Journal of Theoretical Probability, Vol. 7, No. 4, 1994

Law of the Iterated Logarithm for Perturbed Empirical Distribution Functions Evaluated at a Random Point for Nonstationary Random Variables[1]

Michel Harel[2] and Madan L. Puri[3]

Received April 1, 1993; revised October 7, 1993

We consider perturbed empirical distribution functions $\hat{F}_n(x) = 1/n \sum_{i=1}^{n} G_n(x - X_i)$, where $\{G_n, n \geqslant 1\}$ is a sequence of continuous distribution functions converging weakly to the distribution function of unit mass at 0, and $\{X_i, i \geqslant 1\}$ is a nonstationary sequence of absolutely regular random variables. We derive the almost sure representation and the law of the iterated logarithm for the statistic $\hat{F}_n(U_n)$ where U_n is a U-statistic based on $X_1,..., X_n$. The results obtained extend or generalize the results of Nadaraya,(7) Winter,(16) Puri and Ralescu,(9,10) Oodaira and Yoshihara,(8) and Yoshihara,(19) among others.

KEY WORDS: Perturbed empirical distribution functions; absolutely regular processes; strong mixing; almost sure representation; U-statistic; law of the iterated logarithm.

1. INTRODUCTION

Let $\{X_i, i \geqslant 1\}$ be a sequence of absolutely regular r.v.s (random variables) with continuous c.d.f.s (cumulative distribution functions) $\{F_i, i \geqslant 1\}$ defined on the real line $\mathbb{R}$ and let $\tilde{F}_n$ be the corresponding empirical distribution function based on $X_1,..., X_n$, i.e., $\tilde{F}_n(x) = n^{-1} \sum_{i=1}^{n} u(x - X_i)$ where $u(t) = 1$ or 0 according as $t \geqslant 0$ or $t < 0$. Suppose that F_i converges to a fixed d.f. F as $i \to \infty$. If $\{X_i, i \geqslant 1\}$ is a stationary sequence (in which case $F_i \equiv F$ for all i), $\tilde{F}_n$ is a natural estimator of F based on the sample $X_1,..., X_n$. However if F is a smooth distribution function, then it seems

[1] Research supported by the Office of Naval Research Contract N00014-91-J-1020.
[2] U.R.A. C.N.R.S. 743, Orsay, France and Indiana University, Bloomington, Indiana, 47405.
[3] Indiana University, Bloomington, Indiana, 47405.

860/7/4-9

0894-9840/94/1000-0831$07.00/0 © 1994 Plenum Publishing Corporation

reasonable to consider smooth estimators which are better adapted to this situation. For the case $\{X_i, i \geq 1\}$ is a sequence of iid (independent and identically distributed) r.v.s, Watson and Leadbetter,[14] Nadaraya,[7] Yamato,[18] and Winter,[16,17] have considered an estimator of the form

$$\hat{F}_n(x) = n^{-1} \sum_{i=1}^{n} G_n(x - X_i), \qquad x \in \mathbb{R}, \quad n \geq 1 \tag{1.1}$$

where $G_n(x) = \int_{-\infty}^{x} g_n(t)\, dt$, $g_n(t) = \alpha_n^{-1} g(t\alpha_n^{-1})$, $\{\alpha_n; n \geq 1\}$ is a sequence of positive real numbers such that $\alpha_n \to 0$ as $n \to \infty$, and g is a probability density function. Watson and Leadbetter[14] established the asymptotic normality of $\hat{F}_n$, while Nadaraya,[7] Winter,[16] and Yamato,[18] proved the a.s. (almost sure) convergence of $\hat{F}_n$ to F. Later Winter[17] showed that $\hat{F}_n$ has the Chung-Smirnov property, i.e. with probability one,

$$\limsup_{n \to \infty} (2n/\log n)^{1/2} \sup_x |\hat{F}_n(x) - F(x)| \leq 1$$

For some related results dealing with the estimation of probability density function, see Schuster,[11] Scott *et al.*,[12] and Wertz,[15] for a general review in this area.

Let $h(x_1,..., x_m)$, symmetric in its arguments be a measurable kernel (of degree m), and let U_n be the corresponding U-statistics given by

$$U_n = \binom{n}{m}^{-1} \sum_{C_{n,m}} h(X_{i_1},..., X_{i_m}) \tag{1.2}$$

where $C_{n,m}$ denotes the set of all $\binom{n}{m}$ combination of m distinct elements $\{i_1,..., i_m\}$ from $\{1,..., n\}$. Let

$$\theta(F) = \int_{\mathbb{R}^m} h(x_1,..., x_m) \prod_{i=1}^{m} dF(x_i) \tag{1.3}$$

When the r.v.s are i.i.d., Puri and Ralescu[10] derived the asymptotic normality of $\hat{F}_n(U_n)$ which is useful in estimating $\xi = F(\theta(F))$ when F is unknown. Later Sun[13] extended this result when the r.v.s are stationary and absolutely regular. Lea and Puri[6] established for the iid case, (i) the almost sure (a.s.) representation, (ii) the law of the iterated logarithm and (iii) an invariance principle for $\hat{F}_n(U_n)$. In this paper, we extend the results (i) and (ii) when the r.v.s are nonstationary and absolutely regular.

With our general framework (nonstationarity and absolute regularity), the results have applications for a larger class of processes such as ARMA processes and Markov processes for which the initial measure is not necessary to be the invariant measure. For example, suppose that the sequence $\{X_i, i \geqslant 1\}$ of r.v.'s is an ARMA process defined by

$$X_i = aX_{i-1} + b\varepsilon_i + \varepsilon_{i-1}, \qquad i \geqslant 1 \tag{1.4}$$

where X_0 admits a strictly positive density, $\{\varepsilon_i, i \geqslant 1\}$ is a sequence of i.i.d. r.v.'s with strictly positive density such that $E(\varepsilon_i) = 0$ and a and b are real numbers such that $|a| < 1$. Then, from Harel and Puri,[4] this process is absolutely regular and for any initial measure, the conclusions of the theorems and propositions here hold if the sequence of r.v.'s is an ARMA process defined by Eq. (1.4). The results of our paper can also be applied for nonstationary Markov processes (see, e.g. Harel and Puri[5]).

2. SOME PRELIMINARY RESULTS

In the following, we suppose that the sequence $\{X_i, i \geqslant 1\}$ is absolutely regular with rates

$$\beta(m) = O(\rho^m), \qquad 0 < \rho < 1 \tag{2.1}$$

Recall that a nonstationary sequence $\{X_{ni}, 1 \leqslant i \leqslant n, n \geqslant 1\}$ is absolutely regular if

$$\sup_{m \leqslant n} \max_{1 \leqslant j \leqslant n-m} E\{\sup_{A \in \sigma(X_{ni}, i \geqslant j+m)} |P(A \mid \sigma(X_{ni}, 1 \leqslant i \leqslant j)) - P(A)|\}$$
$$= \beta(m) \downarrow 0 \text{ as } m \to \infty$$

Here $\sigma(X_{ni}, 1 \leqslant i \leqslant j)$ and $\sigma(X_{ni}, i \geqslant j+m)$ are the σ-fields generated by $(X_{n1},..., X_{nj})$ and $(X_{n,j+m}, X_{n,j+m+1},..., X_{nn})$ respectively. Also recall that $\{X_{ni}\}$ satisfies the strong mixing condition if $\sup_{m \leqslant n} \sup_{1 \leqslant j \leqslant n-m} \{|P(A \cap B) - P(A)\, P(B)|;\ A \in \sigma(X_{ni}, 1 \leqslant i \leqslant j),\ B \in \sigma(X_{ni}, i \geqslant j+m)\} = \alpha(m) \downarrow 0$ as $m \to \infty$. Since $\alpha(m) \leqslant \beta(m)$, it follows that if $\{X_{ni}\}$ is absolutely regular, then it is also strong mixing.

Let $\{d_n\}$ be a sequence of positive real numbers such that $d_n \sim kn^{-1/2}(\log \log n)^{1/2}$ as $n \to \infty$, for some $k > 0$. Define

$$K_n(t) = \sup_{|x| \leqslant d_n} |\tilde{F}_n(t+x) - \tilde{F}_n(t) - \bar{F}_n(t+x) - \bar{F}_n(t)| \tag{2.1'}$$

where $\bar{F}_n = n^{-1} \sum_{i=1}^n F_i$, and let F_i' denote the density of F_i.

