Sankhyā : The Indian Journal of Statistics
1982, Volume 44, Series A, Pt. 2, pp. 173–212.

ON THE ASYMPTOTIC THEORY OF ESTIMATION WHEN THE LIMIT OF THE LOG-LIKELIHOOD RATIOS IS MIXED NORMAL

By P. JEGANATHAN
Indian Statistical Institute

SUMMARY. Some results concerning the asymptotic theory of estimation are presented when the limit distribution of the log-likelihood ratios is mixed normal. More specifically, the notion of a locally asymptotically mixed normal (LAMN) sequence of families of distributions is introduced, and it is shown that when a certain kind of differentiability in quadratic mean type regularity condition is satisfied the given sequence of families satisfies the LAMN condition. As a consequence of the LAMN condition, it is shown that the limit distribution of a regular sequence of estimators can be conditionally decomposed as a convolution. Using this conditional convolution result, some results concerning the asymptotic lower bound for risk functions of estimators are obtained. Given that the sequence of families satisfies the LAMN condition quite general additional assumptions are sought under which it is shown that the maximum probability estimators, maximum likelihood estimators and a certain class of Bayes estimators are asymptotically optimal in a certain sense. A result concerning the posterior approximation at the true value of the parameter is also presented.

1. Introduction

In one of his fundamental papers LeCam (1960) introduced what is now called locally asymptotically normal (LAN) families of distributions and obtained several basic results regarding the asymptotic theory of estimation and testing. Roughly speaking, a sequence of families is said to satisfy the LAN condition if the corresponding sequence of appropriately normalised log-likelihood ratios is locally approximated with probability tending to one by the sum of two expressions, the first one being a sequence of random linear functions of the normalised parameter and the second one being a non-random quadratic form of the normalised parameter and the sequence of random vectors involved in the linear term of the approximation converges weakly to the normal distribution with mean vector zero and the covariance matrix being the matrix involved in the quadratic form of the approximation.

An important thing to observe regarding the LAN condition of LeCam (1960) is that a large part of asymptotic theory depends only on the approximating form of the log-likelihood ratios, and any specific property such as i.i.d. or any other particular form of dependence is not much relevant.

In recent times, there occur situations, e.g., in Galton-Watson branching processes and pure-birth process as has been discussed in e.g., Keiding (1975), Heyde and Feigin (1975), Basawa and Scott (1976), Heyde (1978) and Bhat (1978), in which LAN condition is not satisfied, but it can be seen that a quite similar and more general condition, which may be called *locally asymptotically mixed normal* (LAMN) condition, is satisfied. Roughly speaking, a sequence of families may be said to satisfy the LAMN condition if the corresponding sequence of appropriately normalised log-likelihood ratios is locally approximated with probability tending to one by the sum of two expressions, the first one being a sequence of random linear functions of the normalised parameter and the second one being a sequence of random quadratic forms of the normalised parameter, the sequence of random matrices involved in the quadratic forms being convergent weakly to an almost surely positive definite *random* matrix and the random vectors involved in the linear terms being convergent weakly to an appropriate mixed normal distribution.

In Section 2, a detailed study of the LAMN condition is carried out under a certain kind of differentiability in quadratic mean type regularity condition. Most of the results of this section were originally obtained for the i.i.d. case by LeCam (1970). LAN-condition for the dependent observations has been studied, among several others, by Roussas (1972, 1979) and for the independent but not necessarily identically distributed case has been studied by Phillippou and Roussas (1973) and Ibragimov and Khasminskii (1975).

As a consequence of the LAMN-condition, in Section 3 it is shown that the limit distribution of any convergent subsequence of estimators satisfying certain invariance restriction can be conditionally decomposed as a convolution. Using this conditional convolution result, some results concerning the asymptotic lower bound for risk functions of estimators are obtained. This result extends and strengthens the convolution result obtained by Hájek (1970) for the LAN case. Convolution result for the LAN case was independently obtained by Inagaki (1970) also under some restrictive assumptions. LeCam (1972) has extended Hájek's convolution result to a much more general situations than that of the LAN case.

In Section 4, an exponential approximation result analogous to Theorem 3.1 of LeCam (1960) is presented. This result is used in Section 5 of the present paper. Also it serves as a powerful tool in serveral other places, see e.g. Jeganathan (1980a).

The main purpose of Section 5 is, under suitable global assumptions, to see what are the minimum possible local regularity conditions needed under which the sequences of maximum likelihood estimators, maximum probability estimators and a certain class of Bayes estimators can be approximated with probability tending to one by the random vectors involved in the linear terms of the LAMN condition. A result concerning the posterior approximation at the true value of the parameter is also presented. Our arguments depend only on the approximating form of the log-likelihood ratios, and they do not in any way depend on any particular nature of the sample space. For example, given that a sequence of maximum probability estimators is consistent at a certain rate, the only additional condition we assume to show that this sequence satisfies the above requirement is the LAMN condition.

In Section 6, we give easily verifiable regularity conditions, in terms of the first derivatives, implying the more general differentiability in quadratic mean type regularity condition of Section 2.

In connection with the present paper it should be mentioned here that Heyde (1978) has obtained under some specific assumptions, an important result that maximum likelihood estimators provide the best asymptotic probability of concentration in symmetric intervals. This result is clarified in Section 3 of the present paper.

The first version of the present paper appeared as a technical report in January, 1979. Since then or at about the same period of time several important results have been obtained. Regarding the further study of the LAMN condition and the asymptotic theory of estimation, mention may be made of Davies (1979), Jeganathan (1981, 1980, 1980a, 1980b) and Swansen (Ph.D. Thesis, September, 1980).

The results, obtained independently, of the first chapter of Swanson's thesis is related to the results of Section 2 of the present paper.

Swanson's thesis further contains some important results concerning the asymptotic theory of testing for the LAMN case. For an earlier treatment of testing problem for the general case, under some specific assumptions, see Basawa and Scott (1977) and Feigin (1978).

The following notations are used throughout. If P and Q are probability measures on a measurable space $(\mathscr{X}, \mathscr{A})$, then dP/dQ denotes the Radon-Nikodym derivative of the Q-continuous part of P with respect to Q. If Y is a random vector its distribution will be denoted by $\mathscr{L}(Y)$ or by $\mathscr{L}(Y \mid P)$ when Y is a random vector defined on $(\mathscr{X}, \mathscr{A}, P)$. For a vector $h \in \mathcal{R}^k$, h' denotes the transpose of h and $|h|$ denotes the euclidean norm. For a square matrix D, $\|D\|$ denotes the norm defined by the square root of the sum of the squares of its elements. '$\Longrightarrow$' denotes the converges in distribution and '$\xrightarrow{P}$' denotes the convergence in probability.

We now introduce the precise definition of the LAMN condition.

Definition 1. For each $n \geqslant 1$, let $\{P_{\theta,n};\ \theta \in \Theta\}$ be a family of probability measures defined on $(\mathcal{X}_n, \mathcal{F}_n)$, where Θ is an open subset of $\mathcal{R}^k$, $k \geqslant 1$. Then the sequence of families $\{P_{\theta,n};\ \theta \in \Theta\}$, $n \geqslant 1$, satisfies the LAMN condition at $\theta = \theta_0 \in \Theta$ if the following two conditions are satisfied.

(A.1). There exists a sequence $\{W_n(\theta_0)\}$ of $\mathcal{F}_n$-measurable k-vectors and a sequence $\{T_n(\theta_0)\}$ of $\mathcal{F}_n$-measurable $k \times k$ symmetric matrices such that $P_{\theta_0,n}\,[T_n(\theta_0)$ is p.d.$] = 1$ for every $n \geqslant 1$ and the difference

$$\log \frac{dP_{\theta_0+\delta_n^{-1}h,n}}{dP_{\theta_0,n}} - [h' T_n^{1/2}(\theta_0) W_n(\theta_0) - \frac{1}{2} h' T_n(\theta_0) h]$$

converges to zero in $P_{\theta_0,n}$-probability for every $h \in \mathcal{R}^k$, where $\{\delta_n\}$ is a sequence of p.d. matrices.

(A.2). There exists an almost surely (a.s.) p.d. random matrix $T(\theta_0)$ such that

$$\mathcal{L}(W_n(\theta_0), T_n(\theta_0) \mid P_{\theta_0,n}) \Longrightarrow \mathcal{L}(W, T(\theta_0))$$

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

Definition 2. Suppose that the sequence of families $\{P_{\theta,n};\ \theta \in \Theta\}$, $n \geqslant 1$, satisfies the LAMN condition at $\theta = \theta_0 \in \Theta$. Then the sequence $\{V_n\}$ of estimators is said to be a sequence of *asymptotically centering sequence* (ACS) of estimators at $\theta = \theta_0 \in \Theta$ if the difference

$$\delta_n(V_n - \theta_0) - T_n^{-1/2}(\theta_0) W_n(\theta_0)$$

converges to zero in $P_{\theta_0,n}$-probability.

2. Differentiability in quadratic mean type regularity condition and the LAMN condition

For $n \geqslant 1$, let $(X_1, X_2, \ldots, X_n)$ be a sequence of random vectors defined on a probability space $(\mathcal{X}, \mathcal{F}, P_\theta)$ where the k-dimensional parameter $\theta \in \Theta$, an open subset of $\mathcal{R}^k$. Let $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$ be the σ-field induced by the random vector $(X_1, X_2, \ldots, X_n)$ and $P_{\theta,n}$ be the restriction of P_θ to $\mathcal{F}_n$. Let θ_0 be the true value of the parameter and we assume that $\theta_0 \in \Theta$. We further assume that, for $j \geq 2$, a regular conditional probability measure of X_j given $(X_1, X_2, \ldots, X_{j-1})$ is absolutely continuous with respect to a σ-finite measure μ_j with a corresponding density $f_j(X_j | X_1, \ldots, X_{j-1}; \theta)$, 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

$$f_j(X_j | X_1, \ldots, X_{j-1}; \theta) = f_j(\theta),\ j \geq 2, \text{ and } f_1(X_1; \theta) = f_1(\theta).$$

Let

$$L_n(X_1, \ldots, X_n; \theta) = \prod_{j=1}^{n} f_j(\theta).$$

To simplify the notation we set

$$L_n(X_1, \ldots, X_n; \theta) = L_n(\theta).$$

We shall call this the likelihood function.

Let

$$\Lambda_n(\theta) = \log \frac{L_n(\theta)}{L_n(\theta_0)} = \sum_{j=1}^{n} \log \frac{f_j(\theta)}{f_j(\theta_0)}$$

which is well defined with $P_{\theta_0,n}$ probability one. We call this the log-likelihood function.

In this section we assume the following set of *assumptions*.

(2.A.1): There are positive definite matrices δ_n, $n \geq 1$, depending neither on θ nor on the observations, and random vectors $\dot{\xi}_j(\theta_0)$, $j \geq 1$, such that, for every $h \in \mathcal{R}^k$,

$$\sum_{j=1}^{n} E\{\int [\xi_{nj}(\theta_0, h) - \tfrac{1}{2} h' \delta_n^{-1} \dot{\xi}_j(\theta_0)]^2 \, d\mu_j\} \to 0$$

as $n \to \infty$, where we set

$$\xi_{nj}(\theta_0, h) = \sqrt{f_j(\theta_0 + \delta_n^{-1} h)} - \sqrt{f_j(\theta_0)}.$$

Define

$$\eta_j(\theta_0) = \begin{cases} \dot{\xi}_j(\theta_0)/\sqrt{f_j(\theta_0)} & \text{if } f_j(\theta_0) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

(2.A.2): $E[\eta_j(\theta_0) | \mathcal{F}_{j-1}] = 0$ for every $j \geq 1$.

(2.A.3) : There exists a measurable function $T(\theta_0)$ mapping $\mathscr{X}$ to the set of $k \times k$ symmetric matrices such that

$$P_{\theta_0}(T(\theta_0) \text{ is p.d.}) = 1$$

and

$$\left| \delta_n^{-1} \sum_{j=1}^{n} E[\eta_j(\theta_0)\eta_j'(\theta_0) \mid \mathscr{F}_{j-1}]\delta_n^{-1} - T(\theta_0) \right| \xrightarrow{P} 0.$$

(2.A.4) : For every $\varepsilon > 0$ and $h \in \mathcal{R}^k$

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

where $I(C)$ denotes the indicator function of the set C.

(2.A.5) : For every $h \in \mathcal{R}^k$, there exists a constant $K > 0$ such that

$$\sup_{n \geqslant 1} \sum_{j=1}^{n} E[\,|h'\delta_n^{-1}\eta_j(\theta_0)|^2] \leqslant K.$$

Remarks (1) : See Section 6 for a discussion of the above assumptions in terms of the first derivatives and for a method of finding out the normalising matrices δ_n, $n \geqslant 1$.

(2) Assumption (2.A.2) is imposed in order to invoke the central limit theorems for martingales. It is possible to relax the assumption (2.A.2) slightly if one uses the central limit theorems for near martingales, as considered by Hall (1977). It is also possible to deduce (2.A.2) from (2.A.1) in some special cases; see e.g. LeCam (1970) and Roussas (1972 and 1979). We were not able to deduce (2.A.2) from (2.A.1) in the general case.

(3) For each $n \geqslant 1$, the quantity $\sum_{j=1}^{n} E[\eta_j(\theta_0)\eta_j'(\theta_0)]$ is called the Fisher information matrix and the quantity $\sum_{j=1}^{n} E[\eta_j(\theta_0)\eta_j'(\theta_0) \mid \mathscr{F}_{j-1}]$ is called the conditional Fisher information matrix.

Following are the main theorems of this section.

