ON SMALL SAMPLE PROPERTIES OF DISTRIBUTION AND DENSITY FUNCTION ESTIMATORS

by

Albert Dionne and C. P. Quesenberry

Institute of Statistics
Mimeograph Series No. 722
Raleigh - November 1970

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

1. INTRODUCTION

The problems of estimating distribution and density functions are important in statistical theory and its applications. For example, the researcher interested in statistical control of items of mass production has to derive estimates of relative frequencies on the basis of sample data. In reliability there is interest in estimating tail probabilities in a distribution belonging to a specified family. The likelihood ratio is often used to partition the sample space in the construction of decision rules; in discrimination procedures, for example, it is sometimes necessary to estimate the likelihood ratio and hence the resultant partition of the sample space.

When different distribution and density estimators are available study is required to choose among them. A large part of this thesis is devoted to studying the performance of minimum variance unbiased estimators relative to maximum likelihood estimators. While the latter do not possess optimal small sample properties they are, nevertheless, widely used in practice even in small sample situations and it would appear reasonable to compare newer competing estimators with them. It is of interest to consider also general families of estimators that depend on arbitrary constants and to choose the constants so as to make the mean square error (or a weighted mean square error) small. The objective of the study is to find at what regions of the relevant parameter space one estimator is superior to the others.

If unbiasedness is considered to be an essential property, then the minimum variance unbiased estimator would, of course, be preferred.

For example, when separate estimates from independent small samples are to be combined, bias in an estimator is an undesirable property; the bias is likely to be in the same direction for every sample and the product rule of biased estimators will exert strong cumulative effects on the calculations. However, in some other estimation problems, it might be appropriate to use biased estimators as well; a biased maximum likelihood estimator with smaller variance can provide lower probability of a large estimation error.

In Chapter 3 estimators of distribution functions for the uniform and exponential (with location and scale parameters) distributions are considered. In the uniform case the minimum variance unbiased and maximum likelihood estimators are studied under various measures of efficiency; the empirical distribution function is compared with the minimum variance unbiased estimator. Tail probabilities and a family of estimators are also considered and a coefficient of resemblance of the competing distribution functions is computed.

In Chapter 4 the estimation of a normal probability density function with known standard deviation is examined. The minimum variance unbiased and maximum likelihood estimators are considered; the empirical distribution function is also compared with the minimum variance unbiased distribution function estimator. A family of density estimators is studied. The same measures of relative efficiency are used as for the uniform distribution previously examined.

Finally, the thesis is concluded with a summary and suggestions for further research.

2. REVIEW OF LITERATURE

Most of the research on distribution and density function estimation that appears in the literature deals with the following problem. Let $\underline{X}_n = (X_1, \ldots, X_n)$ be n independent observations on a real-valued random variable X. The only knowledge of the distribution of X is that its distribution function F is absolutely continuous. F and, hence, the corresponding density function f, are otherwise assumed to be completely unknown. By making use of the observation $\underline{X}_n$, the objective is to develop a procedure for estimating F(x), or f(x), at each point x. This is usually called the nonparametric distribution, or density, estimation problem.

A result of Rosenblatt (1956) showed that in this setting there exists no unbiased estimator of the parent density function. As a result the emphasis has since been on finding density estimators with good large sample properties, such as consistency and asymptotic normality. An expository review of efforts made in nonparametric density estimation (Wegman, 1969) is commended to the reader's attention.

Some work relevant to this thesis appears in papers concerned with estimating a distribution function $F(x;\theta)$ at a point x on the real line. Here, θ is a parameter which varies over some index set Ω, which is usually a Borel subset of a Euclidean space. Most of the writers of these papers consider particular families of distributions on the real line such that θ admits a complete and sufficient statistic T, say. Their objective is then to find unique minimum variance unbiased estimators of $F(x;\theta)$ for these particular families. Their approach is to use the Rao-Blackwell-Lehmann-Scheffé method of conditioning a simple unbiased estimator of $F(x;\theta)$ on a complete and sufficient statistic T.

Barton (1961) obtained the estimators for the binomial, the Poisson and the normal (known standard deviation) density functions. Pugh (1963) derived the estimator for the one-parameter exponential distribution function. Tate (1959) using integral transform methods and Laurent (1963) have considered the gamma and two-parameter exponential distribution functions. Folks _et al._ (1965) found the estimator for the normal distribution function with known mean. Basu (1964) has given a summary of all these methods. Sathe and Varde (1969) discussed the estimation of the reliability function, $1-F(x;\theta)$, and they obtained, again, most of the results mentioned above. Their method is to find a statistic which is stochastically independent of the complete and sufficient statistic and whose distribution can be easily obtained; the unique minimum variance unbiased estimator is based on this distribution. Washio _et al._ (1956), also using integral transform theory, studied the problem of estimating functions of the parameters in the Koopman-Pitman family of densities. In an earlier paper Kolmogorov (1950) derived an estimator of $P_\theta[X \in A]$ for the normal distribution with known variance.

Fisher (1912) proposed the maximum likelihood procedure for parametric density estimation. Robertson (1967) and Wegman (1970) have studied maximum likelihood density estimation for the nonparametric problem. In general, maximum likelihood estimates for density functions do not exist in this setting, but, if some appropriate restriction is placed on the class of densities from which the estimate may be selected, then a maximum likelihood estimate over the restricted class may exist. The restriction considered in these papers is that the estimates are measurable with respect to a σ-lattice. A σ-lattice is a collection of subsets which is closed under countable unions and intersections and contains both the whole set and the empty set.

The problems of existence and uniqueness of an unbiased estimator of a density function $f(x;\theta)$ have been solved by Seheult (1969). In the present notation, his theorems state that an unbiased estimator of $f(x;\theta)$ exists if and only if there exists an unbiased estimator of $F(x;\theta)$ that is absolutely continuous with respect to Lebesgue measure on the real line; and if there exists a complete and sufficient statistic T for the family, the unique unbiased estimator of the density function $f(x;\theta)$ based on T has minimum variance among all unbiased estimators of $f(x;\theta)$.

The research which is most closely related to the present work is that of Zacks and Even (1966). They presented the results of a study concerning the small sample relative efficiency of maximum likelihood and minimum variance unbiased estimators of reliability functions of one-unit systems; three regular cases of estimation were considered: the Poisson, exponential with scale parameter, and normal with known standard deviation. Two kinds of relative efficiency functions were considered. The first one consisted of the ratio of the Cramer-Rao lower bound of the variances of unbiased estimators to the mean square error of the proposed estimators. The second was the closeness relative efficiency function, which will be defined in Chapter 3 of this thesis.

It is a fact of practice that the maximum likelihood estimators of reliability functions are the most commonly used ones. The question Zacks and Even raised was whether the maximum likelihood estimators of reliability functions are better than other asymptotically normal estimators, even in moderate size and small sample problems. One of the objectives of their study was to find at what regions of the relevant parameter space one estimator is superior to the other.

Zacks and Milton (1969) using numerical integration techniques obtained the mean square errors of the maximum likelihood and minimum variance unbiased estimators of tail probabilities for the normal distribution with both parameters unknown.

3. COMPARISONS OF UNIFORM AND EXPONENTIAL DISTRIBUTION FUNCTION ESTIMATORS

3.1 Introduction

The main purpose of this chapter is to study small sample relative efficiency of maximum likelihood (ML) estimators with respect to minimum variance unbiased (MVU) estimators of distribution functions. The following classes of distributions on the real line are treated: the uniform and exponential (with location and scale parameters) distributions. The empirical distribution function, tail probabilities and a general family of estimators are also considered in the uniform case.

3.2 Relative Efficiency of the ML Estimator with Respect to the MVU Estimator for the Uniform Distribution

Let $X_1, \ldots, X_n$ be a random sample of size n (> 2) from a distribution having the probability density function

$$f(x;\theta) = (1/\theta)\ I_{(0,\theta)}(x), \qquad \theta > 0\ , \tag{3.1}$$

where $I_{(0,\theta)}(x)$ is the indicator function of the set $\{x: 0 < x < \theta\}$. The restriction $n > 2$ is merely a matter of uniformity of presentation. The function to be estimated is the cumulative distribution function at an arbitrary point x

$$F(x;\theta) = (x/\theta)\ I_{(0,\theta)}(x) + I_{[\theta,\infty)}(x), \qquad \theta > 0\ . \tag{3.2}$$

3.2.1 Variance of the MVU Estimator

It is well known that the largest order statistic $X_{(n)}$ is complete and sufficient for this family. The MVU estimator of $F(x;\theta)$ is denoted by $\tilde{F}(x;X_{(n)})$ and is given by (Seheult, 1969)

$$\tilde{F}(x;X_{(n)}) = [(n-1)x/nX_{(n)}]I_{(0,X_{(n)})}(x) + I_{[X_{(n)},\infty)}(x) \quad . \tag{3.3}$$

The following figure illustrates the estimator in (3.3).

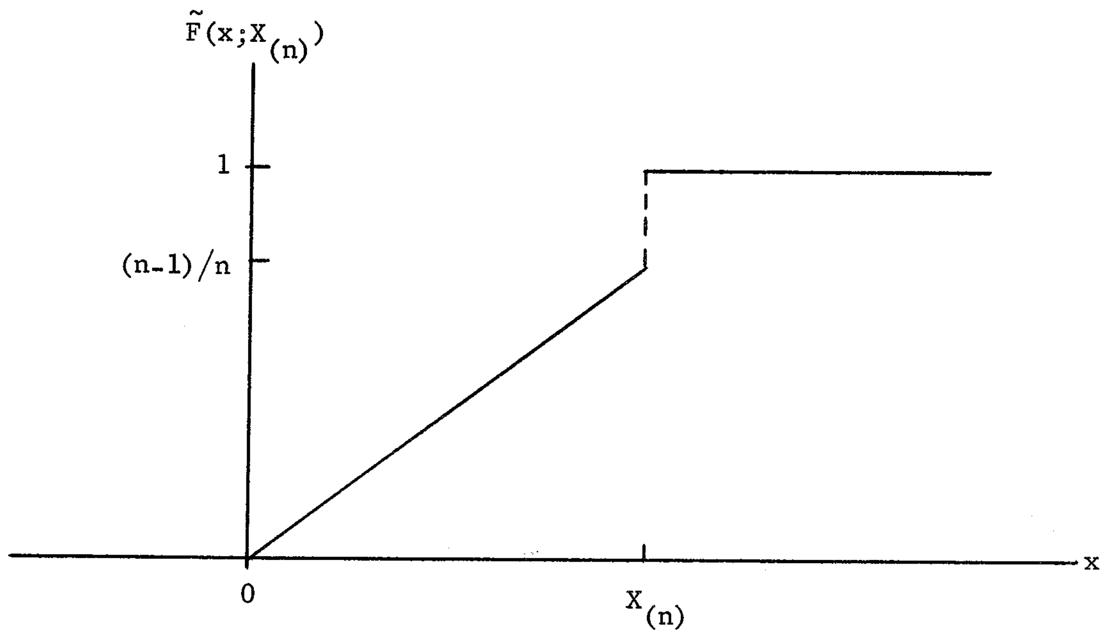

Figure 3.1 The MVU estimator of $F(x;\theta)$ in the uniform distribution

The unbiasedness of $\tilde{F}(x;X_{(n)})$ is easily verified directly by integration. The second moment is computed as follows:

$$E[\tilde{F}(x;X_{(n)})]^2 = [(n-1)x/n]^2 E[(1/X_{(n)})^2 I_{(0,X_{(n)})}(x)]$$
$$+ E[I_{[X_{(n)},\infty)}(x)] \quad . \tag{3.4}$$

For $x \leq 0$,

$$E[\tilde{F}(x;X_{(n)})]^2 = 0 \ . \tag{3.5}$$

For $x \geq \theta$,

$$E[\tilde{F}(x;X_{(n)})]^2 = 1 \ . \tag{3.6}$$

For $0 < x < \theta$, the first member of (3.4) is evaluated by using the iterative property of expectations; thus,

$$\begin{aligned} E[(1/X_{(n)})^2 I_{(0,X_{(n)})}(x)] &= E\{E[(1/X_{(n)})^2 I_{(0,X_{(n)})}(x) \mid I_{(0,X_{(n)})}(x)]\} \\ &= E[(1/X_{(n)})^2 \mid I_{(0,X_{(n)})}(x) = 1] \cdot P_\theta[I_{(0,X_{(n)})}(x) = 1] \\ &= [n/(n-2)][(1/\theta^2) - (x^{n-2}/\theta^n)] \ . \end{aligned} \tag{3.7}$$

The second member of (3.4) is, for $0 < x < \theta$,

$$E[I_{[X_{(n)},\infty)}(x)] = x^n/\theta^n \ . \tag{3.8}$$

Therefore, equation (3.4) can be written as

$$E[\tilde{F}(x;X_{(n)})]^2 = \begin{cases} 0, & x \leq 0, \\ [(n-1)^2x^2]/[n(n-2)\theta^2]-[x^n]/[n(n-2)\theta^n], & 0 < x < \theta, \\ 1, & x \geq \theta. \end{cases} \tag{3.9}$$

Thus, the variance of the MVU estimator is given by

$$\mathrm{Var}[\tilde{F}(x;X_{(n)})] = \{1/[n(n-2)]\}\{(x/\theta)^2 - (x/\theta)^n\} I_{(0,\theta)}(x) . \tag{3.10}$$

This will be used in section 3.2.3 to compute the relative efficiency of the ML estimator with respect to the MVU estimator. The variance of the MVU estimator is a function of the sample size n, and of the ratio (x/θ); it is of order $0(1/n^2)$. The variance function (3.10) is plotted as a function of (x/θ) for n = 3, 8 in Figure 3.2 with the MSE of the ML estimator to be computed in the next section.

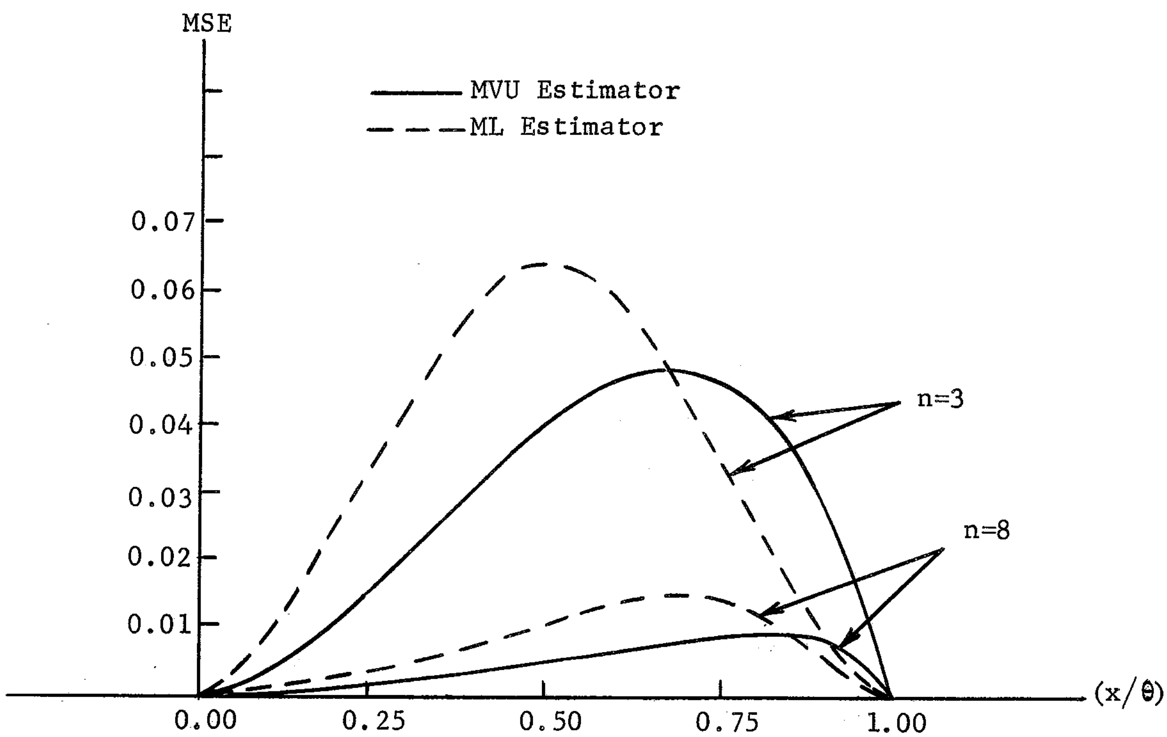

Figure 3.2 Variance function of the MVU estimator and mean square error of the ML estimator in the uniform distribution; n = 3, 8

3.2.2 Mean Square Error of the ML Estimator

It is easy to verify that the largest order statistic $X_{(n)}$ is the ML estimator of θ in the uniform distribution (3.1). The ML estimator of $F(x,\theta)$ is denoted by $\hat{F}(x;X_{(n)})$ at a point x and is given by

$$\hat{F}(x;X_{(n)}) = (x/X_{(n)})I_{(0,X_{(n)})}(x) + I_{[X_{(n)},\infty)}(x) \quad . \tag{3.11}$$

The following figure illustrates this estimator (cf. Figure 3.1).

