The Annals of Statistics
1981, Vol. 9, No. 2, 266–277

ON LOCAL ASYMPTOTIC MINIMAXITY AND ADMISSIBILITY IN ROBUST ESTIMATION[1]

BY HELMUT RIEDER

University of Freiburg

For a particular pseudoloss function, local asymptotic minimaxity and admissibility in the sense of Hajek and Le Cam are studied when probability measures are replaced by certain capacities (ϵ-contamination, total variation). A minimax bound for arbitrary estimator sequences is established, admissibility of minimax estimators is proved, and it is shown that minimax estimators must necessarily have an asymptotic expansion in terms of a truncated logarithmic derivative.

1. Introduction. In [17] an asymptotic minimax result has been obtained for the estimation of the parameter θ of a one real parameter family $\{P_\theta\}$ of probability measures, when the laws of the (independent) observations are allowed to vary over a shrinking neighborhood (ϵ-contamination, total variation) of some P_θ. The particular pseudoloss function employed is based upon the upper probability that an estimator falls below, or exceeds, the true parameter value by a certain decreasing amount. The estimators are restricted to a class of regular estimators that are distinguished by asymptotic expansions.

If the regularity assumptions on the estimators are dropped, similar complications are encountered as in the classical asymptotic variance theory as superefficient estimators can be constructed which render the bound on the asymptotic risk and the optimality result invalid. These difficulties indicate that the previous risk is only an inadequate measure for the asymptotic robustness performance of an estimator—unless the estimator is very regular.

The clue to this problem is the observation that the previously assumed asymptotic expansions imply locally uniform approximation of the extreme limit laws. The idea then is to evaluate the risk uniformly, instead of imposing uniformity requirements on the estimators. By this method, unreasonable estimators will be cut out by a high maximum risk, and not by assumption.

In this way, a local asymptotic minimax bound for arbitrary estimators is established (Theorem 4.1). Local asymptotic admissibility of minimax estimators is proved (Theorem 4.2). Moreover, it is shown that they must necessarily have an asymptotic expansion in terms of a truncated logarithmic derivative (Theorem 4.3).

These results may be viewed as robust analogues of corresponding classical results, for which the reader is referred to Le Cam (1953), Stein and Rubin, cf. Chernoff (1956), Bahadur (1964), Huber (1965), Hájek (1970, 1972), and Le Cam (1972).

Our method of proof utilizes the correspondence between estimators and tests, and this is made possible by the particular choice of pseudoloss functions. It is not clear, and very unlikely, that more general loss functions can be treated this way. Moreover, in contrast to the classical case, our minimax solution generally depends on the loss function.

As for local asymptotic minimax bounds in other robust settings, see the manuscripts of Beran (1979a, b) and Millar (1979).

Received October 1978; revised December 1979.

[1] Research performed while holding a scholarship of Deutsche Forschungsgemeinschaft and visiting the Department of Statistics, University of California, Berkeley.

AMS 1970 *subject classifications*. Primary 62G35; secondary 62E20, 62C15.

Key words and phrases. Local asymptotic minimax bound, local asymptotic admissibility, asymptotic expansions, regular estimators, superefficiency, contiguity, least favorable pairs.

2. The model. Let $(\Omega, \mathscr{B})$ be a measurable space and $\mathscr{M}$ the set of all probability measures on $\mathscr{B}$. Let a family $\{P_\theta\} \subset \mathscr{M}$ be given, which is parametrized by an open subset Θ of the real line, and, furthermore, parameter functions ϵ, δ, and $\tau: \Theta \to [0, \infty)$ such that $\epsilon_\theta + \delta_\theta > 0$ and $\tau_\theta > 0$ for all $\theta \in \Theta$. The family $\{P_\theta\}$ is assumed to be of the following regular kind:

(2.1) for each $\theta \in \Theta$ there exists a neighborhood $U_\theta \subset \Theta$ of θ such that $P_\zeta \ll P_\theta$ for all $\zeta \in U_\theta$.

Let μ be a dominating σ-finite measure, which may be assumed to be equivalent to the family $\{P_\theta\}$, and denote by p_θ a version of $dP_\theta/d\mu$.

(2.2) For each $\theta \in \Theta$ there exists a nondegenerate function $\Lambda_\theta \in L^2(dP_\theta)$ such that

$$\frac{p_\zeta^{1/2} - p_\theta^{1/2}}{(\zeta - \theta)p_\theta^{1/2}} \to \frac{1}{2}\Lambda_\theta \quad \text{in } L^2(dP_\theta) \quad \text{as } \zeta \to \theta.$$

The parameter functions are subject to the condition that for all $\theta \in \Theta$

$$\epsilon_\theta + 2\delta_\theta < 2\tau_\theta \int \Lambda_\theta^+ \, dP_\theta. \tag{2.3}$$

For later reference, let the numbers d_θ', d_θ'', $\sigma_\theta^\sharp$ and the functions $\Lambda_\theta^\sharp$, $IC_\theta^\sharp$ be (uniquely) defined by

$$\int (d_\theta' - \Lambda_\theta)^+ \, dP_\theta = \frac{\epsilon_\theta + 2\delta_\theta}{2\tau_\theta} = \int (\Lambda_\theta - d_\theta'')^+ \, dP_\theta$$

$$\Lambda_\theta^\sharp = d_\theta' \vee \Lambda_\theta \wedge d_\theta'', \qquad \sigma_\theta^\sharp = \left(\int (\Lambda_\theta^\sharp)^2 \, dP_\theta\right)^{1/2}, \qquad IC_\theta^\sharp = \frac{\Lambda_\theta^\sharp}{\int \Lambda_\theta^\sharp \Lambda_\theta dP_\theta}; \tag{2.4}$$

note that $d_\theta' \in (-\infty, 0)$, $d_\theta'' \in (0, \infty)$, $\sigma_\theta^\sharp \in (0, \infty)$, and $\int \Lambda_\theta^\sharp \Lambda_\theta \, dP_\theta \in (0, \infty)$ (cf. also [15], [17]).

Estimators shall be studied locally at an arbitrary parameter value θ. This value is fixed and may therefore be dropped from notation whenever feasible. With this understanding we define (for $\zeta \in U_\theta$, say):

$$\epsilon_N = N^{-1/2}\epsilon_\theta, \quad \delta_N = N^{-1/2}\delta_\theta, \quad \tau_N = N^{-1/2}\tau_\theta,$$

$$\mathscr{P}_{N,\zeta} = \{Q \in \mathscr{M} \mid Q \geqq (1 - \epsilon_N)P_\zeta - \delta_N \text{ on } \mathscr{B}\}, \tag{2.5}$$

$$\mathscr{P}_{N,\zeta}^N = \{\otimes_{i=1}^N Q_i \mid Q_i \in \mathscr{P}_{N,\zeta} \quad \text{for } i = 1, \cdots, N\}.$$

The generic element of $\mathscr{P}_{N,\zeta}^N$ is subsequently denoted by $W_{N,\zeta}$. It stands for the joint law of the independent observations $x_1, \cdots, x_N$ at sample size N, when the single laws are allowed to range over the class $\mathscr{P}_{N,\zeta}$, which is a natural generalization of ϵ-contamination and total variation neighborhoods (cf. [15], page 1082, Remark 5).

