ASYMPTOTIC DISTRIBUTION OF LINEAR COMBINATIONS OF FUNCTIONS OF ORDER STATISTICS WITH APPLICATIONS TO ESTIMATION[1]

By Herman Chernoff, Joseph L. Gastwirth and M. V. Johns, Jr.

Stanford University

1. Introduction. The research leading to the results reported in this paper was originally motivated by a desire to find asymptotically optimal linear combinations of order statistics for estimating location and scale parameters in both the uncensored and censored data cases. It was recognized at the outset that the methods of Chernoff-Savage [2] could be applied to obtain the asymptotic normal distribution of statistics of the form

$$T_n = n^{-1} \sum J[j/(n + 1)]X_{jn} ,$$

where $X_{1n} \leqq X_{2n} \leqq \cdots \leqq X_{nn}$ are the ordered observations of a sample, and where $J(\cdot)$ is a well behaved function. It was also discovered that a simple, standard variational argument could be used to obtain the asymptotically optimal J's explicitly for the case of estimating location and scale parameters.

Subsequently, the attention of the authors was directed to the unpublished dissertation of Carl Bennett [1] where the asymptotically optimal J's had been obtained for both the uncensored and multicensored cases by a tour de force which did not include a derivation of the asymptotic normality of the estimates. Some of Bennett's results were obtained independently by Jung [5] under rather restrictive conditions. Plackett [9], and Weiss [13], independently considered the case where all observations below the pth and above the qth sample percentiles ($0 < p < q < 1$) are censored, and obtained asymptotic normality for suitable linear combinations of the available order statistics. Plackett also characterized the asymptotically optimal weights for this case. Asymptotic normality for the case of uncensored data was not treated by these authors.

The present authors found that, while the Chernoff-Savage approach was adequate for the particular applications initially considered, it nevertheless involved certain objectionable mathematical inelegancies. A different technique based on representing the ordered observations in terms of independent exponentially distributed random variables was therefore selected. This technique yields stronger results and yet involves only arguments which are essentially elementary. Regrettably, these results still seem to fall short of the "best possible" results which may ultimately require a sophisticated "invariance principle" type of argument.

The authors have benefited from helpful discussions with several colleagues. Early in the course of this research two of the authors had several interesting conversations with Z. Govindarajulu who subsequently derived results overlapping some of those of the present paper in [3]. Govindarajulu's technique is based on an unpublished result of LeCam [7], and his main result requires bounds on $J(u)$ and $J'(u)$ as $u \to 0$ or 1, which ours does not. On the other hand, we require a smoothness condition on the tails of the distribution of the observations which Govindarajulu apparently does not need.

In Section 2 we first obtain a quite general theorem concerning the conditions under which statistics of the form

$$T_n = n^{-1} \sum c_{jn} h(X_{jn})$$

are asymptotically normally distributed. This theorem and its corollaries are then specialized to the more useful Theorem 2 and finally to Theorem 3, where the c_{jn}'s are of the form $c_{jn} = J[j/(n+1)]$. These theorems involve the decomposition

$$T_n = \mu_n + Q_n + R_n,$$

where μ_n is non-random,

$$Q_n = n^{-1} \sum \alpha_{jn}(Z_j - 1),$$

where the Z_j's are independent and identically distributed exponential random variables, $n^{\frac{1}{2}} Q_n$ is asymptotically normal, and R_n is asymptotically negligible. Corollaries 3 and 4 of Theorem 3 are concerned with the case where additional weight is given to certain sample percentiles, and useful alternative formulae for the variances of the resulting asymptotic distributions are given. The decomposition referred to above facilitates the consideration of vectors whose components are statistics such as T_n or functions of such statistics.

In Section 3 the results of Section 2 are applied to the problem of obtaining asymptotically efficient estimates for location and scale parameters. It is shown that the covariance matrix of the asymptotic normal distribution of Bennett's estimators coincides with the Cramér-Rao bounds given by the inverse of the Fisher information matrix. Given the estimates and the above demonstration of efficiency it is unnecessary to present the formal variational derivation of the estimates. It should be noted that a heuristic derivation of the estimates is possible based on a linearized approximation to the maximum likelihood estimates for a related "partially grouped data" problem. This derivation is not discussed further in this paper.

Examples of applications of these results to the logistic, Cauchy and normal distributions are given in Section 3.

2. Asymptotic distribution of linear combinations of order statistics. In this section we study the asymptotic distribution of

(2.1) $$T_n = n^{-1} \sum_{j=1}^{n} c_{jn} h(X_{jn})$$

where $X_{1n} \leqq X_{2n} \leqq \cdots \leqq X_{nn}$ are the ordered observations from a random sample of size n from a distribution F. Theorem 3 will specialize earlier results to the case where $c_{jn} = J[j/(n + 1)]$, i.e.

$$(2.1') \qquad T_n = n^{-1} \sum_{j=1}^{n} J[j/(n + 1)]h(X_{jn}).$$

For applications, it is convenient to represent the results in terms of decompositions of the form

$$(2.2) \qquad T_n = \mu_n + Q_n + R_n$$

where μ_n is non random,

$$(2.3) \qquad Q_n = n^{-1} \sum_{j=1}^{n} \alpha_{jn}(Z_j - 1),$$

with $Z_1, Z_2, \cdots$ independently and identically distributed according to the negative exponential distribution $1 - e^{-z}$, $z \geqq 0$ and R_n is a remainder which is asymptotically negligible.

The following notational conventions will be observed. The probability law of a random variable (or random vector) X will be indicated by $\mathcal{L}(X)$. In the special case where X is normally distributed with mean vector μ and covariance matrix Σ, its law is indicated by the symbol $\mathfrak{N}(\mu, \Sigma)$. The Mann-Wald symbols, O_p and o_p are to be interpreted in the usual sense, i.e., if X_n, $n = 1, 2, \cdots$, is a sequence of random variables and $g(\cdot)$ is a positive function then the statement $X_n = O_p[g(n)]$ means that $X_n/g(n)$ is bounded in probability for all n and the statement $X_n = o_p[g(n)]$ means that $X/g(n)$ approaches zero in probability as $n \to \infty$. We shall write

$$\sum_{i=j+1}^{k} a_i = \sum_{i=1}^{k} a_i - \sum_{i=1}^{j} a_i$$

even when $j \geqq k$. Furthermore

$$(2.4) \qquad \mathfrak{M}_u^{u'}\{b_{jn}\} = n^{-1} \sum_{j=[nu]+1}^{[nu']} b_{jn}$$

where the limits u and u' are suppressed when they are 0 and 1 respectively. In particular we write

$$\mathfrak{M}_u^{u'}\{b_{jn}\} = O(\sigma_n) \qquad \textit{absolutely}$$

if $\mathfrak{M}_u^{u'}\{b_{jn}\} = O(\sigma_n)$ and if, for any sub-interval E of (u, u') of length δ, $n\sigma_n^{-1} \sum_{j/n \varepsilon E} |b_{jn}| \leqq \rho(\delta, n)$ where $\rho(\delta, n)$ approaches 0 as $\delta \to 0$ and $n \to \infty$.

Let F^{-1} be an inverse (not necessarily unique) of the cdf F. Then the X_{jn} may be represented from the point of view of distribution by

$$(2.5) \qquad X_{jn} = F^{-1}(U_{jn}),$$

where the U_{jn} are the ordered observations from a sample of size n on the uniform distribution on (0, 1) Similary

$$(2.6) \qquad V_{jn} = -\log(1 - U_{jn})$$

represent the ordered observations from the negative exponential distribution

with cdf $1 - e^{-v}$, $v \geq 0$. We define $\tilde{H}$ and H according to

$$(2.7) \qquad h(x) = H(u) = \tilde{H}(v)$$

where $x = F^{-1}(u)$ and $v = -\log(1 - u)$. It follows from (2.7) that $H'(u)(1 - u) = \tilde{H}'(v)$. We shall observe that the mean of V_{jn} is given by

$$(2.8) \qquad \bar{v}_{jn} = E(V_{jn}) = n^{-1} + (n-1)^{-1} + \cdots + (n-j+1)^{-1}$$

which is close to

$$(2.9) \qquad v_{jn} = -\log[1 - j/(n+1)].$$

Certain elementary and well known propositions which are referred to in subsequent arguments are listed here without proof. Let X_n, Y_n, $n = 1, 2, \cdots$, be sequences of random variables.

