Sankhyā : The Indian Journal of Statistics
1980, Volume 42, Series A, Pts. 3 and 4, pp. 146–160.

AN EXTENSION OF A RESULT OF L. LECAM CONCERNING ASYMPTOTIC NORMALITY

By P. JEGANATHAN
Indian Statistical Institute

SUMMARY. L. LeCam (1960) has shown that a certain kind of asymptotic differentiability of the log-likelihood ratios implies that the limit distribution of the log-likelihood tunction is normal. This result is extended and strengthened to a situation where it is shown that the limit distributiqn is, in general, mixed normal,

1. Introduction

LeCam (1960) has shown that a certain kind of asymptotic differentiability of the log-likelihood ratios together with the contiguity condition implies that the limit distribution of the suitably normalised log-likehood function is normal. A remarkable thing to be noted here is that the asymptotic normality occur through an argument which has nothing to do with sums of independent random variables or martingale differences. It is the purpose of this paper to extend and strengthen this and other related results of LeCam to a situation where the limit distribution turns out to be a mixed normal.

More specifically, in his definition of asymptotic differentiability, LeCam assumed that the sequence of normalised log-likelihood ratios is 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 function of the normalised parameter. In this paper we assume that this second expression is also a sequence of random function of the normalised parameter and then we first establish Theorem 1 that the limit distribution, when it exists, is a mixed normal for almost all points of the parameter space. Secondly we establish Theorem 2 that, without assuming the existence of the limit distribution, the log-likelihood function converges in a certain weak topology to a mixed normal distribution; though the convergence stated here is very much weaker than the convergence stated in Theorem 1, Theorem 1 actually follows from this result and it appears that this result is more important than Theorem 1. In the special case when the second expression mentioned above is assumed to be a non-random function of the normalised parameter, it is possible to obtain the convergence stated in Theorem 1 under the assumptions of Theorem 2 (see the remark following this Theorem 2); thus the conclusion of Theorem 4.1 of LeCam (1960) holds even when the existence of the limit distribution is not assumed. Thirdly, we establish Theorem 3 that, when the second expression mentioned above is a sequence of random quadratic forms of the normalised parameter and when the sequence of random matrices of this quadratic forms satisfies a certain invariance condition, the limit distribution is a mixed normal for all points of the parameter space. It may be mentioned here that the contiguity condition plays a crucial role in establishing all these results.

Our approximation of the log-likelihood ratios, stated in section 2, is slightly weaker them the one assumed in LeCam (1960), and therefore we will have to further assume that the random quantities involved in the approximation are jointly measurable in the observations and the parameter, and that the given sequence of family of probability measures are measurable in a certain sense. In section 4 it is shown that these measurability restrictions can be removed when the approximation of the log-likelihood ratios is analogous to the one assumed in LeCam (1960).

Assumptions and the main results are stated in section 2 and the proofs of the main results are presented in section 3.

After the completion of this work we came to know of a related work Davies (1979) which contains a version of the third result Theorem 3 of the present paper.

When the asymptotic differentiability condition, together with the condition that the limit distribution of the log-likelihood function is mixed normal, is satisfied, the given sequence of families of probability measures may be called locally asymptotically mixed normal (LAMN) families. A detailed study of the LAMN condition under the 'differentiability in quadratic mean'—type regularity conditions, using the recent martingale central limit theorems, can be found in Jeganathan (1979a). Further interesting consequences of the LAMN-condition and extensions of several basic results of LeCam and Hajek can be found in Jeganathan (1979a, 1979b and 1980). For the practical situations where the LAMN condition is satisfied, see e.g. Bhat (1978), Basawa and Prakasa Rao (1979) and Davies (1979).

2. Assumptions and the main results

Let $\mathcal{E}_n = \{\mathcal{X}_n, \mathcal{A}_n, P_{\theta,n}; \theta \in \Theta\}$, $n \geqslant 1$, be a sequence of experiments; through out this paper it will be assumed, without further mentioning, that Θ is an open subset of R^k, $k \geqslant 1$.

We use the following notations. 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 $\mathscr{L}(Y|P)$ when $Y:(\mathscr{X}, \mathscr{A}) \to (R^q, \mathscr{B}^q)$, $q \geqslant 1$, $\mathscr{B}^q$ being the σ-field of Borel subsets of R^q. $\mu^k | \mathscr{B}^k$ denotes the Lebesgue measure. For a vector $h \in R^k$, h' denotes the transpose of h and $|h|$ denotes the euclidean norm. For a matrix D, $|D|$ denotes the norm defined by the square root of the sum of squares of its elements.

Definition: A sequence of experiments $\mathscr{E}_n = \{\mathscr{X}_n, \mathscr{A}_n, P_{\theta,n}; \theta \in \Theta\}$, $n \geqslant 1$, will be called asymptotically differentiable on Θ if the following six assumptions are satisfied.

