JOURNAL OF MULTIVARIATE ANALYSIS 11, 1–16 (1981)

Central Limit Theorems under Weak Dependence*

RICHARD C. BRADLEY, JR.†

Department of Mathematical Statistics, Columbia University, New York, New York 10027

Communicated by M. Rosenblatt

This article is motivated by a central limit theorem of Ibragimov for strictly stationary random sequences satisfying a mixing condition based on maximal correlations. Here we show that the mixing condition can be weakened slightly, and construct a class of stationary random sequences covered by the new version of the theorem but not Ibragimov's original version. Ibragimov's theorem is also extended to triangular arrays of random variables, and this is applied to some kernel-type estimates of probability density.

1. INTRODUCTION AND THEOREMS

Let $(\Omega, \mathscr{F}, P)$ be a probability space. For any collection Y of r.v.'s let $\mathscr{B}(Y)$ denote the Borel field generated by Y. For any $\delta > 0$ and any r.v. X with $EX = 0$ and $0 < \operatorname{Var} X < \infty$ let

$$a(\delta, X) = (E\,|X|^{2+\delta})/(\operatorname{Var} X)^{(2+\delta)/2}$$

and if instead $X = 0$ a.s., then define $a(\delta, X) = 1$.

For any two σ-fields $\mathscr{A}$ and $\mathscr{B}$ let

$$\alpha(\mathscr{A}, \mathscr{B}) = \operatorname*{Sup}_{A \in \mathscr{A}, B \in \mathscr{B}} |P(A \cap B) - P(A)\,P(B)|,$$

$$\rho(\mathscr{A}, \mathscr{B}) = \operatorname*{Sup}_{f \in L^2(\mathscr{A}),\, g \in L^2(\mathscr{B})} |\operatorname{Corr}(f, g)|.$$

Received February 19, 1980; revised May 22, 1980.

AMS 1970 subject Classifications: 60F05, 60G10.

Key words and phrases: Strictly stationary, strong mixing, maximal correlation, kernel-type density estimator.

* This work was partially supported by NSF Grant MCS 79-05811.

† Present address: Department of Mathematics, Indiana University, Bloomington, Indiana 47401.

0047-259X/81/010001-16$02.00/0

Let $(X_k,\ k = \dots, -1, 0, 1, \dots)$ be a strictly stationary sequence of random variables on $(\Omega, \mathscr{F}, P)$. For $n = 1, 2, 3, \dots$, let $S_n = X_1 + X_2 + \cdots + X_n$ and let

$$\alpha_n = \alpha(\mathscr{B}(X_k, k \leqslant 0), \qquad \mathscr{B}(X_k, k \geqslant n)),$$
$$\rho_n = \rho(\mathscr{B}(X_k, k \leqslant 0), \qquad \mathscr{B}(X_k, k \geqslant n))$$

Kolmogorov and Rozanov [10] introduced the condition $\rho_n \to 0$ as $n \to \infty$, which is weaker than "φ-mixing" [6] (see [9, Theorem 17.2.3, p. 309]) and stronger than the "strong mixing" [13] condition $\alpha_n \to 0$ as $n \to \infty$. In their article Kolmogorov and Rozanov showed that for stationary Gaussian sequences $\rho_n \to 0$ is equivalent to strong mixing. Ibragimov [8] proved

THEOREM 0 (Ibragimov). *Suppose (X_k) is strictly stationary with $EX_k = 0$ and $0 < \operatorname{Var} X_k < \infty$. (i) If $\underline{\operatorname{Lim}}_{n\to\infty} \operatorname{Var} S_n = \infty$ and $\operatorname{Lim}_{n\to\infty} \rho_n = 0$, then $\operatorname{Var} S_n = nh_n$, where (h_n) is slowly varying in the sense of Karamata. (ii) If in addition $E|X_k|^{2+\delta} < \infty$ for some $\delta > 0$, then $S_n/(\operatorname{Var} S_n)^{1/2} \to N(0,1)$ in distribution as $n \to \infty$.*

Here $N(0, 1)$ denotes the standard normal distribution, with mean 0 and variance 1. The proof of Theorem 0(ii) consists mainly of showing that $a(\delta, S_n)$ is bounded; then Theorem 0(ii) follows from the known central limit theorems involving the strong mixing condition (for example [9, Theorem 18.4.1, p. 334]).

Theorem 0(ii) fails if its hypothesis $E|X_k|^{2+\delta} < \infty$ is omitted; a counterexample is constructed in [1]. Lifshits [11] proved some central limit theorems on Markov chains under $\rho_n \to 0$ and other slightly weaker conditions. In this article we will extend Theorem 0(ii) to triangular arrays of random variables and also show that $\rho_n \to 0$ can be weakened slightly, and then apply these results to some estimates of probability density.

Consider the following array of random variables:

$$\begin{matrix} X_1^{(1)}, & X_2^{(1)}, \dots, & X_{k(1)}^{(1)} \\ X_1^{(2)}, & X_2^{(2)}, \dots, & X_{k(2)}^{(2)} \\ X_1^{(3)}, & X_2^{(3)}, \dots, & X_{k(3)}^{(3)} \\ \vdots & \vdots & \vdots \end{matrix} \tag{1.1}$$

where $k(1), k(2), k(3), \dots$ are positive integers. Suppose that $\forall n, k$ such that $n \geqslant 1$ and $1 \leqslant k \leqslant k(n)$ one has

$$EX_k^{(n)} = 0 \qquad \text{and} \qquad \operatorname{Var} X_k^{(n)} \leqslant 1. \tag{1.2}$$

For $n = 1, 2, 3, \ldots$ and $0 \leqslant J < L \leqslant k(n)$ define

$$S(n, J, L) = \sum_{k=J+1}^{L} X_k^{(n)}.$$

Assume $\operatorname{Lim}_{n\to\infty} k(n) = \infty$, and for $m = 1, 2, 3, \ldots$ let

$$\begin{aligned}
s_m^2 &= \operatorname{Inf} \operatorname{Var} S(n, J, L), \\
&\quad \{(n, J, L)\colon n = 1, 2, 3, \ldots, 0 \leqslant J < J + m \leqslant L \leqslant k(n)\}, \\
\alpha_m &= \operatorname{Sup} \alpha(\mathscr{B}(X_k^{(n)}, 1 \leqslant k \leqslant J), \mathscr{B}(X_k^{(n)}, J + m \leqslant k \leqslant k(n))), \\
&\quad \{(n, J)\colon n = 1, 2, 3, \ldots, 1 \leqslant J < J + m \leqslant k(n)\}, \\
\rho_m &= \operatorname{Sup} \rho(\mathscr{B}(X_k^{(n)}, 1 \leqslant k \leqslant J), \mathscr{B}(X_k^{(n)}, J + m \leqslant k \leqslant k(n))), \\
&\quad \{(n, J)\colon n = 1, 2, 3, \ldots, 1 \leqslant J < J + m \leqslant k(n)\}.
\end{aligned}$$

Let $\rho^* = \operatorname{Lim}_{m\to\infty} \rho_m$ and for $n = 1, 2, 3, \ldots$ and any $\delta > 0$ let

$$a_n(\delta) = \operatorname*{Sup}_{1 \leqslant k \leqslant k(n)} E\,|X_k^{(n)}|^{2+\delta}.$$

For $\delta > 0$ and $0 \leqslant \rho < 1$ define

$$g(\delta, \rho) = \frac{(1 + \rho^{2\delta/(2+\delta)} + 2\rho^{2/(2+\delta)})}{2^{\delta/2}(1 - \rho)^{(2+\delta)/2}}.$$

We start with a technical statement that will be useful in applying some of the other theorems.

