The Annals of Statistics
1974, Vol. 2, No. 6, 1127–1137

ASYMPTOTIC PROPERTIES OF ESTIMATORS OF A LOCATION PARAMETER

BY CHARLES J. STONE

University of California, Los Angeles

Consider the problem of estimating the location parameter $\theta \in R^d$ based on a sample of size n from $(\theta + X, Y)$, where X is a d-dimensional random vector, Y is a random element of some measure space, and (X, Y) has a known distribution. Let $\mathscr{I}^-$ denote the corresponding inverse Fisher information matrix. We show that there is always an invariant estimator $\hat{\theta}_n$ such that $\mathscr{L}(n^{\frac{1}{2}}(\hat{\theta}_n - \theta)) \to N(0, \mathscr{I}^-)$ as $n \to \infty$. Let ρ be a fixed probability density on R^d, let $\tilde{\theta}_n$ be any estimator of θ and set $R_n(c) = \int \rho(\theta)\, d\theta E_\theta \min(c, n|\tilde{\theta}_n - \theta|^2)$. We show that $\lim_{c\to\infty} \liminf_{n\to\infty} R_n(c) \geqq \operatorname{trace} \mathscr{I}^-$ and that if $\lim_{c\to\infty} \limsup_{n\to\infty} R_n(c) = \operatorname{trace} \mathscr{I}^-$, then $\lim_{n\to\infty} \int \rho(\theta)\, d\theta P_\theta(n^{\frac{1}{2}}|\tilde{\theta}_n - \hat{\theta}_n| \geqq c) = 0$ for all $c > 0$. These results are obtained with *no* regularity conditions imposed on the distribution of (X, Y).

1. Introduction. In this paper, which is an outgrowth of Port and Stone [9], we continue the study of estimators of the location parameter $\theta \in R^d$ based on a random sample of size n from $(\theta + X, Y)$, where X is a d-dimensional random vector, Y is a random element of some measure space $\mathscr{Y}$ having distribution μ_Y, and (X, Y) has a known distribution. Our results are obtained with *no conditions whatsoever* on the distribution (X, Y).

The Fisher information $\mathscr{I} = \mathscr{I}(\theta + X, Y)$ and its inverse $\mathscr{I}^- = \mathscr{I}^-(\theta + X, Y)$, defined and studied in [9], play an important role here. Recall that $\mathscr{I}^-$ is a nonnegative definite symmetric linear transformation from R^d to itself, which we can think of as a $d \times d$ matrix. In the special case $d = 1$ we can think of $\mathscr{I}$ as a number such that $0 < \mathscr{I} \leqq \infty$ and $\mathscr{I}^-$ as the finite nonnegative number $1/\mathscr{I}$. If, in addition, Y is degenerate, then $\mathscr{I}$ agrees with the definition of information given in Huber [4]. There Huber showed that $\mathscr{I} = \int (f'/f)^2 f\, dx$ if X has an absolutely continuous density f such that the indicated integral is finite and $\mathscr{I} = +\infty$ otherwise.

In general, for $e \in R^d$ set $\mu = (e, \theta)$, where $(\ ,\ )$ denotes the usual inner product on R^d. By an estimator T_n of μ or an estimator $\tilde{\theta}_n$ of θ we will always mean a (possibly randomized) estimator based on a sample of size n from $(\theta + X, Y)$. We say that T_n is invariant if $T_n - \mu$ is independent of θ and that $\tilde{\theta}_n$ is invariant if $\tilde{\theta} - \theta$ is independent of θ. In the remainder of the Introduction we will describe the results for estimators of θ. Similar results are obtained for estimators of μ.

In Section 2 we define the Pitman estimator $\hat{\theta}_n$ of θ corresponding to an

Received May 1973; revised January 1974.

[1] The preparation of this paper in part was supported by NSF Grant Number GP-3343IX.

AMS 1970 *subject classifications.* Primary 62F10; Secondary 62F20.

Key words and phrases. Estimation, location parameter, Fisher information, Pitman estimator, asymptotic properties.

appropriate bounded and continuous loss function $L(x)$, $x \in R^d$, that is asymptotic to $|x|^2$ as $x \to 0$. The Pitman estimator is invariant and minimax with respect to L. In Theorem 2.1 we show that

$$\mathscr{L}(n^{\frac{1}{2}}(\hat{\theta}_n - \theta)) \to N(0, \mathscr{I}^-) \qquad \text{as } n \to \infty . \tag{1.1}$$

This result is perhaps best regarded as a constructive proof of the existence in complete generality of an invariant estimator satisfying (1.1). Observe that this result is nonvacuous even if $\mathscr{I}^- = 0$, which is true if and only if $\mathscr{I}$ is infinite in all directions. In that case (1.1) shows that there is an invariant estimator $\hat{\theta}_n$ of θ such that $n^{\frac{1}{2}}(\hat{\theta}_n - \theta) \to 0$ in probability as $n \to \infty$. One could attempt to obtain still stronger results by looking at the asymptotic distribution of $g(n)n^{\frac{1}{2}}(\hat{\theta}_n - \theta)$ where $g(n) \to \infty$ as $n \to \infty$. But this direction is not pursued here.