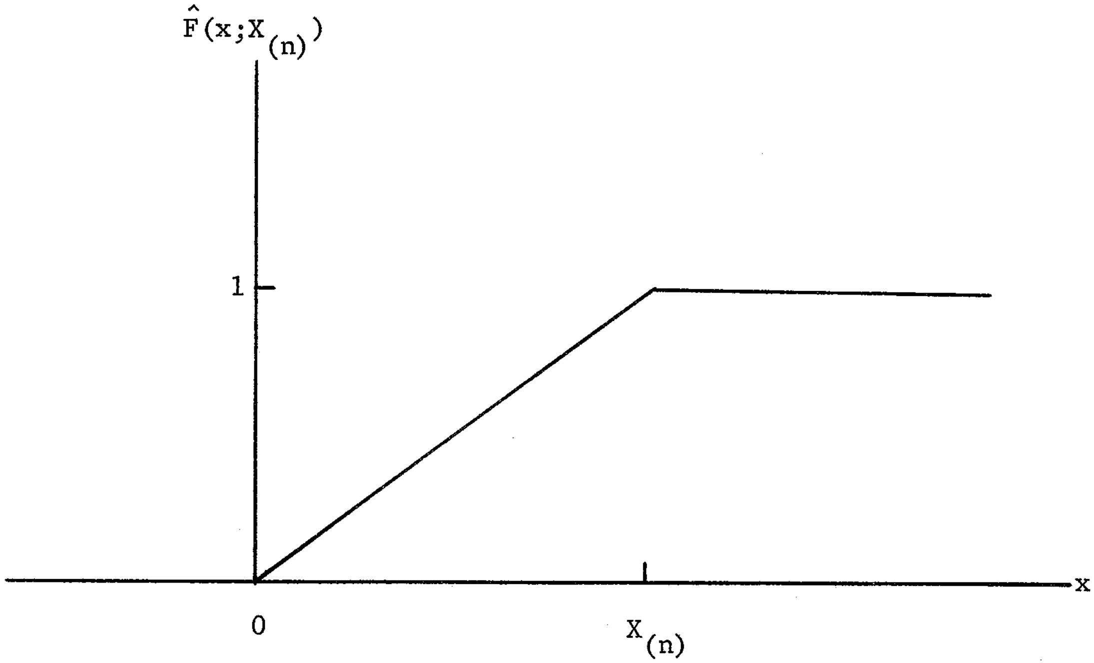

Figure 3.3 The ML estimator of $F(x;\theta)$ in the uniform distribution

The mean of the ML estimator $\hat{F}(x;X_{(n)})$ is

$$E[\hat{F}(x;X_{(n)})] = \begin{cases} 0, & x \le 0, \\ [nx/(n-1)\theta] - [x^n/(n-1)\theta^n], & 0 < x < \theta, \\ 1, & x \ge \theta. \end{cases} \tag{3.12}$$

Thus, $\hat{F}(x;X_{(n)})$ is a biased estimator and its bias function $B_n(x;\theta)$ is

$$B_n(x;\theta) = E[\hat{F}(x;X_{(n)})] - F(x;\theta)$$

$$= [1/(n-1)][(x/\theta) - (x/\theta)^n] \; I_{(0,\theta)}(x) \; , \qquad (3.13)$$

which is non-negative and of order $0(1/n)$.

Using the same technique as in section 3.2.1, the second moment of the ML estimator $\hat{F}(x;X_{(n)})$ is evaluated and found to be

$$E[\hat{F}(x;X_{(n)})]^2 = \begin{cases} 0, & x \leq 0, \\ [nx^2/(n-2)\theta^2] - [2x^n/(n-2)\theta^n], & 0 < x < \theta, \\ 1, & x \geq \theta. \end{cases} \qquad (3.14)$$

Thus, from (3.12) and (3.14)

$$\begin{aligned} \mathrm{Var}[\hat{F}(x;X_{(n)})] = \{ & [nx^2/(n-2)(n-1)^2\theta^2] - [2x^n/(n-2)\theta^n] \\ & + [2nx^{n+1}/(n-1)^2\theta^{n+1}] \\ & - [x^{2n}/(n-1)^2\theta^{2n}] \} \; I_{(0,\theta)}(x) \; . \end{aligned} \qquad (3.15)$$

The MSE of the estimator is required in section 3.2.3 and is computed by adding the square of the bias function to the variance; from (3.13) and (3.15),

$$\begin{aligned} \mathrm{MSE}[\hat{F}(x;X_{(n)})] = \{ & [2x^2/(n-1)(n-2)\theta^2] - [2x^n/(n-2)\theta^n] \\ & + [2x^{n+1}/(n-1)\theta^{n+1}] \} \; I_{(0,\theta)}(x) \; . \end{aligned} \qquad (3.16)$$

This is a function of the sample size and of the ratio (x/θ); it is of order $0(1/n^2)$. The MSE function (3.16) has been plotted as a function of (x/θ) for $n = 3, 8$ in Figure 3.2, section 3.2.1.

3.2.3 Relative Efficiency

Definition 3.1:

$$\text{R.E.}[\hat{F};\tilde{F}] = [\text{Var}(\tilde{F})]/[\text{MSE}(\hat{F})] \; . \tag{3.17}$$

That is, the relative efficiency of the ML estimator with respect to the MVU estimator, denoted by R.E. $[\hat{F};\tilde{F}]$, is defined as the ratio of the variance of the MVU estimator to the mean square error of the ML estimator.

In the uniform distribution, from (3.10) and (3.16)

$$\text{R.E.}[\hat{F};\tilde{F}] =$$

$$\frac{\{1/[n(n-2)]\}\{1-(x/\theta)^{n-2}\}I_{(0,\theta)}(x)}{\{[2/(n-1)(n-2)] - [2x^{n-2}/(n-2)\theta^{n-2}] + [2x^{n-1}/(n-1)\theta^{n-1}]\}I_{(0,\theta)}(x)} \; . \tag{3.18}$$

The relative efficiency function will be taken to be unity outside the range $(0,\theta)$. It is a function of the sample size and the ratio (x/θ); for n = 3, 8 and 16, the MVU estimator $\tilde{F}(x;X_{(n)})$ is superior for (x/θ) less than 0.667, 0.857, and 0.926, respectively. This behavior of the MVU estimator for x values near θ $(0 < x < \theta)$ is related to the fact that $\tilde{F}(x;X_{(n)})$ has a jump of size $(1/n)$ at $x = X_{(n)}$. As the sample size increases R.E.$[\hat{F};\tilde{F}]$ tends to one-half. The relative efficiency function (3.18) is plotted as a function of (x/θ) for n = 3, 8 and 16 in Figure 3.4.

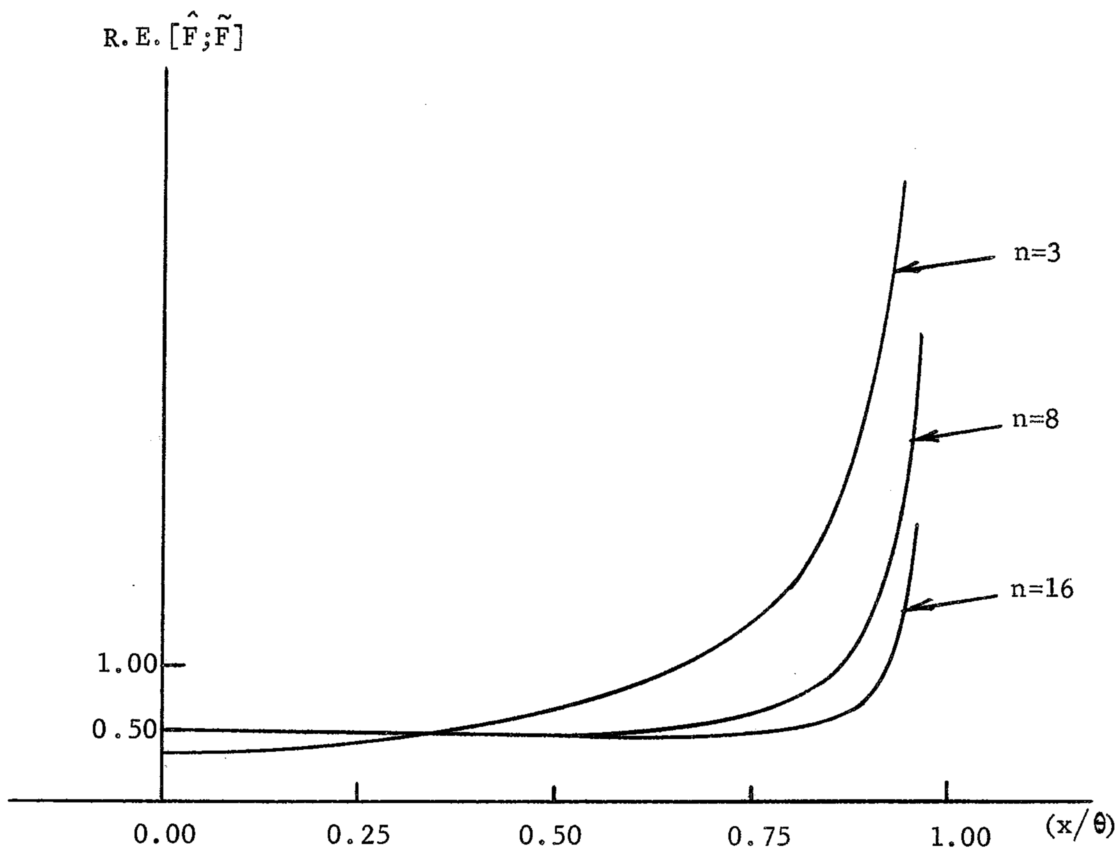

Figure 3.4 Relative efficiency function of the ML estimator with respect to the MVU estimator in the uniform distribution; n = 3, 8, 16

3.2.4 The Empirical Distribution Function

This section considers the gain in information from the additional assumption on the nature of the parent population. It is instructive to compare the MVU estimator with the distribution-free empirical (sample) distribution function $S_n(x)$, where

$$S_n(x) = (1/n) \sum_{i=1}^{n} I_{(0,x]}(X_i) \quad . \tag{3.19}$$

For each fixed x, $S_n(x)$ is an unbiased estimator of $F(x;\theta)$ in (3.2) and since $nS_n(x)$ is a binomial random variable with parameters $[n, F(x;\theta)]$, then

$$\text{Var}[S_n(x)] = (1/n)[(x/\theta) - (x/\theta)^2] I_{(0,\theta)}(x) . \qquad (3.20)$$

This variance function depends on the sample size and the ratio (x/θ), and it is of order $0(1/n)$. $\text{Var}[S_n(x)]$ and $\text{Var}[\tilde{F}(x;X_{(n)})]$ are plotted as functions of (x/θ) for n = 3, 8 in Figure 3.5.

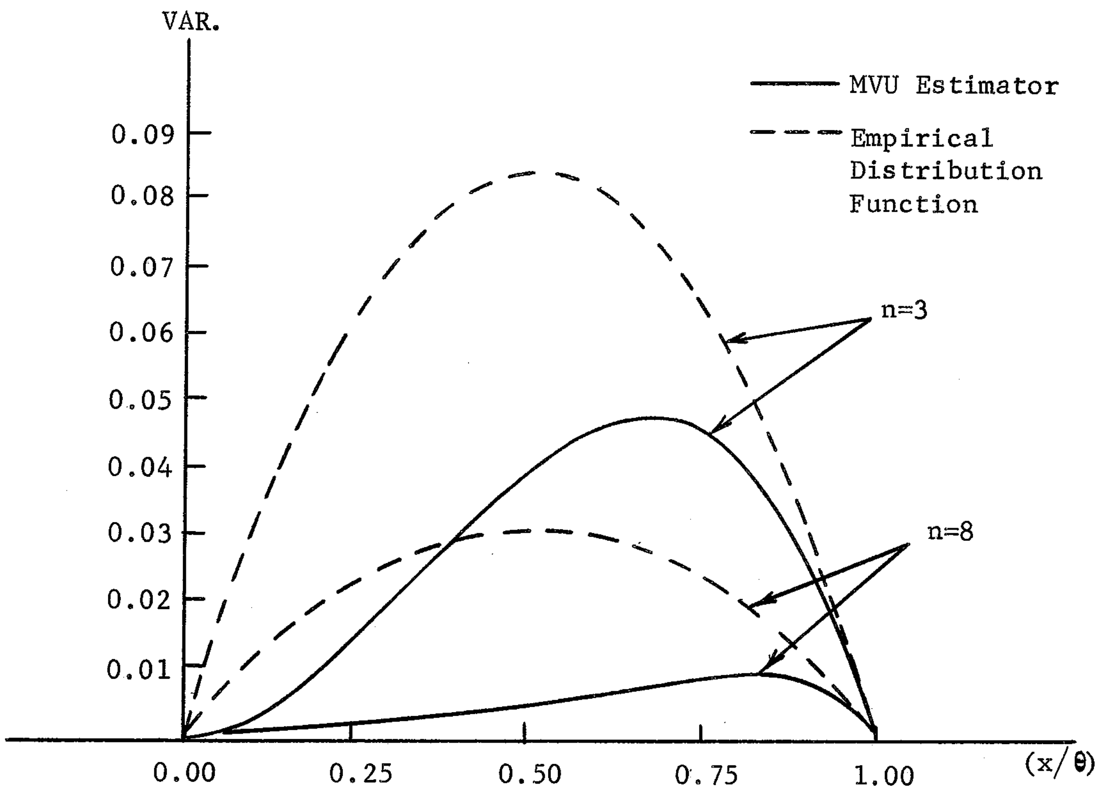

Figure 3.5 Variance functions of the empirical distribution function and of the MVU estimator of $F(x;\theta)$ in the uniform distribution; n = 3, 8

The relative efficiency function R.E. $[S_n;\tilde{F}]$ is computed from (3.10) and (3.20) and is given by

$$R.E.[S_n;\tilde{F}] = \frac{[1/(n-2)][(x/\theta) - (x/\theta)^{n-1}]I_{(0,\theta)}(x)}{[1-(x/\theta)]I_{(0,\theta)}(x)} \quad . \tag{3.21}$$

As in the preceding case, the relative efficiency function (3.21) will be taken to be unity outside the range $(0,\theta)$. It is a function of the sample size and of the ratio (x/θ); for every value of n and $0 < x/\theta < 1$ it is less than unity and tends to zero as n increases. The relative efficiency function (3.21) is plotted as a function of (x/θ) for n = 3, 8, 16 in Figure 3.6.