THEOREM 1. *For the array* (1.1) *of random variables suppose that* (1.2) *is satisfied, that* $k(n) \to \infty$ *as* $n \to \infty$, *and that for some positive integers* K, L, *and* N *with* $\rho_K < 1/2$ *one has that*

$$\left[\begin{matrix} \operatorname{Inf} \operatorname{Var} S(n, J, J + L) \\ \{(n, J)\colon n \geqslant N, 0 \leqslant J < J + L \leqslant k(n)\} \end{matrix}\right] > \left[\frac{K}{\{2(1 - \rho_K)\}^{1/2} - 1}\right]^2.$$

Then $s_m^2 \to \infty$ *as* $m \to \infty$.

THEOREM 2. *For the array* (1.1) *of random variables suppose that* (1.2) *holds,* $k(n) \to \infty$ *as* $n \to \infty$, $s_m^2 \to \infty$ *as* $m \to \infty$, *and* $\rho^* < 1/2$. *Then for any* ρ, $\rho^* < \rho < 1/2$, *there exist positive constants* C_1 *and* C_2 *and a positive integer* M *such that if* $n \geqslant 1$, $m \geqslant M$, *and* $0 \leqslant j < j + m \leqslant k(n)$, *then*

$$C_1[2(1 - \rho)]^{\operatorname{Log}_2 m} \leqslant \operatorname{Var} S(n, j, j + m) \leqslant C_2[2(1 + \rho)]^{\operatorname{Log}_2 m}.$$

THEOREM 3. *For the array* (1.1) *of random variables suppose that* (1.2) *holds,* $k(n) \to \infty$ *as* $n \to \infty$, $\alpha_m \to 0$, *and* $s_m^2 \to \infty$ *as* $m \to \infty$ *and for some* $0 < \delta \leqslant 1$, *one has that* $g(\delta, \rho^*) < 1$, *that* $E\,|X_k^{(n)}|^{2+\delta} < \infty$ $\forall n, k$, *and that for some* $A > 1$,

$$\lim_{n\to\infty} \left([\operatorname{Log}_2 k(n)] - A(\operatorname{Log}_{[2(1-\rho^*)]}[2(1+\rho^*)]) \times [\operatorname{Log}_{1/g(\delta,\rho^*)} \alpha_n(\delta)]\right) = \infty.$$

Then $S(n, 0, k(n))/(\operatorname{Var} S(n, 0, k(n)))^{1/2} \to N(0, 1)$ *in distribution as* $n \to \infty$.

THEOREM 4. *For the array* (1.1) *of random variables suppose that* (1.2) *holds,* $k(n) \to \infty$ *as* $n \to \infty$, *and for each* n, *the sequence* $X_1^{(n)},..., X_{k(n)}^{(n)}$ *is weakly stationary, i.e.,* $\operatorname{Cov}(X_J^{(n)}, X_L^{(n)})$ *depends only on* $|J-L|$ *and* n. *Suppose* $\alpha_m \to 0$ *and* $s_m^2 \to \infty$ *as* $m \to \infty$, *and for some* $0 < \delta \leqslant 1$ *one has that* $g(\delta, \rho^*) < 1$, $E\,|X_k^{(n)}|^{2+\delta} < \infty$ $\forall n, k$, *and for some* $A > 1$,

$$\lim_{n\to\infty} \left([\operatorname{Log}_2 k(n)] - A\,[\operatorname{Log}_{1/g(\delta,\rho^*)} \alpha_n(\delta)]\right) = \infty.$$

Then $S(n, 0, k(n))/(\operatorname{Var} S(n, 0, k(n)))^{1/2} \to N(0, 1)$ *in distribution as* $n \to \infty$.

Let us return to strictly stationary sequences (X_k) and in this context retain the definition $\rho^* = \operatorname{Lim}_{n\to\infty} \rho_n$.

THEOREM 5. *Suppose* (X_k) *is strictly stationary,* $EX_k = 0$, $\operatorname{Var} X_k < \infty$, $\operatorname{Var} S_n \to \infty$ *and* $\alpha_n \to 0$ *as* $n \to \infty$, *and for some* $0 < \delta \leqslant 1$, $E\,|X_k|^{2+\delta} < \infty$ *and* $g(\delta, \rho^*) < 1$. *Then* $S_n/(\operatorname{Var} S_n)^{1/2} \to N(0, 1)$ *in distribution as* $n \to \infty$.

This is a corollary of Theorem 3 or 4. The condition $g(\delta, \rho^*) < 1$ seems artificial. If it is deleted, then Theorem 5 fails, as Davydov [3, 4] showed; yet for his counterexamples, which are Markov chains, one can easily verify that $\rho_n = 1$ for all n. It seems possible that in Theorem 5, $g(\delta, \rho^*) < 1$ can be weakened to $\rho^* < 1$, and perhaps also $\alpha_n \to 0$ can be deleted; another method of proof would be needed. We will now show that Theorem 5 covers some sequences (X_k) not covered by Theorem 0(ii) or by any central limit theorem in which it is assumed that $\alpha_n \to 0$ at a specific rate.

THEOREM 6. *Suppose* (c_n) *and* (d_n), $n = 1, 2, 3,...$, *are each a non-increasing sequence of numbers such that* $\forall n$, $0 \leqslant 4d_n \leqslant c_n \leqslant 1$. *Suppose* (f_n) *is an arbitrary sequence of positive numbers. Then there exists a strictly stationary random sequence* (X_k) *such that for each* n, $\rho_n = c_n$ *and* $d_n \leqslant \alpha_n \leqslant d_n + f_n$.

In Theorem 6 note that for any n such that $4d_n = c_n$ one automatically has $\alpha_n = d_n$, since the inequality $4\alpha_n \leqslant \rho_n$ always holds.

These theorems will be proved in Section 3.

2. Asymptotic Normality of Some Kernel-Type Density Estimates

Let us take a quick look at some kernel-type estimators of probability density (see for example Rosenblatt [14, 15], Parzen [12], or Woodroofe [17]). We will discuss weak consistency and asymptotic normality under the condition $\rho_n \to 0$.

Let (X_k) be strictly stationary, with marginal probability density f, such that $\rho_n \to 0$ as $n \to \infty$. We will consider the problem of estimating $f(x)$ for some fixed x; without losing generality we assume $x = 0$. Assume f is continuous at 0 and that $f(0) > 0$.

Let $(b_n, n = 1, 2, 3,...)$ be a sequence of positive numbers such that $b_n \to 0$ and $nb_n \to \infty$ as $n \to \infty$. Also let w be a real Borel function with these properties:

(i) w is non-negative.

