The Annals of Statistics
1975, Vol. 3, No. 2, 451-456

MINIMAX INTERVAL ESTIMATORS OF LOCATION PARAMETERS

BY BENJAMIN ZEHNWIRTH

University of Melbourne

Minimax interval estimators of location parameters analogous to the Pitman point estimator are found by employing invariance in decision theory. The half normal distribution is used as an illustrative example.

1. Summary and introduction. The present investigation is concerned with the problem of interval estimation from a decision theoretic point of view.

Valand [7] employed the invariance principle in the particular case where the observable random variable X has density function $f_X(x; \theta) = f(x - \theta)$ for some density f. Since it is recognized that more than one observation is usually required to gain substantial knowledge about the parameter θ, X must generally be regarded as a one-dimensional sufficient statistic for θ. We are therefore led to investigate a class of problems for which a one-dimensional sufficient statistic does not necessarily exist; the random sample comes from a distribution with θ a location parameter. Such a situation arises, for example, with the half-normal distribution.

2. Framework and notation. We adopt a formulation similar to that of Valand's paper [7], but with several modifications. Briefly, a decision problem involves the following basic elements:

(i) $\mathscr{X}$—a sample space of a random variable $\mathbf{X}$.
(ii) Θ—a parameter space.
(iii) $\mathscr{A}$—an action space.
(iv) L—a function on $\Theta \times \mathscr{A}$ into the real line called a loss function.
(v) $t(\mathbf{x}; \theta)$—a probability density function for each $\theta \in \Theta$.

The principle of invariance is centred around groups of transformations G, $\bar{G}$, $\tilde{G}$ acting on $\mathscr{X}$, Θ, $\mathscr{A}$ respectively. The reader is referred to Blackwell and Girshick [1] for a detailed discussion. In the present paper we suppose that:

(a) $\mathscr{X} = R^n$, the n-dimensional Euclidean space.
(b) $\Theta = R^1$, the real line.
(c) $\theta \in \Theta$ is a location parameter for the random variable $\mathbf{X} = (X_1, \cdots, X_n)$ with probability density function $t(\mathbf{x}; \theta) = t(x_1 - \theta, \cdots, x_n - \theta)$ for some n-dimensional joint density t. $X_1, \cdots, X_n$ need not be independent.
(d) $\mathscr{A} = \{(a, b) : a \leq b\}$ the set of all open intervals.

Received April 1972; revised April 1974.
AMS 1970 *subject classifications.* Primary 62A05; Secondary 62C99.
Key words and phrases. Minimax estimator, location parameter, interval estimator invariance.

(e) $L(\theta, (a, b)) = g(b - a) + 1 - I_{(a,b)}(\theta)$ where

(i) g is a nonnegative, non-decreasing function on the real line;

(ii) $I_{(a,b)}(\theta)$ is the indicator function of the interval (a, b).

This decision problem is invariant under G, the additive group on R^n.

3. Characterization of the best invariant rule. Let $U(\mathbf{X})$ possess the translation property and let $V(\mathbf{X})$ be a maximal invariant on R^n under G. Then $(U - b_1(V), U - b_2(V))$, where b_1 and b_2 are functions into the real line is an invariant rule. The risk $R(b_1, b_2)$ associated with this rule is given by

$$R(b_1, b_2) = E_0[L(0, (U - b_1(V), U - b_2(V)))], \tag{3.1}$$

where E_0 denotes expectation when $\theta = 0$. The distribution of $U \mid V$ has θ as a location parameter. The best invariant rule for this conditional problem under the additive group on R^1 is given by

$$\delta(b_1^*, b_2^*) = (U - b_1^*(V), U - b_2^*(V)) \tag{3.2}$$

where $b_1^*(V)$, $b_2^*(V)$ are numbers (provided they exist) such that

$$\begin{aligned} R_0(V) &= E_0[L(0, (U - b_1^*, U - b_2^*)) \mid V] \\ &= \inf_{(b_1, b_2)} E_0[L(0, (U - b_1, U - b_2)) \mid V]. \end{aligned} \tag{3.3}$$

Note that $b_1^*(V)$ and $b_2^*(V)$ are each defined separately for each possible value of the random variable $V(\mathbf{X})$.