A new parameter space $\mathscr{S}_\theta$ has to be introduced for the local and asymptotic investigations at θ,

$$\mathscr{S}_\theta = \{\theta_N \mid \sup_N N^{1/2} |\theta_N - \theta| < \infty\}. \tag{2.6}$$

For each such sequence θ_N we define the asymptotic neighborhood H_{θ_N} of the sequence of product measures $(P_{\theta_N}^N)$ by

$$H_{\theta_N} = \{(W_{N,\theta_N}) \mid W_{N,\theta_N} \in \mathscr{P}_{N,\theta_N}^N \quad \text{for all } N\}. \tag{2.7}$$

A mathematically interesting variant is the contiguity-sub-model consisting of the classes

${}^cH_{\theta_N}$,

$$(2.8)\qquad {}^cH_{\theta_N} = \{(W_{N,\theta_N}) \in H_{\theta_N} \mid (W_{N,\theta_N}) \quad \text{contiguous to} \quad (P_\theta^N)\}.$$

In this distributional framework, an estimator $T = T_N$, i.e., a sequence of measurable mappings $T_N\colon \Omega^N \to (-\infty, \infty)$, shall be scrutinized with respect to the limiting upper probability that T_N falls below, or exceeds, the true parameter value by the amount τ_N. Formally, let a sequence of randomization constants $\gamma_N \in [0, 1]$ be given. Define random intervals C_N^T of the following form,

$$(2.9)\qquad \begin{aligned} C_N^T &= (T_N - \tau_N, T_N + \tau_N] && \text{with probability} \quad \gamma_N \\ &= [T_N - \tau_N, T_N + \tau_N) && \text{with probability} \quad 1 - \gamma_N, \end{aligned}$$

where the randomization at the boundaries is independent of the observations. Then, with self-explanatory notation, the risk $r(T; \theta_N)$, respectively ${}^cr(T; \theta_N)$, of T at $\theta_N \in \mathscr{S}_\theta$ is given by

$$(2.10)\qquad \begin{aligned} r(T; \theta_N) &= \limsup_N \sup W_{N,\theta_N}[\theta_N < C_N^T] \vee \sup W_{N,\theta_N}[\theta_N > C_N^T] \\ &= \sup_{H_{\theta_N}} \limsup_N W_{N,\theta_N}[\theta_N < C_N^T] \vee W_{N,\theta_N}[\theta_N > C_N^T], \\ {}^cr(T; \theta_N) &= \sup_{{}^cH_{\theta_N}} \limsup_N W_{N,\theta_N}[\theta_N < C_N^T] \vee W_{N,\theta_N}[\theta_N > C_N^T]. \end{aligned}$$

Occasionally, the corresponding quantities will be used when $\limsup_N$ is exchanged for $\liminf_N$ in the above expressions; they are denoted by $r_0(T; \theta_N)$, respectively by ${}^cr_0(T; \theta_N)$.

The particular choice of randomization constants γ_N is of no importance (e.g., to avoid randomization at all, γ_N could be taken 0 or 1). Moreover, if the estimator T satisfies condition (4.1) below, then the boundary points of C_N^T may be treated completely arbitrarily (e.g., C_N^T could be taken $(T_N - \tau_N, T_N + \tau_N)$ or $[T_N - \tau_N, T_N + \tau_N]$). However, form (2.9) of C_N^T seems to be necessary if no assumptions about T are made.

Some further notation is required. Given $\theta_N \in \mathscr{S}_\theta$ the sequences $\theta_{0N}, \theta_{1N} \in \mathscr{S}_\theta$ are defined by

$$(2.11)\qquad \theta_{0N} = \theta_N - \tau_N, \qquad \theta_{1N} = \theta_N + \tau_N.$$

Let $(Q^*_{N,\theta_{0N}}, Q^*_{N,\theta_{1N}})$ be any least favorable pair for $(\mathscr{P}_{N,\theta_{0N}}, \mathscr{P}_{N,\theta_{1N}})$, cf. [14]. Then the N-fold product of $Q^*_{N,\theta_{jN}}$ is abbreviated by $W^*_{N,\theta_{jN}}$, $j = 0, 1$, L_{N,θ_N} denotes their loglikelihood,

$$(2.12)\qquad L_{N,\theta_N} = \log \frac{dW^*_{N,\theta_{1N}}}{dW^*_{N,\theta_{0N}}},$$

and φ^*_{N,θ_N} stands for the Neymann-Pearson test for $W^*_{N,\theta_{0N}}$ versus $W^*_{N,\theta_{1N}}$, based on L_{N,θ_N},

$$(2.13)\qquad \varphi^*_{N,\theta_N} = (1 - \gamma_N^*) I[L_{N,\theta_N} > l_N] + \gamma_N^* I[L_{N,\theta_N} \geqq l_N],$$

where the critical value l_N and the randomization constant γ_N^* are chosen in such a way that

$$(2.14)\qquad \int \varphi^*_{N,\theta_N}\, dW^*_{N,\theta_{0N}} = \int (1 - \varphi^*_{N,\theta_N})\, dW^*_{N,\theta_{1N}} = \alpha^*_{N,\theta_N}, \quad \text{say}.$$

3. Technicalities. This section supplies the tools for the derivation of the main theorems; perhaps, the subsequent results are even of some independent interest. The first lemma is a generalization of Theorem 4.1 of [15].

LEMMA 3.1. *For $\theta_N \in \mathscr{S}_\theta$ the following asymptotic normality holds*:

$$\mathscr{L}(L_{N,\theta_N} \mid W^*_{N,\theta_{0N}}) \Rightarrow \mathscr{N}(-2\tau_\theta^2(\sigma_\theta^*)^2; 4\tau_\theta^2(\sigma_\theta^*)^2),$$

$$\mathscr{L}(L_{N,\theta_N} \mid W^*_{N,\theta_{1N}}) \Rightarrow \mathscr{N}(2\tau^2(\sigma_\theta^*)^2; 4\tau_\theta^2(\sigma_\theta^*)^2), \quad \textit{as } N \to \infty.$$

PROOF. As for the explicit form of the loglikelihood of least favorable pairs, cf. [14], [15]. Thus $L_{N,\theta_N} = 2\tau_N \sum_{i=1}^N \Lambda_N^*(x_i)$, where $\Lambda_N^* = d_N' \vee \Lambda_N \wedge d_N''$, $\Lambda_N = (1/2\tau_N) \log p_{\theta_{1N}}/p_{\theta_{0N}}$ and $d_N' = (1/2\tau_N) \log\Delta_N'$, $d_N'' = (1/2\tau_N) \log\Delta_N''$, Δ_N' and Δ_N'' being the positive numbers defined according to equations (4.11), (4.12) of [14]. In view of Lemma 4.3 of [15], whose proof can straightforwardly be extended to the present situation, we have that

$$(3.1) \qquad d_N' \to d_\theta', \qquad d_N'' \to d_\theta'' \qquad \text{as} \quad N \to \infty,$$

where d_θ', d_θ'', as well as the quantities Λ_θ^*, σ_θ^* subsequently used, have been introduced in (2.4). Since $\Lambda_N \to_{P_\theta} \Lambda_\theta$ by (2.2), (3.1) and bounded convergence imply that

$$(3.2) \qquad \int (\Lambda_N^* - \Lambda_\theta^*)^2 \, dP_\theta \to 0 \qquad \text{as} \quad N \to \infty.$$