Theorem 1 : *Suppose the assumptions* (2.A.1)–(2.A.5) *are satisfied. Then the sequence* $\{P_{\theta,n};\ \theta \in \Theta\}$, $n \geqslant 1$, *satisfies the LAMN condition at* $\theta = \theta_0$ *with*

$$T_n(\theta_0) = \delta_n^{-1} \sum_{j=1}^{n} E[\eta_j(\theta_0)\eta_j'(\theta_0) \mid \mathscr{F}_{j-1}]\,\delta_n^{-1}$$

and

$$W_n(\theta_0) = T_n^{-1/2}(\theta_0)\delta_n^{-1} \sum_{j=1}^{n} \eta_j(\theta_0).$$

Theorem 2 : *Suppose the sequence $\{P_{\theta,n}; \theta \in \Theta\}$, $n \geq 1$, satisfies the LAMN-condition at $\theta = \theta_0$. Let $\{V_n\}$, $n \geq 1$ be a sequence of estimators satisfying the ACS-condition at $\theta = \theta_0$. Then, for every $h \in \mathcal{R}^k$,*

$$\mathcal{L}(T_n(\theta_0), \delta_n(V_n - \theta_0 - \delta_n^{-1}h) \mid P_{\theta_0+\delta_n^{-1}h,\, n}) \Longrightarrow \mathcal{L}(T(\theta_0), T^{-1/2}(\theta_0)W \mid P_{\theta_0}).$$

The proofs of the above Theorem 1 will be given through a series of lemmas; proofs are based on the ideas of LeCam (1970), Roussas (1972 and 1979) and Ibragimov and Khasminskii (1975). We start with the following lemma, the proof of which is essentially contained in corollary (3.8) of McLeish (1974) (see also Lemma (3.1) of Basawa and Scott, (1977).

Lemma 1 : *Suppose the assumptions* (2.A.2)–(2.A.5) *are satisfied. Then, for every $t \in \mathcal{R}^k$,*

$$\Big| \sum_{j=1}^{n} |t'\delta_n^{-1}\eta_j(\theta_0)|^2 - \sum_{j=1}^{n} E[\,|t'\delta_n^{-1}\eta_j(\theta_0)|^2 \mid \mathcal{F}_{j-1}]\Big| \xrightarrow{P} 0.$$

Lemma 2 : *Suppose the assumptions* (2.A.2)–(2.A.4) *are satisfied. Then*

$$(T_n(\theta_0), W_n(\theta_0)) \Longrightarrow (T(\theta_0), W),$$

where $T_n(\theta_0)$ and $W_n(\theta_0)$, $n \geq 1$, are as defined in Theorem 1, and W is a copy of $N(0, I)$ independent of $T(\theta_0)$.

Proof: Since, for every $\varepsilon > 0$ and $t \in \mathcal{R}^k$,

$$E\Big[\max_{j \leq n} |t'\delta_n^{-1}\eta_j(\theta_0)|^2\Big] \leq \varepsilon^2 + \sum_{j=1}^{n} E[\,|t'\delta_n^{-1}\eta_j(\theta_0)|^2 I(|t\delta_n^{-1}\eta_j(\theta_0)| > \varepsilon],$$

we have, by (2.A.4),

$$E\Big[\max_{j \leq n} |t'\delta_n^{-1}\eta_j(\theta_0)|^2\Big] \to 0. \qquad \dots \quad (2.1)$$

Now Lemma 1 and (2.A.3) implies that, for every $t \in R^k$,

$$\sum_{j=1}^{n} |t'\delta_n^{-1}\eta_j(\theta_0)|^2 \xrightarrow{P} t'T(\theta_0)t. \qquad \dots \quad (2.2)$$

Hence by the corollary of Hall (1977) and the remarks preceding the Theorem 2 of Aldous and Eagleson (1978), we have

$$\delta_n^{-1} \sum_{j=1}^{n} \eta_j(\theta_0) \Longrightarrow T^{1/2}(\theta_0)W \text{ (stably)}, \qquad \ldots \quad (2.3)$$

where W is a copy of $N(0, I)$, independent of $T(\theta_0)$. In particular

$$(\delta_n^{-1} \sum_{j=1}^{n} \eta_j(\theta_0), T(\theta_0)) \Longrightarrow (T^{1/2}(\theta_0)W, T(\theta_0)).$$

In view of (2.A.3) we then have

$$(\delta_n^{-1} \sum_{j=1}^{n} \eta_j(\theta_0), T_n(\theta_0)) \Longrightarrow (T^{1/2}(\theta_0)W, T(\theta_0)).$$

This completes the proof by noting that

$$\delta_n^{-1} \sum_{j=1}^{n} \eta_j(\theta_0) = T_n^{1/2}(\theta_0)W_n(\theta_0).$$

The following lemma may be considered as an obvious generalisation of Lemma 5 of LeCam (1974).

Lemma 3: *Suppose the assumption* (2.A.1) *is satisfied. Then, letting* Z_j *for the indicator of the set* $\{f_j(\theta_0) = 0\}$,

$$\sum_{j=1}^{n} E[\int Z_j f_j(\theta_0+\delta_n^{-1}h)\, d\mu_j] \to 0 \qquad \ldots \quad (2.4)$$

and

$$\sum_{j=1}^{n} E[\int Z_j |h'\delta_n^{-1}\dot{\xi}_j(\theta_0)|^2\, d\mu_j] \to 0 \qquad \ldots \quad (2.5)$$

as $n \to \infty$ *for every* $h \in \mathcal{R}^k$.

Proof: Fix $h \in \mathcal{R}^k$. Let Z_{1j} be the indicator of the set $\{f_j(\theta_0) = 0, h'\delta_n^{-1}\dot{\xi}_j(\theta_0) < 0\}$ and Z_{2j} be the indicator of the set $\{f_j(\theta_0) = 0, h'\delta_n^{-1}\dot{\xi}_j(\theta_0) > 0\}$ so that $Z_j = Z_{1j}+Z_{2j}$. We then have

$$\sum_{j=1}^{n} E\{\int Z_{1j}[\xi_{nj}(\theta_0), h)-\tfrac{1}{2}h'\delta_n^{-1}\dot{\xi}_j(\theta_0)]^2\, d\mu_j\}$$

$$\geq \sum_{j=1}^{n} E[\int Z_{1j} f_j(\theta_0+\delta_n^{-1}h)\, d\mu_j]$$

$$+\frac{1}{4} \sum_{j=1}^{n} E[\int Z_{1j} |h'\delta_n^{-1}\dot{\xi}_j(\theta_0)|^2 d\mu_j]$$

By (2.A.1), the left hand side of the above expression tends to zero. Hence we have

$$\sum_{j=1}^{n} E[\int Z_{1j} | h'\delta_n^{-1}\dot{\xi}_j(\theta_0)|^2 d\mu_j] \to 0. \qquad \ldots \quad (2.6)$$

Now let $t < 0$. Then

$$\sum_{j=1}^{n} E\{\int Z_{2j}[\xi_{nj}(\theta_0, th) - \frac{t}{2} h'\delta_n^{-1}\dot{\xi}_j(\theta_0)]^2 d\mu_j\}$$

$$\geqslant \sum_{j=1}^{n} E[\int Z_{2j} f_j(\theta_0 + \delta_n^{-1} th)\, d\mu_j]$$

$$+\frac{t^2}{4} \sum_{j=1}^{n} E[\int Z_{2j} | h'\delta_n^{-1} \dot{\xi}_j(\theta_0)|^2 d\mu_j].$$

By (2.A.1), this implies that

$$\sum_{j=1}^{n} E[\int Z_{2j} h'\delta_n^{-1} \dot{\xi}_j(\theta_0)|^2 d\mu_j] \to 0. \qquad \ldots \quad (2.7)$$

Combining (2.6) and (2.7) we have

$$\sum_{j=1}^{n} E[\int Z_j h'\delta_n^{-1} \dot{\xi}_j(\theta_0)|^2 d\mu_j] \to 0.$$

This proves (2.5). To prove (2.4) consider the inequality

$$\sum_{j=1}^{n} E\left[\int Z_j |\xi_{nj}^2(\theta_0, h) - \frac{1}{4} | h'\delta_n^{-1} \xi_j(\theta_0)|^2| d\mu_j\right]$$

$$\leqslant 2 \sum_{j=1}^{n} E\{\int Z_j[\xi_{nj}(\theta_0, h) - \frac{1}{2} h'\delta_n^{-1} \dot{\xi}_j(\theta_0)]^2 d\mu_j\}$$

$$+\frac{1}{4} \sum_{j=1}^{n} E[\int Z_j | h'\delta_n^{-1} \dot{\xi}_j(\theta_0)|^2 d\mu_j]; \qquad \ldots \quad (2.8)$$

here we have used the inequality

$$|c^2 - d^2| \leqslant (1+\alpha)|c-d|^2 + d^2/\alpha, \ \alpha > 0 \text{ and } c, d \in R. \qquad \ldots \quad (2.9)$$

The first term of the right hand side of (2.8) tends to zero by (2.A.1) while the second term tends to zero by (2.5).. This completes the proof of the lemma.

To simplify the notation we set, in what follows,

$$\eta_{nj}(\theta_0, h) = \begin{cases} \sqrt{\dfrac{f_j(\theta_0+\delta_n^{-1}h)}{f_j(\theta_0)}}-1 & \text{if } f_j(\theta_0) \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 4: *Suppose the assumptions* (2.A.1) *and* (2.A.5), *are satisfied. Then*

$$\sum_{j=1}^{n} \int |\xi_{nj}^2(\theta_0, h)-\frac{1}{4}|h'\delta_n^{-1}\dot{\xi}_j(\theta_0)|^2|\,d\mu_j \xrightarrow{P} 0 \qquad \ldots \quad (2.10)$$

and

$$\sum_{j=1}^{n} E\left[\,|\eta_{nj}^2(\theta_0, h)-\frac{1}{4}|h'\delta_n^{-1}\eta_j(\theta_0)|^2|\,\right] \to 0. \qquad \ldots \quad (2.11)$$

Proof: Using the inequality (2.9), we have

$$\sum_{j=1}^{n} E\left[\int |\xi_{nj}^2(\theta_0, h)-\frac{1}{4}|h'\delta_n^{-1}\dot{\xi}_j(\theta_0)|^2|\,d\mu_j\right]$$

$$\leqslant (1+\alpha)\sum_{j=1}^{n} E\left\{\int \left[\xi_{nj}(\theta_0, h)-\frac{1}{2}\,h'\delta_n^{-1}\dot{\xi}_j(\theta_0)\right]^2 d\mu_j\right\}$$

$$+\frac{1}{4\alpha}\sum_{j=1}^{n} E\{\int |h'\delta_n^{-1}\dot{\xi}(\theta_0)|^2\,d\mu_j\}. \quad \alpha > 0. \qquad \ldots \quad (2.12)$$

For each fixed α, the first term of the right hand side of (2.12) tends to zero as $n \to \infty$ by (2.A.1). Now consider

$$\sum_{j=1}^{n} E[\int |h'\delta_n^{-1}\dot{\xi}_j(\theta_0)|^2 d\mu_j] = \sum_{j=1}^{n} E[|h'\delta_n^{-1}\eta_j(\theta_0)|^2]$$

$$+\sum_{j=1}^{n} E[\int Z_j\,|h'\delta_n^{-1}\dot{\xi}_j(\theta_0)|^2 d\mu_j],$$

where Z_j is the indicator of the set $\{f_j(\theta_0) = 0\}$. Hence we see that the second term of the right hand side of (2.12) tends to zero by first letting $n \to \infty$ and then $\alpha \to \infty$, by (2.A.5) and (2.5). Thus the right hand side of (2.12) tends to zero by first letting $n \to \infty$ and then $\alpha \to \infty$. This proves

(2.10), since convergence in the first mean implies the convergence in probability. Similarly (2.11) is proved by considering the inequality

$$\sum_{j=1}^{n} E\left[\,|\eta_{nj}^2(\theta_0, h) - \frac{1}{2}\, h'\delta_n^{-1}\,\eta_j(\theta_0)|^2\right]$$

$$\leqslant (1+\alpha)\sum_{j=1}^{n} E\left\{\left[\eta_{nj}(\theta_0, h) - \frac{1}{2}\, h'\delta_n^{-1}\,\eta_j(\theta_0)\right]^2\right\}$$

$$+\frac{1}{4\alpha}\sum_{j=1}^{n} E[\,|h'\delta_n^{-1}\eta_j(\theta_0)|^2], \quad \alpha > 0,$$

and by noting that

$$\sum_{j=1}^{n} E\left\{\left[\eta_{nj}(\theta_0, h) - \frac{1}{2}\, h'\delta_n^{-1}\,\eta_j(\theta_0)\right]^2\right\}$$

$$= \sum_{j=1}^{n} E\left\{\int_{(f_j(\theta_0) \neq 0)} \left[\xi_{nj}(\theta_0, h) - \frac{1}{2}\, h'\delta_n^{-1}\,\dot{\xi}_j(\theta_0)\right]^2 d\mu_j\right\} \to 0$$

$$\text{as } n \to \infty. \qquad \ldots \quad (2.13)$$

This completes the proof of the lemma.

Lemma 5 : *Under the assumptions of Theorem* 1, *we have*

$$\left|\sum_{j=1}^{n} \eta_{nj}^2(\theta_0, h) - \frac{1}{4}\, h\, T(\theta_0) h\right| \xrightarrow{P} 0.$$

Proof : By assumption (2.A.3) and Lemma 1 we have

$$\left|\sum_{j=1}^{n} |h'\delta_n^{-1}\eta_j(\theta_0)|^2 - h' T(\theta_0) h\right| \xrightarrow{P} 0.$$

Hence it is enough to show that

$$\sum_{j=1}^{n}\left|\eta_{nj}^2(\theta_0, h) - \frac{1}{4}\,|h'\delta_n^{-1}\eta_j(\theta_0)|^2\right| \xrightarrow{P} 0.$$

This follows from (2.11) by applying Chebyshev's inequality.

