The Annals of Mathematical Statistics
1971, Vol. 42, No. 4, 1241–1253

THE BERNSTEIN–VON MISES THEOREM FOR MARKOV PROCESSES[1]

BY J. BORWANKER, G. KALLIANPUR AND B. L. S. PRAKASA RAO

Indian Institute of Technology, Kanpur; University of Minnesota

0. Introduction and summary. Since the appearance of P. Billingsley's monograph [2] on the large sample inference in Markov processes in which the weak consistency and asymptotic normality of the maximum likelihood estimate was investigated, there has been considerable interest in the further development of the theory along other directions. Billingsley's work was mainly concerned with extending the results of H. Cramér ([4] page 500). Among more recent developments one might mention the proof of the almost sure consistency of the maximum likelihood estimator following the ideas of A. Wald by G. Roussas [7], and the results on asymptotic Bayes estimates obtained by Lorraine Schwartz [9].

In the present paper we extend to Markov processes one of the fundamental results in the asymptotic theory of inference, viz., the approach of the posterior density (in a sense to be made precise later) to the normal. When the observed chance variables are independent and identically distributed, this result was obtained by L. LeCam in [5] (page 309). The same author offers another derivation of this result in [6]. Special cases of the theorem were first given by S. Bernstein and R. von Mises (for reference see [5]).

The regularity conditions satisfied by the transition probability density of the Markov process are given in Section 1. We prove in Theorem 2.4 of Section 2 those properties of the maximum likelihood estimator that are needed for the proof of the main result of the paper given in Section 3 (Theorem 3.1). Theorem 3.1 is stated in a form which is general enough to include the Bernstein–von Mises theorem as well as the somewhat sharper versions that are available when it is known that the prior probability distribution has a finite absolute moment of order m. Theorem 3.2 deduces these results as a consequence of Theorem 3.1. The latter result also enables us to prove a theorem on the asymptotic behavior of regular Bayes estimates. This is done in Theorem 4.1 of Section 4.

1. Notations and assumptions. Let $X_0, X_1, X_2, \cdots$ be random variables forming a strictly stationary ergodic Markov process and taking values in a measurable space $(S, \mathscr{B}_s)$. The stationary initial probability distribution and the transition probability function of the process will be denoted by $P_\theta(A)$ and $P_\theta(y; A)$ ($y \in S$ and $A \in \mathscr{B}_s$) respectively, where θ is an unknown parameter belonging to a set Θ, assumed here to be an open set of the real line. We suppose that there exists a σ-finite

---

Received August 28, 1969; revised January 4, 1971.

[1] Prepared with the support of a grant from the U.S. Air Force Office of Aerospace Research, Grant No. AF-AFOSR-885-65 and a grant from the U.S. Navy, Grant No. N00014-67-C0307, and partial support from the National Science Foundation Grant GP-7490.

measure μ on $(S, \mathscr{B}_s)$ such that $P_\theta(A)$ and $P_\theta(y; A)$ are both absolutely continuous with respect to μ with densities $f(z; \theta)$ and $f(y, z; \theta)$ respectively. For $\theta \in \Theta$, let P_θ denote the measure on the product measurable space determined by the initial probability distribution $P_\theta(\cdot)$ and the transition probability function $P_\theta(\cdot j \cdot)$. The log likelihood function of the process, given the observations $x_0, x_1, \cdots, x_n$ is defined to be the function,

$$\log L_n(\theta; \mathbf{x}) = \log f(x_0; \theta) + \sum_{i=0}^{n-1} f(x_i, x_{i+1}; \theta)$$

where $\mathbf{x} = (x_0, x_1, \cdots, x_n)$. This definition is meaningful for almost all $\mathbf{x}(P_\theta)$. Since we are concerned only with the large sample theory, we may neglect $\log f(x_0; \theta)$ in the above expression (see [2] page 4). We shall write $L_n(\theta)$ for $L_n(\theta; \mathbf{x})$ for convenience.

ASSUMPTION 1.1. The parameter space Θ is an open interval of the real line. Λ is a prior probability measure on $(\Theta, \mathscr{F})$ where $\mathscr{F}$ is the σ-field of Borel subsets of Θ. Λ has a density λ with respect to the Lebesgue measure.

ASSUMPTION 1.2. Let $h(x_0, x_1; \theta) = \log f(x_1, x_1; \theta)$. $\partial h(x_0. x_1; \theta)/\partial\theta$ and $\partial^2 h(x_0, x_1; \theta)/\partial\theta^2$ exist and are continuous in θ for almost all pairs (x_0, x_1) $(\mu \times \mu)$.

ASSUMPTION 1.3. For each $\theta \in \Theta$, there corresponds $\eta(\theta) > 0$ such that

$$E_\theta[\sup\{|\partial^2 h(X_0, X_1; \theta')/\partial\theta^2| : |\theta - \theta'| < \eta(\theta), \theta' \in \Theta\}]$$

is finite where E_θ denotes the expectation when θ is the true parameter.

ASSUMPTION 1.4. For each $\theta \in \Theta$ and any $\varepsilon > 0$,

$$-\infty < E_\theta[\sup\{h(X_0, X_1; \theta') - h(X_0, X_1; \theta) : |\theta' - \theta| \geqq \varepsilon, \theta' \in \Theta\}] < 0.$$

ASSUMPTION 1.5. Let

$$i(\theta) \equiv -E_\theta[\partial^2 h(X_0, X_1; \theta)/\partial\theta^2]$$

for $\theta \in \Theta$. $i(\theta)$ is finite for all $\theta \in \Theta$ by Assumption 1.2. We shall assume that $i(\theta)$ is continuous and nonzero for all $\theta \in \Theta$.

ASSUMPTION 1.6. Let θ_0 denote the true parameter and $P_0 = P_{\theta_0}$. Let K be a nonnegative measurable function satisfying the following conditions:

There exists a number ε, $0 < \varepsilon < i_0$ $(i_0 = i(\theta_0))$ such that

$$M(\theta_0) = (i_0/(2\pi))^{\frac{1}{2}} \int_{-\infty}^{\infty} K(t) \exp\{-[i_0 - \varepsilon]t^2/2\}\, dt$$

is finite.

ASSUMPTION 1.7. For every $h > 0$ and every $\delta > 0$

$$e^{-\delta n} \int_{|t| > h} K(n^{\frac{1}{2}} t) \lambda(\hat{\theta}_n + t)\, dt \to 0$$

a.s. (P_{θ_0}) as $n \to \infty$, where $\hat{\theta}_n$ denotes a maximum likelihood estimator (MLE) as defined in Section 2.

