1978

Inferences on the shape parameter of the gamma distribution and Cramer-Rao lower bounds from censored data

James Wyckoff

INFERENCES ON THE SHAPE PARAMETER OF THE GAMMA DISTRIBUTION AND CRAMÉR-RAO LOWER BOUNDS FROM CENSORED DATA

BY

JAMES WYCKOFF, 1950-

A DISSERTATION

Presented to the Faculty of the Graduate School of the

UNIVERSITY OF MISSOURI-ROLLA

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

1978

M. E. Engelhardt
Advisor

Lee J. Bain

F. T. Wright

Jagdish K. Patel

Billy E. Gillett

PUBLICATION THESIS OPTION

This dissertation has been prepared in the form of two papers ready for publication. The format for each paper is in the style of the Journal of the American Statistical Association. This dissertation will be submitted as two separate papers consisting of pages 1 - 38 and 39 - 79, respectively.

ABSTRACT

This dissertation is presented in publication form and consists of two articles. The first article considers inferential procedures on the shape parameter of a gamma distribution from censored sampling. Moments for the statistic $T = \log(\bar{x}_r/\tilde{x}_r)$ are found and used to derive a two-moment chi-square approximation for T. This approximation is then used for testing, estimating, and setting confidence bounds on the shape parameter of a gamma distribution. The second article concerns the Cramér-Rao lower bounds for the variances of estimators, where the estimators are based on censored data. Convenient techniques are derived to evaluate the lower bounds in the presence or absence of nuisance parameters.

ACKNOWLEDGMENT

The author wishes to express his sincere appreciation to Dr. Max Engelhardt of the Department of Mathematics for his aid in the selection of the topics and his many suggestions for the preparation of this dissertation.

The author also wishes to express his heartfelt appreciation to his friends and family for their unselfish encouragement and support during his graduate studies.

TABLE OF CONTENTS

TABLE OF CONTENTS, CONTINUED

LIST OF TABLES

LIST OF TABLES, CONTINUED

Inferential Procedures on the Shape Parameter of a Gamma Distribution from Censored Data

James Wyckoff and Max Engelhardt

ABSTRACT

Various techniques are used to find the first two-moments of the statistic $T = \log (\bar{x}_r/\tilde{x}_r)$, where $\bar{x}_r$ and $\tilde{x}_r$ are, respectively, the Windsorized arithmetic and geometric means based on censored data from a gamma distribution. A two-moment chi-square approximation is then derived for T and used for testing, estimating, and setting confidence bounds on the shape parameter of a gamma distribution. Power comparisons are made between optimum tests and the corresponding tests based on T.

KEY WORDS: Gamma distribution; Censored sampling; Tests of hypothesis; Confidence bounds; Median unbiased estimation.

AUTHOR'S FOOTNOTE

*James Wyckoff is a graduate student and Max Engelhardt is Associate Professor both at the University of Missouri-Rolla, Rolla, MO 65401. This paper is a portion of a dissertation presented in partial fulfillment of the requirements for the Ph.D. degree in Mathematics at the University of Missouri-Rolla and was supported in part by the U. S. Army Research Office under Grants DAAG29-77-G-0128 and DAAG29-78-G-0094.

1. INTRODUCTION

Let $x_1 < x_2 < \cdots < x_n$ denote the order statistics of a random sample of size n from the two-parameter gamma distribution having density function

$$f(x;\beta,\theta) = [\beta^{\theta}\Gamma(\theta)]^{-1}x^{\theta-1}\exp[-x/\beta],$$

$$x > 0, \ \beta > 0, \ \theta > 0.$$

The gamma distribution, of which the chi-square and exponential distributions are particular cases, provides a useful population model in many areas of statistics such as life tests, reliability, and acceptance sampling based on life tests (Barlow and Proschan 1965, Gupta 1960, and Gupta and Groll 1961). In particular, a significant property of the gamma distribution is that the hazard rate function, $h(x) = f(x)/[1-F(x)]$, is increasing (decreasing) if $\theta > 1$ ($\theta < 1$) and converges to $1/\beta$ as $x \to \infty$. If $\theta = 1$ then the distribution is exponential and $h(x) = 1/\beta$ for all $x > 0$.

In this paper we are interested in inference procedures for θ when only the first r out of n order statistics, $x_1 < x_2 < \cdots < x_r$, are available. The statistic that we shall base our procedures on is

$T = \log(\bar{x}_r/\tilde{x}_r)$, where $\bar{x}_r = (\sum_{i=1}^{r} x_i + (n-r)x_r)/n$ and $\tilde{x}_r = [(\Pi_{i=1}^{r} x_i)(x_r^{n-r})]^{1/n}$. We note that $\bar{x}_r$ and $\tilde{x}_r$ are the Winsorized arithmetic and geometric means, respectively. The Windsorization method, as discussed by Engelhardt (1975a), substitutes the nearest available observation for the censored observations, in either extreme. Under the full sample case (r=n) various authors have considered inference procedures for θ based on the statistic $S = \log(\bar{x}_n/\tilde{x}_n)$, where $\bar{x}_n$ and $\tilde{x}_n$ are, respectively, the arithmetic and geometric means. It has been shown that the uniformly most powerful scale invariant and the uniformly most powerful unbiased hypothesis tests concerning θ can be based, in the full sample case, solely in terms of S (Glaser 1976, Linhardt 1965, and Shorack 1972).

In this article we derive a two-moment chi-square approximation for the distribution of $W = 2n\theta T$. A similar procedure has been done for the full sample case, $W = 2n\theta S$, by Bain and Engelhardt (1975). In this case it was possible to derive the exact mean and variance for any choice of n and θ, and comparisons could be made with a limited number of exact percentage points given by Bishop and Nair (1939). But in the censored sample case ($r < n$) it has been necessary to find the mean, variance, and distribution of W, for various values of θ, n, and r, by Monte-Carlo techniques. We do, however, find the exact expectation and the exact variance for W when $\theta = 1$, the exponential case, for various values of n and r. The exact asymptotic moments, as $n \to \infty$ with $r/n \to p > 0$, for different censoring levels, and techniques to find the exact moments as $\theta \to \infty$, for different censoring levels, are also given.

The chi-square approximation can be easily obtained from the tables provided. A table comparing exact and approximate percentage points is given, along with tables indicating the accuracy of the Monte-Carlo results. A power comparison is also given, comparing the power of the test based on W to the power of the Neyman-Pearson test of the simple hypothesis $H:\theta = \theta_0$ versus the simple alternative $K:\theta = \theta_1$. Even though the test based on W is a composite test, that is we need only know if $\theta_1 > \theta_0$ or if $\theta_1 < \theta_0$, we show that its power is extremely close to the most powerful test. Finally the usefulness of the approximation is illustrated in setting up hypothesis tests on θ, confidence bounds on θ, and in finding a median unbiased estimate for θ.

2. THE CHI-SQUARE APPROXIMATION

To obtain a two-moment chi-square approximation for the distribution of $W = 2n\theta T$ we set

$$c(\theta,n,r) = 2E(W)/Var(W)$$

and

$$\nu(\theta,n,r) = 2[E(W)]^2/Var(W),$$

where $E(W)$ and $Var(W)$ are the expected value and the variance of W, respectively. Thus, an approximate chi-square distribution with correct first and second moments is given by

$$c(\theta,n,r)W \sim \chi^2[\nu(\theta,n,r)].$$

Tabulated expected values of θT and variances of $n^{\frac{1}{2}}\theta T$ can be found in Table 1 that cover a wide range of θ, censoring levels $r/n = .3, .5, .7, .9$ plus the full sample case, and $n = 10, 20, 40, \infty$. In Table 1 interpolation on r/n, on $1/n$, and on $1/\theta$ is excellent for $\theta E(T)$ and $n\theta^2 Var(T)$.

The values in Table 1 for the full sample case, along with the chi-square approximation for the full sample case, can be found in Bain and Engelhardt (1975).

The tabulated values pertaining to the censored cases, except for the asymptotic results and the values obtained where $\theta = 1$, the exponential case, are Monte-Carlo results. The details of the Monte-Carlo simulation and the accuracy of the results are discussed in Section 4.

The methods of Chernoff, Gastwirth, and Johns (1967) are applied to find the asymptotic joint distribution of $\bar{x}_r$ and log $\tilde{x}_r$. Then proposition (ii), 6a.2 of Rao (1965) is applied to find the asymptotic normal distribution of T. For each n, define random variables T_{1n} and T_{2n} as follows:

$$T_{1n} = (1/n) \sum_{i=1}^{r} x_i + [(n-r)/n]x_r$$

and

$$T_{2n} = (1/n) \sum_{i=1}^{r} \log(x_i) + [(n-r)/n]\log x_r,$$

where $x_1 < x_2 < \cdots < x_r$ are the ordered gamma random variables. If $r/n \to p$ as $n \to \infty$, say $r = [np]+1$, where $[np]$ is the largest integer not exceeding np, then $(n-r)/n \to 1-p$.

Corollary 3 of Chernoff et al. (1967) is applied with $h_1(x) = x$, $H_1(u) = F^{-1}(u)$, $a_1 = (1-p)$, $J_1(u) = 1$ if $0 < u \leq p$ and 0 if $u > p$ for the random variable T_{1n}. The notations $F(\cdot)$ and $f(\cdot)$ are used to signify the cdf and the density of the gamma distribution, respectively. For the random variable T_{2n} we apply the methods of Chernoff et al. (1967) with $h_2(x) = \log x$, $H_2(u) = \log(F^{-1}(u))$, $a_2 = (1-p)$, $J_2(u) = 1$ if $0 < u \leq p$ and 0 if $u > p$. T_{1n} and T_{2n} can be shown to have the same asymptotic distributions as

$$T_{1n}' = (1/n) \sum_{i=1}^{n} J_1(i/(n+1))x_i + a_1 x_r$$

and

$$T_{2n}' = (1/n) \sum_{i=1}^{n} J_2(i/(n+1))\log x_i + a_2 \log x_r,$$

respectively.

The asymptotic mean for T_{1n}' is

$$\mu_1 = \int_0^1 J_1(u)H_1(u)du + a_1 h_1(F^{-1}(p))$$

$$= \int_0^p F^{-1}(u)du + (1-p)F^{-1}(p)$$

and the asymptotic variance is $\sigma_1{}^2/n$ where

$$\sigma_1{}^2 = \int_0^1 \alpha_1{}^2(u)du,$$

with $\alpha_1(u) = 0$ if $u > p$ and

$$\alpha_1(u) = [1/(1-u)]\{\int_u^p H_1'(w)(1-w)dw + (1-p)^2 H_1'(p)\}$$

if $u \le p$, where $H_1'(\cdot)$ denotes the derivative of $H_1(\cdot)$. Finally

$$\mu_1 = \theta p + (1-p)F^{-1}(p) - f(F^{-1}(p))$$

and

$$\begin{aligned}\sigma_1{}^2 = {} & 2(1-p)(1-p-\theta p)F^{-1}(p) \\ & + F^{-1}(p)f(F^{-1}(p))(2\theta p-\theta-1) \\ & - [F^{-1}(p)]^2[f(F^{-1}(p))]^2 \\ & + [F^{-1}(p)]^2[f(F^{-1}(p))](1-2p) \\ & - 2\theta(1-p)^2 p/[f(F^{-1}(p))] \\ & + \theta p(\theta-\theta p+1) \\ & + (1-p)p[F^{-1}(p)+(1-p)/f(F^{-1}(p))]^2.\end{aligned}$$

The asymptotic mean for T_{2n}' is

$$\mu_2 = \int_0^1 J_2(u)H_2(u)du + a_2h_2(F^{-1}(p))$$

$$= \int_0^p \log(F^{-1}(u))du + (1-p)\log(F^{-1}(p))$$

and the asymptotic variance is $\sigma_2^{\,2}/n$ where

$$\sigma_2^{\,2} = \int_0^1 \alpha_2^{\,2}(u)du,$$

with $\alpha_2(u) = 0$ if $u > p$ and

$$\alpha_2(u) = [1/(1-u)]\{\int_u^p H_2'(w)(1-w)dw + (1-p)^2H_2'(p)\}$$

if $u \leq p$. Finally

$$\mu_2 = \log(F^{-1}(p)) - \int_0^{F^{-1}(p)} F(w)/w dw$$

and

$$\sigma_2^{\,2} = \{2\log(F^{-1}(p))+2(1-p)^2/[F^{-1}(p)]f(F^{-1}(p))\}$$

$$\cdot\int_0^{F^{-1}(p)} F(w)/w dw$$

$$+\int_0^{F^{-1}(p)} f(w)\log^2(w)dw$$

$$- p \log^2(F^{-1}(p)) - [\int_0^{F^{-1}(p)} F(w)/wdw]^2$$

$$+ p(1-p)^3/[F^{-1}(p)]^2[f(F^{-1}(p))]^2.$$

The asymptotic covariance of T_{1n}' and T_{2n}' is σ_{12}/n where

$$\sigma_{12} = \int_0^1 \alpha_1(u)\alpha_2(u)du,$$

with $\alpha_1(u)$ and $\alpha_2(u)$ defined as before. Finally we have

$$\sigma_{12} = 1-p(1-p)+p(1-p)^2[F^{-1}(p)-\theta+(1-p)/f(F^{-1}(p))]$$

$$+ \{(1-p)F^{-1}(p)-[F^{-1}(p)]f(F^{-1}(p))-\theta(1-p)$$

$$+ (1-p)^2/f(F^{-1}(p))\}\cdot\int_0^{F^{-1}(p)} F(w)/wdw.$$

The above quantities were found numerically and the results were applied to proposition (ii), 6a.2 of Rao (1965), with $g(T_{1n},T_{2n}) = \log(T_{1n})-T_{2n}$. Thus, the asymptotic mean of T was found to be $\log(\mu_1)-\mu_2$ and the asymptotic variance of T was found to be $(\sigma_2{}^2+\sigma_1{}^2/\mu_1{}^2-2\sigma_{12}/\mu_1)/n$.

