Sankhyā : The Indian Journal of Statistics
1982, Volume 44, Series A, Pt. 2 pp. 213–232.

ON THE CONVERGENCE OF MOMENTS OF STATISTICAL ESTIMATORS

By P. JEGANATHAN
Indian Statistical Institute

SUMMARY. It is shown that when the observations are dependent, including the situations where the limit distributions of estimators are mixed normal, and when the parameter space is of multidimensional, the moments of any order of maximum likelihood estimators, maximum probability estimators and a certain class of Bayes estimators converge to the corresponding moments of a multivariate mixed normal distribution. As a by-product of the given assumptions it is also shown that the sequence of properly normalised likelihood functions, considered as a sequence of random processes in the parameter space, converges weakly to a mixed Gaussian shift process.

1. Introduction

In an important paper Ibragimov and Khasminskii (1972 and 1973) (henceforth this paper will be referred to in short by I.K.) considered the asymptotic behaviour of maximum likelihood estimators (MLE) and a certain class of Bayes estimators when the observations are i.i.d. and when the parameter space is a subset of the real line. Among other things, they proved that the moments of any order of the above mentioned estimators converge to the corresponding moments of a normal distribution. Their investigations are based on a general method which amounts to treat the likelihood function as a random function of the parameter; we would like to mention here that this general method of investigating MLE was first developed in Rubin (1961), and then later utilised in Prakasa Rao (1968) where the weak convergence results for random functions were further employed for the first time for investigating MLE in non-regular cases. A similar method of investigating MLE was also developed by LeCam (1970). This approach offered some fresh insights into the problems and as a result I.K. were able to prove powerful results under quite general assumptions. However, so far as the practical purposes are considered, the situation considered by I.K. was not quite general in the sense that the parameter space was assumed to be a subset of the real line. Though the methods of analysis of I.K. were simple, some of the arguments depend in a very crucial manner on the dimension of the parameter space. For example, they invoke Kolmogorov's sufficient condition for the continuity of a random process to get some estimates of the continuity modulus of the realisations of the likelihood function (see Prokhorov, 1956, p. 180). However, it does not seem to be possible to extend this idea to more than one dimension. In fact LeCam (1970) mentions that some of the arguments about continuity of sample paths for random processes do not extend to more than one-dimension. At the same time there is no reason to suppose that the results on the convergence of moments of statistical estimators would depend on the dimensionality restriction of the parameter space. This suggests that one can obtain the same type of estimates of the continuity modulus by some other methods whose arguments would not depend on the dimension of the parameter space. It is also important to know how far the results on the convergence of moments can be extended to the situation where the observations are not necessarily i.i.d. Thus our aim in this paper is to prove that the moments of any order of MLE, maximum probability estimators (MPE) and a certain class of Bayes estimators converge to the corresponding moments of a mixed normal distribution when the observations are not necessarily i.i.d., including the situations where the limit distributions are mixed normal (see, e.g. Jeganathan, 1979), and the parameter space is a subset of R^k, $k \geqslant 1$; as a by product we also present a weak convergence result for the likelihood ratio random processes.

Majority of the ideas of this paper are either inspired by or adapted from I.K. though we have substantially simplified the proofs. Also we have avoided using weak convergence results for random processes, which I.K. use freely in their paper. We would like to point out that the weak convergence results for the likelihood ratio process have been extended by Inagaki and Ogata (1975 and 1977) to the situations where the parameter space is of multidimension and the observations are from a strictly stationary Markov process, with a number of interesting applications. When the parameter space is a subset of the real line, the results of I.K. on the convergence of moments have been extended to the independent not necessarily identically distributed case by Ibragimov and Khasminskii (1975) and to a certain class of Markov chains by Levit (1974).

In Section 2 we introduce the Assumption (A.1)–(A.11). (A.10) of this section is very direct. The reason is that we are not able to impose satisfactory conditions on the densities implying this assumption in the general case. However, it is possible to verify this assumption directly in some problems (for example, for a certain class of mixed Gaussian processes). In the situations where the observations have a certain 'mild' form of dependence it is possible to impose conditions on the densities implying this assumption. These things are done in Section 5. In section 3 we obtain some preliminary results on the behaviour of the likelihood function which are used in Section 4 where the results on the convergence of moments of MLE, MPE and a certain class of Bayes estimators are obtained. In Section 3 we also show, as a by-product of our assumptions, that the likelihood functions belong to a certain complete separable metric space for all sufficiently large sample size with probability one and the corresponding sequence of induced probability measures on this metric space converge weakly to the probability measure induced by a mixed Gaussian shift process.

A refree has drawn our attention to a treatment of the multidimensional case in a new book by Ibragimov and Khaiminskii : "Asymptotic Estimation Theory", Moscow, 1979, (in Russian)". Unfortunately we were not able to compare our results with those of this book which was not available to us even at the time of the final revision.

2. Assumptions

Let $(X_1, X_2, \dots, X_n)$, $n \geqslant 1$, be a sequence of random vectors defined on a probability space $(\mathscr{X}, \mathscr{A}, P_\theta)$ where the k-dimensional parameter $\theta \in \Theta$, an open subset of R^k, $k \geqslant 1$. Let $\mathscr{A}_n = \sigma(X_1, \dots, X_n)$ be the σ-field induced by the random vector $(X_1, \dots, X_n)$ and $P_{\theta, n}$ be the restriction of P_θ to $\mathscr{A}_n$. Let $\theta_0 \in \Theta$ be the 'true' value of the parameter. We further assume that, for $j \geqslant 2$, a regular conditional probability measure of X_j given $(X_1, \dots, X_{j-1})$ is absolutely continuous with respect to a σ-finite measure μ_j with a corresponding density $f_j(X_j \mid X_1, \dots, X_{j-1}; \theta)$, and the probability measure of X_1 is absolutely continuous with respect to a σ-finite measure μ_1 with a corresponding density $f_1(X_1; \theta)$. For the sake of simplicity we set $f_j(X_j \mid X_1, \dots, X_{j-1}; \theta) = f_j(\theta)$, $j \geqslant 2$, and $f_1(X_1; \theta) = f_1(\theta)$.

In what follows, unless otherwise specified, all the probability concepts and expectations are with respect to P_{θ_0}.

(A.1) : For all $(X_1, \dots, X_j)$ and for every $j \geqslant 1$ the functions $\theta \to f_j(\theta)$, $j \geqslant 1$, are absolutely continuous in θ.

