maximum likelihood estimate of the mean is; $\hat{\mu} = 1.742$ with standard error $\sigma_{\hat{\mu}} = .0268$. The best linear estimate of the mean is $\tilde{\mu} = 1.746$ with standard error $\sigma_{\tilde{\mu}} = .0311$. Below for comparative purposes is Table 2 showing 95 percent confidence limits for these data and the posterior content thereof. The non Bayesian limits are set using a t-statistic with six degrees of freedom, while the posterior limits are taken directly from the posterior c.d.f. Graphs of the posterior distribution and of the approximate distributions of the maximum likelihood estimate and the best linear estimate of the mean are given below. The distributions of the non-Bayesian estimates are approximated by the normal distribution for plotting purposes. These could be thought of as approximate fiducial distributions, giving a description of all possible approximate confidence limits based on the approximate procedures.

Table 2. Confidence limits for Gupta's data.

	Confidence Level .025	.975	Posterior Content	
Bayes posterior limits	1.678	1.824	.025	.975
M.L.E. approx. limits	1.677	1.807	.023	.935
B.L.E. approx. limits	1.670	1.833	.015	.967

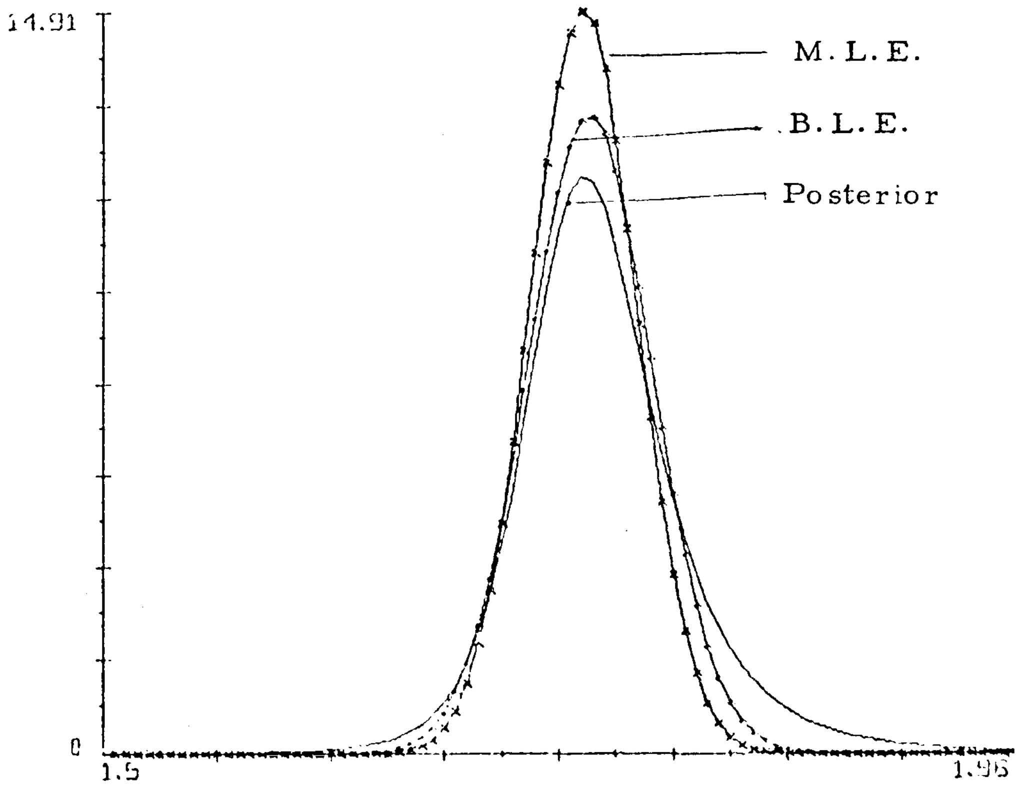

Figure 2. Posterior distribution from the data given by Gupta.

Importance of Ancillary Information

The reasons for the use of ancillary information were discussed in Chapter II where it was pointed out that when sufficient estimators were non-existent one could obtain conditionally sufficient estimators by conditioning on ancillary statistics. As the M.L.E. and B.L.E. are not sufficient estimators for the problem at hand, it is of interest to examine the importance of conditioning on the ancillary statistic. This can be done by examining the variability of the posterior content of confidence points among several samples taken at random from the same normal population and censored by the same censoring scheme. If the amount of ancillary information is considerable, there will be a large amount of variation in the posterior content of the confidence intervals based on the M.L.E. or B.L.E. If the ancillary information contained in the ancillary statistic is slight then the posterior content should have very little variability. If there were no information contained in the ancillary statistics, then the estimates would be sufficient and the posterior content would not vary at all.

Table 3 gives the maximum absolute difference between the posterior content of confidence levels of; .005, .025, .975, and .995 among five samples for many different censor schemes. The difference is generally on the order of .3 percent of the posterior content for the M.L.E. and .5 percent for the B.L.E. at the .975 level and

Table 3. Between sample differences of posterior content.

Censor Scheme	Confidence Points							
	M.L.E.				B.L.E.			
	.005	.025	.975	.995	.005	.025	.975	.995
10, 10, 1	0	0	0	0	.0012	.0043	.0045	.0013
10, 9, 1	.0001	.0001	.0001	.0001	.0009	.0014	.0029	.0013
10, 8, 1	.0015	.0044	.0003	.0002	.0017	.0034	.0073	.0049
10, 7, 1	.0010	.0014	.0030	.0029	0010	.0016	.0088	.0062
10, 6, 1	.0008	.0012	0028	.0026	.0009	.0012	.0071	.0060
10, 5, 1	.0004	.0011	.0080	.0050	.0003	.0012	.0100	.0080
10, 9, 2	.0027	.0077	.0073	.0024	.0013	.0045	.0044	.0023
10, 8, 3	.0005	.0017	.0008	.0004	.0005	.0022	.0026	.0011
10, 7, 4	.0027	.0050	.0041	.0024	.0007	.0016	.0023	.0017
20, 20, 1	.0	.0001	0	0	.0010	.0027	.0028	.0010
20, 18, 1	.0019	.0091	.0039	.0009	.0005	.0027	.0044	.0013
20, 16, 1	.0008	.0021	0006	.0002	.0008	.0022	.0025	.0013
20, 14, 1	.0004	.0010	.0004	.0002	.0007	.0010	.0043	.0023
20, 12, 1	0	.0022	.0010	.0004	0	.0003	.0045	.0011
20, 18, 3	.0022	.0096	.0107	.0022	.0008	.0055	.0046	.0014
20, 16, 5	.0012	.0027	.0024	.0011	.0008	.0024	.0023	.0009
20, 14, 7	.0028	.0054	.0014	.0010	.0009	.0014	.0023	.0013
50, 50, 1	.0001	.0002	.0002	.0001	.0002	.0019	.0018	.0006
50, 45, 1	.0016	.0093	.0051	.0012	.0010	.0094	.0042	.0012
50, 40, 1	0+	.0001	0+	0+	.0001	.0003	.0004	.0001
50, 35, 1	0+	0+	.0001	.0001	.0012	.0021	.0200	.0036
50, 45, 6	.0019	.0095	.0087	.0035	.0025	.0080	.0072	.0022
50, 35, 16	.0006	.0016	0+	.0002	.0002	.0034	.0009	.0001

less than 10 percent at the .025 level. Although there seems to be some rounding error (see censor scheme 20, 20, 1 for example) which contributed to the variation, the total variability is such that we can infer that the amount of information contained in the ancillary statistic is small compared to the information in the sample.

Visual Comparisons of Approximate Distributions of the Maximum Likelihood Estimate, Best Linear Estimate and Posterior Distribution

Given in this section are graphical comparisons of the posterior distribution with the approximate distributions of the M.L.E. and B.L.E. For each censor scheme two plots are given. The approximate distributions are assumed to be normal with parameters $\hat{\mu}$ and $\sigma_{\hat{\mu}}$ for the M.L.E. and $\tilde{\mu}$, $\sigma_{\tilde{\mu}}$ for the B.L.E. Each of the plots was done on a Calcomp plotter utilizing the software package Grafit of the Oregon State University Computer Center. The samples are normal random variates which are ordered prior to censoring. One sided censoring is plotted for one side only as censoring on the opposite side results in posterior distributions which are reflections of those shown.

Examination of these plots reveals several interesting points. First the posterior distribution is skewed in the direction of censoring. If the sample is censored on the right the skewness or heavy tail is to the right. Second, the posterior distribution and the approximate

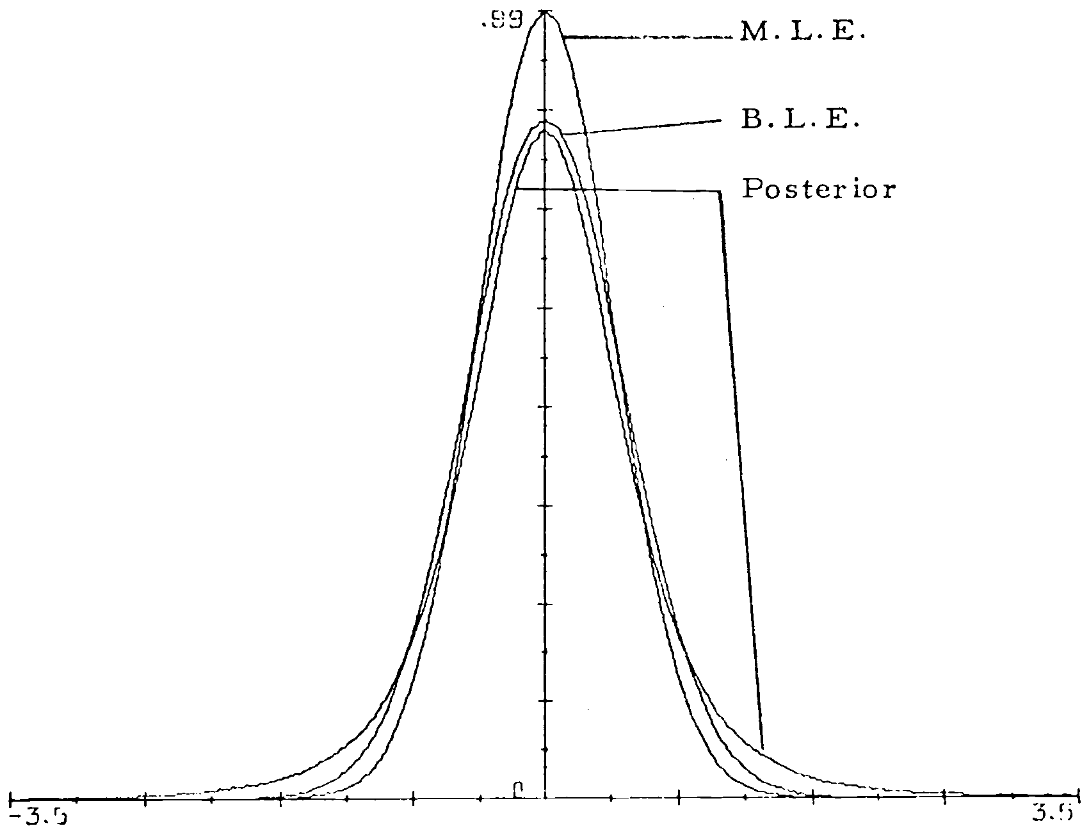

Figure 3. Plot of censor scheme (10, 8, 3) sample 1.