Under the assumption that $E|X|^\delta < \infty$ for some $\delta > 0$, the existence of an invariant estimator satisfying (1.1) follows from Theorem 5.2 of [9]. If $d = 1$ and $\mathscr{I} < \infty$, the existence of a maximum likelihood estimator satisfying (1.1) probably can be obtained by verifying the conditions of Proposition 6 of LeCam [8].

Weiss and Wolfowitz in a series of papers [12], [13], and [14] discuss "maximum probability estimators." If one could show that the conditions they postulate on an appropriate maximum probability estimator hold in the location parameter model considered in this paper, then Theorem 2.1 together with the Weiss–Wolfowitz results would imply that that the maximum probability estimator satisfies (1.1). It would certainly be worthwhile to show that their postulated conditions do hold in our model under no further regularity conditions. But it is not at all clear to this author that this is in fact the case.

Let ρ be a probability density on R^d. Then any estimator $\hat{\theta}_n$ satisfying (1.1), in particular the Pitman estimator, satisfies

$$\lim_{c\to\infty} \lim_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\hat{\theta}_n - \theta|^2) = \operatorname{trace} \mathscr{I}^- . \tag{1.2}$$

In Theorem 3.1 we show that if $\tilde{\theta}_n$ is any estimator of θ, then

$$\lim_{c\to\infty} \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\tilde{\theta}_n - \theta|^2) \geqq \operatorname{trace} \mathscr{I}^- . \tag{1.3}$$

In Theorem 3.2 we show that if

$$\lim_{c\to\infty} \limsup_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\tilde{\theta}_n - \theta|^2) = \operatorname{trace} \mathscr{I}^- , \tag{1.4}$$

then

$$\lim_{n\to\infty} \int \rho(\theta)\, d\theta P_\theta(n^{\frac{1}{2}}|\tilde{\theta}_n - \hat{\theta}_n| \geqq c) = 0 , \qquad 0 < c < \infty . \tag{1.5}$$

A sufficient condition for (1.4) to hold is that $\mathscr{L}_\theta(n^{\frac{1}{2}}(\tilde{\theta}_n - \theta)) \to N(0, \mathscr{I}^-)$ for almost all θ. If we think of ρ as a prior density for the random variable θ, then (1.5) states that $n^{\frac{1}{2}}(\tilde{\theta}_n - \hat{\theta}_n) \to 0$ in probability as $n \to \infty$.

Given the prior density ρ, we say that $\tilde{\theta}_n$ is an *asymptotically optimal estimator* of θ if (1.4) holds. We say that such an estimator is the *essentially unique* asymptotically optimal estimator of θ if whenever θ_n' is any asymptotically optimal estimator of θ, then $n^{\frac{1}{2}}(\theta_n' - \tilde{\theta}_n) \to 0$ in probability as $n \to \infty$. Then Theorem 2.1, Theorem 3.1, and Theorem 3.2 together imply the following result.

THEOREM 1.1. *For every prior density ρ, the Pitman estimator $\hat{\theta}_n$ is the essentially unique asymptotically optimal estimator of θ.*

In [9] we showed that for any estimator $\tilde{\theta}_n$

$$n \sup_\theta E_\theta|\tilde{\theta}_n - \theta|^2 \geqq \text{trace } \mathscr{I}^- \qquad \text{for all } n \geqq 1 . \tag{1.6}$$

This is close to (1.3), but actually neither result implies the other. Theorem 3.1 was motivated by LeCam [7]. Presumably this result follows from [7] under strong enough regularity conditions.

Hájek in Theorem 4.2 of [3] has a result which is closely related to Theorem 3.1 but does not directly contain it. His results do imply that under his LAN (local asymptotic normality) conditions

$$\lim_{c\to\infty} \lim_{\delta\to 0} \liminf_{n\to\infty} \sup_{|t-\theta|<\delta} E_\theta(\min(c, n(T_n - (e, \theta))^2)) \geqq (e, \mathscr{I}^- e) ,$$

where $e \in R^d$ and T_n is any estimator of (e, θ). It should be possible by modifying the proof of Hájek's result to show that under the same conditions for any probability density ρ on R^d

$$\lim_{c\to\infty} \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta(\min(c, n(T_n - (e, \theta))^2 \geqq (e, \mathscr{I}^- e) .$$

The conclusion (1.3) of Theorem 3.1 would follow easily from this result. Hájek's LAN conditions do not hold if $\mathscr{I}^-$ is singular, i.e. if $\mathscr{I}$ is infinite in one or more directions. If $\mathscr{I}^-$ is nonsingular they can probably be verified by using the arguments in the Appendix to Hájek's paper. (This would show that when the information $\mathscr{I}$ defined in [9] is finite, it agrees with the matrix Γ_t defined in Hájek's LAN conditions.) If $\mathscr{I}^-$ is zero, i.e. if $\mathscr{I}$ is infinite in all directions, the conclusion of Theorem 3.1 is trivial. In the mixed case, when $\mathscr{I}$ is infinite in some but not all directions, one can probably use Theorem 2.1 and the properties of $\mathscr{I}$ obtained in [9] to reduce this case to the nonsingular case. This method of obtaining Theorem 3.1 would be worthwhile, but would be no shorter than the direct proof given in Section 3.