PROPOSITION 1. (Slutsky). *If $X_n = Y_n + o_p(1)$ and $\mathcal{L}(Y_n) \to \mathcal{L}(Y)$ for some random variable, Y, then $\mathcal{L}(X_n) \to \mathcal{L}(Y)$ as $n \to \infty$.*

PROPOSITION 2. *For any $c > 0$, $P\{|X_n| > c\} < c^{-1}E\{|X_n|\}$. Consequently, $X_n = O_p\{E\{|X_n|\}\}$, and if $Y_n = \sum_{j=1}^n X_j$, then $Y_n = O_p\{\sum_{j=1}^n [E(X_j^2)]^{\frac{1}{2}}\}$.*

PROPOSITION 3. *The V_{jn} may be represented by*

$$V_{jn} = Z_1/n + Z_2/(n-1) + \cdots + Z_j/(n-j+1), \qquad j = 1, 2, \cdots, n,$$

where $Z_1, Z_2, \cdots, Z_n$ are independent and identically distributed random variables with common cdf $1 - e^{-z}$, $z \geq 0$.

Equation 2.8 follows immediately as does

$$(2.10) \quad \omega_{jn} = E\{(V_{jn} - \bar{v}_{jn})^2\} = n^{-2} + (n-1)^{-2} + \cdots + (n-j+1)^{-2},$$

and, using approximations by integrals it is easy to see that

$$j/(n+1) < 1 - e^{-\bar{v}_{jn}} < j/(n+\tfrac{1}{2}).$$

Thus, since $v_{jn} = -\log(1 - j/(n+1))$, the mean value theorem yields

$$0 \leqq \bar{v}_{jn} - v_{jn} < j/(2n+1)(n-j+\tfrac{1}{2}),$$

and by integral approximations again,

$$j/(n+1)(n-j+1) < \omega_{jn} < j/(n+\tfrac{1}{2})(n-j+\tfrac{1}{2})$$

$$< 4j/(n+1)(n-j+1).$$

Clearly the random variables U_{in} and V_{in} can be simultaneously bounded in probability. That is to say there exist $u_{in}(\epsilon)$, $u^{in}(\epsilon)$, $v_{in}(\epsilon)$, $v^{in}(\epsilon)$, such that

$$P\{u_{in}(\epsilon) < U_{in} < u^{in}(\epsilon), 1 \leqq i \leqq n\} \geqq 1 - \epsilon, \qquad n \geqq 1,$$

$$P\{v_{in}(\epsilon) < V_{in} < v^{in}(\epsilon), 1 \leqq i \leqq n\} \geqq 1 - \epsilon, \qquad n \geqq 1,$$

with

$$v_{in}(\epsilon) = -\log[1 - u_{in}(\epsilon)], \qquad v^{in}(\epsilon) = -\log[1 - u^{in}(\epsilon)],$$

$$u_{in}(\epsilon) < i/(n+1) < u^{in}(\epsilon), \qquad v_{in}(\epsilon) < v_{in} < \bar{v}_{in} < v^{in}(\epsilon).$$

PROPOSITION 4. *There exist $u_{in}(\epsilon)$, $u^{in}(\epsilon)$, $v_{in}(\epsilon)$, $v^{in}(\epsilon)$ as described above such that*

(a) (Glivenko-Cantelli)

$$\sup_{1 \leq i \leq n} [u^{in}(\epsilon) - u_{in}(\epsilon)] = o(1),$$

and

(b) (Based on the Kolmogorov inequality)

$$\sup_{1 \leq i \leq n} |v^{in}(\epsilon) - v_{in}(\epsilon)| \leq K(\epsilon) \qquad (K(\epsilon) \text{ independent of } n).$$

This implies

(c)

$$\sup_{1 \leq i \leq n} (1 - u_{in}(\epsilon))/(1 - u^{in}(\epsilon)) \leq K_1(\epsilon) \qquad (K_1(\epsilon) \text{ independent of } n)$$

and with the use of symmetry

(d) $$\sup_{1 \leq i \leq n} u^{in}(\epsilon)/u_{in}(\epsilon) \leq K_1(\epsilon).$$

PROPOSITION 5.[2] *If $X_{1n} \leq X_{2n} \leq \cdots \leq X_{nn}$ are order statistics from a sample of size n having common cdf $F(x)$ and empirical cdf $F_n(x)$, and if $F(\lambda_p) = p$, $0 < p < 1$, where $F'(\lambda_p)$ exists and is finite and positive, then*

$$(2.11) \qquad F_n(\lambda_p) - p = F'(\lambda_p)(\lambda_p - X_{[np],n}) + o_p(n^{-\frac{1}{2}}).$$

Incidentally Proposition 5 furnishes a useful proof of the asymptotic normality of the sample percentile $X_{[np],n}$, since the left hand side of (2.11) is clearly the sample average of $g(X)$ where $g(x) = 1 - p$ for $x \leq \lambda_p$ and $-p$ otherwise. This representation of $X_{[np],n}$ in terms of a sample average permits the derivation of the joint limiting distribution of several sample percentiles and sample averages. In particular, it quickly yields the result of Mosteller [8] and the result of Sarkadi, Schnell and Vincze [11] on the asymptotic joint distribution of the sample mean and a percentile.

We now present a preliminary lemma which can be derived by applying the Lindeberg-Feller theorem. The proof given here was suggested by R. G. Miller. For any sequence of constants α_{jn}, $j = 1, 2, \cdots, n$, $n = 1, 2, \cdots$, let $Q_n = n^{-1} \sum_{j=1}^{n} \alpha_{jn}(Z_j - 1)$, and $\sigma_n^2 = n^{-1} \sum_{j=1}^{n} \alpha_{jn}^2$.

LEMMA 1.

$$\mathcal{L}(n^{\frac{1}{2}}Q_n/\sigma_n) \to \mathfrak{N}(0, 1) \quad \text{if and only if} \quad \max_{1 \leq j \leq n} |\alpha_{jn}| = o(n^{\frac{1}{2}}\sigma_n)$$

as $n \to \infty$.

PROOF. Let $\varphi_n(t)$ be the characteristic function of $n^{\frac{1}{2}}Q_n/\sigma_n$. Then since the Z_j are independent and exponentially distributed

$$\log \varphi_n(t) = \sum_{j=1}^{n} \{-it \, \alpha_{jn}/n^{\frac{1}{2}}\sigma_n - \log (1 - it \, \alpha_{jn}/n^{\frac{1}{2}}\sigma_n)\}.$$

---

[2] This proposition although familiar to many investigators seems not to have appeared previously in the literature. A proof of a stronger result, however, may be found in "A note on quantiles in large samples" by R. R. Bahadur, *Ann. Math. Statist.* **37** 577–580, which appeared after the present paper was submitted for publication.

But

$$\sum_{j=1}^{n} \log\left[1 - it\,\alpha_{jn}/n^{\frac{1}{2}}\sigma_n\right] = -it\sum_{j=1}^{n}\alpha_{jn}/n^{\frac{1}{2}}\sigma_n + t^2/2 + r_n(t)$$

where $r_n(t) \to 0$ if [and $\mathrm{Re}[r_n(t)] \to 0$ only if] $\max_j |\alpha_{jn}/n^{\frac{1}{2}}\sigma_n| \to 0$ as $n \to \infty$. The desired result follows.

Let

(2.12) $$\mu_n = n^{-1}\sum_{j=1}^{n} c_{jn}\tilde{H}(\tilde{\nu}_{jn}),$$

(2.13) $$\alpha_{jn} = (n - j + 1)^{-1}\sum_{i=j}^{n} c_{in}\tilde{H}'(\tilde{\nu}_{in}),$$

(2.14) $$\sigma_n^{2} = n^{-1}\sum_{j=1}^{n}\alpha_{jn}^{2}.$$

Then

$$T_n = n^{-1}\sum_{j=1}^{n} c_{jn}h(X_{jn}) = n^{-1}\sum_{j=1}^{n} c_{jn}\tilde{H}(V_{jn}) = \mu_n + Q_n + R_n$$

where

(2.15) $$Q_n = n^{-1}\sum_{j=1}^{n}\alpha_{jn}(Z_j - 1),$$

$$R_n = n^{-1}\sum_{j=1}^{n} c_{jn}(V_{jn} - \tilde{\nu}_{jn})G_{jn}(V_{jn}),$$

and

$$G_{jn}(v) = [\tilde{H}(v) - \tilde{H}(\tilde{\nu}_{jn})]/(v - \tilde{\nu}_{jn}) - \tilde{H}'(\tilde{\nu}_{jn}) \quad \text{if} \quad v \neq \tilde{\nu}_{jn}$$

and 0 if $v = \tilde{\nu}_{jn}$.

ASSUMPTION A. $\tilde{H}(v)$ is continuously differentiable for $0 < v < \infty$.