Lemma 2.1. For $\delta > 0$, if $\{F_i', i \geqslant 1\}$ is uniformly bounded in $(\theta(F) - \delta, \theta(F) + \delta)$ and

$$K_n^* = \sup_{|t - \theta(F)| < \delta/2} K_n(t) \tag{2.2}$$

then, for any v $(0<v<\frac{1}{8})$ and for any $s>1$ there exist two positive constants $C_v^{(1)}$ and $C_v^{(2)}$ and an integer n_s such that

$$P[K_n^* \geqslant C_v^{(1)} n^{-3/4+v}] \leqslant C_v^{(2)} n^{-s} \tag{2.3}$$

for all $n \geqslant n_s$. Hence

$$K_n^* = O(n^{-3/4+v}) \text{ a.s. as } n \to \infty$$

Proof. We start in a way similar way to Lemma 2.1 of Lea and Puri.[6] For $\eta \in [\theta(F)-\delta/2, \theta(F)+\delta/2]$, let

$$H_n(x) = \tilde{F}_n(x) - \tilde{F}_n(\eta) - \bar{F}_n(x) + \bar{F}_n(\eta) \tag{2.4}$$

Let $\{b_n, n \geqslant 1\}$ be a sequence of positive integers such that $b_n \sim n^{1/4}$ as $n \to \infty$. Consider a particular n. For any integer $r \in [-b_n, b_n]$, let $\eta_{r,n} = \eta + d_n b_n^{-1} r$ and $\theta_{r,n} = \bar{F}_n(\eta_{r+1,n}) - \bar{F}_n(\eta_{r,n})$. Since $\tilde{F}_n$ and $\bar{F}_n$ are nondecreasing, it follows from Eq. (2.4) that for $x \in [\eta_{r,n}, \eta_{r+1,n}]$,

$$H_n(x) \leqslant \tilde{F}_n(\eta_{r+1,n}) - \tilde{F}_n(\eta) - \bar{F}_n(\eta_{r,n}) + \bar{F}_n(\eta) = H_n(\eta_{r+1,n}) + \theta_{r,n} \tag{2.5}$$

Similarly

$$H_n(x) \geqslant H_n(\eta_{r,n}) - \theta_{r,n} \text{ for } x \in [\eta_{r,n}, \eta_{r+1,n}] \tag{2.6}$$

It follows from Eqs. (2.1), (2.5), and (2.6) that

$$K_n(\eta) \leqslant J_n + \beta_n \tag{2.7}$$

where

$$J_n = \max_{-b_n \leqslant r \leqslant b_n} |H_n(\eta_{r,n})|, \qquad \beta_n = \max_{-b_n \leqslant r \leqslant b_n} \theta_{r,n} \tag{2.8}$$

Now, since $\{F_i, i \geqslant 1\}$ is uniformly bounded in a neighborhood of η, $\bar{F}'_n$ is also uniformly bounded in a neighborhood of η, it follows that for large n

$$\beta_n = O(n^{-3/4}(\log\log n)^{1/2}) \tag{2.9}$$

Let C_2 be the upper bound of $\sup_{n \geqslant 1} \bar{F}'_n(x)$ in $(\xi-\delta, \xi+\delta)$. Since $d_n \to 0$ as $n \to \infty$, there exists a positive integer N_1 such that $\sup_{n \geqslant 1} \bar{F}_n(\eta) - \bar{F}_n(\eta - d_n) \leqslant C_2 d_n$ for all $n \geqslant N_1$. Applying Lemma 5.2 (see Appendix), with $m = n^{2v-\varepsilon}$, $0<\varepsilon<2v$, and $a=\frac{7}{8}$, we have

$$P[|H_n(\eta_{r,n})| \geqslant \delta_n^{(v)}] \leqslant 2n^{2v-\varepsilon} \exp\{-r_n^{(v)}\} + 2n\beta([n^{2v-\varepsilon}]) \tag{2.10}$$

where $\delta_n^{(v)} = C_1 n^{-3/4+v}$, [a] is the integer part of a, C_1 is some constant > 0 and

$$r_n^{(v)} = n^{7/8(1-2v+\varepsilon)} \delta_n^{(v)} / 5M_0 \tag{2.11}$$

After some routine computations, we obtain

$$r_n^{(v)} = O(n^{1/4 - v + \varepsilon})$$

which implies

$$n^{2v-\varepsilon} \exp\{-r_n^{(v)}\} = O(n^{-(s+3/4)}) \tag{2.12}$$

We also have

$$\beta([n^{2v-\varepsilon}]) = O(e^{(\log \rho) n^{2v-\varepsilon}}) \tag{2.13}$$

Thus, from Eq. (2.10), we deduce

$$P[|H(\eta_{r,n})| \geqslant \delta_n^{(v)}] = O(n^{-(s+3/4)}) + O(n e^{n^{(2v-\varepsilon)} \log \rho}) \tag{2.14}$$

It follows from Eqs. (2.8) and (2.14) that

$$P[J_n \geqslant \delta_n^{(v)}] = O(b_n n^{-(s+3/4)}) + O(b_n n e^{n^{(2v-\varepsilon)} \log \rho}) \tag{2.15}$$

It is easy to see that if C_1 is chosen sufficiently large, there exists a positive constant C_2 and a positive integer N_1 such that for $n \geqslant N_1$

$$P[J_n \geqslant \delta_n^{(v)}] \leqslant C_2 n^{-(s+1/2)} \tag{2.16}$$

From Eqs. (2.7), (2.9), and (2.16), it follows that there exists a positive integer N_2 such that

$$P[K_n(\eta) \geqslant 2\delta_n^{(v)}] \leqslant C_2 n^{-(s+1/2)} \qquad \text{for} \quad n \geqslant N_2 \tag{2.17}$$

Now, let

$$r_{j,n} = \theta(F) + \frac{j\delta}{2\sqrt{n}}, \qquad j = -[\sqrt{n}], ..., 0, ..., [\sqrt{n}]$$

Then by Eq. (2.17), we have

$$P[\max_{-\sqrt{n} \leqslant j \leqslant \sqrt{n}} K_n(r_{j,n}) \geqslant 2\delta_n^{(v)}] \leqslant 2C_2 n^{1/2} n^{-(s+1/2)}$$
$$= 2C_2 n^{-s} \qquad \text{for} \quad n \geqslant N_2 \tag{2.18}$$

For any $t \in [\theta(F) - \delta/2, \theta(F) + \delta/2]$ and for any $x \in [-d_n, d_n]$ where $x \neq 0$, there exist two integers j_1 and j_2 such that

$$[-\sqrt{n}] \leqslant j_1 < j_2 \leqslant [\sqrt{n}]$$

and

$$\theta(F) + \frac{(j_1 - 1)\,\delta}{2\sqrt{n}} < \underline{t} \leqslant \theta(F) + \frac{j_1 \delta}{2\sqrt{n}} \leqslant \theta(F) + \frac{(j_2 - 1)\,\delta}{2\sqrt{n}}$$
$$< \bar{t} \leqslant \theta(F) + \frac{j_2 \delta}{2\sqrt{n}}$$

where $\underline{t} = \min\{t, t + x\}$ and $\bar{t} = \max\{t, t + x\}$.

From the triangular inequality, we obtain

$$|V_n(t) - V_n(t+x)| \leqslant \left| V_n(\underline{t}) - V_n\left(\theta(F) + \frac{j_1 \delta}{2\sqrt{n}}\right)\right|$$
$$+ \left| V_n\left(\theta(F) + \frac{j_1 \delta}{2\sqrt{n}}\right) - V_n\left(\theta(F) + \frac{(j_2 - 1)\,\delta}{2\sqrt{n}}\right)\right|$$
$$+ \left| V_n\left(\theta(F) + \frac{(j_2 - 1)\,\delta}{2\sqrt{n}}\right) - V_n(t + x)\right|$$

where $V_n(t) = \tilde{F}_n(t) - \bar{F}_n(t)$.

As $d_n \geqslant \delta/2\sqrt{n}$ for n sufficiently large, we deduce

$$K_n^* \leqslant 3 \max_{-\sqrt{n} \leqslant j \leqslant \sqrt{n}} K_n(r_{j,n}) \tag{2.19}$$

Choosing $C_v^{(1)} = 6C_1$, $C_v^{(2)} = 2C_2$ and $n \geqslant N_2$, it follows, from Eqs. (2.18) and (2.19), that

$$P[K_n^* \geqslant C_v^{(1)} n^{-3/4 + v}] \leqslant C_v^{(2)} n^{-s}$$

Thus Eq. (2.3) follows. Also

$$\sum_{n=1}^{\infty} P[K_n^* \geqslant C_v^{(1)} n^{-3/4 + v}] \leqslant C_v^{(2)} \sum_{n=1}^{\infty} n^{-s} < \infty$$

Consequently, using the Borel-Cantelli Lemma, we obtain

$$K_n^* = O(n^{-3/4 + v}) \text{ a.s.} \qquad \text{as} \quad n \to \infty$$

This completes the proof.

Lemma 2.2. If (i) $\{F'_i, i \geqslant 1\}$ is uniformly bounded in a neighborhood of $\theta(F)$ and (ii) $\int x^2 g(x)\,dx < \infty$, then

$$|\hat{F}_n(\theta(F)) - \tilde{F}_n(\theta(F))| = O(n^{-3/4+v}) + O(n\alpha_n^2(\log\log n)^{-1}) \text{ a.s. } (0 < v < \tfrac{1}{8})$$

Proof. We proceed as in Lemma 2.2 of Lea and Puri.[6] First note that

$$\hat{F}_n(\theta(F)) = n^{-1} \sum_{i=1}^{n} G_n(\theta(F) - X_i) = \int_{-\infty}^{+\infty} G_n(\theta(F) - t)\, d\tilde{F}_n(t)$$
$$= \int_{-\infty}^{+\infty} \tilde{F}_n(\theta(F) - t)\, g_n(t)\, dt \tag{2.20}$$

and

$$\int_{-\infty}^{+\infty} g_n(t)\, dt = 1 \tag{2.21}$$

Condition (i) ensures that there exists $\delta > 0$ and $M > 0$ such that $\sup_{i \in \mathbb{N}} |F_i(\theta(F) - t) - F_i(\theta(F))| \leqslant M\,|t|$ for all $|t| < \delta$.