In the preceding discussion it was tacitly assumed that the distribution of (X, Y) is known. We now consider the important situation wherein this distribution is unknown. Theorem 3.1 is obviously still directly applicable. Theorem 2.1 is not directly applicable since we cannot implement the Pitman estimator without knowing the distribution of (X, Y). This result, however, is strongly suggestive of what should be true, as we now indicate.

If the distribution of (X, Y) is totally unknown it does not make any sense to talk about estimation of θ. It is well known, however, that if the distribution of (X, Y) satisfies appropriate symmetry conditions, then estimation of θ is possible. Stein [10] has shown "formally" that in such cases one should be able to do as well asymptotically, without knowing the distribution of (X, Y), as one could do knowing this distribution. Stein's result together with Theorem

2.1 strongly suggests that there should be an invariant estimator $\hat{\theta}_n$ which is independent of the distribution of (X, Y) and which satisfies (1.1) whenever the distribution of (X, Y) satisfies appropriate symmetry conditions. Such an estimator would presumably be of the "adaptive" type. In Stone [11] it is shown that if $d = 1$ and Y is degenerate, then such an adaptive estimator does indeed exist.

It would be worthwhile to extend the results of [9] and the present paper to more general models, e.g. those involving both location and scale parameters. Similarly it would be worthwhile to extend the results of [11] in this direction.

2. Asymptotic properties of an invariant estimator.

We begin this section by giving a lower bound to the asymptotic mean square error of an invariant estimator. This result will be extended to noninvariant estimators in Theorem 3.1.

PROPOSITION 2.1. (i) *If T_n is an invariant estimator of $\mu = (e, \theta)$, then*

$$\lim_{c\to\infty} \liminf_{n\to\infty} E \min(c, n(T_n - \mu)^2) \geqq (e, \mathscr{I}^{-}e) . \tag{2.1}$$

(ii) *If $\tilde{\theta}_n$ is an invariant estimator of θ, then*

$$\lim_{c\to\infty} \liminf_{n\to\infty} E \min(c, n|\tilde{\theta}_n - \theta|^2) \geqq \operatorname{trace} \mathscr{I}^{-} . \tag{2.2}$$

PROOF. We will prove (i), from which (ii) follows immediately. Set $U_n = T_n - \mu$ and let α^2 denote the left side of (2.1). We can suppose that $\alpha^2 < \infty$. Then there is a sequence $\{n_k\}$ such that $n_k \uparrow \infty$ as $k \to \infty$ and a random variable U such that $\mathscr{L}(n_k^{\frac{1}{2}} U_{n_k}) \to \mathscr{L}(U)$ and $E \min(k, n_k U_{n_k}^2) \to \alpha^2$ as $k \to \infty$. It is easily seen that

$$\alpha^2 \geqq EU^2 \geqq \operatorname{Var} U . \tag{2.3}$$

It follows from Theorems 2.4 and 3.4 and Corollary 3.1 all of [9] that

$$\begin{aligned} \mathscr{I}^{-}(\mu + U) &\geqq \limsup_{k\to\infty} \mathscr{I}^{-}(\mu + n_k^{\frac{1}{2}} U_{n_k}) \\ &= \limsup_{k\to\infty} n_k \mathscr{I}^{-}(\mu + U_{n_k}) . \end{aligned}$$

Now $n\mathscr{I}^{-}(\mu + U_n) \geqq (e, \mathscr{I}^{-}(\theta + X, Y)e)$ by Theorem 3.5 of [9], so that

$$\mathscr{I}^{-}(\mu + U) \geqq (e, \mathscr{I}^{-}(\theta + X, Y)e) . \tag{2.4}$$

According to Theorem 5.1 of [9]

$$\operatorname{Var} U \geqq \mathscr{I}^{-}(\mu + U) . \tag{2.5}$$

From (2.3)—(2.5) we conclude that $\alpha^2 \geqq (e, \mathscr{I}^{-}(\theta + X, Y)e)$ as desired.

COROLLARY 2.1. (i) *If T_n is an invariant estimator of $\mu = (e, \theta)$ such that $\mathscr{L}(n^{\frac{1}{2}}(T_n - \mu)) \to N(0, \sigma^2)$ as $n \to \infty$, then $\sigma^2 \geqq (e, \mathscr{I}^{-}e)$.*

(ii) *If $\tilde{\theta}_n$ is an invariant estimator of θ such that $\mathscr{L}(n^{\frac{1}{2}}(\tilde{\theta}_n - \theta)) \to N(0, \Sigma)$ as $n \to \infty$, then $\Sigma \geqq \mathscr{I}^{-}$.*

We will now define the loss function L and the Pitman estimator $\tilde{\theta}_n$ referred to in the Introduction.