When $\theta = 1$, the exponential case, the exact moments of T can be found since $\bar{x}_r$ is then a complete and sufficient statistic for β (Mann, Schafer, and Singpurwalla 1974). Using the fact that T is distributed independently of β it follows, by the results of Basu (1955), that $\bar{x}_r$ and T are independent statistics. Likewise, $\exp(-T) = \tilde{x}_r/\bar{x}_r$ and $\bar{x}_r$ are independent statistics. Considering the moment generating function of $-T = \log(\tilde{x}_r/\bar{x}_r)$, $M_{-T}(t) = E[(\tilde{x}_r/\bar{x}_r)^t]$, we apply the independence of $(\tilde{x}_r/\bar{x}_r)$ and $\bar{x}_r$. Thus, $E[(\tilde{x}_r/\bar{x}_r)^t(\bar{x}_r)^t]$ $= E[(\tilde{x}_r/\bar{x}_r)^t]E[(\bar{x}_r)^t]$ and $E[(\tilde{x}_r/\bar{x}_r)^t(\bar{x}_r)^t] = E[(\tilde{x}_r)^t]$. Combining these results we have $M_{-T}(t) = E[(\tilde{x}_t)^t]/E[(\bar{x}_r)]^t$. Using the methods of Engelhardt (1975b) to find $E[(\tilde{x}_r)^t]$ and using the fact that, in the exponential case, $n\bar{x}_r/\beta$ has a gamma distribution with scale parameter 1 and shape parameter r (Mann et al. 1974) we take the first two derivatives of $M_{-T}(t)$, set t equal to zero, and thus find the exact moments of T when $\theta = 1$ (see Appendix for details).

For the asymptotic moments of W, as $\theta \to \infty$, we use Taylor's formula for functions of several variables. Using the notations of Apostol (1957) we set

$$f(x_1,\cdots,x_r) = \log(\bar{x}_r/\tilde{x}_r)$$

$$= -\log(n)+\log[x_1+\cdots+x_{r-1}+(n-r+1)x_r]$$

$$-(1/n)[\log x_1+\cdots+\log x_{r-1}+(n-r+1)\log x_r]$$

and let $x = (x_1,\cdots,x_r)$ and $a = (1,\cdots,1)$. Applying Taylor's formula we have

$$f(x)-f(a) = \sum_{k=1}^{2} (1/k!)d^k f(1;x-1) + (1/3!)d^3 f(z;x-1).$$

Evaluating the above derivatives we get

$$f(x) = (1/2)\sum_{i=1}^{r}\sum_{j=1}^{r} c_{ij}(x_i-1)(x_j-1)$$

$$+ (1/6)\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r} D_{ijk}f(z)(x_i-1)(x_j-1)(x_k-1),$$

where

$$c_{ij} = (1-1/n)/n \qquad \text{if } i=j\neq r,$$

$$= -1/n^2 \qquad \text{if } i\neq j,\ i\neq r, \text{ and } j\neq r,$$

$$= (n-r+1)(1-r)/n^2 \qquad \text{if } i=j=r,$$

$$= (r-n-1)/n^2 \qquad \text{otherwise.}$$

Since $\log(\bar{x}_r/\tilde{x}_r) = \log[(\bar{x}_r/\theta)/(\tilde{x}_r/\theta)]$, we can write

$$W = 2n\theta f(x)$$

$$= n \sum_{i=1}^{r} \sum_{j=1}^{r} c_{ij}[\theta^{\frac{1}{2}}(x_i/\theta-1)][\theta^{\frac{1}{2}}(x_j/\theta-1)]$$

$$+ (n/3) \sum_{i=1}^{r} \sum_{j=1}^{r} \sum_{k=1}^{r} D_{ijk} f(z)[\theta^{\frac{1}{2}}(x_i/\theta-1)]$$

$$\cdot [\theta^{\frac{1}{2}}(x_j/\theta-1)][\theta^{\frac{1}{2}}(x_k/\theta-1)]/\theta^{\frac{1}{2}}.$$

It can be easily shown, using the moment generating function, that $\theta^{\frac{1}{2}}(x_i/\theta-1)$ converges in distribution to z_i, the ith ordered standard normal random variable. Consequently, $E(W) \to n\sum_{i=1}^{r}\sum_{j=1}^{r}c_{ij}E(z_iz_j)$ and $\mathrm{Var}(W) \to n^2\sum_{i=1}^{r}\sum_{j=1}^{r}\sum_{k=1}^{r}\sum_{m=1}^{r}c_{ij}c_{km}E(z_iz_jz_kz_m) - [E(W)]^2$. Tables of $E(z_iz_j)$ can be found in Sarhan and Greenberg (1962) and $E(z_iz_jz_kz_m)$ can be found numerically.

Monte-Carlo percentage points of T have been used to study the accuracy of the chi-square approximation for various α levels. The comparison of the approximate and exact levels, for various values of θ, n, and r/n, is given in Table 2. Upper and lower α levels are provided to indicate the usefulness of the chi-square approximation in hypothesis testing and for confidence bounds on θ. The α level at .5 is given to show the accuracy of the approximate median unbiased estimator of θ discussed in Section 3.

3. APPLICATIONS OF THE APPROXIMATION

Since $2n\theta c(\theta,n,r)T$ has an approximate chi-square distribution with $\nu(\theta,n,r)$ degrees of freedom we can construct convenient approximate tests of hypothesis. An approximate unbiased α level test for $H:\theta\leq\theta_0$ versus the alternative $K:\theta>\theta_0$ would reject H if $T < t_\alpha$, where $t_\alpha = \chi^2_\alpha[\nu(\theta_0,n,r)]/2n\theta_0 c(\theta_0,n,r)$ and $P[\chi^2(\nu) < \chi^2_\alpha(\nu)] = \alpha$. Other approximate unbiased tests of hypotheses can be appropriately constructed using the chi-square approximation. For example, an approximate unbiased level α test for $H:\theta=\theta_0$ versus the alternative $K:\theta\neq\theta_0$ can be obtained by following the method outlined by Guenther (1972). This test would fail to reject H if $D_1 \leq 2n\theta_0 c(\theta_0,n,r)T \leq D_2$, where D_1 and D_2 are solutions to $P\{D_1<\chi^2[\nu(\theta_0,n,r)] <D_2\} = 1-\alpha$ and $P\{D_1<\chi^2[\nu(\theta_0,n,r)+2]<D_2\} = 1-\alpha$. Several authors have numerically solved for D_1 and D_2 and have tabulated their results (Pacharas 1961, Tate and Klett 1959, and Lindley, East, and Hamilton 1960). These tests can be used to find approximate $1-\alpha$ level lower (or upper) unbiased confidence bounds and confidence

intervals for θ. An approximate $1-\alpha$ level unbiased lower confidence bound, say θ_L, is found by finding θ_L such that $2n\theta_L c(\theta_L,n,r)T = \chi^2_\alpha[\nu(\theta_L,n,r)]$, where T is observed. To find an approximate $1-\alpha$ level unbiased confidence interval for θ, say $[\theta_L, \theta_U]$, we find θ_L and θ_U such that $2n\theta_L c(\theta_L,n,r)T = D_1$ and $2n\theta_U c(\theta_U,n,r)T = D_2$.

The chi-square approximation also permits a convenient median unbiased point estimate of θ. This is achieved by finding the .5 level approximate lower bound (or upper bound) for θ.

To illustrate the above we use the simulated life-test data of 40 ordered observations from a three-parameter gamma distribution found in Harter and Moore (1965). The scale and shape parameters given are $\beta = 100$ and $\theta = 2$. In addition, they have a location parameter $\eta = 30$. To test the hypothesis $H:\theta=1$ versus $K:\theta>1$ we assume $\eta = 30$ is known and is subtracted from each x_i, but that β is unknown. Assuming that only the first 30 observations are available we have $\bar{x}_{30} = 162.375$, $\tilde{x}_{30} = 135.566$, and $T = \log(162.375/135.566) = .180$. From Table 1, under $\theta = 1$, we obtain $E(T) = .398$ and $40\ \text{Var}(T) = .477$ and consequently $c(1,40,30) = .890$ and $\nu(1,40,30) = 28.350$. For $\alpha = .01$ we have

$t_{.01} = \chi^2_{.01}[28.35]/(80)(.89) = .194$, so H is rejected. To determine the power of the approximate test for a particular $\theta_1 > \theta_0$ we evaluate $P\{\chi^2[\nu(\theta_1,n,r)] < 2n\theta_1 c(\theta_1,n,r)t_\alpha\}$. For the alternative $K:\theta=2$ the power is approximately $P\{\chi^2(27.17)<(80)(2)(.943)(.194)\} = .64$. Then for a .99 lower confidence bound for θ we find θ_L such that $(80)(\theta_L)c(\theta_L,40,30)(.18) = \chi^2_{.01}[\nu(\theta_L,40,30)]$. Iterating on θ_L we find $\theta_L = 1.06$. To find an approximate median unbiased point estimate for θ using the same data we find θ_M such that $(80)(\theta_M)c(\theta_M,40,30)(.18) = \chi^2_{.5}[\nu(\theta_M,40,30)]$. Iterating we find $\theta_M = 1.95$.

Power comparisons for tests of hypothesis based on T versus tests based on the Neyman-Pearson ratio are given in Table 3. The hypothesis $H:\theta=1$ versus $K:\theta=2$ and the hypothesis $H:\theta=2$ versus $K:\theta=1$ were selected since the Neyman-Pearson ratio can be conveniently evaluated in these cases (see Section 4). The tables are rounded to two significant figures, and even though they agree in several instances, we do not mean to imply that the powers are equal. This is clearly not the case since the test based on T is a composite test applied to these hypotheses while the test based on the Neyman-Pearson ratio is the most powerful test for these hypotheses. The values are so close that sometimes, through Monte-Carlo variations, the tabulated power based on T may be slightly larger than the power based on the Neyman-Pearson ratio. Thus, very little power is lost using tests based on the statistic T.

4. MONTE-CARLO SIMULATION

The one-parameter gamma distribution,

$$f(x;\theta) = [\Gamma(\theta)]^{-1}x^{\theta-1}e^{-x},\ x>0,\ \theta>0,$$

was used when simulating results based on T, since T is distributed independently of the scale parameter, β. The details of the Monte-Carlo simulation are as follows:

(i) A random sample from the one-parameter gamma distribution was generated for each of eight values of θ, θ =.5, 1, 1.5, 2, 3, 5, 10, 20. These were obtained by either taking the negative of the natural logarithm of a uniform (0,1) random variable to obtain a random variable which is gamma distributed with θ=1, or taking one-half of the square of a standard normal random variable to obtain a random variable which is gamma distributed with θ=1/2. Using the reproductive property of the gamma distribution these were suitably added to attain a gamma random variable with the desired value of θ (θ a multiple of 1/2). The uniform (0,1) random variable and the standard normal random variable were generated by the packaged routines GGUBF and GGNOR from International Mathematical and Statistical Libraries (1977). All simulation work was done on an IBM 370/58/68 in double precision.

