THE ADMISSIBILITY OF PITMAN'S ESTIMATOR OF A SINGLE LOCATION PARAMETER[1]

By Charles Stein

Stanford University and University of California, Berkeley[2]

1. Introduction. Pitman [1] gave a thorough discussion of the problem of estimating the location and scale parameters of a distribution which is known except for one or both of these parameters. In particular, if $X_1 \cdots X_n$ are real random variables independently and identically distributed according to the density $r(x - \xi)$ (with respect to Lebesgue measure), where ξ is unknown but the function r is known, Pitman shows that the estimator

$$\hat{\xi}_0(X_1, \cdots, X_n) = \frac{\int \xi \prod r(X_i - \xi)\, d\xi}{\int \prod r(X_i - \xi)\, d\xi} \tag{1.1}$$

is the best translation-invariant estimator in the sense that it minimizes $\mathcal{E}_\xi[\hat{\xi}(X_1 \cdots X_n) - \xi]^2$ among all estimators $\hat{\xi}$ for which

$$\hat{\xi}(x_1 + c, \cdots, x_n + c) = \hat{\xi}(x_1, \cdots, x_n) + c \tag{1.2}$$

for all $x_1, \cdots, x_n$ and c. Girshick and Savage [2] showed that $\hat{\xi}_0$ is minimax in the class of all estimators (not restricted by (1.2)) and this also follows from the later more general results of Kudo [3] and Kiefer [4]. Karlin [5] has shown that under certain conditions $\hat{\xi}_0$ is admissible, that is, if $\hat{\xi}$ is any estimator for which

$$E_\xi(\hat{\xi}(X_1, \cdots, X_n) - \xi)^2 \leqq E_\xi(\hat{\xi}_0(X_1, \cdots, X_n) - \xi)^2 \tag{1.3}$$

for all ξ, then equality holds for all ξ. Since his conditions are fairly strong, and his method somewhat special, it seems desirable to present an alternative proof. Theorem 1 of Section 2, when reformulated for the present slightly special case, becomes

Theorem. *If*

$$\int \prod r(x_i) \left\{ \frac{\int \xi^2 \prod r(x_i - \xi)\, d\xi}{\int \prod r(x_i - \xi)\, d\xi} - \left(\frac{\int \xi \prod r(x_i - \xi)\, d\xi}{\int \prod r(x_i - \xi)\, d\xi} \right)^2 \right\}^{3/2} \prod dx_i < \infty \tag{1.4}$$

then $\hat{\xi}_0$ *defined by* (1.1) *is admissible.*

Received August 25, 1958.

[1] This work was sponsored by the Office of Naval Research.

[2] The author is now at Stanford University.

The condition (1.4) is not very strong. For example, if there is any translation-invariant estimator $\hat{\xi}$ for which $\mathcal{E}_\xi |\hat{\xi} - \xi|^3 < \infty$, then (1.4) holds. For the Cauchy distribution $r(x) = 1/\pi(1 + x^2)$ with $n \geqq 7$, this is true with $\hat{\xi}$ equal to the sample median.

The proof is given by a method first used by Blyth [6], and the result seems to be the best possible obtainable by this method. Here, as in Lehmann and Stein [7], roughly speaking, the theorem requires one more moment than is clearly relevant. In [7] a first moment is required, although it is a testing problem, and here, a third moment rather than a second. It would be interesting to know whether conditions of this type are necessary. Essentially the same method will be applied in a paper, now being prepared, to the problem of estimating two unknown location parameters with quadratic loss. There it is necessary to vary the form, as well as the scale, of the *a priori* distribution (see the argument around (2.16)). The bivariate normal case has already been treated by the author in [8]. For three or more translation parameters with positive definite quadratic loss, Pitman's estimator is not admissible. This was proved in the normal case in [8]. While it is of some theoretical interest to prove the admissibility of the natural estimator when it is admissible, the careful study of other estimators when the natural estimator is not admissible has greater practical value.

