ON THE EXPONENTIAL APPROXIMATION OF A FAMILY OF PROBABILITY MEASURES AND A REPRESENTATION THEOREM OF HÁJEK-INAGAKI

G. G. ROUSSAS* AND A. SOMS

(Received Oct. 18, 1971; revised Aug. 14, 1972)

1. Introduction and summary

Let $(\mathcal{X}, \mathcal{A})$ be a measurable space and let Θ be an open subset of the k-dimensional space R^k. For each $\theta \in \Theta$, let P_θ be a probability measure on $\mathcal{A}$. Let $\{X_n, n \geqq 0\}$ be a discrete parameter Markov process defined on $(\mathcal{X}, \mathcal{A}, P_\theta)$, X_n taking values in the Borel real line $(R, \mathcal{B})$. Finally, let $\mathcal{A}_n$ be the σ-field induced by the r.v.'s $X_0, X_1, \cdots, X_n$ and let $P_{n,\theta}$ be the restriction of P_θ to the σ-field $\mathcal{A}_n$.

It has been shown in Johnson and Roussas [5] that for an arbitrary $\theta \in \Theta$, the probability measure $P_{n,\theta}$ may be approximated, in the L_1-norm sense and in the neighborhood of θ, by an exponential probability measure. One of the purposes of the present paper is to present a simpler proof of the exponential approximation just mentioned; this is done by exploiting some ideas taken from Hájek [3]. In the course of the proof, we also establish a lemma of some independent interest. This auxiliary result is discussed in Section 3 and the first main result of this paper is presented in Section 4.

In the reference [3] cited above, Hájek considered a class of (weakly) convergent sequences of properly normalized estimates of θ and showed that, under suitable regularity conditions, the limiting probability measures may be represented as the convolution of two probability measures one of which is normal. He then employed this representation in order to obtain, in a unified and elegant manner, certain results in asymptotic efficiency of estimates discussed by Wolfowitz [12], Kaufman [6], Schmetterer [11] and Roussas [10]. The representation result mentioned above has also been established by Inagaki in a nice paper [4] independently and almost simultaneously with Hájek. In the present paper, it is shown that the assumptions given in Section 2 are sufficient to allow us to establish the Hájek-Inagaki representation theorem; the proof of the

* This research was supported by the National Science Foundation, Grant GP-20036.

theorem itself is entirely different from the ones given by Hájek and Inagaki and is based on an idea due to Bickel [1]. Still another proof of the theorem in question is obtained as a special case of some general results which are discussed by LeCam [7]. The proof of the above mentioned representation theorem is presented in Section 5.

2. Notation and assumptions

The notation and assumptions to be employed here are essentially the same as the ones first used in Roussas [9] and also in Johnson and Roussas [5]. For the sake of completeness, however, they will also be explicitly mentioned below.

Let Θ, $\{X_n, n \geqq 0\}$, $(\mathcal{X}, \mathcal{A}, P_\theta)$ and $P_{n,\theta}$ be as in Section 1. It will be assumed in the following that the probability measures $\{P_{n,\theta}, \theta \in \Theta\}$, $n \geqq 0$, are mutually absolutely continuous. Therefore for any $\theta, \theta^* \in \Theta$, we may set

$$[dP_{0,\theta^*}/dP_{0,\theta}] = q(X_0; \theta, \theta^*) , \qquad [dP_{1,\theta^*}/dP_{1,\theta}] = q(X_0, X_1; \theta, \theta^*) .$$

Furthermore, let

$$f_j(\theta, \theta^*) = [q(X_{j-1}, X_j; \theta, \theta^*)]^{1/2} , \qquad j = 1, \cdots, n ,$$

$$q(X_1 | X_0; \theta, \theta^*) = q(X_0, X_1; \theta, \theta^*)/q(X_0; \theta, \theta^*)$$

and

$$\varphi_j(\theta, \theta^*) = [q(X_j | X_{j-1}; \theta, \theta^*)]^{1/2} , \qquad j = 1, \cdots, n ,$$

so that $\int_{\mathcal{X}} \varphi_1^2(\theta, \theta^*) dP_{1,\theta} = 1$.

ASSUMPTIONS. (A1) For each $\theta \in \Theta$, the Markov process $\{X_n, n \geqq 0\}$ is stationary and metrically transitive (ergodic). (See, e.g., Doob [2], p. 457.)

(A2) The probability measures $\{P_{n,\theta}, \theta \in \Theta\}$, $n \geqq 0$, are mutually absolutely continuous.

(A3) (i) For each $\theta \in \Theta$, the random function $\varphi_1(\theta, \theta^*)$ is differentiable in quadratic mean (q.m.) with respect to θ^* at (θ, θ) when P_θ is employed.

Let $\dot{\varphi}_1(\theta)$ be the derivative of $\varphi_1(\theta, \theta^*)$ with respect to θ^* at (θ, θ). Then

(ii) $\dot{\varphi}_1(\theta)$ is $\mathcal{A}_1 \times \mathcal{C}$-measurable, where $\mathcal{C}$ denotes the σ-field of Borel subsets of Θ.

Let $\Gamma(\theta)$ be the covariance function defined by $\Gamma(\theta) = 4\mathcal{E}_\theta[\dot{\varphi}_1(\theta)\dot{\varphi}_1'(\theta)]$. Then

(iii) $\Gamma(\theta)$ is positive definite for every $\theta \in \Theta$.

