ON THE CENTRAL LIMIT THEOREM FOR ρ-MIXING SEQUENCES OF RANDOM VARIABLES¹

By Magda Peligrad

University of Cincinnati

In this note we establish the central limit theorem for ρ -mixing sequences under combinations of moment assumptions and ρ -mixing rates. These results contain the well-known Ibragimov theorems and answer a problem from recent work of Bradley.

1. Introduction. Mixing sequences of random variables are sequences for which past and distant future are asymptotically independent. In this note we discuss the central limit theorem for ρ -mixing sequences of random variables where the ρ -mixing coefficient is defined to be the maximal coefficient of correlation. The result we obtain bridges the gap between two well-known theorems due to Ibragimov (1975).

First some notation: log denotes the logarithm with base 2 and $\log^+x := \max\{0, \log x\}$. The indicator function of a set A is denoted by I_A . The notation $a \ll b$ means a = O(b). The notation $a \sim b$ means $\lim a/b = 1$. The greatest integer $\leq x$ is denoted [x]. In some places a_n will be written as a(n). The norm in L_p is denoted $\|\cdot\|_p$. N(0,1) denotes the standard normal distribution. For f, $g \in L_2$, we denote by $\operatorname{corr}(f,g) := (Efg - EfEg)/\|f\|_2 \|g\|_2$.

Throughout the paper we suppose that $\{X_k\}_{k\in Z}$ is a strictly stationary sequence of real-valued random variables on a probability space (Ω, F, P) . For $-\infty \leq J \leq L \leq \infty$, let F_J^L denote the σ -field of events generated by the random variables $(X_k, J \leq k \leq L)$. For each natural number $n \geq 1$ define the dependence coefficient

$$\rho(n) := \sup[\operatorname{corr}(f, g)], \qquad f \in L_2(F_{-\infty}^0), g \in L_2(F_n^\infty).$$

The stationary sequence $\{X_k\}$ is said to be " ρ -mixing" [Kolmogorov and Rozanov (1960)] if $\rho(n) \to 0$ as $n \to \infty$.

For each $n \ge 1$ define the partial sum $S_n := X_1 + \cdots + X_n$ and denote by $\sigma_n^2 := \text{var } S_n$. Ibragimov [(1975), Theorems 2.1 and 2.2] proved the following central limit theorem (or CLT for short).

Theorem 0 (Ibragimov). Suppose $\{X_k\}$ is a strictly stationary sequence of random variables such that

(1.1)
$$EX_0 = 0, \quad EX_0^2 < \infty, \quad \sigma_n \to \infty \quad as \quad n \to \infty$$
 and $\rho(n) \to 0$ as $n \to \infty$.

www.jstor.org

Received October 1985; revised May 1986.

¹Supported in part by NSF Grant DMS-8503016.

AMS 1980 subject classifications. Primary 60F05, 60B10.

Key words and phrases. ρ -mixing sequences, central limit theorem, strict stationarity, maximal correlation.

1387

Then $\sigma_n^2 = nh(n)$, where h(n) is slowly varying as $n \to \infty$. Suppose, in addition, at least one of the following two conditions is satisfied:

(i)
$$E|X_0|^{2+\delta} < \infty$$
 for some $\delta > 0$, or

(ii)
$$\sum_{i=1}^{\infty} \rho(2^i) < \infty.$$

Then $S_n/\sigma_n \to N(0,1)$ in distribution as $n \to \infty$.

Bradley (1980) showed that conditions (i) and (ii) cannot be omitted altogether. In a recent paper, Bradley (1987) proved that condition (i) cannot be improved to $Eq(|X_0|) < \infty$, where q(x) is a function such that for every $\delta > 0$, $q(x) = o(x^{2+\delta})$ as $x \to \infty$. On the other hand, he showed that (ii) is in fact the slowest possible rate under which one has the CLT under just the assumption of finite second moment. In order to prove these results Bradley constructed a counterexample, a strictly stationary sequence $\{X_k\}$ satisfying $Eq(|X_0|) < \infty$, where $x^2 \ll q(x) \ll x^{2+\delta}$ as $x \to \infty$ for every $\delta > 0$, with S_n failing to be asymptotically normally distributed and with $\rho(n) \le \tau(n)$. Here $\tau(n)$ is an arbitrary nonincreasing sequence such that $\sum_{n=1}^{\infty} \tau(2^n) = \infty$ and satisfying in connection with the function q the following condition: for some positive number d,