Furthermore, $\int | dP_{\theta_{0N}} - dP_\theta | = O(N^{-1/2})$ by (2.2), and $\sup\{\int | dQ_N - dP_{\theta_{0N}} \parallel Q_N \in \mathscr{P}_{N,\theta_{0N}} \} = O(N^{-1/2})$ by the definition of neighborhoods, (2.5). Hence

$$(3.3) \qquad \int \Lambda_N^* \, dQ_N \to 0, \qquad \int (\Lambda_N^*)^2 \, dQ_N \to (\sigma_\theta^*)^2, \\ \text{uniformly in} \quad Q_N \in \mathscr{P}_{N,\theta_{0N}} \quad \text{as } N \to \infty.$$

Assume that, along some subsequence, $N^{1/2} \int \Lambda_N^* \, dQ_{N,\theta_{0N}}^*$ converges to some a_θ. Then, in view of (3.1) and (3.3), the normal convergence criterion (Loève (1977), page 328) tells us that, along this subsequence, $\mathscr{L}(L_{N,\theta_N} \mid W_{N,\theta_{0N}}^*) \Rightarrow \mathscr{N}(2\tau_\theta a_\theta;\, 4\tau_\theta^2(\sigma_\theta^*)^2)$. Because $(P_{\theta_{0N}}^N)$ and $(P_{\theta_{1N}}^N)$ are mutually contiguous, so are $(W_{N,\theta_{0N}}^*)$ and $(W_{N,\theta_{1N}}^*)$ (Lemma 4.2 of [15]). Thus necessarily

$$(3.4) \qquad N^{1/2} \int \Lambda_N^* \, dQ_{N,\theta_{0N}}^* \to -\tau_\theta(\sigma_\theta^*)^2 \qquad \text{as} \quad N \to \infty,$$

and the first part of the assertion is proved. The second one can again be obtained by a contiguity argument. □

Let Φ denote the standard normal cdf, and let $\theta_N \in \mathscr{S}_\theta$. Then the following immediate consequence of the preceding lemma on the tests φ_{N,θ_N}^* introduced in (2.13), (2.14) will be used later on:

$$(3.5) \qquad \alpha_{N,\theta_N}^* \to \Phi(-\tau_\theta \sigma_\theta^*), \qquad \text{and} \quad l_N \to 0 \qquad \text{as } N \to \infty.$$

Moreover, if ψ_{N,θ_N}^* denotes the Neyman-Pearson test for $W_{N,\theta_{0N}}^*$ versus $W_{N,\theta_{1N}}^*$ which minimizes the sum of the error probabilities, then

$$(3.6) \qquad \int \psi_{N,\theta_N}^* \, dW_{N,\theta_{0N}}^* + \int (1 - \psi_{N,\theta_N}^*) \, dW_{N,\theta_{1N}}^* \to 2\Phi(-\tau_\theta \sigma_\theta^*) \qquad \text{as} \quad N \to \infty.$$

The second lemma shows that the loglikelihoods L_{N,θ_N}, in addition to being asymptotically normal, even have nice asymptotic expansions.

LEMMA 3.2. *If $\theta_N \in \mathscr{S}_\theta$ and $N^{1/2}(\theta_N - \theta) \to h$ then*

$$L_{N,\theta_N} - \{2\tau_N \textstyle\sum_{i=1}^N \Lambda_\theta^*(x_i) - 2\tau_\theta h \displaystyle\int \Lambda_\theta^* \Lambda_\theta \, dP_\theta - \tau_\theta \epsilon_\theta (d_\theta' + d_\theta'')\} \to 0$$

in $L^2(dP_\theta^N)$, as $N \to \infty$.

PROOF. If $a_N = \int \Lambda_N^* \, dP_\theta$ then

$$\int (L_{N,\theta_N} - 2\tau_\theta N^{1/2} a_N - 2\tau_N \textstyle\sum_{i=1}^N \Lambda_\theta^*(x_i))^2 \, dP_\theta^N$$

$$= 4\tau_\theta^2 \int (\Lambda_N^* - a_N - \Lambda_\theta^*)^2 \, dP_\theta$$
$$\leqq 4\tau_\theta^2 \int (\Lambda_N^* - \Lambda_\theta^*)^2 \, dP_\theta \to 0,$$

because of $\int \Lambda_\theta^* \, dP_\theta = 0$ and (3.2). Thus it remains to prove that $N^{1/2} a_N$ tends to the appropriate limit. For this purpose observe that, by Theorem 5.2 of [14],

$$N^{1/2} \int \Lambda_N^* \, d[Q_{N,\theta_{0N}}^* - (1 - \epsilon_N) P_{\theta_{0N}}] = -\delta_\theta d_N' + (\epsilon_\theta + \delta_\theta) \, d_N''. \tag{3.7}$$

Because $\int \Lambda_N^* \, dP_{\theta_{0N}} \to 0$, (3.1), (3.4) and (3.7) entail that

$$N^{1/2} \int \Lambda_N^* \, dP_{\theta_{0N}} \to -\tau_\theta (\sigma_\theta^*)^2 + \delta_\theta d_\theta' - (\epsilon_\theta + \delta_\theta) \, d_\theta'' \quad \text{as} \quad N \to \infty. \tag{3.8}$$

The summands of $S_N' = N^{-1/2} \sum_{i=1}^N \Lambda_N^*(x_i)$ are uniformly bounded and $\operatorname{Var}(\Lambda_N^* \mid P_{\theta_{0N}}) \to (\sigma_\theta^*)^2$, by (3.1), (3.3). So, in view of the normal convergence criterion, (3.8) implies that, on one hand,

$$\mathscr{L}(S_N' \mid P_{\theta_{0N}}^N) \Rightarrow \mathscr{N}(-\tau_\theta(\sigma_\theta^*)^2 + \delta_\theta d_\theta' - (\epsilon_\theta + \delta_\theta) d_\theta''; (\sigma_\theta^*)^2) \quad \text{as} \quad N \to \infty. \tag{3.9}$$

On the other hand, it follows from the asymptotic expansion of $\log dP_{\theta_{0N}}^N / dP_\theta^N$, implied by (2.2), and by Le Cam's third lemma that for $S_N'' = N^{-1/2} \sum_{i=1}^N \Lambda_\theta^*(x_i)$ we have

$$\mathscr{L}(S_N'' \mid P_{\theta_{0N}}^N) \Rightarrow \mathscr{N}((h - \tau_\theta) \int \Lambda_\theta^* \Lambda_\theta \, dP_\theta; (\sigma_\theta^*)^2) \quad \text{as} \quad N \to \infty. \tag{3.10}$$

Since, by the first part of the proof, $S_N' - N^{1/2} a_N - S_N'' \to_{P_{\theta_{0N}}^N} 0((P_{\theta_{0N}}^N)$ contiguous to $(P_\theta^N))$, (3.9) and (3.10) cannot hold unless

$$N^{1/2} a_N \to -h \int \Lambda_\theta^* \Lambda_\theta \, dP_\theta + \tau_\theta \int \Lambda_\theta^* (\Lambda_\theta - \Lambda_\theta^*) \, dP_\theta + \delta_\theta d_\theta' - (\epsilon_\theta + \delta_\theta) \, d_\theta''$$
$$= -h \int \Lambda_\theta^* \Lambda_\theta \, dP_\theta - \frac{1}{2} \epsilon_\theta (d_\theta' + d_\theta'').$$

This proves the assertion. □

The preceding asymptotic expansion will be carried over to minimax estimators by means of the third lemma.