(A4) For every $\theta \in \Theta$, the random function $f_1(\theta, \theta^*)$ is continuous in $P_{1,\theta}$-probability at (θ, θ).

It should be pointed out here that in the case that the r.v.'s involved are i.i.d., Assumption (A1) is automatically satisfied and Assumption (A4) follows from Assumption (A3)(i). Thus in this case, the assumptions made become as follows.

(A1)′ For each $\theta \in \Theta$, the distribution of X_0 under $P_{0,\theta}$ is absolutely continuous with respect to some σ-finite measure μ on $\mathcal{B}$, and if $q(\cdot\,;\theta)$ is a specified version of the Radon-Nikodym derivative involved, then $q(\cdot\,;\theta)$ is positive on a set independent of θ.

Set $\varphi_1(\theta, \theta^*)=[q(X_1;\theta^*)/q(X_1;\theta)]^{1/2}$. Then

(A2)′ (i) For each $\theta \in \Theta$, the random function $\varphi_1(\theta, \theta^*)$ is differentiable in q.m. with respect to θ^* at (θ, θ) when P_θ is employed.

Let $\dot{\varphi}_1(\theta)$ be as above. Then

(ii) $\dot{\varphi}_1(\theta)$ is $X_1^{-1}(\mathcal{B})\times\mathcal{C}$-measurable, where $\mathcal{C}$ is as above.

(iii) For every $\theta \in \Theta$, $\Gamma(\theta)=4\mathcal{E}_\theta[\dot{\varphi}_1(\theta)\dot{\varphi}_1'(\theta)]$ is positive definite.

In all that follows, all limits will be taken as $\{n\}$, or subsequences thereof, converges to infinity unless otherwise explicitly stated, and integrals without limits will be taken over the entire (appropriate) space.

3. Some auxiliary lemmas

In this section, we establish a lemma and a corollary to be used in the proof of the main results in Sections 4 and 5 of this paper. To this end, consider the r.v.'s U_n, $n=1, 2, \cdots$ defined on the probability space $(\Omega, \mathcal{F}, P)$ and recall that the r.v.'s $|U_n|$, $n=1, 2, \cdots$ are said to be *uniformly integrable* if $\int_{(|U|>a)} |U_n|\,dP \to 0$ uniformly in n as $a\to\infty$ (see, e.g., Loève [8], p. 162).

LEMMA 3.1. *For $n=1, 2, \cdots$, let U_n be r.v.'s defined on the probability space $(\Omega, \mathcal{F}, P)$ and let U be a r.v. defined on the probability space $(\Omega', \mathcal{F}', P')$. Suppose that $\mathcal{E}(|U_n||P)=\mathcal{E}|U_n|\to\mathcal{E}|U|=\mathcal{E}(|U||P')$ finite and $\mathcal{L}(U_n|P)\Rightarrow\mathcal{L}(U|P')$. Then the r.v.'s $|U_n|$, $n\geqq1$, are uniformly integrable.*

This lemma is only a slightly different formulation of a theorem in Loève (see A (iii), p. 183) and is taken from that with $g(x)=x$, $x \in R$.

The following lemma relates uniform integrability of the r.v.'s $|U_n-V_n|$, $n\geqq1$, to that of the r.v.'s $|U_n|$, $|V_n|$, $n\geqq1$. More precisely, one has

LEMMA 3.2. *For $n\geqq1$, let U_n, V_n be r.v.'s defined on the probability space $(\Omega, \mathcal{F}, P)$ and suppose that the r.v.'s $|U_n|$, $n\geqq1$, and $|V_n|$, $n\geqq1$,* *are uniformly integrable and that $U_n - V_n \to 0$ in P-probability. Then the r.v.'s $|U_n - V_n|$, $n \geqq 1$, are uniformly integrable.*

PROOF. Clearly,

$$\int_{(|U_n - V_n| \geqq a)} |U_n - V_n| dP \leqq \int_{(|U_n - V_n| \geqq a)} |U_n| dP + \int_{(|U_n - V_n| \geqq a)} |V_n| dP \tag{3.1}$$

and

$$\begin{aligned}\int_{(|U_n - V_n| \geqq a)} |U_n| dP &= \int_{(|U_n - V_n| \geqq a) \cap (|U_n| \geqq c)} |U_n| dP \\ &\quad + \int_{(|U_n - V_n| \geqq a) \cap (|U_n| < c)} |U_n| dP \\ &\leqq \int_{(|U_n| \geqq c)} |U_n| dP + cP(|U_n - V_n| \geqq a) .\end{aligned} \tag{3.2}$$

By the uniform integrability of the r.v.'s $|U_n|$, $n \geqq 1$, one has that $\int_{(|U_n| \geqq c)} |U_n| dP < \frac{\varepsilon}{4}$ for all sufficiently large c, whereas for each such a c and all sufficiently large a, one also has that $P(|U_n - V_n| \geqq a) < \frac{\varepsilon}{4c}$ for all $n \geqq n_1$, say. Therefore for a c and an a as just described, relation (3.2) becomes as follows

$$\int_{(|U_n - V_n| \geqq a)} |U_n| dP < \frac{\varepsilon}{2}, \qquad n \geqq n_1 . \tag{3.3}$$

In a similar manner, one obtains the inequality

$$\int_{(|U_n - V_n| \geqq a)} |V_n| dP < \frac{\varepsilon}{2}, \qquad n \geqq n_2, \text{ say} . \tag{3.4}$$

Increasing a so that (3.3) and (3.4) hold true for $n = 1, 2, \cdots, n_3 = \max(n_1, n_2)$, we obtain the desired result by means of (3.1).

