The Annals of Statistics
1979, Vol. 7, No. 4, 756–771

CONDITIONAL PROPERTIES OF STATISTICAL PROCEDURES FOR LOCATION AND SCALE PARAMETERS

BY G. K. ROBINSON
University of Sheffield

A systematic investigation of the conditional properties of statistical procedures for location and scale parameters is commenced. These conditional properties are compared with the known admissibility properties of point estimators in the same situations.

1. Introduction. This paper may be viewed as being a continuation of Robinson (1979). The definitions and notation are the same as that paper. This paper investigates the conditional properties of some statistical procedures for location and scale parameters more for the purpose of investigating the nature of the conditional properties than for passing judgement on the statistical procedures.

The main reason for choosing location and scale parameter problems is that many admissibility results are known for them. A review of these appears in Zacks (1971). The ones which the reader needs to be familiar with are

(1) The multidimensional mean is admissible as an estimator of the mean of a multivariate normal population for one and for two dimensions but inadmissible for three or more dimensions.

(2) When we have n independent observations on a normal population with unknown mean and unknown variance, σ^2, the minimum-mean-squared-error equivariant estimator of σ^2,

$$\frac{1}{n+1}\Sigma(x_i - \bar{x})^2$$

is inadmissible. This was shown by Stein (1964). These two results remain true with assumptions weaker than normality. We shall investigate conditional properties which may be compared with these admissibility properties.

Many of the estimators discussed in this paper will be described as being Pitman estimators. By a Pitman interval estimator we mean one where the confidence function may be derived using an appropriate improper prior (uniform for location parameters, density θ^{-1} for a scale parameter θ) and Bayes' theorem. As discussed by Pitman (1938), for confidence sets which are as in Fisher's and Neyman's theories these posterior confidence levels agree numerically with those theories. By a Pitman point estimator we mean the Bayesian posterior expectation of a parameter based upon a uniform prior if the parameter is a location parameter or a θ^{-2}

Received September 1975; revised February 1978.
AMS 1970 *subject classifications.* Primary 62A99, 62F99; secondary 62F10, 62F15, 62F25.
Key words and phrases. Conditional properties, relevant subset, Bayesianity, admissibility, location parameters, scale parameters, Pitman estimators, multidimensional normal mean.

prior density for a scale parameter θ. The symbol y will be used to denote an ancillary statistic, and Y will denote the space of its possible values.

2. Relevant betting procedures for Pitman interval estimators for location parameters. Buehler (1959) proved that Pitman estimators based on one observation on a location parameter family of distributions do not allow the existence of relevant subsets, provided that a single moment exists. Stein (1961) extended this result to the case of a single p-dimensional observation depending on a p-dimensional location parameter, but he required the existence of $p + 1$ moments for his proof. Both these results avoid the problems which ancillary statistics seem to bring. In this section we show that no moment condition is necessary to prove Stein's result. When several observations are made, relevant betting procedures may exist (an example is given for two observations on a distribution whose first moment does not exist) but a moment condition is sufficient to ensure the nonexistence of relevant betting procedures.

It should be remembered that every scale parameter family of distributions can be transformed into a location parameter family by a logarithmic transformation. Thus an improper Bayesian interval estimator for a scale parameter θ based upon a θ^{-1} improper prior density is equivalent to a Pitman interval estimator for a location parameter, so the results of this section are applicable. For point estimation the situation is more complicated as we shall see in Section 4.

PROPOSITION 2.1. *Suppose* $\mathbf{X} = (X_1, X_2, \cdots, X_p)$ *is a vector-valued random variable with probability density* $f(\mathbf{x} - \boldsymbol{\theta})$ *depending upon a location parameter* $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_p)$. *Provided*

$$\alpha(\mathbf{x}) = \int_{I(\mathbf{x})} f(\mathbf{x} - \boldsymbol{\theta})\, d\boldsymbol{\theta},$$

there is no relevant betting procedure for the interval estimator $\langle I(x), \alpha(x)\rangle$.

REMARK. It is sufficient to prove that there can be no positively biased relevant selection for any interval estimator of the type described in the proposition. If a relevant betting procedure, $s(\mathbf{x})$, exists for the interval estimator $\langle I(\mathbf{x}), \alpha(\mathbf{x})\rangle$ then the selection $k(\mathbf{x})$ is positively biased relevant for $\langle J(\mathbf{x}), \beta(\mathbf{x})\rangle$ where $k(\mathbf{x}) = s(\mathbf{x})$, $J(\mathbf{x}) = I(\mathbf{x})$ and $\beta(\mathbf{x}) = \alpha(\mathbf{x})$ if $s(\mathbf{x}) \geqslant 0$ while $k(\mathbf{x}) = -s(\mathbf{x})$, $J(\mathbf{x}) = \Theta \setminus I(x)$, the complement of the set $I(\mathbf{x})$, and $\beta(\mathbf{x}) = 1 - \alpha(\mathbf{x})$ if $s(\mathbf{x}) < 0$. This argument is also used in the proofs of Propositions 2.2, 2.3 and 3.1.

PROOF OF PROPOSITION 2.1. Suppose that a positively biased relevant betting procedure does exist for $\langle I(\mathbf{x}), \alpha(\mathbf{x})\rangle$, i.e., there is a number $\varepsilon > 0$ and a nontrivial selection $k(\mathbf{x})$ such that

$$E[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \alpha(\mathbf{X}) - \varepsilon\}k(\mathbf{X})] \geqslant 0 \quad \text{for all} \quad \theta. \tag{2.1}$$

Clearly $\varepsilon < 1$. First, find real numbers A and B such that

$$P[\mathbf{A} < \mathbf{X} < \mathbf{B} | \boldsymbol{\theta} = \mathbf{0}] \geqslant 1 - C\varepsilon$$

where $C = 2^{-p-3}$ and the vector symbol $\mathbf{A}$ denotes $(A, A, \cdots, A)$. Define $D = B - A$.

For $i = 1, 2, 3, \cdots$ consider

$$g_i(\mathbf{t}) = \int_{[\mathbf{t}-(i-1)\mathbf{D}, \mathbf{t}+i\mathbf{D}]} k(\mathbf{x})\, d\mathbf{x}$$

as a function of $\mathbf{t} = (t_1, t_2, \cdots, t_p)$. The function $g_i(\mathbf{t})$ is bounded and continuous so there is a point $\mathbf{a}_i$ such that

$$(1 + \varepsilon/10) g_i(\mathbf{a}_i) \geqslant g_i(\mathbf{t}) \qquad \text{for all} \quad \mathbf{t}.$$

By (2.1)

$$J_i = \int_{[\mathbf{a}_i-\mathbf{A}-i\mathbf{D}, \mathbf{a}_i-\mathbf{A}+i\mathbf{D}]} \int_{\mathfrak{X}} \{\chi_{I(\mathbf{x})}(\boldsymbol{\theta}) - \alpha(\mathbf{x}) - \varepsilon\} f(\mathbf{x} - \boldsymbol{\theta}) k(\mathbf{x})\, d\mathbf{x}\, d\boldsymbol{\theta} \geqslant 0. \tag{2.2}$$