Using Lemma 2.1 and Eqs. (2.20) and (2.21), we have

(a) $|\hat{F}_n(\theta(F)) - \tilde{F}_n(\theta(F))|$

$$\leqslant \int_{-\infty}^{+\infty} |\tilde{F}_n(\theta(F) - t) - \tilde{F}_n(\theta(F))|\, g_n(t)\, dt$$
$$\leqslant \int_{|t| \leqslant d_n} |\tilde{F}_n(\theta(F) - t) - \tilde{F}_n(\theta(F))|\, g_n(t)\, dt + \int_{|t| > d_n} |\tilde{F}_n(\theta(F) - t) - \tilde{F}_n(\theta(F))|\, g_n(t)\, dt$$
$$\leqslant \int_{|t| \leqslant d_n} |\bar{F}_n(\theta(F) - t) - \bar{F}_n(\theta(F))|\, g_n(t)\, dt + \int_{|t| > d_n} g_n(t)\, dt + O(n^{-3/4+v}) \text{ a.s.}$$

For large n, the first term on the right-hand side of the last inequality in (a) is bounded by

(b) $$M \int_{-\infty}^{+\infty} |t|\, g_n(t)\, dt = M\alpha_n \int_{-\infty}^{+\infty} |y|\, g(y)\, dy = O(\alpha_n)$$

and the second term is bounded by

$$\text{(c)}\ \ n\alpha_n^2(\log\log n)^{-1}\int_{|y|>d_n/\alpha_n} y^2 g(y)\,dy = O(n\alpha_n^2(\log\log n)^{-1})$$

since $\int_{-\infty}^{+\infty} y^2 g(y)\,dy < \infty$. Thus, from (a)–(c) we have

$$\begin{aligned} |\hat{F}_n(\theta(F)) - \tilde{F}_n(\theta(F))| &= O(\alpha_n) + O(n\alpha_n^2(\log\log n)^{-1}) + O(n^{-3/4+v}) \\ &= O(n\alpha_n^2(\log\log n)^{-1}) + O(n^{-3/4+v}) \end{aligned}$$

a.s. as $n \to \infty$ and Lemma 2.2 is proved.

Because of the hypothesis of nonstationarity, we have to define some new notations for the U-statistics.

For every $p (1 \leqslant p \leqslant m)$ and $n \geqslant m$, let $1 \leqslant i_1 \neq \cdots \neq i_p \leqslant n$ be arbitrary integers. Put

$$h_{p,n}^{(i_1,\dots,i_p)}(x_1,\dots,x_p) = \sum_{(i_{p+1},\dots,i_m)\in I_{p,n}(i_1,\dots,i_p)} \lambda(x_1,\dots,x_p; i_{p+1},\dots,i_m) \quad (2.22)$$

where

$$\lambda(x_1,\dots,x_p; i_{p+1},\dots,i_m) = \int_{\mathbb{R}^{m-p}} h(x_1,\dots,x_m)\,dF_{i_{p+1}}(x_{p+1})\dots dF_{i_m}(x_m)$$

and $I_{p,n}(i_1,\dots,i_p) = \{(i_1,\dots,i_p);\quad 1 \leqslant i_{p+1} \neq \cdots \neq i_m \leqslant n,\quad i_l \notin (i_1,\dots,i_p),$ $p+1 \leqslant l \leqslant m\}$ and

$$h_{0,n} = \sum_{(i_1,\dots,i_m)\in I_{0,n}} \int_{\mathbb{R}^m} h(x_1,\dots,x_m)\,dF_{i_1}(x_1),\dots, dF_{i_m}(x_m)$$

where $I_{0,n} = \{(i_1,\dots,i_m),\ 1 \leqslant i_1 \neq \cdots \neq i_m \leqslant n\}$.

For every $p (1 \leqslant p \leqslant m)$, set

$$U_n^{(p)} = n^{-[m]} \sum_{1\leqslant i_1\neq\cdots\neq i_p\leqslant n} \int_{\mathbb{R}^p} h_{p,n}^{(i_1,\dots,i_p)}(x_1,\dots,x_p) \times \prod_{j=1}^{p} d(I_{[X_{i_j}\leqslant x_j]} - F_{i_j}(x_j)) \quad (2.23)$$

where $n^{-[m]} = (n(n-1)\cdots(n-m+1))^{-1}$.

Lemma 2.3. If there exists a positive number δ such that for $r = 4 + \delta$

$$\sup_n \max_{(i_1, \ldots, i_m) \in \{1, \ldots, n\}^m} \int_{\mathbb{R}^m} |h(x_1, \ldots, x_m)|^r \, dF_{i_1}(x_1) \times \cdots \times dF_{i_m}(x_m) \leqslant M_0 < \infty \tag{2.24}$$

$$\sup_n \max_{(i_1, \ldots, i_m) \in \{1, \ldots, n\}^m} E(|h(X_{i_1}, \ldots, X_{i_m})|^r) \leqslant M_0 < \infty \tag{2.25}$$

and for some $\delta'(0 < \delta' < \delta)$, $\beta(n) = O(n^{-3(4+\delta')/(2+\delta')})$, then we have

$$E(U_n^{(2)})^4 = O(n^{-3-\gamma}) \tag{2.26}$$

where $\gamma = 6(\delta - \delta')/(4+\delta)(2+\delta') > 0$ and

$$E(U_n^{(p)})^2 = O(n^{-3}), \qquad 3 \leqslant p \leqslant m \tag{2.27}$$

Proof. We first note that

$$U_n^{(2)} = n^{-[m]} \sum_{1 \leqslant i_1 \neq i_2 \leqslant n} \int_{\mathbb{R}^2} h_{2,n}^{(i_1, i_2)}(x_1, x_2) \, d(I_{[X_{i_1} \leqslant x_1]} - F_{i_1}(x_1)) \, d(I_{[X_{i_2} \leqslant x_2]} - F_{i_2}(x_2))$$

So we have

$$E(U_n^{(2)})^4 = (n^{-[m]})^4 \sum_{1 \leqslant i_{11}, i_{12} \leqslant n} \sum_{1 \leqslant i_{21}, i_{22} \leqslant n} \sum_{1 \leqslant i_{31}, i_{32} \leqslant n} \times \sum_{1 \leqslant i_{41}, i_{42} \leqslant n} J((i_{11}, i_{12}), (i_{21}, i_{22}), (i_{31}, i_{32}), (i_{41}, i_{42})) \tag{2.28}$$

where

$$J((i_{11}, i_{12}), (i_{21}, i_{22}), (i_{31}, i_{32}), (i_{41}, i_{42})) = E\left\{ \prod_{j=1}^{4} \int_{\mathbb{R}^2} h_{2,n}^{(i_{j1}, i_{j2})}(x_{j1}, x_{j2}) \, d(I_{[X_{i_{j1}} \leqslant x_{j1}]} - F_{i_{j1}}(x_{j1})) \, d(I_{[X_{i_{j2}} \leqslant x_{j2}]} - F_{i_{j2}}(x_{j2})) \right\} \tag{2.29}$$

and

$$h_{2,n}^{(i_{j1}, i_{j2})}(x_{i1}, x_{i2}) = \sum_{(i_3, \ldots, i_m) \in I_{2,n}(i_{j1}, i_{j2})} g(x_{i1}, x_{i2}, x_3, \ldots, x_m) \, dF_{i_3}(x_3) \cdots dF_{i_m}(x_m)$$

Let $i_{rs}(\leqslant n)(r=1,\ldots,4,\ s=1,2)$ be mutually different. Reorder i_{rs} as $1\leqslant k_1<k_2<\cdots<k_8\leqslant n$ and put

$$J((i_{11},i_{12}),\ldots,(i_{41},i_{42}))=H(k_1,\ldots,k_8)=E[g(X_{k_1},\ldots,X_{k_8})] \tag{2.30}$$

Let $d^{(c)}$ be the cth largest difference among $(k_{j+1}-k_j),\ j=1,\ldots,7$

Since

$$\int_{\mathbb{R}^8} g(x_1,\ldots,x_8)\, P_j^{(k_1,\ldots,k_8)}(x_1,\ldots,x_8)=0\ (j=1,7)\ (\text{see Appendix})$$

It follows from Lemma 5.1 (see Appendix) that

$$H(k_1,\ldots,k_8)\leqslant\left(\frac{n^{[m]}}{n(n-1)}\right)^4 M_0[\beta(k_8-k_7)]^{(2+\delta)/(4+\delta)}$$
$$\text{if}\quad k_8-k_7=d^{(1)} \tag{2.31}$$

and

$$H(k_1,\ldots,k_8)\leqslant\left(\frac{n^{[m]}}{n(n-1)}\right)^4 M_0[\beta(k_2-k_1)]^{(2+\delta)/(4+\delta)}$$
$$\text{if}\quad k_2-k_1=d^{(1)} \tag{2.32}$$

