ON THE RELATIVE EFFICIENCIES OF BAN ESTIMATES BASED ON DOUBLY TRUNCATED AND CENSORED SAMPLES

by B. Raja Rao, *Department of Mathematics and Statistics, Poona University*

(Communicated by V. S. Huzurbazar, F.N.I.)

(*Received July 22 ; read October 17, 1958*)

Abstract

It is shown that the efficiency of BAN (best asymptotically normal) estimates which involve computation in terms of a doubly truncated or censored probability law is, in general, necessarily less than the efficiency of BAN estimates based on the original (untruncated or uncensored) probability law and, further, that the efficiency increases, i.e. the loss of information decreases as $a \leftarrow x_0'$ or $x_0'' \rightarrow b$ where x_0' and x_0'' $(x_0' < x_0'')$ are the points of truncation. This is established for estimates of a single parameter and, also, for joint estimates of several parameters involved in a continuous probability distribution. In the case of a doubly censored distribution, it is shown that certain amount of information is lost due to not knowing the number of missing observations separately for each 'tail'. Thus our results indicate that in any case of regular estimation, the use of a truncated or censored probability law should always be accompanied by an investigation into the loss of efficiency.

1. Introduction

The problem of estimating the parameters of a distribution from truncated and censored samples, assuming various forms of truncation, has been treated by several authors, Cohen (1950), Gupta (1952, and the references cited therein) among others. Populations can be singly or doubly truncated, but samples can be censored in two different ways: (1) Observations below or above a known point may be censored, and (2) Out of a sample of size $r+n+s$, the first r and the last s observations may be censored. According to terminology which has recently come into popular usage, truncated samples are understood to be those from which the number of observations eliminated by the restricting process is unknown, and the censored samples are those in which the total number of specimens is known, but measurements on some known number of these are lacking.

BAN estimates are defined by Neyman (1949) as those functions of the observed sample values which (i) are consistent, (ii) are asymptotically normally distributed and (iii) are asymptotically efficient. It thus becomes important to know the answers to some of the questions in this direction: if a random sample of size n is drawn from a population involving an unknown parameter θ, then Fisher's information index $n\,I(\theta)$, defined in the usual way, measures the amount of information about θ obtainable from the sample or the reciprocal of the asymptotic variance of the BAN (best asymptotically normal) estimate of θ. Similarly, if a random sample of the same size is drawn from a doubly truncated or censored population, the points of truncation being x_0' and x_0'' $(x_0' < x_0'')$, can anything be said as to whether there is loss or gain of information due to truncation or censoring? In other words, does the efficiency of the BAN estimate of θ which involves computation in terms of a truncated or censored sample depend essentially upon the points of truncation? Does the knowledge of the number of missing observations in the regions of truncation improve or weaken the position? The present paper contains some information regarding the relative merits of these estimates based on doubly truncated and censored samples.

We shall treat the three cases, when the number of missing observations is (1) unknown for each tail, (2) known jointly for the two tails, and (3) known separately for each tail. The first characterizes a doubly truncated sample and the last two a doubly censored sample. The corresponding results for the singly truncated samples follow as special cases.

2. Relative Efficiency

Let $f(z)$ be the probability law of a continuously distributed random variable Z, involving an unknown parameter θ, over the range $a \leqslant z \leqslant b$, where a and b are independent of θ. Rao (1945) and simultaneously Cramér (1946) have shown that, under certain mild restrictions on $f(z)$, the variance of an unbiased estimate $\theta^* = \theta^*(z_1, z_2, \ldots, z_n)$, of the parameter θ, where $z_1, z_2, \ldots, z_n$ are the observed sample, satisfies the following inequality for fixed n:

$$\text{Var}\ \theta^* \geqslant \frac{1}{nE\left(\frac{\partial}{\partial\theta}\log f\right)^2} = \left\{nI(\theta)\right\}^{-1}, \qquad (1)$$

where $n\,I(\theta)$ is Fisher's index of information about θ in the sample, the lower bound being attained only by 'efficient' estimates.

When the distribution is truncated or censored, we shall replace Z by X and denote respectively by x_0' and x_0'' the minimum and the maximum values of X that can be observed. x_0' and x_0'' are assumed to be known in advance. Then (a, x_0') and (x_0'', b) respectively form the left and right tails. If $x_1, x_2, \ldots, x_n$ is a random sample from this incomplete population, we shall define $n\,I_{Ti}(\theta) =$ Fisher's index of information about θ in the truncated sample characterized by case (i), $i = 1, 2, 3$.

