Sankhyā : The Indian Journal of Statistics
1989, Volume 51, Series A, Pt. 3, pp. 343-369

LOCAL ASYMPTOTIC MINIMAX ESTIMATION IN NON-REGULAR CASE

By TAPAS SAMANTA

Indian Statistical Institute

SUMMARY. A lower bound to the local asymptotic minimax risk is obtained for a family of non-regular cases using the notion of limiting experiments. A locally asymptotically minimax estimator is suggested using this bound. Also the asymptotic properties of maximum probability estimator (for 0-1 loss function) and Bayes estimators are studied.

1. Introduction

Let $\{f_n(\cdot, \theta)\}$ $(n \geqslant 1)$ be a family of densities depending on a parameter θ taking values in Θ where Θ is an open subset of the real line R. Our problem is to estimate θ efficiently. Let $\{T_n\}$ be a sequence of estimators of θ. Hajek (1972) considered the quantity

$$R(\theta, \{T_n\}) = \lim_{\delta \to 0} \varliminf_{n \to \infty} \sup_{|\theta' - \theta| < \delta} E_{\theta'} L[K_n(T_n - \theta')] \tag{1.1}$$

as a measure of the asymptotic performance at θ (see Ghosh, 1985), where L is an appropriate loss function and $K_n(\uparrow \infty)$ is the normalizing factor (see Weiss and Wolfowitz, 1974) for the given family of distributions (for the usual regular cases $K_n = \sqrt{n}$). An estimator minimizing (1.1) can thus be considered as an efficient estimator. For regular cases a lower bound to the local asymptotic (maximum) risk (1.1) was established in Hajek (1972). It is also proved that the maximum likelihood estimator and Bayes estimators attain this lower bound under certain regularity assumptions (see, for example, Ibragimov and Hasminskii, 1981). In this paper, we consider the non-regular cases where the support of the density depends on the parameter θ. We, however, consider a variant of (1.1) :

$$\rho(\theta, \{T_n\}) = \lim_{A \to \infty} \varliminf_{n \to \infty} \sup_{|\theta' - \theta| \leqslant AK_n^{-1}} E_{\theta'} L[K_n(T_n - \theta')]$$

as the asymptotic performance criterion following Fabian and Hannan (1982) and Millar (1983). Thus, an estimator T_n for which the local asymptotic risk

AMS (1980) *subject classification :* Primary 62F12 ; Secondary 62C20.

Key words and phrases : Local asymptotic minimax risk, lower bound, locally asymptotically minimax estimator, convergence of experiments, convolution theorem, non-regular classes of densities.

$\rho(\theta, \{T_n\})$ (with $\underline{\lim}$ replaced by lim) is equal to the local asymptotic minimax (LAM) risk

$$\rho(\theta) = \lim_{A\to\infty} \underline{\lim}_{n\to\infty} \inf_{T_n} \sup_{|\theta'-\theta|\leqslant AK_n^{-1}} E_{\theta'} L[K_n(T_n-\theta')]$$

may be considered as an efficient estimator. In Section 2 of this paper a lower bound to the LAM risk is obtained for a general class of densities admitting certain local asymptotic expansion of the likelihood ratio. The same bound was obtained in Ibragimov and Hasminskii (1981, Ch. V) using some other approach. We here use the results of Millar (1983) to get a sequence of experiments converging to some exponential shift experiment and then obtain the lower bound using the Hajek-Le Cam asymptotic minimax theorem (Millar, 1983). An estimator is suggested in Section 2 which is shown to be efficient in the above sense. A convolution theorem which gives the decomposition of the limiting distribution of a sequence of estimators is also proved using the notion of convergence of experiments. In Section 3 we consider specific non-regular cases. It is shown that the conditions of the theorems of Section 2 hold for these non-regular cases and hence the estimator suggested in Section 2 is efficient. In Sections 4 and 5 the asymptotic behaviour of the maximum probability estimators and the Bayes estimators are studied for the family of non-regular cases considered in Section 3 and it is shown that these estimators are efficient under certain conditions.

2. LAM estimation under an asymptotic expansion of likelihood ratio

2.1 *Convergence of experiments assuming asymptotic expansion.* Let $\{(\mathscr{X}^n, \mathscr{A}^n), P_\theta^n ; \theta \in \Theta\}, n \geqslant 1$, be a sequence of statistical experiments, where Θ is an open subset of the real line $\boldsymbol{R}$. Let $dP_{\theta_2}^n/dP_{\theta_1}^n$ denote the derivative of the absolutely continuous component of $P_{\theta_2}^n$ with respect to $P_{\theta_1}^n$. Fix $\theta_0 \in \Theta$. We assume that either of the following two conditions holds a.e. $P_{[\theta_0]}^n$.

Condition $(A1)$: For any $\lambda \geqslant 0$ and some sequence $K_n \uparrow \infty$,

$$\frac{dP^n_{\theta_0+\lambda K_n^{-1}}}{dP^n_{\theta_0}} = \begin{cases} \exp\{\lambda\, \Delta_n(\theta_0)+e_n(\lambda, \theta_0)\}, & \text{if } K_n(Z_n-\theta_0) > \lambda \\ 0, & \text{if } K_n(Z_n-\theta_0) < \lambda, \end{cases} \quad \ldots \quad (2.1)$$

where $\Delta_n(\theta_0)$ converges in $P_{\theta_0}^n$-probability to $c(\theta_0)$ for some $c(\theta_0) > 0$,

ϵ_n converges in $P^n_{\theta_0}$-probability to zero and Z_n is a random variable which does not depend on θ_0 and for which

$$\lim_{n\to\infty} P^n_{\theta_0}(K_n(Z_n-\theta_0) > t) = e^{-tc(\theta_0)} \text{ for all } t \geqslant 0.$$

Condition (*A*2) : For any $\lambda \leqslant 0$ and some sequence $K_n \uparrow \infty$,

$$\frac{dP^n_{\theta_0+\lambda K_n^{-1}}}{dP^n_{\theta_0}} = \begin{cases} \exp\{\lambda\,\Delta^*_n(\theta_0)+\epsilon^*_n(\lambda,\theta_0)\}, \text{ if } K_n(Z^*_n-\theta_0) < \lambda \\ 0, \qquad\qquad \text{if } K_n(Z^*_n-\theta_0) > \lambda \end{cases} \quad \dots \quad (2.2)$$

where $\Delta^*_n(\theta_0)$ converges in $P^n_{\theta_0}$-probability to $c^*(\theta_0)$ for some $c^*(\theta_0) < 0$, ϵ^*_n converges in $P^n_{\theta_0}$-probability to zero and Z^*_n is a random variable which does not depend on θ_0 and satisfies

$$\lim_{n\to\infty} P^n_{\theta_0}(K_n(Z^*_n-\theta_0) < t) = e^{-tc^*(\theta_0)} \text{ for all } t \leqslant 0.$$

We now define experiments

$$E^n = \{P^n_{\theta_0+\lambda K_n^{-1}} : \lambda \geqslant 0\} \text{ and } E^{*n} = \{P^n_{\theta_0+\lambda K_n^{-1}} : \lambda \leqslant 0\},\ n \geqslant 1.$$

Let Q_{λ,θ_0} $(\lambda \geqslant 0)$ denote a probability on $\boldsymbol{R}$ with density

$$q_{\lambda,\theta_0}(x) = \begin{cases} c(\theta_0)\, e^{-c(\theta_0)(x-\lambda)}, & \text{for } x > \lambda, \\ 0, & \text{for } x \leqslant \lambda, \end{cases}$$

and Q^*_{λ,θ_0} $(\lambda \leqslant 0)$ denote a probability on $\boldsymbol{R}$ with density

$$q^*_{\lambda,\theta_0}(x) = \begin{cases} -c^*(\theta_0) e^{-c^*(\theta_0)(x-\lambda)}, & \text{for } x < \lambda, \\ 0, & \text{for } x \geqslant \lambda. \end{cases}$$

Then we have the following result :

Theorem 1 : (*i*) *Under condition* (*A*1) *the sequence of experiments* E^n *converges to* $E = \{Q_\lambda : \lambda \geqslant 0\}$.

(*ii*) *Under condition* (*A*2) *the sequence of experiments* E^{*n} *converges to* $E^* = \{Q^*_\lambda : \lambda \leqslant 0\}$.

(We write just Q_λ and Q^*_λ in place of Q_{λ,θ_0} and Q^*_{λ,θ_0}).

Proof : We will give the proof for case (*i*) only. The proof of case (*ii*) is exactly similar.

Set $$Q^n_\lambda = P^n_{\theta_0+\lambda K_n^{-1}}.$$

It is given that for all $\lambda \geqslant 0$,

$$\frac{dQ^n_\lambda}{dQ^n_0} = \begin{cases} \exp(Y_n) & \text{on } B_n, \\ 0, & \text{otherwise,} \end{cases}$$

where $Y_n \xrightarrow{Q^n_0} \lambda c(\theta_0)$ and $Q^n_0(B_n) \to \exp(-\lambda c(\theta_0))$ as $n \to \infty$. This gives us

$$\mathcal{L}\left\{\frac{dQ^n_\lambda}{dQ^n_0}\,\Big|\, Q^n_0\right\} \Rightarrow \mathcal{L}\left\{\frac{dQ_\lambda}{dQ_0}\,\Big|\, Q_0\right\}.$$

Since $E_{Q_0}\left(\frac{dQ_\lambda}{dQ_0}\right) = 1$, by a result on contiguity (referred to as Le Cam's 1st lemma in Hajek and Sidak (1967)) it follows that Q^n_λ is contiguous to Q^n_0 for all $\lambda \geqslant 0$.