The risk $\bar{R}(b_1, b_2)$ associated with an invariant rule in the conditional problem is given by

$$\bar{R}(b_1, b_2) = E_0[L(0, (U - b_1, U - b_2)) \mid V]. \tag{3.4}$$

The following theorem whose proof appears in Ferguson [2] permits us to work with an appropriate one-variable problem instead of the original problem.

THEOREM 1. *The best invariant rule* (3.2) *in the conditional problem is the best invariant rule in the original problem.*

4. Minimax property of the best invariant rule. It is the purpose of this section to investigate the minimax behavior of the best invariant rule $\delta(b_1^*, b_2^*)$. Since $(X_2 - X_1, \cdots, X_n - X_1)$ is a maximal invariant on R^n under G it is reasonable to set $V(\mathbf{X}) = (V_1, \cdots, V_r)$ where $r \leqq n$ and each V_i is a function of the X_i's. See Wijsman [9]. In addition let $F(U)$ designate the distribution function of $U \mid V = v$ when $\theta = 0$.

THEOREM 2. *In the problem of interval estimation described in Section* 2, *if for every $\varepsilon > 0$ and realization v of V there exists a finite number $N = N_\varepsilon(v)$ such that $N_\varepsilon(V)$ is measurable in $(V_1, \cdots, V_r)$ and*

$$R_N(b_1, b_2) = \int_{-N}^{N} L(0, (U - b_1, U - b_2))\, dF(U) \geqq R_0(v) - \varepsilon \tag{4.1}$$

for all $(b_2, b_1) \in \mathscr{A}$ then $(U - b_1^*, U - b_2^*)$ is a minimax rule among all the non-randomized rules.

PROOF. The proof follows by an argument similar to that used by Ferguson [2] page 189.

We are now interested in finding sufficient conditions for expression (4.1) and the measurability hypothesis to hold. The following theorem which is the main result of this paper deals with this.

THEOREM 3. *If $f(U) = F'(U)$ is continuous in $(V_1, \cdots, V_r)$ almost everywhere and there exists an even, nonnegative function $P(x)$ on the real line such that*

(i) *if $0 \leqq x_1 \leqq x_2$, $P(x_1) \leqq P(x_2)$,*
(ii) *$E_0[P(U) \mid V] < \infty$ for all realizations of V,*
(iii) *$P(x)^{-1} = O(x^{-\gamma})$ for some $\gamma > 0$,*
(iv) *$(g(2x) + 1)P(x)^{-1} = O(x^{-\delta})$ for some $\delta > 0$,*

then the conditions of Theorem 2 are satisfied.

The proof of this result relies fundamentally on the properties of g. It is quite long and so will be split into several parts.

PROOF. Recalling the definition of $R_0(V)$ we see that

$$R_0(v) \leqq R(b_1, b_2) \qquad \text{for all} \quad (b_2, b_1) \in \mathscr{A}. \tag{4.2}$$

Now, for any $N > 0$,

$$\textstyle\int_{-N}^{N} dF(U) \geqq 1 - P(N)^{-1}E_0[P(U) \mid V = v]. \tag{4.3}$$

This is a form of Chebyshev inequality. Combining (4.2) and (4.3) leads to the following inequality:

$$\begin{aligned} R_0(v) \leqq {} & (g(b_1 - b_2) + 1)(\textstyle\int_{-N}^{N} dF(U) + P(N)^{-1}E_0[P(U) \mid V = v]) \\ & - \textstyle\int_{b_2}^{b_1} dF(U). \end{aligned} \tag{4.4}$$

The remainder of the proof is divided into several parts.

