Winter 1986

Estimation in Truncated Exponential Family of Distributions

Laxman M. Hegde

Old Dominion University

ESTIMATION IN TRUNCATED EXPONENTIAL FAMILY
OF DISTRIBUTIONS

By

Laxman M. Hegde
M.Sc. 1975, Karnatak University, India
B.Sc. 1973, Karnatak College, India

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY
Computational and Applied Mathematics

OLD DOMINION UNIVERSITY
December, 1986

Approved by:

Ram C. Dahiya (Director)

ABSTRACT

ESTIMATION IN TRUNCATED EXPONENTIAL FAMILY

OF DISTRIBUTIONS

LAXMAN M. HEGDE

Old Dominion University, 1986

Director: Dr. Ram C. Dahiya

Estimating the parameters of a truncated distribution is a well known problem in statistical inference. The non-existence of the maximum likelihood estimator (m.l.e.) with positive probability in certain truncated distributions is not well known. To mention a few results in the literature:

i) Deemer and Votaw [1955] show that the maximum likelihood estimator does not exist in a truncated negative exponential distribution on [0,T], T>0 known, whenever the sample mean $\bar{x} \geq T/2$.

ii) Broeder [1955] shows that the maximum likelihood estimator of the scale parameter of a truncated gamma distribution, with the shape parameter being known, becomes infinite with positive probability whenever the sample mean $\bar{x} \geq \alpha/\alpha+1$, $\alpha > 0$.

iii) Mittal [1984] derives a sufficient condition for the non-existence of the maximum likelihood estimator

in a two parameter doubly truncated normal distribution on [A,B], A < B known. The m.l.e.'s become infinite whenever the sample variance exceeds $(B-A)^2/12$.

iv) Barndoff-Neilsen [1978] (BN) gives a set of general conditions for the existence and uniqueness of a solution to the maximum likelihood equations in a minimal representation of a k-parameter exponential family which depend upon a few results from convex analysis.

Using certain results from BN [1978], we give a unified approach to the problem of maximum likelihood estimation in the two parameter doubly truncated normal, truncated gamma, and singly truncated normal families, and obtain a set of necessary and sufficient conditions in terms of observable sample quantities. This approach basically depends upon characterizing the population and the sample moment spaces using a monotonicity property of the moments.

We also study the Bayes modal estimator introduced by Blumenthal and Marcus [1975] and the harmonic mean estimator introduced by Joe and Reid [1984]. We present certain computational results for solving the maximum likelihood equations in the above families. Simulation results for the probability of non-existence of the m.l.e., for the bias vector, and for the mean square error of the Bayes modal, the harmonic mean and the mixed estimator are presented.

to

my parents

ACKNOWLEDGEMENTS

This dissertation is a result of the inspiring guidance of Professor Ram Dahiya. I am sincerely thankful to him for directing this research work. My special thanks to Dr. N.R. Chaganty who has been my Guru throughout the closure of my graduate work. I am also thankful to Dr. John Morgan for giving me useful suggestions whenever I approached him. I am thankful to Dr. Wilkie Chaffin for being my committee member. I also wish to thank Barbara Jeffrey for typing the dissertation.

Last but not the least, I am grateful to my wife, Annapoorna and my son Chinmay for providing the family support which in fact was a necessary condition for the successful completion of my work.

TABLE OF CONTENTS

Chapter PAGE

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Research articles on inferential procedures based on truncated and censored samples appear in the statistical literature right from the period of R.A. Fisher. There are a number of examples, particularly in life testing and reliability theory, wherein truncation and censoring of observations occur naturally. Typically in life testing problems, n items are put on test for a fixed period of time T and inferences are made on the basis of the random number n of failures during time T under some distributional assumption for failure times (usually the negative exponential, gamma, Weibull or log-normal). Blumenthal and Marcus [1975] consider the above situation in which a particular defect is identifiable only after the item fails and study the problem of estimating N, the total number of items having the defect, under exponential failure times. There are other examples of truncation and censoring in quality control, biological studies, and genetic investigations. We present below a few examples.

i) Cohen [1950a] considers the estimation of parameters of a normal distribution based on truncated samples. He illustrates a doubly truncated normal sample with an example in which a certain bushing is sorted through go, no-go gauges, with the result that the items of diameter in excess of 0.6015 in. and those less than 0.5985 in, are discarded. This is an example in quality control.

ii) Deemer and Votaw [1955] give applications of a a truncated negative exponential distribution in bombing accuracy studies with an example of gun camera missions in which the view angle of the camera is restricted.

iii) The truncated Poisson typically arises when the zero class is unobservable, owing to the fact that only the cases where at least one event occurs are reported while the total number of cases are unknown. Blumenthal, Dahiya, and Gross [1973] consider estimating the complete sample size from an incomplete Poisson sample.

iv) Another interesting example in software reliability is studied by Joe and Reid [1985]. They consider estimating the unknown number of bugs in software based on truncated negative exponential failure times.

v) In general whenever a random variable assumes values over a finite interval, it would be reasonable to model the data by truncated distribution. Student GPA one such example.

In a survey article, Blumenthal [1981] clearly defines the problem of estimation of distributional parameters and the sample sizes from truncated/censored samples and lists a good number of references on this subject.

The censoring or truncation of a random variable X considered here means that there is a fixed region R where if X belongs to R, X is not observable. Though both the concepts have the same definition, they differ in practice as to how the scheme of obtaining a sample is carried out. In censored data, the total sample size and hence the number of missing observations is known, while in truncated data the number of missing observations is unknown. It is clear that in truncation problems, both the distributional parameters and the sample size may be of interest depending upon the situation. In some situations, only one of them is of concern and the other is incidental. Blumenthal [1981] gives an exhaustive account of the problem of estimating the unknown sample size. In estimating the distributional parameters, the conditional approach is followed. That is, the unknown total sample size is eliminated from consideration by assuming the number of observations to be fixed and then examining the conditional distribution of the given observatons, namely, the truncated distribution. We assume the truncation region to be known.

The important papers dealing with the estimation in truncated distributions belong to a general exponential family such as normal, gamma, log-normal and Poisson. A thorough survey of such articles may be found in the book by Johnson and Kotz (JK) [1980]. A partial review of some papers is done here in Chapter two. Derivation of the maximum likelihood estimators is emphasized in most of the articles. In general, the method of maximum likelihood estimation in a truncated exponential family leads to solving a system of non-linear equations. Solving the maximum likelihood equations iteratively is one of the main concerns in those research papers. The question of existence and uniqueness of the m.l.e. in truncated distributions, with all the parameters being unknown, needs further investigation. However in the simple cases of one parameter truncated negative exponential and truncated gamma (shape parameter known) distributions, non-existence of the m.l.e. is discussed by Deemer and Votaw [1955] and Broeder [1955] respectively. Mittal [1984] points out these facts and obtains a sufficient condition for non-existence of the m.l.e. in a doubly truncated normal distribution.

Barndorff-Neilsen (BN) [1978] studies maximum likelihood estimation in exponential families using the concepts of fullness, regularity, and steepness. As a particular case, BN [1978] shows the non-existence of maximum likelihood estimators in the case of a doubly truncated normal distribution. However, BN does not give conditions for the existence or non-existence of m.l.e.'s. Here, we follow BN's exponential theory [1978] and obtain a necessary and sufficient condition for the existence of a solution to the maximum likelihood equations in the cases of the doubly truncated normal, truncated gamma, and truncated log-normal families.

As an alternate method of estimating the parameters, Blumenthal and Marcus [1975] introduce the Bayes modal estimator, which is a mode of the posterior distribution of the parameter given a prior and the data. The Bayes modal estimators do not generally have the problem of non-existence and are asymptotically equivalent to the maximum likelihood estimators. Also the parameters of the conjugate prior distributions are selected to minimize the asymptotic bias. We discuss this estimator for the truncated gamma and singly truncated normal distributions.

Joe and Reid [1984] introduce the notion of likelihood intervals and develop a new estimator based on the harmonic mean of the end points of a likelihood interval. One of the ideas in this method is that the harmonic mean of infinity and a positive real number is again a positive real number. The derivation of the harmonic mean estimator requires unimodality, with a possible maximum at infinity, of the likelihood function which is true for the truncated distributions studied here. We discuss the performance of the harmonic mean estimator over the mixed and the Bayes modal estimators using mean square error and probability of nearness in the case of doubly truncated normal distribution.

The scheme of presentation of this dissertation is as follows:

i) Chapter two contains a review of the literature. Section 2.1 reviews exponential theory. Sections 2.2-2.4 consider estimation in truncated normal, gamma, and log-normal families. Finally Section 2.5a deals with an asymptotic theory of Bayes modal estimators and Section 2.5b introduces the harmonic mean estimator.

ii) Chapter three investigates a two parameter truncated normal distribution from the point of view of BN's exponential theory. Derivation of a necessary and sufficient condition for the existence of maximum likelihood estimators is discussed. The performance of the harmonic mean estimator in relation to the Bayes modal and mixed estimators is evaluated. Certain interesting computational results are also presented.

iii) Chapter four deals with a two parameter truncated gamma family. The question of existence of the m.l.e. and the Bayes modal estimator is discussed. Some computational results pertaining to solving the maximum likelihood equations are presented. Simulation results for the probability of non-existence of the m.l.e., the mean square error, the bias vector (the length and direction), and the probability of nearness are discussed.

iv) In Chapter five, the same problems considered in

Chapter three are studied for thesingly truncated normal distribution. Using the parabolic cylinder function, a recurrence relation for the moments of the singly truncated normal distribution is derived, which in turn simplifies computation of the roots of the maximum likelihood equations.

2. REVIEW

An exclusive review of some important results from BN's exponential theory is presented first in Section 2.1. In Sections 2.2-2.4, a brief discussion of the research work related to the maximum likelihood estimation in truncated normal, gamma, and log-normal families is done. In Section 2.5a, a review of the asymptotic theory of the Bayes modal estimation (Blumental [1982]) is made. Finally in Section 2.5b, the notion of harmonic mean estimation is presented.

2.1 Exponential Theory

We discuss the exponential theory, BN [1978], through the subsections a) - e). As we are concerned with only the two parameter family of densities, the general k-parameter results are summarized for k=2. In further discussions, the following notations are repeatedly used:

$$\underline{X} = (X_1,\ldots,X_n),\ \underline{x} = (x_1,\ldots,x_n),\ \omega = (\omega_1,\omega_2),$$

$$\theta = (\theta_1,\theta_2),\ t = (t_1,t_2),\ \bar{t} = (\bar{t}_1,\bar{t}_2),$$

$$T(\underline{X}) = (T(\underline{X}),T_2((\underline{X}))),\ \tau(\omega) = (\tau_1(\omega),\tau_2(\omega)),$$

$$T(\underline{x}) \cdot \tau(\omega) = T_1(\underline{x})\tau_1(\omega) + T_2(\underline{x})\tau_2(\omega).$$

$$\theta \cdot t = \theta_1 t_1 + \theta_2 t_2, \text{ and } dt = dt_1\, dt_2.$$

a) Formulation of an Exponential Family

Let $\mathscr{B} = [L(\underline{x};\omega) : \underline{x} \in \mathfrak{X} ; \omega \in \Omega]$ be an indexed family of density functions, where $\mathfrak{X}$ is a sample space in R^n, and Ω is an indexing parameter space in R^2. In the subsequent discussions, the indexed family represents a family of joint distributions of n i.i.d. random variables of inferential interest. Then $\mathscr{B}$ is said to be an exponential family provided

$$L(\underline{x};\omega) = a(\omega)b(\underline{x})e^{\tau(\omega)\cdot T(\underline{x})} , \qquad (2.1)$$

where

$\tau(W):\Omega\to\Theta$, $\Theta=[(\theta_1,\theta_2):\theta_1=\tau_1(\omega_1,\omega_2),\ \theta_2 = \tau_2(\omega_1,\omega_2),\ \omega \in \Omega]$,
$T(\underline{x}):\mathfrak{X} \to S$, $S = [(t_1,t_2):t_1 = T_1(\underline{x}),\ t_2 = T_2(\underline{x}),\ \underline{x} \in \mathfrak{X}]$,

and $a(\omega):\Omega \to R$, and $b(\underline{x}): \mathfrak{X} \to R$.

In the representation (2.1), $\tau(\omega)$ is known as a parameterization if and only if τ is a one-one mapping on Ω. Also $T(\underline{x})$ is known as a canonical statistic. If both $t = (t_1,t_2)$ and $\theta = (\theta_1,\theta_2)$ do not satisfy any linear constraint in R^2, then (2.1) is said to be minimal. We assume only the minimal representations. In applications, Ω is usually a parent parameter space of interest and Θ is a parameterization under the exponential representation.

For the purposes of discussing some intrinsic properties of an exponential family, we consider the family of distributions of $T(\underline{X})$ which is again an exponential family. It can be shown that distribution of $T(\underline{X})$ is of the form

$$g(t;\theta) = [c(\theta)]^{-1}h(t)e^{\theta \cdot t}, \tag{2.2}$$

where h(t) is a function of t only and $c(\theta)$ is given by

$$c(\theta) = \int\int h(t)e^{\theta \cdot t}dt.$$

Note that S is the support of $g(t;\theta)$. In (2.2), we may without loss of generality assume, (0,0) in Θ so that h(t) is a proper density function on S. That is, c(0,0)=1. We may rewrite (2.2) in a more convenient form as

$$g(t;\theta) = h(t)e^{\theta \cdot t - K(\theta)}, \tag{2.3}$$

where $K(\theta) = \log c(\theta)$.

In fact $c(\theta)$ is the moment generating function of h(t) and hence $K(\theta)$ is the cumulant function of h(t). It is a well known result that $K(\theta)$ is strictly convex and differentiable on the interior of Θ. Note that $K(\theta)$ plays an important role in discussing the maximum likelihood estimation in an exponential family.

Let

$$\mathscr{A} = [g(t;\theta):\theta \in \Theta,\ g(t;\theta) \text{ as in } (2.3)]. \tag{2.4}$$

Let $\bar{\Theta} = [(\theta_1,\theta_2):\ \int\int h(t)e^{\theta\cdot t}dt < \infty]$. In other words $\bar{\Theta}$ is the set of all possible values of $\theta \in R^2$ for which $g(t;\theta)$ is proper density on S. In the classical exponential theory, $\bar{\Theta}$ is known as the natural parameter space. We also use $\bar{\Theta}$ to mean "the natural parameter space" and Θ to mean "the parameter space of interest". Next, we state some important properties of $\mathscr{A}$.

b) Some Properties of $\mathscr{A}$

In this subsection, we discuss the fullness, the regularity and the steepness of an exponential family $\mathscr{A}$. Also we discuss a monotonicity property of the derivative of $K(\theta)$.

i) <u>Full</u>: An exponential family $\mathscr{A}$ is said to be full if and only if $\Theta = \bar{\Theta}$.

ii) <u>Regular</u>: An exponential family $\mathscr{A}$ is said to be regular if and only if $\Theta = \bar{\Theta}$ and Θ is open in R^2.

iii) <u>Steep</u>: Let Θ^* be the boundary of Θ, Θ° the interior of Θ and $DK(\theta) = (\partial K/\partial\theta_1,\ \partial K/\partial\theta_2)$. Here, $DK(\theta)$ is the partial derivative vector of $K(\theta)$ for θ in Θ°. Then $\mathscr{A}$ is said to be steep whenever $K(\theta)$ is steep. $K(\theta)$ is said to be steep if and only if

$$(\theta-\theta^*)\cdot DK(\lambda\theta + (1-\lambda)\theta^*) \rightarrow \infty, \text{ as } \lambda \longrightarrow 0, \qquad (2.5)$$

for every θ in Θ° and θ^* in Θ^*. The product $(\theta-\theta^*)\cdot DK(\lambda\theta + (i-\lambda)\theta^*)$ in (2.5) be interpreted as the inner product for a given λ. The following Theorem 2.1 proved in BN [1978] states an important result for a regular exponential family.

<u>Theorem 2.1</u>: Let $\mathscr{A}$ be an exponential family as defined in (2.4). If $\mathscr{A}$ is regular, then it is steep.
Note that the converse of Theorem 2.1 is not true (see BN [1978], example 8.4).

iv) <u>Monotonicity of DK(θ)</u>: Let DK(θ) be the partial derivative vector of K(θ) and

$$I = [(m_1,m_2) : m_1 = \frac{\partial K(\theta)}{\partial\theta_1}, m_2 = \frac{\partial K(\theta)}{\partial\theta_2}, \theta \in \Theta^\circ]. \qquad (2.6)$$

Theorem 2.2 below states a monotonicity property of DK(θ).

<u>Theorem 2.2</u>: Let I be as defined in (2.6). Then DK(θ) is a continously differentiable mapping between the two open sets Θ° and I. Furthermore, DK(θ) is strictly increasing in the sense that

$$(\theta_1 - \tilde{\theta}_1)(m_1-\tilde{m}_1) + (\theta_2-\tilde{\theta}_2)(m_2-\tilde{m}_2) > 0, \qquad (2.7)$$

where (m_1,m_2) and $(\tilde{m}_1,\tilde{m}_2)$ are values of $DK(\theta)$ at θ and $\tilde{\theta}$ respectively. Theorem 2.2, implies that the boundary of I occurs on the boundary of Θ. Next, we state a key result of this section given in BN [1978], page 142.

<u>Theorem 2.3</u>: Let C° be the interior of the convex hull of the support of $g(t;\theta)$, where $g(t;\theta)$ is as in (2.3). Then $I=C^{\circ}$ if and only if $K(\theta)$ is steep.

Theorem 2.3 has direct implications on the problem of maximum likelihood estimation in the exponential family.

c) Maximum Likelihood Estimation

A set of general conditions for the uniqueness and the existence of maximum likelihood estimator are presented here. For details, one may refer to BN [1978] Sections 9.5 and 9.6.

By the definition (2.1) of an exponential family, it is clear that the log-likelihood function is

$$l(\theta;t) = \theta\cdot t - K(\theta) + \log h(t).$$

Now for a given t, the maximum likelihood estimator, $\hat{\theta}$, is defined to be the set of $\theta \in \Theta^{\circ}$ which maximizes $l(\theta;t)$. That is,

$$\hat{\theta} = [\theta \in \Theta^{\circ} : l(\theta;t) = \sup(l(\theta;t))]. \qquad (2.8)$$

In Theorems 2.4a and 2.4b below, the existence and the uniqueness properties of $\hat{\theta}$ are stated, for the cases when the family is full and not full repsectively.

Theorem 2.4a: Let $\hat{\theta}$ be the maximum likelihood estimator of θ as defined in (2.8) in a full exponential family. If $K(\theta)$ is steep, then the following holds:

i) $\hat{\theta}$ exists if and only if $t \in C^{\circ}$.

ii) $\hat{\theta}$ is one-one and single valued.

iii) Whenever $\hat{\theta}$ exists, it is the unique solution of $DK(\theta) = t$, for θ in Θ°.

Note that Theorem 2.4a assumes the fullness of a family. If a family is not full, then we actually deal with a proper subfamily. The problem of maximum likelihood estimation in an arbitrary subfamily of a full exponential family requires an advanced theory of convex analysis. However, if the parameter space of interest is an open convex set, BN [1978] shows that a result similar to Theorem 2.4a can be stated for such subfamilies.

Let $\mathscr{A}_1 = [g(t;\theta) : \theta \in \Theta_1]$, $\Theta_1 \subset \overline{\Theta}$ and Θ_1 is an open convex set and

$$I_1 = [(m_1, m_2) : m_1 = \frac{\partial K(\theta)}{\partial \theta_1}, \; m_2 = \frac{\partial K(\theta)}{\partial \theta_2}, \text{ and } \theta \epsilon \Theta_1]. \qquad (2.9)$$

Note that I_1, in (2.9), is the range of $(\frac{1}{n})DK(\theta)$ with the restricted domain Θ_1. The subscrpits are used to denote that these results are applicable only in a subfamily.

Theorem 2.4b: Let $\hat{\theta}$ be the maximum likelihood estimator in the subfamily $\mathscr{A}_1$. Then the following holds:

i) $\hat{\theta}$ exists if and only if $t \in I_1$, I_1 defined in (2.9).

ii) $\hat{\theta}$ is one-one and single valued.

iii) Whenever $\hat{\theta}$ exists, it is the unique solution of the equation $DK(\theta) = t$, $\theta \in \Theta_1$.

Whenever we restrict ourselves to a proper subfamily in the sense mentioned earlier, it is possible that the maximum likelihood equation may not admit a solution, because $I_1 \subset I$ and I may be equal to C° while I_1 may not.

d) Some Special Results

BN [1978] makes two interesting comments (without proof) on a truncated exponential family which we think worth mentioning here in view of its relevence to our problem. Also, a short proof of the same is given here.

These two results are presented through Theorems 2.5a and 2.5b.

<u>Theorem 2.5a</u>: Let $\mathscr{A}$ be an exponential family as defined in (2.4), with S as the support of $g(t;\theta)$. If $\mathscr{A}$ is regular, then every truncation is regular.

Proof: Let S' be the support of the minimal canonical statistic T in the truncated family generated by $\mathscr{A}$. Obviously $S' \subset S$ due to truncation. Hence the result.

<u>Theorem 2.5b</u>: Let $\mathscr{A}$ be an exponential family as defined in (2.4) with S as the support of $g(t;\theta)$. If S is a bounded set in R^2, then the natural parameter space is the whole of R^2. That is, $\mathscr{A}$ is regular.

Proof: Note that the natural parameter space $\bar{\Theta}$ is

$$\bar{\Theta} = [\theta \in R^2 \colon \iint h(t)e^{\theta \cdot t}dt < \infty].$$

Since S is a bounded set, $e^{\theta \cdot t} < M$ (M finite) and h(t) is a density function on S, we have

$$\iint h(t)e^{\theta t}dt < \infty \quad \text{for all } \theta \in R^2.$$

Hence the result.

e) Some Remarks

As a concluding part of the review of the exponential theory discussed above, some relevant points are presented in this section as guidelines while dealing with the maximum likelihood estimation in an exponential family.

i) Given any indexed family of distribution, check whether it is representable in some minimal exponential family.

ii) Let Θ be a parameter space of interest and $\bar{\Theta}$ be the natural parameter space of the minimal representation in (i). Check whether the minimal representation is regular. That is, Θ is open and $\Theta = \bar{\Theta}$. We know, due to Theorem 2.1, that the regularity implies the steepness. Hence in a regular family, the probem of non-existence of the maximum likelihood estimators does not arise.

iii) If the minimal family in (i) is not regular, then either $\Theta \subset \bar{\Theta}$ or Θ is a closed set. If $\Theta \subset \bar{\Theta}$, then the family is not full and Theorem 2.4b is applicable.

iv) If in (iii), $\Theta = \bar{\Theta}$ but Θ is a closed set, then we deal with a full family and Theorem 2.4a is applicable. That is, the equation $DK(\theta) = t$, $\theta \in \Theta$ and $t \in C^{\circ}$, admits a solution if and only if $K(\theta)$ is steep. Therefore we are required to check the steepness of $K(\theta)$ in order to determine the existence of a solution.

As a special remark, the case of a truncated exponential family being not regular may often be due to the parameter space of interest being a proper subset of the natural parameter space. This may happen because the support of T is a bounded set and hence the natural parameter space is the whole of R^2 (Theorem 2.5b) while the parameter space of interest is a proper subset of R^2. Also the natural parameter space associated with the untruncated family may be extendable due to truncation while the parameter space of interest is the same both in the truncated and the untruncated families. In any case, Theorem 2.4b is more likely to be referred to in the maximum likelihood estimation problems of a truncated exponential family.

2.2 Doubly Truncated Normal Distribution

A brief review of the problem of estimating the parameters of doubly truncated normal distribution is presented here.