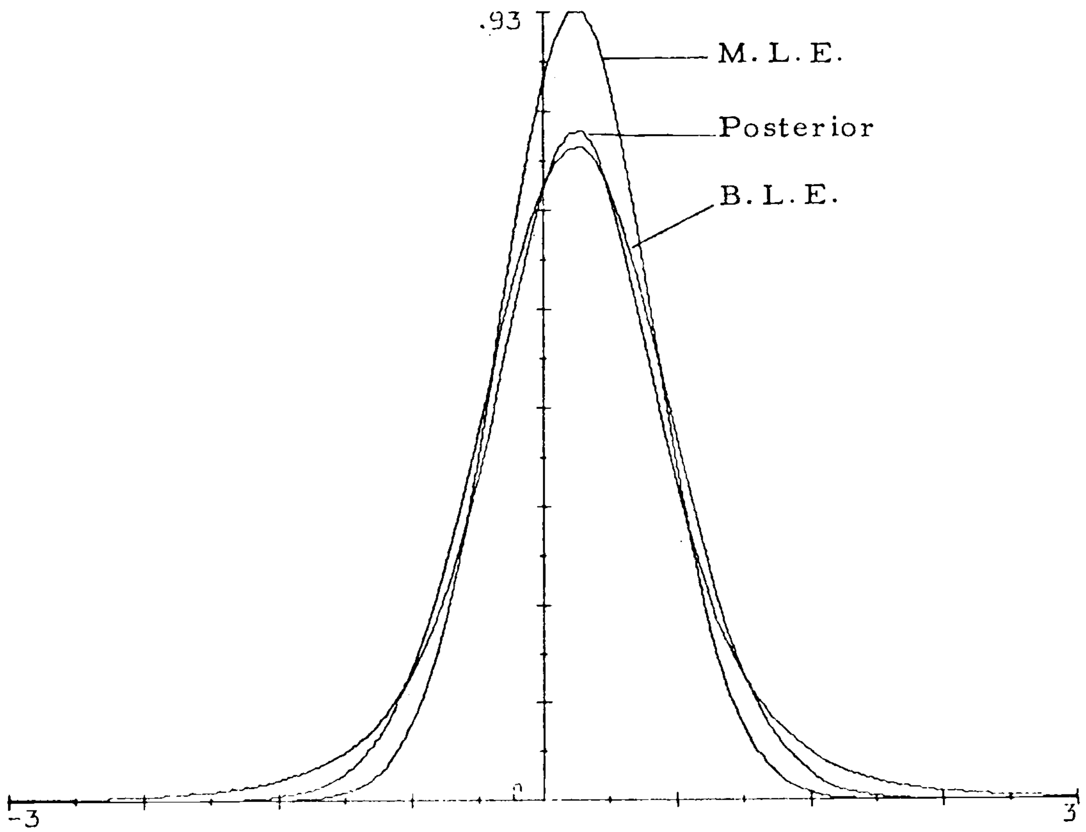

Figure 4. Plot of censor scheme (10, 8, 3) sample 2.

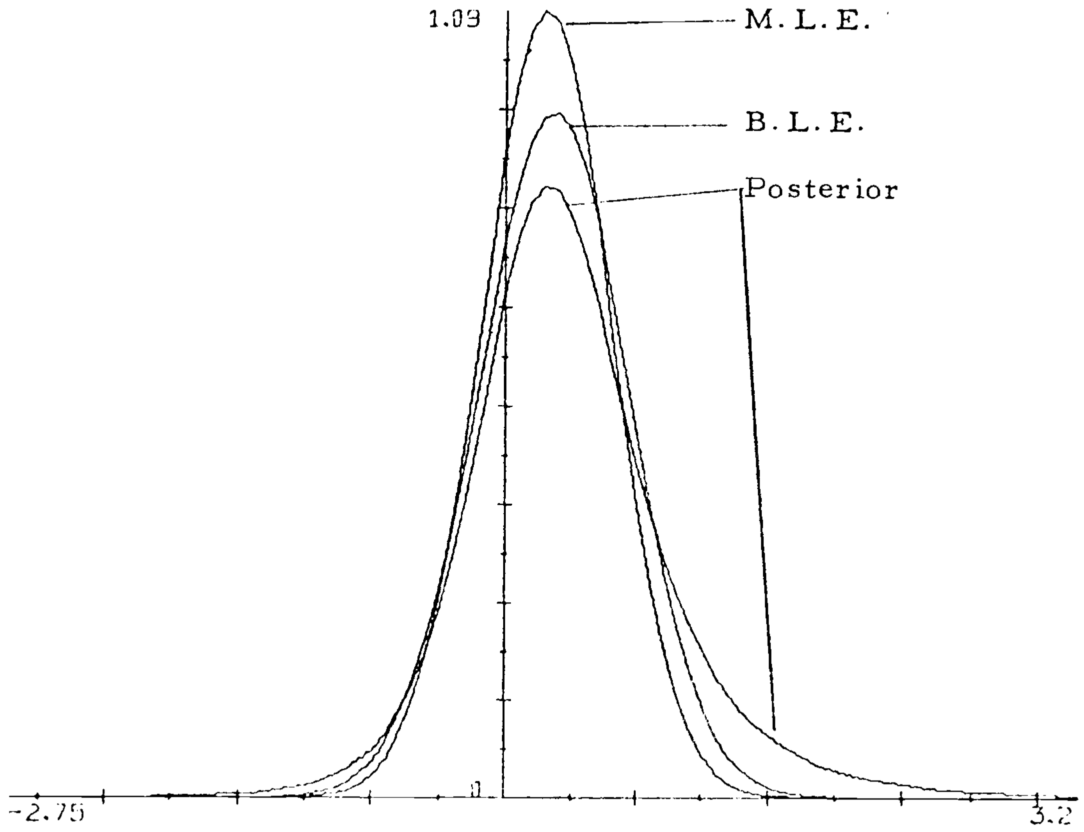

Figure 5. Plot of censor scheme (10, 7, 1) sample 1.

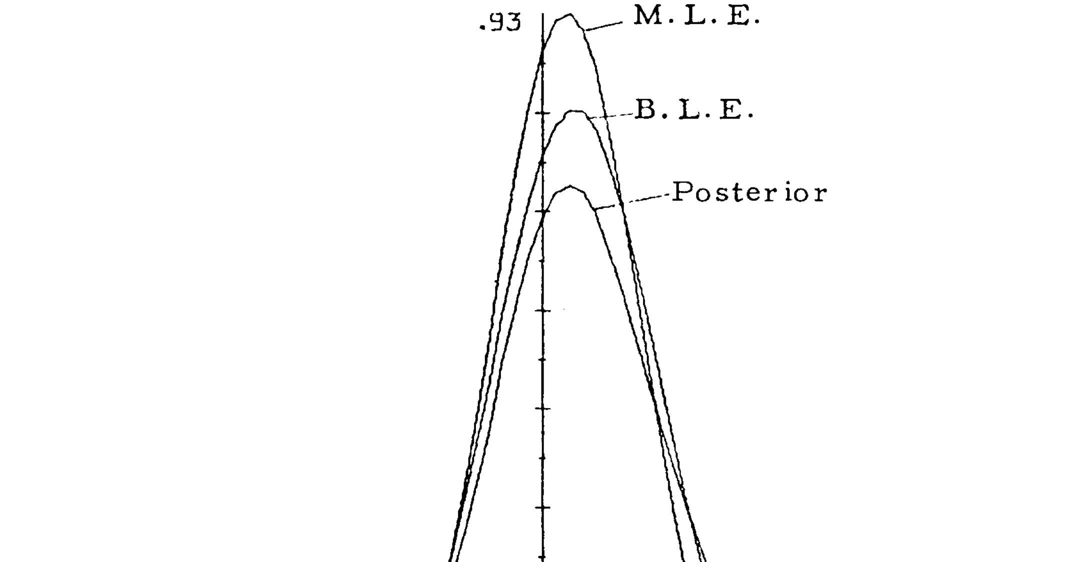

Figure 6. Plot of censor scheme (10, 7, 1) sample 2.

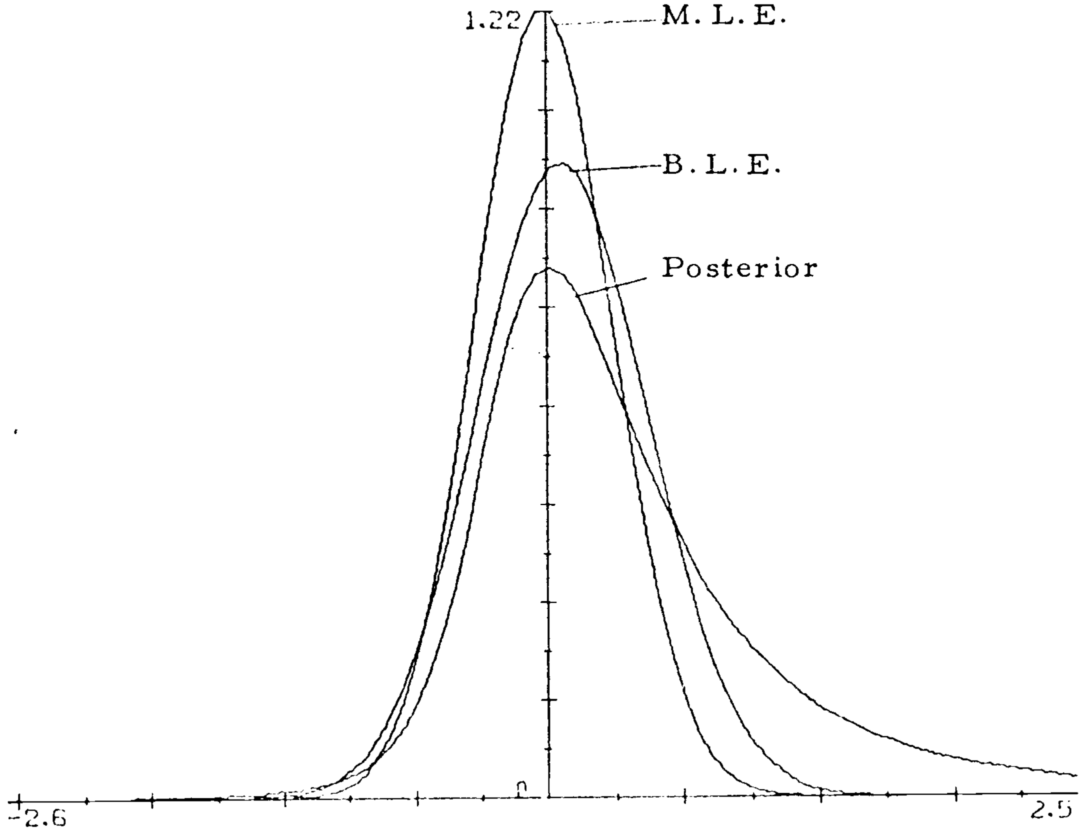

Figure 7. Plot of censor scheme (10, 5, 1) sample 1.

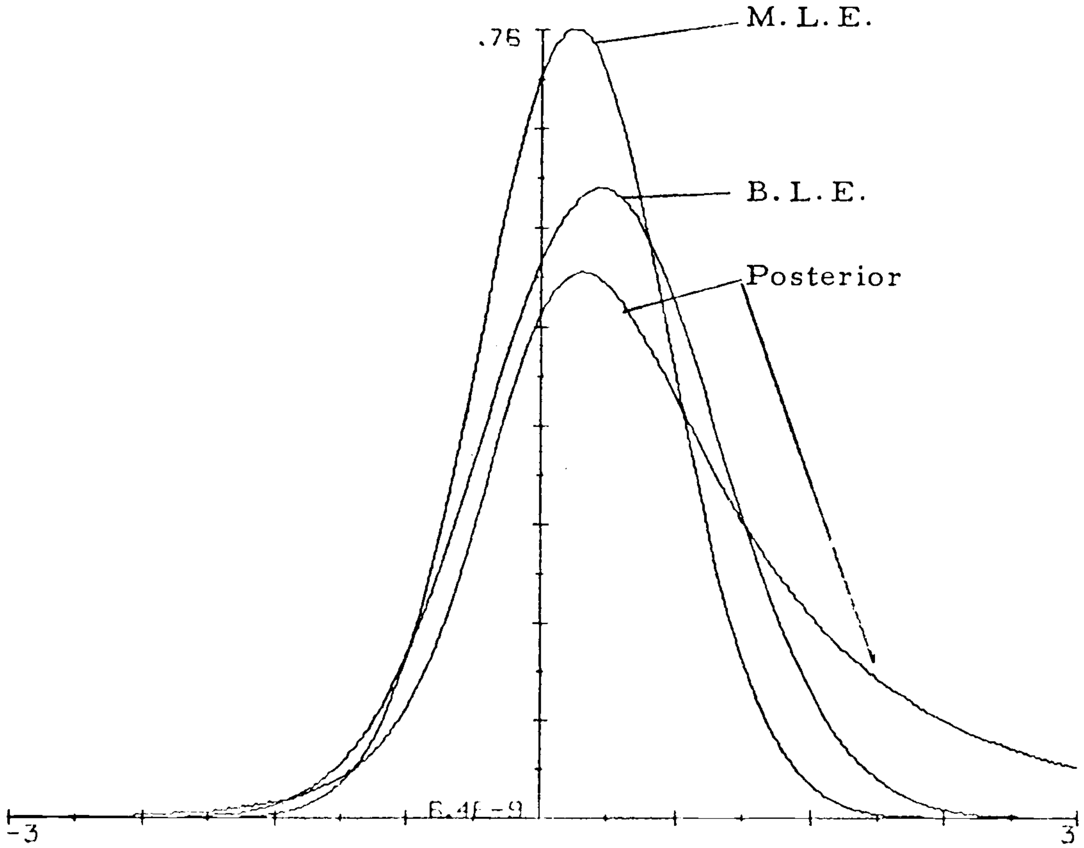

Figure 8. Plot of censor scheme (10, 5, 1) sample 2.