Further, using the asymptotic expansion (2.1) again we can prove that for $0 \leqslant \lambda_1 < \lambda_2 < \ldots < \lambda_k$,

$$\mathcal{L}\left\{\left(\frac{dQ^n_{\lambda_1}}{dQ^n_0}, \frac{dQ^n_{\lambda_2}}{dQ^n_0}, \ldots, \frac{dQ^n_{\lambda_k}}{dQ^n_0}\right)\Big|\, Q^n_0\right\} \Rightarrow \mathcal{L}\left\{\left(\frac{dQ_{\lambda_1}}{dQ_0}, \frac{dQ_{\lambda_2}}{dQ_0}, \ldots, \frac{dQ_{\lambda_k}}{dQ_0}\right)\Big|\, Q_0\right\}.$$

Hence by proposition II.2.3 of Millar (1983) the theorem is proved.

Remark 1.1 : Contiguity plays an important role in the proof of the above therem. Millar's results cannot be applied if $\{P^n_{\theta_0+\lambda K_n^{-1}}\}$ is not contiguous to $\{P^n_{\theta_0}\}$ and it is usually very difficult to solve the problem if contiguity does not hold. In the proof of the above theorem we have seen that condition $(A1)$ implies contiguity. Now suppose (2.1) holds for all $\lambda \geqslant 0$ where $\Delta_n(\theta_0)$ and ϵ_n are as in condition $(A1)$ and Z_n is a sequence of random variables such that $\mathcal{L}\{K_n(Z_n-\theta_0)\,|\,P^n_{\theta_0}\}$ converges weakly to some arbitrary distribution. Then to have contiguity we must have

$$\lim_{n\to\infty} P^n_{\theta_0}(K_n(Z_n-\theta_0) > t) = e^{-tc(\theta_0)} \text{ for all } t \geqslant 0.$$

This follows from a result on contiguity.

Remark 1.2 : If condition $(A1)$ is replaced by the following stronger condition :

$(A1)^*$ For any real u and v such that $u < v$

$$\frac{dP^n_{\theta_0+vK_n^{-1}}}{dP^n_{\theta_0+uK_n^{-1}}} = \begin{cases} \exp\{(v-u)\,\Delta_n(\theta_0)+\epsilon_n\}, & \text{if } K_n(Z_n-\theta_0) > v \\ 0, & \text{otherwise,} \end{cases}$$

where $\Delta_n(\theta_0)$ and ϵ_n are as in ($A1$) but the convergence is with respect to $P^n_{\theta_0+uK_n^{-1}}$ and Z_n is such that

$$P^n_{\theta_0+uK_n^{-1}}[K_n(Z_n-\theta_0)>v]\to e^{-(v-u)c(\theta_0)},$$

then proceeding as above the sequence of experiments $\{P^n_{\theta_0+\lambda K_n^{-1}}, \lambda \epsilon \boldsymbol{R}\}$ may be shown to be converging to the experiment $\{Q_\lambda, \lambda \epsilon \boldsymbol{R}\}$.

From now onwards, we will consider only the case where condition ($A1$) is satisfied. The treatment for the case where condition ($A2$) holds is similar with obvious modifications.

2.2 *Lower bound for asymptotic risk and an efficient estimator.* In this section we obtain a lower bound to the local asymptotic minimax risk using Theorem 1 and the Hajek-Le Cam asymptotic minimax theorem (see Millar (1983), Ch. III).

Definition : A loss function of the form $L(\theta, a) = L(\theta-a)$ is said to be subconvex if L satisfies the following conditions :

(*i*) $L(x) \geqslant 0$ for all x.

(*ii*) $L(x) = L(-x)$ for all x.

(*iii*) $\{x : L(x) \leqslant c\}$ is closed and convex for all $c > 0$.

All the loss functions considered in this paper will be assumed to be subconvex.

Lemma 1 : *Under assumption* ($A1$), *for any subconvex loss function* L,

$$\lim_{A\to\infty}\ \varliminf_{n\to\infty}\ \inf_{T_n}\ \sup_{|\theta-\theta_0|\leqslant AK_n^{-1}} E_\theta L[K_n(T_n-\theta)] \geqslant \inf_{\delta}\ \sup_{0\leqslant\lambda<\infty} R(\delta,\lambda) \ \ldots \quad (2.3)$$

where the infimum in left hand side is over all estimators T_n of θ, the infimum in right hand side is over all randomized (Markov kernel) procedures for the experiment E with decision space as $\mathbf{R}$ *and parameter space as* $[0, \infty)$ *and $R(\delta, \lambda)$ is the risk of the procedure δ at λ with loss function L.*

Proof : The proof is similar to that of Theorem VII .2.6 of Millar (1983). We use Theorem 1 and the asymptotic minimax theorem. The infimum in the right hand side of (2.3) may be taken over all Markov kernels (transition probabilities) because in the decision theoretic structure for the limiting experiment all generalized procedures (see Millar (1983), Ch.II) are given by Markov Kernels.

We will now compute the minimax risk given in the right hand side of (2.3). We will use a well-known technique of finding minimax risk

We assume that

C(i) $E_{Q_0}L(X-a) = \int L(x-a)dQ_0(x)$ exists and is finite for some a and there exists $b = b(\theta_0)$ such that

$$E_{Q_0}L(X-b(\theta_0)) = \inf_a E_{Q_0}L(X-a) = R_{\theta_0}, \text{ say.}$$

C(ii) For every $\epsilon > 0$, there exists $N > 0$ such that for all $a \in R$,

$$\int_0^N L(x-a)\, dQ_0(x) \geqslant R_{\theta_0} - \epsilon.$$

C(iii) $b(\theta)$ is a continuous function of θ.

Lemma 2 : *For any subconvex loss function satisfying conditions C(i) and C(ii), we have*

$$\inf_{\delta} \sup_{0 \leqslant \lambda < \infty} R(\delta, \lambda) = \int_0^\infty L(x-b(\theta_0))c(\theta_0)e^{-c(\theta_0)x}\, dx$$

where the minimax risk in the left hand side is as described in Lemma 1.

Proof : We shall exhibit a sequence τ_M of prior distributions on $[0, \infty)$ and show that

$$\lim_{M \to \infty} \inf_{\delta} r(\delta, \tau_M) \geqslant R_{\theta_0} \quad \ldots \quad (2.4)$$

where the infimum in the left hand side is over all randomized (Markov kernel) procedures and $r(\delta, \tau_M)$ is the Bayes risk of δ with respect to the prior τ_M.

We choose τ_M as the uniform distribution over the interval $(0, M)$. Let $\epsilon > 0$ and N be such that $\int_0^N L(x-a)\, dQ_0(x) \geqslant R_{\theta_0} - \epsilon$ for all a. Proceeding as in Ferguson (1967, Section 4.6, p. 172) we can prove that for any $M > N$ and any nonrandomized decision rule δ,

$$r(\delta, \tau_M) \geqslant (R_{\theta_0} - \epsilon)\, \frac{M-N}{M}.$$

Therefore for any $M > N$, $r(\delta, \tau_M) \geqslant (R_{\theta_0} - \epsilon)\, \frac{M-N}{M}$ for all "randomized" procedures δ which are probabilities over the space of nonrandomized decision rules. This proves (2.4) using a result on equivalence of two methods of randomization (see, for example, Wald and Wolfowitz, 1951). Since $X - b(\theta_0)$ is an equalizer rule with constant risk R_{θ_0}, the lemma is proved.

Now, from Lemma 1 and Lemma 2 we get the following result :

Theorem 2 : *Under assumption (A1), for any subconvex loss function L satisfying C(i) and C(ii),*

$$\lim_{A\to\infty} \underline{\lim}_{n\to\infty} \inf_{T_n} \sup_{|\theta-\theta_0|<AK_n^{-1}} E_\theta L[K_n(T_n-\theta)]$$

$$\geqslant \int_0^\infty L(x-b(\theta_0))\, c(\theta_0)\, e^{-c(\theta_0)x}\, dx.$$

Remark : To prove Theorem 2 we need not assume that Z_n (in Condition (A1)) is independent of θ_0. Indeed, we may replace the set $\{K_n(Z_n-\theta_0)>\lambda\}$ by $\{\tau_n>\lambda\}$, where τ_n is a random variable such that

$$\lim_{n\to\infty} P^n_{\theta_0}\ (\tau_n>t)=e^{-tc(\theta_0)} \qquad \text{for all } t\geqslant 0.$$

Our problem is now to search for an estimator $\hat{\theta}_n$ for which

$$\left.\begin{aligned} &\lim_{A\to\infty} \underline{\lim}_{n\to\infty} \inf_{T_n} \sup_{|\theta-\theta_0|<AK_n^{-1}} E_\theta L[K_n(T_n-\theta)] \\ &=\lim_{A\to\infty} \lim_{n\to\infty} \sup_{|\theta-\theta_0|\leqslant AK_n^{-1}} E_\theta L[K_n(\hat{\theta}_n-\theta)] \text{ for all } \theta_0\in\Theta \end{aligned}\right\} \quad \ldots(2.5)$$

Definition : An estimator $\hat{\theta}_n$ for which (2.5) holds is said to be a locally asymptotically minimax (LAM) estimator of θ.

It follows from Theorem 2 that an estimator $\hat{\theta}_n$ for which

$$\lim_{A\to\infty} \lim_{n\to\infty} \sup_{|\theta-\theta_0|<AK_n^{-1}} E_\theta L[K_n(\hat{\theta}_n-\theta)]$$

$$=\int_0^\infty L(x-b(\theta_0))\, c(\theta_0)\, e^{-c(\theta_0)x}\, dx$$

is a locally asymptotically minimax estimator.