(i) We first show that for every $\varepsilon > 0$, there exists $N_\varepsilon^1(v)$ such that $R_N(b_1, b_2) \geqq R_0(v) - \varepsilon$ if $N \geqq N_\varepsilon^1(v)$ and $-N \leqq b_2 \leqq b_1 \leqq N$.
In view of the monotonic property of g we have from (4.4)

$$R_N(b_1, b_2) \geqq R_0(v) - \varepsilon \qquad \text{if} \quad P(N)^{-1}(g(2N) + 1)E_0[P(U) \mid V = v] \leqq \varepsilon. \tag{4.5}$$

Therefore, by applying condition (iv) we have $R_N(b_1, b_2) \geqq R_0(v) - \varepsilon$ if $-N \leqq b_2 \leqq b_1 \leqq N$ and $N \geqq N_\varepsilon^1(v) = [AE_0[P(U) \mid V = v]/\varepsilon]^{1/\delta}$; A is some constant.

(ii) We now show that $R_N(\bar{b}_1, b_2) \geqq R_0(v) - \varepsilon$ for $-N \leqq b_2 \leqq b_1 \leqq N < \bar{b}_1$ and $N \geqq N_\varepsilon^1(v)$.

By definition and the monotonic properties of g,

$$\begin{aligned} R_N(\bar{b}_1, b_2) &= (g(\bar{b}_1 - b_2) + 1) \textstyle\int_{-N}^{N} dF(U) - \int_{b_2}^{N} dF(U) \\ &\geqq (g(b_1 - b_2) + 1) \textstyle\int_{-N}^{N} dF(U) - \int_{b_2}^{N} dF(U) \\ &\geqq R_0(v) - \varepsilon \qquad \text{from part (i).} \end{aligned}$$

(iii) Similarly it can be shown that $R_N(b_1, \bar{b}_2) \geqq R_0(v) - \varepsilon$ for $\bar{b}_2 < -N \leqq b_2 \leqq b_1 \leqq N$ and $N \geqq N_\varepsilon^1(v)$, and

(iv) it can also be shown that $R_N(\bar{b}_1, \bar{b}_2) \geqq R_0(v) - \varepsilon$ for $\bar{b}_2 < N \leqq b_2 \leqq b_1 \leqq N < \bar{b}_1$ and $N \geqq N_\varepsilon^1(v)$ by using the monotonicity property of g and parts (ii) and (iii) of the proof.

The following two cases are somewhat different from the foregoing.

(v) Let $\bar{b}_2 \leqq \bar{b}_1 < -N \leqq b_2 \leqq b_1 \leqq N$ where N will be chosen appropriately. Now,

$$R_N(\bar{b}_1, \bar{b}_2) = (g(\bar{b}_1 - \bar{b}_2) + 1) \int_{-N}^{N} dF(U) .$$

Applying (4.3) leads to

$$R_N(\bar{b}_1, \bar{b}_2) \geqq (g(b_1 - \bar{b}_2) + 1)(1 - P(N)^{-1}E_0[P(U) \mid V = v]) . \tag{4.6}$$

But in view of (3.3) and Theorem 1,

$$\begin{aligned} R(\bar{b}_1, \bar{b}_2) &= (g(\bar{b}_1 - \bar{b}_2) + 1) - \textstyle\int_{\bar{b}_2}^{\bar{b}_1} dF(U) \\ &\geqq R_0(v) . \end{aligned}$$

So,

$$\begin{aligned} g(\bar{b}_1 - \bar{b}_2) + 1 &\geqq R_0(v) + \textstyle\int_{\bar{b}_2}^{\bar{b}_1} dF(U) \\ &\geqq R_0(v) . \end{aligned}$$

Combining the last inequality with (4.6) we have

$$R_N(\bar{b}_1, \bar{b}_2) \geqq R_0(v) - \varepsilon \qquad \text{if} \quad P(N)^{-1}R_0(v)E_0[P(U) \mid V = v] \leqq \varepsilon .$$

Therefore, by using condition (iii) we have,

$$\begin{aligned} R_N(\bar{b}_1, \bar{b}_2) \geqq R_0(v) - \varepsilon \qquad &\text{if} \quad \bar{b}_2 \leqq \bar{b}_1 < -N \quad \text{and} \\ &N \geqq N_\varepsilon^2(v) = [BR_0(v)E_0[P(U) \mid V = v]/\varepsilon]^{1/\gamma} ; \end{aligned}$$

B is some constant.