The BAN estimate of θ in a complete population has minimum variance given by $\{n\,I(\theta)\}^{-1}$ and similarly the variance of the BAN estimate of θ calculated from the incomplete population is given by $\{n\,I_{Ti}(\theta)\}^{-1}$. This suggests that the optimum variances of the estimates in the two populations in samples of the same size n may differ and, further, $n\,I_{Ti}(\theta)$ may depend essentially upon the choice of x_0' and x_0''.

The information lost due to truncation and censoring may conveniently be measured in terms of Fisher's indices of information. We may consider the quantity

$$\lambda(x_0', x_0'') = I_{Ti}(\theta)/I(\theta) \qquad (2)$$

as a measure of the relative precision attainable in the estimation of the parameter of the truncated probability law in the case (i), $i = 1, 2, 3$. We shall show that only in the cases (2) and (3), when the number of eliminated observations is known jointly and separately for the two tails, $\lambda(x_0', x_0'') \leqslant 1$ for all x_0' and x_0'' and that $\lambda \to 1$ as $a \leftarrow x_0'$ and $x_0'' \to b$. Because of the nature of the results obtained, it is convenient to consider the loss function

$$\frac{1}{n} L_{0i}(x_0', x_0'') = I(\theta) - I_{Ti}(\theta), \quad i = 1, 2, 3. \qquad (3)$$

In the case of functions depending upon several parameters, $f(x/\theta_1, \theta_2, \ldots, \theta_s)$, and unbiased estimates, θ_ν^* ($\nu = 1, 2, \ldots, s$), which are functions of the observed sample values, with non-singular covariance matrix $\| \Lambda_{ik} \|$, Cramér (loc. cit.) showed that the fixed ellipsoid

$$n \sum_{i=1}^{s} \sum_{k=1}^{s} \delta_{ik} t_i t_k = s+2, \qquad (4)$$

where
$$\delta_{ik} = E\left\{\frac{\partial}{\partial\theta_i}\log f\cdot\frac{\partial}{\partial\theta_k}\log f\right\}$$

lies wholly within the concentration ellipsoid

$$\sum_{i=1}^{s}\sum_{k=1}^{s}\Lambda^{ik}t_i t_k = s+2, \qquad \ldots \quad \ldots \quad \ldots \quad (5)$$

where $\|\Lambda^{ik}\| = \|\Lambda_{ik}\|^{-1}$. The two ellipsoids will coincide if, and only if, the θ_ν^* are jointly efficient for the θ_ν. Thus the covariance matrix of a set of jointly efficient estimates is $\|n\delta_{ik}\|^{-1}$. In this case, we may define separately the relative efficiency with respect to each of the parameters as in (2), or we may consider the set of estimates for one function to possess greater concentration than the set for the other function if the fixed ellipsoid (4) for the first lies wholly within the similar ellipsoid for the second function. We shall adopt the latter procedure in § 4.

3. Estimation of a Single Parameter

The Cramér-Rao inequality (1) holds under the regularity conditions which involve the existence of $\frac{\partial}{\partial\theta}f(z)$ for all z in $a \leqslant z \leqslant b$ and the absolute convergence of the integral $\int_a^b \frac{\partial f}{\partial\theta}\,dz$. We may assume that we have a regular estimation case in Cramér's sense so that these conditions hold. The variable of integration in each of the following integrals being x, it may be omitted, where no confusion seems likely to occur, for the sake of convenience.

Case (1): *Number of missing observations unknown for each tail.*

In this case, X has a density, say $g(x)$, which is the conditional density of Z given that $x_0' \leqslant Z \leqslant x_0''$; thus

$$g(x) = f(x) \Big/ \int_{x_0'}^{x_0''} f \qquad (x_0' \leqslant x \leqslant x_0''). \qquad \ldots \quad \ldots \quad (6)$$

Here

$$I_{T_1}(\theta) = \left\{\int_{x_0'}^{x_0''}\frac{1}{f}\left(\frac{\partial f}{\partial\theta}\right)^2\right\}\left\{\int_{x_0'}^{x_0''} f\right\}^{-1} - \left\{\int_{x_0'}^{x_0''}\frac{\partial f}{\partial\theta}\right\}^2\left\{\int_{x_0'}^{x_0''} f\right\}^{-2}.$$

By actually computing the loss function (3), it will be seen that it is not possible to determine the sign of $L_{01}(x_0', x_0'')$ in general. It follows, therefore, that whether there is loss or gain of information due to truncation in this case cannot be said definitely but depends upon the particular density function at hand.

