ON MINIMUM VARIANCE IN NONREGULAR ESTIMATION

By R. C. Davis

U. S. Naval Ordnance Test Station, China Lake, California

1. Summary. A case of nonregular estimation arises in attempting to estimate a single unknown parameter, θ, in the probability distribution of a single chance variable in which one or both of the extremities of the range of the distribution are functions of the unknown parameter. The case treated in this paper is the one in which a probability density of exponential type exists. When one extremity alone of the range depends non-trivially upon θ, a necessary and sufficient condition is given in order that a single order statistic be a sufficient statistic for θ. In this case conditions are given for the existence of a unique unbiased estimate of θ possessing minimum variance uniformly in θ. In the case in which both extremities of the range depend upon θ, a necessary and sufficient condition is given that the smallest and largest order statistics constitute a set of sufficient statistics for θ. In this case Pitman [1] has shown that a single sufficient statistic exists if one extremity of the range is a monotone decreasing function of the other extremity.[1] It is shown that under the above condition a unique unbiased estimate exists possessing minimum variance. Moreover a surmise of Pitman is proved that only under this condition does a single sufficient statistic exist. When a single sufficient statistic does not exist, an unbiased estimate of a known function of θ is obtained which has less variance than any analytic function of the set of sufficient statistics for θ.

2. Introduction. Let X be a chance variable assuming values x in a one-dimensional Euclidean space, R_1, and let X possess a probability density function $f(x, \theta)$ depending on a single unknown parameter θ which lies in Ω, a subset of R_1. Denote by $a(\theta)$ and $b(\theta)$ the lower and upper extremities of the range of $f(x, \theta)$. We treat the cases in which either one or both the extremities of the range depend nontrivially upon θ. For each $\theta \in \Omega$ denote by $R^*(\theta)$ the subset of R_1 satisfying $a(\theta) \leqq x \leqq b(\theta)$, and by $R^{**}(\theta)$ the complement of $R^*(\theta)$ in R_1. We make the following assumptions:

Assumption A.

$$f(x, \theta) = 0 \quad \textit{for all} \quad (x, \theta) \textit{ on } R^{**}(\theta) \times \Omega;$$
$$f(x, \theta) = e^{\theta K(x) + S(x) + T(\theta)} \quad \textit{for all} \quad (x, \theta) \textit{ on } R^*(\theta) \times \Omega,$$

where $T(\theta)$ is a real single-valued continuous function of θ at all points of Ω, and $S(x)$, $K(x)$ are real single-valued continuous functions of x defined almost everywhere on R_1.

Assumption B. *$a(\theta)$ and $b(\theta)$ are continuous functions of θ satisfying for all $\theta \in \Omega$ the inequality $a(\theta) \leqq b(\theta)$.*

[1] The author is deeply indebted to the referee for bringing to his attention the paper by Pitman and for many other helpful suggestions.

The exponential type of frequency function assumed above is the type which Koopman [2] has shown to hold whenever a sufficient statistic for θ exists. We do not require any of his results, however, in this paper.

For convenience in notation we write $P(x) = e^{S(x)}$ and $Q(\theta) = e^{T(\theta)}$, so that obviously we have the relation

$$[Q(\theta)]^{-1} = \int_{a(\theta)}^{b(\theta)} P(\eta)\, d\eta.$$

Furthermore if an estimate of θ is a continuous function of n independent sample values, is unbiased, and possesses minimum variance uniformly in $\theta \in \Omega$, we term this a best estimate of θ.

3. One extremity of the range depending upon θ. First we treat the case in which only one extremity of the range depends upon the unknown parameter θ. To fix the argument we assume that the upper extremity $b(\theta)$ depends upon θ, and the lower extremity is independent of θ. The results of this section are extended in an obvious manner to the case in which the lower extremity alone depends upon θ.

THEOREM 1. *Let $x_1, x_2, \cdots, x_n$ be the values of n independent drawings from a population having the probability density function $f(x, \theta)$ satisfying Assumptions A and B, and in which the upper extremity only of the range depends upon θ. The necessary and sufficient condition that the* nth *order statistic, denoted by $x_{(n)}$, be a sufficient statistic for θ is that*

$$f(x, \theta) = P(x)\, Q(\theta) \quad \textit{for all} \quad (x, \theta) \text{ in } R^*(\theta) \times \Omega.$$

PROOF OF NECESSITY. Suppose that in a sample of n independent observations that the nth order statistic, $x_{(n)}$, is a sufficient statistic for θ. It follows from the definition of sufficiency that

$$f(x_1, \theta) \cdots f(x_n, \theta) = g(x_{(n)}; \theta)\, h(x_{(1)}, \cdots, x_{(n-1)} \mid x_{(n)}; \theta),$$

where $g(x_{(n)}, \theta)$ is the frequency function of $x_{(n)}$, and $h(x_{(1)}, \cdots, x_{(n-1)} \mid x_{(n)}; \theta)$ denotes the conditional frequency function of the order statistics $x_{(1)}, \cdots, x_{(n-1)}$, given a fixed value of $x_{(n)}$, and is independent of θ. It is well known from the theory of order statistics that $g(x_{(n)}; \theta)$ has the form

$$g(x_{(n)}; \theta) = n[F(x_{(n)})]^{n-1} f(x_{(n)}) = nP(x_{(n)})[Q(\theta)]^n e^{\theta K(x_{(n)})} \left[\int_a^{x(n)} P(\eta) e^{\theta K(\eta)}\, d\eta\right]^{n-1},$$

where $\mathrm{F}(x_{(n)}) = \int_a^{x(n)} f(\eta, \theta)\, d\eta$.