Let $Q_0, Q_1 \in \mathscr{M}$, $\Delta = dQ_1/dQ_0$, $S: \Omega \to (-\infty, \infty)$ measurable, $\gamma \in [0, 1]$, $\varphi = (1 - \gamma) I[S > 0] + \gamma I[S \geqq 0]$, and $\gamma^* \in [0, 1]$, $k \in [0, \infty]$, $\varphi^* = (1 - \gamma^*) I[\Delta > k] + \gamma^* I[\Delta \geqq k]$ such that $\int \varphi^* \, dQ_0 = \int (1 - \varphi^*) \, dQ_1 = \alpha^*$, say. Let $\nu^k = kQ_0 - Q_1$, $|\nu^k|$ the total variation measure of ν^k, and $\eta \geqq 0$. The lemma is certainly well known for $\eta = 0$.

LEMMA 3.3. *Assume that* $\int \varphi \, dQ_0 \vee \int (1 - \varphi) \, dQ_1 \leqq \alpha^* + \eta$. *Then*

$$|\nu^k| (A \, \triangle \, [\Delta > k]) \leqq (k + 1)\eta \quad \textit{for} \quad A = [S > 0] \quad \textit{or} \quad A = [S \geqq 0].$$

PROOF. By assumption we have

$$\int \varphi \, dQ_0 \leqq \int \varphi^* \, dQ_0 + \eta, \qquad \int (1 - \varphi) \, dQ_1 \leqq \int (1 - \varphi^*) \, dQ_1 + \eta,$$

which entails that

$$\int \varphi \, d\nu^k \leqq \int \varphi^* \, d\nu^k + (k + 1)\eta,$$

i.e.,

$$\nu^k[S>0]+\gamma\nu^k[S=0]\leqq\nu^k[\Delta>k]+\gamma^*\nu^k[\Delta=k]+(k+1)\eta.$$

Since $\nu^k[\Delta=k]=0$ this implies that

$$\nu^k(A)\leqq\nu^k[\Delta>k]+(k+1)\eta, \tag{3.11}$$

where $A=[S>0]$ if $\nu^k[S=0]\geqq 0$, $A=[S\geqq 0]$ if $\nu^k[S=0]<0$.

Let $D\in\mathscr{B}$, $D\subset[\Delta>k]\backslash A$. Then necessarily $\nu^k(D)\leqq 0$. Assume that $\nu^k(D)<-(k+1)\eta$. Put $E=A\cup D$ and conclude from (3.11) that, under this assumption, $\nu^k(E)<\nu^k[\Delta>k]$. However, $[\Delta>k]$ minimizes ν^k. Thus $\nu^k(D)\geqq-(k+1)\eta$ for all such D.

Similarly it can be shown that $\nu^k(D)\leqq(k+1)\eta$ for all $D\in\mathscr{B}$, $D\subset A\backslash[\Delta>k]$, which completes the proof. □

For the application of this lemma later on, let us mention the following fact. Given two sequences of statistics S_N, $L_N:\Omega^N\to(-\infty,\infty)$ such that

(3.12) for all $h\in(-\infty,\infty)$ and for $A_N=[S_N>h]$ or $A_N=[S_N\geqq h]$, $P_\theta^N(A_N\,\Delta\,[L_N>h])\to 0$ as $N\to\infty$,

(3.13) the laws $\mathscr{L}(L_N\mid P_\theta^N)$ are uniformly tight.

Then $S_N-L_N\to_{P_\theta^N}0$ as $N\to\infty$. (A certain converse holds also.)

Turning again to the sequences of least favorable pairs $(W^*_{N,\theta_{0N}})$, $(W^*_{N,\theta_{1N}})$, the fourth lemma states that in particular $(W^*_{N,\theta_{0N}})\in{}^cH_{\theta_{0N}}$ and $(W^*_{N,\theta_{1N}})\in{}^cH_{\theta_{1N}}$ for $\theta_N\in\mathscr{S}_\theta$; a possibly related result has also been announced by Wang (*Bull. Inst. Math. Statist.* **41,** abstract 78t-187). This lemma, which is an asymptotic version of the fact that the measures $Q^*_{N,\theta_{0N}}$, $Q^*_{N,\theta_{1N}}$ and P_θ are eventually equivalent, must be seen in line with the observation in other areas of robust statistics: the least favorable situation always looks rather innocent (e.g., Fisher information is minimized by a very smooth Lebesgue density, whose tails are unexpectedly slim, rank statistics attain their extreme laws already under continuous i.i.d. distributions, cf. also Lemma 2.1 of [17]).

LEMMA 3.4. *For $\theta_N\in\mathscr{S}_\theta$ the sequences $(W^*_{N,\theta_{jN}})$ and (P_θ^N) are mutually contiguous, $j=0$, 1.*

PROOF. According to Corollary 1 of Oosterhoff and van Zwet (1975), or Lemma A3 of Le Cam and Traxler (1978), the following conditions are necessary and sufficient for mutual contiguity of $(W^*_{N,\theta_{0N}})$ and $(P^N_{\theta_{0N}})$:

(3.14) these measures do not separate entirely, for instance sums of squares of Hellinger distances stay bounded;

$$\lim_{b\to\infty}\lim_N N\cdot Q^*_{N,\theta_{0N}}\left[\frac{dQ^*_{N,\theta_{0N}}}{dP_{\theta_{0N}}}>b\right]=0; \tag{3.15}$$

$$\lim_{b\to\infty}\lim_N N\cdot P_{\theta_{0N}}\left[\frac{dQ^*_{N,\theta_{0N}}}{dP_{\theta_{0N}}}<\frac{1}{b}\right]=0. \tag{3.16}$$

By Theorem 5.2 of [14] we have that $Q^*_{N,\theta_{0N}}\ll P_{\theta_{0N}}+P_{\theta_{1N}}$, and, if we let $q^*_{N,\theta_{0N}}$ denote the μ-density of $Q^*_{N,\theta_{0N}}$, that the following relations must hold μ a.e.

$$q^*_{N,\theta_{0N}}=(1-\epsilon_N)p_{\theta_{0N}}\quad\text{on}\quad\left[\Delta'_N\leqq\frac{p_{\theta_{1N}}}{p_{\theta_{0N}}}\leqq\Delta''_N\right]$$

$$\frac{(1-\epsilon_N)}{\Delta'_N}p_{\theta_{1N}}\leqq q^*_{N,\theta_{0N}}\leqq(1-\epsilon_N)p_{\theta_{0N}}\quad\text{on}\quad\left[\frac{p_{\theta_{1N}}}{p_{\theta_{0N}}}<\Delta'_N\right]$$

$$(1-\epsilon_N)p_{\theta_{0N}}\leqq q^*_{N,\theta_{0N}}\leqq\frac{(1-\epsilon_N)}{\Delta''_N}p_{\theta_{1N}}\quad\text{on}\quad\left[\frac{p_{\theta_{1N}}}{p_{\theta_{0N}}}>\Delta''_N\right],$$

where Δ'_N, Δ''_N are the numbers introduced in the proof to the first lemma; $0 < \Delta'_N < 1 < \Delta''_N < \infty$.