(ii) $\int_{-\infty}^{\infty} w(u)\, du = 1$.

(iii) For some $u_0 > 0$, $w(u) = 0$ for $|u| > u_0$.

(iv) For some $\delta > 0$, $\int_{-\infty}^{\infty} w^{2+\delta}(u)\, du < \infty$.

In the discussion below we will assume $\delta \leqslant 1$.

For each $n = 1, 2, 3,...$, let

$$f_n(0) = (nb_n)^{-1} \sum_{k=1}^{n} w(X_k/b_n).$$

For any $C > 1$ one has that for all n sufficiently large,

$$C^{-1} b_n f(0) < Ew(X_k/b_n) < Cb_n f(0),$$

$$C^{-1} b_n f(0) \int_{-\infty}^{\infty} w^2(u)\, du < \operatorname{Var} w(X_k/b_n) < Ew^2(X_k/b_n) < Cb_n f(0) \int_{-\infty}^{\infty} w^2(u)\, du,$$

$$C^{-1} b_n f(0) \int_{-\infty}^{\infty} w^{2+\delta}(u)\, du < Ew^{2+\delta}(X_k/b_n) < Cb_n f(0) \int_{-\infty}^{\infty} w^{2+\delta}(u)\, du.$$

Theorem 7. *If $b_n^{-1} \leqslant O(n^\beta)$ for some $0 < \beta < 1$, then (under the above assumptions on (X_k), f, w, and (b_n)) one has that $f_n(0)$ is a consistent estimator of $f(0)$ and $[f_n(0) - Ef_n(0)]/(\operatorname{Var} f_n(0))^{1/2} \to N(0, 1)$ in distribution as $n \to \infty$.*

Proof. For each $n \geqslant 1$ and each k let

$$X_k^{(n)} = (\operatorname{Var} w(X_k/b_n))^{-1/2}\, [w(X_k/b_n) - Ew(X_k/b_n)].$$

(Set $X_k^{(n)} \equiv 0$ if $\operatorname{Var} w(X_k/b_n) = 0$ or ∞, but this can happen for at most finitely many n.) It is easily seen that for each $m \geqslant 1$, $\underline{\operatorname{Lim}}_{n\to\infty} \operatorname{Cov}(X_k^{(n)}, X_{k+m}^{(n)}) \geqslant 0$, and hence for each $m \geqslant 1$, $\underline{\operatorname{Lim}}_{n\to\infty} \operatorname{Var}(X_{k+1}^{(n)} + \cdots + X_{k+m}^{(n)}) \geqslant m$. Defining $k(n) \equiv n$ in the array (1.1), we have that $s_m^2 \to \infty$ as $m \to \infty$ by Theorem 1. Also for some $B > 0$ one has $a(\delta, X_k^{(n)}) \leqslant Bb_n^{-\delta/2}$ for all n, and by the hypothesis of Theorem 7 one has that the hypothesis of Theorem 3 (or of Theorem 4) is fulfilled. The asymptotic normality of $f_n(0)$ now follows from Theorem 3. Also, by Theorem 2, $\operatorname{Var} f_n(0) \to 0$ as $n \to \infty$, and consistency follows easily.

(Another article by the author will contain a more extensive discussion of the question of asymptotic normality and joint asymptotic normality for kernel-type estimators of probability density, under certain conditions on ρ_n.)

3. Proofs

Except for Theorem 6, we will prove the theorems by adapting Ibragimov's proof of Theorem 0. First, for any r.v. X and positive number p let $\|X\|_p \equiv E^{1/p}|X|^p$.

Proof of Theorem 1. Let L_m, $m = 0, 1, 2, \ldots$, be defined by $L_0 = L$ and $L_{m+1} = 2L_m + K$, and let q_m, $m = 0, 1, 2, \ldots$, be defined by

$$q_m = \operatorname{Inf} \operatorname{Var} S(n, J, J + L_m), \qquad \{(n, J): n \geqslant N, 0 \leqslant J < J + L_m \leqslant k(n)\}.$$

Let C be such that $1 < C < [2(1-\rho_K)]^{1/2}$ and $q_0 \geqslant [K/([2(1-\rho_K)]^{1/2} - C)]^2$. If $n \geqslant N$ and $0 \leqslant J < J + L_1 \leqslant k(n)$, and letting $U = S(n, J, J + L)$, $V = S(n, J + L, J + L + K)$, and $W = S(n, J + L + K, J + L_1)$, we have

$$\begin{aligned} \operatorname{Var} S(n, J, J + L_1) &\geqslant [\|U + W\|_2 - \|V\|_2]^2 \\ &\geqslant [(2(1 - \rho_K)\, q_0)^{1/2} - K]^2 \geqslant Cq_0 \end{aligned}$$

and thus $q_1 \geqslant Cq_0$. By a similar argument, $q_{m+1} \geqslant Cq_m$ for each m.

Suppose l is sufficiently large that $(1 - \rho_K)\, q_l > K^2$, and that $n \geqslant N$ and $0 \leqslant J \leqslant J + L_l + K + 1 \leqslant J^* \leqslant k(n)$. Then letting $U = S(n, J, J + L_l)$, $V = S(n, J + L_l, J + L_l + K)$, and $W = S(n, J + L_l + K, J^*)$, we have

$$\operatorname{Var} S(n, J, J^*) \geqslant [\|U + W\|_2 - \|V\|_2]^2 \geqslant [((1 - \rho_K)\, q_l)^{1/2} - K]^2.$$

Hence $s_{L(l)+K+1}^2 \to \infty$ as $l \to \infty$, since $q_l \to \infty$. Theorem 1 is proved.

Proof of Theorem 2. The proof of Theorem 2 will include a couple of extra lemmas that are not really needed for Theorem 2 but are needed for Theorems 3 and 4. In effect we are also beginning the proof of Theorems 3 and 4, but (for now) only assuming the hypothesis of Theorem 2.

Assume $\rho^* < \rho < 1/2$. Let the positive integer K be such that $\rho^* \leqslant \rho_K < \rho$. Let the constants C and C^* be such that

$$\begin{aligned} 0 < C < C^* < 1, \qquad & 2(1-\rho) < (1+C)(1-\rho_K), \\ C(1-\rho)(1+C) > 1, \qquad & 2(1+\rho) > (1+1/C)(1+\rho_K). \end{aligned} \tag{3.1}$$

Let $\varepsilon > 0$ be such that

$$\begin{aligned} &\text{(i)} \quad 1 + 4\varepsilon/(1-2\varepsilon) < C^{*-1/2}, \\ &\text{(ii)} \quad (1-\rho_K)(1+C)(1-\varepsilon)^2 > 1, \\ &\text{(iii)} \quad (1-\rho_K)(1-\varepsilon)^4 (1+C) > 2(1-\rho), \\ &\text{(iv)} \quad (1+\rho_K)(1+\varepsilon)^4 (1+1/C) < 2(1+\rho), \\ &\text{(v)} \quad C < (1-\varepsilon)^2 C^* \text{ and } (1+\varepsilon)^2/C^* < 1/C. \end{aligned} \tag{3.2}$$

Let the positive integers $L < L^*$ be such that

$$s_L^2 > K^2/(C\varepsilon^2), \qquad s_{L^*}^2 > 4L^2. \tag{3.3}$$

For $n \geqslant 1$ and $0 \leqslant J \leqslant k(n)$ define $S(n, J, J) \equiv 0$. For $n \geqslant 1$ and $0 \leqslant j \leqslant l \leqslant k(n)$ let $q(n, j, l) \equiv$ greatest integer q, $j \leqslant q \leqslant l$, such that $\operatorname{Var} S(n, j, q) \leqslant \operatorname{Var} S(n, q, l)$.