It follows from the above that

$$(1) \qquad h(x_{(1)}, \cdots, x_{(n-1)} \mid x_n; \theta) = \frac{\left[\exp\left[\theta \sum_{i=1}^{n-1} K(x_{(i)})\right]\right] \prod_{j=1}^{n-1} P(x_j)}{n\left[\int_a^{x(n)} P(\eta) e^{\theta K(\eta)}\, d\eta\right]^{n-1}},$$

where $h(x_{(1)}, \cdots, x_{(n-1)} \mid x_{(n)}; \theta)$ is independent of θ. Differentiating equation (1) partially with respect to θ, substituting the value of $h(x_{(1)}, \cdots, x_{(n-1)} \mid x_{(n)}; \theta)$ from (1) and placing $\frac{\partial h}{\partial \theta} = 0$, we obtain after some simple algebra

$$(2) \qquad \int_a^{x(n)} K(\eta)P(\eta)e^{\theta K(\eta)}\, d\eta = \frac{\left[\sum_{i=1}^{n-1} K(x_{(i)})\right]}{n-1} \int_a^{x(n)} P(\eta)e^{\theta K(\eta)}\, d\eta.$$

Since $f(x, \theta) \geqq 0$ for all x in R_1, it follows that $P(\eta)e^{\theta K(\eta)} \geqq 0$, for $a \leqq \eta \leqq x_{(n)}$. Moreover we obtain from the first mean value theorem for integrals that

$$\int_a^{x(n)} K(\eta)P(\eta)e^{\theta K(\eta)}\, d\eta = K(\xi) \int_a^{x(n)} P(\eta)e^{\theta K(\eta)}\, d\eta,$$

where $a \leqq \xi \leqq x_{(n)}$. Equation (2) reduces then to the form

$$(3) \qquad K(\xi) = \frac{1}{n-1} \sum_{i=1}^{n-1} K(x_{(i)}).$$

It is noted that the only sample value on which ξ is dependent is the $x_{(n)}$. Equation (3) is valid for every $x_{(1)}, \cdots, x_{(n-1)}$, satisfying the inequalities $x_{(1)} \leqq x_{(2)} \leqq \cdots \leqq x_{(n-1)} \leqq x_{(n)}$ with the $x_{(i)}$ assuming values in $R^*(\theta)$. Let $x_{(n)}$ take some fixed value arbitrarily close to $b(\theta)$. (If $f(b(\theta), \theta) \neq 0$, we can of course let $x_{(n)} = b(\theta)$.) Also let x be any number satisfying the inequality $a \leqq x \leqq x_n$. Now let $x_{(1)} = x_{(2)} = \cdots = x_{(n-1)} = x$, and we obtain from (3) the relation (4) $K(x) = K(\xi)$.

Since this relation is true for every x in the interval $a \leqq x < x_{(n)}$, it follows that $K(x)$ is a constant in the interval $a \leqq x < x_{(n)}$. (Again if we assume $f(b(\theta), \theta) \neq 0$, we can let $x_{(n)} = b(\theta)$, and it follows that $K(x)$ would be a constant in the closed interval $a \leqq x \leqq b(\theta)$.) Therefore, necessity is proved.

Proof of sufficiency. This proof is extremely simple. If $f(x, \theta) = P(x)\, Q(\theta)$, we have

$$h(x_{(1)}, \cdots, x_{(n-1)} \mid x_{(n)}; \theta) = \frac{[Q(\theta)]^n P(x_{(1)}) \cdots P(x_{(n)})}{n[Q(\theta)]^n P(x_n) \left[\int_a^{x_n} P(\eta)\, d\eta\right]^{n-1}}$$

and is independent of θ. Hence $x_{(n)}$ is a sufficient statistic for θ. This completes the proof of Theorem 1.

Before proceeding to the problem of constructing a best estimate for θ, we will use a theorem due to Blackwell [3] which will enable us to restrict ourselves to the class of unbiased estimates of θ which are functions of the sufficient statistic for θ. Blackwell's results are applicable to a much more general situation than we are considering here, and the results needed can be obtained in a different manner. Nevertheless we will summarize briefly the result which we need. He has proved that if x is any chance variable and y is any numerical chance variable for which $E(y)$ and $E[y - E(y)]^2$ are finite, and $f(x)$ is any real valued function for which $E[f(x)y]$ is finite, then $\sigma^2 E(y \mid x)$ is finite, where $E(y \mid x)$ denotes the conditional expected value of y given x. Moreover he proves that $E[f(x)\, E(y \mid x)] = E[f(x)y]$ and $\sigma^2 E(y \mid x) \leqq \sigma^2 y$ with equality holding only if $y = E(y \mid x)$ with probability one.

As a particular application of Blackwell's result it follows that if a sufficient statistic S exists, and if t is any unbiased estimate of θ, then $\alpha(S) = E[t \mid S]$ is an unbiased estimate of θ with $\sigma^2[\alpha(S)] \leqq \sigma^2 t$. It follows that we can restrict ourselves (in the case in which only the upper extremity of the range depends on θ) to the class of functions of the sufficient statistic $x_{(n)}$ which yield sufficient statistics. If we can obtain out of this class a unique function of $x_{(n)}$ which is unbiased and possesses minimum variance in this class, we will obtain an unbiased estimate of θ possessing minimum variance.

4. Derivation of the best estimate for θ when the range varies from a to $b(\theta)$. If we make the transformation of parameters $\varphi = [Q(\theta)]^{-1}$, matters are simplified considerably. If we assume that the function $\varphi(\theta)$ possesses a unique inverse $\theta(\varphi)$ and let $c(\varphi) = b[\theta(\varphi)]$, we have the condition that $\alpha(x_{(n)})$ is an unbiased estimate of φ in the form

$$\int_a^{c(\varphi)} \alpha(x_{(n)}) g(x_{(n)}, \varphi)\, dx_{(n)} \equiv \varphi. \tag{5}$$

This reduces to the condition

$$\int_a^{c(\varphi)} \alpha(x_{(n)}) P(x_{(n)}) \left[\int_a^{x(n)} P(\eta)\, d\eta\right]^{n-1} dx_{(n)} = \frac{\varphi^{n+1}}{n}.$$