(A.2) : For $\mu_1 \times \dots \times \mu_j$ almost all $(X_1, \dots, X_j)$ and for every $j \geqslant 1$, the functions $\theta \to \log f_j(\theta)$ are differentiable in θ.

Remark : Note that implicit in (A.2) is the assumption that, for $\mu_1 \times \dots \times \mu_j$ almost all $(X_1, \dots, X_j)$ and for every $j \geqslant 1$, $\log f_j(\theta)$ is finite for all θ.

Set

$$\eta_j(\theta) = \begin{cases} (\partial/\partial\theta) \log f_j(\theta) & \text{if the derivative exists,} \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that we have selected a suitable sequence $\{\delta_n\}$ of normalising matrices such that $\|\delta_n\| \to 0$; one way of selection is to set

$$\delta_n'\delta_n = \Big[\sum_{j=1}^{n} E_\phi[\eta_j(\phi)\eta_j'(\phi)\Big]^{-1}$$

for some fixed $\phi \in \Theta$, where E_ϕ denotes the expectation with respect to P_ϕ; in the i.i.d. case one may set $\delta_n'\delta_n = n^{-1}I$, I is the unit matrix. Further we set

$$\xi_j(\theta) = \eta_j(\theta)f_j^{1/2}(\theta).$$

(A.3) : For every $h \in R^k$

$$E[\int |h'\delta_n\xi_j(\theta_0)|^2 d\mu_j] < \infty,\ 1 \leqslant j \leqslant n < \infty.$$

(A.4) : For every $h \in R^k$, for some $a < 0$ and $b > 1$

$$\sup_{a \leqslant t \leqslant b} \sum_{j=1}^{n} E\{\int |h'\delta_n[\xi_j(\theta_0+\delta_n h)-\xi_j(\theta_0)]|^2 d\mu_j\} \to 0.$$

(A.5) : $E[\eta_j(\theta_0)|A_{j-1}] = 0$ for every $j \geqslant 1$.

(A.6) : For every $\varepsilon > 0$ and $h \in R^k$

$$\sum_{1=j}^{n} E[|h'\delta_n\eta_j(\theta_0)|^2 I(|h'\delta_n\eta_j(\theta_0)| > \varepsilon] \to 0.$$

(A.7) : There exists an a.s. positive definite random matrix $T(\theta_0)$ such that the difference

$$\delta_n \sum_{j=1}^{n} [\eta_j(\theta_0)\eta_j'(\theta_0)|\mathcal{A}_{j-1}]\delta_n - T(\theta_0)$$

converges to zero in probability.

(A.8) : $$\sup_{n \geqslant 1} \|\sum_{j=1}^{n} E[\delta_n\eta_j'(\theta_0)\eta_j(\theta_0)\delta_n]\| \leqslant K \text{ for some } K > 0.$$

(A.9) : $$E\Big\{\sup_{a \leqslant |h| \leqslant a+1} |\delta_n \sum_{j=1}^{n} [\eta_j(\theta_0+\delta_n h)-\eta_j(\theta_0)]|\Big\} \leqslant Ca^p$$

for some constants $C > 0$ and $p > 0$ and for all sufficiently large n.

(A.10) : To any positive N there exists an n_0 and a constant C_N depending only on N such that for every $n \geqslant n_0$

$$P\Big\{ \prod_{j=1}^{n} [f_j(\theta_0+\delta_n h)/f_j(\theta_0)] \geqslant \frac{1}{|h|^N} \Big\} \leqslant \frac{C_N}{|h|^N}.$$

Remark : See Section 3 for a discussion of this condition (A.10).

The next assumption will be used only in proving the results for MPE and Bayes estimators.

(A.11) : There exists an $\varepsilon_0 > 0$ such that for all $0 < \varepsilon \leqslant \varepsilon_0$ and for all sufficiently large n

$$E\left\{\sup_{|h| \leqslant \varepsilon} \left|\delta_n \sum_{j=1}^{n} [\eta_j(\theta_n+\delta_n h)-\eta_j(\theta_0)]\right|\right\} \leqslant C\varepsilon$$

where C is some constant.

3. Study of the likelihood function

Throughout what follows we set

$$Z_n(h) = \prod_{j=1}^{n} [f_j(\theta_0+\delta_n h)/f_j(\theta_0)].$$

Theorem 1 : *Suppose the assumptions* (1–8) *are satisfied. Then for every* $h \in R^k$, *the difference*

$$Z_n(h)-exp[h'T^{1/2}(\theta_0)W_n(\theta_0)-\tfrac{1}{2}\, h'T(\theta_0)h]$$

converges to zero in probability and

$$\mathcal{L}(W_n(\theta_0), T(\theta_0)) \Longrightarrow \mathcal{L}(Z, T(\theta_0))$$

where

$$W_n(\theta_0) = T^{-1/2}(\theta_0)\left[\delta_n \sum_{j=1}^{n} \eta_j(\theta_0)\right]$$

and Z *is a copy of the standard k-variate normal distribution independent of* $T(\theta_0)$.

Proof : The proof follows from Theorem 1 and Proposition 1 of Jeganathan (1979).

The next lemma gives an estimate of the continuity modulus of the processes $h \to \log Z_n(h)$, $n \geqslant 1$.

Lemma 1 : *Suppose the assumptions* (A.2), (A.5), (A.8) *and* (A.9) *are satisfied. Then for some constant* $C > 0$

$$P\left[\sup_{|h_2-h_1|<d} |\log Z_n(h_2, \theta_0)-\log Z_n(h_1, \theta_0)| \geqslant d^{1/2}; h_1, h_2 \in B_a\right] \leqslant Ca^p d^{1/2}$$

where the set $B_a = \{h \in R^k ; a \leqslant |h| \leqslant a+1\}$ *and* p *is the positive constant occuring in the condition* (A.9).