ASSUMPTION 1.8. The prior density λ is continuous and positive in an open neighborhood of the true parameter θ_0.

2. Some preliminary results. Before we state our main theorem we shall prove in Theorem 2.4 the properties of MLE's which will prove useful to us in the next section. Theorem 2.4, in turn, is based on results taken from Billingsley ([2], [3]) which were obtained by him in his own investigation of maximum likelihood estimation in Markov chains. These results of Billingsley's are stated here first for convenience.

LEMMA 2.1 [2]. *Suppose $\{U_n, n \geqq 1\}$ is a sequence of random variables such that $U_n \to_{\mathscr{L}} v$ where v is a probability measure on the real line. Suppose further that $\{V_n, n \geqq 1\}$ is another sequence of random variables such that*

$$|U_n - V_n| \leqq \varepsilon_n |V_n|$$

where ε_n tends to zero in probability as n tends to infinity. Then $V_n \to_{\mathscr{L}} v$. (Here $V_n \to_{\mathscr{L}} v$ denotes that the distribution of U_n converges weakly to v.)

THEOREM 2.1 [2]. *Suppose that for each $\theta \in \Theta$, the stationary distribution exists and is unique and has the property that for each y in the state space S, $P_\theta(y; \cdot)$ is absolutely continuous with respect to the stationary distribution $P_\theta(\cdot)$. Then, for any $\theta \in \Theta$, the Markov process $\{X_n, n \geqq 0\}$ is metrically transitive if the initial distribution is the stationary one. No matter what the initial distribution is, if $g(x_0, x_1)$ is a Borel measurable function of (x_0, x_1) such that $E_\theta\{|g(X_0, X_1)|\} < \infty$, then*

$$\lim_{n\to\infty} n^{-1} \textstyle\sum_{k=0}^{n-1} g(X_k, X_{k+1}) = E_\theta[g(X_0, X_1)]$$

with probability one.

THEOREM 2.2 [3]. *Let $\{\xi_n, n \geqq 1\}$ be a strictly stationary ergodic process such that $E[\xi_1{}^2] < \infty$ and*

$$E[\xi_n \mid \xi_1, \cdots, \xi_{n-1}] = 0; \qquad E[\xi_1] = 0.$$

Then the distribution of $n^{-\frac{1}{2}} \sum_{k=1}^n \xi_k$ approaches the normal distribution with mean zero and variance $E[\xi_1{}^2]$. The next result follows immediately from Theorem 2.2.

THEOREM 2.3. *Let assumptions* 1.1–1.5 *be satisfied. Then for any $\theta \in \Theta$ and for any initial distribution,*

$$n^{-\frac{1}{2}} \textstyle\sum_{k=1}^n \partial \log f(X_k, X_{k+1}; \theta)/\partial\theta \to_{\mathscr{L}} N(0, i(\theta)).$$

Let $x_0, x_1, \cdots, x_n$ be a Markov sequence of observations. Suppose there exists a compact-neighborhood $U(\theta_0)$ of θ_0 and an estimator $\hat\theta_n = \hat\theta_n(x_0, x_1, \cdots, x_n)$ such that

$$L_n(\hat\theta_n) = \max\{L_n(\theta) : \theta \in U(\theta_0)\}$$

then $\hat\theta_n$ is called a Maximum Likelihood Estimator (MLE) of θ. By the compactness of $U(\theta_0)$ and the continuity of $f(x_0, x_1; \theta)$ the maximum is attained and we shall assume that $\hat{\theta}_n$ is measurable. We shall now state and prove a theorem which gives the strong consistency and asymptotic normality of $\hat{\theta}_n$. The first two conclusions of the theorem are similar to the theorem given by Bickel and Yahav [1] and LeCam [6] for the case of independent observations. Since the assumptions we have made are somewhat different from those usually imposed in proving the asymptotic normality of $\hat{\theta}_n$, we give a complete proof here.

THEOREM 2.4. *Under the assumptions* 1.1–1.5, *there exists a compact neighborhood* $U(\theta_0)$ *of* θ_0 *such that*

$$\hat{\theta}_n \to \theta_0 \quad \text{a.s.}; \tag{2.1}$$

here

$$\partial \log L(\theta_n)/\partial\theta = 0 \qquad \text{or } n \geqq N, \tag{2.2}$$

where N depends on the sample of observations;

$$n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0) \to_{\mathscr{L}} N(0, i_0^{-1}), \tag{2.3}$$

where $N(0, \sigma^2)$ *denotes the normal distribution with mean* 0 *and variance* σ^2.

PROOF. Let $U(\theta_0) = \{\theta\colon |\theta-\theta_0| \geqq 2^{-1}\eta(\theta_0)\}$, where $\eta(\theta_0)$ is given by Assumption 1.3. For any $\theta \in U(\theta_0)$,

$$\begin{aligned} h(x_0, x_1; \theta) - h(x_0, x_1, \theta_0), \\ = (\theta-\theta_0)\,\partial h(x_0, x_1; \theta_0)/\partial\theta \\ + (\theta-\theta_0)^2 \textstyle\int_0^1 (1-\omega)[\partial^2 h(x_0, x_1; \theta')/\partial\theta^2]_{\theta'=\theta_0+\omega(\theta-\theta_0)}\, d\omega. \end{aligned} \tag{2.4}$$

Fix any $\theta \neq \theta_0$ in $U(\theta_0)$. Consider

$$g(x_0, x_1; \delta) \equiv \sup\{h(x_0, x_1, \tau) - h(x_0, x_1; \theta_0)\colon |\tau-\theta| < \delta, \tau \in U(\theta_0)\}. \tag{2.5}$$

