BERRY-ESSEEN TYPE BOUND FOR DENSITY ESTIMATORS OF STATIONARY MARKOV PROCESSES

By

B. L. S. Prakasa Rao*

(Received November 1, 1975)

Abstract

Berry-Esseen type bound for the distribution function of a density estimator of kernel type is obtained when the observations are from a stationary Markov process. It is shown that the bound is of the order $0(n^{-1/3+\tau})$ for any $\tau>0$ under some conditions.

1. Introduction

Estimators of the density function of a population based on a sample of independent observations have been considered by several authors. An excellent survey of the results in this area is given in Rosenblatt (1971). Rosenblatt (1970) and Roussas (1967, 1969) considered kernel type of density estimators when the observations are assumed to be sampled from a stationary Markov process. Rosenblatt (1970) has shown that these estimators have the same character as in the case of independence. He proved that the kernel type estimators are asymptotically normal under some regularity conditions. Our aim in this paper is to obtain Berry-Esseen type bound for the density estimator after proper norming. Result obtained here generalizes a similar result of Wertz (1971) in the independent case. Order of the bound obtained here is the same as in the independent case.

2. Preliminaries

Consider a probability space $(R, \mathscr{B}, P)$ and let $\{X_n, n\geq 1\}$ be a Markov process taking values in $(R, \mathscr{B}, P)$ with stationary transition measure $p(\xi, A)=P(X_{n+1}\in A | X_n=\xi)$. Assume that $p(\xi, A)$ is a measurable fuuction of ξ for fixed A and a probability measure on $\mathscr{B}$ for fixed ξ. Such a transition measure gives rise to a Markov process by Doob (1953). Assume that the process $\{X_n, n\geq 1\}$ satisfies a Doeblin's condition (D_0) as given by Doob (1953), p. 221 viz. there is a finite-valued measure λ on $\mathscr{B}$ with $\lambda(R)>0$, an integer $\nu\geq 1$ and $\varepsilon>0$ such that

AMS (1970) subject classification: Primary 62G05, 62M05. Key words and phrases: Density Estimator, Berry-Esseen type bound, Stationary Markov process.

* Centre de Recherches Mathématiques, Université de Montréal.

$$(2.1)\qquad p^{(\nu)}(\xi, A)\leq 1-\varepsilon \text{ if } \lambda(A)\leq\varepsilon$$

and there is only one ergodic set E in R with $\lambda(E)>0$ and this set contains no cyclically moving subsets. (Here $p^{(n)}(\cdot,\cdot)$ is the n-step transition measure.) Under (D_0), it can be shown that there exist positive constants $r_f\geq 1$, $0<\rho_f<1$ and a unique stationary probability distribution $\Pi(\cdot)$ such that

$$(2.2)\qquad |p^{(n)}(\xi, E)-\Pi(E)|\leq r_f\rho_f^n$$

for $n\geq 1$. The distribution $\Pi(\cdot)$ taken as the initial distribution together with the stochastic transition function determines a stationary Markov Process. We shall assume that the initial distribution is always the stationary distribution. Suppose that $p(\cdot,\cdot)$ and $\Pi(\cdot)$ are absolutely continuous with respect to a σ-finite measure μ. Let $f(\cdot,\cdot)$ and $f(\cdot)$ be densities of $p(\cdot,\cdot)$ and $\Pi(\cdot)$ with respect to μ. The problem which we consider in this paper is to find the rate of convergence of distribution of kernel type density estimators of $f(\cdot)$. Let P_f be the probability measure on $(R^\infty, \mathcal{B}^\infty)$ correspondıng to $f(\cdot,\cdot)$ and $f(\cdot)$. Here after, we shall suppose that R is the real line and $\mathcal{B}$ the σ-field of Borel sets of R, μ Lebesgue measure.

We shall now state a Berry-Esseen type theorem for random variables defined on a Markov process satisfying the above conditions.