Let us now consider the case for which condition (A1) is satisfied for all $\theta_0\in\Theta$. Condition (A1) ensures the existence of a sequence of statistics Z_n for which $K_n(Z_n-\theta_0)$ converges in distribution (under $P^n_{\theta_0}$) to a random variable X with distribution Q_0.

Definition : A sequence of estimators T_n is said to be regular at $\theta_0\in\Theta$ if for some probability distribution G,

$$\mathcal{L}\left\{K_n(T_n-\theta_0-\lambda K_n^{-1})\,|\,P^n_{\theta_0+\lambda K_n^{-1}}\right\}\Rightarrow G \text{ as } n\to\infty$$

uniformly in $\{|\lambda|\leqslant c\}$ for any $c>0$.

Theorem 3 : *Suppose condition (A1) holds for all $\theta_0 \in \Theta$ and the sequence of statistics Z_n is regular at all values θ_0 in Θ.*

Set $$\hat{\theta}_n = Z_n - K_n^{-1} b(Z_n)$$

Then the following results hold :

(*i*) *For any bounded subconvex loss function satisfying conditions $C(i)$, $C(iii)$ (condition $C(ii)$ is satisfied for bounded loss function) $\hat{\theta}_n$ is LAM.*

(*ii*) *Suppose that for some $r > 0$,*

$$\lim_{A\to\infty} \lim_{n\to\infty} \sup_{|\theta-\theta_0| \leqslant AK_n^{-1}} E_\theta |K_n(\hat{\theta}_n - \theta)|^r < \infty \qquad \ldots (2.6)$$

for all $\theta_0 \in \Theta$. Then for any subconvex loss function L satisfying conditions $C(i)$, $C(ii)$ and $C(iii)$, for which

$$L(u) \leqslant B(1+|u|^s) \text{ for all } u \in \mathbf{R}, \text{ for some } B > 0 \text{ and } 0 < s < r,$$

we have

$$\lim_{A\to\infty} \lim_{n\to\infty} \sup_{|\theta-\theta_0| \leqslant AK_n^{-1}} E_\theta L[K_n(\hat{\theta}_n - \theta)]$$

$$= \int L(x - b(\theta_0)) dQ_0 \text{ for all } \theta_0 \in \Theta$$

and hence $\hat{\theta}_n$ is LAM.

Proof: Fix $A > 0$. Under the conditions of the theorem, for any $\theta_0 \in \Theta$ and for any sequence $\{\theta_n\}$ satisfying $|K_n(\theta_n - \theta_0)| \leqslant A$

$$\mathcal{L}\{K_n(Z_n - \theta_n) | P^n_{\theta_n}\} \Rightarrow Q_{0,\theta_0}.$$

Since $b(\theta)$ is continuous in θ, $b(Z_n)$ converges in $P^n_{\theta_n}$-probability to $b(\theta_0)$. Thus,

$$\mathcal{L}\{K_n(\hat{\theta}_n - \theta_n) | P^n_{\theta_n}\} \Rightarrow \mathcal{L}\{X - b(\theta_0)\},$$

where X is a random variable with distribution Q_{0,θ_0}.

We shall now prove that

$$\mathcal{L}\{L[K_n(\hat{\theta}_n - \theta_n)] | P^n_{\theta_n}\} \Rightarrow \mathcal{L}\{L(X - b(\theta_0))\}. \qquad \ldots (2.7)$$

Take any $t \geqslant 0$. $B_t = \{x : L(x) \leqslant t\}$ is closed convex subset of $\mathbf{R}$. Since the Lebesgue measure of the boundary of any convex set is zero, B_t is a continuity set with respect to the distribution of $X - b(\theta_0)$ and we have

$$\lim_{n\to\infty} P^n_{\theta_n}(L[K_n(\hat{\theta}_n - \theta_n)] \leqslant t) = \lim_{n\to\infty} P^n_{\theta_n}[K_n(\hat{\theta}_n - \theta_n) \in B_t]$$

$$= Q_{0,\theta_0}(\{x : (x - b(\theta_0)) \in B_t\})$$

$$= Q_{0,\theta_0}(\{x : L(x - b(\theta_0)) \leqslant t\}).$$

Hence (2.7) is proved.

Now
$$\lim_{n\to\infty} \sup_{|\theta-\theta_0|\leqslant AK_n^{-1}} E_\theta L[K_n(\hat{\theta}_n-\theta)]$$
$$= \lim_{n\to\infty} E_{\theta_n} L[K_n(\hat{\theta}_n-\theta_n)]$$

for some sequence $\{\theta_n\}$ satisfying $|K_n(\theta_n-\theta_0)| \leqslant A$. The proof of the 1st part of the theorem is now obvious. We will now prove the 2nd part of the theorem. We are given that

$$\operatorname*{Lim}_{n\to\infty} E_{\theta_n} |K_n(\hat{\theta}_n-\theta_n)|^r < \infty.$$

Since $L(u) \leqslant B(1+|u|^\theta)$ for all $u \in \boldsymbol{R}$, there exists $\epsilon > 0$ such that for some B_1, $B_2 > 0$,

$$[L(u)]^{1+\epsilon} \leqslant B_1+B_2|u|^r \text{ for all } u \in \boldsymbol{R}.$$

Therefore we have

$$\lim_{n\to\infty} E_{\theta_n}|L[K_n(\hat{\theta}_n-\theta_n)]|^{1+\epsilon} \leqslant B_1+B_2 \lim_{n\to\infty} E_{\theta_n}|K_n(\hat{\theta}_n-\theta_n)|^r < \infty.$$

This together with (2.7) proves the theorem.

Remark 3.1 : Part (*i*) of Theorem 3 holds for any sequence of regular estimators $\hat{\theta}_n$ for which $K_n(\hat{\theta}_n-\theta_0)$ converges in distribution (under $P^n_{\theta_0}$) to $X-b(\theta_0)$ ($X \sim Q_0$) for all $\theta_0 \in \Theta$.

Remark 3.2 : It is interesting to note that condition (A1) itself implies that for all $\lambda \geqslant 0$

$$\mathscr{L}(K_n(Z_n-\theta_0-\lambda K_n^{-1})\,|\,P^n_{\theta_0+\lambda K_n^{-1}}) \Rightarrow Q_0 \qquad \dots \quad (2.8)$$

This follows from the fact that $P^n_{\theta_0+\lambda K_n^{-1}}$ is contiguous to $P^n_{\theta_0}$. Similarly condition (A1)* implies that (2.8) holds for all real λ. Moreover, under uniform versions of conditions (A1) and (A1)*, where the convergence of ϵ_n are uniform in λ and (u, v) belonging to compact sets, the convergence (2.8) may be shown to be uniform for all λ in compact sets.

Remark 3.3 : The LAM estimator here depends on the loss function chosen whereas in the regular case it is possible to find LAM estimators not depending on the loss function.

2.3 *A convolution theorem* : We shall now try to characterize the class of possible limiting distributions of appropriately normalized estimators.

We shall consider only the class of estimators T_n (which includes all regular estimators) for which

$$\mathcal{L}\left\{ K_n(T_n-\theta_0-\lambda K_n^{-1}) \mid P^n_{\theta_0+\lambda K_n^{-1}} \right\} \Rightarrow G \quad \text{for all } \lambda \geqslant 0 \qquad \ldots \quad (2.9)$$

where G is some probability distribution not depending on λ.

Theorem 4 : *Suppose condition* (A1) *holds. Then for any estimator* T_n *satisfying* (2.9) *the limiting distribution* G *of* $K_n(T_n-\theta_0)$ *under* $P^n_{\theta_0}$ *is a convolution of* Q_0 *and some probability distribution* μ *depending on* $\{T_n\}$:

$$G = Q_0 * \mu.$$

To prove Theorem 4, we shall use a slightly different version of Millar's Convolution Theorem (Theorem III.2.10 in Millar, 1983) which we state and prove below (For an alternative proof of Theorem 4 see Ibragimov and Hasminskii, 1981, Ch. V.).

Theorem (Millar) : *Let* $E^n = \{(S^n, \mathcal{S}^n), Q^n_\lambda, \lambda \geqslant 0\}$, $n \geqslant 1$, $E = \{(\boldsymbol{R}, \mathcal{B}), Q_\lambda, \lambda \geqslant 0\}$ *be statistical experiments* ($\mathcal{B}$ *denotes the Borel* σ*-field on* $\boldsymbol{R}$). *Assume* E^n *converges to* E. *Suppose* R_n *is a sequence of statistics on* $(S^n, \mathcal{S}^n)$ *taking values in* $\boldsymbol{R}$. *Assume further*

(*i*) *there is a family of probabilities* $\{G_\lambda, \lambda \geqslant 0\}$ *on* $(R, \mathcal{B})$ *such that for each* $\lambda \geqslant 0$,

$$\mathcal{L}\{R_n \mid Q^n_\lambda\} \Rightarrow G_\lambda.$$

(*ii*) Q_0 *is concentrated on* $\boldsymbol{R}^+$ *and is absolutely continuous with respect to Lebesgue measure. Also the number* 0 *belongs to the support of* Q_0.

(*iii*) $Q_\lambda(A) = Q_0(A-\lambda)$, $G_\lambda(A) = G_0(A-\lambda)$ *for all* $\lambda \geqslant 0$ *and all* $A \in \mathcal{B}$.