(ii) For each θ value three sample sizes were generated, n = 10, 20, 40.

(iii) For each value of n the first r order statistics were kept for five values of r corresponding to r/n = .3, .5, .7, .9 and 1.0. These r order statistics were then used to evaluate the statistic $T = \log(\bar{x}_r/\tilde{x}_r)$.

(iv) Twenty thousand random samples were generated for each set of θ, n, and r. The empirical cumulative distribution of T for each group of θ, n, and r was compiled using these 20,000 random samples. The expected value of T and the variance of T, given in Table 1, were also found using these random samples.

(v) An Aitken-Lagrange interpolation technique (Hildebrand 1956) was used on the empirical cumulative distribution of T to obtain the power of the tests of hypothesis based on T, which are provided in Table 3. To obtain the power of the hypothesis tests based on the

Neyman-Pearson ratio we consider the joint probability density function of the first r order statistics, using the two-parameter gamma distribution,

$$f(x_1,\cdots,x_r) = (n!)\beta^{-n\theta}(\prod_{i=1}^{r} x_i)^{\theta-1}\exp(-\sum_{i=1}^{r} x_i/\beta)$$

$$\cdot\ [\int_{x_r}^{\infty} x^{\theta-1}\exp(-x/\beta)dx]/[(n-r)!(\Gamma(\theta))^n],$$

$$\theta>0,\ \beta>0,\ x_i>0,\ i=1,\cdots,r.$$

Making the transformation $y_i = x_i/x_r$, $i=1,\cdots,r-1$, and $y_r = x_r/\beta$ we have

$$f(y_1,\cdots,y_r) = (n!)y_r^{r\theta-1}(\prod_{i=1}^{r-1} y_i)^{\theta-1}\exp[-\sum_{i=1}^{r-1} y_i y_r - y_r]$$

$$\cdot\ [\int_{y_r}^{\infty} y^{\theta-1}e^{-y}dy]^{n-r},$$

$$\theta>0,\ y_i>0,\ i=1,\cdots,r.$$

This transformation eliminates the dependency on β. Setting $\theta=1$ and integrating out y_r we have

$$f(y_1,\cdots,y_{r-1}|\theta=1) = (n!)\Gamma(r)/\{(n-r)![\sum_{i=1}^{r-1} y_i+(n-r+1)]^r\}.$$

Doing the same with $\theta=2$, we have

$$f(y_1,\cdots,y_{r-1}|\theta=2) = (n!)(\prod_{i=1}^{r-1} y_i)$$

$$\cdot\{\int_0^\infty z^{2r-1}e^{-z}(1+z/[\sum_{i=1}^{r-1} y_i+(n-r+1)])dz\}$$

$$/\{(n-r)![\sum_{i=1}^{r-1} y_i+(n-r+1)]^{2r}\}.$$

The integration contained in this last joint probability distribution can be expanded by using the binomial expansion and integrated term by term. However, the integrations contained in the joint probability distributions for other values of θ make other hypothesis tests less tractable. The ratio

$$f(y_1,\cdots,y_{r-1}|\theta=1)/f(y_1,\cdots,y_{r-1}|\theta=2) = [\sum_{i=1}^{r-1} y_i+(n-r+1)]^r\Gamma(r)$$

$$/ \{(\prod_{i=1}^{r-1} y_i)\int_0^\infty z^{2r-1}e^{-z}(1+z/[\sum_{i=1}^{r-1} y_i+(n-r+1)])^{n-r}dz\}$$

was calculated for n = 10, 20, 40 and r/n = .3, .5, .7, .9, 1. The empirical cumulative distribution for this Neyman-Pearson ratio, using the 20,000 random samples described above, for the situation when $\theta=1$ and when $\theta=2$

was thus compiled. The Aitken-Lagrange interpolation technique (Hildebrand 1956) was then used to obtain the powers associated with the Neyman-Pearson based hypothesis tests of $H:\theta=1$ versus $K:\theta=2$ and $H:\theta=2$ versus $K:\theta=1$.

(vi) The accuracy of the Monte-Carlo simulations was checked in different ways. Since the full sample case was simulated along with the censored case we compared the simulated values of E(T) and nVar(T) to the exact values provided by Bain and Engelhardt (1975). These full sample comparisons are found in Table 4. For the case $\theta=1$, E(T) and nVar(T) were both simulated and found exactly (see Section 2). The comparisons of these results are in Table 5. Finally, the linearity of the results in Table 1 implies a consistency in the results. For example, if in the case where $\theta=5$ and $r/n=.5$ a linear fit to the values of E(T) is calculated for n = 10, 20, 40 the extrapolated value of E(T) for $n = \infty$ is .0401, which is the same value found using the methods of Chernoff et al. (1967). Such extrapolated values, for E(T) and nVar(T), using other values of θ and r/n are similarly close.

APPENDIX

The following computations were done to find the exact moments of T when $\theta=1$. As pointed out in Section 2 we have $M_{-T}(t) = E[(\tilde{x}_r)^t]/E[(\bar{x}_r)^t]$. To find $E[(\bar{x}_r)^t]$ we have $[\sum_{i=1}^{r} x_i+(n-r)x_r]/\theta$ is distributed as a gamma random variable with scale parameter 1 and shape parameter r (Mann et al. 1974). Since $\theta=1$ it follows that, with $y = n\bar{x}_r$,

$$E[n^t(\bar{x}_r)^t] = E[y^t]$$

$$= \int_0^\infty y^{t+r-1}e^{-y}/\Gamma(r)dy.$$

Thus, $E[(\bar{x}_r)] = \Gamma(t+r)/n^t\Gamma(r)$.

To find $E[(\tilde{x}_r)^t]$ we note that

$$E[(\tilde{x}_r)^t] = E[\exp(t \log \tilde{x}_r)]$$

$$= [n!/(n-r)!]\int_0^\infty \exp[(n-r+1)(-x_r+t(\log x_r)/n)]$$

$$\cdot H_{r-1}(x_r;t)dx_r,$$

where

$$H_{r-1}(x_r;t) = \int_0^{x_r}\cdots\int_0^{x_2}\prod_{i=1}^{r-1} \exp[-x_i+t(\log x_i)/n]\}dx_1\cdots dx_{r-1}.$$

Using the methods of Engelhardt (1975b) we set

$$H(x;t) = \int_0^{x}\exp[-u+t(\log u)/n]du$$

and, taking successive integrations, we have

$$H_{r-1}(x_r;t) = [H(x_r;t)]^{r-1}/(r-1)!.$$

Subsequently

$$E[(\tilde{x}_r)^t] = K\int_0^{\infty}\exp[(n-r+1)(-x_r+t(\log x)/n)[H(x;t)]^{r-1}dx$$

where $K=n!/(n-r)!(r-1)!$.

To find the expected value and variance of T we take the first and second derivatives of $M_{-T}(t)$ with respect to t and then set $t=0$. Using $M_{-T}(t) = E[(\tilde{x}_r)^t]/E[(\bar{x}_r)^t]$ and setting $A = \partial E[(\tilde{x}_r)^t]/\partial t$, $B = \partial E[(\tilde{x}_r)^t]/\partial t$, $C = \partial^2 E[(\tilde{x}_r)^t]/\partial t^2$, $D = \partial^2 E[(\bar{x}_r)^t]/\partial t^2$ we find, when A, B, C, D are evaluated at $t=0$, $E(T) = B-A$ and $Var(T) = C-D+B^2-A^2$.

Taking the above indicated derivatives and evaluating A, B, C, D at $t=0$ we have the following formulas:

$$A = K_1\int_0^1 \log \lambda_u (1-u)^{n-r} u^{r-1} du + K_2\int_0^1 (1-u)^{n-r} u^{r-2} I_1(\lambda_u) du,$$

and

$$C = K_3\int_0^1 \log^2 \lambda_u (1-u)^{n-r} u^{r-1} du$$

$$+ 2K_4\int_0^1 \log \lambda_u (1-u)^{n-r} u^{r-2} I_1(\lambda_u) du$$

$$+ K_5\int_0^1 (1-u)^{n-r} u^{r-3} [I_1(\lambda_u)]^2 du$$

$$+ K_6\int_0^1 (1-u)^{n-r} u^{r-2} I_2(\lambda_u) du,$$

where $\lambda_u = -\log(1-u)$, $K_1 = [n-r+1]nP(1)$, $K_2 = nP(2)$, $K_3 = [n-r+1]^2P(1)$, $K_4 = 2[n-r+1]P(2)$, $K_5 = P(3)$, and $K_6 = P(2)$ such that $P(i) = (n-1)!/n(n-r)!(r-i)!$. The functions

$$I_1(u) = \int_0^u \log s \exp(-s) ds$$

and

$$I_2(u) = \int_0^u \log^2 s \exp(-s) ds$$

can be evaluated as follows: $I_1(\lambda_u) = u \log(\lambda_u) + Q(\lambda_u)$

and $I_2(\lambda_u) = u[\log(\lambda_u)]^2+2\log(\lambda_u)Q(\lambda_u)+2R(\lambda_u)$ where

$$Q(\lambda_u) = \sum_{i=1}^{\infty} (-1)^i \lambda_u{}^i/i \cdot i!,$$

and

$$R(\lambda_u) = \sum_{i=1}^{\infty} (-1)^{i+1} \lambda_u{}^i/i^2 \cdot i!.$$

Then for B and D we have,

$$B = \psi(r)-\log n$$

and

$$D = \psi^{(1)}(r)+[(\log n)-\psi(r)]^2,$$

where $\psi(\cdot)$ and $\psi^{(1)}(\cdot)$ are the digamma and trigamma functions, respectively, and can be evaluated by the following:

$$\psi(n+1) = -\gamma+\sum_{k=1}^{n} k^{-1} \qquad (n\geq 1),$$

and

$$\psi^{(1)}(n+1) = \zeta(2)-\sum_{i=1}^{n} 1/i^2,$$

where γ is Euler's constant and $\zeta(\cdot)$ is the Reimann zeta function. From Abramowitz and Stegun (1970) we find $\gamma \doteq .5772156649$ and $\zeta(2) = \sum_{i=1}^{\infty} i^{-2} = \pi^2/6$. From these formulas for A, B, C, and D the exact expected value and the exact variance for T, when $\theta=1$, can be obtained.

1. Moments for the Approximation

n	θ	θ E(T)					$n\,\theta^2$ Var(T)				
		r/n									
		.3	.5	.7	.9	1.0	.3	.5	.7	.9	1.0
10	.5	.131	.258	.385	.515	.583	.155	.310	.460	.608	.683
	1.0	.105	.214	.330	.456	.527	.121	.247	.377	.516	.594
	1.5	.0929	.195	.308	.434	.503	.104	.216	.338	.477	.551
	2.0	.0856	.184	.296	.422	.490	.0932	.198	.318	.456	.528
	3.0	.0777	.172	.281	.405	.477	.0776	.179	.295	.431	.504
	5.0	.0705	.160	.266	.394	.467	.0658	.155	.265	.405	.483
	10.0	.0640	.151	.257	.384	.458	.0549	.141	.251	.383	.467
	20.0	.0610	.147	.252	.378	.454	.0496	.130	.234	.365	.456
20	.5	.161	.285	.409	.538	.610	.196	.345	.488	.633	.708
	1.0	.130	.239	.354	.480	.552	.150	.274	.403	.542	.619
	1.5	.114	.218	.330	.456	.528	.125	.241	.365	.497	.578
	2.0	.106	.206	.314	.442	.516	.112	.218	.338	.472	.556
	3.0	.0966	.192	.303	.429	.504	.0963	.194	.313	.450	.529
	5.0	.0885	.181	.289	.417	.492	.0815	.173	.288	.425	.508
	10.0	.0808	.170	.278	.407	.483	.0698	.157	.267	.408	.492
	20.0	.0762	.165	.272	.402	.480	.0624	.148	.259	.399	.484