Lemma 6 : *Under the assumptions of Theorem* 1, *we have*

$$\max_{j \leqslant n} |\eta_{nj}(\theta_0, h)| \xrightarrow{P} 0 \qquad \ldots \quad (2.14)$$

and

$$\sum_{j=1}^{n} |\eta_{nj}(\theta_0, h)|^2 \xrightarrow{P} 0. \qquad \ldots \quad (2.15)$$

Proof: For any $\varepsilon > 0$, consider

$$P[\max_{j\leq n}|\eta_{nj}(\theta_0, h)| > \varepsilon]$$

$$\leq \sum_{j=1}^{n} P[|\eta_{nj}(\theta_0, h)| > \varepsilon]$$

$$\leq \sum_{j=1}^{n} P\left[\left|\eta_{nj}(\theta_0, h)-\frac{1}{2}h'\delta_n^{-1}\eta_j(\theta_0)\right| > \varepsilon/2\right]$$

$$+\sum_{j=1}^{n} P\left[\left|\frac{1}{2}h'\delta_n^{-1}\eta_j(\theta_0)\right| > \varepsilon/2\right]. \qquad \ldots \quad (2.16)$$

Now (2.13) implies that, by applying Chebyshev's inequality,

$$\sum_{j=1}^{n} P\left[\left|\eta_{nj}(\theta_0, h)-\frac{1}{2}h'\delta_n^{-1}\eta_j(\theta_0)\right| > \varepsilon/2\right] \to 0. \qquad \ldots \quad (2.17)$$

It is easily seen that the assumption (2.A.4) implies

$$\sum_{j=1}^{n} P\left[\left|\frac{1}{2}h'\delta_n^{-1}\eta_j(\theta_0)\right| > \varepsilon/2\right] \to 0. \qquad \ldots \quad (2.18)$$

Combining (2.16), (2.17) and (2.18) we see that (2.14) is proved. To prove (2.15) consider

$$\sum_{j=1}^{n}|\eta_{nj}(\theta_0, h)|^3 \leq \max_{j\leq n}|\eta_{nj}(\theta_0, h)| \sum_{j=1}^{n}\eta_{nj}^2(\theta_0, h).$$

Hence (2.15) follows by applying Lemma 5 and (2.14). This completes the proof of the lemma.

Lemma 7: *Under the assumptions of Theorem* 1, *we have*

$$\left|2\sum_{j=1}^{n}\eta_{nj}(\theta_0, h)-h'\delta_n^{-1}\sum_{j=1}^{n}\eta_j(\theta_0)+\frac{1}{4}h'T_n(\theta_0)h\right| \xrightarrow{P} 0.$$

Proof: Consider

$$E[\eta_{nj}(\theta_0, h)|\mathcal{F}_{j-1}]$$

$$=\int\sqrt{f_j(\theta_0+\delta_n^{-1}h)f_j(\theta_0)}\,d\mu_j-1$$

$$=-\frac{1}{2}\int\xi_{nj}^2(\theta_0, h)\,d\mu_j.$$

By (2.10), (2.5) and Lemma 1, this implies that

$$\left| 2 \sum_{j=1}^{n} E[\eta_{nj}(\theta_0, h) | \mathcal{F}_{j-1}] + \frac{1}{4} h' T_n(\theta_0) h \right| \xrightarrow{P} 0.$$

Hence, since $E[\eta_j(\theta_0) | \mathcal{F}_{j-1}] = 0$, $j \geqslant 1$, it is enough to show that

$$\sum_{j=1}^{n} [Y_j - E(Y_j | \mathcal{F}_{j-1})] \xrightarrow{P} 0,$$

where we set

$$Y_j = 2\left[\eta_{nj}(\theta_0, h) - \frac{1}{2} h' \delta_n^{-1} \eta_j(\theta_0)\right].$$

Since this summands are martingale differences, we have

$$E\left\{ \sum_{j=1}^{n} [Y_j - E(Y_j | \mathcal{F}_{j-1})]\right\}^2$$

$$= \sum_{j=1}^{n} E[Y_j - E(Y_j | \mathcal{F}_{j-1})]^2$$

$$\leqslant \sum_{j=1}^{n} E[|Y_j|^2] \to 0 \text{ as } n \to \infty \text{ by (2.13).}$$

This completes the proof of the lemma by applying Chebyshev's inequality.

Proof of Theorem 1 : In view of (2.14) and Taylor's expansion we have

$$\Lambda_n(\theta_0 + \delta_n^{-1} h)$$

$$= 2 \sum_{j=1}^{n} \log (1 + \eta_{nj}(\theta_0, h))$$

$$= 2 \sum_{j=1}^{n} \eta_{nj}(\theta_0, h) - \sum_{j=1}^{n} \eta_{nj}^2(\theta_0, h)$$

$$+ \sum_{j=1}^{n} \alpha_{nj} | \eta_{nj}(\theta_0, h)|^3$$

with probability tending to one, where $|\alpha_{nj}| \leqslant 1$. By (2.15) we then have

$$\Lambda_n(\theta_0 + \delta_n^{-1} h) - 2 \sum_{j=1}^{n} \eta_{nj}(\theta_0, h) + \sum_{j=1}^{n} \eta_{nj}^2(\theta_0, h) \Big| \xrightarrow{P} 0.$$

Further, Lemma 1 and (2.11) implies that

$$\left| \sum_{j=1}^{n} \eta_{nj}^2(\theta_0, h) - \frac{1}{4} h'T_n(\theta_0)h \right| \xrightarrow{P} 0.$$

Hence the result follows by Lemma 2 and Lemma 7.

We shall now prove Theorem 2; the proof will be given through a series of lemmas. We shall first give the following proposition which establishes the contiguity of the sequences $\{P_{\theta_0, n}\}$ and $\{P_{\theta_0+\delta_n^{-1}h, n}\}$, $h \in \mathcal{R}^k$; detailed discussion of contiguity can be found in LeCam (1960).

Proposition 1: *Suppose the sequence* $\{P_{\theta,n} ; \theta \in \Theta\}$, $n \geq 1$, *satisfies the LAMN-condition at* $\theta = \theta_0$. *Then the sequences* $\{P_{\theta_0, n}\}$ *and* $\{P_{\theta_0+\delta_n^{-1}h, n}\}$, $h \in \mathcal{R}^k$, *are contiguous.*

Proof: LAMN-condition at $\theta = \theta_0$ implies that

$$\Lambda_n(\theta_0+\delta_n^{-1}h) \Longrightarrow h'T^{1/2}(\theta_0)W - \frac{1}{2} hT(\theta_0)h$$

for every $h \in \mathcal{R}^k$, where W is a copy of $N(0, 1)$, independent of $T(\theta_0)$. We have to show that

$$E\left[\exp(h'T^{1/2}(\theta_0)W - \frac{1}{2} h'T(\theta_0, h)\right] = 1.$$

Using the independence of $T(\theta_0)$ and W it is easily seen that

$$E^T\left[\exp(h'T^{1/2}(\theta_0)W - \frac{1}{2} h'T(\theta_0)h)\right] = 1,$$

where E^T denotes the conditional expectation given $T(\theta_0)$. This proves the proposition.

Lemma 8: *Suppose that the sequence* $\{P_{\theta, n} ; \theta \in \Theta\}$, $n \geq 1$, *satisfies LAMN-condition at* $\theta = \theta_0$. *Then, for every* $h \in \mathcal{R}^k$,

$$\mathcal{L}(T_n(\theta_0), T_n^{-1/2}(\theta_0)W_n(\theta_0) \mid P_{\theta_0+\delta_n^{-1}h, n}) \Longrightarrow \mathcal{L}(T(\theta_0), T^{-1/2}(\theta_0)W + h \mid P_{\theta_0}).$$

Proof: For simplicity we assume that the parameter space is of one dimension. Then according to the statement (6) of Theorem 2.1 of LeCam (1960), we have for every u, $v \in \mathcal{R}$,

$$E_{\theta_0+\delta_n^{-1}h}[\exp(iuT_n^{-1/2}(\theta_0)W_n(\theta_0)+ivT_n(\theta_0))]$$

$$\to E\Big[\exp\Big(iuT^{-1/2}(\theta_0)W+ivT(\theta_0)+hT^{1/2}(\theta_0)W-\frac{h^2}{2}T(\theta_0)\Big)\Big]$$

$$= E\Big\{\exp(ivT(\theta_0)+iuh)E^T\Big[\exp\Big(iu(T^{-1/2}(\theta_0)W-h)$$

$$+hT^{1/2}(\theta_0)W-\frac{h^2}{2}T(\theta_0)\Big)\Big]\Big\}.$$

Using the independence of $T(\theta_0)$ and W, it follows that

$$E^T\Big[\exp\Big(iu(T^{-1/2}(\theta_0)W-h)+hT^{1/2}(\theta_0)W-\frac{h^2}{2}T(\theta_0)\Big)\Big]$$

$$= E^T[\exp(iuT^{-1/2}(\theta_0)W)].$$

Hence we see that, for every $u, v \in \mathcal{R}$,

$$E_{\theta_0+\delta_n^{-1}h}[\exp(iuT_n^{-1/2}(\theta_0)W_n(\theta_0)+ivT_n(\theta_0))]$$

$$\to E[\exp(iu(T^{-1/2}(\theta_0)W+h)+ivT(\theta_0))].$$

This gives the required result.

Proof of Theorem 2: Since the sequence of estimators $\{V_n\}$, $n > 1$, satisfies ACS condition at $\theta = \theta_0$, we have

$$[\delta_n(V_n-\theta_0)-T_n^{-1/2}(\theta_0)W_n(\theta_0)] \xrightarrow{P} 0.$$

Since the sequence $\{P_{\theta_0 n}\}$ and $\{P_{\theta_0+\delta_n^{-1}h,n}\}$ $h \in \mathcal{R}^k$, are contiguous, $|\delta_n(V_n-\theta_0)-T_n^{-1/2}(\theta_0)W_n(\theta_0)|$ tends to zero in $P_{\theta_0+\delta_n^{-1}h,n}$ probability also. Hence by Lemma 9, we have

$$\mathcal{L}(T_n(\theta_0), \delta_n(V_n-\theta_0) \mid P_{\theta_0+\delta_n^{-1}h,n}) \Longrightarrow \mathcal{L}(T(\theta_0), T_n^{-1/2}(\theta_0)W+h \mid P_{\theta_0}).$$

This completes the proof of the theorem.

3. Decomposition of the limit distribution

We shall first prove a conditional decomposition result; the proof is based on Bickel's simple short proof of Hajek's convolution theorem (See Roussas, 1972 for a published version of Bickel's proof).

It is important to note that we do not assume the existence of the limit distribution and therefore we prove our conditional convolution result for the limit distribution, possibly substochastic, of any convergent subsequence.

Let $\{V_n\}$ be a sequence of estimators satisfying the invariance restriction stated in the theorem given below Let $\{r\}$ be subsequence and H_{θ_0} be a (sub-stochastic) measure such that

$$\mathcal{L}(T_r(\theta_0), \delta_r(V_r-\theta_0) \mid P_{\theta_0, r}) \Longrightarrow H_{\theta_0}$$

Let $\mathcal{L}_{\theta_0}$ be the law of $T(\theta_0)$ and let $\bar{\mathcal{R}}^k$ be one-point compactification of $\mathcal{R}^k$. Define

$$\bar{H}_{\theta_0}(B\times\{\infty\}) = \mathcal{L}_{\theta_0}(B) - H_{\theta_0}(B\times\mathcal{R}^k)$$

and

$$\bar{H}_{\theta_0}(B\times A) = H_{\theta_0}(B\times A)$$

for every Borel sets $B \subseteq \mathcal{R}^{k^2}$ and $A \subseteq \mathcal{R}^k$. Then $\bar{H}_{\theta_0}$ is a probability measure defined on $\mathcal{R}^{k^2}\times\bar{\mathcal{R}}^k$ induced by H_{θ_0}. Let $\bar{\mathcal{L}}_{T(\theta_0)}$ be a regular conditional probability measure (on $\bar{\mathcal{R}}^k$) such that

$$\bar{H}_{\theta_0}(C) = \int I(C)\bar{\mathcal{L}}_t(dx)\mathcal{L}_{\theta_0}(dt)$$

for every Borel set $C \subseteq \mathcal{R}^{k^2}\times\bar{\mathcal{R}}^k$.

Theorem 3: *Suppose that the sequence of families* $\{P_{\theta,n};\theta\in\Theta\}$, $n\geq 1$, *satisfies the LAMN condition at* $\theta=\theta_0\in\Theta$. *Let* $\{V_n\}$ *be a sequence of estimators such that the difference*

$$E[f(T_n(\theta_0), \delta_n(V_n-\theta_0-\delta_n^{-1}h)) \mid P_{\theta_0+\delta_n^{-1}h,n}]$$
$$-E[f(T_n(\theta_0), \delta_n(V_n-\theta_0)) \mid P_{\theta_0,n}]$$

converges to zero for every $h \in \mathcal{R}^k$ and for every continuous functions f vanishing outside compacts. Let the regular conditional probability measure $\bar{\mathcal{L}}_{T(\theta_0)}$ be as above and let $\mathcal{L}_{T(\theta_0)}$ be the restriction of $\bar{\mathcal{L}}_{T(\theta_0)}$ to $\mathcal{R}^k$. Then there exists (sub-stochastic) kernel $K_{T(\theta_0)}$ such that

$$\mathcal{L}_{T(\theta_0)} = K_{T(\theta_0)} \bullet N(0, T^{-1}(\theta_0)) \text{ a.s.}$$

The following familiar version, in which the existence of the limit distribution is assumed, is immediate from the above Theorem 1.