Then there is a probability μ *on* $\boldsymbol{R}$ *such that*

$$G_0 = Q_0 * \mu.$$

Proof : Proceeding as in the proof given in Millar (1983) we can get a Markov kernel K of $(\boldsymbol{R},\mathcal{B})/(\boldsymbol{R},\mathcal{B})$ such that for every $g \geqslant 0$ and $A \in \mathcal{B}$

$$K(x,A) = K(x+g, A+g) \ \text{ a.e. } x \geqslant 0$$

and also

$$G_\lambda = KQ_\lambda \ \text{ for all } \lambda \geqslant 0.$$

Therefore, by Fubini's theorem there exists a null set N such that

for $x \notin N$, $x \geqslant 0$, $\{K(x, A) = K(x+g, A+g)$ for all $A \in \mathcal{B}\}$ a.e. $g \geqslant 0$.

We now choose a sequence $\alpha_n \downarrow 0$, $\alpha_n \in N^c$ for all n.

For all $n \geqslant 1$, there is a null set N_n such that for all $g \notin N_n, g \geqslant 0$

$$K(\alpha_n+g, A+g) = K(\alpha_n, A) \text{ for all } A \in \mathcal{B}.$$

Therefore, for all $x \notin N_n+\alpha_n, x \geqslant \alpha_n$,

$$K(x, A+x) = K(\alpha_n, A+\alpha_n) \text{ for all } A \in \mathcal{B}.$$

Let
$$N_0 = \bigcup_{n \geqslant 1} (N_n+\alpha_n).$$

then N_0 is a null set and for any $x, y > 0$ such that $x \notin N_0, y \notin N_0$ we have

$$K(x, A+x) = K(y, A+y) \text{ for all } A \in \mathcal{B}.$$

To see this choose $\alpha_n < x, y$ and note that $x, y \notin N_0$ implies $x, y \notin N_n+\alpha_n$ and hence $K(x, A+x) = K(\alpha_n, A+\alpha_n) = K(y, A+y)$. Suppose the common value is $\mu(A)$. i.e., for all $x \notin N_0$,

$$K(x, A+x) = \mu(A) \text{ for all } A \in \mathcal{B}.$$

This implies

$$K(x, A) = K(x, (A-x)+x) = \mu(A-x) \text{ for all } A \in \mathcal{B}, \text{ for all } x \notin N_0$$

and therefore,

$$G_\lambda(A) = \int K(x, A)\, dQ_\lambda(x) = \int \mu(A-x)\, dQ_\lambda(x).$$

Since μ is a probability this proves the theorem.

Proof of Theorem 4: Consider $E^n = \{Q_\lambda^n : \lambda \geqslant 0\}, E = \{Q_\lambda : \lambda \geqslant 0\}$, $R_n = K_n(T_n-\theta_0)$, where Q_λ is as defined earlier in this section and

$$Q_\lambda^n = P^n_{\theta_0+\lambda K_n^{-1}} \text{ for } \lambda \geqslant 0.$$

It is easy to see that all the conditions of the above theorem are satisfied and hence

$$G = Q_0 * \mu$$

for some probability μ on R.

Corollary: *Under the conditions of Theorem 4, for any sequence of estimators T_n satisfying* (2.9), *we have*

$$\lim_{n\to\infty} E_{\theta_0} L[K_n(T_n-\theta_0)] \geqslant \lim_{n\to\infty} E_{\theta_0} L[K_n(\hat{\theta}_n-\theta_0)]$$

where L is a loss function satisfying the conditions given in Theorem 3.

Proof. By convolution theorem

$$\mathcal{L}\{L[K_n(T_n-\theta_0)] \mid P^n_{\theta_0}\} \Rightarrow \mathcal{L}\{L(X+\xi)\}$$

where X and ξ are independent random variables and $X \sim Q_0$.

Using Fatou's lemma

$$\varliminf_{n\to\infty} E_{\theta_0} L[K_n(T_n-\theta_0)] \geqslant E\,L(X+\xi)$$
$$= \int E\,L(X+y)dF_\xi(y)$$
$$\geqslant E\,L[X-b(\theta_0)]$$
$$= \lim_{n\to\infty} E_{\theta_0} L[K_n(\hat\theta_n-\theta_0)].$$

3. Examples of non-regular cases

In this section we apply the results of the previous section for the estimation problem in two important classes of non-regular examples. We obtain the asymptotic lower bound and also suggest efficient estimators in each of these cases.

3.1 *I.I.D. observations*: Let $X_1, X_2, \ldots, X_n$ be i.i.d. observations, each X_i having density $f(x, \theta)$ on $\boldsymbol{R}$ with respect to the Lebesgue measure, where $\theta \in \Theta$, an open subset of $\boldsymbol{R}$.

We assume that $f(x, \theta)$ is strictly positive for all x in a closed interval (bounded or unbounded) $S(\theta)$ depending on θ and is zero outside $S(\theta)$. Let $A_1(\theta)$, $A_2(\theta)$, $(A_1 < A_2)$ be the boundaries of $S(\theta)$. We consider the following cases:

Case I. The support $S(\theta)$ is nonincreasing in θ, i.e., $S(\theta_2) \subseteq S(\theta_1)$ wherever $\theta_2 > \theta_1$

Case II. The support $S(\theta)$ is nondecreasing in θ, i.e., $S(\theta_2) \supseteq S(\theta_1)$ whenever $\theta_2 > \theta_1$.

We now make the following assumptions on the density $f(x, \theta)$ (Weiss and Wolfowitz (1974) have similar assumptions when they study properties of maximum probability estimators).

(1) $A_1(\theta)$ and $A_2(\theta)$ are continuously differentiable functions of θ (if not infinity).

(2) On the set $\{(x, \theta) : x \in S(\theta)\}$, $f(x, \theta)$ is jointly continuous in (x, θ).

(3) The derivatives $\dfrac{\partial f(x, \theta)}{\partial\theta}$, $\dfrac{\partial^2 \log f(x, \theta)}{\partial\theta^2}$ exist for all (x, θ) in $\{(x, \theta) : A_1(\theta) < x < A_2(\theta)\}$.

(4) For all $\theta_0 \in \Theta$, there exists a neighbourhood $N(\theta_0)$ of θ_0 and a constant $D(\theta_0) > 0$ such that for $\theta \in N(\theta_0)$

$$\left|\frac{\partial^2 \log f(x, \theta)}{\partial\theta^2}\right| \leqslant D(\theta_0)$$

for all x for which the derivative exists,

(5) For all $\theta \in \Theta$, $E_\theta \dfrac{\partial \log f(X,\theta)}{\partial\theta} = c(\theta)$ is finite and not equal to zero.

In all the above non-regular cases we can obtain an asymptotic expansion of likelihood ratio $\frac{dP^n_{\theta_0+\lambda n^{-1}}}{dP^n_{\theta_0}}$ at any $\theta_0 \in \Theta$ and for all λ in an appropriate subset Λ of $\boldsymbol{R}$. Here P^n_θ is the n-fold product of the measure P_θ with density $f(x, \theta)$. For Case I, $\Lambda = [0, \infty)$ and for Case II, $\Lambda = (-\infty, 0]$. In either of the cases, for all $\theta_0 \in \Theta$ and $\lambda \in \Lambda$, $P^n_{\theta_0+\lambda n^{-1}}$ is absolutely continuous with respect to $P^n_{\theta_0}$.

Expanding at θ_0 by Taylor's theorem we get

$$\log \frac{dP^n_{\theta_0+\lambda n^{-1}}}{dP^n_{\theta_0}} = \lambda \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \log f(X_i, \theta)}{\partial \theta}\bigg|_{\theta_0} + \frac{\lambda^2}{2n^2} \sum_{i=1}^{n} \frac{\partial^2 \log f(X_i, \theta)}{\partial \theta^2}\bigg|_{\theta_n'(\boldsymbol{X})}$$

on

$$B_{n,\lambda} = \left\{X_i \in (A_1(\theta_0), A_2(\theta_0)) \cap \left(A_1\left(\theta_0+\frac{\lambda}{n}\right), A_2\left(\theta_0+\frac{\lambda}{n}\right)\right) \text{ for } i = 1, 2, \ldots, n\right\}$$

$$= \left\{X_i \in \left(A_1\left(\theta_0+\frac{\lambda}{n}\right), A_2\left(\theta_0+\frac{\lambda}{n}\right)\right) \text{ for } i = 1, 2, \ldots, n\right\}$$

(i.e., on the set where the Taylor's expansion is possible)

where $\qquad \theta_n'(\boldsymbol{X})$ lies between θ_0 and $\theta_0 + \frac{\lambda}{n}$,

and
$$\frac{dP^n_{\theta_0+\lambda n^{-1}}}{dP^n_{\theta_0}} = 0 \quad \text{a.e. } [P^n_{\theta_0}] \text{ on } B^c_{n,\lambda}$$

Also,
$$\Delta_n(\theta_0) = \frac{1}{n} \sum_{i=1}^{n} \frac{\partial \log f(X_i, \theta)}{\partial \theta}\bigg|_{\theta_0'} \to c(\theta_0) \text{ a.e. } [P_{\theta_0}]$$

by strong law of large numbers.

By assumption 4,

$$\frac{1}{n^2} \sum_{i=1}^{n} \frac{\partial^2 \log f(X_i, \theta)}{\partial \theta^2}\bigg|_{\theta_n'} \to 0 \text{ a.e. } [P_{\theta_0}]$$

The set $B_{n,\lambda}$ can be expressed as $\{n(Z_n(\boldsymbol{X})-\theta_0) > \lambda\}$ or $\{n(Z_n^*(\boldsymbol{X})-\theta_0) < \lambda\}$ but the form of Z_n or Z_n^* depends on $A_1(\theta)$ and $A_2(\theta)$.