LEMMA 1. *Suppose* $n \geqslant 1$ *and* $0 \leqslant j < j + L \leqslant l \leqslant k(n)$. *Then letting* $q \equiv q(n, j, l)$ *one has*

$$C^* < (\operatorname{Var} S(n, j, q))/(\operatorname{Var} S(n, q, l)) < 1/C^*.$$

Proof. The second inequality is trivial. To prove the first, let $U = S(n, j, q)$, $V = X_{q+1}^{(n)}$, and $W = S(n, q+1, l)$. Then $\|V\|_2 \leqslant 1 < \varepsilon \|U + V + W\|_2$ by (1.2) and (3.3), and also $\operatorname{Var}(U + V) > \operatorname{Var} W$. Hence

$$\begin{gathered} \|U + V + W\|_2 \leqslant 2[\|U\|_2 + \|V\|_2], \\ \|U\|_2 \geqslant (1/2)[\|U + V + W\|_2 - 2\|V\|_2] > ((1-2\varepsilon)/2)\|U + V + W\|_2, \end{gathered}$$

$$\begin{aligned} \|V + W\|_2/\|U\|_2 &\leqslant (2\|V\|_2 + \|U\|_2)/\|U\|_2 \\ &\leqslant 1 + 4(1-2\varepsilon)^{-1}\|V\|_2/\|U + V + W\|_2 < C^{*-1/2} \end{aligned}$$

by (3.2)(i), and Lemma 1 is proved.

LEMMA 2. *Suppose* $n \geqslant 1$ *and* $0 \leqslant j < k < l \leqslant k(n)$ *with* $k - j \geqslant L$ *and* $l - k \geqslant L$. *Letting* $D_1 \equiv \min\{\operatorname{Var} S(n, j, k), \operatorname{Var} S(n, k, l)\}$ *and* $D_2 \equiv \max\{\operatorname{Var} S(n, j, k), \operatorname{Var} S(n, k, l)\}$ we have that $2(1-\rho) D_1 < \operatorname{Var} S(n, j, l) < 2(1+\rho) D_2$. *If also* $C < D_1/D_2$, *then* $2(1-\rho) D_2 < \operatorname{Var} S(n, j, l) < 2(1+\rho) D_1$.

Proof. Let $U = S(n, j, k-K)$, $V = S(n, k-K, k)$, and $W = S(n, k, l)$. Then $\|V\|_2 \leqslant \varepsilon \|U+V\|_2$ by (1.2) and (3.3), and by Minkowski's inequality and (3.2)(ii),

$$\begin{aligned}
\operatorname{Var}(U+W) &\leqslant (1+\rho_K)(\operatorname{Var} U + \operatorname{Var} W) \\
&\leqslant (1+\rho_K)(1+\varepsilon)^2 [\operatorname{Var}(U+V) + \operatorname{Var} W], \\
\operatorname{Var}(U+W) &\geqslant (1-\rho_K)(\operatorname{Var} U + \operatorname{Var} W) \\
&\geqslant (1-\rho_K)(1-\varepsilon)^2 [\operatorname{Var}(U+V) + \operatorname{Var} W] \\
&\geqslant 2(1-\rho_K)(1-\varepsilon)^2 s_L^2 \geqslant s_L^2.
\end{aligned}$$

Hence $\|V\|_2 < \varepsilon \|U+W\|_2$ and

$$\begin{aligned}
\operatorname{Var}(U+V+W) &\leqslant (1+\varepsilon)^2 \operatorname{Var}(U+W) \\
&\leqslant (1+\rho_K)(1+\varepsilon)^4 [\operatorname{Var}(U+V) + \operatorname{Var} W], \\
\operatorname{Var}(U+V+W) &\geqslant (1-\varepsilon)^2 \operatorname{Var}(U+W) \\
&\geqslant (1-\rho_K)(1-\varepsilon)^4 [\operatorname{Var}(U+V) + \operatorname{Var} W].
\end{aligned}$$

Now both parts of Lemma 2 can be deduced from (3.2)(iii–iv).

Theorem 2 now follows from an induction argument using the first part of Lemma 2.

Proof of Theorem 3. For any $d > 0$, $0 \leqslant r < 1$, $R > 0$, and $c > 0$, define

$$h(d, r, R, c) = \frac{2(1 + r^{2d/(2+d)} + 2r^{2/(2+d)} + 3R^{-2d/(2+d)})}{[(1-r)(1+c)]^{(2+d)/2}}.$$

The condition $g(\delta, \rho^*) < 1$ implies $\rho^* < 1/2$. Let γ and ρ be such that

$$\begin{gathered}
0 < \gamma < 1, \qquad \rho^* < \rho < 1/2, \\
g(\delta, \rho^*) < g(\delta, \rho) < \gamma, \\
(\operatorname{Log}_{[2(1-\rho)]}[2(1+\rho)]) \operatorname{Log}_\gamma g(\delta, \rho^*) < A \operatorname{Log}_{[2(1-\rho^*)]}[2(1+\rho^*)],
\end{gathered} \tag{3.4}$$

where A is as in Theorem 3.

Let the positive integer K be such that $\rho^* \leqslant \rho_K < \rho$, as in the proof of Theorem 2. Let $R > 0$ be such that $h(\delta, \rho, R, 1) < \gamma$. Let the constants C and C^* satisfy (3.1) and the condition $h(\delta, \rho, R, C) < \gamma$. Let $\varepsilon > 0$ satisfy (3.2) and the condition

$$[\varepsilon + h(\delta, \rho, R, C)^{1/(2+\delta)} (1+\varepsilon)/(1-\varepsilon)]/(1-\varepsilon) < \gamma^{1/(2+\delta)} \tag{3.5}$$

and let the positive integers $L < L^*$ satisfy (3.3). Let the sequences R_n and J_n, $n = 1, 2, 3, \ldots$, be defined by

$$R_n = \max\{R, L^{*2+\delta} a_n(\delta)/\varepsilon^{2+\delta}\}, \tag{3.6}$$

$$J_n = \text{least positive integer such that } R_n \gamma^{J(n)} \leqslant R.$$

We borrow the notations and Lemmas 1 and 2 from the proof of Theorem 2.