1. Moments for the Approximation, Continued

n	θ	θ E(T)					$n\ \theta^2$ Var(T)				
		r/n									
		.3	.5	.7	.9	1.0	.3	.5	.7	.9	1.0
40	.5	.177	.300	.424	.552	.623	.210	.360	.503	.643	.720
	1.0	.142	.251	.366	.492	.564	.164	.287	.416	.554	.632
	1.5	.126	.228	.341	.467	.540	.141	.254	.378	.513	.590
	2.0	.116	.216	.328	.454	.528	.122	.227	.346	.488	.568
	3.0	.106	.203	.312	.441	.516	.107	.206	.324	.460	.543
	5.0	.0970	.190	.300	.428	.505	.0905	.187	.303	.440	.520
	10.0	.0892	.180	.289	.419	.495	.0763	.165	.278	.417	.503
	20.0	.0848	.174	.284	.414	.492	.0696	.156	.268	.408	.496
∞	.5	.192	.315	.437	.564	.635	.231	.375	.518	.658	.733
	1.0	.154	.263	.378	.504	.577	.178	.301	.429	.567	.645
	1.5	.137	.240	.353	.480	.554	.150	.263	.387	.524	.603
	2.0	.127	.226	.338	.466	.540	.134	.241	.362	.500	.580
	3.0	.116	.213	.324	.453	.528	.115	.217	.335	.473	.555
	5.0	.106	.201	.311	.440	.515	.0988	.195	.310	.450	.533
	10.0	.0976	.190	.300	.430	.508	.0838	.175	.290	.431	.516
	20.0	.0922	.183	.294	.426	.504	.0752	.163	.280	.424	.508

2. Comparison of Exact and Approximate Percentage Points of $W = 2n\theta T$

		r/n = .3							r/n = .7						
n	θ	α							α						
		.01	.05	.10	.50	.90	.95	.99	.01	.05	.10	.50	.90	.95	.99
10	.5	.005	.033	.076	.502	.902	.950	.989	.009	.045	.092	.502	.901	.951	.988
	1.0	.008	.041	.089	.496	.902	.951	.990	.011	.050	.098	.502	.901	.951	.990
	2.0	.008	.046	.095	.491	.896	.951	.990	.011	.048	.096	.492	.900	.952	.990
	5.0	.009	.050	.100	.498	.896	.949	.991	.014	.054	.105	.496	.900	.950	.990
20	.5	.010	.045	.094	.504	.901	.950	.989	.008	.048	.097	.502	.901	.950	.990
	1.0	.012	.049	.097	.503	.902	.951	.990	.010	.047	.096	.502	.901	.950	.989
	2.0	.016	.058	.104	.492	.901	.949	.991	.011	.053	.106	.500	.901	.951	.990
	5.0	.018	.062	.109	.497	.902	.952	.991	.012	.053	.103	.497	.990	.950	.990
40	.5	.009	.047	.095	.508	.902	.948	.989	.010	.046	.099	.506	.901	.949	.990
	1.0	.011	.046	.096	.494	.899	.949	.990	.011	.049	.101	.501	.896	.948	.990
	2.0	.011	.052	.103	.499	.901	.951	.989	.009	.051	.100	.493	.898	.950	.988
	5.0	.013	.056	.105	.502	.900	.951	.991	.011	.054	.105	.498	.899	.950	.990

3. Comparison of Power Based on T and on the Neyman-Pearson Ratio (R)

n	α		H: θ = 1 versus K: θ = 2					H: θ = 2 versus K: θ = 1				
			r/n									
			.3	.5	.7	.9	1.0	.3	.5	.7	.9	1.0
10	.01	T	.026	.046	.071	.11	.13	.13	.20	.26	.32	.35
		R	.026	.050	.073	.11	.13	.13	.20	.26	.32	.35
	.05	T	.12	.19	.27	.35	.40	.27	.38	.45	.52	.55
		R	.12	.19	.27	.35	.40	.27	.38	.45	.52	.55
	.10	T	.23	.35	.44	.54	.57	.37	.48	.56	.63	.66
		R	.23	.34	.44	.53	.57	.37	.48	.56	.63	.66
20	.01	T	.067	.13	.22	.31	.36	.23	.37	.47	.56	.60
		R	.066	.14	.22	.31	.36	.23	.37	.47	.56	.60
	.05	T	.25	.40	.53	.64	.70	.43	.57	.67	.75	.79
		R	.25	.40	.53	.65	.70	.43	.57	.67	.75	.79
	.10	T	.41	.58	.71	.80	.84	.54	.68	.76	.83	.86
		R	.42	.59	.71	.80	.84	.55	.68	.76	.83	.86

4. Full Sample Comparison for Moments of T

θ		E(T)			n Var(T)		
		n					
		10	20	40	10	20	40
.5	M.C.	1.169	1.217	1.245	2.73	2.83	2.87
	Exact	1.166	1.220	1.245	2.72	2.83	2.88
1.0	M.C.	.529	.549	.565	.604	.614	.636
	Exact	.527	.552	.564	.594	.619	.632
1.5	M.C.	.336	.352	.360	.244	.256	.263
	Exact	.335	.352	.360	.245	.257	.262
2.0	M.C.	.247	.257	.264	.134	.138	.141
	Exact	.245	.258	.264	.132	.139	.142
3.0	M.C.	.159	.167	.172	.0563	.0588	.0596
	Exact	.159	.168	.172	.0560	.0588	.0603
5.0	M.C.	.0933	.0981	.101	.0193	.0202	.0209
	Exact	.0934	.0983	.101	.0193	.0203	.0208
10.0	M.C.	.0458	.0483	.0496	.00466	.00491	.00504
	Exact	.0458	.0483	.0495	.00467	.00492	.00503
20.0	M.C.	.0227	.0239	.0246	.00112	.00121	.00124
	Exact	.0227	.0240	.0246	.00114	.00121	.00124

5. Exponential Comparison for Moments of T

n		E(T)					n Var(T)				
		r/n									
		.3	.5	.7	.9	1.0	.3	.5	.7	.9	1.0
10	M.C.	.105	.215	.331	.457	.529	.122	.249	.383	.521	.604
	Exact	.105	.214	.330	.456	.527	.121	.247	.377	.516	.594
20	M.C.	.129	.238	.352	.478	.549	.149	.272	.403	.541	.614
	Exact	.130	.239	.354	.480	.552	.150	.274	.403	.542	.619
40	M.C.	.142	.251	.366	.493	.565	.162	.286	.411	.554	.636
	Exact	.142	.251	.366	.492	.564	.164	.287	.416	.554	.632

REFERENCES

Abramowitz, M., and Stegun, I. A. (1970), Handbook of Mathematics Functions, New York: Dover Publications.

Apostol, Tom M. (1957), Mathematical Analysis, Reading: Addison-Wesley Publishing Co.

Bain, L. J., and Engelhardt, M. E. (1975), "A Two-Moment Chi-Square Approximation for the Statistic Log($\bar{x}/\tilde{x}$)," Journal of the American Statistical Association, 70, 948-950.

Barlow, R. E., and Proschan, F. (1965), Mathematical Theory of Reliability, New York: John Wiley and Sons, Inc.

Basu, D. (1955), "On Statistics Independent of a Complete Sufficient Statistic," Sankya, 15, 377-380.

Bishop, D. J., and Nair, U. S. (1939), "A Note on Certain Methods of Testing for the Homogeneity of a Set of Estimated Variances," Journal of the Royal Statistical Society, Supplement, 6, No. 1, 89-99.

Chernoff, H., Gastwirth, J. L., and Johns, M. V. (1967), "Asymptotic Distribution of Linear Combinations of Functions of Order Statistics with Applications to Estimation," Annals of Mathematical Statistics, 38, 52-72.

Engelhardt, Max (1975a), "Simple Linear Estimation of the Parameters of the Logistic Distribution from a Complete or Censored Sample," Journal of the American Statistical Association, 70, 899-902.

________ (1975b), "On Simple Estimation of the Parameters of the Weibull or Extreme-Value Distribution," Technometrics, 17, 369-374.

Glaser, R. E. (1976), "The Ratio of the Geometric Mean to the Arithmetic Mean for a Random Sample from a Gamma Distribution," Journal of the American Statistical Association, 71, 480-487.

Guenther, William C. (1972), "On the Use of the Incomplete Gamma Table to Obtain Unbiased Tests and Unbiased Confidence Intervals for the Variance of a Normal Distribution," The American Statistician, 26, 31-34.

Gupta, S. S. (1960), "Order Statistics from the Gamma Distribution," Technometrics, 2, 243-262.

________, and Groll, P. A. (1961), "Gamma Distribution in Acceptance Sampling Based on Life Tests," Journal of the American Statistical Association, 56, 942-970.

Harter, H. L., and Moore, A. H. (1965), "Maximum-Likelihood Estimation of the Parameters of Gamma and Weibull Populations from Complete and from Censored Samples," Technometrics, 7, 639-643.

Hildebrand, F. B. (1956), Introduction to Numerical Analysis, New York: McGraw-Hill Book Co.

International Mathematical and Statistical Libraries, Inc. (1977), "IMSL-Library 1," Houston, Texas.

Lindley, D. V., East, D. A., and Hamilton, P. A. (1960), "Tables for Making Inference About the Variance of a Normal Distribution," Biometrika, 47, 433-437.

Linhart, H. (1965), "Approximate Confidence Limits for the Coefficient of Variation of Gamma Distributions," Biometrics, 31, 733-737.

Mann, N. R., Schafer, R. E., and Singpurwalla, N. D. (1974), Methods for Statistical Analysis of Reliability and Life Data, New York: John Wiley and Sons, Inc.

Pachares, J. (1961), "Tables for Unbiased Tests on the Variance of a Normal Population," Annals of Mathematical Statistics, 32, 84-87.

Rao, C. R. (1965), Linear Statistical Inference and Its Applications, New York: John Wiley and Sons, Inc.

Sarhan, A. E., and Greenburg, B. G. (1962), Contributions to Order Statistics, New York: John Wiley and Sons, Inc.

Shorack, G. R. (1972), "The Best Test of Exponentiality Against Gamma Alternatives," Journal of the American Statistical Association, 67, 213-214.

Tate, R. F., and Klett, G. W. (1959), "Optimal Confidence Intervals for the Variance of a Normal Distribution," Journal of the American Statistical Association, 54, 674-682.

Cramér-Rao Lower Bounds for Estimators Based on Censored Data

James Wyckoff and Max Engelhardt

ABSTRACT

The Cramér-Rao lower bounds for the variances of unbiased estimators based on censored data are given. Convenient techniques of evaluation are then derived for these lower bounds. Examples are given to illustrate these techniques. Small-sample comparisons are made between the resulting lower bounds, the variances of the best linear unbiased estimators, and the variances of unbiased estimators which are based on the maximum likelihood estimators.

KEY WORDS: Cramér-Rao lower bound; Censored sampling; Nuisance parameters.

AUTHOR'S FOOTNOTE

*James Wyckoff is a graduate student and Max Engelhardt is Associate Professor both at the University of Missouri-Rolla, Rolla, MO 65401. This paper is a portion of a dissertation presented in partial fulfillment of the requirements for the Ph.D. degree in Mathematics at the University of Missouri-Rolla and was supported in part by the U.S. Army Research Office under Grant DAAG29-78-G-0094.

1. INTRODUCTION

Let $Y_1,\cdots,Y_n$ be a random sample from a one-parameter distribution having density function $f(y;\theta)$, where θ belongs to a subset of the real line. Let $T = t(Y_1,\cdots,Y_n)$ be an unbiased estimator of $\tau(\theta)$. Under certain regularity conditions (see e.g. Wijsman 1973)

$$\mathrm{var}_\theta(T) \geq [\tau'(\theta)]^2/nE_\theta[(\partial \log f(Y;\theta)/\partial\theta)^2] \quad (1.1)$$

where the right-hand side of equation (1.1) is the well-known Cramér-Rao lower bound (CRLB) for the variance of unbiased estimators of $\tau(\theta)$ (Cramér 1951). When a CRLB exists it is divided by the variance of T to measure the efficiency of T. For density functions depending on more than one parameter a generalization of the CRLB with nuisance parameters is available (see e.g. Lehmann 1949 or Wasan 1970), again under the assumption that all n observations of the random sample are available.