Changing the order of integration,

$$J_i = \int_{\mathfrak{X}} \int_{[\mathbf{a}_i-\mathbf{A}-i\mathbf{D}, \mathbf{a}_i-\mathbf{A}+i\mathbf{D}]} \{\chi_{I(\mathbf{x})}(\boldsymbol{\theta}) - \alpha(\mathbf{x}) - \varepsilon\} f(\mathbf{x} - \boldsymbol{\theta})\, d\boldsymbol{\theta} k(\mathbf{x})\, d\mathbf{x}. \tag{2.3}$$

For $\mathbf{x} \in [\mathbf{a}_i - (i-1)\mathbf{D}, \mathbf{a}_i + i\mathbf{D}]$,

$$\int_{[\mathbf{a}_i-\mathbf{A}-i\mathbf{D}, \mathbf{a}_i-\mathbf{A}+i\mathbf{D}]} \{\chi_{I(\mathbf{x})}(\boldsymbol{\theta}) - \alpha(\mathbf{x}) - \varepsilon\} f(\mathbf{x} - \boldsymbol{\theta})\, d\boldsymbol{\theta}$$
$$\leqslant C\varepsilon - \varepsilon(1 - C\varepsilon) < -\varepsilon + 2C\varepsilon < 0$$

since the region of integration for $\mathbf{z} = \mathbf{x} - \boldsymbol{\theta}$ contains at least $1 - C\varepsilon$ of the total probability. Hence the contribution to the integral in (2.3) from this region of $\mathfrak{X}$ is less than

$$g_i(\mathbf{a}_i)(-\varepsilon + 2C\varepsilon) \leqslant -\tfrac{1}{2}\varepsilon g_i(\mathbf{a}_i). \tag{2.4}$$

The contribution to the integral in (2.3) from the region

$$\mathbf{x} \in [\mathbf{a}_i - i\mathbf{D}, \mathbf{a}_i + (i+1)\mathbf{D}] \backslash [\mathbf{a}_i - (i-1)\mathbf{D}, \mathbf{a}_i + i\mathbf{D}]$$

is less than the integral of $k(\mathbf{x})$ over the region which is

$$\int_{[\mathbf{a}_i - i\mathbf{D}, \mathbf{a}_i + (i+1)\mathbf{D}]} k(\mathbf{x})\, d\mathbf{x} - g_i(\mathbf{a}_i) \leqslant (1 + \varepsilon/10) g_{i+1}(\mathbf{a}_{i+1}) - g_i(\mathbf{a}_i). \tag{2.5}$$

To find a bound to the contribution to the integral in (2.3) from the region $\mathbf{x} \notin [\mathbf{a}_i - i\mathbf{D}, \mathbf{a}_i + (i+1)\mathbf{D}]$, divide the region into hypercubes of the form $[\mathbf{t}, \mathbf{t} + \mathbf{S}]$ where $S = (2i+1)D$. Index the hypercubes by a vector $\mathbf{n} = (n_1, n_2, \cdots, n_p)$ so that the $\mathbf{n}$th hypercube is

$$[\mathbf{a}_i - i\mathbf{D} + \mathbf{n}S, \mathbf{a}_i - i\mathbf{D} + (\mathbf{n} + \mathbf{1})S].$$

The contribution to the integral in (2.3) from the $\mathbf{n}$th hypercube is less than

$$\int_{[\mathbf{a}_i - i\mathbf{D} + \mathbf{n}S, \mathbf{a}_i - i\mathbf{D} + (\mathbf{n}+1)S]} k(\mathbf{x}) \int_{[\mathbf{a}_i - \mathbf{A} - i\mathbf{D}, \mathbf{a}_i - \mathbf{A} + i\mathbf{D}]} f(\mathbf{x} - \boldsymbol{\theta})\, d\boldsymbol{\theta}\, d\mathbf{x}$$
$$\leqslant \int_{[\mathbf{a}_i - i\mathbf{D} + \mathbf{n}S, \mathbf{a}_i - i\mathbf{D} + (\mathbf{n}+1)S]} k(\mathbf{x})\, d\mathbf{x} \int_{[\mathbf{n}S - 2i\mathbf{D} + \mathbf{A}, (\mathbf{n}+1)S + \mathbf{A}]} f(\mathbf{z})\, d\mathbf{z}$$
$$\leqslant (1 + \varepsilon/10) g_i(\mathbf{a}_i) \int_{[(\mathbf{n}-1)S + \mathbf{B}, (\mathbf{n}+1)S + \mathbf{A}]} f(\mathbf{z})\, d\mathbf{z}$$

where, again, $\mathbf{z} = \mathbf{x} - \boldsymbol{\theta}$. Now

$$\Sigma_{\mathbf{n} \neq \mathbf{0}} \int_{[(\mathbf{n}-1)S + \mathbf{B}, (\mathbf{n}+1)S + \mathbf{A}]} f(\mathbf{z})\, d\mathbf{z} \leqslant 2^p C\varepsilon = \tfrac{1}{8}\varepsilon$$

because the summation is a sum of probabilities of regions for $\mathbf{z}$; no more than 2^p regions overlap at any point; and no region overlaps the region $[\mathbf{A}, \mathbf{B}]$ which has probability at least $1 - C\varepsilon$. Therefore, the contribution to the integral in (2.3) from $\mathbf{x} \notin [\mathbf{a}_i - i\mathbf{D}, \mathbf{a}_i + (i+1)\mathbf{D}]$ is not more than

$$\tfrac{1}{8}\varepsilon(1 + \varepsilon/10)g_i(\mathbf{a}_i) < \tfrac{1}{4}\varepsilon g_i(\mathbf{a}_i). \tag{2.6}$$

Combining (2.2), (2.3), (2.4), (2.5) and (2.6) we obtain

$$(1 + \varepsilon/10)g_{i+1}(\mathbf{a}_{i+1}) \geqslant (1 + \varepsilon/4)g_i(\mathbf{a}_i).$$

Therefore,

$$g_{i+1}(\mathbf{a}_{i+1}) \geqslant (1 + \varepsilon/8)g_i(\mathbf{a}_i);$$

and, by induction,

$$g_i(\mathbf{a}_i) \geqslant (1 + \varepsilon/8)^{i-1}g_1(\mathbf{a}_1). \tag{2.7}$$

However, from its definition,

$$g_i(\mathbf{a}_i) \leqslant \{(2i + 1)D\}^p. \tag{2.8}$$

The nontriviality of $k(\mathbf{x})$ ensures that $g_1(\mathbf{a}_1) > 0$; so the geometrically increasing lower bound in (2.7) must exceed the more slowly increasing upper bound in (2.8) for some large i. This contradiction establishes the result.

2.1. Examples of relevant betting procedures for Pitman interval estimators. Consider the location parameter family of density functions

$$f(x|\theta) = \tfrac{1}{2}\beta\{1 + |x - \theta|\}^{-1-\beta}$$

where $0 < \beta < 1$ and x and θ are real. For two observations, x_1 and x_2, the interval estimator $\langle(\tfrac{1}{2}x_1 + \tfrac{1}{2}x_2, \infty), \tfrac{1}{2}\rangle$ is a Pitman interval estimator. However, for arbitrary $\delta > 0$ there is a number $\gamma > 0$ such that

$$P\left[\theta > \tfrac{1}{2}X_1 + \tfrac{1}{2}X_2 | 0 < X_2 < \gamma X_1\right] < \delta \qquad \text{for all} \quad \theta. \tag{2.9}$$

Thus the subset $0 < x_2 < \gamma x_1$ of the sample space is a negatively biased relevant subset for this interval estimator and a positively biased relevant subset for the interval estimator $\langle(-\infty, \tfrac{1}{2}x_1 + \tfrac{1}{2}x_2), \tfrac{1}{2}\rangle$.