Proof: Consider, for $\mu_1 \times \mu_2 \times \ldots \times \mu_n$ almost all $(X_1, X_2, \ldots, X_n)$

$$\log Z_n(h_2) - \log Z_n(h_1) = (h_2-h_1)\delta_n \sum_{j=1}^{n} \eta_j(\theta_n^*)$$

$$= (h_2-h_1)'\delta_n \sum_{j=1}^{n} \eta_j(\theta_0) + (h_2-h_1)'\delta_n \sum_{j=1}^{n} [\eta_j(\theta_n^*) - \eta_j(\theta_0)]$$

where $$|\theta_n^* - (\theta_0 + \delta_n h_1)| \leqslant |\delta_n(h_2-h_1)|.$$

Now,

$$P\Big[\sup_{|h_2-h_1|<d} |(h_2-h_1)'\delta_n \sum_{j=1}^{n} \eta_j(\theta_0)| \geqslant d^{1/2}/2\Big]$$

$$\leqslant 4dE\Big[|\phi_n \sum_{j=1}^{n} \eta_j(\theta_0)|^2 \|\delta_n \phi_n^{-1}\|^2\Big] \qquad \ldots (1)$$

where $$\phi_n'\phi_n = \Big\{\sum_{j=1}^{n} E[\eta_n(\theta_0)\eta_j'(\theta_0)]\Big\}^{-1}$$

It can be easily seen, using the fact that

$$E[\eta_j(\theta_0) | \mathscr{A}_{j-1}] = 0 \text{ for all } j \geqslant 1$$

and $$E\Big\{\phi_n \sum_{j=1}^{n} [\eta_j(\theta_0)\eta_j'(\theta_0)]\phi_n\Big\} = I \text{ (unit matrix)}$$

for all n, that

$$E\Big[|\phi_n \sum_{j=1}^{n} \eta_j(\theta_0)|^2\Big] = k. \qquad \ldots (2)$$

Hence from (1), (2) and (A.8) we see that for some constant $C > 0$

$$P\Big[\sup_{|h_2-h_1|\leqslant d} |(h_2-h_1)'\delta_n \sum_{j=1}^{n} \eta_j(\theta_0)| > d^{1/2}/2\Big] \leqslant d4C \qquad \ldots (3)$$

for all sufficiently large n. Next by (A.9) we have

$$P\Big\{\sup_{|h_2-h_1|\leqslant d} |(h_2-h_1)'\delta_n \sum_{j=1}^{n} [\eta_j(\theta_n^*) - \eta_j(\theta_0)]| \geqslant d^{1/2}/2;\, h_1, h_2 \in B_a\Big\}$$

$$\leqslant 2d^{1/2}Ca^p \qquad \ldots (4)$$

for some constants $C > 0$ and $p > 0$. Hence the result follows from (3) and (4).

The results of the next theorem will be instrumental in the next section.

Theorem 2: *Suppose that the assumptions of Lemma 1 and (A.10) are satisfied. Then to any positive N there exist an n_0 and a constant C_N depending only on N such that for $n \geqslant n_0$*

$$P\left[\sup_{|h| \geqslant a} Z_n(h) \geqslant \frac{1}{a^N}\right] \leqslant C_N/a^N, a \geqslant 2 \quad \dots \quad (5)$$

and

$$P\left[\sup_{a \leqslant |h| \leqslant a+1} Z_n(h) \geqslant \frac{1}{a^N}\right] \geqslant C_N/a^N, a \geqslant 2. \quad \dots \quad (6)$$

Proof: By virtue of the inequality

$$P\left[\sup_{|h| \geqslant a} Z_n(h) \geqslant 2\sum_{k=0}^{\infty} \frac{1}{(a+k)^N}\right]$$

$$\leqslant \sum_{k=0}^{\infty} P\left[\sup_{a+k \leqslant |h| \leqslant a+k+1} Z_n(h) > 1/(a+k)^N\right]$$

relation (5) is a consequence of (6), whose derivation we shall now consider. We partition the set $\{h : a \leqslant |h| \leqslant a+1\}$ into cubes of sides of length a^{5N}. Then totally we will have a^{5kN} number of cubes. Denote the i-th cube and its center by D_i and t_i respectively. Then, for $a \geqslant 2$,

$$P\left[\sup_{a \leqslant |h| \leqslant a+1} Z_n(h) > 1/a^N\right] \leqslant P\left[\sup_{i} \sup_{h \epsilon D_i} Z_n(h) > 1/a^N\right]$$

$$\leqslant P\left[\sup_{i} Z_n(t_i) > 1/a^{5Nk+N}\right]$$

$$+P\left[\sup_{i} \sup_{h \epsilon D_i} |Z_n(h)-Z_n(t_i)| > 1/a^{5N}\right]$$

Now using the fact that, for every $x > 0$ and $y > 0$, $x < \delta^{1/2}/2$ and $|\log x - \log y| \leqslant \delta^{1/2}$ implies $|x-y| \leqslant \delta$ whenever $\delta^{1/2} \leqslant \log 2$, we have, for $a \geqslant 2$,

$$P\left[\sup_{i} \sup_{h \epsilon D_i} |Z_n(h)-Z_n(t_i)| > 1/a^{5N}\right]$$

$$\leqslant P\left[\sup_{i} Z_n(t_i) > 1/2a^{5N/2}\right]$$

$$+P\left[\sup \sup_{h \epsilon D_i} |\log Z_n(h) - \log Z_n(t_i)| > 1/a^{5N/2}\right].$$

Further, for $a \geqslant 2$,

$$P\left[\sup_{l} Z_n(t_l) \geqslant 1/2a^{6N/2}\right] \leqslant P\left[\sup_{l} Z_n(t_l) \geqslant 1/a^{6kN+N}\right]$$

$$\leqslant \sum_{l=1}^{a^{6kN}} P\left[Z_n(t_l) > 1/a^{6kN+N}\right] \leqslant C_N/a^N \quad \text{(by (A.10))}.$$

Hence using Lemma 1 we have, for $a \geqslant 2$,

$$P\left[\sup_{a \leqslant |h| \leqslant a+1} Z_n(h) > 1/a^N\right] \leqslant 2C_N/a^N + C_N/a^{6N/2-p} \leqslant C_N/a^N$$

whenever $N > 2p/3$ (for some $C > 0$).

Hence the result follows since if (6) holds for some N_0, then it will hold for $N \leqslant N_0$.

Using the foregoing results we now present a result concerning the weak convergence of the log-likelihood ratio process to a mixed Gaussian shift process. It may, however, be noted that this result is not used anywhere in the rest of the paper.

Let C_0 be the space of functions which are continuous on $\bar{R}^k$ the one point compactification of R^k, and for which $\lim_{|x| \to \infty} f(x) = 0$ endowed with the usual uniform metric. Suppose that $\Omega_n = \{h : \theta_0 + \delta_n h \in \Theta\} = R^k$ for all sufficiently large n. Then the above theorem in particular implies that $h \to Z_n(h) \in C_0$ a.s. for all sufficiently large n. When $\Omega_n \neq R^k$ for all sufficiently large n we make the following modification.