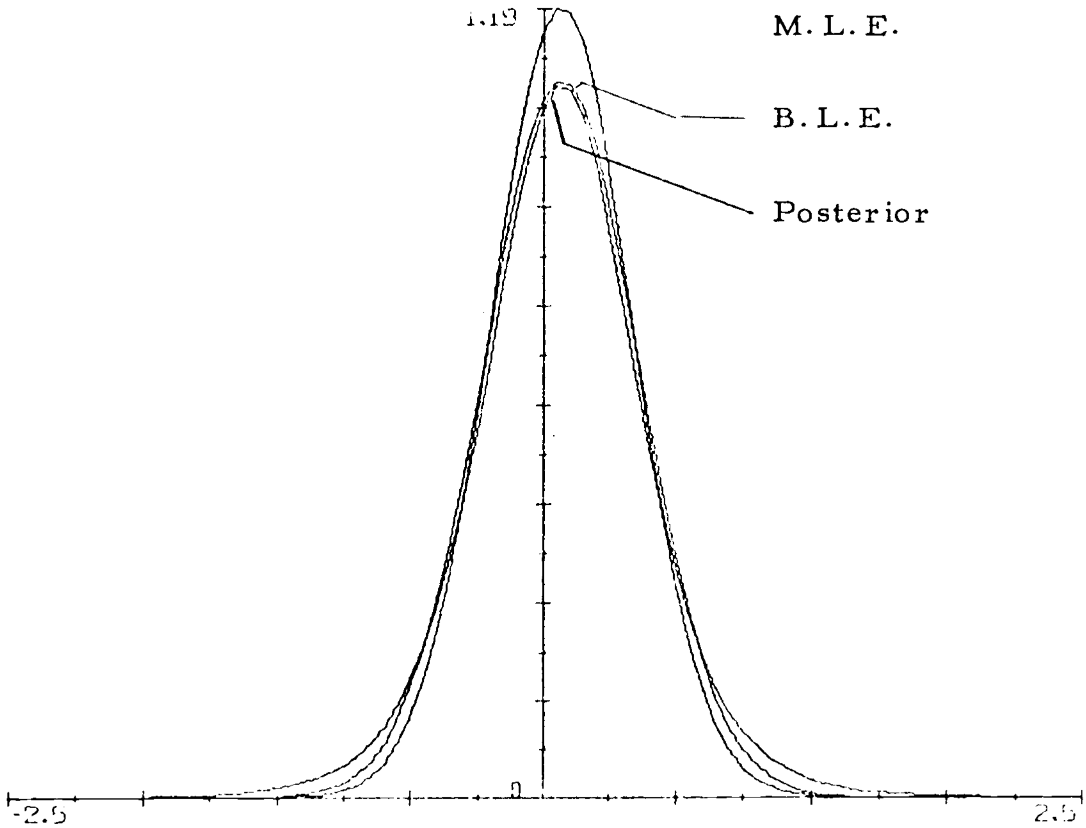

Figure 9. Plot of censor scheme (20, 15, 6) sample 1.

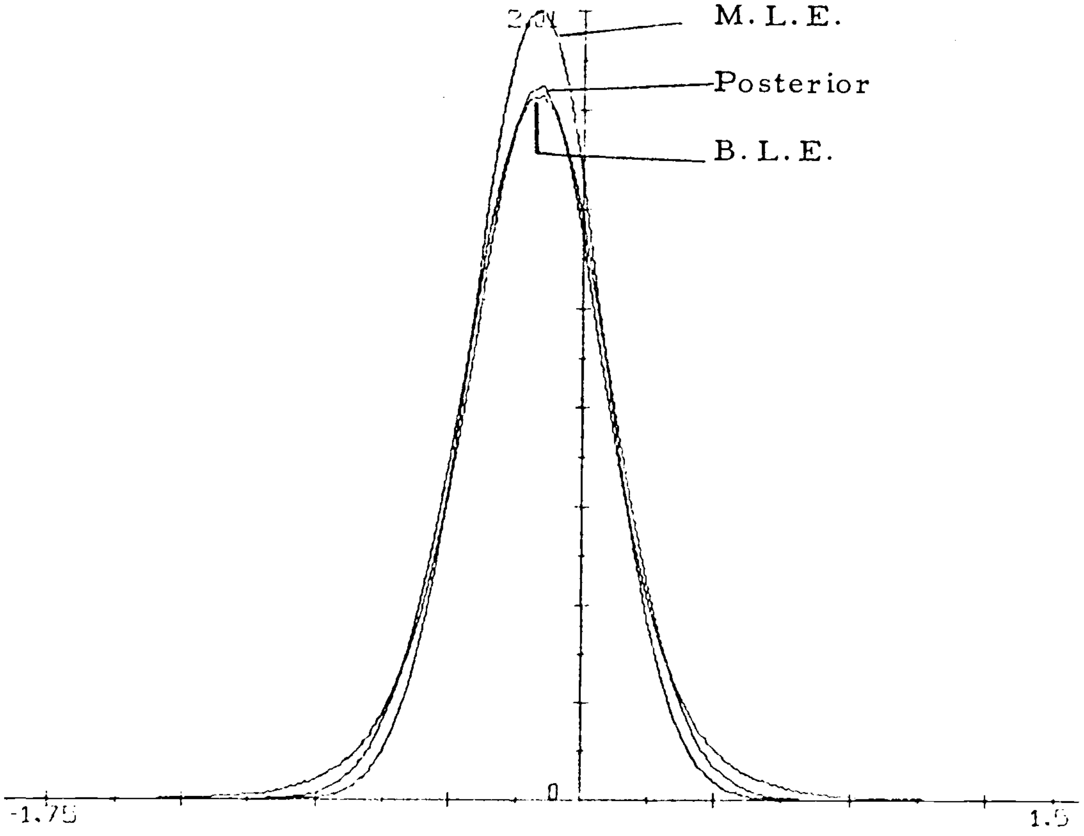

Figure 10. Plot of censor scheme (20, 15, 6) sample 2.

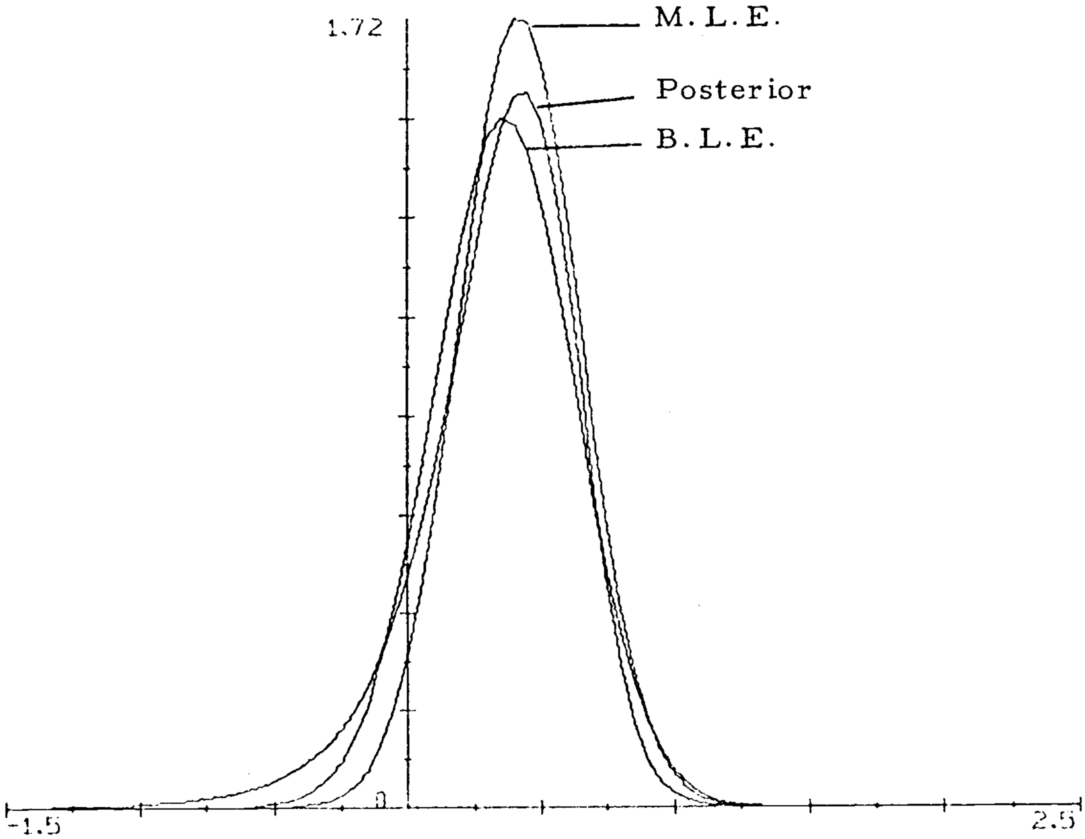

Figure 11. Plot of censor scheme (20, 20, 11) sample 1.

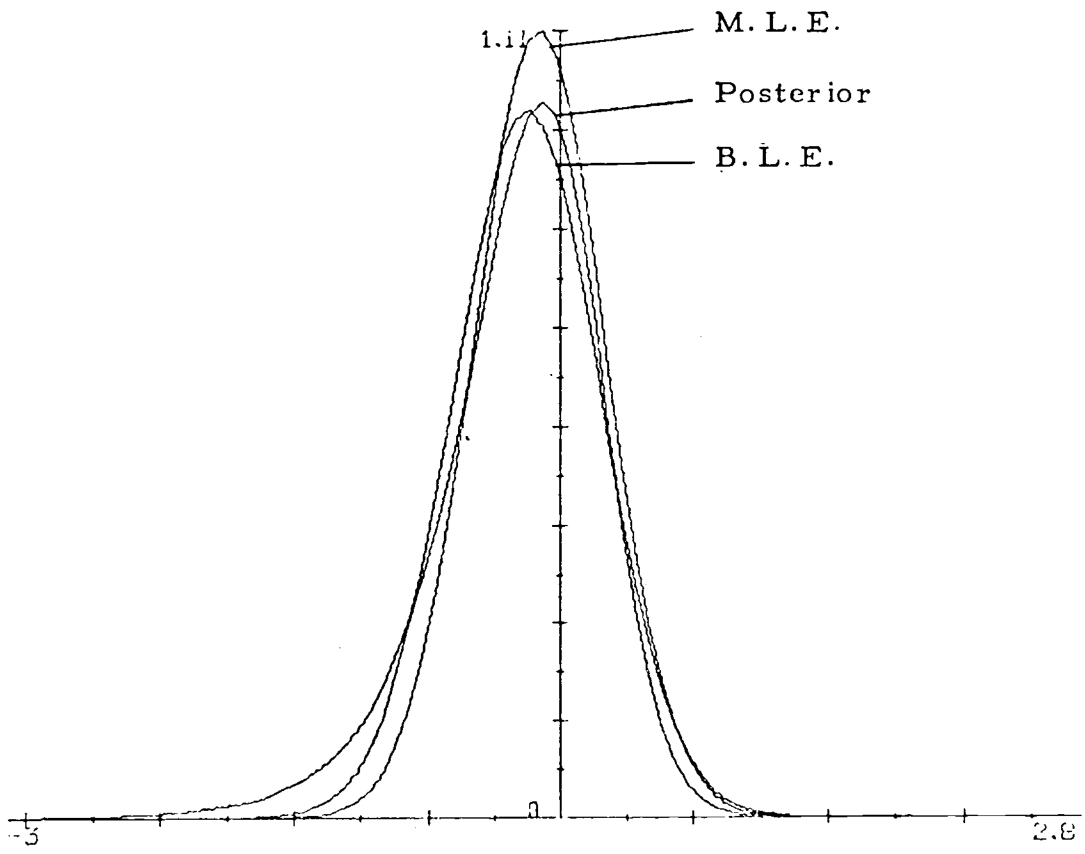

Figure 12. Plot of censor scheme (20, 20, 11) sample 2.