ASSUMPTION B. For each $\epsilon > 0$,

$$\sum_{j=1}^{n} |c_{jn}|\, g_{jn}(\epsilon)[j/(n - j + 1)]^{\frac{1}{2}} = o(n\sigma_n)$$

where

$$g_{jn}(\epsilon) = \sup_{v_{jn}(\epsilon) < v < v i n(\epsilon)} |G_{jn}(v)|.$$

ASSUMPTION C. $\max_{1 \leq j \leq n} |\alpha_{jn}| = o(n^{\frac{1}{2}}\sigma_n)$.

THEOREM 1. If Assumptions A, B, and C are satisfied

$$\lim_{n\to\infty} \mathcal{L}[n^{\frac{1}{2}}(T_n - \mu_n)/\sigma_n] = \lim_{n\to\infty} \mathcal{L}[n^{\frac{1}{2}}Q_n/\sigma_n] = \mathfrak{N}(0, 1).$$

PROOF. By Assumption C and Lemma 1, $\mathcal{L}(n^{\frac{1}{2}}Q_n/\sigma_n) \to \mathfrak{N}(0, 1)$.

It remains only to show that $R_n = o_p(n^{-\frac{1}{2}}\sigma_n)$. Applying Proposition 4, we have

$$P\{|R_n| \leq n^{-1}\sum_{j=1}^{n} |c_{jn}g_{jn}(\epsilon)(V_{jn} - \tilde{\nu}_{jn})|\} \geq 1 - \epsilon$$

for all n. But

$$E\{\sum_{j=1}^{n} |c_{jn}g_{jn}(\epsilon)(V_{jn} - \tilde{\nu}_{jn})|\}$$

$$\leq E\{\sum_{j=1}^{n} |c_{jn}|\, g_{jn}(\epsilon)[E(V_{jn} - \tilde{\nu}_{jn})^2]^{\frac{1}{2}}\}$$

$$= O\{\sum_{j=1}^{n} |c_{jn}|\, g_{jn}(\epsilon)[j/(n + 1)(n - j + 1)]^{\frac{1}{2}}\} = o(n^{\frac{1}{2}}\sigma_n)$$

by Assumption B. Applying Proposition 2 it follows that

$$\sum |c_{jn}g_{jn}(\epsilon)(V_{jn} - \tilde{\nu}_{jn})| = o_p(n^{\frac{1}{2}}\sigma_n)$$

and hence $R_n = o_p(n^{-\frac{1}{2}}\sigma_n)$.

Theorem 1 is our fundamental theorem but its present form, while allowing for flexibility, is too general for most applications. Before commenting on the theorem and specializing it we shall digress somewhat to state two related results which we shall label as corollaries in order to emphasize the general approach. The first corollary deals with T_n equal to the function h evaluated at the sample percentile $X_{[np],n}$.

COROLLARY 1. (*Sample Percentile*). *If* $0 < p < 1$, $F(\lambda_p) = p$ *and* $\tilde{H}'(v)$ *exists and is finite and non-zero at* $v = \nu_p = -\log(1 - p)$ *then*

(i) $$h(X_{[np],n}) = h(\lambda_p) + Q_n + R_n$$

where $R_n = o_p(n^{-\frac{1}{2}})$ *and* $Q_n = n^{-1}\sum_{j=1}^{n}\alpha_{jn}(Z_j - 1)$ *with*

$$\alpha_{jn} = n(n - j + 1)^{-1}\tilde{H}'(\nu_p) = n(n - j + 1)^{-1}(1 - p)H'(p) \quad \text{if} \quad j \leq np,$$
$$= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{otherwise,}$$

and therefore,

(ii) $\mathcal{L}\{n^{\frac{1}{2}}[h(X_{[np],n}) - h(\lambda_p)]\} \to \mathfrak{N}(0, p(1 - p)^{-1}[\tilde{H}'(\nu_p)]^2) = \mathfrak{N}(0, \sigma^2).$

(iii) *If in addition* F' *and* h' *exist and are finite and non-zero at* λ_p ,

$$\tilde{H}'(\nu_p) = (1 - p)h'(\lambda_p)/F'(\lambda_p)$$

and

$$\sigma^2 = p(1 - p)[h'(\lambda_p)/F'(\lambda_p)]^2.$$

(Parts (ii) and (iii) are of course well known).

PROOF. The last two parts follow simply from the first. For that part, Lemma 1 applies to

$$Q_n = (V_{[np],n} - \tilde{\nu}_{[np],n})\tilde{H}'(\nu_p) = n^{-1}\sum_{j=1}^{n}\alpha_{jn}(Z_j - 1)$$

since

$$\sigma_n^2 = n^{-1}\sum_{j=1}^{n}\alpha_{jn}^2 = n\omega_{[np],n}[\tilde{H}'(\nu_p)]^2 \to [p/(1 - p)][\tilde{H}'(\nu_p)]^2.$$

It remains only to show that

$$R_n = \tilde{H}(V_{[np],n}) - \tilde{H}(\nu_p) - (V_{[np],n} - \nu_p)\tilde{H}'(\nu_p) - (\nu_p - \tilde{\nu}_{[np],n})\tilde{H}'(\nu_p)$$
$$= o_p(n^{-\frac{1}{2}}).$$

But $\tilde{\nu}_{[np],n} - \nu_p = o(n^{-\frac{1}{2}})$. The existence of $\tilde{H}'$ at ν_p implies that $\tilde{H}(v_n) - \tilde{H}(\nu_p) - (v_n - \nu_p)\tilde{H}'(\nu_p) = o(n^{-\frac{1}{2}})$, if $(v_n - \nu_p) = O(n^{-\frac{1}{2}})$. Since $V_{[np],n} - \nu_p = O_p(n^{-\frac{1}{2}})$, the desired result follows from the O_p calculus.

Corollary 1 could be treated more directly as a corollary by letting $c_{[np],n} = n$ and $c_{jn} = 0$ otherwise.

ASSUMPTION D. (Virtual Stability). There exists a sequence $\psi_n(p) \neq 0$ such that for b_n bounded away from 0 and ∞,

$$\sum_{j=[np]+1}^{[np+b_n n^{\frac{1}{2}}]} c_{jn} = b_n n^{\frac{1}{2}} \psi_n(p)[1 + o(1)].$$

COROLLARY 2. *For $0 < p < 1$ let $\Delta_p(v) = 1$ if $0 \leq v \leq \nu_p = -\log(1-p)$, and 0 otherwise. If Assumption D is satisfied*

$$T_n = n^{-1} \sum_{j=1}^{n} c_{jn}\Delta_p(V_{jn}) = \mu_n + Q_n + R_n$$

where

$$\mu_n = n^{-1} \sum_{j=1}^{[np]} c_{jn}$$

$$\sigma_n^2 = n(1-p)^2\psi_n^2(p)\omega_{[np],n} = p(1-p)\psi_n^2(p)[1 + O(n^{-1})]$$

$$Q_n = -(1-p)\psi_n(p) \sum_{j=1}^{[np]} (n-j+1)^{-1}(Z_j - 1)$$

$$R_n = o_p[n^{-\frac{1}{2}}\psi_n(p)]$$

and $\mathcal{L}\{n^{\frac{1}{2}}(T_n - \mu_n)/\sigma_n\} \to \mathfrak{N}(0,1)$.

(This corollary is relevant to functions H which have a finite number of jump discontinuities.)

PROOF. Referring to the above definitions of T_n, μ_n, Q_n,

$$R_n = T_n - \mu_n - Q_n = n^{-1} \sum_{j=[np]+1}^{B_n} c_{jn} - Q_n$$

where $B_n = \sum_{j=1}^{n} \Delta_p(V_{jn})$ is the number of $V_{jn} \leq \nu_p$. It is convenient for us to apply Proposition 5 to the exponential distribution to obtain

$$B_n - np = n(1-p)[\nu_p - V_{[np],n}] + o_p(n^{\frac{1}{2}})$$

$$= -n(1-p)[(\tilde{\nu}_{[np],n} - \nu_p) + \sum_{j=1}^{[np]}(n-j+1)^{-1}(Z_j - 1)]$$

$$\qquad + o_p(n^{\frac{1}{2}})$$

$$= n[\psi_n(p)]^{-1}Q_n + o_p(n^{\frac{1}{2}}).$$

Since B_n has the binomial distribution, $n^{-\frac{1}{2}}(B_n - np)$ is bounded away from 0 and ∞ in probability in the special sense that, given $\epsilon > 0$, there are positive $b_0(\epsilon)$ and $b_1(\epsilon)$ such that

$$P\{b_0(\epsilon) < n^{-\frac{1}{2}}|B_n - np| < b_1(\epsilon)\} \geqq 1 - \epsilon$$

for all n sufficiently large. Applying Assumption D, and the O_p calculus

$$\sum_{j=[np]+1}^{B_n} c_{jn} = (B_n - np)\psi_n(p)[1 + o_p(1)] = nQ_n[1 + o_p(1)].$$

Thus

$$R_n = Q_n o_p(1) = o_p[n^{-\frac{1}{2}}\psi_n(p)].$$

Lemma 1 applies trivially to Q_n and the rest of the corollary follows immediately.