Suppose the first r out of n order statistics, $X_1<\cdots<X_r$, are available. This situation, known as type II censoring from above, is commonly encountered in life testing and reliability studies. Under these conditions various authors have measured the

efficiency of unbiased estimators with respect to two basic norms, depending on the sample size n. Such efficiencies have typically been based on comparisons with the variance of the best linear unbiased estimator (BLUE) for small sample sizes ($n \leq 20$) and the variance of the asymptotic maximum-likelihood estimator (MLE) for large sample sizes (see e.g. Gupta 1952, Sarhan and Greenberg 1955 and 1957, Saw 1959, Saleh 1966, Engelhardt and Bain 1974, and Engelhardt 1975a). In this paper we are interested in establishing a convenient numerical method of finding the CRLB, both with and without nuisance parameters, for unbiased estimators of $T(\theta)$ with censored data. This would be more appropriate than the BLUE variance when the estimator is not a linear function of order statistics and preferable to the asymptotic variance of the MLE for small samples.

Harter and Moore (1968a) found, for the type II asymptotic distribution of smallest values having distribution function

$$F(x;\nu,K) = 1 - \exp[-(x/\nu)^{-K}], \quad x < 0, \ \nu < 0, \ K > 0,$$

the CRLB for the variance of unbiased estimators of ν, given K, to be ν^2/rK^2, where only the first r observations are available. Harter and Moore (1969) also found, for the Pareto distribution having distribution function

$$F(x;\lambda,L) = 1 - (x-\lambda)^{-L}, \quad x \geq 1+\lambda, \quad L > 0,$$

the CRLB for the variance of unbiased estimators for L, given λ, to be L^2/r. Govindarajulu (1968) gives lower bounds, in the censored case, for unbiased estimators of the location and scale parameters α and β, when the other is unknown, for the two-parameter distributions having distribution functions of the form $F[(x-\alpha)/\beta]$. These bounds, as presented by Govindarajulu and discussed in Section 3 of this paper, are difficult to numerically evaluate. Also in the censored case, Cramér-Rao type bounds have been derived by Mann (1969) for the mean squared error of regular invariant estimators for parametric functions of location and scale parameters.

In this article we derive CRLB's for unbiased estimators of $T(\theta)$, both with and without nuisance parameters, in Section 3. These bounds are based on a simple set of integrals that are obtained through a moment generating function technique discussed in Section 2. Finally the CRLB of the location and scale parameters, with and without the other being known, for the type I extreme-value distribution of smallest values and the normal distributions are found and tabled for sample sizes n = 10,20,30,40,50,60 and censoring levels r/n = .3,.4,.5,.6,.7,.8,.9 and the full sample case. For sample sizes n = 10,20 these are then compared to the variances of the associated BLUE's and the variances of unbiased estimators which are based on the MLE's. Bounds for the full sample case and asymptotic variances of the MLE's have been included for the sake of comparison.

2. DERIVATIONS

To derive the CRLB for censored data we will need the following two lemmas.

Lemma 1: Let $X_1<\cdots<X_n$ denote the order statistics for a random sample of size n from a distribution whose density function is f(x) and whose cumulative distribution is F(x). When g(x), h(x), and k(x) are measurable functions such that the joint moment generating function of $\sum_{i=1}^{r-1}g(X_i)$, $h(X_r)$, and $\sum_{j=r+1}^{n}k(X_j)$ exists, then

$$M_{\sum_{i=1}^{r-1}g(X_i),h(X_r),\sum_{j=r+1}^{n}k(X_j)}(a,b,c) = C_{1,0}\int_{-\infty}^{\infty}$$

$$\cdot\{\exp[b\ h(x)]f(x)[G(x,g,a)]^{r-1}[H(x,k,c)]^{n-r}\}dx,$$

where $C_{i,j} = n!/[(r-i)!(n-r+j)!]$,

$$G(x,g,a) = \int_{-\infty}^{x} \exp[a\ g(u)]f(u)du,$$

and

$$H(x,k,c) = \int_{x}^{\infty} \exp[c\ k(u)]f(u)du.$$

Proof: Denote $M_{\sum_{i=1}^{r-1} g(X_i), h(X_r), \sum_{j=r+1}^{n} k(X_j)}(a,b,c)$ by $M(a,b,c)$, then

$$M(a,b,c) = E\{\exp[a\sum_{i=1}^{r-1} g(X_i) + b\, h(X_r) + c \sum_{j=r+1}^{n} k(X_j)]\}$$

$$= n!\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{x_{r+1}} \exp[b\ h(x_r)] \prod_{j=r+1}^{n} \{f(x_j)$$

$$\cdot \exp[c\ k(x_j)]\}f(x_r)G_{r-1}(x_r,g,a)dx_r\cdots dx_n,$$

where

$$G_{r-1}(x_r,g,a) = \int_{-\infty}^{x_r}\cdots\int_{-\infty}^{x_2} \prod_{i=1}^{r-1} \{f(x_i)\exp[a\, g(x_i)]\}$$

$$dx_1\cdots dx_{r-1}.$$

Using the methods of Engelhardt (1975b) we integrate successively and have

$$G_{r-1}(x_r,g,a) = [\int_{-\infty}^{x_r} \exp[a\ g(u)]f(u)du]^{r-1}/(r-1)!$$

$$= [G(x_r,g,a)]^{r-1}/(r-1)!.$$

Changing the order of integration it follows that

$$M(a,b,c) = C_{1,0}\int_{-\infty}^{\infty} \exp[b\ h(x_r)]f(x_r)[G(x_r,g,a)]^{r-1}$$

$$\cdot\ H_{r-1}(x_r,k,c)dx_r,$$

where

$$H_{r-1}(x_r,k,c) = \int_{x_r}^{\infty} \cdots \int_{x_{n-1}}^{\infty} \prod_{j=r+1}^{n} f(x_j)\exp[c\ k(x_j)]$$

$$\cdot\ dx_n \cdots dx_{r+1}.$$

Applying the same methods of integration (Engelhardt 1975a) we have

$$H_{r-1}(x_r,k,c) = [\int_{x_r}^{\infty} \exp[c\ k(u)]f(u)du]^{n-r}/(n-r)!$$

$$= [H(x_r,k,c)]^{n-r}/(n-r)!,$$

the result then follows.

Before proceeding to Lemma 2 we note that $G(x,g,0) = F(x)$ and $H(x,k,0) = 1 - F(x)$.

Lemma 2: For the functions $g(x)$, $h(x)$, and $k(x)$, under the assumptions of Lemma 1, we have the following expected values:

(i) $$E[\sum_{i=1}^{r-1} g(X_i)] = C_{2,0}\int_{-\infty}^{\infty} f(x)[F(x)]^{r-2} \cdot [1-F(x)]^{n-r} I_1(g,x)dx,$$

(ii) $$E[(\sum_{i=1}^{r-1} g(X_i))^2] = C_{3,0}\int_{-\infty}^{\infty} f(x)[F(x)]^{r-3} \cdot [1-F(x)]^{n-r}[I_1(g,x)]^2 dx + C_{2,0}\int_{-\infty}^{\infty} f(x)[F(x)]^{r-2} \cdot [1-F(x)]^{n-r} I_2(g,x)dx,$$

(iii) $$E[h(X_r)\sum_{i=1}^{r-1} g(X_i)] = C_{2,0}\int_{-\infty}^{\infty} h(x)f(x)[F(x)]^{r-2} \cdot [1-F(x)]^{n-r} I_1(g,x)dx,$$

where $I_1(g,x) = \int_{-\infty}^{x} g(u)f(u)du$ and $I_2(g,x) = \int_{-\infty}^{x}[g(u)]^2 \cdot f(u)du$.

Proof: Denote $M_{\sum_{i=1}^{r-1} g(X_i), h(X_r), \sum_{j=r+1}^{n} k(X_j)}(a,b,c)$ by $M(a,b,c)$. Then

$$\partial M(a,b,c)/\partial a = C_{1,0}\int_{-\infty}^{\infty}\{\exp[b\ h(x)]f(x)(r-1) \cdot [G(x,g,a)]^{r-2}[H(x,k,c)]^{n-r} \cdot [\int_{-\infty}^{x} g(u)\exp[a\ g(u)]f(u)du]\}dx.$$

Setting a = b = c = 0 result (i) follows.

For (ii) we have

$$\partial^2 M(a,b,c)/\partial a^2 = C_{2,0}\int_{-\infty}^{\infty}\{\exp[b\ h(x)]f(x)(r-2)$$

$$\cdot [G(x,g,a)]^{r-3}[H(x,k,c)]^{n-r}$$

$$\cdot [\int_{-\infty}^{x} g(u)\exp[a\ g(u)]f(u)du]^2\}dx$$

$$+ C_{2,0}\int_{-\infty}^{\infty}\{\exp[b\ h(x)]f(x)[G(x,g,a)]^{r-2}$$

$$\cdot [H(x,k,c)]^{n-r}[\int_{-\infty}^{x}(g(u))^2\exp[a\ g(u)]$$

$$\cdot f(u)du]\}dx.$$

Setting a = b = c = 0 the result follows.

Taking the partial of $\partial M(a,b,c)/\partial a$ with respect to b we have

$$\partial^2 M(a,b,c)/\partial b\partial a = C_{2,0}\int_{-\infty}^{\infty}\{h(x)\exp[b\ h(x)]f(x)$$

$$\cdot [G(x,g,a)]^{r-2}[H(x,k,c)]^{n-r}$$

$$\cdot [\int_{-\infty}^{x} g(u)\exp[a\ g(u)]f(u)du]\}dx.$$

Again setting a = b = c = 0 result (iii) follows. Other expected values, such as $E[\sum_{i=1}^{r-1} g(X_i)\sum_{j=r+1}^{n} h(X_j)]$, can be found similarly.

3. CRAMÉR-RAO LOWER BOUNDS

Let $f(x;\theta)$ be a one-parameter density function and let $X_1<\cdots<X_r$ be the r smallest order statistics of a random sample of size n. Let $T_r = t(X_1,\cdots,X_r)$ be an unbiased estimator of $\mathcal{T}(\theta)$ and set $f(x_1,\cdots,x_r;\theta) = [n!/(n-r)!]\Pi_{i=1}^{r} f(x_i;\theta)[1-F(x_r;\theta)]^{n-r}$, the joint density function of the first r order statistics. We make the following assumptions:

(i) θ lies in an open interval Ω of the real line.

(ii) $\partial f(x;\theta)/\partial\theta$ and $\partial F(x;\theta)/\partial\theta$ exists for all θ in Ω.

(iii) $\int\cdots\int f(x_1,\cdots,x_r;\theta)dx_1\cdots dx_r$ can be differentiated under the integral signs.

(iv) $\int\cdots\int T_r f(x_1,\cdots,x_r;\theta)dx_1\cdots dx_r$ can be differentiated under the integral signs.

(v) $E_\theta\{[\partial \log f(X_1,\cdots,X_r;\theta)/\partial\theta]^2\} > 0$ for every θ in Ω.

These assumptions are similar to those of the complete sample case except for the difference in the joint density function. If T_r, $f(x_1,\cdots,x_r;\theta)$, and θ meet the above regularity conditions then a proof similar to the proof of the full sample case of the CRLB (see e.g. Lehmann 1949 or Wasan 1970) yields the following theorem.

Theorem 1: Under the above assumptions

$$\mathrm{Var}[T_r] \geq [T'(\theta)]^2/E([\partial \log f(X_1,\cdots,X_r;\theta)/\partial\theta]^2). \quad (3.1)$$

The following theorem gives a convenient evaluation of the denominator in equation (3.1).

Theorem 2: $E([\partial \log f(X_1,\cdots,X_r;\theta)/\partial\theta]^2)$ is equal to

$$C_{3,0}\int_{-\infty}^{\infty} f(x;\theta)[F(x;\theta)]^{r-3}[1-F(x;\theta)]^{n-r}[I_1(g,x)]^2dx$$
$$+ C_{2,0}\int_{-\infty}^{\infty} f(x;\theta)[F(x;\theta)]^{r-2}[1-F(x;\theta)]^{n-r}I_2(g,x)dx$$
$$+ 2C_{2,0}\int_{-\infty}^{\infty} f(x;\theta)[F(x;\theta)]^{r-2}[1-F(x;\theta)]^{n-r}h(x)I_1(g,x)dx$$
$$+ C_{1,0}\int_{-\infty}^{\infty} f(x;\theta)[F(x;\theta)]^{r-1}[1-F(x;\theta)]^{n-r}[h(x)]^2dx,$$

where

$$h(x) = \partial[\log f(x;\theta) + (n-r)\log(1-F(x;\theta))]/\partial\theta,$$

and where $I_1(g,x)$ and $I_2(g,x)$ are defined in Lemma 2 if the function g is given by

$$g(u) = \partial[\log f(u;\theta)]/\partial\theta.$$

Proof: Replacing $f(X_1,\cdots,X_r;\theta)$ with the joint density function we have

$$E([\partial \log f(X_1,\ldots,X_r;\theta)/\partial\theta]^2) = E\{[\sum_{i=1}^{r-1} \partial \log f(X_i;\theta)/\partial\theta$$

$$+ \partial \log f(X_r;\theta)/\partial\theta + (n-r)\partial \log(1-F(X_r;\theta))/\partial\theta]^2\}$$

$$= E\,[\sum_{i=1}^{r-1} g(X_i) + h(X_r)]^2\}$$

$$= E[\sum_{i=1}^{r-1} g(X_i)]^2 + 2E[h(X_r)\sum_{i=1}^{r-1} g(X_i)] + E[h(X_r)]^2.$$

Using the density function of the rth order statistic, $f_r(x;\theta) = C_{1,0}f(x;\theta)[F(x;\theta)]^{r-1}[1-F(x;\theta)]^{n-r}$, and Lemma 2 the result follows.