In terms of the notations of Section 2.1, the indexed family of doubly truncated normal distributions is

$$\mathscr{B}_1=[L(\underline{x};\mu,\sigma^2) \;:\; \underline{x}=(x_1 \ldots X_n) \in \mathfrak{X} \;,\; (\mu,\sigma^2)\in\Omega],$$

$$\mathfrak{X} = [\underline{x}:A\le x_i\le B \;,\; i=1,2,\ldots,n,\; A<B \text{ known}],$$

$$\Omega = [(\mu,\sigma^2) \;:\; -\infty<\mu<\infty,\; 0<\sigma^2<\infty],$$

and

$$L(\underline{x};\mu,\sigma^2) = \frac{e^{-\frac{1}{2}\sum_{i=1}^{n}\left[\frac{x_i-\mu}{\sigma}\right]^2}}{\left[\int_A^B e^{-\frac{1}{2}\left[\frac{u-\mu}{\sigma}\right]^2} du\right]^n} . \qquad (2.10)$$

Here, with $L(\underline{x};\mu,\sigma^2)$ as in (2.10) is an exponential family. The details of the representation are shown in Chapter three. It is also shown that the distribution of the canonical statistic $T(\underline{X}) = (\Sigma X_i, \Sigma X_i^2)$ with the parameterization $\theta_1 = \frac{\mu}{\sigma^2}$, $\theta_2 = 1 - \frac{1}{2\sigma^2}$ is of the form

$$g(t;\theta) = h(t)e^{\theta \cdot t - K(\theta)}, \; t \in S, \; \theta \in \Theta, \qquad (2.11)$$

where $K(\theta) = \log c(\theta)$, $c(\theta) = \left[\int_A^B e^{\theta_1 u + (\theta_2 - 1)u^2} du\right]^n$,

$\Theta = [(\theta_1,\theta_2): -\infty < \theta_1 < \infty, -\infty < \theta_2 < 1]$,

$\overline{\Theta} = [(\theta_1,\theta_2): -\infty < \theta_1 < \infty, -\infty < \theta_2 < \infty]$,

and $S = [(t_1,t_2): t_1 = \Sigma x_i, t_2 = \Sigma x_i^2, (x_1 \ldots x_n) \in \mathfrak{X}]$.

In (2.11), Θ is the parameter space of interest and $\overline{\Theta}$ is the natural parameter space.

There are many research articles, listed by JK [1970], on the estimation problems in a truncated normal distribution. We discuss here the work of Cohen [1950a], Crain [1979], Shah and Jaiswal [1964], Bar-lev [1984], and BN [1978]. The work of Cohen seems to be one of the most pioneering and is frequently referred to by many authors. Mittal [1984] also studies this problem in detail and gives a good review of the work of Cohen and Crain in particular. Cohen, Shah and Jaiswal, and Mittal study with respect to an indexed family as in (2.10) while Crain, Bar-lev, and BN study with respect to exponential representation similar to (2.11).

Cohen [1950a] deals with estimating (μ,σ^2) by the method of maximum likelihood estimation which is the same as the method of moments. It is known that the maximum likelihood equations are complicated non-linear equations in functions of (μ,σ^2). Cohen simplifies these equations from a computational point of view, by computing the cumulative distribution function of the complete normal distribution. The maximum likelihood estimates are obtained by Newton-Raphson method iteratively. In a subsequent paper, Cohen [1957] pursues the work done in Cohen [1950a] with a slight modification of the likelihood equations. The new equations as reviewed by Mittal [1984] are given below:

$$[(h_0-z_1)/(z_2-z_1)] - [\bar{x}-A)/(B-A)] = 0$$

$$[(1+h_1-h_0^2)/(z_2-z_1)^2] - [s^2/(B-A)^2] = 0 \qquad (2.12)$$

where $z_1=(A-\mu)/\sigma$, $z_2=(B-\mu)/\sigma$, $h_1= \dfrac{\left[z_1 e^{-z_1^2/2} - z_2 e^{-z_2^2/2}\right]}{\Phi(z_2) - \Phi(z_1)}$

$h_0 = \dfrac{e^{-z_1^2/2} - e^{-z_2^2/2}}{\Phi(z_2) - \Phi(z_1)}$, $s^2 = \Sigma(x_i - \bar{x})^2/n$, $\bar{x} = \Sigma x_i/n$,

and Φ is the cumulative distribution function of the complete normal distrtibution.

Given the sample quantities $s^2/(B-A)^2$ and $(\bar{x}-A)/(B-A)$, Cohen uses a graphical method of obtaining quick estimates of the roots of equations (2.12). However the graphical chart for reading the estimates of (μ,σ^2) does not contain any information whenever $s^2 > (B-A)^2/12$. Cohen [1957] does not address the problem of non-existence at all. Mittal [1984] mentions this fact and obtains a sufficient condition that the maximum likelihood equations in the indexed family do not admit a solution whenever the sample variance is greater than $(B-A)^2/12$.

Crain [1979] studies this problem in the family $\mathscr{B}_1$ using a minimal exponential representation after transforming $L(\underline{x};\mu,\sigma^2)$ as in (2.10) to $L(\underline{y};\mu,\sigma^2)$ with

$$Y=[2(X-A)/(B-A)]-1,$$

so that the truncation interval [A,B] transforms into [-1,1]. Crain [1979] uses a complicated exponential representation of (2.10) with the canonical statistics

$$T_1(\underline{Y}) = (3/2)^{1/2}\Sigma Y_i, \quad T_2(\underline{Y}) = (45/8)^{1/2}\Sigma Y_i^2 - (5n/8)$$

and the parameterization

$$\frac{-1}{2\sigma^2} = (45/8)^{1/2}(2/B-A)^2\theta_2,$$

$$\frac{\mu}{\sigma^2} = (3/2)^{1/2}(2/B-A)\theta_1-(45/8)^{1/2}\ \frac{4(A+B)\theta_2}{(B-A)^2}\ .$$

Using BN's exponential theory, without mentioning the applicable results, Crain [1979] proves the existence of maximum likelihood estimators of (θ_1,θ_2) with probability one under the assumption of the natural parameter space (R^2). Finally, Crain (1979) erroneously concludes the same result for the (μ,σ^2) parameterization also.

As can be seen in (2.11) that the natural parameter space is the whole of R^2 and hence family is regular implying the existence of maximum likelihood estimators with probability one. However the important point Crain [1979] misses is that the problem of maximum likelihood estimation in a truncated normal distribution with the parameter space of interest Ω, comes under the purview of Theorem 2.4b of Chapter two. Hence, Crain's conclusion about (μ,σ^2) is not valid.

Bar-lev [1984] deals with only the asymptotic properties of the maximum likelihood estimator (cmle) in a full truncated exponential family with the truncation region unknown. Shah and Jaiswal [1964] try to estimate (μ, σ^2) using three moment equations involving the first four moments. It is interesting to note that Shah and Jaiswal approach does not require solving any non-linear equations.

From the review, it is observed that the problem of non-existence of maximum likelihood estimators in $\mathscr{B}_1$ needs further investigation. Mittal [1984] only obrtains a sufficient condition for the non-existence of m.l.e. BN [1978] does not investigate any condition in terms of the sample quantities for the existence of m.l.e. In this dissertation, we study the same problem and obtain a necessary and sufficient condition for the existence of m.l.e.'s. It is important to note that a necessary and sufficient condition is useful in developing some improved estimators of which mixing of the m.l.e. with the Bayes modal estimator is one proposed by Mital [1984]. Also the harmonic mean estimator studied here depends on a necessary and sufficient condition.

2.3 Truncated Gamma Distribution.

The truncated gamma distribution is another important distribution of interest due to its applications in life testing and reliability problems. We discuss a few research papers in this family. The indexed family of truncated gamma distributions is

$\mathcal{B}_2 = [L(\underline{x};\beta,\alpha):\underline{x} = (x_1 \ldots x_n) \epsilon\ \mathfrak{X},\ (\beta,\alpha) \epsilon\ \Omega],$

$\mathfrak{X} = [\underline{x}:\ 0 \le x_i \le U,\ i=1,2,\ldots,n\ ,\ U > 0 \text{ known}],$

$\Omega = [(\beta,\alpha):\ 0 < \beta < \infty,\ 0 < \alpha < \infty],$

and

$$L(\underline{x};\beta,\alpha) = \frac{e^{-\frac{1}{\beta}\Sigma x_i + (\alpha-1)\Sigma \log x_i}}{\left[\int_0^U e^{-\frac{1}{\beta}u + (\alpha-1)\log u} du\right]^n} . \tag{2.13}$$

Note that (2.13) is an exponential family (details are shown in Chapter four). The distribution of the canonical statistic $T(\underline{X}) = (T_1(\underline{X}) = \Sigma X_i, T_2(\underline{X}) = \Sigma \log X_i)$ with the parameterization $\theta_1 = 1-U/\beta$, and $\theta_2 = \alpha-1$ is of the form

$$g(t;\theta) = h(t)e^{\theta \cdot t - K(\theta)},\ t \ \epsilon\ S,\ \theta \ \epsilon\ \Theta, \tag{2.14}$$

where $K(\theta) = \log c(\theta)$ and $c(\theta)$ is given by

$$c(\theta) = \left[\int_0^1 e^{(\theta_1 - 1)u + \theta_2 \log u}\right]^n ,$$

with

$$\Theta = [(\theta_1,\theta_2):\ -\infty < \theta_1 < 1,\ -1 < \theta_2 < \infty],$$

$$\bar{\Theta} = [(\theta_1,\theta_2): -\infty < \theta_1 < \infty,\ -1 < \theta_2 < \infty],$$

and $S = [(t_1,t_2): t_1 = \Sigma x_i,\ t_2 = \Sigma \log x_i, (x_1 \ldots x_n) \in \mathcal{X}]$.

In (2.14), Θ is the parameter space of interest and $\bar{\Theta}$ is the natural parameter space of interest.

Broeder [1955], Chapman [1956], and Gross [1971] deal with the estimation in the truncated gamma distribution. Deemer and Votaw [1955] study a truncated negative exponential distribution which is a special case of the truncated gamma family with $\alpha = 1$. Cohen [1950b], and Des Raj [1953] deal with estimating the parameters of the truncated Pearson Type III distributions of which the truncated gamma is a member.

Broeder [1955] considers the maximum likelihood estimator of β, with α-known, in the truncated gamma distribution with the density function

$$f(x;\beta,\alpha) = \frac{e^{-\beta x} x^{\alpha-1}}{\left[\int_0^1 e^{-\beta u} u^{\alpha-1} du\right]},$$

where $0 \le x \le 1$, $\beta > 0$ and $\alpha > 0$.

The maximum likelihood estimator of β is the solution of $F(\beta) = 0$, where

$$F(\beta) = \frac{\int_0^1 x^{\alpha} e^{-\beta x} dx}{\int_0^1 x^{\alpha-1} e^{-\beta x} dx} - \bar{x} . \tag{2.15}$$

As Broeder [1955] mentions, it is easy to see that the range of $F(\beta)$ is $(0, \alpha/\alpha+1)$ while the range of $\bar{x}$ is $[0,1]$. Hence (2.15) does not admit a solution whenever $\bar{x} \geq \alpha/\alpha+1$. The same result also indicates the non-existence of m.l.e. in a truncated negative exponential distribution (with $\alpha=1$). Deemer and Votaw (DV) [1955] consider the problem of estimation in a negative exponential deistribution and define the maximum likelihood estimator of β as zero whenever $\bar{x} \geq 1/2$ which is undesirable since zero does not belong to the parameter space of interest.

Gross [1971] proves a monotonicity of r^{th} moment, about the origin, of the truncated gamma distribution. Note that monotonicity of the first moment of X and log X follows from the monotonicity result (2.7) stated in Section 2.1. Des Raj [1953] and Cohen [1950b] deal with obtaining a solution to the maximum likelihood equations in the truncated pearson Type III distribution. They do not investigate the question of existence/non-existence of m.l.e. Chapman [1956] proposes a method of estimating the parameters of a truncated gamma distribution with the density

$$f(x;\beta,\alpha) = [H(\beta,\alpha)]^{-1}e^{-\beta x}x^{\alpha-1} ,$$

where $0\le x\le U$, $U>0$ is known, and $H(\beta,\alpha) = \int_0^U e^{-\beta x}x^{\alpha-1}dx$.

The method proposed by Chapman [1956] requires subdividing the truncation interval $[0,U]$ into r subintervals $[\xi_i \pm h_i]$, with $i=1,2,\ldots,r$ and $\xi_1-h_1 = 0$, $\xi_r-h_r = U$. Let v_i be the frequency of the original data in the i^{th} class interval. Then $f(x;\beta,\alpha)$ is approximated in $[\xi_i \pm h_i]$ by p_i, $i=1,2,\ldots,r$, where

$$p_i = e^{-\beta\xi_i}[\xi_i]^{\alpha-1}(2h_i).$$

Consider $q_i = v_i/n$ (n is sample size) as an estimate of p_i. It is easy to see that

$$\ell n\, p_i - \ell n\, p_{i+1} = \beta(\xi_{i+1}-\xi_i)+(\alpha-1)\ell n(\xi_i/\xi_{i+1}) + \ell n(h_i/h_{i+1}). \tag{2.16}$$

Replacing p_i by q_i in (2.16) and defining $y_i=\ell n q_{i+1}-\ell n\, q_i$, we have a system of linear equations

$$y_i=\beta(\xi_{i+1}-\xi_i) + (\alpha-1)\ell n(\xi_i/\xi_{i+1}) + \ell n(h_i/h_{i+1}),$$

for $i=1,\ldots,r$. Hence Chapman [1956] proposes a least square estimator of (β,α) by minimizing the quadratic form

$$(Y-EY)'\Sigma_Y^{-1}(Y-EY)$$

where Σ_Y is the variance-covariance matrix of Y.

The method proposed by Chapman [1956] is an adhoc procedure of estimating the parameters (β,α). Using y_i in lieu of x_i causes a loss of the information in the original data. Also $H(\beta,\alpha)$ is completely ignored.

It is clear from the review that the estimation problem in the two-parameter truncated gamma family $(\mathscr{B}_2)$ is not studied in greater depth. However, Mittal [1984] addresses the problem with a partial treatment and leaves it open due to the intractability of the mathematics involved. In Chapter four of this dissertation, we study the two-parameter truncated gamma family with the representation (2.11) and obtain a necessary and sufficient condition for the existence of solution of the maximum likelihood equations. Solving the m.l. equations in the truncated gamma family is complicated (compared to the truncated normal) due to the singularity of the functions involved at $x=0$ whenever $0<\alpha<1$. A simple method for handling the same is shown in Chapter four.

2.4 Truncated Log-normal Distribution.

A singly truncated log-normal distribution is equivalent to a singly truncated normal and a doubly truncated log-normal is equivalent to a doubly truncated normal distribution. This may perhaps be one reason that we see a few research papers exclusively on the truncated log-normal distribution. We consider the indexed family of singly truncated log-normal distributions as

$$\mathcal{B}_3 = [L(\underline{y};\mu,\sigma^2):(y_1 \ldots y_n)\epsilon\, \bar{\underline{Y}},\ (\mu,\sigma^2)\epsilon\ \Omega],$$

$$\bar{\underline{Y}} = [\underline{y}:\ 0 \le y_i \le U,\ i = 1,2,\ldots,n,\ U > 0 \text{ known}],$$

$$\Omega = [(\mu,\sigma^2):\ -\infty < \mu < \infty,\ 0 < \sigma^2 < \infty],$$

and

$$L(\underline{y};\ \mu,\sigma^2) = \frac{e^{\frac{1}{2}\sum_{i=1}^{n}\left[\frac{\log y_i-\mu}{\sigma}\right]^2} \prod_{i=1}^{n}\left[\frac{1}{y_i}\right]}{\left[\int_0^U \frac{1}{u}\, e^{-\frac{1}{2}\left[\frac{\log u-\mu}{\sigma}\right]^2} du\right]^n} \ . \qquad (2.17)$$

By changing X = log Y , it is obvious that X has the singly truncated normal distribution. We designate $\mathcal{B}_4$ as the family of singly truncated normal distributions. That is,

$$\mathcal{B}_4 = [L(\underline{x};\ \mu,\sigma^2):\ \underline{x} = (x_1,\ldots,x_n)\ \epsilon\, \bar{\underline{X}};\ (\mu,\sigma^2)\epsilon\ \Omega],$$

$$\bar{\underline{X}} = [\underline{x}:\ -\infty < x_i \le \log U,\ i = 1,2,\ldots,n,\ U > 0 \text{ known}],$$

and

$$L(\underline{x};\ \mu,\sigma^2) = \frac{e^{-\frac{1}{2}\Sigma\left[\frac{x_i-\mu}{\sigma}\right]^2}}{\left[\int_{-\infty}^{\log U} e^{-\frac{1}{2}\left[\frac{u-\mu}{\sigma}\right]^2} du\right]^n} . \tag{2.18}$$

As in the case of doubly truncated normal, $\mathscr{B}_4$ is also an exponential family with the canonical statistic $T(\underline{X}) = (T_1(\underline{X}) = \Sigma X_i, T_2(\underline{X}) = \Sigma X_i^2)$ and the parameterization $\theta_1 = \frac{\mu}{\sigma^2}$, $\theta_2 = 1 - \frac{1}{2\sigma^2}$. Then the distribution of $T(\underline{X})$ is

$$g(t;\theta) = h(t)e^{\theta\cdot t-K(\theta)}, t \in S,\ \theta \in \Theta, \tag{2.19}$$

where $K(\theta) = \log c(\theta)$, $c(\theta) = \left[\int_{-\infty}^{\log U} e^{\theta_1 u+(\theta_2-1)u^2} du\right]^n$,

$\Theta = [(\theta_1,\theta_2):\ -\infty < \theta_1 < \infty\ ,\ -\infty < \theta_2 < 1]$,

$\bar{\Theta} = [(\theta_1,\theta_2):-\infty<\theta_1\leq 0,-\infty<\theta_2<1]\cup[0<\theta_1<\infty,-\infty<\theta_2\leq 1]$,

$S = [(t_1,t_2):\ t_1 = \Sigma x_i, t_2 = \Sigma x_i^2, (x_1,\ldots,x_n)\in \mathfrak{X}]$.

In (2.19), Θ is the parameter space of interest and $\bar{\Theta}$ is the natural parameter space.

Cohen [1950a, 1957], and Thompson [1951] are the two prominent authors who deal with estimation in a singly truncated normal distribution. Cohen deals with the single truncation as a special case of the double truncation with the right truncation equal to infinity. Though the title of the paper is "Truncated Log-normal Distributions", Thompson [1951] presents the main results for the singly truncated normal distribution only. Thompson considers the density function of the singly truncated normal as

$$
\begin{aligned}
f(x;u,\sigma^2) &= 0, \quad x < A, \\
&= \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left[\frac{u-\mu}{\sigma}\right]^2} du\,, \quad x = A, \\
&= \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}, \qquad x > A, \qquad (2.20)
\end{aligned}
$$

where $A > -\infty$ is known.

Thompson assumes positive mass at $x = A$ and considers estimating (μ,σ^2) based on the moment equations, the moments being about the truncation point $x = A$. It is easy to see that the moment equations are

$$
\begin{aligned}
\mu_1'(A) &= (z-uq)\sigma = (\bar{x}-A)/n = \eta_1 \\
\mu_2'(A) &= (q-uz + u^2q)\sigma^2 = \Sigma\ (x_i-A)^2/n = \eta_2 \qquad (2.21)
\end{aligned}
$$

where $u = (A-\mu)/\sigma$, $z = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$, $q = \int_u^\infty z\,du$.

Define,

$$\phi_2 = [\mu_1'(A)]^2/[\mu_2'(A)], \text{ and } \phi_1 = z-uq.$$

It is obvious that ϕ_2 is independent of σ^2. Hence a table for ϕ_1 and ϕ_2 is given for different values of u in [-4,2] with an increment of 0.1. Given the sample quantity η_1^2/η_2, the value of u and hence ϕ_1 can be read from the table. With such estimates of u and ϕ_1, it is easy to see that $\hat{\sigma} = \eta_1/\phi_1$, and $\hat{\mu} = A - \hat{\sigma}u$. Note that the estimates $(\hat{\mu},\hat{\sigma})$ are the only rough estimates which can be obtained quickly. With the assumption of positive mass at $x = A$, in the definition of truncated density, solution to moment equations (2.21) always exists. However such a definition of giving positive mass at a point of truncation is not considered in our study. In fact (2.20) does not belong to $\mathscr{B}_4$. If we take the usual definition as in (2.18), the Thompson method does not always work. It is easy to see that ϕ_1 and ϕ_2 with respect to (2.18) are

$$\phi_2 = (z-uq)^2/(q-uz + u^2q)q \quad \text{and} \quad \phi_1 = z-uq.$$

If we compute ϕ_2 now for u in $(-\infty,\infty)$, we observe that $\phi_2>1/2$. Hence if a sample quantity of η_1^2/η_2 is less than 0.5, which can occur with positive probability, we cannot obtain u from the table and hence (μ,σ^2) cannot be estimated.

Cohen [1950a, 1957] with the definition (2.18), proposes a similar estimating procedure by defining

$$\phi_2' = (\mu_2'(A))/\ (\mu_1'(A))^2\ ,\ \text{and}\ \phi_1' = z - uq, \tag{2.22}$$

which is known as the Pearson-Lee-Fisher equation. Note that ϕ_2' is the reciprocal of ϕ_2. It is obvious that $\phi_2'<2$. Hence it is clear that the singly truncated normal distribution does have the problem of non-existence.

In Chapter five, we study the singly truncated log-normal distribution, equivalently the singly truncated normal distribution, using the representation (2.19) and prove the non-existence of a solution to the maximum likelihood equations. In fact, the result that $\phi_2' < 2$ follows from the result $E(X^2) < 2(E(X))^2$ which we prove in Theorem 5.1 of Chapter five. We also derive an easily computable recurrence relation for the moments of the singly truncated normal distribution (2.19) which simplifies a computational work.

2.5a Bayes Modal Estimation

From the review of the estimation in truncated normal (Section 2.2) and truncated gamma (Section 2.3), we know that the maximum likelihood estimator has the problem of non-existence. Blumenthal and Marcus [1975] encounter a similar problem while dealing with the estimation of distributional parameter (θ) and total sample size (N) based on a sample from truncated negative exponential distribution. They resolve this problem by introducing a new method of estimation called "Bayes Modal Estimation" which is based on the principle of maximizing the posterior distribution of the parameter vector, given a sample and a prior distribution.

Subsequent to the paper by Blumenthal and Marcus [1975], we see the Bayes modal estimation being employed in the work of Blumenthal [1981], Blumenthal, Dahiya, and Gross [1978], and Mittal [1984] in a variety of situations. Mittal [1984], studies the behavior of Bayes modal estimators in relation to the maximum likelihood and the mixed estimators in the cases of the doubly truncated normal, one parameter truncated gamma, and the truncated Weibull distributions. We discuss the Bayes modal estimators in the cases of two parameter truncated gamma and two parameter singly truncated normal distributions.

It is clear that choosing a proper prior distribution and the optimum parameters of the prior distribution is an important part of the Bayes modal estimation. Usually the conjugate prior distribution is assumed to be a realistic choice. Blumenthal [1984] outlines a method of choosing the parameters of a prior distribution. This method depends upon minimizing the asymptotic bias by assuming the stochastic expansion of the type

$$\tilde{\theta} = \theta + A/\sqrt{n} + b/n + 0(n^{-3/2}) \ , \qquad (2.23)$$

where A, B are polynomials in sums of i.i.d. random variabales and θ is a scalar parameter. We now give a brief summary of the work of Blumenthal [1982] and utilize these results in Chapters four and five.

Let $\underset{\sim}{x} = (x_1,\ldots,x_n)$ be i.i.d. observations with the density function $f(x;\theta)$. Then the modified likelihood of the sample $\underset{\sim}{x} = (x_1,\ldots,x_n)$ is

$$L^*(\underset{\sim}{x};\theta) = h(\theta) \prod_{i=1}^{n} f(x;\theta),$$

where $h(\theta)$ is some appropriate prior distribution. Then the Bayes modal estimator $(\tilde{\theta})$ is a solution of

$$0 = \frac{\partial \log L^*(\underset{\sim}{x};\theta)}{\partial \theta} = \sum S(x_p;\theta) + \xi(\theta), \qquad (2.24)$$

where

$$S(x;\theta) = \frac{\partial \log f(x;\theta)}{\partial \theta} = \frac{f'(x;\theta)}{f(x;\theta)} \ ,$$

and

$$\xi(\theta) = \frac{\partial \log h(\theta)}{\partial \theta} = h'(\theta)/h(\theta).$$

Note that here the prime notation denotes the differentiation with respect to θ. Sufficient regularity conditions like differentiation under the integral sign, and measuribility conditions are assumed in further discussions. Let $i,j,k,\ell=0,1,2,\ldots$ and

$$L_{ijk\ell} = \int S_p^i (S_p')^j (S_p'')^k (S_p''')^\ell f(x;\theta)dx,$$

$$Z_{ijk\ell} = \frac{\Sigma\, S_p^i (S_p')^j\, (S_p'')^k\, (S_p''')^\ell - n\, L_{ijk\ell}}{\sqrt{nV_{ijk\ell}}},$$

$$V_{ijk\ell} = L_{2i,2j,2k,2\ell} - (L_{ijk\ell})^2,$$

and $S_p = S(x_p;\theta)$.

Generally, we encounter $i,j,k,\ell=0,1$. In the above notations, trailing zeros are omitted. That is, $L_{ij00} = L_{ij}$.