Since $h(x_0, x_1; \theta)$ is a continuous function of θ almost surely, $g(x_0, x_1; \delta)$ tends to $h(x_0, x_1; \theta) - h(x_0, x_1; \theta_0)$ almost surely as δ approaches zero. We shall now show that $g(x_0, x_1; \delta)$ is uniformly bounded in δ by an integrable function. Choose $\delta < 2^{-1}\eta(\theta_0)$. Now from (2.4), it follows that

$$\begin{aligned} |g(x_0, x_1; \delta)| &\leqq |\partial h(x_0, x_1; \theta_0)/\partial\theta| \cdot \sup\{|\tau-\theta_0|\colon |\tau-\theta| < \delta, \tau \in U(\theta_0)\} \\ &\quad + \sup\{|\tau-\theta_0|^2\colon |\tau-\theta| < \delta, \tau \in U(\theta_0)\} \\ &\quad \cdot \sup\{\textstyle\int_0^1 |\partial^2 h(x_0, x_1; \theta')/\partial\theta^2]_{\theta'=\theta_0+\omega(\tau-\theta_0)}|\, d\omega : |\tau-\theta| < \delta, \\ &\qquad\qquad \tau \in U(\theta_0)\} \\ &\leqq |\partial h(x_0, x_1; \theta_0)/\partial\theta|\eta(\theta_0) \\ &\quad + \eta^2(\theta_0) \sup\{|\partial^2 h(x_0, x_1; \theta')/\partial\theta^2|\colon |\theta'-\theta_0| \\ &\quad < \eta(\theta_0)\} \end{aligned} \tag{2.6}$$

since

$$\begin{aligned} |\theta'-\theta_0| < |\theta_0-\tau| &< |\theta_0-\theta| + |\tau-\theta| \\ &< \tfrac{1}{2}\eta(\theta_0) + \delta < \eta(\theta_0). \end{aligned}$$

The second term on the right-hand side of (2.6) has finite expectation by Assumption 1.3. The first term is integrable with finite expectation by Assumptions 1.2 and 1.3. Hence $g(X_1, X_2; \delta)$ is bounded by an integrable function. It now follows by the dominated convergence theorem that for any $\theta \in U(\theta_0)$ different from θ_0,

$$\lim_{\delta\to 0} E_{\theta_0}[g(X_0, X_1; \delta)] = E_{\theta_0}[h(X_0, X_1; \theta) - h(X_0, X_1; \theta_0)] \tag{2.7}$$

which is less than zero by Assumption 1.4 since $|\theta - \theta_0| > 0$. Let us now consider $U(\theta_0)$ as our new parameter space. $U(\theta_0)$ is compact and any $\theta \in U(\theta_0)$ different from θ_0 satisfies (2.7). Hence by using arguments similar to Wald's on strong consistency of MLE, we obtain that $\hat{\theta}_n \to \theta_0$ a.s. which proves (2.1). Since θ_0 is an interior point of Θ and since $\hat{\theta}_n \to \theta_0$ a.s., $\hat{\theta}_n$ will also be an interior point of Θ for large n and hence by Assumption 1.2, $\partial \log L_n(\theta)/\partial\theta$ at $\theta = \hat{\theta}_n$ exists and is zero for sufficiently large $n > N$, N possibly depending on the sample. This proves (2.2).

Since $\hat{\theta}_n$ is a solution of the likelihood equation for $n > N$, we get that

$$\begin{aligned} &\partial \log L_n(\hat{\theta}_n)/\partial\theta \\ &\quad = \partial \log L_n(\theta_0)/\partial\theta + (\hat{\theta}_n - \theta_0)\, \partial^2 \log L_n(\theta_n')/\partial\theta^2 \\ &\quad = 0 \end{aligned} \tag{2.8}$$

where $|\theta_n' - \theta_0| \leqq |\hat{\theta}_n - \theta_0|$ a.s. Rewriting (2.8), it is easy to see that

$$Y_n + W_n Z_n + \varepsilon_n Z_n = 0 \tag{2.9}$$

where

$$Y_n = n^{-\frac{1}{2}} \partial \log L_n(\theta_0)/\partial\theta, \tag{2.10}$$

$$Z_n = n^{\frac{1}{2}}(\hat{\theta}_n - \theta_0), \tag{2.11}$$

$$W_n = n^{-1}\, \partial^2 \log L_n(\theta_0)/\partial\theta^2, \tag{2.12}$$

and

$$\varepsilon_n = n^{-1}\{\partial^2 \log L_n(\theta')/\partial\theta^2 - \partial^2 \log L_n(\theta_0)/\partial\theta^2\}. \tag{2.13}$$

It is clear from (2.9) that

$$|Y_n + W_n Z_n| \leqq |\varepsilon_n|\,|Z_n|. \tag{2.14}$$

Since $E_{\theta_0}[\partial \log f(X_0, X_1; \theta_0)/\partial\theta]^2 < \infty$ by Assumptions 1.2–1.5 it follows from Theorem 2.3 that

$$Y_n \to_{\mathscr{L}} N(0, i_0). \tag{2.15}$$

Suppose it is shown that $\varepsilon_n \to 0$ in probability as n approaches infinity. Then it follows that

$$-W_n Z_n \to_{\mathscr{L}} N(0, i_0) \tag{2.16}$$

by (2.14), (2.15) in view of Lemma 2.1. But

$$W_n \to -i_0 \text{ a.s.} \tag{2.17}$$

by Theorem 2.1 since $E_{\theta_0}[|\partial^2 \log f(X_0, X_1, \theta_0)/\partial\theta^2|] < \infty$. Combining (2.16) and (2.17), we obtain that

$$Z_n \rightarrow_{\mathscr{L}} N(0, 1/i_0) \tag{2.18}$$

by Slutsky's Theorem (Cramér [4]). This proves (2.3).

We shall now show that $\varepsilon_n \to 0$ a.s. as n approaches infinity. Choose an $\varepsilon > 0$. It is easy to see from the continuity of $\partial^2 h(x_0, x_1; \theta)/\partial\theta^2$, the lower semi-continuity of

$$\sup\{|\partial^2 h(x_0, x_1; \theta)/\partial\theta^2 - \partial^2 h(x_0, x_1; \theta_0)/\partial\theta^2| : |\theta - \theta_0| < \delta\}$$

and from Assumption 1.3, that for sufficiently small $\delta > 0$ and less than $\eta(\theta_0)$,

$$E_{\theta_0}[\sup|\partial^2 h(X_0, X_1; \theta)/\partial\theta^2 - \partial^2 h(X_0, X_1; \theta_0)/\partial\theta^2| : |\theta - \theta_0| < \delta] \tag{2.19}$$

is less than $\varepsilon/2$. Let us choose such a δ. Since $\hat{\theta}_n \to \theta_0$ a.s. there exists an integer $N_1 > 0$ such that, for $n > N_1$, $|\hat{\theta}_n - \theta_0| < \delta$, N_1 possibly depending on the sample. Now for $n > N_1$,

$$\begin{aligned}
|\varepsilon_n| &= n^{-1}|\partial^2 \log L_n(\theta')/\partial\theta^2 - \partial^2 \log L_n(\theta_0)/\partial\theta^2| \\
&\leqq n^{-1}[\textstyle\sum_{k=0}^{n-1} |\partial^2 h(x_k, x_{k+1}; \theta')/\partial\theta^2 - \partial^2 h(x_k, x_{k+1}; \theta_0)/\partial\theta^2|] \\
&\leqq n^{-1}\textstyle\sum_{k=0}^{n-1} \sup\{|\partial^2 h(x_k, x_{k+1}; \theta)/\partial\theta^2 - \partial^2 h(x_k, x_{k+1}; \theta_0)/\partial\theta^2| \\
&\qquad\qquad : |\theta - \theta_0| < |\hat{\theta}_n - \theta_0|\} \\
&\leqq n^{-1}\textstyle\sum_{k=0}^{n-1} \sup\{|\partial^2 h(x_k, x_{k+1}; \theta)/\partial\theta^2 - \partial^2 h(x_k, x_{k+1}; \theta_0)/\partial\theta^2| \\
&\qquad\qquad : |\theta - \theta_0| < \delta\}.
\end{aligned} \tag{2.20}$$