It is interesting to note, however, that

$$I(\theta) - \left(\int_{x_0'}^{x_0''} f\right) I_{T_1}(\theta) = \int_a^{x_0'}\frac{1}{f}\left(\frac{\partial f}{\partial\theta}\right)^2 + \int_{x_0''}^{b}\frac{1}{f}\left(\frac{\partial f}{\partial\theta}\right)^2 + \left\{\int_{x_0'}^{x_0''}\frac{\partial f}{\partial\theta}\right\}^2\left\{\int_{x_0'}^{x_0''} f\right\}^{-1}$$

is positive for all x_0' and x_0''. This is true even when $a \leftarrow x_0'$ or $x_0'' \rightarrow b$, separately.

CASE (2): *Number of unmeasured observations known jointly for the two tails.*

Here when $x_0' \leqslant Z \leqslant x_0''$, $X = Z$; when $Z < x_0'$ or $Z > x_0''$, the only information obtained about X is simply that $X < x_0'$ or $X > x_0''$. Thus X can be regarded as having a density, say $g(x)$ when $x_0' \leqslant x \leqslant x_0''$;

$$g(x) = f(x) \quad (x_0' \leqslant x \leqslant x_0'')$$

$$\text{Prob}\left(X < x_0', \text{ or } X > x_0''\right) = \int_a^{x_0'} f + \int_{x_0''}^b f. \quad \ldots \quad (7)$$

In this case the loss function becomes

$$\frac{L_{02}}{n}(x_0', x_0'') = \int_a^b \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2 - \int_{x_0'}^{x_0''} \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2 - \left\{\int_a^{x_0'} \frac{\partial f}{\partial \theta} + \int_{x_0''}^b \frac{\partial f}{\partial \theta}\right\}^2 \left\{\int_a^{x_0'} f + \int_{x_0''}^b f\right\}^{-1}$$

$$= \int_a^{x_0'} \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2 + \int_{x_0''}^b \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2 - \left\{\int_a^{x_0'} \frac{\partial f}{\partial \theta} + \int_{x_0''}^b \frac{\partial f}{\partial \theta}\right\}^2 \left\{\int_a^{x_0'} f + \int_{x_0''}^b f\right\}^{-1}$$

so that

$$\frac{L_{02}}{n}\left\{\int_a^{x_0'} f + \int_{x_0''}^b f\right\} = K_1^2 + K_2^2 + \left(\int_a^{x_0'} f\right)\left\{\int_{x_0''}^b \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2\right\}$$

$$+\left(\int_{x_0''}^b f\right)\left\{\int_a^{x_0'} \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2\right\} - 2\left(\int_a^{x_0'} \frac{\partial f}{\partial \theta}\right)\left(\int_{x_0''}^b \frac{\partial f}{\partial \theta}\right),$$

where

$$\left.\begin{aligned} K_1^2 &= \left(\int_a^{x_0'} f\right)\int_a^{x_0'} \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2 - \left(\int_a^{x_0'} \frac{\partial f}{\partial \theta}\right)^2 \geqslant 0 \\ K_2^2 &= \left(\int_{x_0''}^b f\right)\int_{x_0''}^b \frac{1}{f}\left(\frac{\partial f}{\partial \theta}\right)^2 - \left(\int_{x_0''}^b \frac{\partial f}{\partial \theta}\right)^2 \geqslant 0, \end{aligned}\right\} \quad \ldots \quad (7a)$$

by Schwartz's inequality for integrals. Thus it is apparently not possible to determine the sign of L_{02} in this case also. But it is shown in CASE (3) below that

$$I(\theta) - I_{T_3}(\theta) \geqslant 0$$

and in § 5 that $I_{T_3}(\theta) - I_{T_2}(\theta) \geqslant 0$. Combining these two results, it follows that $\frac{1}{n}L_{02} = I(\theta) - I_{T_2}(\theta) \geqslant 0$, so that in the above case, there is some loss of information. We shall now establish some monotonicity properties of L_{02}. For, we note that

$$\frac{\left\{\int_a^{x_0'} f+\int_{x_0''}^b f\right\}^2}{n\left\{\frac{\partial}{\partial\theta}f(x_0')\right\}^2}\frac{\partial L_{02}}{\partial x_0'} = U_{x_0'}-2V_{x_0'}+\frac{V^2_{x_0'}}{U_{x_0'}} = \frac{\left(U_{x_0'}-V_{x_0'}\right)^2}{U_{x_0'}} \geqslant 0$$

and that

$$\frac{\left\{\int_a^{x_0'} f+\int_{x_0''}^b f\right\}^2}{n\left\{\frac{\partial}{\partial\theta}f(x_0'')\right\}^2}\frac{\partial L_{02}}{\partial x_0''} = -T_{x_0''}+2W_{x_0''}-\frac{W^2_{x_0''}}{T_{x_0''}} = -\frac{\left(T_{x_0''}-W_{x_0''}\right)^2}{T_{x_0''}} \leqslant 0,$$