For $1 \leqq k \leqq d$, let $L_k(x)$, $x \in R$, be a real-valued function satisfying the following properties:

(i) L_k is bounded and continuous;
(ii) $L_k(x) \sim x^2$ as $x \to 0$;
(iii) $L_k(x) > 0$ for $x \neq 0$;
(iv) $\lim_{|x|\to\infty} L_k(x) = \sup_x L_k(x)$.

Examples of functions that satisfy (i)—(iv) are $\min(c, x^2)$ and $x^2/(1 + cx^2)$ for c a positive constant. The "loss function" $L(x)$, $x \in R^d$, is defined by setting

$$L(x) = \sum_{k=1}^{d} L_k(x_k)\,, \qquad x = (x_1, \cdots, x_d)\,.$$

We use this loss function instead of the usual quadratic loss in order to avoid having to make the assumption $E|X|^\delta < \infty$ for some $\delta > 0$ that was required for some of the results in [9].

We now define the Pitman estimator $\hat{\theta}_n$ of θ corresponding to the loss function L and based on the sample

$$\theta + X_1, Y_1, \cdots, \theta + X_n, Y_n$$

of size n from $(\theta + X, Y)$. Fix k, $1 \leqq k \leqq d$, and let θ_k and X_{1k} denote respectively the kth coordinates of θ and X_1. Set $Z = (Y_1, X_2 - X_1, Y_2, \cdots, X_n - X_1 Y_n)$ and let φ be a real-valued measurable function on the range of Z such that

$$E[L_k(X_{1k} - \varphi(Z)) \mid Z] = \inf_v E[L_k(X_{1k} - v) \mid Z]\,.$$

It follows from (i)—(iv) that such a function φ does indeed exist. (It can be defined as a lower semicontinuous function of the regular conditional distribution of X_{1k} given Z—this guarantees measurability.) The Pitman estimator $\hat{\theta}_{nk}$ of θ_k is given as $\theta_k + X_{1k} - \varphi(Z)$, so that

$$\hat{\theta}_{nk} = \theta_k + X_{1k} - \varphi(Y_1, X_2 - X_1, Y_2, \cdots, X_n - X_1, Y_n)\,.$$

It is clearly invariant and is well known to be minimax with respect to the loss function L_k. That is, if $\tilde{\theta}_{nk}$ is any, possibly randomized, estimator of θ_k based on the same sample, then

$$\sup_\theta E_\theta L_k(\tilde{\theta}_{nk} - \theta_k) \geqq EL_k(\hat{\theta}_{nk} - \theta_k)\,.$$

The Pitman estimator $\hat{\theta}_n$ of θ is given by $\hat{\theta}_n = (\hat{\theta}_{n1}, \cdots, \hat{\theta}_{nd})$. It is invariant and minimax with respect to the loss function L. For discussions of Pitman estimators and their minimax properties in various levels of generality see Girshick and Savage [2], Blackwell and Girshick [1], Kudō [6], and Kiefer [5].

Let P_n be the estimator of $\mu = (e, \theta)$ given by $P_n = (e, \hat{\theta}_n)$.

THEOREM 2.1. (i) *$\mathscr{L}(n^{\frac{1}{2}}(P_n - \mu)) \to N(0, (e, \mathscr{I}^- e))$ as $n \to \infty$ and for all $c > 0$*

$$\lim_{n\to\infty} nE \min(c, (P_n - \mu)^2) = (e, \mathscr{I}^- e)\,.$$

(ii) *$\mathscr{L}(n^{\frac{1}{2}}(\hat{\theta}_n - \theta)) \to N(0, \mathscr{I}^-)$ as $n \to \infty$ and*

$$\lim_{n\to\infty} nEL(\hat{\theta}_n - \theta) = \operatorname{trace} \mathscr{I}^-\,. \tag{2.6}$$

PROOF. We will prove (ii), from which (i) follows immediately. Let W be independent of (X, Y) and have the standard normal distribution on R^d. We know from [9] that

$$\lim_{\sigma\to 0} \mathscr{I}^-(\theta + X + \sigma W, Y) = \mathscr{I}^-(\theta + X, Y) . \tag{2.7}$$

In [9] we verified the existence of estimators $\tilde{\theta}_n(\sigma, \nu)$, $\sigma > 0$, and $\nu = 1, 2, \cdots$, based on a sample of size n from $(\theta + X + \sigma W, Y)$ which satisfy the following properties:

(i) $\tilde{\theta}_n(\sigma, \nu)$ is an invariant estimator of θ;
(ii) $\lim_{n\to\infty} \mathscr{L}(n^{\frac{1}{2}}(\tilde{\theta}_n(\sigma, \nu) - \theta)) = N(0, \Sigma(\sigma, \nu))$;
(iii) $\lim_{n\to\infty} nE \min (c, |\tilde{\theta}_n(\sigma, \nu) - \theta|^2) = \operatorname{trace} \Sigma(\sigma, \nu)$ for all $c > 0$;
(iv) $\lim_{\nu\to\infty} \Sigma(\sigma, \nu) = \mathscr{I}^{-1}(\theta + X + \sigma W, Y)$.