This implies that

$$\int ((q^*_{N,\theta_{0N}})^{1/2} - (p_{\theta_{0N}})^{1/2})^2 \, d\mu$$

$$\leqq 2(1-\epsilon_N) \int ((p_{\theta_{1N}})^{1/2} - (p_{\theta_{0N}})^{1/2})^2 \, d\mu + 2(1-(1-\epsilon_N)^{1/2})^2. \tag{3.17}$$

Moreover, because $\Delta'_N, \Delta''_N \to 1$ by (3.1), (3.17) further entails that for $b > 1$ eventually

$$Q^*_{N,\theta_{0N}}\left[\frac{q^*_{N,\theta_{0N}}}{p_{\theta_{0N}}} > b\right] \leqq P_{\theta_{1N}}\left[\frac{p_{\theta_{1N}}}{p_{\theta_{0N}}} > b\right] \tag{3.18}$$

$$P_{\theta_{0N}}\left[\frac{q^*_{N,\theta_{0N}}}{p_{\theta_{0N}}} < \frac{1}{b}\right] \leqq P_{\theta_{0N}}\left[\frac{p_{\theta_{1N}}}{p_{\theta_{0N}}} < \frac{1}{b}\right]. \tag{3.19}$$

From (3.17) – (3.19) it follows that conditions (3.14) – (3.16) are satisfied whenever $(P^N_{\theta_{0N}})$ and $(P^N_{\theta_{1N}})$ are mutually contiguous. The assertion concerning $(W^*_{N,\theta_{1N}})$ can be obtained by similar arguments, or directly from Lemma 4.2 of [15]. □

4. The main results. Local uniformity considerations play an important role in modern asymptotic estimation theory. On one hand, locally uniform weak convergence of estimators is a basic requirement for the Hajek-Le Cam convolution representation theorem; on the other hand, the local asymptotic minimax criterion has proved to cut out the artificial, formerly superefficient, estimators, in an intuitively appealing and elegant way; cf. Hajek (1970, 1972) and Le Cam (1972). Although not very explicitly, uniformity has also been made use of in [17]. If this is spelled out, a more natural and more general proof of this asymptotic minimax result is possible, which may serve as a motivation for the approach to follow.

As for the notions of stochastic limit superior and stochastic limit inferior, the reader is referred to [16]. The estimators T now considered are supposed to have extreme limit laws with respect to ${}^cH_\theta$ (θ also denotes the sequence identically θ) that are normal and locally uniformly approached; i.e., there exist $s'_\theta\, s''_\theta \in [0,)$ and $\sigma_\theta \in (0, \infty)$ such that for all $\theta_N \in \mathscr{S}_\theta$

$$\begin{aligned} &\lim \sup\{(\mathscr{L}(N^{1/2}(T_N - \theta_N)|\, W_{N,\theta_N}))|(W_{N,\theta_N}) \in {}^cH_{\theta_N}\} = \mathscr{N}(s'_\theta; \sigma^2_\theta) \\ &\lim \inf\{(\mathscr{L}(N^{1/2}(T_N - \theta_N)|(W_{N,\theta_N}))\,|\,(W_{N,\theta_N}) \in {}^cH_{\theta_N} = \mathscr{N}(-s''_\theta; \sigma^2_\theta). \end{aligned} \tag{4.1}$$

Evaluation of (4.1) at the sequence identically θ shows that the risk defined by (2.10) is of the form

$${}^cr(T;\theta) = \Phi\left(\frac{s'_\theta - \tau_\theta}{\sigma_\theta}\right) \vee \Phi\left(\frac{s''_\theta - \tau_\theta}{\sigma_\theta}\right). \tag{4.2}$$

The assertion then is that

$$\Phi\left(\frac{s'_\theta - \tau_\theta}{\sigma_\theta}\right) \vee \Phi\left(\frac{s''_\theta - \tau_\theta}{\sigma_\theta}\right) \geqq \Phi(-\tau_\theta\sigma^*_\theta). \tag{4.3}$$

To prove it we may pick any $\theta_N \in \mathscr{S}_\theta$ and employ T_n as a test statistic for $W^*_{N,\theta_{0N}}$ versus $W^*_{N,\theta_{1N}}$; $\varphi_N = I[N^{1/2}(T_N - \theta_N) > 0]$. We choose η such that $\Phi\,((s'_\theta - \tau_\theta)/\sigma_\theta) = \Phi(-\tau_\theta\sigma^*_\theta - \eta)$ and let χ^*_N be the Neyman-Pearson tests for $W^*_{N,\theta_{0N}}$ versus $W^*_{N,\theta_{1N}}$ such that $\lim_N \int \chi^*_N \, dW^*_{N,\theta_{0N}} = \Phi(-\tau_\theta\sigma^*_\theta - \eta)$. By Lemma 3.1, $\lim_N \int (1 - \chi^*_N)\, dW^*_{N,\theta_{1N}} = \Phi(-\tau_\theta\sigma^*_\theta + \eta)$. As $\lim\sup_N \int \varphi_N \, dW^*_{N,\theta_{0N}} \leqslant \Phi((s'_\theta - \tau_\theta)/\sigma_\theta)$ it follows that $\Phi((s''_\theta - \tau_\theta)/\sigma_\theta) \geqslant \Phi(-\tau_\theta\sigma^*_\theta + \eta)$. So $\Phi((s'_\theta - \tau_\theta)/\sigma_\theta) \vee \Phi((s''_\theta - \tau_\theta)/\sigma_\theta) \geqslant \Phi(-\tau_\theta\sigma^*_\theta - \eta) \vee \Phi(-\tau_\theta\sigma^*_\theta + \eta)$ which is $\geqslant \Phi(-\tau_\theta\sigma^*_\theta)$; this proves (4.3).

The estimators $T = T(IC)$ considered in [17] satisfy (4.1) with $s'_\theta = s'_\theta(IC_\theta)$, $s''_\theta = s''_\theta(IC_\theta)$

and $\sigma_\theta^2 = \int IC_\theta^2 \, dP_\theta$; cf. [17]. Indeed, the proof to Lemma 2.1 of [17] is still applicable in the case of moving centers P_{θ_N}; moreover, we can make use of $\int IC_\theta \Lambda_\theta \, dP_\theta = 1$. Thus it appears that the essential purpose of the asymptotic expansions assumed in [17] has been to guarantee (4.1). Besides being more general, the present proof also clarifies the relation to testing which has only formally shown up in [17].

REMARK. Condition (4.1) appears to be a natural, though more restrictive, analog of the regularity assumption used for the convolution theorem. However, since the present risk not only involves asymptotic variance but simultaneously bias, a convolution representation cannot be expected in this framework.