The assertion of the last paragraph is clearly correct when $\theta < 0$. To prove it when $\theta > 0$ let us define

$$P_1 = P\left[\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2 < \theta \text{ and } 0 < X_2 < \gamma X_1\right],$$

$$P_2 = P\left[\tfrac{1}{2}X_1 + \tfrac{1}{2}X_2 > \theta \text{ and } 0 < X_2 < \gamma X_1\right].$$

Considering the regions of the sample space of which these are probabilities, it is clear that

$$P_1 \leqslant P\left[0 < X_1 < 2\theta \text{ and } 0 < X_2 < 2\gamma\theta\right],$$

and that, provided $\gamma < \tfrac{1}{2}$,

$$P_2 \geqslant P\left[X_1 > \theta/\gamma \text{ and } 0 < X_2 < \theta\right].$$

Using the formula

$$P[X - \theta > t] = \tfrac{1}{2}|1 + t|^{-\beta};$$

and noting that the probability density of X_2 is less than $\frac{1}{2}\beta(1 + \frac{1}{2}\theta)^{-1-\beta}$ throughout $0 < X_2 < 2\gamma\theta$ provided $\gamma < \frac{1}{4}$:

$$P_1 \leqslant \{1 - (1 + \theta)^{-\beta}\}\gamma\theta\beta(1 + \tfrac{1}{2}\theta)^{-1-\beta},$$

$$P_2 \geqslant \tfrac{1}{4}(1 + \theta/\gamma - \theta)^{-\beta}\{1 - (1 + \theta)^{-\beta}\}.$$

Hence

$$\begin{aligned} P_1/P_2 &\leqslant 4\gamma\theta\beta(1 + \tfrac{1}{2}\theta)^{-1-\beta}(1 + \theta/\gamma - \theta)^{\beta} \\ &\leqslant 4\gamma\theta\beta(1 + \tfrac{1}{2}\theta)^{-1-\beta}\{(1 + \tfrac{1}{2}\theta)(1 + 2/\gamma)\}^{\beta} \\ &\leqslant 8\beta\gamma(1 + 2/\gamma)^{\beta} \\ &\to 0 \quad \text{as} \quad \gamma \to 0. \end{aligned}$$

The probability in (2.9) is smaller than P_1/P_2 so the assertion is proved.

This class of example shows that if only moments of order less than $\beta < 1$ exist then there may be relevant betting procedures. The next result shows that the existence of a first order moment is sufficient to ensure that there are no relevant betting procedures.

PROPOSITION 2.2. *If $\langle I(\mathbf{x}), \alpha(\mathbf{x})\rangle$ is a Pitman interval estimator for a real location parameter, θ, then no relevant betting procedure exists provided that $E[|\xi|]$ exists where ξ is the Pitman point estimator of θ.*

PROOF. From Pitman (1938) we see that $\mathbf{x}$ can be written as an ordered pair, (ξ, y), where ξ is the Pitman estimator of θ and y is an ancillary statistic. We will sometimes use alternative notation to take advantage of this: $I(\xi, y)$ for set function, $\alpha(\xi, y)$ for the confidence function, $g(\xi - \theta|y)$ for the density of $\xi - \theta$ given y (it is independent of θ) and $G(y)$ for the distribution function of y so that integrals over all y values are written in the form $\int_Y \cdots dG(y)$.

By the remark after Proposition 2.1, we only need to prove that no positively biased relevant selection exists.

Suppose that for some $\varepsilon > 0$ there is a nontrivial selection $k(\mathbf{x}) = k(\xi, y)$ such that

$$E[\{\chi_{I(\mathbf{X})}(\theta) - \alpha(\mathbf{X}) - \varepsilon\}k(\mathbf{X})] \geqslant 0 \quad \text{for all} \quad \theta. \tag{2.10}$$

If $\int_{\mathfrak{X}} k(\mathbf{x})\, d\mathbf{x}$ is finite, then we can change the order of integration to show that

$$\begin{aligned} &\int_\Theta E[\{\chi_{I(\mathbf{X})}(\theta) - \alpha(\mathbf{X}) - \varepsilon\}k(\mathbf{X})]\, d\theta \\ &\qquad = \int_{\mathfrak{X}} k(\mathbf{x}) \int_\Theta \{\chi_{I(\mathbf{x})}(\theta) - \alpha(\mathbf{x}) - \varepsilon\} f(\mathbf{x}|\theta)\, d\theta\, d\mathbf{x} \\ &\qquad = -\varepsilon \int_{\mathfrak{X}} k(\mathbf{x})\, d\mathbf{x} < 0. \end{aligned}$$

This is in contradiction with (2.10).

Therefore, we suppose that $\int_{\mathfrak{X}} k(\mathbf{x})\, d\mathbf{x}$ is infinite, and consider

$$h(R) = \int_{-R}^{R} E\left[\{\chi_{I(\mathbf{X})}(\theta) - \alpha(\mathbf{X}) - \varepsilon\} k(\mathbf{X})\right] d\theta$$
$$= \int \int_{-\infty}^{\infty} k(\xi, y) \int_{-R}^{R} \{\chi_{I(\xi, y)}(\theta) - \alpha(\xi, y) - \varepsilon\} g(\xi - \theta | y)\, d\theta\, d\xi\, dG(y).$$

Defining $A(y)$ by

$$P[|\xi - \theta| > A(y) | y] = \tfrac{1}{2}\varepsilon; \tag{2.11}$$

and using the bounds

$$\int_{-R}^{R} \{\chi_{I(\xi, y)}(\theta) - \alpha(\xi, y) - \varepsilon\} g(\xi - \theta | y)\, d\theta \leqslant -\tfrac{1}{4}\varepsilon \qquad \left(\text{provided } \varepsilon < \tfrac{1}{2}\right)$$

for $|\xi| < R - A(y)$;

$$\int_{-R}^{R} \{\chi_{I(\xi, y)}(\theta) - \alpha(\xi, y) - \varepsilon\} g(\xi - \theta | y)\, d\theta \leqslant 1$$

for $R - A(y) < |\xi| < R$; and

$$k(\xi, y) \int_{-R}^{R} \{\chi_{I(\xi, y)}(\theta) - \alpha(\xi, y) - \varepsilon\} g(\xi - \theta | y)\, d\theta \leqslant \int_{-R}^{R} g(\xi - \theta | y)\, d\theta$$

for $|\xi| > R$; we find that

$$\begin{aligned} h(R) \leqslant\ & \int_Y \int_{-\infty}^{-R} \int_{-R}^{R} g(\xi - \theta | y)\, d\theta\, d\xi\, dG(y) \\ & + \int_Y 2A(y)\, dG(y) \\ & + \int_Y \int_{-R+A(y)}^{R-A(y)} k(\xi, y)\left(-\tfrac{1}{2}\varepsilon\right) d\xi\, dG(y) \\ & + \int_Y \int_R^{\infty} \int_{-R}^{R} g(\xi - \theta | y)\, d\theta\, d\xi\, dG(y). \end{aligned}$$

Now

$$\begin{aligned} \int_{-\infty}^{-R} \int_{-R}^{R} g(\xi - \theta | y)\, d\theta\, d\xi &\leqslant \int_{-\infty}^{-R} \int_{-R}^{\infty} g(\xi - \theta | y)\, d\theta\, d\xi \\ &= \int_{-\infty}^{0} \int_{\frac{1}{2}z - R}^{-\frac{1}{2}z - R} g(z | y)\, du\, dz \\ &= \int_{-\infty}^{0} |z|\, g(z | y)\, dz \end{aligned}$$

where $z = \xi - \theta$ and $u = \frac{1}{2}(\xi + \theta)$.

