Z. Wahrscheinlichkeitstheorie verw. Gebiete
59, 459–478 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On Estimation and Adaptive Estimation for Locally Asymptotically Normal Families*

Václav Fabian[1,2] and James Hannan[2]

[1] University of Bern, Switzerland
[2] Michigan State University, Dept. of Statistics and Probability, East Lansing MI 48824, USA

Summary. Locally asymptotically minimax (LAM) estimates are constructed for locally asymptotically normal (LAN) families under very mild additional assumptions. Adaptive estimation is also considered and a sufficient condition is given for an estimate to be locally asymptotically minimax adaptive. Incidently, it is shown that a well known lower bound due to Hájek (1972) for the local asymptotic minimax risk is not sharp.

1. Introduction

Elaborating on work of others, Hájek (1972, Theorems 4.1 and 4.2) established a lower bound for the local asymptotic minimax risk of a sequence of estimates $\langle Z_n \rangle$ under the Condition LAN (cf. our Theorem 2.5) and showed that, in the one-dimensional case, a condition close to our regularity (cf. Definition 6.2) is necessary for $\langle Z_n \rangle$ to be locally asymptotically minimax (HLAM) in the rough sense that, for $\langle Z_n \rangle$, the lower bound is attained (cf. Definition 3.1).

Sharper inequalities were obtained by Ibragimov and Has'minskij (1979, Remark 12.2, Chap. II), Le Cam (1979, p. 134) and we give such a sharper inequality in Theorem 2.6 below. We shall say that a sequence of estimates $\langle Z_n \rangle$ is LAM if it attains the lower bound given in Theorem 2.6 (cf. Definition 3.1).

The paper begins with three simple results.

(i) Hájek's inequality is not sharp, i.e., there are cases in which Condition LAN holds and no $\langle Z_n \rangle$ is HLAM.

(ii) The sharper inequality of Theorem 2.6 involves a bound which is attainable: under Condition LAN, there is always an LAM sequence of estimates.

(iii) Under LAN, if $\langle Z_n \rangle$ is a regular sequence of estimates then $\langle Z_n \rangle$ is LAM (Theorem 6.3).

* Research partially supported by NSF grants no. MCS 78-02846 and MCS 77-03493-01

0044-3719/82/0059/0459/$04.00

Each of these results is easy to prove, nevertheless seems to be proven here for the first time.

Some details were left unexplained above and will be attended to now: for the attainability, we consider only bounded loss functions. A sequence of estimates $\langle Z_n \rangle$ may be LAM (HLAM) at a θ or at each θ in Θ. We shall show that Condition LAN at θ does not imply the existence of an HLAM at θ sequence of estimates (Remark 3.2) but does imply the existence of an LAM at θ sequence of estimates (Remark 6.4). Sequences $\langle Z_n \rangle$ which are LAM (θ) for each θ are constructed under mild additional conditions (Theorem 6.15 and Remarks 6.17 and 6.16).

We have talked about the LAN Condition, but there is a complication in that Hâjek's LAN Condition is weaker than the condition considered by Le Cam (1960) (see Remark 2.4). The usual sufficient conditions for LAN yield the stronger version of LAN (cf. Fabian and Hannan (1980)). We shall refer to the weaker of the two LAN conditions by HLAN, and reserve LAN for the stronger condition.

Under certain rather general conditions, Le Cam (1969, proof of Theorem 4) constructs a regular sequence of estimates which by our result (iii) is LAM. We obtain regular and LAM estimates under weaker conditions than Le Cam; we were influenced by the discussions in Le Cam (1974, Sect. 12).

Particular estimates have been proved HLAM (and thus LAM) (e.g. Levit (1974, 1975), Koshevnik and Levit (1976) and Has'minskij and Ibragimov (1979)) by ad hoc arguments showing the asymptotic distribution of the estimate is approached uniformly with respect to the parameter values in a given neighborhood.

The results on LAM estimates are used to reformulate Stein's (1956) heuristic arguments and to obtain a definition of LAM adaptivity of estimates, a necessary condition for the existence of an LAM adaptive estimate and a sufficient condition for an estimate to be LAM adaptive. The results are also related to the results by Levit, Koshevnik, Has'minskij and Ibragimov (cf. references above) who use the LAM criterion in a non-parametric situation. Our sufficient condition is used in Fabian (1980) to prove the LAM-adaptivity of a recursive location parameter estimate and its validity was proved by Beran (1978) for his estimate.

The LAM-adaptivity is a stronger property than the adaptivity hitherto considered, since it again involves uniform convergence. As a consequence, the result in Fabian (1980) refutes an opinion widely spread in the statistical folklore that adaptive procedures converge to the asymptotic distribution *non-uniformly* and are therefore of little practical use (for a written comment of this type, see, e.g., Hâjek's (1971, Example 7) remarks on his own adaptive test; it is unclear whether Hâjek had a proof of the non-uniformity he claimed). Objections can be made that the neighborhoods over which the supremum is taken are small but that is true for the LAM considerations in general, not only for the LAM adaptive procedures.

In a paper published after the submission of the present paper, Beran (1980) treats a related but different situation and presents an attainable asymptotic lower bound. In his conditions, Θ is not a subset of Euclidean space, but a specialized parameterization is used and the loss function is more restricted than in our considerations.

Some notation will be introduced next. If P, Q are probabilities on a σ-algebra $\mathbf{X}$ and Q_+ is the absolutely continuous, with respect to P, part of Q, then any Radon-Nikodym derivative of Q_+ with respect to P will be called a pseudodensity of Q with respect to P, and also, a pseudodensity of $\int \cdot dQ$ with respect to $\int \cdot dP$. We shall talk frequently about expectations using terms which make sense when applied to the probabilities. If E and F are expectations on a σ-algebra $\mathbf{X}$, then dF/dE denotes the set of all pseudodensities of F with respect to E which are non-negative and finite valued.

R^k denotes the k-dimensional Euclidean space, $\mathbf{B}_k$ the σ-algebra of the Borel subsets of R^k, $R=R^1$, $\mathbf{B}=\mathbf{B}_1$, $\mathscr{B}_k$ is the family of all real valued Borel functions on R^k, $\mathscr{B}_{k+}=\{f; f\in\mathscr{B}_k, f\geqq 0\}$. $\langle\ \rangle$ will be used to denote finite or infinite sequences, and, in particular, points in R^k. In matrix calculations, points in R^k are columns. By $\mathscr{C}_k$ we denote the set of all orthogonal $k\times k$ matrices.

If P is a probability on $\mathbf{X}$, $E=\int \cdot dP$, h an $\langle \mathbf{X}, \mathbf{Y}\rangle$ measurable transformation then E^h denotes the expectation (on $\mathbf{Y}$) induced by h, i.e., $E^h g = E g\circ h = \int g dP h^{-1}$ for each g non-negative and $\mathbf{Y}$-measurable. $\mathscr{N}_t$ denotes the integral with respect to the normal $(t, \mathbf{1})$ distribution for $t\in R^k$; $\mathbf{1}$ denotes the identity matrix. We shall abbreviate $\mathscr{N}_0$ to $\mathscr{N}$. The dimension is not displayed in $\mathscr{N}$ but will be clear from the context. We define a function ρ_t on R^k by $\rho_t(x) = e^{t'x-\|t\|^2/2}$. Note that $\rho_t\in d\mathscr{N}_t/d\mathscr{N}$.

The symbol $\Rightarrow$ denotes the vague convergence of expectations (i.e., the pointwise convergence on the set of all bounded continuous functions).

If H_n is an expectation on a σ-algebra $\mathbf{X}_n$, g_n an $\langle \mathbf{X}_n, \mathbf{B}_k\rangle$ measurable transformation for each $n=1,2,\ldots$, and if $c\in R^k$, then (i) we write $g_n\to c$ in $\langle H_n\rangle$-prob. if $H_n \chi_{\{\|g_n-c\|>\varepsilon\}}\to 0$ for every $\varepsilon>0$ and (ii) we say that $\langle g_n\rangle$ is bounded in $\langle H_n\rangle$-prob. if $\langle H_n^{g_n}\rangle$ is tight.

2. Locally Asymptotically Normal Families of Expectations and Lower Bounds

Throughout the paper we suppose the following assumption holds:

1. Assumption. k is a positive integer, for each $n=1,2,\ldots$, $\Theta_n\subset R^k$, $\mathbf{X}_n$ is a σ-algebra, $E_{n,\delta}$ an expectation on $\mathbf{X}_n$ for each δ in Θ_n.

2. Definition. We shall say that Condition LAN $\langle \theta, M_n, \gamma_n\rangle$ holds if, for each $n=1,2,\ldots$, θ is in Θ_n, M_n is a $k\times k$ positive definite matrix, γ_n a k-dimensional random vector on $\mathbf{X}_n$ such that

$$E_{n,\theta}^{\gamma_n}\Rightarrow \mathscr{N} \tag{1}$$

and if, for each bounded sequence $\langle t_n\rangle$ in R^k, $\delta_n=\theta+M_n^{-1/2}t_n$ is eventually in Θ_n, and

$$g_n\in dE_{n,\delta_n}/dE_{n,\theta} \tag{2}$$

implies

$$g_n/\rho_{t_n}(\gamma_n)\to 1 \quad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{3}$$

We shall say that Condition HLAN $\langle\theta, M_n, \gamma_n\rangle$ holds if the above conditions hold but with $\langle t_n\rangle$ restricted to constant sequences $\langle t\rangle$.

3. *Definition.* l is called a *loss function on* R^s if l is in $\mathscr{B}_{s+}$, $l(x)=l(-x)$ for each $x\in R^s$ and if $\{x;\, l(x)\leqq u\}$ is convex for each $u\in(0,\infty)$.