Let $U_n = \left\{x : d(x, \Omega_n^c) \geqslant \frac{1}{n}\right\}$ and $V_n = \left\{x : d(x, \Omega_n^c) \geqslant \frac{1}{1+n}\right\}$ where Ω^c denotes the complement of Ω_n and $d(x, A)$ means the distance between x and the set A in the usual sense. Then U_n is closed and $U_n \subseteq V_n^0 \subset \Omega$, where A^0 denotes the interior of the set A. Let $f_n : R \to [0,1]$ be a continuous function such that

$$f(x) = \begin{cases} 1 & \text{if } x \in U_n \\ 0 & \text{if } x \in (V_n^0)^c \end{cases}$$

(Such functions always exist). Then define

$$\tilde{Z}_n(h) = \begin{cases} f_n(h) Z_n(h) & \text{if } h \in \Omega_n \\ 0 & \text{otherwise.} \end{cases}$$

It is easy to see from the construction that the function $\bar{Z}_n : R^k \to [0, \infty)$ is continuous. Also note that

$$\bar{Z}_n(h) \begin{cases} \leqslant Z_n(h) & \text{if } h \in \Omega_n \\ = 0 & \text{otherwise.} \end{cases}$$

Hence theorem 2 holds for the sequence $\{\bar{Z}_n(h)\}$, and hence $\bar{Z}_n \in C_0$ a.s. for all sufficiently large n. Further note that, since $\|\delta_n\| \to 0$ as $n \to \infty$, for every $b > 0$, there exists on $n_0(b)$, possibly depending on b, such that $\{h : |h| \leqslant b\} \subseteq U_n$ for all $n \geqslant n_0(b)$. Hence

$$\sup_{|h| \leqslant b} |\bar{Z}_n(h) - Z_n(h)| \xrightarrow{p} 0 \text{ for all } b > 0.$$

Thus it follows from Lemma 1 that the sequence of processes $\bar{Z}_n(h)$ is uniformly equicontinuous in probability, under the assumptions of Theorem 2. In other words, for every $\varepsilon > 0$

$$\lim_{\delta \to 0} \limsup_{n \to \infty} P\left[\sup_{|h_2 - h_1| < \delta} |\bar{Z}_n(h_2) - \bar{Z}_n(h_1)| > \varepsilon;\ h_1, h_2 \in \bar{R}^k\right] = 0.$$

We thus have the following theorem by invoking appropriate theorems in Prakasa Rao (1975) or Straf (1972).

Theorem 3 : *Suppose the Assumptions* (A.1)–(A.10) *are satisfied. Then the distribution in C_0 generated by the process $\bar{Z}_n(h)$ converge as $n \to \infty$ to the distribution generated by the process*

$$R(h) = exp(h'T^{1/2}(\theta_0)Z - \tfrac{1}{2}\, h'(\theta_0)h)$$

where Z is a copy of the standard normal distribution $N(0, 1)$ independent of $T(\theta_0)$. In particular, if f is a continuous functional on C_0, then for all $x \in R$.

$$\lim_{n \to \infty} P\{f[Z_n(h)] < x\} = P\{f[R(h)] < x\}.$$

4. Convergence of moments

(a) *Maximum likelihood estimators.* A Borel measurable function $\hat{\theta}_n = \hat{\theta}_n(X_n)$ is called a maximum likelihood estimator if

$$L_n(X_n; \hat{\theta}_n) \geqslant L_n(X_n; \theta)$$

for all $\theta \in \Theta$, where $L_n(X_n; \theta) = \prod_{j=1}^{n} f_j(\theta)$.

Theorem 4 ; *Suppose the assumptions* (A.1)-(A.10) *are satisfied. Then for any* $m > 0$

$$\lim_{n\to\infty} E[|\delta_n^{-1}(\hat\theta_n-\theta_0)|^m] = E[|T^{-1/2}(\theta_0)Z|^m],$$

Z is a copy of the standard k-variate normal distribution independent of $T(\theta_0)$.

Before giving the proof of this theorem we first prove the following

Lemma 2 : *Suppose the assumptions of Theorem 2 are satisfied. Then for any given* $N > 0$ *there exists an* n_0 *and a constant* C_N *depending only on N such that*

$$P[|\delta_n^{-1}(\hat\theta_n-\theta_0)| > x] \leqslant C_N/x^N$$

for all $n \geqslant n_0$ *and for all sufficiently large* $x > 0$.

Proof : Consider, for $x > 0$,

$$P[|\delta_n^{-1}(\hat\theta_n-\theta_0)| > x] \leqslant P\left[\sup_h Z_n(h) \leqslant \sup_{|h|>x} Z_n(h)\right]$$

$$\leqslant P\left[\sup_h Z_n(h) \leqslant \sup_{|h|>x} Z_n(h);\ \sup_{|h|>x} Z_n(h) \leqslant 1/x^N\right]$$

$$+P\left[\sup_{|h|>x} Z_n(h) > 1/x^N\right].$$

Now note that $\sup_h Z_n(h) \geqslant Z_n(0) = 1$. Hence for all x such that $x^{-N} < 1$ we have by Theorem 2

$$P[|\delta_n^{-1}(\hat\theta_n-\theta_0)| > x] \leqslant P\left[\sup_{|h|>x} Z_n(h) > 1/x^N\right] < C_N/x^N.$$

Hence the proof of the lemma is complete.

Proof of Theorem 4 : Lemma 2 in particular entails that the sequence $\{\delta_n^{-1}(\hat\theta_n-\theta_0)\}$ is relatively compact. Hence in view of Lemma 1 and Theorem 7 of Jeganathan (1979) it follows first that

$$\mathcal{L}(\delta_n^{-1}(\hat\theta_n-\theta_0)) \Longrightarrow N(0, T^{-1}(\theta_0)).$$

Secondly, it is easily seen from Lemma 2 that the sequence $\{E[|\delta_n^{-1}(\hat\theta_n-\theta_0)|^{m+1}\}$, $m > 0$, is uniformly bounded for all sufficiently large n and hence the sequence $\{|\delta_n^{-1}(\hat\theta_n-\theta_0)|^m\}$ is uniformly integrable. Now the proof can be easily concluded from these two facts.

(b) *Maximum probability estimators.* Let $B = \{t \in R^k : |t| \leqslant a, a > 0\}$. A maximum probability estimator with respect to B (to be denoted by $\bar{\theta}_n(\bar{a})$) is one which maximises with respect to d,

$$\int L_n(X_n : \theta)d\theta$$

the integral being over the set $\{d - \delta_n B\}$.