where

$$\left.\begin{aligned} U_{x_0'} &= \left\{\int_a^{x_0'} f+\int_{x_0''}^b f\right\}\left\{f(x_0')\right\}^{-1}, \quad V_{x_0'} = \left\{\int_a^{x_0'}\frac{\partial f}{\partial\theta}+\int_{x_0''}^b\frac{\partial f}{\partial\theta}\right\}\left\{\frac{\partial}{\partial\theta}f(x_0')\right\}^{-1} \\ T_{x_0''} &= \left\{\int_a^{x_0'} f+\int_{x_0''}^b f\right\}\left\{f(x_0'')\right\}^{-1}, \quad W_{x_0''} = \left\{\int_a^{x_0'}\frac{\partial f}{\partial\theta}+\int_{x_0''}^b\frac{\partial f}{\partial\theta}\right\}\left\{\frac{\partial}{\partial\theta}f(x_0'')\right\}^{-1}\end{aligned}\right\}. \quad (7b)$$

It follows, therefore, that L_{02} is monotone increasing in x_0' when x_0'' is fixed, and is monotone decreasing in x_0'' when x_0' is fixed, so that when $a \leftarrow x_0'$, $L_{02}(x_0', x_0'')$ decreases monotonically to $L(x_0'')$, which is the loss function due to truncation at x_0'' and when $x_0'' \to b$, $L_{02}(x_0'', x_0')$ decreases monotonically to $L(x_0')$, which represents the loss function due to truncation at x_0'.

CASE (3): *Number of eliminated observations known separately for each tail.*

In this case, X can be regarded as having the density

$$\begin{aligned} \text{Prob}(X < x_0') &= \int_a^{x_0'} f \\ g(x) &= f(x), \; x_0' < x < x_0'', \qquad \ldots \quad \ldots \quad (8) \\ \text{Prob}(X > x_0'') &= \int_{x_0''}^b f \end{aligned}$$

and the loss function becomes

$$\frac{L_{03}}{n}(x_0', x_0'') = I(\theta)-I_{T_3}(\theta)$$

$$= \int_a^{x_0'}\frac{1}{f}\left(\frac{\partial f}{\partial\theta}\right)^2 - \left(\int_a^{x_0'}\frac{\partial f}{\partial\theta}\right)^2 \Bigg/ \int_a^{x_0'} f + \int_{x_0''}^b \frac{1}{f}\left(\frac{\partial f}{\partial\theta}\right)^2 - \left(\int_{x_0''}^b\frac{\partial f}{\partial\theta}\right)^2 \Bigg/ \int_{x_0''}^b f \quad (9)$$

$$= \frac{1}{n} \cdot L_1(x_0') + \frac{L_2}{n}(x_0'')$$

so that

$$\frac{1}{n} L_{03} = K_1^2 \left(\int_a^{x_0'} f \right)^{-1} + K_2^2 \left(\int_{x_0''}^{b} f \right)^{-1},$$

where K_1^2, and K_2^2 are the same as those defined in (7*a*). It follows that $L_{03}(x_0', x_0'') \geqslant 0$, since the right member is non-negative. Equality holds if, and only if, $\frac{\partial f}{\partial \theta}$ is proportional to $f(x)$ for all $x \leqslant x_0'$ and $x \geqslant x_0''$.

In such a case,

$$f(x) = \Theta\, e^{v(x)},$$

where Θ is a constant depending only on θ. If now $v(x)$ is a constant, $f(x)$ is a rectangular probability density function. On the other hand, if $v(x)$ is not a constant, two cases arise, namely

$$\left. \begin{array}{lll} (a) & f(x) = \Theta\, e^{v(x)}, & a \leqslant x \leqslant b, \\ (b) & f(x) = \Theta\, e^{v(x)}, & x \leqslant x_0', \; x \geqslant x_0''. \end{array} \right\} \quad \ldots \quad \ldots (10)$$

In the first case, $\Theta = \left(\int_a^b e^{v(x)} dx \right)^{-1}$ and is independent of θ. Thus $f(x)$ does not contain θ, so that we do not have a case of estimation at all. In the second case, $f(x)$ is known to within a multiplicative constant depending upon θ and, hence, no essential information is lost in truncation. Thus except in these trivial cases, the relative efficiency is less than unity.