Since

$$E_{\theta_0}[\sup|\partial^2 h(X_0, X_1; \theta)/\partial\theta^2 - \partial^2 h(X_0, X_1; \theta_0)/\partial\theta^2| : |\theta - \theta_0| < \delta]$$

has finite expectation, it follows by Theorem 2.1 that

$$n^{-1}\textstyle\sum_{k=0}^{n-1} \sup\{|\partial^2 h(X_k, X_{k+1}; \theta)/\partial\theta^2 - \partial^2 h(X_k, X_{k+1}; \theta_0)/\partial\theta^2| : |\theta - \theta_0| < \delta\}$$

converges almost surely to

$$E_{\theta_0}[\sup\{|\partial^2 h(X_0, X_1; \theta)/\partial\theta^2 - \partial^2 h(X_0, X_1; \theta_0)/\partial\theta^2| : |\theta - \theta_0| < \delta\}].$$

Hence there exists an integer N_2 such that for $n > N_2$,

$$\begin{aligned}
&n^{-1}\textstyle\sum_{k=0}^{n-1} \sup\{|\partial^2 h(x_k, x_{k+1}; \theta)/\partial\theta^2 - \partial^2 h(x_k, x_{k+1}; \theta_0)/\partial\theta^2| : |\theta - \theta_0| < \delta\} \\
&\qquad < 2E_{\theta_0}[\sup\{|\partial^2 h(X_0, X_1; \theta)/\partial\theta^2 - \partial^2 h(X_0, X_1; \theta_0)/\partial\theta^2| : |\theta - \theta_0| < \delta\}],
\end{aligned} \tag{2.21}$$

N_2 possibly depending on the sample. Combining (2.19), (2.20), and (2.21), we obtain that for $n > \max(N_1, N_2)$, $|\varepsilon_n| < \varepsilon$ which proves that ε_n tends to zero almost surely. This completes the proof of this theorem.

3. The Bernstein–von Mises theorem. We shall now prove the main result of this paper—Theorem 3.1, which may be regarded as a generalized version of the

Bernstein-von Mises theorem for Markov chains. First, let us denote by $f_n(\theta \mid x_0, \cdots, x_n)$ the posterior density based on the observations $x_0, \cdots, x_n$ from the chain and corresponding to the prior probability density λ. Let

$$f_n^*(t \mid x_0, \cdots, x_n) = n^{-\frac{1}{2}} f_n(\theta \mid x_0, \cdots, x_n),$$

i.e. $f_n^*(t \mid x_0, \cdots, x_n)$ is the posterior density of $n^{\frac{1}{2}}(\theta - \hat{\theta}_n)$.

THEOREM 3.1. *Under Assumptions* 1.1–1.8.

$$\lim \int_{-\infty}^{\infty} K(t) | f_n^*(t \mid x_0, x_1, \cdots, x_n) - (i_0/(2\pi))^{\frac{1}{2}} \exp(-\tfrac{1}{2} i_0 t^2) | \, dt = 0 \text{ a.s. } P_0. \tag{3.1}$$

The proof of this result will be based on the following lemmas: Define

$$v_n(t) = \exp[\textstyle\sum_{i=0}^{n-1} \{h(x_i, x_{i+1}; \hat{\theta}_n + tn^{-\frac{1}{2}}) - h(x_i, x_{i+1}; \hat{\theta}_n)\}] \tag{3.2}$$

and

$$C_n = \int_{-\infty}^{\infty} v_n(t) \lambda(\hat{\theta}_n + tn^{-\frac{1}{2}}) \, dt. \tag{3.3}$$

It is easy to see that

$$f_n^*(t \mid x_0, x_1, \cdots, x_n) = C_n^{-1} v_n(t) \lambda(\hat{\theta}_n + tn^{-\frac{1}{2}}). \tag{3.4}$$

LEMMA 3.1. *Let Assumptions* 1.1–1.5 *be satisfied. Then the following conclusions hold. For every* $\varepsilon(0 < \varepsilon < i_0)$ *there exists a* $\delta_0 > 0$ *and an integer* N *such that*

$$v_n(t) \leqq \exp[-\tfrac{1}{2} t^2 (i_0 - \varepsilon)], \tag{3.5}$$

for $|t| \leqq \delta_0 n^{\frac{1}{2}}$ *and* $n \geqq N$; *for every* $\delta > 0$ *there exists a positive* ε *and an integer* N *such that*

$$\sup_{|t| > \delta n^{1/2}} v_n(t) \leqq \exp(-\tfrac{1}{4} n\varepsilon) \tag{3.6}$$

for $n \geqq N$; *for every fixed* t

$$\lim_{n \to \infty} v_n(t) = \exp[-\tfrac{1}{2} i_0 t^2] \tag{3.7}$$

a.s. P_0.

PROOF.

$$\begin{aligned} \log v_n(t) &= \textstyle\sum_{i=0}^{n-1} [h(x_i, x_{i+1}; \hat{\theta}_n + tn^{-\frac{1}{2}}) - h(x_i, x_{i+1}; \hat{\theta}_n)] \\ &= n^{-\frac{1}{2}} t \sum_{i=0}^{n-1} \frac{\partial h(x_i, x_{i+1}; \hat{\theta}_n)}{\partial \theta} + \frac{t^2}{2n} \sum_{i=0}^{n-1} \frac{\partial^2 h(x_i, x_{i+1}; \theta_n')}{\partial \theta^2} \end{aligned} \tag{3.8}$$

where $|\theta_n' - \hat{\theta}_n| \leqq tn^{-\frac{1}{2}}$.