Let us now review Theorem 1 and its proof. The following remarks are relevant to its effective utilization and comprehension.

REMARK 1. One may regard the theorem as having reduced the stochastic problem of the study of $\mathcal{L}(T_n)$ to the analytic problem of studying $g_{jn}(\epsilon)$, c_{jn} and α_{jn}.

REMARK 2. Assumption A is not mentioned in the proof of Theorem 1. It occurs only implicitly in the use of $\tilde{H}'$. However, in applications where $\tilde{H}'$ is not defined everywhere one may replace $\tilde{H}'$ by some other function as long as Assumptions B and C still hold in terms of the appropriately modified G and g. For example we could replace $\tilde{H}(\tilde{\nu}_{jn})$ by $\tilde{H}(\nu_{jn}) = H[j/(n+1)]$ and we could replace $\tilde{H}'(\tilde{\nu}_{jn})$ by $\tilde{H}'(\nu_{jn}) = [1 - j/(n+1)]H'[j/(n+1)]$ if H' exists at $j/(n+1)$ and by 0 otherwise. Then G_{jn} could be replaced by

$$\{[\tilde{H}(v) - \tilde{H}(\nu_{jn})]/(v - \nu_{jn})\} - \tilde{H}'(\nu_{jn}).$$

The proof of Theorem 1 then requires very little modification. One observes that

$$E(V_{jn} - \nu_{jn})^2 = E(V_{jn} - \tilde{\nu}_{jn})^2 + (\tilde{\nu}_{jn} - \nu_{jn})^2 = E[(V_{jn} - \tilde{\nu}_{jn})^2[1 + O(1)]$$

uniformly in j. One must also add the assumption that

$$\sum_{j=1}^n c_{jn}(\nu_{jn} - \tilde{\nu}_{jn})\tilde{H}'(\nu_{jn}) = o(n^{\frac{1}{2}}\sigma_n).$$

REMARK 3. In its present form Theorem 1 has an apparent asymmetry derived from the use of the exponential distribution. This asymmetry is more apparent than real and is essentially dispelled under the transformation $H(u) = \tilde{H}(v)$.

REMARK 4. Squaring and adding the α_{jn} given by (2.13) we have

$$(2.16) \qquad n\sigma_n^2 = \sum_{i,j=1}^n c_{in}c_{jn}\tilde{H}'(\tilde{\nu}_{in})\tilde{H}'(\tilde{\nu}_{jn})K_{ijn}$$

where $K_{ijn} = K_{jin} = \omega_{in} \approx i/(n+1)(n-i+1)$ for i/n bounded away from 1. Thus σ_n^2 is the analogue of a Riemann sum over the unit square.

REMARK 5. If σ_n in Assumption B is replaced by some other sequence τ_n, the proof of Theorem 1 yields $|R_n| = o_p(n^{-\frac{1}{2}}\tau_n)$.

REMARK 6. Suppose

$$(2.17) \qquad T_n^{(i)} = \mu_n^{(i)} + Q_n^{(i)} + R_n^{(i)}, \qquad 1 \leq i \leq r,$$

where

$$(2.18) \qquad Q_n^{(i)} = n^{-1} \sum_{j=1}^n \alpha_{jn}^{(i)}(Z_i - 1), \qquad 1 \leq i \leq r,$$

and $R_n^{(i)} = o_p(n^{-\frac{1}{2}}\tau_n)$. Let $[\sigma_n^{(i)}]^2 = n^{-1} \sum_{j=1}^n [\alpha_{jn}^{(i)}]^2$, $T_n = \sum_{i=1}^r T_n^{(i)}$, $\mu_n = \sum_{i=1}^r \mu_n^{(i)}$, $\alpha_{jn} = \sum_{i=1}^r \alpha_{jn}^{(i)}$, $\sigma_n^2 = n^{-1} \sum_{j=1}^n \alpha_{jn}^2$, and $\tau_n^2 = \sum_{i=1}^r [\sigma_n^{(i)}]^2$. With natural conditions on the $\alpha_{jn}^{(i)}$ and τ_n, one can easily establish the approximate normality of T_n and of the vector whose coordinates are $T_n^{(i)}$. In particular if,

$$\max_{1 \leq j \leq n} |\alpha_{jn}^{(i)}| = o_p(n^{\frac{1}{2}}\sigma_n^{(i)}), \qquad 1 \leq i \leq r,$$

and

$$\sigma_n^{(i)} = O(\sigma_n), \qquad 1 \leq i \leq r,$$

then

$$\mathcal{L}[n^{\frac{1}{2}}(T_n - \mu_n)/\sigma_n] \to \mathfrak{N}(0, 1).$$

Alternatively if

(2.19) $$\tau_{ij} = \lim_{n \to \infty} \{n^{-1}\tau_n^{-2} \sum_{k=1}^{n} \alpha_{kn}^{(i)} \alpha_{kn}^{(j)}\}$$

exists finite for all i, j, $1 \leq i, j \leq r$, then

$$\mathcal{L}\{n^{\frac{1}{2}}(T_n^{(1)} - \mu_n^{(1)})/\tau_n, \cdots, n^{\frac{1}{2}}(T_n^{(r)} - \mu_n^{(r)})/\tau_n\} \to \mathfrak{N}(0, \|\tau_{ij}\|).$$

The asymptotic normality of the sample percentile can be derived by a variety of standard methods. In view of Remark 6, our form of Corollary 1 facilitates the study of the joint asymptotic behavior of several percentiles and other linear functions of the order statistics.

Now we shall replace Theorem 1 by an alternative in which the assumptions are modified to a more useful form.

ASSUMPTION A*. $H(\cdot)$ is continuous on $(0, 1)$ and satisfies a first order Lipschitz condition[3] in every interval bounded away from 0 and 1. H' exists and is continuous except on a set S of Jordan content 0. (We shall replace H' by 0 wherever H' is not defined.)

ASSUMPTION B*. (Recalling the definition of $\mathfrak{M}$ (in (2.4))

$$\mathfrak{M}\{|c_{jn}H'[j/(n + 1)][j(n + 1)^{-1}(1 - j(n + 1)^{-1})]^{\frac{1}{2}}|\} = O(\sigma_n) \qquad \text{absolutely,}$$

$$\mathfrak{M}_\delta^{1-\delta}\{|c_{jn}|\} = O(\sigma_n)$$

absolutely for each $\delta > 0$.

ASSUMPTION E. (Tail Smoothness). There exists a δ_0, $0 < \delta_0 < 1$ such that (a) either $c_{jn} = 0$ for $j \leq n\delta_0$ for all sufficiently large n, or for each $K > 0$ there exists a finite M such that if $0 < u_1, u_2 < \delta_0$ and $K^{-1} < u_1/u_2 < K$ then $M^{-1} < H'(u_1)/H'(u_2) < M$, and (b) either $c_{jn} = 0$ for $j \geq n(1 - \delta_0)$ for all sufficiently large n or for each $K > 0$ there exists a finite M such that if $1 - \delta_0 < u_1, u_2 < 1$ and $K^{-1} < (1 - u_1)/(1 - u_2) < K$ then $M^{-1} < H'(u_1)/H'(u_2) < M$.

THEOREM 2. *If Assumptions* A*, B* *and* E *are satisfied, then the results of Theorem 1 apply with equations* (2.12) *and* (2.13) *replaced by*

(2.12′) $$\mu_n = n^{-1} \sum_{j=1}^{n} c_{jn}H[j/n + 1)]$$

and

(2.13′) $$\alpha_{jn} = (n - j + 1)^{-1} \sum_{i=j}^{n} c_{in}H'[i/(n + 1)][1 - i/(n + 1)].$$

PROOF. Recalling Remark 2, it suffices to prove the following two lemmas.

LEMMA 2. *Assumption* B* *implies Assumption* C *and* $\sum_{j=1}^{n} c_{jn}(\nu_{jn} - \tilde{\nu}_{jn})\tilde{H}'(\nu_{jn}) = o(n^{\frac{1}{2}}\sigma_n).$

---

[3] The function $H(\cdot)$ satisfies a first order Lipschitz condition on S if $(H(x) - H(y))/(x - y)$ is bounded for $x, y \, \varepsilon \, S$, $x \neq y$.

