WEAK CONVERGENCE OF MULTIDIMENSIONAL EMPIRICAL PROCESSES FOR STATIONARY ϕ-MIXING PROCESSES

By

Pranab Kumar Sen

Department of Biostatistics
University of North Carolina, Chapel Hill, N. C.

Institute of Statistics Mimeo Series No. 846

OCTOBER 1972

WEAK CONVERGENCE OF MULTIDIMENSIONAL EMPIRICAL PROCESSES FOR STATIONARY ϕ-MIXING PROCESSES[1]

BY PRANAB KUMAR SEN

University of North Carolina, Chapel Hill

For a stationary ϕ-mixing sequence of stochastic $p(\geq 1)$-vectors, weak convergence of the empirical process (in the J_1-topology on $D^p[0,1]$) to an appropriate Gaussian process is established under a simple condition on the mixing constants $\{\phi_n\}$. Weak convergence for random number of stochastic vectors is also studied.

AMS 1970 Classification: 60F05, 60B10.

Key Words and Phrases: $D^p[0,1]$ space, empirical processes, Gaussian process, random sample size, Skorokhod J_1-topology, tightness, weak convergence.

1) Research sponsored by the Aerospace Research Laboratories, Air Force Systems Command, U. S. Air Force, Contract F33615-71-C-1927. Reproduction in whole or in part permitted for any purpose of the U. S. Government.

1. Introduction. Let $\{\underset{\sim}{X}_i=(X_{i1},\ldots,X_{ip})', -\infty<i<\infty\}$ be a stationary ϕ-mixing sequence of stochastic vectors defined on a probability space (Ω,A,P) with each $\underset{\sim}{X}_i$ having (marginally) a continuous distribution function (df) $F(\underset{\sim}{x})$, $\underset{\sim}{x}\varepsilon R^p$, the $p(\geq 1)$ dimensional Euclidean space. Thus, if $M_{-\infty}^{k}$ and M_{k+n}^{∞} be respectively the σ-fields generated by $\{\underset{\sim}{X}_i, i\leq k\}$ and $\{\underset{\sim}{X}_i, i\geq k+n\}$, and if, $A\varepsilon M_{-\infty}^{k}$ and $B\varepsilon M_{k+n}^{\infty}$, then for all $k(-\infty<k<\infty)$ and non-negative n,

$$(1.1) \qquad |P(A\cap B)-P(A)P(B)| \leq \phi_n P(A),\ \phi_n\geq 0,$$

where ϕ_n is $\downarrow$ in n and $\lim_{n\to\infty}\phi_n=0$. We denote the marginal df of X_{ij} by $F_{[j]}$, let $Y_{ij} = F_{[j]}(X_{ij})$, $1\leq j\leq p$, $\underset{\sim}{Y}_i = (Y_{i1},\ldots,Y_{ip})'$, $-\infty<i<\infty$, and denote the df of $\underset{\sim}{Y}_i$ by

$$(1.2) \qquad G(\underset{\sim}{t}) = P\{\underset{\sim}{Y}_i \leq \underset{\sim}{t}\}, \quad \underset{\sim}{t}\varepsilon E^p,$$

where $E^p = \{\underset{\sim}{t}: \underset{\sim}{0}\leq\underset{\sim}{t}\leq\underset{\sim}{1}\}$ is the unit p-dimensional cube, $\underset{\sim}{0}=(0,\ldots,0)$, $\underset{\sim}{1}=(1,\ldots,1)$, and $\underset{\sim}{a}\leq\underset{\sim}{b}$ means that $a_j\leq b_j$, $1\leq j\leq p$. Note that all the p univariate marginals of G are rectangular [0,1] df, while the bivariate or higher order dfs depend on F, unless $X_{i1},\ldots,X_{ip}$ are independent. Also, note that $G(\underset{\sim}{t})=0$ if at least one of its coordinates is zero. For a sample $\underset{\sim}{X}_1,\ldots,\underset{\sim}{X}_n$ of size n, we define the <u>empirical df</u> for $\underset{\sim}{Y}_1,\ldots,\underset{\sim}{Y}_n$ by

$$(1.3) \qquad G_n(\underset{\sim}{t}) = n^{-1}\sum_{i=1}^{n} c(\underset{\sim}{t}-\underset{\sim}{Y}_i),\ \underset{\sim}{t}\varepsilon E^p,\ n\geq 1,$$