LEMMA 3. *Suppose* $n \geqslant 1$ *and* $0 \leqslant j < k < l \leqslant k(n)$ *with* $k - j \geqslant L$ *and* $l - k \geqslant L$. *Suppose* $R^* \geqslant R$ *and that*

$$a(\delta, S(n, j, k)) \leqslant R^*,$$

$$a(\delta, S(n, k, l)) \leqslant R^*,$$

$$C^* < (\operatorname{Var} S(n, j, k))/(\operatorname{Var} S(n, k, l)) < 1/C^*,$$

$$[\operatorname{Var} S(n, j, k)]^{(2+\delta)/2} R^* \geqslant R_n.$$

Then

$$[\operatorname{Var} S(n, j, l)]^{(2+\delta)/2} R^* \gamma \geqslant R_n,$$

$$a(\delta, S(n, j, l)) \leqslant R^* \gamma.$$

Proof. Let $U = S(n, j, k - K)$, $V = S(n, k - K, k)$, and $W = S(n, k, l)$. By (3.4) and Lemma 2,

$$[\operatorname{Var}(U + V + W)]/[\operatorname{Var}(U + V)] > 2(1 - \rho) > 2^{\delta/2}(1 - \rho)^{(2+\delta)/2}$$
$$> 1/g(\delta, \rho) > 1/\gamma$$

and we have the first conclusion of Lemma 3.

Let $R^{**} \equiv R^*(1 + \varepsilon)^{2+\delta}/(1 - \varepsilon)^{2+\delta}$. Now $\|V\|_2 \leqslant \varepsilon \|U + V\|_2$ by (1.2) and (3.3), and

$$\|U\|_2 \geqslant (1 - \varepsilon)\|U + V\|_2, \qquad C < (\operatorname{Var} U)/(\operatorname{Var} W) < 1/C \tag{3.7}$$

by Minkowski and (3.2)(v). Since

$$\|V\|_{2+\delta} \leqslant K a_n(\delta)^{1/(2+\delta)} < \varepsilon R_n^{1/(2+\delta)} \leqslant \varepsilon \|U + V\|_2 R^{*1/(2+\delta)}, \tag{3.8}$$

we have $\|U\|_{2+\delta} \leqslant \|U + V\|_2 (1 + \varepsilon) R^{*1/(2+\delta)}$ by Minkowski and the hypothesis of Lemma 3, and hence $a(\delta, U) \leqslant R^{**}$. Now

$$E|U + W|^{2+\delta} \leqslant E(|U| + |W|)^2 (|U|^\delta + |W|^\delta) \tag{3.9}$$
$$= E|U|^{2+\delta} + E|W|^{2+\delta} + EU^2|W|^\delta + E|U|^\delta W^2$$
$$+ 2E|U|^{1+\delta}|W| + 2E|W|^{1+\delta}|U|.$$

We will use Holder's and Minkowki's inequalities to get bounds on these terms. First let $D = \max\{EU^2, EW^2\}$.

Then

$$\begin{aligned}
EU^2|W|^\delta &\leqslant [E|U|^{2+\delta}]^{(2-\delta)/(2+\delta)} \\
&\quad \cdot \begin{bmatrix} E|U|^{(2+\delta)/2}|W|^{(2+\delta)/2} - E|U|^{(2+\delta)/2}E|W|^{(2+\delta)/2} \\ + E|U|^{(2+\delta)/2}E|W|^{(2+\delta)/2} \end{bmatrix}^{2\delta/(2+\delta)} \\
&\leqslant [R^{**}(EU^2)^{(2+\delta)/2}]^{(2-\delta)/(2+\delta)} \\
&\quad \cdot [\rho_K(E|U|^{2+\delta}E|W|^{2+\delta})^{1/2} + (\|U\|_2\|W\|_2)^{(2+\delta)/2}]^{2\delta/(2+\delta)} \\
&\leqslant R^{**(2-\delta)/(2+\delta)}D^{(2-\delta)/2}[\rho_K^{2\delta/(2+\delta)}R^{**2\delta/(2+\delta)}D^\delta + D^\delta] \\
&= R^{**}D^{(2+\delta)/2}[\rho_K^{2\delta/(2+\delta)} + R^{**-2\delta/(2+\delta)}]
\end{aligned}$$

and

$$\begin{aligned}
E|U|^{1+\delta}|W| &\leqslant [E|U|^{2+\delta}]^{\delta/(2+\delta)} \\
&\quad \cdot \begin{bmatrix} E|U|^{(2+\delta)/2}|W|^{(2+\delta)/2} - E|U|^{(2+\delta)/2}E|W|^{(2+\delta)/2} \\ + E|U|^{(2+\delta)/2}E|W|^{(2+\delta)/2} \end{bmatrix}^{2/(2+\delta)} \\
&\leqslant [R^{**}(EU^2)^{(2+\delta)/2}]^{\delta/(2+\delta)} \\
&\quad \cdot [\rho_K(E|U|^{2+\delta}E|W|^{2+\delta})^{1/2} + (\|U\|_2\|W\|_2)^{(2+\delta)/2}]^{2/(2+\delta)} \\
&\leqslant R^{**\delta/(2+\delta)}D^{\delta/2}[\rho_K^{2/(2+\delta)}R^{**2/(2+\delta)}D + D] \\
&= R^{**}D^{(2+\delta)/2}[\rho_K^{2/(2+\delta)} + R^{**-2/(2+\delta)}].
\end{aligned}$$

Estimating the other terms in (3.9) similarly, we get

$$E|U+W|^{2+\delta} \leqslant 2D^{(2+\delta)/2}R^{**}[1 + \rho_K^{2\delta/(2+\delta)} + 2\rho_K^{2/(2+\delta)} + 3R^{**-2\delta/(2+\delta)}].$$

By (3.7) and the third inequality in (3.1),

$$\operatorname{Var}(U+W) \geqslant (1-\rho_K)(1+C)D \geqslant (1-\rho_K)(1+C)\operatorname{Var} W > \operatorname{Var}(U+V)$$

and we get both $a(\delta, U+W) \leqslant R^{**}h(\delta, \rho_K, R^{**}, C) < R^{**}h(\delta, \rho, R, C)$ and $\operatorname{Var} V < \varepsilon^2 \operatorname{Var}(U+V) < \varepsilon^2 \operatorname{Var}(U+W)$ (by (1.2) and (3.3) again). Hence $\operatorname{Var}(U+V+W) \geqslant (1-\varepsilon)^2 \operatorname{Var}(U+W)$ and also by (3.8),

$$\begin{aligned}
&\|U+V+W\|_{2+\delta} \\
&\quad \leqslant \|U+W\|_{2+\delta} + \|V\|_{2+\delta} \\
&\quad \leqslant \|U+W\|_2\,[R^{**}h(\delta,\rho,R,C)]^{1/(2+\delta)} + \varepsilon\,\|U+V\|_2\,R^{*\,1/(2+\delta)} \\
&\quad \leqslant \|U+W\|_2\,R^{*\,1/(2+\delta)}[\varepsilon + h(\delta,\rho,R,C)^{1/(2+\delta)}(1+\varepsilon)/(1-\varepsilon)]
\end{aligned}$$

and now by (3.5) we have Lemma 3.