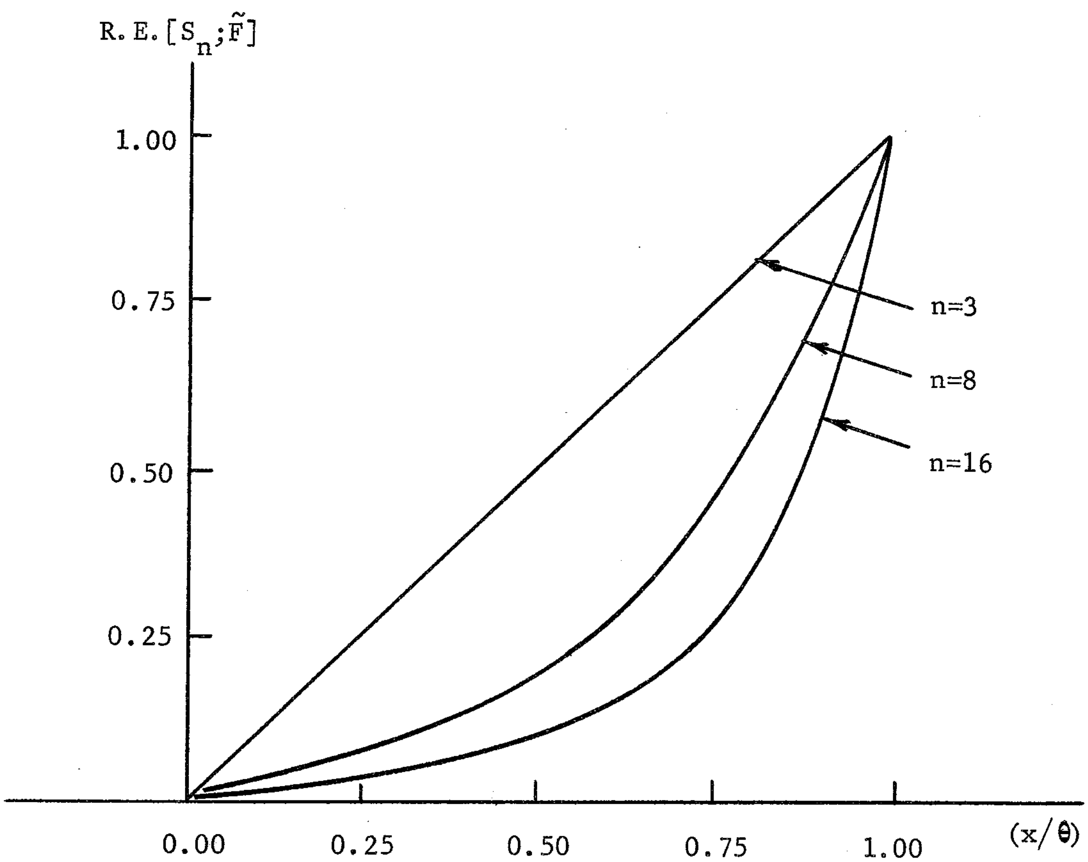

Figure 3.6 Relative efficiency function of the empirical distribution function with respect to the MVU estimator in the uniform case; n = 3, 8, 16

Under the sole assumption of absolute continuity of the distribution function, $S_n(x)$ is the MVU estimator of the parent distribution function; however, with the additional assumption that the sample is from the uniform distribution (3.2), $S_n(x)$ is an unbiased estimator and

$$\tilde{F}(x;X_{(n)}) \equiv E[S_n(x)\,|X_{(n)}] \tag{3.22}$$

is the MVU estimator of $F(x;\theta)$. Thus, Figure 3.6 reflects in terms of the variance functions the gain in information from the additional assumption.

The previous results show that when the observations are taken from a uniform distribution an empirical distribution function estimator has a considerably larger variance than the MVU estimator.

3.2.5 Closeness Relative Efficiency of the MVU Estimator with Respect to the ML Estimator

The next measure of relative efficiency to be considered is based on the notion of a closeness function of an estimator. A definition introduced by Zacks and Even (1966) will be recalled and some of its properties examined in detail.

Let $\underline{X}_n = (X_1, \ldots, X_n)$ be an observed vector of random variables having a joint distribution $F(\underline{X}_n;\theta)$. Let $u(\underline{X}_n)$ be an estimator of a parametric function $g(\theta)$.

Definition 3.2: The closeness function of $u(\underline{X}_n)$ to $g(\theta)$ is the probability function

$$\psi_n(\varepsilon;\theta) = P_\theta\{|u(\underline{X}_n)-g(\theta)| \leq \varepsilon\cdot|g(\theta)|\}, \quad 0 < \varepsilon < 1 . \tag{3.23}$$

In words, this is the probability that the magnitude of the absolute deviation of the estimator from the true value of $g(\theta)$ will not exceed a given proportion, ϵ, of the magnitude of $g(\theta)$.

Let $u_1(\underline{X}_n)$ and $u_2(\underline{X}_n)$ be two estimators of $g(\theta)$ having closeness functions $\psi_n^{(1)}(\epsilon;\theta)$ and $\psi_n^{(2)}(\epsilon;\theta)$, respectively.

Definition 3.3: The estimator $u_1(\underline{X}_n)$ is more efficient than $u_2(\underline{X}_n)$ at θ_0 if

$$\psi_n^{(1)}(\epsilon;\theta_0) > \psi_n^{(2)}(\epsilon;\theta_0) \quad . \tag{3.24}$$

Definition 3.4: If $\psi_n^{(1)}(\epsilon;\theta) \geq \psi_n^{(2)}(\epsilon;\theta)$ for all θ,

and if $\psi_n^{(1)}(\epsilon;\theta) > \psi_n^{(2)}(\epsilon;\theta)$ for at least one θ,

(3.25)

then $u_2(\underline{X}_n)$ is an inadmissible estimator in the sense of this closeness criterion.

In order to apply these notions to the uniform distribution, the closeness functions of the MVU estimator and ML estimator are computed as follows. The closeness function of the MVU estimator is denoted by $\tilde{\psi}_n(\epsilon;\theta)$; for $x \leq 0$ or $x \geq \theta$, equations (3.2) and (3.3) imply that

$$\tilde{\psi}_n(\epsilon;\theta) = 1 \; ; \tag{3.26}$$

for $0 < x < \theta$,

$$\begin{aligned}\tilde{\psi}_n(\epsilon;\theta) &= P_\theta\{\,|[(n-1)x/nX_{(n)}]I_{(0,X_{(n)})}(x) - (x/\theta)\,| \leq \epsilon\cdot(x/\theta)\} \\ &= P_\theta\{[n(1-\epsilon)/(n-1)\theta] \leq (1/X_{(n)})I_{(0,X_{(n)})}(x) \\ &\qquad \leq [n(1+\epsilon)/(n-1)\theta]\} \; .\end{aligned} \tag{3.27}$$

Since a probability function can be evaluated as the expected value of a conditional probability function, equation (3.27) is written as

$$\tilde{\psi}_n(\epsilon;\theta) = E\, P_\theta\{[n(1-\epsilon)/(n-1)\theta] \le (1/X_{(n)})\, I_{(0,X_{(n)})}(x) \le [n(1+\epsilon)/(n-1)\theta] \,|\, I_{(0,X_{(n)})}(x)\}$$

$$= P_\theta\{[n(1-\epsilon)/(n-1)\theta] \le 1/X_{(n)} \le [n(1+\epsilon)/(n-1)\theta] \,|\, I_{(0,X_{(n)})}(x) = 1\} \cdot P_\theta\{I_{(0,X_{(n)})}(x) = 1\} \;. \qquad (3.28)$$

To simplify the writing, the proportionality factor ϵ is chosen large enough $[\epsilon \ge (1/n)]$ so that

$$\theta \le [(n-1)\theta/n(1-\epsilon)] \;. \qquad (3.29)$$

Evaluating the closeness function (3.28) for different values of x, from (3.26) and (3.29),

$$\tilde{\psi}_n(\epsilon;\theta) = \begin{cases} 1 - [(n-1)/n(1+\epsilon)]^n, & 0 < x < [(n-1)\theta/n(1+\epsilon)], \\ 1 - [x/\theta]^n, & [(n-1)\theta/n(1+\epsilon)] \le x < \theta, \\ 1, & x \le 0 \text{ or } x \ge \theta. \end{cases} \qquad (3.30)$$

Thus, the closeness function of the MVU estimator is a function of the sample size, the ratio (x/θ) and the proportionality factor ϵ which is chosen greater than or equal to $(1/n)$ for simplicity; it tends to unity for all x as n increases.

Similarly, the closeness function of the ML estimator, denoted by $\hat{\psi}_n(\epsilon;\theta)$, is found to be

$$\hat{\psi}_n(\epsilon;\theta) = \begin{cases} 1 - [1/(1+\epsilon)]^n, & 0 < x < \theta/(1+\epsilon), \\ 1 - [x/\theta]^n, & \theta/(1+\epsilon) \leq x < \theta, \\ 1, & x \leq 0 \text{ or } x \geq \theta. \end{cases} \tag{3.31}$$

This is also a function of the sample size, the ratio (x/θ) and the proportionality factor $(0 < \epsilon < 1)$; it tends to unity for every x as n increases.

The notion of closeness relative efficiency of the MVU estimator with respect to the ML estimator, denoted by C.R.E. $[\tilde{F};\hat{F}]$, is introduced by the following definition.

Definition 3.5:

$$\text{C.R.E.}[\tilde{F};\hat{F}) = [\tilde{\psi}_n(\epsilon;\theta)]/[\hat{\psi}_n(\epsilon;\theta)] \; . \tag{3.32}$$

That is, the closeness relative efficiency of the MVU estimator with respect to the ML estimator is defined as the ratio of their closeness functions. Thus, from (3.30) and (3.31)

$$\text{C.R.E.}[\tilde{F};\hat{F}] = \begin{cases} \{1-[(n-1)/n(1+\epsilon)]^n\}/\{1-[1/(1+\epsilon)]^n\}, & 0<x<[(n-1)\theta/n(1+\epsilon)], \\ \{1-[x/\theta]^n\}/\{1-[1/(1+\epsilon)]^n\}, & [(n-1)\theta/n(1+\epsilon)] \leq x < [\theta/(1+\epsilon)], \\ 1, & x \leq 0 \text{ or } x \geq [\theta/(1+\epsilon)]. \end{cases} \tag{3.33}$$

This function depends on the sample size, the ratio (x/θ) and the proportionality factor ϵ [$\epsilon \geq (1/n)$ for simplicity]. It is greater than one for $0 < x < [\theta/(1+\epsilon)]$ and unity elsewhere; it tends to unity for all x as the sample size increases $[\epsilon \geq (1/n)]$.

On the basis of the closeness criterion, for the range $0 < x < [\theta/(1+\epsilon)]$, the MVU estimator is a more efficient estimator of $F(x;\theta)$ than the ML estimator which is therefore inadmissible by Definition 3.4; for the rest of the range, these estimators are equivalent. A weaker conclusion was reached at the end of section 3.2.3 where the relative efficiency function was the criterion of comparison.

The closeness relative efficiency function (3.33) is plotted as a function of (x/θ) for arbitrary values of n and $\epsilon[\epsilon \geq (1/n)]$ in Figure 3.7.

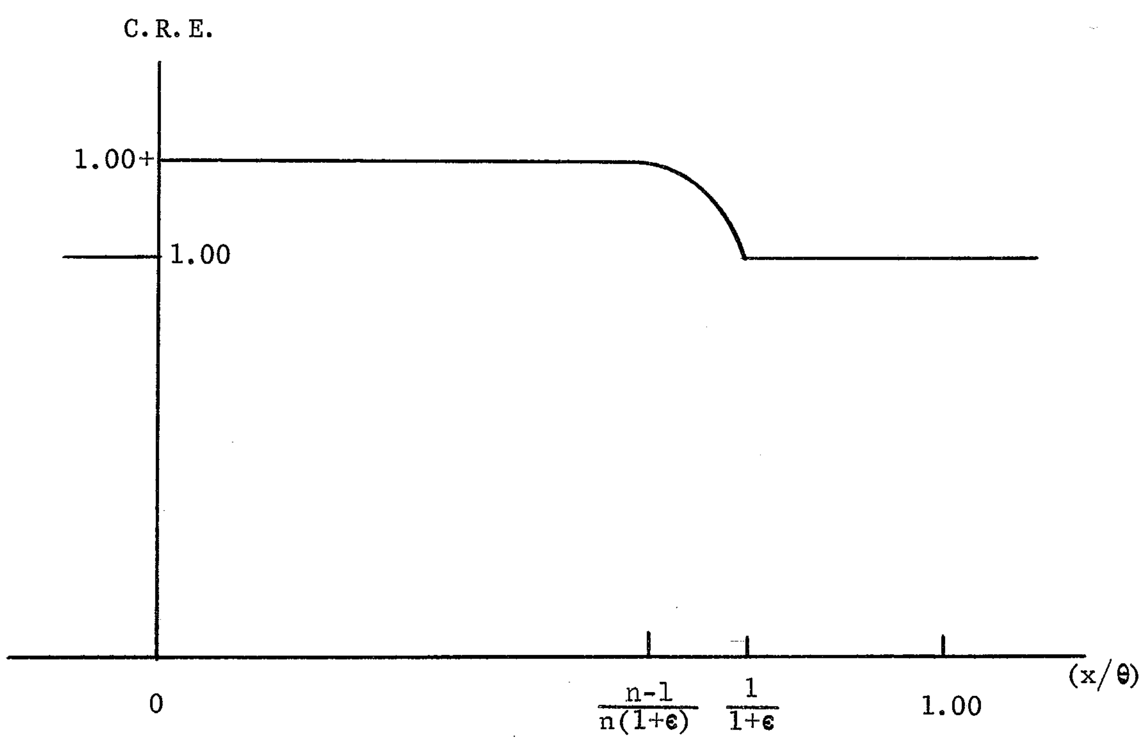

Figure 3.7 Closeness relative efficiency function in the uniform distribution; arbitrary values of n and ϵ $[\epsilon \geq (1/n)]$

For $0 < \varepsilon < (1/n)$ the functional forms of (3.30) and (3.33) change but it is easy to verify that the closeness relative efficiency function is greater than one for $0 < x < \theta$ and tends to unity as the sample size increases.

3.2.6 Tail Probabilities

In the preceding sections properties of the MVU estimator and ML estimator have been considered which are functions of the parent distribution. The approach in this section is different in that it simply considers how much the estimators differ for various sample sizes. The point here is to see how large n must be before the two estimators can be considered to be essentially the same.

For $n > 2$ and $0 < p < 1$, the smallest root of the equation

$$\hat{F}(x;X_{(n)}) = p \tag{3.34}$$

is

$$\hat{\xi}_p = p \cdot X_{(n)} \; . \tag{3.35}$$

In order to compare the estimators $\tilde{F}(x;X_{(n)})$ and $\hat{F}(x;X_{(n)})$ it is appropriate to evaluate the probability

$$\tilde{F}(\hat{\xi}_p;X_{(n)}) = \alpha \; , \tag{3.36}$$

$0 < \alpha < [(n-1)/n]$, and to compute the ratio of probabilities (α/p) for fixed $p(0 < p < 1)$ and different values of sample sizes. This ratio is used as a coefficient of resemblance of the estimating distribution functions and is denoted by C.R.$[n;p]$.

The following figure illustrates the two steps in the computation of C.R.[n;p].