Expanding $S(x;\tilde{\theta})$ and $\xi(\tilde{\theta})$ in Taylor series around θ in (2.24), using (2.23), leads to

$$0 = \sqrt{n}\,(Z_1\sqrt{L_2} - A\, L_2)$$
$$+ (-BL_2 + A\, Z_{01}\sqrt{V_{01}} + (A^2 L_{001}/2) + \xi(\theta)$$
$$+ o(n^{-1/2}). \tag{2.25}$$

Equating to zero the coefficients of $\sqrt{n}$ and the constant term in (2.25) and solving for A, and B, we get

$A = Z_1/(\sqrt{L_2})$, and

$$B = L_2^{-2}[Z_1 Z_{01}\sqrt{V_{01}L2} + (Z_1^2 L_{001}/2) + L_2\xi(\theta)]. \qquad (2.26)$$

From (2.23) and (2.26), we can see that

$$E(\tilde{\theta}) = \theta + E(B)/n + O(n^{-3/2}), \qquad (2.27)$$

where E(B) is regarded as the asymptotic bias of $\tilde{\theta}$ and it is easy to see that

$$E(B) = L_2^{-2}[L_{11} + (L_{001}/2) + L_2\xi(\theta)]. \qquad (2.28)$$

We may express L's with the more fundamental quantities

$$\mu_{ijk} = \int (f'/f)^i \, (f''/f)^j \, (f'''/f)^k \, f \, dx,$$

where i,j,k = 0,1,2,... Hence (2.28) can be rewritten in terms of μ's as

$$E(B) = \mu_2^{-2}\left[-(\mu_{11}/2) + \mu_2\xi(\theta)\right], \qquad (2.29)$$

by noting $L_2=\mu_2$, $L_{01}= -\mu_2$, $L_{11}=\mu_{11}-\mu_3$, $L_{001}=-3\mu_{11}+2\mu_3$. We use (2.29) in determining the optimum parameters of a prior distribution in a single parameter family. In other words, the optimum parameters of a prior are chosen such that E(B) is minimum.

The above analysis is applicable to one parameter families only. Here we deal with two parameter truncated normal and gamma families. Note that we modify the likelihood function with respect to only one parameter by properly chosen prior while using a noninformative prior for the other. In this way we may still use the above analysis in finding the optimum parameters of a prior distribution. We mainly present a proof of the existence of the Bayes modal estimator for the truncated gamma and singly truncated normal distributions in Chapters four and five respectively. Also we present simulation results for comparing the Bayes modal and the mixed estimators.

2.5b Harmonic Mean Estimation

Joe and Reid [1984] consider estimating the number of faults (N) of a reliability system under a truncated negative exponential failure times which is similar to the problem studied by Blumenthal and Marcus [1975]. However in the paper of Joe and Reid [1984], blowing up of the maximum likelihood estimator N is remedied by the harmonic mean estimator which always exists and has finite expectation. The harmonic mean estimator can be studied only in the particular situations wherein the parameter of interest is non-negative and the likelihood function is unimodal, possibly maximum occuring at infinity. The concepts of deriving the harmonic mean estimator is explained next in one parameter family with the density $f(x;\theta)$.

Let $\underline{x} = (x_1 \ldots x_n)$ be a random sample with the common density $f(x;\theta)$ where θ is a scalar parameter. Then the likelihood functin of $\underline{x}$ is

$$L(\underline{x};\theta) = \prod_{i=1}^{n} f(\underline{x};\theta).$$

Let $\hat{\theta}$ be the unique maximum likelihood estimator of θ. Then assuming that $L(\underline{x};\theta)$ is finite, define

$$\lambda(\underline{x};\theta) = L(\underline{x};\theta)/L(\underline{x};\hat{\theta}).$$

Given c in (0,1), the likelihood interval is defined to be

$$I_c = \left[\theta_c^{(1)}, \theta_c^{(2)}\right],$$

where $\theta_c^{(1)}=\inf[\theta:\lambda(\underline{x};\theta)\geq c]$, and $\theta_c^{(2)}=\sup[\theta:\lambda(\underline{x};\theta)\geq c]$. We call c the likelihood coefficient in further references. Then the harmonic mean estimator $\hat{\theta}_c$ is defined to be the harmonic mean of endpoints of I_c. That is,

$$\hat{\theta}_c = \frac{2\theta_c^{(1)} \cdot \theta_c^{(2)}}{\theta_c^{(1)} + \theta_c^{(2)}}, \quad \text{if } \hat{\theta} < \infty, \tag{2.30}$$

$$= 2\,\theta_c^{(1)} \qquad \text{if } \hat{\theta} = \infty.$$

The subscript c denotes the dependency on c. In simulation studies, we may choose any c in (0,1) and generate a sequence of harmonic mean estimators. In fact, as $c \to 1$, $\hat{\theta}_c \to \hat{\theta}$. The key idea of using the harmonic estimator is that $\hat{\theta}_c < \infty$ even if $\hat{\theta} = \infty$.

Note that the above discussion is valid in one parameter families only. Though we are dealing with two parameter families, we are mainly interested in modifying one of the parameters which is found to have the problem of non-existence as mentioned in the review of the Bayes modal estimators. Hence we reduce our case to one parameter family by considering $L(\underline{x};\hat{\theta}_1,\theta_2)$ where $\hat{\theta}_1$ is the m.l.e. of θ_1. In other words, we study the harmonic mean estimator of one of the parameters while keeping the m.l.e. of the other. As mentioned in Section 2.2, we study the harmonic mean estimator of θ_2 in the doubly truncated distribution while keeping the maximum likelihood estimator of θ_1.

3. DOUBLY TRUNCATED NORMAL DISTRIBUTION

We discuss here the two parameter doubly truncated normal distribution. The results are mainly presented in an exponential family representation. We deal with the family representation in Section 3.1, derivation of the maximum likelihood equations and the question of existence of m.l.e. in Section 3.2, computational results in Section 3.3, and simulation results in Section 3.4.

3.1 Family Representation

A doubly truncated normal population is described by a random variable with density

$$f(y;\mu,\sigma^2) = \frac{e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}}{\left[\int_A^B e^{-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2} du\right]}, \tag{3.1}$$

where $A \le y \le B$, $A < B$ known, and $(\mu,\sigma^2) \in \Omega$ with

$$\Omega = [(\mu,\sigma^2): -\infty < \mu < \infty,\ 0 < \sigma^2 < \infty].$$

Since the truncation interval $[A,B]$ is assumed to be known, we may change $X=2[(Y-A)/(B-A)] - 1$ so that the truncation interval becomes $[-1,1]$. In other words, (3.1) may be rewritten as

$$f(x;\mu,\sigma^2) = \frac{e^{\frac{c_1(\mu-c_2)}{\sigma^2} x - \frac{c_1^2}{2\sigma^2} x^2}}{\left[\int_{-1}^{1} e^{\frac{c_1(\mu-c_2)u}{\sigma^2} - \frac{c_1^2 u^2}{2\sigma^2}} du\right]}, \tag{3.2}$$

where $c_1 = (B-A)2$, $c_2 = (A+B)/2$.

Here, (μ,σ^2) is the parameter vector of interest, given a sample of n i.i.d. observations from (3.2). Hence the indexed family of truncated normal distribution

$\mathscr{B}_1 = [L(\underline{x};\mu,\sigma^2)$: $\underline{x} = (x_1,\ldots,x_n) \in \mathfrak{X}$, $(\mu,\sigma^2)\in\Omega]$,

$\mathfrak{X} = [\underline{x}$: $-1 \le x_i \le 1$, $i = 1,2,\ldots n]$,

and

$$L(\underline{x};\mu,\sigma^2) = \frac{e^{\frac{c_1(\mu-c_2)}{\sigma^2}\Sigma x_i - \frac{c_1^2}{2\sigma^2}\Sigma x_i^2}}{\left[\int_{-1}^{1} e^{\frac{c_1(\mu-c_2)u}{\sigma^2} - \frac{c_1^2}{2\sigma^2}u^2} du\right]^n}. \tag{3.3}$$

In (3.3), $L(x;\mu,\sigma^2)$ is known as the likelihood of a sample. Now it is easy to see that $\mathscr{B}_1$ is an exponential family since $L(\underline{x};\mu,\sigma^2)$ can be represented as in the definition (2.1) of Chapter two with

$$a(\mu,\sigma^2):\Omega\longrightarrow R, a(\mu,\sigma^2)=\left[\int_{-1}^{1} e^{\tau_1(\mu,\sigma^2)u+(\tau_2(\mu,\sigma^2)-1)u^2} du\right]^{-n},$$

$$b(\underline{x}):\ \mathfrak{X}\longrightarrow R, \quad b(\underline{x}) = e^{-\Sigma x_i^2},$$

$T(\underline{x}):\mathfrak{X}\to S$, $S=[(t_1,t_2):t_1=\Sigma\, x_i,\ t_2=\Sigma\, x_i^2,(x_1\ldots x_n)\in\mathfrak{X}]$,

$\tau(\mu,\sigma^2):\Omega\to\Theta, \Theta=[(\theta_1,\theta_2):\theta_1=\tau_1(\mu,\sigma^2),\theta_2=\tau_2(\mu,\sigma^2),(\mu,\sigma^2)\in\Omega]$

and $\tau_1(\mu,\sigma^2) = \frac{c_1(\mu-c_2)}{\sigma^2}$, and $\tau_2(\mu,\sigma^2) = 1 - \frac{c_1^2}{2\sigma^2}$.

Note that τ is one-one on Ω with the inverse mapping

$$\mu = (c_1\theta_1/2(1-\theta_2)) + c_2, \text{ and } \sigma^2 = c_1^2/2(1-\theta_2).$$

We may express the density function of X in the (θ_1,θ_2) parameterization as

$$f(x;\theta_1,\theta_2) = \frac{e^{\theta_1 x+(\theta_2-1)x^2}}{\left[\int_{-1}^{1} e^{\theta_1 u+(\theta_2-1)u^2} du\right]} . \tag{3.4}$$

The representation (3.4) is frequently referred to in further discussions. The distribution of $T(\underline{X})=(\Sigma X_i,\Sigma\, X_i^2)$ is of the form

$$g(t;\theta) = h(t)e^{\theta\cdot t - K(\theta)}, \quad t\in S,\ \theta\in\Theta. \tag{3.5}$$

In (3.5), $K(\theta) = \log c(\theta)$ and $c(\theta)$ is given by

$$c(\theta) = \left[\int_{-1}^{1} e^{\theta_1 u+(\theta_2-1)u^2} du\right]^n .$$

Since Θ is one-one mapping of Ω, it is easy to see that

$$\Theta = [(\theta_1,\theta_2):\ -\infty < \theta_1 < \infty\ ,\ -\infty < \theta_2 < 1].$$

We can easily see that the natural parameter space, $\bar{\Theta}$, is

R^2 since the integral

$$\left[\int_{-1}^{1} e^{\theta_1 u+(\theta_2-1)u^2} du\right]$$

is finite for all θ in R^2.

Since the parameter space of interest is a proper subset of the natural parameter space,

$$\mathscr{A} = [g(t;\theta) : \theta \in \Theta] \tag{3.6}$$

is not full. In (3.6), $\mathscr{A}$ is a proper subfamily of a full exponential family. Next, we discuss monotonicity property of $DK(\theta) = \left[\frac{\partial K(\theta)}{\partial \theta_1}, \frac{\partial K(\theta)}{\partial \theta_2}\right]$, the partial derivative vector of $K(\theta)$. By the definition of $K(\theta)$ in (3.5), it is easy to check that

$$\frac{\partial K(\theta)}{\partial \theta_1} = n\ E(X)\ ,\ \text{and}\ \frac{\partial K(\theta)}{\partial \theta_2} = n\ E(X^2),$$

where $E(X)$ and $E(X^2)$ are expected values of X and X^2 with respect to (3.4).

Let

$$I = [(m_1,m_2): m_1 = E(X), m_2 = E(X^2),\ \theta \in \Theta]. \tag{3.7}$$

In (3.7), I is the range of $(1/n)DK(\theta)$. It may be appropriate to call I the "population moment space". By the exponential theory as reviewed in Chapter two, it is clear that $(1/n)DK(\theta)$ satisfies the monotonicity result

(2.7). That is,

$$(\theta_1-\tilde{\theta}_1)(m_1-\tilde{m}_1) + (\theta_2-\tilde{\theta}_2)(m_2-\tilde{m}_2) > 0 \qquad (3.8)$$

where (m_1,m_2) and $(\tilde{\theta}_1,\tilde{\theta}_2)$ are values of $\frac{1}{n}DK(\theta)$ at θ and $\tilde{\theta}$ respectively. The said monotonicity result (3.8) also implies the following:

i) The strict monotonicity of $E(X)$ for a fixed θ_2.

ii) The strict monotonicity of $E(X^2)$ for a fixed θ_1.

(3.9)

Next, we discuss the maximum likelihood estimation in $\mathscr{A}$.

3.2 Maximum Likelihood Estimation

Given a sample quantity t, the log-likelihood function is

$$l(\theta;t) = \theta\cdot t - K(\theta) + \log h(t)$$

where $K(\theta)$ and $h(t)$ are as defined in (3.5). Hence by the definition (2.8), the maximum likelihood estimator, $\hat{\theta}$, is the solution of

$$DK(\theta) = t, \quad \theta \in \Theta.$$

Let $\bar{t}_1 = (\frac{1}{n})(t_1)$, $\bar{t}_2 = (\frac{1}{n})(\bar{t}_2)$, and

$$\bar{S} = [(\bar{t}_1,\bar{t}_2) : (t_1,t_2) \in S].$$

It may be appropriate to call $\bar{S}$ the "sample moment space" as a counterpart of the set I defined in (3.7). Hence the equation $DK(\theta) = t$ can equivalently be stated as

$$E(X) = \bar{t}_1,$$

$$E(X^2) = \bar{t}_2, \tag{3.10}$$

where $E(X) = \int_{-1}^{1} x\, f(x;\theta)$, $E(X^2) = \int_{-1}^{1} x^2\, f(x;\theta)$ and $f(x;\theta)$ is as defined in (3.4). In further discussion, we refer to (3.10) as the maximum likelihood equations. Now the question of uniqueness and existence of a solution to (3.10) is of concern to us and the same is discussed next.

Since we are dealing with a proper subfamily of a full exponential family and Θ is an open convex subset of $\bar{\Theta}$, Theorem 2.4b is applicable. That is,

i) $\hat{\theta}$ exists if and only if $\bar{t} \in I$.

ii) $\hat{\theta}$ is one-one and single valued.

iii) $\hat{\theta}$ is the unique solution of (3.10) whenever $\bar{t} \in I$.

Hence the existence question needs further investigation. The condition (i) of Theorem 2.4b requires determining the set I. That is, we must express the range of $E(X^2)$ as a function of the range of $E(X)$. This is done in Theorem 3.1.

<u>Theorem 3.1</u> Let $D_{m_1} = [(\theta_1,\theta_2) : E(X) = m_1]$, $-1 < m_1 < 1$, and Θ^* be the boundary of Θ. Then for (θ_1,θ_2) in D_{m_1}

$$m_1^2 < E(X^2) < 1 - (2/\theta_1^*)m_1, \tag{3.11}$$

where θ_1^* is such that $H(\theta_1{}^*) = m_1$ and

$$H(\theta_1^*) = \frac{e^{\theta_1^*}+e^{-\theta_1^*}}{e^{\theta_1^*}-e^{-\theta_1^*}} - \frac{1}{\theta_1^*} .$$

Proof: It is easy to see that

$\Theta^* = [(\theta_1^*,1) : -\infty < \theta_1^* < \infty]$.

Given $(\theta_1^*,1)$ in Θ^*, it can be shown that

$$\lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to 1}} E(X) = \lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to 1}} \int_{-1}^{1} x\, f(x;\theta)dx,$$

$$= H(\theta_1^*), \tag{3.12a}$$

and

$$\lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to 1}} E(X^2) = \lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to 1}} \int_{-1}^{1} x^2 f(x;\theta)dx,$$

$$= 1 - (2/\theta_1^*)H(\theta_1^*). \tag{3.12b}$$

where $f(x;\theta)$ is as defined in (3.4) and $H(\theta_1^*)$ as in (3.11). By elementary calculus, we can verify that

i) $H(\theta_1^*)$ is a strictly increasing function of θ_1^*,

ii) $\lim_{\theta_1^*\to\pm\infty} H(\theta_1^*) = \pm 1$,

iii) Continuity of $H(\theta_1^*)$ together with (ii) imply that $-1 < H(\theta_1^*) < 1$.

Due to i) - iii) above, given $-1 < m_1 < 1$, there exist a unique θ_1^* such that $H(\theta_1^*) = m_1$. In other words, $(\theta_1^*, 1)$ is the boundary point of D_{m_1}. Due to the continuity and the monotonicity property (3.8) and the limiting quantity (3.12b),

$$\sup_{D_{m_1}} E(X^2) = 1 - (2/\theta_1^*)m_1.$$

Note that $-\infty < \theta_2 < 1$ on D_{m_1}. The lower inequality, $E(X^2) > m_1^2$ follows from the well known result $E(X^2) > (E(X))^2$. Hence the result.

As a consequence of Theorem 4.1, we may write the set I explicitly as

$$I = [(m_1, m_2): -1 < m_1 < 1,\ m_1^2 < m_2 < 1 - (2/\theta_1^*)m_1], \quad (3.13)$$

where $m_1 = H(\theta_1^*)$.

Theorem 3.1 gives the required necessary and sufficient condition for the existence of a solution for the m.l. equations (3.10), and is stated in Corollary 3.1 below.

<u>Corollary 3.1</u>: Given $\bar{t} = (\bar{t}_1, \bar{t}_2)$, $-1 < \bar{t}_1 < 1$, the m.l. equations (3.10) admit a solution in θ if and only if

$$\bar{t}_1^2 < \bar{t}_2 < 1 - (2/\hat{\theta}_1^*)\bar{t}_1, \tag{3.14}$$

and $\hat{\theta}_1^*$ is the unique solution of $H(\theta_1^*) - \bar{t}_1 = 0$.

Proof: Given $(\bar{t}_1, \bar{t}_2)$, $-1<\bar{t}_1<1$, it is obvious that the solution to the equations

$$E(X) = \bar{t}_1$$
$$E(X^2) = \bar{t}_2$$

must lie in the set $D_{\bar{t}_1}$. By Theorem 3.1, $E(X^2)$ satisfies the inequality (3.11). Hence (3.14) follows.

It may be worth noting that the m.l. equations (3.10) in the regular family $\mathscr{A} = [g(t;\theta) : \theta \in \bar{\Theta} = R^2]$ admit the solution with probability one.

As a by-product result, we may answer the question of existence of the m.l.e.'s in the subfamilies

$$\mathscr{A}_{\theta_2} = [g(t;\theta) : -\infty < \theta_1 < \infty],\ \theta_2 < 1 \text{ is known,}$$

and

$\mathscr{A}_{\theta_1} = [g(t;\theta) : -\infty < \theta_2 < 1],\ |\theta_1| < \infty$ is known.

Note that $\mathscr{A}_{\theta_2}$ and $\mathscr{A}_{\theta_1}$ are not exactly the same as μ-known and σ^2-known with respect to the representation (3.1). However the cases of μ-known and σ^2-known can be dealt with similarly using a slightly modified canonical statistics. We discuss the m.l. equation only in $\mathscr{A}_{\theta_2}$ and $\mathscr{A}_{\theta_1}$ here. It is easy to see that the maximum likelihood estimator of θ_1 in $\mathscr{A}_{\theta_2}$ is the solution of

$$E(X) = \bar{t}_1, \tag{3.15}$$

and the maximum likelihood estimator of θ_2 in $\mathscr{A}_{\theta_1}$ is the solution of

$$E(X^2) = \bar{t}_2 \tag{3.16}$$

Due to the monotonicity results stated in (3.9), we have the following results:

i) $E(X)$ is strictly increasting in $\mathscr{A}_{\theta_2}$ with the range $(-1,1)$,

ii) $E(X^2)$ is strictly increasing in $\mathscr{A}_{\theta_1}$ with the range $(0,1-(2/\theta_1)H(\theta_1))$.

where $H(\theta_1)$ is the same as defined in (3.11). Hence

(3.15) always admit a solution while (3.16) does admit a solution if and only if $0 < \bar{t}_2 < 1 - (2/\theta_1)H(\theta_1)$. Note that if $\theta_1 = 0$ in $\mathscr{A}_{\theta_1}$, this states that (3.16) does not admit a solution whenever $\bar{t}_2 \geq 1/3$ since the limit of $H(\theta_1)/\theta_1$ is 1/3 as θ_1 goes to zero. These results are equivalent to the results obtained by Mittal [1984] in the case of a single parameter doubly truncated normal distribution.

Next we consider characterizing the set $\bar{S}$, similar to the set I in Theorem 3.2 below.

<u>Theorem 3.2</u>. Let $-1 \leq x_i \leq 1$, $i = 1,2,\ldots,n$ and $(\bar{t}_1, \bar{t}_2)$ be as defined earlier. Then for $\frac{2(k-1)}{n} - 1 \leq \bar{t}_1 \leq \frac{2k}{n} - 1$, $k = 1,2,\ldots,n$

$$\bar{t}_1^2 \leq \bar{t}_2 \leq \bar{t}_2^*, \tag{3.17}$$

where $\bar{t}_2^* = (n-1)/n + (1/n)[n\bar{t}_1 - \{2(k-1) - (n-1)\}]^2$.

Proof: Let $N_k = [\frac{2(k-1)}{n} - 1, \frac{2k}{n} - 1]$, $k = 1,2,\ldots,n$ be subintervals dividing $[-1,1]$. Let $\mathfrak{X} = [x_1,\ldots,x_n) : -1 \leq x_i \leq 1, i = 1,\ldots,n]$, and $\mathcal{X}_{\bar{t}_1} = [(x_1,\ldots,x_n)\epsilon\ \mathcal{X}: \frac{1}{n}\Sigma\ x_i = \bar{t}_1]$, $\bar{t}_1$ in N_k for some k. We know that $\mathcal{X}_{\bar{t}_1}$ is a hyperplane intersecting $\mathcal{X}$.

Since $\bar{t}_2 \geq \bar{t}_1^{\,2}$ is a well known result, we need to show only the upper inequality in (3.17). Now $\bar{t}_2 = \Sigma\ x_i^2$ is a convex function on $\mathcal{X}_{\bar{t}_1}$ and hence attains its maximum on the set of extreme points of $\mathcal{X}_{\bar{t}_i}$. The set of extreme point of $\mathcal{X}_{\bar{t}_1}$ consists of just the permutations of $(x_1..x_j,..x_n)$ where, for $k > 1$,

$$
\begin{aligned}
x_j &= 1, && \text{for } j = 1,2,\ldots,k-1,\\
&= n\,\bar{t}_1 - (k-1)+(n-k) && \text{for } j = k,\\
&= -1 && \text{for } j = k+1,\ldots,n.
\end{aligned}
$$

Since $\bar{t}_2$ is invariant under permutation, it attains the same maximum value for all the extreme points of $\mathcal{X}_{\bar{t}_1}$. Now it can be checked that

$$\underset{\mathcal{X}_{\bar{t}_1}}{\mathrm{Max}}\ \bar{t}_2 = \frac{(n-1)}{n} + \frac{1}{n}\,[n\bar{t}_1 - \{2(k-1)-(n-1)\}]^2 = \bar{t}_2{}^*.$$

Hence the result.

Using Theorem 3.2, we may express $\bar{S}$ explicitely as

$$\bar{S} = \bigcup_k [\bar{t}_1, \bar{t}_2) : \bar{t}_1 \in N_k \, , \, \bar{t}_1^2 \le \bar{t}_2 \le \bar{t}_2^*] \, . \qquad (3.18)$$

Let C denote the convex hull of $\bar{S}$. It is easy to check that

$$C = [(\bar{t}_1, \bar{t}_2) : -1 \le \bar{t}_1 \le 1 \, , \, \bar{t}_1^2 \le \bar{t}_2 \le 1].$$

In the general review of exponential theory, BN [1978] considers convex hull C of $\bar{S}$ for discussion of the existence and the uniqueness of maximum likelihood estimators. However for a given n, $\bar{S}$ being the support of an absolutely continuous function, the observed value of $\bar{t}$ belongs to $\bar{S}$ with probability one. Also it is interesting to note that $\bar{S}$ tends to C as n tends to ∞. Hence, the results discussed in BN [1978] hold for all n.

As a corollary to Theorem 3.2, we state a result for the sample variance below.

