Ann. Inst. Statist. Math.
Vol. 53, No. 3, 447–468 (2001)

ASYMPTOTIC NORMALITY OF KERNEL DENSITY ESTIMATORS UNDER DEPENDENCE

ZUDI LU

Insititut de Statistique, Université Catholique de Louvain, B-1348, Belgium and Laboratory of Management Decision and Information Systems, Institute of Systems Science, Chinese Academy of Sciences, Beijing 100080, China

(Received May 20, 1997; revised September 21, 1999)

Abstract. In this paper, we study the kernel methods for density estimation of stationary samples under generalized conditions, which unify both the linear and α-mixing processes discussed in the literature and also adapt to the non-linear or/and non-α-mixing processes. Under general, mild conditions, the kernel density estimators are shown to be asymptotically normal. Some specific theorems are derived within various contexts, and their applications and relationship with the relevant references are considered. It is interesting that the conditions on the bandwidth may be very simple, even in the generalized context. The stationary sequences discussed cover a large number of (linear or nonlinear) time series and econometric models (such as the ARMA processes with ARCH errors).

Key words and phrases: Asymptotic normality, α-mixing, linear process, kernel density estimators, stable stationary process, time series.

1. Introduction

Density estimation for dependent observations has received extensive attention in the literature. A number of papers have investigated it under various mixing conditions, including Robinson (1983) , Györfi *et al.* (1989), Roussas (1988), Tran (1989) and Irle (1997). Due to the fact that the mixings cannot cover all linear processes, Chanda (1983), Tran (1992) and Hallin and Tran (1996) explored the consistency and asymptotic normality of density estimators for some linear processes. In this paper, we consider kernel density estimators within a more general context which cover most of the time series models in the literature, unify both the stationary mixing and the linear processes and adapt to the nonlinear or/and non-mixing processes in practice.

We assume that $\{X_t\}$ is a stationary sequence defined on a probability space $(\Omega, \mathcal{F}, P)$. As may be known, α-mixing is the weakest among the widely used ϕ-, ρ-, β- and α-mixings. For later reference, its definition is given below.

DEFINITION 0. The stationary sequence $\{X_t, t = 0, \pm 1, \dots\}$ is α-mixing if

$$\alpha(k) = \sup_{A \in \mathcal{F}_{-\infty}^{n}, B \in \mathcal{F}_{n+k}^{\infty}} |P(AB) - P(A)P(B)| \to 0 \tag{1.1}$$

as $k \to \infty$, where $\mathcal{F}_{-\infty}^{n}$ and $\mathcal{F}_{n+k}^{\infty}$ are two σ-fields generated by $\{X_t, t \le n\}$ and $\{X_t, t \ge n + k\}$, respectively. The $\alpha(k)$ is called the mixing coefficient.

Under some suitable conditions, the stationary solutions of many time series (linear and nonlinear) models are α-mixing (cf., for example, Goródetskii (1977), Pham (1986) and Tong (1990)). From a practical point of view, however, the mixing concept may lead to some undesirable features. Firstly, it is not easy to check the α-mixing of the realizations of a stationary process in practice. Secondly, the requirement of the mixing dependence imposed upon all events (past and future) rules out some stationary processes of interests. As emphasized by Tran *et al.* (1996) and Hallin and Tran (1996), the mixing does not cover all linear processes. Andrews (1984) showed that the stationary solution of a simple linear AR(1) model —$X_t = 1/2X_{t-1}+e_t$ with e_t's being independent symmetric Bernoulli random variables (r.v.'s) taking values -1 and 1— is not α-mixing. Chanda (1983), Tran (1992) and Hallin and Tran (1996) thus developed the density estimation for time series under linear processes covering many important time series models (e.g., linear ARMA models and some long-memory fractional processes),

$$X_t = \sum_{r=0}^{\infty} a_r Z_{t-r}, \tag{1.2}$$

where $\{Z_t\}$ are i.i.d. r.v.'s with $EZ_t = 0$ and $EZ_t^2 = \sigma^2 < \infty$. Finally, a function of a mixing process is not readily the corresponding mixing. For example, even if $\{Z_t\}$ in (1.2) is α -mixing, $\{X_t\}$ in (1.2) needs not be α-mixing. In econometrics, the process $\{X_t\}$ in (1.2), with $\{Z_t\}$ being α-mixing, is of great interest. The ARMA process with ARCH errors discussed in Weiss (1984) and Engle (1982) is, for instance, of this form (note that the ARCH model proposed by Engle (1982) is α-mixing under mild conditions, cf., Lu (1996*a*, 1996*b*)). In addition, it is also noticed that the process in (1.2) with $\{Z_t\}$ being α-mixing may include some long-memory fractional processes with ARCH type errors (cf., Lin and Li (1997)).