Set $\alpha_k^2(\sigma, \nu) = (\Sigma(\sigma, \nu))_{k,k}$ and $\alpha_k^2 = (\mathscr{I}^-(\theta + X, Y))_{k,k}$. Then

$$\lim_{n\to\infty} nEL_k(\tilde{\theta}_{nk}(\sigma, \nu) - \theta_k) = \alpha_k^2(\sigma, \nu) \tag{2.8}$$

and, by (2.7) and (iv),

$$\lim_{\sigma\to 0} (\lim_{\nu\to\infty} \alpha_k^2(\sigma, \nu)) = \alpha_k^2 . \tag{2.9}$$

LEMMA 2.1. $\lim_{n\to\infty} nEL_k(\hat{\theta}_{nk} - \theta_k) = \alpha_k^2$.

PROOF. It follows easily from Proposition 2.1 that

$$\liminf_{n\to\infty} nEL_k(\hat{\theta}_{nk} - \theta_k) \geqq \alpha_k^2 . \tag{2.10}$$

On the other hand, since $\hat{\theta}_{nk}$ is a minimax estimator of θ_k, it is clear that

$$EL_k(\hat{\theta}_{nk} - \theta_k) \leqq EL_k(\tilde{\theta}_{nk}(\sigma, \nu) - \theta_k) . \tag{2.11}$$

By (2.8), (2.9), and (2.11)

$$\limsup_{n\to\infty} nEL_k(\hat{\theta}_{nk} - \theta_k) \leqq \alpha_k^2 . \tag{2.12}$$

The lemma follows from (2.10) and (2.12).

Observe that (2.6) is an immediate consequence of Lemma 2.1.

LEMMA 2.2. *For* $0 < b < \infty$

$$\limsup_{n\to\infty} E \min \left(b, n \left(\frac{\tilde{\theta}_{nk}(\sigma, \nu) - \hat{\theta}_{nk}}{2}\right)^2\right) \leqq \frac{\alpha_k^2(\sigma, \nu) - \alpha_k^2}{2} .$$

PROOF. We can assume that $\theta = 0$. Then

$$\lim_{c\to\infty} \lim_{n\to\infty} E \min (c, n\tilde{\theta}_{nk}^2(\sigma, \nu)) = \alpha_k^2(\sigma, \nu) \tag{2.13}$$

and by the previous lemma

$$\lim_{c\to\infty} \limsup_{n\to\infty} E \min (c, n\hat{\theta}_{nk}^2) \leqq \alpha_k^2 . \tag{2.14}$$

It follows from (2.13) and (2.14) that

$$\lim_{c\to\infty} \limsup_{n\to\infty} E \min \left(c, n \left(\frac{\hat{\theta}_{nk}^2 + \tilde{\theta}_{nk}^2(\sigma, \nu)}{2}\right)^2\right) \leqq \frac{\alpha_k^2 + \alpha_k^2(\sigma, \nu)}{2} \tag{2.15}$$

and hence

$$(2.16)\qquad \lim_{c\to\infty} \limsup_{n\to\infty} P\left(n \frac{\hat{\theta}_{nk}^2 + \tilde{\theta}_{nk}^2(\sigma, \nu)}{2} > c\right) = 0 .$$

By Proposition 2.1

$$(2.17)\qquad \lim_{c\to\infty} \liminf_{n\to\infty} E \min\left(c, n\left(\frac{\hat{\theta}_{nk} + \tilde{\theta}_{nk}(\sigma, \nu)}{2}\right)^2\right) \geqq \alpha_k^2 .$$

We conclude from (2.16) and (2.17) that

$$(2.18)\qquad \lim_{c\to\infty} \liminf_{n\to\infty} E\left[n\left(\frac{\hat{\theta}_{nk} + \tilde{\theta}_{nk}(\sigma, \nu)}{2}\right)^2 ;\ n \frac{\hat{\theta}_{nk}^2 + \tilde{\theta}_{nk}^2(\sigma, \nu)}{2} \leqq c\right] \geqq \alpha_k^2 .$$

Now

$$(2.19)\qquad n\left(\frac{\hat{\theta}_{nk} - \tilde{\theta}_{nk}(\sigma, \nu)}{2}\right)^2 = n \frac{\hat{\theta}_{nk}^2 + \tilde{\theta}_{nk}^2(\sigma, \nu)}{2} - n\left(\frac{\tilde{\theta}_{nk}(\sigma, \nu) + \hat{\theta}_{nk}}{2}\right)^2 .$$

It follows from (2.15), (2.18), and (2.19) that

$$(2.20)\qquad \lim_{c\to\infty} \limsup_{n\to\infty} E\left[n\left(\frac{\tilde{\theta}_{nk}(\sigma, \nu) - \hat{\theta}_{nk}}{2}\right)^2 ;\ n \frac{\hat{\theta}_{nk}^2 + \tilde{\theta}_{nk}^2(\sigma, \nu)}{2} \leqq c\right] \leqq \frac{\alpha_k^2(\sigma, \nu) - \alpha_k^2}{2} .$$

Lemma 2.2 follows from (2.16) and (2.20).