COROLLARY 3.1. *Under the assumptions of Lemma 3.2, one has that $\mathcal{E}|U_n - V_n| \to 0$.*

PROOF. It is an immediate consequence of the assumption $U_n - V_n \to 0$ in P-probability, the uniform integrability of the r.v.'s $|U_n - V_n|$, $n \geqq 1$, as concluded in the lemma, and the L_r-convergence theorem (see, e.g., Loève [8], p. 163).

4. Exponential approximation

The purpose of the present section is to show that the probability measure $P_{n,\theta}$ may be approximated in the L_1-norm sense and in the neighborhood of θ by an exponential probability measure. In carrying out this approximation, we are going to use some of the results obtained in Johnson and Roussas [5]. To start with, let θ be an arbitrarily chosen by fixed point of Θ and consider the k-dimensional random vector $\Delta_n(\theta)$ defined by

$$\Delta_n = \Delta_n(\theta) = 2n^{-1/2} \sum_{j=1}^{n} \dot{\varphi}_j(\theta) . \tag{4.1}$$

In the reference last cited, θ was taken to be the origin in Θ. Since the results obtained there do not hinge on this choice of θ, we will not insist on it here. Next, let $\Delta_n^* = \Delta_n^*(\theta)$ be the truncated version of Δ_n defined by (4.6) in Johnson and Roussas [4] and consider the (exponential with parameter $h \in R^k$) probability measure $R_{n,h}$ defined by (5.1) and (5.2) in the reference just cited, namely,

$$R_{n,h}(A) = \exp[-B_n(h)] \int_A \exp(h'\Delta_n^*) dP_{n,\theta} , \qquad A \in \mathcal{A}_n , \tag{4.2}$$

where

$$\exp B_n(h) = \mathcal{E}_\theta \exp h'\Delta_n^* . \tag{4.3}$$

Let $\{h_n\}$ be a bounded sequence of points in R^k and set $\theta_n = \theta + h_n n^{-1/2}$. Then the main result of this section is the following

THEOREM 4.1. *In terms of the notation introduced so far, one has*

$$\|R_{n,h_n} - P_{n,\theta_n}\| = 2 \sup [|R_{n,h_n}(A) - P_{n,\theta_n}(A)| ; A \in \mathcal{A}_n] \to 0 . \tag{4.4}$$

PROOF. The proof is by contradiction. Assume that (4.4) is not true. Then there exists a subsequence $\{m\} \subseteq \{n\}$ and $\{h_m\} \subseteq \{h_n\}$ with $h_m \to h$ such that

$$\|P_{m,\theta_m} - P_{m,h_m}\| \not\to 0 . \tag{4.5}$$

From Theorem 3.1.1 in Roussas [9], one, clearly, has that

$$[dP_{m,\theta_m}/dP_{m,\theta}] = L_m(h_m) = \exp[h_m'\Delta_m - A(h_m) + Z_m(h_m)] , \tag{4.6}$$

where

$$A(h_m) = \frac{1}{2} h_m'\Gamma(\theta)h_m \quad \text{and} \quad Z_m(h_m) \to 0 \text{ in } P_{m,\theta}\text{-probability} . \tag{4.7}$$

From (4.2), one has

$$[dR_{m,h_m}/dP_{m,\theta}] = L_m^*(h_m) = \exp[-B_m(h_m) + h_m'\Delta_m^*] . \tag{4.8}$$

Then from (4.6) and (4.8), it follows that

$$(4.9)\qquad |L_m(h_m)-L_m^*(h_m)|\leqq[|h_m'(\Delta_m-\Delta_m^*)|+|Z_m(h_m)| \\ +|B_m(h_m)-A(h_m)|]\exp T_m\ ,$$

where T_m is a r.v. lying between the r.v.'s $h_m'\Delta_m-A(h_m)+Z_m(h_m)$ and $-B_m(h_m)+h_m'\Delta_m^*$. The second term on the right-hand side in (4.9) converges to 0 in $P_{m,\theta}$-probability by (4.7), whereas the first term on the same side does likewise by means of Proposition 4.1 in Johnson and Roussas [5]. Finally, by (5.17) in the last reference, one has that $\exp B_m(h_m)-\exp A(h_m)\to 0$. (For a proof of this fact somewhat simpler than the one referred to, see Lemma 4.1 below.) By virtue of (4.7), $A(h_m)\to A(h)$, so that $B_m(h_m)-A(h_m)\to 0$. Therefore (4.9) implies that

$$(4.10)\qquad L_m(h_m)-L_m(h_m^*)\to 0 \quad \text{in } P_{m,\theta}\text{-probability}\ .$$