Theorem 5: *Suppose the assumptions* (A.1)-(A.11) *are satisfied. Then for any* $m > 0$

$$\lim_{n \to \infty} E[|\delta_n^{-1}(\bar{\theta}_n - \theta_0)|^m] = E[|T^{-1/2}(\theta_0)Z|^m].$$

Before giving the proof of this theorem we shall first present some preliminary lemmas.

Lemma 3: *Suppose the assumptions* (A.1) (A.2) *and* (A.11) *are satisfied. Then there are positive constants* C *and* ε_0 *such that for all* $0 < \varepsilon \leqslant \varepsilon_0$

$$P\left[\int_{|u| \leqslant \varepsilon} Z_n(u)du \leqslant \varepsilon^k\right] \leqslant C\varepsilon^2,$$

where k *is the dimension of* Θ.

Proof: Because of the assumed continuity there exists an u^* which may depend on the observations $|u^*| \leqslant \varepsilon$, such that

$$P\left[\inf_{|u| \leqslant \varepsilon} Z_n(u) = Z_n(u^*)\right] = 1.$$

Hence

$$P\left[\int_{|u| \leqslant \varepsilon} Z_n(u)du \leqslant \varepsilon^k\right] \leqslant P[Z_n(u^*) < \tfrac{1}{2}]$$

$$= P[|\log Z_n(u^*)| > |\log(\tfrac{1}{2})|]$$

$$\leqslant P\left[\sup_{|u| \leqslant \varepsilon} |\log Z_n(u)| \geqslant |\log(\tfrac{1}{2})|\right]$$

In view of the condition (A.11) the last term of the above expression can be shown to be less than or equal to $C\varepsilon^2$ by following the arguments similar to the proof of Lemma 1. Hence the proof of the lemma is complete.

Lemma 4: *Suppose that the assumptions of Theorem 2 are satisfied. Then for any given* $N > 0$ *and* $a \geqslant 0$ *there exists an* n_0 *and a constant* C_N *depending only on* N *such that for all* $n \geqslant n_0$

$$P\left[\int_{|h| > M} |h|^a Z_n(h)dh > M^{-N}\right] \leqslant C_N/M^N \quad M \geqslant 1.$$

Proof: It is enough to prove the result for $N > N_0$ for some $N_0 > 0$ since if it is true for N_0 then it will be true for $N \leqslant N_0$. Consider for $M \geqslant 1$

$$P\left[\int_{|h|>M} |h|^a Z_n(h)dh > 1/M^N\right] \leqslant P\left[\int_{|h|>M} |h|^a Z_n(h)dh > 1/M^{N+a+1}\right]$$

$$\leqslant P\left[\int_{|h|>M} |h|^a Z_n(h)dh > \sum_{k=0}^{\infty} 1/(M+k)^{N+a+2}\right] \text{ (for } N > 2)$$

$$\leqslant \sum_{k=0}^{\infty} P\left[\int_{M+k<|h|\leqslant M+k+1} |h|^a Z_n(h)dh > 1/(M+k)^{N+a+2}\right]$$

$$\leqslant \sum_{k=0}^{\infty} P\left[\int_{M+k<|h|<M+k+1} Z_n(h) > 1/(M+k)^{N+2}\right]$$

$$< \sum_{k=0}^{\infty} C_N/(M+k)^{N+2+d} \text{ (by Theorem 2) (for } N > (2^d-k-1)$$

for all $n \geqslant n_0$ where n_0 and C_N are as in Theorem 2 and $d = \dim \Theta$. This completes the proof of the lemma.

Lemma 5: *Suppose the assumptions of Theorem 2 are satisfied. Then there exist an n_0 and a constant C_N depending only on N such that for all $n \geqslant n_0$*

$$P[|\delta_n^{-1}(\theta_n-\theta_0)| > x] \leqslant C_N/x^N,\ a > 0$$

for all sufficiently large $x > 0$.

Proof: Consider setting $D_a = \{h \epsilon R^k : |h| \leqslant a\}$,

$$P[|\delta_n^{-1}(\bar{\theta}_n(a)-\theta_0)| > x]$$

$$\leqslant P\left[\sup_{|u|>x} \int_{u-D_a} Z_n(h)dh \geqslant \sup_{|u|<x} \int_{u-D_a} Z_n(h)dh\right]$$

$$\leqslant P\left[\int_{|h|\geqslant(x-a)} Z_n(h)dh \geqslant \int_{D_a} Z_n(h)dh\right]$$

$$\leqslant P\left[\int_{|h|\geqslant(x-a)} Z_n(h)dh \geqslant \int_{D_a} Z_n(h)dh;\right.$$

$$\int_{|h|\geqslant(x-a)} Z_n(h)dh \leqslant 1/(x-a)^{Nk};$$

$$\int_{D_a} Z_n(h)dh \geqslant \frac{2^{Nk/2}}{x^{Nk/2}}\Big]$$

$$+P\Big[\int_{|h|>(x-a)} Z_n(h)dh > 1/(x-a)^{Nk}\Big]$$

$$+P\Big[\int_{D_a} Z_n(h)dh < \frac{2^{Nk/2}}{x^{Nk/2}}\Big]$$

$$= I_1+I_2+I_3 \quad \text{say. } (k = \dim\Theta).$$

Assume in what follows that $x > \max(2a, 2)$. Then we have $1/(x-a)^{Nk} < \frac{2^{Nk}}{x^{Nk}} < \frac{2^{Nk/2}}{x^{Nk/2}}$ and so $I_1 = 0$.

By Lemma 4 $\quad I_2 \leqslant C_N(2^N/x^N)$ (for some $C_N > 0$)

and by Lemma 3 $\quad I_3 \leqslant 2^N/x^N$.

Hence $\quad P[|\delta_n^{-1}(\bar{\theta}_n(a)-\theta_0)| > x] \leqslant C_N(1+2^N/x^N$ (for some $C_N > 0$).

This completes the proof of the lemma.

Proof of Theorem 5: Lemma 5 in particular implies that the sequence $\{\delta_n^{-1}(\bar{\theta}_n(a)-\theta_0)\}$ is relatively compact. Hence in view of Theorem 1 of the present chapter and Theorem 4 of Jeganathan (1982)* it follows that

$$\mathcal{L}(\delta_n^{-1}(\bar{\theta}_n(a)-\theta_0)) \Longrightarrow N(0, T^{-1}(\theta_0)).$$

Now proceeding as in the proof of Theorem 4 the proof of the theorem is completed.