4. *Remark.* Hájek (1972) considers Condition HLAN $\langle\theta, nM, \gamma_n\rangle$ with a positive definite matrix M and with Θ_n independent of n. Le Cam (1960, 1969) considers Condition LAN $\langle\theta, nM, \gamma_n\rangle$. Hájek (1972) gives conditions under which HLAN $\langle\theta, nM, \gamma_n\rangle$ holds in the case of independent and identically distributed random variables. Ibragimov and Has'minskij (1975, 1979) give a sufficient condition for Condition HLAN $\langle\theta, M_n, \gamma_n\rangle$ for the case of non-identically distributed random variables. Fabian and Hannan (1980) show that a weaker condition is still sufficient for Condition LAN $\langle\theta, M_n, \gamma_n\rangle$, and give additional references to results on sufficient conditions for LAN.

Theorem 5 below is the lower bound result of Hájek (1972, Theorem 4.1). Hájek gives the proof only for case $k=1$. Theorem 6 gives a strengthening of Theorem 5 and will be proved in Sect. 5. Ibragimov and Has'minskij (1979, Sect. II.12) prove the case (ii) of Theorem 6 with additional assumptions on l and, implicitly, with an additional assumption of the measurability of $\delta\rightsquigarrow E_{n,\delta}\, l(M_n^{1/2}(Z_n-\delta))$. Le Cam (1979, p. 134) obtains a similar result.

Let e_i denote the vector $\langle 0,\dots,0,1,0,\dots,0\rangle$ with the i-th coordinate 1.

5. Theorem (Hájek). *Let Condition HLAN $\langle\theta, nM, \gamma_n\rangle$ hold, let $\langle Z_n\rangle$ be a sequence of estimates, l a loss function on R. Then*

$$\lim_{\varepsilon\to 0}\liminf_{n\to\infty}\sup_{\|\delta-\theta\|\leqq\varepsilon} E_{n,\delta}\, l\left(\sqrt{n}\,\frac{e_1'(Z_n-\delta)}{\sqrt{e_1' M^{-1} e_1}}\right)\geqq \mathscr{N} l. \tag{1}$$

6. Theorem. *Let $\langle Z_n\rangle$ be a sequence of estimates, l a loss function on R^k, $\langle Q_n\rangle$ a sequence in $\mathscr{C}_k$. Let either* (i) *Condition LAN $\langle\theta, M_n, \gamma_n\rangle$ hold or* (ii) *Condition HLAN $\langle\theta, M_n, \gamma_n\rangle$ hold and $Q_n=\mathbf{1}$ for all n.*

Then

$$\lim_{K\to\infty}\liminf_{n}\sup_{\|M_n^{1/2}(\delta-\theta)\|\leqq K} E_{n,\delta}\, l(Q_n M_n^{1/2}(Z_n-\delta))\geqq \mathscr{N} l. \tag{1}$$

7. *Remark.* The assertion in case (ii) may be unsatisfactory if M_n are not of the form nM. In such a case, e.g., it may be impossible to choose l such that $l(M_n^{1/2}(Z_n-\delta))=(Z_{n,1}-\delta_1)^2/c_n$ with norming constants c_n. In this sense, the assertion in case (ii) does not give a satisfactory generalization of Theorem 5.

This difficulty does not arise in case (i), as shown in the following Corollary.

8. Corollary. *Let $\langle Z_n\rangle$ be a sequence of estimates, l a loss function on R, $\langle a_n\rangle$ a sequence in $R^k-\{0\}$, let Condition LAN $\langle\theta, M_n, \gamma_n\rangle$ hold. Then*

$$\lim_{K\to\infty}\liminf_{n}\sup_{\|M_n^{1/2}(\delta-\theta)\|\leqq K} E_{n,\delta}\, l\left(\frac{a_n'(Z_n-\delta)}{\sqrt{a_n' M_n^{-1} a_n}}\right)\geqq \mathscr{N} l. \tag{1}$$

Proof. Obtains from (6.1), applied with the first row of Q_n equal to the transpose of $M_n^{-1/2} a_n(a_n' M_n^{-1} a_n)^{-1/2}$ and to $\bar{l}(x) = l(e_1' x)$.

3. The Non-Attainability of (2.5.1)

1. Definition. Let Condition HLAN $\langle \theta, nM, \gamma_n \rangle$ hold. Then we say that $\langle Z_n \rangle$ is HLAM (θ) (locally asymptotically minimax at θ in Hájek's sense) if $\langle Z_n \rangle$ is a sequence of estimates for which

$$\lim_{\varepsilon \to 0} \lim_{n \to \infty} \sup_{\|\delta - \theta\| \leqq \varepsilon} E_{n,\delta}\, l\left(\sqrt{n}\, \frac{e_i'(Z_n - \delta)}{\sqrt{e_i' M^{-1} e_i}}\right) = \mathcal{N} l \tag{1}$$

holds for every bounded loss function l on R and every i.

If Condition LAN $\langle \theta, M_n, \gamma_n \rangle$ holds then $\langle Z_n \rangle$ is LAM (θ) (locally asymptotically minimax at θ) if $\langle Z_n \rangle$ is a sequence of estimates for which

$$\lim_{K \to \infty} \lim_{n \to \infty} \sup_{\|M_n^{1/2}(\delta - \theta)\| \leqq K} E_{n,\delta}\, l(Q_n M_n^{1/2}(Z_n - \delta)) = \mathcal{N} l \tag{2}$$

holds for every sequence $\langle Q_n \rangle$ in $\mathscr{C}_k$ and for every bounded loss function l on R^k.

2. Remark. That Condition LAN $\langle \theta, nM, \gamma_n \rangle$ does not imply the existence of an HLAM (θ) sequence of estimates $\langle Z_n \rangle$ is almost obvious. Indeed, the LAN Condition does not require anything about the large neighborhood $\|\delta - \theta\| \leqq \varepsilon$ appearing in (1.1).

In the following example, for every sequence of estimates $\langle Z_n \rangle$ and any loss function l, the left hand side in (1.1) is equal to $\sup_x l(x)$. This is larger than $\mathcal{N} l$ if l is not a constant function. Consequently no HLAM sequence $\langle Z_n \rangle$ exists in the example.

Let $k=1$, $\Theta_n = R$, $\mathbf{X}_n = \mathbf{B}$, $\langle b_n \rangle$ be a sequence of positive numbers such that $b_n \to 0$, $n^{1/2} b_n \to \infty$. For $n = 1, 2, \ldots$, let E_n be an expectation on $\mathbf{B}$, let $E_{n,\delta}$ be the expectation induced by a normal $(\delta, 1/n)$ random variable if $\delta < b_n$, and let $E_{n,\delta} = E_n$ if $\delta \geqq b_n$.

It is easy to verify that Condition LAN $\langle 0, n, \gamma_n \rangle$ holds with $\gamma_n(x) = n^{1/2} x$. Indeed $E_{n,0}^{\gamma_n} = \mathcal{N}$ so that (2.2.1) holds, and if $\delta_n = n^{-1/2} t_n$ with $\langle t_n \rangle$ bounded, then, eventually, $\delta_n < b_n$ and any g_n in $dE_{n,\delta_n}/dE_{n,0}$ satisfies $g_n = \rho_{t_n}(\gamma_n)$ a.e. $(E_{n,0})$.

Let $\langle Z_n \rangle$ and l be as above, P_n the probability distribution of $\sqrt{n}\, Z_n$ under E_n.

Notice that if P is a probability on $\mathbf{B}$, $[a, b] \subset R$ and $0 < c$, then there is a $t \in [a, b]$ such that $P(t - c, t + c) \leqq \frac{2c}{b - a}$; indeed, there are at least $\frac{b - a}{2c}$ such intervals which are disjoint.

Let $c > 0$. For $\delta \in [b_n, \varepsilon]$, $E_{n,\delta} = E_n$,

$$E_n\, l(\sqrt{n}(Z_n - \delta)) \geqq l(c)[1 - P_n(-c + \sqrt{n}\,\delta, c + \sqrt{n}\,\delta)]. \tag{1}$$

By the preceding remark, there is δ in $[b_n, \varepsilon]$ (alternatively, $\sqrt{n}\,\delta$ in $[\sqrt{n}\,b_n, \sqrt{n}\,\varepsilon]$) for which the right hand side in (1) is at least

$$l(c)\left[1-\frac{2c}{\sqrt{n}(\varepsilon-b_n)}\right]. \tag{2}$$

Thus the left-hand side in (1.1) is at least $l(c)$, for every $c \in (0, \infty)$.

3. *Remark.* It is possible to construct a slightly different example in which Condition LAN holds at *every* point and yet, for a θ, no HLAM (θ) sequence of estimates exists.

4. Preparatory Results

1. *Definition.* A sequence $\langle E_n, F_n\rangle$ will be called *quasinormal* $\langle t_n, \gamma_n\rangle$ if E_n, F_n are expectations on a σ-algebra $\mathbf{X}_n$, γ_n is a k-dimensional random vector on $\mathbf{X}_n$ for $n=1, 2, \ldots$, if $\langle t_n\rangle$ is a bounded sequence in R^k, if

$$E_n^{\gamma_n} \Rightarrow \mathcal{N}, \tag{1}$$

and if

$$f_n \in dF_n/dE_n \tag{2}$$

implies

$$f_n/\rho_{t_n}(\gamma_n) \to 1 \quad \text{in } \langle E_n\rangle\text{-prob.} \tag{3}$$

2. Lemma. *Let $\langle E_n, F_n\rangle$ be quasinormal $\langle t_n, \gamma_n\rangle$. Then*

$$\langle E_n\rangle, \langle F_n\rangle \quad \textit{are contiguous} \tag{1}$$

and

$$F_n^{\gamma_n - t_n} \Rightarrow \mathcal{N}. \tag{2}$$

Proof. Without loss of generality, assume $t_n \to t$. Let (1.2) hold and $\chi_n = \log f_n$ a.e. (E_n) on $\{0 < f_n\}$. The Slutsky theorem gives

$$E_n^{\chi_n} \Rightarrow \mathcal{N}^{\log \rho_t} \tag{3}$$

and

$$E_n^{\langle \chi_n, \gamma_n - t_n\rangle} \Rightarrow \mathcal{N}^{\langle \log \rho_t, \iota - t\rangle} \tag{4}$$

with ι the identity function on R^k.