Due to the above shortcomings, in this paper we first extend the concept of the α-mixing to a more general context in which a nonlinear form includes the linear process in (1.2) with $\{Z_t\}$ being α-mixing. That is,

$$X_t = g(Z_t, Z_{t-1}, Z_{t-2}, \ldots), \tag{1.3}$$

where $g : R^\infty \longrightarrow R^1$ is a Borel measurable function and $\{Z_t\}$ may be vector-valued. The i.i.d. process $\{Z_t\}$ is one of the simplest α-mixing stationary processes. To our knowledge, the idea of extending a mixing process to a function of the entire mixing process goes back to Ibragimov (1962) and was also formalized by Billingsley (1968) and Bierens (1983). We introduce the formal definition as follows:

DEFINITION 1. The stationary process $\{X_t\}$ is stable in L_2 norm (L_2-stable for simplicity) with respect to (hereafter w.r.t.) the stationary α-mixing process $\{Z_t\}$, if

$$v(m) = E|X_t - X_t^{(m)}|^2 \to 0 \tag{1.4}$$

as $m \to \infty$, where $X_t^{(m)} = g_m(Z_t, \ldots, Z_{t-m+1})$, g_m is a Borel function with m arguments involved and $v(m)$ will be called the stable coefficients.

Remark 1. (a) This definition follows Bierens (1983) except that he defined a stable stationary process w.r.t. ϕ-mixing, a special case of the context here.

(b) As done in Bierens (1983), it is usually taken that $X_t^{(m)} = E(X_t \mid Z_t, \ldots, Z_{t-m+1})$. Clearly, $\{X_t^{(m)}\}$ is α-mixing with mixing coefficients $\alpha^*(k) \le 1$ for $k = 0, 1, \ldots, m$, and equals $\alpha(k-m)$ for $k \ge m+1$, where $\alpha(\cdot)$ is the mixing coefficient of $\{Z_t\}$.

(c) Clearly, we may define a L_p-stable process w.r.t. $\{Z_t\}$ with a p-th-order expectation instead of the second order in (1.4). The results in this paper can be easily adapted to such a situation and so we don't pursue this generality here.

(d) If g in (1.3) is of a linear form and $\{Z_t\}$ is i.i.d. as in (1.2), then $\alpha(k) \le 1$ for $k = 0$, and $= 0$ for $k \ge 1$. Further, assuming that $\sum_{r=0}^{\infty} a_r^2 < \infty$, the stable coefficient $v(m) = E|X_t - X_t^{(m)}|^2 = E|\sum_{r=m}^{\infty} a_r Z_{t-r}|^2 = \sigma^2 \sum_{r=m}^{\infty} a_r^2$. This is also considered in Chanda (1983), Tran (1992) and Hallin and Tran (1996). When $\{Z_t\}$ is α-mixing, our model framework covers the ARMA process with ARCH errors discussed in Weiss (1984) and Engle (1982) as well as some long-memory fractionally integrated processes with ARCH type errors.

(e) If g is of a nonlinear form, then many nonlinear time series models fall into our category, e.g., the bilinear and the random coefficient models in references such as Granger and Andersen (1978), Nicholls and Quinn (1982), Tjøstheim (1986) and Tong (1990).

(f) If $\{X_t\}$ itself is α-mixing stationary with $EX_t = 0$ and $EX_t^2 < \infty$, then setting $Z_t = X_t$ and g the identity function leads to the stable coefficient being $v(m) = 0$ for $m \ge 1$.

In Section 2, we define the kernel density estimators based on the realization of the L_2-stable stationary process and give a general result which shows the mild conditions under which the asymptotic normality of the density estimators is ensured. To shed light on the wide applications of this general result, some specific theorems are derived in Sections 3 and 4. In Section 3, we are concerned with the specific theorems under two contexts related with the relevant references, one covering the α-mixing processes considered in Robinson (1983) and the other deducing a result under linear processes better than Hallin and Tran (1996)'s (cf. Remark 6 below). It should be pointed out that the results obtained in Section 2 can also be applied to density estimation for the stationary sequences which are neither α-mixing nor linear processes. In this case, some mild specific conditions are also derived in Section 4 to guarantee the asymptotic normality of the density estimators. To our knowledge, no one has examined this case. The proofs of the theorems in Section 2 are postponed to Section 5.