PROOF. Let

$$a_{jn} = |c_{jn}H'[j(n + 1)][j(n + 1)^{-1}(1 - j(n + 1)^{-1})]^{\frac{1}{2}}|.$$

Considering the behavior of $u(1 - u)$,

$$(n + 1)^{-1}\mathfrak{M}\{|c_{jn}H'[j/(n + 1)]|\}$$

$$\leqq n^{-\frac{1}{2}}[\mathfrak{M}_0{}^{\delta}\{a_{jn}\} + \mathfrak{M}_{1-\delta}^1\{a_{jn}\}] + (n + 1)^{-1}[\delta(1 - \delta)]^{-\frac{1}{2}}\mathfrak{M}_{\delta}{}^{1-\delta}\{a_{jn}\}.$$

Assumption B* implies that the right hand side is $o(\sigma_n/n^{\frac{1}{2}})$. But then for all j

$$|\alpha_{jn}| \leqq (n + 1)^{-1} \sum_{i=1}^n |c_{in}H'[i/(n + 1)]| = o(n^{\frac{1}{2}}\sigma_n)$$

and Assumption C is satisfied. Furthermore

$$\sum_{j=1}^n c_{jn}(\nu_{jn} - \tilde{\nu}_{jn})\tilde{H}'[j/(n + 1)]$$

$$= \sum_{j=1}^n c_{jn}(\nu_{jn} - \tilde{\nu}_{jn})H'[j/(n + 1)][1 - j/(n + 1)]$$

$$= O\{\sum_{j=1}^n |c_{jn}H'[j/(n + 1)]j/(n + 1)^2|\} = o(n^{\frac{1}{2}}\sigma_n),$$

which establishes the second assertion of the lemma.

LEMMA 3. *Assumption* A*, B* *and* E *imply Assumption* B.

PROOF. We modify $g_{jn}(\epsilon)$ according to Remark 2. In view of the assumptions and Proposition 4, it is possible for each $\epsilon > 0, 0 < \delta_1 < \delta_0 , \delta_2 > 0$ to decompose $(0, 1)$ into three non overlapping sets $S_1 , S_2 ,$ and S_3 with the following properties: $S_1 = (0, \delta_1) \cup (1 - \delta_1, 1), H'(u)$ is defined if $u = 1 - e^{-v} \varepsilon S_1$ unless the corresponding c_{jn} are zero, and if$(j/n) \varepsilon S_1$,

$$|g_{jn}(\epsilon)| \leqq \sup_{v_{jn}(\epsilon) \leqq v \leqq v^{jn}(\epsilon)} |\tilde{H}'(v) - \tilde{H}'(\nu_{jn})|$$

$$\leqq [M_1(\epsilon) + 1] |H'[j/(n + 1)][1 - j/(n + 1)]|$$

for n sufficiently large, by Assumption E. The set S_2 is a finite union of open intervals, with total measure less than δ_2, which contains $[u_{jn}(\epsilon), u^{jn}(\epsilon)]$ if $[u_{jn}(\epsilon), u^{jn}(\epsilon)]$ contains a point of S when n is sufficiently large. If $(j/n) \varepsilon S_2$, $|g_{jn}(\epsilon)| \leqq M_2(\delta_1)$ which is determined by the Lipschitz condition bound for u in $[\delta_1/2, 1 - \delta_1/2]$. On the remaining set S_3, $H'(u)$ exists and is uniformly continuous, and $\max_{(j/n)\varepsilon S_3} g_{jn}(\epsilon) = o(1)$ as $n \to \infty$.

Let

$$b_{jn} = |c_{jn}g_{jn}(\epsilon)[j/(n - j + 1)]^{\frac{1}{2}}|.$$

Applying Assumption B*, we have,

$$\sum_{(j/n)\varepsilon S_1} b_{jn} \leqq [M_1(\epsilon) + 1]M_3(\delta_1, n)(n + 1)\sigma_n$$

where $M_3 \to 0$ as $\delta_1 \to 0, \quad n \to \infty,$

$$\sum_{(j/n)\varepsilon S_2} b_{jn} \leqq M_2(\delta_1)M_4(\delta_1, \delta_2, n)[(1 - \delta_1)/\delta_1]^{\frac{1}{2}}n\sigma_n$$

where $M_4 \to 0$ as $\delta_2 \to 0, \quad n \to \infty,$

$\sum_{(j/n)\,\epsilon S_3} b_{nj} \leqq M_5(n,\,\delta_1\,,\,\delta_2)[(1-\delta_1)/\delta_1]^{\frac{1}{2}}n\sigma_n$

where $M_5 \to 0$ as $n \to \infty$.

Given $\epsilon_1 > 0$, we select δ_1, δ_2 and N in order, so as to obtain $\mathfrak{M}\{b_{jn}\} \leqq \epsilon_1\sigma_n$ for $n \geqq N$. But this is the desired result.

REMARK 7. The version of this theorem using $\tilde{H}'(\tilde{\nu}_{jn})$ in place of $H'[j/(n+1)]$ requires less effort. The present version has some minor advantage in appearance which facilitates applications where it is most natural to deal with $H(u)$ rather than $\tilde{H}(v)$.

We now turn to the major source of applications. This is the case of $c_{jn} = J[j/(n+1)]$. Let

(2.20) $\mu = \int_0^1 J(u)H(u)\,du$,

(2.21) $\alpha(u) = (1-u)^{-1} \int_u^1 J(w)H'(w)(1-w)\,dw$,

(2.22) $\sigma^2 = \int_0^1 \alpha^2(u)\,du = \int_0^1 \int_0^1 J(u)H'(u)J(w)H'(w)K(u,\,w)\,du\,dw$,

where

(2.23) $K(u,\,w) = K(w,\,u) = u(1-w)$ for $0 \leqq u \leqq w \leqq 1$,

(2.24) $B_1 = \int_0^1 J(u)H'(u)[u(1-u)]^{\frac{1}{2}}\,du$,

(2.25) $B_2 = \int_{\delta}^{1-\delta} J(u)\,du$, $0 < \delta < \frac{1}{2}$.

These expressions are the continuous analogues of those of Theorem 2 and (2.16), i.e.,

(2.12″) $\mu_n = \mathfrak{M}\{J[j/(n+1)]H[j/(n+1)]\}$,

(2.13″) $\alpha_{jn} = (n-j+1)^{-1} \sum_{i=j}^n J[i/(n+1)]H'[i/(n+1)][1-i/(n+1)]$,

(2.14′) $\sigma_n^2 = \mathfrak{M}\{\alpha_{jn}^2\} = \mathfrak{M}\{J[i/(n+1)]H'[i/(n+1)]J[j/(n+1)]$

$\cdot H'[j/(n+1)]K[i/(n+1),\,j/(n+1)]\}$,

and those appearing in Assumption B*. (The last expression for σ_n^2 corresponds to the average over the square. See Remark 4.) We now introduce

ASSUMPTION B**. The Riemann integrals B_1, B_2 converge absolutely.

Let $Q_n = n^{-1} \sum_{j=1}^n \alpha_{jn}(Z_j - 1)$ as before and let

(2.26) $Q_n^* = n^{-1} \sum_{j=1}^n \alpha[j/(n+1)](Z_j - 1)$.

THEOREM 3. (i) *Under Assumptions* A*, B** *and* E, *if* σ_n^2 *is bounded away from* 0, *then*

$$T_n = \mu_n + Q_n + o_p(n^{-\frac{1}{2}}),$$

with $\mathfrak{L}\{n^{\frac{1}{2}}(T_n - \mu_n)/\sigma_n\} \to \mathfrak{N}(0,\,1)$.

(ii) *If also, the double Riemann integral* (2.22) *for* $\sigma^2 > 0$ *converges absolutely,*

then

$$Q_n = Q_n{}^* + o_p(n^{-\frac{1}{2}}) \quad and \quad \sigma_n{}^2 \to \sigma^2 = \int_0^1 \alpha^2(u) \, du,$$

so that σ_n can be replaced by σ in part (i) and $\mathcal{L}(n^{\frac{1}{2}}Q_n{}^*/\sigma) \to \mathfrak{N}(0, 1)$.

(iii) *If $\mu_n = \mu + o(n^{-\frac{1}{2}})$, then μ_n can be replaced by μ in part* (i).