(A.1). The functions $\theta \to P_{\theta,n}(A)$, $A \in \mathscr{A}_n$, $n \geqslant 1$, are Borel measurable.

(A.2). There exist $\mathscr{A}_n \times \mathscr{B}^k$-measurable functions

$$W_n(\cdot): \mathscr{X}_n \times \Theta \to R^k \text{ and } A_n(h, .): \mathscr{X}_n \times \Theta \to R,\ n \geqslant 1,\ h \in R^k,$$

such that the difference

$$\frac{dP_{\theta+\delta_n h, n}}{dP_{n,}} - \exp[h' W_n(\theta) - A_n(h, \theta)]$$

converges to zero in $P_{\theta,n}$-probability for every $h \in R^k$ and $\theta \in \Theta$, where $\{\delta_n\}$ is a sequence of positive definite (p.d.) matrices such that $\|\delta_n\| \to 0$ as $n \to \infty$.

(A.3). The sequences $\{P_{\theta+\delta_n h, n}\}$ and $\{P_{\theta,n}\}$ are contiguous for every $h \in R^k$ and $\theta \in \Theta$.

(A.4). For very $\theta \in \Theta$, there exist a random function $h \to A(h, \theta)$ and a random vector $W(\theta)$ defined on some probability space $(\mathscr{Y}, \mathscr{F}, \lambda_\theta)$ such that for every finite $\{h_i; i = 1, 2, \ldots, m\}$

$$\mathscr{L}(W_n(\theta), A_n(h_i, \theta), i = 1, 2, \ldots, m | P_{\theta,n})$$
$$\Longrightarrow \mathscr{L}(W(\theta), A(h_i, \theta), i = 1, 2, \ldots, m | \lambda_\theta).$$

(A.5). For every $\theta \in \Theta$, there exists a set $N_\theta \in \mathscr{F}$ of λ_θ-measure zero such that the functions $h \to A(h, \theta)$ are continuous for all points outside the set N_θ.

(A.6). For every $s, h \in R^k$ and $\theta \in \Theta$, the difference

$$A_n(h, \theta + \delta_n s) - A_n(h, \theta)$$

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

Theorem 1 : *Suppose that the sequence* $\{\mathcal{E}_n\}$ *of experiments satisfies the conditions* (A.1)–(A.6). *Then there are* $\mathcal{A}_n \times \mathcal{B}^k$*-measurable k-vectors* $\gamma_n(\theta)$ *and posive semi-definite (p.s.d.)* $k \times k$ *matrices* $T_n(\theta)$, $n \geqslant 1$, $\mathcal{F} \times \mathcal{B}^k$*-measurable k-vector* $\gamma(\theta)$ *and a p.s.d.* $k \times k$ *matrix* $T(\theta)$, *and a Lebesgue null set* $N \subsetneq \Theta$ *such that for every* $\theta \in \Theta - N$

(i) *The difference*

$$A_n(h, \theta) - [h'\gamma_n(\theta) + \tfrac{1}{2} h'T_n(\theta)h]$$

converges to zero in $P_{\theta,n}$*-probability, and*

$$\text{(ii)} \qquad \mathcal{L}(W_n(\theta),\ \gamma_n(\theta)\ T_n(\theta) \mid P_{\theta,n})$$

$$\Longrightarrow \mathcal{L}(T^{1/2}(\theta)Z + \gamma(\theta),\ \gamma(\theta),\ T(\theta) \mid \lambda_\theta)$$

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

Corollary : *Suppose that the sequence* $\{\mathcal{E}_n\}$ *satisfies the conditions* (A.1)–(A.6). *Further assume that* $\mathcal{L}(W(\theta), A(h_i, \theta), i = 1, 2, \ldots, m \mid \lambda_\theta)$ *is a continuous function of* θ *for every finite* $\{h_i;\ i = 1, 2, \ldots, m\}$. *Then the statements* (i) *and* (ii) *of Theorem* 1 *hold for every* $\theta \in \Theta$.

To state the next result, we need a weak topology, introduced by LeCam (1973), on the space of all sub-stochastic kernels on $\mathcal{B}^q$, $q \geqslant 1$, which may be described as follows.

Let $G(\mathcal{B}^q)$ be the space of all sub-stochastic measures on $\mathcal{B}^q$, $q \geqslant 1$. Let $(\mathcal{S}, \mathcal{E}, v)$ be a σ-finite measure space. Consider the sub-stochastic kernels $P : \mathcal{S} \to G(\mathcal{B}^q)$. Let $C_{00}(R^q)$ be the space of all continuous functions vanishing outside compacts. Denote the $C_{00}(R^q) \otimes L_1(v)$ topology of the set of all sub-stochastic kernels to be the smallest topology such that all functions

$$P \to \int [\int f(x)P(t)(dx)g(t)v(dt)],$$

$f \in C_{00}$, $g \in L_1(v)$, are continuous. It is known that the set of all sub-stochastic kernels endowed with this topology is metrisable and compact; a proof can be found in LeCam (1979, Ch. 8).