Corollary 1 : *Suppose that the sequence of families $\{P_{\theta,n};\ \theta \in \Theta\}$, $n \geqslant 1$ satisfies the LAMN-condition at $\theta = \theta_0$. Let $\{V_n\}$ be a sequence of estimators such that, for every $h \in \mathcal{R}^k$,*

$$\mathcal{L}(T_n(\theta_0), \delta_n(V_n - \theta_0 - \delta_n^{-1}h) \mid P_{\theta_0+\delta_n^{-1}h, n}) \Longrightarrow \mathcal{L}(T(\theta_0), V(\theta_0))$$

for some random k-vector $V(\theta_0)$. Let $\mathcal{L}_{T(\theta_0)}$ be a regular conditional probability measure of $V(\theta_0)$ given $T(\theta_0)$ Then there exists a stochastic kernel $K_{T(\theta_0)}$ such that

$$\mathcal{L}_{T(\theta_0)} = K_{T(\theta_0)} \bullet N(0, T^{-1}(\theta_0)). \text{ a.s.}$$

Proof of Theorem 3 : For simplicity assume that $\dim \Theta = 1$. Let

$$f(u, z, x, y) = (e^{iux} - 1)(e^{izy} - 1)/ixiy, \quad u, z, x, y \in R.$$

Note that

$$f(u, z, x, y) \to 0 \text{ as } |(x, y)| \to \infty \text{ for every } u, z \in R.$$

Hence, we have, in view of the invariance restriction

$$\int_{\mathcal{X}_r} f(u, z, T_r(\theta_0), \delta_r(V_r - \theta_0 - \delta_r^{-1}h))\, dP_{\theta_0+\delta_r^{-1}h, r}$$
$$\to \int_{R^2} f(u, z, t, v) \Pi_{\theta_0}(dt, dv) \qquad \ldots \quad (3.1)$$

for every $u, z \in R$.

Now there exists a subsequence $\{m\} \subseteq \{r\}$ and a (sub-stochastic) measure Q_{θ_0} such that

$$\mathcal{L}(T_m(\theta_0), W_m(\theta_0), \delta_m(V_m - \theta_0) \mid P_{\theta_0, m}) \Longrightarrow Q_{\theta_0}.$$

Without loss of generality assume that $\{m\} = \{r\}$. In view of contiguity we can further assume, without loss of generality, that

$$P_{\theta_0+\delta_n^{-1}h,n} \ll P_{\theta_0,n} \text{ for every } n \geq 1 \text{ and } h \in R.$$

Hence the l.h.s. of (3.1) can be written as

$$\int_{\mathcal{X}_r} f(u, z, T_r(\theta_0), \delta_r(V_r-\theta_0)-h) \frac{dP_{\theta+\delta_n^{-1}h,r}}{dP_{\theta_0,r}} dP_{\theta_0,r}.$$

and it is not difficult to see from the LAMN condition and the corresponding contiguity condition, that this converges to

$$\int_{R^3} f(u, z, t, (v-h)) \exp\left(ht^{1/2}w-\frac{h^2}{2}t\right) Q_{\theta_0}(dt, dw, dv) \qquad \ldots \quad (3.2)$$

for every $u, z, h \in R$. That is, we have the equality

$$\int_{R^2} f(u, z, t, v) \Pi_{\theta_0}(dt, dv)$$

$$= \int_{R^3} f(u, z, t, (v-h)) \exp\left(ht^{1/2}w-\frac{h^2}{2}t\right) Q_{\theta_0}(dt, dw, dv) \qquad \ldots \quad (3.3)$$

for every $u, z, h \in R$.

Now define

$$\bar{Q}_{\theta_0}(A \times B \times \{\infty\}) = F_{\theta_0}(A \times B) - Q_{\theta_0}(A \times B \times \mathcal{R}^k)$$

and

$$\bar{Q}_{\theta_0}(A \times B \times C) = Q_{\theta_0}(A \times B \times C)$$

for every Borels subsets A, B and C of R, where F_{θ_0} is the law of $(T(\theta_0), W)$. Then $\bar{Q}_{\theta_0}$ is a probability measure on $R^2 \times R$. Let Π_{θ_0} be as defined earlier. Note that $\bar{\Pi}_{\theta_0}$ and F_{θ_0} are marginals of $\bar{Q}_{\theta_0}$. Now (3.3) can be written as

$$\int_{R^2} f(u, z, t, v) \bar{\Pi}_{\theta_0}(dt, dv)$$

$$= \int_{R^3} f(u, z, t, (v-h)) \exp\left(ht^{1/2}w-\frac{h^2}{2}t\right) \bar{Q}_{\theta_0}(dt, dw, dv)$$

for every $u, z, h \in R$. This implies that (cf. Loève (1963, p. 189))

$$\int_{R^2} \exp(iut+iuz)\bar{U}_{\theta_0} d(t, dv)$$

$$= \int_R \exp(iu+iz(v-h)) \exp\left(ht^{1/2}w-\frac{h^2}{2}t\right)\bar{Q}_{\theta_0}(dt, dw, dv) \quad \ldots \quad (3.4)$$

for every $u, z, h \in R$. Let $\bar{\mathcal{L}}_{T(\theta_0)}$ be the regular conditional probability measure as defined earlier and let $\mathcal{L}_{\theta_0}$ be the law of $T(\theta_0)$. Then (3.4) can be written as, for every $u, z, h \in R$,

$$\int_R \exp(iut)\left[\int_R \exp(izv)\bar{\mathcal{L}}_t(dv)\right]\mathcal{L}_{\theta_0}(dt)$$

$$= \int_R \exp(iut)\left[\int_{R^2} \exp(iz(v-h) \exp\left(dt^{1/2}w-\frac{h^2}{2}t\right)\bar{\mathcal{L}}_t(dw, dv)\right]\mathcal{L}_{\theta_0}(dt) \quad \ldots \quad (3.5)$$

for some regular conditional probability measure $\bar{\mathcal{L}}_{T(\theta_0)}$

Now, a simple continuity argument shows that (3.5) entails that

$$\int_R \exp(izv)\bar{\mathcal{L}}_{T(\theta_0)}(dv)$$

$$= \int_{R^2} \exp(iz(v-h) \exp\left(hT^{1/2}(\theta_0)w-\frac{h^2}{2}T(\theta_0)\right)\bar{\mathcal{L}}'_{T(\theta_0)}(dw, dv) \text{ a.s.} \quad \ldots \quad (3.6)$$

for every $u, h \in R$. In what follows assume that $T(\theta_0)$ is fixed. It can be shown that the r.h.s. of (3.6) is analytic in h. Hence, by replacing h by ih, we have

$$\int_R \exp(izv)\bar{\mathcal{L}}_{T(\theta_0)}(dv)$$

$$\int_{R^2} \exp(izv+zh) \exp\left(ihT^{1/2}(\theta_0)w+\frac{h^2}{2}T(\theta_0)\right)\bar{\mathcal{L}}'_{T(\theta_0)}(dw, dv)$$

for every $z, h \in R$. Setting $h = -T^{-1}(\theta_0)z$ in this equality we have

$$\int_R \exp(izv)\bar{\mathcal{L}}_{T(\theta_0)}(dv)$$

$$= \exp\left(-\frac{1}{2}T^{-1}(\theta_0)z^2\right)\int_{R^2} \exp[iz(v-T^{-1/2}(\theta_0)w)]\bar{\mathcal{L}}'_{T(\theta_0)}(dw, dv)$$

for every z. This proves the result.

By applying the above conditional convolution result we obtain the following two propositions.

Proposition 2 : *Assume that the sequence of families* $\{P_{\theta, n}, \theta \in \Theta\}$, $n \geq 1$, *satisfies the LAMN condition at* $\theta = \theta_0 \in \Theta$. *Let* $\{V_n\}$ *be a sequence of estimators satisfying the invariance condition of Theorem 3. Let* l $\mathcal{R}^k \to R$ *be a loss function of the form* $l(0) = 0$, $l(x) = l(|x|)$ *and* $l(|x|) \leq l(|y|)$ *if* $|x| \leq |y|$. *Then*

$$\liminf_{n \to \infty} E_{\theta_0}[l(\delta_n(V_n - \theta_0))] \geq E[l(T^{-1/2}(\theta_0)W)].$$

Proof: The proof is an easy consequence of Theorem 3.

The proof of the statement (i) of the following Proposition 3 is immediate from Corollary 1. The proof of the statement (ii) is also a consequence of Corollary 1 Its proof is essentially contained in Roussas (1972), pp 141–147).

Proposition 3 : *Assume that the sequence of families* $\{P_{\theta, n}; \theta \in \Theta\}$, $n \geq 1$, *satisfies the LAMN condition at* $\theta = \theta_0 \in \Theta$ *Let* $\{V_n\}$ *be a sequence of estimators satisfying the invariance condition of Corollary* 1. *Let* l *be a loss function as in Proposition* 1 *Then*

(i) $$E[l(V(\theta_0))] \geq E[l(T^{-1/2}(\theta_0)W)],$$

and

(ii) *for every* $q \in \mathcal{R}^k$ *and* $t_1, t_2 > 0$,

$$P[-t_1 < q'V(\theta_0) < t_2] \leq P[-t_1 < q'T^{-1/2}(\theta_0)W < t_2]$$

provided, for every $q \in \mathcal{R}^k$,

$$P[q'V(\theta_0) \geq 0 \,|\, T(\theta_0)] \geq \frac{1}{2}$$

and

$$P[q'V(\theta_0) \leq 0 \,|\, T(\theta_0)] \geq \frac{1}{2}.$$

Remark 1. Note that the usual examples (see e.g., LeCam, 1953)) show that the invariance restriction cannot be relaxed in Theorem 3 if one tries to establish the conditional convolution result for all points of the parameter space. However, it is possible to obtain the conditional convolution result, without the invariance restriction, for almost all points of the parameter space; this is done in Jeganathan (1981).

Remark 2. A result similar to the result of Proposition 2 was earlier obtained by Heyde (1978) by different arguments under some specific assumptions and with special reference to maximum likelihood estimators. See also Basawa and Scott (1979).

Note that in Proposition 2 we have imposed the invariance restriction on the sequence $\{T_n(\theta_0), \delta_n(V_n-\theta_0)\}$. It is enough to impose the invariance restriction on $\delta_n(V_n-\theta_0)$; for this and several other results concerning the asymptotic properties of risk functions, see Jeganathan (1980a).

4. An exponential approximation result

In this section we establish an exponential approximation result analogous to Theorem 3.1 of LeCam (1960). This result will be used repeatedly in Section 5 of the present paper. Also it serves as a powerful tool in several other places; see, e.g., Jeganathan (1980a) where it is used to extend certain basic results of LeCam and Hajek concerning asymptotic properties of risk functions and a certain kind of posterior approximations.

Further, this result in particular implies that the sequence of random vectors and matrices of the LAMN families forms a sequence of locally asymptotically sufficient statistics.

Proposition 4 : *Assume that the sequence of families* $\{P_{\theta,n};\theta\in\Theta\}$, $n\geq 1$, *satisfies the LAMN condition at* $\theta=\theta_0\in\Theta$. *Then there exist*

(i) *an increasing sequence* $\{k_n\}$ *tending to infinity as* $n\to\infty$,

(ii) *functions* $C_n:\Theta\times\mathcal{R}^k\to R$ *such that*

$$\sup_{|h|\leq a}|C_n(\theta_0,h)-1|\to 0 \text{ as } n\to\infty$$

for every $a>0$, *such that the measures* $Q_n(\theta_0,h)|\mathcal{A}_n$, $Q_n(\theta_0,h)\ll P_{\theta_0,n}$ *defined by*

$$\frac{dQ_n(\theta_0,h)}{dP_{\theta_0,n}}=C_n(\theta_0,h)\exp\left[h'T_n^{\frac{1}{2}}(\theta_0)W_n^*(\theta_0)-\frac{1}{2}h'T_n(\theta_0)h\right]$$

with

$$W_n^*(\theta_0)=W_n(\theta_0)I(|T_n^{\frac{1}{2}}(\theta_0)W_n(\theta_0)|\leq k_n),$$

are probability measures and satisfy

$$\|P_{\theta_0+\delta_n^{-1}h,n}-Q_n(\theta_0,h)\|\to 0 \text{ as } n\to\infty$$

for every $h\in\mathcal{R}^k$.

