NONPARAMETRIC ESTIMATION IN MARKOV PROCESSES*

GEORGE G. ROUSSAS

(Received March 6, 1967)

1. Introduction and summary

The problem of statistical inferences in Markov processes has received considerable attention during the last fifteen years. Much of the work consists in carrying over to the Markov case the maximum likelihood and chi-square methods from processes with independent identically distributed random variables. (See, for example, [1] and other references cited there.) Alternative approaches have also been adopted [11], some of which [7] refer to statistical inferences in more general processes.

It is not long ago that presumably the first paper [10] appeared on nonparametric estimation of the density in the case of independent identically distributed random variables. Soon a number of others ([14], [8], [13], [3], [6]) followed, which by using either similar or different methods obtained further results.

The purpose of the present paper is to consider the nonparametric estimation of densities in the case of Markov processes. The methods being used and results being obtained here are similar to those in [9]. What we do specifically here is this: We first construct asymptotically unbiased estimates for the initial and (two-dimensional) joint densities. This is done in section 2. In section 3 these estimates are shown to be consistent in quadratic mean, and furthermore a consistent, in the probability sense, estimate for the transition density is obtained. Finally, it is proved in section 4 that, under suitable conditions, all three estimators mentioned, properly normalized, are asymptotically normal. The appropriate versions of the Central Limit Theorem which are used for this purpose are stated and proved in an appendix, so that the continuity of the paper will not be interrupted.

2. Asymptotically unbiased estimation of the initial and (two-dimensional) joint densities

The results of this paper, like those of [9] and [3], rely heavily on

* This paper was prepared with the support of the National Science Foundation, Grant GP-6242.

a slight variation of a theorem of Bochner [2] that we formulate here. By $C(f)$ we will denote the set of continuity points of the function f.

THEOREM 2.1. *Let* $(\mathcal{E}_m, \mathcal{B}^{(m)})$ *be the m-dimensional Euclidean space with the corresponding Borel σ-field and* $(R, \mathcal{B})$ *the Borel real line, and let* $K\colon (\mathcal{E}_m, \mathcal{B}^{(m)}) \to (R, \mathcal{B})$ *be measurable and such that*

$$(2.1) \qquad |K(z)| \leqq M_1(<\infty)\,, \qquad z \in \mathcal{E}_m\,; \qquad \int |K(z)|\,dz < \infty\,,$$

$$(2.2) \qquad \|z\|^m\,|K(z)| \to 0\,, \qquad as \quad \|z\| \to \infty\,,$$

where $\|\cdot\|$ *is the usual norm in* $\mathcal{E}_m$, *and integrals without limits here and thereafter are assumed to be taken over the whole space.*

Furthermore, let $g\colon (\mathcal{E}_m, \mathcal{B}^{(m)}) \to (R, \mathcal{B})$ *be measurable and such that*

$$(2.3) \qquad \int |g(z)|\,dz < \infty\,.$$

Define

$$(2.4) \qquad g_n(x) = h^{-m}(n) \int K(zh^{-1}(n))\,g(x-z)\,dz\,,$$

where $\{h(n)\}$, $n=1, 2, \cdots$ *is a sequence of positive constants such that*

$$(2.5) \qquad h(n) \to 0\,, \qquad as \quad n \to \infty\,.$$

Then for $x \in C(g)$,

$$(2.6) \qquad \lim g_n(x) = g(x) \int K(z)\,dz\,, \qquad as \quad n \to \infty\,.$$

If g *is continuous on* $\mathcal{E}_m$, *then the convergence* (2.6) *is uniform on compact subsets of* $\mathcal{E}_m$.

For the proof of the first part of the theorem the reader is referred to [3]. The uniform convergence assertion on compact subsets of $\mathcal{E}_m$ follows from the uniform continuity and boundedness of g on compact subsets of $\mathcal{E}_m$.

For later use, we now impose on the function K the further condition

$$(2.7) \qquad \int K(z)\,dz = 1\,.$$

Examples of functions satisfying conditions (2.1), (2.2) and (2.7) are given in [9] and [3].

Let $\{X_j\}$, $j=1, 2, \cdots$ be a stationary Markov process defined on the probability space $(\Omega, \mathcal{A}, P)$, and let p and q be, respectively, the initial and (two-dimensional) joint density with respect to Lebesgue measures. It is further assumed that p is strictly positive on R. Then $(q/p)=t$ is a transition density of the process.

For $i=1, 2$, we consider two functions K_i such that K_i: $(\mathcal{E}_i, \mathcal{B}^{(i)}) \rightarrow (R, \mathcal{B})$ are measurable and satisfy conditions (2.1), (2.2) and (2.7). On the basis then of the first $n+1$ random variables X_j, $j=1, \cdots, n+1$ of the Markov process, we define the following random variables (suppressing the random element $\omega \in \Omega$)

$$p_n(x)=(nh_1(n))^{-1}\sum_{j=1}^{n} K_1((x-X_j)h_1^{-1}(n)) , \qquad x \in \mathcal{E}_1 , \tag{2.8}$$

$$q_n(y)=(nh_2^2(n))^{-1}\sum_{j=1}^{n} K_2((y-Y_j)h_2^{-1}(n)) , \qquad y \in \mathcal{E}_2 , \tag{2.9}$$

where $Y_j=(X_j, X_{j+1})$, $j=1, \cdots, n$, and $h_1(n)$, $h_2(n)$ satisfy (2.5). For convenient reference we will denote by (C_i') the assumption that K_i and h_i satisfy (2.1), (2.2), (2.7) and (2.5), $i=1, 2$. We intend to show that p_n and q_n are asymptotically unbiased estimates of p and q, respectively. More precisely, we have