We will now consider cases with different possible A_1, A_2.

Case I(*a*). $S(\theta)$ is an unbounded interval

$$\text{i.e.,}\quad S(\theta) = [A_1(\theta), \infty) \text{ or } S(\theta) = (-\infty, A_2(\theta)]$$

where A_1 is a monotonic nondecreasing function of θ and A_2 is a monotonic nonincreasing function of θ. For simplicity, let us first consider the simple case where $S(\theta) = [\theta, \infty)$.

$$\text{In this case, } B_{n,\lambda} = \left\{X_i \,\epsilon\left(\theta_0+\frac{\lambda}{n}, \infty\right) \text{ for } i = 1, 2, \ldots, n\right\}$$

$$= \{n(W_n-\theta_0) > \lambda\},\ \lambda \geqslant 0.$$

where $W_n = \min(X_1, X_2, \ldots, X_n)$. Thus the asymptotic expansion (2.1) holds. Also for any sequence $\{\theta_n\}$ satisfying $|n(\theta_n-\theta_0)| \leqslant C$ and for any $t > 0$,

$$P^n_{\theta_n}\{n(W_n-\theta_n) > t\} = \left[1-\int_{\theta_n}^{\theta_n+\frac{t}{n}} f(x, \theta_n)dx\right]^n$$

and

$$n\int_{\theta_n}^{\theta_n+\frac{t}{n}} f(x, \theta_n)dx \to f(\theta_0, \theta_0)t.$$

Thus assumption (A1) holds with $Z_n = W_n$ which is regular and $c(\theta) = f(\theta, \theta)$ and hence conditions of all the theorems in Section 2 are satisfied. For arbitrary A_1, A_2 we can define A_1^{-1}, A_2^{-1} as in case I(*b*) or II(*b*) and proceed in a similar manner.

Case I(*b*). $S(\theta) = [A_1(\theta), A_2(\theta)]$ with $A_1'(\theta) \geqslant 0$ and $A_2'(\theta) \leqslant 0$.
Here

$$B_{n,\lambda} = \left\{A_1\left(\theta_0+\frac{\lambda}{n}\right) < X_i < A_2\left(\theta_0+\frac{\lambda}{n}\right) \text{ for } i = 1, 2, \ldots, n\right\}$$

$$= \left\{W_n > A_1\left(\theta_0+\frac{\lambda}{n}\right),\ V_n < A_2\left(\theta_0+\frac{\lambda}{n}\right)\right\},\ \lambda \geqslant 0$$

where $W_n = \min(X_1, X_2, \ldots, X_n)$, $V_n = \max(X_1, X_2, \ldots, X_n)$.

If A_1, A_2 are strictly monotonic functions, they possess unique inverse A_1^{-1}, A_2^{-1} and $B_{n,\lambda}$ can be expressed as $\{n(Z_n-\theta_0) > \lambda\}$ with

$$Z_n = \min\{A_1^{-1}(W_n), A_2^{-1}(V_n)\}$$

Here $c(\theta) = A_1'(\theta)f(A_1(\theta), \theta) - A_2'(\theta)f(A_2(\theta), \theta) > 0$.

For arbitrary A_1, A_2 we define $A_1^{-1}(w) = \sup\{\theta : A_1(\theta) \leqslant w\}$
and $A_2^{-1}(v) = \sup\{\theta : A_2(\theta) \geqslant v\}$.

Then $$B'_{n,\lambda} = \left\{A_1\left(\theta_0+\frac{\lambda}{n}\right) \leqslant X_i \leqslant A_2\left(\theta_0+\frac{\lambda}{n}\right) \text{ for } i = 1, 2, \ldots, n\right\}$$

$$= \left\{W_n \geqslant A_1\left(\theta_0+\frac{\lambda}{n}\right), V_n \leqslant A_2\left(\theta_0+\frac{\lambda}{n}\right)\right\}$$

$$= \left\{A_1^{-1}(W_n) \geqslant \theta_0+\frac{\lambda}{n}, A_2^{-1}(V_n) \leqslant \theta_0+\frac{\lambda}{n}\right\}$$

$$= \{n(Z_n-\theta_0) \geqslant \lambda\} \text{ where } Z_n = \min\{A_1^{-1}(W_n), A_2^{-1}(V_n)\}$$

Thus the asymptotic expansion (2.1) holds a.e. $[P^n_{\theta_0}]$.

For arbitrary A_1, A_2, $c(\theta)$ may not be nonzero for all θ. We consider only the case where $c(\theta) > 0$ for all θ. If, for example, at least for one i, $A_i'(\theta) > 0$ for all θ, this condition is satisfied.

Now for any sequence $\{\theta_n\}$ satisfying $|n(\theta_n-\theta_0)| \leqslant C$ for any $C > 0$, and any $t \geqslant 0$,

$$P^n_{\theta_n}[n(Z_n-\theta_n) \geqslant t]$$

$$= \left[1-\int_{A_1(\theta_n)}^{A_1\left(\theta_n+\frac{t}{n}\right)} f(x, \theta_n)dx - \int_{A_2\left(\theta_n+\frac{t}{n}\right)}^{A_2(\theta_n)} f(x, \theta_n)dx\right]^n$$

$$\to e^{-c(\theta_0)t} \text{ as } n \to \infty,$$

because $$\lim_{n\to\infty} n\left[\int_{A_1(\theta_n)}^{A_1\left(\theta_n+\frac{t}{n}\right)} f(x, \theta_n)dx + \int_{A_2\left(\theta_n+\frac{t}{n}\right)}^{A_2(\theta_n)} f(x, \theta_n)dx\right]$$

$$= t\,A_1'(\theta_0)f(A_1(\theta_0), \theta_0) - t\,A_2'(\theta_0)f(A_2(\theta_0), \theta_0)$$

$$= t\,c(\theta_0).$$

Thus assumption (A1) and the assumption of regularity of Z_n hold and hence the conclusion of all the theorems in Section 2 hold.

Case II(a) : $S(\theta)$ is an unbounded interval, i.e., $S(\theta) = [A_1(\theta), \infty)$ or $S(\theta) = (-\infty, A_2(\theta)]$ and $A_1'(\theta) \leqslant 0$, $A_2'(\theta) \geqslant 0$ for all $\theta \in \Theta$.

Proceeding as in case I(a) we can prove that condition ($A2$) is satisfied for some Z_n which is regular

Case II(b) : $S(\theta) = [A_1(\theta), A_2(\theta)]$ with $A_1'(\theta) \leqslant 0$ and $A_2'(\theta) \geqslant 0$ for all $\theta \in \Theta$. Here $B_{n,\lambda} = \left\{ W_n > A_1\left(\theta_0+\frac{\lambda}{n}\right),\ V_n < A_2\left(\theta_0+\frac{\lambda}{n}\right)\right\}$, $\lambda \leqslant 0$,

where W_n, V_n are as defined earlier and

$$c(\theta) = A_1'(\theta)f(A_1(\theta), \theta) - A_2'(\theta)f(A_2(\theta), \theta) \leqslant 0 \text{ for all } \theta \in \Theta.$$

We consider only the cases where $c(\theta) < 0$ for all $\theta \in \Theta$. We define

$$A_1^{-1}(w) = \inf\{\theta : A_1(\theta) \leqslant w\}$$
$$A_2^{-1}(v) = \inf\{\theta : A_2(\theta) \geqslant v\}.$$

Then proceeding as in Case I(b) we can prove that condition $A(2)$ is satisfied with $Z_n = \max\{A_1^{-1}(W_n), A_2^{-1}(V_n)\}$ and this Z_n is regular.

3.2 *Regression type model.* We now consider an example where the observations $X_1, X_2, \ldots, X_n$ are independent but not identically distributed. We consider the model

$$X_t = g(t)\,\theta + e_t, \quad t = 1, 2, \ldots$$

where e_t's are i.i.d. random variables having a common density $f(x)$ such that $f(x) > 0$ for $x \geqslant 0$ and $f(x) = 0$ for $x < 0$, and $g(t)$, $t = 1, 2, \ldots$ are values of a non-stochastic variable. We consider only the case where $g(t)$'s are positive. Let $K_n = \sum_{t=1}^{n} g(t)$. We make the following assumptions :

R1 $f(x)$ is continuous on $[0, \infty)$ and twice differentiable on $(0, \infty)$.

R2 (*a*) $\int |(\log f)'(x)| f(x)dx < \infty$

(*b*) $\int |(\log f)''(x)| f(x)dx < \infty$

R3 For all $\lambda \geqslant 0$,

$$\frac{1}{K_n^2}\sum_{t=1}^{n} g^2(t) \sup\left\{ |(\log f)''(e_t+\alpha)-(\log f)''(e_t)| : 0 \leqslant \alpha \leqslant \lambda \max_{1\leqslant t\leqslant n} g(t)K_n^{-1}\right\}$$

converges to 0 in probability.

R4 As $n \to \infty$

(*a*) $\max_{1\leqslant t\leqslant n} g(t) \Big/ \sum_{t=1}^{n} g(t) \to 0$

and (*b*) $K_n^{-2}\sum_{t=1}^{n} g^2(t) \to 0.$

Assumption R4 is satisfied if, for example, we take $g(t) = t$ or any polynomial in t. Assumption R3 is satisfied for almost all the usual cases.