We shall now extend this one-parameter case of the CRLB to the case with nuisance parameters. That is, let $f(x;\theta_1,\cdots,\theta_m)$ be a multi-parameter density function and let $T_r = t(X_1,\cdots,X_r)$ be an unbiased estimator of θ_1, based on the first r order statistics. We want to establish a convenient numerical method of finding the CRLB for the variances of unbiased estimators of θ_1, when the other parameters are unknown. If T_r, $f(x_1,\cdots,x_r;\theta_1,\cdots,\theta_m)$, and θ_i $(i = 1,\cdots,m)$ meet assumptions (i),$\cdots$,(v) for the one-parameter CRLB for censored data, where differentiation is with respect to any $\theta_i (i = 1,\cdots,m)$ and if $\partial \log f(x_1,\cdots,x_r;\theta_1,\cdots,\theta_m)/\partial\theta_i$ $(i = 1,\cdots,m)$ are linearly independent, then a similar proof as in the full sample case of the CRLB with nuisance parameters (see e.g. Lehmann 1949 or

Wasan 1970) yields the following theorem. We note that the above assumptions are essentially those for the full sample case except for the difference in the joint density function.

Theorem 3: If $P = [p_{ij}]$ is the symmetric $m \times m$ matrix where

$$p_{ij} = E\{[\partial \log f(X_1,\cdots,X_r;\theta_1,\cdots,\theta_m)/\partial\theta_i] \cdot [\partial \log f(X_1,\cdots,X_r;\theta_1,\cdots,\theta_m)/\partial\theta_j]\} \quad (3.2)$$

then

$$\mathrm{var}(T_r) \geq \text{cofactor of } p_{11}/|P|.$$

To evaluate equation (3.2) set

$$g_i(x) = \partial \log f(x;\theta_1,\cdots,\theta_m)/\partial\theta_i$$

and

$$h_i(x) = \partial(\log f(x;\theta_1,\cdots,\theta_m) + (n-r)\log[1-F(x,\theta_1,\cdots,\theta_m)])/\partial\theta_i$$

then

$$\partial \log f(x_1,\cdots,x_r;\theta_1,\cdots,\theta_m)/\partial\theta_i = \sum_{j=1}^{r-1} g_i(x_j) + h_i(x_r)$$

and equation (3.2) becomes

$$p_{ij} = E\{[\sum_{m=1}^{r-1} g_i(X_m)+h_i(X_r)][\sum_{m=1}^{r-1} g_j(X_n)+h_j(X_r)]\}$$

$$= E_1 + E_2 + E_3 + E_4,$$

where

$$E_1 = E\{[\sum_{m=1}^{r-1} g_i(X_m)][\sum_{n=1}^{r-1} g_j(X_n)]\},$$

$$E_2 = E\{h_j(X_r)[\sum_{m=1}^{r-1} g_i(X_m)]\},$$

$$E_3 = E\{h_i(X_r)[\sum_{n=1}^{r-1} g_j(X_n)]\},$$

$$E_4 = E\{h_i(X_r)h_j(X_r)\}.$$

It should be noted that if $i = j$ then p_{ij} is the denominator of equation (2.1), which is evaluated by Theorem 2.

Theorem 4: Element p_{ij} of matrix P in Theorem 3 is equal to

$$C_{2,0}\int_{-\infty}^{\infty} h_j(x)f(x)[F(x)]^{r-2}[1-F(x)]^{n-r}I_1(g_i,x)dx$$

$$+ C_{2,0}\int_{-\infty}^{\infty} h_i(x)f(x)[F(x)]^{r-2}[1-F(x)]^{n-r}I_1(g_j,x)dx$$

$$+ C_{1,0}\int_{-\infty}^{\infty} h_i(x)h_j(x)f(x)[F(x)]^{r-1}[1-F(x)]^{n-r}dx$$

$$+ C_{2,1}\int_{-\infty}^{\infty} g_i(x)g_j(x)f(x)[F(x)]^{r-2}[1-F(x)]^{n-r+1}dx$$

$$+ C_{3,1}\int_{-\infty}^{\infty} g_i(x)f(x)[F(x)]^{r-3}[1-F(x)]^{n-r+1}I_1(g_j,x)dx$$

$$+ C_{3,1}\int_{-\infty}^{\infty} g_j(x)f(x)[F(x)]^{r-3}[1-F(x)]^{n-r+1}I_1(g_i,x)dx$$

$$+ C_{4,1}\int_{-\infty}^{\infty} f(x)[F(x)]^{r-4}[1-F(x)]^{n-r+1}I_1(g_i,x)I_1(g_j,x)dx$$

$$+ C_{3,1}\int_{-\infty}^{\infty} f(x)[F(x)]^{r-3}[1-F(x)]^{n-r+1}[\int_{-\infty}^{x} g_i(y)g_j(y)f(y)dy]dx.$$

Proof: To evaluate E_1 we denote the moment generating function $M_{\sum_{m=1}^{r-1}g_i(X_m),\sum_{n=1}^{r-1}g_j(X_n)}(a,b)$ by $M(a,b)$ and using a technique similar to the one used in Lemmas 1 and 2 we have

$$M(a,b) = [n!/(n-r+1)!]\int_{-\infty}^{\infty} \exp[a\ g_i(x_{r-1})$$
$$+ b\ g_j(x_{r-1})]f(x_{r-1})[1-F(x_{r-1})]^{n-r+1}$$
$$\cdot K_{r-2}(x_{r-1},g_i,g_j,a,b)dx_{r-1},$$

where

$$K_{r-2}(x_{r-1},g_i,g_j,a,b) = \int_{-\infty}^{x_{r-1}}\cdots\int_{-\infty}^{x_2} \prod_{k=1}^{r-2}\{f(x_k)\exp$$
$$[a\,g_i(x_k)+b\,g_j(x_k)]\}dx_1\cdots dx_{r-2}.$$

Using the methods of Engelhardt (1975b) we integrate successively and have

$$K_{r-2}(x_{r-1},g_i,g_j,a,b) = [K(x_{r-1},g_i,g_j,a,b]^{r-2}/(r-2)!,$$

where

$$K(x_{r-1},g_i,g_j,a,b) = \int_{-\infty}^{x_{r-1}} f(y)\exp[a\,g_i(y)+b\,g_j(y)]dy.$$

Thus,

$$M(a,b) = C_{2,1}\int_{-\infty}^{\infty} \exp[a\,g_i(x)+b\,g_j(x)]$$

$$\cdot\ f(x)[1-F(x)]^{n-r+1}[K(x,g_i,g_j,a,b)]^{r-2}dx.$$

Taking the partial with respect to a we have

$$M(a,b)/\partial a = C_{2,1}\int_{-\infty}^{\infty} g_i(x)\exp[a\,g_i(x)+b\,g_j(x)]$$

$$\cdot\ f(x)[1-F(x)]^{n-r+1}[K(x,g_i,g_j,a,b)]^{r-2}dx$$

$$+\ C_{2,1}\int_{-\infty}^{\infty} \exp[a\,g_i(x)+b\,g_j(x)]f(x)$$

$$\cdot\ [1-F(x)]^{n-r+1}(r-2)[K(x,g_i,g_j,a,b)]^{r-3}$$

$$\cdot\ [K_i^{*}(x,g_i,g_j,a,b)]dx,$$

where

$$K_p^{*}(x,g_i,g_j,a,b) = \int_{-\infty}^{x} g_p(y)\exp[a\,g_i(y)+b\,g_j(y)]f(y)dy.$$

Taking the second partial with respect to b we have

$$\partial^2 M(a,b)/\partial b\partial a = C_{2,1}\int_{-\infty}^{\infty} g_i(x)g_j(x)\exp[a\,g_i(x)+b\,g_j(x)]$$

$$\cdot\ f(x)[1-F(x)]^{n-r+1}[K(x,g_i,g_j,a,b)]^{r-2}dx$$

$$+\ C_{2,1}\int_{-\infty}^{\infty} g_i(x)\exp[a\, g_i(x)+b\, g_j(x)]f(x)$$

$$\cdot\ [1-F(x)]^{n-r+1}(r-2)[K(x,g_i,g_j,a,b)]^{r-3}$$

$$\cdot\ [K_j^{*}(x,g_i,g_j,a,b)]dx$$

$$+\ C_{3,1}\int_{-\infty}^{\infty} g_j(x)\exp[a\, g_i(x)+b\, g_j(x)]f(x)$$

$$\cdot\ [1-F(x)]^{n-r+1}[K(x,g_i,g_j,a,b)]^{r-3}$$

$$\cdot\ [K_i^{*}(x,g_i,g_j,a,b)]dx$$

$$+\ C_{3,1}\int_{-\infty}^{\infty} \exp[a\, g_i(x)+b\, g_j(x)]f(x)[1-F(x)]^{n-r+1}$$

$$\cdot\ (r-3)[K(x,g_i,g_j,a,b)]^{r-4}$$

$$\cdot\ [K_j^{*}(x,g_i,g_j,a,b)][K_i^{*}(x,g_i,g_j,a,b)]dx$$

$$+\ C_{3,1}\int_{-\infty}^{\infty} \exp[a\, g_i(x)+b\, g_j(x)]f(x)]1-F(x)]^{n-r+1}$$

$$\cdot\ [K(x,g_i,g_j,a,b)]^{r-3}$$

$$\cdot\ \{\int_{-\infty}^{x} g_i(y)g_j(y)\exp[a\, g_i(y)+b\, g_j(y)]$$

$$\cdot\ f(y)dy\}\,dx.$$

Setting a = b = 0 the evaluation of E_1 follows.

The evaluation of E_2 and E_3 follows directly from Lemma 2 and E_4 is seen to be the expectation of a function of the rth order statistic.

The above theorems, as illustrated in Section 4, give a convenient numerical method of evaluating the CRLB, with and without nuisance parameters. If we restrict ourselves to the two-parameter cumulative distribution function of the form $F[(x-\mu)/\sigma]$, that is, μ and σ are location and scale parameters, respectively, then under similar regularity conditions Govindarajulu (1968) has developed CRLB's for unbiased estimators of μ and σ, with or without the other parameter known. However, expected values similar to those of equation (3.2) of this paper are needed in Govindarajulu's lower bound but a convenient evaluation scheme is not provided. The above theorems of this article extend the results of Govindarajulu and provide a numerical scheme of evaluation.

4. EXAMPLES

To illustrate Theorems 2 and 4 we shall evaluate the CRLB, with and without nuisance parameters, for unbiased estimators of the parameters in the normal distribution and the type I extreme-value distribution of smallest values.

Let the density function of the normal distribution be denoted by

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

$$-\infty < x < \infty,\ -\infty < \mu < \infty,\ \sigma > 0.$$

To evaluate the CRLB for unbiased estimators of σ with μ known, denoted by $\sigma_c|\mu$, we have, from Theorem 2,

$$I_1(g,x) = \int_{-\infty}^{x} \partial f(y;\mu,\sigma)/\partial\sigma dy$$

$$= -(x-\mu)f(x;\mu,\sigma)/\sigma,$$

$$I_2(g,x) = \int_{-\infty}^{x} [\partial f(y;\mu,\sigma)/\partial\sigma]^2 f(y;\mu,\sigma)dy$$

$$= 2F(x;\mu,\sigma)/\sigma^2 - (x-\mu)^3 f(x;\mu,\sigma)/\sigma^4$$

$$- (x-\mu)f(x;\mu,\sigma)/\sigma^2,$$

and

$$h(x) = -1/\sigma + (x-\mu)^2/\sigma^3 + (n-r)(x-\mu)f(x;\mu,\sigma)$$

$$/ [1-F(x;\mu,\sigma)]\sigma.$$

The integrals in Theorem 2 were numerically evaluated and applied to Theorem 1, with the results being shown in Table 6. All numerical integrations in this paper were done on an IBM 370/58/68 in double precision using the package routine DCADRE from International Mathematical and Statistical Libraries (1977), DCADRE uses a cautious adaptive Romberg extrapolation.