where $c(\underset{\sim}{u})=1$ iff $\underset{\sim}{u}\geq\underset{\sim}{0}$, and 0, otherwise. Also, $G_n(\underset{\sim}{t})=0$, when at least one of $t_1,\ldots,t_p$ is equal to 0. The <u>empirical process</u> $W_n = \{W_n(\underset{\sim}{t}),\ \underset{\sim}{t}\varepsilon E^p\}$ is then defined by

$$(1.4) \qquad W_n(\underset{\sim}{t}) = n^{\frac{1}{2}}[G_n(\underset{\sim}{t})-G(\underset{\sim}{t})],\ \underset{\sim}{t}\varepsilon E^p,\ n\geq 1.$$

For every $n\geq 1$, the process W_n belongs to the space $D^p[0,1]$ of all real valued functions on E^p with no discontinuities of the second kind, and with $D^p[0,1]$, we associate the (extended) Skorokhod J_1-topology. For an excellant exposition of weak convergence of processes on $D^p[0,1]$, we may refer to Neuhaus (1971). Also, for p=1, a detailed account is given in Billingsley (1968).

When the $\underset{\sim}{X}_i$ are independent and identically distributed (iid), i.e., $\phi_n=0$, $\forall\, n\geq 1$, W_n converges in distribution (in the J_1-topology on $D^p[0,1]$) to an appropriate Gaussian process (Brownian bridge for p=1); we may again refer to Neuhaus (1971) who also cites other references. For ϕ-mixing processes, satisfying (1.1), Billingsley (1968, p. 197) established the weak convergence of W_n to a Gaussian function when p=1 and $A_2(\phi)<\infty$, where

(1.5) $$A_k(\phi) = \sum_{n=0}^{\infty}(n+1)^k\, \phi_n^{\frac{1}{2}},\ k\geq 0.$$

Later, Sen (1971) showed that for p=1, $A_1(\phi)<\infty$ insures the weak convergence of W_n. Our first objective is to show that for general $p\geq 1$, the weak convergence of W_n to an appropriate Gaussian process holds under $A_1(\phi)<\infty$.

Let now $\{N_\nu,\ \nu\geq 1\}$ be a sequence of positive integer valued random variables, such that

(1.6) $$\nu^{-1}N_\nu\to\xi,\ \text{in probability, as}\ \nu\to\infty,$$

where ξ is a positive random variable defined on the same probability space (Ω,A,P). For independent $\{\underset{\sim}{X}_i\}$ and p=1, Pyke (1968) has shown that W_{N_ν} converges in law to a Brownain bridge as $\nu\to\infty$. Using certain sub-martingale properties of the Kolmogorov sumpremum of W_n, Sen (1972b) has shown that for general $p\geq 1$ the iid vectors, W_{N_ν} weakly converges to an appropriate Gaussian process. Our second objective is to show that for ϕ-mixing processes, satisfying (1.1), $A_1(\phi)<\infty$ again insure the weak convergence of W_{N_ν} to an appropriate Gaussian process, for every $p\geq 1$.

2. Weak convergence of W_n. Let us write

$$A_k(\phi^2) = \sum_{n=0}^{\infty}(n+1)^k\phi_n, \; k \geq 0. \tag{2.1}$$

Then, $[A_k(\phi)<\infty] \Rightarrow [A_\ell(\phi)<\infty], \forall \ell\leq k$, and $[A_\ell(\phi^2)<\infty], \forall \ell\leq 2k$. Consider now a p-dimensional Gaussian process $W = \{W(\underset{\sim}{t}), \underset{\sim}{t}\varepsilon E^p\}$, where $E[W(\underset{\sim}{t})] = 0$, $\underset{\sim}{t}\varepsilon E^p$, and for every $\underset{\sim}{s}$, $\underset{\sim}{t}\varepsilon E^p$,

$$E[W(\underset{\sim}{s})W(\underset{\sim}{t})] = E\{[c(\underset{\sim}{s}-\underset{\sim}{Y}_1)c(\underset{\sim}{t}-\underset{\sim}{Y}_1)]-G(\underset{\sim}{s})G(\underset{\sim}{t})\} +$$

$$\sum_{k=2}^{\infty}\{c(\underset{\sim}{s}-\underset{\sim}{Y}_1)c(\underset{\sim}{t}-\underset{\sim}{Y}_k) + c(\underset{\sim}{s}-\underset{\sim}{Y}_k)c(\underset{\sim}{t}-\underset{\sim}{Y}_1)-G(\underset{\sim}{s})G(\underset{\sim}{t})\}. \tag{2.2}$$

Note that by Theorem 20.1 of Billingsley (1968, p. 174), the series on the right hand side (rhs) of (2.2) converges when $A_0(\phi)<\infty$.