(vi) If $-N \leqq b_2 \leqq b_1 \leqq N < \bar{b}_2 \leqq \bar{b}_1$, then $R_N(\bar{b}_1, \bar{b}_2) \geqq R_0(v) - \varepsilon$ so long as $N \geqq N_\varepsilon^2(v)$. The proof follows from a similar argument to that in part (v). Now, let $N_\varepsilon(v) = \max(N_\varepsilon^1(v), N_\varepsilon^2(v))$; then a simple argument will convince the reader that if $N \geqq N_\varepsilon(v)$, then

$$R_N(b_1, b_2) \geqq R_0(v) - \varepsilon \qquad \text{for all} \quad (b_2, b_1) \in \mathscr{A} .$$

Finally let us turn to the measurability part of the proof. $R_0(V)$ is measurable since it is the infimum of a limit of measurable functions. Similarly, $E_0[P(U) \mid V]$ is measurable since it is the limit of measurable functions. It therefore follows that $N_\varepsilon(V)$ is measurable.

This completes the proof of the theorem.

At this point the reader may note that although the present paper falls within the framework of Kiefer's [5] or Wesler's [8] general considerations, their regularity conditions are not satisfied under our specific formulation. Theorem 3 replaces a verification of (2.4) of Kiefer's paper, the other assumptions there being trivial in this case, but this condition requires verification since our problem does not satisfy condition 4(b). As for Wesler's work [8], our problem does not meet his condition (ii) of the generalized Hunt–Stein theorem.

The reader should now focus attention on Valand's paper [6]. The function

h is more general than g but unfortunately there is an error in the proof of Lemma 2, page 197. It is unreasonable to assume that "a compact set A exists." In fact, if h is taken to be the function $(b - a)$ then it can be shown that such a set A cannot exist. However, the technique used in the proof of Theorem 3 can be applied to verify Lemma 2 provided that h is replaced by g.

We now turn to an example which is quite appealing from the point of view of the preceding theory.

5. Minimax interval estimator for the location parameter of the half-normal distribution. Consider the invariant decision problem described in Section 2 with

(i) $t(\mathbf{x}; \theta) = (2/\pi)^{n/2} \exp[-\sum_{i=1}^{n} (x_i - \theta)^2/2] I_{(\theta,\infty)}(\min x_i)$
(ii) $g(b - a) = b - a$.

Further, let

$$V(\mathbf{X}) = (X_2 - X_1, \cdots, X_n - X_1), \qquad U(\mathbf{X}) = X_1$$

and put $V_j = X_j - X_1, j = 2, \cdots, n$ with $V_1 = 0$. When $\theta = 0$ the probability density function $f(U)$ of $U \mid V$ is given by

$$\begin{aligned} f(U) &= t(U, U + V_2, \cdots, U + V_n)/\textstyle\int_{-\infty}^{\infty} t(U, U + V_2, \cdots, U + V_n)\, dU \\ &= \exp[-n(U + \bar{V})^2/2] I_{(-V_{(1)},\infty)}(U) / \textstyle\int_{-\infty}^{\infty} \exp[-n(U + \bar{V})^2/2] \\ &\quad \times I_{(-V_{(1)},\infty)}(U)\, dU \end{aligned} \tag{5.1}$$

where

$$\bar{V} = \sum_{i=1}^{n} V_i/n\,; \qquad V_{(1)} = \min_{i=1}^{n} V_i\,.$$

Let Φ be the distribution function of the standard normal distribution. Let $\alpha = [1 - \Phi(n^{\frac{1}{2}}(\bar{V} - V_{(1)}))]^{-1}$; then (5.1) can be re-cast in the form

$$f(U) = \alpha(n/2\pi)^{\frac{1}{2}} \exp[-n(U + \bar{V})^2/2] I_{(-V_{(1)},\infty)}(U)\,. \tag{5.2}$$

To find the best invariant rule we minimize

$$\bar{R}(b_1, b_2) = \textstyle\int_{-\infty}^{\infty} \{(b_1 - b_2) + 1 - I_{(U-b_1, U-b_2)}(0)\} f(U)\, dU\,.$$