We apply now Theorem 2.1 in Le Cam (1960) with its six conditions $\mathcal{L}1$ to $\mathcal{L}6$. By (3) and since $\mathcal{N}\rho_t = \mathcal{N}_t 1 = 1$, $\mathcal{L}5$ and thus (1) hold. Since (4) holds, $\mathcal{L}6$ implies $F_n^{\langle \chi_n, \gamma_n - t_n\rangle} \Rightarrow \mathcal{N}_t^{\langle \log \rho_t, \iota - t\rangle}$ which in turn implies (2).

3. Lemma. *Let $\langle u_n\rangle$, $\langle v_n\rangle$ be bounded sequences in R^k, let Condition LAN $\langle \theta, M_n, \gamma_n\rangle$ hold and let*

$$E_n = E_{n, \theta + M_n^{-1/2} u_n}, \qquad F_n = E_{n, \theta + M_n^{-1/2} v_n}.$$

Then $\langle E_n, F_n \rangle$ *is quasinormal* $\langle v_n - u_n, \gamma_n - u_n \rangle$.

Proof. From the definition of Condition LAN, we obtain that $\langle E_{n,\theta}, E_n \rangle$ is quasinormal $\langle u_n, \gamma_n \rangle$. By (2.2),

$$E_n^{\gamma_n - u_n} \Rightarrow \mathcal{N} \tag{1}$$

and by (2.1), $\langle E_{n,\theta} \rangle$ and $\langle E_n \rangle$ are contiguous.

Let $e_n \in dE_n/dE_{n,\theta}$, $f_n \in dF_n/dE_{n,\theta}$. Since $(\rho_{v_n}/\rho_{u_n})(\gamma_n) = \rho_{v_n - u_n}(\gamma_n - u_n)$, we obtain

$$(f_n/e_n)/\rho_{v_n - u_n}(\gamma_n - u_n) \to 1 \quad \text{in } \langle E_n \rangle\text{-prob.} \tag{2}$$

If $g_n \in dF_n/dE_n$ then $g_n \neq f_n/e_n$ on $A_n \cup B_n$ with $E_{n,\theta}\chi_{A_n} + E_n \chi_{B_n} = 0$. By the contiguity, $E_n \chi_{A_n \cup B_n} \to 0$ and (2) gives

$$g_n/\rho_{v_n - u_n}(\gamma_n - u_n) \to 1 \quad \text{in } \langle E_n \rangle\text{-prob.} \tag{3}$$

This and (1) imply the assertion of the lemma.

5. The New Lower Bound

1. Remark. The proof below is close to that of Ibragimov and Has'minskij (1979), except that we avoid integration which would require measurability properties we do not assume. The slightly more general loss l and the inclusion of Q_n in the LAN case do not cause any problem.

2. Proof of Theorem 2.6.

Write $\delta = \theta + M_n^{-1/2} Q_n' t$ in (2.6.1) and denote $Q_n M_n^{1/2}(Z_n - \theta)$ by $\tilde{Z}_n$, $E_{n,\delta}$ by $\tilde{E}_{n,t}$. It is straightforward that $\langle \tilde{E}_{n,t} \rangle$ satisfy Condition LAN $\langle 0, \mathbf{1}, Q_n \gamma_n \rangle$ in case (i) and Condition HLAN $\langle 0, \mathbf{1}, \gamma_n \rangle$ in case (ii). The asymptotic normality of $\tilde{E}_{n,0}^{Q_n \gamma_n}$ follows easily by an application of the Slutsky theorem to convergent subsequences of $\langle Q_n \rangle$.

Since $Q_n M_n^{1/2}(Z_n - \delta) = \tilde{Z}_n - t$, it is enough to prove (2.6.1) with $E_{n,\delta}$, Z_n, M_n, Q_n replaced by $\tilde{E}_{n,t}$, $\tilde{Z}_n$, $\mathbf{1}$, $\mathbf{1}$. Simpler yet, it is enough to prove (2.6.1) for case (ii) with $\theta = 0$, $M_n = \mathbf{1}$.

For $x \in R^k$, denote $\max|x_i|$ by $|x|$. For $M > 0$, $q \in \{1, 2, \dots\}$, denote the cube $\{x; x \in R^k, |x| \leqq M\}$ by C_M, its indicator function by χ_M and the grid $\{x; x \in C_M$, the qx_i are integers$\}$ by $C_{M,q}$.

Assume K, J are positive numbers, $J < K$, denote by $\mathcal{U}_q$ and $\mathcal{U}$ the expectations with respect to the uniform distributions on $C_{K,q}$ and C_K.

Relation (2.6.1) will be proved when we have proved

$$\liminf_q \liminf_n \mathcal{U}_q E_{n\delta}\, l(Z_n - \delta) \geqq \left(\frac{J}{K}\right)^k \mathcal{N}(\chi_{K-J}\, l). \tag{1}$$

Since $l \geqq h_m = m^{-1} \sum_{i=1}^{m^2} \chi_{\{ml \geqq i\}}$ and $\mathcal{N}(\chi_{K-J}\, h_m) \to \mathcal{N}(\chi_{K-J}\, l)$, it is enough to prove (1) assuming $l = 1 - \chi_A$ with A a convex Borel subset of R^k, $A = \{-x; x \in A\}$.

We shall call a function f on R^k quasiconcave if $\{f \geqq c\}$ is a convex set for every c in R. For an $a \in (0, \infty)$, let $\mathscr{A}$ be the family of all quasi-concave functions on R^k into $[0, a]$.

We shall show that, as $q \to \infty$,

$$\mathscr{U}_q \to \mathscr{U} \text{ uniformly on } \mathscr{A}. \tag{2}$$

That (2) holds with $\mathscr{A}$ replaced by the family of all indicator functions of convex sets follows from results on uniformity of vague convergence; see, e.g. Theorem 4.2 in Ranga Rao (1962) or Theorem 4.1 in Fabian (1970). Hence, (2) follows by the representation $f = \int_0^a \chi_{\{f \geqq c\}} dc$ for f in $\mathscr{A}$.

For b in R^k, *let* φ_b, l_b denote the translates of φ, the standard normal density on R^k, and of l by $-b$; e.g., $\varphi_b(x) = \varphi(x-b)$. With $a = (2\pi)^{-k/2}$, $(1-l_z)\varphi_\gamma$ and φ_γ are in $\mathscr{A}$ for every γ, z in R^k and, by (2),

$$\mathscr{U}_q \varphi_\gamma l_z \to \mathscr{U} \varphi_\gamma l_z \quad \text{uniformly in } \gamma, z. \tag{3}$$

But $(2K)^k\ \mathscr{U}\varphi_\gamma l_z = \mathscr{N}_\gamma \chi_K l_z \geqq \mathscr{N} \chi_{K-|\gamma|} l_{z-\gamma}$ (interpret χ_M as 0 for $M \leqq 0$). By Theorem 1 in Anderson (1955), the last term above is at least $\mathscr{N}\chi_{K-|\gamma|} l$ and thus, for all $z \in R^k$,

$$\mathscr{U}\varphi_\gamma l_z \geqq (2K)^{-k} \mathscr{N} \chi_{K-|\gamma|} l. \tag{4}$$

Write a pseudodensity of $E_{n,\delta}$ with respect to $E_{n,0}$ as $\rho_\delta(\gamma_n)(1+\varepsilon_{n,\delta})$. On C_J, $\rho_\delta \leqq \exp(kJ^2/2)$ for every δ, and thus

$$E_{n,\delta}\, l(Z_n - \delta) \geqq E_{n,0}\, \chi_J(\gamma_n)\, \rho_\delta(\gamma_n)\, l(Z_n-\delta) - e^{kJ^2/2} E_{n,0}(\varepsilon_{n\delta})_- . \tag{5}$$

Since the pseudodensities are non-negative a.e. with respect to $E_{n,0}$, we obtain $(\varepsilon_{n\delta})_- \leqq 1$ and, by Condition HLAN, the $\mathscr{U}_q$ expectation of the last term in (5) goes to 0 as $n \to \infty$. The $\mathscr{U}_q$ expectation of the second term in (5) is

$$E_{n,0} \frac{\chi_J(\gamma_n)}{\varphi(\gamma_n)} \mathscr{U}_q \varphi_{\gamma_n} l_{Z_n}. \tag{6}$$

By (3) and (4), if $\eta \in (0,1)$ then for q large enough, on $\{\gamma_n \in C_J\}$,

$$\mathscr{U}_q \varphi_{\gamma_n} l_{Z_n} \geqq \eta (2K)^{-k} \mathscr{N} \chi_{K-J} l. \tag{7}$$

By Condition HLAN, $E_{n,0}^{\gamma_n} \Rightarrow \mathscr{N}$ and $E_{n,0} \dfrac{\chi_J(\gamma_n)}{\varphi(\gamma_n)} \to (2J)^k$. Thus the $\liminf_q \liminf_n$ of the second term in (5) is at least the right hand side of (1), (1) holds and the proof is completed.

6. Locally Asymptotically Minimax Estimates

The following assumption will be assumed in §§ 6.2 to 6.16.