(1.2)
$$q\left(\left\lceil n\exp\left(-d\sum_{i=1}^{\lceil\log n\rceil}\tau(2^i)\right)\right\rceil^{1/2}\right)=o(n)\quad\text{as }n\to\infty.$$

This counterexample suggests that there might be a positive result, a more general CLT than Theorem 0, that bridges the gap between (i) and (ii). Based on the hope that the example is essentially sharp, Bradley conjectured that if $\{X_k\}$ is strictly stationary and satisfies

$$(1.3a) Eq(|X_0|) < \infty$$

and

(1.3b)
$$n \ll q \left(\left[n \exp \left(-d \sum_{i=1}^{\lceil \log n \rceil} \rho(2^i) \right) \right]^{1/2} \right)$$

as $n \to \infty$, for every d > 0, then $\{S_n/\sigma_n\}$ is asymptotically normally distributed as $n \to \infty$.

REMARK 1. It is easy to see that $\exp(d\sum_{i=1}^{\lceil \log n \rceil} \rho(2^i))$ is a slowly varying function when $n \to \infty$, for every real d.

We shall denote $q(x) = x^2 g(x)$, where $g: [0, \infty) \to [0, \infty)$ is supposed to be a nondecreasing function. One can see that (1.3b) implies

(1.4)
$$g(n^{1/2}) \gg \exp\left(d\sum_{i=1}^{\lceil \log n \rceil} \rho(2^i)\right)$$

for every d > 0.

Our result is somehow more precise than that conjectured by Bradley. We shall establish the following.

Theorem 1. Suppose $\{X_k\}$ is a strictly stationary sequence satisfying (1.1) and

$$(1.5a) EX_0^2 g(|X_0|) < \infty$$

and

(1.5b)
$$g(n^{1/2}) \gg \exp\left(2\sum_{i=1}^{\lceil \log n \rceil} \rho(2^i)/(1-\eta)\right)$$
 for some $0 < \eta < 1$.

Then $S_n/\sigma_n \to N(0,1)$ in distribution as $n \to \infty$.

This theorem contains Theorem 0. By taking g(x) = constant for every x > 0 we get the conclusion of Theorem 0 under (ii). By taking $g(x) = x^{\delta}$ with $\delta > 0$ and using Remark 1, we obtain the conclusion of Theorem 0 under (i). By simple computations we get the following corollaries.

COROLLARY 1. Assume $\{X_k\}$ is strictly stationary satisfying (1.1) and for some $0 < \varepsilon < 1$ and c > 0,

$$EX_0^2(\log^+|X_0|)^{2c/(1-\varepsilon)} < \infty$$

and

(1.6b)
$$\rho(n) \le c(\log n)^{-1} \quad \text{for every n sufficiently large.}$$

Then the CLT holds.

REMARK 2. Bradley [(1987), Corollary 1] established that there is a strictly stationary sequence satisfying (1.6a) for which $\rho(n) \ll (\log n)^{-1}$ and the sequence does not satisfy the CLT. This result shows that our Corollary 1 is sharp. Moreover, our corollary specifies that the mentioned counterexample has to satisfy in addition $\limsup_{n} \rho(n) \log n > c$, where c is the constant from (1.6a).

COROLLARY 2. Assume $\{X_k\}$ is strictly stationary satisfying (1.1) and for some $0 < \beta < 1$, $0 < \varepsilon < 1$ and c > 0,

(1.7)
$$EX_0^2 \Big[\exp(2\log^+|X_0|)^{1-\beta} \Big]^{2c/(1-\beta)(1-\epsilon)} < \infty$$

and

$$\rho(n) \leq c(\log n)^{-\beta}$$
 for every n sufficiently large.

Then the CLT holds.