Corollary 3.2: Let $-1 \le x_i \le 1$, $i = 1,2,\ldots,n$ and $(\bar{t}_1, \bar{t}_2)$ be as defined in Theorem 3.2 and s^{2**} be the maximum of the sample variance, $s^2 = \bar{t}_2 - \bar{t}_1^2$, over the sample space $\mathfrak{X}$. Then

$$s^{2^{**}} = 1 \text{ if } n \text{ is even (when } \bar{t}_1 = 0),$$
$$= 1 - 1/n^2 \text{ if } n \text{ is odd (when } \bar{t}_1 = \pm 1/n). \qquad (3.19)$$

Proof: Let N_k and $\bar{t}_2^*$ be as defined in Theorem 3.2. From Theorem 3.2 we know that, on $\mathfrak{X}_{\bar{t}_1}$,

$$\bar{t}_1^2 \leq \bar{t}_2 \leq \bar{t}_2^* \text{ for a given } \bar{t}_1 \text{ on } N_k,\ k = 1,2,\ldots,n.$$

Hence $s^{2^*} = \bar{t}_2^* - \bar{t}_1^2$ is the maximum of s^2 on $\mathfrak{X}_{\bar{t}_1}$. By elementary calculus,we can show that s^{2^*} attains its local maximum at the end points of N_k, hence

$$s^{2^{**}} = \max s^{2^*},$$
$$= \max_{j=0,\ldots,n} [1-(2j/n - 1)^2].$$

Now evaluating $\max_{j=0,\ldots,n}[1-(2j/n-1)^2]$, we get the desired result.

We can obtain a similar result about the population variance of the doubly truncated normal population (3.4) and is given below.

<u>Corollary 3.3</u>: Let X be the doubly truncated normal random variable with the density function $f(x;\theta)$ as defined in (3.4). Then the supremum of variance of X over

θ is equal to 1/3.

Proof: Let D_{m_1} be as defined in Theorem 3.1. Then by Theorem 3.1, for (θ_1,θ_2) in D_{m_1}, we have

$$m_1^2 < E(X^2) < 1-(2/\theta_1^*)m_1,\ -1 < m_1 < 1,$$

where $H(\theta_1^*) = m_1$ and $H(\theta_1^*)$ is defined in (3.11). Hence

$$\begin{aligned} V^* &= \sup_{D_{m_1}} Var(X), \\ &= \sup_{D_{m_1}} (E(X^2) - m_1^2), \\ &= \frac{-4}{\left[e^{\theta_1^*} - e^{-\theta_1^*}\right]^2} + \frac{2}{\theta_1^{*2}} = F(\theta_1^*) \text{ (Say)}. \end{aligned}$$

Note that V^* is the supremum of variance of X in the subfamily $[f(x;\theta) : \theta \in D_{m_1}]$. Now it is easy to see that

$$\begin{aligned} V^{**} &= \sup Var(X), \\ &= \sup V^*, \\ &= \operatorname*{Max}_{\theta_1^*} F(\theta_1^*). \end{aligned}$$

But Max $F(\theta_1^*) = 1/3$ which follows by noting that

i) $H(\theta_1^*)$ is an odd function,

ii) $F(\theta_1^*) = \dfrac{dH(\theta_1^*)}{d\theta_1^*}$ is an even function,

iii) $\dfrac{dF(\theta_1^*)}{d\theta_1^*}$ is an odd function. That is $F(\theta_1^*)$ is strictly monotone for $\theta_1^* \neq 0$,

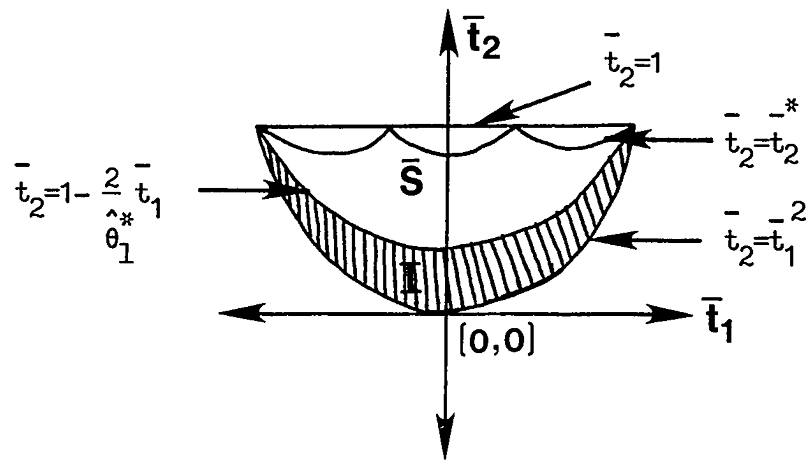
iv) $\underset{\theta_1^* \to \pm\infty}{\text{Limit}}\ F(\theta_1^*) = 0$ and $\underset{\theta_1^* \to 0}{\text{Limit}}\ F(\theta_1^*) = 1/3$.

Hence the result.

Note that $V^* = 1/3$ corresponds to the variance of a uniform random variable on $[-1,1]$. It may be of interest to know that a sufficient condition for the non-existence of the m.l.e. derived by Mittal [1984] with respect to the representation (3.1) is equivalent to saying that the maximum likelihood estimators of (μ,σ^2) do not exist whenever $s^2 > 1/3$ which follows by Corollary 3.3.

Contents of Theorem 3.1 through Corollary 3.3 are presented graphically in Figure 3.1. Note that in Figure 3.1, the set I as in (3.12) is superimposed on the set $\bar{S}$ as in (3.17). Hence, the m_1 axis is the same as $\bar{t}_1$ axis and the m_2 axis is the same as $\bar{t}_2$ axis.

In the following graph, the solution to (3.10) exists if and only if $(\bar{t}_1, \bar{t}_2)$ is in the shaded area.



Population and Sample Moment Spaces

Figure 3.1

3.3 Computational Results

In this section, we present a brief account of computational aspects in the truncated normal distribution. Note that the necessary and sufficient condition (3.14) requires solving the non-linear equation $H(\theta_1^*) = \bar{t}_1$. Also obtaining the m.l.e. of θ, whenever it exists, requires solving two simultaneous non-linear equations. Hence we discuss these aspects next in subsections (a) - (d).

(a) Computing Necessary and Sufficient Condition

We know that the m.l. equations (3.10) admits a solution in θ if and only if $\bar{t}_1^2 < \bar{t}_2 < 1 - (2/\hat{\theta}_1^*)\bar{t}_1$, where $\hat{\theta}_1^*$ is the solution of

$$H(\theta_1^*) - \bar{t}_1 = 0 \tag{3.20}$$

The equation (3.20) can be solved by Newton's method or by IMSL subroutines ZSPOW or ZSCNT. Next in Theorem 3.3, we state a result about a root of (3.20).

<u>Theorem 3.3</u>: Let $\hat{\theta}_1^*$ be a solution of the equation $H(\theta_1^*) - \bar{t}_1 = 0$. Then the following results are true:

i) $\hat{\theta}_1^*$ is unique and exists for all $-1 < \bar{t}_1 < 1$,

ii) $\hat{\theta}_1^*$ has the same sign as $\bar{t}_1$,

iii) $|\hat{\theta}_1^*| > 3|\bar{t}_1|$.

Proof: The result in i) follows from the fact that $H(\theta_1^*)$ is a strictly increasing function with the range $(-1,1)$, and ii) follows from i) by noting H(0) = 0. Finally, iii) follows from the following arguments.

Let

$$G(\theta_1^*) = 1 - (2/\theta_1^*)H(\theta_1^*).$$

It is easy to check that $\frac{dG(\theta_1^*)}{d\theta_1^*}$ has the same sign as θ_1^*.

Also $\lim_{\theta_1^* \to \pm\infty} G(\theta_1^*) = 1$ and $\lim_{\theta_1^* \to 0} G(\theta_1^*) = 1/3$. Hence the result.

Theorem 3.4 warns that the root of (3.20) must be accepted only if ii) and iii) are met. In actual computation, whenever $\bar{t}_1$ is in a small neighborhood of zero, we observe a violation of these conditions in some cases. In such cases, $\hat{\theta}_1^*$ can be rejected or re-tried with a new initial guess.

Giving a good starting guess while solving non-linear equations is an important step. Due to Theorem 3.4, $\theta_1^* = 3\bar{t}_1$ can be one starting guess. Next we state a result about generating a good guess for (3.20). One simple idea behind generating a starting value is to use the inverse function of some function that behaves similar to $H(\theta_1^*)$. It is observed that

$$h(\theta_1^*) = \text{sign}(\theta_1^*)\, e^{-\frac{1}{|\theta_1^*|}} , \quad -\infty < \theta_1^* < \infty,$$

is one such function which approximates $H(\theta_1^*)$ closely. It is easy to check that the inverse function of $h(\theta_1^*)$ is

$$h^{-1}(\bar{t}_1) = -\text{sign}(\bar{t}_1)/\log(|\bar{t}_1|) \;, \quad |\bar{t}_1| < 1.$$

It is observed that $\theta_1^* = h^{-1}(\bar{t}_1)$ is quite efficient in generating a starting value to (3.20). In fact when $|\bar{t}_1| > 0.5$, θ_1^* agrees with the root up to 3 decimal places.

b) Solving M.L. Equation

The m.l. equation (3.10) may be rewritten as

$$U(\theta) = 0, \text{ and } U(\theta) = \begin{bmatrix} U_1(\theta) \\ U_2(\theta) \end{bmatrix}. \tag{3.21}$$

In (3.21), $U_1(\theta) = \frac{1}{n}\frac{\partial K(\theta)}{\partial \theta_1} - \bar{t}_1$ and $U_2(\theta) = \frac{1}{n}\frac{\partial K(\theta)}{\partial \theta_2} - \bar{t}_2$. We already know that (3.21) can be solved only when (3.14) is satisfied. We may solve (3.21) by two-dimensional version of Newton's method or by IMSL subroutine ZSPOW. By Newton's method, iterative solution at $(i+1)^{th}$

iteration is

$$\theta^{(i+1)} = \theta^{(i)} - [J(\theta^{(i)})]^{-1}(U(\theta^{(i)})),\ i=0,1,2,\ldots \quad (3.22)$$

where $\theta^{(0)}$ is an initial guess and

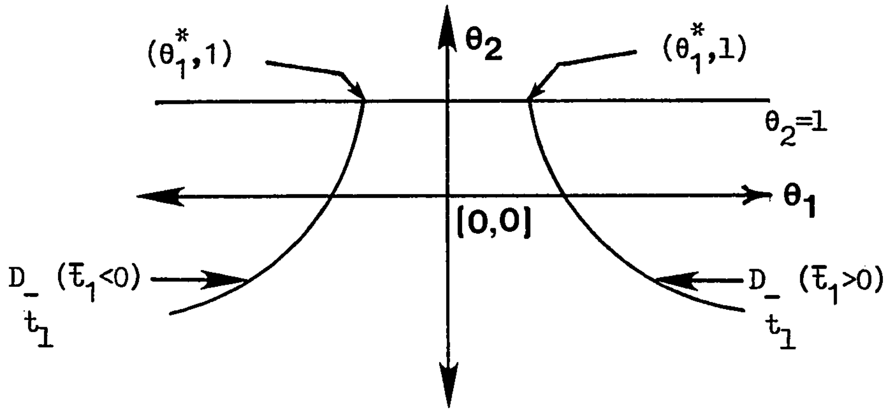
$$J(\theta) = \frac{1}{n}\begin{bmatrix} \frac{\partial^2 K}{\partial\theta_1^2}, & \frac{\partial^2 K}{\partial\theta_2 \partial\theta_1} \\ \frac{\partial^2 K}{\partial\theta_1 \partial\theta_2}, & \frac{\partial^2 K}{\partial\theta_2^2} \end{bmatrix}.$$

By using the definiton of $K(\theta)$ in (3.5), we can check that $\frac{\partial^2 K(\theta)}{\partial\theta_1^2} = Var(X)$, $\frac{\partial^2 K(\theta)}{\partial\theta_2^2} = Var(X^2)$, and $\frac{\partial^2 K(\theta)}{\partial\theta_1 \partial\theta_2} = Cov(X,X^2)$.

In the next paragraph, we mention an interesting result which is observed while solving (3.21).

We know that, given $(\bar{t}_1,\bar{t}_2)$, the solution to (3.21) must lie in $D_{\bar{t}_1} = [(\theta_1,\theta_2) : E(X) = \bar{t}_1]$, $-1 < \bar{t}_1 < 1$. The graph of $D_{\bar{t}_1}$ is observed to be like in Figure (3.2).

In the following graph, the point $(\hat{\theta}_1^*,1)$ denotes the boundary point of $D_{\bar{t}_1}$ and $\hat{\theta}_1^*$ is the solution of (3.20). Though not theoretically estabalished, $(\hat{\theta}_1^*,1)$ is found to be a good starting value for solving (3.21).



Solution Set

Figure 3.2

c) The Information Matrix

Let $f(x;\theta)$ be as defined in (3.4). Then the information matrix (in one observation) is defined to be

$$I(\theta) = (I_{ij}(\theta)) \ , \ i,j = 1,2, \tag{3.23}$$

and

$$I_{ij}(\theta) = \text{Cov}\left[\frac{\partial \log f(x;\theta)}{\partial \theta_i} \ , \ \frac{\partial \log f(x;\theta)}{\partial \theta_j}\right].$$

Using (3.4), it can be checked that

$$I(\theta) = \begin{bmatrix} \text{Var}(X) \ , & \text{Cov}(X,X^2) \\ \text{Cov}(X,X^2) \ , & \text{Var}(X^2) \end{bmatrix} .$$

Now we can see that $I(\theta)$ is the same as $J(\theta)$ defined in (3.22) and hence $I(\theta)$ can be obtained as a by-product result in the process of solving m.l. equation in simulation studies. It can be noted that the information matrix $I(\mu,\sigma^2)$ with respect to the representation (3.1) can be obtained as

$$I(\mu,\sigma^2) = Q\ I(\theta)\ Q', \tag{3.24}$$

where

$$Q = \begin{bmatrix} \frac{\partial\theta_1}{\partial\mu}, & \frac{\partial\theta_1}{\partial\sigma^2} \\ \frac{\partial\theta_2}{\partial\mu}, & \frac{\partial\theta_2}{\partial\sigma^2} \end{bmatrix}.$$

With θ_1 and θ_2 as defined in (3.4), it is easy to see that

$$Q = \begin{bmatrix} c_1/\sigma^2 , & \frac{-c_1(\mu-c_2)}{\sigma^4} \\ 0 , & c_1^2/2\sigma^4 \end{bmatrix}.$$

Now given the information matrix $I(\theta)$, the asymptotic variance-covariance matrix of the m.l.e.'s can be computed as

$$\sum_{\hat{\theta}} = (1/n)[I^{-1}(\theta)]. \tag{3.25}$$

d) Some Advantages of the (θ_1,θ_2) Parameterization over (μ,σ^2).

i) The necessary and sufficient condition (3.14) is computable only in the (θ_1,θ_2) parameterization since the inverse mapping of $(\theta_1,1)$ is $(\pm\infty, \infty)$. Hence it is felt that deriving a necessary and sufficient condition directly under the (μ,σ^2) parameterization would be difficult.

ii) As mentioned in the subsection b) above, we could give $(\hat{\theta}_1^*,1)$ as a starting value to the m.l. equations (3.21) while giving such a starting value is not possible under the (μ,σ^2) parameterization, particularly in the case of real life data.

iii) The derivation and computation of the information matrix, under the (θ_1,θ_2) parameterization is relatively easy due to the fact that the integrals involved in $J(\theta)$ are easily computable even when θ_2 is near 1, while the same is not possible under the (μ,σ^2) parameterization since $\sigma^2 \to \infty$ as $\theta_2 \to 1$.

3.4 Simulation Results

In this section, we discuss the relative performances of the Bayes modal, the harmonic mean and the mixed estimators. Since the maximum likelihood estimator does not exist with probability one, its performance is not considered for comparison. The performances of these estimators are judged with respect to the simulated values of the bias vector (the length and the direction), the mean square error (MSE), and the probability of nearness. Since the behaviour of the harmonic mean estimator of (μ,σ) is of interest, the results here are presented for the (μ,σ) parameterization. However the actual computational work is done in the (θ_1,θ_2)

parameterization. This study is based upon 500 random samples of sizes 10, 20, and 40. The details are discussed in subsections (a) - (c).

a) Plan of Simulation

All the computation work is carried out in Fortran interactively with IMSL subroutines. Since the simulated values of the bias, the MSE, and the probability of nearness do not vary with μ, we present the results for

i) $\mu = -5$, $\sigma = 0.5$,

ii) $\mu = -5$, $\sigma = 2.0$, and

iii) $\mu = -5$, $\sigma = 4.0$.

For a given (μ,σ), we consider the cases of left truncation probability (q_1) and right truncator probability (q_2) as

i) $q_1 = 0.05$, $q_2 = 0.05$,

ii) $q_1 = 0.05$, $q_2 = 0.10$, and

iii) $q_1 = 0.10$, $q_2 = 0.10$.

Then, using IMSL routine GGNML, we obtain a random observation (z) from Normal (0,1) and transform to $y=\mu+\sigma z$. We retain y if it lies in [A,B], otherwise GGNML is called again. For a given value of (q_1,q_2), we obtain $[z_1,z_2]$ from IMSL routine MDNRIS such that

$$q_1 = \int_{-\infty}^{z_1} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \text{ , and } q_2 = \int_{z_2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du,$$

and hence $A = \mu+\sigma z_1$, $B = \mu+\sigma z_2$. The above procedure is continued until we get a required sample $(y_1,\ldots,y_n)$. Then, $(y_1,\ldots,y_n)$ are transformed into $(x_1,\ldots,x_n)$ using the transformation $X = \frac{2(Y-B)}{(B-A)} - 1$. We then compute the sample statistics $\bar{t}_1 = \frac{1}{n} \Sigma x_i$, and $\bar{t}_2 = \frac{1}{n}\Sigma x_i^2$. Once $(\bar{t}_1,\bar{t}_2)$ are computed, we obtain the different estimates as follows. All the non-linear equations involved are solved using IMSL routine ZSPOW.

i) Maximum Likelihood Estimates: The maximum likelihood estimates $(\hat{\theta}_1,\hat{\theta}_2)$ are the solutions of (3.21) whenever $\bar{t}_2 < 1 - (2/\hat{\theta}_1^*)\bar{t}_1$ where $\hat{\theta}_1^*$ is the solution of the equation defined in (3.20).

ii) Bayes Modal Estimates: The Bayes modal estimates $(\tilde{\theta}_1,\tilde{\theta}_2)$ are the solutions of

$$E(X) = \bar{t}_1 \;,$$

$$E(X^2) - a/n + b/(n(1-\theta_2)) = \bar{t}_2, \quad a = 1/2, \quad b = 1,$$

where $E(X)$ and $E(X^2)$ are evaluated with respect to (3.4).

iii) Mixed Estimates: The mixed estimates $(\hat{\theta}_{1m},\hat{\theta}_{2m})$ are obtained as

$$\begin{aligned} \hat{\theta}_{1m} &= \hat{\theta}_1 \\ \hat{\theta}_{2m} &= \hat{\theta}_2 \end{aligned} \qquad \text{if (3.14) if true,}$$

and

$$\hat{\theta}_{1m} = \tilde{\theta}_1 \qquad\qquad \text{otherwise.}$$
$$\hat{\theta}_{2m} = \tilde{\theta}_2$$

Note that the mixed estimator studied by Mittal [1984] depends upon a mixing criterion as a function of the sufficient condition for the non existence of the m.l.e.

iv) Harmonic Mean Estimates: Before obtaining the harmonic mean estimates, we need to solve the equation

$$(\theta_2-\hat{\theta}_2)\bar{t}_2 - K(\hat{\hat{\theta}}_1,\theta_2) + K(\hat{\hat{\theta}}_1,\hat{\hat{\theta}}_2) - (1/n)\log c = 0, \qquad (3.26)$$

where $K(\theta_1,\theta_2)$ is as defined in (3.5) and c is a likelihood coefficient as mentioned in Section 2.5b. In (3.26) above, we take

$$\hat{\hat{\theta}}_1 = \hat{\theta}_1 \qquad\qquad \text{if (3.14) is true,}$$
$$\hat{\hat{\theta}}_2 = \hat{\theta}_2$$

and

$$\hat{\hat{\theta}}_1 = \hat{\theta}_1^* \qquad\qquad \text{otherwise,}$$
$$\hat{\hat{\theta}}_2 = 1$$

where $(\hat{\theta}_1,\hat{\theta}_2)$ are the m.l.e.'s and $\hat{\theta}_1^*$ is the solution of (3.20). Due to unimodality of the likelihood function, we can show that the equation (3.26) admits two solutions with the possibility of one of the solutions on the boundary, $\theta_2 = 1$, for certain samples. Let $(\theta_2)_c^+$ and $(\theta_2)_c^{++}$ be the two solutions of (3.26) such that $(\theta_2)_c^+ < (\theta_2)_c^{++}$. Given $[(\theta_2)_c^+, (\theta_2)_c^{++}]$, we first invert the same to $[(\sigma^2)_c^+, (\sigma^2)_c^{++}]$ using the inverse relation $\sigma^2 = 1 - \dfrac{c_1^2}{2(1-\theta_2)}$. Note that whenever $\theta_2 = 1$, $\sigma^2 = \infty$.

Hence we obtain the harmonic mean estimate $(\hat{\sigma}^2)_c$ of σ^2 as

$$(\hat{\sigma}^2)_c = \frac{2(\hat{\sigma}^2)_c^+ (\sigma^2)_c^{++}}{(\sigma^2)_c^+ + (\sigma^2)_c^{++}} \qquad \text{if } (\sigma^2)_c^{++} < \infty$$

$$= 2(\sigma^2)_c^+ \qquad \text{if } (\sigma^2)_c^{++} = \infty .$$

Once $(\hat{\sigma})_c$ is computed, we obtain the harmonic mean estimate of μ as $(\hat{\mu}_c)$

$$\hat{\mu}_c = \frac{\hat{\theta}_1(\hat{\sigma}^2)c}{c_1} + c_2 \ , \ c_1 = \frac{B-A}{2} \text{ and } c_2 = \frac{A+B}{2} .$$

Different values of c in (0,1) are tried and c = 0.01 is observed to be the optimum in the sense of minimum mean square error. Hence, we use the same for comparing with the Bayes modal and the mixed estimators.

b) Bias, MSE, and Probability of Nearness

Let $T = (T_1,T_2)$ and $T^* = (T_1^*,T_1^*)$ be any two estimators of $\theta = (\theta_1,\theta_2)$. Let (t_{1i},t_{2i}) and (t_{1i}^*,t_{2i}^*) be the values of (T_1,T_2) and (T_1^*,T_2^*) respectively for the i^{th} sample where $i = 1,2,\ldots,M$. We use M=500 in this study. Then the simulated values of the bias vector of T, the MSE of T, the probability of nearness of T with respect to T^* are obtained as follows.

Bias: The magnitude of the bias of T in estimating θ is defined to be $\|b\| = \sqrt{(b_1^2 + b_2^2)}$ where

$$b_1 = (1/M)\left[\sum_{i=1}^{M} t_{1i}\right] - \theta_1, \text{ and } b_2 = (1/M)\left[\sum_{i=1}^{M} t_{2i}\right] - \theta_2.$$

Hence, we compute the direction angle of b as

$$\text{Dir} = \cos(b_1/\|b\|).$$

MSE: The simulated mean square error (MSE) of T is defined to be

$$\text{MSE}(T) = (1/M) \sum_{i=1}^{M} [(t_{1i} - \theta_1)^2 + (t_{2i} - \theta_2)^2].$$

Hence, we define the efficiency of T^* with respect to T as

$$\text{EFF}(T^*,T) = \text{MSE}(T^*)/\text{MSE}(T).$$

Probability of Nearness: The probability of nearnes of T with respect to T^*, $p(T,T^*)$, is defined as follows. Let E represent the event

$$\left[\sqrt{(T_1 - \theta_1)^2 + (T_2 - \theta_2)^2} < \sqrt{(T_1^* - \theta_1)^2 + (T_2^* - \theta_2)^2}\right].$$

Then, $p(T,T^*) = m/M$, where m is the number of times E occurs.

The Probability of Non-existence: Let n_0 be the number of times (3.14) is not true. Then the probability of non-existence of the m.l.e. is $p_0 = n_0/M$.

c) Comments on Table 3.1 and Table 3.2

Table 3.1 contains the asymptotic variances of the maximum likelihood estimators $(\hat{\mu},\hat{\sigma})$ for different values of (μ,σ) and the truncation probabilities. Table 3.2

contains the simulated values of the bias, the MSE, the efficiency, and the probability of nearness for different estimators. Next, a brief account of the performances of these estimators is presented.

Bias: The magnitude of bias of the Bayes modal is relatively smaller compared to the mixed and the harmonic mean estimators for all the chosen levels of the parameters. The harmonic mean estimator has relatively larger length of the bias vector. By looking at the directional angle, we can see that the Bayes modal, and the harmonic mean estimators have approximately the same angles in the range of 250-280 degrees. In other words both the components of the bias vector are negative. On the contrary, the directional angle for the mixed estimator lies in the first quadrant and hence it has the tendency of overestimating both the parameters. Another important point observed is that the length of bias of the Bayes modal estimator converges to zero faster while the harmonic mean estimator has a slower rate of convergence.