PROOF. The first part follows since Assumption B^{**} implies B^*. For the second part, the convergence of B_1 implies that of the integral for $(1 - u)\alpha(u)$ and also the continuity of $\alpha(u)$. It follows that $\alpha_{jn} - \alpha[j/(n + 1)] = o(1)$ uniformly in any interval bounded away from 0 and 1. The absolute convergence of the double Riemann integral for σ^2 implies, with Remark 4, that

$$n^{-1} \sum_{j=1}^n \alpha_{jn}^2 = \sigma_n{}^2 \to \sigma^2 = \int_0^1 \alpha^2(u) \, du = \lim_{n\to\infty} n^{-1} \sum_{j=1}^n \alpha^2[j/(n + 1)].$$

Thus the tails contribute "little" to $\mathfrak{M}\{\alpha^2[j/(n + 1)]\}$, the middle contributes little to $\mathfrak{M}\{\alpha_{jn}^2 - \alpha^2[j/(n + 1)]\}$ and hence the tails contribute little to $\mathfrak{M}\{\alpha_{jn}^2\}$. Hence $\mathfrak{M}\{[\alpha_{jn} - \alpha[j/(n + 1)]]^2\} = o(1)$ and $Q_n - Q_n{}^* = n^{-1} \sum [\alpha_{jn} - \alpha[j/(n + 1)]](Z_j - 1) = o_p(n^{-\frac{1}{2}})$. The rest is trivial.

A simple application of Corollary 1, Theorem 3 and Remark 6 yields the following corollary which corresponds to the case where a finite number of sample percentiles are given special weight.

COROLLARY 3. *If*

$$T_n = n^{-1} \sum_{j=1}^n J[j/(n + 1)]h(X_{jn}) + \sum_{i=1}^r a_i h(X_{[np_i],n})$$

where J and h satisfy the conditions of Theorem 3, $0 < p_1 < p_2 < \cdots < p_r < 1$, F' and h' exist and are finite and F' is non zero at the corresponding population percentiles $\lambda_1, \lambda_2, \cdots, \lambda_r$, then

$$\mathcal{L}\{n^{\frac{1}{2}}(T_n - \mu)/\sigma\} \to \mathfrak{N}(0, 1),$$

where

(2.27) $$\mu = \int_0^1 J(u)H(u) \, du + \sum_{i=1}^r a_i h(\lambda_i),$$

(2.28) $$\sigma^2 = \int_0^1 \alpha^2(u) \, du,$$

and

(2.29) $$\alpha(u) = (1 - u)^{-1}\{\int_u^1 J(w)H'(w)(1 - w) \, dw$$
$$+ \sum_{p_i \geq u} a_i(1 - p_i)H'(p_i)\}.$$

REMARK 8. If H has a finite number of jump discontinuities, it may be decomposed into the form $H(u) = H_0(u) + \sum a_i \Delta_{p_i}(u)$ where H_0 is continuous and Theorems 1, 2, or 3 can be combined with Corollary 2. An alternative is to show that the results of using a smooth approximation to H approximates that of using H. This can be done by studying the effect of functions $\delta_p(v) - \Delta_p(v)$ where δ_p is a smooth function rising rapidly from 0 to 1. Thus we see that Theorem 3 may be regarded as applicable if H has a finite number of jump discontinuities and the integrals for α and σ^2 in (2.21), (2.22) are replaced by Stieltjes

integrals. This comment leads naturally to a conjecture that the "natural" theorem would involve the condition that H is of bounded variation.

REMARK 9. If Theorem 3 applies to $T_n^{(i)}$ obtained from J_i and H_i for $1 \leq i \leq r$, then it applies to the vector $(T_n^{(1)}, T_n^{(2)}, \cdots, T_n^{(r)})$. Here the relevant covariance matrix is $\|\sigma_{ij}\|$ where

$$(2.30) \quad \sigma_{ij} = \int_0^1 \alpha_i(u)\alpha_j(u) \ du$$

$$= \int_0^1 \int_0^1 J_i(u)H_i'(u)J_j(w)H_j'(w)K(u, w) \ du \ dw,$$

where $\alpha_i(u)$ is obtained by inserting subscripts in (2.21).

REMARK 10. Let $A_{jn} = n^{-1} \sum_{i=1}^j a_{in}$ and $B_{jn} = n^{-1} \sum_{i=1}^j b_{in}$. Then

$$(2.31) \quad n^{-2} \sum_{i,j=1}^n a_{in}b_{jn}K[i/(n+1), j/(n+1)] = (n+1)^{-1} \sum_{j=1}^n A_{jn}B_{jn}$$

$$- [(n+1)^{-1} \sum_{j=1}^n A_{jn}][(n+1)^{-1} \sum_{j=1}^n B_{jn}]$$

where the right-hand side resembles a sample covariance. Analogously, if we introduce the indefinite integrals $A(u) = \int a(u) \ du$ and $B(u) = \int b(u) \ du$, it can be seen that

$$(2.32) \quad \int_0^1 \int_0^1 a(u)b(w)K(u, w) \ du \ dw$$

$$= \int_0^1 A(u)B(u) \ du - [\int_0^1 A(u) \ du][\int_0^1 B(u) \ du].$$

Note that adding constants to $A(u)$ and $B(u)$ does not affect the right hand side of equation (2.32) which is valid even if $A(u)$ and/or $B(u)$ diverge as u approaches zero or one so long as the Lebesgue integrals on either side converge absolutely. The connection of (2.31) and (2.32) with (2.22), (2.14′) and (2.30) is evident. Hence we have

COROLLARY 4. *Let*

$$(2.33) \quad A(u) = \int J(u)H'(u) \ du, \qquad A_i(u) = \int J_i(u)H_i'(u) \ du.$$

Then (2.22) *and* (2.30) *become*

$$(2.22') \quad \sigma^2 = \int_0^1 \alpha^2(u) \ du = \text{Var}\ \{A(U)\} = \text{Var}\ \{A[F(X)]\},$$

$$(2.30') \quad \sigma_{ij} = \int_0^1 \alpha_i(u)\alpha_j(u) \ du = \text{Covar}\ \{A_i(U)A_j(U)\}$$

$$= \text{Covar}\ \{A_i[F(X)], A_j[F(X)]\},$$

where X has cdf F and U is uniformly distributed on $(0, 1)$. *Moreover, these results apply to the case covered in Corollary 3 if $J(u)$ is treated as a delta function, i.e., if $A(u)$ has a jump of $a_iH'(p_i)$ at $u = p_i$.*

3. Applications to asymptotically efficient linear estimation for the uncensored and multicensored cases.

In this section we apply the results of Section 2 to demonstrate the asymptotic efficiency of Bennett's estimates of location and scale parameters in the uncensored and multicensored cases. Throughout this section we shall assume that the standard conditions for the validity of the Cramér-Rao bounds are satisfied, and that the related integrations by parts are

permitted, and also that the conditions of Corollaries 3 and 4 are satisfied. To indicate asymptotic efficiency we shall demonstrate that the asymptotic covariance matrix of our estimates coincides with the inverse of the Fisher information matrix.

Let $X_{1n} \leqq X_{2n} \leqq \cdots \leqq X_{nn}$ be the ordered observations from a distribution of the form

$$F(x; \theta_1, \theta_2) = F[(x - \theta_1)/\theta_2]$$

with density

$$f(x; \theta_1, \theta_2) = \theta_2^{-1} f[(x - \theta_1)/\theta_2].$$

We consider estimates of the form

$$(3.1) \qquad T_n = n^{-1} \sum J[j/(n + 1)] X_{jn}.$$

Hence in this section $H(u)$ will always be the inverse function of $F(x; \theta_1, \theta_2)$. That is

$$(3.2) \qquad H(u) = \theta_1 + \theta_2 F^{-1}(u)$$

and hence $H'(u)\, du = \theta_2\, dy$ and $du = f(y)\, dy$ where $y = F^{-1}(u)$.

We shall begin with the uncensored case, and there we shall first treat the situations where only one of the parameters is unknown.

Uncensored case (complete sample). The Fisher information matrix[4] is defined by

$$\| E\{ [\partial \log f(X; \theta_1, \theta_2)/\partial\theta_i] \cdot [\partial \log f(X; \theta_1, \theta_2)/\partial\theta_j] \} \|$$

and under conditions, to be discussed shortly, which permit integration by parts, it is equal to

$$\| E\{ - \partial^2 \log f(X; \theta_1, \theta_2)/\partial\theta_i \partial\theta_j \} \|.$$

Let

$$(3.3) \qquad L_1(y) = -f'(y)/f(y),$$

and

$$(3.4) \qquad L_2(y) = -\{ 1 + y f'(y)/f(y) \}.$$

Then

$$\partial \log f(x; \theta_1, \theta_2)/\partial\theta_1 = \theta_2^{-1} L_1(y), \qquad y = (x - \theta_1)/\theta_2,$$

$$\partial \log f(x; \theta_1, \theta_2)/\partial\theta_2 = \theta_2^{-1} L_2(y),$$

$$\int_{-\infty}^{\infty} L_1(y) f(y)\, dy = \int_{-\infty}^{\infty} L_2(y) f(y)\, dy = 0.$$

---

[4] To avoid confusion we note that I represents an information matrix and not the identity matrix.