REMARK 3. (1.7) is implied by

(1.8)
$$EX_0^2 \exp(\log^+|X_0|)^{\alpha} < \infty \quad \text{and} \quad \rho(n) \ll (\log n)^{-\beta},$$
 where $\alpha > 0, 0 < \beta < 1 \text{ and } \alpha + \beta > 1.$

Therefore, (1.8) implies that $\{S_n/\sigma_n\}$ is asymptotically normally distributed as $n\to\infty$. This result complements Corollary 2 from Bradley (1987) that states that for each $\alpha>0$ and $\beta>0$ such that $\alpha+\beta\leq 1$, there exists a strictly stationary sequence $\{X_k\}$ satisfying (1.1) such that $EX_0^2\exp(\log^+|X_0|^\alpha)<\infty$ and $\rho(n)\ll(\log n)^{-\beta}$ and the sequence does not satisfy the CLT.

2. Proving Theorem 1. We shall give first three preliminary lemmas followed by the proof of Theorem 1.

LEMMA 1. Suppose $\{X_k\}$ satisfies (1.1). Let $0 < \varepsilon < \varepsilon^* < 1$. Then there exist two positive constants $C_1 = C_1(\{\rho_k\}, \varepsilon)$ and $C_2 = C_2(\{X_k\}, \varepsilon, \varepsilon^*)$ such that for every $n \ge 1$,

(2.1)
$$\sigma_n^2 \le C_1 n E X_0^2 \exp \left(\sum_{i=1}^{\lfloor (1-\varepsilon) \log n \rfloor} \rho(2^i) / (1-\varepsilon) \right)$$

and .

(2.2)
$$\sigma_n^2 \ge C_2 n \exp \left(-\sum_{i=1}^{\left[(1-\varepsilon)\log n \right]} \rho(2^i) / (1-\varepsilon^*) \right).$$

PROOF. Let us notice first that by the proof of Lemma 3.4 of Peligrad (1982), for every $n \ge 1$,

$$ES_n^2 \le 8000 \prod_{i=1}^{\lceil \log n \rceil} (1 + \rho(\lceil 2^{i/3} \rceil)) nEX_0^2.$$

Because $\rho(n) \to 0$, for every $\eta > 0$ we can find a positive constant $C = C(\eta, \{\rho_k\})$ such that

(2.3)
$$ES_n^2 \le C^2 n^{1+\eta} EX_0^2$$
, for every $n \ge 1$.

Denote $S_n(m) = X_{m+1} + \cdots + X_{n+m}$. Let $0 < \eta < 1$. By (2.3),

$$|\sigma_{2m} - ||S_m(0) + S_m(m+p)||_2| \le 2Cp^{(1+\eta)/2}\sigma_1$$

for every integer $m \ge 1$, $p \ge 1$. By the definition of the ρ -mixing coefficient,

$$|E(S_m(0) + S_m(m+p))^2 - 2\sigma_m^2| \le 2\rho(p)\sigma_m^2.$$

Therefore, from the last two inequalities we deduce

(2.4)
$$\sigma_{2m} \le 2^{1/2} (1 + \rho(p))^{1/2} \sigma_m + 2Cp^{(1+\eta)/2} \sigma_1$$

and

(2.5)
$$\sigma_m \le 2^{-1/2} (1 - \rho(p))^{-1/2} (\sigma_{2m} + 2Cp^{(1+\eta)/2} \sigma_1).$$

Let $0 < \varepsilon < 1$. We take in (2.4) and (2.5), $p = \lfloor m^{1-\varepsilon} \rfloor$ and $\eta = \varepsilon/(2-2\varepsilon)$. By (2.4), by recurrence, for every $r > k \ge 0$ we have

$$\sigma(2^r) \leq \prod_{i=k}^{r-1} \left(1 + \rho([2^{i(1-\varepsilon)}])\right)^{1/2} \left[2^{(r-k)/2}\sigma(2^k) + \overline{C}\sum_{i=1}^{r-k} 2^{(i-1)/2 + (r-i)(1-\varepsilon/2)/2}\sigma_1\right].$$

So there is a constant $\tilde{C} = \tilde{C}(\varepsilon, \{\rho_k\})$ such that

$$\sigma(2^r) \leq \tilde{C} \prod_{i=0}^{r-1} (1 + \rho([2^{i(1-\varepsilon)}]))^{1/2} 2^{r/2} \sigma_1.$$

By the relation $1 + x \le \exp(x)$ for every x, we obtain

$$\sigma^2(2^r) \leq \tilde{C}^2 2^r E X_0^2 \exp\left(\sum_{i=0}^{r-1} \rho(\left[2^{i(1-\epsilon)}\right])\right).$$

Relation (2.1) follows now by writing n in binary form and by a simple computation.