The first term on the right-hand side of (3.8) equals zero a.s. for $n \geqq N_1$ (say) by (2.3) of Theorem 2.4. For the second term we have

$$\frac{t^2}{2n} \sum_{i=0}^{n-1} \frac{\partial^2}{\partial \theta^2} h(x_i, x_{i+1}; \theta_n') \tag{3.9}$$

$$= \frac{t^2}{2n}\sum_{i=0}^{n-1}\frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta_0) + \frac{t^2}{2n}\left[\sum_{i=0}^{n-1}\left\{\frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta_n') - \frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta_0)\right\}\right].$$

Since the first term on the right side of (3.9) converges a.s. (P_0) to $-\frac{1}{2}i_0 t^2$, it follows that for a positive ε, ($\varepsilon < i_0$)

$$\frac{t^2}{2n}\sum_{i=0}^{n-1}\frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta_0) < \frac{t^2}{2}\left(-i_0 + \frac{\varepsilon}{2}\right) \tag{3.10}$$

for $n \geqq N_2$ where we may take $N_2 \geqq N_1$. Now choose a positive δ less than $\frac{1}{2}\eta(\theta_0)$ where $\eta(\theta_0)$ is the number defined in Assumption 1.3. From (2.1) of Theorem 2.4 we have $|\hat{\theta}_n - \theta_0| < \delta$ and $|\theta_n' - \hat{\theta}_n| \leqq tn^{-\frac{1}{2}} \leqq \delta$ for $n \geqq N_3$ (say). Hence, if $n \geqq N_3$, $|\theta_n' - \theta_0| \leqq 2\delta$ and

$$n^{-1}\sum_{i=0}^{n-1}\left\{\frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta_n') - \frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta_0)\right\} \tag{3.11}$$

$$\leqq n^{-1}\sum_{i=0}^{n-1}\sup_{|\theta-\theta_0|\leqq 2\delta}\left\{\frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta) - \frac{\partial^2}{\partial\theta^2}h(x_i, x_{i+1}; \theta_0)\right\}$$

which tends (a.s. P_0) to

$$E\left\{\sup_{|\theta-\theta_0|\leqq 2\delta}\left[\frac{\partial^2}{\partial\theta^2}h(X_0, X_1; \theta) - \frac{\partial^2}{\partial\theta^2}h(X_0, X_1; \theta_0)\right]\right\}$$

by Assumption 1.3 and Theorem 2.1. From condition 1.2 the quantity inside the expectation tends to zero as $\delta \to 0$ for each (X_0, X_1) and is, furthermore, bounded above by the integrable function $2\sup_{|\theta-\theta_0|\leqq\eta(\theta_0)}\partial^2/(\partial\theta^2)h(x_0, x_1, \theta)$. Hence there exists a $\delta_0(0 < \delta < \frac{1}{2}\eta(\theta_0))$ such that

$$E\left\{\sup_{|\theta-\theta_0|\leqq 2\delta_0}\left[\frac{\partial^2}{\partial\theta^2}h(X_0, X_1; \theta) - \frac{\partial^2}{\partial\theta^2}h(X_0, X_1; \theta_0)\right]\right\} < \varepsilon/4. \tag{3.12}$$

Combining (3.8), (3.10) and (3.11), we get

$$\log v_n(t) < -\tfrac{1}{2}t^2(i_0 - \varepsilon), \qquad |t| \leqq n^{\frac{1}{2}}\delta_0, n > N \text{ (say)} \tag{3.13}$$

which proves (3.5). To prove (3.6), we have

$$n^{-1}\log v_n(t) = n^{-1}\textstyle\sum_{i=0}^{n-1}[h(x_i, x_{i+1}; \hat{\theta}_n + tn^{-\frac{1}{2}}) - h(x_i, x_{i+1}; \theta_0)] \tag{3.14}$$
$$+ n^{-1}\textstyle\sum_{i=0}^{n-1}[h(x_i, x_{i+1}; \theta_0) - h(x_i, x_{i+1}; \hat{\theta}_n)].$$

If $|\hat{\theta}_n - \theta_0| < \delta/2$ for $n \geqq N_4$, then $|tn^{-\frac{1}{2}}| > \delta$ implies that $|\hat{\theta}_n + tn^{-\frac{1}{2}} - \theta_0| > \delta/2$. Hence ($n \geqq N_4$)

$$n^{-1}\textstyle\sum_{i=0}^{n-1}[h(x_i, x_{i+1}; \hat{\theta}_n + t_n^{-\frac{1}{2}}) - h(x_i, x_{i+1}; \theta_0)]$$

$$\begin{aligned}(3.15)\qquad &\leqq n^{-1}\textstyle\sum_{i=0}^{n-1}\sup_{|\theta-\theta_0|>\delta/2}[h(x_i,x_{i+1};\theta)-h(x_i,x_{i+1};\theta_0)]\\ &\to E[\sup_{|\theta-\theta_0|>\delta/2}\{h(X_0,X_1,\theta)-h(X_0,X_1;\theta_0)\}]\\ &<0.\end{aligned}$$

from Assumption 1.4 and Theorem 2.1.

Moreover,

$$\begin{aligned}(3.16)\quad &\frac{1}{n}\sum_{i=0}^{n-1}\{h(x_i,x_{i+1};\theta_0)-h(x_i,x_{i+1};\hat{\theta}_n)\}\\ &=\frac{1}{n}(\theta_0-\hat{\theta}_n)\sum_{i=0}^{n-1}\frac{\partial}{\partial\theta}h(x_i,x_{i+1};\hat{\theta}_n)+\left(\frac{\theta_0-\hat{\theta}_n}{2n}\right)^2\sum_{i=0}^{n-1}\frac{\partial^2}{\partial\theta^2}h(x_i,x_{i+1};\theta_n'),\end{aligned}$$

which (by arguments given in (3.8) through (3.11) and the fact that $\hat{\theta}_n \to \theta_0$ a.s. P_0) converges to zero a.s. P_0. Now choose ε such that

$$0<\varepsilon< -E[\sup_{|\theta-\theta_0|}\{h(X_0,X_1;\theta)-h(X_0,X_1;\theta_0)\}].$$

Combining (3.15) and (3.16), it follows that for $|t|>n^{\frac{1}{2}}\delta$ and $n \geqq N = \max(N_1, N_2, N_3, N_4)$

$$(3.17)\qquad \log v_n(t) \leqq -\tfrac{1}{4}n\varepsilon,$$

which proves (3.6). Next, for a fixed t and $\varepsilon>0$ choose an ε_1 such that $0<(t^2/2)\varepsilon_1<\varepsilon$. By (3.8), (3.9), (3.10), and (3.11),

$$(3.18)\qquad |\log v_n(t)+\tfrac{1}{2}t^2 i_0|<\tfrac{1}{2}t^2\varepsilon_1<\varepsilon \qquad \text{for } n\geqq \max(N, t^2/\delta_0{}^2).$$

(3.7) is thus proved.