It may be useful to indicate the correspondence between the notation used in this introduction and that of the slightly more general problem treated in the remainder of the paper. Let $\mathcal{Y}$ be the $n - 1$ dimensional real coordinate space, $\mathcal{C}$ the σ-algebra of all Borel subsets of $\mathcal{Y}$ and ν the distribution of Y defined by (1.9).

$$f(y) = \frac{\int xr(x)r(x + y_1) \cdots r(x + y_{n-1})\, dx}{\int r(x)r(x + y_1) \cdots r(x + y_{n-1})\, dx} \tag{1.5}$$

and

$$g(y) = \int r(x)r(x + y_1) \cdots r(x + y_{n-1})\, dx, \tag{1.6}$$

where

$$y = (y_1, \cdots, y_{n-1}). \tag{1.7}$$

Also, let

$$p(x, y) = \frac{r(x + f(y))r(x + f(y) - y_1) \cdots r(x + f(y) - y_{n-1})}{g(y)}. \tag{1.8}$$

Then conditions (2.1) and (2.2) are satisfied by p, and (1.4) reduces to (2.3). If we define the random point (X, Y) by

$$\begin{aligned} Y_1 &= X_2 - X_1, \\ &\vdots \\ Y_{n-1} &= X_n - X_1, \end{aligned} \tag{1.9}$$

$$X = X_1 - f(Y_1, \cdots, Y_{n-1}), \tag{1.10}$$

then the estimate X, proved admissible in Section 2, is seen to reduce to $\hat{\xi}_0(X_1, \cdots, X_n)$.

2. The results. Let $\mathcal{B}$ be the σ-algebra of all Borel subsets of the real line $\mathfrak{X}$, and $\mathcal{C}$ a σ-algebra of subsets of a set $\mathcal{Y}$. Let μ be Lebesgue measure on $\mathcal{B}$ and ν a probability measure on $\mathcal{C}$. Let p be a nonnegative valued $\mathcal{BC}$ measurable function on $\mathfrak{X} \times \mathcal{Y}$ such that

$$\left.\begin{aligned} &\int p(x, y)\, dx = 1 \qquad (2.1)\\ &\int xp(x, y)\, dx = 0 \qquad (2.2)\end{aligned}\right\} \text{ for all } y$$

$$\int d\nu(y) \left(\int x^2 p(x, y)\, dx\right)^{3/2} < \infty, \tag{2.3}$$

where we write dx instead of $d\mu(x)$. Then of course p is a probability density with respect to $\mu\nu$. We shall prove

THEOREM 1. *Under the above hypotheses, if we observe (X, Y) distributed so that, for some unknown ξ, $(X - \xi, Y)$ has probability density p with respect to $\mu\nu$, then X is an admissible estimator of ξ with squared error as loss.*

In other words, if φ is any $\mathcal{BC}$ measurable function on $\mathfrak{X} \times \mathcal{Y}$ such that

$$\begin{aligned} &\int d\nu(y) \int [\varphi(x, y) - \xi]^2 p(x - \xi, y)\, dx \\ &\leqq \int d\nu(y) \int (x - \xi)^2 p(x - \xi. y)\, dx = \int d\nu(y) \int x^2 p(x, y)\, dx \end{aligned} \tag{2.4}$$

for all ξ, then the two sides of (2.4) are identically equal. Actually we prove the trivially stronger result that $\varphi(x, y) = x$ almost everywhere $(\mu\nu)$. One might hope to prove this result under the condition

$$\int d\nu(y) \int x^2 p(x, y)\, dx < \infty, \tag{2.3$'$}$$

which is weaker than (2.3). Of course (2.3′) is necessary, for otherwise we could take $\varphi(x, y) \equiv 0$.

We shall derive Theorem 1 from a slightly more general but weaker theorem. With $\mathfrak{X}$, $\mathcal{Y}$, $\mathcal{B}$, $\mathcal{C}$, μ, ν as before, let P be a nonnegative valued $\mathcal{BC}$ measurable function on $\mathfrak{X} \times \mathcal{Y}$ such that, for each y, $P(\cdot, y)$ is a cumulative distribution function and

$$\int x\, d_x P(x, y) = 0 \tag{2.5}$$

$$\int d\nu(y) \left(\int x^2\, d_x P(x, y)\right)^{3/2} < \infty \tag{2.6}$$

THEOREM 2. *Under the above hypotheses, if we observe* (X, Y) *distributed so that, for some unknown* ξ, Y *is distributed according to* ν *and the conditional cumulative distribution function of* $X - \xi$ *given* Y *is* $P(\cdot, Y)$, *then* X *is an almost admissible estimator of* ξ *with squared error as loss. That is, if* φ *is any* $\mathcal{B}\mathcal{C}$ *measurable function on* $\mathfrak{X} \times \mathcal{Y}$ *such that*

$$\begin{aligned}\int d\nu(y) \int [\varphi(x, y) - \xi]^2 \, d_x P(x - \xi, y) \\ \leqq \int d\nu(y) \int (x - \xi)^2 \, d_x P(x - \xi, y) = \int d\nu(y) \int x^2 \, d_x P(x, y)\end{aligned} \tag{2.7}$$

for all ξ, *then the two sides are equal for almost all* ξ.