In order to prove (2.2), let us mention that for every $\beta > 0$ there exists x_{β} , such that $(1-x) > \exp(-x/(1-\beta))$ for every $0 < x < x_{\beta}$.

Let $\varepsilon < \varepsilon^* < 1$, $\beta > 0$, such that $(1 - \varepsilon)(1 - \beta) = 1 - \varepsilon^*$, and let k^* be an integer such that $\rho([2^{k^*(1-\varepsilon)}]) < \min((1-2^{-\varepsilon/2}), x_\beta)$. By (2.5), by recurrence, for every $r > k \ge k^*$,

$$\begin{split} \sigma(2^k) &\leq 2^{(k-r)/2} \prod_{i=k}^{r-1} \left(1 - \rho(\left[2^{i(1-\varepsilon)}\right]) \right)^{-1/2} \sigma(2^r) \\ &+ \bar{C} 2^{k(1-\varepsilon/2)/2} \sigma_1 \sum_{i=0}^{r-k} \left(2^{\varepsilon/2} \left(1 - \rho(\left[2^{k(1-\varepsilon)}\right]) \right) \right)^{-i/2}. \end{split}$$

Therefore, for every $r > k \ge k^*$ we have

$$\sigma(2^{k}) \leq 2^{(k-r)/2} \sigma(2^{r}) \exp \left(\sum_{i=k}^{r-1} \rho([2^{i(1-\epsilon)}])/2(1-\beta) \right) + C_{3} 2^{k(1-\epsilon/2)/2} \sigma_{1},$$

where $C_3 = C_3(\varepsilon, \{\rho_k\})$.

Now, by Theorem 0, $\sigma(2^k) = 2^{k/2}h^{1/2}(2^k)$, where h(x) is a slowly varying function on R^+ , when $x \to \infty$. Therefore, there is a positive constant $C_4 = C_4(\varepsilon, \beta, \{X_k\})$ such that, for every $r \ge 1$,

$$\sigma^2(2^r) \geq C_4 2^r \exp \left(-\sum_{i=1}^{r-1} \rho(\left[2^{i(1-\epsilon)}\right])/(1-\beta)\right).$$

After a simple computation, we apply relation (4.4) of Peligrad (1982) and we get (2.2). \square

REMARK 4. Bradley (1985) noticed that relation (2.2) cannot be obtained with a constant C_2 depending only on ε , ε^* and $\{\rho_k\}$, even if a factor EX_0^2 is included in the r.h.s. of (2.2) [as in the r.h.s. of (2.1)]. It also has to depend on the sequence $\{X_k\}$ as the following example shows.

Suppose $\{Y_k\}$ and $\{Z_k\}$ are i.i.d. N(0,1) random variables and consider the sequence

$$X_k^{(\alpha)} = Y_k - Y_{k-1} + \sqrt{\alpha} Z_k.$$

Because $\rho(2) = 0$, this sequence is ρ -mixing with $\sigma_n^2 = 2 + n\alpha$, $\inf_n \sigma_n^2/n = \alpha$ and

 $E(X_k^{(\alpha)})^2 = 2 + \alpha$. It is obvious we cannot have a constant C_2 , such that (2.2) applies to all sequences $\{(X_k^{(\alpha)}), \alpha > 0\}$, even if we include EX_0^2 as a factor in the r.h.s. of (2.2).

However, when we need a lower bound of type (2.2), with a constant which does not depend on the sequence, we shall use the following lemma.

LEMMA 2. Suppose $\{X_k\}$ satisfies (1.1). Let $0 < \varepsilon < 1$. Then there exists a positive constant $C_3 = C_3(\{\rho_k\}, \varepsilon)$ such that

(2.6)
$$\sigma(2^{k}) \le 2^{l(\varepsilon-1)/2}\sigma(2^{l+k}) + C_3\sigma_1$$

for every integer $k \ge 1$, $l \ge 1$.