LEMMA 2.1. Under (1.1) and $A_0(\phi)<\infty$, the finite dimensional distributions of $\{W_n\}$ converge (as $n\to\infty$) to those of W.

Proof. For any (fixed) $m(\geq 1)$ and given $\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_m$ (εE^p), consider an arbitrary linear compound $Z_n(\underset{\sim}{\lambda}) = \sum_{j=1}^{m} \lambda_j W_n(\underset{\sim}{t}_j)$, where $\underset{\sim}{\lambda}=(\lambda_1,\ldots,\lambda_m)\neq\underset{\sim}{0}$. Then, by (1.3) and (1.4),

$$Z_n(\underset{\sim}{\lambda}) = n^{-\frac{1}{2}}\sum_{i=1}^{n}\{\sum_{j=1}^{m} \lambda_j[c(\underset{\sim}{t}_j-\underset{\sim}{Y}_i)-G(\underset{\sim}{t}_j)]\}$$

$$= n^{-\frac{1}{2}}\sum_{i=1}^{n} U_i(\underset{\sim}{\lambda},\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_m), \text{ say,}$$

where the U_i are bounded random variables (with zero expectations), and by (1.1), $\{U_i, -\infty<i<\infty\}$ is also a ϕ-mixing sequence. Therefore by Theorem 20.1 of Billingsley (1968, p. 174), under (1.1) and $A_0(\phi)<\infty$, $Z_n(\underset{\sim}{\lambda})$ is asymptotically normally distributed with 0 mean and variance

$$\sum_{j=1}^{m} \sum_{\ell=1}^{m} \lambda_j\lambda_\ell E[W(\underset{\sim}{t}_j)W(\underset{\sim}{t}_\ell)]. \tag{2.4}$$

Since this holds for all $\underset{\sim}{\lambda}\neq\underset{\sim}{0}$, $[W_n(\underset{\sim}{t}_1),\ldots,W_n(\underset{\sim}{t}_m)]$ has asymptotically a m-variate normal distribution with null mean vector and dispersion matrix $((E[W(\underset{\sim}{t}_j)W(\underset{\sim}{t}_\ell)]))_{j,\ell=1,\ldots,m}$. Since m and $\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_m$ are arbitrarily fixed, the lemma follows.

By virtue of Lemma 2.1, we need to establish only the tightness of $\{W_n\}$ which will ensure then the weak convergence of $\{W_n\}$ to W. In this context, the following lemma will be useful. Let $\{T_i = T(\underset{\sim}{X}_i),\ -\infty<i<\infty\}$ be a ϕ-mixing sequence of bounded random variables, such that (1.1) holds and

(2.5) $$ET_i=0,\ ET_i^2 = \tau:\ 0\leq\tau\leq 1,\ P\{|T_i|>1\}=0 \text{ and } E|T_i|<c\tau,$$

where $0<c<\infty$. In particular, for centered Bernoullian random variables, c=2. Let then $S_n = T_1+\ldots+T_n$, $n\geq 1$.

LEMMA 2.2. Under (1.1), (2.5) and $A_k(\phi)<\infty$ for some $k\geq 0$, for every $n\geq 1$,

(2.6) $$E(S_n^{2(k+1)}) \leq K_\phi\{n\tau+\ldots+(n\tau)^{k+1}\},\ K_\phi<\infty,$$

where K_ϕ depends only on $\{\phi_n\}$.