By a familiar argument, we observe that Theorem 1 follows from Theorem 2, if we put

$$P(x, y) = \int_{-\infty}^{x} p(t, y) \, dt. \tag{2.8}$$

If φ satisfies the hypotheses of Theorem 1, it also satisfies those of Theorem 2 and we conclude that in (2.4) equality holds for almost all ξ. Now suppose that contrary to the conclusion of Theorem 1,

$$\mu\nu(S) > 0, \tag{2.9}$$

where

$$S = \{(x, y) : \varphi(x, y) \neq x\}. \tag{2.10}$$

Then, for all ξ in a set T of positive measure,

$$\int d\nu(y) \int_{S_y} p(x - \xi, y) \, dx > 0, \tag{2.11}$$

where $S_y = \{x : \varphi(x, y) \neq x\}$, since

$$\begin{aligned}\int d\xi \int d\nu(y) \int_{S_y} p(x - \xi, y) \, dx \\ = \int d\nu(y) \int_{S_y} dx \int p(x - \xi, y) \, d\xi = \int d\nu(y) \int_{S_y} dx = \mu\nu(S).\end{aligned} \tag{2.12}$$

Let

$$\varphi_0(x, y) = \tfrac{1}{2}(x + \varphi(x, y)). \tag{2.13}$$

Then

$$[\varphi_0(x, y) - \xi]^2 \leqq \tfrac{1}{2}\{[\varphi(x, y) - \xi]^2 + (x - \xi)^2\} \tag{2.14}$$

with strict inequality whenever $\varphi(x, y) \neq x$. It follows that we have strict inequality in (2.4) and thus in (2.7) for all $\xi \varepsilon T$ contradicting the conclusion of Theorem 2. An example given by Blackwell [9] with $\mathcal{Y}$ reducing to a point and P concentrated on a finite set shows that in Theorem 2 we cannot conclude admissibility.

To prove Theorem 2 we suppose the conclusion does not hold, that is, we suppose (2.7) holds with strict inequality for ξ in a set S having positive Lebesgue measure. For $\epsilon > 0$, let S_ϵ be the set of ξ for which

$$\int d\nu(y) \int [\varphi(x, y) - \xi]^2 \, dP(x - \xi, y) \leqq \int d\nu(y) \int x^2 \, dP(x, y) - \epsilon. \tag{2.15}$$

Since $S = US_\epsilon$, S_ϵ will have positive Lebesgue measure for sufficiently small ϵ, and we suppose ϵ chosen so that $\mu(S_\epsilon) > 0$. Since S_ϵ (like any measurable set) is of density 1 at almost all points of itself (see for example Titchmarsh [10], p. 371), there exists $\kappa > 0$ and an interval $I = (a - \kappa, a + \kappa)$ such that the set of $\xi \,\varepsilon\, I$ for which (2.15) holds has Lebesgue measure $\geqq \kappa$. There is no real loss of generality in assuming $I = (-\kappa, \kappa)$. Now we assign to ξ an *a priori* density $(1/\sigma)q(\xi/\sigma)$, taking for simplicity of computation

$$q(\xi) = \frac{1}{\pi(1 + \xi^2)}. \tag{2.16}$$

From (2.7), and the fact that (2.15) holds for a set of measure $\geqq \kappa$ in $(-\kappa, \kappa)$, it follows that

$$\mathcal{E}[\varphi(X, Y) - \xi]^2 \leqq \int d\nu(y) \int x^2 \, dP(x, y) - \frac{\epsilon\kappa}{2\pi\sigma} \tag{2.17}$$

for sufficiently large σ, where ξ has the indicated *a priori* distribution, and the conditional distribution of (X, Y) given ξ is that indicated before Theorem 2. However, we shall show that under the same distribution

$$\inf_{\psi} E[\psi(X, Y) - \xi]^2 \geqq \int d\nu(y) \int x^2 \, dP(x, y) - \frac{f(\sigma)}{\sigma}, \tag{2.18}$$

where

$$\lim_{\sigma \to \infty} f(\sigma) = 0. \tag{2.19}$$

For sufficiently large σ this contradicts (2.17).