1. Assumption. Condition LAN $\langle \theta, \mathrm{M}_n, \gamma_n \rangle$ holds.

2. Definition. A sequence $\langle Z_n \rangle$ of estimates is called *regular* (θ) if

$$M_n^{1/2}(Z_n-\theta)-\gamma_n\to 0 \quad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{1}$$

3. Theorem. *Let $\langle Z_n\rangle$ be a sequence of estimates. Then the regularity (θ) of $\langle Z_n\rangle$ implies*

$$E_{n,\delta_n}^{M_n^{1/2}(Z_n-\delta_n)}\Rightarrow \mathcal{N} \tag{1}$$

for every sequence $\delta_n=\theta+M_n^{-1/2}t_n$ such that $\langle t_n\rangle$ is bounded; the latter property, in turn, implies that $\langle Z_n\rangle$ is LAM (θ).

Proof. Let $\langle t_n\rangle$ be a bounded sequence, $\tilde{Z}_n=M_n^{1/2}(Z_n-\theta)$, $\delta_n=\theta+M_n^{-1/2}t_n$, so that $M_n^{1/2}(Z_n-\delta_n)=\tilde{Z}_n-t_n$.

Let (2.1) hold, i.e., $\tilde{Z}_n-\gamma_n\to 0$ in $\langle E_{n,\theta}\rangle$-prob. By Lemmas 4.3 and 4.2, $E_{n,\delta_n}^{\gamma_n-t_n}\Rightarrow\mathcal{N}$ and $\langle E_{n,\delta_n}\rangle$ is contiguous to $\langle E_{n,\theta}\rangle$, which implies (1).

Let (1) hold, let $\langle Q_n\rangle$ be a sequence in $\mathscr{C}_k$. From (1) we obtain

$$E_{n,\delta_n}^{Q_n(\tilde{Z}_n-t_n)}\Rightarrow\mathcal{N} \tag{2}$$

by an application of the Slutsky theorem to subsequences for which $\langle Q_{n_i}\rangle$ converges. From (2), if l is a bounded loss function,

$$E_{n,\delta_n}\, l(Q_n(\tilde{Z}_n-t_n))\to\mathcal{N}\, l, \tag{3}$$

since the discontinuity set of l is covered by the boundaries of the convex sets $\{x;\, l(x)\leqq y\}$ with y rational and is therefore $\mathcal{N}$-null (cf. Eggleston, 1958, proof of Theorem 35). (3) implies (3.1.2) and therefore the last assertion of the theorem.

4. Remark. According to Theorem 4.1 in Hájek (1972), for k = 1 the regularity (θ) of $\langle Z_n\rangle$ is necessary, under the HLAN $\langle\theta, nM, \gamma_n\rangle$ Condition, for the HLAM (θ) property of $\langle Z_n\rangle$; we have proved it sufficient for the LAM (θ) property of $\langle Z_n\rangle$ under the LAN $\langle\theta, M_n, \gamma_n\rangle$ Condition.

It follows from Theorem 3 that $\langle\theta+M_n^{-1/2}\gamma_n\rangle$ is an LAM (θ) sequence of estimates. This is non-trivial since $\langle\theta\rangle$ is not an LAM (θ) sequence of estimates. Next we shall study estimates which do not depend on θ.

5. Definition. A sequence $\langle U_n\rangle$ will be called an *auxiliary estimate* if $\langle U_n\rangle$ is a sequence of estimates and if $\langle\|M_n^{1/2}(U_n-\theta)\|\rangle$ is bounded in $\langle E_{n,\theta}\rangle$-prob.

A sequence $\langle m_n\rangle$ will be called a *rate estimate* if each m_n is an $\mathbf{X}_n$-measurable function with values $k\times k$ positive definite matrices and if $\langle\|M_n^{-1}m_n\|+\|m_n^{-1}M_n\|\rangle$ is bounded in $\langle E_{n,\theta}\rangle$-prob.

A sequence $\langle W_n\rangle$ will be called a *consistent estimate of* $\langle M_n\rangle$ if W_n are positive definite matrix valued random variables on $\mathbf{X}_n$ and if

$$M_n^{-1/2}W_nM_n^{-1/2}\to\mathbf{1} \quad \text{in } \langle E_{n\theta}\rangle\text{-prob.} \tag{1}$$

6. Definition. A sequence $\langle U_n\rangle$ of $\langle\mathbf{X}_n,\mathbf{Y}_n\rangle$-measurable transformations is called *discrete* if for every positive number ε there is an integer q and sets C_n in $\mathbf{X}_n$ such that $\liminf_n P_{n,\theta}C_n\geqq 1-\varepsilon$ and such that for each n, $U_n[C_n]$ has at most q elements.

7. Remark. Discrete auxiliary estimates, discrete rate estimates and consistent estimates of $\langle M_n \rangle$ will be used to construct regular and LAM estimates.

Of course, to be practically useful, these estimates should not depend on θ, which can be formalized by asking that they have the required properties for every θ in Θ, not only for one point θ; see more in Remark 17.

Auxiliary estimates are easy to obtain in many practical situations and we shall not concern ourselves with their construction here (see, e.g., Le Cam, 1956).

Similarly, in many situations, rate estimates are easy to obtain. For example, in the i.i.d. case $M_n = nM$ and $\langle n\mathbf{1} \rangle$ is a rate estimate (cf. Remark 11).

If Condition LAN $\langle \delta, \overline{M}_n(\delta), \gamma_{n,\delta} \rangle$ is satisfied at each δ, $\overline{M}_n(\theta) = M_n$, $\overline{M}_n$ are $\mathbf{B}_k$-measurable, and $\langle U_n \rangle$ is an auxiliary estimate then $\langle W_n \rangle = \langle \overline{M}_n(U_n) \rangle$ is a consistent estimate of $\langle M_n \rangle$, provided that $M_n^{-1/2} \overline{M}_n(\delta_n) M_n^{-1/2} \to \mathbf{1}$ if $M_n^{1/2}(\delta_n - \theta)$ is bounded. If $\langle U_n \rangle$ is discrete then so is $\langle W_n \rangle$.

8. Remark. Discretization of auxiliary estimates was suggested and used by Le Cam (1960, Appendix 1 and 1969, Theorem 4, Chap. III) to extend convergence (4.1.3) to the case of t_n replaced by $M_n^{1/2}(U_n - \theta)$ with $\langle U_n \rangle$ an auxiliary estimate.

We shall use the discreteness in Lemma 12. If condition LAN holds in a strengthened form (as in Proposition 3 in Le Cam, 1970), the assumption of discreteness can be omitted in Lemma 12 and thereafter.

9. Assumption. Let $g_{n,\delta_1,\delta_2} \in dE_{n,\delta_1}/dE_{n,\delta_2}$ for all n and all δ_1, δ_2 in Θ_n. Let $\langle U_n \rangle$ be a discrete auxiliary sequence of estimates, $\langle m_n \rangle$ a discrete rate estimate.

10. Notation. If Assumption 9 holds, the following notation will be used for $a \in R^k$:

$$\lambda_{n,a} = \log g_{n,\, U_n + m_n^{-1/2} a,\, U_n} \tag{1}$$

if the argument of the log in (1) is defined and in $(0, \infty)$; set $\lambda_{n,a} = 0$ for the other cases.

Finally, with $e_j \in R^k$, $(e_j)_i = \chi_{\{j\}}(i)$, set

$$\lambda_n = \langle \lambda_{n,e_1} - \lambda_{n,-e_1}, \ldots, \lambda_{n,e_k} - \lambda_{n,-e_k} \rangle. \tag{2}$$

11. Remark. It may be useful to consider now the special i.i.d. case in which $\Theta_n = \Theta$, E_δ are expectations on a common σ-algebra $\mathbf{\Omega}$, $Y_1, Y_2, \ldots$ are independent and identically distributed generalized random variables (with values in a measurable space) under each E_δ. Let $\mathbf{X}_n$ denote the σ-algebra generated by $\langle Y_1, \ldots, Y_n \rangle$ and $E_{n,\delta}$ the restriction of E_δ to $\mathbf{X}_n$-measurable functions.

Suppose that for an integral J, each $E_\delta^{Y_1}$ has a density f_δ with respect to J, that the function $\delta \rightsquigarrow f_\delta^{1/2}$ has a derivative q in $L_2(J)$ at θ and that $M = 4Jqq'$, is non-singular. Then Condition LAN (θ, nM, γ_n) holds for a sequence $\langle \gamma_n \rangle$ (due to Le Cam, 1970, Lemma 4; cf. also Fabian and Hannan, 1980, Theorem 4.8).

Write $l_{n,\delta} = \frac{1}{n} \sum_{i=1}^{n} \log f_\delta(Y_i)$ (set $l_{n,\delta} = 0$ on the set where $\prod_{i=1}^{n} f_\delta(Y_i)$ is not in $(0, \infty)$).