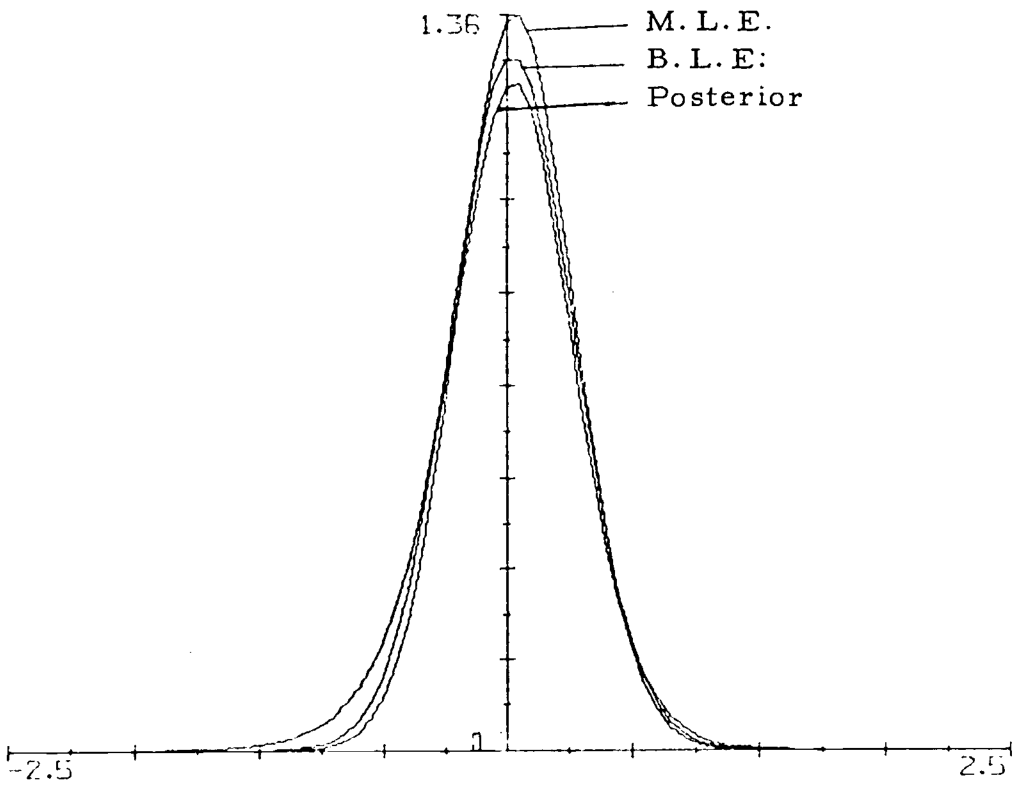

Figure 13. Plot of censor scheme (20, 20, 6) sample 1.

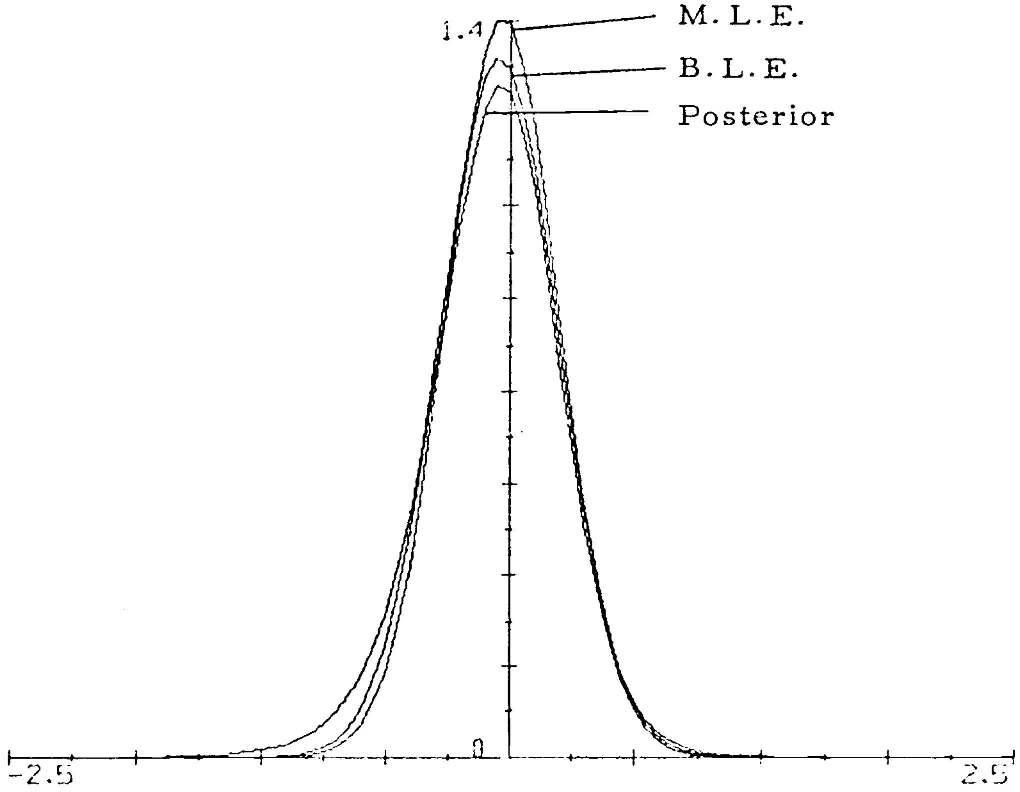

Figure 14. Plot of censor scheme (20, 20, 6) sample 2.

distribution of the B.L.E. have heavier tails than the M.L.E. This is particularly evident with symmetric censoring, and indicates that the standard error of the M.L.E. has a negative bias.

Numerical Comparison of Confidence Intervals

As shown in Chapter II, the Bayes-fiducial method has exact frequency interpretation and the Bayesian intervals have the confidence interval property. Let μ_a be the point on the posterior distribution where

$$\int_{-\infty}^{\mu_a} g(\mu/\underline{x})d\mu = a.$$

If we calculate values of $K = (\mu_a - \mu^*)/\sigma_{\mu^*}$, where μ^* is the M.L E. or the B.L.E., we have a useful numerical method of comparing the approximate distributions of the M.L.E. and B.L.E. with the posterior distribution. For example a value of $K = 6.477$ indicates that the standard error of the mean has been vastly underestimated and that the confidence intervals calculated using a normal approximation will not have proper posterior content.

Table 4 is a compilation of the values of K for both M.L.E. and B.L.E. Gupta's approximation of $V = I$ was used to calculate the best linear estimates for $N = 50$.

Table 4. Values of K for M.L.E. and B.L.E.

Censor Scheme	.005	.025	.975	.995
A. M.L.E.				
50,50, 1	2.769	2.031	2.031	2.709
50,45, 1	2.765	2.051	2.522	2.841
50,40, 1	2.721	2.008	2.138	2.872
50,35, 1	2,560	1.920	2.320	3.090
50,30, 1	2.430	1.830	2.360	3.177
50,45, 6	2.790	2.108	2.108	2.790
50,40,11	2.861	2.136	2.136	2.861
50,35,16	3.140	2.279	2.279	3.140
50,30,21	4.604	3.179	3.179	4.604
20,20, 1	2.894	2.136	2.136	2.894
20,18, 1	2.883	2.137	2.216	3.010
20,16, 1	2.830	2.098	2.315	3.181
20,14, 1	2.739	2.004	2.517	3.370
20,12, 1	2.571	1.874	2.936	4.302
20,10, 1	2.360	1.713	3.413	5.089
20,18, 3	3.037	2.258	2.258	3.037
20,16, 5	3.446	2.419	2.419	3.446
20,14, 7	4.307	2.929	2.929	4.307
10,10, 1	3.425	2.383	2.383	3.425
10, 9, 1	3.801	2.578	2.872	4.280
10, 8, 1	3.240	2.350	2.874	4.340
10, 7, 1	3.796	2.480	4.036	6.040
10, 6, 1	3.180	2.100	3.960	6.010
10, 9, 2	3.883	2.602	2.602	3.883
10, 8, 3	4.789	3.038	3.038	4.789
10, 7, 4	6.477	4.236	4.236	6.477
B. B.L.E.				
20,20, 1	2.790	2.060	2.060	2.790
20,18, 1	2.800	2.080	2.090	2.860
20,16, 1	2.750	2.050	2.190	2.980
20,14, 1	2.690	2.010	2.300	3.080
20,12, 1	2.460	1.830	2.570	3.820
20,10, 1	2.260	1.700	2.850	4.320
20,18, 3	2.870	2.110	2.110	2.870
20,16, 5	3.170	2.230	2.230	3.170
20,14, 7	3.660	2.410	2.410	3.680

Table 4. Continued.

Censor Scheme	.005	.025	.975	.995
10, 10, 1	3.160	2.260	2.260	3.160
10, 9, 1	3.530	2.390	2.590	3.900
10, 8, 1	3.120	2.170	2.590	3.830
10, 7, 1	3.430	2.260	3.440	5.205
10, 6, 1	2.622	1.922	3.167	4.892
10, 9, 2	3.515	2.360	2.360	3.515
10, 8, 3	4.022	2.550	2.550	4.022
10, 7, 4	6.448	4.198	4.198	6.448

The K_i developed in Tables 4 and 5 portray the values needed in order to obtain proper posterior content for given estimates of μ and σ. An informative method of comparison of confidence intervals calculated in natural ways, is examination of the Bayes posterior content of the intervals Let the intervals be,

$$\mu^* \pm t_{\alpha/2}(N-K_1-K_2-1)\sigma_{\mu^*}$$

where μ^* and σ_{μ^*} are the estimate and standard error of the estimate of the mean. The posterior content of the upper limit A and the lower limit B may be found by interpolation of the posterior C.D.F. That is

$$\Pr(\mu \leq A) = \int_{-\infty}^{A} \pi(\mu/\underline{x})d\mu = \pi(A/\underline{x}) .$$

Tables 6 and 7 give posterior content of censoring schemes for sample size 10, 20 and 50. As previously noted for sample size 50 the B.L.E. is calculated with $V = I$.

The posterior content in Table 5 and 6 and the values of K_i in Table 4 were both found using simulation. As previously noted the small input of the ancillary information to the posterior density leads to very nearly constant posterior content between samples with the same censoring pattern. Thus only five samples were generated at each point in order to determine the K_i and the posterior content of the interval of interest.

From these tables one can make several observations. First, for symmetric censoring the t-statistic with K_2-K_1-1 degrees of freedom, does well as a multiplier for finding confidence intervals. When censoring is greater than about 20 percent, values larger than this t value should be used. When censoring is not symmetric, the t-statistic tends to lead to an overestimate on the non-censored side while underestimating on the censored side.

Table 5. Approximate posterior content of confidence intervals calculated from Maximum Likelihood Estimates.