Hence

$$\sum_{\substack{1\leqslant k_1<\cdots<k_8\leqslant n\\ k_8-k_7=d^{(1)}\text{ or }k_2-k_1=d^{(1)}}} H(k_1,\ldots,k_8)$$
$$\leqslant\left(\frac{n^{[m]}}{n(n-1)}\right)^4 M_0 n^4\sum_{j=1}^{n}(j+1)^3\,[\beta(j)]^{(2+\delta)/(4+\delta)} \tag{2.33}$$

If for some $j_\alpha(2\leqslant j_\alpha\leqslant 6,\ 1\leqslant\alpha\leqslant 4)$, $k_{j_\alpha+1}-k_{j_\alpha}=d^{(\alpha)}(1\leqslant\alpha\leqslant 4)$, then from Lemma 5.1 (see Appendix)

$$H(k_1,\ldots,k_8)\leqslant\left(\frac{n^{[m]}}{n(n-1)}\right)^4 M\sum_{\alpha=1}^{4}[\beta(k_{j_\alpha+1}-k_{j_\alpha})]^{(2+\delta)/(4+\delta)} \tag{2.34}$$

and hence

$$\sum_{\substack{1\leqslant k_1<\cdots<k_8\leqslant n\\ k_{j_\alpha+1}-k_{j_\alpha}=d^{(\alpha)}(1\leqslant\alpha\leqslant 4)}} H(k_1,\ldots,k_8)$$
$$\leqslant 4\left(\frac{n^{[m]}}{n(n-1)}\right)^4 M_0 n^4\sum_{j=1}^{n}(j+1)^3\,[\beta(j)]^{(2+\delta)/(4+\delta)} \tag{2.35}$$

Consequently

$$\sum_{1 \leqslant k_1 < \cdots < k_8 \leqslant n} H(k_1, ..., k_8)$$

$$\leqslant M \left(\frac{n^{[m]}}{n(n-1)} \right)^4 n^4 \sum_{j=1}^{n} (j+1)[\beta(j)]^{(2+\delta)/(4+\delta)} = O(n^{4m-3-\gamma}) \quad (2.36)$$

where $M > 0$ is some constant.

We can use a similar method to estimate the sums in the other cases and so we have Eq. (2.26).

The proof of Eq. (2.27) is analogous and so it is omitted.

Lemma 2.4. Under the conditions of Lemma 2.3, we have

$$U_n - \theta(F) = mU_n^{(1)} + R_n \quad (2.37)$$

where $R_n = O((n^{-1} \log \log n)^{1/2})$ a.s. as $n \to \infty$.

Proof. We rewrite Eq. (1.2) as

$$U_n = \theta(F) + \sum_{p=1}^{m} \binom{m}{p} U_n^{(p)} = \theta(F) + mU_n^{(1)} + \sum_{p=2}^{m} \binom{m}{p} U_n^{(p)}$$

then we have to prove that for every $\varepsilon > 0$

$$P[|W_n| > \varepsilon(n \log \log n)^{1/2} \text{ i.o.}] = 0 \quad (2.38)$$

where $W_n = n(U_n - \theta(F) - mU_n^{(1)})$. Let $n_k = [k^{(2+\delta')(4+\delta)/3(\delta-\delta')}]$ and $n_{k_0} \geqslant m$. Then, from Lemma 2.3 and the Bonferroni inequality

$$\sum_{k=k_0}^{\infty} P[\max_{m \leqslant n \leqslant n_k} |W_n| \geqslant \varepsilon(n_k \log \log n_k)^{1/2}]$$

$$\leqslant \sum_{k=k_0}^{\infty} \left(\sum_{n=m}^{n_k} P[|W_n| \geqslant \varepsilon(n_k \log \log n_k)^{1/2}] \right)$$

$$\leqslant M \sum_{k=k_0}^{\infty} k^{-2} < \infty \quad (2.39)$$

where $M > 0$ is some constant. Consequently, from the Borel-Cantelli Lemma, we have

$$P[|W_n| \geqslant \varepsilon(n \log \log n)^{1/2} \text{ i.o.}]$$

$$\leqslant P[\max_{n_k \leqslant n \leqslant n_{k+1}} |W_n| \geqslant \varepsilon(n_k \log \log n_k) \text{ i.o.}]$$

$$\leqslant P\left[\max_{m \leqslant n \leqslant n_{k+1}} |W_n| > \frac{\varepsilon}{4} (n_{k+1} \log \log n_{k+1}) \text{ i.o.}\right] = 0$$

3. A LAW OF THE ITERATED LOGARITHM FOR THE NONSTATIONARY U-STATISTIC

$$\sigma^2(\{F_l^*\}) = \sigma^2 = \left[\int_{\mathbb{R}^m} h^2(x_1,\ldots,x_m) \prod_{l=1}^{k} dF(x_l) - \theta^2(F)\right]$$

$$+2\sum_{i=1}^{\infty}\left[\int_{\mathbb{R}^{2m}} h(x_1,\ldots,x_m)\, h(x_{m+1},\ldots,x_{2m})\, dF_i^*(x_1, x_{m+1})\right.$$

$$\left.\times \prod_{l=2}^{m} dF(x_l) \prod_{p=m+2}^{2m} dF(x_p) - \theta^2(F)\right] \tag{3.1}$$

if the limit exists and is finite.

Let $F_{i,j}$ be the distribution function of (X_i, X_j), $1 \leqslant i < j$.

Theorem 3.1. Suppose the sequence $\{X_i, i \geqslant 1\}$ is absolutely regular with rates satisfying Eq. (2.1). Furthermore, assume that for any $l > 1$, there exists a continuous d.f. F_l^* on $\mathbb{R}^2$ with marginals F such that

$$\|F_{i,j} - F_{j-i}^*\| = O(\rho_0^{\max(i,j-i)}), \qquad 1 \leqslant i < j \tag{3.2}$$

for some $0 < \rho_0 < 1$ where $\|\ \|$ denotes the norm of total variation.

Suppose also, that there exists a δ' such that Eqs. (2.24) and (2.25) are satisfied. Then we have

$$\limsup_{n\to\infty} \frac{n^{1/2}\,|U_n - \theta(F)|}{2^{1/2} m\sigma(\log\log n)^{1/2}} = 1 \text{ a.s.} \tag{3.3}$$

Proof. From Lemma 2.4, we have only to prove that

$$\limsup_{n\to\infty} \frac{|U_n^{(1)}|}{2^{1/2}\sigma(\log\log n)^{1/2}} = 1 \text{ a.s.} \tag{3.4}$$

where $U_n^{(1)} = n^{-1}\sum_{i=1}^{n} X_{ni}^*$ and

$$X_{ni}^* = n^{-[m-1]}\left(h_{1,n}^{(i)}(X_i) - \int_{\mathbb{R}} h_{1,n}^{(i)}(x)\, dF_i(x)\right) \tag{3.5}$$

From Lemma 5.6 (in the Appendix), we have to verify that the sequence X_{ni}^* satisfies the conditions in Eqs. (5.12)–(5.14). Condition in Eq. (5.13) follows from condition in Eq. (2.24). Condition in Eq. (5.14) follows from condition in Eq. (2.1).

We have to prove condition in Eq. (5.12), that is, for any n, m such that $n \geqslant m$ and any $J \subset \{1,..., n\}$ such that $card\ J = m$

$$E\left(\sum_{i \in J} X^*_{ni}\right)^2 = m\sigma^2(1 + o(1)) \tag{3.6}$$

We prove Eq. (3.6) for $card\ J = n$, because for $card\ J < n$, the proof is similar. Denote

$$\rho(1) = \int_{\mathbb{R}^m} h^2(x_1,..., x_m) \prod_{l=1}^{m} dF(x_l) - \theta^2(F)$$

$$\rho(i) = 2\left[\int_{\mathbb{R}^m} h(x_1,..., x_m)\, h(x_{m+1},..., x_{2m})\, dF^*_i(x_1, x_{m+1}) \times \prod_{l=2}^{m} dF(x_l) \prod_{p=m+2}^{2m} dF(x_p) - \theta^2(F)\right], \qquad \forall i \geqslant 2$$

$$\psi(i, i) = \int_{\mathbb{R}} (h^{(i)}_{1,n}(x))^2\, dF_i(x) - \left(\int_{\mathbb{R}} h^{(i)}_{1,n}(x)\, dF_i(x)\right)^2, \qquad \forall i \geqslant 1$$

$$\psi(i, j) = 2\int_{\mathbb{R}^2} (h^{(j)}_{1,n}(x)\, h^{(j+i)}_{1,n}(y))\, dF_{i,i+j}(x, y) - \left(\int_{\mathbb{R}} h^{(j)}_{1,n}(x)\, dF_j(x)\right)\left(\int_{\mathbb{R}} h^{(i+j)}_{1,n}(y)\, dF_{i+j}(y)\right), \qquad i < j$$

Then

$$\begin{aligned}
&\left|E\left(\left(\sum_{i=1}^{n} X^*_{ni}\right)^2 \Big/ n\right) - \sigma^2\right| \\
&= \left|\frac{1}{n}(n^{-[m-1]})^2 \sum_{i=0}^{n-1} \sum_{j=1}^{n-i} \psi(i, j) - \sum_{i=1}^{\infty} \rho(i)\right| \\
&\leqslant \left|\frac{1}{n}(n^{-[m-1]})^2 \sum_{i=0}^{n-1} \sum_{j=1}^{n-i} \psi(i, j) - \frac{1}{n} \sum_{i=1}^{\infty} (n-i)\, \rho(i)\right| \\
&\quad + \sum_{i=n+1}^{\infty} |\rho(i)| + \sum_{i=1}^{n} \sum_{k=i}^{\infty} |\rho(k)| \\
&= |A_n| + B_n + C_n.
\end{aligned}$$

From condition in Eq. (3.2), it follows that $|A_n| = o(1)$.