Next by Theorem 3.2.1 in Roussas [9],

$$(4.11)\qquad \mathcal{L}[L_m(h_m)\,|\,P_{m,\theta}]\Rightarrow\mathcal{L}(\exp\Delta^*|Q^*) \quad\text{and}\quad \mathcal{E}_{Q^*}\exp\Delta^*=1\ ,$$

where Δ^* is the identity mapping on R and $Q^*=N(-(1/2)h'\Gamma h, h'\Gamma h)$, $\Gamma=\Gamma(\theta)$. Similar results hold also true for $L_m^*(h_m)$, as follows from (4.8) and the fact that $\Delta_m-\Delta_m^*\to 0$ in $P_{m,\theta}$-probability and $\exp B_m(h_m)\to\exp A(h)$. Therefore Lemma 3.1 applies with $\{n\}$ replaced by $\{m\}$, $(\Omega, \mathcal{F}, P)=(\mathcal{X}, \mathcal{A}, P_\theta)$, $(\Omega', \mathcal{F}', P')=(R, \mathcal{B}, Q^*)$ and $U_m=L_m(h_m)(L_m^*(h_m))$, $U=\exp\Delta^*$ and gives that the r.v.'s $|L_m(h_m)|=L_m(h_m)(|L_m^*(h_m)|=L_m^*(h_m))$ corresponding to the sequence $\{m\}$ are uniformly integrable. This result, together with (4.10) and Corollary 3.1, gives that $\int|L_m(h_m)-L_m^*(h_m)|\,dP_{m,\theta}\to 0$ which amounts to a contradiction to (4.5). The proof of the theorem is completed.

COROLLARY 4.1. *Let B be any bounded set in R^k. Then one has that*

$$\sup\,(\|P_{n,\theta_n}-R_{n,h}\|\,;\,h\in B,\ \theta_n=\theta+hn^{-1/2})\to 0\ .$$

PROOF. It follows by a contradiction argument.

This section is closed with the following lemma referred to above.

LEMMA 4.1 *With the quantities $B_n(h)$ and $A(h)$ defined by* (4.3) *and* (4.7), *respectively, one has*

$$(4.12)\qquad \exp B_m(h_m)-\exp A(h_m)\to 0\ ,$$

where $\{m\}\subseteq\{n\}$ and $h_m\to h\in R^k$.

PROOF. Let Δ be the identity mapping on R^k and let $Q=N(0,\Gamma)$, where we recall that Γ stands for $\Gamma(\theta)$. Then $\int\exp h'\Delta dQ$ is the moment generating function of $h'\Delta$ evaluated at $t=1$. Since this is equal to $\exp A(h)$ and $A(h_m)\to A(h)$, in order to prove (4.12) it suffices to show that

$$\exp B_m(h_m)\to\int \exp h'\Delta dQ \quad \left(=\int \exp h'zdQ\right) . \tag{4.13}$$

For simplicity, set $\mathcal{L}_n=\mathcal{L}(\Delta_n^*\,|\,P_{n,\theta})$. Then $\mathcal{L}_n\Rightarrow Q$ by means of Theorem 3.2.1 in Roussas [8] and the fact that $\Delta_n-\Delta_n^*\to 0$ in $P_{n,\theta}$-probability. Next, let S_r be the (closed) sphere centered at the origin and having radius r, and let M_r be a constant such that

$$|\exp h_m'z-\exp h'z|\leqq M_r\|h_m-h\| , \qquad z\in S_r . \tag{4.14}$$

The fact that $Q(\partial S_r)=0$, the convergence $\mathcal{L}_m\Rightarrow Q$ and an easy elaboration on the definition of weak convergence of probability measures, imply that

$$\int_{S_r} \exp h'zd\mathcal{L}_m\to\int_{S_r} \exp h'zdQ . \tag{4.15}$$

Therefore

$$\begin{aligned}&\left|\int_{S_r} \exp h_m'zd\mathcal{L}_m-\int_{S_r} \exp h'zdQ\right|\\ &\quad\leqq\int_{S_r} |\exp h_m'z-\exp h'z|\,d\mathcal{L}_m+\left|\int_{S_r} \exp h'zd\mathcal{L}_m-\int_{S_r} \exp h'zdQ\right|\\ &\quad\leqq M_r\|h_m-h\|+\left|\int_{S_r} \exp h'zd\mathcal{L}_m-\int_{S_r} \exp h'zdQ\right|\end{aligned}$$

and this converges to 0 by means of (4.14) and (4.15). Thus

$$\int_{S_r} \exp h_m'zd\mathcal{L}_m\to\int_{S_r} \exp h'zdQ . \tag{4.16}$$

Let λ_0 be such that

$$\|h_n\|\leqq\lambda_0 , \qquad \|h\|\leqq\lambda_0 . \tag{4.17}$$

Then

$$\begin{aligned}&\left|\int_{S_r^c} \exp h_m'zd\mathcal{L}_m-\int_{S_r^c} \exp h'zdQ\right|\\ &\quad\leqq\int_{S_r^c} \exp |h_m'z|\,d\mathcal{L}_m+\int_{S_r^c} \exp |h'z|\,dQ\\ &\quad\leqq\int_{S_r^c} \exp \lambda_0\|z\|\,d\mathcal{L}_m+\int_{S_r^c} \exp \lambda_0\|z\|\,dQ .\end{aligned} \tag{4.18}$$

Since $\int \exp \lambda_0\|z\|\,dQ<\infty$ by Lemma 4.1 in Johnson and Roussas [5], we may choose r sufficiently large, so that

$$\int_{S_r^c} \exp \lambda_0 \|z\| dQ < \varepsilon \; . \tag{4.19}$$

In a manner similar to the one used in order to establish (4.15), one shows that

$$\int_{S_r} \exp \lambda_0 \|z\| d\mathcal{L}_m \to \int_{S_r} \exp \lambda_0 \|z\| dQ \; . \tag{4.20}$$