THEOREM 2.1. *Let* $Z_{jn}=a_nM(b_nX_j+c_n)$, $1\leq j\leq n$, $n\geq 1$ *be a double sequence of random variables defined on a stationary Markov process* $\{X_n, n\geq 1\}$ *as given above where* $\{a_n\}$, $\{b_n\}$ *and* $\{c_n\}$ *are sequences of non-negative constants such that* $a_n\to 0$ *and* $M(\cdot)$ *is a bounded real valued function. Further suppose that* $\mathrm{E}(Z_{jn})=0$ *for all* j *and* n. *Define*

$$(2.3)\qquad \beta_{in}=\mathrm{E}|Z_{1n}|^i,\ \mathrm{n}\geq 1$$

$$(2.4)\qquad \sigma_n^2=n^{-1}\,\mathrm{Var}\Big[\sum_{j=1}^n Z_{jn}\Big].$$

Suppose that $\sigma_n^2>0$. *Let*

$$(2.5)\qquad S_n=n^{-1/2}\sigma_n^{-1}\sum_{j=1}^n Z_{jn}.$$

Then there exist absolute constants C_i, $i=1,2,3$ *such that for every* n,

$$(2.6)\qquad \sup_{-\infty<u<+\infty}|P(S_n\leq u)-\Phi(u)|$$
$$\leqq C_1n^{-1/2}\frac{\beta_{3n}}{\sigma_n^3}\Big\{1+\Big[\frac{r_f^2}{1-\rho_f}\Big]^2\Big\}$$
$$+C_2n^{-1/2}\frac{\beta_{1n}}{\sigma_n}\Big[\frac{r_f^2}{1-\rho_f}\Big]$$
$$+C_3\Big[\frac{\rho+\sqrt{2r_f}}{1+\sqrt{2r_f}}\Big]^n\frac{\beta_{1n}\sigma_n^2}{\beta_{3n}}\,\frac{r_f^2}{(1-\rho_f)^2}$$

where Φ *is the standard normal distribution.*

Lemma 3.4 of Prakasa Rao (1973) is similar to this theorem and it is based on a result of Aleskevicus (1966). We omit the oroof this result. We shall now state a lemma due to Wertz (1971) which will be used in the sequel.

LEMMA 2.1. *Let* a, b *and* c *be positive real numbers,* $(\Omega_n, \mathcal{B}_n, P_n)$, $n\geq 1$ *be a sequence of probability spaces, and suppose that* $\{\zeta_n\}$ *and* $\{\eta_n\}$ *are sequences of real valued* *random variables defined on* $(\Omega_n, \mathscr{B}_n, P_n)$. *Further suppose that there exists positive constants* C, C' *such that for every* $y \in R$ *and for every* $n \geq 1$,

$$|P_n(\eta_n \leq y) - \Phi(y)| \leq C\, n^{-c}, \tag{2.7}$$

$$P_n(|\zeta_n| \geq n^{-a}) \leq C'\, n^{-b}. \tag{2.8}$$

Then there exists $C'' > 0$ *such that for every* $y \in R$ *and for every* $n \geq 1$

$$|P_n(\zeta_n + \eta_n \leq y) - \Phi(y)| \leq C''\, n^{-min(a,b,c)}. \tag{2.9}$$

We shall state few more lemmas which give upper and lower bounds for the variance of sums of random variables $f(X_j)$ when $\{X_j\}$ is a stationary Markov process satisfying Doeblin's condition (D_0).

LEMMA 2.2. (Doob (1953), p. 222). *Let* $g(\cdot)$ *be* $\mathscr{B}$*-measurable with* $\mathrm{E}[g(X_1)]^2 < \infty$. *Then*

$$\begin{aligned} &|\mathrm{E}(g(X_1)g(X_{k+1})) - \mathrm{E}(g(X_1))\mathrm{E}(g(X_{k+1}))| \\ &\qquad \leq 2r_f^{1/2}\rho_f^{k/2}\mathrm{E}[g(X_1)]^2. \end{aligned} \tag{2.10}$$

LEMMA 2.3. (Prakasa Rao (1973), p. 144). *Let* $g(\cdot)$ *be* $\mathscr{B}$*-measurable with* $\sigma^2 = E[g(X_1)]^2 < \infty$. *Then*

$$\mathrm{Var}\Big[\sum_{i=1}^{n} g(X_i)\Big] \leq n\sigma^2\{1 + 4r_f^{1/2}(1-\rho_f^{1/2})^{-1}\}. \tag{2.11}$$