From condition in Eq. (2.24), we deduce that $|\rho(i)| \leqslant$

$(\alpha(i))^{\delta'/(2+\delta')}M_0^{2/(2+\delta')}$ which implies, by using condition in Eq. (2.1), that $B_n \to 0$ and $C_n \to 0$ as $n \to \infty$. Hence we have $|E(\sum_{i=1}^n X_{ni}^*)^2/n - \sigma^2| = o(1)$ and Theorem 3.1 is proved.

4. A LAW OF THE ITERATED LOGARITHM AND ALMOST SURE REPRESENTATION FOR $\hat{F}_n(U_n)$

In what follows we make the following assumptions on F and g.

$$\{F_i', 1 \geqslant 1\} \text{ is uniformly bounded in a neighborhood of } \theta(F) \tag{4.1}$$

$$\int_{-\infty}^{\infty} |t|\ g(t)\, dt < \infty \tag{4.2}$$

Proposition 4.1. Under assumptions in Eqs. (4.1) and (4.2), we have for any $v(0 < v < \frac{1}{8})$

$$\sup_{|x-\theta(F)| \leqslant d_n} |\hat{F}_n(x) - \hat{F}_n(\theta(F)) - \bar{F}_n(x) + \bar{F}_n(\theta(F))| = O(n^{-3/4+v}) + O(\alpha_n) \tag{4.3}$$

a.s. as $n \to \infty$.

Proof. We proceed as in Theorem 3.2 of Lea and Puri.[6] Suppose that $\{F_i', i \geqslant 1\}$ is uniformly bounded in $(\theta(F)-\delta, \theta(F)+\delta)$ for some $\delta > 0$. Note that $|x-\theta(F)| \leqslant d_n \Rightarrow |x-\theta(F)| \leqslant \frac{1}{2}\delta$ for large n. Now,

$$\begin{aligned} |E(\hat{F}_n(x)) - \bar{F}_n(x)| &\leqslant \int_{-\infty}^{+\infty} |\bar{F}_n(x-t) - \bar{F}_n(x)|\ g_n(t)\, dt \\ &\leqslant \int_{|t|<\delta/2} |\bar{F}_n(x-t) - \bar{F}_n(x)|\ g_n(t)\, dt + \int_{|t| \geqslant \delta/2} g_n(t)\, dt \end{aligned}$$

But

$$\begin{aligned} \int_{|t| \geqslant \delta/2} g_n(t)\, dt &= \int_{|y| \geqslant \delta\alpha_n^{-1}/2} g(y)\, dy \\ &\leqslant \frac{2\alpha_n}{\delta} \int_{|y| \geqslant \delta\alpha_n^{-1}/2} |y|\ g(y)\, dy = O(\alpha_n) \qquad \text{as} \quad n \to \infty \end{aligned} \tag{4.4}$$

and by Taylor's theorem

$$\int_{|t|<\delta/2} |\bar{F}_n(x-t)-\bar{F}_n(x)|\; g_n(t)\,dt$$

$$= \int_{|t|<\delta/2} |\bar{F}'_n(x_t)|\; |t|\; g_n(t)\,dt$$

$$\leqslant \text{const.}\; \alpha_n \int_{-\infty}^{+\infty} |y|\; g(y)\,dy = O(\alpha_n) \qquad \text{as} \quad n\to\infty \tag{4.5}$$

where x_t is between x and t. Combining Eqs. (4.4) and (4.5), we have

$$|E(\hat{F}_n(x))-\bar{F}_n(x)| = O(\alpha_n) \qquad \text{as} \quad n\to\infty \tag{4.6}$$

Moreover, for $|x|\leqslant d_n$

$$|\hat{F}_n(x)-\hat{F}_n(\theta(F))-E(\hat{F}_n(x))+E(\hat{F}_n(\theta(F)))|$$

$$= \left|\int_{-\infty}^{+\infty} (\tilde{F}_n(x-t)-\tilde{F}_n(\theta(F)-t)-\bar{F}_n(x-t)+\bar{F}_n(\theta(F)-t))\; g_n(t)\,dt\right|$$

$$\leqslant \int_{|t|\geqslant\delta/2} g_n(t)\,dt$$

$$+\int_{|t|<\delta/2} |\tilde{F}_n(x-t)-\tilde{F}_n(\theta(F)-t)-\bar{F}_n(x-t)+\bar{F}_n(\theta(F)-t)|\; g_n(t)\,dt \tag{4.7}$$

But

$$\int_{|t|\geqslant\delta/2} g_n(t)\,dt = \int_{|y|\geqslant\delta/2} g(y)\,dy$$

$$\leqslant \frac{2\alpha_n}{\delta}\int_{|y|\geqslant\delta/2\alpha_n} |y|\; g(y)\,dy = O(\alpha_n) \qquad \text{as} \quad n\to\infty \tag{4.8}$$

and the second term of the last inequality in Eq. (4.7) is bounded by

$$\sup_{|t-\theta(F)|\leqslant\delta/2} K_n(t) = O(n^{-3/4+v}) \text{ a.s.} \qquad \text{as} \quad n\to\infty, \tag{4.9}$$

by Lemma 2.1.

From Eqs. (4.7)–(4.9), it follows that

$$\sup_{|x-\theta(F)|\leqslant d_n} |\hat{F}_n(x)-\hat{F}_n(\theta(F))-E(\hat{F}_n(x))+E(\hat{F}_n(\theta(F))|$$

$$= O(n^{-3/4+v})+O(\alpha_n) \text{ a.s.} \qquad \text{as} \quad n\to\infty \tag{4.10}$$

The proof now follows from Eqs. (4.6), (4.10), and the following inequality

$$\begin{aligned}|\hat{F}_n(x) - \hat{F}_n(\theta(F)) - \bar{F}_n(x) + \bar{F}_n(\theta(F))| \\ \leqslant |\hat{F}_n(x) - \hat{F}_n(\theta(F)) - E(\hat{F}_n(x)) + E(\hat{F}_n(\theta(F)))| \\ + |E(\hat{F}_n(x)) - \bar{F}_n(x) + E(\hat{F}_n(\theta(F)))| - \hat{F}_n(\theta(F))|\end{aligned}$$

Theorem 4.1. Suppose that $F_i''(\theta(F))\ \forall i \geqslant 1$ exists and is finite. Then, under the conditions of Theorem 3.1 and Eq. (4.2), we have

$$\hat{F}_n(U_n) = \hat{F}_n(\theta(F)) + m\,\frac{\bar{F}'_n(\theta(F))}{n} \sum_{i=1}^{n} X^*_{ni} + R_n \tag{4.11}$$

where X^*_{ni} is defined in Eq. (3.5) and

$$R_n = O(n^{-3/4+v}) + O(\alpha_n) \text{ a.s.}$$

as $n \to \infty$ and $0 < v < \frac{1}{8}$.

Proof. First, using Theorem 3.1, we obtain

$$|U_n - \theta(F)| \leqslant (1+\varepsilon)\, 2^{1/2} m\sigma n^{-1/2} (\log\log n)^{1/2} \text{ a.s.} \quad \text{as} \quad n \to \infty, \quad \text{for some} \quad \varepsilon > 0 \tag{4.12}$$

Next, by Proposition 4.1, we get

$$\hat{F}_n(U_n) - \hat{F}_n(\theta(F)) = \bar{F}_n(U_n) - \bar{F}_n(\theta(F)) + O(n^{-3/4+v}) + O(\alpha_n) \tag{4.13}$$

a.s. as $n \to \infty$. Now using Young's form of Taylor's theorem, and Eq. (4.12), we obtain,

$$\bar{F}_n(U_n) - \bar{F}_n(\theta(F)) = (U_n - \theta(F))\, \bar{F}'_n(\theta(F)) + O(n^{-1} \log\log n). \tag{4.14}$$

The proof follows by using Eqs. (4.13), (4.14), and Lemma 2.4.

Theorem 4.2. Suppose that $F_i''(\theta(F))$ exists for $\forall i \geqslant 1$ and is finite, the conditions of Theorem 3.1 and condition Eq. (4.2) are satisfied, $\theta(F) > 0$, and $\int_{-\infty}^{\infty} y^2 g(y)\, dy < \infty$. Then

$$\hat{F}_n(U_n) = n^{-1} \sum_{i=1}^{n} Y_{ni} + R_n \tag{4.15}$$

where $Y_{ni} = u(\theta(F) - X_i) + m\bar{F}'_n(\theta(F))\, X^*_{ni}$ and

$$R_n = O(n^{-3/4+v}) + O(n\alpha_n^2 \log\log n)^{-1} \text{ a.s.}$$

as $n \to \infty$ and $0 < v < \frac{1}{8}$.