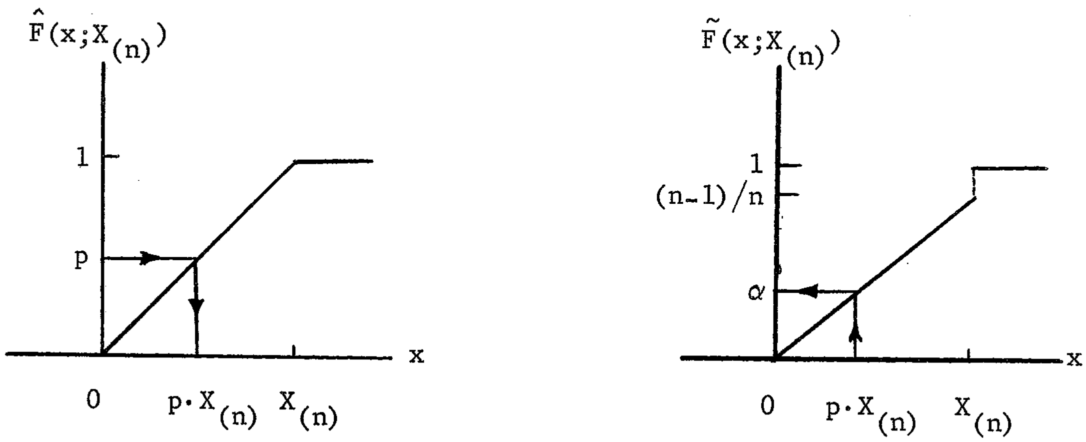

Figure 3.8 Computation of the coefficient of resemblance in the uniform distribution

From (3.3), (3.35) and (3.36), it follows that

$$\alpha = [(n-1)/n]p \ . \tag{3.37}$$

Thus, the coefficient of resemblance is given by

$$C.R.[n;p] = [(n-1)/n], \quad 0 < p < 1 \ . \tag{3.38}$$

Values of the coefficient of resemblance (3.38) are given in Table 3.1 for various sample sizes.

Table 3.1 Values of the coefficient of resemblance in the uniform case, $0 < p < 1$

n	C.R.[n;p]
3	0.666
4	0.750
5	0.800
10	0.900
20	0.950
30	0.966

It takes a large number of observations before one can consider the estimating distribution functions as being approximately equivalent. The discrepancy is due to the facts that the ML estimator is positively biased and that the MVU estimator has a jump of size $(1/n)$ at $x = X_{(n)}$.

3.2.7 Families of Estimators

Having studied small sample behavior of ML and MVU estimators it seems desirable to see if one can improve in performance by considering general families of estimators which contain the previous ones as special cases and select from these families a particular member which is best according to minimum mean square error criterion.

The first such family consists of those estimators of the uniform distribution function that have an arbitrary jump of size 1-b at $x = X_{(n)}$:

$$F_b(x;X_{(n)}) = (bx/X_{(n)})I_{(0,X_{(n)})}(x) + I_{[X_{(n)},\infty)}(x), \quad 0 < b \leq 1. \tag{3.39}$$

This class of estimators includes the ML estimator $\hat{F} = F_1$ and the MVU estimator $\tilde{F} = F_{(n-1)/n}$.

The mean square error of $F_b(x;X_{(n)})$ is directly computed and is given by

$$\begin{aligned} MSE[F_b(x;X_{(n)})] &= (x/\theta)^2[nb^2/(n-2) - 2nb/(n-1) + 1] \\ &+ (x/\theta)^n[1 - nb^2/(n-2)] \\ &+ (x/\theta)^{n+1}[2nb/(n-1) - 2], \quad 0 < x < \theta. \end{aligned} \tag{3.40}$$

Choosing

$$b = (n-2)/(n-1) \tag{3.41}$$

minimizes the coefficient of the dominant term $(x/\theta)^2$ in (3.40); a consideration of the other terms is less important as the sample size increases.

It is interesting and instructive to compare this new estimator $F_{(n-2)/(n-1)}$ with the MVU estimator previously studied. This result, as will soon become apparent, strengthens the case for using the MVU estimator.

The mean square error of $F_{(n-2)/(n-1)}$ and the variance of the MVU estimator $\tilde{F}$ are plotted as functions of (x/θ) for $n = 3, 8$ in Figure 3.9.

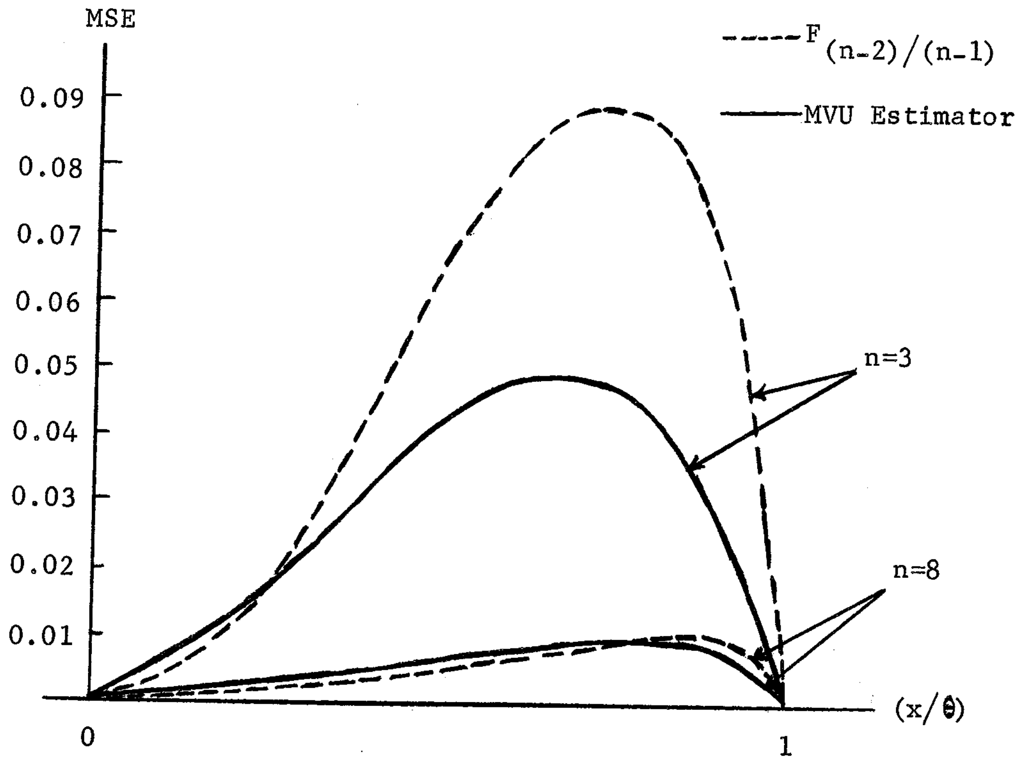

Figure 3.9 Variance of the MVU estimator and mean square error of $F_{(n-2)/(n-1)}$ in the uniform distribution; n = 3, 8

The relative efficiency of $F_{(n-2)/(n-1)}$ with respect to the MVU estimator $\tilde{F}$ is given by:

$$R.E.[F_{(n-2)/(n-1)};\ \tilde{F}] = \frac{(n-1)^2[1-(x/\theta)^{n-2}]}{n(n-2)[1+(x/\theta)^{n-2} - 2(x/\theta)^{n-1}]}, \quad 0 < x < \theta . \tag{3.42}$$

This function will be taken to be unity outside (0, θ). It is plotted as a function of (x/θ) for n = 3, 8, 16 in Figure 3.10.

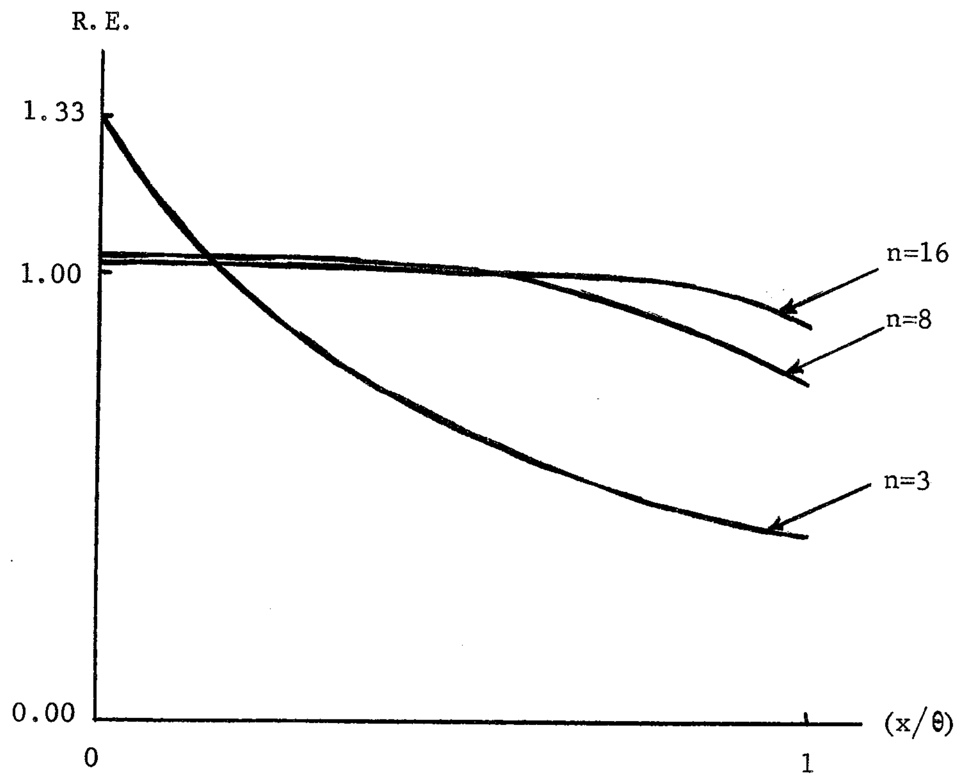

Figure 3.10 Relative efficiency function of $F_{(n-2)/(n-1)}$ with respect to the MVU estimator in the uniform distribution; n = 3, 8, 16

For n = 3, the MVU estimator is preferred if $0.166 < (x/\theta) < 1.00$; for n = 8, 16, and 24, $F_{(n-2)/(n-1)}$ is slightly more efficient if (x/θ) is positive and less than 0.527, 0.705 and 0.780, respectively. From these results and Figure 3.10 it appears that the MVU estimator should be used in practice. While the range over which $F_{(n-2)/(n-1)}$ is superior increases, the degree of superiority is very small and decreases as the sample size increases. Even for n moderately small the estimators can be taken to be equivalent since the relative efficiency function (3.42) tends to unity very rapidly.

It is of interest to consider larger classes of distribution function estimators by introducing more constants. The first such family consists of statistics of the form

$$F_{a,b}(x;X_{(n)}) = \begin{cases} 0, & x \leq 0, \\ bx/X_{(n)}, & 0 < x < (X_{(n)}/a), \\ 1, & (X_{(n)}/a) \leq x, \end{cases} \tag{3.43}$$

where a and b ($0 < b \leq a < 1$) are arbitrary positive real numbers less than unity which can be chosen so as to give the estimator optimal properties. The restriction $b \leq a$ ensures that the estimator (3.43) is everywhere smaller than or equal to unity; when $a = b$, $F_{a,a}(x;X_{(n)})$ belongs to the subfamily of continuous estimators.

Moments of the statistic $F_{a,b}(x;X_{(n)})$ are directly computed. The mean square error MSE $[F_{a,b}(x;X_{(n)})]$, abbreviated MSE, is given by the following expressions; its functional form changes depending on x.

For $0 < x < \theta$,

$$\begin{aligned} \text{MSE} = {} & (x/\theta)^2[nb^2/(n-2) - 2nb/(n-1) + 1] \\ & + (x/\theta)^n[a^n - na^{n-2}b^2/(n-2)] \\ & + (x/\theta)^{n+1}[2na^{n-1}b/(n-1) - 2a^n]; \end{aligned} \tag{3.44}$$

for $\theta \leq x < (\theta/a)$,

$$\begin{aligned} \text{MSE} = {} & 1 - (x/\theta)[2nb/(n-1)] + (x/\theta)^2[nb^2/(n-2)] \\ & + (x/\theta)^n[2na^{n-1}b/(n-1) - a^n - na^{n-2}\, b^2/(n-2)] . \end{aligned} \tag{3.45}$$

Note that the coefficient $(x/\theta)^2$ in (3.44) depends on b and n only and the same is true for the coefficient of (x/θ) in (3.45).

To make the MSE in (3.44) small one might consider choosing b so as to minimize the coefficient of $(x/\theta)^2$; this is done by selecting

$$b = (n-2)/(n-1) \ . \tag{3.46}$$

To minimize the coefficient of (x/θ) in (3.45) a reasonable choice is again

$$b = (n-2)/(n-1) \ . \tag{3.47}$$

Moreover, choosing

$$a = b = (n-2)/(n-1) \tag{3.48}$$

will make the estimator continuous.

A different class of estimators consists of those of the form

$$F_{a,b}(x;X_{(n)}) = \begin{cases} 0, & x \leq 0 \ , \\ bx/X_{(n)} \ , & 0 < x < (X_{(n)}/a), \\ 1, & (X_{(n)}/a) \leq x, \end{cases} \tag{3.49}$$

where a and b ($0 < b \leq a$, $a \geq 1$) are arbitrary positive constants; a is greater than or equal to unity in this case.

For $0 < x < (\theta/a)$, the mean square error of $F_{a,b}(x;X_{(n)})$ is given by (3.44); for $(\theta/a) \leq x < \theta$,

$$\text{MSE} = [1 - (x/\theta)]^2 \ . \tag{3.50}$$

This function is constant in b and n and depends on a only through its range. It is zero if a is taken to be unity.

Thus, choosing the constants

$$a = 1, \quad b = (n-2)/(n-1) \tag{3.51}$$

will make the mean square error small. The relative efficiency of $F_{1,(n-2)/(n-1)}$ with respect to the MVU estimator $\tilde{F}$ is given in (3.42) and is plotted in Figure 3.10.

3.2.8 Summary

As seen in previous sections the MVU estimator of the uniform distribution function is superior to the ML estimator for most of the values of x and θ; for x values near θ only $(0 < x < \theta)$ the ML estimator is more efficient.

On the basis of the closeness relative efficiency criterion the MVU estimator is preferable over most of the range and is, otherwise, equivalent to the ML estimator. Also when the empirical distribution function is compared with the unbiased estimator based on the complete and sufficient statistic it is evident that the gain in information from the additional assumption on the nature of the parent population is quite important.

Finally, the case for the MVU estimator has been strengthened by consideration of general families of estimators. It has been found that the MVU estimator is superior over most of the range in small sample situations and that other estimators are not appreciably more efficient; in any case these statistics become equivalent to the MVU estimator very quickly as the sample size increases.

3.3 Relative Efficiency of the ML Estimator with Respect to the MVU Estimator in the Exponential Distributions with Location and Scale Parameters

The purpose of this section is to study the relative efficiency of the maximum likelihood estimator with respect to the minimum variance unbiased estimator in the exponential distributions with location and scale parameters.

Results for the exponential distribution with location parameter can be immediately obtained from the corresponding results (section 3.2) on the uniform family by making use of the following transformation:

If X is uniformly distributed on the interval $(0, \theta)$, then $Y = -\ln(X)$ has the exponential distribution with location parameter $\tau = -\ln(\theta)$, where $\ln(u)$ is the natural logarithm of the positive real number u.

Some of the results on the exponential distribution with scale parameter have been derived by Zacks and Even (1966) and are repeated here to allow an easy comparison with the location parameter case.

3.3.1 Variance and Mean Square Error in the Exponential Distribution with Location Parameter

Let $X_1, \ldots, X_n$ be n (> 2) independent and identically distributed random variables with common density function

$$f(x;\theta) = \{\exp[-(x-\theta)]\}\, I_{(\theta,\infty)}(x) , \tag{3.52}$$

where θ ($-\infty < \theta < \infty$) is a location parameter. The cumulative distribution function $F(x;\theta)$ at an arbitrary point x is to be estimated, where

$$F(x;\theta) = \{1 - \exp[-(x-\theta)]\}\, I_{(\theta,\infty)}(x) . \tag{3.53}$$

It is well known that the first order statistic $X_{(1)}$ is a sufficient and complete statistic for this family. The MVU estimator of $F(x;\theta)$, denoted by $\tilde{F}(x;X_{(1)})$, is given by (Seheult, 1969)

$$\tilde{F}(x;X_{(1)}) = \{1 - [(n-1)/n] \exp[-(x-X_{(1)})]\} \, I_{(X_{(1)},\infty)}(x) \, . \tag{3.54}$$

From the result (3.10) on the uniform distribution, it follows that

$$\mathrm{Var}[\tilde{F}(x;X_{(1)})] = [1/n(n-2)]\{\exp[-2(x-\theta)] - \exp[-n(x-\theta)]\} \, I_{(\theta,\infty)}(x). \tag{3.55}$$

This is of order $O(1/n^2)$; it is plotted as a function of $(x-\theta)$ for $n = 4, 8$ in Figure 3.11 with the mean square error of the ML estimator.