LEMMA 4. *Suppose* $n \geqslant 1$, $0 \leqslant j < j + L \leqslant l \leqslant k(n)$, *and* $a(\delta, S(n, j, l)) > R_n$. *Then for some* j^* *and* l^* *these three inequalities are satisfied*:

$$0 \leqslant j^* < j^* + L \leqslant l^* \leqslant k(n),$$
$$l^* - j^* < l - j,$$
$$a(\delta, S(n, j^*, l^*)) > R_n.$$

Proof. Since $l - j \geqslant L$, we have $\operatorname{Var} S(n, j, l) \geqslant 1$ trivially, and $E\,|S(n, j, l)|^{2+\delta} > R_n \geqslant L^{*2+\delta} a_n(\delta)/\varepsilon^{2+\delta}$.

If $l - j < L^*$, then $\|S(n, j, l)\|_{2+\delta} \leqslant L^* a_n(\delta)^{1/(2+\delta)}$ which gives a contradiction. So $l - j \geqslant L^*$.

Let $q = q(n, j, l)$. Then

$$1 \leqslant [\operatorname{Var} S(n, q, l)]/[\operatorname{Var} S(n, j, q)] < 1/C^*$$

by Lemma 1.

If $q - j < L$, then $L^2 \geqslant \operatorname{Var} S(n, j, j + L) > \operatorname{Var} S(n, j + L, l)$ and $\operatorname{Var} S(n, j, l) < 4L^2 < s^2_{L^*}$. So $q - j \geqslant L$.

If $l - q < L$, then $\operatorname{Var} S(n, j, q) \leqslant \operatorname{Var} S(n, q, l) < L^2$ and $\operatorname{Var} S(n, j, l) < 4L^2 \leqslant s^2_{L^*}$. So $l - q \geqslant L$.

If $a(\delta, S(n, j, q)) \leqslant R_n$ and $a(\delta, S(n, q, l)) \leqslant R_n$, then by Lemma 3, $a(\delta, S(n, j, l)) \leqslant R_n$, which contradicts the hypothesis of Lemma 4.

If $a(\delta, S(n, j, q)) > R_n$, then let $j^* = j$ and $l^* = q$. If $a(\delta, S(n, q, l)) > R_n$ we can let $j^* = q$ and $l^* = l$. Either way we get Lemma 4.

LEMMA 5. *Suppose* $n \geqslant 1$, $0 \leqslant j < l \leqslant k(n)$, *and* $\operatorname{Var} S(n, j, l) \geqslant [2(1 + \rho)]^{J(n)} s^2_{L^*}$.

Then $a(\delta, S(n, j, l)) \leqslant R$.

Proof. Let $M = 2^{J(n)}$. We define the integers $j = Q_0 \leqslant Q_1 \leqslant Q_2 \leqslant \cdots \leqslant Q_M = l$ as follows: For each m, $1 \leqslant m \leqslant M - 1$, represent $m = 2^t u$, where t is a non-negative integer and u is an odd integer, and let $Q_m = q(n, Q(2^t(u - 1)),\ Q(2^t(u + 1)))$. (Here $Q(m)$ means Q_m, for typographical convenience.)

Since $\operatorname{Var} S(n, Q_0, Q_m) \geqslant s^2_{L^*}$ we can show, as in the proof of Lemma 4, that $Q_{m/2} - Q_0 \geqslant L$ and $Q_m - Q_{m/2} \geqslant L$. Then by Lemma 2 we have

$$\operatorname{Var} S(n, Q_0, Q_{m/2}) \geqslant [2(1 + \rho)]^{J(n)-1} s^2_{L^*},$$
$$\operatorname{Var} S(n, Q_{m/2}, Q_m) \geqslant [2(1 + \rho)]^{J(n)-1} s^2_{L^*}.$$

Proceeding inductively in this manner we can show for each m, $1 \leqslant m \leqslant M$, that $\operatorname{Var} S(n, Q_{m-1}, Q_m) \geqslant s^2_{L^*}$ and hence $Q_m - Q_{m-1} \geqslant L$, and hence $a(\delta, S(n, Q_{m-1}, Q_m)) \leqslant R_n$, for otherwise repeated applications of Lemma 4 would lead to a contradiction. Using induction with Lemma 3, we can establish Lemma 5.

LEMMA 6. *If* $\varepsilon > 0$, $1 < \mu < 2(1-\rho)$, *and* $M = 2^N$ *for some positive integer* N, *then for all* n *sufficiently large there exist integers* $0 = Q(0) < Q(1) < Q(2) < \cdots < Q(2M) = k(n)$ (*depending on* n) *such that*

$$\alpha_{Q(2m)-Q(2m-1)} < \varepsilon \qquad \text{for} \quad m = 1,\ldots, M,$$

$$\max_{m\,\text{even}} \operatorname{Var} S(n, Q(m-1), Q(m)) < \varepsilon \min_{m\,\text{odd}} \operatorname{Var} S(n, Q(m-1), Q(m)),$$

$$\max_{m\,\text{odd}} \operatorname{Var} S(n, Q(m-1), Q(m)) < \mu^{-N} \operatorname{Var} S(n, 0, k(n)),$$

$$\max_{m\,\text{odd}} a(\delta, S(n, Q(m-1), Q(m))) \leqslant R,$$

$$\max_{m \neq l\,\cdot\,\text{both odd}} |\operatorname{Corr}(S(n, Q(m-1), Q(m)), S(n, Q(l-1), Q(l)))| < \varepsilon.$$

Proof. By Theorem 2, the last inequality in (3.4), and the hypothesis of Theorem 3, $[\operatorname{Var} S(n, 0, k(n))]/[2(1+\rho)]^{J(n)} \to \infty$ as $n \to \infty$. Thus for all n sufficiently large, the following procedure will suffice for defining the integers $Q(m)$:

If $m = 2^t u$, where $1 \leqslant t \leqslant N-1$ and u is odd, then let $Q(m) = q(n, Q(2^t(u-1)), Q(2^t(u+1)))$. Let H be an integer such that $(4+6R)\,\alpha_H^{1-2/(2+\delta)} < \varepsilon$. For m odd, let $Q(m) = Q(m+1) - H$.

By Lemma 2 (with induction), if n is sufficiently large and m even, then $\operatorname{Var} S(n, Q(m), Q(m+2)) \geqslant [2(1+\rho)]^{-N} \operatorname{Var} S(n, 0, k(n))$ and with the aid of [9, Theorem 17.2.2, p. 307] we can verify the conclusions of Lemma 6.

Theorem 3 can now be proved by a standard "big block, small block" argument. In fact, we can simply apply a theorem of Dvoretzky [5, Theorem 5.1, p. 528].