It then appears that, in every case of regular estimation, the variance of the BAN estimate of the parameter θ of the truncated distribution (8) is greater than the corresponding variance for the complete population. Thus it follows that the maximum likelihood estimate (*m.l.e.*) $\hat{\theta}$ of θ in the complete population is asymptotically more efficient than the *m.l.e.* $\hat{\theta}_T$ of θ for fixed n.

It is curious to note that the total loss function

$$\frac{1}{n} L_{03}(x_0', x_0'') = \frac{L_1(x_0')}{n} + \frac{L_2(x_0'')}{n}$$

is the sum of the loss functions due to truncation on the left at $x = x_0'$ and on the right at $x = x_0''$. For the monotonicity properties of $L_{03}(x_0', x_0'')$ we note that

$$\frac{1}{n} \frac{\partial L_{03}}{\partial x_0'} = \frac{1}{n} \frac{d}{dx_0'} L_1(x_0') = \left\{ \frac{\partial}{\partial \theta} f(x_0') \right\}^2 \left\{ \int_a^{x_0'} f \right\}^{-1} \frac{\left(P_{x_0'} - Q_{x_0'}\right)^2}{P_{x_0'}} \geqslant 0,$$

and

$$\frac{1}{n} \frac{\partial L_{03}}{\partial x_0''} = \frac{1}{n} \frac{d}{dx_0''} L_2(x_0'') = -\left\{ \frac{\partial}{\partial \theta} f(x_0'') \right\}^2 \left\{ \int_{x_0''}^{b} f \right\}^{-1} \frac{\left(R_{x_0''} - S_{x_0''}\right)^2}{R_{x_0''}} \leqslant 0,$$

where

$$\left.\begin{aligned} P_{x_0'} &= \left(\int_a^{x_0'} f\right)\left\{f(x_0')\right\}^{-1},\ Q_{x_0'} = \left(\int_a^{x_0'} \frac{\partial f}{\partial \theta}\right)\left\{\frac{\partial}{\partial \theta} f(x_0')\right\}^{-1} \\ R_{x_0''} &= \left(\int_{x_0''}^{b} f\right)\left\{f(x_0'')\right\}^{-1},\ S_{x_0''} = \left(\int_{x_0''}^{b} \frac{\partial f}{\partial \theta}\right)\left\{\frac{\partial}{\partial \theta} f(x_0'')\right\}^{-1} \end{aligned}\right\}. \qquad \ldots (10a)$$

Observe that $R_{x_0''}$ is the Mills' ratio for the distribution with the density $f(x)$ and $S_{x_0''}$ can be regarded as the Mills' ratio for the distribution with the density = const. $\left|\left(\frac{\partial f}{\partial \theta}\right)\right|$, $(a \leqslant x \leqslant b)$. [See Raja Rao (*in press*)].

It follows, therefore, that $L_{03}(x_0', x_0'')$ is increasing in x_0' whatever be x_0'' and is decreasing in x_0'' whatever be x_0', so that $L_{03}(x_0', x_0'') \xrightarrow{\text{mon}} 0$ as $a \leftarrow x_0'$, and $x_0'' \rightarrow b$.

In other words, $\lambda(x_0', x_0'') \leqslant 1$ for all x_0' and x_0'' and $\lambda(x_0', x_0'') \xrightarrow{\text{mon}} 1$ as $a \leftarrow x_0'$ and $x_0'' \rightarrow b$. Thus the variance of the BAN estimate of the parameter θ of the truncated distribution (8) depends essentially upon the points of truncation and decreases monotonically to the variance of the BAN estimate of the parameter in a complete population.

Note that (9) may be written as

$$\frac{1}{n} L_{03}(x_0', x_0'') = \left(\int_a^{x_0'} f\right) \underset{x \leqslant x_0'}{E} \left\{\frac{\partial}{\partial \theta} \log\left(f \Big/ \int_a^{x_0'} f\right)\right\}^2 + \left(\int_{x_0''}^{b} f\right) \underset{x \geqslant x_0''}{E} \left\{\frac{\partial}{\partial \theta} \log\left(f \Big/ \int_{x_0''}^{b} f\right)\right\}^2, \qquad \ldots \quad \ldots (9')$$

which means that the loss of information due to truncation is equal to $\left(\int_a^{x_0'} f\right)$ times the amount of information obtainable from a sample of size n from the left tail, together with $\left(\int_{x_0''}^{b} f\right)$ times that obtainable from a sample of the same size from the right tail.