If we use a new variable of integration u, where $u = \int_a^{x(n)} P(\eta)\, d\eta$, and let $\alpha(x_{(n)}) = \psi(u)$, the condition of unbiasedness becomes

$$\int_0^{\varphi} \psi(u) u^{n-1}\, du = \frac{\varphi^{n+1}}{n}.$$

Clearly the only solution of this integral equation which is an analytic function of u is given by

$$\psi(u) = \left(1 + \frac{1}{n}\right) u.$$

Since this is the unique solution for all finite φ, it follows that

$$\left(1 + \frac{1}{n}\right) \int_a^{x(n)} P(\eta)\, d\eta$$

is the only unbiased estimate of φ. Its variance can be obtained by a simple integration, and we obtain

$$\sigma_\alpha^2 = \frac{\varphi^2}{n(n+2)}.$$

If we wish to obtain an estimate for θ directly, the analysis is somewhat more complicated. Moreover it is necessary to make a further assumption to insure that the unique unbiased estimate of θ among the class of functions of $x_{(n)}$ is also a sufficient statistic. We may state this assumption as follows:

ASSUMPTION C. *$b(\theta)$ is a strictly monotone function of θ. If we define the following well defined functions*

$$u(x_{(n)}) = \int_a^{x(n)} P(\eta)\, d\eta, \qquad \beta(x_{(n)}) = b^{-1}(x_{(n)}),$$

the functions $u(x_{(n)})$ and $\beta(x_{(n)})$ satisfy the following condition:

$$u \frac{d}{du}\left[\ln\left(\frac{d\beta}{du}\right)\right] > -2 \qquad (\textit{if } b(\theta) \textit{ is strictly monotone increasing}),$$

$$u \frac{d}{du}\left[\ln\left(\frac{d\beta}{du}\right)\right] < -2 \qquad (\textit{if } b(\theta) \textit{ is strictly monotone decreasing}).$$

Moreover, the parameter set Ω is the interval defined by $\theta \geqq \theta_0$ when $b(\theta)$ is strictly monotone increasing and the interval $\theta \leqq \theta_0$ when $b(\theta)$ is strictly monotone decreasing. θ_0 satisfies the equation $b(\theta) = \theta$, so that $b(\theta_0) = a$.

Let $\alpha(x_{(n)})$ represent now a function of the sufficient statistic $x_{(n)}$. The condition that α be an unbiased estimate is expressed in the form

$$\int_a^{b(\theta)} \alpha(x_{(n)}) g(x_{(n)}, \theta)\, dx_{(n)} \equiv \theta \tag{6}$$

for every $\theta \in \Omega$. This reduces to the condition

$$\int_a^{b(\theta)} \alpha(x_{(n)}) P(x_{(n)}) \left[\int_a^{x(n)} P(\eta)\, d\eta\right]^{n-1} dx_{(n)} \equiv \frac{\theta}{n[Q(\theta)]^n}.$$

If we make the same substitution used before; namely, $u = \int_a^{x(n)} P(\eta)\, d\eta$, and let $\alpha(x_{(n)}) = \psi(u)$, the condition of unbiasedness becomes

$$\int_0^{1/Q(\theta)} \psi(u) u^{n-1}\, du \equiv \frac{\theta}{n[Q(\theta)]^n}. \tag{7}$$

It follows from Assumptions B and C that $\frac{db}{d\theta}$ and hence $\frac{dQ}{d\theta}$ exist almost everywhere in Ω. Hence differentiating (7), we obtain after simple algebra,

$$\psi\left[\frac{1}{Q(\theta)}\right] = \theta - \frac{1}{n \frac{d}{d\theta} \ln Q(\theta)}$$

for $\theta \in \Omega$. Since Ω is an interval having θ_0 as one end point, we obtain after some manipulation the expression

$$\alpha(x_{(n)}) = \beta(x_{(n)}) + \frac{u}{n} \frac{d\beta(x_{(n)})}{du(x_{(n)})}, \tag{8}$$

where $\beta(x_{(n)})$ is the function inverse to $b(x_{(n)})$, denoted in Assumption C as $b^{-1}(x_{(n)})$. $\alpha(x_{(n)})$ is the only continuous function of $x_{(n)}$ which is an unbiased estimate of θ. In order to insure that $\alpha(x_{(n)})$ is also a sufficient statistic we must be certain that $\alpha(x_{(n)})$ has a unique inverse $\alpha^{-1}(x_{(n)})$. If we take the case in which $b(\theta)$ is strictly monotone increasing, this condition becomes

$$\frac{d\alpha}{dx_{(n)}} = (n + 1) \frac{d\beta}{dx_{(n)}} + u \frac{du}{dx_{(n)}} \frac{d^2 \beta}{d^2 u} > 0. \tag{9}$$

If Assumption C holds, $\alpha(x_{(n)})$ is a sufficient statistic for $n \geqq 1$. Finally applying Blackwell's theorem we conclude that $\alpha(x_{(n)})$ given by (8) is the best estimate of θ. From (9) it is obvious that if the function $u \frac{d}{du}\left[\ln\left(\frac{d\beta}{du}\right)\right]$ is a bounded function of $x_{(n)}$ for $a \leqq x_{(n)} \leqq b(\theta)$, $\theta \in \Omega$, then for n sufficiently large $\alpha(x_{(n)})$ is a sufficient statistic and hence is the best estimate of θ assuming only the strict monotonicity of the function $b(\theta)$.

4a. *Examples.*

Rectangular Distribution. Let

$$\begin{aligned} f(x, \theta) &= \frac{1}{\theta}, \qquad 0 \leqq x \leqq \theta, \\ &= 0, \qquad \text{otherwise.} \end{aligned}$$

Since $P(x) = 1$, and $b(\theta) = \theta$, we obtain $u = x_{(n)}$ and $\beta = x_{(n)}$. Hence

$$\alpha(x_{(n)}) = \beta(x_{(n)}) + \frac{u}{n}\frac{d\beta}{du} = \left(1 + \frac{1}{n}\right) x_{(n)}.$$

Its variance is given by the expression $\sigma_\alpha^2 = \dfrac{\theta^2}{n(n+2)}$.