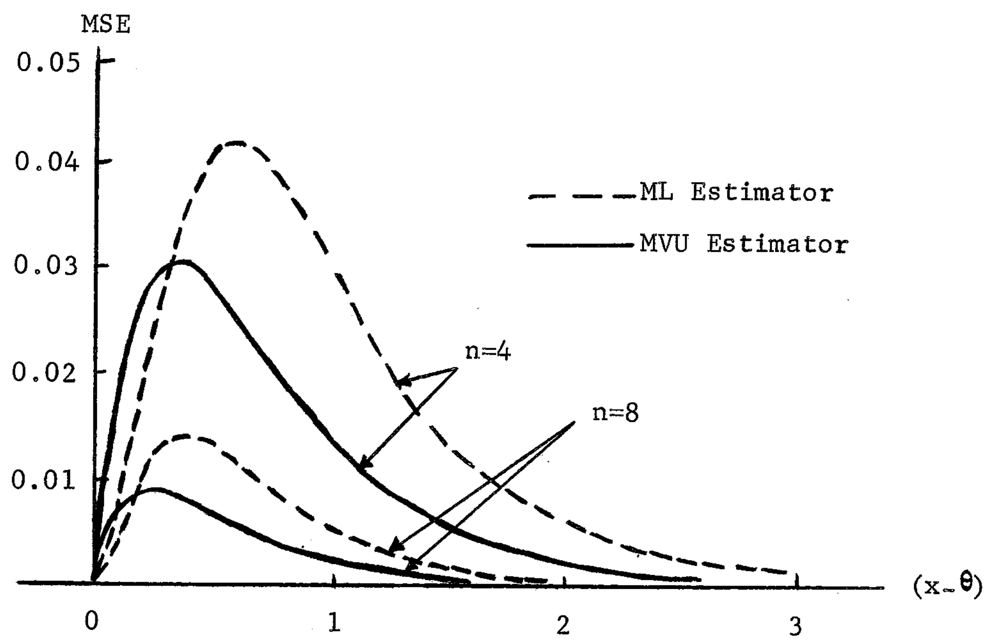

Figure 3.11 Variance function of the MVU estimator and mean square error of the ML estimator in the exponential distribution with location parameter θ; $n = 4, 8$

The first order statistic $X_{(1)}$ is also the ML estimator of θ in the exponential distribution (3.52). The ML estimator of $F(x;\theta)$ is denoted by $\hat{F}(x;X_{(1)})$ at a point x and is

$$\hat{F}(x;X_{(1)}) = \{1 - \exp[-(x-X_{(1)})]\} \; I_{(X_{(1)},\infty)}(x) \; . \tag{3.56}$$

The MSE of this statistic can be obtained from (3.16); in fact,

$$\begin{aligned} MSE[\hat{F}(x;X_{(1)})] = \{&[2/(n-1)(n-2)]\exp[-2(x-\theta)] \\ &- [2/(n-2)]\exp[-n(x-\theta)] \\ &+ [2/(n-1)]\exp[-(n+1)(x-\theta)]\} \; I_{(\theta,\infty)}(x) \; . \end{aligned} \tag{3.57}$$

It is also of order $0(1/n^2)$ and has been plotted in Figure 3.11.

3.3.2 Variance and Mean Square Error in the Exponential Distribution with Scale Parameter

Let $X_1, \ldots, X_n$ be n (> 2) independent and identically distributed random variables with common density function

$$f(x;\theta) = \{(1/\theta) \exp[-(x/\theta)]\} \; I_{[0,\infty)}(x) \; , \tag{3.58}$$

where θ ($\theta > 0$) is a scale parameter. The function

$$F(x;\theta) = \{ 1 - \exp[-(x/\theta)]\} \; I_{[0,\infty)}(x) \tag{3.59}$$

is to be estimated.

It is known that $\sum_{i=1}^{n} X_i$ is a complete and sufficient statistic for the exponential distribution (3.58). The MVU estimator of $F(x;\theta)$ is given by (Pugh, 1963)

$$\tilde{F}(x; \sum_{i=1}^{n} X_i) = \{1 - [1-(x/\sum_{i=1}^{n} X_i)]^{n-1}\} \, I_{(0, \sum_{i=1}^{n} X_i)}(x)$$

$$+ I_{[\sum_{i=1}^{n} X_i, \infty)}(x) \; . \tag{3.60}$$

On the other hand the sample mean $(1/n) \sum_{i=1}^{n} X_i$ is the ML estimator of θ in the distribution (3.58) and therefore, the ML estimator of $F(x,\theta)$ is

$$\hat{F}(x; \sum_{i=1}^{n} X_i) = \{1 - \exp[-(nx/\sum_{i=1}^{n} X_i)]\} \, I_{[0,\infty)}(x) \; . \tag{3.61}$$

Computations of moments of these estimators require in particular tables of the exponential integral

$$Ei(-x) = - \int_{x}^{\infty} u^{-1} e^{-u} \, du \; , \tag{3.62}$$

and tables of the modified Bessel functions; they are carried out in detail in Pugh (1963), and in Zacks and Even (1966). To allow for a comparison with the preceding case the graphs alone are reproduced from Zacks and Even (p. 1039).

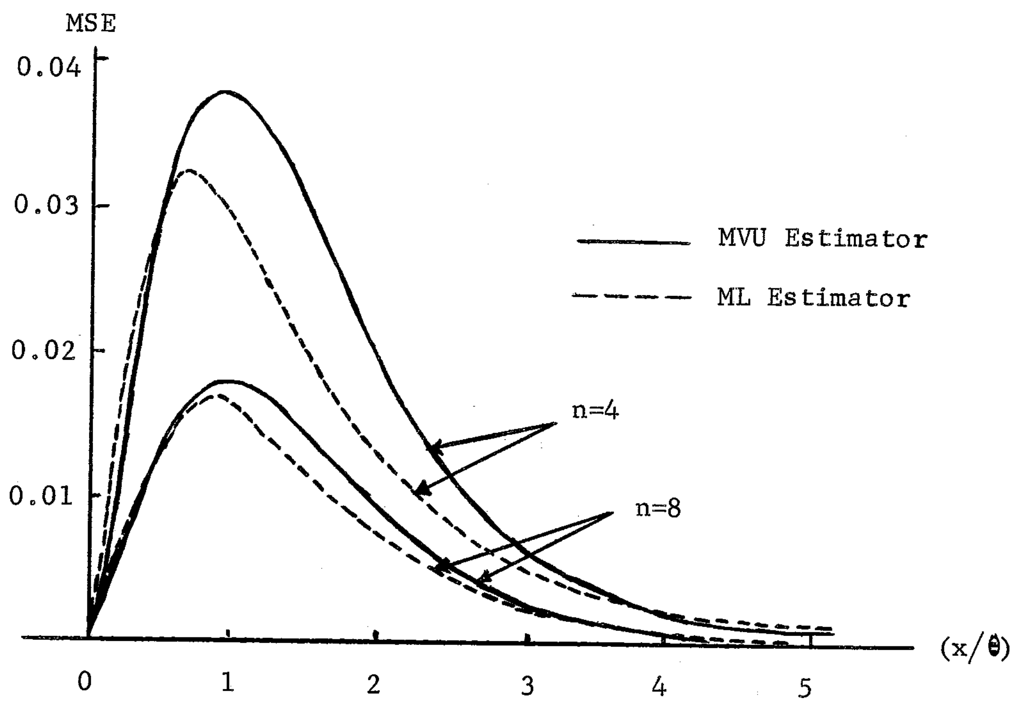

Figure 3.12 Variance of the MVU estimator and mean square error of the ML estimator in the exponential distribution with scale parameter θ; $n = 4, 8$

3.3.3 Discussion

The relative efficiency function R.E.$[\hat{F};\tilde{F}]$ of the ML estimator with respect to the MVU estimator has been defined in section 3.2.3 as the ratio of the variance of the MVU estimator to the mean square error of the ML estimator.

In the exponential case with location parameter it can be argued from (3.55), (3.57) that for $n = 4$ and 8, R.E.$[\hat{F};\tilde{F}]$ is less than unity for $(x-\theta)$ greater than 0.305 and 0.154, respectively; it tends to one-half as the sample size increases. Therefore, the MVU estimator is more efficient as an estimator of $F(x;\theta)$ in (3.53); this is not true, however, for x values near θ.

In the exponential case with scale parameter, the ML estimator is to be preferred if $0.497 < (x/\theta) < 3.769$ ($n=4$) and if $0.475 < (x/\theta) < 3.678$ ($n=8$); the MVU estimator is better, otherwise. The region where the ML estimator is more efficient is slowly contracting as the sample size increases.

Thus, under the relative efficiency criterion the ML estimator is preferred for estimating lower tails in the location parameter case whereas the MVU estimator is more efficient for lower and upper tails of the exponential distribution with scale parameter.

4. COMPARISONS OF NORMAL DENSITY AND DISTRIBUTION FUNCTION ESTIMATORS

4.1 Introduction

The main purpose of this chapter is to study small sample relative efficiency of ML estimators with respect to MVU estimators of normal density and distribution functions. The normal distribution with unknown mean and unit standard deviation is considered. The measures of efficiency introduced in Chapter 3 are used here also.

4.2 Relative Efficiency of the ML Estimator with Respect to the MVU Estimator

Let $X_1, \ldots, X_n$ be a random sample of size n (≥ 2) from a distribution having the probability density function

$$f(x;\theta) = [1/(2\pi)]^{1/2} \exp[-(x-\theta)^2/2] , \qquad (4.1)$$

where θ ($-\infty < \theta < \infty$) denotes the unknown mean of the distribution, and $x \in R$ is any real number; for simplicity the known standard deviation is chosen to be unity. This section considers estimating the density function (4.1) and the corresponding distribution function $F(x;\theta)$ at an arbitrary point x.

4.2.1 Variance of the MVU Estimator

As is well known the sample mean $\bar{X}$ is a complete and sufficient statistic for this family, and its distribution law is normal with mean θ and variance $(1/n)$. The MVU estimator of $f(x;\theta)$ is denoted by $\tilde{f}(x;\bar{X})$ and is given by Seheult (1969)

$$\tilde{f}(x;\bar{X}) = [n/2\pi(n-1)]^{1/2} \exp[-n(x-\bar{X})^2/2(n-1)], \ x \in R . \qquad (4.2)$$

That is, the MVU estimator of a normal density with mean θ and unit variance is itself a normal density with mean $\bar{X}$ and variance $[(n-1)/n]$.

By direct integration it is easy to verify that $\tilde{f}(x;\bar{X})$ is an unbiased estimator of $f(x;\theta)$, $x \in R$; its second moment is found to be

$$E[\tilde{f}(x;\bar{X})]^2 = [n/2\pi(n^2-1)^{1/2}] \exp[-n(x-\theta)^2/(n+1)], \quad x \in R . \tag{4.3}$$

Thus,

$$\begin{aligned} \mathrm{Var}[\tilde{f}(x;\bar{X})] &= (1/2\pi)\exp[-(x-\theta)^2]\{[n/(n^2-1)^{1/2}] \\ &\quad \cdot \exp[(x-\theta)^2/(n+1)]-1\}, \quad x \in R . \end{aligned} \tag{4.4}$$

The variance function (4.4) depends on the sample size and the quadratic function $(x-\theta)^2$; it is symmetric about $x = \theta$ and is of order $O(1/n)$. It will be used in section 4.2.3 to compute the relative efficiency function. The variance of the MVU estimator is plotted as a function of $(x-\theta)$ for $n = 2, 6$ in Figure 4.1 with the mean square error of the ML estimator to be computed in the next section.

4.2.2 Mean Square Error of the ML Estimator

It is well known that the sample mean $\bar{X}$ is the ML estimator of θ in the normal distribution having density function (4.1). The ML estimator of $f(x;\theta)$ at a point x is denoted by $\hat{f}(x;\bar{X})$ and is given by

$$\hat{f}(x;\bar{X}) = [1/(2\pi)]^{1/2} \exp[-(x-\bar{X})^2/2], \quad x \in R . \tag{4.5}$$

That is, the ML estimator of a normal density with mean θ and unit variance is itself a normal density with mean $\bar{X}$ and variance one.

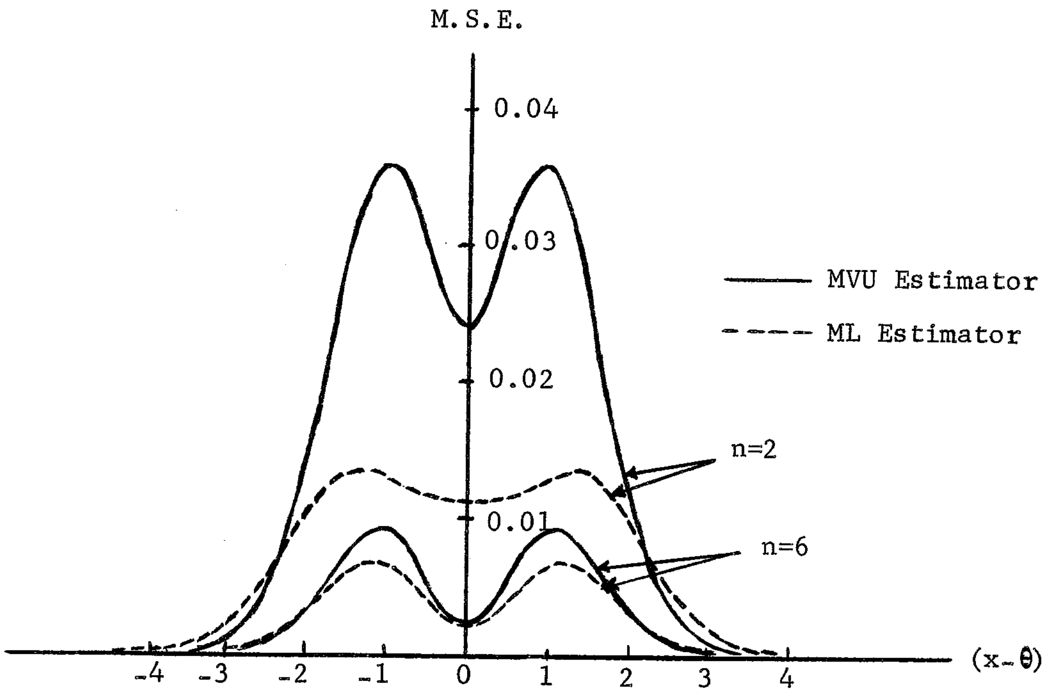

Figure 4.1 Variance function of the MVU estimator and mean square error of the ML estimator in the normal distribution with unit standard deviation; n = 2, 6

The first moment of the ML estimator is given by

$$E[\hat{f}(x;\bar{X})] = \{n/[2\pi(n+1)]\}^{1/2} \exp\{-[n(x-\theta)^2/2(n+1)]\}, \quad x \in R . \tag{4.6}$$

Thus, the ML estimator is biased and its bias function is

$$\begin{aligned} B_n(x;\theta) &= E[\hat{f}(x;\bar{X})] - f(x;\theta) \\ &= [1/(2\pi)]^{1/2}\exp[-(x-\theta)^2/2]\{[n/(n+1)]^{1/2} \\ &\quad \cdot \exp[(x-\theta)^2/2(n+1)] - 1\} \end{aligned} \tag{4.7}$$

which is non-negative everywhere and of order $0(1/n)$. The second moment is found to be

$$E[\hat{f}(x;\bar{X})]^2 = [1/(2\pi)][n/(n+2)]^{1/2} \exp[-n(x-\theta)^2/(n+2)], \quad x \in R . \tag{4.8}$$

Therefore, from (4.6) and (4.8),

$$\begin{aligned} \mathrm{Var}[\hat{f}(x;\bar{X})] = [1/(2\pi)]\exp[-(x-\theta)^2]\{[n/(n+2)]^{1/2} \\ \cdot \exp[2(x-\theta)^2/(n+2)] \\ - [n/(n+1)]\exp[(x-\theta)^2/(n+1)]\} , \quad x \in R . \end{aligned} \tag{4.9}$$

The mean square error of the ML estimator is deduced from (4.7) and (4.9) and is

$$\begin{aligned} \mathrm{MSE}[\hat{f}(x;\bar{X})] = [1/(2\pi)]\exp[-(x-\theta)^2]\{[n/(n+2)]^{1/2} \\ \cdot \exp[2(x-\theta)^2/(n+2)] \\ + 1 - 2[n/(n+1)]^{1/2} \exp[(x-\theta)^2/2(n+1)]\}, \quad x \in R . \end{aligned} \tag{4.10}$$

It is a function of the sample size and the quadratic function $(x-\theta)^2$, is symmetric about $x = \theta$, and of order $0(1/n)$. The mean square error of the ML estimator has been plotted as a function of $(x-\theta)$ for $n = 2, 6$ in Figure 4.1.