We shall find the formula

$$\begin{aligned} &\int d\nu(y) \int x^2 \, dP(x, y) - \inf_{\psi} E[\psi(X, Y) - \xi]^2 \\ &\qquad = \frac{1}{\sigma} \int d\nu(y) \int dx \frac{\left[\int \eta q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)\right]^2}{\int q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)} \end{aligned} \tag{2.20}$$

useful in proving (2.18). To prove (2.20) we first observe that

$$\begin{aligned}\inf_{\psi} E[\psi(X, Y) - \xi]^2 &= \inf_{\psi} EE\{[\psi(X, Y) - \xi]^2 \mid X, Y\} \\ &= E\{E[\xi \mid (X, Y)] - \xi\}^2,\end{aligned} \tag{2.21}$$

so that

$$\begin{aligned}
&\int d\nu(y) \int x^2\, dP(x, y) - \inf E[\psi(X, Y) - \xi]^2 \\
&= E(X - \xi)^2 - E\{E[\xi \mid (X, Y)] - \xi\}^2 \\
&= E\{X^2 - 2\xi X + (E[\xi \mid (X, Y)])^2\} = E\{X - E[\xi \mid (X, Y)]\}^2 \\
&= \frac{1}{\sigma} \int q\left(\frac{\xi}{\sigma}\right) d\xi \int d\nu(y) \int \\
&\qquad \cdot \left\{ x - \frac{\int \xi' q\left(\frac{\xi'}{\sigma}\right) d_{\xi'}\, P(x - \xi', y)}{\int q\left(\frac{\xi'}{\sigma}\right) d_{\xi'}\, P(x - \xi', y)} \right\}^2 d_x\, P(x - \xi, y) \\
&= \frac{1}{\sigma} \int d\nu(y) \int q\left(\frac{\xi}{\sigma}\right) d\xi \int \left[\frac{\int \eta q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)}{\int q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)} \right]^2 d_x\, P(x - \xi, y) \\
&= \frac{1}{\sigma}\, d\nu(y) \int q\left(\frac{\xi}{\sigma}\right) d\xi \int \left[\frac{\int \eta q\left(\frac{\xi + \eta' - \eta}{\sigma}\right) dP(\eta, y)}{\int q\left(\frac{\xi + \eta' - \eta}{\sigma}\right) dP(\eta, y)} \right]^2 dP(\eta', y) \qquad (2.22) \\
&= \frac{1}{\sigma} \int d\nu(y) \int dP(\eta', y) \int \left[\frac{\int \eta q\left(\frac{\xi + \eta' - \eta}{\sigma}\right) dP(\eta, y)}{\int q\left(\frac{\xi + \eta' - \eta}{\sigma}\right) dP(\eta, y)} \right]^2 q\left(\frac{\xi}{\sigma}\right) d\xi \\
&= \frac{1}{\sigma} \int d\nu(y) \int dP(\eta', y) \int \left[\frac{\int \eta q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)}{\int q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)} \right]^2 q\left(\frac{x - \eta'}{\sigma}\right) dx \\
&= \frac{1}{\sigma} \int d\nu(y) \int dx \frac{\left[\int \eta q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)\right]^2}{\int q\left(\frac{x - \eta}{\sigma}\right) dP(\eta, y)}.
\end{aligned}$$

Comparing this result with (2.17) and (2.18), we see that, in order to complete the proof of Theorem 2, we need only show that

$$\lim_{\sigma\to\infty} \int d\nu(y) \int \frac{\left[\int \eta q\left(\frac{x-\eta}{\sigma}\right) dP(\eta, y)\right]^2}{\int q\left(\frac{x-\eta}{\sigma}\right) dP(\eta, y)} dx = 0. \tag{2.23}$$

In order to prove (2.23) we consider the integral

$$\Phi(P, \sigma) = \int dx \frac{\left[\int \eta q\left(\frac{x-\eta}{\sigma}\right) dP(\eta)\right]^2}{\int q\left(\frac{x-\eta}{\sigma}\right) dP(\eta)} \tag{2.24}$$

and the function

$$\psi(\lambda, \sigma) = \sup_{P \varepsilon \mathfrak{U}_\lambda} \Phi(P, \sigma), \tag{2.25}$$

where $\mathfrak{U}_\lambda$ is the set of probability measures P for which

$$\int \eta \, dP(\eta) = 0 \tag{2.26}$$

$$\int \eta^2 \, dP(\eta) = \lambda. \tag{2.27}$$

As indicated earlier, we take

$$q(\xi) = \frac{1}{\pi(1+\xi^2)}, \tag{2.28}$$

but the basic formulas hold for an arbitrary q. We first observe that

$$\psi(\lambda, \sigma) = \sigma^3 \psi\left(\frac{\lambda}{\sigma^2}, 1\right), \tag{2.29}$$