We now state

Theorem 2 : *Suppose that the sequence* $\{\mathcal{E}_n\}$ *of experiments satisfies, in addition to* (A.1)–(A.3), (A.5) *and* (A.6), *the condition* $(*)$: *for every finite* $\{h_i; 1, 2, \ldots, m\}$ *and almost all* $\theta \in \Theta$

$$\mathcal{L}(A_n(h_i, \theta), i = 1, 2, \ldots, m \mid P_{\theta,n})$$

$$\Longrightarrow \mathcal{L}(A(h_i, \theta), i = 1, 2, \ldots, m \mid \lambda_\theta).$$

Let the random functions $\gamma_n(\theta)$, $T_n(\theta)$, $n \geqslant 1$, $\gamma(\theta)$ *and* $T(\theta)$ *be as in Theorem* 1. *Then there exists a Lebesgue null set* $N \subseteq \Theta$ *such that for every* $\theta \in \Theta - N$, *the statement* (i) *of Theorem* 1 *is satisfied, and* (ii) *the sequence* $\{\mathcal{L}(W_n(\theta), \gamma_n(\theta), T_n(\theta) \mid P_{\theta,n})\}$ *is* $C_{00}(R^{2k+k^2}) \otimes L_1(\mu^k)$ *convergent to the stochastic kernel* $\mathcal{L}(T^{1/2}(\theta)Z + \gamma(\theta), \gamma(\theta), T(\theta) \mid \lambda_\theta)$, *where* Z *is a copy of the standard* k*-variate normal distribution independent of both* $\gamma(\theta)$ *and* $T(\theta)$.

Note that in Theorem 2, the existence of the limit distribution of the sequence $\{W_n(\theta)\}$ is not assumed.

Remark : In the special case when $A_n(h, \theta) = A(h, \theta)$ for every $n \geqslant 1$, where the function $h \rightarrow A(h, \theta)$ is non-random, it is easy to see directly from the statement (i) of Theorem 2 that there exists a Lebesgue null set $N \subseteq \Theta$ such that for every $\theta \in \Theta - N$, $\mathcal{L}(W_n(\theta) \mid P_{\theta,n})$ converges weakly to the k-variate normal distribution with mean vector $\gamma(\theta)$ and covariance matrix $T(\theta)$. Thus the conclusion of Theorem 4.1 of LeCam (1960) holds even when the existence of the limit distribution is not assumed.

Theorem 3 : *Assume that, for* $\theta_0 \in \Theta$,

(i) *there exists a sequence* $\{W_n(\theta)\}$ *of* $\mathcal{A}_n$*-measurable* k*-vectors and a sequence* $\{T_n(\theta_0)\}$ *of* $\mathcal{A}_n$*-measurable* $k \times k$*-symmetric matrices such that the difference*

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

converges to zero in $P_{\theta_0,n}$*-probability where* $\{\delta_n\}$ *is a sequence of p.d. matrices,*

(ii) *there exists an almost surely p.s.d. matrix* $T(\theta_0)$ *such that*

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

Then,

$$\text{(iii)} \qquad \mathcal{L}(W_n(\theta_0), T_n(\theta_0) \mid P_{\theta_0,n}) \Longrightarrow \mathcal{L}\,(T^{1/2}(\theta_0)Z, T(\theta_0))$$

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

(iv) *the sequence* $\{P_{\theta_0,n}\}$ *and* $\{P_{\theta_0+\delta_n h,n}\}$, $n \geqslant 1$, *are contiguous for every* $h \in R^k$, *and*

(v) $\mathcal{L}(T_n(\theta_0) \mid P_{\theta_0+\delta_n h,n}) \Longrightarrow \mathcal{L}(T(\theta_0))$ *for every* $h \in R^k$.

3. Proofs of the results

Note that Theorem 1 actually follows from Theorem 2. However, we will present the proof of Theorem 1 only for the following reasons. Firstly, the proofs of both the theorems are essentially identical. Secondly the arguments in the proof of Theorem 1 are more transparent and notationally less cumbersome.

We first present the following invariance lemma, a proof of which can be found in Jeganathan (1979b). A much more detailed and deep discussions on this type of invariance results can be found in LeCam (1974, Ch. 11) and LeCam (1979, Ch. 8).