The Fisher information matrix is given by $\theta_2^{-2}I$ where

$$(3.5) \qquad I = \| \int_{-\infty}^{\infty} L_i(y)L_j(y)f(y) \; dy \|,$$

$$(3.5') \qquad I = \left\| \begin{matrix} \int_{-\infty}^{\infty} L_1'(y)f(y) \; dy & \int_{-\infty}^{\infty} L_2'(y)f(y) \; dy \\[2mm] \int_{-\infty}^{\infty} yL_1'(y)f(y) \; dy & \int_{-\infty}^{\infty} yL_2'(y)f(y) \; dy \end{matrix} \right\|,$$

and $I_{12} = I_{21}$. To insure these relations an elementary argument shows that it suffices to have the existence of $f''(y)$ and

$$(3.6) \qquad y^2 f'(y) \to 0 \quad \text{as} \quad y \to \pm \infty.$$

(This condition can be reduced to $yf'(y) \to 0$ as $y \to \pm \infty$ unless we use I_{22} which is not required in the estimation of θ_1 when θ_2 is known.)

   Location parameter: (scale parameter known). The estimator T_{1n} is determined as in (3.1) by the function

$$(3.7) \qquad J_1(u) = I_{11}^{-1}L_1'(y), \qquad y = F^{-1}(u).$$

Bennett's estimator is T_{1n} minus the known bias correction $I_{11}^{-1}I_{12}\theta_2$. Applying Corollary 3 we have for T_{1n},

$$(3.8) \qquad \mu_1 = \int_0^1 J_1(u)H(u) \; du = \int_{-\infty}^{\infty} I_{11}^{-1}L_1'(y)[\theta_1 + \theta_2 y]f(y) \; dy$$

$$= \theta_1 + I_{11}^{-1}I_{12}\theta_2 .$$

Next

$$A_1(u) = \int J_1(u)H'(u) \; du = \theta_2 I_{11}^{-1} \int L_1'(y) \; dy = \theta_2 I_{11}^{-1}L_1(y).$$

Since $A_1(U) = \theta_2 I_{11}^{-1}L_1(Y)$ where Y has cdf $F(y)$, $A_1(U)$ has variance $\theta_2^2 I_{11}^{-2}I_{11}$ and

$$(3.9) \qquad \sigma^2 = \theta_2^2 I_{11}^{-1} .$$

Combining (3.8) and (3.9) we see that $T_{1n} - I_{11}^{-1}I_{12}\theta_2$ *is an asymptotically efficient estimator of the location parameter* θ_1.

   Scale parameter: (location parameter known). The estimator T_{2n} is based on the function

$$(3.10) \qquad J_2(u) = I_{22}^{-1}L_2'(y), \qquad y = F^{-1}(u).$$

Bennett's estimator is T_{2n} minus the known bias correction $I_{22}^{-1}I_{12}\theta_1$. For the estimator T_{2n} we have

$$(3.11) \qquad \mu_2 = \int_0^1 J_2(u)H(u) \; du = I_{22}^{-1}I_{12}\theta_1 + \theta_2 ,$$

$$A_2(u) = \int J_2(u)H'(u) \; du = \theta_2 I_{22}^{-1}L_2(y)$$

and $A_2(U)$ has variance

$$(3.12) \qquad \sigma_2^2 = \theta_2^2 I_{22}^{-1}$$

which establishes the asymptotic efficiency of $T_{2n} - I_{22}^{-1}I_{12}\theta_1$ as an estimator of the scale parameter θ_2.

Both parameters unknown. Here Bennett's estimators (T_{3n}, T_{4n}) are based on (3.1) with

$$(3.13) \qquad \{J_3(u), J_4(u)\} = \{L_1'(y), L_2'(y)\}\Gamma^{-1}, \qquad y = F^{-1}(u).$$

Then the mean vector (μ_3, μ_4) is obtained by integrating each term of

$$(\theta_1, \theta_2)\begin{pmatrix}1 \\ y\end{pmatrix}\{L_1'(y), L_2'(y)\}\Gamma^{-1}$$

yielding

$$(3.14) \qquad\qquad (\mu_3, \mu_4) = (\theta_1, \theta_2).$$

Furthermore, $\{A_3(u), A_4(u)\} = \theta_2\{L_1(y), L_2(y)\}\Gamma^{-1}$ and $\{A_3(U), A_4(U)\}$ has the covariance matrix

$$(3.15) \qquad\qquad \|\sigma_{ij}\| = \theta_2^2 I^{-1},$$

establishing the asymptotic optimality of the estimator (T_{3n}, T_{4n}).

Censored Case. We now study the case of multiple Type II censoring where the observations which lie in the sample percentile ranges $(p_1, p^{(1)})$, $(p_2, p^{(2)}), \cdots, (p_r, p^{(r)})$, $0 \leq p_1 < p^{(1)} < p_2 < p^{(2)} < \cdots < p_r < p^{(r)} \leq 1$, are censored (i.e., unavailable). Thus, for each i, the observations lying between the sample $p^{(i)}$-percentile and the sample p_{i+1}-percentile (inclusive) are available. If $p_1 = 0$ [$p_r = 1$], then the smallest [largest] segment of observations is censored. We first present the information matrix for a related problem involving partially grouped data (see Kulldorff [6]). Here we start with $-\infty \leq \xi_1 < \xi^{(1)} < \cdots < \xi_r < \xi^{(r)} \leq \infty$. If an observation falls in one of the intervals $(\xi_i, \xi^{(i)})$, the interval in which it falls is identified. Otherwise, the observation itself is recorded. To exhibit the information matrix, let

$$(3.16) \qquad \begin{aligned} \lambda_i &= (\xi_i - \theta_1)/\theta_2, & \lambda^{(i)} &= (\xi^{(i)} - \theta_1)/\theta_2, \\ f_i &= f(\lambda_i), & f^{(i)} &= f(\lambda^{(i)}), \\ F_i &= F(\lambda_i), & F^{(i)} &= F(\lambda^{(i)}), \end{aligned}$$

where f_i and $f^{(i)}$ are taken to be zero when $F_i = 0$ or $F^{(i)} = 1$ respectively. Let E be the complement of the union of the intervals $(\lambda_i, \lambda^{(i)})$. For this problem, the information matrix is readily seen to be $\theta_2^{-2}I$ where

$$I_{11} = \int_E L_1^2(y)f(y)\,dy + \sum_{i=1}^r [f^{(i)} - f_i]^2/(F^{(i)} - F_i);$$
$$(3.17) \quad I_{12} = \int_E L_1(y)L_2(y)f(y)\,dy$$
$$\qquad\qquad + \sum_{i=1}^r (f^{(i)} - f_i)(\lambda^{(i)}f^{(i)} - \lambda_i f_i)/(F^{(i)} - F_i);$$
$$I_{22} = \int_E L_2^2(y)f(y)\,dy + \sum_{i=1}^r (\lambda^{(i)}f^{(i)} - \lambda_i f_i)^2/(F^{(i)} - F_i).$$

It can be seen that when λ_i, $\lambda^{(i)}$ are such that $F_i = p_i$ and $F^{(i)} = p^{(i)}$, the in-

formation in the partially grouped data matches that for the case of multiple Type II censoring and $\theta_2^2 I^{-1}$ corresponds to the Cramér-Rao bound for this case as well.

For this case it will be convenient to estimate (θ_1, θ_2) by $(T_{5n}^*, T_{6n}^*) = (T_{5n}, T_{6n})I^{-1}$, where T_{5n} and T_{6n} are determined by weight functions $J(u)$ and by additional discrete weights given to certain sample percentiles as treated in Corollary 3. The "continuous" part is determined by

$$(3.18) \qquad \{J_5(u), J_6(u)\} = \{L_1'(y), L_2'(y)\}, \qquad y = F^{-1}(u) \ \varepsilon \ E$$
$$= 0 \qquad\qquad \text{elsewhere.}$$

The "discrete" contribution attaches additional weights

$$(a_{5i}, a_{6i}) = f_i\{- L_1(\lambda_i) - (f^{(i)} - f_i)/(p^{(i)} - p_i),$$
$$(3.19) \qquad\qquad - L_2(\lambda_i) - (\lambda^{(i)}f^{(i)} - \lambda_i f_i)/(p^{(i)} - p_i)\},$$
$$(a_5^{(i)}, a_6^{(i)}) = f^{(i)}\{L_1(\lambda^{(i)}) + (f^{(i)} - f_i)/(p^{(i)} - p_i),$$
$$L_2(\lambda^{(i)}) + (\lambda^{(i)}f^{(i)} - \lambda_i f_i)/(p^{(i)} - p_i)\}$$

to the p_i and $p^{(i)}$ sample percentiles, respectively.