Censor Scheme	Confidence Level .005	.025	.975	.995
50, 50, 1	.0056	.0270	.973	.994
50, 45, 1	.006	.032	.973	.992
50, 40, 1	.005	.024	.969	.993
50, 35, 1	.004	.019	959	.991
50, 30, 1	.002	.016	.960	.987
50, 45, 6	.006	.040	.974	.994
50, 40, 11	.007	.032	.969	.993
50, 35, 16	.004	.026	.972	.995
50, 30, 21	.003	.008	.916	.975
20, 20, 1	.007	.032	.968	.993
20, 18, 1	.006	.030	.966	.992
20, 18, 1	.006	.030	.966	.992
20, 16, 1	.005	.028	.959	.989
20, 14, 1	.004	.022	.950	.986
20, 12, 1	.002	.016	.938	.982
20, 10, 1	.001	.009	.910	.963
20, 18, 3	.007	.036	.964	.992
20, 16, 5	.006	.035	.962	.993
20, 14, 7	.013	.048	.949	.986
20, 12, 9	.025	.099	.898	.974
10, 10, 1	.006	.034	.966	.994
10, 9, 1	.008	.040	.949	.989
10, 8, 1	.004	.030	.920	.985
10, 7, 1	.004	.030	.920	.985
10, 6, 1	.001	.016	.933	.993
10, 5, 1	.002	.011	.839	.940
10, 9, 2	.006	.038	.962	.994
10, 8, 3	.004	.041	.958	.996
10, 7, 4	.012	.058	.904	.977

Table 6. Approximate posterior content of confidence intervals calculated from Best Linear Estimates.

Censor Scheme	Confidence Level .005	.025	.975	.995
50, 50, 1	.005	.025	.975	.995
50, 45, 1	.004	.022	.985	.998
50, 40, 1	.001	.009	.985	.998
50, 30, 1	.0001	.0002	.9926	.999
50, 40, 11	.001	.010	.993	.999
50, 30, 21	.000	.0006	.999	1.000
20, 20, 1	.005	.026	.974	.995
20, 18, 1	.005	.027	.971	.994
20, 16, 1	.004	.025	.966	.992
20, 14, 1	.003	.020	.961	.991
20, 14, 1	.002	.014	.955	.989
20, 10, 1	.001	.008	.935	.975
20, 18, 3	.006	.031	.969	.994
20, 16, 5	.004	.027	.970	.996
20, 14, 7	.007	.031	.966	.992
20, 12, 9	.007	.020	.980	.992
10, 10, 1	.003	.024	.976	.996
10, 9, 1	.004	.031	.962	.994
10, 8, 1	.002	.021	.969	.997
10, 7, 1	.001	.027	.952	.990
10, 6, 1	.000	.010	.969	.999
10, 5, 1	.001	.007	.902	.975
10, 9, 2	.003	.026	.973	.997
10, 8, 3	.000	.022	.977	.999
10, 7, 4	.005	.037	.940	.991

IV. NUMERICAL METHODS

A significant amount of time was devoted to development of the numerical procedures used in this thesis. The largest portion of this time was spent on the development and assessment of techniques used in computing posterior distributions. Much of this study was devoted to the assessment of the accuracy of the numerical procedures presented in this chapter. This included development of a rescaling procedure for use in the numerical quadrature techniques chosen for the numerical integration and studies on the size of the numerical quadrature needed to give satisfactory results.

Posterior Distributions

Let us assume that a sample of size N is observed and that the observations follow a normal distribution with mean μ and variance σ^2. Let the observations be censored by a Type II censoring scheme so that the smallest K_1-1 and the largest $N-K_2$ observations are censored. That is the censor scheme is (N, K_2, K_1). The likelihood function is then of the form (1.1), that is:

$$L(\underline{x}/\mu,\sigma) \propto (1/\sigma)^{K_2-K_1+1} \exp\left(-\frac{1}{2}\sum_{K_1}^{K_2}(x_i-\mu)^2/\sigma^2\right)(F(x_{K_1}))^{K_1}(1-F(x_{K_2}))^{N-K_2})$$

where

$$F(x) = 1/(\sqrt{2\pi}\,\sigma) \int_{-\infty}^{x} \exp - \frac{1}{2}(t-\mu)^2/\sigma^2 \, dt \; .$$

By applying Bayes formula with a prior density of the form $\pi(\mu, \sigma) \propto 1/\sigma$, the posterior density may be written as

(4.1) $$\pi(\mu, \sigma/\underline{x}) \propto \pi(\mu, \sigma) L(x/\mu, \sigma)$$

$$= (1/\sigma)^{K_2 - K_1 + 2} \exp - \frac{1}{2} \sum_{K_1}^{K_2} (x_i - \mu/\sigma)^2$$

$$\times (F(x_{K_1}))^{K_1} (1 - F(x_{K_2}))^{N-K_2})$$

Inferences about the population parameters μ and σ may be made from the marginal posterior distributions, that is, from;

(4.2) $$h(\sigma/\mu, \underline{x}) \propto \int H(\mu, \sigma/\underline{x}) d\mu$$

(4.3) $$g(\mu/\sigma, \underline{x}) \propto \int H(\mu, \sigma/\underline{x}) d\sigma$$

For the non-censored normal case and for several other distributions this integration can be done in "closed form". In the censored normal case the form of the posterior distribution is such that numerical integration techniques are needed to evaluate the marginal posterior distributions. Several different techniques could be used to accomplish this integration. One could use Simpson's rule, Romberg integration, Gaussian quadrature or other numerical procedures for the integration.

The Numerical Integration

The Gauss-Laguerre numerical integration technique was chosen as the method to evaluate the integral (4.3), because it is known to work well over a wide range of values and is a proven, highly accurate technique. The Gauss-Laguerre method of numerical integration involves approximating an integral of the form;

$$\int_0^\infty \exp -\theta \cdot f(\theta) d\theta \tag{4.4}$$

with the sum $\sum_{i=1}^{M} w_i f(a_i)$.

The base points, a_i, used in the M-point integration are the roots of the M-th degree Laguerre polynomial

$$\mathcal{L}_M(x) = (2M-x-1)\mathcal{L}_{M-1}(x) - (M-1)^2\mathcal{L}_{M-2}(x)$$

where

$$\mathcal{L}_0(x) = 1; \quad x \in (0,\infty).$$

The corresponding weight factors w_i are given by

$$w_i = \int_0^{\infty} \exp - z \prod_{\substack{j=0 \\ j \neq i}}^{M} (z - z_j)/(z_i - z_j) dz \ .$$

The integral (4.4) can be shown to be exact when $f(\theta)$ is a polynomial of degree $(2M-1)$ or less (see Carnaham, Luther, Wilks, pp. 100-115).

Because the interpolating polynomial evaluates the function $\exp -\theta \cdot f(\theta)$ at fixed points, a_i, we felt that the domain of the function $\exp -\theta \cdot f(\theta)$ should be adjusted in order for the interpolating polynomial and hence the numerical integration to obtain maximum accuracy. There is remarkably little comment on this idea in the textbooks on numerical integration. One is by Hammings (1962, p. 163).

The joint posterior distribution (4. 1) may be written in the proper form for the Gaussian integration by using the substitution;

(4.5)
$$\theta = \sum_{i=K_1}^{K_2} (x_i - \mu)^2 / \sigma^2 = SSx/\theta^2$$

The posterior distribution may then be written as

$$\text{(4.6)} \qquad \pi(\mu, \sigma/\underline{x}) \propto (2\theta/SSx)^{(K_2-K_1+1)/2} e^{-\theta} F\left(\frac{x_{K_1}-\mu}{\frac{\sqrt{SSx}}{2\theta}}\right)^{K_1}$$

$$\times \left(1-F\left(\frac{x_{K_2}-\mu}{\frac{\sqrt{SSx}}{2\theta}}\right)\right)^{N-K_2} SSx/\theta^{-3/2}$$

$$= (2\theta/SSx)^{(K_2-K_1+1)/2} \exp - \theta \cdot f(\theta)$$

As mentioned previously this function must be rescaled in order to achieve maximum accuracy. In order to achieve the rescaling let $\theta^* = \theta/S$, where θ^* has the same range as the interpolating polynomial. Then let $L^*(\theta) = \exp - \theta/S \cdot f(\theta/S)$, so that

$$g(\mu/x) = \int_0^\infty \exp - \theta^* f(\theta^*) d\theta^* = \int_0^\infty L^*(\theta) d\theta/S$$

$$= \int_0^\infty \exp - \theta[(\exp{+\theta}) L^*(\theta)] d\theta/S$$

$$\doteq 1/S \sum_1^M w_i \cdot \exp(a_i) \cdot L^*(a_i)$$

$$= 1/S \sum_1^M w_i(\exp(a_i) \cdot (1-1/S)) \cdot f(a_i/S)$$

This rescaled function has the property that the domain of interest of the approximated function is the same as that of the approximating polynomial, that is, for proper choice of S. In order to find an appropriate value of S, one needs to determine the range of θ. Recall from (4.5) that

$$\theta = \sum^{M} (x_i - \mu)^2 / 2\sigma^2 .$$

If we determine a lower limit for σ^2, say σ_L^2, we could then find an upper limit for θ, say θ_U. For a given size of interpolating polynomial the maximum value of a_i is known and $S = \max(a_i)/\theta_U$ is an appropriate choice for the scale factor.

The use of the scale factor is illustrated in Figure 15 below where a sample of size 20 was taken from a $N(0, 1)$ population. The sample was censored using a (20, 15, 1), Type II censoring scheme. Depicted is a plot of $\exp - \theta \cdot f(\theta)$ in 15A, and $1/S \cdot \exp - \theta/S \cdot f(\theta/S)$ in 15B.

Figure 15B takes into account that most of the area under the curve is prior to θ_U, which is taken to be 17.90.

Better accuracy should be obtained by using the scale factor in the calculation of the posterior distribution. In order to determine the magnitude of the improvement, the same posterior distributions were calculated with and without the scale factor. Comparisons of upper and lower confidence points were done for several different sample sizes and several censoring schemes. These comparisons are shown in Table 7 below. All data is generated from $N(0, \sigma^2)$ distributions; posterior distributions are calculated using Gaussian quadratures of order five and ten.

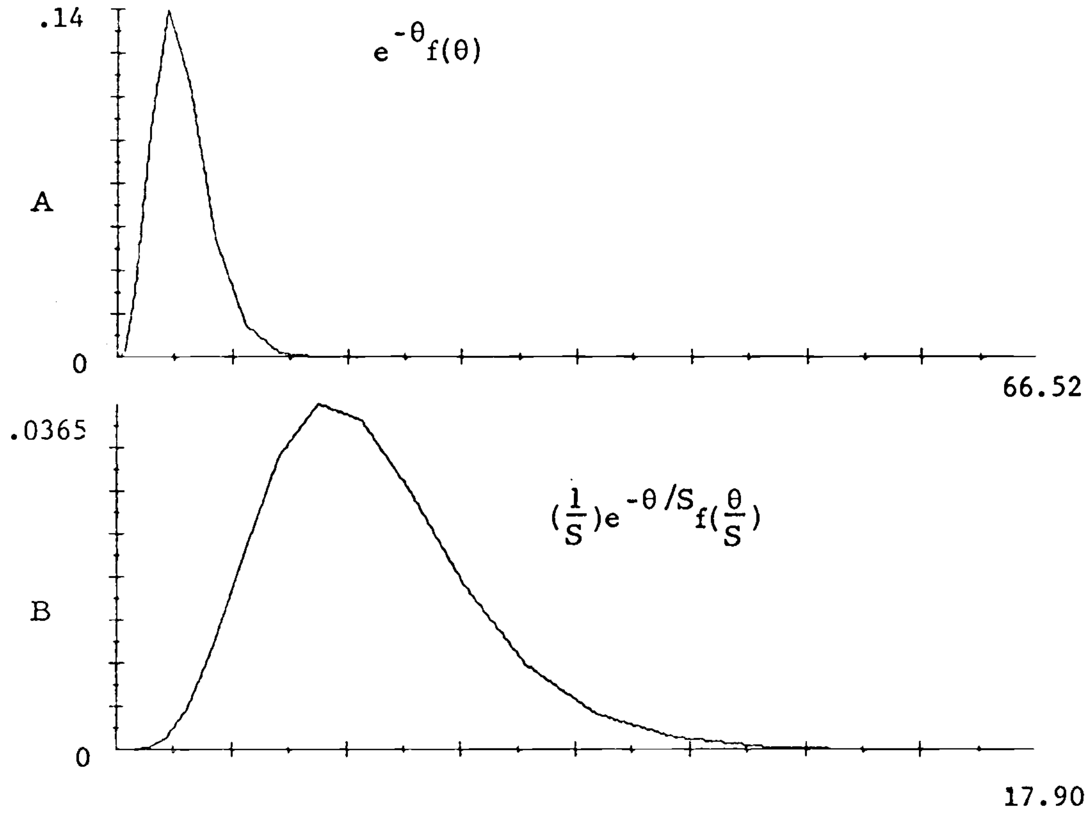

Figure 15. Plot of $e^{-\theta}f(\theta)$ and $(1/S)e^{-\theta/S}f(\theta/S)$.