Efficiency: We compare the MSE's of the Bayes modal and the harmonic mean estimators with that of the mixed estimator. By studying the efficiencies reported in Table 3.2, it is strikingly obvious that the harmonic mean estimator has uniformly smaller MSE compared to the Bayes modal estimator. Note that both the estimators (Bayes modal and harmonic mean) have efficiencies larger than one with respect to the mixed estimator. Another important point noted from Table 3.1 and Table 3.2 together is that the variances of all the estimators are close to the asymptotic variance of the maximum likelihood estimator for n=40.

Probability of Nearness: In Table 3.2, p_1 represents the probability of nearness of the mixed estimator with respect to the Bayes modal estimator, p_2 represents the probability of nearness of the mixed estimator with respect to the harmonic mean estimator, and p_3 is the probability of nearness of the Bayes modal with respect to the harmonic mean estimator. It is clear that $p_1 < 0.5$, $p_2 > 0.5$, and $p_3 > 0.5$. In other words, the Bayes modal is relatively near to the parameter more often while the harmonic mean estimator does not behave well with respect to this criterion.

Probability of Non-existence: In Table 3.2, p_0 represents the probability of non-existence of the maximum likelihood estimator. We can see that p_0 is as high as 20%, particularly when the truncation probabilities are high. Note that the probability of non-existence of the m.l.e. as reported by Mittal [1984] are based opon the sufficient condition while the ones reported here are based upon the necesary and sufficient condition and hence the figures reported here are more accurate.

As a concluding remark to the above comments,

i) The Bayes modal estimator is preferable to the mixed and the harmonic mean estimators if the bias or the probability of nearness is a criterion,

ii) The harmonic mean estimator is preferable if MSE is a criterion.

Table 3.1: Asymptotic Variance of the Maximum Likelihood Estimators of (μ,σ) for Doubly Truncated Normal Distribution

μ	σ	n	v_1	v_2	v_1+v_2	μ	σ	n	v_1	v_2	v_1+v_2
$q_1 = 0.05$, $q_2 = 0.05$											
-5.0	0.5	10	0.029	0.058	0.087	-5.0	4.0	10	2.086	6.077	8.162
		20	0.015	0.029	0.044			20	1.042	3.038	4.080
		40	0.007	0.014	0.021			40	0.521	1.519	2.040
						$q_1 = 0.10$, $q_2 = 0.10$					
-5.0	2.0	10	0.468	0.921	1.389	5.0	0.5	10	0.037	0.178	0.215
		20	0.234	0.460	0.694			20	0.018	0.088	0.107
		40	0.117	0.230	0.347			40	0.009	0.044	0.053
-5.0	4.0	10	1.872	3.683	5.555	-5.0	2.0	10	0.584	2.844	3.428
		20	0.936	1.842	2.788			20	0.292	1.422	1.714
		40	0.468	0.921	1.389			40	0.146	0.711	0.857
$q_1 = 0.05$, $q_2 = 0.10$											
-5.0	0.5	10	0.033	0.095	0.128	-5.0	4.0	10	2.336	11.374	13.710
		20	0.016	0.047	0.063			20	1.168	5.687	6.855
		40	0.008	0.024	0.032			40	0.584	2.844	3.428
-5.0	2.0	10	0.521	1.519	2.040						
		20	0.261	0.760	1.021						
		40	0.130	0.380	0.510						

v_1=variance of $(\hat{\mu})$ q_1=left trunc. probability

v_2=variance of $(\hat{\sigma})$ q_2=right trunc. probability

Note: v_1+v_2 is comparable with MSE in Table 3.2

Table 3.2 Simulated Expected Bias and MSE of the Mixed, the Bayes Modal, and the Harmonic Mean Estimators of (μ,σ) for Doubly Truncated Normal Distribution
Prior Parameters : $a = 1/2$, $b = 1$; Likelihood Coefficient: $c = 0.01$

		Mixed Estimates					Bayes Modal Estimates					Harmonic Mean Estimates			
n	p_0	Bias	Dir	p_1	p_2	MSE	Bias	Dir	p_3	MSE	EFF	Bias	Dir	MSE	Eff
10	0.07	0.028	94	0.21	0.54	0.255	0.056	269	0.64	0.057	4.467	0.190	270	0.049	5.175
20	0.03	0.043	102	0.23	0.55	0.139	0.028	269	0.62	0.033	4.204	0.119	270	0.027	5.205
40	0.01	0.046	56	0.30	0.53	0.409	0.010	282	0.58	0.021	13.046	0.063	270	0.014	28.811
$\mu = -5.00$, $\sigma = 2.00$, $q_1 = 0.05$, $q_2 = 0.05$															
10	0.09	0.308	149	0.22	0.55	61.526	0.182	276	0.65	0.935	65.818	0.701	270	0.757	81.317
20	0.04	0.262	77	0.23	0.54	2.997	0.059	278	0.60	0.628	4.776	0.440	270	0.413	7.261
40	0.00	0.043	93	0.32	0.59	1.350	0.087	275	0.63	0.312	4.329	0.281	271	0.236	5.717
$\mu = -5.00$, $\sigma = 4.00$, $q_1 = 0.05$, $q_2 = 0.05$															
10	0.08	0.605	49	0.24	0.56	54.251	0.396	275	0.66	3.714	14.606	1.409	270	2.993	18.124
20	0.03	0.442	94	0.22	0.53	8.729	0.162	277	0.60	2.124	4.110	0.910	271	1.616	5.403
40	0.01	0.204	87	0.30	0.55	2.743	0.079	269	0.58	1.237	2.217	0.500	269	0.867	3.164
$\mu = -5.00$, $\sigma = 0.50$, $q_1 = 0.50$, $q_2 = 0.10$															
10	0.11	0.041	14	0.33	0.60	0.331	0.082	262	0.72	0.062	5.364	0.205	256	0.057	5.781
20	0.06	0.124	36	0.35	0.58	0.867	0.048	270	0.68	0.039	22.063	0.145	256	0.034	25.693
40	0.03	0.120	38	0.42	0.58	0.955	0.016	279	0.63	0.026	36.065	0.085	254	0.018	53.484
$\mu = -5.00$, $\sigma = 2.00$, $q_1 = 0.50$, $q_2 = 0.10$															
10	0.12	0.385	21	0.32	0.58	23.032	0.311	256	0.72	0.923	24.340	0.815	254	0.893	25.806
20	0.08	0.558	49	0.36	0.56	17.688	0.155	259	0.66	0.653	27.213	0.564	253	0.525	33.672
40	0.02	0.270	99	0.42	0.57	9.585	0.091	249	0.62	0.401	23.876	0.359	250	0.293	32.757

Table 3.2 (Continued)

n	p_0	Mixed Estimates Bias	Dir	p_1	p_2	MSE	Bayes Modal Estimates Bias	Dir	p_3	MSE	EFF	Harmonic Mean Estimates Bias	Dir	MSE	EFF
= -5.0,		= 4.0,		q_1 = 0.50,		q_2 = 0.10									
10	0.13	1.964	11	0.36	0.59	838.053	0.623	253	0.71	4.176	200.665	1.630	254	3.697	226.663
20	0.07	1.489	13	0.39	0.61	528.645	0.439	274	0.72	2.542	207.953	1.190	257	2.213	238.885
40	0.02	0.350	56	0.48	0.64	8.182	0.250	266	0.67	1.532	5.341	0.749	254	1.194	6.850
= -5.00,		= 0.50,		q_1 = 0.10,		q_2 = 0.10									
10	0.20	0.040	206	0.45	0.62	0.284	C.110	265	0.73	0.065	4.350	0.224	268	0.064	4.439
20	0.12	0.039	70	0.47	0.58	0.275	C.073	273	0.70	0.043	6.361	0.168	270	0.040	6.945
40	0.04	0.072	72	0.51	0.58	0.268	C.043	279	0.69	0.027	0.947	0.114	272	0.022	12.085
= -5.00,		= 2.00,		q_1 = 0.10,		q_2 = 0.1C									
10	0.17	0.139	344	0.41	0.62	5.365	0.451	278	0.75	0.943	5.686	0.900	272	0.005	5.338
20	0.11	0.498	49	0.49	0.58	17.116	0.279	278	0.73	0.623	27.462	0.663	272	0.609	28.082
40	0.08	0.457	101	0.47	0.54	50.174	0.107	269	0.66	0.470	106.819	0.425	270	0.340	147.691
= -5.00,		= 4.00,		q_1 = 0.10,		q_2 = 0.10									
10	0.18	0.656	9	0.45	0.58	119.737	0.831	271	0.70	3.877	30.880	1.760	270	3.928	30.484
20	0.12	0.649	60	0.49	0.55	95.341	0.528	262	0.71	2.587	36.847	1.312	268	2.401	39.715
40	0.06	0.596	55	0.57	0.62	51.477	0.352	269	0.73	0.706	30.180	0.923	269	1.424	36.147

Index: p_0: Prob of non-existence of the m.l.e. q_1: Left Truncation, q_2 : Right Trunc.

p_1: Prob. of Nearnes of (Mixed, Bayes modal)

p_2: Prob. of Nearness of (Mixed, Harmonic mean)

p_3: Prob. of Nearness of (Bayes modal, Harmonic mean)

4. TRUNCATED GAMMA DISTRIBUTION

In this chapter, we discuss a two parameter gamma distribution. The scheme of presentation is as follows:

i) Section 4.1: Family representation.

ii) Section 4.2: Derivation of the maximum likelihood equations and discussion of the existence of a solution.

iii) Section 4.3: Computational results for solving the maximum likelihood equations.

iv) Section 4.4: Derivation of the Bayes modal estimator.

v) Section 4.5: Discussion of simulation results.

The discussion is mainly in an exponential family representation. Theorem 4.1 is the key result of this chapter.

4.1 Family Representation

A truncated gamma population is described by a random variable Y with density function

$$f(y;\beta,\alpha) = \frac{e^{-\frac{1}{\beta}y}\, y^{\alpha-1}}{\left[\int_0^U e^{-\frac{1}{\beta}u}\, u^{\alpha-1}du\right]}, \qquad (4.1)$$

where $0 \le y \le U$, $U > 0$ known, and (β,α) in Ω with

$$\Omega = [(\beta,\alpha) : 0 < \beta < \infty, \quad 0 < \alpha < \infty] .$$

Since the truncation point U is assumed to be known, we can change $X = Y/U$ in (4.1) so that the truncation interval becomes [0,1]. That is, we may rewrite (4.1) as

$$f(x;\beta,\alpha) = \frac{e^{-\frac{1}{\beta}y} y^{\alpha-1}}{\left[\int_0^1 e^{-\frac{1}{\beta}u} u^{\alpha-1} du\right]} . \tag{4.2}$$

Given n i.i.d. observations from (4.2), (β,α) is the parameter of interest. In other words, the indexed family of interest is

$$\mathcal{B}_2 = [L(\underline{x}:\beta,\alpha) : \bar{x} = (x_1 \ldots x_n) \in \mathfrak{X} , (\beta,\alpha)\in\Omega],$$

$$\mathfrak{X} = [\underline{x} : 0 \le x_i \le 1, \quad i = 1,2,\ldots,n],$$

and

$$L(\underline{x}\ ;\ \beta,\alpha) = \frac{e^{-\frac{U}{\beta}\sum x_i + (\alpha-1)\sum \log x_i}}{\left[\int_0^1 e^{-\frac{U}{\beta}u + (\alpha-1)\log u} du\right]^n} . \qquad (4.3)$$

In (4.3), $L(x;\beta,\alpha)$ is the likelihood of the sample. Here $\mathfrak{B}_2$ is an exponential family. It is easy to see that $L(\underline{x};\beta,\alpha)$ can be expressed in a standard minimal exponential representation of the type (2.1) with

$$a(\beta,\alpha)\ :\ \Omega \longrightarrow R,\ a(\beta,\alpha) = \left[\int_0^1 e^{(\tau_1(\beta,\alpha)-1)u+\tau_2(\beta,\alpha)\log u}\right]^{-n}$$

$b(\underline{x})\ :\ \mathfrak{X} \longrightarrow R,\quad b(\underline{x}) = e^{-\Sigma\, x_i}$,

$T(\underline{x})\ :\ \mathfrak{X} \longrightarrow S,\ S = [(t_1,t_2)\ :\ t_1 = \Sigma\, x_i,\ t_2 = \Sigma \log x_i,\ \underline{x}\in \mathfrak{X}]$,

$\tau(\beta,\alpha):\Omega \longrightarrow \Theta,\ \Theta=[(\theta_1,\theta_2):\ \theta_1=\tau_1(\beta,\alpha),\ \theta_2=\tau_2(\beta,\alpha),(\beta,\alpha)\in\Omega]$,

and $\tau_1(\beta,\alpha)=1 - \frac{U}{\beta}$, $\tau_2(\beta,\alpha)=\alpha-1$.

It is easy to see that τ is one-one with the inverse mapping

$$\beta = U/(1-\theta_1), \quad \text{and} \quad \alpha = \theta_2 + 1.$$

We may represent the truncated gamma population (4.2) under the (θ_1, θ_2) parameterization as

$$f(x;\theta_1,\theta_2) = \frac{e^{(\theta_1-1)x+\theta_2 \log x}}{\left[\int_0^1 e^{(\theta_1-1)u+\theta_2 \log u} du\right]} . \tag{4.4}$$

In our later discussions, we refer (4.4) as the truncated gamma distribution and find it to be useful.

The distribution of $T(\underline{X})=(T_1(\underline{X})=\Sigma\ X_i, T_2(\underline{X})=\Sigma \log X_i)$ with the parameterization (θ_1,θ_2) is of the form

$$g(t;\theta) = h(t)e^{\theta \cdot t-K(\theta)}, t \in S, \quad \theta \in \Theta. \tag{4.5}$$

In (3.5), $K(\theta) = \log c(\theta)$, and $c(\theta)$ is given by

$$c(\theta) = \left[\int_0^1 e^{(\theta_1-1)u+\theta_2 \log u} du\right]^n .$$

Note that Θ is one-one mapping of Ω and hence

$$\Theta = [(\theta_1,\theta_2): -\infty < \theta_1 < 1, -1 < \theta_2 < \infty].$$

Here Θ is parameter space of interest. It is easy to check that the natural parameter space $\bar{\Theta}$ is

$$\bar{\Theta} = \left[(\theta_1,\theta_2):\ \int_0^1 e^{(\theta_1-1)u+\theta_2 \log u} du < \infty\right],$$

$$= \left[(\theta_1,\theta_2):\ -\infty < \theta_1 < \infty\ ,\ -1 < \theta_2 < \infty\right].$$

Now it is clear that the parameter space of interest is a proper subset of the natural parameter space. In other words,

$$\mathscr{A} = [g(t;\theta):\ \theta \in \Theta] \tag{4.6}$$

is a proper subfamily of the full family. Now, we discuss a monotonicity property of $DK(\theta)$ where $DK(\theta) = \left[\frac{\partial K}{\partial\theta_1}, \frac{\partial K}{\partial\theta_2}\right]$.

From the definition of $K(\theta)$ in (4.5), we can verify that

$$\frac{\partial K}{\partial\theta_1} = nE(X)\ ,\quad \text{and}\quad \frac{\partial K}{\partial\theta_2} = nE(\log X),$$

where $E(X)$ and $E(\log X)$ are expected values of X and $\log X$ with respect to (4.4) respectively.

Let

$$I = [(m_1,m_2):\ m_1 = E(X),\ m_2 = E(\log X),\quad \theta \in \Theta]. \tag{4.7}$$

In (4.7), I is the range of $(\frac{1}{n})DK(\theta)$ on Θ. As mentioned in Chapter three, I may be called the population moment space. Note that $(\frac{1}{n})DK(\theta)$ satisfies the monotonicity result (2.7). That is,

$$(\theta_1-\tilde{\theta}_1)(m_1-\tilde{m}_1) + (\theta_2-\tilde{\theta}_2)(m_2-\tilde{m}_2) > 0 \tag{4.8}$$

where (m_1,m_2) and $(\tilde{m}_1,\tilde{m}_2)$ are the values of $(\frac{1}{n})DK(\theta)$ at θ and $\tilde{\theta}$ respecttively. It can be seen that the result (4.8) also implies the following:

i) The strict montonicity of $E(X)$ for fixed θ_2,
ii) The strict montonicity of $E(\log X)$ for fixed θ_1.

(4.9)

These monotonicity properties of $E(X)$ and $E(\log X)$ are useful in discussing the maximum likelihood estimation in certain subfamilies of $\mathscr{A}$.

4.2 Maximum Likelihood Estimation

The log likelihood of the sample $t = (t_1,t_2)$ is

$$l(\theta;t) = \theta.t - K(\theta) + \log h(t),$$

where $K(\theta)$ and $h(t)$ are as defined in (4.5). Hence the maximum likelihood estimator of θ is a solution of

$$DK(\theta) = t, \quad \theta \in \Theta.$$

Let $\bar{t}_1 = t_1/n$, $\bar{t}_2 = t_2/n$ and $\bar{S} = [(\bar{t}_1,\bar{t}_2):(t_1,t_2) \in S]$. We may call $\bar{S}$ the sample moment space. Hence the equation, $DK(\theta) = t$, can equivalently be written as

$$\begin{aligned} E(X) &= \bar{t}_1, \\ E(\log X) &= \bar{t}_2. \end{aligned} \tag{4.10}$$

The equation (4.10) is referred to as the maximum likelihood equations. We now discuss the question of existence of a solution to (4.10).

Since Θ is an open convex subset of $\bar{\Theta}$, Theorem 2.4b is applicable. Hence the question of existence of a solution to the m.l. equations requires determining the set I. It means that the range of $E(\log X)$ be expressed as an explicit function of m_1. This is done in Theorem 4.1 below.

<u>Theorem 4.1</u>: Let $D_{m_1} = [(\theta_1,\theta_2):E(X) = m_1]$, $0 < m_1 < 1$, and Θ^* be the boundary of Θ. Then for θ in D_{m_1}, the following result holds:

$$1 - 1/m_1 < E(\log X) < \log(m_1). \qquad (4.11)$$

Proof: The upper inequality in (4.11) follows from the well known result that $E(\log X) < \log(E(X))$. Hence we need to show only the lower inequality in (4.11). It is clear that $\Theta^* = \Theta_1^* \cup \Theta_2^*$ where

$$\Theta_1^* = [(\theta_1^*,-1):\ -\infty<\theta_1^*\le 1], \text{ and } \Theta_2^* = [(1,\theta_2^*):\ -1\le\theta_2^*<\infty].$$

Note that $\{1,-1\}$ is the only common point in Θ_1^* and Θ_2^*. Let $(\theta_1^*,-1)$ and $(1,\theta_2^*)$ be any given points in Θ_1^* and Θ_2^* respectively. Then the following results hold:

i) $$\lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to -1}} E(X) = \lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to -1}} \left[\int_0^1 x\, f(x;\theta)dx\right], = 0,$$

and

$$\lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to -1}} E(\log X) = \lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to -1}} \left[\int_0^1 \log x\, f(x;\theta)dx\right] = -\infty.$$

ii) $$\lim_{\substack{\theta_1 \to 1 \\ \theta_2 \to \theta_2^*}} E(X) = \lim_{\substack{\theta_1 \to 1 \\ \theta_2 \to \theta_2^*}} [\int_0^1 x\, f(x;\theta)dx],$$

$$= (\theta_2^*+1)/(\theta_2^*+2),$$

$$= F_1(\theta_2^*), \quad \text{say,} \tag{4.12a}$$

and

$$\lim_{\substack{\theta_1 \to 1 \\ \theta_2 \to \theta_2^*}} E(\log X) = \lim_{\substack{\theta_1 \to 1 \\ \theta_2 \to \theta_2^*}} [\int_0^1 \log x\, f(x;\theta)dx],$$

$$= -1/(\theta_2^*+1),$$

$$= F_2(\theta_2^*), \quad \text{say.} \tag{4.12b}$$

The direct proof of (i) is difficult. However, the same follows by noting that the limit of $f(x;\theta)$, as $\theta_1 \to \theta_1^*$ and $\theta_2 \to -1$, is a degenerate function at $x = 0$. The proof of (ii) follows by a straight forward evaluation of the limiting integrals. It is easy to check that in 4.12a and 4.12b, the following are true:

i) $F_1(\theta_2^*)$ is strictly increasing function of θ_2^* with the range $(0,1)$.

ii) $F_2(\theta_2^*)$ is strictly increasing function of θ_2^* the with range $(-\infty,0)$.

Due to the montonicity of $F_1(\theta_2^*)$, for a given $0 < m_1 < 1$, there exists a unique θ_2^* such that $F_1(\theta_2^*) = m_1$. In other words, $(1,\theta_2^*)$ is the boundary point of D_{m_1}. Note that D_{m_1} does not have any boundary point in θ_1^*. In fact

$$\theta_2^* = (2m_1-1)/(1-m_1).$$

We may check that

$$F_2(\theta_2^*) = 1 - 1/F_1(\theta_1^*) = 1 - 1/m_1.$$

Now the monotonicity property (4.8) of E(log X) in D_{m_1} and the fact that $\theta_2 > \theta_2^*$ in D_{m_1} together imply that

$$\operatorname*{Inf}_{D_{m_1}} E(\log X) = 1 - 1/m_1.$$

Hence the result.

Theorem 4.1 characterizes the set I completely and hence we may write the set I (population moment space) explicitely as

$$I = [(m_1,m_2): 0 < m_1 < 1,\ 1 - 1/m_1 < m_2 < \log m_1]. \qquad (4.13)$$

Now using Theorem 4.1, we derive a necessary and sufficient condition for the existence of a solution to the equation (4.10) and it is stated in Corollary 4.1 below.

<u>Corollary 4.1</u>: Given $\bar{t} = (\bar{t}_1,\bar{t}_2)$, $0 < \bar{t}_1 < 1$, the maximum likelihood equation (4.10) admits a solution in Θ if and only if

$$1 - \frac{1}{\bar{t}_1} \le \bar{t}_2 \le \log \bar{t}_1. \qquad (4.14)$$

Proof: Given $(\bar{t}_1,\bar{t}_2)$, $0 < \bar{t}_1 < 1$, it is obvious that the solution to equations (4.10) must lie in

$$D_{\bar{t}_1} = [\theta_1,\theta_2):E(X) = \bar{t}_1].$$

By Theorem 4.1, on $D_{\bar{t}_1}$, $1 - \frac{1}{\bar{t}_1} < E(\log X) < \log \bar{t}_1$. Hence the equations $E(X) = \bar{t}_1$ and $E(\log X) = \bar{t}_2$ admit a solution in Θ if and only if (4.14) is true.

As an important remark, the equation (4.10) admit a solution with probability one in the full family $\mathscr{A} = [g(t;\theta):\theta \in \bar{\Theta}]$. Next we discuss the existence of a solution to the m.l. equations in subfamilies

$$\mathscr{A}_{\theta_2} = [g(t;\theta): -\infty < \theta_1 < 1], \quad \theta_2 > -1 \text{ is known},$$

$$\mathscr{A}_{\theta_1} = [g(t;\theta): -1 < \theta_2 < \infty], \quad \theta_1 < 1 \text{ is known},$$

where $g(t;\theta)$ is as defined in (4.5). Note that $\mathscr{A}_{\theta_2}$ is the same as α -known and $\mathscr{A}_{\theta_1}$ is the same as β-known in the representation (4.1). It may be checked that the m.l. estimator of θ_1 in $\mathscr{A}_{\theta_2}$ is a solution of

$$E(X) = \bar{t}_1, \tag{4.15}$$

and the m.l. estimator of θ_2 in $\mathscr{A}_{\theta_1}$ is a solution of

$$E(\log X) = \bar{t}_2. \tag{4.16}$$

The results in (4.9) imply the following:

i) $E(X)$ is strictly increasing in $\mathscr{A}_{\theta_2}$ with the range $(0,(\theta_2+1)/(\theta_2+2))$,

ii) $E(\log X)$ is strictly increasing in $\mathscr{A}_{\theta_1}$ with the range $(-\infty,0)$.

Now it is obvious that (4.15) admits a solution if and only if $\bar{t}_1 < (\theta_2+1)/(\theta_2+2)$ while (4.16) always admits a solution. The subfamily $\mathscr{A}_{\theta_2}$ with $\theta_2 = 0$ is the truncated negative exponential distribution. These results about $\mathscr{A}_{\theta_2}$ are equivalent to results obtained by Broeder [1955].

Next we discuss figuring out the set $\bar{S}$ (sample moment space) on the similar lines of Theorem 3.2.