If condition (4.1) is dropped then the bound ${}^c r(T; \theta) \geqq \Phi(-\tau_\theta \sigma_\theta^*)$ is no longer valid. Superefficient estimators can be constructed in the following way. Given an estimator T^0 such that for every $\zeta \in \Theta$ the laws $\mathscr{L}(N^{1/2}(T_N^0 - \zeta) | P_\zeta^N)$ are uniformly tight. Let $T_N = \theta + (T_N^0 - \theta) I[|T_N^0 - \theta| > N^{-1/4}]$. Then $W_{N,\zeta}[T_N = T_N^0] \to 1$ as $N \to \infty$ for all $(W_{N,\zeta}) \in {}^c H_\zeta$ and so ${}^c r(T; \zeta) = {}^c r(T^0; \zeta)$ if $\zeta \neq \theta$; whereas ${}^c r(T; \theta) = 0$ because $W_{N,\theta}[T_N = \theta] \to 1$ as $N \to \infty$ for all $(W_{N,\theta}) \in {}^c H_\theta$. Thus, in the absence of uniformity, the risk ${}^c r(T; \theta)$ is not a meaningful quantity.

REMARKS. 1. With obvious modifications, the preceding considerations carry over to the full neighborhood model, i.e., with the superscripts ^{c} dropped throughout.

2. The problem whether the set of superefficiency must necessarily have Lebesgue measure zero is still unsolved. Bahadur's (1964) most lucid proof in the classical case does not seem to carry over due to measurability difficulties.

When no restrictions shall be imposed on the estimators, the appropriate method will be to evaluate the risk uniformly. That is, we consider the local asymptotic maximum risk of estimators T at θ which in view of the definitions in Section 2 is naturally given by

$$\begin{aligned} R(T; \theta) &= \sup_{\mathscr{S}_\theta} r(T; \theta_N) \\ &= \sup_{\theta_N} \sup_{H_{\theta_N}} \limsup_N W_{N,\theta_N}[\theta_N < C_N^T] \vee W_{N,\theta_N}[\theta_N > C_N^T] \\ &= \lim_{K\to\infty} \limsup_N \sup_{|\zeta - \theta| < K_N} \sup W_{N,\zeta}[\zeta < C_N^T] \vee W_{N,\zeta}[\zeta > C_N^T], \end{aligned} \tag{4.4}$$

where $K_N = N^{-1/2} K$, respectively in the contiguity-submodel by

$$\begin{aligned} {}^c R(T; \theta) &= \sup_{\mathscr{S}_\theta} {}^c r(T; \theta_N) \\ &= \sup_{\theta_N} \sup_{{}^c H_{\theta_N}} \limsup_N W_{N,\theta_N}[\theta_N < C_N^T] \vee W_{N,\theta_N}[\theta_N > C_N^T]. \end{aligned} \tag{4.5}$$

The corresponding quantities when the sign $\limsup_N$ is replaced by $\liminf_N$ in the above expressions are denoted by $R_0(T; \theta)$, respectively by ${}^c R_0(T; \theta)$.

THEOREM 4.1. *For any estimator T, the quantities $R(T; \theta)$, ${}^c R(T; \theta)$, $R_0(T; \theta)$ and ${}^c R_0(T; \theta)$ are bounded from below by the number $\Phi(-\tau_\theta \sigma_\theta^*)$.*

PROOF. Subsequently, if T is an estimator, γ_N the randomization constants occurring in (2.9), and $\zeta \in \Theta$, we shall employ the notation

$$\varphi_{N,\zeta}^T = (1 - \gamma_N) I[N^{1/2}(T_N - \zeta) > 0] + \gamma_N I[N^{1/2}(T_N - \zeta) \geqq 0]. \tag{4.6}$$

Let $\theta_N \in \mathscr{S}_\theta$ and $(W_{N,\theta_{jN}}) \in H_{\theta_{jN}}, j = 0, 1$. Then, by the definition of C_N^T,

$$\begin{aligned} W_{N,\theta_{0N}}[\theta_{0N} < C_N^T] &= \int \varphi_{N,\theta_N}^T \, dW_{N,\theta_{0N}} \\ W_{N,\theta_{1N}}[\theta_{1N} > C_N^T] &= \int (1 - \varphi_{N,\theta_N}^T) \, dW_{N,\theta_{1N}}. \end{aligned} \tag{4.7}$$

Now, in view of Lemma 3.4 and Lemma 3.1, (3.5), we may conclude that

$$\sup_{{}^cH_{\theta_{0N}}, {}^cH_{\theta_{1N}}} \liminf_N W_{N,\theta_{0N}}[\theta_{0N} < C_N^T] \vee W_{N,\theta_{1N}}[\theta_{1N} > C_N^T]$$

$$\geqq \liminf_N W^*_{N,\theta_{0N}}[\theta_{0N} < C_N^T] \vee W^*_{N,\theta_{1N}}[\theta_{1N} > C_N^T]$$

$$= \liminf_N \int \varphi^T_{N,\theta_N}\, dW^*_{N,\theta_{0N}} \vee \int (1 - \varphi^T_{N,\theta_N})\, dW^*_{N,\theta_{1N}}$$

$$\geqq \lim_N \int \varphi^*_{N,\theta_N}\, dW^*_{N,\theta_{0N}} \vee \int (1 - \varphi^*_{N,\theta_N})\, dW^*_{N,\theta_{1N}}$$

$$= \Phi(-\tau_\theta \sigma^*_\theta).$$

So the theorem follows if we can show that

$$(4.8) \qquad {}^cR_0(T;\theta) \geqq \sup_{\theta_N} \sup_{{}^cH_{\theta_{0N}}, {}^cH_{\theta_{1N}}} \liminf_n W_{N,\theta_{0N}}[\theta_{0N} < C_N^T] \vee W_{N,\theta_{1N}}[\theta_{1N} > C_N^T].$$

Given $\theta_N \in \mathscr{S}_\theta$ and $(W_{N,\theta_{jN}}) \in {}^cH_{\theta_{jN}}, j = 0, 1$ define another sequence ζ_N as follows,

$$\zeta_N = \theta_{0N} \quad \text{if} \quad W_{N,\theta_{0N}}[\theta_{0N} < C_N^T] \geqq W_{N,\theta_{1N}}[\theta_{1N} > C_N^T]$$

$$= \theta_{1N} \quad \text{if} \quad W_{N,\theta_{0N}}[\theta_{0N} < C_N^T] < W_{N,\theta_{1N}}[\theta_{1N} > C_N^T].$$

Then $\zeta_N \in \mathscr{S}_\theta$, $(W_{N,\zeta_N}) \in {}^cH_{\zeta_N}$ and

$$W_{N,\zeta_N}[\zeta_N < C_N^T] \vee W_{N,\zeta_N}[\zeta_N > C_N^T] \geqq W_{N,\theta_{0N}}[\theta_{0N} < C_N^T] \vee W_{N,\theta_{1N}}[\theta_{1N} > C_N^T].$$

(By a similar argument we even have equality in (4.8)). □

The possibly most suggestive version of the local asymptotic minimax bound is the following, which says that the limit of the minimax risks of the experiments at time N is at least $\Phi(-\tau_\theta\sigma^*_\theta)$.