LEMMA 3.2. *Under the Assumptions* 1.1–1.8, *there exists a positive* δ_0 *such that*

$$(3.19)\quad \lim_{n\to\infty}\textstyle\int_{|t|\leqq\delta_0 n^{1/2}}K(t)|v_n(t)\lambda(\hat{\theta}_n+t_n{}^{-\frac{1}{2}})-\lambda(\theta_0)\exp(-\tfrac{1}{2}i_0t^2)|\,dt=0\,(\text{a.s. } P_0).$$

PROOF.

$$\begin{aligned}(3.20)\quad &\textstyle\int_{|t|\leqq\delta_0 n^{1/2}}K(t)|v_n(t)\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})-\lambda(\theta_0)\exp(-\tfrac{1}{2}i_0t^2)|\,dt\\ &\leqq\textstyle\int_{|t|\leqq\delta_0 n^{1/2}}K(t)\lambda(\theta_0)|v_n(t)-\exp(-\tfrac{1}{2}i_0t^2)|\,dt\\ &\quad+\textstyle\int_{|t|\leqq\delta_0 n^{1/2}}K(t)v_n(t)|\lambda(\theta_0)-\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})|\,dt.\end{aligned}$$

Choose an $\varepsilon>0$ such that $\int K(t)\exp(-\frac{1}{2}(i_0-\varepsilon)t^2)\,dt<\infty$.

This is possible because of Assumption 1.6. Then there exists a δ_1 and an N such that

$$v_n(t)\leqq\exp[-(i_0-\varepsilon)\tfrac{1}{2}t^2] \qquad |t|\leqq\delta_1 n^{\frac{1}{2}}, \qquad n\geqq N$$

by Lemma 3.1. Hence, using (3.17), we have, by the dominated convergence theorem

$$(3.21)\quad \textstyle\int_{|t|\leqq\delta_1 n^{1/2}}K(t)\lambda(\theta_0)|v_n(t)-\exp(-\tfrac{1}{2}i_0t^2)|\,dt\to 0 \quad\text{as}\quad n\to\infty, \text{a.s. } P_0.$$

Again,

$$\int_{|t|\leqq\delta_1 n^{1/2}} K(t)v_n(t)|\lambda(\theta_0)-\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})|\,dt$$
$$\leqq \sup_{|\theta-\theta_0|\leqq\delta_1}|\lambda(\theta)-\lambda(\theta_0)|\int_{|t|\leqq\delta_1 n^{1/2}} K(t)\exp[-(i_0-\varepsilon)\tfrac{1}{2}t^2]\,dt.$$

For a given η, choose $\delta_0 \leqq \delta_1$ such that

$$(3.22)\qquad \sup_{|\theta-\theta_0|<\delta_0}|\lambda(\theta)-\lambda(\theta_0)|\int_{|t|\leqq\delta_0 n^{1/2}} K(t)\exp[-(i_0-\varepsilon)\tfrac{1}{2}t^2]\,dt<\eta.$$

Combining (3.21) and (3.22), we get (3.19).

LEMMA 3.3. *Under Assumptions* 1.1–1.8, *for every* $\delta > 0$,

$$(3.23)\quad \lim_{n\to\infty}\int_{|t|>\delta n^{1/2}} K(t)|v_n(t)\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})-\lambda(\theta_0)\exp(-\tfrac{1}{2}i_0t^2)|\,dt=0\text{ a.s. } P_0.$$

PROOF.

$$\int_{|t|>\delta n^{1/2}} K(t)|v_n(t)\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})-\lambda(\theta_0)\exp(-\tfrac{1}{2}i_0t^2)|\,dt$$
$$\leqq \int_{|t|>\delta n^{1/2}} K(t)v_n(t)\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})\,dt+\int_{|t|>\delta n^{1/2}} K(t)\lambda(\theta_0)\exp(-\tfrac{1}{2}i_0t^2)\,dt$$
$$\leqq e^{-n\varepsilon/4}\int_{|t|>\delta n^{1/2}} K(t)\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})+\lambda(\theta_0)\int_{|t|>\delta n^{1/2}} K(t)\exp(-\tfrac{1}{2}i_0t^2)\,dt \quad \text{by (3.6)}$$
$$\to 0 \qquad \text{as } n\to\infty, \text{ by Assumptions 1.7 and 1.6.}$$

PROOF OF THEOREM 3.1. From Lemmas 3.2 and 3.3 we obtain

$$(3.24)\qquad \lim_{n\to\infty}\int K(t)|v_n(t)\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})-\lambda(\theta_0)\exp(-\tfrac{1}{2}i_0t^2)|\,dt=0\ (\text{a.s. } P_0).$$

Putting $K(t)\equiv 1$, which satisfies the assumptions on the function K trivially, we get

$$(3.25)\qquad C_n=\int v_n(t)\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})\,dt\to\lambda(\theta_0)\int\exp(-\tfrac{1}{2}i_0t^2)\,dt$$
$$=\lambda(\theta_0)((2\pi)/i_0)^{\frac{1}{2}}.$$

Hence

$$\int K(t)|f_n^*(t\mid x_0,\cdots,x_n)-(i_0/(2\pi))^{\frac{1}{2}}\exp(-\tfrac{1}{2}i_0t^2)|\,dt$$
$$\leqq\int K(t)|C_n^{-1}\lambda(\hat{\theta}_n+tn^{-\frac{1}{2}})v_n(t)-C_n^{-1}\lambda(\theta_0)\exp(-\tfrac{1}{2}i_0t^2)|\,dt$$
$$+\int K(t)|C_n^{-1}\lambda(\theta_0)-(i_0/(2\pi))^{\frac{1}{2}}|\exp(-\tfrac{1}{2}i_0t^2)\,dt$$
$$\to 0\ (\text{a.s. } P_0) \qquad \text{by (3.24) and (3.25).}$$

As a corollary to Theorem 3.1 we give below a result which includes, besides the more traditional form of the theorem of Bernstein and von Mises, other interesting variations.