Fix $\theta_0 \in \Theta$, the parameter space. Let P_θ^n be the joint probability distribution of $X_1, \ldots, X_n$ under θ. Expanding at θ_0 by Taylor's theorem we get for all $\lambda \geqslant 0$,

$$\log \frac{dP^n_{\theta_0+\lambda K_n^{-1}}}{dP^n_{\theta_0}} = \frac{\lambda}{K_n} \sum_{t=1}^{n} (-g(t)) (\log f)'(X_t - g(t)\theta_0)$$

$$+ \frac{\lambda^2}{2K_n^2} \sum_{t=1}^{n} g^2(t) (\log f)''(X_t - g(t)\theta_n')$$

$$= \lambda \Delta_n + \epsilon_n, \text{ say}$$

$$\text{on } B_{n,\lambda} = \{X_t > g(t)(\theta_0 + \lambda K_n^{-1}), t = 1, 2, \ldots, n\}$$

$$= \left\{ K_n \left(\min_{1 \leqslant t \leqslant n} \frac{X_t}{g(t)} - \theta_0 \right) > \lambda \right\}$$

where θ_n' lies between θ_0 and $\theta_0 + \lambda K_n^{-1}$, and

$$\frac{dP^n_{\theta_0+\lambda K_n^{-1}}}{dP^n_{\theta_0}} = 0 \text{ a.e. } [P^n_{\theta_0}] \text{ on } B^c_{n,\lambda}$$

We will now verify the following :

(*A*) $\Delta_n \xrightarrow{P^n_{\theta_0}} f(0)$

(*B*) $\epsilon_n \xrightarrow{P^n_{\theta_0}} 0$

(*C*) $P^n_{\theta_0}(B_{n,\lambda}) \to e^{-\lambda f(0)}$ for all $\lambda \geqslant 0$.

where Δ_n, ϵ_n and $B_{n,\lambda}$ are as above.

(A) follows from condition R2(a), the law of large numbers for weighted average (see, for example, Jamison, Orey and Pruitt (1965)), condition R4(b) and the fact that $-\int (\log f)'(x)f(x)dx = f(0)$.

Condition R2(b) implies that

$$\frac{1}{K_n^2} \Sigma\, g^2(t) (\log f)'' (X_t - g(t)\,\theta_0) \xrightarrow{P^n_{\theta_0}} 0.$$

(B) now follows from condition R3.

To prove (C) we use the following result :

Lemma : *Consider a double sequence of real numbers* $\{a_{in}\}$. *If* (*i*) $\sup_{1 \leqslant i \leqslant n} |a_{in}| \to 0$, (*ii*) $\sum_{i=1}^{n} |a_{in}|$ *is bounded and* (*iii*) $\sum_{i=1}^{n} a_{in} \to a$.

Then $\prod_{i=1}^{n} (1-a_{in}) \to e^{-a}$.

Now, $P^n_{\theta_0}(B_{n,k}) = \prod_{t=1}^{n} \left[1-F\left(\frac{g(t)u}{\Sigma g(t)}\right)\right]$ where F is the distribution function for f. Using continuity of f at 0 and condition R4(a) we can prove that

$$\sum_{t=1}^{n} F\left(\frac{g(t)}{\Sigma g(t)}\, u\right) - u\, f(0) \to 0.$$

Thus (C) is verified to be true.

Also the random variable $Z_n = \min_{1 \leqslant t \leqslant n} X_t/g(t)$ is obviously regular since in this case the distribution of $Z_n - \theta$ does not depend on θ.

4. Asymptotic properties of maximum probability estimators

Weiss and Wolfowitz (1974) studied the efficiency of maximum probability estimators (m.p.e.) for many non-regular cases. They also considered a a general case and indeed proved that the m.p. estimator is LAM under certain reasonable assumptions. In this section we will first prove the same result for the above family of non-regular cases (given in Section 3) by showing that the lower bound to the local asymptotic minimax risk is attained by the m.p. estimators. We will consider only $0-1$ loss functions :

$$L(x) = \begin{cases} 0 & \text{if } |x| \leqslant r \\ 1 & \text{othersise} \end{cases} \qquad \ldots \quad (4.1)$$

where r is some positive number.

For all the nonregular cases given in Section 3 the set on which the joint density of the observations $X_1, \ldots, X_n$ under θ is positive can be expressed as either (*a*) $\{\mathbf{X}_n : Z_n \geqslant \theta\}$ or (*b*) $\{\mathbf{X}_n : Z^*_n \leqslant \theta\}$. Proceeding as in Weiss and Wolfowitz (1974) we can find statistics $\tilde{\theta}_n$ which are asymptotically "equivalent" (see discussion following (3.4) in Weiss and Wolfowitz (1974)) to the m.p. estimators.

$$\text{For Case } (a), \quad \tilde{\theta}_n = Z_n - r\, K_n^{-1},$$

and $$\text{for Case } (b), \quad \tilde{\theta}_n = Z^*_n + r\, K_n^{-1}$$

where K_n is the normalizing factor.

We will consider only case (a). For case (a), $\int_0^\infty L(x-b)c(\theta_0)e^{-c(\theta_0)x}\,dx$ is minimized at $b = r$. Thus, using results of Section 3, for all $\theta_0 \in \Theta$ and all $A > 0$,

$$\lim_{n\to\infty} \sup_{|\theta-\theta_0| < AK_n^{-1}} E_\theta L[K_n(Z_n - rK_n^{-1} - \theta)]$$

$$= \int_0^\infty L(x-r)c(\theta_0)e^{-c(\theta_0)x}\,dx$$

$$= \int_{2r}^\infty c(\theta_0)e^{-c(\theta_0)x}\,dx$$

and hence the estimator $\tilde{\theta}_n$ is LAM. The treatment of Case (b) is similar.

We shall now prove that the m.p.e. $\bar{\theta}_n(r)$, if it exists, is equivalent to the estimator $\tilde{\theta}_n (= Z_n - rK_n^{-1})$ suggested in Section 2 in the sense that their difference converges to zero in probability (see Theorem 5, below).

We consider the set up of Section 2. Let $f_n(x, \theta)$ be the density of P_θ^n with respect to some dominating σ-finite measure on $\mathscr{A}^n$. We consider only Case (a) and assume that the following condition holds a.s. $[P_{\theta_0}^n]$

(A1)* For any $\lambda \in \mathbf{R}$,

$$L_n(\lambda) = \begin{cases} \exp \lambda\Delta_n(\theta_0) + \epsilon_n(\lambda, \theta_0), & \text{if } K_n(Z_n - \theta_0) > \lambda, \\ 0, & \text{if } K_n(Z_n - \theta_0) < \lambda, \end{cases}$$

where
$$L_n(\lambda) = \frac{dP^n_{\theta_0 + \lambda K_n^{-1}}}{dP^n_{\theta_0}} \text{ for } \lambda \in \mathbf{R},$$

Z_n is a random variable satisfying

$$Z_n \geqslant \theta_0 \text{ a.s. } [P^n_{\theta_0}]$$

and
$$\lim_{n\to\infty} P^n_{\theta_0}[K_n(Z_n - \theta_0) > t] = e^{-tc(\theta_0)} \text{ for all } t \geqslant 0,$$

$$\Delta_n(\theta_0) \xrightarrow{P^n_{\theta_0}} c(\theta_0) \text{ for some } c(\theta_0) > 0$$

and
$$\epsilon_n \xrightarrow{P^n_{\theta_0}} 0.$$

Thus under assumption (A1)*, for all $\lambda > 0$,

$$L_n(\lambda) = \exp\{\lambda\,\Delta_n(\theta_0) + c_n(\lambda, \theta_0)\} \quad \text{a.s.} \quad [P^n_{\theta_0}]$$

and hence for all $\lambda < 0$, $L_n(\lambda) \xrightarrow{P^n_{\theta_0}} e^{\lambda c(\theta_0)}$

We here assume that

$$(B1) \quad E_{\theta_0}(L_n(\lambda)) \to e^{\lambda c(\theta_0)} \quad \text{for all } \lambda < 0.$$

It is to be noted that the above conditions hold for all the non-regular cases considered in Section 3.

Now, the maximum probability estimator $\bar{\theta}_n(r)$ with respect to the loss function (4.1) is defined as that value of d for which the integral

$$\int f_n(\mathbf{X}_n, \theta)d\theta$$

over the set $[d - rK_n^{-1},\ d + rK_n^{-1}]$ is a maximum. Here $\mathbf{X}_n$ denotes the observation at the n^{th} stage. We assume that $f_n(x, \theta)$ is jointly measurable in (x, θ) and a measurable m.p.e. $\bar{\theta}_n(r)$ exists.

Theorem 5 : *Suppose that the sequence* $\{K_n(\bar{\theta}_n - \theta_0)\}$ *is relatively compact for* $\{P^n_{\theta_0}\}$. *Then under assumptions* (A1)* *and* (B1),

$$K_n(\bar{\theta}_n - \theta_0) - K_n(Z_n - rK_n^{-1} - \theta_0) \xrightarrow{P^n_{\theta_0}} 0 \text{ as } n \to \infty.$$

To prove this theorem we need the following lemma.

Lemma : *Set for* $\lambda \in \mathbf{R}$

$$L_n^*(\lambda) = \begin{cases} e^{c(\theta_0)\lambda}, & \text{if } K_n(Z_n - \theta_0) > \lambda \\ 0, & \text{if } K_n(Z_n - \theta_0) < \lambda. \end{cases}$$

Then for any $\lambda \in \mathbf{R}$,

$$E_{\theta_0}\,|L_n(\lambda) - L_n^*(\lambda)| \to 0 \text{ as } n \to \infty.$$

Proof of the lemma : The result follows from the fact that

$$|L_n(\lambda) - L_n^*(\lambda)| \xrightarrow{P^n_{\theta_0}} 0$$

and $L_n(\lambda)$ and $L_n^*(\lambda)$ are uniformly integrable.