Recalling that $\mathcal{L}_m = \mathcal{L}(\Delta_m^* | P_{m,\theta})$, relation (4.20) herein, along with relations (4.1), (4.5) (with λ replaced by λ_0) and (4.6) in Johnson and Roussas [5], implies that

$$\int_{S_r^c} \exp \lambda_0 \|z\| d\mathcal{L}_m \to \int_{S_r^c} \exp \lambda_0 \|z\| dQ \; .$$

Hence for all sufficiently large m, one has

$$\int_{S_r^c} \exp \lambda_0 \|z\| d\mathcal{L}_m < 2\varepsilon \; . \tag{4.21}$$

By (4.19) and (4.21), relation (4.18) gives

$$\int_{S_r^c} \exp h_m' z d\mathcal{L}_m \to \int_{S_r^c} \exp h' z dQ \; .$$

This result, together with (4.16), implies that

$$\int \exp h_m' z d\mathcal{L}_m \to \int \exp h' z dQ \; . \tag{4.22}$$

Finally, $\exp B_m(h_m) = \int \exp h_m' \Delta_m^* dP_{m,\theta} = \int \exp h_m' z d\mathcal{L}_m \to \int \exp h' z dQ$ by (4.22). Then (4.13) completes the proof of the lemma.

5. A representation theorem of Hájek-Inagaki

Define the class C of sequences of estimates of θ, $\{T_n\}$, as follows

$$C = \{\{T_n\}; \mathcal{L}[n^{1/2}(T_n - \theta_n) | P_{n,\theta_n}] \Rightarrow \mathcal{L}(\theta)\} \; , \tag{5.1}$$

a probability measure, where $\mathcal{L}(\theta)$, in general, depends on $\{T_n\}$ and $\theta_n = \theta + hn^{-1/2}$, so that $\theta_n \in \Theta$ for all sufficiently large n. The main result of the present section is that the probability measure $\mathcal{L}(\theta)$ can be represented as the convolution of two probability measures one of which is normal. More precisely, one has the following

THEOREM 5.1. *Consider the class C defined by* (5.1). *Then one has that $\mathcal{L}(\theta) = \mathcal{L}_1(\theta) * \mathcal{L}_2(\theta)$, where $\mathcal{L}_1(\theta) = N(0, \Gamma^{-1}(\theta))$ and $\mathcal{L}_2(\theta)$ is defined* *by* (5.15) *below.*

PROOF. In all that follows θ is kept fixed. Therefore for the sake of simplicity, we may omit θ from our notation. Thus, for example, we shall write $\mathcal{L}$, $\mathcal{L}_1$, $\mathcal{L}_2$, Γ, Δ_n^* etc. rather than $\mathcal{L}(\theta)$, $\mathcal{L}_1(\theta)$, $\mathcal{L}_2(\theta)$, $\Gamma(\theta)$, $\Delta_n^*(\theta)$ etc.

Set $\mathcal{L}_n^*=\mathcal{L}_n^*(\theta)=\mathcal{L}[(n^{1/2}(T_n-\theta), \Delta_n^*)\,|\,P_{n,\theta}]$. Then, by the Weak compactness theorem for probability measures, there exists $\{m\}\subseteq\{n\}$ such that $\mathcal{L}_m^*\Rightarrow\mathcal{L}^*$, a measure. Then the marginal measures $\mathcal{L}[m^{1/2}(T_m-\theta)\,|\,P_{m,\theta}]$ and $\mathcal{L}(\Delta_m^*\,|\,P_{m,\theta})$ of $\mathcal{L}[(m^{1/2}(T_m-\theta), \Delta_m^*)\,|\,P_{m,\theta}]$ converge (weakly) to the corresponding marginal measures of $\mathcal{L}^*$. These latter marginal measures are then probability measures since both $\mathcal{L}[m^{1/2}(T_m-\theta)\,|\,P_{m,\theta}]$ and $\mathcal{L}(\Delta_m^*\,|\,P_{m,\theta})$ converge to probability measures. It follows that $\mathcal{L}^*$ itself is a probability measure. Setting (T, Δ) for the identity mapping on $R^k\times R^k$, we have then that $\mathcal{L}[(T, \Delta)\,|\,\mathcal{L}^*]=\mathcal{L}^*$, so that

$$\mathcal{L}[(m^{1/2}(T_m-\theta), \Delta_m^*)\,|\,P_{m,\theta}]\Rightarrow\mathcal{L}[(T, \Delta)\,|\,\mathcal{L}^*]\ . \tag{5.2}$$

It is shown in Lemma 5.1 below that

$$\begin{aligned}\mathcal{E}_\theta \exp\,&[iu'm^{1/2}(T_m-\theta)+\Lambda_m]\\ &-\mathcal{E}_\theta \exp\left[iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*-\frac{1}{2}h'\Gamma h\right]\to 0\end{aligned} \tag{5.3}$$

and

$$\begin{aligned}\mathcal{E}_\theta \exp&\left[iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*-\frac{1}{2}h'\Gamma h\right]\to\\ &\mathcal{E}_{\mathcal{L}^*} \exp\left(iu'T+h'\Delta-\frac{1}{2}h'\Gamma h\right)\ .\end{aligned} \tag{5.4}$$