(c) *Bayes estimators.* Assume that we are given a prior density $\pi(\theta)$ such that $\pi(\theta) > 0$ for all $\theta \in \Theta$, $\pi(\theta)$ is continuous and for every $\theta_0 \in \Theta$ there exist $h_0 > 0$ and $p > 0$ such that

$$\pi(\theta_0+\delta_n h) \leqslant c|h|^p$$

for all $|h| \geqslant h_0$. We also assume in what follows that the assumptions (A.1)–(A.11) are satisfied.

Define a regular Bayes estimator $t_n = t_n(X_n)$ as an estimator which minimises $B_n(\phi) = \int l_n(\theta, \phi)f_n(\theta|X_n)d\theta$ for all sequences $(X_1, X_2 \ldots)$ where $l_n(\theta, \phi)$ is a loss function defined on $\Theta\times\Theta$ and $f_n(\theta|X_n)$ is the posterior density. We shall assume that a measurable regular Bayes estimator t_n

* Published in this issue, pp. 173–212.

exists and we consider the Bayes estimators with respect to the loss functions $|\delta_n^{-1}(\theta-\phi)|^a$ $a \geqslant 1$. We shall suppose for brevity that $\pi(\theta) = 1$ since the passage to the general case causes no difficulties.

Theorem 6 : *Suppose that the Assumptions* (A.1)–(A.11) *are satisfied. Further assume that*

(A.12) *the largest eigen value of $T^{-1}(\theta_0)$ has finite moments of all orders. Then for every $m > 0$*

$$\lim_{n \to \infty} E[|\delta_n^{-1}(t_n-\theta_0)|^m] = E[|T^{-1/2}(\theta_0)Z|^m].$$

Before giving the proof of this theorem we present some preliminary lemmas.

Lemma 6 : *Let $a \geqslant 1$. Set*

$$D_n(h) = \frac{Z_n(h)}{\int Z_n(h)dh} - \frac{|\det T(\theta_0)|^{1/2}}{(2\pi)^{k/2}} exp[-\tfrac{1}{2}|T^{1/2}(\theta_0)(h-h_n^*)|^2]$$

where $h_n^* = \delta_n^{-1}(\hat{\theta}_n-\theta_0)$. Suppose that the assumptions of Theorem 6 are satisfied. Then for any $N > 0$ there exists an n_0 and a constant C_N depending only on N such that for all $n \geqslant n_0$

$$P\Big[\int_{|h|>M} |h|^a|D_n(h)|dh > M^{-N}\Big] \leqslant C_N M^{-N}, \ M \geqslant 1.$$

Proof :

$$P\Big[\sup_{|h|>M} Z_n(h)/\int Z_n(h)dh > 1/2M^N\Big] \leqslant P\Big[\int_{|h|>M} Z_n(h) > M^{-3N/2}\Big]$$
$$+P[\int Z_n(h)dh < 2M^{-N/2}]$$
$$\leqslant C_N M^{-N}$$

by Theorem 2 and Lemma 3 for all sufficiently large n. Hence from the arguments of Lemma 4 it follows that for all sufficiently large n

$$P\Big\{\int_{|h|>M} |h|^a Z_n(h)dh/Z_n(h)dh > 1/2M^N\Big\} \leqslant C_N/M^N \text{ (for some } C_N > 0).$$

Denote the eigen values of $T(\theta_0)$ by $\lambda_1 < \ldots < \lambda_k$. Now $|h_n^*| \leqslant M/2$ and $1/\lambda_1 < M$ implies that

$$\frac{|\det T(\theta_0)|^{1/2}}{(2\pi)^{k/2}} \int_{|h|>M} |h|^a \exp[-\tfrac{1}{2}|T^{1/2}(\theta_0)(h-h^*)|^2]dh$$

$$\leqslant \frac{K|\det T(\theta_0)|^{1/2}}{(2\pi)^{k/2}} \int_{|h-h_n^*|>M/2} |h-h_n^*|^a \exp(-\tfrac{1}{2}|T^{1/2}(\theta_0)(h-h_n^*)|^2]dh$$

(for some $K > 0$)

$$= \frac{K(\lambda_1 \ldots \lambda_n)^{1/2}}{(2\pi)^{k/2}} \int_{|h|>M/2} (\Sigma h_i^2)^{a/2} \exp(-\tfrac{1}{2} \sum_{i=1}^{k} \lambda_i h_i^2) dh$$

$$\leqslant \frac{(\lambda_1 \ldots \lambda_k)^{1/2}}{(2\pi)^{k/2}} \frac{K2^{4N}}{M^{4N}} \int \left(\sum_{i=1}^{k} h_i^2 \right)^{(a/2)+2N} \exp\left(-\tfrac{1}{2} \sum_{i=1}^{k} \lambda_i h_i^2\right) dh$$

where h_i's are the components of the vector h. In what follows we assume without loss of generality that a is an integer. Now using the fact that

$$\frac{(\lambda_1 \ldots \lambda_k)^{1/2}}{(2\pi)^{k/2}} \int \prod_{i=1}^{k} (\lambda_i h_i)^{l_i} \exp\left(-\frac{1}{2} \sum_{i=1}^{k} \lambda_i h_i^2\right) dh < \infty$$

for every $l_i > 0$ $i = 1$ 2 $\ldots$, k, we see that the above integral is bounded by

$$(K2^{4N}/M^{4N}) \left(\sum_{i=1}^{k} 1/\lambda_i\right)^{a/2+2N} \text{(for some } K > 0)$$

$$\leqslant K2^{4N}/M^{4N-a/2-N} \leqslant K2^N/M^N$$

for some $K > 0$ and for all $N \geqslant 2$, since $1/\lambda_1 < M$. Therefore for every $N > 0$

$$P\left[\frac{|\det T(\theta_0)|^{1/2}}{(2\pi)^{k/2}} \int_{|h|>M} |h|^a \exp(-\tfrac{1}{2}|T^{1/2}(\theta_0)(h-h_n^*)|^2)dh > 1/2M^N\right]$$

$$\leqslant P[|h_n^*| > M/2] + P[1/\lambda_1 < M] \leqslant C_N/M^N \quad \text{(for some } C_N)$$

for all sufficiently large n, by Theorem 4 and the given assumption (A.12).