4. Estimation of Several Parameters

Consider the probability law $f(x/\theta_1, \theta_2, \ldots, \theta_s)$ depending upon s unknown parameters with ellipsoid of concentration for a set of jointly efficient estimates given by (4). For brevity we shall treat only the Case (3), so that the probability law (8) has the corresponding ellipsoid of concentration

$$n \sum_{i=1}^{s} \sum_{k=1}^{s} \delta'_{ik} t_i t_k = s+2, \qquad \ldots \quad \ldots \quad \ldots \quad \ldots (11)$$

where

$$\delta'_{ik} = \int_{x_0'}^{x_0''} \frac{1}{f}(f_i f_k) + \left(\int_a^{x_0'} f_i\right)\left(\int_a^{x_0'} f_k \Big/ \int_a^{x_0'} f\right) + \left(\int_{x_0''}^{b} f_i\right)\left(\int_{x_0''}^{b} f_k \Big/ \int_{x_0''}^{b} f\right),$$

$(i, k = 1, 2, \ldots s)$

and $f_i = \dfrac{\partial f}{\partial \theta_i}$ and $f_k = \dfrac{\partial f}{\partial \theta_k}$. We shall show in this section that the ellipsoid (4) lies entirely within (11); to this end, it is enough if it is shown that the left member of (4) is uniformly greater than the left member of (11), for every choice of the t_i, $i = 1, 2, \ldots, s$. By direct substitution we obtain

$$\frac{1}{n} L(x_0', x_0'') = \sum_{i,k} (\delta_{ik} - \delta'_{ik}) t_i t_k$$

$$= \sum_{i,k} \left[\int_a^{x_0'} \frac{1}{f}(f_i f_k) - \left(\int_a^{x_0'} f_i\right)\left(\int_a^{x_0'} f_k \Big/ \int_a^{x_0'} f\right) + \int_{x_0''}^{b} \frac{1}{f}(f_i f_k) - \left(\int_{x_0''}^{b} f_i\right)\left(\int_{x_0''}^{b} f_k \Big/ \int_{x_0''}^{b} f\right)\right] t_i t_k \quad \ldots \quad (12)$$

Terms in the rectangular brackets may be written as

$$\int_a^{x_0'} \left[\frac{f_i f_k}{f} - \left(\int_a^{x_0'} f_i \Big/ \int_a^{x_0'} f\right) f_k - \left(\int_a^{x_0'} f_k \Big/ \int_a^{x_0'} f\right) f_i + \left(\int_a^{x_0'} f_i\right)\left(\int_a^{x_0'} f_k \Big/ \int_a^{x_0'} f\right) f(x) \quad \int_a^{x_0'} f\right] dx +$$

$$\int_{x_0''}^{b} \left[\frac{f_i f_k}{f} - \left(\int_{x_0''}^{b} f_i \Big/ \int_{x_0''}^{b} f\right) f_k - \left(\int_{x_0''}^{b} f_k \Big/ \int_{x_0''}^{b} f\right) f_i + \left(\int_{x_0''}^{b} f_i\right)\left(\int_{x_0''}^{b} f_k \Big/ \int_{x_0''}^{b} f\right) f(x) \Big/ \int_{x_0''}^{b} f\right] dx.$$

This again can be written as

$$\left(\int_a^{x_0'} f\right)\int_a^{x_0'} \left[\frac{f_i}{f} - \left(\int_a^{x_0'} f_i \Big/ \int_a^{x_0'} f\right)\right]\left[\frac{f_k}{f} - \left(\int_a^{x_0'} f_k \Big/ \int_a^{x_0'} f\right)\right]\left(f(x) \Big/ \int_a^{x_0'} f\right) dx +$$

$$\left(\int_{x_0''}^{b} f\right)\int_{x_0''}^{b} \left[\frac{f_i}{f} - \left(\int_{x_0''}^{b} f_i \Big/ \int_{x_0''}^{b} f\right)\right]\left[\frac{f_k}{f} - \left(\int_{x_0''}^{b} f_k \Big/ \int_{x_0''}^{b} f\right)\right]\left(f(x) \Big/ \int_{x_0''}^{b} f\right) dx.$$

Thus the loss function can be written as

$$\frac{1}{n}L(x_0', x_0'') = \sum_{i,k} (\epsilon_{ik}'+\epsilon_{ik}'') \quad \ldots \quad (13)$$

where

$$\epsilon_{ik}' = \left(\int_a^{x_0'} f\right) \underset{x \leqslant x_0'}{E} \left\{ \frac{f_i}{f} - \left(\int_a^{x_0'} f_i \Big/ \int_a^{x_0'} f\right)\right\}\left\{\frac{f_k}{f} - \left(\int_a^{x_0'} f_k \Big/ \int_a^{x_0'} f\right)\right\} t_i t_k \quad (13a)$$