since

$$\begin{aligned} \Phi(P, \sigma) &= \sigma^3 \int d\left(\frac{x}{\sigma}\right) \frac{\left[\int \frac{\eta}{\sigma} q\left(\frac{x-\eta}{\sigma}\right) dP(\eta)\right]^2}{\int q\left(\frac{x-\eta}{\sigma}\right) dP(\eta)} \\ &= \sigma^3 \int dx \frac{\left[\int \eta q(x-\eta) \, dP(\eta\sigma)\right]^2}{\int q\,(x-\eta) \, dP(\eta\sigma)}. \end{aligned} \tag{2.30}$$

We observe also that

$$\psi(\lambda, 1) \leqq \lambda, \tag{2.31}$$

a bound which will be useful only for large λ. This follows from the convexity of Φ, or from

$$\Phi(P, 1) = \int x^2 \, dP(x) - \inf_{\psi} \mathcal{E}[\psi(X) - \xi]^2, \tag{2.32}$$

with ξ distributed according to q, and $X - \xi$ given ξ according to P, which is essentially (2.22).

For $\lambda \leqq \frac{1}{2}$,

$$\begin{aligned} \int q(x - \eta) \, dP(\eta) &\geqq P[-1, 1] \inf_{z \varepsilon [-1,1]} q(x - z) \\ &\geqq \frac{3}{4} \cdot \frac{2}{5} \frac{1}{\pi(1 + x^2)} = \frac{3}{10\pi(1 + x^2)} \end{aligned} \tag{2.33}$$

by Chebyshev's inequality. Also

$$\begin{aligned} \left[\int \eta q(x - \eta) \, dP(\eta)\right]^2 &= \left\{\int \eta[q(x - \eta) - q(x)] \, dP(\eta)\right\}^2 \\ &\leqq \int \eta^2 \, dP(\eta) \int [q(x - \eta) - q(x)]^2 \, dP(\eta) \end{aligned} \tag{2.34}$$

by Schwarz's inequality. Thus

$$\begin{aligned} \Phi(P, 1) &= \int dx \frac{\left[\int \eta q(x - \eta) \, dP(\eta)\right]^2}{\int q(x - \eta) \, dP(\eta)} \\ &\geqq \frac{10}{3} \int dx (1 + x^2) \left(\int \eta^2 \, dP(\eta)\right) \int \left[\frac{1}{1 + (x - \eta)^2} - \frac{1}{1 + x^2}\right]^2 dP(\eta) \\ &= \frac{10}{3} \left(\int \eta^2 \, dP(\eta)\right) \int dP(\eta) \int \left[\frac{1}{1 + (x - \eta)^2} - \frac{1}{1 + x^2}\right]^2 (1 + x^2) \, dx. \end{aligned} \tag{2.35}$$

But

$$\begin{aligned} \int \left[\frac{1}{1 + (x - \eta)^2} - \frac{1}{1 + x^2}\right]^2 (1 + x^2) \, dx \\ = \int \left[\frac{1 + x^2}{[1 + (x - \eta)^2]^2} - \frac{1}{1 + (x - \eta)^2}\right] dx \\ = \int \left[\frac{1 + (x + \eta)^2}{(1 + x^2)^2} - \frac{1}{1 + x^2}\right] dx = \int \frac{2x + \eta^2}{(1 + x^2)^2} dx = \eta^2 \int \frac{dx}{(1 + x^2)^2}, \end{aligned} \tag{2.36}$$

so that

$$\Phi(P, 1) \leqq c\left[\int \eta^2 \, dP(\eta)\right]^2 \tag{2.37}$$

i.e.,

$$\psi(\lambda, 1) \leqq c\lambda^2 \quad \text{for} \quad \lambda \leqq \tfrac{1}{2}. \tag{2.38}$$

Combining (2.29), (2.31), and (2.38), we have finally

$$\psi(\lambda,\sigma) \leqq \begin{cases} \frac{c}{\sigma}\lambda^2 \text{ for } \lambda \leqq \frac{1}{2}\sigma^2 \\ \sigma\lambda \text{ for } \lambda \geqq \frac{1}{2}\sigma^2. \end{cases} \tag{2.39}$$