<u>Theorem 4.2</u>: Let $0 \le x_i \le 1$, $\bar{t}_1 = \frac{1}{n}\Sigma x_i$, $\bar{t}_2 = \frac{1}{n}\Sigma \log x_i$.

Then

$$-\infty < \bar{t}_2 \le \log \bar{t}_1 \text{ , for } 0 \le \bar{t}_1 \le (n-1)/n,$$

and

$$\frac{1}{n}\log(n\bar{t}_1-(n-1))<\bar{t}_2\le\log \bar{t}_1, \qquad \text{for } \frac{n-1}{n} <\bar{t}_1\le 1. \tag{4.17}$$

Proof: Let $N_k = [(k-1)/n, (k/n)]$, $k = 1,2,\ldots,n$ be subintervals dividing $[0,1]$. Let

$$\mathfrak{X} = [(x_1\ldots x_n): 0 \le x_i \le 1, i=1,\ldots,n] \text{ and}$$

$$\mathcal{X}_{\bar{t}_1} = [(x_1 \ldots x_n) \in \mathcal{X}\colon (1/n)\Sigma\, x_i = \bar{t}_1],\ 0 \le \bar{t}_1 \le 1.$$

Given $\bar{t}_1$, it is well known that $\bar{t}_2 \le \log \bar{t}_1$ due to $E(\log X) \le \log(E(X))$. Hence we need to show only the lower inequality in (4.17). Since $\bar{t}_2 = \frac{1}{n}\left(\sum_{i=1}^{n} \log x_i\right)$ is a concave function, it attains its minimum on the set of extreme points of $\mathcal{X}_{\bar{t}_1}$. As stated in Theorem 3.2, the set of extreme points of $\mathcal{X}_{\bar{t}_1}$, $\bar{t}_1$ in N_k for some k, consists of all the permutations of $(x_1 .. x_j \ldots x_n)$, where

$$\begin{aligned} x_j &= 1\ , & j &= 1,2,\ldots,k, \\ &= n\bar{t}_1 - (k-1), & j &= k, \\ &= 0\ , & j &= k+1,\ldots,n. \end{aligned}$$

Since $\bar{t}_2$ is invariant under permutation, it attains the same minimum for every point in the set of extreme points of $\mathcal{X}_{\bar{t}_1}$. Another important point to be noted is that at least one of the coordinates of any extreme points of $\mathcal{X}_{\bar{t}_1}$ is zero for $\bar{t}_1$ in $[0, \frac{(n-1)}{n}]$ and none of the cordinates of any extreme point of $\mathfrak{X}_{\bar{t}_1}$ is zero for $\bar{t}_1$ in $((n-1)/n,1]$.

Hence,

$$\min \bar{t}_2 = -\infty, \text{ for } \bar{t}_1 \text{ in } [0,\ (n-1)/n],$$
$$= (1/n)(\log(n\bar{t}_1-(n-1)) \quad \text{for } \bar{t}_1 \text{ in } ((n-1)/n,\ 1].$$

Hence the result,

Now it is easy to characterize the set $\bar{S}$ (sample moment space) as

$$\bar{S} = \bar{S}_1 \cup \bar{S}_2, \tag{4.18}$$

where

$$\bar{S}_1 = [(\bar{t}_1;\bar{t}_2):0\le\bar{t}_1\le(n-1)/n,\ -\infty<\bar{t}_2\le\log \bar{t}_1],$$
$$\bar{S}_2 = [(\bar{t}_1;\bar{t}_2):(n-1)/n<\bar{t}_1\le 1,\ \frac{1}{n}\ (\log(n\bar{t}_1-(n-1))<\bar{t}_2\le\log \bar{t}_1].$$

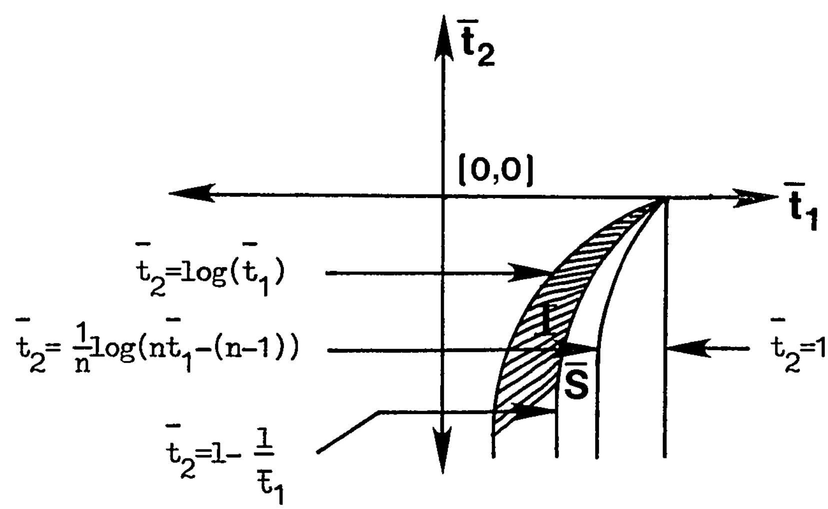
Let C denote the convex hull of $\bar{S}$. We can check that

$$C = [(\bar{t}_1,\bar{t}_2):0 \le \bar{t}_1 \le 1,\ -\infty < \bar{t}_1 \le \log \bar{t}_1].$$

Note that $\bar{S}\rightarrow C$ as $n\rightarrow\infty$ as noted in the case of doubly truncated normal distribution.

Next we show a graphical presentation of the discussions in Theorem 4.1 and Theorem 4.2 through Figure 4.1. Note that the set I as of (4.13) is superimposed on the set $\bar{S}$ as of (4.18) and hence the m_1-axis is the same as the $\bar{t}_1$-axis and the m_2-axis is the same as the $\bar{t}_2$-axis.

In the following Figure 4.1, the solution to the maximum likelihood equations (4.10) exists if and only if $(\bar{t}_1,\bar{t}_2)$ lies in the shaded area.



Population and Sample Moment Spaces

Figure 4.1

4.3 Computational Results

The main computatinal interest in the truncated gamma family ($\mathcal{B}_2$) is in solving the maximum likelihood equations. Recall that the m.l. equations are given by

$$\int_0^1 x\ f(x;\theta)dx = \bar{t}_1,$$

$$\int_0^1 \log x\ f(x;\theta)dx = \bar{t}_2.$$

For computational convenience, we may introduce the following notations:

$$A = \theta_1 - 1\ ,\ B = \theta_2 \text{ and hence } -\infty < A < 0,\ -1 < B < \infty,$$

$$V(A,B,J) = \int_0^1 e^{Ax} x^{B+J} dx,$$

$$W(A,B,J) = \int_0^1 \log x\ e^{Ax} x^{B+J} dx, \qquad (4.19)$$

$$Z(A,B,J) = \int_0^1 (\log x)^2 e^{Ax} x^{B+J} dx.$$

In (4.19), $J = 0,1,2,\ldots.$ With the notations (4.19), it is easy to see that the m.l. equations may be written as

$$U(A,B) = \begin{bmatrix} U_1(A,B) \\ U_2(A,B) \end{bmatrix} = 0, \qquad (4.20)$$

where

$$U_1(A,B) = \frac{V(A,B,1)}{V(A,B,0)} - \bar{t}_1, \text{ and } \quad U_2(A,B) = \frac{W(A,B,0)}{V(A,B,0)} - \bar{t}_2.$$

We may solve (4.20) by Newton's method or by IMSL routine ZSPOW. By Newton's method, a solution at (i+1)th iteration is

$$\begin{bmatrix} A^{(i+1)} \\ B^{(i+1)} \end{bmatrix} = \begin{bmatrix} A^{(i)} \\ B^{(i)} \end{bmatrix} - [J(A^{(i)})]^{-1} U(A^{(i)}, B^{(i)}),$$

where $i=0,1,2,\ldots,$ and $(A^{(0)}, B^{(0)})$ is an initial guess and $J(A,B)$ is given by

$$J(A,B) = \begin{bmatrix} \partial U_1/\partial A \ , & \partial U_1/\partial B \\ \partial U_2/\partial A \ , & \partial U_2/\partial B \end{bmatrix}. \tag{4.21}$$

It is easy to see that

$$\frac{\partial U_1}{\partial A} = \frac{V(A,B,2)}{V(A,B,0)} - \left[\frac{V(A,B,1)}{V(A,B,0)}\right]^2,$$

$$\frac{\partial U_2}{\partial A} = \frac{\partial U_1}{\partial B} = \frac{W(A,B,1)}{V(A,B,0)} - \frac{V(A,B,1)\cdot W(A,B,0)}{[V(A,B,0)]^2},$$

and

$$\frac{\partial U_2}{\partial B} = \frac{Z(A,B,0)}{V(A,B,0)} - \left[\frac{W(A,B,0)}{V(A,B,0)}\right]^2.$$

We can check that

$$\frac{\partial U_1}{\partial A} = \mathrm{Var}(X); \quad \frac{\partial U_1}{\partial B} = \mathrm{Cov}(X, \log X); \quad \text{and } \frac{\partial U_2}{\partial B} = \mathrm{Var}(\log X).$$

Hence at each iteration, we need to evaluate the integrals (V(A,B,J), j=0,1,2), (W(A,B,J), j=0,1), and Z(A,B,0). Note that zero is the point of singularity of the integrals (V(A,B,0), W(A,B,0), Z(A,B,0)) whenever $-1 < B < 0$. The integration by parts would remove the singularity and the same can be seen in the following recurrence relation. We obtain the following results by integrating (4.19) by parts for j=0,1,2,....

$$\text{i)}\quad V(A,B,J) = [e^{A} - A\ V(A,B,J+1)]/(B+J+1), \tag{4.22}$$

$$\text{ii)}\quad W(A,B,J) = -\ [V(A,B,J) + A\ W(A,B,J+1)]/(B+J+1),$$

$$\text{iii)}\quad Z(A,B,J) = -\ [2W(A,B,J) + A\ Z(A,B,J+1)]/(B+J+1).$$

We first compute V(A,B,J), for some $j > 1$, and hence we can compute the other integrals using the recurrence relation (4.22). The key integral in solving the m.l. equations is V(A,B,J). It may be noted that IMSL routine MDGAM computes V(A,B,J), $J > 0$, efficiently using Chebychev's polynomials. We use the same in computing the integrals (4.22). IMSL routine MDGAM provides

$$p_0(\beta,\alpha) = \frac{1}{\Gamma\alpha}\int_0^\beta e^{-u}u^{\alpha-1}du,$$

and hence we can check that

$$V(A,B,J) = (-1/A)^{(B+J+1)}\Gamma(B+J+1)p_0(-A,B+J+1). \qquad (4.23)$$

c) The Information Matrix

The information matrix I(A,B) can be obtained as done in the case of doubly truncated normal distribution in Chapter three. That is, I(A,B) = J(A,B) where J(A,B) is the same as in (4.21). Given I(A,B), we can easily see that

$$I(\theta_1,\theta_2) = R\ I(A,B)R'$$

where

$$R = \begin{bmatrix} \frac{\partial A}{\partial\theta_1} , & \frac{\partial A}{\partial\theta_2} \\ \frac{\partial B}{\partial\theta_1} , & \frac{\partial B}{\partial\theta_2} \end{bmatrix},$$

$$= \begin{bmatrix} 1 , & 0 \\ 0 , & 1 \end{bmatrix}.$$

Hence $I(A,B) = I(\theta_1,\theta_2)$. Now given I(A,B), we can derive the information matrix under the (β,α) parameterization as

$$I(\beta,\alpha) = Q\, I(\beta,\alpha) Q',$$

where

$$Q = \begin{bmatrix} U/\beta^2 & , & 0 \\ 0 & & 1 \end{bmatrix} .$$

Since $\sqrt{n}\,(\hat{\theta} - \theta) \sim$ Multivariate Normal $(0, I^{-1}(\theta))$, the asymptotic variance - covariance matrix of the m.l.e. is

$$\sum_{\hat{\theta}} = \frac{1}{n}\,[I^{-1}(\theta)] \;. \qquad (4.24)$$

4.4 Bayes Modal Estimation

Here we consider the Bayes modal estimation in two parameter truncated gamma family $(\mathcal{B}_2)$. We derive the Bayes modal estimator as defined and discussed in Section 2.5a of Chapter two. For the sake of simplicity of the discussion, we rewrite the truncated gamma distribution (4.4) as

$$f(x;A,B) = \frac{e^{Ax+B\log x}}{\left[\int_0^1 e^{Ax+B\log x}\right]}, \qquad (4.25)$$

where $A = \theta_1 - 1$, $B = \theta_2$ and hence $-\infty < A < 0$, $-1 < B < \infty$. In further discussions in this section, $E(X)$ and $E(\log X)$ are assumed to be evaluated with respect to (4.25). We consider the conjugate prior, the gamma prior, for A and the noninformative prior for B. In other words, prior for (A,B) is

$$p(A,B) \propto e^{aA+b\log(-A)},$$

where $0 < a < \infty$; $-1 < b < \infty$;, are the parameters of the prior. Hence the modified likelihood of the sample of i.i.d. observations from (4.25) is

$$L^*(A,B,t_1,t_2) \propto e^{(t_1+a)A+Bt_2-K(A,B)+b\log(-A)},$$

where $t_1 = \Sigma x_i$, $t_2 = \Sigma \log x_i$, and $K(A,B) = \log\left[\int_0^1 e^{Ax+B\log x}\right]^n$.

The Bayes modal estimator of (A,B) is the solution of

$$E(X) - a/n - (b/n)/A = \bar{t}_1, \qquad \text{and}$$

$$E(\log X) = \bar{t}_2 , \tag{4.26}$$

where $E(X) = \frac{1}{n} \frac{\partial K}{\partial A}$, $E(\log X) = \frac{1}{n} \frac{\partial K}{\partial B}$.

Except for the term $(a/n + b/nA)$ in the first equation, (4.26) is the same as the m.l. equations (4.10). Hence the question of existence of a solution to (4.26) can easily be studied using Theorem 4.1. Let

$$C_{m_2} = [(A,B) : E(\log X) = m_2] , -\infty < m_2 < 0.$$

Here C_{m_2} is the subset of the parameter space of (A,B) with $E(\log X)$ being constant. Again arguing on similar lines of Theorem 4.1, we can show that, for (A,B) in C_{m_2},

$$e^{m_2} < E(X) < 1/(1-m_2). \tag{4.27}$$

Also it is important to note that $-\infty < A < 0$ in C_{m_2}. Using these two facts, we state a result about the existence of a solution to equations (4.26) next.

<u>Theorem 4.3</u>: Given $(\bar{t}_1,\bar{t}_2)$, $0 < \bar{t}_1 < 1$, a solution to the Bayes modal equations (4.26) exists with probability one provided $b > 0$.

Proof: Given $(\bar{t}_1, \bar{t}_2)$, there exists a set $C_{\bar{t}_2}$ in the space of (A,B) and the solution to (4.26) lies in $C_{\bar{t}_2}$ where

$$C_{\bar{t}_2} = [(A,B) : E(\log X) = \bar{t}_2], \ -\infty < \bar{t}_2 < 0.$$

Since $-\infty < A < 0$ on $C_{\bar{t}_2}$, using (4.27), we can show that

$$\lim_{A \to -\infty} [E(X) - a/n - b/(nA)] = e^{\bar{t}_2} - a/n,$$

and

$$\lim_{A \to 0} [E(X) - a/n - b/(nA)] = \infty, \text{ provided } b > 0.$$

Before concluding the result, we should note the following:

i) $E(X)$ is strictly monotone on $C_{\bar{t}_2}$ due to (4.8),

ii) The range of $\bar{t}_1$ is $[e^{\bar{t}_2}, 1]$, for a given $\bar{t}_2$.

Hence the result.

The previous analysis holds for any $a > 0$, and $b > 0$ in the prior distribution. However, the question is how to choose the optimum values of a and b. As mentioned in Section 2.5a, we may use Blumenthal's analysis of one parameter family since we are using noninformative prior for B. In fact, Mittal [1984] shows that $a = \frac{1}{2}$, $b = 1$ are the optimum values in the sense of asymptotic minimum bias as derived in (2.29). Hence we use these values in simulation studies.

4.5 Simulation Results

Here, we discuss the comparative behaviour of the Bayes modal and the mixed estimators based on 500 random samples of sizes 10, 20, and 40. The mixed estimator is defined to be the mixture of the maximum likelihood estimator and the Bayes modal estimator based on the necessary and sufficient condition (4.14). All the simulated results are presented for the (θ_1, θ_2) parameterization. The details are discussed next in subsections a) - b).

a) Plan of Simulation: The general sketch of the simulation work in this family is similar to the one described in Section 3.4 except that the truncation interval here is fixed to be [0,1] in advance. The parameters chosen for simulations are:

$$\theta_1 = -3, -1, -0.3, 0, \text{ and } \theta_2 = -0.6, -0.4, 0, 1.0.$$

The parameters are chosen such that the truncation probability (q_0) ranges approximately in the interval [0.01, 0.25]. The samples are drawn using IMSL subroutine GGAMR. Here, the sample statistics of interst are

$$\bar{t}_1 = \frac{1}{n} \Sigma\, x_i \ , \ \bar{t}_2 = \frac{1}{n} \Sigma \log x_i \ , \text{ and } 0 \le x \le 1.$$

Once $(\bar{t}_1,\bar{t}_2)$ are computed, we obtain the maximum likelihood estimates $(\hat{\theta}_1,\hat{\theta}_2)$, the Bayes modal estimates $(\tilde{\theta}_1,\tilde{\theta}_2)$, and the mixed estimates $(\hat{\theta}_{1m},\hat{\theta}_{2m})$ using the appropriate equations as done in Chapter three. Note that the computation of the bias vector, the probability of nearness, the mean square error (MSE), and the probability of non-existence of m.l.e. are as defined in Section 3.4.

b) Comments on Tabale 4.1: Table 4.1 contains the simulated values of the bias (length and direction), the MSE, the probability of nearness of the mixed and the Bayes modal estimators. We next present the relative performances of the Bayes modal and the mixed estimators with respect to these criteria.

i) Bias: In Table 4.1, "Bias" represents the length of the bias vector and "Dir" denotes the direction.

We can see that bias of the Bayes modal is less than bias of the mixed estimator in majority of the cases. The directional angles of both these estimators vary in a small neighborhood of 170 degrees. In other words, both the estimators tend to underestimate θ_1 and overestimate θ_2. From the direction angle, it is interesting to note that θ_2 component of bias vector is relatively smaller than θ_1 component. Also, bias (of both the estimators) is small for $n = 40$ indicating the consistency.

ii) Probability of Nearness: In Tabale 4.1, p_1 represents the probability of nearness of the Bayes modal with respect to the mixed estimator. It is clear that $p_1 > 0.50$ always. Hence the Bayes modal has the property of being nearer to the true parameter more often compared to the mixed estimator.

iii) Efficiency: The MSE's of the Bayes modal and the mixed estimators are compared through the efficiency as defined in Section 3.4. From Table 4.1, we can see that the efficiency of the Bayes modal is greater than one for all values of (θ_1,θ_2) and for all sample sizes considered here. However, as n becomes large, both the estimators tend to have the same MSE's. In the table, v_1 and v_2 represent the asymptotic variances of the maximum likelihood estimators of θ_1 and θ_2 respectively and

$(v_1 + v_2)$ is comparabale with the MSE for large n. We can see from the table that the MSE's of both the estimators are close to $(v_1 + v_2)$ for $n = 40$.

iv) The Probability of Non-existence: In Table 4.1, p_0 represents the probability of non-existence of the m.l.e. Note that p_0 is computed using the result (4.14) We can see that p_0 ranges in the interval (0,0.18). It is important to note that p_0 is an increasing function of the truncation probability (q_0).

We conclude the discussion by noting that the Bayes modal estimator is preferable to the mixed estimator with respect to all the criteria chosen for comparison.

Table 4.1 Simulated Expected Bias and MSE of the Mixed, and the Bayes Modal Estimators of (θ_2, θ_2) for Truncated Gamma Distribution Prior Parameters: $a = \frac{1}{2}$, $b = 1$

n	q_0	p_0	Mixed Estimates Bias	Dir	MSE	Bayes Modal Estimates Bias	Dir	p_1	MSE	EFF	Asy. Variances v_1	v_2	v_1+v_2
$\theta_1 = -1.00$, $\theta_2 = -0.6$													
10	0.03	0.06	1.607	175	11.631	0.640	172	0.99	1.944	5.982	3.853	0.024	3.877
20		0.02	0.785	176	4.082	0.503	175	1.00	1.658	2.462	1.927	0.012	1.939
40		0.00	0.295	174	1.294	0.240	174	1.00	0.819	1.580	0.963	0.006	0.969
$\theta_1 = -0.30$ $\theta_2 = -0.6$													
10	0.08	0.11	1.278	174	8.799	0.841	173	0.80	2.029	4.337	3.145	0.026	3.171
20		0.09	0.593	175	2.644	0.586	175	0.80	1.358	1.947	1.572	0.013	1.585
40		0.04	0.262	173	1.034	0.359	174	0.70	0.725	1.427	0.786	0.007	0.793
$\theta_1 = 0.00$ $\theta_{23} = -0.6$													
10	0.12	0.17	1.488	173	7.531	1.097	172	0.74	2.463	3.058	2.869	0.027	2.896
20		0.13	0.609	174	1.931	0.706	174	0.62	1.264	1.528	1.434	0.013	1.447
40		0.10	0.253	175	0.693	0.410	175	0.54	0.581	1.193	0.717	0.007	0.724
$\theta_1 = -1.00$ $\theta_2 = -0.4$													
10	0.06	0.08	1.381	171	9.973	0.567	168	1.00	1.889	5.284	3.858	0.064	3.922
20		0.03	0.631	171	2.903	0.432	169	1.00	1.319	2.200	1.929	0.032	1.961
40		0.01	0.244	171	1.169	0.215	171	1.00	0.764	1.529	0.965	0.016	0.981
$\theta_1 = -0.30$ $\theta_2 = -0.4$													
10	0.13	0.15	1.390	170	7.661	0.909	170	0.85	2.147	3.568	3.355	0.070	3.425
20		0.09	0.695	170	2.872	0.674	170	0.76	1.540	1.865	1.678	0.035	1.713
40		0.04	0.310	170	1.040	0.409	171	0.70	0.755	1.378	0.839	0.018	0.857

Table 4.1 (Continued)

n	q_0	p_0	Mixed Estimates Bias	Dir	MSE	Bayes Modal Estimates Bias	Dir	p_1	MSE	EFF	Asy. Variances v_1	v_2	v_1+v_2
$\theta_1 = -1.00$, $\theta_2 = -0.2$													
10	.06	.08	1.530	168	13.005	0.613	165	0.99	2.158	6.026	4.250	0.133	4.383
20		.03	0.660	169	3.201	0.443	169	0.99	1.416	2.261	2.125	0.066	2.191
40		.01	0.228	165	1.207	0.201	165	1.00	0.764	1.571	1.062	0.033	0.095
$\theta_1 = -1.00$, $\theta_2 = 0.0$													
10	.14	.08	1.709	164	13.862	0.657	160	0.99	2.189	6.333	4.872	0.239	5.111
20		.06	0.775	164	4.555	0.473	163	0.99	1.793	2.540	2.436	0.119	2.555
40		.01	0.191	157	1.522	0.175	157	0.99	0.904	1.683	1.218	0.060	1.278
$\theta_1 = -3.00$, $\theta_2 = 1.0$													
10	.09	.01	2.735	159	28.112	0.498	355	0.70	2.434	11.557	10.465	1.114	11.579
20		.00	1.125	158	9.004	0.327	327	0.60	2.988	3.918	5.233	0.557	5.790
40		.00	0.527	159	4.227	0.177	177	0.54	2.107	2.006	2.616	0.227	2.843
$\theta_1 = -1.00$, $\theta_2 = -1.0$													
10	0.40	0.16	2.998	157	35.026	0.588	148	0.98	1.710	20.485	11.395	1.791	13.186
20		0.12	1.423	153	13.934	0.492	151	0.99	2.165	6.434	5.697	0.896	6.593
40		0.06	0.896	155	5.488	0.551	154	0.99	2.173	2.526	2.849	0.448	3.297

Index: q_0: Truncation Probability

p_0: Probability of non-existence (m.l.e.)

p_1: Probability of nearness (Bayes modal, Mixed)

v_1: Asy. Variance of $\hat{\theta}_1$(m.l.e.)

v_2: Asy. Variance of $\hat{\theta}_2$(m.l.e.)

5. SINGLY TRUNCATED LOG-NORMAL (or NORMAL) DISTRIBUTION

This chapter deals with a singly truncated log-normal which is equivalent to a singly truncated normal distribution. The scheme of presentation is quite the same as in Chapters three and four.

i) Section 5.1 deals with family representation.

ii) Section 5.2 deals with a derivation of the maximum likelihood equations and the question of existence of a solution.

iii) Section 5.3 deals with computational results.

iv) Section 5.4 deals with a derivation of the Bayes modal equations and the question of existence of a solution.

v) Section 5.5 deals with simulation results.