and

$$\epsilon_{ik}'' = \left(\int_{x_0''}^{b} f\right) \underset{x \geqslant x_0''}{E} \left\{ \frac{f_i}{f} - \left(\int_{x_0''}^{b} f_i \Big/ \int_{x_0''}^{b} f\right)\right\}\left\{\frac{f_k}{f} - \left(\int_{x_0''}^{b} f_k \Big/ \int_{x_0''}^{b} f\right)\right\} t_i t_k, \quad (13b)$$

where $\underset{x \leqslant x'_0}{E}$ denotes mathematical expectation over the set $x \leqslant x'_0$. Finally, since the (finite) sum of the expected values is equal to the expected value of the sum, we have

$$\sum_{i,k} \epsilon_{ik}' = \left(\int_a^{x_0'} f\right) \underset{x \leqslant x_0'}{E} \left[\sum_{i=1}^{s} t_i \left\{\frac{f_i}{f} - \left(\int_a^{x_0'} f_i \Big/ \int_a^{x_0'} f\right)\right\}\right]^2 \geqslant 0 \quad \ldots \quad (14')$$

$$= \left(\int_a^{x_0'} f\right) \underset{x \leqslant x_0'}{E} \left[\sum_{i=1}^{s} t_i \left\{\frac{\partial}{\partial\theta_i} \log \left(f \Big/ \int_a^{x_0'} f\right)\right\}\right]^2 \geqslant 0$$

and

$$\sum_{i,k} \epsilon_{ik}'' = \left(\int_{x_0''}^{b} f\right) \underset{x \geqslant x_0''}{E} \left[\sum_{i=1}^{s} t_i \left\{\frac{f_i}{f} - \left(\int_{x_0''}^{b} f_i \Big/ \int_{x_0''}^{b} f\right)\right\}\right]^2 \geqslant 0 \quad (14'')$$

$$= \left(\int_{x_0''}^{b} f\right) \underset{x \geqslant x_0''}{E} \left[\sum_{i=1}^{s} t_i \left\{\frac{\partial}{\partial\theta_i} \log \left(f \Big/ \int_{x_0''}^{b} f\right)\right\}\right]^2 \geqslant 0.$$

It follows that since $\left(\int_a^{x_0'} f\right)$ and $\left(\int_{x_0''}^{b} f\right)$ are positive, $L(x_0', x_0'') \geqslant 0$. We need only note that $L(x_0', x_0'') = 0$ only if the linear form in each of the equations (14′) and (14″) is identically zero, i.e. if each coefficient of t_i vanishes. This can happen only in the trivial cases analogous to those described in § 3.

Thus it is shown that the ellipsoid of concentration of a set of jointly efficient estimates of the parameters of a complete population lies wholly within the corresponding ellipsoid for the truncated population. Therefore, the best procedure for estimating the parameters of a truncated population cannot attain the precision of an efficient procedure for estimating those of a complete population. Note that with ϵ_{ik}' and ϵ_{ik}'' as defined in (13a) and (13b)

$$\left(n \Big/ \int_a^{x_0'} f\right) \sum_i \sum_k \epsilon_{ik}' t_i t_k = s+2,$$

$$\left(n \Big/ \int_{x_0''}^{b} f\right) \sum_i \sum_k \epsilon_{ik}'' t_i t_k = s+2$$

are respectively the ellipsoids of concentration of a set of jointly efficient estimates of the parameters of the density functions of the left tail, $l(x) = f \Big/ \int_a^{x_0'} f$, $(a \leqslant x \leqslant x_0')$ and the right tail, $\gamma(x) = f \Big/ \int_{x_0''}^{b} f$, $(x_0'' \leqslant x \leqslant b)$.

To complete the argument for the multiparametric case, we shall show that $L(x_0', x_0'') \to 0$ monotonically when $a \leftarrow x_0'$ and $x_0'' \to b$. This follows if it is shown that $L(x_0', x_0'')$ is increasing in x_0' whatever be x_0'' and decreasing in x_0'', whatever be x_0'. From (13), we have, with notation $f_i' = \frac{\partial}{\partial \theta_i} f(x_0')$, $f_i'' = \frac{\partial}{\partial \theta_i} f(x_0'')$,

$$\frac{1}{n}\frac{\partial L}{\partial x_0'} = f(x_0') \left[\sum_{i=1}^{s} t_i \left\{ \frac{f_i'}{f(x_0')} - \left(\int_a^{x_0'} f_i \Big/ \int_a^{x_0'} f \right) \right\} \right]^2 \geqslant 0$$

and

$$\frac{1}{n}\frac{\partial L}{\partial x_0''} = -f(x_0'') \left[\sum_{i=1}^{s} t_i \left\{ \frac{f_i''}{f(x_0'')} - \left(\int_{x_0''}^{b} f_i \Big/ \int_{x_0''}^{b} f \right) \right\} \right]^2 \leqslant 0$$

which proves the assertion. Thus the BAN estimates of the parameters θ_ν ($\nu = 1, 2, \ldots, s$) of the truncated population are asymptotically jointly less efficient than the corresponding estimates for the complete population and the efficiency of the former set increases monotonically to that of the latter as $a \leftarrow x_0'$, $x_0'' \to b$.