Exponential Distribution. Let

$$\begin{aligned} f(x, \theta) &= e^{x-\theta}, \qquad -\infty \leqq x \leqq \theta, \\ &= 0, \qquad x > \theta. \end{aligned}$$

Since $P(x) = e^x$, and $b(\theta) = \theta$, we obtain $u = e^{x_{(n)}}$, $\beta = x_{(n)}$. Hence

$$\alpha(x_{(n)}) = \beta(x_{(n)}) + \frac{u}{n}\frac{d\beta}{du} = x_{(n)} + \frac{1}{n}.$$

5. Both extremities of the range depending upon θ.

THEOREM 2. *Let $x_1, x_2, \cdots, x_n$ be the values of n independent drawings from a population having the probability density function $f(x, \theta)$ satisfying Assumptions A and B, and in which both extremities of the range depend upon θ. The necessary and sufficient condition that the first and* nth *order statistics, $x_{(1)}$ and $x_{(n)}$, be jointly sufficient statistics for θ is that*

$$f(x, \theta) = P(x)\, Q(\theta) \quad \textit{for all} \quad (x, \theta) \textit{ in } R^*(\theta) \times \Omega.$$

Proof of necessity. Suppose that in a sample of n independent observations that the first and nth order statistics, $x_{(1)}$ and $x_{(n)}$, are jointly sufficient for θ. It follows from the definition of joint sufficiency that

$$f(x_1, \theta) \cdots f(x_n, \theta) = g(x_{(1)}, x_{(n)}; \theta)h(x_{(2)}, \cdots, x_{(n-1)} \mid x_{(1)}, x_{(n)}; \theta),$$

where $g(x_{(1)}, x_{(n)}; \theta)$ is the joint frequency function of $x_{(1)}$ and $x_{(n)}$, and

$$h(x_{(1)}, \cdots, x_{(n-1)} \mid x_{(1)}, x_{(n)}; \theta)$$

denotes the conditional frequency function of the order statistics $x_{(2)}, \cdots, x_{(n-1)}$, given fixed values of $x_{(1)}$ and $x_{(n)}$, and is independent of θ. It is well known from the theory of order statistics that $g(x_{(1)}, x_{(n)}; \theta)$ has the form

$$g(x_{(1)}, x_{(n)}; \theta) = n(n-1)[F(x_{(n)}) - F(x_{(1)})]^{n-2} f(x_{(1)})\, f(x_{(n)}),$$

where $F(x_{(n)}) - F(x_{(1)}) = \int_{x(1)}^{x(n)} f(\eta, \theta)\, d\eta$. It follows from the above that

$$h(x_{(2)}, \cdots, x_{(n-1)}, \mid x_{(1)}, x_{(n)}; \theta) = \frac{\left[\exp\left[\theta \sum_{i=2}^{n-1} K(x_{(i)})\right]\right] \prod_{j=2}^{n-1} P(x_{(j)})}{n(n-1)\left[\int_{(x(1))}^{x(n)} P(\eta)e^{\theta K(\eta)}\, d\eta\right]^{n-2}}.$$

The proof proceeds similarly to the one in Theorem 1, and we end up with a similar equation

$$K(\xi) = \frac{1}{n-2} \sum_{i=2}^{n-1} K(x_{(i)}), \tag{10}$$

where $x_{(1)} \leqq \xi \leqq x_{(n)}$. Hence by a similar argument $K(x)$ is a constant in the open interval $a(\theta) < x < b(\theta)$. If $f(a(\theta), \theta)$ and $f(b(\theta), \theta)$ are unequal to zero, we can make the stronger statement that $K(x)$ is a constant in the closed interval $a(\theta) \leqq x \leqq b(\theta)$.

Proof of sufficiency. Suppose that $f(x, \theta) = P(x)\, Q(\theta)$. Then

$$h(x_{(2)}, \cdots, x_{(n-1)}, \mid x_{(1)}, x_{(n)}; \theta) = \frac{[Q(\theta)]^{n-2} \prod_{i=2}^{n-1} P(x_{(i)})}{n(n-1)[Q(\theta)]^{n-2}\left[\int_{x(1)}^{x(n)} P(\eta)\, d\eta\right]^{n-2}}$$

and is independent of θ. Hence $x_{(1)}$ and $x_{(n)}$ are jointly sufficient statistics for θ. This completes the proof of Theorem 2.

Blackwell's theorem is applicable again to this case and enables us to restrict ourselves to the class of unbiased estimates which are sufficient statistics for θ. Any unbiased sufficient statistic is a solution of the integral equation

$$\int_{a(\theta)}^{b(\theta)} dx_{(n)} \int_{a(\theta)}^{x(n)} \alpha(x_{(1)}, x_{(n)})\, g(x_{(1)}, x_{(n)})\, dx_{(1)}\, dx_{(n)} \equiv \theta \tag{11}$$

for $\theta \in \Omega$.

Pitman has shown [1] that in the particular case $a(\theta) = \theta$, $b(\theta)$ a strictly monotone decreasing function of θ, a sufficient statistic for θ exists. An independent proof is given of this statement. Moreover, the distribution of this sufficient statistic is derived, and it is shown that there exists a unique unbiased estimate of θ in the class of all functions of the sufficient statistic.

Following Pitman we simplify the discussion considerably by assuming $a(\theta) = \theta$. On the basis of Theorem 2 and Blackwell's result we need only consider functions of the smallest and largest order statistics in our search for a best estimate. First we derive Pitman's result independently. Let us consider the sample statistic

$$T = \min. \{x_{(1)}, b^{-1}(x_{(n)})\}.$$

We proceed first to find its probability distribution and then show that it is a sufficient statistic for θ. Figure 1 shows a typical contour of constant T in the $x_{(1)}, x_{(n)}$ plane.