Proof. Applying Theorem 4.1, we obtain

$$\hat{F}_n(U_n) = \hat{F}_n(\theta(F)) + m\frac{\bar{F}_n(\theta(F))}{n}\sum_{i=1}^{n} X^*_{ni} + R_n \tag{4.16}$$

where

$$R_n = O(n^{-3/4+v}) + O(\alpha_n) \text{ a.s. as } n \to \infty$$

Next using Lemma 2.2, we obtain

$$\hat{F}_n(\theta(F)) = \tilde{F}_n(\theta(F)) + O(n^{-3/4+v}) + O(n\alpha_n^2(\log\log n)^{-1}) \text{ a.s.} \tag{4.17}$$

The proof follows using Eqs. (4.16) and (4.17).

Theorem 4.3. Under the conditions of Theorem 4.2, if $\alpha_n = o(n^{-3/4+v}(\log\log n)^{1/2})$ for $0 < v < \frac{1}{8}$ and $F'(\theta(F))$ exists and is finite, then

$$\overline{\lim}\, n^{1/2}\frac{\hat{F}_n(U_n) - F(\theta(F))}{\sqrt{2\sigma^2_* \log\log n}} = 1 \text{ a.s.} \tag{4.18}$$

where $\sigma^2_* = \int_{\mathbb{R}} A^2(x)\, F(dx) + 2\sum_{k=1}^{\infty} \int_{\mathbb{R}^2} A(x)\, A(y)\, F^*_k dx, dy)$ and $A(x) = u(\theta(F) - x) + mF'(\theta(F))[h_1(x) - \theta(F)]$, and $h_1(x) = \int h(x, x_2,..., x_m) \prod_{i=2}^{m} dF(x_i)$.

Proof. The proof follows from Theorem 4.2 and Lemma 5.6 (in the Appendix).

5. APPENDIX

Let $p \geq 2$ and $1 \leq i_1 < i_2 < \cdots < i_p$ be arbitrary integers. For any $j (1 \leq j \leq p-1)$, $P_j^{(i_1,...,i_p)}$ is the probability measure defined by

$$\begin{aligned} &P_j^{(i_1,...,i_p)}(A^{(j)} \times B^{(p-j)}) \\ &\quad = P[(X_{i_1},..., X_{i_j}) \in A^{(j)}]\, P[(X_{i_{j+1}},..., X_{i_p}) \in B^{(p-j)}] \end{aligned}$$

and $P_0^{(i_1,...,i_p)}$ is the probability measure defined by

$$P_0^{(i_1,...,i_p)}(A^{(p)}) = P[(X_{i_1},..., X_{i_p}) \in A^{(p)}]$$

for any $A^{(j)} \in \sigma(X_{i_1},..., X_{i_j})$ $(1 \leqslant j \leqslant p)$ and any $B^{(p-j)} \in \sigma(X_{i_{j+1}},..., X_{i_p})$, $(1 \leqslant j \leqslant p-1)$.

Lemma 5.1. For every $p \geqslant 1$ and $(i_1,..., i_p)$ such that $1 \leqslant i_1 < i_2 < \cdots < i_p$ and any $j (1 \leqslant j \leqslant p-1)$, let $h(x_1,..., x_p)$ be a Borel function such that

$$\int_{\mathbb{R}^p} |h(x_1,..., x_p)|^{1+\delta} \, dP_j^{(i_1,..., i_p)} \leqslant M$$

for some $\delta > 0$, then

$$\left| \int_{\mathbb{R}^p} h(x_1,..., x_p) \, dP_0^{(i_1,..., i_p)} - \int_{\mathbb{R}^p} h(x_1,..., x_p) \, dP_j^{(i_1,..., i_p)} \right| \leqslant 4M^{1/(2+\delta)} \beta^{\delta/(1+\delta)} (i_{j+1} - i_j) \tag{5.1}$$

As a special case, if $h(x_1,..., x_p)$ is bounded, say, $|h(x_1,..., x_p)| \leqslant M^*$, then we can replace the right side of Eq. (5.1) by $2M^* \beta(i_{j+1} - i_j)$.

Proof. Follows from Lemma 1 of Yoshihara.[19]

Lemma 5.2. Let $\{Y_i, i \geqslant 1\}$ be a sequence of absolutely regular random variables with mean 0. Assume that $\sup_{i \geqslant 1} |Y_i| \leqslant M_0$ and $\sup_{i \geqslant 1} [\text{Var } Y_i]^{1/2} \leqslant M$. Then, for any a $(0 < a < 1)$ and $m \leqslant n$ sufficiently large,

$$P\left[n^{-1} \left| \sum_{i=1}^{n} Y_i \right| \geqslant u \right] \leqslant 2m \exp\left\{ -\frac{ak}{5M_0} \right\} + 2n\beta(m) \tag{5.2}$$

where $k = [n/m]$.

Proof. Let $\{\tilde{Y}_i, i \geqslant 1\}$ be a sequence of independent random variables such that for any i, $\tilde{Y}_i$ has the same distribution as Y_i. Then, from Collomb's[2b] extension of the Bernstein inequality (see Bennett[1]), we have for any $I \subset \{1,..., n\}$ such that *card* $I = m$

$$P\left[\left| \sum_{i \in I} \tilde{Y}_i \right| \geqslant v \right] \leqslant 2 \exp\left\{ -\frac{vk^{-a}}{5M_0} \right\} \tag{5.3}$$

Let now $S^{(j)} = Y_j + Y_{j+m} + \cdots + Y_{j+k_j m}$ $(j = 1,..., m)$ where for each j $(1 \leqslant j \leqslant m)$ $k_j = k_{j,n}$ is the largest integer for which $j + k_j m \leqslant n$. Then

$$P\left[\left| \sum_{i=1}^{n} Y_i \right| \geqslant v \right] \leqslant P\left[\sum_{j=1}^{m} |S^{(j)}| \geqslant v \right] \leqslant \sum_{j=1}^{m} P[|S^{(j)}| \geqslant m^{-1} v] \tag{5.4}$$

For each j $(1 \leqslant j \leqslant m-1)$, let A_j be the Borel subset of $\mathbb{R}^{k_j+1}$ defined by

$$A_j = \{(y_0,..., y_{k_j}): |y_0 + \cdots + y_{k_j}| \geqslant m^{-1}v\}$$

and put

$$g_j(y_0,..., y_{k_j}) = \begin{cases} 1 & \text{if } (y_0,..., y_{k_j}) \in A_j \\ 0 & \text{otherwise} \end{cases}$$

Then $|g_j(y_0,..., y_{k_j})| \leqslant 1$ and after iterating $k_j + 1$ times the inequality of Eq. (5.1) for the particular case $M^* = 1$, we get

$$\begin{aligned} P[|S^{(j)}| \geqslant m^{-1}v] &= E(g_j(Y_j,..., Y_{j+k_jm})) \\ &\leqslant \int_{\mathbb{R}^{k_j+1}} g_j(y_0,..., y_{k_j})\, dF_j(y_0) \cdots dF_{j+k_jm}(y_{k_j}) \\ &\quad + 2(k_j+1)\,\beta(m) \\ &= P\left[\left|\sum_{i \in I} \tilde{Y}_i\right| \geqslant m^{-1}v\right] + 2(k_j+1)\,\beta(m) \\ &\leqslant 2\exp\left\{-\frac{m^{-1}vk_j^{-a}}{5M_0}\right\} \\ &\quad + 2(k_j+1)\,\beta(m) \text{ (from Eq. (5.3))}. \end{aligned} \tag{5.5}$$

Now from Eqs. (5.4) and (5.5), we get

$$P\left[\left|\sum_{i=1}^{n} Y_i\right| \geqslant v\right] \leqslant 2m \exp\left\{-\frac{uk_j^{1-u}}{5M_0}\right\} + 2n\beta(m) \tag{5.6}$$

If in Eq. (5.6), we replace v by un, we obtain Eq. (5.2).

The following lemma is the Lemma 3 of Chow and Teicher[2a] [p. 298].

Lemma 5.3. Let $\varphi_n^*(t)$ be the characteristic function of $S_n = \sum_{i=1}^n Z_j$ where $\{Z_j, j \geqslant 1\}$ are independent (not necessarily stationary) r.v.'s with zero means, and variances $\sigma_j^2, j \geqslant 1$. If $\Gamma_n^{2+\delta} = \sum_{j=1}^n \gamma_j^{2+\delta}$ and $s_n^2 = \sum_{j=1}^n \sigma_j^2$ where $\gamma_j^{2+\delta} = E\,|Z_j - E(Z_j)|^{2+\delta}$, then for $0 \leqslant \delta \leqslant 1$

$$\left|\varphi_n^*\left(\frac{t}{s_n}\right) - e^{-t^2/2}\right| \leqslant 3\left|\frac{t\Gamma_n}{s_n}\right|^{2+\delta} e^{-t^2/2} \qquad \text{for} \quad |t| < \frac{s_n}{2\Gamma_n} \tag{5.7}$$

Lemma 5.4. Let $\{Y_{ni}, 1 \leqslant i \leqslant n, n \geqslant 1\}$ be a sequence of absolutely regular random variables (with means 0) satisfying the strong mixing condition. Suppose that $\sum_{m \geqslant 1} \alpha(m) < \infty$ and there exists $\sigma^2 > 0$ such that

$$\operatorname{Var}\left(\sum_{i=1}^{n} Y_{ni}\right) = n\sigma^2(1 + o(1)) \tag{5.8}$$

Then the process $\{Y_{ni}, 1 \leqslant i \leqslant n, n \geqslant 1\}$ obeys the law of the iterated logarithm, that is $\limsup_{n \to \infty} |\sum_{i=1}^{n} (Y_{ni}/\chi(n))| = 1$ a.s., if the following requirements are fulfilled for some $\rho > 0$ and for all sufficiently large n:

$$\sup_{x \in \mathbb{R}} \left| P\left(\sum_{i=1}^{n} Y_{ni} < x\sigma\sqrt{n}\right) - \Phi(x)\right| = O\left(\frac{1}{(\log n)^{1+\rho}}\right) \tag{5.9}$$

where

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2}\, dt$$

and

$$P\left[\max_{1 \leqslant j \leqslant n} \left|\sum_{i=1}^{j} Y_{ni}\right| \geqslant b\chi(n)\right] = O\left(\frac{1}{(\log n)^{1+\rho}}\right) \tag{5.10}$$

where $b > 1$ is an arbitrary number and $\chi(n) = (2\sigma^2 n \log\log n)^{1/2}$.