(Invariance) Lemma 1 : *Assume that* $\Theta \subseteq R^k$ *is a measurable subset. Let* $\{F_n\}$ *be a sequence of (sub-stochastic) kernels* $F_n : \Theta \times R^k \to G(\mathcal{B}^q)$, $k, q \geqslant 1$, *satisfying*

$$F_n(\theta, h) = F_{\theta+\delta_n h, n}$$

for every $\theta \in \Theta$ *and* $h \in R^k$, *where* $\{\delta_n\}$ *is a sequence of p.d. matrices such that* $\|\delta_n\| \to 0$ *as* $n \to \infty$. *Then the following two statements hold.*

(i) *The sequence* $\theta \to F_n(\theta, h)$, $h \in R^k$, $n \geqslant 1$, *is* $C_{00}(R^q) \otimes L_1(\mu^k)$ *convergent to a kernel* $F(\theta, h)$ *if and only if the sequence* $\theta \to F_n(\theta, 0)$, $n \geqslant 1$, *is* $C_{00}(R^q) \otimes L_1(\mu^k)$ *convergent to the kernel* $F(\theta, 0) = F(\theta)$.

(ii) *The kernel* $F(\theta, h)$ *satisfies the invariance condition*

$$\int_{\Theta} \int_{R^q} f(x) F(\theta)(dx) g(\theta) d\theta = \int_{\Theta} \int_{R^q} f(x) F(\theta, h)(dx) g(\theta) d\theta$$

for every $f \in C_{00}$, $g \in L_1(\mu^k)$ *and* $h \in R^k$.

We will assume in what follows, for the sake of simplicity only, that $\dim \Theta = 1$. The proof of the next lemma is based on the proposition 1 of LeCam (1974, Ch. 11).

Lemma 2 : *Suppose that the assumptions of Theorem 1 are satisfied. Then there exist random variables* $\gamma(\theta)$ *and* $T(\theta)$, *and a Lebesgue null set* $N \subseteq \Theta$ *such that*

$$A(h, \theta) = h\gamma(\theta) + \tfrac{1}{2} h^2 T(\theta) \text{ a.s.}$$

for every $h \in R^k$ *and* $\theta \in \Theta - N$.

Proof: Denote, for $s, h \in R$ and $\theta \in \Theta$,

$$Z_{\theta,n}(h \mid s) = \frac{dP_{\theta+\delta_n h, n}}{dP_{\theta+\delta_n s, n}}.$$

Further, the vector whose elements are $A_n(s, \theta)$, $A_n(h, \theta)$, $A_n\left(\frac{s+h}{2}, \theta\right)$, $A_n(s+u, \theta)$, $A_n(h+u, \theta)$ and $A_n\left(\frac{s+h}{2}+u, \theta\right)$ will be denoted by $V_n(s, h, u, \theta)$ and the vector whose elements are $A(s, \theta)$, $A(h, \theta)$, $A\left(\frac{s+h}{2}, \theta\right)$, $A(s+u, \theta)$, $A(h+u, \theta)$ and $A\left(\frac{h+s}{2}+u, \theta\right)$ will be denoted by $V(s, h, u, \theta)$. In view of the statement (6) of Theorem 2.1 of LeCam (1960), it follows from the given conditions that

$$\mathcal{L}(V_n(s, h, u, \theta), Z_{\theta,n}(h+w \mid s+w) \mid P_{\theta+\delta_n(s+w),n})$$

$$\Longrightarrow \mathcal{L}\left(V(s, h, u, \theta), \frac{dG_{\theta,h+w}}{dG_{\theta,s+w}} \mid G_{\theta,s+w}\right) \qquad \dots (1)$$

for every $\theta \in \Theta$ and $s, h, u, w \in R$, where

$$dG_{\theta,h} = \exp[hW(\theta) - A(h, \theta)]d\lambda, \; h \in R.$$

(Note that in view of the contiguity condition $G_{\theta,h}$ is a probability measure for every $\theta \in \Theta$ and $h \in R$).

Further, the contition (A.6) and the invariance lemma 1 implies that both the sequences

$$\mathcal{L}(V_n(s, h, u, \theta), Z_{\theta,n}(h \mid s) P_{\theta+\delta_n s,n})$$

and

$$\mathcal{L}(V_n(s, h, u, \theta), Z_{\theta,n}(h+w \mid s+w) \mid P_{\theta+\delta_n(s+w),n})$$

are $C_{00}(R^7) \otimes L_1(\mu)$ convergent to a same kernel for every $s, h, u, w \in R$. Hence, this fact together with (1) implies that for every $s, h, u, w \in R$, there exists a Lebesgue null set $N(s, h, u, w)$ possibly depending on $(s, h. u, w)$ such that

$$\mathcal{L}\left(V(s, h, u, \theta), \frac{dG_{\theta,h+w}}{dG_{\theta,s+w}} \mid G_{\theta,s+w}\right)$$

$$= \mathcal{L}\left(V(s, h, u, \theta), \frac{dG_{\theta,h}}{dG_{\theta,s}} \mid G_{\theta,s}\right) \qquad \dots (2)$$

whenever $\theta \in \Theta - N(s, h, u, w)$. Let D be the set of all points in R^4 with rational co-ordinates, and let