2. General result

In the following, we let $X_1, X_2, \ldots, X_n$ be a realization of size n from the L_2-stable stationary process $\{X_t\}$ defined in Definition 1. In time-series analysis, estimating a one-dimensional marginal density is of certain interest (cf., Hallin and Tran (1996)), but even more interesting is to consider the estimation of a marginal joint density because a one-dimensional marginal density cannot capture the dependence of the stationary sequence. For background on estimating the marginal joint density, the reader is referred to, for example, Tjøstheim (1996). Generally speaking, we may consider estimating a d-dimensional marginal joint density of $(X_{i-\tau_{d-1}}, \ldots, X_{i-\tau_1}, X_i)'$, say, where $0 < \tau_1 < \cdots < \tau_{d-1} < +\infty$ are positive integers and Y' denotes the transpose of vector Y. For simplicity of notation, we are concerned with $Y_i = (X_{i-d+1}, \ldots, X_{i-1}, X_i)'$ in this paper (the results obtained can be adapted for general τ_k's). Let $f(y)$ be the marginal joint density of Y_i. The kernel density estimator of $f(y)$ is therefore defined by

$$f_n(y) = (nh_n^d)^{-1} \sum_{i=d}^{n} K((y - Y_i)/h_n). \tag{2.1}$$

Here $K(\cdot)$ is a kernel function defined on R^d, and $h_n > 0$ are bandwidths tending to zero as $n \to \infty$.

Our purpose in this paper is to investigate the asymptotic normality of the kernel density estimators for stable stationary sequences. In this section, we first give a general result which will be applied specifically in Sections 3 and 4, and its usefulness will be seen later on. Throughout the paper, C will denote a generic constant which may differ at different places.

ASSUMPTION 1. $\{Z_t\}$ is α-mixing stationary with the mixing coefficients $\alpha(\cdot)$ satisfying

$$k \sum_{j=k}^{\infty} \alpha(j) \to 0 \quad \text{as} \quad k \to \infty. \tag{2.2}$$

ASSUMPTION 2. $\{X_t\}$ is L_2-stable stationary w.r.t. $\{Z_t\}$ with stable coefficients $v(m)$ as defined in Definition 1. Let $Y_t = (X_{t-d+1}, \ldots, X_{t-1}, X_t)'$ have marginal joint density $f(y)$.

ASSUMPTION 3. The kernel function K is a bounded density function with an integrable radial majorant $Q(x)$, that is $Q(x) = \sup\{K(y) : \|y\| \geq \|x\|\}$ is integrable. Assume, in addition, that K satisfies the following Lipschitz condition:

$$|K(x) - K(y)| \leq C\|x - y\|, \tag{2.3}$$

where $\|\cdot\|$ is an Euclidean norm of R^d.

ASSUMPTION 4. (i) The density function f of Y_t is Lipschitz continuous, i.e., for any x, $y \in R^d$,

$$|f(x) - f(y)| \leq C\|x - y\|; \tag{2.4}$$

(ii) The joint density $f_j(x, y)$ of (Y_0, Y_j) is bounded uniformly in j (> 0), that is

$$\sup_j \sup_{(x,y)\in R^d \times R^d} f_j(x, y) \leq C, \tag{2.5}$$

where if $j < d$, (Y_0, Y_j) reads as $(X_{-d+1}, \ldots, X_{j-d+1}, \ldots, X_0, \ldots, X_j) \in R^{j+d}$.

ASSUMPTION 5. The bandwidth h_n satisfies that, as $n \to \infty$,
(i) $h_n \to 0, \quad nh_n^d \to \infty, \quad nh_n^{2d} \to 0$,
(ii) $nh_n^{2+d} \to 0$.

ASSUMPTION 6. There exist two sequences of positive integers, $p(n)$ and $q(n)$, such that as $n \to \infty$,

(i) $p(n) \to \infty, \quad p^3(n)/n \to 0,$
(ii) $q(n) \to \infty, \quad q(n)/p(n) \to 0,$
(iii) $q(n)h_n^d \to 0,$
(iv) $n\alpha(q(n))/p(n) \to 0,$
(v) $v(q(n)) = O(h_n^{3d+2}).$

Remark 2. Among the above assumptions, (i) Assumption 1 is the condition on the mixing dependence by Robinson ((1983), A3.1, p. 189) for density estimation; (ii) Assumptions 3, 4(i) and 5(ii) with $d = 1$ are the same as Assumptions 1, 4 and 5 in Hallin and Tran ((1996), pages 432, 443) respectively; (iii) Assumption 4(ii) on the joint density is often assumed in the literature (cf., Robinson (1983), A4.5, p. 191) and is easily satisfied by time series models (cf., Lemma 5.1 of Hallin and Tran 1996, p. 446, for the linear processes (1.2)). We assume, without loss of generality, that the mixing and the stable coefficients, $\alpha(\cdot)$ and $v(\cdot)$, in Assumptions 1 and 2 are monotonously decreasing.