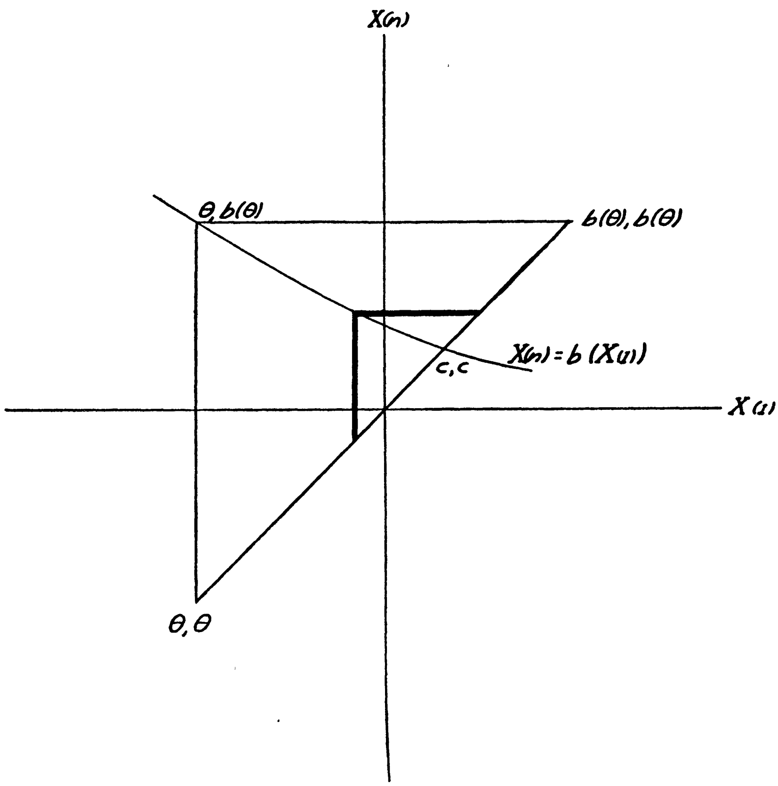

FIGURE 1

First it is clear from Assumption A that we may confine ourselves to the interior of the triangle shown in Fig. 1. Moreover, it is clear from the continuity and monotony of the function $b(\theta)$ that there exists a point with coordinates c, c (where $b(c) = c$) which is independent of θ and is such that

$$\theta \leqq c \leqq b(\theta) \text{ for all } \theta \in \Omega.$$

From Assumption B, $\Omega \subseteqq I$, where I is the interval in R_1 given by $\theta \leqq c$. It it clear from the definition of T that

$$\begin{aligned} T &\equiv b^{-1}(x_{(n)}) && \text{for all points above the curve } x_{(n)} = b(x_{(1)}), \\ T &\equiv x_{(1)} && \text{for all points below the curve } x_{(n)} = b(x_{(1)}), \\ T &\equiv x_{(1)} \equiv b^{-1}(x_{(n)}) && \text{for all points on the curve } x_{(n)} = b(x_{(1)}). \end{aligned}$$

A typical contour of constant T is shown in the figure. If we denote as before by $g(x_{(1)}, x_{(n)})$ the joint frequency function of the order statistics $x_{(1)}$ and $x_{(n)}$, it follows that

$$\text{(12)} \qquad Pr.\{t < T < t + dt\} = \left[\int_t^{b(t)} \underset{[x_{(1)}=t]}{g(x_{(1)}, x_{(n)})\, dx_{(n)}}\right] dt + \left[\int_t^{b(t)} \underset{[x_{(n)}=b(t)]}{g(x_{(1)}, x_{(n)})\, dx_{(1)}}\right] [b(t) - b(t + dt)],$$

where the first integral is evaluated holding $x_{(1)} = t$ and the second integral holding $x_{(n)} = b(t)$. It follows from the continuity and monotony of $b(\theta)$ that if we restrict the parameter set Ω to be a bounded interval in R_1, $\frac{db}{d\theta}$ will exist everywhere except on a set of points having probability measure zero. In this case T possesses a frequency function $w(t)$ almost everywhere. After performing the elementary integrations in (12) by noting that the integrands can be expressed as perfect differentials, we obtain

$$\text{(13)} \qquad w(t) = n[Q(\theta)]^n \left[\int_t^{b(t)} P(\eta)\, d\eta\right]^{n-1} \left[P(t) - \frac{db}{dt} P(b(t))\right].$$

To prove that T is a sufficient statistic for θ, we must prove that the conditional frequency function of $x_{(1)}, x_{(2)}, \cdots, x_{(n)}$, given T, is independent of θ. To do this we show that this property holds in each of the two regions indicated in Figure 1; namely in the regions below and above the curve $x_{(n)} = b(x_{(1)})$. In the region below the curve, we have

$$h(x_{(1)}, x_{(2)}, \cdots, x_{(n)} \mid T) = \frac{P(x_{(1)})P(x_{(2)}) \cdots P(x_{(n)})[Q(\theta)]^n}{w(t)}.$$

Obviously this conditional frequency function is independent of θ. In the region above the curve, $x_{(n)} = b(x_{(1)})$, we make the following transformation in the sample space: Let $\rho_1 = x_{(1)}, \rho_2 = x_{(2)}, \cdots, \rho_{n-1} = x_{(n-1)}, \rho_n = T$. Since

$$\frac{\partial(\rho_1, \rho_2, \cdots, \rho_n)}{\partial(x_{(1)}, x_{(2)}, \cdots, x_{(n)})} = \frac{db^{-1}(x_{(n)})}{dx_{(n)}},$$

the transformed likelihood function becomes

$$f(x_{(1)}, \theta)f(x_{(2)}, \theta) \cdots f(x_{(n)}, \theta) \cdot \left(\frac{db^{-1}}{dx_{(n)}}\right)^{-1}.$$

If we now assume that $b^{-1}(x_{(n)})$ is a strictly monotone decreasing function of $x_{(n)}$, the transformation is one-to-one and $\frac{db^{-1}}{dx_{(n)}}$ is unequal to zero except possibly at a set of points in the $x_{(1)}$, $x_{(n)}$ plane of probability measure zero. We may state then that

$$h(x_{(1)}, x_{(2)}, \cdots, x_{(n)} \mid T) = \frac{P(x_{(1)})P(x_{(2)}) \cdots P(x_{(n)})[Q(\theta)]^n\left(\frac{db^{-1}}{dx_{(n)}}\right)^{-1}}{w(t)}.$$

Again this conditional frequency function is independent of θ, so that this property holds throughout the triangle in Figure 1. Hence T is a sufficient statistic for θ.