Similarly,

$$\int_R^{\infty} \int_{-R}^{R} g(\xi - \theta | y)\, d\theta\, d\xi \leqslant \int_0^{\infty} |z|\, g(z | y)\, dz.$$

Therefore,

$$\begin{aligned} h(R) \leqslant\ & \int_Y \int_{-\infty}^{\infty} |z|\, g(z | y)\, dz\, dG(y) + (2 + \varepsilon) \int_Y A(y)\, dG(y) \\ & - \tfrac{1}{4}\varepsilon \int_Y \int_{-R}^{R} k(\xi, y)\, d\xi\, dG(y). \end{aligned}$$

However,

$$\int_Y \int_{-\infty}^{\infty} |z|\, g(z | y)\, dz\, dG(y) = E[|\xi - \theta|] < \infty;$$

from (2.11)

$$A(y) \leqslant (2/\varepsilon) E[|\xi - \theta|\ | y],$$

so that

$$\int_Y A(y)\, dG(y) \leqslant (2/\varepsilon)E[|\xi - \theta|] < \infty;$$

but $\int_{\mathfrak{X}} k(\mathbf{x})\, d\mathbf{x}$ is infinite. Hence $h(R) \to -\infty$ as $R \to \infty$ in contradiction with (2.10).

For interval estimation of a p-dimensional location parameter, $\boldsymbol{\theta}$, I suspect that the moment condition necessary to ensure that no relevant betting procedure exists is either that $E[|\boldsymbol{\xi}|]$ exists or that $E[|\boldsymbol{\xi}|^p]$ exists where $|\boldsymbol{\xi}|$ denotes the norm of the Pitman estimator, $\boldsymbol{\xi}$, of $\boldsymbol{\theta}$. It would be of some interest to see how the moment condition varied with p but this is not very important.

It can be shown that, no matter what the moment condition, any Pitman interval estimator can be changed so that there are no relevant betting procedures by altering it to make no conclusion (i.e., to quote either $I(\mathbf{x}) = \Theta$ and $\alpha(\mathbf{x}) = 1$ or $I(\mathbf{x}) = \phi$ and $\alpha(\mathbf{x}) = 0$) whenever any two observations differ by more than a preassigned amount. The only result which I have is the following one which concerns p separate location parameters, not the general p-dimensional situation.

PROPOSITION 2.3. *Suppose that we observe* $\mathbf{X}$ *whose distribution depends upon a vector* $\boldsymbol{\theta} = (\theta_1, \theta_2, \cdots, \theta_p)$ *of location parameters. Suppose that the situation is really one of p separate location parameter problems, so that*

$$f(\boldsymbol{\xi} - \boldsymbol{\theta}|\mathbf{y}) = \Pi_{i=1}^p g_i(\xi_i - \theta_i|y_i)$$

where $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_p)$ *and* $\mathbf{y} = (y_1, y_2, \cdots, y_p)$, *where* ξ_i *is the Pitman estimator of* θ_i *and* y_i *is the ancillary statistic containing the remainder of the information in the observations depending on* θ_i. *The function* $g_i(\xi_i - \theta_i|y_i)$ *is the density of* $\xi_i - \theta_i$ *given* y_i. *Provided* $E[|\xi_i - \theta_i|^p]$ *exists for* $i = 1, 2, \cdots, p$ *there is no relevant betting procedure for any Pitman interval estimator for* $\boldsymbol{\theta}$.

PROOF. Suppose that there is a positively biased relevant selection $k(\mathbf{x})$ for some Pitman interval estimator $\langle I(x), \alpha(x)\rangle$ so that

$$E[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \alpha(\mathbf{X}) - \varepsilon\}k(\mathbf{X})] \geqslant 0 \quad \text{for all } \boldsymbol{\theta}. \tag{2.12}$$

Define a sequence of unnormalized prior densities

$$h_m(\boldsymbol{\theta}) = \Pi_{i=1}^p(1 + |\theta_i|/m)^{-2}$$

and calculate the corresponding posterior probabilities that $\boldsymbol{\theta} \in I(\mathbf{x})$

$$\beta_m(\mathbf{x}) = \frac{\int_{I(\mathbf{x})} h_m(\boldsymbol{\theta})\Pi_{i=1}^p g_i(\xi_i - \theta_i|y_i)\, d\boldsymbol{\theta}}{\int_\Theta h_m(\boldsymbol{\theta})\Pi_{i=1}^p g_i(\xi_i - \theta_i|y_i)\, d\boldsymbol{\theta}}.$$

Using the inequalities

$$\begin{aligned}\{1 + |x/m|\}^{-2}\{1 - (2/m)|x - a|\} &\leqslant \{1 + |a/m|\}^{-2} \\ &\leqslant \{1 + |x/m|\}^{-2}\{1 + (2/m)|x - a| + (3/m^2)(x - a)^2\}\end{aligned}$$

$$\beta_m(\mathbf{x}) = \frac{\int_{I(\mathbf{x})}\Pi_{i=1}^p[\{1 + (2/m)|\xi_i - \theta_i| + (6/m^2)(\xi_i - \theta_i)^2\} g_i(\xi_i - \theta_i|y_i)]\, d\boldsymbol{\theta}}{\int_\Theta \Pi_{i=1}^p[\{1 - (2/m)|\xi_i - \theta_i|\} g_i(\xi_i - \theta_i|y_i)]\, d\boldsymbol{\theta}}.$$

Hence, using the first of the lemmas in the appendix,

$$\beta_m(\mathbf{x}) \leqslant \alpha(\mathbf{x}) + 3\int_\Theta\left[\Pi_{i=1}^p\{1 + (2/m)|t_i| + (6/m^2)t_i^2\} - 1\right]\Pi_{i=1}^p g_i(t_i|y_i)\,d\mathbf{t},$$

this integral contains $3^p - 1$ terms and it will be less than $\frac{1}{2}\varepsilon$ at least whenever

$$(2.13)\qquad (6/m)\int_{-\infty}^{\infty}|t_i|\,g_i(t_i|y_i)\,dt_i \leqslant 3^{-p}\tfrac{1}{2}\varepsilon \qquad \text{for}\quad i = 1, 2, \cdots, p;$$

and

$$(2.14)\qquad (18/m^2)\int_{-\infty}^{\infty}t_i^2 g_i(t_i|y_i)\,dt_i \leqslant 3^{-p}\tfrac{1}{2}\varepsilon \qquad \text{for}\quad i = 1, 2, \cdots, p.$$

Now Proposition 2.2 implies the case $p = 1$ of this result, so we may assume that $p \geqslant 2$. Since $E[|\xi_i - \theta_i|^p]$ exists, the probabilities (in the probability distribution of y_i) that (2.13) and (2.14) are violated must both decrease more rapidly than m^{-p} as $m \to \infty$.