This Lemma was proved by Oodaira and Yoshihara[8] [Th. 1] for the stationary case. The proof is similar for the nonstationary case and is therefore omitted.

Lemma 5.5. For a sequence of random variables satisfying the strong mixing condition, the condition in Eq. (5.10) holds if Eq. (5.9) holds and there exists a function $r = r(n)$ such that $r(n) \leqslant n (n \geqslant 1)$, $r(n) \to \infty$ as $n \to \infty$ and

$$\max\left\{\frac{n}{r} P\left[\sum_{i=1}^{r} |Y_{ni}| \geqslant \varepsilon\chi(n)\right],\ \frac{n}{r}\alpha(r)\right\} = O\left(\frac{1}{(\log n)^{1+\rho}}\right) \tag{5.11}$$

for any ε $(0 < \varepsilon < (b-1)/b$ where $b > 1$ is an arbitrarily fixed number).

Proof. From Oodaira and Yoshihara[8][Remark 1, Th. 1].

Lemma 5.6. Let $\{Y_{ni}, 1 \leqslant i \leqslant n, n \geqslant 1\}$ be a sequence of absolutely regular random variables satisfying the strong mixing condition with means 0. Suppose for any n, m such that $n \geqslant m$ and any $J \subset \{1,..., n\}$ such that $card\ J = m$

$$E\left(\sum_{j \in J} Y_{nj}\right)^2 = m\sigma^2(1 + o(1)) \tag{5.12}$$

Then the process $\{Y_{ni},\ 1 \leqslant i \leqslant n, n \geqslant 1\}$ obeys the law of the iterated logarithm if the following requirements are fulfilled for some δ and δ' such that

$$\sup_{n \geqslant 1} \max_{1 \leqslant i \leqslant n} E\,|Y_{ni}|^{2+\delta} = M < \infty \tag{5.13}$$

$$\sum_{m=1}^{\infty} \{\alpha(m)\}^{\delta'/(2+\delta')} < \infty \tag{5.14}$$

Proof. Define

$$f_N(x) = \begin{cases} x & \text{if } |x| \leqslant N \\ 0 & \text{otherwise} \end{cases}$$

We define $\bar{f}_N(x) = x - f_N(x)$. Let $N = n^{1/2(1+\delta')}(\log n)^{-3}$ and

$$r(n) = [n^{\delta'/2(1+\delta')}(\log n)^3]$$

For convenience and without loss of generality, we denote Y_i for Y_{ni}. Then for any $b > 0$

$$\begin{aligned}
&\frac{n}{r} P[|Y_1| + \cdots + |Y_r| \geqslant b_\chi(n)] \\
&= \frac{n}{r} P\left[\sum_{i=1}^{r} |\bar{f}_N(Y_i)| \geqslant b\chi(n)\right] \\
&\leqslant \frac{n}{b^2\{\chi(n)\}^2 r} E\left(\sum_{i=1}^{r} |\bar{f}_N(Y_i)|\right)^2 \\
&\leqslant \frac{n}{b^2\{\chi(n)\}^2 r}\left\{\sum_{1 \leqslant i,j \leqslant r} E[|\bar{f}_N(Y_i)| \cdot |\bar{f}_N(Y_j)|]\right\} \\
&\leqslant \frac{n}{b^2\{\chi(n)\}^2 r}\left\{2\sum_{i=0}^{r-1}\sum_{j=1}^{r-i} E[|\bar{f}_N(Y_i)| \cdot |\bar{f}_N(Y_{j+i})|]\right\} \\
&\leqslant \frac{n}{b^2\{\chi(n)\}^2 r}\left\{2\sum_{i=1}^{r} E[|\bar{f}_N(Y_i)|^2]\right. \\
&\quad \left. + 2\sum_{i=1}^{r-1}\sum_{j=1}^{r-i} E[|\bar{f}_N(Y_i)| \cdot |\bar{f}_N(Y_{j+i})|]\right\}
\end{aligned}$$

$$\leqslant \frac{2n}{b^2\{\chi(n)\}^2 r}\left\{\frac{1}{N^\delta}\sum_{i=1}^{r} E[|\bar{f}_N(Y_i)|^{2+\delta}]\right.$$

$$+12\sum_{i=1}^{r-1}(\alpha(i))^{\delta'/(2+\delta')}\sum_{j=1}^{r-i} E[|\bar{f}_N(Y_i)|^{2+\delta'}]^{1/(2+\delta')}$$

$$\left.\times E[|\bar{f}_N(Y_{j+i})|^{2+\delta'}]^{1/(2+\delta')}\right\} \text{(from Lemma 5.7)}$$

$$\leqslant \frac{2n}{b^2\{\chi(n)\}^2 r}\left\{\frac{r}{N^\delta}M+12\sum_{i=1}^{r-1}(\alpha(i))^{\delta'/(2+\delta')}\sum_{j=1}^{r-i}\frac{1}{N^{2(\delta-\delta')}}\right.$$

$$\left.\times E[|\bar{f}_N(Y_i)|^{2+\delta'}]^{1/(2+\delta')} E[|\bar{f}_N(Y_{j+i})|^{2+\delta'}]^{1/(2+\delta')}\right\}$$

$$\leqslant \frac{2n}{b^2\{\chi(n)\}^2 r}\left\{\frac{rM}{N^\delta}+\frac{12r\sum_{i=1}^{r-1}(\alpha(i))^{\delta'/(2+\delta')}}{N^{2(\delta-\delta')}}M^{2/(2+\delta')}\right\}$$

$$= O(n^{-\gamma}) \tag{5.15}$$

holds for some $\gamma > 0$ and

$$\frac{n}{r}\alpha(r) = \frac{n}{r} O\,(r^{-(2+\delta')/\delta'}) = O\left(\frac{1}{(\log n)^3}\right)$$

From Lemmas 5.4 and 5.5 it suffices to show

$$\sup_{x\in\mathbb{R}}\left|P\left(\sum_{i=1}^{n} Y_i < \sqrt{n}\,\sigma x\right) - \Phi(x)\right| = O\left(\frac{1}{(\log n)^3}\right) \tag{5.16}$$

Define p, q and k by $p(n) = [n^{1/2+\alpha}]$, $q(n) = [n^{1/2-\alpha}]$ and $k(n) = [n/(p+q)]$ where α is a small number. Let $N' = n^{\delta'/16(1+\delta')}$ if $0 \leqslant \delta \leqslant 2$ and $N' = n^{1/16(1+\delta')}$ if $\delta > 2$. Put

$$S'_n = \frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{n}(f_{N'}(Y_j) - E(f_{N'}(Y_j)))$$

$$S'' = \frac{1}{\sigma\sqrt{n}}\sum_{j=1}^{n}(\bar{f}_{N'}(Y_j) - E(\bar{f}_{N'}(Y_j)))$$

$$Z_i = \sum_{j=1}^{p}(f_{N'}(Y_{(i-1)(p+q)+j})) - E(f_{N'}(Y_{(i-1)(p+q)+j})), \qquad i = 1,\dots,k$$

$$T'_n = \frac{1}{\sqrt{n}\,\sigma}\sum_{i=1}^{k} Z_i$$

$$T''_n = S'_n - T'_n$$

Then, it can easily be proved that for some $\gamma > 0$

$$E\,|S''_n|^2 \leqslant \frac{1}{\sigma^2}\left(\frac{1}{N'^{\delta}}\,M + \frac{K}{N'^{2(\delta-\delta')(2+\delta')}}\,M^{2/(2+\delta')}\right) = O(n^{-\gamma}) \qquad (5.17)$$

(where M is defined in Eq. (5.13)), K is some constant > 0. $E\,|T''_n|^2 = O(n^{-\gamma})$, $|E(T'^2_n) - 1| = O(n^{-\gamma})$.