PROOF. This lemma follows by recurrence from (2.5). Let p be an integer such that $\rho(p) < 1 - 2^{-\epsilon}$. By (2.5) there is a $\overline{C} = \overline{C}(\{\rho_k\}, \epsilon)$ such that for every $m \ge 1$,

$$\sigma(m) \leq 2^{(\varepsilon-1)/2} (\sigma(2m) + 2\overline{C}p\sigma_1),$$

whence, by recurrence, we get (2.6). \square

LEMMA 3. Suppose $\{X_k\}$ satisfies (1.1) and $E|X_0|^4 < \infty$. Then for every $\varepsilon > 0$ there is a positive constant $C = C(\varepsilon, \{\rho_k\})$ such that for every $n \ge 1$,

(2.7)
$$E|S_n|^4 \le C(n^{1+\epsilon}E|X_0|^4 + \sigma_n^4).$$

PROOF. Denote by $a_m = ||S_m||_4$. It is easy to prove [see Lemma (3.6) of Peligrad (1982)] that for all integers $m \ge 1$, $k \ge 1$, we have

$$a_{2m} \leq 2^{1/4} \big(1 \, + \, 7 \rho^{1/2} \big(\, k\, \big) \big)^{1/4} a_m + \, 2 \sigma_m + \, 2 \, k a_1.$$

Let $0 < \varepsilon < 1/3$, and k sufficiently large such that $1 + 7\rho^{1/2}(k) \le 2^{\varepsilon}$. By recurrence, for every $r \ge 1$,

$$a(2^r) \le 2^{r(1+\varepsilon)/4}a_1 + 2\sum_{i=1}^r 2^{(i-1)(1+\varepsilon)/4} (\sigma(2^{r-i}) + ka_1).$$

By (2.6), for every $r \geq 1$,

$$a(2^r) \le C_4 2^{r(1+\varepsilon)/4} a_1 + 2\sigma(2^r) \sum_{i=1}^r 2^{i[(1+\varepsilon)/4 - (1-\varepsilon)/2]},$$

where $C_4 = C_4(\varepsilon, \{\rho_k\})$. So

$$a(2^r) \leq C_5 \left(2^{r(1+\varepsilon)/4}a_1 + \sigma(2^r)\right).$$

By writing n in binary form and by (2.6), we get

$$a_n \leq C_6 (n^{(1+\varepsilon)/4}a_1 + \sigma(2^r)).$$

The result follows now by (4.4) of Peligrad (1982). \square

PROOF OF THEOREM 1. By a theorem of Denker (1986) the CLT for strictly stationary sequences satisfying (1.1) is equivalent with the uniform integrability

of $\{S_n^2/\sigma_n^2\}_n$. In Lemma 3.5 of Peligrad (1982), the uniform integrability of $\{S_n^2/\sigma_n^2\}$ under the assumption $\sum_{i=1}^{\infty}\rho(2^i)<\infty$ was established. We shall treat here the case when $\sum_{i=1}^{\infty}\rho(2^i)=\infty$, where we shall consider that $g(x)\to\infty$ as $x\to\infty$. It is easy to see that under this assumption, (1.5b) implies the existence of ε^* , $0<\varepsilon^*\leq\eta$, such that

$$(2.8) \ \ g(n^{1/2}) \ge \exp\left(2\sum_{i=1}^{\lceil \log n \rceil} \rho(2^i)/(1-\epsilon^*)\right), \ \ \text{for every n sufficiently large.}$$

In order to establish the uniform integrability of $\{S_n^2/\sigma_n^2\}$, we shall truncate at the level T:

$$T = g^{\mathrm{inv}} \left(\exp \left(2 \sum_{i=1}^{\left[(1-arepsilon)r
ight]}
ho(2^i) / (1-arepsilon^*) \right) \right).$$

Here $g^{\text{inv}}(x)$ denotes the inverse function of g(x), $r = \log n$ and $0 < \varepsilon < \varepsilon^* < 1$. We put

$$egin{aligned} X_{i1} &= X_i I_{\{|X_i| \le T\}} - E X_i I_{\{|X_i| \le T\}}, \\ X_{i2} &= X_i I_{\{|X_i| > T\}} - E X_i I_{\{|X_i| > T\}}, \\ S_{n1} &= \sum_{i=1}^n X_{i1}, \qquad S_{n2} = \sum_{i=1}^n X_{i2}, \\ \sigma_{n1}^2 &= \operatorname{Var} S_{n1}, \qquad \sigma_{n2}^2 &= \operatorname{Var} S_{n2}. \end{aligned}$$