Define $q(m)$ to be the probability that one of the inequalities (2.13) and (2.14) is violated for some i. Now $\beta_m(\mathbf{x}) \leqslant \alpha(\mathbf{x}) + \frac{1}{2}\varepsilon$ except with probability $q(m)$, so

$$\begin{aligned}\lim_{m\to\infty}&\int_\Theta h_m(\boldsymbol{\theta})E\left[\{\chi_{I(\mathbf{X})}(\theta) - \alpha(\mathbf{X}) - \varepsilon\}k(\mathbf{X})\right]d\boldsymbol{\theta}\\ &\leqslant \lim_{m\to\infty}\int_\Theta h_m(\boldsymbol{\theta})E\left[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \beta_m(\mathbf{X}) - \tfrac{1}{2}\varepsilon\}k(\mathbf{X})\right]d\boldsymbol{\theta}\\ &\quad + \lim_{m\to\infty}\int_\Theta h_m(\boldsymbol{\theta})q(m)\,d\boldsymbol{\theta}\\ &= -\tfrac{1}{2}\varepsilon\int_\Theta E\left[k(\mathbf{X})\right]d\boldsymbol{\theta}\end{aligned}$$

since

$$\lim_{m\to\infty}\int_\Theta h_m(\theta)q(m)\,d\boldsymbol{\theta} = \lim_{m\to\infty}(2m)^p q(m) = 0$$

and a change of order of integration shows that

$$\int_\Theta h_m(\boldsymbol{\theta})E\left[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \beta_m(\mathbf{X})\}k(\mathbf{X})\right]d\boldsymbol{\theta} = 0.$$

However, since $k(\mathbf{X})$ is a nontrivial selection,

$$\int_\Theta E\left[k(\mathbf{X})\right]d\boldsymbol{\theta} > 0$$

so we have a contradiction with (2.12).

3. Semirelevant betting procedures for Pitman interval estimators for location parameters. As Example 4.2 of Robinson (1979) illustrates, semirelevant betting procedures may exist for Pitman interval estimators even if all moments of the underlying distribution exist. However, if the confidence set, $I(x)$, is always balanced (in a sense to be defined) about the Pitman point estimator of a real location parameter, then no semirelevant betting procedure exists. This will be proved as Proposition 3.1. For a 2-dimensional location parameter it seems natural, in the light of the known results on the admissibility of Pitman point estimators, to conjecture that there can again be no semirelevant betting procedures for "balanced" interval estimators. For 3-dimensional location parameters there may be semirelevant betting procedures even for "balanced" interval estimators, as is shown below.

Suppose that θ is a real location parameter. Using ξ for the Pitman estimator of θ and y for the maximal ancillary statistic, let $g(\xi - \theta | y)$ denote the density function of $\xi - \theta$ for given y.

PROPOSITION 3.1. *Suppose $\langle I(\mathbf{x}), \alpha(\mathbf{x})\rangle$ is a Pitman interval estimator of θ which is balanced about ξ in the sense that*

$$(3.1) \qquad \int_{I(\mathbf{x})}(\xi - \theta) g(\xi - \theta | y)\, d\theta = 0 \qquad \text{for all} \quad \mathbf{x}.$$

Provided that $E[\xi^2]$ exists, that the function $g(\xi - \theta|y)$ is continuous in $\xi - \theta$ and that $I(\mathbf{x})$ is such that the measure of $J(\theta) \Delta J(\theta')$ tends to zero as $\theta \to \theta'$ where $J(\theta) = \{\mathbf{x} \in \mathfrak{X}: \theta \in I(\mathbf{x})\}$, there is no semirelevant betting procedure for $\langle I(\mathbf{x}), \alpha(\mathbf{x})\rangle$.

PROOF. Suppose that a selection $k(\mathbf{x})$ is such that

$$E[\{\chi_{I(\mathbf{X})}(\theta) - \alpha(\mathbf{X})\}k(\mathbf{X})] \geq 0 \qquad \text{for all} \quad \theta.$$

Consider the sequence of prior densities proportional to

$$h_m(\theta) = (1 + \theta^2/m^2)^{-1} \qquad m = 1, 2, 3, \cdots,$$

and let $\beta_m(\mathbf{x})$ denote the posterior confidence in $I(\mathbf{x})$ based on the mth prior.

Using Lemma 2, condition (3.1) and Lemma 1:

$$\beta_m(\mathbf{x}) = \frac{\int_{I(\mathbf{x})}(1 + \theta^2/m^2)^{-1} g(\xi - \theta|y)\, d\theta}{\int_{-\infty}^{\infty}(1 + \theta^2/m^2)^{-1} g(\xi - \theta|y)\, d\theta}$$

$$\leq \frac{\int_{I(\mathbf{x})}\left\{1 + \frac{\xi^2}{m^2}\right\}^{-1}\left\{1 - \frac{2\xi}{m^2 + \xi^2}(\theta - \xi) + \frac{3}{m^2}(\theta - \xi)^2\right\} g(\xi - \theta|y)\, d\theta}{\int_{-\infty}^{\infty}\left\{1 + \frac{\xi^2}{m^2}\right\}^{-1}\left\{1 - \frac{2\xi}{m^2 + \xi^2}(\theta - \xi) - \frac{3}{m^2}(\theta - \xi)^2\right\} g(\xi - \theta|y)\, d\theta}$$

$$\leq \alpha(\mathbf{x}) + (9/m^2)V(y)$$

where $V(y) = E[(\xi - \theta)^2|y]$. Note that $E[V(y)] = E[(\xi - \theta)^2]$ which is finite. Now

$$\begin{aligned} L &= \lim_{m\to\infty}\int_{-\infty}^{\infty} h_m(\theta) E[\{\chi_{I(\mathbf{X})}(\theta) - \alpha(\mathbf{X})\}k(\mathbf{X})]\, d\theta \\ &\leq \lim_{m\to\infty}\int_{-\infty}^{\infty} h_m(\theta) E[\{\chi_{I(\mathbf{X})}(\theta) - \beta_m(\mathbf{X})\}k(\mathbf{X})]\, d\theta \\ &\quad + \lim_{m\to\infty}\int_{-\infty}^{\infty} h_m(\theta) E[(9/m^2)V(y)k(\mathbf{X})]\, d\theta \\ &= 0 \end{aligned}$$

since a change of order of integration shows that

$$\int_{-\infty}^{\infty} h_m(\theta) E[\{\chi_{I(\mathbf{X})}(\theta) - \beta_m(\mathbf{X})\}k(\mathbf{X})]\, d\theta = 0$$

and the bound $k(\mathbf{X}) \leq 1$ shows that the second limit is zero.

If $k(\mathbf{X})$ were to be a positively biased relevant selection then L would have to be strictly positive (by arguments like those used in the proof of Proposition 7.3 of

Robinson (1979)). Therefore there can be no positively biased relevant selection and hence, by the argument accompanying Proposition 2.1, there can be no semirelevant betting procedure.