4.2.3 Relative Efficiency

The notion of relative efficiency introduced in section 3.2.3 can also be used in connection with density estimation.

Definition 4.1:

$$R.E.[\hat{f};\tilde{f}] = Var(\tilde{f}) \ / \ MSE(\hat{f}). \tag{4.11}$$

That is, the relative efficiency of the ML estimator with respect to the MVU estimator of a density function $f(x;\theta)$ is defined as the ratio of the variance of the MVU estimator to the mean square error of the ML estimator.

Thus, in the normal distribution (unit standard deviation), from (4.4) and (4.10),

$$R.E.[\hat{f};\tilde{f}] = \frac{[n/(n^2-1)^{1/2}]\exp[(x-\theta)^2/(n+1)]-1}{[n/(n+2)]^{1/2}\exp[2(x-\theta)^2/(n+2)]+1-2[n/(n+1)]^{1/2}\exp[(x-\theta)^2/2(n+1)]}, \quad x \in R. \tag{4.12}$$

This relative efficiency function depends on the sample size and the quadratic function $(x-\theta)^2$; its graph is symmetric about $x = \theta$. The following table gives the regions of the parameter space where it is greater than unity for selected values of n. Outside these intervals

Table 4.1 Ranges of the function $|x-\theta|$ for which the ML estimator is superior to the MVU estimator; n = 2, 6, 12, 20

n	Range
2	$0 < \lvert x-\theta\rvert < 2.103$
6	$0 < \lvert x-\theta\rvert < 1.794$
12	$0.206 < \lvert x-\theta\rvert < 1.723$
20	$0.236 < \lvert x-\theta\rvert < 1.695$

the function is less than unity. As the number of observations increases it can be verified that the function (4.12) tends to unity for all $x-\theta \neq 0$, but it tends to two-thirds at $x-\theta = 0$. Therefore, on the basis of the relative efficiency criterion the ML estimator is superior to the MVU estimator for the ranges given in Table 4.1.

A similar conclusion was also reached in Zacks and Even (1966) when estimating the reliability function, $1 - F(x;\theta)$.

The relative efficiency function (4.12) is plotted as a function of $(x-\theta)$ for $n = 2, 6, 12$ in Figure 4.2.

4.2.4 Closeness Relative Efficiency of the MVU Estimator with Respect to the ML Estimator

This section contains derivations of closeness functions of the MVU estimator and ML estimator, and of the closeness relative efficiency function in the case of a normal density with unit standard deviation. These notions have been defined in section 3.2.5 of the previous chapter.

The closeness function of the MVU estimator is denoted by $\tilde{\psi}_n(\epsilon;\theta)$, $0 < \epsilon < 1$, and is computed as follows:

$$\begin{aligned}\tilde{\psi}_n(\epsilon;\theta) = P_\theta\{&\,|[n/2\pi(n-1)]^{1/2} \exp[-n(x-\bar{X})^2/2(n-1)] \\ &- [1/(2\pi)]^{1/2} \exp[-(x-\theta)^2/2]| \\ &\leq \epsilon\cdot[1/(2\pi)]^{1/2} \exp[-(x-\theta)^2/2]\}\ , \quad x \in R\ . \end{aligned} \tag{4.13}$$

This equation can be written after some manipulations as

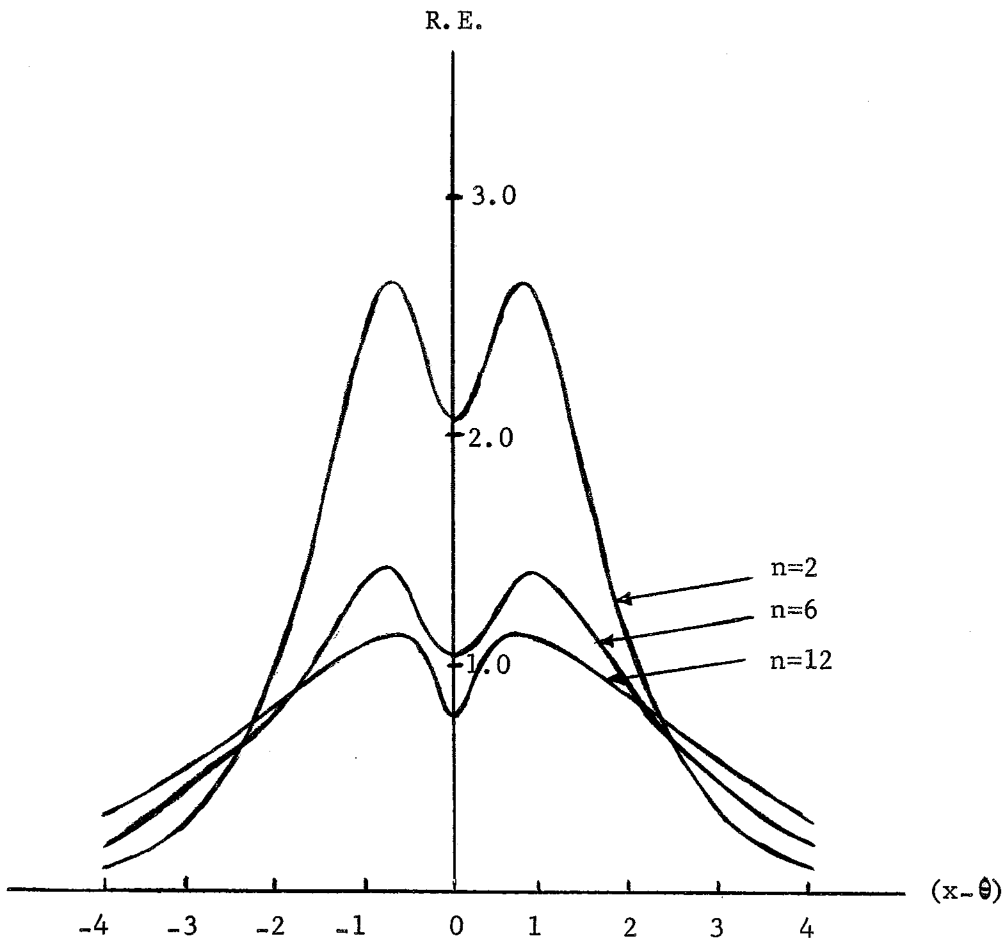

Figure 4.2 Relative efficiency function of the ML estimator with respect to the MVU estimator of the normal density function with unit standard deviation; n = 2, 6, 12

$$\tilde{\psi}_n(\epsilon;\theta) = P_\theta\{(n-1)(x-\theta)^2 - 2(n-1)\ln\{[(n-1)/n]^{1/2}(1+\epsilon)\} \leq n(\bar{X}-x)^2$$

$$\leq (n-1)(x-\theta)^2 - 2(n-1)\ln\{[(n-1)/n]^{1/2}(1-\epsilon)\}\}, \quad x \in R \, . \tag{4.14}$$

The sample mean $\bar{X}$ has a normal distribution with mean θ and variance $(1/n)$; therefore, the statistic $n(\bar{X}-x)^2$ is distributed as a non-central chi-square variate with one degree of freedom and noncentrality parameter (Graybill, 1961)

$$\lambda = n(x-\theta)^2/2 \ . \tag{4.15}$$

Thus, equation (4.14) can be evaluated with tables of a noncentral chi-square variate with one degree of freedom and noncentrality parameter λ given by (4.15).

The closeness function of the ML estimator denoted by $\hat{\psi}_n(\epsilon;\theta)$, $0 < \epsilon < 1$, is computed in a similar way and is given by

$$\hat{\psi}_n(\epsilon;\theta) = P_\theta\{n(x-\theta)^2 - 2n\ln(1+\epsilon) \le n(\bar{X}-x)^2 \le n(x-\theta)^2 - 2n\ln(1-\epsilon)\} \ , \quad x \in R \ , \tag{4.16}$$

which is also evaluated by means of tables of a noncentral chi-square variate with one degree of freedom and noncentrality parameter λ given by (4.15).

The closeness relative efficiency function has been defined as the ratio of the closeness functions. Thus, from (4.14) and (4.16),

$$\text{C.R.E.}[\tilde{f};\hat{f}] = \frac{P_\theta\{(n-1)(x-\theta)^2 - 2(n-1)\ln\{[(n-1)/n]^{1/2}(1+\epsilon)\} \le n(\bar{X}-x)^2 \le (n-1)(x-\theta)^2 - 2(n-1)\ln\{[(n-1)/n]^{1/2}(1-\epsilon)\}\}}{P_\theta\{n(x-\theta)^2 - 2n\ln(1+\epsilon) \le n(\bar{X}-x)^2 \le n(x-\theta)^2 - 2n\ln(1-\epsilon)\}} \ ,$$

$$x \in R \ . \tag{4.17}$$

This function depends on the sample size, the quadratic function $(x-\theta)^2$ and the proportionality factor ϵ. For a given $\epsilon (0 < \epsilon < 1)$, the

ML estimator is superior to the MVU estimator if x satisfies

$$2\ln(1-\epsilon) - (n-1)\ln[(n-1)/n] < (x-\theta)^2 < 2\ln(1+\epsilon)$$

$$- (n-1)\ln[(n-1)/n] \;. \qquad (4.18)$$

Some numerical values of this expression are given in Table 4.2 for different sample sizes and proportionality factors.

Table 4.2 Ranges of the quadratic function $(x-\theta)^2$ for which the ML estimator is superior to the MVU estimator; ϵ = 0.25, 0.50, 0.75; n = 2, 6, 12

	n = 2	n = 6	n = 12
ϵ = 0.25	(0.117, 1.139)	(0.336, 1.358)	(0.381, 1.403)
ϵ = 0.50	[0, 1.503)	[0, 1.722)	[0, 1.767)
ϵ = 0.75	[0, 1.813)	[0, 2.032)	[0, 2.077)

Moreover, a conjecture that the ML estimator is more efficient in most cases can be based on a comparison of the lower limits of integration in equations (4.14) and (4.16). The lower limit of integration in (4.16) is smaller than the lower limit in (4.14) if

$$(x-\theta)^2 < 2\ln(1+\epsilon) - (n-1)\ln[(n-1)/n] \;. \qquad (4.19)$$

An approximation of this inequality is, for large n,

$$(x-\theta)^2 < 1 + 2\ln(1+\epsilon) \;. \qquad (4.20)$$

A similar comparison of the upper limits of integration is less relevant since the amount of probability mass involved is smaller.

In some cases a final decision about the choice of the estimator can be made only by using tables of a noncentral chi-square distribution with one degree of freedom and noncentrality parameter λ given by (4.15). However, as the number of observations increases, the closeness relative efficiency function (4.17) tends to unity over its entire range and then one can use either one of these estimators of the normal density (4.1).

4.2.5 Tail Probabilities

The interesting feature of the approach here is that it considers the variability in the estimators themselves for various sample sizes.

Let $\tilde{F}(x;\bar{X})$ and $\hat{F}(x;\bar{X})$ be the cumulative distribution functions at x corresponding to the density estimators $\tilde{f}(x;\bar{X})$ and $\hat{f}(x;\bar{X})$, respectively. Also, let $\Phi(u)$ denote the standard normal integral at u $(-\infty < u < \infty)$, and $\Phi^{-1}(v)$ be the v^{th} quantile $(0 < v < 1)$ of the standard normal distribution.

For $n \geq 2$ and $0 < p < 1$, the smallest root of the equation

$$\hat{F}(x;\bar{X}) = p \tag{4.21}$$

is

$$\hat{\xi}_p = \bar{X} + \Phi^{-1}(p) \ . \tag{4.22}$$

As has been done before to compare the estimators $\tilde{F}(x;\bar{X})$ and $\hat{F}(x;\bar{X})$ it is instructive to evaluate the probability

$$\tilde{F}(\hat{\xi}_p;\bar{X}) = \alpha \ , \tag{4.23}$$

$[0 < \alpha < 1]$, and to compute the ratio of probabilities (α/p) for various sample sizes. This ratio has been used in the preceding chapter as a coefficient of resemblance of the estimating distribution functions; it is denoted by C.R.[n;p]. From equations (4.2), (4.22) and (4.23), it follows that

$$\alpha = \Phi\{[n/(n-1)]^{1/2}\ \Phi^{-1}(p)\}\ . \tag{4.24}$$

Thus, the coefficient of resemblance is given by

$$C.R.[n;p] = (1/p)\,\Phi\{[n/(n-1)]^{1/2}\ \Phi^{-1}(p)\}\ . \tag{4.25}$$

Numerical values of the coefficient of resemblance (4.25) are given in Table 4.3 for various values of p and n. For small p the distribution functions can be taken to be approximately equivalent only when n is large. When p is large, however, the coefficient of resemblance is very close to unity even in small sample situations.

Table 4.3 Values of the coefficient of resemblance in the normal distribution (known standard deviation); p = 0.01, 0.05, 0.95, 0.99

n	C.R.[n;0.01]	C.R.[n;0.05]	C.R.[n;0.95]	C.R.[n;0.99]
2	0.050	0.200	1.042	1.010
3	0.220	0.440	1.029	1.008
4	0.370	0.578	1.022	1.006
10	0.700	0.830	1.009	1.003
20	0.855	0.918	1.004	1.001
30	0.910	0.950	1.003	1.001

4.2.6 The Empirical Distribution Function

This section considers the gain in information from the additional assumption on the nature of the parent population.

It is known (Zacks and Even, p. 1042) that the variance of the MVU estimator $\tilde{F}(x;\bar{X})$ at a point x is given by

$$\mathrm{Var}[\tilde{F}(x;\bar{X})] = \Phi[(x-\theta),\ (x-\theta);\ (1/n)] - \Phi^2(x-\theta), \tag{4.26}$$

where $\Phi[u_1,u_2;\rho]$ denotes the standard bivariate normal integral at the limits u_1 and u_2 $(-\infty < u_1,\ u_2 < \infty)$ and a coefficient of correlation ρ $(-1 < \rho < 1)$, and $\Phi(u)$ is as before the standard normal integral at u $(-\infty < u < \infty)$.

On the other hand the variance of the empirical distribution function

$$S_n(x) = (1/n) \sum_{i=1}^{n} I_{(-\infty,x]}(X_i) \tag{4.27}$$

is given by

$$\mathrm{Var}[S_n(x)] = (1/n)[\Phi(x-\theta) - \Phi^2(x-\theta)] . \tag{4.28}$$

Expressions (4.26) and (4.28) are both of order $O(1/n)$ and are plotted as functions of $(x-\theta)$ for n = 2, 5 in Figure 4.3. Calculations are based on the values of the tails of the bivariate standard normal distribution; tables of the National Bureau of Standards are used.