THEOREM 3.2. *Let Assumptions* 1.1–1.5 *be satisfied. Let the prior density* λ *satisfy Assumption* 1.8 *and*

$$(3.26)\qquad \int_{-\infty}^{\infty}|\theta|^m\lambda(\theta)\,d\theta<\infty$$

for some nonnegative integer m. *Then*

$$(3.27)\quad \lim_{n\to\infty}\int_{-\infty}^{\infty}|t|^m|f_n^*(t\mid x_0,\cdots,x_n)-(i_0/(2\pi))^{\frac{1}{2}}e^{-\frac{1}{2}i_0t^2}|\,dt=0\ (\text{a.s. } P_0).$$

PROOF. For $m \geqq 1$ set $K(\theta) = |\theta|^m$ in Theorem 3.1. From the elementary identity $|a+b|^m \leqq 2^{m-1}(|a|^m + |b|^m)$ we have (for positive numbers h and δ)

$$\begin{aligned} e^{-\delta n} & \int_{|t|>h} K(n^{\frac{1}{2}}t)\lambda(\hat{\theta}_n + t)\,dt \\ &= n^{m/2} e^{-\delta n} \int_{|t-\hat{\theta}_n|>h} \lambda(t)|t-\hat{\theta}_n|^m\,dt \\ &\leqq 2^{m-1} n^{m/2} e^{-\delta n}\Big[\int_{|t-\hat{\theta}_n|>h} \lambda(t)|t|^m\,dt + |\hat{\theta}_n|^m \int_{|t-\hat{\theta}_n|>h} \lambda(t)\,dt\Big] \\ &\leqq 2^{m-1} n^{m/2} e^{-\delta n}\Big[\int_{-\infty}^{\infty} |t|^m \lambda(t)\,dt + |\hat{\theta}_n|^m\Big] \end{aligned}$$

which tends to zero a.s. P_0 from (2.1) of Theorem 2.4 and assumption (3.26). Hence condition 1.7 holds. Condition 1.6 is easily verified for $K(t) = |t|^m$. For $m = 0$, the verification of Assumptions 1.6 and 1.7 follows even more simply. The conclusion (3.27) now follows immediately from Theorem 3.1.

For $m = 0$ the assertion of Theorem 3.2 is the classical form of the Bernstein–von Mises result, while Theorem 3.2 itself is an extension to Markov chains of the corresponding result for the case of independent random variables due to Bickel and Yahav ([1], Theorem 2.2. Note that Theorem 2.2 is given for θ a vector parameter. The extension to a vector parameter of Theorems 3.1 and 3.2 can be carried out without difficulty).

4. Some asymptotic properties of regular Bayes estimates. In this section we shall be concerned with showing one application of Theorem 3.1 in the theory of asymptotic Bayesian inference for Markov chains. We shall derive results similar to those in [1].

Following LeCam [5], we define a regular Bayes estimate $T_n = T_n(x_0, \cdots, x_n)$ as an estimate which minimizes $B_n(\beta) = \int l(\theta, \beta)\, f_n(\theta \mid x_0, \cdots, x_n)\, d\theta$ for all $(x_0, x_1, \cdots, x_n)$ and all n, where $l(\theta, \beta)$ is a loss function defined on $\Theta \times \Theta$. In the sequel we shall assume that a measurable, regular Bayes estimate T_n exists. We now prove the main result of this section.

THEOREM 4.1. *Let a Markov chain $\{X_n, n \geqq 0\}$ satisfy the assumptions of Section* 1. *Let T_n be a measurable regular Bayes estimate of θ with respect to a loss function $\bar{l}(\theta, \beta)$ which satisfies the following conditions:*

(4.1) $$\bar{l}(\theta, \beta) = l(\theta - \beta) \geqq 0$$

(4.2) $$l(t_1) \geqq l(t_2) \qquad \text{if } t_1 > t_2 \geqq 0 \text{ or if } t_1 < t_2 \leqq 0.$$

There exist constants $\{a_n\}$ and functions $K(t)$, and $G(t)$ such that

(4.3) $$a_n \geqq 0.$$

(4.4) $G(t)$ *satisfies Assumptions* 1.6–1.7 *and* $a_n l(t/n^{\frac{1}{2}}) \leqq G(t)$ *for all* n.

(4.5) $$a_n l(t/n^{\frac{1}{2}}) \to K(t) \quad \text{uniformly on compact sets.}$$

(4.6) $$\int K(t+m)\exp(-\tfrac{1}{2} i_0 t^2)\,dt \quad \text{has a strict minimum at} \quad m = 0.$$

Then

$$T_n \to \theta_0 \qquad \text{a.s. } P_0. \tag{4.7}$$

$$n^{\frac{1}{2}}(\theta - T_n) \to_{\mathscr{L}} N(0, i_0^{-1}). \tag{4.8}$$

$$a_n B_n(T_n) \to (i_0/(2\pi))^{\frac{1}{2}} \int K(t) \exp(-\tfrac{1}{2} i_0 t^2)\, dt \text{ as } n \to \infty \text{ (a.s. } P_0). \tag{4.9}$$

PROOF. We shall show that $n^{\frac{1}{2}}(\theta_n - T_n) \to 0$ a.s. P_0 and that $a_n B_n(\hat{\theta}_n) \to (i_0/(2\pi))^{\frac{1}{2}} \int K(t) \exp(-\frac{1}{2} i_0 t^2)\, dt$, from which (4.7)–(4.9) follow easily, due to Theorem 2.4. First note that

$$\begin{aligned}
\limsup_n a_n B_n(T_n) &\leqq \limsup_n a_n B_n(\hat{\theta}_n) \\
&= \limsup_n \int a_n l(\theta - \hat{\theta}_n) f_n(\theta \mid x_0, \cdots, x_n)\, d\theta \\
&= \limsup_n \int a_n l(t/n^{\frac{1}{2}}) f_n^*(t \mid x_0, \cdots, x_n)\, dt \\
&\leqq \limsup_n \int |a_n l(t/n^{\frac{1}{2}}) - K(t)|\, |f_n^*(t) \\
&\quad - \exp(-\tfrac{1}{2} i_0 t^2)(i_0/(2\pi))^{\frac{1}{2}}|\, dt \\
&\quad + \limsup_n (i_0/(2\pi))^{\frac{1}{2}} \int |a_n l(t n^{-\frac{1}{2}}) - K(t)| \exp(-\tfrac{1}{2} i_0 t^2)\, dt \\
&\quad + \limsup_n \int K(t) f_n^*(t \mid x_0, \cdots, x_n)\, dt.
\end{aligned} \tag{4.10}$$

The first term on the right-hand side of (4.10) is $\leqq \limsup_n \int 2G(t) |f_n^*(t) - (i_0/(2\pi))^{\frac{1}{2}} \exp(-\frac{1}{2} i_0 t^2)|\, dt \to 0$ by Theorem 3.1.