This is proved in Ibragimov and Hasminskii (1981) for all $\lambda \geqslant 0$. Using condition (B1) it can be proved for all $\lambda < 0$ in a similar maner.

Proof of Theorem 5: We use the idea of the proof of Theorem 4 in Jogenathan (1982). We shall prove that for all $\delta > 0$,

$$\lim_{n\to\infty} P^n_{\theta_0} [|K_n(\bar{\theta}_n-\theta_0)-K_n(Z_n-rK_n^{-1}-\theta_0)| > \delta] = 0.$$

Given $\epsilon > 0$, we choose $K > 0$ sufficiently large such that for all n

$$P^n_{\theta_0} [|K_n(\bar{\theta}_n-\theta_0)| > K-r] < \frac{\epsilon}{4}$$

and
$$P^n_{\theta_0} [|K_n(Z_n-rK_n^{-1}-\theta_0)| > K-r] < \frac{\epsilon}{4}.$$

Thus it is enough to prove that for all sufficiently large n,

$$P^n_{\theta_0} (A_n) < \frac{\epsilon}{2} \qquad \dots \quad (4.2)$$

where
$$A_n = \{|K_n(\bar{\theta}_n-\theta_0)-K_n(Z_n-rK_n^{-1}-\theta_0)| > \delta,$$
$$|K_n(\bar{\theta}_n-\theta_0)| \leqslant K-r,\ |K_n(Z_n-rK_n^{-1}-\theta_0)| \leqslant K-r\}.$$

Since
$$E_{\theta_0}|L_n(\lambda)-L_n^*(\lambda)| \leqslant 1+e^{c(\theta_0)\lambda}$$

the above lemma implies that

$$\int_{-K}^{K} E_{\theta_0}|L_n(\lambda)-L_n^*(\lambda)|\,d\lambda \to 0.$$

Therefore,
$$E_{\theta_0}\left[\int_{-K}^{K}|L_n(\lambda)-L_n^*(\lambda)|\,d\lambda\right] \to 0. \qquad \dots \quad (4.3)$$

Now, if we set

$$B_1 = [K_n(\bar{\theta}_n-\theta_0)-r,\ K_n(\bar{\theta}_n-\theta_0)+r]$$

and
$$B_2 = [K_n(Z_n-rK_n^{-1}-\theta_0)-r,\ K_n(Z_n-rK_n^{-1}-\theta_0)+r]$$

we have $B_1 \subset [-K, K]$, $B_2 \subset [-K, K]$ whenever A_n occurs and hence (4.3) implies that

$$\int_{A_n}\int_{B_i} |L_n(\lambda)-L_n^*(\lambda)|\,d\lambda\, dP^n_{\theta_0} \to 0 \quad \text{for } i = 1, 2. \qquad \dots \quad (4.4)$$

Now suppose that (4.2) is not true. Then

$$P^n_{\theta_0}(A_n) \geqslant \frac{\epsilon}{2} \qquad \dots \quad (4.5)$$

for infinitely many values of n.

From the definition of $L_n^*(\lambda)$ it can be shown that when the event A_n occurs we have

$$a+\int_{B_1} L_n^*(\lambda)d\lambda < \int_{B_2} L_n^*(\lambda)d\lambda$$

where a is a positive real number not depending on n.

Then (4.5) implies that for some $a_0 > 0$

$$a_0 + \int_{A_n} \int_{B_1} L_n^*(\lambda) d\lambda < \int_{A_n} \int_{B_2} L_n^*(\lambda) d\lambda$$

for infinitely many values of n. This together with (4.4) implies that

$$\int_{A_n} \int_{B_1} L_n(\lambda) d\lambda < \int_{A_n} \int_{B_2} L_n(\lambda) d\lambda \qquad \dots \quad (4.6)$$

for infinitely many values of n.

On the other hand, from the definition of m.p.e

$$\int_{B_1} \frac{f_n(X_n, \theta_0 + \lambda K_n^{-1})}{f_n(X_n, \theta_0)} d\lambda \geqslant \int_{B_2} \frac{f_n(X_n, \theta_0 + \lambda K_n^{-1})}{f_n(X_n, \theta_0)} d\lambda$$

i.e.,

$$\int_{A_n} \int_{B_1} L_n(\lambda) d\lambda \geqslant \int_{A_n} \int_{B_2} L_n(\lambda) d\lambda$$

for all n, contradicting (4.6). Thus (4.2) is true and hence the theorem is proved.

Remark : The result (Theorem 5) can also be proved for any loss function of the form

$$\begin{aligned} L(x) = L(|x|) &= M, \text{ if } |x| > r, \\ &\leqslant M, \text{ if } |x| \leqslant r, \end{aligned}$$

for any $M, r > 0$. The maximum probability estimate for such a loss function is defined to be that value of d for which the integral

$$\int [M - L(K_n(d - \theta))] fn(X_n, \theta) d\theta$$

is a maximum. The proof follows the same lines as the proof of Theorem 5.

5. Asymptotic properties of Bayes estimators

The asymptotic properties of Bayes estimators were studied in Ibragimov and Hasminskii (1981) for a large family of non-regular cases when the observations are independently and identically distributed. In this section we consider the regression model of Section 3 and using a general result on the asymptotic behaviour of the Bayes estimators (Theorem I.10.2 in Ibragimov and Hasminskii (1981, Ch. I)) we prove the efficiency of the Bayes estimators. For this we make the following assumptions in addition to the assumptions R1—R4 made in Section 3.

R5 There exist constants a, M_1, $M_2 > 0$ such that for all $x \geqslant 0$

$$f(x) \leqslant M_1 + M_2 x^a$$

R6 There exists a constant $C^* > 0$ such that for all $n \geqslant 1$

$$\left[\prod_{t=1}^{n} g(t)\right]^{1/n} \Big/ \max_{1 \leqslant t \leqslant n} g(t) \geqslant C^*.$$

(Condition R6 is satisfied, for example, when $g(t)$ is some polynomial in t)

When the parameter set Θ is unbounded we make the following assumption :

R7 $$\int f^{1/2}(x-h) f^{1/2}(x) dx \leqslant [C_1 |h|]^{-\alpha}$$

for all h and for some $\alpha > 0$, $C_1 > 0$.

We now consider the family $\{\tilde{\theta}_n\}$ of Bayes estimators with respect to the loss function $L(K_n^{-1}(\theta - a))$ and some prior density q. We assume that L is a subconvex loss function possessing a polynomial majorant and satisfying the following condition.

(C) $\varphi(b) = \int_0^\infty L(x-b) f(0) e^{-f(0)x} dx$ is finite for some b and attains its minimum at the unique point b_0.

Let Q be the set of continuous positive functions on $\boldsymbol{R}$ possessing a polynomial majoriant.

Theorem 6 : *Let $\tilde{\theta}_n$ be a Bayes estimator with respect to a prior density $q \in Q$ and the loss function $L(K_n^{-1}(\theta - a))$, where L is a continuous subconvex function possessing a polynomial majorant and satisfying condition (C). Then under conditions R1—R7, the Bayes estimator $\tilde{\theta}_n$ is asymptotically efficient for estimating θ in the sense that uniformly in θ belonging to any compact subset of Θ,*

$$\lim_{n \to \infty} E_\theta L[K_n(\tilde{\theta}_n - \theta)] = \int_0^\infty L(x - b_0) f(0) e^{-f(0)x} dx$$

where the right hand side is the lower bound to the asymptotic risk of an estimator obtained in Theorem 2.

Proof : We verify all the conditions of Theorem I.10.2 of Ibragimov and Hasminskii (1981). Fix some $\theta \in \Theta$. For $u \in \boldsymbol{R}$ we set

$$\Lambda_{n,\theta}(u) = \frac{\prod_{t=1}^{n} f(X_t - g(t)\theta - g(t) K_n^{-1} u)}{\prod_{t=1}^{n} f(X_t - g(t).\theta)}.$$

First of all note that the marginal distributions of the process $\Lambda_{n,\theta}(u)$ do not depend on θ. Then for any $u_1 < u_2$

$$E_\theta[\Lambda^{1/2}_{n,\theta}(u_2)-\Lambda^{1/2}_{n,\theta}(u_1)]^2$$

$$\leqslant 2\left[1-\prod_{t=1}^{n}\int f^{1/2}(x_t-g(t)K_n^{-1}u_2)f^{1/2}(x_t-g(t)K_n^{-1}u_1)dx_t\right]$$

$$\leqslant 2\sum_{t=1}^{n}[1-\int f^{1/2}(x-g(t)K_n^{-1}u_2)f^{1/2}(x-g(t)K_n^{-1}u_1)dx]$$

$$\left[\text{since for } 0 \leqslant \rho_1, \rho_2, \ldots, \rho_n \leqslant 1,\ 1-\rho_1\rho_2\ldots\rho_n \leqslant \sum_{t=1}^{n}(1-\rho_t)\right]$$

$$= \sum_{t=1}^{n}\int[f^{1/2}(x-g(t)K_n^{-1}u_2)-f^{1/2}(x-g(t)K_n^{-1}u_1)]^2dx \qquad \ldots \quad (5.1)$$

Now, $$\int[f^{1/2}(x-g(t)K_n^{-1}u_2)-f^{1/2}(x-g(t)K_n^{-1}u_1)]^2dx$$

$$\leqslant \int|f(x-g(t)K_n^{-1}u_2)-f^{1/2}(x-g(t)K_n^{-1}u_1)|dx$$

$$[\text{since for any } \alpha, \beta > 0,\ (\sqrt{\alpha}-\sqrt{\beta})^2 \leqslant |\alpha-\beta|]$$

$$= \int_{g(t)K_n^{-1}u_1}^{g(t)K_n^{-1}u_2} f(x-g(t)K_n^{-1}u_1)dx + \int_{g(t)K_n^{-1}u_2}^{\infty}\left|\int_{g(t)K_n^{-1}u_1}^{g(t)K_n^{-1}u_2} f'(x-s)ds\right|dx$$

$$= I_1+I_2, \text{ say.}$$

$$I_2 \leqslant \int_{g(t)K_n^{-1}u_1}^{g(t)K_n^{-1}u_2}\left\{\int_{g(t)K_n^{-1}u_2-s}^{\infty}|f'(x)|dx\right\}ds$$

$$\leqslant g(t)K_n^{-1}(u_2-u_1)\int_0^\infty |f'(x)|dx.$$

$$= g(t)K_n^{-1}(u_2-u_1)\, M, \text{ say}$$

where $M = \int_0^\infty |f'(x)|dx < \infty$ by assumption $R2\,(a)$.