It follows from properties (ii) and (iv) of $\tilde{\theta}_n(\sigma, \nu)$, (2.9) and Lemma 2.2 that $\mathscr{L}(n^{\frac{1}{2}}(\hat{\theta}_n - \theta)) \to N(0, \mathscr{I}^-)$ as $n \to \infty$. This completes the proof of the theorem.

PROPOSITION 2.2. (i) *If T_n is an invariant estimator of $\mu = (e, \theta)$ such that*

$$(2.21)\qquad \lim_{c\to\infty} \limsup_{n\to\infty} E \min(c, n(T_n - \mu)^2) = (e, \mathscr{I}^- e) ,$$

then $n^{\frac{1}{2}}(T_n - P_n) \to 0$ in probability as $n \to \infty$.

(ii) *If $\tilde{\theta}_n$ is an invariant estimator of θ such that*

$$(2.22)\qquad \lim_{c\to\infty} \limsup_{n\to\infty} E \min(c, n|\tilde{\theta}_n - \theta|^2) = \operatorname{trace} \mathscr{I}^- ,$$

then $n^{\frac{1}{2}}(\tilde{\theta}_n - \hat{\theta}_n) \to 0$ in probability as $n \to \infty$.

PROOF. If (2.21) holds we can argue as in the proof of Lemma 2.2 to conclude that

$$\lim_{n\to\infty} E \min(b, n(T_n - P_n)^2) = 0 , \qquad 0 < b < \infty ,$$

and hence that $n^{\frac{1}{2}}(T_n - P_n) \to 0$ in probability as $n \to \infty$. This proves (i). Statement (ii) follows immediately from (i) and the first part of Proposition 2.1.

Proposition 2.2 will be extended to noninvariant estimators in Theorem 3.2.

3. Asymptotic properties of any estimator. In this section we extend Propositions 2.1 and 2.2 to estimators which are not necessarily invariant.

THEOREM 3.1. *Let ρ be a probability density on R^d.* (i) *If T_n is any estimator of $\mu = (e, \theta)$, then*

$$(3.1)\qquad \lim_{c\to\infty} \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n(T_n - \mu)^2) \geqq (e, \mathscr{I}^- e) .$$

(ii) *If* $\tilde{\theta}_n$ *is any estimator of* θ, *then*

$$\lim_{c\to\infty} \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\tilde{\theta}_n - \theta|^2) \geqq \operatorname{trace} \mathscr{I}^- . \tag{3.2}$$

PROOF. We first prove (ii). Let α^2 denote the left side of (3.2). Let $\hat{\theta}_n$ denote the Pitman estimator of θ corresponding to the loss function L. For $b > 0$ define the estimator θ_n^* of θ by setting

$$\begin{aligned} \theta_{nk}^* &= \tilde{\theta}_{nk} && \text{if } n^{\frac{1}{2}}|\tilde{\theta}_{nk} - \hat{\theta}_{nk}| \leqq 2b^{\frac{1}{2}} , \\ &= \hat{\theta}_{nk} && \text{if } n^{\frac{1}{2}}|\tilde{\theta}_{nk} - \hat{\theta}_{nk}| > 2b^{\frac{1}{2}} . \end{aligned}$$

Then

$$n|\theta_n^* - \theta|^2 \leqq 2n|\hat{\theta}_n - \theta|^2 + 8bd \tag{3.3}$$

and

$$n|\theta_n^* - \theta|^2 \leqq n|\tilde{\theta}_n - \theta|^2 \quad \text{if } n|\hat{\theta}_n - \theta|^2 \leqq b . \tag{3.4}$$

We conclude from (3.3) and Theorem 2.1 that

$$\lim_{c\to\infty} \limsup_{n\to\infty} \sup_\theta E_\theta \min(c, n|\theta_n^* - \theta|^2) < \infty . \tag{3.5}$$

It follows from (3.3) and (3.4) that

$$\begin{aligned} \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\theta_n^* - \theta|^2) \\ \leqq \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\tilde{\theta}_n - \theta|^2) \\ + 2 \lim_{n\to\infty} E[\min(c, n|\hat{\theta}_n - \theta|^2); n|\hat{\theta}_n - \theta|^2 > b] \\ + 8bd \lim_{n\to\infty} P(n|\hat{\theta}_n - \theta|^2 > b) . \end{aligned}$$

Choose $\varepsilon > 0$. It now follows from Theorem 2.1 that we can choose b such that

$$\lim_{c\to\infty} \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\theta_n^* - \theta|^2) \leqq \alpha^2 + \varepsilon . \tag{3.6}$$

From (3.5) we see that in (3.6) ρ can be chosen to be a continuously differentiable function having compact support, since the collection of such functions is dense in $\mathscr{L}_1(R^d)$.