COROLLARY. *There are no semirelevant betting procedures for the usual interval estimators for the mean of a normal distribution with known variance.*

3.1. An example of a selection which is semirelevant for a class of Pitman interval estimators. Consider the selection

$$k(\mathbf{x}) = (1 + R^2)^{-\frac{1}{2}}$$

where R is the distance from the 3-dimensional sample mean to the origin. Let μ denote the distance of the population mean, $\boldsymbol{\theta}$, from the origin, and consider a sphere of radius r centered at $\boldsymbol{\theta}$. The average of the probability of selection over the surface of the sphere is

$$K(r, \mu) = \frac{\int_0^\pi 2\pi r \sin\phi(1 + \mu^2 + r^2 - 2\mu r \cos\phi)^{-\frac{1}{2}} r\, d\phi}{\int_0^\pi 2\pi r \sin\phi r\, d\phi}$$

where ϕ is the angle between $\boldsymbol{\theta}$ and the vector from a variable point on the surface to $\boldsymbol{\theta}$. Now

$$\begin{aligned} K(r, \mu) &= \tfrac{1}{2}\int_0^\pi \sin\phi(1 + \mu^2 + r^2 - 2\mu r\cos\phi)^{-\frac{1}{2}}\, d\phi \\ &= (1/2\mu r)\left[\{1 + (\mu + r)^2\}^{\frac{1}{2}} - \{1 + (\mu - r)^2\}^{\frac{1}{2}}\right] \end{aligned}$$

which is monotone decreasing in r since

$$\begin{aligned} \frac{\partial K(r, \mu)}{\partial r} &= -(1/2\mu r^2)\left[\{1 + (\mu + r)^2\}^{\frac{1}{2}} - \{1 + (\mu - r)^2\}^{\frac{1}{2}}\right] \\ &\quad + (1/2\mu r)\left[(\mu + r)\{1 + (\mu + r)^2\}^{-\frac{1}{2}} - (\mu - r)\{1 + (\mu - r)^2\}^{-\frac{1}{2}}\right] \\ &= (1/2\mu r^2)\left[(-1 - \mu^2 - \mu r)\{1 + (\mu + r)^2\}^{-\frac{1}{2}}\right. \\ &\quad \left. + (1 + \mu^2 - \mu r)\{1 + (\mu - r)^2\}^{-\frac{1}{2}}\right] \\ &\leqslant (1 + \mu^2)\left[\{1 + (\mu - r)^2\}^{-\frac{1}{2}} - \{1 + (\mu + r)^2\}^{-\frac{1}{2}}\right] / (2\mu r^2) \\ &\leqslant 0. \end{aligned}$$

Thus, for all values of $\boldsymbol{\theta}$ the selection $k(\mathbf{x})$ includes points near to $\boldsymbol{\theta}$ more often than points far from $\boldsymbol{\theta}$. Therefore this selection will be a positively biased semirelevant selection for any Pitman interval estimator $\langle I(\mathbf{x}), \alpha(\mathbf{x})\rangle$ such that

$$I(\mathbf{x}) = \{\boldsymbol{\theta}: |\boldsymbol{\theta} - \boldsymbol{\xi}| \leqslant c\}$$

where $\boldsymbol{\xi}$ is the Pitman estimator of $\boldsymbol{\theta}$ and c is a constant.

4. Admissibility of confidence regions for the multidimensional normal mean. From Proposition 8.2 of Robinson (1979), which showed that the absence of semirelevant betting procedures implies admissibility with respect to squared-error loss, and Proposition 3.1, it follows that the usual interval estimates for the one-dimensional normal mean are admissible with respect to squared-error loss. I have tried to use the selection of Example 3.1 to prove the inadmissibility of the usual confidence regions for the 3-dimensional normal mean but have not been successful. For the 5-dimensional case I can demonstrate inadmissibility.

PROPOSITION 4.1. *Spherical Neyman confidence regions for the 5-dimensional normal mean are inadmissible with respect to squared-error loss in the case where the covariance matrix is known.*

PROOF. Let $\boldsymbol{\theta}$ denote the population mean, let $\mathbf{X}$ denote the sample mean, and suppose that the confidence region is $|\mathbf{x} - \boldsymbol{\theta}| < a$ with associated confidence level α. Let $R = |\mathbf{x}|$ and $\mu = |\boldsymbol{\theta}|$ and consider the selection

$$k(\mathbf{x}) = (1 + R^2)^{-1}$$

The average value of $k(\mathbf{x})$ over the surface $|\mathbf{x} - \boldsymbol{\theta}| = r$ is

$$K(r, \mu) = \frac{\int_0^\pi \sin^3\phi(1 + \mu^2 + r^2 - 2\mu r \cos\phi)^{-1} d\phi}{\int_0^\pi \sin^3\phi d\phi}$$

$$= (3/8\mu r)\left[2x - (x^2 - 1)\log\{(x + 1)/(x - 1)\}\right]$$

where $x = (1 + \mu^2 + r^2)/(2\mu r)$. Note that $x > 1$.

For $r \geqslant \mu$ differentiation under the integral sign shows that $K(r, \mu)$ is decreasing as a function of r. For $r < \mu$ differentiate the closed form:

$$\frac{\partial K(r, \mu)}{\partial r} = \frac{-3}{8\mu r^2}$$

$$\times\left[2x - (x^2 - 1)\log\left(\frac{x+1}{x-1}\right) + \frac{1 + \mu^2 - r^2}{\mu r}\left\{2 - x\log\left(\frac{x+1}{x-1}\right)\right\}\right].$$

For given x, the extreme values of the expression in square brackets are at $(1 + \mu^2 - r^2)/(\mu r) = 0$ and $(1 + \mu^2 - r^2)/(\mu r) = 2x$. The first of these extreme values is positive since $K(r, \mu)$ is positive. The second of the extreme values is

$$2x - (x^2 - 1)\log\left(\frac{x+1}{x-1}\right) + 4x - 2x^2\log\left(\frac{x+1}{x-1}\right)$$

$$= 6x + (1 - 3x^2)\log\{1 + 2/(x - 1)\}$$

$$> 6x + (1 - 3x^2)\{2(x - 1)^{-1} - 2(x - 1)^{-2}\}$$

(using the Taylor series for $\log(1 + z)$)

$$= (x - 1)^{-2}\{12x^2 - 10x + 2\}$$

$$> 0 \quad \text{since} \quad x > 1.$$

Thus $K(r, \mu)$ is a decreasing function of r in this region also. Therefore the selection $k(\mathbf{x})$ is positively biased semirelevant. Further, since

$$\frac{d^3[2x - (x^2 - 1)\log\{(x+1)/(x-1)\}]}{dx^3} = -\frac{8}{(x^2-1)^2}$$
$$= -8(x^{-4} + 2x^{-6} + 3x^{-8} + 4x^{-10} + \cdots),$$

(4.1)
$$K(r, \mu) = \left(\tfrac{3}{8}\mu r\right)\left\{\left(\tfrac{4}{3}\right)x^{-1} + \left(\tfrac{4}{15}\right)x^{-3} + \cdots\right\}$$
$$= \mu^{-2} - \left\{1 + \left(\tfrac{1}{5}\right)r^2\right\}\mu^{-4} + \cdots,$$

the series in x converging for $x > 1$ and the last series being asymptotic as $\mu \to \infty$.