$$N = \bigcup_{(s, h, u, w) \in D} N(s, h, u, w).$$

Then it follows that whenever $\theta \in \Theta - N$, the equality in the expression (2) holds for every $(s, h, u, w) \in D$. In particular it easily follows that, for every $(s, h, u, w) \in D$ and $\theta \in \Theta - N$,

$$E^V \left\{ \sqrt{\frac{dG_{\theta, h+u}}{d\lambda_\theta} \frac{dG_{\theta, s+u}}{d\lambda_\theta}} \right\}$$

$$= E^V \left\{ \sqrt{\frac{dG_{\theta, h}}{d\lambda_\theta} \frac{dG_{\theta, s}}{d\lambda_\theta}} \right\} \text{a.s.} \qquad \dots (3)$$

and

$$E^V \left(\frac{dG_{0, w}}{d\lambda_\theta} \right) = 1 \quad \text{a.s.} \qquad \dots (4)$$

where E^V denotes the conditional exprectation given $V(s, h, u, \theta)$ with the underlying probability space being $(\mathcal{Y}, \mathcal{F}, \lambda_\theta)$. In what follows assume that $\theta \in \Theta - N$ is fixed. Since we have, as is easily checked using (4),

$$\log E^V \left\{ \sqrt{\frac{dG_{\theta, h+u}}{d\lambda_\theta} \frac{dG_{\theta, s+u}}{d\lambda_\theta}} \right\}$$

$$= -\frac{1}{2} \left[A(s+u, \theta) + A(h+u, \theta) - 2A\left(\frac{s+h}{2} + u, \theta\right) \right]$$

for every $s, h, u \in R$, it follows from (3) that for every rational $s, h, u \in R$

$$A(s+u, \theta) + A(h+u, \theta) - 2A\left(\frac{s+h}{2} + u, \theta\right)$$

$$= A(s, \theta) + A(h, \theta) - 2A\left(\frac{s+h}{2}, \theta\right) \text{ a.s.}$$

Hence in view of the condition (A.5) there exists a set of λ_θ-measure zero such that outside this null set

$$A(s+u, \theta) + A(h+u, \theta) - 2A\left(\frac{s+h}{2} + u, \theta\right)$$

$$= A(s, \theta) + A(h, \theta) - 2A\left(\frac{s+h}{2}, \theta\right)$$

for every $s, h, u \in R$, i.e., the random function $h \to A(h, \theta)$ has constant second differences outside a set of λ_θ-measure zero. This implies that there exist random variables $\gamma(\theta)$ and $T(\theta)$ such that, (note that $A(0, \theta) = 0$ a.s.),

$$A(h, \theta) = h\gamma(\theta) + \frac{1}{2} h^2 T(\theta) \text{ a.s.}$$

for every $\theta \in \Theta - N$. This completes the proof of the lemma.

Proof of Theorem 1 : First note that, for almost all Θ, $T(\theta)$ is the second difference of the random function $h \to A(h, \theta)$ at $h = 0$ and that $\gamma(\theta)$ can be expressed in terms of the first and second differences of the function $h \to A(h, \theta)$ at $h = 0$. Hence it follows from the relation (4) and lemma 2 together with a simple continuity argument that there exists a Lebesgue null set $N \subseteq \Theta$ such that

$$E^{(\gamma(\theta), T(\theta))}[\exp hW(\theta)] = \exp[h\gamma(\theta) + \tfrac{1}{2} h^2 T(\theta)] \text{ a.s.}$$

for every $h \in R$ and $\theta \in \Theta - N$. Hence it follows that $W(\theta)$ is distributed as $T^{1/2}(\theta)Z + \gamma(\theta)$ where Z is a copy of the standard normal distribution independent of both $\gamma(\theta)$ and $T(\theta)$. Now let $T_n(\theta)$ be the second difference of $h \to A_n(h, \theta)$ at $h = 0$, $\theta \in \Theta$, $n \geqslant 1$, and define $\gamma_n(\theta)$ similarly. This $T_n(\theta)$ need not be non-negative, but this can be easily remedied since $T(\theta)$ is non-negative. This completes the proof of Theorem 1.

Proof of Theorem 3 : It is clear that the statements (i) and (iii) imply the statement (iv). In view of the statement (6) of Theorem (2.1) of LeCam (1960), to prove that (i) and (iii) implies (v) it is enough to show that

$$\begin{aligned} &\int \exp[itT(\theta_0) + hT^{1/2}(\theta_0)Z - \tfrac{1}{2} h^2 T(\theta_0)] d\, \mathcal{L}(T(\theta_0), Z) \\ &= \int \exp(itT(\theta_0)) d\, \mathcal{L}(T(\theta_0)) \end{aligned}$$

for every $t, h \in R$. Using the independence of Z and $T(\theta_0)$ it is easy to verify this equality. This proves the necessary part of the theorem. To prove the sufficiency part, first note that, in view of the statement (4) of Theorem 2.1 of LeCam (1960), the sequence $\{W_n(\theta_0), T_n(\theta_0)\}$ is relatively compact for the sequence $\{P_{\theta_0, n}\}$. Hence for every sub-sequence there exists a further subsequence $\{m\} \subset \{n\}$ and a random vector (T', W) such that

$$\mathcal{L}(T_m(\theta_0), W_m(\theta_0) \mid P_{\theta_0, m}) \Longrightarrow \mathcal{L}(T', W).$$