Lemma 7 : *Let $a \geqslant 1$. Set*

$$D_n^*(h) = f_n^*(\theta_n + \delta_n h) - \frac{|T(\theta_0)|^{1/2}}{(2\pi)^{k/2}} \exp(-\tfrac{1}{2} h' T(\theta_0) h),$$

where $$f_n^*(\theta_n+\delta_n h) = \prod_{j=1}^{n} f_j(\theta_n+\delta_n h) \bigg/ \int \prod_{j=1}^{n} f_j(\theta_n+\delta_n h) dh.$$

Suppose the assumptions of Theorem 6 are satisfied. Then for any given $N > 0$, there exist an n_0 and constant $C_N^{(1)}$ and $C_N^{(2)}$ depending only on N such that for all $n \geqslant n_0$ and $M \geqslant 1$

$$P\left[\int_{|h|>M} |h|^a |D_n^*(h)| dh > C_N^{(1)} M^{-N}\right] \leqslant C_N^{(2)} M^{-N}.$$

Proof: Let $g_n = h+h_n^*$.

Then $$|h| \leqslant d_a|g_n|^a+d_a|h_n^*|^a$$

where $d_a = 2^{a-1}$ Using this inequality we have

$$\int_{|h|>M} |h|^a D_n^*(h)|dh \leqslant d_a \int_{|g-h_n^*|>M} |g|^a D_n(g)|dg$$
$$+d_a|h_n^*|^a \int_{|g-h_n^*>M} |D_n(g)|dg,$$

where $D_n(g)$ is as defined in Lemma 6. Consider

$$P\Big[\int_{|g-h_n^*|>M} |g|^a|D_n(g)|dg > 2^N M^{-N}\Big]$$

$$\leqslant P\Big[\int_{|g-h_n^*|>M} |g|^a|D_n(g)|dg > 2^N M^{-N};\ |h_n^*| \leqslant M/2\Big]+P[|h_n^*|>M/2]$$

$$\leqslant P\Big[\int_{|g|>M/2} |g|^a|D_n(g)|dg > 2^N M^{-N}\Big]+P[|h_n^*|>M/2]$$

$$\leqslant C_N^{(2)}M^{-N} \quad \text{(for some } C_N^{(2)}>0) \qquad \ldots (7)$$

by the previous lemma 6 and Theorem 4. Similarly it can be shown that, for some $C_N^{(1)}>0$ and $C_N^{(2)}>0$,

$$P\Big[|h_n^*|^a \int_{|g-h_n^*|\geqslant M} |D_n(h)|dh > C_N^{(1)}M^{-N}\Big] \leqslant C_N^{(2)}M^{-N} \qquad \ldots (8)$$

Hence the result follows from (7) and (8).

Lemma 8: *Suppose that the assumptions of Theorem 6 are satisfied. Then for every $N>0$, there exist an n_0 and a constant C_N depending only on N such that*

$$P[|\delta_n^{-1}(t_n-\theta_0)| > M] \leqslant C_N M^{-N}$$

for all $n \geqslant 0$ and $M \geqslant 1$.

Proof: Since t_n is a Bayes estimator with respect to the loss function $|\delta_n^{-1}(\theta-\phi)|^a$, $a \geqslant 1$, we have putting $u_n = \delta_n^{-1}(t_n-\hat{\theta}_n)$,

$$\int |h|^a f_n^*(\hat{\theta}_n+\delta_n h)dh \geqslant \int |h+u_n|^a f_n^*(\hat{\theta}_n+\delta_n h)dh \geqslant \int_{|h|<M/4} |h+u_n|^a f_n^*(\hat{\theta}_n+\delta_n h)dh$$

Hence, since $|h| \leqslant M/4$ and $|u_n| > M$ implies that $|h+u_n| > |h|+M/2$,

$$P[|\delta_n^{-1}(t_n-\theta_n)| > M]$$

$$\leqslant P\left\{\int |h|^a f_n^*(\theta_n+\delta_n h)dh \geqslant \int_{|h|<M/4} |(|h|+M/2)^a f_n^*(\theta_n+\delta_n h)dh\right\}$$

$$\leqslant P\left\{\int |h|^a f_n^*(\theta_n+\delta_n h)dh \geqslant \int_{|h|<M/4} |h|^a f_n^*(\theta_n+\delta_n h)dh + K\int_{|h|<M/4} f_n^*(\theta_n+\delta_n h)dh\right\}$$

(for some $K > 0$)

$$\leqslant P\left\{\int_{|h|>M/4} |h|^a f_n^*(\theta_n+\delta_n h)dh \geqslant K\int_{|h|<M/4} f_n^*(\theta_n+\delta_n h)dh\right\} \quad \ldots \quad (9)$$

Now note that

$$\int f_n^*(\theta_n+\delta_n h)dh = 1 = \frac{|\det T(\theta_0)|^{1/2}}{(2\pi)^{1/2}} \int \exp[-\tfrac{1}{2}\, h'T(\theta_0)h]dh.$$

Hence by the arguments of the proof of the lemma 6 and by the previous lemma 7, for any given $N > 0$, there exist constants $C_N^{(1)}$ and $C_N^{(2)}$ and an n_0 depending only on N such that for all $n \geqslant n_0$

$$P\left[\int_{|h|>M/4} |h|^a f_n^*(\theta_n+\delta_n h)dh > C_N^{(1)}M^{-N}\right] \leqslant C_N^{(2)}M^{-N}$$

and

$$P\left\{\left|\int_{|h|<M/4} f_n^*(\theta_n+\delta_n h)dh - \frac{|\det T(\theta_0)|^{1/2}}{(2\pi)^{k/2}} \int_{|h|<M/4} \exp[-\tfrac{1}{2}h'T(\theta_0)h]dh\right| \geqslant C_N^{(1)}M^{-N}\right\} \leqslant C_N^{(2)}M^{-N}.$$

Hence it is easily seen that the last term of the above inequality (9) is less than or equal to $2C_N^{(2)}M^{-N}$ for all $n \geqslant n_0$, that is,

$$P[\delta_n^{-1}(t_n-\theta_n)| > M] \leqslant 2C_N^{(2)}M^{-N} \quad \ldots \quad (10)$$

for all $n \geqslant n_0$. Now

$$\begin{aligned} &P[|\delta_n^{-1}(t_n-\theta_0)| > M] \\ &\leqslant P[|\delta_n^{-1}(t_n-\theta_n)| > M/2] \\ &+P[|\delta_n^{-1}(\theta_n-\theta_0)| > M/2] \\ &\leqslant 2^{N+1}C_N^{(2)}M^{-N}+2^N M^{-N}E[|\delta_n^{-1}(\theta_n-\theta_0)|^N] \end{aligned}$$

by (10). By Theorem 4, the proof of the lemma is complete.