If $c \in (0, \infty)$ then $\langle cn\mathbf{1}\rangle$ is a discrete rate estimate,

$$\lambda_{n,a} = n[l_{n,U_n+(cn)^{-1/2}a} - l_{n,U_n}], \tag{1}$$

and the i-th coordinate of λ_n is

$$n[l_{n,U_n+(cn)^{-1/2}e_i} - l_{n,U_n-(cn)^{-1/2}e_i}]. \tag{2}$$

If, additionally, the matrices $\ddot{f}_\delta$, $(\log \ddot{f}_\delta)$ of second order partial derivatives exist and are continuous in a neighborhood of θ, and dominated, respectively, by a function in $L_1(J)$ and a function in $L_1(E_\theta)$, then as a well known and easy result we obtain

$$\ddot{l}_{n,U_n} \to -M \quad \text{in } E_\theta\text{-prob.} \tag{3}$$

12. Lemma. *If Assumption 9 holds and*

$$a \in R^k, \qquad Q_n = m_n^{-1/2} M_n^{1/2}, \qquad \tilde{U}_n = M_n^{1/2}(U_n - \theta), \tag{1}$$

then

$$\lambda_{n,a} - a' Q_n(\gamma_n - \tilde{U}_n) + \tfrac{1}{2}\|Q_n' a\|^2 \to 0 \quad \text{in } \langle E_{n\theta}\rangle\text{-prob.} \tag{2}$$

Proof. By Lemma 4.3

$$g_{n,\theta+M_n^{-1/2}v_n,\theta+M_n^{-1/2}u_n} = \rho_{v_n-u_n}(\gamma_n - u_n)\, k_n(u_n, v_n) \tag{4}$$

with

$$k_n(u_n, v_n) \to 1 \quad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{5}$$

for all bounded sequences $\langle u_n\rangle$, $\langle v_n\rangle$. This holds (cf. Remark 8) even with $\tilde{U}_n$, $\tilde{V}_n$ substituted for u_n, v_n, where

$$V_n = U_n + m_n^{-1/2} a, \qquad \tilde{V}_n = M_n^{1/2}(V_n - \theta) \tag{6}$$

so that

$$\tilde{V}_n = \tilde{U}_n + Q_n' a. \tag{7}$$

Take logarithm in (4) to obtain (2).

13. Remark. Consistent estimates of $\langle M_n\rangle$. There are various possibilities of choosing a consistent estimate $\langle W_n\rangle$ of $\langle M_n\rangle$.

Note first that if W_n are symmetric $k \times k$ matrices then (i) $M_n^{-1} W_n \to \mathbf{1}$ implies (ii) $M_n^{-1/2} W_n M_n^{-1/2} \to \mathbf{1}$. Indeed, the non-zero roots of $M_n^{-1} W_n$ and $M_n^{-1/2} W_n M_n^{-1/2}$ are the same, (i) implies that they all converge to 1, and for the symmetric matrices in (ii) this implies (ii) holds.

Consequently (5.1) is a weaker property than $M_n^{-1} W_n \to \mathbf{1}$ in $\langle E_{n,\theta}\rangle$-prob. for symmetric valued W_n.

Secondly, suppose W_n are symmetric valued and satisfy (5.1). Then a simple change of $\langle W_n\rangle$ to $\langle \overline{W}_n\rangle$ yields a consistent estimate of $\langle M_n\rangle$: Set $\overline{W}_n = W_n$ on the set where W_n is positive definite and $\overline{W}_n = \mathbf{1}$ on the complement. Indeed

$$\left|\frac{x' W_n x}{x' M_n x} - 1\right| \leqq \|M_n^{-1/2} W_n M_n^{-1/2} - \mathbf{1}\| \quad \text{for } x \neq 0$$

and the assertion follows easily.

In view of these remarks, we shall describe W_n symmetric valued but not necessarily positive definite, and satisfying $M_n^{-1} W_n \to \mathbf{1}$ in $\langle E_{n,\theta}\rangle$-prob.

In the i.i.d. case, if (11.3) holds, we can set $W_n = -n\ddot{l}_{n,U_n}$. Another possible choice of W_n is mentioned at the end of Remark 7.

Finally, we shall describe a choice of W_n which does not need any assumptions additional to Assumption 9. Define the $k \times k$-matrix Λ_n by

$$2\Lambda_{n,ij} = -\lambda_{n,e_i+e_j} - \lambda_{n,-e_i-e_j} + \lambda_{n,e_i} + \lambda_{n,-e_i} + \lambda_{n,e_j} + \lambda_{n,-e_j}; \tag{1}$$

then we may choose

$$W_n = m_n^{1/2} \Lambda_n m_n^{1/2}. \tag{2}$$

Indeed, by Lemma 12, with $A_n = Q_n Q_n'$

$$\Lambda_{n,ij} = \tfrac{1}{2}[(e_i+e_j)' A_n (e_i+e_j) - e_i' A_n e_i - e_j' A_n e_j] + \eta_{nij} = e_i' A_n e_j + \eta_{nij}$$

with $\eta_n \to 0$ in $\langle E_{n\theta}\rangle$-prob. Thus

$$M_n^{-1} W_n = M_n^{-1} m_n^{1/2} Q_n Q_n' m_n^{1/2} + M_n^{-1} m_n^{1/2} \eta_n m_n^{1/2}.$$

The first term on the right-hand side is $\mathbf{1}$, the second converges to 0 in $\langle E_{n,\theta}\rangle$ by properties of m_n and η_n.

14. Remark. Under similar but stronger conditions than Assumptions 1 and 9, Le Cam (1969, proof of Theorem 4, Chap. III) proves that an estimate, similar to the one described in our Theorem 15, is regular (θ) and thus, by our Theorem 3, LAM (θ).

Le Cam's proof is, however, inadequate for the theorem, unless an assumption is added in the theorem that Δ_n have the property described in Le Cam's Proposition 1.

15. Theorem. *Let Assumption 9 hold, let $\langle W_n\rangle$ be a consistent estimate of $\langle M_n\rangle$ and let*

$$Z_n = U_n + \tfrac{1}{2} W_n^{-1} m_n^{1/2} \lambda_n. \tag{1}$$

Then $\langle Z_n\rangle$ is a regular (θ) and LAM (θ) *sequence of estimates.*

Proof. Set $\tilde{Z}_n = M_n^{1/2}(Z_n - \theta)$, $V_n = M_n^{1/2} W_n^{-1} M_n^{1/2}$ so that $V_n \to \mathbf{1}$ in $\langle E_{n,\theta}\rangle$-prob. From Lemma 12, using its notation,

$$\tilde{Z}_n - \gamma_n = \tilde{U}_n - \gamma_n + \tfrac{1}{2} V_n Q_n^{-1} \lambda_n = (\mathbf{1} - V_n)(\tilde{U}_n - \gamma_n) + V_n Q_n^{-1} \varepsilon_n$$

with $\varepsilon_n \to 0$ in $\langle E_{n,\theta}\rangle$-prob. Because $\langle \tilde{U}_n\rangle$, $\langle \gamma_n\rangle$, $\langle Q_n^{-1}\rangle$ are bounded in $\langle E_{n,\theta}\rangle$-prob., we obtain that $\tilde{Z}_n - \gamma_n \to 0$ in $\langle E_{n,\theta}\rangle$-prob., the regularity (θ). Theorem 3 then implies the LAM (θ) property.

16. Remark. Under Assumption 9, Theorem 15 allows for any choice $\langle W_n\rangle$ of a consistent estimate of $\langle M_n\rangle$. But by Remark 13 there always exists such a $\langle W_n\rangle$ (cf. 13.2); with that choice, (15.1) becomes

$$Z_n = U_n + \tfrac{1}{2} m_n^{-1/2} \Lambda_n^{-1} \lambda_n \tag{1}$$

(replace Λ_n by m_n^{-1} on the set where it fails to be positive definite).

Consider the i.i.d. case (cf. Remark 11) with $k=1$. We can choose $m_n = cn$ with any $c \in (0, \infty)$ and, if no simpler choice for W_n is available than (13.2), we may use (1) with

$$\Lambda_n = \lambda_{n,1} + \lambda_{n,-1} - \tfrac{1}{2}\lambda_{n,2} - \tfrac{1}{2}\lambda_{n,-2}; \tag{2}$$

the $\lambda_{n,i}$ and λ_n are given by (11.1) and (11.2).

We see that the estimate Z_n is quite explicit and fully determined by U_n and the densities f_δ. In particular, a computer program may be written which would calculate Z_n given in (1), based on the values of Y_i, U_n and a subroutine evaluating f_δ.

17. Remark. Of interest are estimates which are LAM (θ) at every θ in Θ (let $\Theta = \Theta_n$ for all n). If Condition LAN $\langle \theta, M_n(\theta), \gamma_{n,\theta} \rangle$ is satisfied for every θ in Θ and if $\langle m_n \rangle$, $\langle U_n \rangle$ are discrete rate and auxiliary sequences for every θ in Θ then such $\langle Z_n \rangle$ are described in Theorem 15 and Remark 16. So if the LAN $\langle \theta, M_n(\theta), \gamma_{n,\theta} \rangle$ holds for every θ, under mild conditions there are what may be called global locally asymptotically minimax estimates.

7. On Adaptive Estimates

1. Introduction. Stein (1956) discusses conditions under which so called adaptive estimation may be possible. The discussion is heuristic in some parts. Related to Stein's paper are results on adaptive estimators of a median of a symmetric density (see Stone (1975) for a rather general estimate of this type and for bibliography) or of other parameters (e.g. Weiss and Wolfowitz (1970a, b)). Pfanzagl (1976) showed non-existence of adaptive estimators of a quantile in case of asymmetric densities.

We shall reformulate some of Stein's (1956) results in terms of LAM estimates. This is rather easy since Stein's arguments concerned mostly properties of the Fisher information, with which the LAM criterion (in contrast with the classical theory) is solidly tied.

As Stein did, we base our considerations on subproblems of the original problem (cf. Definition 3), and define LAM adaptivity in Definition 6. Theorem 9 then gives a necessary condition for the existence of an LAM adaptive estimate and Theorem 10 gives a sufficient condition for an estimate to be LAM adaptive. (As mentioned in Sect. 1, Fabian (1980) obtains from Theorem 10 the LAM adaptivity of a recursive estimate of the location parameter with an unknown symmetric density.) Example 12 reformulates a result due to Pfanzagl (1976) which shows that, in estimating a quantile of an unknown density, and if the class of the densities is rich enough then no translation invariant, asymptotically uniformly median unbiased sequence of estimates is asymptotically more concentrated around the quantile than the sample quantile. In our reformulation, with the LAM risk used, the assertion is true for every sequence of estimates.