Therefore by setting

$$\phi(u, h)=\mathcal{E}_{\mathcal{L}^*} \exp\,(iu'T+ih'\Delta) \tag{5.5}$$

and replacing h by zero, we obtain by means of (5.4) that

$$\mathcal{E}_\theta \exp\,[iu'm^{1/2}(T_m-\theta)]\to\mathcal{E}_{\mathcal{L}^*} \exp iu'T=\mathcal{E}_{\mathcal{L}} \exp iu'T=\phi(u, 0)\ . \tag{5.6}$$

Also from (5.1) and (5.5), it follows that

$$\mathcal{E}_{\theta_m} \exp\,[iu'm^{1/2}(T_m-\theta_m)]\to\phi(u, 0)\ . \tag{5.7}$$

Next,

$$\begin{aligned}\mathcal{E}_{\theta_m} \exp\,[iu'm^{1/2}(T_m-\theta_m)]&=\mathcal{E}_\theta \exp\,[iu'm^{1/2}(T_m-\theta_m)+\Lambda_m]\\ &=\mathcal{E}_\theta \exp\,[iu'm^{1/2}(T_m-\theta)-iu'h+\Lambda_m]\\ &=\exp\,(-iu'h)\mathcal{E}_\theta \exp\,[iu'm^{1/2}(T_m-\theta)+\Lambda_m]\end{aligned}$$

and this last expression converges to $\exp(-iu'h)\mathcal{E}_{\mathcal{L}^*}\exp(iu'T+h'\Delta-(1/2)h'\Gamma h)$ by means of (5.3) and (5.4). That is,

$$\mathcal{E}_{\theta_m}\exp[iu'm^{1/2}(T_m-\theta_m)]\to \exp(iu'h)\mathcal{E}_{\mathcal{L}^*}\exp\left(iu'T+h'\Delta-\frac{1}{2}h'\Gamma h\right). \tag{5.8}$$

From (5.7) and (5.8), we obtain

$$\phi(u, 0)=\exp(-iu'h)\mathcal{E}_{\mathcal{L}^*}\exp\left(iu'T+h'\Delta-\frac{1}{2}h'\Gamma h\right). \tag{5.9}$$

It is shown in Lemma 5.2 below that in (5.9) we may replace h by ih. By doing so, we obtain

$$\phi(u, 0)=\exp u'h\mathcal{E}_{\mathcal{L}^*}\exp\left(iu'T+ih'\Delta+\frac{1}{2}h'\Gamma h\right). \tag{5.10}$$

By means of (5.5), relation (5.10) may also be written as follows

$$\phi(u, 0)=(\exp u'h)\phi(u, h)\exp\frac{1}{2}h'\Gamma h,$$

so that

$$\phi(u, h)=\phi(u, 0)\exp(-u'h)\exp\left(-\frac{1}{2}h'\Gamma h\right). \tag{5.11}$$

Next, it is easily seen that

$$-u'h-\frac{1}{2}h'\Gamma h=-\frac{1}{2}(h'+u'\Gamma^{-1})\Gamma(h+\Gamma^{-1}u)+\frac{1}{2}u'\Gamma^{-1}u,$$

so that (5.11) becomes

$$\phi(u, h)=\exp\left[-\frac{1}{2}(h'+u'\Gamma^{-1})\Gamma(h+\Gamma^{-1}u)\right]\phi(u, 0)\exp\frac{1}{2}u'\Gamma^{-1}u. \tag{5.12}$$

Setting successively $h=0$ and $h=-\Gamma^{-1}u$ in (5.12), we obtain

$$\phi(u, 0)=\exp\left(-\frac{1}{2}u'\Gamma^{-1}u\right)\phi(u, 0)\exp\frac{1}{2}u'\Gamma^{-1}u \tag{5.13}$$

and

$$\phi(u, -\Gamma^{-1}u)=\phi(u, 0)\exp\frac{1}{2}u'\Gamma^{-1}u, \tag{5.14}$$

respectively.

From (5.5) and (5.14), it follows that $\phi(u, 0)\exp 1/2\cdot u'\Gamma^{-1}u$ is a characteristic function (under $\mathcal{L}^*$), namely, the characteristic function of the random vector $T-\Gamma^{-1}\Delta$. Define $\mathcal{L}_2$ by

$$\mathcal{L}_2=\mathcal{L}(T-\Gamma^{-1}\Delta\,|\,\mathcal{L}^*)\ . \tag{5.15}$$

Also $\exp(-(1/2)u'\Gamma^{-1}u)$ is a characteristic function (under $\mathcal{L}^*$), namely, the characteristic function of the random vector $\Gamma^{-1}\Delta$ which is distributed as $N(0,\Gamma^{-1})$. Set $\mathcal{L}_1=N(0,\Gamma^{-1})$. Then from relation (5.13) and the Composition theorem (see, e.g., Loève [8], p. 193), it follows that $\mathcal{L}=\mathcal{L}_1*\mathcal{L}_2$, as was to be seen.

We now proceed to establish the lemmas used in the proof of the theorem.