LEMMA 2.4. *Let* $g(\cdot)$ *be* $\mathscr{B}$*-measurable with* $\mathrm{E}[g(X_1)] = 0$, $\sigma^2 = \mathrm{E}[g(X_1)]^2 < \infty$. *Then*

$$\mathrm{Var}\Big[\sum_{i=1}^{n} g(X_i)\Big] \geq n\sigma^2\{1 - \rho_f^{1/2}(1 + 4r_f^{1/2})\}(1-\rho_f^{1/2})^{-1}. \tag{2.12}$$

PROOF: By the stationarity of the process $\{X_j\}$ and the fact that $\mathrm{E}[g(X_j)] = 0$,

$$\begin{aligned} \mathrm{Var}\Big[\sum_{i=1}^{n} g(X_i)\Big] &= \mathrm{E}\Big[\sum_{i=1}^{n} g(X_i)\Big]^2 \\ &= n\sigma^2 + 2\sum_{k=1}^{n-1}(n-k)\mathrm{E}[g(X_1)g(X_{k+1})] \\ &\geq n\sigma^2 - 4r_f^{1/2}\sigma^2\sum_{k=1}^{n-1}(n-k)\rho_f^{k/2} \quad \text{(By Lemma 2.2)} \\ &= n\sigma^2\Big\{1 - 4r_f^{1/2}\sum_{k=1}^{n-1}\frac{n-k}{n}\rho_f^{k/2}\Big\} \\ &\geq n\sigma^2\Big\{1 - 4r_f^{1/2}\sum_{k=1}^{n-1}\rho_f^{k/2}\Big\} \\ &= n\sigma^2\Big\{1 - 4r_f^{1/2}\frac{\rho_f^{1/2}(1-\rho_f^{\frac{n-1}{2}})}{1-\rho_f^{1/2}}\Big\} \\ &\geq n\sigma^2\Big\{1 - \frac{4r_f^{1/2}\rho_f^{1/2}}{1-\rho_f^{1/2}}\Big\} \\ &= n\sigma^2\{1 - \rho_f^{1/2}(1 + 4r_f^{1/2})\}(1-\rho_f^{1/2})^{-1}. \end{aligned}$$

3. Main Result

Let $\{X_n, n\geq 1\}$ be a stationary Markov process satisfying Doeblin's condition (D_0) and other assumptions at the begining of section 2. Let f denote the density of X_j and r_f and ρ_f be the Doeblin's constants in (2.2).

Let I be any interval on R and $0<\varepsilon_1<\varepsilon_2$, $A>0$, $r\geq 1$, $0<\rho, h, q<1$. Let $F_0=F_0(A,\rho,h,q,r;\ I,\varepsilon_1,\varepsilon_2)$ be the class of all stationary Markov processes with marginal density f and the Doeblin's constants r_f and ρ_f satisfynig the following conditions:

(3.1) (i) $0<\varepsilon_1\leq \inf_{x\in I} f(x)\leq \sup_{x\in I} f(x)\leq \varepsilon_2$,

(ii) $f''(x)$ exists for every $x\in R$ and $\sup\{|f''(x)|: x\in R\}\leq A<\infty$,

(iii) $1\leq r_f\leq r<\infty$, $0<\rho_f\leq\rho<1$,

$0\leq(\rho_f+2^{1/2}r_f^{1/2})(1+2^{1/2}r_f^{1/2})^{-1}\leq h<1$, and

$0\leq \rho_f^{1/2}(1+4r_f^{1/2})\leq q<1$.