THEOREM 2.2. *Asymptotic unbiasedness. Under (C_1') and (C_2'), respectively, the random variables defined by* (2.8) *and* (2.9) *are asymptotically unbiased estimates of p and q, respectively, in the sense that*

$$Ep_n(x) \rightarrow p(x) , \qquad as \quad n \rightarrow \infty, \quad x \in C(p) ,$$

and

$$Eq_n(y) \rightarrow q(y) , \qquad as \quad n \rightarrow \infty, \quad y \in C(q) .$$

Furthermore these estimates are uniformly asymptotically unbiased on compact subsets of $\mathcal{E}_i$, $i=1, 2$ if p and q are continuous on $\mathcal{E}_1$ and $\mathcal{E}_2$, respectively.

PROOF. The proof is an immediate application of Theorem 2.1. In fact, writing h_1 and h_2 instead of $h_1(n)$ and $h_2(n)$, we get

$$\begin{aligned} Ep_n(x) &= h_1^{-1}\int K_1((x-z)h_1^{-1})\, p(z)\, dz \\ &= h_1^{-1}\int K_1(zh_1^{-1})\, p(x-z)\, dz , \end{aligned}$$

and as $n \rightarrow \infty$, this converges to $p(x)$, provided $x \in C(p)$; this convergence is uniform on compact subsets of $\mathcal{E}_1$ if p is continuous. Similarly,

$$\begin{aligned} Eq_n(y) &= h_2^{-2}\int K_2((y-z)h_2^{-1})\, q(z)\, dz \\ &= h_2^{-2}\int K_2(zh_2^{-1})\, q(y-z)\, dz , \end{aligned}$$

and as $n \to \infty$, this converges to $q(y)$, provided $y \in C(q)$; this convergence is uniform on compact subsets of $\mathcal{E}_2$ if q is continuous.

3. Consistent and uniform consistent estimation

The results of this section as well as those of the next one are derived under the additional assumption that the process satisfies hypothesis (D_0) ([5], p. 221). Namely,

HYPOTHESIS (D_0). (a) Condition (D) (Doeblin's condition) is satisfied, and (b) there is only a single ergodic set and this set contains no cyclically moving subsets.

We first prove consistency in quadratic mean. We have

$$E[p_n(x)-p(x)]^2=\sigma^2[p_n(x)]+[Ep_n(x)-p(x)]^2 ,$$

and the second term converges to zero, as $n \to \infty$, provided $x \in C(p)$; this convergence is uniform on compact subsets of $\mathcal{E}_1$ if p is continuous by Theorem 2.2. Next,

$$\sigma^2[p_n(x)]=n^{-1}h_1^{-2}\sigma^2[K_1((x-X_1)h_1^{-1})] \\ +2(nh_1)^{-2}\sum_{i<j}\operatorname{Cov}[K_1((x-X_i)h_1^{-1}),\ K_1((x-X_j)h_1^{-1})] ,$$

where the summation extends over all i's and j's such that $1 \leqq i < j \leqq n$. But

$$h_1^{-1}\sigma^2[K_1((x-X_1)h_1^{-1})]=h_1^{-1}\int K_1^2((x-z)h_1^{-1})p(z)\,dz \\ -h_1\Big[h_1^{-1}\int K_1((x-z)h_1^{-1})p(z)\,dz\Big]^2 ,$$

and for $x \in C(p)$ this tends to $p(x)\int K_1^2(z)\,dz$ as $n \to \infty$ by Theorem 2.1, since $\int K_1^2(z)\,dz$ is finite, as is easily seen from (2.1). This convergence is uniform on compact subsets of $\mathcal{E}_1$ if p is continuous. As for the covariance we have:

Under hypothesis (D_0), Lemma 7.1 in ([5], p. 222) applies and gives

$$|\operatorname{Cov}[K_1((x-X_1)h_1^{-1}),\ K_1((x-X_{j+1})h_1^{-1})]| \leqq 2\gamma^{1/2}\rho^{j/2}[E|K_1((x-X_1)h_1^{-1})|^2]^{1/2}$$

for some γ, ρ such that $\gamma>0$, $0<\rho<1$. Therefore

$$|(nh_1)^{-1}\sum_{i<j}\operatorname{Cov}[K_1((x-X_i)h_1^{-1}),\ K_1((x-X_j)h_1^{-1})]| \\ \leqq (nh_1)^{-1}\sum_{j=1}^{n-1}(n-j)2\gamma^{1/2}\rho^{j/2}E|K_1((x-X_1)h_1^{-1})|^2 \\ \leqq (nh_1)^{-1}n\rho^{1/2}(1-\rho^{1/2})^{-1}2\gamma^{1/2}E|K_1((x-X_1)h_1^{-1})|^2 \\ =2\gamma^{1/2}\rho^{1/2}(1-\rho^{1/2})^{-1}h_1^{-1}E|K_1((x-X_1)h_1^{-1})|^2 ,$$

and this last expression converges as $n \to \infty$ to

$$2\gamma^{1/2}\rho^{1/2}(1-\rho^{1/2})^{-1}p(x)\int K_1^2(z)\,dz \qquad \text{for} \quad x \in C(p)\ ,$$

and the convergence is uniform on compact subsets of $\mathcal{E}_1$ if p is continuous. Thus, if we assume that $h_1 = h_1(n)$ can be chosen so that

$$nh_1(n) \to \infty\ , \qquad \text{as} \quad n \to \infty\ , \tag{3.1}$$

it follows that

$$\sigma^2[p_n(x)] \to 0\ , \qquad \text{as} \quad n \to \infty, \quad x \in C(p)$$

and this convergence is uniform on compact subsets of $\mathcal{E}_1$ if p is continuous. Denoting by (C_1), for convenience, the assumption that both (C_1') and (3.1) are satisfied, we then get:
Under (C_1) and (D_0), $E[p_n(x)-p(x)]^2 \to 0$ as $n \to \infty$ provided $x \in C(p)$, and this convergence is uniform on compact subsets of $\mathcal{E}_1$ if p is continuous.