In view of the statement (6) of Theorem (2.1) of LeCam (1960) we then have from the statement (iv) that

$$\mathcal{L}(T') = \mathcal{L}(T' \mid R_h) \quad \ldots \quad (5)$$

for every $h \in R$, where the probability measure R_h is defined by

$$dR_h = \exp\left(hW - \frac{h^2}{2}T'\right) d\,\mathcal{L}(T', W).$$

In particular (5) implies that

$$E^{T'}[\exp(hW)] = \exp(\tfrac{1}{2}h^2T')$$

and hence

$$E[\exp(itT'+iuW]$$

$$= E[\exp(itT' - \tfrac{1}{2}u^2T']$$

$$= E[\exp(itT(\theta_0) - \tfrac{1}{2}u^2T(\theta_0))].$$

This proves the sufficiency part.

4. Discussion on the measurability conditions

The purpose of this section is to show that under a condition which is slightly stronger than (A.2), the joint measurability of $W_n(\theta)$ and $A_n(h, \theta)$, $n \geqslant 1$, can be removed in both the theorems 1 and 2 and further that the condition (A.1) can be removed in Theorem 1. Consider the following condition :

(A.2′). There exist $\mathcal{A}_n$-measurable functions $W_n(\theta)$, $A_n(h, \theta)$, $n \geqslant 1$, $h \in R^k$, $\theta \in \Theta$ such that the difference

$$\frac{dP_{\theta+\delta_n h_n, n}}{dP_{\theta, n}} - \exp[h_n' W_n(\theta) - A_n(h_n, \theta)]$$

converges to zero in $P_{\theta, n}$ probability for every bounded sequence $\{h_n\}$ of R^k and that the difference

$$A_n(h_n, \theta) - A_n(h_n^*, \theta)$$

converges to zero for every sequence $\{h_n^*\}$ of R^k satisfying $|h_n - h_n^*| \to 0$, where the sequence $\{\delta_n\}$ of p.d. matrices is such that $\|\delta_n\| \to 0$.

Lemma 3 : *Suppose that the conditions* (A.2′) *and* (A.3) *are satisfied. Then the quantity*

$$\left\| P_{\theta+\delta_n h_n, n} - P_{\theta+\delta_n h_n^*, n} \right\| \to 0 \qquad \ldots \quad (6)$$

for every bounded sequences $\{h_n\}$ *and* $\{h_n^*\}$ *of* R^k *satisfying* $|h_n - h_n^*| \to 0$.

Proof : It is easy to see that the given conditions entails that the sequence $\{P_{\theta+\delta_n h_n, n}\}$, $\{P_{\theta+\delta_n h_n^*, n}\}$ and $\{P_{\theta, n}\}$ are contiguous, and hence we can assume without loss of generality that $P_{\theta+\delta_n h_n, n} \approx P_{\theta, n} \approx P_{\theta+\delta_n h_n^*, n}$ for every $\theta \in \Theta$ and $n \geqslant 1$, where the symbol $\approx$ denotes mutual absolute continuity. Set, for every $h \in R^k$, $\theta \in \Theta$ and $n \geqslant 1$.

$$Z_{\theta, n}(h) = \frac{dP_{\theta+\delta_n h, n}}{dP_{\theta, n}}.$$

Now note that the difference

$$Z_{\theta, n}(h_n) - Z_{\theta, n}(h_n^*) \qquad \ldots \quad (7)$$

converges to zero in $P_{\theta, n}$ probability for every $\theta \in \Theta$. Next note that

$$\int_{\{|Z_{\theta, n}(h_n)| > \alpha\}} Z_{\theta, n}(h_n) dP_{\theta, n} = P_{\theta+\delta_n h_n, n}[\,|Z_{\theta, n}(h_n)| > \alpha]$$

and hence, in view of contiguity,

$$\lim_{\alpha \to \infty} \limsup_{n \to \infty} \int_{\{|Z_{\theta, n}(h_n)| > \alpha\}} Z_{\theta, n}(h_n) dP_{\theta, n} = 0. \qquad \ldots \quad (8)$$

Similarly

$$\lim_{\alpha \to \infty} \limsup_{n \to \infty} \int_{\{|Z_{\theta, n}(h_n^*)| > \alpha\}} Z_{\theta, n}(h_n^*) dP_{\theta, n} = 0. \qquad \ldots \quad (9)$$

Now combining (7), (8) and (9) we have

$$\int |Z_{\theta, n}(h_n) - Z_{\theta, n}(h_n^*)|\, dP_{\theta, n} \to 0.$$

Hence the proof of the lemma is complete.