Now let $\langle I(\mathbf{x}), \alpha\rangle$ denote the given interval estimator for $\boldsymbol{\theta}$ and define

$$\beta_\delta(\mathbf{x}) = \alpha + \delta(1 + R^2)^{-1}$$

where $\delta > 0$. The difference

$$E\left[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \alpha\}^2\right] - E\left[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \beta_\delta(\mathbf{X})\}^2\right]$$

is a function of μ since all the functions are spherically symmetrical. Denoting the difference by $d_\delta(\mu)$,

$$d_\delta(\mu) = E\left[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \alpha\}2\delta(1 + R^2)^{-1} - \delta^2(1 + R^2)^{-2}\right].$$

From (4.1) and the shape of $I(\mathbf{x})$ it is clear that $E[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \alpha\}(1 + R^2)^{-1}]\mu^4$ tends to a strictly positive finite limit as $\mu \to \infty$. It is also clear that $E[(1 + R^2)^{-2}]\mu^4$ tends to a finite limit as $\mu \to \infty$. Hence there is a number $\delta_0 > 0$ and a number μ_0 such that $d_\delta(\mu) > 0$ for all $\delta < \delta_0$ and all $\mu > \mu_0$.

Now for $\mu \leqslant \mu_0$, $E[\{\chi_{I(\mathbf{X})}(\boldsymbol{\theta}) - \alpha\}(1 + R^2)^{-1}]$ is strictly positive (since $k(\mathbf{x})$ is positively biased semirelevant) and, therefore, bounded away from zero. However $E[(1 + R^2)^{-2}]$ is bounded, so, for sufficiently small δ, $d_\delta(\mu) > 0$ for $\mu \leqslant \mu_0$. This completes the proof.

In Section 6. of Robinson (1979) the concept of a relevant betting procedure for multidimensional point estimation was discussed. Since the p-dimensional sample mean is a uniform limit of proper Bayesian estimators, there are no relevant betting procedures for this sample mean as an estimator of the mean of a p-dimensional normal population with known covariance matrix no matter what the value of p. From Proposition 2.3, we know that the usual confidence intervals for the p-dimensional normal mean do not allow relevant betting procedures for any p. The admissibility results in this situation seem to be more closely related to semirelevant selections than to relevant selections.

5. Relevant betting procedures for Pitman point estimators of location and scale parameters. It seems intuitively reasonable that the conditional properties of Pitman point estimators should closely parallel those of Pitman interval estimators. Regrettably, it is not trivial to construct proofs by analogy to the proofs of Sections 2 and 3.

An additional difficulty with point estimation is that the problems of location and scale must be considered separately. If θ is a location parameter then e^{θ} is a scale parameter where an observation x on the location parameter family is transformed to the observation e^x. However, if the estimator T has good conditional properties as an estimator of θ then it is unlikely that e^T has good conditional properties as an estimator of e^{θ}.

For a location parameter, θ, the Pitman point estimator of θ, denoted ξ, is the improper Bayesian estimator for a uniform prior distribution and is unbiased. The method of proof of Proposition 2.3 is sufficient to prove that there is no relevant betting procedure for ξ provided $E[|\xi|^3]$ exists. Semirelevant betting procedures always exist since the betting procedure

$$s(x) = 1 \quad \text{if} \quad \xi < a \tag{5.1}$$
$$= -1 \quad \text{if} \quad \xi \geqslant a$$

is semirelevant for any real number a. This betting procedure guesses that if $\xi \geqslant a$, it is too large, while if $\xi < a$, it is too small.

For a scale parameter, θ, Pitman (1938) has shown that the improper Bayesian estimator of θ^m based on a θ^{-1-m} prior density is unbiased provided $m > 0$. We shall denote this estimator by T_m and call it the Pitman estimator of θ^m. The betting procedure given in (5.1) is semirelevant for T_m provided $a > 0$ and there should again be results showing that relevant betting procedures do not exist if certain conditions hold.

6. Conditional properties of statistical procedures in situations involving both location and scale parameters. In this section we make no attempt to discuss general situations involving location and scale parameters. We restrict ourselves to one or two normal distributions with unknown location and scale parameters.

For a single normal distribution with unknown mean, μ, and variance, σ^2, Stein (1961) showed that the usual interval estimators for μ based on the t-distribution allow the existence of positively biased relevant selections but not negatively biased relevant selections. Buehler and Fedderson (1963) exhibited a positively biased relevant subset for an interval estimator based on two observations and Brown (1967) has shown that positively biased relevant subsets exist for any number of observations. A published proof of the nonexistence of negatively biased relevant selections appears in Robinson (1976). Cohen (1972) has discussed the length of interval estimators of σ^2, but nobody has investigated their conditional properties.

Even less is known about the conditional properties of statistical procedures in situations involving two normal populations. Fisher (1956) showed that negatively biased relevant subsets exist for tests of hypotheses and interval estimators based on the ideas of Welch (1947). Robinson (1976) has shown that there are no negatively biased relevant selections for interval estimators based on the Behrens-Fisher solution to the two means problem. Presumably positively biased relevant selections exist for the Behrens-Fisher solution (possibly based on the relevant subsets for the t-distribution), but no other conditional properties seem to have ever been discussed for this situation.

The only definite result of this section concerns

$$s^2 = \frac{1}{n-1}\Sigma_{i=1}^n (x_i - \bar{x})^2$$

as a point estimator of σ^2. Here $X_1, X_2, \cdots, X_n$ are $n(\geqslant 2)$ independent observations from a normal population with mean μ and variance σ^2, both unknown, and $\bar{X}$ is the average of $X_1, X_2, \cdots, X_n$. We show the relevant subsets exist for s^2.

PROPOSITION 6.1. *For any number δ such that $0 < \delta < 1/(n-1)$ there is a number a such that*

$$E\left[s^2 \mid |\bar{X}|/s < a\right] \geqslant (1+\delta)\sigma^2.$$

PROOF. It is clear that $\sigma^{-2}E[s^2 \mid |\bar{X}/s| < a]$ is dependent only upon the ratio μ/σ. Therefore, we can assume without loss of generality that $\sigma = 1$. Denoting the conditional density of s given $|\bar{X}|/s < k, \mu, \sigma^2 = 1$ by $f(s \mid |\bar{X}|/s < k, \mu)$, equation (6.1) of Brown (1968) states that $f(s \mid |\bar{X}|/s < k, \mu)$ has strict monotone likelihood ratio with respect to $f(s \mid |\bar{X}|/s < k, 0)$. Hence

$$E\left[s^2 \mid |\bar{X}|/s < k, \mu, \sigma = 1\right] \geqslant E\left[s^2 \mid |\bar{X}|/s < k, \mu = 0, \sigma = 1\right].$$

Now $E[s^2 \mid |\bar{X}|/s < k, \mu = 0, \sigma = 1]$ is a continuous function of k for $0 < k < \infty$;

$$\lim_{k\to\infty} E\left[s^2 \mid |\bar{X}|/s < k, \mu = 0, \sigma = 1\right] = 1$$

since s^2 is an unbiased estimator of σ^2; and

$$\lim_{k\to 0} E\left[s^2 \mid |\bar{X}|/s < k, \mu = 0, \sigma = 1\right] = \frac{E\left[s^3 \mid \mu = 0, \sigma = 1\right]}{E\left[s \mid \mu = 0, \sigma = 1\right]} = \frac{n}{n-1},$$

where we have used the independence of $\bar{X}$ and s for $\mu = 0$ and that the range of values of $\bar{X}$ for given s is of length $2ks$. The result follows.