To evaluate the CRLB for unbiased estimators of μ with σ known, denoted by $\mu_c|\sigma$, we have

$$I_1(g,x) = \int_{-\infty}^{x} \partial f(y;\mu,\sigma)/\partial\mu dy$$

$$= -f(x;\mu,\sigma),$$

$$I_2(g,x) = \int_{-\infty}^{x} [\partial f(y;\mu,\sigma)/\partial\mu]^2/f(y;\mu,\sigma)dy$$

$$= F(x;\mu,\sigma)/\sigma^2 - (x-\mu)f(x;\mu,\sigma)/\sigma^2,$$

and

$$h(x) = (x-\mu)/\sigma^2 + (n-r)f(x;\mu,\sigma)/[1-F(x;\mu,\sigma)].$$

After the numerical integrations of Theorem 2 were evaluated, Theorem 1 was applied and the resulting CRLB's are shown in Table 7.

For the CRLB for unbiased estimators of μ with σ unknown, denoted by μ_c, or for unbiased estimators of σ with μ unknown, denoted by σ_c, we only need the off diagonal element of the matrix P of Theorem 3. The diagonal elements of P are just the denominators of the CRLB without nuisance parameters. To evaluate the element p_{12} of P we note that in Theorem 4 $I_1(g_1,x)$, $I_1(g_2,x)$, $h_1(x)$, and $h_2(x)$ are those in the CRLB without nuisance parameters and that $g_1(x)$ and $g_2(x)$ are direct evaluations of partial derivatives. Then with

$$\int_{-\infty}^{x} g_1(y)g_2(y)f(y;\mu,\sigma)dy = -(x-\mu)^2 f(x;\mu,\sigma)/\sigma^3$$
$$- f(x;\mu,\sigma)/\sigma$$

the integrals of Theorem 4 are numerically evaluated and applied to Theorem 3. The CRLB's for unbiased estimators of μ and σ, with the other unknown, are found in Tables 8 and 9, respectively. We also note that the CRLB's for unbiased estimators of σ^2, with or without knowing μ, can be found by multiplying those of unbiased estimators of σ by $4\sigma^2$. The full sample case in Tables 6, 7, 8, and 9 are well-known results and the asymptotic values are the variances of the asymptotic MLE's and are tabled by Harter and Moore (1966).

For the next example we consider the type I extreme-value distribution of smallest values, having distribution function

$$F(x;\beta,\theta) = 1 - \exp\{-\exp[(x-\beta)/\theta]\},$$
$$-\infty < x < \infty, \ -\infty < \beta < \infty, \ \theta > 0.$$

Applying Theorem 2 to derive the CRLB for unbiased estimators of θ with β known, denoted by $\theta_c|\beta$, we have

$$I_1(g,x) = \int_{-\infty}^{x} \partial f(y;\beta,\theta)/\partial\theta dy$$

$$= -(x-\beta)f(x;\beta,\theta)/\theta,$$

$$I_2(g,x) = \int_{-\infty}^{x} [\partial f(y;\beta,\theta)/\partial\theta]^2/f(y;\beta,\theta)dy$$

$$= F(x;\beta,\theta)/\theta^2 + (x-\beta)^2 f(x;\beta,\theta)/\theta^3$$

$$- (x-\beta)^2 \exp[(x-\beta)/\theta]f(x;\beta,\theta)/\theta^3$$

$$+ J_1(x),$$

and

$$h(x) = -1/\theta - (x-\beta)/\theta^2 + (n-r+1)(x-\beta)\exp[(x-\beta)/\theta]/\theta^2,$$

where

$$J_1(x) = \theta^{-2}\int_0^{\exp[(x-\beta)/\theta]} u[\log u]^2 \exp[-u]du.$$

Letting $w = 1 - \exp\{-\exp[(x-\beta)/\theta]\}$ a change of variable was made and the resulting integrals of Theorem 2 were numerically integrated and applied to Theorem 1, the results are in Table 10.

To evaluate the CRLB for unbiased estimators of β with θ known, denoted by $\beta_c|\theta$, we have

$$I_1(g,x) = \int_{-\infty}^{x} f(y;\beta,\theta)/\partial\beta dy$$

$$= -f(x;\beta,\theta),$$

$$I_2(g,x) = \int_{-\infty}^{x} [\partial f(y;\beta,\theta)/\partial\beta]^2/f(y;\beta,\theta)dy$$

$$= F(x;\beta,\theta)/\theta^2 - \exp[(x-\beta)/\theta]f(x;\beta,\theta)/\theta,$$

and

$$h(x) = -1/\theta + (n-r+1)\exp[(x-\beta)/\theta]/\theta.$$

In this situation the integrals of Theorem 2 can be integrated by parts and, after like terms are collected, only the integral

$$(r/\theta^2)C_{1,0}\int_{-\infty}^{\infty} f(x;\beta,\theta)[F(x;\beta,\theta)]^{r-1}[1-F(x;\beta,\theta)]^{n-r}dx$$

$$= (r/\theta^2)C_{1,0}[(r-1)!(n-r)!/n!]$$

remains. Thus the CRLB for unbiased estimators of β, when θ is known, is θ^2/r.

To find the CRLB for unbiased estimators of β with θ unknown, denoted by β_c, or for unbiased estimators of θ with β unknown, denoted by θ_c, we need the off diagonal element of matrix P from Theorem 3. For the element p_{12} of P we note that $I_1(g,x), I_1(g_2,x)$, $h_1(x)$, and $h_2(x)$ of Theorem 4 are those of the CRLB without nuisance parameters and that $g_1(x)$ and $g_2(x)$ are direct evaluations of partial derivatives. With

$$\int_{-\infty}^{x} g_1(y)g_2(y)f(y;\beta,\theta)dy = F(x;\beta,\theta)/\theta^2 - (x-\beta) \cdot \exp[(x-\beta)/\theta]f(x;\beta,\theta)/\theta^2 + J_2(x),$$

where

$$J_2(x) = \theta^{-2}\int_0^{\exp[(x-\beta)/\theta]} [\log u]\exp[-u]du,$$

the integrals of Theorem 4 are numerically evaluated and applied to Theorem 3. The CRLB's for unbiased estimators of β and θ, with the other unknown, are found in Tables 11 and 12, respectively. The asymptotic values in Tables 10, 11, and 12 are the variances of the asymptotic MLE's and are tabled by Harter and Moore (1968b). The full sample cases in Tables 10, 11, and 12 are derived by Chan and Kabir (1969) and are evaluated by

$$\theta_c|\beta = \theta^2/n[(\pi^2/6) + (1-\gamma)^2],$$

$$\theta_c = \theta^2/n[\pi^2/6],$$

and

$$\beta_c = \theta^2[1+6(1-\gamma)^2/\pi^2]/n,$$

where γ is Euler's constant.

Tables 13 and 14 compare the CRLB's for unbiased estimators of the parameters in the normal distribution and the type I extreme-value distribution of smallest values to the variances of the associated BLUE's and the variances of unbiased estimators which are based on the MLE's. The headings of Tables 13 and 14 are defined by:

s_c = CRLB for unbiased estimators of s with t unknown,

$\tilde{s}_v$ = variance of the unbiased estimator of s based on the MLE with t unknown,

s^*_v = variance of the BLUE of s with t unknown,

$s_c|t$ = CRLB for unbiased estimators of s with t known,

$\tilde{s}_v|t$ = variance of the unbiased estimator of s based on the MLE with t known,

where s and t are suitable replaced by the parameters μ, σ, β, and θ.

The unbiased estimator of μ, with σ unknown, based on the MLE is denoted by $\tilde{\mu}$ and is found by setting

$$\tilde{\mu} = \hat{\mu} - \hat{\sigma}E[(\hat{\mu}-\mu)/\sigma]/E[\hat{\sigma}/\sigma]$$

and the unbiased estimator of σ, with μ unknown, based on the MLE is denoted by $\tilde{\sigma}$ and is found by setting

$$\tilde{\sigma} = \hat{\sigma}/E[\hat{\sigma}/\sigma],$$

where $\hat{\mu}$ and $\hat{\sigma}$ denote the MLE's when the other parameter is unknown. Likewise, the unbiased estimators of μ and σ, when the other parameter is known, are denoted by $\tilde{\mu}|\sigma$ and $\tilde{\sigma}|\mu$, respectively, and are found by setting

$$\tilde{\mu}|\sigma = (\hat{\mu}|\sigma) - E[(\hat{\mu}|\sigma)-\mu]$$

and

$$\tilde{\sigma}|\mu = (\hat{\sigma}|\mu)/E[(\hat{\sigma}|\mu)/\sigma],$$

where $\hat{\mu}|\sigma$ and $\hat{\sigma}|\mu$ denote the MLE's when the other parameter is known.

In a similar manner unbiased estimators of β and θ, with and without the other parameter known, based on the MLE's are found by setting

$$\tilde{\beta} = \hat{\beta} - \hat{\theta}E[(\hat{\beta}-\beta)/\theta]/E[\hat{\theta}/\theta],$$

$$\tilde{\beta}|\theta = (\hat{\beta}|\theta) - E[(\hat{\beta}|\theta)-\beta],$$

$$\tilde{\theta} = \hat{\theta}/E[\hat{\theta}/\theta],$$

and

$$\tilde{\theta}|\beta = (\hat{\theta}|\beta)/E[(\hat{\theta}|\beta)/\theta],$$

where $\hat{\beta}$, $\hat{\beta}|\theta$, $\hat{\theta}$, and $\hat{\theta}|\beta$ denote the appropriate MLE's.

Harter and Moore (1966 and 1968b) obtained Monte-Carlo results for the expected values and variances of $\hat{\mu}$, $\hat{\mu}|\sigma$, $\hat{\sigma}$, $\hat{\sigma}|\mu$, $\hat{\beta}$, $\hat{\beta}|\theta$, $\hat{\theta}$, and $\hat{\theta}|\beta$, and the covariances of $\hat{\mu}$ with $\hat{\sigma}$ and $\hat{\beta}$ with $\hat{\theta}$. These results were for sample sizes n = 10,20 and censoring levels r/n = .2,.3,.4,.5,.6,.7,.8,.9 and the full sample case. Harter and Moore also tabled the exact variances of the BLUE's that are in Tables 13 and 14. The Monte-Carlo results were used to find the variances of the unbiased estimators based on the MLE's. Some Monte-Carlo error is apparent since the variance of the unbiased estimators based on the MLE is less than the CRLB for some censored and full sample cases.

In the above examples the joint moment generating function of $\sum_{i=1}^{r-1} g(X_i)$, $h(X_r)$, and $\sum_{j=r+1}^{n} k(X_j)$ exists. For examples where this joint moment generating function does not exist but where $E\{[\sum_{i=1}^{r-1} g(X_i)]^2\}$, $E\{[h(X_r)]^2\}$, and $E\{[\sum_{j=r+1}^{n} k(X_j)]^2\}$ are all finite, then a characteristic function approach could be used to derive the indicated CRLB's. This approach is quite similar to that used with the joint moment generating function.