Proof: Define, for $\alpha > 0$,

$$W_n^\alpha(\theta_0) = W_n(\theta_0) I(|T_n^{\frac{1}{2}}(\theta_0) W_n(\theta_0)| < \alpha)$$

and

$$W^\alpha(\theta_0) = W I(|T^{1/2}(\theta_0) W| < \alpha).$$

There is a dense set of values of α for which

$$\mathcal{L}(T_n(\theta_0), W_n^\alpha(\theta_0) | P_{\theta_0, n}) \Longrightarrow \mathcal{L}(T(\theta_0), W^\alpha(\theta_0)).$$

For any such α, we have

$$\sup_{|h| \leq \alpha} \Big| E\left[\exp\left(h' T_n^{1/2}(\theta_0) W_n^\alpha(\theta_0) - \frac{1}{2} h' T_n(\theta_0) h\right)\right]$$

$$- E\left[\exp\left(h' T^{1/2}(\theta_0) W^\alpha(\theta_0) - \frac{1}{2} h' T(\theta_0) h\right)\right]\Big| \to 0 \text{ as } n \to \infty,$$

since the family of functions, defined on the space of k-vectors and $k \times k$ p. matrices,

$$\left\{(x, D) \to \exp\left(h'x - \frac{1}{2} h' D h\right) : |h| \leqslant \alpha\right\}$$

is uniformly bounded and equicontinuous whenever the domain of x is bounded Hence by a standard diagonal argument one can choose an increasing sequence $\{k_n\}$ tending to infinity such that

$$\sup_{|h| < k_n} \Big| E\left[\exp\left(h' T_n^{1/2}(\theta_0) W_n^*(\theta_0) - \frac{1}{2} h' T_n(\theta_0) h\right)\right]$$

$$- E\left[\exp\left(h' T^{1/2}(\theta_0) W^{k_n}(\theta_0) - \frac{1}{2} h' T(\theta_0) h\right)\right]\Big| \to 0 \text{ as } n \to \infty,$$

where $W_n^*(\theta_0) = W_n^{k_n}(\theta_0)$, and, hence

$$\sup_{|h| \leq b} \Big| E\left[\exp\left(h' T_n^{1/2}(\theta_0) W_n^*(\theta_0) - \frac{1}{2} h' T_n(\theta_0) h\right)\right]$$

$$- E\left[\exp\left(h' T^{1/2}(\theta_0) W^{k_n}(\theta_0) - \frac{1}{2} h' T(\theta_0) h\right)\right]\Big| \to 0, \quad \dots \quad (4.1)$$

as $n \to \infty$ for every $b > 0$.

We now show that, for every $b > 0$,

$$\sup_{|h| \leq b} \Big| E\left[\exp\left(h' T^{1/2}(\theta_0) W^{k_n}(\theta_0) - \frac{1}{2} h' T(\theta_0) h\right)\right] - 1 \Big| \to 0 \quad \dots \quad (4.2)$$

as $n \to \infty$. Now note that

$$\sup_{|h| \leq b} \left| E\left[\exp\left(h'T^{1/2}(\theta_0)W^{k_n}(\theta_0) - \frac{1}{2} h'T(\theta_0)h\right)\right] - 1 \right|$$

$$\leq E\left[\sup_{|h| \leq b} \left| E^T\left[\exp(h'T^{1/2}(\theta_0)W^{k_n}(\theta_0) - \frac{1}{2} h'T(\theta_0)h\right] - 1\right|\right]$$

and

$$\exp\left(h'T^{1/2}(\theta_0)W^{k_n}(\theta_0) - \frac{1}{2} h'T(\theta_0)h\right)$$

$$\leq \exp\left(h'T^{1/2}(\theta_0)W - \frac{1}{2} h'T(\theta_0)h\right) + 1$$

for every $n \geqslant 1$ and $h \in \mathcal{R}^k$ and hence, using the independence of W and $T(\theta_0)$,

$$\sup_{|h| < b} E^T\left[\exp(h'T^{1/2}(\theta_0)W^{k_n}(\theta_0) - \frac{1}{2} h'T(\theta_0)h)\right] \leqslant 2$$

for every $b > 0$ and $n \geqslant 1$. Hence (4.2) will follow if we show that, for each fixed $T(\theta_0)$,

$$\sup_{|h| < b} \left| E^T \exp\left[(h'T^{1/2}(\theta_0)W^{k_n}(\theta_0) - \frac{1}{2} h'T(\theta_0)h)\right] - 1\right| \to 0.$$

This is quite easy to see. From (4.1) and (4.2) we now have, for every $b > 0$,

$$\sup_{|h| < b} \left| E\left[\exp(h'T_n^{1/2}(\theta_0)W_n^*(\theta_0) - \frac{1}{2} h'T_n(\theta_0)h)\right] - 1\right| \to 0 \quad \dots \quad (4.3)$$

as $n \to \infty$. Set

$$C_n(\theta_0, h) = 1/E[\exp(h'T_n^{1/2}(\theta_0)W_n^*(\theta_0) - \frac{1}{2} h'T_n(\theta_0)h)].$$

From (4.3) it follows that

$$\sup_{|h| < b} |C_n(\theta_0, h) - 1| \to 0 \text{ as } n \to \infty \quad \dots \quad (4.4)$$

for every $b > 0$. To complete the proof of the statements (i) and (ii) it remains to show that

$$\|P_{\theta_0 + \delta_n^{-1}h, n} - Q_n(\theta_0, h)\| \to 0 \text{ as } n \to \infty$$

for every $h \in \mathcal{R}^k$. Without loss of generality we can assume that $P_{\theta_0+\delta_n h, n} \approx P_{\theta_0, n}$ for every $n \geqslant 1$, $h \in \mathcal{R}^k$. In view of (4.4) and since $|W_n^*(\theta_0) - W_n(\theta_0)| \to 0$ in $P_{\theta_0, n}$-probability, we see that the difference

$$Z_{n,\theta_0}(h) - Z'_{n,\theta_0}(h) \quad \dots \quad (4.5)$$

converges to zero in $P_{\theta_0,n}$-probability for every $h \in R^k$, where we set

$$Z_{n,\theta_0}(h) = \frac{dP_{\theta_0+\delta_n^{-1}h,n}}{dP_{\theta_0,n}}$$

and

$$Z'_{n,\theta_0}(h) = \frac{dQ_n(\theta_0, h)}{dP_{\theta_0,n}}.$$

Further, in view of contiguity,

$$\lim_{a\to\infty} \limsup_{n\to\infty} \int_{(|Z_{n,\theta_0}(h)|>a)} Z_{n,\theta_0}(h) dP_{\theta_0}$$

$$= \lim_{a\to\infty} \limsup_{n\to\infty} P_{\theta_0+\delta_n h,n}[|Z_{n,\theta_0}(h)| > a] = 0. \quad \ldots \quad (4.6)$$

Similarly we see that, since the sequence $\{Q_n(\theta_0, h)\}$ and $\{P_{\theta,n}\}$ are contiguous,

$$\lim_{a\to\infty} \limsup_{n\to\infty} \int_{\{|Z'_{n,\theta_0}(h)|>a\}} Z'_{n,\theta_0}(h) dP_{\theta_0,n}$$

$$= \lim_{a\to\infty} \limsup_{n\to\infty} Q_n(\theta_0, h)[|Z'_{n,\theta_0}(h)| > a) = 0. \quad \ldots \quad (4.7)$$

Combining (4.5), (4.6) and (4.7) we see that

$$\int |Z_{n,\theta_0}(h) - Z'_{n,\theta_0}(h)| dP_{\theta_0,n} \to 0$$

as $n \to \infty$ for every $h \in \mathcal{R}^k$. This completes the proof of the statements (i) and (ii).

5. Asymptotic behaviour of estimators

In this section, the sequence $\{W_n^*(\theta_0)\}$ of random vectors and the sequence of probability measures $\{Q_n(\theta_0, h)\}$, $h \in \mathcal{R}^k$, constructed in Proposition 4 of Section 4 will be used without any further mentioning.

In view of the Proposition 4, the sequence of σ-fields generated by the sequence $\{T_n(\theta_0), \delta_n(V_n - \theta_0)\}$, $n \geqslant 1$, where $\{V_n\}$ is any sequence of estimators satisfying the ACS-condition at $\theta = \theta_0$, is locally asymptotically sufficient* at $\theta = \theta_0$ in the sense of LeCam (1960, p. 49). Furthermore, the bounds of Propositions 2 and 3 of Section 3 are attained for any sequence of estimators satisfying the ACS-condition at $\theta = \theta_0$. The purpose of this section is to

* In fact, it can be shown that any sequence of ACS estimators together with a sequence of estimates of the matrices of the LAMN condition satisfies the "global" asymtotic sufficiency criteria of LeCam (1960). See Davies (1979) and Jeganathan (1980b) for the details.

show that maximum probability estimators, maximum likelihood estimators and a certain class of Bayes estimators satisfy the ACS-condition. Throughout the following subsections (a) and (b) we assume the set up of Section 2 and that the likelihood function $L_n(X_1, \ldots, X_n : \theta)$ is jointly measurable in $(X_1, X_2, \ldots, X_n, \theta)$.

(a) *Maximum probability estimators* : A maximum probability estimator $\bar{\theta}_n(a)$ with respect to the set $D_a = \{h \in \mathcal{R}^k : |h| \leqslant a\}$, $a > 0$, is defined as that value of d for which the integral

$$\int L_n(X_1, \ldots, X_n; \theta)d\theta$$

over the set $\{d - \delta_n^{-1} D_a\}$, is maximum.

We assume that a measurable maximum probability estimator exists. Detailed discussion of maximum probability estimators can be found in Weiss and Wolfowitz (1974).

Theorem 4 : *Suppose that* (i) *the sequence* $\{P_{\theta, n}; \theta \in \Theta\}$ $n \geqslant 1$, *satisfies the LAMN-condition at* $\theta = \theta_0$ *and* (ii) *the sequence* $\delta_n(\bar{\theta}_n(a) - \theta_0)$, $n \geqslant 1$, *is relatively compact. Then the sequence* $\delta_n(\theta_n(a) - \theta_0)$, $n \geqslant 1$, *satisfies the ACS-condition at* $\theta = \theta_0$.

Proof : First note that

$$|W_n(\theta_0) - W_n^*(\theta_0)| \xrightarrow{P} 0.$$

Hence it is enough to show that, for every $\delta > 0$,

$$P[|\delta_n(\bar{\theta}_n - \theta_0) - T_n^{-1/2}(\theta_0)W_n^*(\theta_0)| > \delta] \to 0. \qquad \ldots \quad (5.1)$$

Select α sufficiently large such that, for a given $\varepsilon > 0$,

$$\overline{\lim_n}\, P[|\delta_n(\bar{\theta}_n - \theta_0)| > \alpha - a] < \varepsilon/2$$

and

$$\overline{\lim_n}\, P[|T_n^{-1/2}(\theta_0)W_n^*(\theta_0)| > \alpha - a] < \varepsilon/2.$$

Hence (5.1) will follow if we show that for every given $\varepsilon > 0$ and $\delta > 0$, there exists an n_0 such that

$$P(A_n) < \varepsilon/2 \text{ for all } n \geqslant n_0, \qquad \ldots \quad (5.2)$$

where we set

$$A_n = \{|\delta_n(\bar\theta_n-\theta_0)-T_n^{-1/2}(\theta_0)W_n^*(\theta_0)| > \delta,$$
$$|\delta_n(\bar\theta_n-\theta_0)| \leqslant \alpha-a,$$
$$|T_n^{-1/2}(\theta_0)W_n^*(\theta_0)| \leqslant \alpha-a\}.$$

Now, since

$$\int\left|\exp\Lambda_n(\theta_0+\delta_n^{-1}h)-\frac{dQ_n(\theta_0)}{dP_{\theta_0,n}}\right|dP_{\theta_0,n} \leqslant 2,$$

Proposition 4 implies using dominated convergence theorem, that

$$\int_{D_\alpha}\int\left|\exp\Lambda_n(\theta_0+\delta_n^{-1}h)-\frac{dQ_n(\theta_0,h)}{dP_{\theta_0,n}}\right|dP_{\theta_0,n}\,dh \to 0 \qquad \dots (5.3)$$

for every $\alpha > 0$, where we set $D_\alpha = \{h \in \mathcal{R}^k : |h| \leqslant \alpha\}$. Since

$$\sup_{|h|\leqslant\alpha} |C_n(\theta_0,h)-1| \to 0 \text{ for every } \alpha > 0,$$

(5.3) implies that

$$E\Big[\int_{D_\alpha} |\exp\Lambda_n(\theta_0+\delta_n^{-1}h)-S_n(h)|\,dh\Big] \to 0, \qquad \dots (5.4)$$

where we set

$$S_n(h) = \exp\Big(h'T_n^{1/2}(\theta_0)W_n^*(\theta_0)-\frac{1}{2}\,h'T_n(\theta_0)h\Big).$$

Now $|\delta_n(\bar\theta_n-\theta_0)| \leqslant \alpha-a$ implies that, setting $H_1 = \{\delta_n(\bar\theta_n-\theta_0)-D_a\}$, $H_1 \subseteq D_\alpha$. Hence (5.4) implies that

$$\int_{A_n}\int_{H_1} \exp|\Lambda_n(\theta_0+\delta_n^{-0}h)-S_n(h)|\,dh\,dP_{\theta_0,n} \to 0. \qquad \dots (5.5)$$

Similarly, setting $H_2 = \{T_n^{-1/2}(\theta_0)W_n^*(\theta^0)-D_a\}$,

$$\int_{A_n}\int_{H_2} |\exp\Lambda_n(\theta_0+\delta_n h)-S_n(h)|\,dh\,dP_{\theta_0,n} \to 0. \qquad \dots (5.6)$$

Now suppose that (5.2) is not true. Then for every n_0 there exists a $\delta > 0$ such that

$$P(A_n) > \delta \text{ for some } n \geqslant n_0.$$

It can be easily checked when the event A_n is true, that

$$\eta+\int_{H_1} S_n(h)dh < \int_{H_2} S_n(h)dh, \text{ for some r.v. } \eta > 0.$$

Since $P(A_n) > \delta > 0$, this implies that

$$\eta' + \int_{A_n} \int_{H_1} S_n(h) dh \, dP_{\theta_0, n}$$

$$< \int_{A_n} \int_{H_2} S_n(h) dh \, dP_{\theta_0, n}, \text{ for some } \eta' > 0.$$

In view of (5.5) and (5.6), this implies for all sufficiently large n_0, that

$$\eta' + \int_{A_n} \int_{H_1} [\exp \Lambda_n(\theta_0 + \delta_n^{-1} h) dh \, dP_{\theta_0, n}$$

$$< \int_{A_n} \int_{H_2} \exp \Lambda_n(\theta_0 + \delta_n^{-1} h) dh \, dP_{\theta_0, n}$$

for some $n > n_0$ and $\eta' > 0$. On the other hand, the definition of maximum probability estimator gives us

$$\int_{A_n} \int_{H_1} \exp \Lambda_n(\theta_0 + \delta_n^{-1} h) dh \, dP_{\theta_0, n}$$

$$\geqslant \int_{A_n} \int_{H_2} \exp \Lambda_n(\theta_0 + \delta_n^{-1} h) dh \, dP_{\theta_0, n}$$

for every n. Thus we have arrived at a contradiction. This completes the proof.