Stein's positively biased relevant selection and Buehler and Fedderson's and Brown's relevant subsets for interval estimators for μ are also based on the ratio $|\bar{X}|/s$. The lengths of confidence intervals based on the t-distribution are always multiples of s, so it seems reasonable that in a subset where s^2 seems to be too large the probability of covering μ seems also to be too large.

It is natural to conjecture that there is no $\varepsilon > 0$ and conditioning set C such that

$$E\left[s^2 \mid \mathbf{X} \in C\right] \leqslant (1-\varepsilon)\sigma^2 \qquad \text{for all} \quad \mu, \sigma^2.$$

A proof along the lines of the proof of Theorem 1 in Robinson (1976) should be fairly straightforward.

It is also natural to conjecture that there is no $\delta > 1/(n-1)$ and conditioning set C such that

$$E[s^2|\mathbf{X} \in C] \geqslant (1+\delta)\sigma^2 \qquad \text{for all} \quad \mu, \sigma^2,$$

since

$$\frac{n-1}{n}s^2 = \frac{1}{n}\Sigma_{i=1}^n(x_i - \bar{x})^2 \leqslant \frac{1}{n}\Sigma_{i=1}^n x_i^2$$

and $(1/n)\Sigma x_i^2$, being the improper Bayesian estimator of σ^2 based on a σ^{-2} prior density for σ^2 with $\mu = 0$, should have good conditional properties as an estimator of σ^2.

7. Discussion. A major issue raised by this paper and its sister paper, Robinson (1979), may be most clearly seen in the light of the results on admissibility and conditional properties of the multidimensional normal mean. Within decision theory, Stein's (1956) result that the three-dimensional normal mean is inadmissible seems unavoidably applicable, so that if one rejects the idea that all decision-making ought to be Bayesian then one must search the complexities of Efron and Morris (1973) and similar work for methods of estimation. However, within inference it may be argued that admissibility is not an applicable concept, that conditional properties, like the nonexistence of negatively biased relevant selections for interval estimation, are more appropriate, that the multidimensional normal mean is a perfectly reasonable statistical procedure and that there is no need to combine unrelated estimation problems.

APPENDIX

Here two elementary results are established.

LEMMA 1. *If* $0 \leqslant x \leqslant 1$ *and* $0 \leqslant (x+z)/(1+y) \leqslant 1$, *then*

$$|x - (x+z)/(1+y)| \leqslant 3\max\{|y|, |z|\}.$$

PROOF. If $|y| < \frac{1}{3}$ then

$$\begin{aligned} |x - (x+z)/(1+y)| &\leqslant |z| + |x + z - (x+z)/(1+y)| \\ &= |z| + |(x+z)y/(1+y)| \leqslant |z| + \tfrac{3}{2}|xy + zy| \\ &\leqslant |z| + \tfrac{3}{2}|y| + \tfrac{1}{2}|z| \leqslant 3\max\{|y|, |z|\}. \end{aligned}$$

If $|y| \geqslant \frac{1}{3}$ then

$$3\max\{|y|, |z|\} \geqslant 1 \geqslant |x - (x+z)/(1+y)|.$$

LEMMA 2. *For all real* m, x *and* θ *such that* $m \neq 0$

$$\begin{aligned} &\left\{1 + \frac{x^2}{m^2}\right\}^{-1}\left\{1 - \frac{2x}{m^2+x^2}(\theta - x) - \frac{3}{m^2}(\theta - x)^2\right\} \\ &\qquad \leqslant \left\{1 + \frac{\theta^2}{m^2}\right\}^{-1} \leqslant \left\{1 + \frac{x^2}{m^2}\right\}^{-1}\left\{1 - \frac{2x}{m^2+x^2}(\theta - x) + \frac{3}{m^2}(\theta - x)^2\right\}. \end{aligned}$$

PROOF. Subtracting $\{1 + \theta^2/m^2\}^{-1}$ from each of the three expressions, multiplying by $(m^2 + x^2)^2(m^2 + \theta^2)$ and rearranging terms shows that the pair of inequalities is equivalent to

$$- 3x^2\theta^2(x - \theta)^2 - m^2(x - \theta)^2(4x^2 + 2x\theta + 3\theta^2) - 2m^4(x - \theta)^2$$
$$\leqslant 0 \leqslant 3x^2\theta^2(x - \theta)^2 + m^2(x - \theta)^2(2x^2 - 2x\theta + 3\theta^2) + 4m^2(x - \theta)^2.$$

This pair of inequalities is clearly valid.

REFERENCES

[1] BROWN, L. D. (1967). The conditional level of Student's test. *Ann. Math. Statist.* **38** 1068–1071.

[2] BROWN, L. D. (1968). Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Ann. Math. Statist.* **39** 29–48.

[3] BUEHLER, R. J. (1959). Some validity criteria for statistical inference. *Ann. Math. Statist.* **30** 845–863.

[4] BUEHLER, R. J. and FEDDERSON, A. P. (1963). Note on a conditional property of Student's *t*. *Ann. Math. Statist.* **34** 1098–1100.

[5] COHEN, A. (1972). Improved confidence intervals for the variance of a normal distribution. *J. Amer. Statist. Assoc.* **67** 382–387.

[6] EFRON, B. and MORRIS, C. (1973). Combining possibly related estimation problems. *J. Roy. Statist. Soc. Ser. B* **35** 379–421

[7] FISHER, R. A. (1956). On a test of significance in Pearson's Biometrika Table No. 11. *J. Roy. Statist. Soc. Ser. B* **18** 56–60.

[8] PITMAN, E. J. G. (1938). The estimation of location and scale parameters of a continuous population of any given form. *Biometrika* **30** 391–421.

[9] ROBINSON, G. K. (1976). Conditional properties of Student's *t* and of the Behrens-Fisher solution to the two means problem. *Ann. Statist.* **4** 963–971.

[10] ROBINSON, G. K. (1979). Conditional properties of statistical procedures. *Ann. Statist.* **7** 742–755.

[11] STEIN, C. (1956). Inadmissibility of the usual estimation for the mean of a multivariate normal distribution. *Proc. Third Berkeley Symp. Math. Statist. Probability* 197–206.

[12] STEIN, C. (1961). Approximation of prior measures by probability measures. *Inst. Math. Statist., Wald Lectures. Unpublished manuscript.*

[13] STEIN, C. (1964). Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Ann. Inst. Statist. Math.* **16** 155–160.

[14] WELCH, B. L. (1947). Generalization of Student's problem. *Biometrika* **34** 28–35.

[15] ZACKS, S. (1971). *The Theory of Statistical Inference.* Wiley, New York.

AUSTRALIAN ROAD RESEARCH BOARD
P.O. BOX 156, BAG 4,
NUNANADING 3131
VICTORIA, AUSTRALIA