Now, let $\varphi_n(t)$ be the characteristic function of $S_n/\sigma\sqrt{n}$ where $S_n = \sum_{i=1}^{n} Y_i$. Then

$$\begin{aligned} |\varphi_n(t) - e^{-t^2/2}| &\leqslant |\varphi_n(t) - E(e^{itS'_n})| + |E(e^{itS'_n}) - E(e^{itT'_n})| \\ &\quad + \left| E(e^{itT'_n}) - \prod_{j=1}^{k} E(e^{itZ_j/\sqrt{kE(Z_j)^2}}) \right| \\ &\quad + \left| (e^{-t^2/2}) - \prod_{j=1}^{k} E(e^{itZ_j/\sqrt{kE(Z_j)^2}}) \right| \\ &= I_1 + I_2 + I_3 + I_4, \text{ say.} \end{aligned} \qquad (5.18)$$

Then, we have

$$I_1 \leqslant |t|\; E(|S''_n|) = O(n^{-\gamma/2}), \qquad I_2 \leqslant |t|\; E(|T''_n|) = O(n^{-\gamma/2}) \qquad (5.19)$$

From Lemma 5.3, we have

$$\begin{aligned} &\left| e^{-t^2/2} - \prod_{j=1}^{k} E(e^{itZ_j/\sqrt{kE(Z_j)^2}}) \right| \\ &\quad \leqslant 3\,\frac{|t|^{2+\delta}}{k^{\delta/2}}\, e^{-t^2/2} \sup_{1\leqslant j\leqslant k} E\,|Z_j|^{2+\delta} \min_{1\leqslant j\leqslant k} \frac{1}{(E\,|Z_j|^2)^{(2+\delta)/2}} \\ &\quad \text{for } |t| \leqslant \frac{(\sum_{j=1}^{k} E(Z_j)^2)^{1/2}}{2\sum_{j=1}^{k} E\,|Z_j|^{2+\delta}} \end{aligned} \qquad (5.20)$$

Since

$$\begin{aligned} E(Z_j^4) &\leqslant K(N')^4\, p^2 \sum_{j=1}^{p} j\alpha(j) \leqslant K(N')^4\, p^2 \sum_{j=1}^{p} j^{-2/\delta'} \\ &\leqslant K(N')^4\, p^2 \max(1, p^{-2/\delta}) \end{aligned}$$

and $E(Z_j^2) = p\sigma^2(1 + o(1))$ (from condition Eq. (5.12)), for all sufficiently large n, we have

$$\begin{aligned} &\sup_{1\leqslant j\leqslant k} E\,|Z_j|^{2+\delta} \frac{1}{k^{\rho/2} \min_{1\leqslant j\leqslant k}(E(Z_j^2))^{(2+\rho)/2}} \\ &\quad \leqslant \sup_{1\leqslant j\leqslant k} (E(Z_j^4))^{(2+\delta)/4} \frac{2^{(2+\rho)/2}}{k^{\rho/2}(p\sigma^2)^{(2+\rho)/2}} = O(n^{-\gamma}) \end{aligned} \qquad (5.21)$$

holds for all sufficiently large n where $\rho = \min(2, \delta)$ and γ is a positive number.

We deduce that

$$\left| e^{-t^2/2} - \prod_{j=1}^{k} E(e^{itZ_j/\sqrt{kE(Z_j)^2}}) \right| \leqslant K' n^{-\gamma} |t|^{2+\delta} e^{-t^2/2} \tag{5.22}$$

holds for all sufficiently large n and for all t such that

$$|t| \leqslant \frac{1}{3} n^{1/2 - 1/(1+2\delta)} \sigma \times \frac{1}{\min_{1 \leqslant i \leqslant n} E\,|Y_i|^{2+\delta}}$$

Now from Eq. (5.14) and using Lemma 5.7, we obtain

$$\left| E(e^{itT'_n}) - \prod_{j=1}^{k} E(e^{itZ_j/\sqrt{kE(Z_j^2)}}) \right| \leqslant k\alpha(q) = n^{1/2-\alpha} o((n^{1/2-\alpha})^{-(2+\delta')/\delta'}) \tag{5.23}$$

From Eqs. (5.17)–(5.23) it follows that

$$\begin{aligned}
&\left| P\left(\sum_{i=1}^{n} Y_i < \sqrt{n}\,\sigma x \right) - \Phi(x) \right| \\
&\quad \leqslant \int_{-(\log n)^3}^{(\log n)^3} \left| \frac{\varphi_n(t) - e^{-t^2/2}}{t} \right| dt + \frac{K_1}{(\log n)^3} \\
&\quad \leqslant \int_{-(\log n)^3}^{(\log n)^3} K_2 n^{-\gamma} |t|^{1+\delta}\, dt + \int_{-(\log n)^3}^{(\log n)^3} \{E(|S''_n|) + E(|T''_n|)\}\, dt \\
&\qquad + K_3 \left\{ \int_{0 \leqslant |t| \leqslant n^{-1/4}} dt + \int_{n^{-1/4} \leqslant |t| \leqslant (\log n)^3} \frac{k\alpha(q)}{|t|}\, dt + \frac{K_4}{(\log n)^3} \right\} \\
&\quad = O\left(\frac{1}{(\log n)^3} \right)
\end{aligned} \tag{5.24}$$

where K_i, $i = 1, \dots, 4$ are constants >0, and Lemma 5.6 is proved.

Lemma 5.7. Let $\{Y_{ni},\ 1 \leqslant i \leqslant n, n \geqslant 1\}$ be a sequence of random variables satisfying the strong mixing condition with coefficient $\alpha(m)$. Let U be $\sigma(Y_{ni}, 1 \leqslant i \leqslant j \leqslant n)$-measurable and V be $\sigma(Y_{ni}, j+m \leqslant i \leqslant n)$-measurable. If $E\,|U|^2 < \infty$ and $E\,|V|^2 < \infty$ where $r^{-1} + p^{-1} + s^{-1} = 1$ $(1 \leqslant r, s, p < \infty)$ then,

$$|E(UV) - E(U)\,E(V)| \leqslant 12(\alpha(m))^{p^{-1}} \{E\,|U|^r\}^{r^{-1}} \{E\,|V|^s\}^{s^{-1}} \tag{5.25}$$

Proof. This is Proposition 2.8 of Doukhan and Portal.[3]

REFERENCES

1. Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57**, 33–45.
2a. Chow, Y. S. and Teicher, H. (1978). *Probability Theory*. Springer Verlag.
2b. Collomb, G. (1984). Propriétés de convergence presque complète du prédicteur à noyau. *Z. Wahrsch. Verw. Gebiete* **66**, 441–460.
3. Doukhan, P. and Portal, F. (1987). Principe d'invariance faible pour la fonction de répartition empirique dans un cadre multidimensionnel et mélangeant. *Prob. and Math. Stat.* **8**, (2), 117–132.
4. Harel, M. and Puri, M. L. (1991). Théorème central limite de l'estimateur de densité et de l'erreur quadratique intégrée pour des variables aléatoires nonstationnaires. *C.R.Academy of Sciences* Série 1, **312**, 145–148.
5. Harel, M. and Puri, M. L. (1989). Weak convergence of the *U*-statistic and weak invariance of the one sample rank order statistic for Markov processes and ARMA models. *J. of Mult. Anal.* **31**, 258–265.
6. Lea, C. and Puri, M. L. (1988). Asymptotic properties of perturbed empirical distribution functions evaluated at a random point. *J. Statistical Planning and Inference* **19**, 201–215.
7. Nadaraya, E. A. (1964). Some new estimates for distribution function. *Theory Probab. Appl.* **9**, 497–500.
8. Oodaira, H. and Yoshihara, K. (1971). The law of the iterated logarithm for stationary processes satisfying mixing conditions. *Kōdai Math. Sem. Rep.* **23**, 311–334.
9. Puri, M. L. and Ralescu, S. (1983). On Berry-Esséen rates, a law of the iterated logarithm and an invariance principle for the proportion of the sample below the sample mean. *J. Multivariate Anal.* **14**, 231–247.
10. Puri, M. L. and Ralescu, S. (1986). Central limit theorem for perturbed empirical distribution functions evaluated at a random point. *J. Multivariate Anal.* **16**, 273–279.
11. Schuster, E. F. (1969). Estimation of probability density function and its derivatives. *Ann. Math. Statist.* **40**, 1187–1195.
12. Scott, D. W., Tapia, R. A. and Thompson, J. R. (1977). Kernel density estimation revisited. *J. Nonlinear Analysis Theory Meth. Applic.* **1**, 339–372.
13. Sun, S. (1993). Asymptotic behavior of the perturbed empirical distribution function evaluated at a random point for absolutely regular sequences *J. Multivariate Anal.* (to appear).
14. Watson, G. S. and Leadbetter, M. R. (1964). Hazard analysis II. Ser. A, *Sankhyā*, **26**, 101–116.
15. Wertz, W. (1978). *Statistical Density Estimation: A Survey*. Vandenhoeck and Ruprecht, Göttingen.
16. Winter, B. B. (1973). Strong uniform consistency of integrals of density estimators. *Canad. J. Statist.* **1**, 247–253.
17. Winter, B. B. (1979). Convergence rate of perturbed empirical distribution functions. *J. Appl. Probab.* **16**, 163–173.
18. Yamato, H. (1973). Uniform convergence of an estimator of a distribution function. *Bull. Math. Statist.* **15**, 69–78.
19. Yoshihara, K. (1976). Limiting behavior of *U*-statistics for stationary absolutely regular processes. *Z. Wahrsch. Verw. Gebiete* **35**, **237–252**.