LEMMA 5.1 *With the notation employed in the theorem, one has that*

(i) $\mathcal{E}_\theta \exp[iu'm^{1/2}(T_m-\theta)+\Lambda_m]$

$$-\mathcal{E}_\theta \exp\left[iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*-\frac{1}{2}h'\Gamma h\right]\rightarrow 0$$

and

(ii) $\mathcal{E}_\theta \exp\left[iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*-\frac{1}{2}h'\Gamma h\right]\rightarrow$

$$\mathcal{E}_{\mathcal{L}^*}\exp\left(iu'T+h'\Delta-\frac{1}{2}h'\Gamma h\right)\ .$$

PROOF. (i) For simplicity, set $U_n=\exp\Lambda_n$, $V_n=\exp(h'\Delta_n^*-(1/2)h'\Gamma h)$ and $Q^*=N(-(1/2)h'\Gamma h,\ h'\Gamma h)$. Then the convergence $\mathcal{L}(\Lambda_n|P_{n,\theta})\Rightarrow\mathcal{L}(\Delta|Q^*)$, implies that $\mathcal{L}(U_n|P_{n,\theta})\Rightarrow\mathcal{L}(\exp\Delta|Q^*)$, whereas $\mathcal{E}_\theta U_n=\mathcal{E}_{Q^*}\exp\Delta=1$. Also the convergence $\mathcal{L}(\Delta_n^*|P_{n,\theta})\Rightarrow N(0,\Gamma)$ implies that $\mathcal{L}(h'\Delta_n^*-(1/2)h'\Gamma h|P_{n,\theta})\Rightarrow\mathcal{L}(\Delta|Q^*)$, so that $\mathcal{L}(V_n|P_{n,\theta})\Rightarrow\mathcal{L}(\exp\Delta|Q^*)$. Furthermore, $\mathcal{E}_\theta\exp h'\Delta_n^*\rightarrow\mathcal{E}_Q\exp h'\Delta=\exp(1/2)h'\Gamma h$ by (4.3) and (4.13), where, we recall that, $Q=N(0,\Gamma)$. Thus $\mathcal{E}_\theta V_n\rightarrow 1$, so that Lemma 3.1 applies and gives that U_n, V_n, $n\geqq 1$, are uniformly integrable. Next, $U_n-V_n\rightarrow 0$ in $P_{n,\theta}$-probability by arguing as in (4.9) and using the fact that $\Lambda_n-h'\Delta_n^*\rightarrow-(1/2)h'\Gamma h$ in $P_{n,\theta}$-probability. Then, by Lemma 3.2, it follows that $|U_n-V_n|$, $n\geqq 1$, are uniformly integrable, and Corollary 3.1 gives that $\mathcal{E}_\theta|U_n-V_n|\rightarrow 0$. The proof of (i) is then completed by observing that the left-hand side of (i) is bounded in absolute value by $\mathcal{E}_\theta|U_n-V_n|$.

(ii) Clearly, it suffices to show that

$$\mathcal{E}_\theta[\exp iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*]\rightarrow\mathcal{E}_{\mathcal{L}^*}\exp(iu'T+h'\Delta)\ . \tag{5.16}$$

Set $A_m=(|h'\Delta_m^*|>c)$. Then one has that

$$|\mathcal{E}_\theta[\exp iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*]-\mathcal{E}_{\mathcal{L}^*}\exp(iu'T+h'\Delta)| \tag{5.17}$$

$$\leqq\int_{A_m}\exp h'\Delta_m^* dP_{m,\theta}+\int_{A_m}\exp h'\Delta d\mathcal{L}^*$$

$$+\left|\int_{\Delta_m^c} \exp\left[iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*\right]dP_{m,\theta} - \int_{\Delta_m^c} \exp(iu'T+h'\Delta)d\mathcal{L}^*\right| .$$

As was seen in the proof of (i), $\mathcal{E}_\theta \exp h'\Delta_m^* \to \exp(1/2)h'\Gamma h$. Also $\mathcal{L}(h'\Delta_m^* \mid P_{n,\theta}) \Rightarrow \mathcal{L}(h'\Delta \mid Q)$ implies that $\mathcal{L}(\exp h'\Delta_m^* \mid P_\theta) \Rightarrow \mathcal{L}(\exp h'\Delta \mid Q)$. Therefore Lemma 3.1 gives that $\exp h'\Delta_m^*$, for all m, are uniformly integrable. Thus one may choose c sufficiently large, so that

$$\int_{\Delta_m} \exp h'\Delta_m^* dP_{m,\theta} < \varepsilon \quad \text{and} \quad \int_{\Delta_m} \exp h'\Delta d\mathcal{L}^* = \int_{\Delta_m} \exp h'\Delta dQ < \varepsilon . \tag{5.18}$$

Furthermore,

$$\begin{aligned} &\int_{\Delta_m^c} \exp\left[iu'm^{1/2}(T_m-\theta)+h'\Delta_m^*\right]dP_{m,\theta} \\ &\quad = \int_{(|h'\Delta|\leq c)} \exp(iu'T+h'\Delta)d\mathcal{L}_m^* \to \int_{(|h'\Delta|\leq c)} \exp(iu'T+h'\Delta)d\mathcal{L}^* \end{aligned} \tag{5.19}$$

by the fact that $\mathcal{L}_m^* \Rightarrow \mathcal{L}^*$ and $\mathcal{L}^*(R^k \times B) = Q(B) = 0$, where $B = \{\Delta \in R^k;\ |h'\Delta| = c\}$. The relations (5.17)–(5.19) complete the proof of (5.16) and hence that of (ii).