By assumption R5, for all $u_1 < u_2$ such that $|u_1| \leqslant R$, $|u_2| \leqslant R$ we have

$$I_1 \leqslant g(t)K_n^{-1}(u_2-u_1)[M_1+M_2(2R)^a].$$

Therefore, from (5.1)

$$E_\theta[\Lambda^{1/2}_{n,\theta}(u_2)-\Lambda^{1/2}_{n,\theta}(u_1)]^2 \leqslant (u_2-u_1)[M+M_1+M_2 2^a R^a]$$

i.e., $$\sup_{|u_1|<R,|u_2|<R} |u_2-u_1|^{-1} E_\theta|\Lambda^{1/2}_{n,\theta}(u_2)-\Lambda^{1/2}_{n,\theta}(u_1)|^2 \leqslant B(1+R^a)$$

for some $B > 0$ and for all $\theta \in \Theta$.

Thus condition (1.1) of Theorem I.10.2 of Ibragimov and Hasminskii (1981) is satisfied.

Now,

$$E_\theta \Lambda_{n,\theta}^{1/2}(u) = \prod_{t=1}^{n} \left\{1-\frac{1}{2}\int |f^{1/2}(x-g(t)K_n^{-1}u)-f^{1/2}(x)|^2 dx\right\}$$

$$\leqslant \exp\left\{-\frac{1}{2}\sum_{t=1}^{n}\int |f^{1/2}(x-g(t)K_n^{-1}u)-f^{1/2}(x)|^2\,dx\right\}$$

[since $1-\rho \leqslant e^{-\rho}$].

We choose $A > 0$ sufficiently small such that whenever $0 \leqslant x \leqslant A$ we have $f(x) \geqslant \frac{1}{2} f(0)$.

For $u \geqslant 0$, $\int |f^{1/2}(x-g(t)K_n^{-1}u)-f^{1/2}(x)|^2\;dx$

$$\geqslant \int_0^{g(t)K_n^{-1}u} f(x)\;dx$$

and for $u \leqslant 0$, $\int |f^{1/2}(x-g(t)K_n^{-1}(u) - f^{1/2}(x)|^2\;dx$

$$\geqslant \int_{-g(t)K_n^{-1}|u|}^{0} |f^{1/2}(x+g(t)K_n^{-1}|u|)-f^{1/2}(x)|^2\,dx.$$

$$= \int_{-g(t)K_n^{-1}|u|}^{0} f(x+g(t)K_n^{-1}|u|)\,dx.$$

Thus for $\max_{1\leqslant t\leqslant n} g(t)K_n^{-1}|u| \leqslant A$ we have

$$E_\theta \Lambda_{n,\theta}^{1/2}(u) \leqslant \exp\left\{-\frac{1}{4}f(0)|u|\right\}. \qquad \ldots \quad (5.2)$$

Also by assumption R7, for all $u \in \boldsymbol{R}$,

$$E_\theta \Lambda_{n,\theta}^{1/2}(u) \leqslant [C_1|u|]^{-n\alpha}\left[\left(\prod_{t=1}^{n} g(t)\right)^{1/n} K_n^{-1}\right]^{-n\alpha} \qquad \ldots \quad (5.3)$$

Fix any $r > 0$. We want to prove that

$$\lim_{\substack{|u|\to\infty \\ n\to\infty}} |u|^r\, E_\theta \Lambda_{n,\theta}^{1/2}(u) = 0. \qquad \ldots \quad (5.4)$$

From (5.3), for $\max_{1\leqslant t\leqslant n} g(t)K_n^{-1}|u| > A$,

$$|u|^r E_\theta \Lambda_{n,u}^{1/2}(u)$$

$$\leqslant |u|^r \left[C_1|u|K_n^{-1}\max_{1\leqslant t\leqslant n} g(t)\right]^{-n\alpha}\left[(\Pi g(t))^{1/n}/\max_{1\leqslant t\leqslant n} g(t)\right]^{-n\alpha}$$

$$= C_1^{-r}\left[C_1|u|K_n^{-1}\max_{1\leqslant t\leqslant n} g(t)\right]^{-n\alpha+r}\left[K_n/\max_{1\leqslant t\leqslant n} g(t)\right]^r\left[(\Pi g(t))^{1/n}/\max_{1\leqslant t\leqslant n} g(t)\right]^{-n\alpha}$$

$$< C_1^{-r}(C_1A)^{-n\alpha+r}\left[K_n/\max_{1\leqslant t\leqslant n} g(t)\right]^r\left[\left(\prod_1^n g(t)\right)^{1/n}/\max_{1\leqslant t\leqslant n} g(t)\right]^{-n\alpha}$$

(we choose n so large that $-n\alpha+r < 0$)

$$= A^r \frac{n^r}{\left[C_1A\dfrac{(\Pi g(t))^{1/n}}{\max\limits_{1\leqslant t\leqslant n} g(t)}\right]^{n\alpha}} \quad \left[\text{since } K_n \leqslant n.\max_{1\leqslant t\leqslant n} g(t)\right]$$

and this converges to 0 (as $n\to\infty$) by assumption $R6$. This result and (5·2) give (5.4).

Now proceeding as in Section 3 we can express $\Lambda_{n,\theta}(u)$ for all $u \in \boldsymbol{R}$ as

$$\Lambda_{n,\theta}(u) = \begin{cases} \exp\{f(0)u+\epsilon_n\} & \text{if } \tau_n > u, \\ 0, & \text{if } \tau_n \leqslant u, \end{cases}$$

where $\epsilon_n \xrightarrow{P_\theta} 0$ and τ_n is a random variable converging in distribution to a random variable τ with density $f(0)e^{-f(0)x}$ on $(0, \infty)$. This is proved for all $u \geqslant 0$ in Section 3. The proof for $u < 0$ is similar to that for the case $u \geqslant 0$. Then it can be easily shown that the marginal distributions of the process $\Lambda_{n,\theta}(u)$, $u \in \boldsymbol{R}$ converge to the marginal distributions of the process

$$\Lambda(u) = \begin{cases} e^{f(0)u}, & \text{if } \tau > u, \\ 0, & \text{if } \tau \leqslant u. \end{cases}$$

Also the random function

$$\psi(s) = \int L(s-u)\,\Lambda(u)du$$

attains its minimum value at the unique point $s = \tau - b_0$. Thus all the conditions of Theorem I.10.2 of Ibragimov and Hasminskii (1981, Ch. I) are satisfied and Theorem 6 is proved.

Acknowledgement. The author is indebted to Professor J. K. Ghosh for suggesting the problem and to Professor K. K. Roy for his guidance. He is also grateful to the referee for useful comments and suggestions which improved the paper.

References

FABIAN, V. and HANNAN, J. (1982) : On estimation and adaptive estimation for locally asymptotically normal families. *Z. Wahr*, **59**, 459-478.

FERGUSON, T. S. (1967) : *Mathematical Statistics*, Academic Press.

GHOSH, J. K. (1985) : Efficiency of estimates—Part I, *Sankhyā*, **47**, Ser. A. 310-325.

HÁJEK, J. (1972) : Local asymptotic minimax and admissibility in estimation, *Proc. 6th Berkeley Symposium*, Vol. I, 175-194 (L. M. LeCam, J. Neyman and E. J. Scott, eds.) University of California Press, Berkeley and Los Angeles.

HÁJEK, J. and SIDAK, Z. (1967) : *Theory of Rank Tests*, Academic Press, New York.

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1981) : *Statistical Estimation*, Springer-Verlag, New York.

JAMISON, B., OREY, S., and PRUITT, W. (1965) : Convergence of weighted averages of independent random variables. *Z. Wahr*, **4**, 40–44.

JEGANATHAN, P. (1982) : On the asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. *Sankhyā*, **44**, Ser. A, 173-212.

MILLAR, P. W. (1983) : *The Minimax Principles in Asymptotic Statistical Theory*. Ecole d'Ete de Probabilites de Saint-Flour XI (P. L. Hennequin Ed.) (Lecture Notes in Mathematics). Springer-Verlag, Berlin.

WALD, A. and WOLFOWITZ, J. (1951) : Two methods of randomization in statistics and the theory of games. *Ann. Math.* **53**, 581-586.

WEISS, L. and WOLFOWITZ J. (1974) : Maximum probability estimators and related topics