Remark 3. Assumption 6 is general but looks cumbersome. In Sections 3 and 4 below, we will specify some more explicit and easily verifiable conditions to ensure Assumption 6. From them, the existence of $p(n)$ and $q(n)$ will become clear.

The asymptotic normality of the kernel density estimators was studied by Robinson (1983) under α-mixing and Hallin and Tran (1996) under linear processes. The following theorems extend their contexts to stable stationary processes w.r.t. α-mixing:

Corresponding to Theorem 3.1 of Hallin and Tran ((1996), p. 443), we have

THEOREM 1. *Suppose that Assumptions* 1–3, 4(ii), 5(i) *and* 6 *hold and* $y_1, \dots, y_k$ *are* k *distinct points of* R^d, *then*

$$(nh_n^d)^{1/2}(f_n(y_1) - Ef_n(y_1), \dots, f_n(y_k) - Ef_n(y_k))' \xrightarrow{\mathcal{L}} N(\mathbf{0}, \boldsymbol{C}) \tag{2.6}$$

where $\mathbf{0} = (0, \dots, 0)' \in R^k$, $\boldsymbol{C}$ *is a diagonal matrix with elements* $C_{ii} = f(y_i)\int_{R^d} K^2(u)du$, $i = 1, \dots, k$, *and* "$\to^{\mathcal{L}}$" *denotes the convergence in distribution.*

The next theorem corresponds to Theorem 3.2 of Hallin and Tran (1996, p. 443).

THEOREM 2. *If Assumptions* 1–6 *hold and, in addition,*

$$\int_{R^d} \|u\| K(u)du < \infty, \tag{2.7}$$

then for any k and any distinct points $y_1, \dots, y_k \in R^d$,

$$(nh_n^d)^{1/2}(f_n(y_1) - f(y_1), \dots, f_n(y_k) - f(y_k))' \xrightarrow{\mathcal{L}} N(\mathbf{0}, \boldsymbol{C}). \tag{2.8}$$

3. Specific theorems related with relevant references

In Section 2, we gave a general result in which $p(n)$ and $q(n)$ are not defined explicitly in Assumption 6. Notice that Assumptions 6(iv) and 6(v) are related to the mixing coefficients, $\alpha(\cdot)$, and the stable coefficients, $v(\cdot)$, respectively. Different definitions of $p(n)$ and $q(n)$ will lead to different conditions on the mixing and the stable coefficients. In this section, we shall derive some specific conditions which will cover the contexts considered in Robinson (1983) (α-mixing) and in Hallin and Tran (1996) (linear processes), respectively. An easily verifiable condition within a generalized context will be provided in Section 4.

3.1 *Case* I *for Robinson* (1983)

Robinson (1983) considered the setting where $X_t = Z_t$ with the α-mixing coefficients satisfying Assumption 1 in Section 2. To keep this condition on the mixing coefficients hold, we may choose $p(n)$ and $q(n)$, as done in the proof of Lemma 7.1 of Robinson ((1983), p. 199), that is

$$p(n) \sim n^{1/3}\eta_n, \qquad q(n) \sim n^{1/3}\eta_n^2, \quad r(n) \sim n/(p(n)+2q(n)),$$

where $a_n \sim b_n$ means that $\lim_{n\to\infty} a_n/b_n = 1$, and η_n is a positive sequence such that $p(n)$, $q(n)$ and $r(n)$ are nondecreasing, $\eta_n \to 0$ and

$$\eta_n > \max(n^{-1/24}, \tilde{\varepsilon}_{n^{1/4}}^{1/6}), \qquad \text{where} \quad \tilde{\varepsilon}_n = \sup_{N\geq n} \varepsilon_N, \ \ \varepsilon_N = N\sum_{j=N}^{\infty} \alpha(j).$$

Thus, as Robinson ((1983), p. 199) proved, Assumptions 6(i), 6(ii) and 6(iv) hold; $q(n)h_n^d \sim n^{1/3}\eta_n^2 h_n^d \to 0$ follows from Assumption 5(i) and, hence, Assumption 6(iii) holds.