THEOREM 4.1 *A. The following bound holds,*

$$\lim_{K\to\infty} \liminf_N \inf_T \sup_{|\zeta-\theta|<K_N} \sup W_{N,\zeta}[\zeta < C_N^T] \vee W_{N,\zeta}[\zeta > C_N^T] \geqq \Phi(-\tau_\theta\sigma^*_\theta).$$

PROOF. Since for every estimator T

$$\sup_{|\zeta-\theta|\leqq\tau_N} \sup W_{N,\zeta}[\zeta < C_N^T] \vee W_{N,\zeta}[\zeta > C_N^T]$$
$$\geqq W^*_{N,\theta-\tau_N}[\theta - \tau_N < C_N^T] \vee W^*_{N,\theta+\tau_N}[\theta + \tau_N > C_N^T] \geqq \alpha^*_{N,\theta},$$

in view of (4.6), (4.7), it follows that an even stronger statement is true, namely

$$\liminf_N \inf_T \sup_{|\zeta-\theta|\leqq\tau_N} \sup W_{N,\zeta}[\zeta < C_N^T] \vee W_{N,\zeta}[\zeta > C_N^T] \geqq \Phi(-\tau_\theta\sigma^*_\theta). \qquad \square$$

By an estimator which is locally asymptotically minimax at θ we shall mean any estimator T that satisfies $R(T;\theta) = \Phi(-\tau_\theta\sigma^*_\theta)$, respectively ${}^cR(T;\theta) = \Phi(-\tau_\theta\sigma^*_\theta)$, depending on the model considered. The next theorem asserts that such estimators are automatically also locally asymptotically admissible at θ. In view of the infinite dimensional aspects of the problem the admissibility result is fairly surprising even though we deliberately employ such pseudoloss functions that essentially enable us to treat estimators as test statistics.

THEOREM 4.2. *Let the estimator T be locally asymptotically minimax at θ. Then T is locally asymptotically admissible at θ.*

PROOF. The theorem follows if estimators T that are locally asymptotically minimax at θ necessarily have constant local risk at θ, i.e., $r(T;\theta_N) = \Phi(-\tau_\theta\sigma_\theta^*)$, respectively ${}^c r(T;\theta_N) = \Phi(-\tau_\theta\sigma_\theta^*)$, for all $\theta_N \in \mathscr{S}_\theta$. This in turn is a consequence of the following implication, which holds for any estimator T:

If $r_0(T;\zeta_N) < \Phi(-\tau_\theta\sigma_\theta^*)$ for some $\zeta_N \in \mathscr{S}_\theta$ then

$r(T;\theta_N) > \Phi(-\tau_\theta\sigma_\theta^*)$ for another $\theta_N \in \mathscr{S}_\theta$. (Similarly, if superscripts ^{c} are added.)

To prove this, assume on the contrary that $\limsup_N \sup W_{N,\theta_N}[\theta_N < C_N^T] \vee W_{N,\theta_N}[\theta_N > C_N^T] \leqq \Phi(-\tau_\theta\sigma_\theta^*)$, for all $\theta_N \in \mathscr{S}_\theta$. Using the sequence $\xi_N = \zeta_{1N}$ and (4.6), (4.7), we conclude that

$$\liminf_N \int \varphi_{N,\zeta_N}^T \, dW_{N,\xi_{0N}}^* < \Phi(-\tau_\theta\sigma_\theta^*),$$

$$\limsup_N \int (1-\varphi_{N,\zeta_N}^T)\, dW_{N,\xi_{1N}}^* \leqq \Phi(-\tau_\theta\sigma_\theta^*).$$

However, this is a contradiction to Lemma 3.1, (3.6). □

A closer analysis allows the following sharper version of this theorem that seems particularly suited to demonstrate the erratic behavior of formerly superefficient estimators. Note that $\Phi(-\tau_\theta\sigma_\theta^*) - \Phi(-\tau_\theta\sigma_\theta^* - \eta) < \Phi(-\tau_\theta\sigma_\theta^* + \eta) - \Phi(-\tau_\theta\sigma_\theta^*)$ for all $\eta > 0$. Thus it will follow that the amount by which the local risk can fall below the minimax risk at one local parameter point is necessarily more than offset by the increase of the risk at another point. For example, if ${}^c r_0(T;\theta) = 0$ then ${}^c r(T;\theta_N) = 1$ for the two sequences $\theta_N = \theta - 2\tau_N$ and $\theta_N = \theta + 2\tau_N$.

THEOREM 4.2 A. *Let T be an estimator such that $r_0(T;\zeta_N) \leqq \Phi(-\tau_\theta\sigma_\theta^* - \eta)$ for some $\zeta_N \in \mathscr{S}_\theta$ and some $\eta > 0$. Then there is another sequence $\theta_N \in \mathscr{S}_\theta$ such that $r(T;\theta_N) \geqq \Phi(-\tau_\theta\sigma_\theta^* + \eta)$. (Similarly, if superscripts ^{c} are added.)*

PROOF. According to the preceding proof, $\liminf_N \int \varphi_{N,\xi_N}^T \, dW_{N,\xi_{0N}}^* \leqq \Phi(-\tau_\theta\sigma_\theta^* - \eta)$ for $\xi_N = \zeta_{1N}$. Let χ_{N,ξ_N}^* denote the Neyman-Pearson test for $W_{N,\xi_{0N}}^*$ versus $W_{N,\xi_{1N}}^*$ such that $\lim_N \int \chi_{N,\xi_N}^* \, dW_{N,\xi_{0N}}^* = \Phi(-\tau_\theta\sigma_\theta^* - \eta)$. Then Lemma 3.1 implies that $\lim_N \int (1-\chi_{N,\xi_N}^*)\, dW_{N,\xi_{1N}}^* = \Phi(-\tau_\theta\sigma_\theta^* + \eta)$. Thus $\limsup_N \int (1-\varphi_{N,\xi_N}^T)\, dW_{N,\xi_{1N}}^* \geqq \Phi(-\tau_\theta\sigma_\theta^* + \eta)$, and so $r(T;\xi_{1N}) \geqq \Phi(-\tau_\theta\sigma_\theta^* + \eta)$. □

The last theorem gives the necessary asymptotic expansion at θ of estimators that are locally asymptotically minimax at θ. This expansion uniquely determines the asymptotic behavior of these estimators under each $(W_{N,\theta_N}) \in {}^c H_{\theta N}$, $\theta_N \in \mathscr{S}_\theta$; in particular, another admissibility proof for the contiguity-submodel is implied.

Recall the definition of IC_θ^*, (2.4).