In their studies of non-parametric problems, Levit (1974, 1975), Koshevnik and Levit (1976), Has'minskij and Ibragimov (1978) also combine the LAM concept and Stein's approach, but they use larger neighborhoods than ours, arriving that way at a stronger LAM property. They prove their LAM property for some estimates. However, at our level of generality, it is impossible to obtain our results with their LAM property. Indeed, specialized to an LAN situation, their LAM becomes the HLAM property and it follows from our Sect. 3 that there are cases in which no HLAM estimates exist.

The mathematics of the results below is very simple and we treat the questions with some restrictions to retain the simplicity. In particular, we treat families satisfying the LAN $\langle\theta, M_n, \gamma_n\rangle$ condition with $M_n = nM$ and we do not include parameters which make the problem more difficult in the sense of Stein's remark following his Lemma in Sect. 3.

We consider (local) adaptivity at a point θ, because this simplifies the treatment; statements about global attainability follow immediately (see also Remark 11).

In this section, a sequence $\langle Z_n\rangle$ of estimates will mean m-dimensional Z_n, unless specified otherwise. Subscripts will be used to indicate coordinates or elements of matrices. If M is a $k_1 \times k_2$ matrix with $k_i \geqq m$ we shall partition M as $\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22}\end{bmatrix}$ with M_{11} an $m\times m$ matrix and we shall denote

$$M_{1.} = [M_{11}, M_{12}], M_{.1} = \begin{bmatrix} M_{11} \\ M_{21}\end{bmatrix}.$$

In the subsequent assumptions and definitions the motivation is to estimate θ_1, a component of the unknown parameter $\langle\theta_1, \theta_2\rangle$. Roughly speaking, a sequence of estimates is adaptive if it is asymptotically as good as a sequence of estimates in case the nuisance parameter θ_2 is known. Various possible definitions of adaptivity are discussed in Remark 5.

We shall consider subproblems obtainable by restricting the nuisance parameter to a subset which can be reparametrized as a subset of R^{k-m}. The nuisance parameter δ_2 in the subset is then labeled $\beta(\delta_2)$, and the parameter $\langle\delta_1, \delta_2\rangle$ is labeled $\alpha\langle\delta_1, \delta_2\rangle = \langle\delta_1, \beta(\delta_2)\rangle$.

2. *Assumption.* Assumption 2.1 holds with $\Theta_n = \Theta$ independent of n, but not necessarily a subset of R^k. The set Θ satisfies $\Theta = \Theta_1 \times \Theta_2$ for some sets Θ_1, Θ_2 with $\Theta_1 \subset R^m$ for an integer m. $\theta = \langle\theta_1, \theta_2\rangle$ is a point in Θ.

3. *Definition.* If k is an integer, $k \geqq m$ then by a *k-dimensional subproblem* we mean a pair $\langle\Theta_0, \alpha\rangle$ with the following properties.

Θ_0 is a subset of Θ, containing $D = \Theta_1 \times \{\theta_2\}$, α is a one-to-one function on Θ_0 into R^k. If $k = m$ then $\Theta_0 = D$ and $\alpha = (\pi_1)_D$, the restriction to D of π_1. For $i = 1, 2$, the function π_i is $\langle\delta_1, \delta_2\rangle \in \Theta_1 \times \Theta_2 \rightsquigarrow \delta_i$. If $k > m$, then $\alpha = \langle(\pi_1)_{\Theta_0}, \beta\circ(\pi_2)_{\Theta_0}\rangle$ with β a function on $\pi_2[\Theta_0]$ to R^{k-m}.

A *subproblem* is a k-dimensional subproblem for some k. The m-dimensional subproblem is called *singular*.

A subproblem $\langle\Theta_0, \alpha\rangle$ is LAN $\langle nM, \gamma_n\rangle$ if, with $\tilde{E}_{n,\tilde{\delta}}=E_{n,\alpha^{-1}(\tilde{\delta})}$, the family $\langle\tilde{E}_{n,\tilde{\delta}}; \tilde{\delta}\in\alpha[\Theta_0]\rangle$ satisfies Condition LAN $\langle\alpha(\theta), nM, \gamma_n\rangle$. A subproblem is LAN if it is LAN $\langle nM, \gamma_n\rangle$ for some M, γ_n.

The following is assumed for the rest of this section.

4. *Assumption.* Assumption 2 holds, the singular subproblem is LAN $\langle nM, \gamma_n\rangle$ and $\mathscr{A}$ is a family of LAN subproblems.

5. *Remark.* The *adaptivity property.* A sequence $\langle Z_n\rangle$ of estimates is LAM for the singular problem, if (cf. Definition 3.1)

$$\lim_{K\to\infty}\lim_{n\to\infty}\sup_{\|\delta_1-\theta_1\|\leq K/\sqrt{n}} E_{n,\langle\delta_1,\theta_2\rangle}\, l(Q_n(nM)^{1/2}(Z_n-\delta_1))=\mathcal{N}l \tag{1}$$

for every sequence $\langle Q_n\rangle$ in $\mathscr{C}_m$ and every bounded loss function l on R^m. Call the property (1) of $\langle Z_n\rangle$, property $\mathscr{P}_2(\theta)$. This property can be weakened to $\mathscr{P}_1(\theta)$, the classical asymptotic efficiency, by changing δ_1 to θ_1 and omitting $\lim_{K\to\infty}$ and sup in (1). Property $\mathscr{P}_2(\theta)$ can be strengthened to property $\mathscr{P}_3(\theta)$ by enlarging the neighborhood over which the supremum is taken to include points $\langle\delta_1, \delta_2\rangle$ with $\delta_2\neq\theta_2$. This will be done in Definition 6 below.

If a sequence of estimates has property $\mathscr{P}_i(\theta)$ for all $\theta\in\Theta$, we say it has property $\mathscr{P}_i$.

The adaptivity considered by Stone and others (ref. above) was $\mathscr{P}_1$. A strengthening of that property is $\mathscr{P}_2$, but we shall consider the strongest of these properties, $\mathscr{P}_3$.

Properties $\mathscr{P}_1$ and $\mathscr{P}_2$ are of interest because it is not easy to construct estimates with property $\mathscr{P}_1(\theta)$ (or $\mathscr{P}_2(\theta)$) for all θ; there is no problem if $\mathscr{P}_1(\theta)$ or $\mathscr{P}_2(\theta)$ are required for only one θ. In contrast, even the local property $\mathscr{P}_3(\theta)$ is of interest, since (cf. (6.1)) it reflects on the behavior of the estimate with the nuisance parameter close, rather than equal to θ_2. In this sense $\mathscr{P}_3(\theta)$ is an *asymptotic robustness* property at the point θ.

We shall study property $\mathscr{P}_3$ by studying $\mathscr{P}_3(\theta)$ (cf. Remark 11).

6. *Definition.* A sequence $\langle Z_n\rangle$ of estimates is locally asymptotically minimax $\mathscr{A}$-adaptive at θ (LAMA $(\mathscr{A},\theta)$) if, for any subproblem $\langle\Theta_0,\alpha\rangle$ in $\mathscr{A}$,

$$\lim_{K\to\infty}\lim_{n\to\infty}\sup_{\substack{\delta\in\Theta_0\\ \|\alpha(\delta)-\alpha(\theta)\|\leq K/\sqrt{n}}} E_{n,\delta}\, l(Q_n(nM)^{1/2}(Z_n-\delta_1))=\mathcal{N}l \tag{1}$$

holds for any sequence $\langle Q_n\rangle$ in $\mathscr{C}_m$ and every bounded loss function l on R^m.

$\langle Z_n\rangle$ is called LAMA (θ) if it is LAMA $(\mathscr{A},\theta)$ and if $\mathscr{A}$ is the family of all LAN subproblems.

7. Lemma. *If* $\langle\Theta_0,\alpha\rangle$ *is an* LAN $\langle n\tilde{M}, \tilde{\gamma}_n\rangle$ *subproblem then*

$$M=\tilde{M}_{11}, \quad \gamma_n-M^{-1/2}(\tilde{M}^{1/2})_{1.}\tilde{\gamma}_n\to 0 \quad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{1}$$

Proof. Let $t\in R^m$, $\delta_n=\theta_1+(Mn)^{-1/2}t$. $g_n\in dE_{n,\langle\delta_n,\theta_2\rangle}/dE_{n,\theta}$. By the LAN property of the singular subproblem,

$$g_n/\rho_t(\gamma_n)\to 1 \quad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{2}$$

Next, set $\tilde{\theta}=\alpha(\theta)$, $\tilde{\delta}_n=\alpha\langle\delta_n,\theta_2\rangle=\langle\delta_n,\beta(\theta_2)\rangle$. Then $\tilde{\delta}_n=\tilde{\theta}+(\tilde{M}n)^{-1/2}\tilde{t}$ with

$$\tilde{t}=(\tilde{M}n)^{1/2}\langle(Mn)^{-1/2}t,0\rangle=(\tilde{M}^{1/2})_{.1}M^{-1/2}t.$$

By the LAN condition for the subproblem

$$g_n/\rho_{\tilde{t}}(\tilde{\gamma}_n)\to 1 \qquad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{3}$$

This and (2) gives

$$(t'\gamma_n-\tfrac{1}{2}\|t\|^2)-(\tilde{t}'\tilde{\gamma}_n-\tfrac{1}{2}\|\tilde{t}\|^2)\to 0 \qquad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{4}$$

Thus, a comparison of the limiting normal distributions of the two terms in (4) gives $\|\tilde{t}\|=\|t\|$ and then

$$t'(\gamma_n-M^{-1/2}(\tilde{M}^{1/2})_{1.}\tilde{\gamma}_n)\to 0 \qquad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{5}$$

The first property, valid for all $t\in R^k$, implies $\mathbf{1}=M^{-1/2}(\tilde{M}^{1/2})_{1.}(\tilde{M}^{1/2})_{.1}M^{-1/2}$ $=M^{-1/2}\tilde{M}_{11}M^{-1/2}$. This and (5) imply (1).