Theorem 4 is proved in the same manner, but with this change of definition:

$$q(n,j,l) = \begin{cases} (l+j)/2 & \text{if} \quad l-j \text{ is even,} \\ (l+j-1)/2 & \text{if} \quad l-j \text{ is odd.} \end{cases}$$

Then Lemmas 1, 2, and 3 hold and instead of Lemmas 4 and 5 we can simply show that if $n \geqslant 1$, $J \geqslant 0$, and $0 \leqslant j < j + 2^J L \leqslant l \leqslant k(n)$, then $a(\delta, S(n,j,l)) \leqslant \max\{R, R_n \gamma^J\}$. Under the hypothesis of Theorem 4, $k(n)/[2^{J(n)}L] \to \infty$ as $n \to \infty$ and we can carry out the argument of Lemma 6 (with the new definition of $q(n,j,l)$) and establish Theorem 4.

It was noted earlier that Theorem 5 is a corollary of Theorem 3 or 4. Before proving Theorem 6 we will need some lemmas and a definition.

LEMMA 7 (Csaki and Fischer [2, Theorem 6.2]). *Suppose* $\mathscr{A}_n$ and $\mathscr{B}_n$, $n = 1, 2, 3,\ldots$, *are σ-fields, and the σ-fields* $(\mathscr{A}_n \vee \mathscr{B}_n)$, $n = 1, 2, 3,\ldots$, *are independent. Then* $\rho(\bigvee_{n=1}^{\infty} \mathscr{A}_n, \bigvee_{n=1}^{\infty} \mathscr{B}_n) = \operatorname{Sup}_n \rho(\mathscr{A}_n, \mathscr{B}_n)$.

Witsenhausen [16, Theorem 1] gives a proof of Lemma 7.

LEMMA 8. *Under the hypothesis of Lemma 7,* $\alpha(\bigvee_{n=1}^{\infty} \mathscr{A}_n, \bigvee_{n=1}^{\infty} \mathscr{B}_n) \leqslant \sum_{n=1}^{\infty} \alpha(\mathscr{A}_n, \mathscr{B}_n)$.

Proof. By induction, it suffices to prove $\alpha(\mathscr{A}_1 \vee \mathscr{A}_2, \mathscr{B}_1 \vee \mathscr{B}_2) \leqslant \alpha(\mathscr{A}_1, \mathscr{B}_1) + \alpha(\mathscr{A}_2, \mathscr{B}_2)$. Suppose $\{A_1,\ldots, A_I\}$ and $\{B_1,\ldots, B_J\}$ are each a partition of Ω, such that each $A_i \in \mathscr{A}_1$ and each $B_j \in \mathscr{B}_1$. Let $C_1,\ldots, C_I$ be arbitrary elements of $\mathscr{A}_2$ and let $D_1,\ldots, D_J$ be arbitrary elements of $\mathscr{B}_2$. Let $A \equiv \bigcup_{i=1}^{I} (A_i \cap C_i)$ and $B \equiv \bigcup_{j=1}^{J} (B_j \cap D_j)$. It suffices to prove $|P(A \cap B) - P(A)\,P(B)| \leqslant \alpha(\mathscr{A}_1, \mathscr{B}_1) + \alpha(\mathscr{A}_2, \mathscr{B}_2)$. Now

$$
\begin{aligned}
&|P(A \cap B) - P(A)\,P(B)| \\
&\quad \leqslant \sum_{i=1}^{I} \sum_{j=1}^{J} |P(C_i \cap D_j) - P(C_i)\,P(D_j)|\,P(A_i \cap B_j) \\
&\qquad + \left| \sum_{i=1}^{I} \sum_{j=1}^{J} [P(A_i \cap B_j) - P(A_i)\,P(B_j)]\,P(C_i)\,P(D_j) \right|
\end{aligned}
$$

and the first term on the right is bounded by $\alpha(\mathscr{A}_2, \mathscr{B}_2)$. We wish to show that the second term on the right is bounded by $\alpha(\mathscr{A}_1, \mathscr{B}_1)$.

For each i and j let $f_{ij} \equiv P(A_i \cap B_j) - P(A_i)\,P(B_j)$, $c_i \equiv P(C_i)$, and $d_j \equiv P(D_j)$. Define the function h: $[0, 1]^{I+J} \to R$ by $h(u_1,\ldots, u_I, v_1,\ldots, v_J) \equiv \sum_{i=1}^{I} \sum_{j=1}^{J} f_{ij} u_i v_j$. Now h is linear in each variable u_i or v_j separately; therefore on $[0, 1]^{I+J}$, h achieves its maximum at some point $(u_1,\ldots, u_I, v_1,\ldots, v_J)$ for which each coordinate u_i or v_j is 0 or 1. With respect to this point let us define

$$
\begin{aligned}
A^* &\equiv \bigcup A_i, \qquad \{i: u_i = 1\}; \\
B^* &\equiv \bigcup B_j, \qquad \{j: v_j = 1\},
\end{aligned}
$$

and then we have

$$
\begin{aligned}
h(c_1,\ldots, c_I, d_1,\ldots, d_J) &\leqslant h(u_1,\ldots, u_I, v_1,\ldots, v_J) \\
&= P(A^* \cap B^*) - P(A^*)\,P(B^*) \leqslant \alpha(\mathscr{A}_1, \mathscr{B}_1).
\end{aligned}
$$

A similar argument yields $h(c_1,\ldots, c_I, d_1,\ldots, d_J) \geqslant -\alpha(\mathscr{A}_1, \mathscr{B}_1)$, and hence $|h(c_1,\ldots, c_I, d_1,\ldots, d_J)| \leqslant \alpha(\mathscr{A}_1, \mathscr{B}_1)$, which is what was needed to complete the proof of Lemma 8.

DEFINITION 1. Suppose N is a positive integer, $0 < q < 1$, and $0 \leqslant r \leqslant 1$. A random sequence $(W_k, k = ..., -1, 0, 1, ...)$ is said to have the "$\mathscr{S}(N, q, r)$-distribution" if the following statements hold:

(i) The σ-fields $\mathscr{B}(W_k, k \equiv J \bmod N)$, $J = 1, 2, ..., N$, are independent.

(ii) For each $J = 1, 2, ..., N$, $(W_{kN+J}, k = ..., -1, 0, 1, ...)$ has the same probability distribution as a strictly stationary Markov chain $(V_k, k = ..., -1, 0, 1, ...)$ with state space $\{1, 2, 3, 4\}$, with invariant marginal probability distribution $\mu_i = P(V_0 = i)$ given by $(\mu_1, \mu_2, \mu_3, \mu_4) = ((1-q)^2, q(1-q), q(1-q), q^2)$, and with one-step transition probability matrix $p_{ij} = P(V_1 = j \mid V_0 = i)$ given by

$$(p_{ij}) = (1-r)\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} [\mu_1 \mu_2 \mu_3 \mu_4] + r \begin{bmatrix} 1-q & q & 0 & 0 \\ 0 & 0 & 1-q & q \\ 1-q & q & 0 & 0 \\ 0 & 0 & 1-q & q \end{bmatrix}.$$

For any strictly stationary (W_k) and any $n \geqslant 1$ let

$$\alpha_n((W_k)) \equiv \alpha(\mathscr{B}(W_k, k \leqslant 0), \mathscr{B}(W_k, k \geqslant n)),$$
$$\rho_n((W_k)) \equiv \rho(\mathscr{B}(W_k, k \leqslant 0), \mathscr{B}(W_k, k \geqslant n)).$$

LEMMA 9. *Suppose N is a positive integer, $0 < q < 1$, and $0 \leqslant r \leqslant 1$. If (W_k) is strictly stationary with the $\mathscr{S}(N, q, r)$-distribution, then*

(i) *(W_k) is N-dependent,*

(ii) $\rho_1((W_k)) = \rho_N((W_k)) = r$,

(iii) $\alpha_1((W_k)) \leqslant Nqr$,

(iv) $\alpha_N((W_k)) = rq(1-q)$.