To compute the asymptotic mean corresponding to (T_{5n}, T_{6n}) we apply Corollary 3. Noting that

$$\int_a^b L_1'(y)f(y) \, dy = f(y)L_1(y)|_a^b + \int_a^b L_1^2(y)f(y) \, dy,$$
$$\int_a^b L_1'(y)yf(y) \, dy = yf(y)L_1(y)|_a^b + \int_a^b L_1(y)L_2(y)f(y)dy,$$
$$\int_a^b L_2'(y)f(y) \, dy = f(y)L_2(y)|_a^b + \int_a^b L_1(y)L_2(y)f(y) \, dy,$$
$$\int_a^b L_2'(y)yf(y) \, dy = yf(y)L_2(y)|_a^b + \int_a^b L_2^2(y)f(y) \, dy$$

a direct computation gives $(\mu_5, \mu_6) = (\theta_1, \theta_2)I$, so that (T_{5n}^*, T_{6n}^*) is asymptotically unbiased.

Applying Corollary 4, we interpret $A_5(u) = \int J_5(u)H'(u) \, du = \theta_2 \int J_5(u) \, dy$, so that it takes jumps of $\theta_2 a_{5i}/f_i$ and $\theta_2 a_5^{(i)}/f^{(i)}$ at $u = p_i$ and $u = p^{(i)}$, respectively. In this way we have

$$A_5(u) = \theta_2 L_1(y), \qquad\qquad\qquad y = F^{-1}(u) \ \varepsilon \ E,$$
$$= -\theta_2(f^{(i)} - f_i)/(p^{(i)} - p_i), \qquad \lambda_i < y < \lambda^{(i)},$$
$$A_6(u) = \theta_2 L_2(y), \qquad\qquad\qquad y \ \varepsilon \ E,$$
$$= -\theta_2(\lambda^{(i)}f^{(i)} - \lambda_i f_i)/(p^{(i)} - p_i), \qquad \lambda_i < y < \lambda^{(i)}.$$

But then $[A_5(U), A_6(U)]$ has covariance matrix $\theta_2^2 I$. It follows that

$$(T_{5n}^*, T_{6n}^*) = (T_{5n}, T_{6n})I^{-1}$$

is an asymptotically efficient estimator of (θ_1, θ_2).

Remarks and examples. If F is symmetric and we have uncensored or two-

sided symmetrically censored samples, then $I_{12} = 0$ and the efficient estimate of θ_1 is the same regardless of whether or not θ_2 is known. The same is true for the optimum estimate of θ_2. Some examples of efficient estimates from complete samples are given below. In each of these the required regularity assumptions are easily verified.

EXAMPLE 1. The efficient estimate of θ_1 for the logistic distribution,

$$F(x; \theta_1, \theta_2) = [1 + \exp\{-(x - \theta_1)/\theta_2\}]^{-1}$$

is determined by the weight function $J_1(u) = 6u(1 - u)$. The efficient estimator of θ_2 is determined by

$$J_2(u) = 9(\pi^2 + 3)^{-1}\{2u - 1 + 2u(1 - u) \log [u/(1 - u)]\}.$$

EXAMPLE 2. The efficient estimate of θ_1 for the Cauchy distribution,

$$F(x; \theta_1, \theta_2) = \pi^{-1}\{\tan^{-1}[(x - \theta_1)/\theta_2] + \pi/2\}$$

is determined by the weight function

$$J_1(u) = \sin 4\pi(u - \tfrac{1}{2})/\tan \pi(u - \tfrac{1}{2}).$$

In this example it is interesting to note that $J_1(u)$ is negative for $(u - \tfrac{1}{2}) > \tfrac{1}{4}$.

The efficient estimate of θ_2 for the Cauchy distribution is specified by the weight function

$$J_2(u) = 8 \tan \pi(u - \tfrac{1}{2})/\sec^4 \pi(u - \tfrac{1}{2}).$$

The formulas for the t-distribution with three or more degrees of freedom were obtained by Jung. The preceding formulas are consistent with his.

EXAMPLE 3. The efficient estimator of θ_2 for the normal distribution,

$$F(x; \theta_1, \theta_2) = \Phi[(x - \theta_1)/\theta_2],$$

where

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\, dt = \int_{-\infty}^{x} (2\pi)^{-\frac{1}{2}} e^{-t^2/2}\, dt$$

is determined by the inverse function

$$J_2(u) = \Phi^{-1}(u).$$

In the normal case, of course, the efficient estimator θ_1 is obtained by setting $J_1(u) = 1$.

Since one of the authors [4] has already applied these results to one-sided censoring we shall discuss estimating the location parameter of a normal distribution in the case of symmetric two-sided censoring. If upper and lower $100p$ percent of the observations are censored, then the asymptotically optimum estimate is formed by using weights specified by $J(u) = I_{11}^{-1} L_1'(y)$ for the uncensored observations. The largest and smallest available order statistics are given the additional weight,

$$a = I_{11}^{-1}\{f^2(\lambda_p)/p - f'(\lambda_p)\}, \qquad F(\lambda_p) = p.$$

For the normal problem, $\Phi(\lambda_p) = p$,

$$I_{11} = \int_{\lambda_p}^{-\lambda_p} y^2 \varphi(y)\, dy + 2\varphi^2(\lambda_p)/p$$

$$= 1 - 2p + 2\lambda_p\varphi(\lambda_p) + 2\varphi^2(\lambda_p)/p,$$

while $J(u) = I_{11}^{-1}$, $p < u < 1 - p$, and

$$a = I_{11}^{-1}\{\varphi^2(\lambda_p)/p + \lambda_p\varphi(\lambda_p)\}.$$

When $p = .05$, $I_{11} = .986$ and $a = .0437$. When $p = .10$, $I_{11} = .966$, and $a = .086$. For estimating the mean of a normal distribution in the uncensored case, the optimal weight function is $J(u) = 1$. In this censored case the optimum estimate is

$$T_n = a[X_{ln} + X_{sn}] + n^{-1}I_{11}^{-1} \sum_{j=s}^{l} X_{jn}$$

where s and l denote the indices of the smallest and largest available order statistics.

The extra weight given to each of the extreme available order statistics is only slightly less than that suggested by Winsor and advocated by Tukey [12]. This fact is of theoretical interest as the asymptotically optimum estimate given above is slightly better than the Winsorized mean for normal data and may well be slightly less sensitive to contamination. Finally for $p = .05$ and .10, the asymptotically optimum weights given here are quite close (within 1.5 percent) of the weights given for sample size 20 by Sarhan-Greenberg [10], p. 248. The latter weights are those which provide the minimum variance unbiased estimate based on linear combinations of order statistics in the censored sample.

REFERENCES

[1] BENNETT, C. A. (1952). *Asymptotic properties of ideal linear estimators.* Unpublished Dissertation, Univ. of Michigan.

[2] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29** 972–994.

[3] GOVINDARAJULU, Z. (1965). Asymptotic normality of linear functions of order statistics in one and multi-samples. Technical Report. U.S. Air Force Grant No. AF-AFOSR-741-65, 1–26.

[4] JOHNS, M. V. and LIEBERMAN, J. G. (1966) An exact asymptotically efficient confidence bound for the reliability in the case of the Weibull distribution. *Technometrics* **8** 135–175.

[5] JUNG, JAN (1955). On linear estimates defined by a continuous weight function. *Arkiv. für Matematik* Bd **3** nr **15** 199–209.

[6] KULLDORFF, GUNNAR (1961). *Estimation From Grouped and Partially Grouped Samples.* Almquist & Wiksell, Stockholm.

[7] LECAM, LUCIEN: Some theorems on the distribution of sums of independent linear stochastic processes. Unpublished.

[8] MOSTELLER, F. (1946). On some useful inefficient statistics. *Ann. Math. Statist.* **17** 377–408.

[9] PLACKETT, R. L. (1958). Linear estimation from censored data. *Ann. Math. Statist.* **29** 131–142.

[10] SARHAN, A. E. and GREENBERG, B. G. (1962). *Contributions to Order Statistics*. Wiley, New York.

[11] SARKADI, K., SCHNELL, E., and VINCZE, I. (1962). On the position of the sample mean among the ordered sample elements. *Publ. Math. Inst. of Hung. Acad. Sci. A*, **7** 239–254.

[12] TUKEY, J. W. (1962). The future of data analysis. *Ann. Math. Statist.* **33** 1–67.

[13] WEISS, L. (1963). On the asymptotic distribution of an estimate of a scale parameter. *Nav. Res. Logistics Quart.* **10** 1–11.