5.1 Family Representation

A singly truncated log-normal population is described by a random variable Y with density

$$f(y;\mu,\sigma^2) = \frac{\left(\frac{1}{y}\right) e^{-\frac{1}{2}\left(\frac{\log y-\mu}{\sigma}\right)^2}}{\left[\int_0^U \frac{1}{u} e^{-\frac{1}{2}\left(\frac{\log u-\mu}{\sigma}\right)^2} du\right]}, \qquad (5.1)$$

where $0 \leq y \leq U$, $U > 0$ is known, and $(\mu,\sigma^2) \epsilon \Omega$ with

$$\Omega = [(\mu,\sigma^2) : -\infty < \mu < \infty, \quad 0 < \sigma^2 < \infty].$$

Hence the indexed family of joint distributions of i.i.d. random variables $(Y_1 \ldots Y_n)$ is

$$\mathcal{B}_3 = [L(\underline{y};\ \mu,\sigma^2) : \underline{y} = (y_1 \ldots y_n) \ \epsilon\ \bar{\underline{Y}},\ (\mu,\sigma^2)\epsilon\Omega],$$

where

$$\bar{\underline{Y}} = [\underline{y}\text{: } 0 \leq y_i \leq U,\ i=1,2,\ldots,n,\ U \text{ as in } (5.1)],$$

and

$$L(\underline{y};\mu,\sigma^2) = \frac{\left[e^{\frac{\mu}{\sigma^2}\Sigma \log y_i - \frac{1}{2\sigma^2}\Sigma(\log y_i)^2}\right]\prod_{i=1}^{n}\left[\frac{1}{y_i}\right]}{\left[\int_{-\infty}^{\log U} e^{\left[\frac{\mu - \log U}{\sigma^2}\right]U - \frac{1}{2\sigma^2}u^2} du\right]^n} \tag{5.2}$$

By changing $X = Y - \log U$ in (5.1), the singly truncated log-normal becomes the singly truncated normal distribution with density

$$f(x;\mu,\sigma^2) = \frac{e^{\left[\frac{\mu-\log U}{\sigma^2}\right]x - \frac{1}{2\sigma^2}x^2}}{\left[\int_{-\infty}^{\log U} e^{\left[\frac{\mu-\log U}{\sigma^2}\right]u - \frac{1}{2\sigma^2}u^2} du\right]}, \tag{5.3}$$

where $-\infty < x_i \leq 0$, and $(\mu,\sigma^2)\epsilon\,\Omega$.

Now the indexed family of joint distributions of i.i.d. random variables $(X_1 \ldots X_n)$ with the density (5.3) is

$$\mathscr{B}_4 = [L(\underline{x};\mu,\sigma^2) : \underline{x} = (x_1 \ldots x_n)\epsilon\,\mathfrak{X}, \quad (\mu,\sigma^2)\epsilon\,\Omega],$$

$$\mathfrak{X} = [\underline{x}: -\infty < x_i \leq 0, \quad i=1,2,\ldots,n],$$

and

$$L(\underline{x};\mu,\sigma^2) = \frac{e^{\left[\frac{\mu-\log U}{\sigma^2}\Sigma x_i - \frac{1}{2\sigma^2}\Sigma x_i^2\right]}}{\left[\int_{-\infty}^{0} e^{\left[\frac{\mu-\log U}{\sigma^2}\right]u - \frac{1}{2\sigma^2}u^2} du\right]^n} . \tag{5.4}$$

We can express $\mathscr{B}_4$ in a minimal exponential family as defined in (2.1) with

$$a(\mu,\sigma^2):\Omega\longrightarrow R,\ a(\mu,\sigma^2)=\left[\int_{-\infty}^{0} e^{\tau_1(\mu,\sigma^2)u+(\tau_2(\mu,\sigma^2)-1)u^2} du\right]^{-n},$$

$$b(\underline{x}):\ \mathfrak{X}\longrightarrow R,\quad b(\underline{x}) = e^{-\Sigma x_i^2},$$

$$T(\underline{x}):\mathfrak{X}\rightarrow S,\ S = [(t_1,t_2):t_1=\Sigma x_i,\ t_2=\Sigma x_i^2,\ (x_1\ldots x_n)\in\mathfrak{X}],$$

$$\tau(\mu,\sigma^2):\Omega\rightarrow\Theta,$$

$$\Theta=[(\theta_1,\theta_2):\theta_1=\tau_1(\mu,\sigma^2),\ \theta_2=\tau_2(\mu,\sigma^2),\ (\mu,\sigma^2)\in\Omega],$$

and

$$\tau_1(\mu,\sigma^2) = \frac{\mu-\log U}{\sigma^2},\quad \tau_2(\mu,\sigma^2) = 1 - \frac{1}{2\sigma^2} .$$

Note that τ is one-one with the inverse mapping

$$\mu = \frac{\theta_1}{2(1-\theta_2)} + \log U, \quad \text{and} \quad \sigma^2 = \frac{1}{2(1-\theta_2)} .$$

We use this inverse mapping in getting the maximum likelihood estimators of (μ,σ^2) from the maximum likelihood estimators of (θ_1,θ_2). We may represent (5.3) under the (θ_1,θ_2) parameterization as

$$f(x;\theta_1,\theta_2) = \frac{e^{\theta_1 x+(\theta_2-1)x^2}}{\left[\int_{-\infty}^{0} e^{\theta_1 u+(\theta_2-1)u^2} du\right]} . \tag{5.5}$$

The representation (5.5) is quite useful in our later discussions.

The distribution of $T(\underset{\sim}{X}) = (\Sigma X_i, \Sigma X_i^2)$ is of the form

$$g(t;\theta) = h(t)e^{\theta\cdot t-K(\theta)}, \ t \in S, \ \theta \in \Theta, \tag{5.6}$$

where $K(\theta) = \log c(\theta)$, and $c(\theta)$ is given by

$$c(\theta) = \left[\int_{-\infty}^{0} e^{\theta_1 u+(\theta_2-1)u^2} du\right]^n .$$

Here, Θ being one-one mapping of Ω, we can easily see that

$$\Theta = [(\theta_1,\theta_2): -\infty < \theta_1 < \infty, -\infty < \theta_2 < 1].$$

It can be verified that the natural parameter space is

$$\bar{\Theta} = \left[(\theta_1,\theta_2): \int_{-\infty}^{0} e^{\theta_1 x+(\theta_2-1)x^2} dx < \infty\right],$$

$$= \left[(\theta_1,\theta_2): -\infty<\theta_1\leq 0, -\infty<\theta_2<1\right] \cup \left[(\theta_1,\theta_2): 0<\theta_1<\infty, -\infty<\theta_2\leq 1\right].$$

It is interesting to note that the natural parameter space $(\bar{\Theta})$ and the parameter space of interest (Θ) differ only in the partial boundary, $[(\theta_1,\theta_2):\theta_1>0, \theta_2 = 1]$, of Θ. Hence we may study

$$\mathcal{A} = [g(t;\theta): \theta \in \bar{\Theta}], \tag{5.7}$$

which is a full exponential family in contrast to the cases we study in Chapters three and four where it is a proper subfamily.

Next we discuss monotonicity property of the partial derivative vector $DK(\theta) = \left(\frac{\partial K}{\partial \theta_1}, \frac{\partial K}{\partial \theta_2}\right)$ of $K(\theta)$. By the definition of $K(\theta)$, we can see that

$$\frac{\partial K}{\partial \theta_1} = nE(X), \quad \text{and} \quad \frac{\partial K}{\partial \theta_2} = nE(X^2),$$

where $E(X)$ and $E(X^2)$ are evaluated with respect to (5.5). Let

$$I = [(m_1, m_2): m_1 = E(X),\ m_2 = E(X^2), \text{ and } \theta \in \Theta]. \qquad (5.8)$$

Then I is the range of $(\frac{1}{n})DK(\theta)$. I is called the population moment space. Hence from (2.7), $\frac{1}{n}\ DK(\theta)$ satisfies the montonicity property

$$(\theta_1 - \tilde{\theta}_1)(m_1 - \tilde{m}_1) + (\theta_2 - \tilde{\theta}_2)(m_2 - \tilde{m}_2) > 0 \qquad (5.9)$$

where (m_1, m_2), $(\tilde{m}_1, \tilde{m}_2)$ are values of $(\frac{1}{n})DK(\theta)$ at θ and $\tilde{\theta}$ respectively.

The said montonicity result (5.9) implies the following:

i) The strict montonicity of $E(X)$ for fixed θ_2.
ii) The strict montonicity of $E(X^2)$ for fixed θ_1.
(5.10)

Above results are used in discussing the maximum likelihood estimators in certain subfamilies of $\mathscr{A}$.

5.2 Maximum Likelihood Estimation

Given t, the log-likelihood function is

$$l(\theta;t) = \theta\cdot t - K(\theta) + \log h(t),$$

where $K(\theta)$ and $h(t)$ are defined in (5.7). By the definition (2.8), the maximum likelihood estimator of θ is a solution of

$$DK(\theta) = t, \quad \theta \in \Theta.$$

Let $\bar{t}_1 = t_1/n$, $\bar{t}_2 = t_2/n$ and define

$$\bar{S} = [(\bar{t}_1,\bar{t}_2): (t_1,t_2) \in S].$$

We may call $\bar{S}$ the sample moment space as in earlier chapters. Hence we may equivalently rewrite the equation $DK(\theta) = t$ as

$$E(X) = \bar{t}_1,$$
$$E(X^2) = \bar{t}_2, \qquad (5.11)$$

where $E(X) = \int_{-\infty}^{0} xf(x;\theta)dx$, and $E(X^2) = \int_{-\infty}^{0} x^2 f(x;\theta)dx$ and

$f(x;\theta)$ is defined in (5.5). The equations (5.11) are referred to as the maximum likelihood equations.

Our main concern is the question of existence of a solution to (5.11). Since we are dealing with a full exponential family, Theorem 2.4a of Chapter two is applicable. As stated in Theorem 2.4a, we are required to check the steepness of $\mathscr{A}$, the steepness of $K(\theta)$, in order to know the existence of a solution. Next in Proposition 5.1, we state the result about the steepness of $K(\theta)$.

<u>Proposition 5.1</u>: Let $\mathscr{A} = [g(t;\theta):\theta \text{ in } \bar{\Theta}]$, $g(t;\theta)$ as defined in (5.6). Then $\mathscr{A}$ is not steep.

Proof: Let $(\theta_1^*,1)$, $\theta_1^*>0$, be a boundary point of $\bar{\Theta}$. Then we can check that the following results hold:

$$\lim_{\substack{\theta_1\to\theta_1^* \\ \theta_1\to 1}} E(X) = \lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to 1}} \int_{-\infty}^{0} x\, f(x;\theta)\,dx = -\frac{1}{\theta_1^{*2}},$$

and

$$\lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to 1}} E(X) = \lim_{\substack{\theta_1\to\theta_1^* \\ \theta_2\to 1}} \int_{-\infty}^{0} x^2 f(x;\theta)\,dx = \frac{2}{\theta_1^{*2}}.$$

In order that $K(\theta)$ is steep, at least one of the above limits must be infinite. Since both the limiting quantities are finite for any $\theta_1^* > 0$, the result follows.

By Propositon 5.1, we know that $K(\theta)$ is not steep and hence the interior of convex hull of $\bar{S}$ is not equal to the set I (Theorem 2.3) implying that the equations (5.11) do not admit a solution with probability one. The fact that these equations do not admit a solution is not useful in itself for practical applications unless we obtain a necessary and sufficient condition in terms of $(\bar{t}_1, \bar{t}_2)$. This result is shown in Theorem 5.1 and Corollary 5.1.

Theorem 5.1: Let $D_{m_1} = [(\theta_1, \theta_2): E(X) = m_1]$, $-\infty < m_1 < 0$. Then for (θ_1, θ_2) in D_{m_1}, the following inequality holds:

$$m_1^2 < E(X^2) < 2m_1^2. \tag{5.12}$$

Proof: Let Θ^* be the boundary of $\bar{\Theta}$. It is easy to see that

$$\Theta^* = [(\theta_1^*, 1): -\infty < \theta_1^* < \infty].$$

Since $E(X^2) > (E(X))^2$ is the well known result, we only need to show the upper inequality in (5.12). Given any boundary point $(\theta_1^*, 1)$ in Θ^*, we can obtain the following:

Let

$$F_1(\theta_1^*) = \lim_{\substack{\theta_1 \to \theta_1^* \\ \theta_2 \to 1}} \frac{\int_{-\infty}^{0} x\, e^{\theta_1 + (\theta_2 - 1)x^2} dx}{\left[\int_{-\infty}^{0} e^{\theta_1 x + (\theta_2 - 1)x^2} dx\right]}, \tag{5.13a}$$

and

$$F_2(\theta_1^*) = \lim_{\substack{\theta_1 \to \theta_1^* \\ \theta_2 \to 1}} \frac{\int_{-\infty}^{0} x^2 e^{\theta_1 x + (\theta_2 - 1)x^2} dx}{\left[\int_{-\infty}^{0} e^{\theta_1 x + (\theta_2 - 1)x^2} dx\right]}. \tag{5.13b}$$

By taking the limit under the intergal sign, we see that,

$$\begin{aligned} F_1(\theta_1^*) &= -1/\theta_1^{*2}, \quad \theta_1^* > 0, \\ &= -\infty \quad\;\;\;, \quad \theta_1^* \le 0, \end{aligned}$$

and

$$\begin{aligned} F_2(\theta_1^*) &= 2/\theta_1^{*2} \;\;, \quad \theta_1^* > 0, \\ &= \infty \quad\;\;\;\;, \quad \theta_1^* \le 0. \end{aligned}$$

It is easy to see that

i) $F_1(\theta_1^*)$ is strictly increasing in $(-\infty\ ,\ \infty)$, with the range $(-\infty, 0)$

ii) $F_2(\theta_1^*)$ is strictly decreasing in $(-\infty\ ,\ \infty)$, with the range $(0,\infty)$.

Given any m_1 in $(-\infty, 0)$, due the above monotonicity properties of $F_1(\theta_1^*)$ and $F_2(\theta_1^*)$, there exists a unique $\theta_1^*>0$ such that

$$F_1(\theta_1^*) = m_1, \quad \text{and} \quad F_2(\theta_1^*) = 2\ m_1^2.$$

In other words, $(\theta_1^*,1)$ is the boundary point of D_{m_1}. Hence using the monotonicity result (5.9) and the result (5.13b), we have

$$\sup_{D_{m_1}} E(X^2) = 2\ m_1^2,$$

which proves the result.

Theorem 5.1 gives an explicit expression for the set I, the population moment space. That is,

$$I = [(m_1,m_2):\ -\infty<m_1<0,\ m_1^2<m_2< 2m_1^2]. \tag{5.14}$$

Having proved Theorem 5.1, obtaining a necessary and sufficient condition for the existence of a solution to the equation(5.11) is straight forward and stated in Corollary 5.1 below.

<u>Corollary 5.1</u>: Given $(\bar{t}_1,\bar{t}_2)$, $-\infty < \bar{t}_1 < 0$, the maximum likelihood equations (5.11) admit a solution in the interior of $\bar{\Theta}$ if and only if

$$\bar{t}_1^2 < \bar{t}_2 < 2\,\bar{t}_1^2. \tag{5.15}$$

Proof: Given $(\bar{t}_1,\bar{t}_2)$, it is clear that a solution to the equation (5.11) must lie in

$$D_{\bar{t}_1} = [(\theta_1,\theta_2): E(X) = \bar{t}_1].$$

By Theorem 5.1, on $D_{\bar{t}_1}$, for $\bar{t}_1 < 0$,

$$\bar{t}_1^2 < E(X^2) < 2\,\bar{t}_1^2 .$$

Hence the result.

It may be worth noting that the m.l. equations (5.11) in singly truncated normal distribution do not admit a solution, whenever $2\bar{t}_1^2 \leq \bar{t}_2$, in the full family

$$\mathscr{A} = [g(t;\theta): \theta \text{ in } \bar{\Theta}],$$

unlike the cases of doubly truncated normal and truncated gamma distributions wherein the non-existence is only in a proper subfamily. Next, we deal with the question of existence of the maximum likelihood estimator in the following subfamilies

i) $\mathscr{A}_{\theta_2} = [g(t;\theta): -\infty < \theta_1 < \infty], -\infty < \theta_2 < 1$ is known,

ii) $\mathscr{A}_{\theta_1} = [g(t;\theta): -\infty < \theta_2 < 1], -\infty < \theta_1 < \infty$ is known.

These subfamilies have the same meaning as explained in the case of doubly truncated normal distributions. Hence the cases of μ-known and σ^2-known can be dealt in a slightly different representation and we skip them here. It is easy to check that the m.l. estimator of θ_1 in $\mathscr{A}_{\theta_2}$ is the solution of

$$E(X) = \bar{t}_1 , -\infty < \bar{t}_1 < 0, \tag{5.16}$$

and the m.l. estimator of θ_2 in $\mathscr{A}_{\theta_1}$ is the solution of

$$E(X^2) = \bar{t}_2 , 0 < \bar{t}_2 < \infty. \tag{5.17}$$

The monotonicity result (5.10) implies the following:

i) $E(X)$ is strictly increasing in $\mathscr{A}_{\theta_2}$ with the range $(-\infty, 0)$ for all $-\infty < \theta_2 < 1$.

ii) $E(X^2)$ is strictly increasing in $\mathscr{A}_{\theta_1}$ with the range

$(0,\infty)$ for all $\theta_1 \le 0$ and $E(X^2)$ is strictly increasing in $\mathcal{A}_{\theta_1}$ with the range $(0, 2/\theta_1^2)$ for all $\theta_1 > 0$.

Hence equation (5.16) always admits the solution in $\mathcal{A}_{\theta_2}$. But the equation (5.17) admits the solution in $\mathcal{A}_{\theta_1}$ with probability one provided $\theta_1 \le 0$ and does not admit a solution whenever $\bar{t}_2 \ge 2/\theta_1^2$ for $\theta_1 > 0$.

In Theorem 5.2 below, we state a result on characterizing the set $\bar{S}$, the sample moment space.

<u>Theorem 5.2</u>: Let $-\infty < x_i \le 0$, $i=1,2,\ldots n$, $\bar{t}_1 = \frac{1}{n}\sum x_i$, $\bar{t}_2 = \frac{1}{n}\sum x_i^2$. Then,

$$\bar{t}_1^2 \le \bar{t}_2 \le n\bar{t}_1^2. \tag{5.18}$$

Proof: Let $\mathcal{X} = [(x_1 \ldots x_n): -\infty < x_i \le 0,\ i = 1,2,\ldots n]$, and

$$\mathcal{X}_{\bar{t}_1} = [(x_1 \ldots x_n) \in \mathcal{X}: \frac{1}{n} \sum x_i = \bar{t}_1].$$

Note that $\mathcal{X}_{\bar{t}_1}$ is a hyper plane intersecting $\mathcal{X}$. Since

$\bar{t}_2 \geq \bar{t}_1^2$ is the known result (due to $E(X^2) \geq (EX^2)$), we only need to show the upper inequality in (5.18). Since $\bar{t}_2 = (\frac{1}{n})\Sigma x_i^2$ is a convex function, it attains its maximum on the set of extreme points of $\mathfrak{X}_{\bar{t}_1}$. The set of extreme points of $\mathfrak{X}_{\bar{t}_1}$ is just the permutations of $(n\bar{t}_1, 0, \ldots 0)$. Hence on the similar lines of arguments of Theorem 3.2, we get

$$\begin{aligned} \max_{\substack{\mathfrak{X} \\ \bar{t}_1}} \bar{t}_2 &= 1/n[n^2\bar{t}_1{}^2], \\ &= n\, \bar{t}_1{}^2. \end{aligned}$$

Hence the result.

Theorem 5.2 figures out the set $\bar{S}$ explicitely as

$$\bar{S} = [(\bar{t}_1, \bar{t}_2) : -\infty < \bar{t}_1 < 0,\ \bar{t}_1^2 \leq \bar{t}_2 \leq n\ \bar{t}_1^2]. \qquad (5.19)$$

Let C denote the convex hull of $\bar{S}$. Then, it is easy to check that

$$C = [(\bar{t}_1, \bar{t}_2) : -\infty < \bar{t}_1 \leq 0,\ \bar{t}_1{}^2 \leq \bar{t}_2 < \infty].$$

As in the cases of doubly truncated normal and truncated gamma, limit $\bar{S} = C$ as n goes to infinity. A natural question now to ask is, if the limit $\bar{S} = C$ is true in general? The answer is not investigated here.

Note that the results stated in Corollaries 3.2 and 3.3 are not of interest here because of the following:

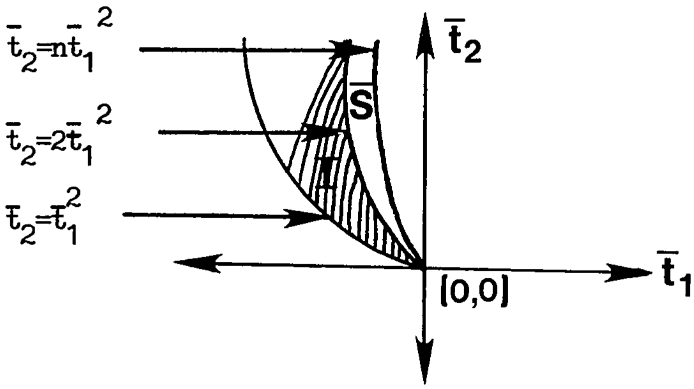
i) The supremum of the sample variance over the sample sapce is infinity.

ii) The supremum of the population variance over the parameter space of interest is infinity.

Hence we cannot state any sufficient condition for the existence or the non-existence of a solution to the m.l. equations in this family similar to the one obtained by Mittal [1984] for the case of doubly truncated normal distribution. As a final remark of this section, it is interesting to note that the sample moment space ($\bar{S}$) is the same as the population moment space (I) for n = 2. That is, the equations (5.11) admit a solution with probability one for n = 2.

Next in Figure 5.1, we show a graphical presentation of the contents of discussions in Theorem 5.1 and Theorem 5.2. As done in Chapters three and four, in Figure 5.1, the set I is superimposed on the set $\bar{S}$ and hence the

m_1-axis is the same as the $\bar{t}_1$-axis and the m_2-axis is the same as the $\bar{t}_2$-axis.



Population and Sample Moment Spaces

Figure 5.1

We may note that in the above graph, the equations (5.11) admit a solution in θ if and ony if $(\bar{t}_1, \bar{t}_2)$ belong to the shaded area.

5.3 Computational Results

Recall that the singly truncated normal density is

$$f(x;\theta) = \frac{e^{\theta_1 x + (\theta_2 - 1)x^2}}{\left[\int_{-\infty}^{0} e^{\theta_1 x + (\theta_2 - 1)x^2} dx\right]} \, .$$

The maximum likelihood estimators of (θ_1, θ_2) are the solution of

$$\int_{-\infty}^{0} x\ f(x;\theta)dx = \bar{t}_1, \tag{5.20}$$

and

$$\int_{-\infty}^{0} x^2 f(x;\theta)dx = \bar{t}_2.$$

By Corollary 5.1, we know that (5.20) admits a solution if and only if $\bar{t}_1^2 < \bar{t}_2 < 2\ \bar{t}_1^2$. Whenever the sample quantities $(\bar{t}_1, \bar{t}_2)$ satisfy the said necessary and sufficient condition, we solve (5.20) iteratively by Newton's method. We need to compute the integral

$$\int_{-\infty}^{0} x^r e^{\theta_1 x + (\theta_2 - 1)x^2} dx \tag{5.21}$$

for r = 0,1,2,3,4 at each iteration. It is possible to derive a recurrence relation to evaluate the integral (5.21) for r > 2 and we discuss the same next. Let z be any real, and

$$D_{-p}(z) = e^{-z^2/2}\ ,\ p = 0 \tag{5.22}$$

$$= \frac{e^{-z^2/4}}{\Gamma p} \int_{-\infty}^{0} x^{p-1} e^{-zx-x^2/2} dx,\ p > 0,$$

and

$$ERF(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-u^2} du.$$

In the table of integrals by Gradshteyn and Ryzhik [1980] $D_{-p}(z)$ is known as the parabolic cylinder function and ERF(z) is known as the error function. In the said table, certain results are stated about $D_{-p}(z)$ and we state them in Theorem 5.3 giving a brief sketch of the proof.