Write the estimator θ_n^* as

$$\theta_n^* = \varphi_n(\theta + X_1, Y_1, \cdots, \theta + X_n, Y_n) .$$

For any $t \in R^d$ let $\theta_n^*(t)$ be the estimator of θ defined as

$$\theta_n^*(t) = \varphi_n(\theta + t + X_1, Y_1, \cdots, \theta + t + X_n, Y_n) - t .$$

Then

$$\mathscr{L}_\theta(\theta_n^*(t) - \theta) = \mathscr{L}_{\theta+t}(\theta_n^* - \theta - t) , \qquad t \in R^d . \tag{3.7}$$

Since $\hat{\theta}_n$ is invariant, it follows from (3.3) that

$$n|\theta_n^*(t) - \theta|^2 \leqq 2n|\hat{\theta}_n - \theta|^2 + 8bd , \qquad t \in R^d . \tag{3.8}$$

Define the estimator $\bar{\theta}_n$ of θ as

$$\bar{\theta}_n = \int \rho(t)\, dt \theta_n^*(t - \hat{\theta}_n) .$$

Since $\hat{\theta}_n$ is invariant, so is $\bar{\theta}_n$. Let C be a compact subset of R^d whose interior contains the support of ρ. Then

$$\begin{aligned}\bar{\theta}_n - \theta &= \int \rho(t + \hat{\theta}_n - \theta)\, dt(\theta_n{}^*(t - \theta) - \theta) \\ &\doteq \int \rho(t)\, dt(\theta_n{}^*(t - \theta) - \theta) + O_p(1)|\hat{\theta}_n - \theta|) \int_C |\theta_n{}^*(t - \theta) - \theta|\, dt\,.\end{aligned}$$

From (3.8) and Theorem 2.1 we conclude that

$$(3.9) \qquad \bar{\theta}_n - \theta = \int \rho(t)\, dt(\theta_n{}^*(t - \theta) - \theta) + \frac{1}{n} O_p(1)\,.$$

Using (3.7), (3.8), and Schwarz's inequality, we see that

$$\begin{aligned}E_\theta[\min (c, n|\textstyle\int \rho(t)\, dt(\theta_n{}^*(t - \theta) - \theta)|^2); n|\hat{\theta}_n - \theta|^2 \leqq a] \\ \leqq E_\theta[\textstyle\int \rho(t)\, dtn|\theta_n{}^*(t - \theta) - \theta|^2; n|\hat{\theta}_n - \theta|^2 \leqq a] \\ \leqq \textstyle\int \rho(t)\, dtE_\theta \min (2a + 8bd, n|\theta_n{}^*(t - \theta) - \theta|^2) \\ = \textstyle\int \rho(t)\, dtE_t \min (2a + 8bd, n|\theta_n{}^* - t|^2)\,.\end{aligned}$$

Thus by (3.6)

$$\liminf_{n\to\infty} E_\theta[\min (c, n|\textstyle\int \rho(t)\, dt(\theta_n{}^*(t - \theta) - \theta)|^2; n|\hat{\theta}_n - \theta|^2 \leqq a] \leqq \alpha^2 + \varepsilon\,.$$

Since $n|\hat{\theta}_n - \theta|^2 = O_p(1)$ it follows that

$$(3.10) \qquad \lim_{c\to\infty} \liminf_{n\to\infty} E_\theta \min (c, n|\textstyle\int \rho(t)\, dt(\theta_n{}^*(t - \theta) - \theta)|^2) \leqq \alpha^2 + \varepsilon\,.$$

We conclude (3.9) and (3.10) that

$$\lim_{c\to\infty} \liminf_{n\to\infty} E \min (c, n|\bar{\theta}_n - \theta|^2) \leqq \alpha^2 + \varepsilon\,.$$

Since $\bar{\theta}_n$ is invariant, Proposition 2.1 implies that $\alpha^2 + \varepsilon \geqq \operatorname{trace} \mathscr{I}^-$. By letting $\varepsilon \to 0$ we see that (3.2) holds.

In proving (i) we can assume by Corollary 3.1 of [9] that $e = (1, 0, \cdots, 0)$. We have to show that if $\tilde{\theta}_{n1}$ is any estimator of θ_1, then

$$(3.11) \qquad \lim_{c\to\infty} \liminf_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min (c, n(\tilde{\theta}_{n1} - \theta_1)^2) \geqq (\mathscr{I}^-)_{1,1}\,.$$

But (3.11) follows by applying Theorem 3.1 (ii) to the estimator $\tilde{\theta}_n = (\tilde{\theta}_{n1}, \hat{\theta}_{n2}, \cdots, \hat{\theta}_{nd})$ and using Theorem 2.1.