In a similar fashion we get that:
Under (C_2) and (D_0), $E[q_n(y)-q(y)]^2 \to 0$, as $n \to \infty$, provided $y \in C(q)$, and this convergence is uniform on compact subsets of $\mathcal{E}_2$ if q is continuous. Here by (C_2) we denote the assumption that both (C_2') and (3.2) are satisfied, where

$$nh_2^2(n) \to \infty\ , \qquad \text{as} \quad n \to \infty\ . \tag{3.2}$$

Putting together these results we have the following theorem:

THEOREM 3.1. *Consistency in quadratic mean (q.m.). Under* (D_0) *and* (C_1) *and* (C_2), *respectively, the random variables defined by* (2.8) *and* (2.9) *are consistent in q.m. estimates of* p *and* q, *respectively, for* $x \in C(p)$, $y \in C(q)$; *and they are uniformly consistent in q.m. estimates on compact subsets of* $\mathcal{E}_i$, $i=1,2$ *if* p *and* q *are continuous.*

Of course, consistency in q.m. (and Tchebichev inequality) implies consistency in the probability sense for $x \in C(p)$, $y \in C(q)$, and this consistency is uniform on compact subsets of $\mathcal{E}_i$, $i=1,2$ if p and q are continuous.

By taking into account now that the random variables (2.8) are to be used in order to estimate the positive quantity p, one may assume that K_1 is strictly positive. Under this condition a consistent estimate of the transition density can be constructed. More precisely, we have

COROLLARY 3.1. *Let*

$$y=(x, x') \in C(q) \qquad \textit{and} \quad x \in C(p).$$

We set

$$t_n(x'|x)=[q_n(y)/p_n(x)]\ ,$$

and

$$t(x'|x)=[q(y)/p(x)]\ .$$

Then, as $n\to\infty$, $t_n(x'|x)\to t(x'|x)$ *in probability and this convergence is uniform on compact subsets of* $\mathcal{E}_2$ *if* p *and* q *are continuous.*

4. Asymptotic normality

In this section asymptotic normality of the estimators p_n, q_n and t_n will be obtained under some further restrictions on the process. Actually, these results are merely an application of the results presented in the appendix, and have also served as a motivation for the type of assumption being made there.

In (A1) of the appendix we take $h_n=nh_1(n)$. Then (A1) is satisfied on account of (3.1) herein. Next for $r=1$ in (A2) and with $L_n(z)$ being replaced by $K_1((x-z)h_1^{-1})$, (A2) (i) and (A2) (iv) are automatically satisfied on the basis of Theorem 2.1 here with $\sigma_1^2(x)=p(x)\int K_1^2(z)\,dz$, $x\in C(p)$. As for (A2) (ii) and (A2) (iii), they clearly follow from the assumption being made below.

The joint densities of X_1, X_i and X_1, X_i, X_j are bounded by $M_2(<\infty)$ for all i, j such that

$$1<i\leqq n\ ,\qquad 1<i<j\leqq n\ ,\qquad n=2,3,\cdots\ . \tag{4.1}$$

In the appendix the positive integers α,β, and μ are introduced with the property that they tend to infinity together with n and also satisfy the properties:

$$\beta\mu\alpha^{-1}\to 0\quad \text{and}\quad \alpha h_n n^{-1}\to 0\ ,\qquad \text{as}\quad n\to\infty\ .$$

With the above choice of h_n these relations become

$$\beta\mu\alpha^{-1}\to 0\quad \text{and}\quad \alpha h_1(n)\to 0\ ,\qquad \text{as}\quad n\to\infty\ . \tag{4.2}$$

Theorem 1 in the appendix then becomes

THEOREM 4.1. *Let assumptions* (D_0), (C_1), (4.1) *and* (4.2) *be satisfied. Then for* $x\in C(p)$,

$$\mathcal{L}\{(nh_1)^{1/2}[p_n(x)-Ep_n(x)]\}\to N(0,\sigma_1^2(x))\ ,\qquad \textit{as}\quad n\to\infty\ ,$$

where

$$\sigma_1^2(x)=p(x)\int K_1^2(z)\,dz\ .$$

We next choose $h_n = nh_2^2(n)$, and then (A1) in the appendix is again satisfied by (3.2) herein. For $s=2$ in (A2)* and with $L_n^*(z)$ being replaced by $K_2((y-z)h_2^{-1})$, (A2)* (i) and (A2)* (iv) are automatically satisfied on account of Theorem 2.1 of this paper with $\sigma_2^2(y) = q(y)\int K_2^2(z)\,dz$, $y \in C(q)$. As for (A2)* (ii) and (A2)* (iii), they follow in an obvious way from (4.3) below.