Let $K(\cdot)$ be any bounded continuous symmetric density function such that $c_{i,k}<\infty$ for $0\leq i\leq 3$, $0\leq k\leq 2$ where

$$c_{i,k}=\int_R t^k K^i(t)\,dt. \tag{3.2}$$

Let $k_n\to\infty$ as $n\to\infty$ such that $nk_n\to\infty$ as $n\to\infty$. Define

$$h_n(X^n;x)=\frac{1}{nk_n}\sum_{i=1}^n k\Big[\frac{x-X_i}{k_n}\Big],\ (x^n,x)\in R^n\times R.$$

Rosenblatt (1970) proved that

$$\sqrt{nk_n}\,[h_n(X^n;x)-\mathrm{E}(h_n(X^n;x))] \tag{3.3}$$

is asymptotically normal with mean 0 and variance

$$f(x)\int_{-\infty}^{\infty}K^2(t)\,dt=f(x)c_{2,0} \tag{3.4}$$

under some conditions. Let $\Delta_n^2(f;x)=\mathrm{Var}(h_n(X^n;x))$.

We shall now state and prove the main theorem of this paper.

THEOREM 3.1. *Let $k_n=n^{-1/3}$. Given any $0<\tau<\frac{1}{3}$, there exists n_0 depending on F_0 such that*

$$\sup_{F_0}\sup_{x\in I}\sup_{n\geqq n_0}\sup_{y\in R} n^{1/3-\tau}\Big|P_f\Big[\frac{h_n(X^n;x)-f(x)}{\Delta_n(f;x)}\leq y\Big]-\Phi(y)\Big|<\infty. \tag{3.5}$$

PROOF: Let

$$\eta_{jn}=\frac{1}{nk_n}K\Big[\frac{x-X_j}{k_n}\Big],\ 1\leq j\leq n. \tag{3.6}$$

Since $\{X_j\}$ is stationary, η_{jn}, $1\leq j\leq n$ are identically distributed. Let

$$\sigma_n^2=\sigma_n^2(f;x)=\mathrm{Var}_f[\eta_{jn}]. \tag{3.7}$$

Lemma 2.3 implies that

$$\mathrm{Var}_f\Big[\sum_{j=1}^n\eta_{jn}\Big]\leq n\sigma_n^2[1+4r_f^{1/2}(1-\rho_f^{1/2})^{-1}]. \tag{3.8}$$

Hence

$$\text{(3.9)} \qquad \mathrm{Var}_f[h_n(X^n;x)] \equiv \Delta_n^2(f;x) = 0(n\sigma_n^2)$$

uniformly in F_0. Let $\nu_{ni}(f;x)=\mathrm{E}_f|\eta_{jn}-E(\eta_{jn})|^i$ and

$$\text{(3.10)} \qquad Z_{jn}=\eta_{jn}-E(\eta_{jn}),\ 1\leqq j\leqq n.$$

Equations (10) and (11) of Wertz (1971) show that there exists an integer N_1 depending only on F_0 such that

$$\text{(3.11)} \qquad n^{-3/2}\sigma_n^{-3}(f;x)=0((nk_n)^{3/2})$$

and

$$\text{(3.12)} \qquad \nu_{n3}(f;x)\leqq \frac{f(x)c_{3,0}}{n^3k_n^2}+\frac{C_4}{n^3k_n}$$

for every $n>N_1$, uniformly for F_0 and $x\in I$ where C_4 is again a constant depending only on F_0. Applying Theorem 2.1, we obtain that there exists absolute constants C_i, $1\leqq i\leqq 3$ such that

$$\text{(3.13)} \qquad \sup_{y\in R}\left|P_f\left[\frac{\sum_{j=1}^n Z_{jn}}{\Delta_n(f;x)}\right]-\Phi(y)\right|$$

$$\leqq C_1 n^{-1/2}\frac{\nu_{n3}(f;x)}{(n^{-1/2}\Delta_n(f;x))^3}\left\{1+\left[\frac{r_f^2}{1-\rho_f}\right]^2\right\}$$

$$+C_2 n^{-1/2}\frac{\nu_{n1}(f;x)}{n^{-1/2}\Delta_n(f;x)}\left[\frac{r_f^2}{1-\rho_f}\right]$$

$$+C_3\frac{\nu_{n1}(f;x)\,(n^{-1/2}\Delta_n(f;x))^2}{\nu_{n3}(f;x)}\left[\frac{\rho_f+\sqrt{2r_f}}{1+\sqrt{2r_f}}\right]^n\frac{r_f^2}{(1-\rho_f)^2}$$