Setting $\partial\bar{R}(b_1, b_2)/\partial b_1 = 0$, $\partial\bar{R}(b_1, b_2)/\partial b_2 = 0$, we have

$$f(b_1) = 1 \qquad \text{and} \qquad f(b_2) = 1\,. \tag{5.3}$$

Solutions to (5.3) will be denoted by (b_1^*, b_2^*) where $b_1^* \geqq b_2^*$. Now, if $f(b) = 1$ then

$$(n/2\pi)^{\frac{1}{2}} \exp[-n(b + \bar{V})^2/2] I_{(-V_{(1)},\infty)}(b) = \alpha^{-1}\,. \tag{5.4}$$

It follows that

$$b_1^* = -\bar{V} + [(2/n) \log(n\alpha/2\pi)]^{\frac{1}{2}} \tag{5.5}$$

and

$$b_2^* = -\bar{V} - [(2/n) \log(n\alpha/2\pi)]^{\frac{1}{2}}\,. \tag{5.6}$$

To verify that (b_1^*, b_2^*) given by (5.5) and (5.6) is the value of (b_1, b_2) that minimizes $\bar{R}(b_1, b_2)$ we proceed as follows:

From (5.3)

$$\frac{\partial^2 \bar{R}(b_1, b_2)}{\partial b_1^2} = -\frac{\partial f(b_1)}{\partial b_1}; \qquad \frac{\partial^2 \bar{R}(b_1, b_2)}{\partial b_2^2} = \frac{\partial f(b_2)}{\partial b_2}; \qquad \frac{\partial^2 \bar{R}(b_1, b_2)}{\partial b_1 \partial b_2} = 0.$$

By substituting expressions (5.5) and (5.6) into $\partial f(b_1^*)/\partial b_1$ and $-\partial f(b_2^*)/\partial b_2$ respectively, we see that $\partial^2 \bar{R}(b_1^*, b_2^*)/\partial b_1^2 > 0$ and $\partial^2 \bar{R}(b_1^*, b_2^*)/\partial b_2^2 > 0$. So (b_1^*, b_2^*) gives the minimum point for $\bar{R}(b_1, b_2)$.
For this specific example $g(x) = x$ and by letting $P(x) = x^2$, say, all the conditions of Theorem 3 are clearly satisfied. Hence

$$\bar{X} - [(2/n) \log (n\alpha/2\pi)]^{\frac{1}{2}}, \qquad \bar{X} + [(2/n) \log (n\alpha/2\pi)]^{\frac{1}{2}})$$

is a minimax interval estimator for θ.

In conclusion we remark that only a minor change is required for the case $g(x) = cx$, $c > 0$.

Acknowledgment. I am indebted to the referee for some valuable suggestions.

REFERENCES

[1] BLACKWELL, D. and GIRSHICK, M. A. (1954). *Theory of Games and Statistical Decisions.* Wiley, New York.

[2] FERGUSON, T. S. (1967). *Mathematical Statistics, A Decision Theoretic Approach.* Academic Press, New York.

[3] GIRSHICK, M. A. and SAVAGE, L. J. (1951). Bayes and minimax estimates for quadratic loss functions. *Proc. Second Berkeley Symp. Math. Statist. Prob.* Univ. of California Press. 53–73.

[4] HALL, W. J., WIJSMAN, R. A. and GHOSH, J. K. (1965). The relationship between sufficiency and invariance with applications in sequential analysis. *Ann. Math. Statist.* **36** 575–614.

[5] KIEFER, J. (1954). Invariance, minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* **27** 887–906.

[6] PITMAN, E. J. G. (1939). The estimation of location and scale parameters of a continuous population. *Biometrika* **30.**

[7] VALAND, R. S. (1968). Invariant interval estimation of a location parameter. *Ann. Math. Statist.* **39** 193–199.

[8] WESLER, O. (1959). Invariance theory and a modified minimax principle. *Ann. Math. Statist.* **30** 1–20.

[9] WIJSMAN, R. A. (1967). Cross-sections of orbits and their application to densities of maximal invariants. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* Univ. of California Press. 389–400.

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE
ROYAL MELBOURNE INSTITUTE OF TECHNOLOGY
124 LATROBE STREET
MELBOURNE, C1, AUSTRALIA