The joint densities of Y_1, Y_i and Y_1, Y_i, Y_j are bounded by $M_3(<\infty)$ for all i, j such that

$$(4.3) \qquad 1 < i \leqq n\,, \qquad 1 < i < j \leqq n\,, \qquad n = 2, 3, \cdots\,.$$

We finally require α, β and μ to tend to infinity as $n \to \infty$, and be such that

$$(4.4) \qquad \beta\mu\alpha^{-1} \to 0 \quad \text{and} \quad \alpha h_2^2(n) \to 0\,, \qquad \text{as} \quad n \to \infty\,.$$

Then Theorem 1* in the appendix becomes as follows.

THEOREM 4.2. *Let assumptions* (D_0), (C_2), (4.3) *and* (4.4) *be satisfied. Then for* $y \in C(q)$ *and such that* $q(y) > 0$, *we have*

$$\mathcal{L}\{(nh_2^2)^{1/2}[q_n(y) - Eq_n(y)]\} \to N(0, \sigma_2^2(y))\,, \qquad as \quad n \to \infty\,,$$

where $\sigma_2^2(y) = q(y)\int K_2^2(z)\,dz$.

Finally we will examine the estimator of the transition density from the point of view of asymptotic normality.

In the first place we take $h_1(n) = h_2^2(n) = h(n)$, $n = 1, 2, \cdots$ for simplicity. Thus h_n in the appendix is now $h_n = nh_1(n) = nh_2^2(n)$. Next (A1)** (i) is again clearly true, and so is (A1)** (ii) with l being $p(x)$, $x \in C(p)$ because of Theorem 3.1 herein. Furthermore (A3)** (i) follows from (4.1) and (4.3), (A3)** (ii) is true with $v(x, y) = -[q(y)/p(x)]$ by Theorem 2.2 herein, provided $x \in C(p)$, $y \in C(q)$, and (A3)** (iii) is also valid with $\sigma = 0$ on account of (4.1). Therefore Theorem 2 of the appendix becomes as follows.

THEOREM 4.3. *Let assumptions* (D_0), (C_1), (C_2), (4.1), (4.2) *and* (4.3) *be satisfied. Then for* $y = (x, x') \in C(q)$ *such that* $x \in C(p)$, *we have that the law of*

$$(nh)^{1/2}\{t_n(x'|x) - [EK_2((y - Y_1)h^{-1/2})/EK_1((x - X_1)h^{-1})]\}$$

converges to

$$N(0, \sigma_0^2(x, y)l^{-2}(x))\,, \qquad as \quad n \to \infty\,,$$

where

$$\sigma_0^2(x, y) = \sigma_2^2(y) + v^2(x, y)\sigma_1^2(x)$$

and

$$\sigma_1^2(x)=p(x)\int K_1^2(z)\,dz\ ,\qquad \sigma_2^2(y)=q(y)\int K_2^2(z)\,dz\ ,$$

$$v(x,y)=-[q(y)/p(x)]\ ,\qquad l(x)=p(x)\ ,$$

provided $q(y)>0$.

Appendix

In this appendix some minor generalizations of known results are presented for the sake of completeness of the paper.

It is assumed throughout that the Markov process $\{X_n\}$, $n=1,2,\cdots$ satisfies hypothesis (D_0), and we set

$$Y_j=(X_j,\cdots,X_{j+r-1})\ ,\qquad Z_j=(X_j,\cdots,X_{j+s-1})\ ,\qquad j=1,2,\cdots\ ,$$

where r and s are two integers greater than or equal to 1. We note then that the processes $\{Y_j\}$, $\{Z_j\}$, $j=1,2,\cdots$ are Markov process which also satisfy hypothesis (D_0) ([5], p. 231).

For the formulation and proof of the first result here we need to introduce some additional notation and make the following assumptions:

(A1) $\{h_n\}$, $n=1,2,\cdots$ is a sequence of positive constants such that $h_n\to\infty$, as $n\to\infty$.

(A2) For $n=1,2,\cdots$, $\{L_n\}$ is a sequence of uniformly bounded real-valued measurable functions on $(\mathcal{E}_r,\mathcal{B}^{(r)})$ such that

(i) $E|L_n(Y_1)|^2$ is $O(h_n n^{-1})$,

(ii) $E|f_n(Y_1)f_n(Y_j)|$ are $O(h_n^2 n^{-2})$ uniformly in j, $1<j\leqq n$,

(iii) $E|f_n(Y_1)f_n(Y_i)f_n(Y_j)|$ are $O(h_n^3 n^{-3})$ uniformly in i and j, $1<i<j\leqq n$, $n=2,3,\cdots$,

(iv) $h_n^{-1}n\sigma^2[L_n(Y_1)]\to\sigma_1^2$ (for some $\sigma_1^2<\infty$), as $n\to\infty$,

where f_n is defined by

$$f_n(Y_j)=L_n(Y_j)-EL_n(Y_j)\ .$$

From (A2) (iv) it follows that $E|f_n(Y_1)|^2$ is $O(h_n n^{-1})$ and hence so is also $E|f_n(Y_1)|^3$ by the boundedness assumption of L_n. The same boundedness assumption and (A2) (ii) imply that $E|f_n^2(Y_i)L_n(Y_j)|$ are $O(h_n^2 n^{-2})$ uniformly in i and j with $i,j=1,\cdots,n$, $i\neq j$.

Under the regularity assumptions (A2) and an additional one which we will make later on, the asymptotic normality of

$$(1)\qquad h_n^{-1/2}\sum_{j=1}^{n} f_n(Y_j)$$

will be established. In discussing the asymptotic normality of (1) we follow a method parallel to the one used in proving Theorem 7.5 in ([5], p. 228).