By (2.1) and the fact that g(x) is an increasing function, we have

$$\sigma_{n2}^2 \leq C_1(n/g(T))EX_0^2g(|X_0|)I_{\{|X_0| > T\}} \exp\Biggl(\sum_{i=1}^{\lceil (1-\epsilon)r \rceil} \rho(2^i)/(1-\epsilon)\Biggr),$$

where C_1 does not depend on n. By (2.2) and the definition of T it follows that

$$\sigma_{n2}^2 \leq (C_1/C_2)\sigma_n^2 E X_0^2 g(|X_0|) I_{\{|X_0| > T\}}.$$

From this and because $g^{inv}(x) \to \infty$ when $x \to \infty$ we deduce that

(2.9)
$$\sigma_{n2} = o(\sigma_n) \text{ as } n \to \infty,$$

whence it is easy to see that

$$(2.10) \sigma_{n1} \sim \sigma_n as n \to \infty.$$

By applying Lemma 3 to the sequence $\{X_{n1}\}$, we can find a constant $K_1 = K_1(\{\rho_k\}, \varepsilon)$ such that for every $n \ge 1$,

$$E|S_{n1}|^4 \leq K_1(n^{1+\varepsilon/2}T^2EX_0^2 + \sigma_{n1}^4).$$

By (2.2) and Remark 1, we can find $C = C(\{X_k\}, \epsilon)$ such that $\sigma_n^4 \ge Cn^{2-\epsilon/2}$ for every $n \ge 1$, whence, by (2.10),

(2.11)
$$E(|S_{n1}|/\sigma_n)^4 \le K_2(T^2/n^{1-\epsilon}+1)$$
, for every $n \ge 1$,

where K_2 is a constant that does not depend on n.

By (2.8), for every n sufficiently large

$$g([n^{1-\varepsilon}]^{1/2}) \ge \exp\left(2\left(\sum_{i=1}^{[(1-\varepsilon)r]} \rho(2^i)/(1-\varepsilon^*)\right)\right).$$

This implies, by the definition of T, that $T/n^{(1-\epsilon)/2}$ is bounded and by (2.11) we get

$$\sup_{n} E(|S_{n1}|/\sigma_n)^4 < \infty.$$

Now, from (2.9) and (2.12), we can conclude that $\{S_n^2/\sigma_n^2\}$ is uniformly integrable and so we have established the CLT. \Box

Acknowledgments. I am grateful to Richard Bradley for his clarifying comments on the treatment of the constants in Lemmas 1 and 3, which helped us to fill some gaps in their statements. I would like also to thank the referee and Thomas Liggett for carefully reading the manuscript and for many suggestions that improved the presentation of this paper.

REFERENCES

Bradley, R. C. (1980). A remark on the central limit question for dependent random variables.

J. Appl. Probab. 17 94-101.

BRADLEY, R. C. (1985). Personal communication.

BRADLEY, R. C. (1987). The central limit question under ρ -mixing. Rocky Mountain J. Math. 17 95-114.

DENKER, M. (1986). Uniform integrability and the central limit theorem. In *Dependence in Probability and Statistics* (E. Eberlein and M. S. Taqqu, eds.) 269–274. Birkhäuser, Boston.

IBRAGIMOV, I. A. (1975). A note on the central limit theorem for dependent random variables. Theory Probab. Appl. 20 135-141.

IBRAGIMOV, I. A. and LINNIK, Yu. V. (1971). Independent and Stationary Sequences of Random Variables. Walters-Noordhoff, Groningen.

Kolmogorov, A. N. and Rozanov, Y. A. (1960). On strong mixing conditions for stationary Gaussian processes. *Theory Probab. Appl.* 5 204-208.

Peligrad, M. (1982). Invariance principles for mixing sequences of random variables. *Ann. Probab.* **10** 968–981.

DEPARTMENT OF MATHEMATICAL SCIENCES University of Cincinnati Cincinnati, Ohio 45221-0025