Proof. Note that for $k\geq 0$,

(2.7) $$E(S_n^{2(k+1)}) \leq [(2k+2)!]\ n\ \Sigma_{n,2k+1}|E(T_1T_{i_1}\cdots T_{i_{2k+1}})|,$$

where the summation $\Sigma_{n,2k+1}$ extends over all $1\leq i_1\leq\ldots\leq i_{2k+1}\leq n$. Also, note that if ξ and η be $\mathcal{M}_{-\infty}^{k}$ and $\mathcal{M}_{k+n}^{\infty}$ measurable, $E|\xi|<\infty$ and $P\{|\eta|>1\} = 0$, then [cf. Billingsley (1968, p. 171)]

(2.8) $$|E(\xi\eta)-E(\xi)E(\eta)| \leq 2\phi_n E|\xi|,\ \forall\, n.$$

Proceeding as in the proof of Lemma 2.1 of Sen (1971) and using (2.8), we obtain that if $A_0(\phi)<\infty$, under (1.1) and (2.5)

$$(2.9)\qquad n\,\Sigma_{n,1}|E(T_1T_{i_1})| \leq [2cA_0(\phi^2)](n\tau),$$

and if $A_1(\phi)<\infty$ [$\Rightarrow A_2(\phi^2)<\infty$], then

$$n\,\Sigma_{n,2}|E(T_1T_{i_1}T_{i_2})| \leq 6nc\tau\textstyle\sum_{i=0}^{n-1}(i+1)^2\phi_i$$

$$(2.10)\qquad < [6cA_2(\phi^2)](n\tau),$$

$$(2.11)\qquad n\,\Sigma_{n,3}|E(T_1T_{i_1}\cdots T_{i_3})| \leq K_\phi[n\tau+(n\tau)^2],\ K_\phi<\infty.$$

Let us now assume that for $1\leq a\leq 2k-1$, $k\geq 1$, $n\geq 1$,

$$(2.12)\qquad n\,\Sigma_{n,a}|E(T_1T_{i_1}\cdots T_{i_a})| \leq K_{\phi,a}\{n\tau+\ldots+(n\tau)^{a^*}\},\ K_{\phi,a}<\infty,$$

where $a^*=t$ for $a=2t$ or $2t-1$, $t\geq 1$. Then, we shall show that $A_k(\phi)<\infty$ implies that (2.12) also holds for $a=2k$ and $2k+1$. We consider only the case of $a=2k+1$ (as the other case follows similarly). For this, we let $i_0=1$, $i_j=i_{j-1}+r_j$, $r_j\geq 0$, $1\leq j\leq 2k+1$, and let $\sum_{n,2k+1}^{(j)}$ be the summation over all $1\leq i_1\leq\ldots\leq i_{2k+1}\leq n$ for which $r_j=\max\{r_1,\ldots,r_{2k+1}\}$ for $j=1,\ldots,2k+1$. Then

$$(2.13)\qquad n\,\Sigma_{n,2k+1}|E(T_1T_{i_1}\cdots T_{i_{2k+1}})| \leq \textstyle\sum_{j=1}^{2k+1}\{n\sum_{n,2k+1}^{(j)}|E(T_1\cdots T_{i_{2k+1}})|\},$$

where by (2.5) and (2.8), for each j: $1\leq j\leq 2k+1$,

$$n\textstyle\sum_{n,2k+1}^{(j)}|E(T_1\cdots T_{i_{2k+1}})| \leq n\sum_{n,2k+1}^{(j)}|E(T_1\cdots T_{i_{j-1}})E(T_{i_j}\cdots T_{i_{2k+1}})|$$

$$(2.14)\qquad + 2\,n\textstyle\sum_{n,2k+1}^{(j)}\phi_{r_j}E|T_1\cdots T_{i_{j-1}}|,$$

where the second term on the rhs of (2.14) is bounded by

$$2n\,E|T_1|\textstyle\sum_{n,2k+1}^{(j)}\phi_{r_j} \leq 2nc\tau\sum_{r_j=0}^{n-1}(r_j+1)^{2k}\phi_{r_j}$$

$$(2.15)\qquad < 2nc\tau(\textstyle\sum_{i=0}^{\infty}(i+1)^{2k}\phi_i) = [2cA_{2k}(\phi^2)](n\tau),$$

and $A_{2k}(\phi^2)<\infty$. For $j=1$ or $2k+1$, the first term on the rhs of (2.14) vanishes [by (2.5)], while for $2\leq j\leq 2k$, we have

$$n \sum_{n,2k+1}^{(j)} |E(T_1 \cdots T_{i_{j-1}}) E(T_{i_j} \cdots T_{i_{2k+1}})|$$