(b) *Bayes estimators.* Throughout this section we assume that we are given a prior density $\pi(\theta)$ such that $\pi(\theta)$ is continuous, $\sup_{\theta \in \Theta} \pi(\theta) < \infty$ and $\pi(\theta) > 0$ for all $\theta \in \Theta$. We define a regular Bayes estimator $t_n = t_n(X_1, \ldots, X_n)$ as an estimator which minimises

$$B_n(\phi) = \int l_n(\theta, \phi) f_n(\theta \mid X_1, \ldots, X_n) \, d\theta \quad \ldots \quad (5.7)$$

for all sequence $(X_1, X_2, \ldots)$ where $l_n : \Theta \times \Theta \to R$, $n \geqslant 1$, are loss functions and

$$f_n(\theta \mid X_1, \ldots, X_n) = \frac{\pi(\theta) L_n(X_1, \ldots, X_n; \theta) \, d\theta}{\int \pi(\theta) L_n(X_1, \ldots, X_n; \theta) \, d\theta}$$

is the posterior density. We also assume a regular, measurable Bayes estimator exists.

In this section we consider Bayes estimators with respect to the loss function $l_n(\theta, \phi) = |\delta_n(\theta - \phi)|^a$, $a \geqslant 1$. Important works dealing with the asymptotic behaviour of Bayes estimators are, among others, LeCam (1953, 1958), Bickel and Yahav (1969), Ibragimov and Khasminskii (1972 and 1973), Borwankar, Kallianpur and Prakasa Rao (1971), Levit (1974) and Prakasa Rao (1974); it may be mentioned here that the results of these papers and the present paper are not entirely in the Bayesian spirit since the results are obtained at the true value of the parameter.

Following are the theorems of this sub-section.

Theorem 5 : *Suppose that* (i) *the sequence* $\{P_{\theta,n};\ \theta \in \Theta\}$, $n \geqslant 1$, *satisfies the LAMN-condition at* $\theta = \theta_0$ *and* (ii) *for every* $\varepsilon > 0$ *and for some* $a \geqslant 0$

$$\lim_{a\to\infty} \limsup_{n\to\infty} P\Big[\int_{|h|>a} |h|^a \exp \Lambda_n(\theta_0+\delta_n^{-1}h)dh > \varepsilon\Big] = 0. \quad \dots \quad (5.8)$$

Then for any sequence V_n, $n \geqslant 1$, *of estimators satisfying the ACS-condition at* $\theta = \theta_0$ *we have, for every* $0 \leqslant a' \leqslant a$.

$$\int |h|^{a'} \Big| f_n^*(V_n+\delta_n^{-1}h) - J \exp\Big(-\frac{1}{2}\, hT(\theta_0)h\Big)\Big| dh \xrightarrow{P} 0,$$

where we set

$$f_n^*(V_n+\delta_n^{-1}h) = \frac{\pi(V_n+\delta_n^{-1}h)\ exp\ \Lambda_n(V_n+\delta_n^{-1}h)}{\int \pi(V_n+\delta_n^{-1}h)\ exp\ \Lambda_n(V_n+\delta_n^{-1}h)dh}$$

and

$$J = \frac{|det\ T(\theta_0)|^{1/2}}{(2\pi)^{k/2}}.$$

Theorem 6 : *Suppose that* (i) *the sequence* $\{P_{\theta,n};\theta \in \Theta\}$, $n \geqslant 1$, *satisfies the LAMN-condition at* $\theta = \theta_0$ *and* (ii) *for every* $\epsilon > 0$ *and some* $\varepsilon \geqslant 1$

$$\lim_{a\to\infty} \limsup_{n\to\infty} P\Big[\int_{|h|>a} |h|^a \exp \Lambda_n(\theta_0+\delta_n^{-1}h)\, dh > \varepsilon\Big] = 0.$$

Then the sequence t_n, $n \geqslant 1$, *of Bayes estimators with respect to the loss functions* $|\delta_n(\theta-\phi)|^a$, $n \geqslant 1$, *satisfies the ACS-condition at* $\theta = \theta_0$, *and*

$$B_n(t_n) \xrightarrow{P} J \int |h|^a \exp\Big(-\frac{1}{2}\, h'T(\theta_0)h\Big) dh$$

where $B_n(t_n)$ *is the posterior risk as defined in* (5.7).

Remark : Theorem 5 is known as the Bernstein-von Mises theorem.

Since the proofs of these theorems are long, we split the proofs into several lemmas. To simplify the notations we set

$$\eta_n(\theta_0) = T^{1/2}(\theta_0)\delta_n(V_n-\theta_0)$$

and

$$R_n(h) = \exp\Big(h'T^{1/2}(\theta_0)\eta_n(\theta_0) - \frac{1}{2}\, h'T(\theta_0)h\Big),$$

where $\{V_n\}$, $n \geqslant 1$, is the sequence of estimators as considered in Theorem 5.

Lemma 10 : *For every* $\varepsilon > 0$ *and* $a \geqslant 0$,

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

Proof : Consider

$$\int |h|^{a+1} R_n(h) dh = Q(a, \eta_n(\theta_0), T(\theta_0))$$

where $Q(a, \ldots)$ is a continuous function. Since $\eta_n(\theta_0)$ and $T(\theta_0)$ are bounded in probability we see that, for any given $\varepsilon > 0$, there exists a constant $A > 0$, such that

$$P[Q(a, \eta_n, T(\theta_0)) \geqslant A] \leqslant 1-\varepsilon$$

for every n. Let $\alpha_0 = A/\varepsilon$. Since

$$\int_{|h|>\alpha} |h|^a R_n(h) dh \leqslant \alpha^{-1} \int |h|^{a+1} R_n(h)\, dh,$$

we then have

$$P\left[\int_{|h|>\alpha} |h|^a R_n(h) dh \leqslant A/\alpha \leqslant \varepsilon\right] \geqslant 1-\varepsilon$$

for every $\alpha \geqslant \alpha_0$ and all $n \geqslant 1$. This proves the result

Lemma 11 : *Suppose the assumptions of the Theorem 5 are satisfied. Then, for every* $0 \leqslant a' \leqslant a$,

$$\int |h|^{a'} |\pi(\theta_0+\delta_n^{-1}h) \exp \Lambda_n(\theta_0+\delta_n^{-1}h) - \pi(\theta_0) R_n(h)|\, dh \xrightarrow{P} 0.$$

Proof : (5.4) implies that, for every $\alpha > 0$,

$$\int_{|h|\leqslant\alpha} |\exp \Lambda_n(\theta_0+\delta_n^{-1}h) - S_n(h)| \xrightarrow{P} 0. \quad \ldots \quad (5.9)$$

Since the sequence $\{V_n\}$, $n \geqslant 1$, satisfies the ACS-condition at $\theta = \theta_0$,

$$|T^{1/2}(\theta_0)\eta_n(\theta_0) - T_n^{1/2}(\theta_0) W_n^*(\theta_0)| \xrightarrow{P} 0.$$

Hence it is not difficult to see that, for every $\alpha > 0$,

$$\int_{|h|\leqslant\alpha} |S_n(h) - R_n(h)|\, dh \xrightarrow{P} 0. \quad \ldots \quad (5.10)$$

Combining (5.9) and (5.10) we get

$$\int_{|h|\leqslant\alpha} |\exp \Lambda_n(\theta_0+\delta_n^{-1}h) - R_n(h)|\, dh \xrightarrow{P} 0$$

for every $a > 0$. Hence it follows that, since $\pi(\theta)$ is continuous at $\theta = \theta_0$

$$\int_{|h|\leqslant a} |h|^{a'} \left| \pi(\theta_0+\delta_n^{-1}h) \exp \Lambda_n(\theta_0+\delta_n^{-1}h) - \pi(\theta_0) R_n(h) \right| \xrightarrow{P} 0 \qquad \ldots \quad (5.11)$$

for every $a' \geqslant 0$ and $a > 0$. Now (5.8) implies that, since $\sup_{\theta \in \Theta} \pi(\theta) < \infty$, for every $\epsilon > 0$ and $0 \leqslant a' < a$,

$$\lim_{a} \limsup_{n\to\infty} P\left[\int_{|h|>a} |h|^{a'} \pi(\theta_0+\delta_n^{-1}h) \exp \Lambda_n(\theta_0+\delta_n^{-1}h) > \epsilon \right] = 0. \qquad \ldots \quad (5.12)$$

Similarly Lemma 10 implies that for every $\epsilon > 0$ and $a' \geqslant 0$,

$$\lim_{a\to\infty} \limsup_{n\to\infty} P\left[\int_{|h|>a} |h|^{a'} \pi(\theta_0) R_n(h) > \epsilon \right] = 0. \qquad \ldots \quad (5.13)$$

The results follows from (5.11), (5.12) and (5.13).

Proof of Theorem 5 : Let

$$(\theta_0+\delta_n^{-1}h) - V_n = \delta_n^{-1} g.$$

Then

$$|g|^{a'} \leqslant c_{a'} |h|^{a'} + c_{a'} |\delta_n(V_n - \theta_0)|^{a'}$$

where $c_{a'} = 1$ or $2^{a'-1}$ according as $a' \leqslant 1$ or $a' \geqslant 1$. Using this inequality we have

$$\int |g|^{a'} \left| f_n^*(V_n+\delta_n^{-1}g) - J \exp\left(-\frac{1}{2}\, g' T(\theta_0) g\right) \right| dg$$

$$\leqslant c_{a'} \int |h|^{a'} \Big| f_n^*(\theta_0+\delta_n^{-1}h) - J$$

$$\times \exp\left(-\frac{1}{2} (\eta_n(\theta_0) - T^{1/2}(\theta_0)h)'(\eta_n(\theta_0) - T^{1/2}(\theta_0)h)\right) \Big| \, dh$$

$$+ c_{a'} |\delta_n(V_n - \theta_0)|^{a'} \int \Big| f_n^*(\theta_0+\delta_n^{-1}h) - J$$

$$\times \exp\left(-\frac{1}{2} (\eta_n(\theta_0) - T^{1/2}(\theta_0)h)'(\eta_n(\theta_0) - T^{1/2}(\theta_0)h)\right) \Big| \, dh,$$

where

$$f_n^*(\theta_0+\delta_n^{-1}h) = \frac{\pi(\theta_0+\delta_n^{-1}h) \exp \Lambda_n(\theta_0+\delta_n^{-1}h)}{\int \pi(\theta_0+\delta_n^{-1}h) \exp \Lambda_n(\theta_0+\delta_n^{-1}h) dh}.$$

Hence, since $|\delta_n(V_n-\theta_0)|^{a'}$ is bounded in probability, it is enough to show that

$$\int |h|^{a'} \Big| f_n^*(\theta_0+\delta_n^{-1}h)-J$$

$$\times \exp\Big(-\frac{1}{2}(\eta_n(\theta_0)-T^{1/2}(\theta_0)h)'(\eta_n(\theta_0)-T^{1/2}(\theta_0)h)\Big)\Big| \, dh \xrightarrow{P} 0$$

for every $0 \leqslant a' \leqslant a$. Let

$$Y_n = \int \pi(\theta_0+\delta_n^{-1}h) \exp \Lambda_n(\theta_0+\delta_n^{-1}h) \, dh.$$

Now consider the following inequality

$$\int |h|^{a'} | f_n^*(\theta_0+\delta_n^{-1} h)-J$$

$$\exp\Big(-\frac{1}{2}(\eta_n(\theta_0)-T^{1/2}(\theta_0)h)'(\eta_n(\theta_0-T^{1/2}(\theta_0)h)\Big)\Big| \, dh$$

$$\leqslant Y_n^{-1}\int |h|^{a'} | \pi(\theta_0+\delta_n^{-1}h) \exp \Lambda_n(\theta_0+\delta_n^{-1}h)-\pi(\theta_0)R_n(h)| \, dh$$

$$+\Big| Y_n^{-1}\pi(\theta_0)-J \exp\Big(-\frac{1}{2}\eta_n'(\theta_0)\eta_n(\theta_0)\Big)\Big| \int |h|^{a'} R_n(h) \, dh$$

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

Setting $a'=0$ in Lemma 11, we have

$$\Big| Y_n-\pi(\theta_0)J^{-1} \exp\Big(\frac{1}{2}\eta_n'(\theta_0)\eta_n(\theta_0)\Big)\Big| \xrightarrow{P} 0.$$

Hence $I_2 \xrightarrow{P} 0$, since the proof of the Lemma 10 shows that $\int |h|^{a'}(R_n h) \, dh$ is bounded in probability. Similarly, by Lemma 11 and the fact that Y_n^{-1} is bounded in probability, $I_1 \xrightarrow{P} 0$. This proves the theorem.