Lemma 4 : *Suppose that the sequence of experiments* $\mathcal{E}_n = \{\mathcal{X}_n, \mathcal{A}_n, P_{\theta, n}; \theta \in \Theta\}$, $n \geqslant 1$, *satisfies the condition* (6) *of lemma* 3. *Then there exists* *a construction of another sequence of experiments* $\mathcal{E}_n^* = \{\mathcal{X}_n, \mathcal{A}_n, P_{\theta,n}^*; \theta \in \Theta\}$, $n \geqslant 1$, *with the following properties.*

(i) *The functions* $\theta \to P_{\theta,n}^*(A)$, $A \in \mathcal{A}_n$, $n \geqslant 1$, *are Borel measurable,*

(ii) *the functions*

$$(\mathcal{X}_n \times \Theta \times R^k) \to \log \frac{dP^*_{\theta+\delta_n h, n}}{dP^*_{\theta,n}}, \quad n \geqslant 1,$$

are jointly measurable,

(iii) *for every* $\alpha > 0$ *and* $\theta \in \Theta$

$$\sup_{|h| \leqslant \alpha} \|P_{\theta+\delta_n h, n} - P^*_{\theta+\delta_n h, n}\| \to 0; \quad and$$

(iv) *when the sequence* $\{\mathcal{E}_n\}$ *further satisfies the condition that the sequences* $\{P_{\theta+\delta_n h_n, n}\}$ *and* $\{P_{\theta,n}\}$ *are contiguous for every bounded sequence* $\{h_n\}$ *of* R^k *the difference*

$$\log \frac{dP_{\theta+\delta_n h_n, n}}{dP_{\theta,n}} - \log \frac{dP^*_{\theta+\delta_n h_n, n}}{dP^*_{\theta,n}}$$

converges to zero in $P_{\theta,n}$ *probability for every* $\theta \in \Theta$ *and for every bounded sequence* $\{h_n\}$ *of* R^k.

Proof: See LeCam (1974, pp. 153–155).

Proposition: *Suppose that the sequence* $\{\ _n\}$ *of experiments satisfies the conditions* (A.2′), (A.3), (A.5) *and* (A.6) *and the condition* (•) *of Theorem* 2. *Then there exist* $\mathcal{A}_n \times \mathcal{B}^k$*-measurable functions* $W_n^*(\theta): \mathcal{X}_n \times \Theta \to R^k$ *and* $\mathcal{A}_n \times \mathcal{B}^k \times \mathcal{B}^k$*-measurable functions* $A_n^*(h, \theta): \mathcal{X}_n \times R^k \times \Theta \to R$, $n \geqslant 1$, *such that*

(i) *the condition* (A.2′) *is satisfied with the functions* $W_n(\theta)$ *and* $A_n(h, \theta)$ *are replaced by* $W_n^*(\theta)$ *and* $A_n^*(h, \theta)$ *respectively,*

(ii) *for every* $h \in R^k$ *and* $\theta \in \Theta$, *the differences*

$$A_n^*(h, \theta) - \left[A_n(h, \theta) - \sum_{i=1}^{k} h_i A_n(\varepsilon_i, \theta)\right]$$

and

$$h' W_n^*(\theta) - \left[h' W_n(\theta) - \sum_{i=1}^{k} h_i A_n(\varepsilon_i, \theta)\right]$$

converges to zero in $P_{\theta,n}$-probability where $\{\varepsilon_i;\ i = 1, 2, \ldots, k\}$ is a basis of R^k and h_i's are such that $h = \sum_{i=1}^{k} h_i \varepsilon_i$,

(iii) *for every $\varepsilon > 0$, $h \in R^k$ and $g \in L_1(\mu^k)$,*

$$\int_{\Theta} \int_{\mathscr{X}_n} I[\,|A_n^*(s, \theta+\delta_n h) - A_n^*(s, \theta)| > \varepsilon]\, dP_{\theta+\delta_n h, n}\, g(\theta) d\theta \to 0 \qquad \ldots \quad (10)$$

and

(iv) *the condition* (A.5) *is satisfied for the corresponding limit of the sequence* $\{A_n^*(h, \theta)\}$.

Remark : Note that the above condition (10) is weaker than the condition (A.6), but what we have really used in the proof of Theorem 1 is the above condition (10). It is possible to show, under the condition (A.4) which is stronger than the condition (*) of Theorem 2, that there exists a Lebesgue null set $N \subseteq \Theta$ such that the condition (A.6) is satisfied for the sequence $\{A_n^*(h, \theta)\}$ whenever $\theta \in \Theta - N$; the proof of this statement will not be presented here though the arguments of the proof seem to be somewhat non-trivial.