8. Condition. For every LAN$\langle n\tilde{M},\tilde{\gamma}_n\rangle$ subproblem in $\mathscr{A}$, $\tilde{M}_{12}=0$.

9. Theorem. *Condition 8 is necessary for the existence of an* LAMA$(\mathscr{A},\theta)$ *sequence of estimates.*

Proof. Let $\langle Z_n\rangle$ be an LAMA$(\mathscr{A},\theta)$ sequence of estimates, $\langle\Theta_0,\alpha\rangle$ a k-dimensional LAN$\langle n\tilde{M},\gamma_n\rangle$ subproblem in $\mathscr{A}$, l a bounded loss function on R^m.

Apply (6.1) with the loss function $l(M^{-1/2}\iota_m)$, ι_s the identity function on R^s, and then use the lower bound given by Theorem 2.6 for the subproblem $\langle\Theta_0,\alpha\rangle$ with the loss function $l(\pi_1\tilde{M}^{-1/2}\iota_k)$ and the estimate $\langle Z_n,0\rangle$. Use tilde in $\mathscr{N}$ to refer to R^k. We obtain

$$\tilde{\mathscr{N}}l(\pi_1\tilde{M}^{-1/2}\iota_k)\leqq\mathscr{N}l(M^{-1/2}\iota_m). \tag{1}$$

Applied with $l(x)=\|a'x\|^2\wedge c$, $c\to\infty$, this yields

$$(\tilde{M}^{-1})_{11}\leqq(\tilde{M}_{11})^{-1} \tag{2}$$

since $M=\tilde{M}_{11}$ and in the sense that the appropriate difference of the matrices is positive semidefinite. Thus

$$\tilde{M}_{11}\leqq((\tilde{M}^{-1})_{11})^{-1}=\tilde{M}_{11}-\tilde{M}_{12}(\tilde{M}_{22})^{-1}\tilde{M}_{12}',$$

implying $\tilde{M}_{12}=0$. (We are indebted to Professor Peter J. Bickel for a comment which lead to a shortening of this proof.)

10. Theorem. *Let Condition 8 hold. Let $\langle Z_n\rangle$ be a sequence of estimates which is regular in the singular problem, i.e., which satisfies*

$$n^{1/2}(Z_n-\theta_1)-M^{-1/2}\gamma_n\to 0 \qquad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{1}$$

Then $\langle Z_n\rangle$ is LAMA $(\mathscr{A},\theta)$.

Proof. Let $\langle\Theta_0,\alpha\rangle$ be an LAN $\langle n\tilde{M},\tilde{\gamma}_n\rangle$ subproblem in $\mathscr{A}$. Set

$$\tilde{Z}_n=\langle Z_n,\pi_2[\theta+n^{-1/2}\tilde{M}^{-1/2}\gamma_n]\rangle. \tag{2}$$

Then the part of the regularity condition for $\tilde{Z}_n$ which refers to the second ($k-m$ dimensional) components is automatically satisfied and the regularity condition holds if

$$n^{1/2}(Z_n-\theta_1)-(\tilde{M}^{-1/2})_{1.}\tilde{\gamma}_n\to 0 \quad \text{in } \langle E_{n,\theta}\rangle\text{-prob.} \tag{3}$$

By Condition 8, $(\tilde{M}^{1/2})_{12}=0$ and thus $(\tilde{M}^{\pm 1/2})_{1.}\tilde{\gamma}_n=(\tilde{M}_{11})^{\pm 1/2}\tilde{\gamma}_{n1}$. Then, by Lemma 7, $\gamma_n-\tilde{\gamma}_{n1}\to 0$ in $\langle E_{n,\theta}\rangle$-prob. and (3) follows from (1).

11. Remark. Global adaptivity. Of main interest are sequences of estimates which are LAMA $(\mathscr{A}_\theta, \theta)$ for each θ in Θ. Of course, a necessary condition for such a global adaptivity is obtained directly from Theorem 9, while Theorem 10 gives a sufficient condition.

Note that if the whole problem satisfies Condition LAN $\langle\theta, nM_\theta, \gamma_{n,\theta}\rangle$ and if there exist an auxiliary, at each θ, sequence of estimates then there exists a sequence of estimates which is LAM (θ) at every θ. That sequence of estimates is then LAMA (θ) at every θ if and only if $(M_\theta)_{12}=0$ for each θ.

Whether Condition 8, satisfied for all θ, implies the existence of a globally adaptive sequence of estimates in general, without assuming Condition LAN for the whole problem, is an open question (analogous to an open question in Stein (1956)). However, Theorem 10 gives a simple sufficient condition which has given the desired result in a case considered by Fabian (1980).

12. Example. For symmetric unknown densities, adaptive (classical sense) estimates of location parameter have been constructed by several authors (see §1). It is rather easy to see that in this situation the symmetry of the densities implies Condition 8. On the other hand, if symmetry is not required, the densities involved may be chosen in such a way that Condition 8 fails. Then there is no LAMA (θ) sequence of estimates by Theorem 9.

Pfanzagl (1976) has shown that if the family of densities is rich enough then, in fact, one cannot obtain estimates asymptotically better than the sample quantile. Pfanzagl's result is limited to a certain class of estimates (cf. §1). Here, as an example, we reformulate his result in terms of the LAM property for all possible estimates. (The densities g_δ in the example are chosen essentially as in Pfanzagl (1976).)

Suppose α is a number in $(0, 1)$, φ a density with respect to the Lebesgue measure λ on R. We shall make mild assumptions about φ later.

We shall construct a two parameter family of densities $g_{d,\Delta}$ such that $g_{d,0}=\varphi(\iota-d)$ and show that, roughly speaking, no estimate of the α-quantile can be better than the sample α-quantile (above, ι is the identity function on R).

Denote by r the α-quantile of φ, by β the value $\varphi(r)$, by Φ the distribution function corresponding to φ, by ψ_d the pseudodensity $\varphi\circ(\iota-d)/\varphi$ of $\Phi\circ(\iota-d)$ with respect to Φ.

We assume $\varphi(x)>0$ for all $x\in R$, $\varphi(r)$ is the derivative of Φ at r, $d\rightsquigarrow\psi_d$ has an $L_2(\Phi)$ derivative $\dot{\psi}$ at 0, and $\dot{\psi}\circ(\iota-d)\to\dot{\psi}$ in $L_2(\Phi)$ as $d\to 0$.

If Y_n is the sample α-quantile of n independent random variables, each with density φ, then $\sqrt{n}(Y_n-r)$ is asymptotically normal $(0, \sigma^2)$ with

$$\sigma^2=\frac{\beta^2}{\alpha(1-\alpha)}; \tag{1}$$

this follows from the assumption on $\varphi(r)$.

We shall construct a family of densities g_δ for $\delta=\langle d, \Delta\rangle \in R\times(-1,1)$ with the following properties: (i) $g_{\langle d,0\rangle}=\varphi\circ(\imath-d)$ for all d, (ii) the α-quantile of $g_{\langle d,\Delta\rangle}$ is $r+d$ and (iii) if $E_{n,\delta}$ is the expectation induced by n independent random variables, each with density g_δ then for every sequence $\langle Z_n\rangle$ of estimates, and every loss function l on R,

$$\lim_{K\to\infty}\ \liminf_{n}\ \sup_{\|\delta\|\leqq K/\sqrt{n}} E_{n,\delta}\, l\left(\frac{\sqrt{n}(Z_n-r-d)}{\sigma}\right)\geqq \mathcal{N} l. \tag{2}$$

Let $\|\ \|$, $(,)$ refer to $L_2(\Phi)$. Denote $\|\dot{\psi}\|^2$ by M. Consider the subspace $\mathcal{L}_0$ generated by $\{\chi, 1-\chi\}$, where $\chi=\chi_{(-\infty,r)}$, and the orthogonal complement $\mathcal{L}_1$ to $\mathcal{L}_0$. We have

$$\|\chi\|^2=\alpha, \qquad \|1-\chi\|^2=1-\alpha, \qquad (\chi,\dot{\psi})=-\beta, \qquad (1-\chi,\dot{\psi})=\beta; \tag{3}$$

the first two relations follow from the meaning of α and r, and the last two relations follow by using $x=r$ and $x=\infty$ in

$$(\chi_{(-\infty,x)},\dot{\psi})=\lim_{h\to 0}\frac{1}{h}(\chi_{(-\infty,x)},\psi_h-\psi_0)=\lim_{h\to 0}\frac{1}{h}[\Phi(x-h)-\Phi(x)].$$

Let $\Delta\in(-1,1)$. If $\Delta\neq 0$, set $K=\left(2|\Delta|\left[1+\frac{1}{\sqrt{\alpha}}+\frac{1}{\sqrt{1-\alpha}}\right]\right)^{-1}$; if $\Delta=0$, set $K=\infty$. Denote by $\dot{\psi}_K$ the truncation $(\dot{\psi}\vee(-K))\wedge K$ and by k_Δ the projection of $\dot{\psi}_K$ on $\mathcal{L}_1$; abbreviate k_0 to k. Because of (3),

$$k_\Delta=\dot{\psi}_K-\frac{(\dot{\psi}_K,\chi)}{\alpha}\chi-\frac{(\dot{\psi}_K,1-\chi)}{1-\alpha}(1-\chi). \tag{4}$$

We shall show that

$$|\Delta k_\Delta|\leqq\tfrac{1}{2} \quad \text{for all } \Delta\in(-1,1), \tag{5}$$

and

$$\|k_\Delta\circ(\imath-d)-k\|\to 0 \quad \text{as } \langle d,\Delta\rangle\to 0. \tag{6}$$

Relation (5) obtains from (4) since $|(\dot{\psi}_K,\chi)|\leqq K\|\chi\|$, $|(\dot{\psi}_K,1-\chi)|\leqq K\|1-\chi\|$ and the right hand side is bounded, in absolute value, by $K\left[1+\frac{1}{\sqrt{\alpha}}+\frac{1}{\sqrt{1-\alpha}}\right]$ which is $|\Delta|^{-1}/2$ if $\Delta\neq 0$.