6. Coefficients of σ^2/n for $\sigma_c|\mu$ (Normal Distribution)

r/n	n						
	10	20	30	40	50	60	∞
.3	1.1053	1.1174	1.1220	1.1259	1.1278	1.1292	1.1364
.4	1.0413	1.0614	1.0692	1.0733	1.0759	1.0777	1.0868
.5	.9575	.9770	.9842	.9880	.9903	.9919	1.0000
.6	.8604	.8756	.8812	.8841	.8858	.8870	.8930
.7	.7603	.7707	.7744	.7763	.7775	.7783	.7823
.8	.6649	.6711	.6733	.6744	.6751	.6756	.6779
.9	.5780	.5809	.5820	.5825	.5829	.5831	.5843
1.0	.5000	.5000	.5000	.5000	.5000	.5000	.5000

7. Coefficients of σ^2/n for $\mu_c|\sigma$ (Normal Distribution)

r/n	n						
	10	20	30	40	50	60	∞
.3	1.5974	1.5621	1.5503	1.5444	1.5408	1.5385	1.5266
.4	1.3816	1.3605	1.3534	1.3499	1.3478	1.3464	1.3393
.5	1.2492	1.2356	1.2311	1.2288	1.2275	1.2266	1.2220
.6	1.1605	1.1515	1.1485	1.1470	1.1461	1.1455	1.1425
.7	1.0981	1.0922	1.0902	1.0892	1.0886	1.0882	1.0862
.8	1.0532	1.0495	1.0483	1.0476	1.0472	1.0470	1.0457
.9	1.0212	1.0193	1.0186	1.0182	1.0180	1.0179	1.0172
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

8. Coefficients of σ^2/n for μ_c (Normal Distribution)

r/n	n						
	10	20	30	40	50	60	∞
.3	2.8625	2.9308	2.9579	2.9724	2.9814	2.9875	3.0199
.4	1.9671	1.9774	1.9815	1.9837	1.9851	1.9860	1.9909
.5	1.5252	1.5211	1.5197	1.5191	1.5187	1.5184	1.5171
.6	1.2861	1.2795	1.2773	1.2761	1.2754	1.2750	1.2727
.7	1.1498	1.1441	1.1422	1.1412	1.1406	1.1402	1.1383
.8	1.0702	1.0664	1.0650	1.0644	1.0640	1.0637	1.0623
.9	1.0243	1.0223	1.0216	1.0212	1.0210	1.0208	1.0201
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

9. Coefficients of σ^2/n for σ_c (Normal Distribution)

r/n	n						
	10	20	30	40	50	60	∞
.3	1.9807	2.0964	2.1424	2.1670	2.1823	2.1927	2.2480
.4	1.4825	1.5427	1.5654	1.5773	1.5847	1.5896	1.6155
.5	1.1691	1.2027	1.2150	1.2214	1.2253	1.2279	1.2415
.6	.9536	.9730	.9800	.9836	.9858	.9872	.9948
.7	.7961	.8073	.8114	.8134	.8146	.8155	.8197
.8	.6756	.6818	.6841	.6852	.6859	.6863	.6887
.9	.5797	.5826	.5837	.5842	.5846	.5848	.5859
1.0	.5000	.5000	.5000	.5000	.5000	.5000	.5000

10. Coefficients of θ^2/n for $\theta_c|\beta$ (Type I Extreme-Value Distribution)

r/n	n						
	10	20	30	40	50	60	∞
.3	1.4336	1.4258	1.4238	1.4230	1.4225	1.4222	1.4211
.4	1.3736	1.3914	1.3983	1.4021	1.4044	1.4059	1.4141
.5	1.2961	1.3284	1.3406	1.3470	1.3509	1.3536	1.3673
.6	1.1825	1.2191	1.2328	1.2399	1.2443	1.2473	1.2627
.7	1.0367	1.0692	1.0814	1.0877	1.0916	1.0943	1.1080
.8	.8737	.8976	.9065	.9112	.9141	.9161	.9263
.9	.7089	.7226	.7280	.7308	.7326	.7337	.7400
1.0	.5483	.5483	.5483	.5483	.5483	.5483	.5483

11. Coefficients of θ^2/n for β_C (Type I Extreme-Value Distribution)

r/n	n						
	10	20	30	40	50	60	∞
.3	5.7712	6.3515	6.5951	6.7291	6.8139	6.8723	7.1904
.4	3.4588	3.5735	3.7433	3.7871	3.8145	3.8332	3.9330
.5	2.3459	2.4180	2.4462	2.4612	2.4705	2.4768	2.5102
.6	1.7602	1.7831	1.7920	1.7967	1.7996	1.8016	1.8120
.7	1.4370	1.4414	1.4431	1.4441	1.4447	1.4451	1.4473
.8	1.2559	1.2542	1.2536	1.2534	1.2532	1.2531	1.2526
.9	1.1567	1.1544	1.1535	1.1531	1.1528	1.1526	1.1517
1.0	1.1087	1.1087	1.1087	1.1087	1.1087	1.1087	1.1087

12. Coefficients of θ^2/n for θ_c (Type I Extreme-Value Distribution)

r/n	n						
	10	20	30	40	50	60	∞
.3	2.4821	2.7167	2.8170	2.8726	2.9078	2.9322	3.0655
.4	1.9004	2.0380	2.0937	2.1239	2.1427	2.1557	2.2247
.5	1.5203	1.6061	1.6397	1.6576	1.6687	1.6763	1.7162
.6	1.2489	1.3043	1.3255	1.3367	1.3436	1.3483	1.3728
.7	1.0428	1.0788	1.0924	1.0995	1.1039	1.1069	1.1224
.8	.8778	.9005	.9092	.9137	.9165	.9184	.9282
.9	.7380	.7508	.7557	.7584	.7600	.7612	.7670
1.0	.6079	.6079	.6079	.6079	.6079	.6079	.6079

13. Precision Comparisons of MLE and BLUE Estimators

to the CRLB (Normal Distribution with $\sigma = 1$)

n	r/n	μ_c	$\tilde{\mu}_v$	μ^*_v	$\mu_c\|\sigma$	$\tilde{\mu}_v\|\sigma$	σ_c	$\tilde{\sigma}_v$	σ^*_v	$\sigma_c\|\mu$	$\tilde{\sigma}_v\|\mu$
10	.3	.286	.420	.417	.160	.160	.198	.356	.354	.111	.121
	.4	.197	.229	.237	.138	.138	.148	.213	.225	.104	.114
	.5	.153	.161	.166	.125	.124	.117	.154	.161	.096	.101
	.6	.129	.130	.134	.116	.116	.095	.116	.124	.086	.088
	.7	.115	.116	.117	.110	.109	.080	.101	.099	.076	.082
	.8	.107	.104	.107	.105	.103	.068	.083	.081	.066	.071
	.9	.102	.101	.102	.102	.101	.058	.067	.068	.058	.059
	1.0	.100	.100	.100	.100	.100	.050	.057	.058	.050	.050
20	.3	.147	.170	.175	.078	.074	.105	.138	.139	.056	.060
	.4	.099	.103	.108	.068	.064	.077	.095	.094	.053	.055
	.5	.076	.076	.079	.062	.059	.060	.069	.070	.049	.049
	.6	.064	.062	.065	.058	.055	.049	.052	.055	.044	.043
	.7	.057	.054	.058	.055	.052	.040	.043	.045	.039	.037
	.8	.053	.050	.053	.052	.049	.034	.035	.037	.034	.032
	.9	.051	.048	.051	.051	.048	.029	.029	.032	.029	.028
	1.0	.050	.048	.050	.050	.048	.025	.024	.027	.025	.022

14. Precision Comparisons of MLE and BLUE Estimators to the CRLB (Type I Extreme-Value Distribution with $\theta = 1$)

n	r/n	β_c	$\tilde{\beta}_v$	β^*_v	$\tilde{\beta}_c \mid \theta$	$\tilde{\beta}_v \mid \theta$	θ_c	$\tilde{\theta}_v$	θ^*_v	$\theta_c \mid \beta$	$\tilde{\theta}_v \mid \beta$
10	.3	.577	1.238	1.204	.333	.392	.248	.456	.461	.143	.143
	.4	.346	.570	.559	.250	.269	.190	.302	.298	.137	.136
	.5	.235	.322	.321	.200	.217	.152	.208	.215	.130	.124
	.6	.176	.221	.214	.167	.182	.125	.163	.166	.118	.114
	.7	.144	.160	.162	.143	.151	.104	.127	.132	.104	.099
	.8	.126	.133	.134	.125	.131	.088	.104	.107	.087	.085
	.9	.116	.118	.120	.111	.116	.074	.083	.088	.071	.068
	1.0	.111	.110	.113	.100	.104	.061	.067	.072	.055	.055
20	.3	.318	.465	.456	.167	.186	.136	.193	.184	.071	.076
	.4	.179	.236	.232	.125	.136	.102	.133	.127	.070	.074
	.5	.121	.143	.141	.100	.107	.080	.098	.096	.066	.070
	.6	.089	.098	.098	.083	.087	.065	.080	.075	.061	.063
	.7	.072	.076	.076	.071	.074	.054	.063	.061	.053	.056
	.8	.063	.065	.065	.063	.065	.045	.052	.050	.045	.048
	.9	.058	.059	.059	.056	.058	.038	.043	.041	.036	.039
	1.0	.055	.055	.056	.050	.050	.030	.035	.033	.027	.030

REFERENCES

Chan, L. K., and Kabir, A. B. M. L. (1969), "Optimum Quantiles for the Linear Estimation of the Parameters of the Extreme Value Distribution in Complete and Censored Samples," Naval Research Logistics Quarterly, 16, 381-404.

Cramér, H. (1951), Mathematical Methods of Statistics, Princeton: Princeton University Press.

Engelhardt, Max (1975a), "Simple Linear Estimation of the Parameters of the Logistic Distribution from a Complete or Censored Sample," Journal of the American Statistical Association, 70, 899-902.

________ (1975b), "On Simple Estimation of the Parameters of the Weibull or Extreme-Value Distribution," Technometrics, 17, 369-374.

________, and Bain, Lee J. (1974), "Some Results on Point Estimation for the Two-Parameter Weibull or Extreme-Value Distribution," Technometrics, 16, 49-56.

Govindarajulu, Zakkula (1968), "Certain General Properties of Unbiased Estimates of Location and Scale Parameters Based on Ordered Observations," SIAM Journal on Applied Mathematics, 16, 533-551.

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

Harter, H. Leon, and Moore, Albert H. (1968a), "Conditional Maximum-Likelihood Estimators, from Singly Censored Samples, of the Scale Parameters of Type II Extreme-Value Distributions," Technometrics, 10, 349-359.

________ (1968b), "Maximum-Likelihood Estimation, from Doubly Censored Samples, of the Parameters of the First Asymptotic Distribution of Extreme Values," Journal of the American Statistical Association, 63, 889-901.

________ (1966), "Iterative Maximum-Likelihood Estimation of the Parameters of Normal Populations from Singly and Doubly Censored Samples," Biometrika, 53, 205-213.

________ (1969), "Conditional Maximum-Likelihood Estimation, from Singly Censored Samples, of the Shape Parameters of Pareto and Limited Distributions," IEEE Transactions on Reliability, R-18, 76-78.

International Mathematical and Statistical Libraries, Inc. (1977), "IMSL-Library 1," Houston, Texas.

Lehmann, E. L. (1949), Notes on the Theory of Estimation, Berkeley: University of California Press.

Mann, Nancy R. (1969), "Cramér-Rao Efficiencies of Best Linear Invariant Estimators of Parameters of the Extreme-Value Distribution under Type II Censoring from Above," SIAM Journal on Applied Mathematics, 17, 1150-1162.

Saleh, A. K. (1966), "Estimation of the Parameters of the Exponential Distribution Based on Optimum Order Statistics in Censored Samples," Annals of Mathematical Statistics, 37, 1717-1735.

Sarhan, A. E., and Greenberg, B. G. (1955), "Estimation of the Mean and Standard Deviation by Order Statistics. Part III," Annals of Mathematical Statistics, 26, 576-592.

________ (1957), "Tables for Best Linear Estimates by Order Statistics of the Parameters of Single Exponential Distributions from Singly and Doubly Censored Samples," Journal of the American Statistical Association, 52, 58-87.

Saw, J. G. (1959), "Estimation of the Normal Population Parameters Given a Singly Censored Sample," Biometrika, 46, 150-158.

Wasan, M. T. (1970), Parametric Estimation, New York: McGraw-Hill Book Co.

VITA

James Wyckoff was born April 29, 1950, in Coffeyville, Kansas. He attended elementary and high school at Independence, Kansas, and graduated from high school in May, 1968. He has received his college education from Independence Community Junior College, in Independence, Kansas; Kansas State University, in Manhattan, Kansas; Emporia Kansas State College, in Emporia, Kansas; and the University of Missouri-Rolla, in Rolla, Missouri. He received a B.S.E. degree in Mathematics Education from Emporia Kansas State College in May, 1972, and in August, 1973, received the M.A. degree in Mathematics from the same institution.

He has been enrolled in the Graduate School of the University of Missouri-Rolla since August, 1973, and held graduate teaching and research assistantships during the time of his tenure there.