Now let ν^* be the distribution of $\int q^2 \, dP(q, Y)$, i.e.,

$$\nu^*(S) = \nu\left\{y: \int \eta^2 \, dP(\eta, y) \, \varepsilon S\right\}. \tag{2.40}$$

Then

$$\int d\nu(y) \int dx \frac{\left[\int \eta q\left(\frac{x-\eta}{\sigma}\right) dP(\eta, y)\right]^2}{\int q\left(\frac{x-\eta}{\sigma}\right) dP(\eta, y)} \leqq \frac{c}{\sigma}\int_0^{\frac{1}{2}\sigma^2} \lambda^2 \, d\nu^*(\lambda) + \sigma \int_{\frac{1}{2}\sigma^2}^{\infty} \lambda \, d\nu^*(\lambda). \tag{2.41}$$

For any ϵ between 0 and 1, choose σ_0 so large that

$$\int_{\frac{\epsilon}{2}\sigma_0^2}^{\infty} \lambda^{3/2} \, d\nu^*(\lambda) < \epsilon. \tag{2.42}$$

Then, for $\sigma \geqq \sigma_0$,

$$\begin{aligned} \frac{1}{\sigma}\int_0^{\frac{1}{2}\sigma^2} \lambda^2 \, d\nu^*(\lambda) &= \frac{1}{\sigma}\int_0^{\frac{\epsilon}{2}\sigma^2} \lambda^2 \, d\nu^*(\lambda) + \frac{1}{\sigma}\int_{\frac{\epsilon}{2}\sigma^2}^{\frac{1}{2}\sigma^2} \lambda^2 \, d\nu^*(\lambda) \\ &\leqq \sqrt{\frac{\epsilon}{2}}\int_0^{\frac{\epsilon}{2}\sigma^2} \lambda^{3/2} \, d\nu^*(\lambda) + \frac{1}{\sqrt{2}}\int_{\frac{\epsilon\sigma^2}{2}}^{\infty} \lambda^{3/2} \, d\nu^*(\lambda) \\ &\leqq \sqrt{\frac{\epsilon}{2}}\int_0^{\infty} \lambda^{3/2} \, d\nu^*(\lambda) + \frac{\epsilon}{\sqrt{2}}. \end{aligned} \tag{2.43}$$

$$\sigma \int_{\frac{1}{2}\sigma^2}^{\infty} \lambda \, d\nu^*(\lambda) \leqq \sqrt{2}\int_{\frac{1}{2}\sigma^2}^{\infty} \lambda^{3/2} \, d\nu^*(\lambda) \leqq \sqrt{2}\,\epsilon. \tag{2.44}$$

Thus the right-hand side of (2.41) approaches 0 as $\sigma \to \infty$, which completes the proof of (2.23), and thus the proofs of Theorems 2 and 1.

REFERENCES

[1] E. J. G. Pitman, "The estimation of location and scale parameters of a continuous population of any given form," *Biometrika*, Vol. 30 (1939), pp. 391–421.

[2] M. A. GIRSHICK AND L. J. SAVAGE, "Bayes and minimax estimates for quadratic loss functions," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1951, pp. 53–73.

[3] HIROKICHI KUDO, "On minimax invariant estimates of the transformation parameter," *Natural Science Report of the Ochanomizu University*, Vol. 6 (1955), No. 1, pp. 31–73.

[4] J. KIEFER, "Invariance, minimax sequential estimation, and continuous time processes," *Ann. Math. Stat.*, Vol. 28 (1957), pp. 573–601.

[5] SAMUEL KARLIN, "Admissibility for estimation with quadratic loss," *Ann. Math. Stat.*, Vol. 29 (1958), pp. 406–436.

[6] COLIN R. BLYTH, "On minimax statistical decision procedures and their admissibility," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 22–42.

[7] E. L. LEHMANN AND C. M. STEIN, "The admissibility of certain invariant statistical tests involving a translation parameter," *Ann. Math. Stat.*, Vol. 24 (1953), pp. 473–479.

[8] CHARLES STEIN, "Inadmissibility of the usual estimator for the mean of a multivariate normal distribution," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley and Los Angeles, University of California Press, 1956, Vol. 1, pp. 197–206.

[9] DAVID BLACKWELL, "On the translation parameter problem for discrete variables," *Ann. Math. Stat.*, Vol. 22 (1951), pp. 393–399.

[10] E. C. TITCHMARSH, *Theory of Functions*, Second edition, Clarendon Press, Oxford, 1939.