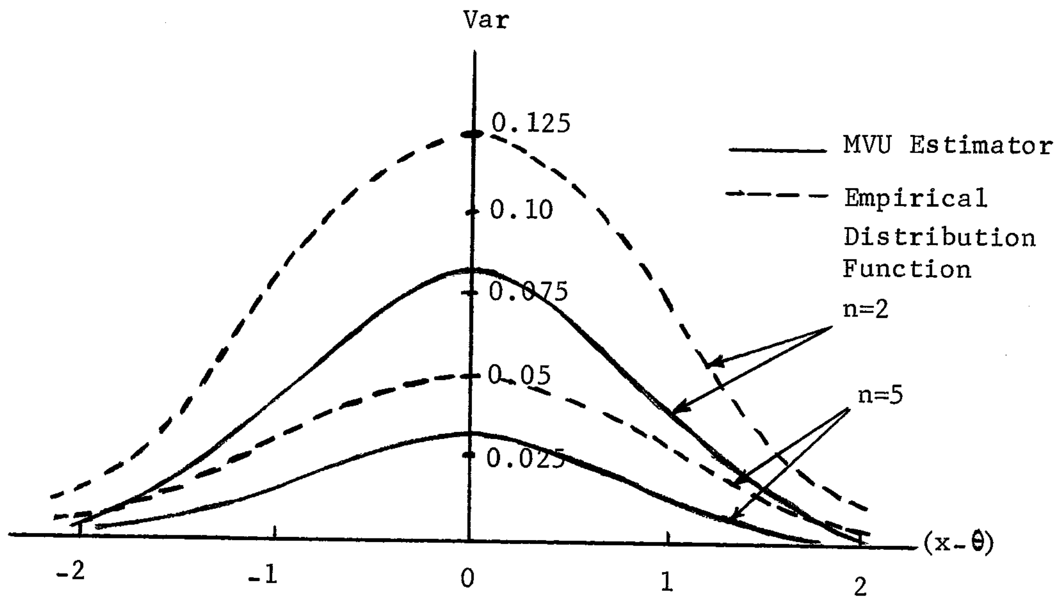

Figure 4.3 Variance functions of the empirical distribution function and of the MVU estimator in the normal distribution with unit standard deviation; n = 2, 5

The relative efficiency function R.E.$[S_n;\ \tilde{F}]$ is computed from (4.26) and (4.28); it is:

$$\text{R.E.}[S_n;\ \tilde{F}] = \frac{\Phi[(x-\theta),(x-\theta);(1/n)]-\Phi^2(x-\theta)}{(1/n)[\Phi(x-\theta)-\Phi^2(x-\theta)]}, \quad x \in R. \tag{4.29}$$

This is a (symmetric) function of $(x-\theta)$ and of the sample size; it is everywhere less than unity. By expanding the bivariate normal distribution function in powers of the coefficient of correlation $(1/n)$, it is easy to verify that R.E.$[S_n;\tilde{F}]$ tends to

$$[\varphi(x-\theta)]^2/[\Phi(x-\theta)-\Phi^2(x-\theta)] \tag{4.30}$$

as n increases, where $\varphi(t)$ is the standard normal density at t. The relative efficiency function (4.29) is plotted as a function of $(x-\theta)$ for n = 2, 5, 10 in Figure 4.4.

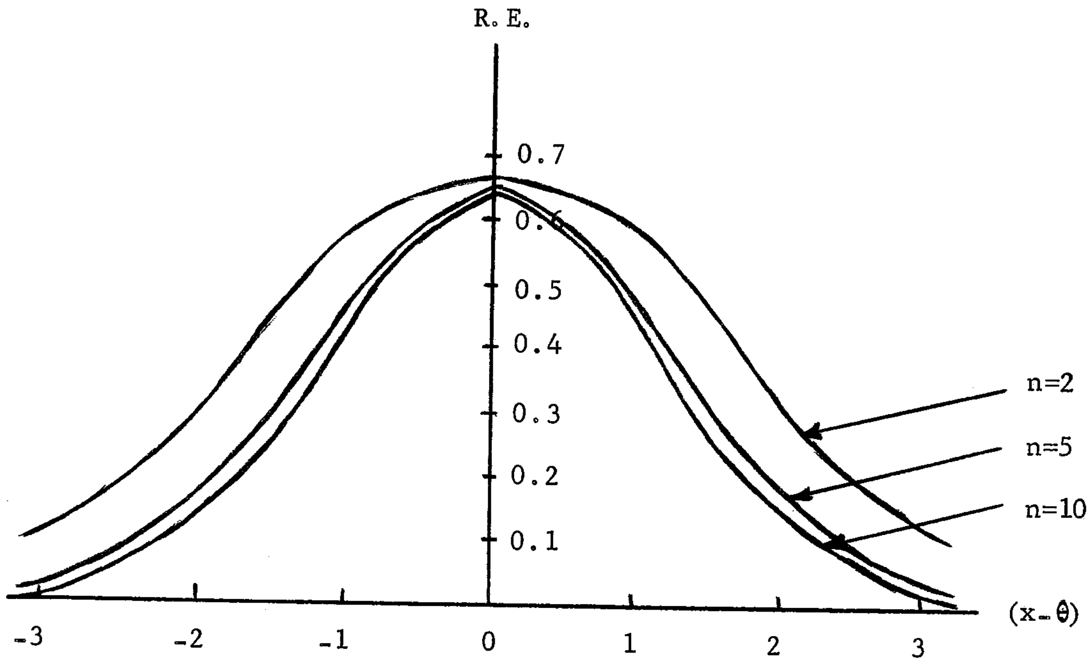

Figure 4.4 Relative efficiency of the empirical distribution function with respect to the MVU estimator in the normal case with unit standard deviation; n = 2, 5, 10

$S_n(x)$ in (4.27) is the MVU estimator of the parent distribution function under the assumption that it is absolutely continuous, whereas

$$\tilde{F}(x;\bar{X}) \equiv E[S_n(x)\,|\bar{X}] \tag{4.31}$$

is the MVU estimator of $F(x;\theta)$ under the additional assumption that the observations are taken from the normal population (4.1). Therefore, Figure 4.4 illustrates the gain in information in terms of the variance functions due to using the further assumption of normality.

From the foregoing results it is clear that when the sample comes from a normal population with known standard deviation, an empirical distribution function estimator has a much larger variance than the MVU estimator.

4.2.7 A Family of Estimators

For estimating the normal density function (4.1) this section considers general estimators of the form

$$f_c(x;\bar{X}) = [1/2\pi c^2]^{1/2} \exp[-(x-\bar{X})^2/2c^2], \quad x \in R, \tag{4.32}$$

where c is an arbitrary positive constant. The ML estimator $\hat{f} \equiv f_1$ and the MVU estimator $\tilde{f} \equiv f_{[(n-1)/n]^{1/2}}$ belong to this class of estimators.

By direct integration it is easy to verify that

$$\begin{aligned} MSE[f_c(x;\bar{X})] &= (1/2\pi)\exp[-(x-\theta)^2] \\ &+ [(n)^{1/2}/2\pi c(nc^2+2)^{1/2}] \exp[-n(x-\theta)^2/(nc^2+2)] \\ &- [(n)^{1/2}/\pi(nc^2+1)^{1/2}] \\ &\cdot \exp\{-(1/2)(x-\theta)^2[1+n/(nc^2+1)]\}, \quad x \in R. \end{aligned} \tag{4.33}$$

It is desired to choose the constant c so as to make the mean square error (4.33) small; no solutions however are acceptable in this case since they are in terms of n, x and θ. The following figure shows that for fixed sample size no choice of $c = c(x,\theta)$ can produce a best estimator over the entire range.

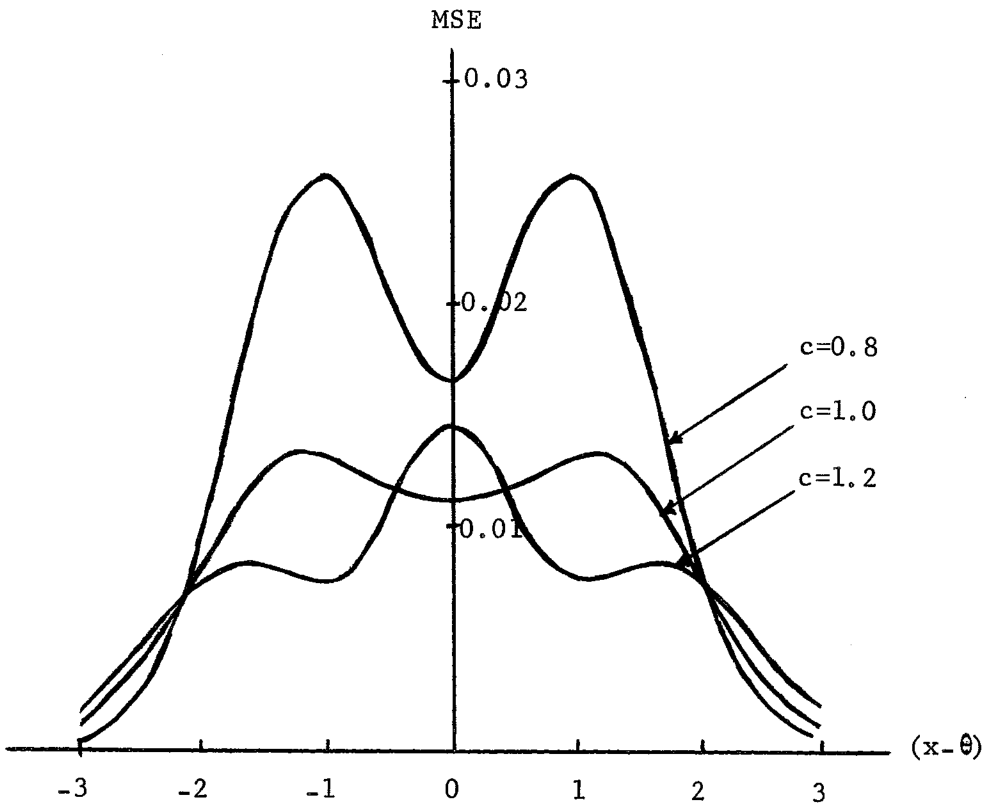

Figure 4.5 Mean square error of $f_c(x;\bar{X})$ in the normal distribution; $c = 0.8, 1.0, 1.2; n = 2$

Therefore, a global measure of goodness of an estimator is needed; the integrated mean square error IMSE can be used for this purpose.

From (4.33) one obtains

$$\text{IMSE} = \int_{-\infty}^{\infty} \text{MSE}[f_c(x;\bar{X})]dx = [1/2c(\pi)^{1/2}] + [1/(4\pi)^{1/2}] - [2n/(nc^2 + n + 1)\pi]^{1/2} . \qquad (4.34)$$

This expression is minimized when

$$c = [(n+1)/n]^{1/2} . \tag{4.35}$$

Thus, the minimum integrated mean square error (MIMSE) estimator is of the form

$$f_{[(n+1)/n]^{1/2}}(x;\bar{X}) = [n/2\pi(n+1)]^{1/2}$$

$$\cdot \exp[-n(x-\bar{X})^2/2(n+1)], \quad x \in R . \tag{4.36}$$

The mean square error of this estimator, computed from (4.33), is plotted in Figure 4.6 as a function of $(x-\theta)$ for $n = 2, 6$ with the mean square error of the ML estimator.

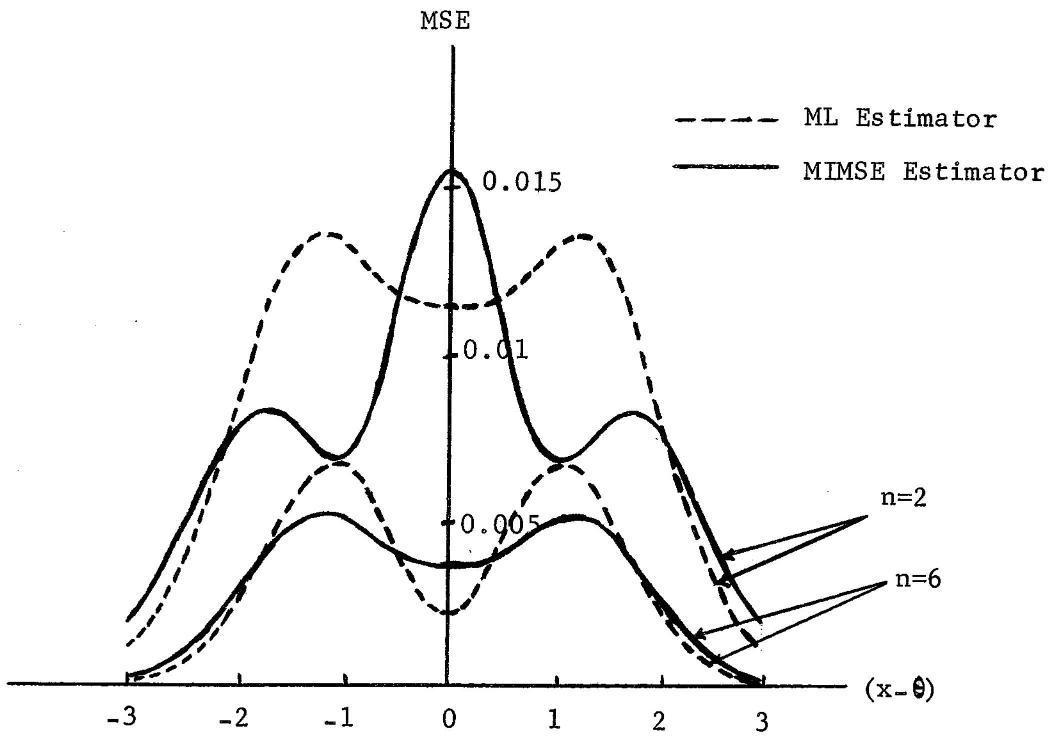

Figure 4.6 Mean square error functions of the ML estimator and MIMSE estimator of the normal density function with unit standard deviation; n = 2, 6

The relative efficiency of the MIMSE estimator with respect to the ML estimator is the ratio of the MSE of the ML estimator to the MSE of the other estimator; these are obtained from (4.33) by letting $c = 1$ and $c = [(n+1)/n]^{1/2}$, respectively. The resulting function is plotted in Figure 4.7 for $n = 2$, 6 and 12.

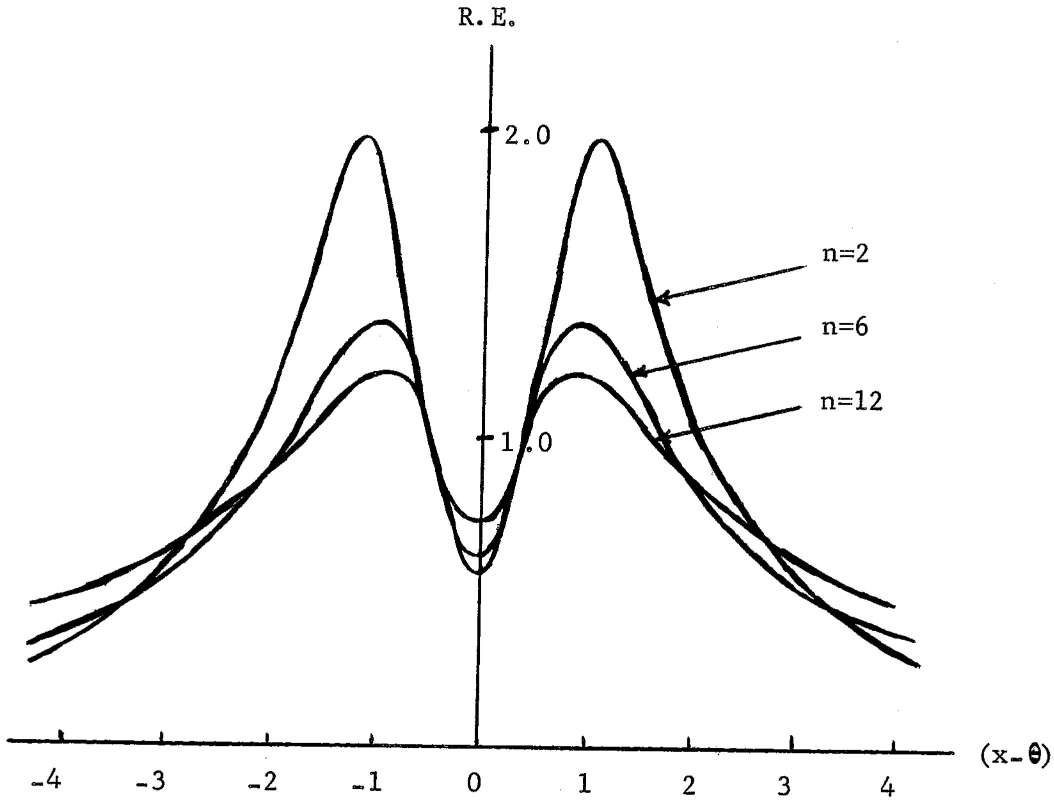

Figure 4.7 Relative efficiency of the MIMSE estimator with respect to the ML estimator of the normal density function with unit standard deviation; $n = 2, 6, 12$

It is to be observed that the ML estimator is less efficient than the MIMSE estimator if $0.467 < |x-\theta| < 2.095$, $n = 2$;

$0.447 < |x-\theta| < 1.748$, $n = 6$; and $0.437 < |x-\theta| < 1.651$, $n = 12$. As the sample size increases it can be shown that the relative efficiency function tends to unity for all $x - \theta \neq 0$, but it tends to one-half at $x - \theta = 0$.