To prove (6) note that $\|\dot{\psi}_K-\dot{\psi}\|\to 0$ and therefore $\|k_\Delta-k\|\to 0$ as $\Delta\to 0$. On the other hand, by (4), and since $(\dot{\psi}_K,\chi)\to(\dot{\psi},\chi)$, $(\dot{\psi}_K,1-\chi)\to(\dot{\psi},1-\chi)$, we obtain

$$|k_\Delta\circ(\imath-d)-k_\Delta|\leqq|\dot{\psi}\circ(\imath-d)-\dot{\psi}|+C\chi_{(r-|d|,\,r+|d|)}$$

with a constant C. Thus (6) follows from the assumed $L_2(\Phi)$ continuity of $\dot{\psi}$ and the continuity of Φ.

Define, for $\delta=\langle d,\Delta\rangle\in R\times(-1,1)$,

$$g_\delta=[\varphi(1+\Delta k_\Delta)]\circ(\imath-d). \tag{7}$$

By (5) we obtain $g_\delta(x)>0$ for all x; since k_Δ is in $\mathcal{L}_1$, and $1\in\mathcal{L}_0$, we have $(k_\Delta, 1)=0$ and $\int g_\delta \, d\lambda=1$. Thus each g_δ is a density; the mass it gives to $(-\infty, r+d)$ is $(1+\Delta k_\Delta, \chi)=\alpha$ so that g_δ has properties (i) and (ii) promised above.

Consider now the pseudodensity

$$q_\delta=\psi_d(1+\Delta k_\Delta \circ (\iota-d)) \tag{8}$$

of $E_{1,\delta}$ with respect to $E_{1,0}$.

Express ψ_d as $1-d\varepsilon(d)$ with $\|\varepsilon(d)-\dot{\psi}\|\to 0$ as $d\to 0$. Then

$$q_\delta-1=\delta'\langle -\varepsilon(d), k_\Delta\circ(\iota-d)\rangle - d\Delta\,\varepsilon(d)\,k_\Delta\circ(\iota-d)$$

so that $r_\delta=q_\delta-1-\delta'\langle -\dot{\psi}, k\rangle$ satisfies

$$r_\delta=\delta'\langle \dot{\psi}-\varepsilon(d), k_\Delta\circ(\iota-d)-k\rangle - d\Delta\,\varepsilon(d)\,k_\Delta\circ(\iota-d)$$

and

$$\|\delta\|^{-1}\|r_\delta\|\leqq\|\dot{\psi}-\varepsilon(d)\|+\|k_\Delta\circ(\iota-d)-k\|+|\Delta|\,\|\varepsilon(d)\|\,\|k_\Delta\circ(\iota-d)\|$$

with the right-hand side converging to 0 as $\delta\to 0$ by properties of $\varepsilon(d)$ and by (6).

We have shown that q_δ has at 0 an $L_2(E_{1,0})$ differential $\dot{q}_0=\langle -\dot{\psi}, k\rangle$. By Remark 2.5 and Theorem 4.8 in Fabian and Hannan (1980), the family $\langle E_{n,\delta}\rangle$ satisfies condition LAN $\langle n\tilde{M}, \tilde{\gamma}_n\rangle$ with $\tilde{M}=E_{1,0}\dot{q}_0(\dot{q}_0)'$. An application of (3) and (4) with $\Delta=0$ gives

$$\|\dot{\psi}-k\|^2=\left\|-\frac{\beta}{\alpha}\chi+\frac{\beta}{1-\alpha}(1-\chi)\right\|^2=\sigma^2.$$

and thus $\|k\|^2=\|\dot{\psi}\|^2-\sigma^2=M-\sigma^2$.

Consequently,

$$\tilde{M}=\begin{bmatrix} M & -M+\sigma^2 \\ -M+\sigma^2 & M-\sigma^2 \end{bmatrix}, \qquad (\tilde{M}^{-1})_{11}=\sigma^{-2}$$

and (2) follows from Theorem 2.6 applied with Z_n-r instead of Z_n. The proof of (i), (ii) and (iii) is now complete.

References

Anderson, T.W.: The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. Proc. Amer. Math. Soc. **6**, 170–176 (1955)

Beran, R.: An efficient and robust adaptive estimator of location. Ann. Statist. **6**, 292–313 (1978)

Beran, R.: Asymptotic lower bounds for risk in robust estimation. Ann. Statist. **8**, 1252–1264 (1980)

Eggleston, H.G.: Convexity. Cambridge: (1958)

Fabian, V.: On uniform convergence of measures. Z. Wahrscheinlichkeitstheorie verw. Gebiete **15**, 139–143 (1970)

Fabian, V.: Asymptotically efficient stochastic approximation; the RM case. Ann. Statist. **1**, 485–495 (1973)

Fabian, V., Hannan, J.: Sufficient conditions for local asymptotic normality. RM-403. Department of Statistics and Probability, Michigan State University. [Submitted to Z. Wahrscheinlichkeitstheorie verw. Gebiete. 1980]

Fabian, V.: Local asymptotic minimax properties of some recursive estimates. RM-409. Department of Statistics and Probability, Michigan State University. [Submitted for publication [1980]]

Hájek, J.: Limiting properties of likelihoods and inference. Foundations of Statistical Inference (Godambe, V.P. and D.A. Sprott eds.) 142-162. New York: Holt, Rinehart and Winston 1971

Hájek, J.: Local asymptotic minimax and admissibility in estimation. Proc. Sixth Berkeley Sympos. Math. Statist. Probab. I, 175-194. Univ. Calif. Press (1972)

Has'minskij, R.Z., Ibragimov, I.A.: On the nonparametric estimation of functionals. Proc. Second Prague Symp. Asympt. Statist. 41-51. Charles University, Prague (1979)

Ibragimov, I.A., Has'minskij, R.Z.: Local asymptotic normality for non-identically distributed observations. Theory Probab. Appl. **20**, 246-260 (1975)

Ibragimov, I.A., Has'minskij, R.Z.: Asymptotic theory of estimation. Nanka, Moskva (1979). (Russian; English translation announced)

Koshevnik, Yu.A., Levit, B.Ya.: On a non-parametric analogue of the information matrix. Theory Probab. Appl. **21**, 738-753 (1976)

Le Cam, L.: On the asymptotic theory of estimation and testing hypotheses. Proc. Third Berkeley Sympos. Math. Statist. Probab. **I**, 129-156. Univ. Calif. Press (1956)

Le Cam, L.: Locally asymptotically normal families of distributions. Univ. Calif. Publ. Statist. **3**, 37-98 (1960)

Le Cam, L.: Théorie asymptotique de la décision statistique. Séminaire de mathématiques supérieures-été 1968. Les presses de l'Université de Montréal (1969)

Le Cam, L.: On the assumptions used to prove asymptotic normality of maximum likelihood estimates. Ann. Math. Statist. **41**, 802-828 (1970)

Le Cam, L.: Limits of experiments. Proc. Sixth Berkeley Sympos. Math. Statist. Probab. **I**, 245-261. Univ. Calif. Press (1972)

Le Cam, L.: Notes on Asymptotic Methods in Statistical Decision Theory. Centre de Recherches Mathématiques. Universite de Montréal, Montréal (1974)

Le Cam, L.: On a theorem of J. Hájek. Contribution to Statistics, Jaroslav Hájek Memorial Volume, Jana Jurečková, ed. Dordrecht, Holland: Reidel Publ. 1979

Levit, B.: On optimality of some statistical estimates. Proc. Prague Symp. Asympt. Statist. **II**, 215-238. Universita Karlova. Praha (1974)

Levit, B.Ya.: On the efficiency of a class of non-parametric estimates. Theory Probab. Appl. **20**, 723-740 (1975)

Pfanzagl, J.: Investigating the quantile of an unknown distribution. Contrib. to Appl. Statist. (Ziegler, W.J., ed.). Basel: Birkhäuser 1976

Rao, R.R.: Relations between weak and uniform convergence of measures with applications. Ann. Math. Statist. **33**, 659-680 (1962)

Stein, C.: Efficient nonparametric testing and estimation. Proc. Third Berkeley Sympos. Math. Statist. and Probab. **I**, 187-195. Univ. Calif. Press (1956)

Stone, C.J.: Adaptive maximum likelihood estimators of a location parameter. Ann. Statist. **3**, 267-284 (1975)

Weiss, L., Wolfowitz, J.: Asymptotically efficient non-parametric estimators of location and scale parameters. Z. Wahrscheinlichkeitstheorie verw. Gebiete **16**, 134-150 (1970a)

Weiss, L., Wolfowitz, J.: Asymptotically efficient estimation of non-parametric regression coefficients. Statistical Decision Theory (S.S. Gupta and J. Yaeckel, eds.) 29-39 (1970b)

Received June 23, 1980; in revised form September 4, 1981