First, $\sum_{j=1}^{n} f_n(Y_j)$ is split up as follows.

Define

$$y_m(n)=\sum_j f_n(Y_j)\ ,$$

where the summation extends from $(m-1)(\alpha+\beta)+1$ to $(m-1)(\alpha+\beta)+\alpha$, $m=1,\cdots,\mu$, $y'_m(n)=\sum_j f_n(y_j)$, where the summation extends from $(m-1)(\alpha+\beta)+\alpha+1$ to $m(\alpha+\beta)$, $m=1,\cdots,\mu$, $y'_{\mu+1}=\sum_j f_n(Y_j)$, where the summation extends from $\mu(\alpha+\beta)+1$ to n. The numbers α, β and μ are positive integers which tend to infinity, as $n\to\infty$, and are such that $\mu(\alpha+\beta)$ is the largest multiple of $\alpha+\beta$ which is $\leqq n$.

Clearly, we get

$$h_n^{-1/2}\sum_{j=1}^{n} f_n(Y_j)=h_n^{-1/2}\sum_{m=1}^{\mu} y_m(n)+h_n^{-1/2}\sum_{m=1}^{\mu+1} y'_m(n)\ .$$

It is first proved that

$$(2)\qquad h_n^{-1/2}\sum_{m=1}^{\mu+1} y'_m(n)\to 0$$

in probability, as $n\to\infty$ $(\mu\to\infty)$.

By the Tchebichev inequality, it suffices to prove that

$$(3)\qquad h_n^{-1}E\left|\sum_{m=1}^{\mu+1} y'_m(n)\right|^2\to 0\ ,\qquad \text{as}\quad n\to\infty\ (\mu\to\infty)\ .$$

Under the hypothesis (D_0), Lemma 7.1 in ([5], p. 222) applies and gives

$$E|y'_m(n)|^2\leqq\beta c_1 EL_n^2(Y_1)\qquad \text{for}\quad m=1,\cdots,\mu$$

and

$$E|y'_{\mu+1}(n)|^2\leqq|n-\mu(\alpha+\beta)|\, c_1 EL_n^2(Y_1)\ ,$$

where

$$c_1=4\gamma_1^{1/2}\rho_1^{1/2}(1-\rho_1^{1/2})^{-1}+1\ ,$$

the constants γ_1 and ρ_1 corresponding to the process $\{Y_j\}$, $j=1,2,\cdots$.

The Minkowski inequality gives

$$h_n^{-1/2}E^{1/2}\left|\sum_{m=1}^{\mu+1} y'_m(n)\right|^2 \leqq h_n^{-1/2}\mu E^{1/2}|y'_1(n)|^2 + h_n^{-1/2}E^{1/2}|y'_{\mu+1}(n)|^2 .$$

Using then the previous two inequalities we get

$$h_n^{-1/2}\mu E^{1/2}|y'_1(n)|^2 \leqq (\beta\mu^2 h_n^{-1})^{1/2} c_1^{1/2} E^{1/2} L_n^2(Y_1)$$

and

$$h_n^{-1/2}E^{1/2}|y'_{\mu+1}(n)|^2 \leqq [n-\mu(\alpha+\beta)]^{1/2} h_n^{-1/2} c_1^{1/2} E^{1/2} L_n^2(Y_1) .$$

Now

$$\beta\mu^2 n^{-1} \leqq \beta\mu\alpha^{-1} ,$$

as is easily seen, and hence

$$\beta\mu^2 h_n^{-1} = (nh_n^{-1})(\beta\mu^2 n^{-1}) \leqq (nh_n^{-1})(\beta\mu\alpha^{-1}) .$$

By choosing α, β, and μ to tend to infinity as $n \rightarrow \infty$ so that

$$(4) \qquad \beta\mu\alpha^{-1} \rightarrow 0 ,$$

we then get

$$(5) \qquad h_n^{-1/2}\mu E^{1/2}|y'_1(n)|^2 \rightarrow 0 , \qquad \text{as} \quad n \rightarrow \infty \ (\mu \rightarrow \infty) ,$$

by means of (A2) (i).

Next,

$$[n-\mu(\alpha+\beta)]h_n^{-1} = (nh_n^{-1})[n-\mu(\alpha+\beta)]n^{-1} \leqq (nh_n^{-1})\mu^{-1} ,$$

as is easily seen, and hence

$$(6) \qquad h_n^{-1/2}E^{1/2}|y'_{\mu+1}(n)|^2 \rightarrow 0 , \qquad \text{as} \quad n \rightarrow \infty \ (\mu \rightarrow \infty) ,$$

again because of (A2) (i). Relations (5) and (6) taken together imply (3) and hence (2).

Next, we prove the asymptotic normality of

$$(7) \qquad h_n^{-1/2} \sum_{m=1}^{\mu} y_m(n) .$$

Setting

$$\Phi_m(t;\, n) = E\left\{\exp\left[it\sum_{j=1}^{n} f_n(Y_j)\right]\right\}$$

and repeating the arguments used in ([5], p. 229), we get

$$E\left\{\exp\left[it\sum_{m=1}^{\mu} y_m(n)\right]\right\} = \Phi_\alpha^\mu(t;\, n) + \zeta_\mu , \qquad |\zeta_\mu| < 2\gamma_1\mu\rho_1^{\beta+1} .$$

Again, α, β and μ are chosen so that they tend to infinity, as $n\to\infty$, and such that

$$\mu\rho_1^\beta \to 0 . \tag{8}$$

Then the characteristic function of (7) is essentially,

$$\Phi_\alpha^\mu(th_n^{-1/2};\, n) , \tag{9}$$

since $\zeta_\mu \to 0$, as $\mu\to\infty$, by (8).