$$\text{(3.14)} \qquad \leqq C_1 n\left\{\frac{f(x)c_{3,0}}{n^3k_n^2}+\frac{C_4}{n^3k_n}\right\}\Delta_n^{-3}(f;x)\,\{1+r^4(1-\rho)^{-2}\}$$

$$+C_2\left\{\frac{f(x)c_{3,0}}{n^3k_n^2}+\frac{C_4}{n^3k_n}\right\}^{1/3}\Delta_n^{-1}(f;x)\,\{r^2(1-\rho)^{-1}\}$$

$$+C_3 n^{-1}\left\{\frac{f(x)c_{3,0}}{n^3k_n^2}+\frac{C_4}{n^3k_n}\right\}^{1/3}\frac{\Delta_n^2(f;x)}{\nu_{n3}(f;x)}h^n r^2(1-\rho)^{-2}.$$

The last inequality follows from the fact that $\nu_{n1}^3\leqq\nu_{n3}$. Equation (13) of Wertz (1971) shows that there exists N_2 depending on F_0 and a constant C_5 depending on F_0 such that

$$\text{(3.15)} \qquad n^{-1/2}\sigma_n^{-1}(f;x)\leqq C_5(nk_n)^{1/2}$$

for all $n>N_2$, uniformly for F_0 and $x\in I$. Lemma 2.4 implies that

$$\text{(3.16)} \qquad \Delta_n^2(f;x)\geqq C_6 n\sigma_n^2(f;x) \geqq C_6C_5(nk_n)^{-1}$$

for some C_6 depending on F_0. (3.12), (3.14) and (3.16) together prove that R.H.S. of (3.14) is less than or equal to

(3.17)
$$C_7(nk_n)^{-1/2} + C_8(nk_n)^{-1/2}k_n^{1/3} + C_9(nk_n)^{-1/2}n^{-3/2}k_n^{-1/6}h^n\Delta_n^2(f;x)\nu_{n3}^{-1}(f;x)$$

where C_7, C_8 and C_9 are constants depending only on F_0. Note that, by Lemma 2.3,

(3.18)
$$\Delta_n^2(f;x)\leqq C_{10}n\sigma_n^2(f;x)$$

where C_{10} depends only on F_0 and

(3.19)
$$\sigma_n^2(f;x)\leqq C_{11}n^{-2}k_n^{-2}.$$

Hence

(3.20)
$$\Delta_n^2(f;x)\leqq C_{10}C_{11}n^{-1}k_n^{-2}.$$

Furthermore

(3.21)
$$\nu_{n3}(f;x)\geqq[\sigma_n^2(f;x)]^{3/2} \geqq[C_5^{-2}n^{-2}k_n^{-1}]^{3/2}=C_{12}n^{-3}k_n^{-3/2}$$

where C_{12} depends only on F_0 by (3.15)· Combining (3.20) and (3.21), it follows that

(3.22)
$$\Delta_n^2(f;x)\nu_{n3}^{-1}(f;x) \leqq C_{10}C_{11}C_{12}^{-1}n^3k_n^{3/2} =C_{13}n^2k_n^{-1/2}$$

wwhere C_{13} depends only on F_0. Observe that the last term in (3.17) is still $0((nk_n)^{-1/2})$ since h^n with $0<h<1$ is involved in it. (3.17) proves that there exists a constant C_{14} depending only on F_0 such that

(3.23)
$$\sup_{y\in R}\left|P_f\left[\frac{\sum_{j=1}^{n}Z_{jn}}{\Delta_n(f;x)}\leqq y\right]-\Phi(y)\right|\leqq C_{14}(nk_n)^{-1/2}$$

uniformly for F_0 and $x\in I$ and $n>N_3$ depending only on F_0. Note that $\sum_{j=1}^{n}Z_{jn}=h_n(X^n;x)-\mathrm{E}_f[h_n(X^n;x)]$. Hence

(3.24)
$$\sup_{y\in R}\left|P_f\left[\frac{h_n(X^n;x)-\mathrm{E}_f[h_n(X^n;x)]}{\Delta_n(f;x)}\leqq y\right]-\Phi(y)\right|\leqq C_{14}(nk_n)^{-1/2}$$

uniformly for F_0, $x\in I$ and $n>N_3$ depending only on F_0.