LEMMA 5.2. *With the notation employed in the theorem, consider the expectation $\mathcal{E}_{\mathcal{L}^*} \exp(iu'T+h'\Delta)$ as a function of h, call it $g(h)$, where $h=(h_1,\cdots,h_k)'$, $h_j \in R$, $j=1,\cdots,k$. Then, for $j=1,\cdots,k$, $g(h)$ is analytic in the jth coordinate h_j when the first $j-1$ coordinates h_r, $r=1,\cdots,j-1$ are any complex numbers and the last $k-j$ coordinates h_r, $r=j+1,\cdots,k$ are any real numbers.*

PROOF. By setting $\underline{h}=(h_2,\cdots,h_k)'$, $\Delta=(\Delta_1,\cdots,\Delta_k)'$ and $\underline{\Delta}=(\Delta_2,\cdots,\Delta_k)'$, one has that

$$\begin{aligned} \mathcal{E}_{\mathcal{L}^*} \exp(iu'T+h'\Delta) &= \mathcal{E}_{\mathcal{L}^*} \exp\left[(iu'T+\underline{h}'\underline{\Delta})+h_1\Delta_1\right] \\ &= \mathcal{E}_{\mathcal{L}^*}\left[\exp(iu'T+\underline{h}'\underline{\Delta}) \sum_{j=0}^{\infty} \frac{(h_1\Delta_1)^j}{j!}\right] \\ &= \mathcal{E}_{\mathcal{L}^*} \sum_{j=0}^{\infty} \left[\exp(iu'T+\underline{h}'\underline{\Delta})\frac{\Delta_1^j}{j!}\right]h_1^j . \end{aligned}$$

Now at this point we observe that for $n \geqq 0$, one has

$$\begin{aligned} &\left|\sum_{j=0}^{n}\left[\exp(iu'T+\underline{h}'\underline{\Delta})\frac{\Delta_1^j}{j!}\right]h_1^j\right| \\ &\quad = \exp \underline{h}'\underline{\Delta}\left|\sum_{j=0}^{n}\frac{\Delta_1^j}{j!}h_1^j\right| \leqq \exp \underline{h}'\underline{\Delta}\sum_{j=0}^{n}\frac{|\Delta_1|^j}{j!}|h_1|^j \\ &\quad \leqq \exp(\underline{h}'\underline{\Delta}+|h_1\Delta_1|) \end{aligned} \tag{5.20}$$

and this last expression is, clearly, Q-integrable and hence $\mathcal{L}^*$-integrable. Then by the Dominated convergence theorem, we get

$$\mathcal{E}_{\mathcal{L}^*} \exp(iu'T+h'\Delta)=\sum_{j=0}^{\infty}\left\{\mathcal{E}_{\mathcal{L}^*}\left[\exp(iu'T+\underline{h}'\underline{\Delta})\frac{\Delta_1^j}{j!}\right]\right\}h_1^j$$

which shows that $g(h)=g(h_1,\cdots,h_k)$ is analytic in h_1 when the other coordinates are kept fixed. Thus, h_1 may be replaced by a complex variable. Making this replacement and working as above, we have that the left-hand side in (5.20) is $\sum_{j=0}^{n}\left[\exp(iu'T+ih_{12}\Delta_1+h_{11}\Delta_1+\underline{\underline{h}}'\underline{\underline{\Delta}})\frac{\Delta_2^j}{j!}\right]h_2^j$, where $h_1=h_{11}+ih_{12}$, $\underline{\underline{h}}=(h_3,\cdots,h_k)'$ and $\underline{\underline{\Delta}}=(\Delta_3,\cdots,\Delta_k)'$, whereas the last bound on the right-hand side of the same relationship is equal to exp $(h_{11}\Delta_1+\underline{\underline{h}}'\underline{\underline{\Delta}}+|h_2\Delta_2|)$ which is Q-integrable and hence $\mathcal{L}^*$-integrable. Therefore h_2 may be replaced by a complex variable. In a similar fashion each one of the remaining coordinates may be replaced by complex variables and this completes the proof of the lemma.

University of Wisconsin

References

[1] Bickel, P. Personal communication.

[2] Doob, J. L. (1953). *Stochastic Processes*, Wiley, New York.

[3] Hájek, J. (1970). A characterization of limiting distributions of regular estimates, *Z. Wahrscheinlichkeitstheorie and Verw. Gebiete*, **14**, 323-330.

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

[5] Johnson, R. A. and Roussas, G. G. (1970). Asymptotically optimal tests in Markov processes, *Ann. Math. Statist.*, **41**, 918-938.

[6] Kaufman, S. (1966). Asymptotic efficiency of the maximum likelihood estimator, *Ann. Inst. Statist. Math.*, **18**, 155-178.

[7] LeCam, L. (1972). Limits of experiments and a theorem of J. Hájek, *Proc. Sixth Berkeley Symp. Math. Statist. Prob.*, **1**, 245-261.

[8] Loève, M. (1963). *Probability Theory*, 3rd ed. Van Nostrand, Princeton.

[9] Roussas, G. G. (1965). Asymptotic inference in Markov processes, *Ann. Math. Statist.*, **36**, 978-992.

[10] Roussas, G. G. (1968). Some applications of the asymptotic distribution of likelihood functions to the asymptotic efficiency of estimates, *Z. Wahrscheinlichkeitstheorie and Verw. Gebiete*, **10**, 252-260.

[11] Schmetterer, L. (1966). On asymptotic efficiency of estimates, *Research Papers in Statistics*, F. N. David (editor), Wiley, New York.

[12] Wolfowitz, J. (1965). Asymptotic efficiency of the maximum likelihood estimator, *Theor. Probab. Appl.*, **10**, 247-260.