We proceed to prove that there exists a unique continuous function of T which is an unbiased estimate of θ. This will involve no additional assumptions not made already. If $\psi(t)$ is an unbiased estimate of θ, we have from (13)

$$(14) \qquad E[\psi(t)] = n[Q(\theta)]^n \int_\theta^c \psi(t)\left[\int_t^{b(t)} P(\eta)\, d\eta\right]^{n-1}\left[P(t) - \frac{db}{dt} P(b(t))\right] dt \equiv \theta$$

for $\theta \in \Omega$. Differentiating (14) with respect to θ, we obtain

$$\psi(\theta) = \theta - \frac{1}{n \frac{d}{d\theta} [\ln Q(\theta)]}.$$

Since Ω is the interval $\theta \leqq c$, we obtain

$$(15) \qquad \psi(T) = T - \frac{1}{n \frac{d}{dt} [\ln Q(T)]}.$$

Hence (15) with $T = \min. \{x_{(1)}, b^{-1}(x_{(n)})\}$ is the unique continuous function of T which is an unbiased estimate of θ.

We now require an additional assumption to insure that $\psi(T)$ given by (15) is a sufficient statistic for θ.

ASSUMPTION D. *For almost all T satisfying $\theta \leqq T \leqq c$, and for all $\theta \in \Omega$, the function $\ln Q(T)$, where $[Q(T)]^{-1} = \int_T^{b(T)} P(\eta)\, d\eta$, satisfies the inequality*

$$-1 < \frac{\frac{d^2}{dt^2} [\ln Q(T)]}{\left[\frac{d}{dt} \ln Q(T)\right]^2} < M,$$

where M is some fixed constant.

The following theorem can be established:

THEOREM 3. *If a probability distribution with range from θ to $b(\theta)$ satisfies Assumptions A, B, and D, with $K(x) \equiv 0$, and if the functions $b(\theta)$ and $b^{-1}(\theta)$ are strictly monotone decreasing for all $\theta \in \Omega$, then the function $\psi(T)$ given by* (15), *where $T = \min. \{x_{(1)}, b^{-1}(x_{(n)})\}$, is the unique best estimate for the unknown parameter θ.*

PROOF. Under the above assumptions (minus Assumption D) we have proved that $\psi(T)$ given by (15) is (among all continuous functions of the sufficient statistic T) the unique unbiased estimate of θ. However, in order to apply Blackwell's theorem, we must show that $\psi(T)$ is also a sufficient statistic. From (15) we obtain

$$\frac{d\psi}{dT} = 1 + \frac{\frac{d^2}{dt^2}[\ln Q(T)]}{n\left[\frac{d}{dt}(\ln Q(T))\right]^2}.$$

From Assumption D it follows that for all sample sizes $n \geqq 1$ we have $1 + \frac{M}{n} > \frac{d\psi}{dT} > 0$. Hence the function $\psi(T)$ establishes a one-to-one correspondence between T and $\psi(T)$ except possibly at a set of points of probability measure zero. Therefore $\psi(T)$ as defined in (15) is a sufficient statistic. It follows immediately from Blackwell's theorem and the existence of a unique unbiased estimate among all functions of T that $\psi(T)$ is the best estimate of the unknown parameter θ.

THEOREM 4. *If a probability distribution with range from θ to $b(\theta)$ satisfies Assumptions A and B with $K(x) \equiv 0$, and if the upper extremity of the range, $b(\theta)$, is not a strictly monotone decreasing function of θ, there exists no single sufficient statistic for θ, which is a single valued function of the values of n independent drawings from the population.*

PROOF. Under the assumptions of the Theorem to be established we have proved in Theorem 2 that $x_{(1)}$ and $x_{(n)}$ are a sufficient set of statistics for θ. We may therefore confine our attention to a search for a single valued function $T(x_{(1)}, x_{(n)})$. It is clear that

$$Pr\{t < T < t + dt\} = n(n-1)[Q(\theta)]^n \iint_{t<T<t+dt} P(x_{(1)})P(x_{(n)}) \cdot \left[\int_{x_{(1)}}^{x_{(n)}} P(\eta)\, d\eta\right]^{n-2} dx_{(1)}\, dx_{(n)}. \tag{16}$$

Since the likelihood function of the ensemble of n independent observations taken from the distribution has (under our assumptions as to its form) the factor $[Q(\theta)]^n$ as the sole term involving θ, it is evident from the definition of sufficiency that the integral

$$\iint_{t<T<t+dt} P(x_{(1)})P(x_{(n)}) \left[\int_{x_{(1)}}^{x_{(n)}} P(\eta)\, d\eta\right]^{n-2} dx_{(1)}\, dx_{(n)}, \tag{17}$$

when evaluated over the region common to the strip $t < T < t + dt$ and the triangle $\theta \leqq x_{(1)} \leqq x_{(n)}$, $\theta \leqq x_{(n)} \leqq b(\theta)$ in the $x_{(1)}$, $x_{(n)}$ plane must be independent of θ except in the case in which the strip includes a finite length of either the line $x_{(1)} = \theta$ or the line $x_{(n)} = b(\theta)$. Moreover this restriction must be satisfied uniformly in θ for $\theta \in \Omega$. The situation is clarified by looking at Figure 2.