The second term on the right-hand side of (4.10) converges to zero by the dominated convergence theorem, whereas the last term converges to $(i_0/(2\pi))^{\frac{1}{2}} \int K(t) \exp(-\frac{1}{2} i_0 t^2)\, dt$. Thus

$$\limsup_n a_n B_n(T_n) \leqq (i_0/(2\pi))^{\frac{1}{2}} \int K(t) \exp(-\tfrac{1}{2} i_0 t^2)\, dt. \tag{4.11}$$

Next, we show that $n^{\frac{1}{2}}(\hat{\theta}_n - T_n) = U_n < \infty$ a.s. For, if not, for every $M > 0$, there exists a set A_M with $P_\theta(A_M) > 0$ such that $|U_n(x)| > M$ infinitely often for x in A_M. Without loss of generality assume that $U_n(x) > M$ infinitely often. Then, for the subsequence $\{n_j\}$ for which the inequality holds,

$$\begin{aligned}
a_{n_j} B_{n_j}(T_{n_j}) &= \int a_{n_j} l\left(\frac{t + U_{n_j}}{n_j^{\frac{1}{2}}}\right) f_{n_j}^*(t \mid x_0, \cdots, x_{n_j})\, dt \\
&\geqq \int_{|t| \leqq M} a_{n_j} l\left(\frac{t + U_{n_j}}{n_j^{\frac{1}{2}}}\right) f_{n_j}^*(t \mid x_0, \cdots, x_{n_j})\, dt \\
&\geqq \int_{|t| \leqq M} a_{n_j} l\left(\frac{t + M}{n_j^{\frac{1}{2}}}\right) f_{n_j}^*(t \mid x_0, \cdots, x_{n_j})\, dt \qquad (\text{since } t + U_{n_j} \geqq t + M \geqq 0) \\
&\to \int_{|t| \leqq M} K(t + M)(i_0/(2\pi))^{\frac{1}{2}} \exp(-\tfrac{1}{2} i_0 t^2)\, dt.
\end{aligned}$$

Since $K(t+M) I_{[|t| \leqq M]}$ is a non-decreasing function of M for each fixed t,

$$\begin{aligned}
\lim_{M \to \infty} \int_{|t| \leqq M} K(t+M)(i_0/(2\pi))^{\frac{1}{2}} \exp(-\tfrac{1}{2} i_0 t^2)\, dt &= K(\infty) \\
&> (i_0/(2\pi))^{\frac{1}{2}} \int K(t) \exp(-\tfrac{1}{2} i_0 t^2)\, dt.
\end{aligned}$$

Hence for a large enough M, for a set of positive probability

$$(4.12)\qquad \liminf_n a_n B_n(T_n) > (i_0/(2\pi))^{\frac{1}{2}} \int K(t) \exp(-\tfrac{1}{2} i_0 t^2) > \limsup_n a_n B_n(\hat{\theta}_n)$$

which contradicts the definition of T_n. Thus $\limsup_n |U_n| < \infty$ a.s. P_0.

Let, for an arbitrary $\varepsilon > 0$, B_M be the set such that for x in B_M, $|U_n(x)| \leqq M$ for every n and $P(B_M) > 1-\varepsilon$. For a fixed x in B_M, $U_n(x)$ is a bounded sequence, hence has a limit point m. Suppose $m \neq 0$. Then, for the subsequence $\{n_j\}$ for which $U_{n_j}(x) \to m$, we have

$$\begin{aligned}\liminf_j a_{n_j} B_{n_j}(T_{n_j}) &\geqq \liminf_j \int_{-T_0}^{T_0} a_{n_j} l\left(\frac{t+U_{n_j}}{n_j^{\frac{1}{2}}}\right) f_{n_j}^*(t \mid x_0, \cdots, x_{n_j})\, dt \\ &\geqq \int_{-T_0}^{T_0} \liminf_j a_{n_j} l\left(\frac{t+U_{n_j}}{n_j^{\frac{1}{2}}}\right) f_{n_j}^*(t \mid x_0, \cdots, x_{n_j})\, dt \\ &= (i_0/(2\pi))^{\frac{1}{2}} \int_{-T_0}^{T_0} K(t+M) \exp(-\tfrac{1}{2} t^2 i_0)\, dt \quad \text{(due to (4.5))}.\end{aligned}$$

Thus, by choosing a large enough T_0, we get

$$(4.13)\qquad \liminf_j a_{n_j} B_{n_j}(T_{n_j}) > (i_0/(2\pi))^{\frac{1}{2}} \int_{-\infty}^{\infty} K(t) \exp(-\tfrac{1}{2} t^2 i_0)\, dt - \varepsilon \quad \text{(due to (4.6))}.$$

Since (4.13) holds for every ε, we have

$$(4.14)\qquad \liminf_n a_{n_j} B_{n_j}(T_{n_j}) \geqq (i_0/(2\pi))^{\frac{1}{2}} \int_{-\infty}^{\infty} K(t) \exp(-\tfrac{1}{2} t^2 i_0)\, dt > \limsup_n a_n B_n(T_n)$$

which is impossible. Hence $m = 0$ and $n^{\frac{1}{2}}(T_n - \hat{\theta}_n) \to 0$. Moreover, (4.12) and (4.14) give

$$\lim a_n B_n(T_n) = \lim a_n B_n(\hat{\theta}_n) = (i_0/(2\pi))^{\frac{1}{2}} \int K(t) \exp(-\tfrac{1}{2} i_0 t^2)\, dt.$$

REFERENCES

[1] Bickel, P. and Yahav, J. (1969). Some contributions to the asymptotic theory of Bayes solutions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **11** 257–276.

[2] Billingsley, P. (1961a). *Statistical Inference for Markov Processes.* Univ. of Chicago Press.

[3] Billingsley, P. (1961b). The Lindeberg-Lévy theorem for martingales. *Proc. Amer. Math. Soc.* **12** 788–792.

[4] Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton Univ. Press.

[5] LeCam, L. (1955). On the asymptotic theory of estimation and testing of hypotheses. *Proc. Third Berkeley Symp. Math. Statist. Prob.* Univ. of California Press.

[6] LeCam, L. (1958). Les proprietes asymptotiques des solutions de Bayes. *Publ. Inst. Statist. Univ. Paris* **7** 18–35.

[7] Roussas, G. G. (1965a). Extension to Markov processes of a result by A. Wald about the consistency of the maximum likelihood estimate. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **4** 69–73.

[8] Roussas, G. G. (1965b). Asymptotic Inference in Markov processes. *Ann. Math. Statist.* **36** 978–992.

[9] Schwartz, L. (1965). On Bayes procedures. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **4** 10–26.