Proof of Theorem 6 : First note that, since $T(\theta_0) > 0$ with probability one, for a given $\epsilon > 0$, there exists a $\delta > 0$ such that $P[\lambda \geqslant \delta] \geqslant 1-\epsilon$, where λ is the smallest eigen value of $T(\theta_0)$. Hence we shall assume without loss of generality that $\lambda \geqslant \delta > 0$ always. Now define

$$g(x) = |x|^a$$

and, for a given $\alpha > 0$,

$$g_\alpha(x) = \begin{cases} |x|^a & \text{if } |x|^a \leqslant \alpha \\ \alpha & \text{if } |x|^a > \alpha. \end{cases}$$

Then select x_0 so large such that, for a given $\varepsilon > 0$,

$$J\int g_{x_0}(h)\exp\left(-\frac{1}{2}h'T(\theta_0)h\right)dh$$

$$> J\int g(h)\exp\left(-\frac{1}{2}h'T(\theta_0)h\right)dh-\varepsilon/2.$$

According to Theorem 5, we have, setting

$$V_n^* = \delta_n^{-1}T^{-1/2}(\theta_0)W_n(\theta_0)+\theta_0$$

$$\int g(h)\left|f_n^*(V_n^*+\delta_n^{-1}h)-J\exp\left(-\frac{1}{2}h'T(\theta_0)h\right)\right|dh \xrightarrow{P} 0$$

and since $g_{x_0}(x)$ is bounded by x_0,

$$\int g_{x_0}(h+u_n)\left|f(V_n^*+\delta_n^{-1}h)-J\exp\left(-\frac{1}{2}h'T(\theta_0)h\right)\right|dh \xrightarrow{P} 0,$$

where we set $u_n = \delta_n(t_n-V_n^*)$. Hence, for any given $\varepsilon > 0$, there exists an n_0 such that

$$P(A_{n\varepsilon}^{(1)}\cap A_{n\varepsilon}^{(2)}) \geqslant 1-\varepsilon \text{ for all } n \geqslant n_0, \qquad \dots \quad (5.14)$$

where

$$A_{n\varepsilon}^{(1)} = \left\{\int g(h)f_n^*(V_n^*+\delta_n^{-1}h)\,dh \leqslant J\int g(h)\exp\left(-\frac{1}{2}h'T(\theta_0)h\right)dh+\varepsilon\right\}$$

and

$$A_{n\varepsilon}^{(2)} = \left\{\int g_{x_0}(h+u_n)f_n^*(V_n^*+\delta_n^{-1}h)\,dh\right.$$

$$\left.> J\int g_{x_0}(h+u_n)\exp\left(-\frac{1}{2}h'T(\theta_0)h\right)dh-\varepsilon/2\right\}.$$

First we shall prove that the sequence $\{t_n\}$ $n \geqslant 1$, satisfies the ACS-condition at $\theta = \theta_0$, i.e., we want to prove that, for every $\delta > 0$, $P[|u_n| > \delta] \to 0$. In view of (5.14) it is enough to prove that the event $A_{n\varepsilon}^{(1)}\cap A_{n\varepsilon}^{(2)}\cap\{|u_n|>\varepsilon\}$ is impossible for all $n \geqslant n_0$. Now suppose that the event $A_{n\varepsilon}^{(1)}\cap A_{n\varepsilon}^{(2)}\cap\{|u_n| > \delta\}$ is true. Using the definition of Bayes estimators we then have on the set $A_{n\varepsilon}^{(1)}$, for every $n \geqslant n_0$,

$$\begin{aligned}\int g_{x_0}(h+u_n)f_n^*(V_n^*+\delta_n^{-1}h)dh &\leqslant \int g(h+u_n)f_n^*(V_n^*+\delta_n^{-1}h)\,dh\\ &= B_n(t_n) \leqslant B_n(V_n^*)\\ &= \int g(h)f_n^*(V_n^*+\delta_n^{-1}h)\,dh\\ &\leqslant J\int g(h)\exp\left(-\frac{1}{2}h'T(\theta_0)h\right)dh+\varepsilon. \qquad \dots \quad (5.15)\end{aligned}$$

Now note that $g_{\epsilon_0}(h)$ is a non-constant function and satisfies $g_{\epsilon_0}(0) = 0$, $g_{\epsilon_0}(|h|) = g_{\epsilon_0}(h)$ and $g_{\epsilon_0}(h_1) \leqslant g_{\epsilon_0}(h_2)$ if $|h_1| \leqslant |h_2|$. Hence we have, for some $\epsilon > 0$,

$$J \int g_{\epsilon_0}(h+u_n) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh > J \int g_{\epsilon_0}(h)$$

$$\times \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh + 2\epsilon$$

whenever $|u_n| > \delta > 0$. Thus on the set $A^{(2)}_{n\epsilon} \cap \{|u_n| > \delta\}$, we have for every $n \geqslant n_0$,

$$\int g_{\epsilon_0}(h+u_n) f_n^*(V_n^* + \delta_n^{-1}h) \geqslant J \int g_{\epsilon_0}(h+u_n) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh - \epsilon/2$$

$$> J \int g_{\epsilon_0}(h) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh - \epsilon/2 + 2\epsilon$$

$$\geqslant J \int g(h) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh + \epsilon. \qquad \ldots \quad (5.16)$$

From (5.15) and (5.16), we thus see that the event $A^{(1)}_{n\epsilon} \cap A^{(2)}_{n\epsilon} \cap \{|u_n| > \delta\}$ is impossible for every $n \geqslant n_0$. This proves that the sequence $\{t_n\}$, $n \geqslant 1$, satisfies the ACS-condition at $\theta = \theta_0$.

Now it follows from the previous arguments, that

$$J \int g_{\epsilon_0}(h+u_n) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh - \epsilon/2 \leqslant \int g_{\epsilon_0}(h+u_n) f_n^*(V_n^* + \delta_n^{-1}h)\, dh$$

$$\leqslant \int g(h+u_n) f_n^*(V_n^* + \delta_n^{-1}h)\, dh$$

$$\leqslant J \int g(h) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh + \epsilon \qquad \ldots \quad (5.17)$$

with probability tending to one. Since $u_n \xrightarrow{P} 0$, it easily follows that

$$J \int g_{\epsilon_0}(h+u_n) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh \xrightarrow{P} J \int g_{\epsilon_0}(h) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh$$

$$\geqslant J \int g(h) \exp\left(-\frac{1}{2} h'T(\theta_0)h\right) dh - \epsilon/2. \qquad \ldots \quad (5.18)$$

Combining (5.17) and (5.18), we see that

$$B_n(t_n) = \int g(u+u_n) f_n^*(V_n^* + \delta_n^{-1} h)\, dh$$

$$\xrightarrow{P} J \int g(h) \exp\left(-\frac{1}{2} h' T(\theta_0) h\right) dh.$$

This proves the theorem.

(c) *Maximum likelihood estimators.* A measurable function $\hat{\theta}_n = \hat{\theta}_n(X_1, \ldots, X_n)$ is called a maximum likelihood estimator if

$$L_n(X_1, \ldots, X_n; \hat{\theta}_n) \geqslant L_n(X_1, \ldots, X_n; \theta)$$

for all $\theta \in \Theta$, where $L_n(X_1, \ldots, X_n; \theta)$ is the likelihood function as defined in (1.1). We assume that a maximum likelihood estimator exists.

Theorem 7 : *Suppose that* (i) *the sequence* $\{P_{\theta, n}; \theta \in \Theta\}$, $n \geqslant 1$, *satisfies the LAMN-condition at* $\theta = \theta_0$, (ii) *for every* $\varepsilon > 0$, *and* $\alpha > 0$, *setting* $D_\alpha = \{h \in \mathcal{R}^k : |h| \leqslant \alpha\}$,

$$\lim_{\delta \to 0} \limsup_{n \to \infty} P\left[\sup_{|h_2 - h_1| \leqslant \delta} |\Lambda_n(\theta_0 + \delta_n^{-1} h_2) - \Lambda_n(\theta_0 + \delta_n^{-1} h_1)| > \varepsilon,\ h_1, h_2 \in D_\alpha\right\} = 0$$

and (iii) *the sequence* $\{\delta_n(\hat{\theta}_n - \theta_0)\}$, $n \geqslant 1$, *is relatively compact. Then the sequence* $\{\delta_n(\hat{\theta}_n - \theta_0)\}$, $n \geqslant 1$, *satisfies the ACS-condition at* $\theta = \theta_0$.

First we shall prove the following lemma.

Lemma 12 : *Suppose that the sequence* $\{P_{\theta, n}; \theta \in \Theta\}$, $n \geqslant 1$, *satisfies the LAMN-condition at* $\theta = \theta_0$. *Then*

$$\sup_{|h| < \alpha} |\Lambda_n(\theta_0 + \delta_n^{-1} h) - Q_n(h)| \xrightarrow{P} 0,\ \alpha > 0,$$

where we set $Q_n(h) = h' T^{1/2}(\theta_0) W_n(\theta_0) - \frac{1}{2} h' T(\theta_0) h$, *if and only if for every* $\varepsilon > 0$,

$$\lim_{\delta \to \infty} \limsup_{n \to \infty} P\left[\sup_{|h_2 - h_1| < \delta} |\Lambda_n(\theta_0 + \delta_n^{-1} h_2) - \Lambda_n(\theta_0 + \delta_n^{-1} h_1)| > \varepsilon,\ h_1, h_2 \in D_\alpha\right] = 0. \qquad \ldots (5.19)$$

Proof : Using the fact that the sequence $\{W_n(\theta_0), T(\theta_0)\}$ is bounded in probability, it is easily seen that for every $\varepsilon > 0$ and $\alpha > 0$,

$$\lim_{\delta \to \infty} \limsup P\left[\sup_{|h_2 - h_1| < \delta} |Q_n(h_2) - Q_n(h_1)| > \varepsilon,\ h_1, h_2 \in D_\alpha\right] = 0. \quad \ldots (5.20)$$

Set $Y_n(h) = \Lambda_n(\theta_0+\delta_n^{-1}h)-Q_n(h)$. From (5.19) and (5.20) it follows that, for a given $\varepsilon > 0$ and $\eta > 0$ there exist a $\delta > 0$ and n_0 such that

$$P\left[\sup_{|h_2-h_1|<\delta} |Y_n(h_2)-Y_n(h_1)| > \varepsilon,\ h_1, h_2 \in D_\alpha\right] \leqslant \eta \quad \dots \quad (5.21)$$

for all $n \geqslant n_0$. We then partition the set D_α into cubes of sides of length δ (without loss of generality we assume that $\delta^{-1}\alpha$ is an integer). Then totally there are $(\delta^{-1}\alpha)^k = m$ cubes. Now

$$\sup_{|h|\leqslant\alpha}|Y_n(h)| \leqslant \sup_{i\leqslant m} Y_n(l_i)$$
$$+\sup_{|h_2-h_1|<\delta} |Y_n(h_2)-Y_n(h_1)|$$

where l_i is a fixed point in the i-th cube, $i = 1, 2, \dots, m$. Then

$$P\left[\sup_{|h|\leqslant\alpha} |Y_n(h)| > \varepsilon\right] \leqslant \sum_{i=1}^{m} P[Y_n(l_i) > \varepsilon/2]$$

$$+P\left[\sup_{|h_2-h_1|<\delta} |Y_n(h_2)-Y_n(h_1)| > \varepsilon/2,\ h_1, h_2 \in D_\alpha\right].$$

Since the sequence $\{P_{\theta,n};\ \theta \in \Theta\}$, $n \geqslant 1$, satisfies the LAMN-condition at $\theta = \theta_0$ it follows that, for every fixed integer m,

$$\sum_{i=1}^{m} P[|Y_n(l_i)| > \varepsilon/2] \to 0 \text{ as } n \to \infty.$$

Hence, by (5.21) it follows that

$$P\left[\sup_{|h|\leqslant\alpha}|Y_n(h)| > \varepsilon\right] \to 0 \text{ as } n \to \infty.$$

The other part of the proof follows easily by noting that

$$\Lambda_n(\theta_0+\delta_n^{-1}h) = Y_n(h)+Q_n(h).$$

This completes the proof of the lemma.