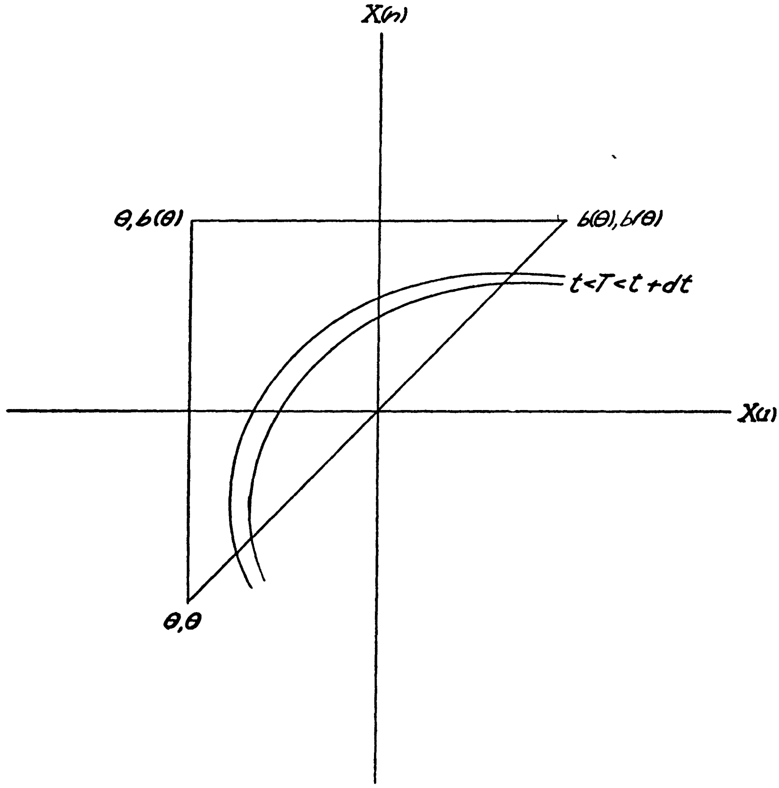

FIGURE 2

It is clear from Figure 2 that if the strip $t < T < t + dt$ does not enter and leave the triangle along the line $x_{(1)} = x_{(n)}$ without crossing either of the other two sides *for every* θ in Ω, the integral in (17) will be a function of θ. Suppose that the statistic T is of such a form that one of its strips $t < T < t + dt$ does not consist of the portions of two straight lines as was the case in Figure 1. Then for some $\theta_1 \in \Omega$ this strip $t < T < t + dt$ will intersect the triangle corresponding to the value θ_1 somewhere along at least one of the lines $x_{(1)} = \theta_1$ or $x_{(n)} = b(\theta_1)$. It follows that the contours T = constant must be of the same type as shown in Figure 1 *regardless* of the nature of the function $b(\theta)$.

Next we proceed to show that if $b(\theta)$ is not strictly monotone decreasing, the assumption that T is a single valued function of $x_{(1)}$ and $x_{(n)}$ is violated. The argument proceeds as follows: under the assumptions of the theorem $b(\theta)$ is a continuous function of θ which is not strictly monotone decreasing. Hence there exist at least two values of $\theta \in \Omega$, say θ_1 and θ_2, such that the corresponding contours of fixed T, say T_1 and T_2 intersect at least in one point P. The situation is shown in Figure 3. Now obviously $T_1 = T_2$, since otherwise $T(x_{(1)}, x_{(n)})$ would not be a single valued function of $x_{(1)}$ and $x_{(n)}$. From the properties of

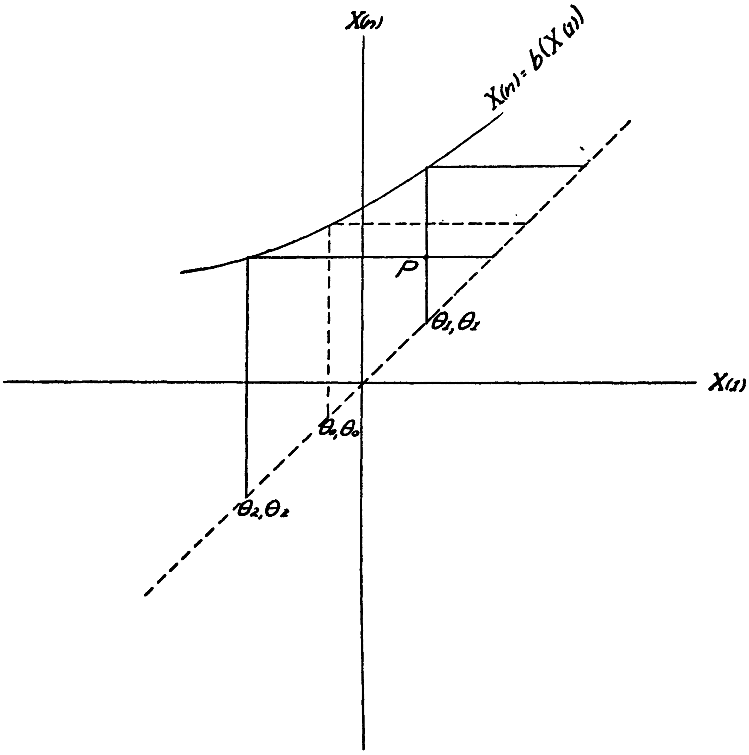

FIGURE 3

the function $b(\theta)$ there exists a $\theta_0 \in \Omega$ such that the triangle defined by $\theta_0 \leqq x_{(1)} \leqq x_{(n)}$, $\theta_0 \leqq x_{(n)} \leqq b(\theta_0)$ includes a finite length of the contour $T_1 = T_2 =$ constant. Moreover since this contour cuts the above triangle at one or more points whose coordinates depend upon the value of θ_0, it follows that if the true value of θ is θ_0, the integral defined in (17) will be a function of θ_0. Hence T is not a sufficient statistic for θ for the true value lying in the parameter set Ω.