The largest variation shown in Table 7 is seen to be approximately six percent where the censor scheme is (20, 16, 1), with a quadrature of order five from a $N(0, 1)$ distribution. This is a remarkably small effect, which we do not really understand. This may or may not be an important improvement depending on the nature of the problem. It is felt that particularly where lower order quadratures are used the scale factor should be used.

Table 7. Differences in posterior content due to scale factor.

Censor Scheme $(N\ K_2\ K_1)$	σ^2		Posterior Confidence Limit .005	.025	.975	.995
Quadrature Order 10						
(100,60,1)	1	Scale	- .2480	- .1828	.2731	.3570
		No scale	- .2485	- .1833	.2662	.3479
(10, 5,1)	1	Scale	-1.1660	- .7049	1.9409	2.4141
		No scale	-1.1790	- .7134	1.9400	2.4138
(5, 3,1)	1	Scale	-2.6001	-1.8826	2.1266	2.3113
		No scale	-2.5977	-1.8741	2.1290	2.3113
(20,15,1)	.01	Scale	- .0936	- .0749	.0376	.0578
		No scale	- .0934	- .0749	.0376	.0578
(20,15,1)	100	Scale	-6.7173	-5.6853	.8764	2.0277
		No scale	-6.6806	-5.6175	.8791	2.0297
Quadrature Order 5						
(20,16,1)	1	Scale	- .5783	- .3834	.7717	.9907
		No scale	- .5945	- .4058	.7720	.9910
(10, 7,1)	1	Scale	-1.4700	- .9018	1.8720	2.733
		No scale	-1.4786	- .9074	1.8698	2.7314

Simpson's rule was used for the integration over μ to find the normalizing constant and the posterior C.D.F. In order to do this the posterior density was calculated numerically at 81 values of μ, the lower limit of which was $\hat{\mu} - 4\sigma_{\hat{\mu}}$ where $\hat{\mu}$ is the M.L.E. When the 81 values of $\pi(\mu/\underline{x})$ have been determined the application of

Simpson's rule is straight-foreward.

Size of the Gaussian Quadrature

Tables of Gauss-Laguerre integration are available up to 50 points. As the posterior distribution must be calculated using 81 Gaussian quadrature numerical integrations (one for each value of μ) plus an additional integration to determine the constant of integration, which is found using Simpson's rule, a significant amount of computer time may be saved by using the smallest Gaussian quadrature possible. A quadrature of order 20 requires 1215 more calculations of (4.5) than a quadrature of order 5. If a two parameter problem were to be proposed the double integration needed would require calculation of the posterior distribution 32,400 times for a quadrature of order 20. While if a quadrature of order five were used the number of calculations would be reduced to 2025, a considerable savings of computer time. Thus it was of interest to compare accuracy of numerical procedures using different sizes of quadrature.

Numerical integration schemes of order 3, 5, 10, 15 and 20 were compared graphically by calculating the posterior distribution for each integration size and then plotting each posterior on the same coordinate axia.

The censor schemes checked are representative of those which might be used in practice. Plots of several of the resulting posterior distributions are shown below.

The experimental plots showed that a quadrature of order five is sufficient in this application. This means that a polynomail of degree less than or equal to nine could be used to fit the points of the posterior distribution. Quadrature of order ten was used throughout this thesis as a conservative measure.

The Computer Program

A computer program developed for this thesis is presented in the Appendix. Features of this program are; joint estimation of M.L.E., B.L.E. and Bayesian confidence intervals; choice of size of the Gauss-Laguerre numerical integration procedure; ability to simulate data or to read from a file. With only minimal changes the program can be used with Type I censoring schemes. It is hoped that this program will be found useful by experimenters and statisticians who are engaged in research where data are approximately normal but are censored.

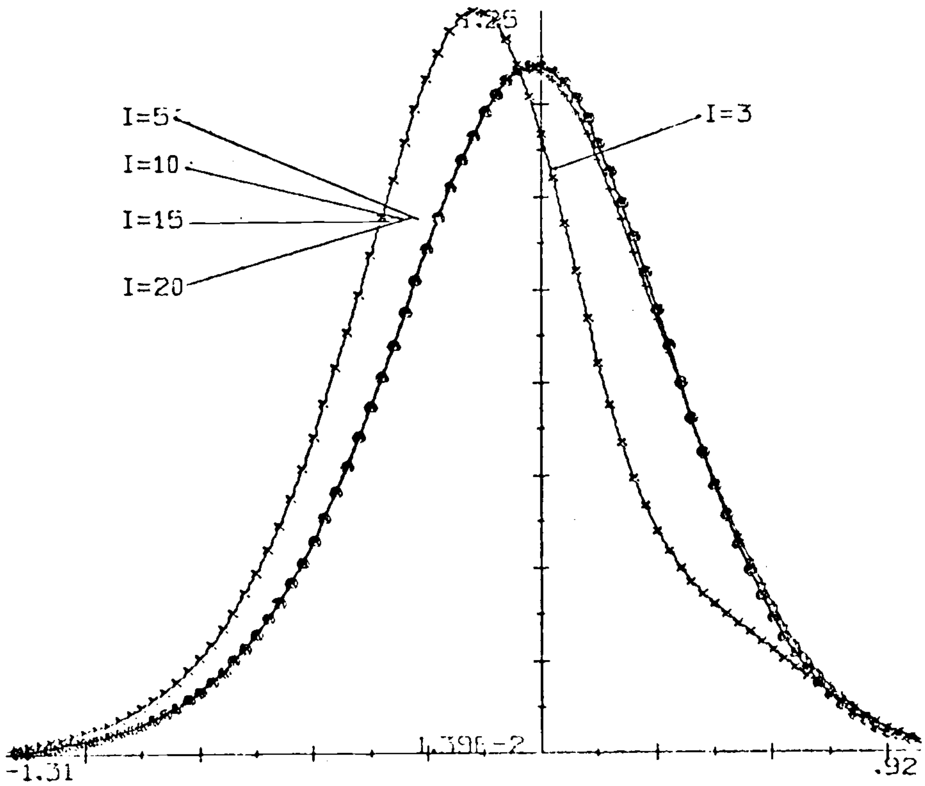

Figure 16. Quadratures of order 3, 5, 10, 15, 20 with censor scheme (20, 20, 6).

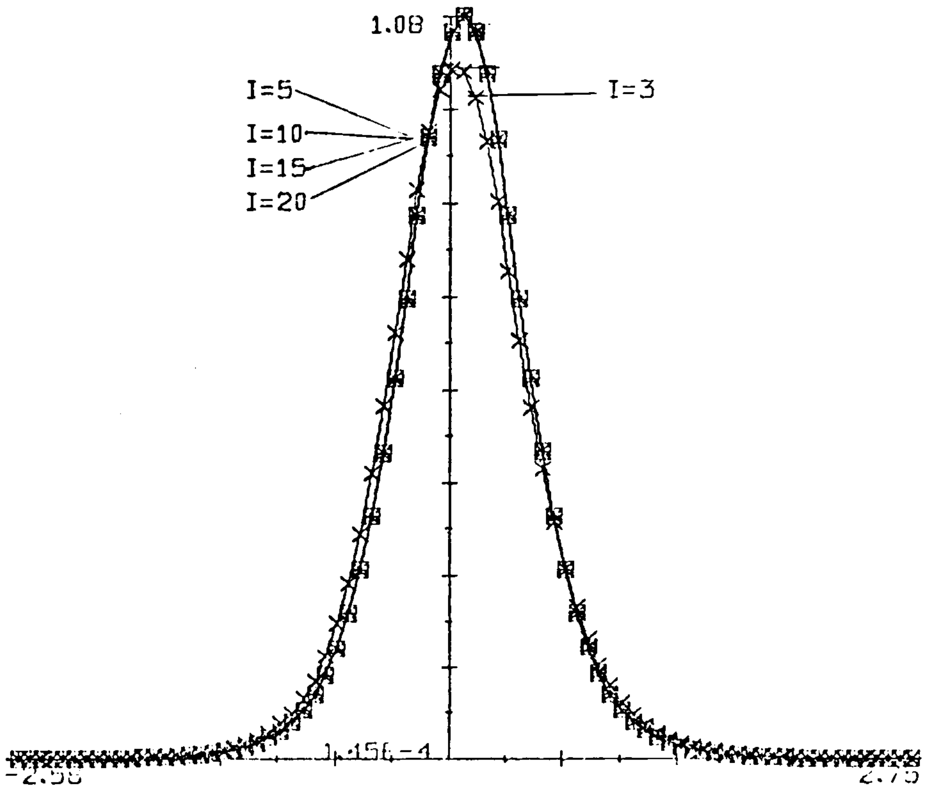

Figure 17. Quadratures of order 3, 5, 10, 15, 20 with censor scheme (20, 15, 6).

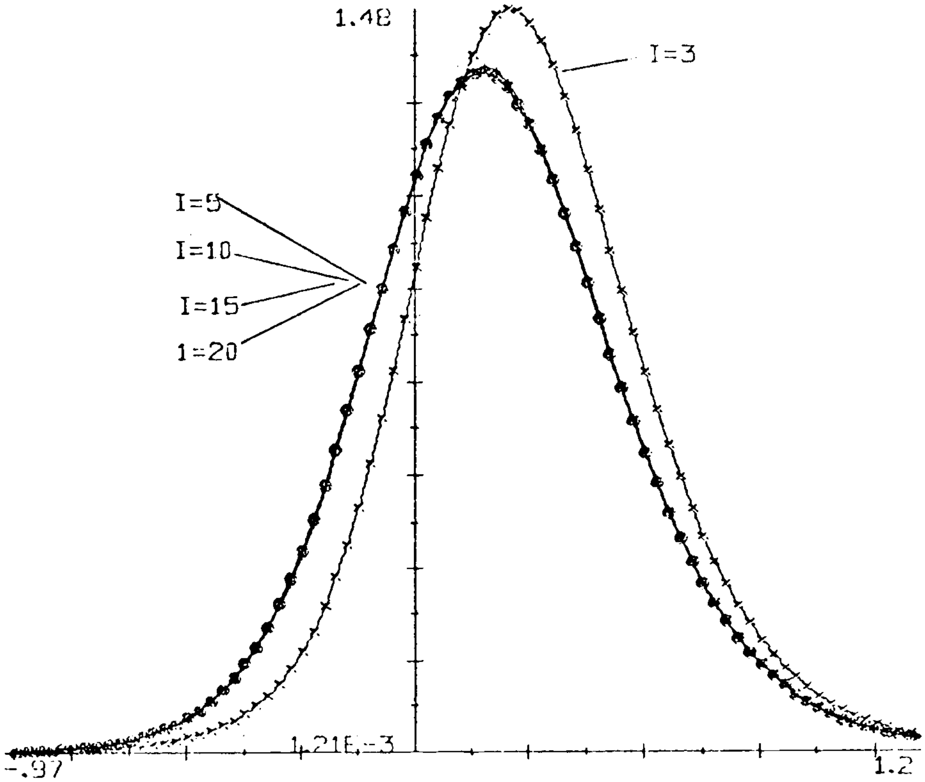

Figure 18. Quadratures of order 3, 5, 10, 15, 20 with censor scheme (20, 15, 1).

V. SUMMARY AND CONCLUSIONS

This study is directed toward comparing the Bayes-Fiducial method of analysis with Type II censored data from a normal distribution with the more common approximate methods of maximum likelihood and linear estimates based on order statistics. Supplemental to this study is the determination of the frequency properties of the Bayes-Fiducial method through application of the theorems of Hora and Buehler (1966).