Proof of Theorem 7: Let λ be the smallest eigen value of $T(\theta_0)$. Since $T(\theta_0)$ is positive definite with probability one, for a given $\varepsilon > 0$, there exist a $\gamma > 0$ such that

$$P[\lambda > \gamma] \geqslant 1-\varepsilon/2.$$

Further, for a given $\varepsilon > 0$ and $\delta > 0$ there exists an $\alpha > 0$ such that

$$\limsup_{n\to\infty} \{P[|\hat{h}_n| > \alpha]+P[|T^{-1/2}(\theta_0)W_n(\theta_0)| > \alpha-\delta]\} \leqslant \varepsilon/3,$$

where we set $\hat{h}_n = \delta_n(\hat{\theta}_n-\theta_0)$. Hence it is enough to show that, for every $\alpha > 0$, $\delta > 0$ and $\varepsilon > 0$, there exists an n_0 such that

$$P[|\hat{h}_n-T^{-1/2}(\theta_0)W_n(\theta_0)| > \delta,$$

$$|\hat{h}_n| \leqslant \alpha,\ |T^{-1/2}(\theta_0)W_n(\theta_0) \leqslant \alpha-\delta, \lambda > \gamma] \leqslant \varepsilon/3$$

for every $n \geqslant n_0$. Equivalently, we shall prove that

$$P\Big[\sup_{h\in D_a} \exp\Lambda_n(\theta_0+\delta_n^{-1}h)$$

$$\leqslant \sup_{h\in E} \exp\Lambda_n(\theta_0+\delta_n^{-1}h),\ |\hat{h}_n| \leqslant \alpha,$$

$$|T^{-1/2}(\theta_0)W_n(\theta_0)| \leqslant \alpha-a, \lambda > \gamma\Big] \leqslant \varepsilon/3 \qquad \ldots \quad (5.22)$$

for every $n \geqslant n_0$, where we set

$$E = \{h \in \mathcal{R}^k : |h-T^{-1/2}(\theta_0)W_0(\theta_0)| > \delta,\ h \in D_a\}.$$

Denote the event inside the bracket of (5.22) by B_n. Let

$$A_n = \Big\{\sup_{|h|\leqslant a} |\exp\Lambda_n(\theta_0+\delta_n^{-1}h)-\exp Q_n(h)| \leqslant \eta\Big\}, \eta > 0. \quad \ldots \quad (5.23)$$

By Lemma 12, there exists an n_0 such that

$$P[A_n^c] \leqslant \varepsilon/3, \text{ for all } n \geqslant n_0.$$

Where A_n^c denotes the complement of the set A_n, $n \geqslant 1$. Now

$$P[B_n] \leqslant P[B_n \cap A_n]+P[A_n^c].$$

We shall now show that $P[B_n \cap A_n] = 0$ for all $n \geqslant n_0$, which will prove $P[B_n] \leqslant \varepsilon/3$ for all $n \geqslant n_0$. Suppose that the event $A_n \cap B_n$ is true. Then, by (5.23),

$$\sup_{|h|\leqslant a} \exp\Lambda_n(\theta_0+\delta_n^{-1}h) \geqslant \sup_{|h|\leqslant a} \exp Q_n(h)-\eta \qquad \ldots \quad (5.24)$$

and

$$\sup_{h\in E} \exp\Lambda_n(\theta_0+\delta_n^{-1}h) \leqslant \sup_{h\in E} \exp Q_n(h)+\eta. \qquad \ldots \quad (5.25)$$

Also note that (since when the event A_n is true,

$$|T^{-1/2}(\theta_0)W_n(\theta_0)| \leqslant \alpha-\delta)$$

$$\sup_{|h|\leqslant a} \exp Q_n(h) = \exp\left(\frac{1}{2} W_n'(\theta_0)W_n(\theta_0)\right) \quad \dots \quad (5.26)$$

and

$$\sup_{h \in B} \exp Q_n(h) \leqslant \exp\left[-\frac{\gamma\delta^2}{2}+\frac{1}{2} W_n'(\theta_0)W_n(\theta_0)\right]. \quad \dots \quad (5.27)$$

Since $\gamma > 0$, there exists an η_0 such that

$$\exp \frac{1}{2} W_n'(\theta_0)W_n(\theta_0)-\eta_0 > \exp\left[-\frac{\gamma\delta^2}{2}+\frac{1}{2} W_n'(\theta_0)W_n(\theta_0)\right]+\eta_0. \quad \dots \quad (5.28)$$

We thus see from (5.24)-(5.26), that, when $\eta \leqslant \eta_0$ and the event $A_n \cap B_n$ is true

$$\sup_{|h|\leqslant a} \exp \Lambda_n(\theta_0+\delta_n^{-1}h) > \sup_{h\in B} \exp \Lambda_n(\theta_0+\delta_n^{-1}h)$$

for all $n \geqslant n_0$. But this is a contradiction since on B_n we have

$$\sup_{|h|\leqslant a} \exp \Lambda_n(\theta_0+\delta_n^{-1}h) \leqslant \sup_{h\in B} \exp \Lambda_n(\theta_0+\delta_n^{-1}h).$$

Hence $P[A_n \cap B_n] = 0$ for a $n \geqslant n_0$. This proves the result.

6. Discussions on the assumption (A.1) of section 2

The arguments of this section are based on LeCam (1970) and Hájek (1972).

Consider the following set of assumptions.

(6.A.6). The functions $f_j(X_j | X_1, \dots, X_{j-1}; \theta) = f_j(\theta) : \Theta \to R$ are absolutely continuous in θ for all $(X_1, \dots, X_j)$, $j \geqslant 1$.

(6.A.7). For every $\theta \in \Theta$ the θ derivative $\dot{f}_j(\theta) = (\partial/\partial\theta)f_j(\theta)$ exists for $\mu_1 \times \dots \times \mu_j$ almost all $(X_1, \dots, X_j)$, $j \geqslant 1$. Define for every $\theta \in \Theta$ and $j \geqslant 1$

$$\dot{\xi}_j(\theta) = \begin{cases} \dot{f}_j(\theta)f^{-1/2}(\theta) & \text{if the derivative exists and } f_j(\theta) > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Suppose that we have selected the sequence $\{\delta_n\}$; one way of selection is to define

$$\delta_n\delta_n' = \left\{\sum_{j=1}^{n} E_\theta[\int \dot{\xi}_j(\phi)\dot{\xi}_j'(\phi)d\mu_j]\right\} \text{ for some fixed } \phi \in \Theta.$$

(6.A.8). For every $h \in \mathcal{R}^k$ and $\theta \in \Theta$

$$E[\int |h'\delta_n\dot{\xi}_j(\theta)|^2 d\mu_j] < \infty, \quad 1 \leqslant j \leqslant n < \infty.$$

(6.A.9). For every $h \in \mathcal{R}^k$ and for every $\theta \in \Theta$

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

Proposition 5: *Suppose the assumptions* (6.A.6)–(6.A.9) *are satisfied. Then the assumption* (2.A.1) *is satisafied for every* $\theta \in \Theta$.

Proof: An application of the inequality (2.6) and (6.A 8) shows that (6.A.9) in particular entails, for every $h \in \mathcal{R}^k$ and $P_{\theta,j-1} \times \mu_j$ almost all $(X_1, \ldots, X_j)$,

$$\int_a^b |h'\delta_n\dot{\xi}_j(\theta+t\delta_n h)|^2 dt < \infty, \quad 1 \leqslant j \leqslant n < \infty$$

Hence according to Lemma (A.1) of Hájek (1972, p. 189), for every $\theta \in \Theta$ and $h \in \mathcal{R}^k$ the functions $t \to f_j^{1/2}(\theta+t\delta_n h)$, $1 \leqslant j \leqslant n \leqslant \infty$ are absolutely continuous in the interval (a, b) for $P_{\theta,j-1} \times \mu_j$ almost all $(X_1, \ldots, X_t)$. Hence we can write for $P_{\theta,j-1} \times \mu_j$ almost all $(X_1, \ldots, X_j)$, for all $h \in \mathcal{R}^k$ and $\theta \in \Theta$

$$f_j^{1/2}(\theta+t_2\delta_n h) - f_j^{1/2}(\theta+t_1\delta_n h)$$
$$= \frac{1}{2}\int_{t_1}^{t_2} h'\delta_n\dot{\xi}_j(\theta+t\delta_n h)\, dt$$

for every t_1 and t_2 such that $a < t_1 < t_2 < b$. Hence

$$\sum_{j=1}^{n} E\left[\int \left| f_j^{1/2}(\theta+\delta_n h) - f_j^{1/2}(\theta) - \frac{1}{2} h'\dot{\xi}_j(\theta)\right|^2 d\mu_j\right]$$

$$= \frac{1}{4}\sum_{j=1}^{n} E\left\{\int \left|\int_0^1 h'\delta_n\left[\dot{\xi}_j(\theta+t\delta_n h) - \dot{\xi}_j(\theta)\right] dt\right|^2 d\mu_j\right\}$$

$$\leqslant \frac{1}{4}\sum_{j=1}^{n} E\left\{\int_0^1 dt \int |h'\delta_n[\dot{\xi}_j(\theta+t\delta_n h) - \dot{\xi}_j(\theta)]|^2 d\mu_j\right\}$$

$$\leqslant \frac{1}{4}\sup_{0\leqslant t\leqslant 1} E\{\int |h'\delta_n[\dot{\xi}_j(\theta+t\delta_n h) - \dot{\xi}_j(\theta_0)]|^2 d\mu_j\}$$

$\to 0$ by (6.A.9).

The proof is complete.

Remark: In connection with the above result it should be mentioned here that LeCam (1974) has given some results, based on Lusin's (N)-condition (cf. Hewitt and Stromberg (1965, p. 288)) instead of absolute continuity, which are applicable to more general situations; LeCam's arguments are restricted to the i.i.d. case but the above discussion shows that his arguments are applicable to the general case also.

Acknowledgement

The first version of this work was written under the supervision of Dr. B. L. S. Prakasa Rao. The final version has benefitted from comments of Professor L. LeCam and Professor J. K. Ghosh.

References

Aldous, D. J. and Eagleson, G. K. (1978): On mixing and stability of limit theorems. *Ann. Probability*, **6**, No. 2, 325-331.

Basawa, I. V. and Scott, D. J. (1976): Efficient tests for branching processes. *Biometrica*, **63**, 531-536.

——— (1977): Efficient tests for stochastic processes. *Sankhyā*, **32**, Ser. A, 21-31.

——— (1979): Efficient estimation for stochastic processes. *J. Indian Statist. Assoc.* (To appear).

Bhat, B. R. (1978): Some problems of estimation for dependent observations. Tech. Report, Univ. of Poona, India.

Bhat, B. R. and Prasad, M. S. (1978): *W*-efficiency of maximum likelihood estimates. Tech. Report, Univ. of Poona, India.

Bickel, P. and Yahav, J. (1969): Some contributions to the asymptotic theory of Bayes solutions. *Z. Wahrscheinlich-keitstheorie und Verw Gebiete.* **11**, 257-279.

Borwanker, J. D., Kallianpur, G. and Prakasa Rao, B. L. S. (1970): The Bernstein-von Mises theorem for Markov processes. *Ann. Math. Statist.* **42**, 1241-1253.

Davies, R. B. (1978): Asymptotic inference when the amount of information is random. Tech. Report, Wellington Univ., New Zealand.

Feigin, P. D. (1978): The efficiency criteria problem for stochastic processes. *Stochastic processes and their applications*, **6**, 115-127.

Hall, P. G. (1977): Martingale invariance principles. *Ann. Probability*, **5**, 875-87.

Hajek, J. (1970): A characterization of limiting distributions of regular estimates. *Z. Wahrscheinlichkeitstheorie verw. Gebiete.*, **14**, 323-30.

——— (1972): Local asymptotic minimax and admissibility in estimation. *Proc. 6th Berkeley Symp.* Vol. **7**, 175-194.

Heyde, C. C. (1978): On an optimal asymptotic property of the maximum likelihood estimator of parameters from a stochastic process. *Stochastic processes and their applications*. **8**, 1-9.

Heyde, C. C. and Feigin, P. D. (1975): On efficiency and exponential families in stochastic process estimation. *Statistical Distributions in Scientific Work*. Vol. 1, Patil, G. P., Kotz, S. and Ord, J. K. editors, Reidel, Dordrech, 227-230.

IBRAGIMOV, I. A. and KHASMINSKII, R. Z. (1972 and 1973): Asymptotic behaviour of some statistical estimators in the smooth case. I and II. *Theory Prob. Applications*. 17, 445-62; 18, 76-91.

——— (1975): Local asymptotic normality for non-identically distributed observations. *Theory of Probability Applications*, 20.

INAGAKI, N. (1970): On the limiting distribution of a sequence of estimators with uniformity property. *Ann. Inst. Statist. Math.* 22, 1-13.

JEGANATHAN, P. (1981): On a decomposition of the limit distribution of a sequence of estimators. *Sankhyā*, 43, A, 26-36

——— (1980): An extension of a result of L. LeCam concerning asymptotic normality. *Sankhyā*, 42, A, 146-160

——— (1980a): Some asymptotic properties of risk functions when the limit of the experiments is mixed normal. *Sankhyā* (To appear).

——— (1980b): Asymptotic theory of estimation when the limit of the log-likelihood ratios is mixed normal. Ph.D. thesis submitted to Indian Statistical Institute, Calcutta.

KEIDING, N. (1975): Maximum likelihood estimation in the birth and death process. *Ann. Statist.* 3, 360-372.

LECAM, L. (1953): On some asymptotic properties of maximum likelihood and related Bayes estimators. *University of Calif. Publ. in Statist.*, 1, 277-330.

——— (1958): Les Proprietes asymptotiques des solutions de Bayes. *Publ. Inst. Statist. Univ. Paris*, 8, 17-35.

——— (1960): Locally asymptotically normal families of distributions. *Univ. of Calif. Publ. in Stat.* 3, 37-98.

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

——— (1972): Limits of experiments. *Proc. 6th Berkeley Symp.* Vol. 1, 245-61.

——— (1974): On some results of J. Hajek concerning asymptotic normality (pre-print).

LEVIT, V. YA. (1974): On the behaviour of generalised Bayes estimates of Markov observations. *Theory Prob. Applications*, 19, 332-346.

MCLEISH, D. L. (1974): Dependent central limit theorems and invariance principles. *Ann. Probability*, 2, 620-28.

PHILLIPPOU, A. N. and ROUSSAS, G. G. (1973): Asymptotic distribution of the likelihood function in the independent not identically distributed case. *Ann. Statist.* 1, 545-71.

PRAKASA RAO, B. L. S. (1974): Statistical inference for stochastic processes. Technical Report CRM 465, Centre de Recherches Mathematiques, Universite de Montreal.

ROUSSAS, G. G. (1972): *Contiguity of Probability Measures*, Cambridge University Press.

——— (1979): Asymptotic distribution of the log-likelihood function for stochastic processes. *Z. Wahrsheinlichkeitstheorie verw. Gebiete*, 47, 31-46.

SWENSEN, A. R. (1980): Asymptotic inference for a class of stochastic processes. Ph.D. thesis, University of Calif., Berkeley.

WEISS L. and WOLFOWITZ, J. (1974): *Maximum Probability Estimators and Related Topics* (Lecture notes in mathematics) Springer—Verlag.

Paper received: February, 1979.
Revised: December, 1980.