6. An alternative approach when a single sufficient statistic does not exist. It follows from Theorem 4 that if $b(\theta)$ is not a strictly decreasing monotone function of θ that no single sufficient statistic exists. The question remains as to what to do to obtain an estimate for θ. The following procedure yields an unbiased estimate for a certain function of θ which is "best" only in the sense that it has minimum variance among the class of all analytic functions of two prescribed functions of $x_{(1)}$ and $x_{(n)}$. The fact that the sufficient statistic first derived by Pitman; i.e., $T = \min. \{x_{(1)}, b^{-1}(x_{(n)})\}$ is not an analytic function of $x_{(1)}$ and $x_{(n)}$ throughout the triangle $\theta \leqq x_{(1)} \leqq x_{(n)}$, $\theta \leqq x_{(n)} \leqq b(\theta)$ suggests that perhaps the best estimate may always be a non-analytic function. In any case the following procedure is suggested for lack of a better one.

Make the transformation of parameter $\varphi = [Q(\theta)]^{-1}$ and the coordinate transformation

$$u = \int_{x(1)}^{x(n)} P(\eta)\, d\eta, \qquad v = \int_{c}^{x(1)} P(\eta)\, d\eta,$$

where c is any fixed point whatsoever in R_1; i.e., c is independent of the value of θ for any $\theta \in \Omega$. First we will prove a lemma concerning fixed points of the nature of c.

LEMMA 1. *For a distribution satisfying Assumptions A and B with $K(x) \equiv 0$ and with the additional restriction that the functions $a(\theta)$ and $b(\theta)$ possess first derivatives ($a(\theta)$ and $b(\theta)$ depending non-trivially upon θ), there exists a point c satisfying for all $\theta \in \Omega$ the conditions*

1.) $a(\theta) \leqq c \leqq b(\theta)$,

2.) *c is a fixed p-quantile $(0 < p < 1)$ of the distribution, if and only if*

$$P[a(\theta)] \neq 0, \qquad P[b(\theta)] \neq 0, \qquad \frac{P[b(\theta)]}{P[a(\theta)]}\frac{db(\theta)}{da(\theta)} = \rho < 0.$$

for all $\theta \in \Omega$.

PROOF. If there exists a fixed point c which is a p-quantile, the

$$Q(\theta) \int_{a(\theta)}^{c} P(\eta)\, d\eta = p. \tag{18}$$

Writing $q(\theta) = \dfrac{1}{Q(\theta)} = \displaystyle\int_{a(\theta)}^{b(\theta)} P(\eta)\, d\eta$,

and differentiating (18) with respect to θ,

$$-\frac{da}{d\theta} P[a(\theta)] = p\frac{dq}{d\theta} = p\left\{\frac{db}{d\theta} P[b(\theta)] - \frac{da}{d\theta} P[a(\theta)]\right\}.$$

Solving for p, we obtain

$$p = \frac{1}{1 - \dfrac{P[b(\theta)]}{P[a(\theta)]}\dfrac{db}{da}}.$$

Since there is at most one value of p obtained from (19), and since $P(x) > 0$, it follows from (18) that c is a single valued function of p. This completes the proof of the lemma.

It is clear from Lemma 1 that in the case we are now considering there exists no fixed point c which is a p-quantile of the distribution, since $\frac{db}{da}$ is not negative for all $\theta \in \Omega$. We are now ready to prove the following theorem:

THEOREM 5. *For a distribution satisfying Assumptions A and B with* $K(x) \equiv 0$ *and with the additional restriction that* $b(\theta)$ *is not a strictly monotone decreasing function of* θ *for all* $\theta \in \Omega$, *there exists among the class of all analytic functions of* $u = \int_{x_{(1)}}^{x_{(n)}} P(\eta)\, d\eta$ *and* $v = \int_{c}^{x_{(1)}} P(\eta)\, d\eta$ *a unique function of* u *and* v; *namely* $\left(\frac{n+1}{n-1}\right)u$, *which is an unbiased estimate for* φ.

PROOF. Under our coordinate transformation to u and v as new variables of integration, $g(x_{(1)}, x_{(n)}; \varphi)\, dx_{(1)}\, dx_{(n)} = n(n-1)\varphi^{-n}u^{n-2}\, du\, dv$. Introducing a new function of θ; namely, $\beta = \int_{c}^{a(\theta)} P(\eta)\, d\eta$ the condition (11) for unbiasedness in θ becomes for the new parameter and in terms of the new variables u and v,

$$\int_0^{\varphi} du \int_{\beta}^{\varphi - u + \beta} n(n-1)\varphi^{-n}\, u^{n-2}\, \psi(u, v)\, dv \equiv \varphi \tag{20}$$

for all φ for which θ lies in Ω, where $\psi(u, v)$ is an estimate of φ. If we expand $\psi(u, v)$ in a double Taylor series about the point $u = 0$, $v = 0$, it is clear that the only terms which satisfy (20) identically in φ are

$$\psi(u, v) = au + bv,$$

where a and b are constants. We will now derive a relationship between a and b by integrating (20). After some easy algebra we obtain the relationship

$$a + b\,\frac{\left[\frac{\beta}{\varphi}(n+1) + 1\right]}{n-1} = \frac{n+1}{n-1}. \tag{21}$$

Under the conditions of the Theorem it is clear from Lemma 1 that the point c is not a p-quantile uniformly in φ and hence $\frac{\beta}{\varphi}$ is not a constant independent of φ. Hence the only solution of (21) is given by $a = \frac{n+1}{n-1}$, $b = 0$; and the only unbiased estimate of φ is

$$\psi = \frac{n+1}{n-1} \int_{x_{(1)}}^{x_{(n)}} P(\eta)\, d\eta. \tag{22}$$

REFERENCES

[1] E. J. G. PITMAN, "Sufficient statistics and intrinsic accuracy," *Proc. Camb. Phil. Soc.*, Vol. 32 (1936), pp. 567–579.

[2] B. O. KOOPMAN, "On distributions admitting a sufficient statistic," *Trans. Am. Math. Soc.*, Vol. 39 (1936), pp. 399–409.

[3] DAVID BLACKWELL, "Conditional expectation and unbiased sequential estimation," *Annals of Math. Stat.*, Vol. 18 (1947), pp. 105–110.