Proof. Let Y_0, Y_1, Z_0, and Z_1 be identically distributed r.v.'s such that $P(Y_0 = 1) = 1 - P(Y_0 = 0) = q$, the σ-fields $\mathscr{B}(Y_0, Y_1)$, $\mathscr{B}(Z_0)$, $\mathscr{B}(Z_1)$ are independent, and $P(Y_0 = Y_1 = 1) = (1-r)q^2 + rq$. Note that $\alpha(\mathscr{B}(Y_0), \mathscr{B}(Y_1)) = rq(1-q)$ and $\rho(\mathscr{B}(Y_0), \mathscr{B}(Y_1)) = r$. Let $U_0 = 1 + Y_0 + 2Z_0$ and $U_1 = 1 + 2Y_1 + Z_1$. Then by Lemmas 7 and 8, $\alpha(\mathscr{B}(U_0), \mathscr{B}(U_1)) = rq(1-q)$ and $\rho(\mathscr{B}(U_0), \mathscr{B}(U_1)) = r$.

Let (V_k) be the Markov chain in Definition 1, for our given parameters N, q, and r. Then the joint distribution of (V_0, V_1) is the same as that of (U_0, U_1), and hence by the Markov property $\alpha_1((V_k)) = rq(1-q)$ and $\rho_1((V_k)) = r$.

Now (V_k) is 1-dependent, and Lemma 9 follows from Lemmas 7 and 8.

Proof of Theorem 6. We shall assume (f_n) is non-increasing and $f_1 \leqslant 1$. For each $n = 1, 2, 3,\ldots$, let $(X_k^{(n)},\ k = \ldots, -1, 0, 1,\ldots)$ be strictly stationary with the $\mathscr{S}(n, 2^{-n}n^{-1}f_n, c_n)$-distribution, and let $(Y_k^{(n)},\ k = \ldots, -1, 0, 1,\ldots)$ be strictly stationary with the $\mathscr{S}(n, 1/2, 4d_n)$-distribution. Assume that these random sequences are all independent of each other. Let (X_k) be defined by $X_k \equiv \sum_{n=1}^{\infty} 25^{-n}(X_k^{(n)} + 5Y_k^{(n)})$.

For $-\infty \leqslant J \leqslant L \leqslant \infty$ define

$$\mathscr{G}_J^L \equiv \mathscr{B}(X_k^{(n)}, J \leqslant k \leqslant L, n = 1, 2, 3,\ldots),$$
$$\mathscr{H}_J^L \equiv \mathscr{B}(Y_k^{(n)}, J \leqslant k \leqslant L, n = 1, 2, 3,\ldots).$$

For each fixed k, $\mathscr{B}(X_k) = \mathscr{G}_k^k \vee \mathscr{H}_k^k$, and hence for each fixed n, $\rho_n((X_k)) = c_n$ and

$$\begin{aligned} d_n &= \alpha_n((Y_k^{(n)})) \leqslant \alpha_n((X_k)) \\ &\leqslant \alpha(\mathscr{G}_{-\infty}^0, \mathscr{G}_n^\infty) + \alpha(\mathscr{H}_{-\infty}^0, \mathscr{H}_n^\infty) \\ &\leqslant \left[\sum_{m=n}^{\infty} m \cdot (2^{-m}m^{-1}f_m) \cdot c_m\right] + (1/4) \cdot \rho(\mathscr{H}_{-\infty}^0, \mathscr{H}_n^\infty) \\ &\leqslant f_n + d_n. \end{aligned}$$

by Lemmas 7, 8, and 9 and the general inequality $\alpha(\mathscr{A}, \mathscr{B}) \leqslant (1/4)\,\rho(\mathscr{A}, \mathscr{B})$. Theorem 6 is proved.

Acknowledgments

The author is indebted to Professor Murray Rosenblatt for suggesting this topic and for many fruitful conversations. Professor Yue-Pok Mack and the referee made some helpful comments.

References

[1] Bradley, R. C. (1980). A remark on the central limit question for dependent random variables. *J. Appl. Probability* **17** 94–101.

[2] Csaki, P., and Fischer, J. (1963). On the general notion of maximal correlation. *Magyar Tud. Akad. Mat. Kutato Int. Kozl.* **8** 27–51.

[3] Davydov, Yu. A. (1969). On the strong mixing property for Markov chains with a countable number of states. *Soviet Math.* **10** 825–827.

[4] Davydov, Yu. A. (1973). Mixing conditions for Markov chains. *Theor. Probability Appl.* **18** 312–328.

[5] Dvoretzky, A. (1972). Asymptotic normality for sums of dependent random variables. In *Sixth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 513–535.

[6] IBRAGIMOV, I. A. (1959). Some limit theorems for strict-sense stationary stochastic processes (in Russian). *Dokl. Akad. Nauk SSR* **125** 711–714.

[7] IBRAGIMOV, I. A. (1962). Some limit theorems for stationary processes. *Theor. Probability Appl.* **7** 349–382.

[8] IBRAGIMOV, I. A. (1975). A note on the central limit theorem for dependent random variables. *Theor. Probability Appl.* **20** 135–141.

[9] IBRAGIMOV, I. A., AND LINNIK, YU. V. (1971). *Independent and Stationary Sequences of Random Variables.* Wolters, Groningen.

[10] KOLMOGOROV, A. N., AND ROZANOV, YU. A. (1960). On strong mixing conditions for stationary Gaussian processes. *Theor. Probability Appl.* **5** 204–208.

[11] LIFSHITS, B. A. (1978). On the central limit theorem for Markov chains. *Theor. Probability Appl.* **23** 279–296.

[12] PARZEN, E. (1962). On estimation of a probability density and mode. *Ann. Math. Statist.* **33** 1065–1076.

[13] ROSENBLATT, M. (1956). A central limit theorem and a strong mixing condition. *Proc. Nat. Akad. Sci. USA* **42** 43–47.

[14] ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of a density. *Ann. Math. Statist.* **27** 832–835.

[15] ROSENBLATT, M. (1970). Density estimates and Markov sequences. *Nonparametric Techniques in Statistical Inference* (M. L. Puri, Ed.), pp. 199–210. Cambridge Univ. Press, London/New York.

[16] WITSENHAUSEN, H. S. (1975). On sequences of pairs of dependent random variables. *SIAM J. Appl. Math.* **28** 100–113.

[17] WOODROOFE, M. (1970). Discussion on Rosenblatt's paper. *Nonparametric Techniques in Statistical Inference*, (M. L. Puri, Ed.), pp. 211–213. Cambridge Univ. Press, London/New York.