In the same manner the relative efficiency of the MIMSE estimator with respect to the MVU estimator is computed as the ratio of the variance of the MVU estimator to the mean square error of the other estimator; these are again derived from (4.33) by letting $c = [(n-1)/n]^{1/2}$ and $c = [(n+1)/n]^{1/2}$, respectively. This function is plotted in Figure 4.8 for n = 2, 6 and 12.

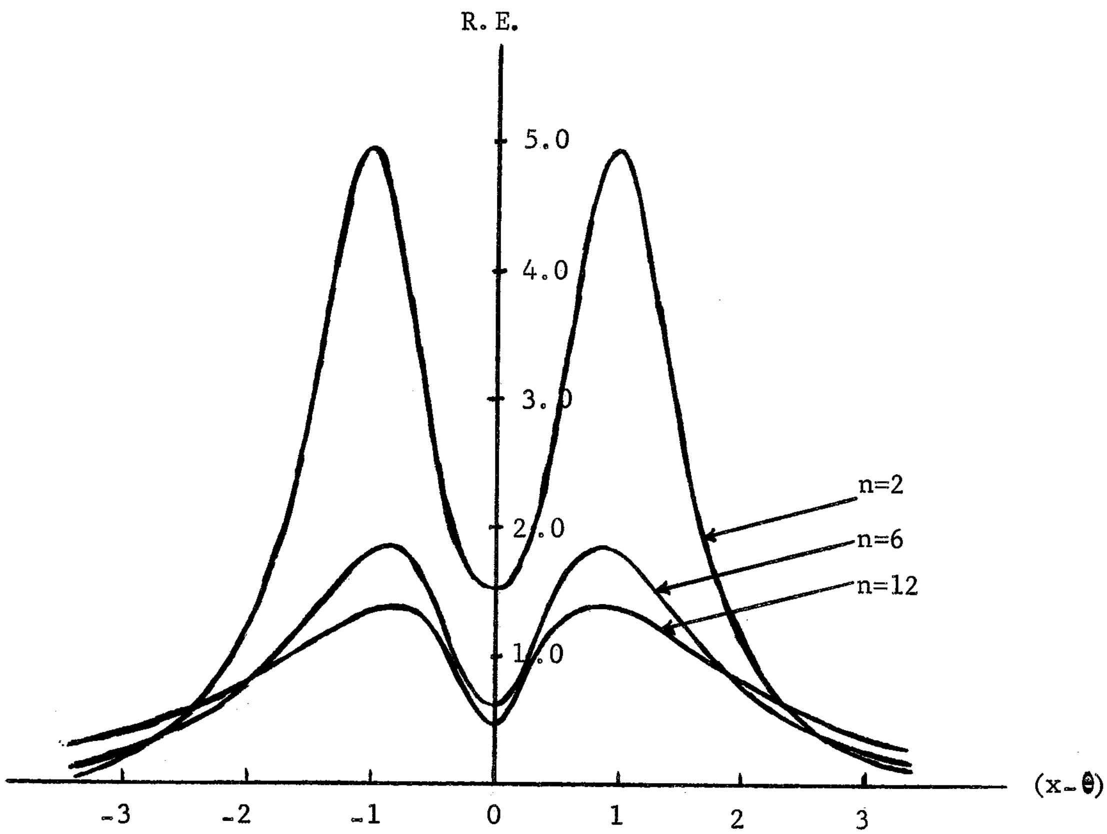

Figure 4.8 Relative efficiency of the MIMSE estimator with respect to the MVU estimator; n = 2, 6, 12

In this case the MVU estimator is less efficient that the MIMSE estimator if

$$0 \leq |x - \theta| \leq 2.100,\ n = 2; \quad 0.299 \leq |x - \theta| \leq 1.771,\ n = 6;$$

and $0.339 \leq |x - \theta| \leq 1.685$, $n = 12$. This relative efficiency function also tends to one-half at $x - \theta = 0$ as the sample size increases and to unity for all $x - \theta \neq 0$.

4.2.8 Summary

In estimating the normal density function (known standard deviation) it follows from the results of section 4.2.3 that the MVU estimator is more efficient than the ML estimator for estimating tails of the distribution; a similar conclusion is also reached under the closeness relative efficiency criterion.

When dealing with the general family of density estimators in section 4.2.7 a global measure of goodness is needed and integrated mean square error is chosen. Both the ML and MVU estimators are superior to the minimum integrated mean square error estimator for tails of the distribution.

In another section the use of the additional assumption on the nature of the parent population results in a significant gain in information when estimating the normal distribution function (known standard deviation); this is especially true in the tails of the distribution.

5. CONCLUSION

The work in this thesis is concerned mainly with the problem of small sample relative efficiency of ML estimators with respect to MVU estimators of distribution and density functions. The following classes of distributions on the real line are assumed as a probabilistic model for the observations: the uniform, exponential with location and scale parameters, and normal (known standard deviation).

In each case various kinds of relative efficiency functions are considered. The first one is the ratio of the mean square error of the estimators. The closeness relative efficiency function (Zacks and Even, 1966) is defined in Chapter 3. Comparisons of tail probabilities provide another approach to comparing estimators. In estimating the uniform and normal distribution functions the relative efficiency of the empirical distribution function with respect to the MVU estimator is computed; it measures the gain in information in terms of the variance functions due to using an additional assumption on the nature of the parent population. For these two distributions general families of estimators are also considered.

In Chapter 3 the distribution function is estimated for the uniform distribution and exponential with location and scale parameters. In the uniform case estimators are examined under various criteria of relative efficiency.

When sampling from the uniform distribution on $(0, \theta)$ one will prefer the MVU estimator to estimate the parent distribution function.

This result has been strengthened by consideration of general classes of estimators.

When the sample is taken from an exponential distribution with location parameter θ different recommendations are made depending on the (anticipated) distance $(x-\theta)$. For small sample sizes the ML estimator of the distribution function is preferable on the basis of the relative efficiency criterion at those x values near the parameter θ (lower tail); for x values away from θ (upper tail) the MVU estimator is more efficient.

In the exponential case with scale parameter the MVU estimator is better for estimating lower and upper tails; the ML estimator is more efficient, otherwise.

The normal density and distribution functions (known standard deviation) are considered in Chapter 4. The results of sections 4.2.3 and 4.2.7 suggest that no density estimator is superior to the others over the entire range; the choice of a best estimator depends on $|x - \theta|$. However, the MVU estimator is in any case most efficient for the tails of the distribution. In the estimation of the parent distribution function the empirical distribution function is significantly less efficient than the unbiased estimator based on the complete sufficient statistic $\bar{X}$.

When unbiasedness is regarded as an essential property of the estimator then the MVU estimator should be used. However, in some situations it might be appropriate to consider biased estimators; for example, in the estimation of a normal density (known standard deviation) the biased ML estimator can sometimes have a small probability of a large estimation error.

In some cases a choice between the MVU estimator and other estimators based upon these relative efficiency criteria can only be made if some idea of the region of the parameter space is available. However, in any practical estimation problem a small departure from the additional assumption on the parent population might affect the selection of an estimator.

Other classes of distributions on the real line could be considered but in most cases only computer approximations could be obtained. For example, the normal distribution with both parameters unknown has been studied by Zacks and Milton (1969).

Additional areas for future research include the following:

1. Study inference procedures based on minimum variance unbiased estimators.
2. Study other criteria of selection of the estimators.
3. Study other classes of estimators.

6. LIST OF REFERENCES

Barton, D. E. 1961. Unbiased estimation of a set of probabilities. Biometrika 48:227-229.

Basu, A. P. 1964. Estimates of reliability for some distributions useful in life testing. Technometrics 6:215-219.

Epstein, B. and Sobel, M. 1954. Some theorems relevant to life testing from an exponential distribution. Ann. Math. Stat. 25:373-381.

Fisher, R. A. 1912. On an absolute criterion for fitting frequency curves. Mess. of Math. 41:155-160.

Folks, J. L., Pierce, D. A., and Stewart, C. 1965. Estimating the fraction of acceptable product. Technometrics 7:43-50.

Ghurye, S. G. and Olkin, Ingram. 1969. Unbiased estimation of some multivariate probability densities and related functions. Ann. Math. Stat. 40:1261-1271.

Graybill, F. A. 1961. An Introduction to Linear Statistical Models. McGraw-Hill Book Company, Inc., New York, New York.

Kolmogorov, A. N. 1950. Unbiased estimates. Izvestia Akad. Nauk SSSR, Seriya Matematiceskaya 14:303-326. Amer. Math. Soc. Translation No. 98 (1953).

Laurent, A. G. 1963. Conditional distribution of order statistics and distribution of the reduced i-th order statistic of the exponential model. Ann. Math. Stat. 34:652-657.

Lieberman, G. J. and Resnikoff, G. J. 1955. Sampling plans for inspection by variables. J. Amer. Stat. Assoc. 50:457-516.

Noack, A. 1950. A class of random variables with discrete distributions. Ann. Math. Stat. 21:127-132.

Patil, G. P. 1963. Minimum variance unbiased estimation and certain problems of additive number theory. Ann. Math. Stat. 34:1050-1056.

Pugh, E. L. 1963. The best estimate of reliability in the exponential case. Opns. Res. 11:57-61.

Robertson, T. 1967. On estimating a density which is measurable with respect to a σ-lattice. Ann. Math. Stat. 38:482-493.

Rosenblatt, M. 1956. Remarks on some nonparametric estimates of a density function. Ann. Math. Stat. 27:832-837.

Roy, J. and Mitra, S. K. 1957. Unbiased minimum variance estimation in a class of discrete distributions. Sankhya 18:371-378.

Sathe, Y. S., and Varde, S. D. 1969. On minimum variance unbiased estimation of reliability. Ann. Math. Stat. 40:710-714.

Seheult, A. H. 1969. On unbiased estimation of density functions. PhD thesis, Department of Experimental Statistics, North Carolina State University at Raleigh.

Tate, R. F. 1959. Unbiased estimation: functions of location and scale parameters. Ann. Math. Stat. 30:341-366.

U. S. Department of Commerce. 1959. Tables of the Bivariate Normal Distribution Function and Related Functions. U. S. Department of Commerce, National Bureau of Standards, Applied Mathematics Series 50.

Washio, Y., Morimoto, H., and Ikeda, N. 1956. Unbiased estimation based on sufficient statistics. Bull. Math. Stat. 6:69-94.

Wegman, E. J. 1969. Nonparametric probability density estimation. Institute of Statistics Mimeo Series # 638, University of North Carolina at Chapel Hill.

Wegman, E. J. 1970. Maximum likelihood estimation of a unimodal density function. Ann. Math. Stat. 41:457-471.

Zacks, S. and Even, M. 1966. The efficiencies in small samples of the maximum likelihood and best unbiased estimators of reliability functions. J. Amer. Stat. Assoc. 61:1033-1051.

Zacks, S. and Milton, R. C. 1969. Mean square errors of the best unbiased and maximum likelihood estimators of tail probabilities in normal distributions. Technical Report 185, Department of Mathematics and Statistics, University of New Mexico, Albuquerque.

NORTH CAROLINA STATE UNIVERSITY
INSTITUTE OF STATISTICS
(Mimeo Series available at 1¢ per page)

652. Morimura, Hidenori. On Some Preventive Maintenance Policies for IFR.

653. Fuguira, Nariaki. Asymptotic Expansions of the Distribution of $S_1S_2^{-1}$ and $S_2(S_1 + S_2)^{-1}$ for Testing the Equality of Two Covariance Matrices.

654. Hayakawa, Takesi. The Asymptotic Distribution of Statistics Based on the Complex Gaussian Distribution.

655. Joshi, Prakash. Efficient Estimation of the Mean of an Exponential Distribution When an Outlier is Present.

656. Talwar, P. P. Age Patterns of Fertility.

657. Smith, W. L. and W. E. Wilkinson. Branching Processes in Markovian Environments.

658. Sen, P. K. and Malay Ghosh. Sequential Confidence Interval for the Regression Coefficient Based on Kendall's Tau.

659. Sen, P. K. Robust Statistical Procedures in Quantitative Bio-assays, I.

660. Kupper, Lawrence L. The Admissibility of a Response Surface Design.

661. Cleveland, W. S. Fitting Time Series Models for Prediction.

662. Gould, F. J. and Jon W. Tolle. A Necessary and Sufficient Qualification for Constrained Optimization.

663. Davis, Clarence E. Some Measures of Skewness.

664. Weissner, Edward N. Multitype Branching Processes in Random Environments.

665. Johnson, N. L. Comparison of Some Tests of Sample Censoring of Extreme Values.

666. Fergany, Nader. On the Macro-Dynamic Stochastic Treatment of the Size and Age Structure of a Human Population.

667. Hayakawa, Takesi. On the Distribution of the Latent Roots of a Complex Wishart Matrix (Non-Central Case).

668. Kelly, Douglas G. A Condition Equivalent to Ferromagnetism for a Generalized Ising Model.

669. Kleinbaum, David G. Estimation and Hypothesis Testing for Generalized Multivariate Linear Models.

670. Ridgeway, Linda. A Generalized Computer Program for the Tabulation of Data.

671. Haseman, Joseph. The Genetic Analysis of Quantitative Traits Using Twin and Sib Data.

672. Koch, Gary and Don Reinfurt. The Analysis of Categorical Data from Mixed Models.

673. Cleveland, W. S. Projection with the Wrong Inner Product and Its Application to Regression with Correlated Errors and Linear Filtering of Time Series.

674. Cambanis, S. and E. Masry. On the Representation of Weakly Continuous Stochastic Processes.

675. Johnson, W. D. and Gary Koch. Some Aspects of a Statistical Analysis of Qualitative Data from Sample Surveys. Part I. Proportions Mean Scores and Other Linear Functions.

676. Lee, Eun Sul. Aspects of Institutional Mobility Patterns of Chemists in Higher Education.

677. Osborn, J. K. and G. D. Simons. A More Precise Estimate of the Cost of Not Knowing the Variance When Making a Fixed-Width Confidence Interval for the Mean.

678. Kupper, Lawrence L. Optimal Response Surface Techniques Using Fourier Series and Spherical Harmonies.

679. Hoeffding, Wassily. The L_1 Norm of the Approximation Error for Bernstein-Type Polynomials.

680. Sen, P. K. and Malay Ghosh. On Bounded Length Confidence Intervals from the Regression Coefficient Based on a Class of Rank Statistics.

681. Simons, Gordon. Identifying Probabilities Limits.

682. Hayakawa, Takesi. On the Distribution of a Trace of a Multivariate Quadratic Form in the Multivariate Normal Sample.

683. Johnson, N. L. Tables to Facilitate Fitting S_B Curves.
Kitchen, J. O. Appendix: Calculation of the Tables.