Now (9) is the characteristic function of $\sum_{m=1}^{\mu} z_m$ where z_m, $m=1, \cdots, \mu$ are independent random variables with their common distribution being that of $h_n^{-1/2}y_1(n)$. Thus, the asserted normality of (7) will follow if we prove that

$$(C_\mu/B_\mu^{1+1/2}) \to 0 , \qquad \text{as} \quad n\to\infty \ (\mu\to\infty) , \tag{10}$$

by Theorem 4.4 in ([5], p. 141), where

$$B_\mu = \sum_{m=1}^{\mu} E(z_m^2) , \qquad C_\mu = \sum_{m=1}^{\mu} E|z_m|^3 , \qquad (E|z_1|^3 < \infty) .$$

Now,

$$E(z_m^2) = h_n^{-1} E\Big[\sum_{j=1}^{\alpha} f_n(Y_j)\Big]^2 ,$$

and

$$E\Big[\sum_{j=1}^{\alpha} f_n(Y_j)\Big]^2 = \alpha\sigma^2[L_n(Y_1)] + 2\sum_{i<j} E[f_n(Y_i)f_n(Y_j)] .$$

Thus,

$$B_\mu = (\alpha\mu n^{-1})nh_n^{-1}\,\sigma^2[L_n(Y_1)] + 2(\alpha\mu n^{-1})(\alpha h_n n^{-1}) \sum_{i<j} E[f_n(Y_i)f_n(Y_j)] .$$

But

$$\begin{aligned} \Big|\sum_{i<j} E[f_n(Y_i)f_n(Y_j)]\Big| &\leqq \alpha \sum_{j=1}^{\alpha-1} |E[f_n(Y_1)f_n(Y_{j+1})]| \\ &= (\alpha h_n n^{-1})^2\,(\alpha^{-1}h_n^{-2}n^2) \sum_{j=1}^{\alpha-1} |E[f_n(Y_1)f_n(Y_{j+1})]| . \end{aligned}$$

Therefore, by means of (A2) (ii), (A2) (iv) and the fact that $\alpha\mu n^{-1}\to 1$, as $n\to\infty$ ($\mu\to\infty$), as is easily seen, we obtain

$$B_\mu \to \sigma_1^2 , \qquad \text{as} \quad n\to\infty \ (\mu\to\infty) ,$$

provided that there is a choice of α satisfying (4) and also

$$\alpha h_n n^{-1} \to 0, \qquad \text{as} \quad n\to\infty , \tag{11}$$

for some choice of h_n satisfying (A1).

It remains for us to prove that $C_\mu \to 0$ as $n \to \infty$ $(\mu \to \infty)$. We have

$$C_\mu = \sum_{m=1}^{\mu} E|z_m|^3 = \mu h_n^{-3/2} E\left|\sum_{j=1}^{\alpha} f_n(Y_j)\right|^3 \leqq \mu h_n^{-3/2}\{\alpha E|f_n(Y_1)|^3$$
$$+3\sum_{i,j} E|f_n^2(Y_i)f_n(Y_j)| + 6\sum_{i<j<k} E|f_n(Y_i)f_n(Y_j)f_n(Y_k)|\} .$$

Now,

$$\mu h_n^{-3/2}\alpha E|f_n(Y_1)|^3 = (\alpha\mu n^{-1}) n h_n^{-3/2} E|f_n(Y_1)|^3 \to 0 , \qquad \text{as} \quad n \to \infty ,$$

by (A1) and the remark following (A2). In a similar fashion,

$$h_n^{-3/2}\sum_{i,j} E|f_n^2(Y_i)f_n(Y_j)| = (\alpha\mu n^{-1})(\alpha h_n n^{-1}) h_n^{-1/2}$$
$$\cdot\, n^2(\alpha h_n)^{-2}\sum_{i,j} E|f_n^2(Y_i)f_n(Y_j)| \to 0 , \qquad \text{as} \quad n \to \infty ,$$

by (A1), the remark following (A2) and (11). Finally,

$$\mu h_n^{-3/2}\sum_{i<j<k} E|f_n(Y_i)f_n(Y_j)f_n(Y_k)| = (\alpha\mu n^{-1})(\alpha h_n n^{-1})^2 h_n^{-1/2} n^3(\alpha h_n)^{-3}$$
$$\cdot \sum_{i<j<k} E|f_n(Y_i)f_n(Y_j)f_n(Y_k)| \to 0 , \qquad \text{as} \quad n \to \infty ,$$

on account of (A1), (A2) (iii) and (11). Therefore

$$C_\mu \to 0 , \qquad \text{as} \quad n \to \infty \ (\mu \to \infty) ,$$

and this establishes (10). Hence the following theorem has been proved.

THEOREM 1. *Let assumptions* (A1) *and* (A2) *be satisfied. We assume that a choice of* α *satisfying* (4) *can be made such that* (11) *is also satisfied. Then*

$$\mathcal{L}\left\{h_n^{-1/2}\sum_{j=1}^{n}[L_n(Y_j) - EL_n(Y_j)]\right\} \to N(0, \sigma_1^2) , \qquad as \quad n \to \infty ,$$

provided $\sigma_1^2 > 0$ *where* $\sigma_1^2 = \lim h_n^{-1} n\sigma^2[L_n(Y_1)]$ *as* $n \to \infty$.