It is shown that the Bayesian method of analysis provides a way of using all information contained in the sample. This was accomplished by showing equality of the Bayes-posterior with the fiducial density, conditioned on ancillary statistics. The comparison between the Bayes-Fiducial method and the approximate maximum likelihood and linear estimates based on the order statistics was thus between one which used all information afforded by the sample and one whose estimates are non-sufficient. Through this study it was determined that the lack of sufficiency is not nearly as important as the failure of the M.L.E. and B.L.E. to meet the usual distributional assumption of normality. It was shown that the assumption of asymptotic normality would lead to confidence intervals with substantially less than the desired posterior content. In fact a much better confidence interval estimate in all cases is found using a t statistic with $K_2 - K_1 - 1$ degrees of freedom.

Within the framework of the thesis topic the consequences of using low-order Gauss-Laguerre quadrature was investigated. It was determined that quadrature of order five is necessary to give the desired numerical accuracy. With two sample problems where double numerical integration is necessary, a reduction in the size of the Gaussian numerical integration will result in significant savings in computer time. The effect of a change of scale within the numerical integration was discussed. It was found that for quadratures of order ten scaling did not apparently increase accuracy, for quadrature of order five accuracy increases of at least 5.8 percent were realized.

Other Possible Topics

Several related topics which were not covered in the bulk of this thesis may be pertinent. The effect of another prior distribution was not considered for several reasons. We were interested mainly in comparing approximate methods with the Bayes-fiducial method which was shown in Chapter II to have an exact frequency interpretation. In the true Bayesian sense, if we have prior information about the parameters we should use it. We wished however to assume that the prior information was negligible in relation to the sample. Additional study on other priors is a good choice for future research.

Point estimation is another topic which was not considered. The posterior mode can be read directly from the posterior distribution. When a point estimate of the mean is needed the mode would be the easiest to determine. Computer calculation of the posterior mean is only slightly more difficult. It was felt however that the major usage of censored data demanded that we consider only comparisons of confidence intervals and not point estimates.

A third topic not discussed is that of the two sample problem. Here we would like information about the distribution of $\phi = \mu_1 - \mu_2$. For simplicity one could study the case where $\sigma_1^2 = \sigma_2^2$. Then to find the posterior density of ϕ we could use the transformation of variables,

$$\phi_1 = \mu_1 - \mu_2$$

$$\phi_2 = \mu_1 + \mu_2$$

and find the joint posterior distribution $\pi(\phi_1, \phi_2, \sigma / \underline{x})$ where a prior of the form $\pi(\mu_1, \mu_2, \sigma) \propto 1/\sigma$ is assumed. A double numerical integration is then necessary to find the posterior density of ϕ_1. This posterior distribution might then be used to test the null hypothesis, $\mu_1 = \mu_2$.

The last point to be mentioned is that of Bayesian tolerance regions. If $\zeta_p = \mu + K_p \sigma$ is a quantile, then a confidence region of ζ_p may be seen to be a tolerance region.

Let

$$\Pr(\zeta_p \le U) = \alpha = \Pr(\mu + K_p \sigma \le U)$$

$$= \Pr(\mu + F^{-1}(p)\sigma \le U) \quad \text{if} \quad K_p = F^{-1}(p)$$

$$= \Pr(F(\frac{U-\mu}{\sigma}) \le p)$$

$$= \Pr(F(U) \le p) = \alpha .$$

It would be a straightforward procedure to compute the posterior density of ζ_p for given p.

BIBLIOGRAPHY

Bogdanoff, D.A. and Pierce, D.A. (1973) Bayes-Fiducial Inference for the Weibull Distribution. Journal of the American Statistical Association. 68, 659-664.

Cohen, A.C. (1959) Simplified Estimators for the Normal Distribution When Samples are Singly Censored or Truncated. Technometrics 1, 217-237.

Cohen, A.C. (1961) Tables for Maximum Likelihood Estimates; Singly Truncated and Singly Censored Samples. Technometrics 3, 535-541.

Cox, D.R. and Hinkley, D.V. (1974) Theoretical Statistics, Chapman and Hall, London.

David, H.A. (1970) Order Statistics, John Wiley and Sons, New York.

Dixon, W.J. (1960) Simplified Estimation From Censored Normal Samples. Annals of Mathematical Statistics. 31, 385-391.

Fisher, R.A. (1930) Inverse Probability. Proceedings, Cambridge Philosophical Society. 26, 528-535.

Fisher, R.A. (1934) Two New Properties of Mathematical Likelihood. Proceedings of the Royal Society. A, 144, 285-307.

Fisher, R.A. (1958) Statistical Methods and Scientific Inference, Oliver and Boyd, Edinburgh.

Fraser, D.A.S. (1961) The Fiducial Method and Invariance. Biometrika. 48, 261-280.

Fraser, D.A.S. (1964) Fiducial Inference for Location and Scale Parameters. Biometrika. 51, 17-24.

Gupta, A.K. (1952) Estimation of the Mean and Standard Deviation of a Normal Population from a Censored Sample. Biometrika. 39, 260-273.

Halperin, M. (1952) Maximum Likelihood Estimation in Truncated Samples. Annals of Mathematical Statistics. 23, 226-238.

Halperin, M. (1961) Confidence Intervals from Censored Samples. Annals of Mathematical Statistics. 32, 828-837.

Hamming, R.W. (1962) Numerical Methods for Scientists and Engineers, McGraw Hill, New York.

Harter, H.L. and Moore, A.H. (1966) Iterative Maximum Likelihood Estimation of the Parameters of Normal Populations from Singly and Doubly Censored Samples. Biometrika. 53, 205-213.

Hora, R.B. and Buehler, R.J. (1966) Fiducial Theory and Invariant Estimation. Annals of Mathematical Statistics. 37, 643-656.

Lindley, D.V. (1965) Introduction to Probability and Statistics, Cambridge University Press, Cambridge.

Lindley, D.V. (1958) Fiducial Distribution and Bayes Theorem. Journal of the Royal Statistical Society, Series B. 20, 102-107.

Lloyd, E.H. 1965) Least Squares Estimation of Location and Scale Parameters using Order Statistics. Biometrika. 39, 88-95.

Pierce, D.A. (1973) Fiducial, Frequency, and Bayesian Inference on Reliability for the Two Parameter Negative Exponential Distribution. Technometrics. 15, 249-253.

Plackett, R.L. (1957) Linear Estimation from Censored Data. Annals of Mathematical Statistics. 29, 131-142.

Rao, U.V.R., Savage, I.R. and Sobel, M. (1960) Contributions to the Theory of Rank Order Statistics: The Two-Sample Censored Case. Annals of Mathematical Statistics. 31, 415-426.

Sarhan, A.E. and Greenbert, B.G. (1956) Estimation of Location and Scale Parameters by Order Statistics from Singly and Doubly Censored Samples. Annals of Mathematical Statistics. 25, 317-328.

Savage, L.J., et al. (1962) The Foundation of Statistical Inference, Methuen London.

Saw, J.G. (1959) Estimation of the Normal Population Parameters Given a Singly Censored Sample. Biometrika. 46, 150-159.

Stroud, A.H. and Secrest, D. (1966) Gaussian Quadrature Formulas. Prentice Hall Inc. Englewood Cliffs, New Jersey.

Tiku, M.L. (1967) Estimating the Mean and Standard Deviation from a Censored Normal Sample. Biometrika. 54, 155-165.

Welch, B.L. and Peers, H.W. (1963) On Formulae for Confidence Points Based on Integrals of Weighted Likelihoods. Journal of the Royal Statistical Society, Series B. 25, 318-329.

APPENDIX

Computer Program

The computer program is written in Fortran IV which allows it to be easily converted to any of a number of scientific computers. The main points to the program are; the calculation of the posterior distribution; the calculation of the Maximum likelihood estimates and the approximate distribution thereof; the calculation of the best linear estimates and the approximate distribution thereof; the calculation of 99% and 95% confidence intervals for each of the above.

In order to run the program the following files need to be read by the program; from LUN 12 the coefficients for the Gaussian numerical integration; from LUN 13, the covariance matrix for normal order statistics of sample size N; from LUN 14 the expected values of the normal order statistics for sample size N.