Theorem 5.3: Let $D_{-p}(z)$ be as defined in (5.22). Then the following results hold:

i) $D_{-1}(z) = \sqrt{\frac{\pi}{2}}\ e^{\frac{z^2}{4}} [1 - ERF \frac{z}{\sqrt{2}}]$,

ii) $D_{-(p+1)}(z) = (-1/p)[z\ D_{-p}(z) - D_{-(p-1)}(z)]$,

$p \geq 1$ and $D_0(z) = e^{-z^2/4}$.

Proof: By the definition of $D_{-p}(z)$ in (5.23),

$$D_{-1}(z) = e^{-z^2/4} \int_0^\infty e^{-zx-x^2/2} dx.$$

By completing the square and substituting $u = (z+x)/2$, we get

$$D_{-1}(z) = \sqrt{2}\; e^{z^2/4}\left[\int_{z/\sqrt{2}}^{\infty} e^{-u^2} du\right],$$

$$= \sqrt{2}\; e^{z^2/4}\left[\frac{\sqrt{\pi}}{2} - \frac{\sqrt{\pi}}{2}\, \mathrm{ERF}\left(\frac{z}{\sqrt{2}}\right)\right],$$

$$= \sqrt{\frac{\pi}{2}}\; e^{z^2/4}\left[1 - \mathrm{ERF}\left(z/\sqrt{2}\right)\right].$$

Hence (i) follows.

Replacing p by (p+1) in (5.23), it is obvious that

$$D_{-(p+1)}(z) = \frac{e^{-z^2/4}}{\Gamma(p+1)} \int_0^{\infty} x^p e^{-zx-x^2/2} dx,$$

$$= \frac{e^{z^2/4}}{\Gamma(p+1)} \left[\int_0^{\infty} x^{p-1}(x+z)\, e^{-\frac{(x+z)^2}{2}} dx - z\int_0^{\infty} x^{p-1} e^{-\left(\frac{z+x}{2}\right)^2} dx\right].$$

Integrating by parts and simplifying the results, we have

$$D_{-(p+1)}(z) = \frac{1}{\Gamma(p+1)}\left[(p-1)\Gamma(p-1)D_{-(p-1)}(z) - z\Gamma p\; D_{-p}(z)\right],$$

$$= -\frac{1}{p}\left[z\, D_{-p}(z) - D_{-(p-1)}(z)\right].$$

Hence (ii) follows.

Theorem 5.3 is useful for solving the equations (5.20) and it is shown next. It is obvious that

$$\int_{-\infty}^{0} x^r e^{\theta_1 x + (\theta_2 - 1)x^2} dx = (-1)^r \int_0^{\infty} x^r e^{-\theta_1 x - (1-\theta_2)x^2} dx.$$

Then using the definition (5.22) of $D_{-p}(z)$, we get

$$\int_{-\infty}^{0} x^r e^{\theta_1 x - (1-\theta_2)x^2} dx$$

$$= (-1)^r [2(1-\theta_2)]^{-\frac{r+1}{2}} \Gamma(r+1)\, e^{\frac{\theta_1^2}{8(1-\theta_2)}} D_{-(r+1)}\left[\frac{\theta_1}{\sqrt{2(1-\theta_2)}}\right]. \tag{5.23}$$

Now applying Theorem 5.3, we state a recurrence relation for the moments of singly truncated normal distribution in Theorem 5.4 below.

<u>Theorem 5.4</u>: Let X be distributed with the density function $f(x;\theta)$ as defined in (5.5) and $E(X^r)$ be the r^{th} moment of X about the origin. Then $E(X^r)$ satisfies the recurrence relation

$$E(X^r) = \frac{1}{2(1-\theta_2)}\left[\theta_1 E(X^r) + (n-1)E(X^{r-2})\right], \quad r \geq 2. \qquad (5.24)$$

Proof: Using the equation (5.23), we can check that

$$E(X^r) = \frac{(-1)^r[2(1-\theta_2)]^{-r/2}\Gamma(r+1)D_{-(r+1)}(z)}{D_{-1}(z)},$$

where $z = \theta_1/\sqrt{(2(1-\theta_2))}$.

Now using the recurrence relation (ii) of Theorem 5.3,

$$E(X^r) = (-1)^{r+1}[2(1-\theta_2)]^{-r/2}\Gamma r\left[\frac{zD_{-r}(z)}{D_{-1}(z} - \frac{D_{-(r+1)}(z)}{D_{-1}(z)}\right],$$

$$= (-1)^{r+1}[2(1-\theta_2)]^{-\frac{r}{2}}\Gamma r\left[\left\{\frac{(-1)^{r-1}}{\Gamma r}(2(1-\theta_2))^{\frac{r-1}{2}} zE(X^{r-1})\right.\right.$$

$$\left.\left. - \frac{(-1)^{r-2}}{\Gamma(r-1)}(2(1-\theta_2))^{\frac{r-2}{2}} E(X^{r-2})\right\}\right], \qquad r \geq 2.$$

By simplifying the above expresdsion, we get the required recurrence relation (5.24)

Theorem 5.4 simplifies solving the maximum likelihood equations a great deal. Let us rewrite the m.l. equations (5.20) as

$$U(\theta_1,\theta_2) = \begin{bmatrix} U_1(\theta_1,\theta_2) \\ U_2(\theta_1,\theta_2) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad (5.25)$$

where $U_1(\theta_1,\theta_2) = E(X) - \bar{t}_1$, and $U_2(\theta_1,\theta_2) = E(X^2) - \bar{t}_2$. By Newton's method, a solution to (5.25) at $(i + 1)^{th}$ iteration is

$\theta^{(i+1)} = \theta^{(i)} - [J(\theta^{(i)})]^{-1}U(\theta^{(i)})$, $i = 0,1,2,\ldots$, where $\theta^{(0)}$ is an initial guess, and

$$J(\theta) = \begin{bmatrix} \frac{\partial U_1}{\partial \theta_1} , & \frac{\partial U_1}{\partial \theta_2} \\ \frac{\partial U_2}{\partial \theta_1} , & \frac{\partial U_2}{\partial \theta_2} \end{bmatrix} .$$

It can be verified that

$$\frac{\partial U_1}{\partial \theta_1} = E(X^2) - (E(X))^2 ,$$

$$\frac{\partial U_1}{\partial \theta_2} = E(X^2) - E(X)\ E(X^2) = \frac{\partial U_2}{\partial \theta_1} ,$$

$$\frac{\partial U_2}{\partial \theta_2} = E(X^4) - (E(X^2))^2,$$

where all the expectations are with respect to $f(x;\theta)$. Now it is obvious that we need to evaluate $E(X^r)$, r = 1,2,3,4, at each iteration. Due to Theorem 5.3 and Theorem 5.4, we may state a simple algorithm to compute the above expected values as follows.

I Given (θ_1,θ_2), $z = \theta_1/\sqrt{2(1-\theta_2)}$,

II $C1 = ERF(z/\sqrt{2})$, (ERF from IMSL),

III $D(0) = Exp(-z^2/4)$,

IV $D(-1) = 1/D(0)\sqrt{\frac{\pi}{2}}\,[1 - C1]$,

V $D(-2) = -1/2[zD(-1) - D(0)]$,

VI $E(X) = -1/(2\sqrt{(1-\theta_2)})D(-2)/D(-1)$,

VII For J = 2,...M (M is a fixed integer),

$$E(X^J) = \frac{1}{2(1-\theta_2)}[\theta_1 E(J-1) + (J-1)E(J-2)],$$

Where $E(J) = E(X^J)$, $E(0) = 1$, and

$D(-J) = D_{-J}(z)$.

We can see that the entire computation work starts with the calling of IMSL subroutine ERF(z) and computing D(0). Note that ERF(z) is a highly efficient routine in calculating the integral (5.21).

b) The Information Matrix

Computation of the information matrix is similar to the case discussed in the doubly truncated normal distribution. That is, $I(\theta) = J(\theta)$ where $I(\theta)$ is the information matrix with respect to (5.5) and $J(\theta)$ is defined in (5.26). Hence the information matrix with respect to the (μ,σ^2) parameterization can be obtained as

$$I(\mu,\sigma^2) = Q\, I(\theta)\, Q' , \tag{5.26}$$

where

$$Q = \begin{bmatrix} \frac{1}{\sigma^2} , & -\frac{\mu}{\sigma^4} \\ 0 , & \frac{1}{2\sigma^4} \end{bmatrix} .$$

Hence the asymptotic variance-covariance matrix $\Sigma_{\hat{\theta}}$ of the maximum likelihood estimator can be computed as follows:

$$\Sigma_{\hat{\theta}} = \frac{1}{n}\left[I(\theta)\right]^{-1} . \tag{5.27}$$

5.4 Bayes Modal Estimation

This section deals with derivation of the Bayes modal equations and the question of existence of a solution. Analysis of the Bayes modal equations is similar to the one in the case of two parameter truncated gamma distribution. We may rewrite the singly truncated normal density (5.5) as

$$f(x;A,B) = \frac{e^{Ax+Bx^2}}{\int_{-\infty}^{0} e^{Ax+Bx^2}dx}, \tag{5.28}$$

where $A = \theta_1$, $B = (\theta_2-1)$ and hence $-\infty < A < 0$, $-\infty < B < 0$.

We consider a gamma prior for B and a noninformative prior for A. In other words, a joint prior for (A,B) is

$$p(A,B) \propto e^{aB+b\log(-B)},$$

where $a > 0$, $b > -1$, are the prior parameters.

Given a sample $\underset{\sim}{x} = (x_1,\ldots x_n)$ of i.i.d. observations from (5.28), the modified likelihood function is

$$L^*(\underset{\sim}{x};\ A,B) \propto e^{At_1+(B+a)t_2-K(A,B)+b\log(-B)},$$

where $K(A,B) = \log\left[\int_{-\infty}^{0} e^{Ax+Bx^2}dx\right]^n$.

Hence the Bayes modal estimator of (A,B) is a solution of

$$E(X) = \bar{t}_1,$$

$$E(X^2) - a/n - b/nB = \bar{t}_2, \tag{5.29}$$

where $E(X) = \frac{1}{n} \frac{\partial K}{\partial A}$, and $E(X) = \frac{1}{n} \frac{\partial K}{\partial B}$.

As in the case of truncated gamma family, the Bayes modal equations (5.29) are similar to the maximum likelihood equations except for the term (a/n + b/nB) in the second equation. Now it is easy to analyse the equations (5.29) with the help of Theorem 5.1. Let

$$D_{m_1} = [(A,B): E(X) = m_1], \ -\infty < m_1 < 0.$$

Note that D_{m_1} is equivalent to D_{m_1} of Theorem 5.1. We already know from Theorem 5.1 that for (A,B) in D_{m_1},

$$m_1^2 < E(X^2) < 2\, m_1^2. \tag{5.30}$$

Hence using the equation (5.30), it is easy to prove the existence of a solution to the Bayes modal equations and it is stated in Theorem 5.5 below.

<u>Theorem 5.5</u>: Given a sample quantity $(\bar{t}_1, \bar{t}_2)$, $\bar{t}_1 < 0$, the Bayes modal equations (5.29) admit a unique solution with probability one, provided $b > 0$.

Proof: Given $(\bar{t}_1, \bar{t}_2)$, there exist a set $D_{\bar{t}_1}$ in the space of (A,B) and a solution to (5.29) must lie in $D_{\bar{t}_1}$. Also it is important to note that $-\infty < B < 0$ in $D_{\bar{t}_1}$. Hence from the montonicity property (5.9) of $E(X^2)$ in $D_{\bar{t}_1}$ along with (5.30), we can see that

$$\lim_{B \to -\infty} [E(X^2) - a/n - b/nB] = \bar{t}_1^2 - a/n,$$

and

$$\lim_{B \to 0} [E(X^2) - a/n - b/nB] = \infty, \quad \text{if } b > 0.$$

Also we know that $\bar{t}_1^2 \le \bar{t}_2 \le n\bar{t}_1^2$. Hence the result.

Note that a = 1/2 and b = 1 are the optimum prior parameters (in the sense of asymptotic bias) which can be shown using the analysis presented in Section 2.5a. We use these values in simulation studies.

5.5 Simulation Results:

Here, we compare the simulated values of the bias, the probability of nearness, and the MSE of the Bayes modal and the mixed estimators. The results are presented for the (θ_1, θ_2) parameterization. The plan of simulation is the same as in the case of doubly truncated normal distribution. The simulation is based on 500 random samples (of sizes 10, 20, and 40) and the samples are drawn using IMSL subroutine GGNML. We choose the values of truncation probability (q_0) as

i) $q_0 = 0.25$, ii) $q_0 = 0.15$, and iii) $q_0 = 0.05$.

The truncation interval $(-\infty, U)$ is transformed to $(-\infty, 0)$ where U is such that

$$\int_{-\infty}^{U} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = q_0.$$

The sample statistics of interest are

$$\bar{t}_1 = \frac{1}{n} \Sigma x_i \ , \text{ and } \bar{t}_2 = \frac{1}{n} \Sigma x_i^2.$$

Given $(\bar{t}_1, \bar{t}_2)$, we obtain the maximum likelihood estimates $(\hat{\theta}_1, \hat{\theta}_2)$, the Bayes modal estimates $(\tilde{\theta}_1, \tilde{\theta}_2)$ and the mixed estimates $(\hat{\theta}_{1m}, \hat{\theta}_{2m})$ using the appropriate equations as done in Section 3.4. Also note that the computational definitions of the bias, the probability of nearness, and the mean square error (MSE) and the probability of non-existence of the maximum likelihood estimator are as defined in Section 3.4.

It may be noted that obtaining the Bayes modal and the m.l. estimators requires to evaluate $E(X^r)$, $(r = 1,2,3,4)$, at each iteration. Due to the fact that $|E(X^r)| \longrightarrow \infty$ as $\theta_2 \longrightarrow 1$, we have a computational instability near $\theta_2=1$. It is observed tnat $|E(X^r)| \longrightarrow \infty$ faster when $\theta_1 < -30$. Hence, we could carry out the actual computational work only for $-30 \le \theta_1 \le 5$ and $-20 \le \theta_2 < b(\theta_1)$, where $b(\theta_1)$ is a computational upper bound such that $E(X^r)$ is computable.

Actual simulation results are presented in Table 5.1. In Table 5.1, p_0 represents the probability of non-existence of the m.l.e. and p_1 represents the probability of nearness of the Bayes modal with respect to the mixed estimator. A brief discussion of the relative performances of the Bayes modal and the mixed estimators is made in the next paragraph.

Before actually presenting the final results in Table 5.1, several combinations of the parameter values of (θ_1,θ_2) are tried and interestingly, it is observed that the Bayes modal and the mixed estimators behave distinctly different for $0 \le \theta_2 < 1$ as compared to $\theta_2 < 0$. Note that $\theta_2 \ge 0$ is equivalent to $\sigma^2 \ge 1$. From the Table 5.1, we can see that the Bayes estimator has smaller magnitude of bias and MSE whenever $\theta_2 = -1$, and otherwise the mixed estimator has smaller values. Further, it is observed that the direction angle of the bias vector for both the estimators lies in the third quadrant. In other words, both the estimators tend to under estimate both θ_1 and θ_2. Also with regard to the probability of nearness, the same pattern exists. That is, the Bayes modal is more often (about 90% of the time) nearer the true values whenever $\theta_2 < 0$, otherwise the mixed is nearer. Unlike we observed in the cases of truncated gamma and doubly truncated normal, the probability of non-existence of the m.l.e. is relatively very small (about 0.01 to 0.02 for n= 10) even when the truncation probability is as high as 0.25. It is clear from Table 5.1 that p_0 is zero up to 3 decimals whenever the sample size is 20 or more. In Table 5.1, $(v_1 + v_2)$ is the sum of the asymptotic variances of the maximum likelihood estimators of θ_1 and θ_2. Note that $(v_1 + v_2)$ is comparable with the MSE for large n. In fact in Table 5.1, we can see that these two quantities are close to each other for n = 40.

Table 5.1 Simulated Expected Bias, MSE of the Mixed and the Bayes Modal Estimators of (θ_1,θ_2) Singly Truncated Normal Distribution

Prior Parameters: a = 1/2, b = 1

			Mixed Estimators			Bayes Modal Estimates					Asy. Variance		
n	q_0	p_o	Bias	Dir	MSE	Bias	Dir	p_1	MSE	EFF	v_1	v_2	v_1+v_2
$\theta_1 = -1.35$, $\theta_2 = -1.00$													
10	0.25	0.016	2.116	218	26.486	0.711	226	0.91	4.986	5.312	6.885	3.298	10.183
20		0.008	1.032	221	8.287	0.653	220	0.93	3.829	2.164	3.443	1.649	5.091
40		0.000	0.455	224	2.904	0.355	225	0.92	1.953	1.487	1.721	0.824	2.546
$\theta_1 = -0.35$, $\theta_2 = 0.85$													
10	0.25	0.024	0.526	192	1.197	0.899	192	0.24	1.799	0.665	0.430	0.013	0.443
20		0.000	0.203	193	0.316	0.424	192	0.31	0.458	0.690	0.215	0.006	0.222
40		0.000	0.073	193	0.123	0.195	191	0.40	0.152	0.810	0.108	0.003	0.111
$\theta_1 = -0.15$, $\theta_2 = 0.95$													
10	0.25	0.012	0.283	186	0.294	0.485	186	0.22	0.505	0.581	0.108	0.001	0.109
20		0.000	0.109	186	0.081	0.224	185	0.29	0.123	0.653	0.054	0.000	0.054
40		0.000	0.036	186	0.033	0.099	185	0.37	0.041	0.793	0.027	0.000	0.027
$\theta_2 = -2.00$, $\theta_2 = -1.00$													
10	0.15	0.008	2.042	218	21.905	0.918	218	0.91	5.814	3.767	6.249	2.367	8.616
20		0.000	0.862	218	7.213	0.571	220	0.90	3.649	1.976	3.124	1.183	4.308
40		0.000	0.522	216	2.903	0.439	216	0.93	2.178	1.333	1.562	0.592	2.154
$\theta_1 = -0.50$, $\theta_2 = 0.85$													
10	0.15	0.012	0.548	190	1.178	0.905	190	0.23	1.797	0.655	0.391	0.009	0.400
20		0.000	0.155	190	0.232	0.348	190	0.34	0.332	0.701	0.195	0.005	0.200
40		0.000	0.097	191	0.137	0.200	190	0.34	0.165	0.831	0.098	0.002	0.100

Table 5.1 (Continued)

n	q_0	p_0	Mixed Estimates Bias	Mixed Estimates Dir	Mixed Estimates MSE	Bayes Modal Estimates Bias	Bayes Modal Estimates Dir	p_1	Bayes Modal Estimates MSE	EFF	Asy. Variances v_1	Asy. Variances v_2	v_1+v_2
$\theta_1 = -.25$, $\theta_2 = 0.95$													
10	0.15	0.004	0.237	185	0.208	0.432	185	0.20	0.337	0.550	0.098	0.001	0.098
20		0.000	0.093	185	0.070	0.197	185	0.31	0.102	0.679	0.049	0.000	0.049
40		0.000	0.046	185	0.310	0.100	184	0.37	0.039	0.799	0.024	0.000	0.025
$\theta_1 = -3.30$, $\theta_2 = -1.00$													
10	0.05	0.000	0.877	210	20.213	1.067	210	0.91	7.709	2.622	5.838	1.472	7.310
20		0.000	0.646	213	5.595	0.479	214	0.92	3.645	1.535	2.919	0.736	3.655
40		0.000	0.398	211	2.371	0.349	211	0.93	1.957	1.212	1.460	0.368	1.828
$\theta_1 = -0.80$, $\theta_2 = 0.85$													
10	0.05	0.000	0.389	188	0.894	0.718	188	0.31	1.369	0.653	0.365	0.006	0.371
20		0.000	0.173	188	0.311	0.339	188	0.33	0.409	0.760	0.182	0.003	0.185
40		0.000	0.061	188	0.101	0.143	187	0.40	0.119	0.849	0.091	0.001	0.093
$\theta_1 = -0.40$, $\theta_2 = 0.95$													
10	0.05	0.000	0.197	184	0.232	0.377	184	0.29	0.386	0.601	0.091	0.000	0.092
20		0.000	0.108	184	0.075	0.196	184	0.28	0.107	0.702	0.046	0.000	0.046
40		0.000	0.023	184	0.026	0.066	183	0.42	0.030	0.852	0.023	0.000	0.023

Index: q_0: Trunc. Probability

p_0: Prob. of non-existence (m.l.e.)

p_1: Prob. of nearess (Bayes Modal, Moxed)

v_1 = Asy. Variance of θ_1(m.l.e.)

v_2 : Asy. Variances of θ_2(m.l.e.)

REFERENCES

[1] Barndorff-Nielsen, O. (1978). (BN) Information and Exponential Families in Statistical Theory, John Wiley.

[2] Bar-lev, S.K. (1984). Large Sample Properties of the MLE and MCLE for the Natural Parameter of a Truncated Exponential Family. Ann. Inst. Stat. Math. 36, A, 217-222.

[3] Blumenthal, S. (1981). A survey of estimating distributional parameters and sample sizes from truncated samples. In C. Taillie et al (eds), Statistical Distributions in Scientific Work 5, 75-86.

[4] Blumenthal S. (1982). Stochastic expansions for point estimation from complete, censored, and truncated samples. Technical report, Department of Mathematics, university of Illinois, Urbana-Champaign.

[5] Blumenthal, S., and Marcus, R. (1975). Estimating population size with exponential failure. J. Amer. Statist. Assoc. 70, 913-922.

[6] Blumenthal, S., Dahiya, R.C. and Gross, A.J. (1978). Estimating the complete sample size from an incomplete Poisson sample. J. Amer. Statist. Assoc. 73, 182-187.

[7] Broeder, G.G. den (1955). On parameter estimation for truncated Pearson Type III distributions. Ann. Math. Statist. 26, 659-663.

[8] Chapman, D.G. (1956). Estimating the parameters of a truncated gamma distribution. Ann. Math. Statist. 27, 487-506.

[9] Cohen, A.C. (1950a). Estimating the mean and variance of populations from singly truncated and doubly truncated samples. Ann. Math. Statist. 21, 557-569.

[10] Cohen, A.C. (1950b). Estimating parameters of Pearson Type III populations from truncated samples. J. Amer. Statist. Assoc. 45, 411-423.

[11] Cohen, A.C. (1957). On the solution of estimating equations for truncated and censored samples from normal populations. Biometrica 44, 225-236.

[12] Crain, B.R. (1979). Estimating the parameters of a

truncated normal distribution. App. Math. and Comp. 5, 149-156.

[13] Dahiya, R.C. and Gross, A.J. (1973). Estimating the zero class from a truncated Poisson sample. J. Amer. Statist. Assoc. 68, 731-733.

[14] Deemer, W.L. and Votaw, D.F. (1955) (DV). Estimation of parameters of truncated or censored exponential distributions. Ann. Math. Statist. 26, 498-504.

[15] Des Raj (1953). Estimation of the parameters of Type III populations from truncated samples. J. Amer. Statist. Assoc. 48, 336-349.

[16] Gross, A.J. (1971). Montonicity properties of the moments of truncated gamma and Weibull density functions. Technometrics 13, 851-857.

[17] IMSL Library: Referene Manual (1982). International Mathematical and Statistical Libraries, Inc. Houston, TX.

[18] Johnson, N.L. and Kotz, S. (1970). (JK). Continuous Univariate Distributions-1. Houghton Mifflin, Boston Mass.

[19] Joe, H. and Reid, N. (1985). Estimating the number of faults in system. J. Amer. Statist. Assoc. 80, 222-226.

[20] Mittal, M. (1984). Ph.D. Thesis. Estimating the Parameters of Truncated Distributions. Old Dominion University.

[21] Shah, S.M. and Jaiswal, M.C. (1964). Estimation of Parameters of Doubly Truncated Normal Distributions. Ann. Inst. Stat. Math. Tokyo, 18, 107-111.

[22] Thomson, H.R. (1951). Truncated log normal distributions. I. Solution by Moments. Biometrica 38, 414-422.

[23] Gradshteyn, I.S. and Ryzhik, I.M. (1980). Table of Integrals, Series, and Products. Academic Press, NY.

Autobiographical Statement

Place and Date of Birth	:	Umblekoppa (India) January 12, 1952
Educational	:	* M.Sc. (Stat.) 1975 Karnatak University, India
	:	* B.Sc., 1975 Karnatak College, India
	:	* Certificate Course in French, 1972
Employment	:	* Research Assistant (January 1976 - September 1976) University of Agricultural Sciences Bangalore, India
	:	* Assistant Director of Statistics (October 1976-July 1982) Bureau of Economics and Statistics Bangalore, India
	:	* Teaching Assistant (August 1982 - August 1986) Old Dominion University Norfolk, VA
Academic	:	* Gold Medal from Karnatak University for securing highest marks in M.Sc.
	:	* Govt. of India Merit Scholarship for securing second rank in B.Sc.