Now let us replace the sequence $\{L_n\}$, $n=1, 2, \cdots$ by a sequence $\{L_n^*\}$, $n=1, 2, \cdots$ of uniformly bounded real-valued measurable functions on $(\mathcal{E}_s, \mathcal{B}^{(s)})$, and let g_n be defined by

$$g_n(Z_j) = L_n^*(Z_j) - EL_n^*(Z_j) .$$

If we impose upon L_n^* and g_n the same conditions we used in connection with L_n and f_n, which we denote by (A2)* here, then, under (A1) and (A2)*, we have a theorem analogous to Theorem 1. We will refer to it as Theorem 1*. The variance of the limiting normal distribution in this theorem will be denoted by σ_2^2.

Remarks. In the various derivations in proving Theorem 1* we will use the constant c_2 rather than c_1, where $c_2=4\gamma_2^{1/2}\rho_2^{1/2}(1-\rho_2^{1/2})^{-1}+1$, the constants γ_2 and ρ_2 corresponding to the process $\{Z_j\}$, $j=1, 2, \cdots$. There is always a choice of α, β and μ with the property that α, β and μ are positive integers tending to infinity with n, such that $\mu(\alpha+\beta)$ is the largest multiple of $\alpha+\beta$ which is $\leqq n$ and for which both conditions (4) and (8) (and the corresponding property: $\mu\rho_2^{\beta}\to 0$ as $n\to\infty$ $(\mu\to\infty)$), are satisfied. This is explained in ([5], p. 230). That is, it suffices to take β to be the largest integer which is $\leqq n^{1/4}$ and $\alpha=\beta^3$. It follows then that μ is approximately β and all required conditions are satisfied.

We now proceed in proving asymptotic normality for a certain quotient. For this purpose it is assumed that

$$EL_n(Y_1)\neq 0\ , \qquad n=1, 2, \cdots\ ,$$

and

$$h_n^{-1}\sum_{j=1}^{n} L_n(Y_j)\to l \ (\neq 0 \text{ constant}) \quad \text{in probability, as } n\to\infty .$$

Then

$$h_n^{1/2}\left\{\left[\sum_{j=1}^{n} L_n^*(Z_j)\Big/\sum_{j=1}^{n} L_n(Y_j)\right]-[EL_n^*(Z_1)/EL_n(Y_1)]\right\}$$

is well defined and we intend to prove its asymptotic normality, under some additional assumptions. It is easily seen that

$$\begin{aligned}
h_n^{1/2}&\left\{\left[\sum_{j=1}^{n} L_n^*(Z_j)\Big/\sum_{j=1}^{n} (Y_j)\right]-[EL_n^*(Z_1)/EL_n(Y_1)]\right\}\\
&=\left[h_n^{-1}\sum_{j=1}^{n} L_n(Y_j)\right]^{-1} h_n^{-1/2}\sum_{j=1}^{n}[L_n^*(Z_j)-EL_n^*(Z_j)]\\
&\qquad +v_n h_n^{-1/2}\sum_{j=1}^{n}[L_n(Y_j)-EL_n(Y_j)]\\
&=\left[h_n^{-1}\sum_{j=1}^{n} L_n(Y_j)\right]^{-1} h_n^{-1/2}\sum_{j=1}^{n}[\varphi_n(W_j)-E\varphi_n(W_j)]\\
&=\left[h_n^{-1}\sum_{j=1}^{n} L_n(Y_j)\right]^{-1} h_n^{-1/2}\sum_{j=1}^{n}\Psi_n(W_j)\ ,
\end{aligned}$$

where

$$\begin{aligned}
&v_n=-[EL_n^*(Z_1)][EL_n(Y_1)]^{-1}\ ,\\
&\varphi_n(W_j)=L_n^*(Z_j)+v_nL_n(Y_j)\ ,\\
&\Psi_n(W_j)=\varphi_n(W_j)-E\varphi_n(W_j)\ ,\\
&W_j=(X_j, \cdots, X_{j+t-1}) \qquad (t=\max(r, s))\ .
\end{aligned}$$

By a well-known result (see, for example, ([4], p. 254)), it suffices then to prove asymptotic normality for

$$h_n^{-1/2} \sum_{j=1}^{n} [\varphi_n(W_j) - E\varphi_n(W_j)] .$$

This last expression will clearly be asymptotically normal, provided φ_n and Ψ_n satisfy a condition analogous to (A2). Below, a theorem referring to the asymptotic normality of the expression in question is formulated, and a set of sufficient conditions for this theorem to be true is given. The conditions to be used in this subsection are

(A1)** (i) $EL_n(Y_1) \neq 0$, $n=1, 2, \cdots$,

(ii) $h_n^{-1} \sum_{j=1}^{n} L_n(Y_j) \to l$ ($\neq 0$ constant) in probability, as $n \to \infty$.

(A2)** For $n=1, 2, \cdots$, $\{\varphi_n\}$ is a sequence of uniformly bounded real-valued measurable functions on $(\mathcal{E}_t, \mathcal{B}^{(t)})$ such that the relations we get if L_n and f_n are replaced by φ_n and Ψ_n, respectively, in (A2) are true. (The relation corresponding to (A2) (iv) may be valid with a different constant $\sigma_0^2 < \infty$).

(A3)** (i) Both (A2) (ii) and (A2) (iii) remain true if any one or two f's are replaced by the corresponding g's.