Let

(3.25)
$$\eta_n=\frac{h_n(X^n;x)-\mathrm{E}_f[h_n(X^n;x)]}{\Delta_n(f;x)}$$

and

(3.26)
$$\zeta_n=\frac{\mathrm{E}_f[h_n(X^n;x)]-f(x)}{\Delta_n(f;x)}.$$

Note that

$$|\mathrm{E}_f[h_n(X^n;x)]-f(x)| =\left|\int_R K(t)f(x+k_nt)\,dt-f(x)\right|$$

$$\leqq 1/2 k_n^2 \|f''\|_\infty c_{1,2}$$

where $\|f''\|_\infty = \sup \{|f''(x)| : x \in R\}$ and

$$\Delta_n(f;x) \geqq (C_6 C_5)^{1/2} (nk_n)^{-1/2}$$

by (3.16). Hence

(3.27) $$|\zeta_n| \leqq C_{15} (nk_n^5)^{1/2}$$

for $n > N_4$ depending on F_0 and for some constant C_{15} depending on F_0. Hence, by Tchebyshev's inequality, for any $a>0$ and $p>0$ and $k_n = n^{-\alpha}$,

(3.28) $$\begin{aligned} P_f(|\zeta_n| \geqq n^{-a}) &\leqq n^{ap} \mathrm{E}_f |\zeta_n|^p \\ &\leqq C_{15}^p n^{ap} (nk_n^5)^{p/2} \\ &= C_{15}^p n^{-p(5\alpha - 2a - 1)/2} \end{aligned}$$

uniformly for F_0 and $x \in I$. (3.24) shows that for $n > N_3$ depending on F_0

(3.29) $$\sup_{F_0} \sup_{x \in I} \sup_{y \in R} |\mathrm{P}_f(\eta_n \leqq y) - \Phi(y)| \leqq C_{14} n^{-(1-\alpha)/2}.$$

(3.28) and (3.29) hold for $n > N$ depending on F_0. Let $a = a$, $b = \frac{p}{2}(5\alpha - 2a - 1)$, $p > \frac{1}{3\tau} - 1$ and $c = (1-\alpha)/2$ and choose α such that $5\alpha - 2a - 1 > 0$ and $0 < \alpha < 1$. In particular, choosing $a = \frac{1}{3} - \tau$ for any $\tau < \frac{1}{3}$, and applying Lemma 2.1, we obtain (3.5).

REMARK: The result can be extended to estimators of two-dimensional density of the stationary Markov process in an analogous manner.

References

[1] ALESKEVICUS, G. J. *Some limit theorems for random variables defined on a homogeneous regular Markov chain.* Litovsk. Math. Sb. 6 (1966), 297–311. (In Russian.)
[2] DOOB, J. L. Stochastic Processes. Wiley, New York, 1953.
[3] PRAKASA RAO, B. L. S. *On the rate of convergence of estimators for Markov processes.* Z. Wahrscheinlichkeitstheorie und Verw. Gebiete. 26 (1973), 141–152.
[4] ROSENBLATT, M. *Density estimates anc Markov sequences.* In Nonparametric Techniques in Statistical Inference (ed. Madan Lal Puri). Cambridge Univ. Press (1970).
[5] ROSENBLATT, M. *Curve Estimates.* Ann. Math. Statist. 42 (1971), 1815–1842.
[6] ROUSSAS, G. G. *Nonparametric estimation in Markov processes.* Technibal Report No. 110. Department of Statistics, Univ. of Wisconsin, Madison, Wisconin (1967).
[7] ROUSSAS, G. G. *Nonparametric estimation of the transition distribution function of a Markov process.* Ann. Math. Statist. 40 (1969), 1386–1400.
[8] WERTZ, W. *Empirische betrachtungen und normalapproximation bei dichteschätzungen.* Operations Research-Verfahren, 8 (1971), 430–448.

B. L. S. Prakasa Rao
Indian Statistical Institute
7, S. J. S. Sansanwal Marg
New Delhi 110029
India