5. Information lost due to not knowing the number of missing observations separately for each tail can conveniently be measured. Denoting respectively by $nI_{T_2}(\theta)$ and $nI_{T_3}(\theta)$ the amounts of information about the parameter θ in the cases (2) and (3), we observe that, defining $l(x_0', x_0'') = nI_{T_3}(\theta) - nI_{T_2}(\theta)$,

$$\frac{l}{n} = \left(\int_a^{x_0'} \frac{\partial f}{\partial \theta}\right)^2 \Big/ \int_a^{x_0'} f + \left(\int_{x_0''}^{b} \frac{\partial f}{\partial \theta}\right)^2 \Big/ \int_{x_0''}^{b} f - \left\{ \int_a^{x_0'} \frac{\partial f}{\partial \theta} + \int_{x_0''}^{b} \frac{\partial f}{\partial \theta} \right\}^2 \times \left\{ \int_a^{x_0'} f + \int_{x_0''}^{b} f \right\}^{-1}$$

$$= \left[\left(\int_a^{x_0'} \frac{\partial f}{\partial \theta} \right) \left(\int_{x_0''}^b f \Big/ \int_a^{x_0'} f \right)^{\frac{1}{2}} - \left(\int_{x_0''}^b \frac{\partial f}{\partial \theta} \right) \left(\int_a^{x_0'} f \Big/ \int_{x_0''}^b f \right)^{\frac{1}{2}} \right]^2$$

$$\times \left\{ \int_a^{x_0'} f + \int_{x_0''}^b f \right\}^{-1} \geqslant 0.$$

Thus we find that some amount of information is lost due to not knowing the number of missing observations separately for each tail. Further, since when $a \leftarrow x_0'$ or $x_0'' \rightarrow b$, Case (2) coincides with Case (3) thus resulting in singly censored samples, the above loss of information must tend to zero when either $a \leftarrow x_0'$ or $x_0'' \rightarrow b$, the other point of truncation remaining fixed. We note, however, that

$$\frac{1}{n} f(x_0') \left\{ \frac{\partial}{\partial \theta} f(x_0') \right\}^{-2} \frac{\partial l}{\partial x_0'} = \left(\frac{V_{x_0'}}{U_{x_0'}} - 1 \right)^2 - \left(\frac{Q_{x_0'}}{P_{x_0'}} - 1 \right)^2 \quad \ldots \quad \ldots \quad (15)$$

and

$$\frac{1}{n} f(x_0'') \left\{ \frac{\partial}{\partial \theta} f(x_0'') \right\}^{-2} \frac{\partial l}{\partial x_0''} = -\left(\frac{W_{x_0''}}{T_{x_0''}} - 1 \right)^2 + \left(\frac{S_{x_0''}}{R_{x_0''}} - 1 \right)^2, \quad \ldots \quad (16)$$

where the various quantities are defined in § 4. It has not been possible to determine the signs of the expressions in (15) and (16).

Acknowledgements

The author wishes to thank Dr. V. S. Huzurbazar, F.N.I., for helpful advice in the preparation of this paper, and the Government of India for awarding him a Senior Research Scholarship.

References

Cohen, Jr., A. C. (1950). Estimating mean and variance of normal populations from singly truncated and doubly truncated samples. *Ann. Math. Stat.*, **21**, 557.

Cramér, H. (1946). Contributions to the theory of statistical estimation. *Skand. Aktuari*, **29**, 85.

Gupta, A. K. (1952). Estimation of the mean and standard deviation of a normal population from a censored sample. *Biometrika*, **39**, 260.

Neyman, J. (1949). Proc. Berkeley Symp. Math. Stat. Prob. University of California Press, p. 239.

Raja Rao, B. (1959). A double inequality on Mills' ratio for the class of distributions admitting sufficient statistics. *Jour. Univ., Poona, Sci. & Tech.*, **13**. (To appear.)

Rao, C. R. (1945). Information and accuracy attainable in the estimation of statistical parameters. *Bull. Cal. Math. Soc.*, **37**, 81.