Proof of Theorem 6 : First note that the lemma 4 implies that for every $\varepsilon > 0$ and $a > 1$

$$\lim_{a \to \infty} \limsup_{n \to \infty} P\left[\int_{|h| > a} |h|^a Z_n(h) dh > \varepsilon \right] = 0.$$

Hence by Theorem 1 of the present chapter and Theorem 6 of Jeganathan (1970) it follows that

$$\mathcal{L}(\delta_n^{-1}(t_n - \theta_0)) \Longrightarrow N(0, T^{-1}(\theta_0)).$$

Now proceeding as in the proof of Theorem 4, the proof of the theorem is completed.

5. Discussions on the assumption (A.10) and some examples

Consider a class of mixed Gaussian processes having the following form

$$Z_n(h) = \exp(h'U_n(\theta_0) - \tfrac{1}{2} h'T_n(\theta_0)h) \qquad \dots \quad (11)$$

where, for every $n \geq 1$, U_n is a random k-vector and $T_n(\theta_0)$ is a p.d. random $k \times k$ matrix. We further assume that the moments of any order of the largest eigen value of the matrices $T_n^{-1}(\theta_0)$ are uniformly bounded for all large n. Let λ_n be the smallest eigen value of $T_n(\theta_0)$. In order to verify (A.10) it is enough to show that for every $N > 0$ there exists an n_0 and C depending only on N such that

$$P[Z_n^{1/2}(h) > |h|^{-N};\ \lambda_n^{-1} < |h|] \leq C_N |h|^{-N}$$

since we have assumed that for every $m > 0$, $\sup_{n \geq n_0} E(\lambda_n^{-m}) < \infty$ for some n_0.

Now

$$\begin{aligned} & P[Z_n^{1/2}(h) > |h|^{-N};\ \lambda_n^{-1} < |h|] \\ & \leq |h|^N E\{I(\lambda_n^{-1} < |h|) Z_n^{1/2}(h)\} \\ & \leq |h|^N E\{I(\lambda_n^{-1} < |h|) Z_n(h/2) \exp(-(\lambda_n/8)|h|^2)\} \\ & \leq |h|^N E\{I(\lambda_n^{-1} < |h|) Z_n(h/2) K_N \lambda_n^{-2N} |h|^{-4N}\} \\ & \qquad\qquad\qquad\qquad\qquad \text{(for some } K_N > 0) \\ & \leq |h|^{2N} |h|^{-4N} K_N E[Z_n(h/2)] \\ & \leq K_N |h|^{-N} \quad (\text{since } E[Z_n(h/2)] = 1). \end{aligned}$$

Thus we see that the assumption (A.10) is satisfied in this case.

In some cases it is possible to verify that

$$\sum_{j=1}^{n} \inf_{X_{j-1}} \int [f_j^{1/2}(\theta_0+\delta_n h)-f^{1/2}(\theta_0)]^2 d\mu_j \geq C|h|^2 \qquad \dots \quad (12)$$

for some $C>0$ and for all sufficiently large n, where $X_j=(X_1, X_2, \dots, X_j)$. Then putting

$$a_j = \inf_{X_{j-1}} \int [f_j^{1/2}(\theta_0+\delta_n h)-f^{1/2}(\theta_0)]^2 d\mu_j,$$

$$E[Z_n^{1/2}(h)] \leq \prod_{j=1}^{n} (1-a_j/2).$$

Note that $|a_i| \leq 1$, since

$$0 \leq \sup_{X_{j-1}} \int f_j^{1/2}(\theta_0+\delta_n h) f_j^{1/2}(\theta_0) d\mu_j \leq 1.$$

Hence

$$E[Z_n^{1/2}(h)] \leq \exp\left[-\frac{1}{2}\sum_{j=1}^{n} a_j\right] \leq \exp[-C|h|^2]$$

for all sufficiently large n, when (12) holds.

A situation where (12) can be easily verified, with $\delta_n'\delta_n = n^{-1}I$, is the following example.

Example : Let $X_0, X_0, \dots,$ be a sequence of Markov chain for which the state space consist of the numbers 0 and 1; the transition matrix is

X_0 \ X_1	0	1
0	$(1-p)+\pi p$	$(1-\pi)p$
1	$(1-\pi)(1-p)$	$\pi+(1-\pi)p$

and the initial distribution of $f(1, p, \pi) = 1-f(0, p, \pi) = p$ where

$$(p, \pi) \in \Theta = (0, 1)\times(0, 1).$$

We finally point out that the assumptions (A.1)–(A.11) can be considerably simplified when the observations are i.i.d.

Acknowledgements. The present work was done when the author was working under the supervision of Dr. B. L. S. Prakasa Rao. The author wishes to express his gratitude to Prof. J. K. Ghosh for his kind help in correcting a number of mistakes in the original version of the paper. The author further wishes to express his thanks to the referee for suggesting some important corrections, and to Dr. S. Bagchi for his kind help in one of these corrections.

References

Ibragimov, I. A. and Khasminskii, R. Z. (1972-73): Asymptotic behaviour of some statistical estimators in the smooth case I and II. *Theory Prob. Applications*, **17**, 445-462, **18**, 76-91.

——— (1975): Properties of maximum likelihood and Bayes estimators for non-identically distributed observations. *Theory Prob. Applications*, **20**, 689-697.

Inagaki, V. and Ogata, Y. (1975): The weak convergence of likelihood ratio random fields and its applications. *Ann. Inst. Statist. Math.*, **27**, 391-419.

——— (1977): The weak convergence of the likelihood ratio random fields for Markov observations. *Ann. Inst. Statist. Math.* **29**, 165-187.

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

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

Levit, V. Ya. (1974): On the behaviour of generalised Bayes estimate for markov observations. *Theory Prob. Applications*. **19**, 332-346.

Prakasa Rao, B. L. S. (1968): Estimation of the location of the cusp of a continuous density. *Ann. Math. Statist.*, **39**, 76-87.

——— (1975): Tightness of probability measures generated by stochastic processes on metric spaces. *Bull. Inst. Math. Academic.* Sinica, **3**, 353-367.

Prokorov, Yu. V. (1966): Convergence of random processes and limit theorems in probability theory. *Theory Prob. Applications*, **1**, 157-214.

Rubin, H. (1961): The estimation of discontinuities in multivariate densities and related problems in stochastic processes. *Proc. Fourth Berkeley Sympos.*, **1**, 563-574.

Straf, M. L. (1970): Weak convergence of random processes with several parameters. *Proc. 6th Berkeley Symp. Math. Statist. Prob.* **2**, 187-221.

Paper received : February, 1979.
Revised : September, 1981.