THEOREM 4.3. *If T is an estimator such that ${}^c R(T;\theta) = \Phi(-\tau_\theta\sigma_\theta^*)$ then*

$$N^{1/2}(T_N - \theta) - \{N^{-1/2}\textstyle\sum_{i=1}^N IC_\theta^*(x_i) - \tfrac{1}{2}\epsilon_\theta(\inf IC_\theta^* + \sup IC_\theta^*)\} \to_{P_\theta^N} 0 \qquad \textit{as } N \to \infty. \tag{4.9}$$

PROOF. Analogously to (4.8) we have

$${}^c R(T;\theta) \geqq \sup_{\theta_N} \sup_{{}^c H_{\theta_{0N}}, {}^c H_{\theta_{1N}}} \limsup_N W_{N,\theta_{0N}}[\theta_{0N} < C_N^T] \vee W_{N,\theta_{1N}}[\theta_{1N} > C_N^T].$$

Thus, employing the notation introduced in (2.13), (2.14), (4.6), the assumption entails: for each $\theta_N \in \mathscr{S}_\theta$ there exists a nullsequence η_N such that for all N

$$\int \varphi_{N,\theta_N}^T \, dW_{N,\theta_{0N}}^* \vee \int (1-\varphi_{N,\theta_N}^T)\, dW_{N,\theta_{1N}}^* \leqq \alpha_{N,\theta_N}^* + \eta_N.$$

Hence, by virtue of Lemma 3.3,

$$| \mu^{l_N}_{N,\theta_N} |(D^{l_N}_{N,\theta_N}) \leqq (e^{l_N} + 1)\eta_N,$$

where $\mu^{l_N}_{N,\theta_N} = e^{l_N} \cdot W^*_{N,\theta_{0N}} - W^*_{N,\theta_{1N}}$, $D^{l_N}_{N,\theta_N} = A_{N,\theta_N}\Delta[L_{N,\theta_N} > l_N]$, and $A_{N,\theta_N} = [N^{1/2}(T_N - \theta_N) > 0]$ or $A_{N,\theta_N} = [N^{1/2}(T_N - \theta_N) \geqq 0]$. Because $l_N \to 0$ (Lemma 3.1, (3.5)) this further implies that

$$(W^*_{N,\theta_{0N}} - W^*_{N,\theta_{1N}})(B_N) \to 0 \qquad \text{as} \quad N \to \infty$$

for every sequence of measurable subsets B_N of $D^{l_N}_{N,\theta_N}$. We insert $B_N = D^{l_N}_{N,\theta_N} \cap [L_{N,\theta_N} \geqq t]$, and $B_N = D^{l_N}_{N,\theta_N} \cap [L_{N,\theta_N} \leqq -t]$, $t > 0$, and use the asymptotic normality of L_{N,θ_N} under $W^*_{N,\theta_{0N}}$ (Lemma 3.1), in order to conclude that

$$W^*_{N,\theta_{0N}}(A_{N,\theta_N}\Delta[L_{N,\theta_N} > 0]) \to 0 \qquad \text{as} \quad N \to \infty.$$

Now the asymptotic expansion of L_{N,θ_N} (Lemma 3.2), again the asymptotic normality of $\mathscr{L}(L_{N,\theta_N} \mid W^*_{N,\theta_{0N}})$, and the contiguity of $(W^*_{N,\theta_{0N}})$ to (P^N_θ) (Lemma 3.4), can be invoked to the effect that, for all sequences $\theta_N = \theta + N^{-1/2} h$, $h \in (-\infty, \infty)$,

$$P^N_\theta(A_{N,\theta_N}\Delta[L_N > h]) \to 0 \qquad \text{as} \quad N \to \infty,$$

where $L_N = N^{-1/2} \sum_{i=1}^N IC^*_\theta(x_i) - \frac{1}{2}\epsilon_\theta(\inf IC^*_\theta + \sup IC^*_\theta)$. Thus (3.12), (3.13) are fulfilled for this L_N and $S_N = N^{1/2}(T_N - \theta)$. The assertion of the theorem follows. □

Assume now that T is an estimator such that

$$(4.10) \qquad N^{1/2}(T_N - \theta) - \{N^{-1/2} \textstyle\sum_{i=1}^N IC^*_\theta(x_i) - \frac{1}{2}\epsilon_\theta(\inf IC^*_\theta + \sup IC^*_\theta)\} \to_{W_{N,\theta_N}} 0 \qquad \text{as} \quad N \to \infty$$

for all $(W_{N,\theta_N}) \in H_{\theta_N}$, respectively $(W_{N,\theta_N}) \in {}^cH_{\theta_N}$, $\theta_N \in \mathscr{S}_\theta$. (Given (4.9), this is no extra requirement in the contiguity-sub-model.) Then, as already employed at the beginning of this section, the results of [15], [17] show that (4.10) is sufficient for $r(T; \theta_N) = \Phi(-\tau_\theta\sigma^*_\theta)$, respectively $cr(T; \theta_N) = \Phi(-\tau_\theta\sigma^*_\theta)$, for all $\theta_N \in S_\theta$, i.e., for T to be locally asymptotically minimax at θ. As for the construction of estimators that satisfy (4.10) for all $\theta \in \theta$, see the corresponding discussion in [17].

Acknowledgment. The author would like to express his gratitude to Professors R. J. Beran, P. J. Bickel, and L. Le Cam for stimulating and helpful discussions during his stay at Berkeley.

REFERENCES

[1] BAHADUR, R. R. (1964). On Fisher's bound for asymptotic variances. *Ann. Math. Statist.* **35** 1545–1552.

[2] BERAN, R. J. (1979). Efficient robust estimates in parametric models. Unpublished manuscript.

[3] BERAN, R. J. (1980). Asymptotic lower bounds for risk in robust estimation. *Ann. Statist.* **8** 1252–1264.

[4] CHERNOFF, H. (1956). Large sample theory: parametric case. *Ann. Math. Statist.* **27** 1–22.

[5] HÁJEK, J. (1970). A characterization of limiting distributions of regular estimates. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **14** 323–330.

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

[7] HUBER, P. J. (1965). Strict efficiency excludes superefficiency. Unpublished manuscript.

[8] LE CAM, L. (1953). On some asymptotic properties of maximum likelihood estimates and related Bayes' estimates. *Univ. California Publications in Statist.* **1** 277–330.

[9] LE CAM, L. (1972). Limits of experiments. *Proc. Sixth Berkeley Symp. Math. Statist. Prob.* **1** 245–261. Univ. California Press.

[10] LE CAM, L. and TRAXLER, R. (1978). On the asymptotic behavior of mixtures of Poisson distributions. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **44** 1–45.

[11] LOÈVE, M. (1977). *Probability Theory I.* 4th ed. Springer, New York, Heidelberg, Berlin.

[12] MILLAR, P. W. (1979). Robust estimation via minimum distance methods. Unpublished manuscript.
[13] OOSTERHOFF, J. and VAN ZWET, W. R. (1975). A note on contiguity and Hellinger distance. Technical report, Mathematisch Centrum, Amsterdam.
[14] RIEDER, H. (1977). Least favorable pairs for special capacities. *Ann. Statist.* **5** 909–921.
[15] RIEDER, H. (1978). A robust asymptotic testing model. *Ann. Statist.* **6** 1080–1094.
[16] RIEDER, H. (1980). Estimates derived from robust tests. *Ann. Statist.* **8** 106–115.
[17] RIEDER, H. (1981). Robustness of one and two sample rank tests against gross errors. *Ann. Statist.* **9**.

INSTITUT FÜR MATHEMATISCHE STOCHASTIK
DER UNIVERSITÄT FREIBURG
HERMANN-HERDER-STRASSE 10
D-7800 FREIBURG I. BR.
WEST GERMANY