The program as written is an interactive program requiring input from a data terminal. Before using the program, one should consult the author on the necessary input from a terminal. The output is written on several files as follows; LUN 15 is the approximate distribution of the M.L.E.; LUN 16 contains the posterior distribution; LUN 17 has the approximate distribution of the B.L.E. The confidence limits and posterior content thereof are written on LUNs 30 through 34. LUN 30 contains the posterior confidence limits, LUN 31 confidence limits using the M.L.E. LUN 33 confidence intervals using the B.L.E. LUN's 32 and 34 contain the posterior content of the intervals written on LUN 31 and 33 respectively.

```
      PROGRAM QUANTILES
        DIMENSION A(20),H(20),F(20),SIGMA(20),G(20)
      COMMON T1,T2,Z(81),PAZ(81),AU(81),HAP,XBAR,NZ,KCP,K2,ZOR(3),SS
     1,X(100),AA(50,50),BB(50,2),ALL
      IRAND=56321
      WRITE(61,1001)
1001    FORMAT(1H ,≠INSERT SIZE OF INTEGRATION-TABLE 20,15,10,5,3≠)
      MNK=TTYIN(4HSIZE)
      DO 3 I=1,MNK
        READ(12,12)A(I),H(I)
3     CONTINUE
      NZ=TTYIN(4H N= )
      K2=TTYIN(4HK2= )
      K1=TTYIN(4HK1= )
      KT=NZ/2
      T1=TTYIN(4HT95=)
      T2=TTYIN(4HT99=)
      IF(NZ.GT.20)GO TO 10
      DO 9 I=1,KT
      READ(13,13)(AA(I,J),J=1,NZ)
      K=NZ-I+1
      DO 9 J=1,NZ
      LX=NZ-J+1
9     AA(K,LX)=AA(I,J)
      GO TO 21
10    DO 22 I=1,NZ
      AA(I,I)=1.0
22    CONTINUE
21    CONTINUE
13    FORMAT(F6.5)
      DO 70 I=1,NZ
      READ(14,14)(BB(I,J),J=1,2)
70     CONTINUE
14    FORMAT(F1.0,2X,F8.5)
12    FORMAT(E18.6,3X,E16.6)
      KP=0
      KCP=K1-1
      G1=0.$SS=0.
      WRITE(61,172)
172   FORMAT(1H ,≠FOR SIMULATION ENTER 1,IF NOT ENTER 0≠)
      DAT=TTYIN(4H ^  )
      IF(DAT.EQ.1)GO TO 5
      WRITE(61,173)
173   FORMAT(1H ,≠DATA FILE MUST BE EQUIPED TO LUN 25≠)
174    FORMAT(F13.6)
      DO 63 I=1,NZ
      READ(25,174)X(I)
63    CONTINUE
      GO TO 7
5     WRITE(61,1200)
1200   FORMAT(1H ,≠INSERT NUMBER OF SAMPLES YOU WISH TO GENERATE≠)
      NG=TTYIN(4HSAMP)
      DO 180 NIG=1,NG
      DO 6 I=1,NZ
      ALPHA=0.
      DO 32 J=1,12
```

```
      ZB=UNIFORM(IRAND)
32    ALPHA=ALPHA+ZB
      X(I)=ALPHA-6.
6     CONTINUE
7     CONTINUE
      DO 41I=1,NZ
      DO 40 J=I,NZ
      IF(X(I).LE.X(J))40,31
31    AX=X(I)$D=X(J)
      X(I)=D$X(J)=AX
40    CONTINUE
41    CONTINUE
      SCALE=0.
      SS=0.
      KK=(K2-KCP)
      ASUM=XBAR=0.
      DO 101 M=K1,K2
      ASUM=ASUM+X(M)
101   SS=SS+X(M)*X(M)
      D=KK
      XBAR=ASUM/D
      JN=NZ/2
      ZOR(1)=X(K1)
      ZOR(2)=X(JN)
      ZOR(3)=X(K2)
      CALL MLE(VARUP,UHAT,UL,ALL,US)
      HAP=(UL-ALL)/80
      DO 102 M=K1,K2
102   SCALE=SCALE+(X(M)-UHAT)**2
      SCALE=SCALE/(2.*VARUP)
      SCALE=A(MNK)/SCALE
      XI1=ALL
       SS1=SS2=0.
      DO 170 NI=1,81
      SS1=KK*(XBAR-XI1)
      SS2=(SS-2*KK*XBAR*XI1+KK*XI1*XI1)
      GAM=0.
      ZINT=0.
      DO 60 N=1,MNK
      DORT=SQRT(SS2*.5/(A(N)/SCALE))
      GAM=(KP*SS1/DORT)
      F(N)=(EXP(-GAM))*((A(N)/SCALE)**(KK/2-1))
      EZ=((X(K2)-XI1)/DORT)+KP
      EX=(X(K1)-XI1)/DORT+KP
      IF(K2.EQ.NZ.AND.K1.EQ.0)GO TO 57
      F(N)=F(N)*((1.-CDFN(EZ))**(NZ-K2))*(CDFN(EX)**KCP)
57    F(N)=F(N)*EXP(A(N)*(1-1/SCALE))
      F(N)=(1/SCALE)*F(N)
      AZG=A(N)/SCALE
      G(N)=(EXP(-A(N)))*F(N)
      ZINT=F(N)*H(N)+ZINT
60    CONTINUE
      Z(NI)=((1/SS2)**(KK/2))*ZINT
      AU(NI)=XI1
      XI1=XI1+HAP
170   CONTINUE
```

```
      CALL SIMPSON(ZAP)
      DO 179 NI=1,81
      Z(NI)=ZAP*Z(NI)
      PAZ(NI)=PAZ(NI)*ZAP
      WRITE(16,190)AU(NI),Z(NI)
190   FORMAT(1X,F12.6,F13.10,F13.10)
179   CONTINUE
108   FORMAT(1X,4F10.4)
      AL99=UHAT-T2*US
      AL95=UHAT-T1*US
      UL95=UHAT+T1*US
      UL99=UHAT+T2*US
      P1=PCONT(AL99)
      P2=PCONT(AL95)
      P3=PCONT(UL95)
      P4=PCONT(UL99)
      WRITE(32,108)P1,P2,P3,P4
      WRITE(31,108)AL99,AL95,UL95,UL99
      CALL MEDIAN
      CALL LSQ
180   CONTINUE
      STOP
      END
      FUNCTION CDFN(X)
      COMMON/DATA/A(6),D(6),C(6,10)
      DATA(A=.625,1.25,2.0,2.45,3.5,4.62)
     1,(D=.3,.925,1.625,2.225,2.95,4.15)
      DATA(C= 6.7982403291E-04,-1.2709753598E-03, 6.7964525797E-04,
     1       -8.8500314300E-05, 1.4791554152E-07, 5.4879072878E-07,
     2        5.1028640888E-03,-1.4822649744E-03,-3.4589883732E-04,
     3        5.6826070991E-04,-6.1965013125E-05,-8.6656125340E-07,
     4       -6.0160862379E-03, 8.7718965600E-03,-3.0883401043E-03,
     5       -4.5015531032E-04, 2.2266710841E-04, 6.6877769441E-07,
     6       -2.2725611373E-02,-2.0200415601E-03, 7.2141082950E-03)
      DATA(C(22)=-1.3872389725E-03,-4.4390018986E-04,-6.4084653798E-07,
     8        3.3555772127E-02,-3.4681674488E-02,-1.0677039440E-03,
     9        5.5422185916E-03, 6.2978750848E-04, 6.2471782478E-07,
     1        7.2703311820E-02, 4.7642403207E-02,-2.4859750804E-02,
     2       -1.0221129286E-02,-6.6382630882E-04,-4.2105004094E-07,
     3       -1.4093895854E-01, 5.6850293056E-02, 5.7420315819E-02,
     4        1.1850195831E-02, 5.1265678121E-04, 2.0742133302E-07)
      DATA(C(43)=-1.5468912970E-01,-2.2180615138E-01,-6.5383705170E-02,
     6       -8.8864030980E-03,-2.7657162532E-04,-7.6856040220E-08,
     7        5.1563045460E-01, 2.3979043073E-01, 4.0236129479E-02,
     8        3.9938885701E-03, 9.3751413490E-05, 1.8722020822E-08,
     9        1.6431337971E-01, 4.0458836515E-01, 4.8922186659E-01,
     X        4.9917416726E-01, 4.9998489849E-01, 4.9999999779E-01)
      Y=X*0.7071067812
      SGNY=1.
      IF(Y)2,1,3
    1 CDFN=.5
      RETURN
    2 SGNY=-1.
      Y=-Y
    3 DO 4 I=1,6
      IF(Y.LE.A(I)) GO TO 5
```

```
    4 CONTINUE
      Z=.5
       GO TO 7
    5 Y=Y-D(I)
      Z=C(I,1)
      DO 6 J=2,10
    6 Z=Z*Y+C(I,J)
    7 CDFN=.5+SGNY*Z
      RETURN
      END
      FUNCTION UNIFORM(IRAND)
      IRAND=AND(AND(4099*IRAND,37777777B)+1220519,37777777B)
      UNIFORM=IRAND/8388607.
      RETURN
      END
C*****************************************************************CSIM00
      SUBROUTINE MEDIAN
      DIMENSION BZ(4)
      COMMON T1,T2, PAZ(81),AU(81),Z(81),HAP
      COMMON/DATA/V(4)
      DATA(V=.005,.025,.975,.995)
      I1=3
      DO 20 J=1,4
      DO 10 I=I1,80,2
      IF(AU(I).LT.V(J))GO TO 10
      GO TO 15
10    CONTINUE
      WRITE(61,1001)V(J)
1001  FORMAT(1X,≠CANNOT FIND A VALUE FOR ≠,F6.3)
15    IK=I1=I
      A=(AU(IK+2)-2.*AU(IK)+AU(IK-2))/2.
      B=(-AU(IK+2)+4*AU(IK)-3*AU(IK-2))/2.
      C=(AU(IK-2)-V(J))
      CON=B*B-4*A*C
      ALPHA=0.
      IF(CON.GT.0.)ALPHA=SQRT(CON)
      BZ(J)=Z(IK-2)+2.*HAP*(-B+ALPHA)/(2*A)
20    CONTINUE
      AL1=BZ(1)$AL2=BZ(2)$UL1=BZ(3)$UL2=BZ(4)
      WRITE(30,5)AL1,AL2,UL1,UL2
5     FORMAT(1X,4F10.4)
      RETURN
      END
      SUBROUTINE SIMPSON(ZAF)
      COMMON T1,T2,Z(81),PAZ(81),AU(81),H
      AREA=0.
      SUM=0.
      ZAP=0
      PAZ(1)=PAZ(2)=0.
      DO 40 I=2,80,2
      J=I-1$K=I+1
      AREA=(H/3.)*(Z(J)+4*Z(I)+Z(K))+AREA
       PAZ(K)=AREA
  40  CONTINUE
      ZAP=1/AREA
      RETURN
```

```
      END
      SUBROUTINE MLE(VARUP,UHAT,UL,ALL,US)
      COMMON T1,T2,Z(81),PAZ(81),AU(81),HAP,XBAR,N,K1,K2,ZORK(3),SS
      KK=K2-K1
      U=ZORK(2)
       SIG=(ZORK(3)-ZORK(1))/3
      DO 55 MNI=1,6
      Z1=(ZORK(1)-U)/SIG $Z2=(ZORK(3)-U)/SIG
      SQ=SIG**2 $SS2=(SS-2*KK*XBAR*U+KK*U*U)
      SSX=KK*(XBAR-U)
      A=PDFN(Z1)/CDFN(Z1)
      B=PDFN(Z2)/(1-CDFN(Z2))
CCCCCCC FIRST DERIVATIVES
      UPRM=SSX/SQ-K1*A/SIG +(N-K2)*B/SIG
      SPRM=(1/SIG)*(-KK+(1/SQ)*SS2-K1*A*Z1+(N-K2)*B*Z2)
CCCCCCC SECOND DERIVATIVES
      T=-(1/SQ)*(-KK-K1*A*(Z1+A)+(N-K2)*B*(Z2-B))
      R=-(1/SQ)*(KK-3*SS2/SQ-K1*A*(Z1*Z1+Z1*A-2)*Z1+
     1(N-K2)*Z2*B*(Z2*Z2-Z2*B-2))
      S=-1/SQ*(-2*SSX/SIG-K1*A*(Z1**2+Z1*A-1)+(N-K2)*B*(Z2
     1**2-Z2*B-1))
CCCCCCCSOLUTIONS TO MLE
      U1=U
      U=U1+(1/(R*T-S**2))*(R*UPRM-S*SPRM)
       S1=SIG
      SIG=S1+(1/(R*T-S**2))*(T*SPRM-S*UPRM)
      IF(SIG.LE.0)SIG=S1/2.0
55     CONTINUE
      UHAT=U
      VA=(1/(T*R-S**2))*T
      SES=SQRT(VA)
      VARUP=(SIG**2)*EXP(-6*SES/SIG)
32    FORMAT(1H ,≠LOWER LIMIT ON VARIANCE=≠,F12.6)
      US=(1/(R*T-S**2))*R
      IZ=4
      IF(K2-K1.LE.13)IZ=8
      UL=UHAT+IZ*SQRT(US)
      ALL=UHAT-IZ*SQRT(US)
      AXE=ALL
       DO 33  I=1,81
      XYZ=(AXE-UHAT)/SQRT(US)
      FGZ=PDFN(XYZ)
      WRITE(15,41)AXE,FGZ
33    AXE=AXE+HAP
41    FORMAT(1X,F12.6,3X,F15.6)
      US=SQRT(US)$VA=SQRT(VA)
      AXL=ALL
      HAP=(UL-AXL)/80.
      RETURN
      END
      FUNCTION PDFN(X)
      PDFN=1/SQRT(6.2831853)*EXP(-.5*X*X)
      RETURN
      END
      FUNCTION PCONT(ELIM)
      COMMON T1,T2,Z(81),AL(81),PAZ(81),HAP
```

```
      I1=2
      IF(AU(I1).EQ.0.)I1=I1+1
      DO 10 I=I1,80,2
      IF(ELIM.LT.PAZ(I))GO TO 7
10    CONTINUE
 7     IK=I1=I
      A=(AU(IK+2)-2.*AU(IK)+AU(IK-2))/2.
      B=(-AU(IK+2)+4*AU(IK)-3*AU(IK-2))/2.
      C=AU(IK-2)
      X=(ELIM-PAZ(IK-2))/(2*HAP)
      PCONT=A*(X**2)+B*X+C
      RETURN
      END
      SUBROUTINE LSQ
      DIMENSION R(40),T(40),L(20),M(20),S(400)
      DIMENSION W(40),Y(4),LL(2),MM(2),Z(40),U(2),COV(4)
      DIMENSION CX(20)
      COMMON T1,T2,ZZ(81),PAZ(81),AU(81),HAP,XBAR,N,KCP,K2,ZOR(3),SS
     1,X(100),A(50,50), B(50,2),ALL
      K1=KCP+1
      KT=N/2
CCCCCCC   ELIMINATION OF ROWS AND COLUMNS
      K=0
      KK=K2-K1+1$KQ=KK**2
      DO 17 J=K1,K2
      DO 17 I=K1,K2
      K=K+1
17    S(K)=A(I,J)
      K=0
      DO 8 J=1,2
      DO 8 I=K1,K2
      K=K+1
8     T(K)=B(I,J)
      CALL MTRA(T,R,KK,2,0)
      CALL MINV(S,KK,D,L,M)
      CALL MPRD(R,S,W,2,KK,0,0,KK)
      CALL MPRD(W,T,Y,2,KK,0,0,2)
      CALL MINV(Y,2,D1,LL,MM)
      CALL MPRD(Y,W,Z,2,2,0,0,KK)
      J=1
      DO 200 I=K1,K2
      CX(J)=X(I)
      J=J+1
200   CONTINUE
      CALL MPRD(Z,CX,U,2,KK,0,0,1)
      WRITE(61,30)U
30    FORMAT(1X,≠MEAN=≠,F12.6,≠ST.DEV.=≠,F12.6)
      V=U(2)**2
      CALL SMPY(Y,V,COV,2,2,0)
CCCCCC  TO SET CONFIDENCE LIMITS ASSUME NORMAL DISTRIBUTION
        STG=SQRT(COV(1))
       ADE=ALL
       DO 31 I=1,81
      XYZ=(ADE-U(I))/SIG
       GLB=PDFN(XYZ)
       WRITE(17,42)ADE,GLB
```