To verify Assumption 6(v), we impose the following condition on the stable coefficient and the bandwidth:

ASSUMPTION 2'. The stable coefficients, $v(\cdot)$, in Assumption 2 satisfy

$$v(n^{1/4}) = O(h_n^{3d+2}). \tag{3.1}$$

Clearly, if we choose $q(n) \sim n^{1/3}\eta_n^2 > n^{1/3}(n^{-1/24})^2 = n^{1/4}$, then $v(q(n)) \leq v(n^{1/4})$ (recall we assume $v(\cdot)$ is monotonously decreasing). Assumption 6(v) thus follows from Assumption $2'$.

THEOREM 3. *If Assumptions* 1, 2' *and* 3–5 *hold and, in addition,* (2.7) *in Theorem* 2 *is satisfied, then the conclusion of Theorem* 2 *holds.*

If $\{X_t\}$ is α-mixing, then from Remark 1(f) in Section 1, it follows that $v(j) = 0$ for $j \geq 1$. Thus, Assumption $2'$ holds naturally.

COROLLARY 1. *If $\{X_t\}$ is an α-mixing stationary sequence with mixing coefficients, $\alpha(\cdot)$, satisfying* (2.1) *in Assumption* 1 *and, in addition, Assumptions* 3–5 *and* (2.7) (*in Theorem* 2) *all are satisfied, then the conclusion of Theorem* 2 *holds.*

Remark 4. From Corollary 1, Theorem 3 clearly covers the context of α-mixing for density estimation considered in Robinson ((1983), Theorem 4.1, p. 191) with a positive definite matrix-valued bandwidth instead of a scalar bandwidth.

3.2 *Case* II *for Hallin and Tran* (1996)

Hallin and Tran (1996) explored the asymptotic normality of the kernel density estimators under linear processes (1.2). Their results cannot be derived well from Theorem 3 above because Assumption $2'$ does not always hold under their conditions (see Remark

7 below). In this subsection, we slightly strengthen the condition imposed on the mixing coefficients but weaken the condition on the stable coefficients. Thus, the specific theorem obtained will easily produce a better result for linear processes than Hallin and Tran (1996)'s.

ASSUMPTION 1°. The condition on the mixing coefficients, $\alpha(\cdot)$, in Assumption 1 is strengthened to

$$k^a\alpha(k) \to 0 \quad (k \to \infty) \tag{3.2}$$

for some $2 < a < \infty$.

ASSUMPTION 2°. The stable coefficients, $v(\cdot)$, in Assumption 2 satisfy

$$v(n^{3/(1+4a)}) = O(h_n^{3d+2}). \tag{3.3}$$

THEOREM 4. *If Assumptions* 1°, 2° *and* 3–5 *hold and, in addition,* (2.7) *in Theorem* 2 *is satisfied, then the conclusion of Theorem* 2 *holds.*

PROOF. First, take $n/p(n) \sim q^a(n)$, then Assumption 6(iv) follows from (3.2) in Assumption 1°. To make $p^3(n)/n \to 0$ and $q(n)/p(n) \to 0$ in Assumptions 6(i) and 6(ii) hold simultaneously, we let $q(n)/p(n) \sim p^3(n)/n$. Thus, we may take $q(n) \sim n^{3/(1+4a)}$, $p(n) \sim n/q^a(n) \sim n^{(1+a)/(1+4a)}$. By a simple calculation, $q(n)/p(n) \sim p^3(n)/n \sim n^{-(a-2)/(1+4a)} \to 0$ for $2 < a < \infty$ and, hence, Assumptions 6(i) and 6(ii) clearly hold. Assumption 6(v) follows from Assumption 2°. Finally, $q(n)h_n^d \sim (nh_n^{(1+4a)d/3})^{3/(1+4a)} \to 0$ follows easily from Assumption 5(i). Now Theorem 4 is clear from Theorem 2 in Section 2.

Hallin and Tran (1996) considered the estimation of one-dimensional marginal density ($d = 1$ in the setting of this paper) for the linear processes (1.2) with $|a_r| = O(r^{-(4+\delta)})$ for some $\delta > 0$ (as $r \to \infty$) in their Assumption 2.

HT'S ASSUMPTION 2. The coefficients of the linear process X_t (in (1.2)) tend to zero sufficiently fast that $|a_r| = O(r^{-(4+\delta)})$ for some $\delta > 0$ as $r \to \infty$. In addition, Z_t (in (1.2)) has mean zero and finite variance and an absolutely integrable characteristic function.