COROLLARY 3.1. *Let A be a Borel set in R^d having positive measure.* (i) *If T_n is any estimator of $\mu = (e, \theta)$, then*

$$\lim_{c\to\infty} \liminf_{n\to\infty} \sup_{\theta\in A} E_\theta \min (c, n(T_n - \mu)^2 \geqq (e, \mathscr{I}^- e)\,.$$

(ii) *If $\tilde{\theta}_n$ is any estimator of θ, then*

$$\lim_{c\to\infty} \liminf_{n\to\infty} \sup_{\theta\in A} E_\theta \min (c, n|\tilde{\theta}_n - \theta|^2) \geqq \operatorname{trace} \mathscr{I}^-\,.$$

COROLLARY 3.2. (i) *If T_n is any estimator of $\mu = (e, \theta)$ such that $\mathscr{L}_\theta(n^{\frac12}(T_n - \mu)) \to N(0, \sigma_\theta^2)$ as $n \to \infty$, then $\sigma_\theta^2 \geqq (e, \mathscr{I}^- e)$ for almost all θ.*

(ii) *If $\tilde{\theta}_n$ is any estimator of θ such that $\mathscr{L}_\theta(n^{\frac12}(\tilde{\theta}_n - \theta)) \to N(0, \Sigma_\theta)$ as $n \to \infty$, then $\Sigma_\theta \geqq \mathscr{I}^-$ for almost all θ.*

COROLLARY 3.3. (i) *Let T_n be any estimator of $\mu = (e, \theta)$ such that $\mathscr{L}_\theta(n^{\frac{1}{2}}(T_n - \mu)) \to N(0, \sigma_\theta^2)$ as $n \to \infty$, where σ_θ^2 depends continuously on θ. Then $\sigma_\theta^2 \geqq (e, \mathscr{I}^- e)$ for all θ.*

(ii) *Let $\tilde{\theta}_n$ be any estimator of θ such that $\mathscr{L}_\theta(n^{\frac{1}{2}}(\tilde{\theta}_n - \theta)) \to N(0, \Sigma_\theta)$ as $n \to \infty$, where Σ_θ depends continuously on θ. Then $\Sigma_\theta \geqq \mathscr{I}^-$ for all θ.*

THEOREM 3.2. *Let ρ be a probability density on R^d.* (i) *If T_n is an estimator of $\mu = (e, \theta)$ such that*

$$\lim_{c\to\infty} \limsup_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n(T_n - \mu)^2) = (e, \mathscr{I}^- e)\,, \tag{3.12}$$

then

$$\lim_{n\to\infty} \int \rho(\theta)\, d\theta P_\theta(n^{\frac{1}{2}}|T_n - P_n| \geqq c) = 0\,, \qquad 0 < c < \infty\,. \tag{3.13}$$

(ii) *If $\tilde{\theta}_n$ is any estimator of θ such that*

$$\lim_{c\to\infty} \limsup_{n\to\infty} \int \rho(\theta)\, d\theta E_\theta \min(c, n|\tilde{\theta}_n - \theta|)^2 = \operatorname{trace} \mathscr{I}^-\,, \tag{3.14}$$

then

$$\lim_{n\to\infty} \int \rho(\theta)\, d\theta P_\theta(n^{\frac{1}{2}}|\tilde{\theta}_n - \hat{\theta}_n| \geqq c) = 0\,, \qquad 0 < c < \infty\,. \tag{3.15}$$

PROOF. Suppose that (3.14) holds. By applying Theorem 3.1 to the estimator $(\tilde{\theta}_n + \hat{\theta}_n)/2$ of θ, using Theorem 2.1 and arguing as in the proof of Lemma 2.2 we conclude that (3.15) holds. This proves (ii). Statement (i) easily reduces to (ii) as it did in the proof of Theorem 3.1.

REFERENCES

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

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

[3] HÁJEK, J. (1972). Local asymptotic minimax and admissibility in estimation. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 175–194. Univ. of California Press.

[4] HUBER, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.

[5] KIEFER, J. (1957). Invariance, minimax sequential estimation and continuous time processes. *Ann. Math. Statist.* **28** 573–601.

[6] KUDŌ, H. (1955). On minimax invariant estimates of the transformation parameter. *Natur. Sci. Rep. Ochanomizu Univ.* **6** 31–73.

[7] LECAM, L. (1953). On the asymptotic properties of maximum likelihood estimates and related Bayes estimates. *Univ. Calif. Publ. Statist.* **1** 277–330.

[8] LECAM, L. (1970). On the assumptions used to prove asymptotic normality of the maximum likelihood estimates. *Ann. Math. Statist.* **41** 802–828.

[9] PORT, S. C. and STONE, C. J. (1974). Fisher information and the Pitman estimator of a location parameter. *Ann. Statist.* **2** 225–247.

[10] STEIN, C. (1956). Efficient nonparametric testing and estimation. *Proc. Third Berkeley Symp. Math. Statist. Prob.* **1** 187–196. Univ. of California Press.

[11] STONE, C. J. (1975). Adaptive maximum likelihood estimators of a location parameter. To appear *Ann. Statist.* **3** No. 2.

[12] WEISS, L. and WOLFOWITZ, J. (1967). Maximum probability estimators. *Ann. Inst. Statist. Math.* **19** 193–206.

[13] WEISS, L. and WOLFOWITZ, J. (1969). Maximum probability estimators with a general loss function. *Probability and Information Theory*, Proceedings of the International Symposium at McMaster University, Canada, 1968; *Lecture Notes in Mathematics*, **89** Springer-Verlag, Berlin.

[14] WEISS, L. and WOLFOWITZ, J. (1970). Maximum probability estimators and asymptotic sufficiency. *Ann. Inst. Statist. Math.* **22** 225–244.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA
LOS ANGELES, CALIFORNIA 90024