(ii) $[EL_n^*(Z_1)][EL_n(Y_1)]^{-1} = -v_n \to -v$ (finite), as $n \to \infty$.

(iii) $h_n^{-1} n E[f_n(Y_1) g_n(Z_1)] \to \sigma$ (finite), as $n \to \infty$.

THEOREM 2. *Let assumptions* (A1), (A1)** *and* (A2)** *be satisfied. We assume that a choice of* α *which satisfies* (4) *also satisfies* (11). *Then, as* $n \to \infty$, *the law of*

$$h_n^{1/2}\left\{\left[\sum_{j=1}^{n} L_n^*(Z_j) \Big/ \sum_{j=1}^{n} L_n(Y_j)\right] - [EL_n^*(Z_1)/EL_n(Y_1)]\right\}$$

converges to $N(0, \sigma_0^2 l^{-2})$, *by a theorem in* ([4], p. 254), *provided* $\sigma_0^2 > 0$ *where* $\sigma_0^2 = \lim h_n^{-1} n \sigma^2[\varphi_n(Z_1)]$, *as* $n \to \infty$. *Furthermore,* (A3)**, (A2) *and* (A2)* *form a set of sufficient conditions for* (A2)** *to be true, and therefore under* (A1), (A2), (A2)*, (A1)**, (A3)** *and a choice of* α *satisfying both* (4) *and* (11), *the theorem is true. In this case* $\sigma_0^2 = \sigma_2^2 + v^2\sigma_1^2 + 2v\sigma$.

PROOF. Clearly, for the first part of the theorem there is nothing to be proved. As for the second part, we have to show that (A2), (A2)* and (A3)** imply (A2)**. The uniform boundedness of $\{\varphi_n\}$, $n=1, 2, \cdots$ follows from that of $\{L_n\}$, $\{L_n^*\}$, $n=1, 2, \cdots$ and (A3)** (ii).

Next, $E|\varphi_n(W_1)|^2$ is $O(h_n n^{-1})$ by Minkowski inequality, (A2)(i), (A2)*(i) and (A3)**(ii). We also have

$$E[\Psi_n(W_1)\Psi_n(W_j)]=E[g_n(Z_1)g_n(Z_j)]+v_n^2E[f_n(Y_1)f_n(Y_j)] \\ +v_nE[g_n(Z_1)f_n(Y_j)]+v_nE[f_n(Y_1)g_n(Z_j)]$$

from which it follows that $E[\Psi_n(W_1)\Psi_n(W_j)]$ are $O(h_n^2n^{-2})$ uniformly in j, $1<j\leqq n$, by means of (A2)(ii), (A2)*(ii), the first part of (A3)**(i) and (A3)**(ii). In a similar fashion replacing the Ψ_n's by what are equal to in $E[\Psi_n(W_1)\Psi_n(W_i)\Psi_n(W_j)]$ and using (A2)(iii), (A2)*(iii), the second part of (A3)**(i) and (A3)**(ii), we see that $E[\Psi_n(W_1)\Psi_n(W_i)\Psi_n(W_j)]$ are $O(h_n^3n^{-3})$ uniformly in i and j, $1<i<j\leqq n$. Finally,

$$h_n^{-1}n\sigma^2[\varphi_n(W_1)]=h_n^{-1}n\sigma^2[L_n^*(Z_1)]+v_n^2h_n^{-1}n\sigma^2[L_n(Y_1)] \\ +2v_nh_n^{-1}nE[f_n(Y_1)g_n(Z_1)]$$

and this converges to $\sigma_2^2+v^2\sigma_1^2+2v\sigma$ as $n\to\infty$ by (A2)(iv), (A2)*(iv), (A3)**(ii) and (A3)**(iii). This completes the proof of the theorem.

UNIVERSITY OF WISCONSIN, MADISON

REFERENCES

[1] P. Billingsley, *Statistical Inference for Markov Processes*, University of Chicago Press, 1961.
[2] S. Bochner, *Harmonic Analysis and the Theory of Probability*, University of California Press, 1955.
[3] Theo. Cacoullos, "Estimation of a multivariate density," *Ann. Inst. Statist. Math.*, 18 (1964), 179-189.
[4] H. Cramér, *Mathematical Methods of Statistics*, Princeton University Press, 1946.
[5] J. Doob, *Stochastic Processes*, Wiley, New York, 1953.
[6] U. Grenander, "Some direct estimates of the mode," *Ann. Math. Statist.*, 36 (1965), 131-138.
[7] L. LeCam, "Locally asymptotically normal families of distributions," *University of California Publications in Statistics*, 3 (1960), 37-98.
[8] M. Loève, *Probability Theory*, 3rd ed., Van Nostrand, N. J., 1963.
[9] E. Parzen, "On estimation of a probability density function and mode," *Ann. Math. Statist.*, 33 (1962), 1065-1076.
[10] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Ann. Math. Statist.*, 27 (1956), 832-837.
[11] G. Roussas, "Asymptotic inference in Markov processes," *Ann. Math. Statist.*, 36 (1965), 978-992.
[12] G. Roussas, "Extension to Markov Processes of a result by Wald about the consistency of the maximum likelihood estimate," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 4 (1965), 69-73.
[13] G. S. Watson and M. R. Leadbetter, "On the estimation of the probability density, I," *Ann. Math. Statist.*, 34 (1963), 480-491.
[14] P. Whittle, "On the smoothing of probability density functions," *J. R. Statist. Soc.*, Ser. B, 20 (1958), 334-343.