From Theorem 4, we may easily derive the following corollary:

COROLLARY 2. *If HT's Assumption* 2 *above and our Assumptions* 3, 4(i) *and* 5 *and* (2.7) *in Section* 2 *with* $d = 1$ *hold and, in addition,*

$$\liminf_{n\to\infty} n^{3(7+2\delta)/(1+4a)}h_n^5 > 0 \quad \text{for some} \tag{3.4}$$
$$a \in (2, (37+12\delta)/20) \text{ and some } \delta > 1/4,$$

then the conclusion of Theorem 2 with $d = 1$ *holds.*

PROOF. First, from Remark 1(d) and HT's Assumption 2, it follows that

$$v(m) = \sigma^2 \sum_{r=m}^{\infty} O(r^{-2(4+\delta)}) = O(m^{-(7+2\delta)}) \tag{3.5}$$

as $m \to \infty$ and, hence, together with (3.4),

$$v(n^{3/(1+4a)}) = O(n^{-3(7+2\delta)/(1+4a)}) = O(h_n^5).$$

Thus, Assumption 2^o with $d = 1$ is met. Assumption 1^o holds naturally
i.i.d. Assumption 4(ii) is clear by Lemma 5.1 of Hallin and Tran ((1996),
Remark 2 above). Hence, Corollary 2 follows from Theorem 4.

Remark 5. First, we point out a negligence in Hallin and Tran (1996)'s
5 and Remark 3.1. From the proof of their Theorem 3.2, it is clear that their
5 should be $nh_n^3 \to 0$, not $nh_n^3 \to \infty$, as $n \to \infty$. Hence, the assertion in t
3.1 that "their Assumption 3 (i.e., $nh_n^{(13+2\delta)/(3+2\delta)}(\log\log n)^{-1} \to \infty$ as n –
their Assumption 5 when $\delta < 1$" is not right.

Remark 6. From Remark 5 above and the conditions of Hallin and T
Theorem 3.2, we know that $(13 + 2\delta)/(3 + 2\delta) < 3$, i.e., $\delta > 1$, shoul
to guarantee their Assumptions 3 and 5 simultaneously. However, in our
when $\delta > 1/4$, (3.4) may be met for some a (since $\delta > 1/4$, $2 < (37 + 1$
may be chosen between 2 and $(37 + 12\delta)/20$, and thus $5(1 + 4a)/[3(7 + 2\delta$
Assumption 5 with $d = 1$ and (3.4) may hold simultaneously). Hence, our
improves Theorem 3.2 of Hallin and Tran (1996).

Remark 7. From (3.5), $v(n^{1/4}) = O(n^{-(7+2\delta)/4})$. To make Assump
5 with $d = 1$ hold simultaneously, it is necessary that $4 \times 5/(7 + 2\delta)$
$\delta > 3/2$. Hence, the result following from Theorem 3 cannot always deduc
Tran (1996)'s Theorem 3.2 ($\delta > 1$, cf., Remark 6).

4. Explicit conditions under the generalized context

Insofar as the incremental contributions cover nonlinear processes, t
interest. In this section, under the generalized context, we will derive s
conditions to ensure Assumption 6, thus the results obtained will apply we
for the nonlinear or/and non-mixing processes.

By choosing $p(n)$ and $q(n)$ in (iv) and (v) of Assumption 6 to satisfy

$$n/p(n) = q(n)^a, \qquad h_n^{-(3d+2)} = q(n)^b, \tag{4.1}$$

where a and b are some positive constants to be specified below, we found

LEMMA 1. *The conditions in Assumption* 6' *below are sufficient to ens*
tion 6.

ASSUMPTION 6'. (i) $j^a\alpha(j) \to 0$, as $j \to \infty$, for some $a > 2$;
(ii) $j^b v(j) = O(1)$, as $j \to \infty$, for some $b > (3d + 2)/d$;
(iii) $nh^{(1+a)(3d+2)/b} \to \infty$, $nh^{3a(3d+2)/(2b)} \to 0$, as $n \to \infty$.

PROOF. First, Assumptions 6(iv) and 6(v) are satisfied by (4.1) and
6'(i) and 6'(ii). Next, it follows from (4.1) that $q(n) = h_n^{-(3d+2)/b}$ and $p(n) =$
Thus Assumptions 6(i), 6(ii) and 6(iii) are easily checked by Assumptions 6'(