Proof : Let $\{\varepsilon_i;\ i = 1, 2, \ldots, k\}$ be a basis of R^k. Define $W_n^*(\theta)$ by

$$h' W_n^*(\theta) = \sum_{i=1}^{k} h_i\, \Lambda_n^*(\theta + \delta_n \varepsilon_i, \theta)$$

where h_i's are such that

$$h = \sum_{i=1}^{k} h_i\, \varepsilon_i,\ h \in R^k,$$

and

$$\Lambda_n^*(\theta+\delta_n h, \theta) = \log \frac{dP^*_{\theta+\delta_n h, n}}{dP^*_{\theta, n}}.$$

Now define $A_n^*(h, \theta)$ by

$$A_n^*(h, \theta) = h' W_n^*(\theta) - \Lambda_n^*(\theta + \delta_n h, \theta).$$

It is clear from the statements (ii) and (iv) of lemma 4, that the condition (A.2′) is satisfied with the functions $W_n(\theta)$ and $A_n(h, \theta)$ are replaced by $W_n^*(\theta)$

and $A_n^*(h, \theta)$ respectively and further that the functions $W_n^*(\theta)$ and $A_n^*(h, \theta)$ are jointly measurable. Now note that

$$A_n^*(h, \theta) = h' W_n^*(\theta) - \Lambda_n^*(\theta + \delta_n h, \theta)$$

$$= \sum_{i=1}^{k} h_i \Lambda_n^*(\theta + \delta_n \varepsilon_i, \theta) - \Lambda_n^*(\theta + \delta_n h, \theta)$$

and in view of the statement (iv) of lemma 4, this can be approximated with $P_{\theta, n}$-probability tending to one by $A_n(h, \theta) - \sum_{i=1}^{k} h_i A_n(\varepsilon_i, \theta)$. Hence the statements (ii) and (iv) follow. Now note that, since for every $s \in R^k$ and $\theta \in \Theta$

$$\Lambda_n^*(\theta + \delta_n s, \theta) - [s' W_n(\theta) - A_n(s, \theta)]$$

converges to zero in $P_{\theta, n}$-probability, the invariance lemma 1 and the condition (A.6) implies that, for every $\varepsilon > 0$, $h, s \in R^k$ and $g \in L_1(\mu^k)$

$$\int_{\Theta} \int_{\mathcal{X}_n} I[\,|\,\Lambda_n^*(\theta + \delta_n(s+h), \theta + \delta_n h)$$

$$-[s' W_n(\theta + \delta_n h) - A_n(s, \theta)]\,| > \varepsilon] dP_{\theta + \delta_n h, n}\, g(\theta) d\theta \to 0. \quad \ldots \quad (11)$$

Now, writing

$$s = \sum_{i=1}^{k} s_i \varepsilon_i,$$

$$A_n^*(s, \theta + \delta_n h) = \sum_{i=1}^{k} s_i \Lambda_n^*(\theta + \delta_n(\varepsilon_i + h),\ \theta + \delta_n h)$$

$$-\Lambda_n^*(\theta + \delta_n(s+h), \theta + \delta_n h),$$

and hence the statement (iii) follows by using (11) and the statement (ii). This completes the proof of the proposition.

Acknowledgement

I am greatly indebted to Prof. L. LeCam for having suggested the problem and for helpful suggestions. I further wish to express my deep gratitude to my supervisor Prof. J. K. Ghosh for his encouragement and for his careful reading of the manuscript and suggesting several improvements.

References

Basawa, I. V. and Prakasa Rao, B. L. S. (1979) : Asymptotic inference for stochastic processes. To appear in *J. Stoch. Proc.* Appl.

Bhat, B. R. (1978) : Some problems of estimation for dependent observations. *Technical Report*, Dept. of Statistics, Univ. of Poona.

Davies, R. B. (1979) : Asymptotic inference when the amount of information is random. Preprint, DSIR, Wellington, New Zealand.

Jeganathan, P. (1979a) : On the asymptotic theory of statistical estimation for dependent observations. Discussion Paper No. 7906, January 1979, ISI, New Delhi. (Revised version, August 1979).

Jeganathan, P. (1979b) : On a decomposition of the limit distribution of a sequence of estimators. To appear in *Sankhya*.

——— (1979a) : On a decomposition of the limit distribution of a sequence of estimators.

——— (1980) : Some asymptotic properties of risk functions in estimation when the limit of the experiment is mixed normal.

LeCam, L. (1960) : Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.* **3**, 27-98.

——— (1973) : Sur les contraintes imposes par les passages a la limite usuels en statistique. *Proc. 39th Session of International Statistical Institute*, **4**, 169-177.

——— (1974) : *Notes on asymptotic methods in statistical decision theory*. Centre de Recherches Mathematiques, Universate de Montreal.

——— (1979) : Asymptotic methods in statistical decision theory. Manuscript for book in preparation.

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