(2.16)
$$\leq n \sum_{i_j=1}^{n} \{\Sigma_{i_j,j-1} |E(T_1 \cdots T_{i_{j-1}})|\}$$

$$\{\Sigma_{n-i_j+1,2k+1-j} |E(T_{i'_0} \cdots T_{i'_{2k+1-j}})|\},$$

where $i'_\ell = i_{j+\ell} - i_j + 1$, $\ell=0,\ldots,2k+1-j$. Since for $2\leq j\leq 2k$, $1\leq j-1, 2k+1-j\leq 2k-1$, and by assumption, (2.12) holds for $a\leq 2k-1$, we obtain from (2.12), (2.16) and the inequality that for $a\geq 0$, $b\geq 0$,

$$\sum_{i=1}^{n} i^a (n-i+1)^b \leq c(n+1)^{a+b+1} \leq c^* n^{a+b+1}, \quad c^*<\infty, \tag{2.17}$$

that the rhs of (2.16) is bounded by

$$K_{\phi,j}[n\tau + \ldots + (n\tau)^{k^*}], \quad K_{\phi,j}<\infty, \tag{2.18}$$

where

$$k^* = \begin{cases} k, & j = \text{odd} \\ k+1, & j = \text{even}. \end{cases} \tag{2.19}$$

Thus, from (2.13) through (2.19) we conclude that (2.12) holds for $a=2k+1$. Using then (2.9)-(2.11) and the method of induction, the proof follows for general $k\geq 1$. Q.E.D.

Let now $\{T^*_i = T^*(\underset{\sim}{X}_i), -\infty<i<\infty\}$ be another ϕ-mixing sequence satisfying (1.1) and (2.5) with τ being replaced by τ^*: $0\leq\tau^*\leq 1$, and let $S^*_n = T^*_1+\ldots+T^*_n$, $n\geq 1$.

<u>LEMMA 2.3</u>. <u>Under</u> (1.1), (2.5) <u>and</u> $A_1(\phi)<\infty$, <u>for every</u> $n\geq 1$,

(2.20) $$E[(S_n S_n^*)^2] \le K_\phi\{n^2\tau\tau^* + n\tau + n\tau^*\},$$

where $K_\phi(<\infty)$ depends only on $\{\phi_n\}$.

The proof follows along the lines of Lemma 2.2, and hence, is omitted. In passing, we may remark that Lemma 2.2 extends Lemma 5.2 of Neuhaus (1971) to more general random variables and to ϕ-mixing processes. For independent processes (i.e., for $\phi_n=0$, $\forall\ n\ge 1$), the rhs of (2.20) can be replaced by $3n^2\tau\tau^*$, for every $0\le\tau\le 1$, $0\le\tau^*\le 1$. On the other hand, for general ϕ-mixing sequence, the rhs of (2.20) can not be replaced by $K_\phi n^2\tau\tau^*$ for every $0\le\tau$, $\tau^*\le 1$. This replacement is possible for more restricted ϕ-mixing processes, where in (1.1), the rhs is $\phi_n P(A)P(B)$, i.e., for the so called *-mixing sequences; we may refer to Lemma 3.1 of Sen (1973).

For iid stochastic vectors, the tightness of $\{W_n\}$ has been studied in section 5 of Neuhaus (1971). On replacing his Lemma 5.2 by our Lemma 2.2 and then using his Lemma 5.1, we are able to prove the tightness of $\{W_n\}$ in a similar way. However, in view of the fact that his Lemma 5.2 involves the central moment of order 2(p+1), we end up with the condition that $A_p(\phi)<\infty$, which for $p>1$, appears to be more stringent than our desired condition $A_1(\phi)<\infty$.

Let now $B=\{\underset{\sim}{t}: \underset{\sim}{b}_0\le\underset{\sim}{t}\le\underset{\sim}{b}_1\}$ be a p-dimensional block contained in E^p (i.e., $\underset{\sim}{0}\le\underset{\sim}{b}_0\le\underset{\sim}{b}_1\le\underset{\sim}{1}$), $\mu(B)=P\{\underset{\sim}{Y}_1\epsilon B\}$ and for every $n\ge 1$, let $\mu_n(B)=n^{-1}[\#\text{ of }\underset{\sim}{Y}_i\epsilon B,\ 1\le i\le n]$. Let us then define

(2.21) $$V_n(B) = n^{\frac{1}{2}}[\mu_n(B)-\mu(B)],\ B\epsilon E^p,\ n\ge 1.$$

Now, by definition, $G(\underset{\sim}{t})$ and $G_n(\underset{\sim}{t})$ are equal to 0 if at least one of the p coordinates of $\underset{\sim}{t}$ is equal to 0. Thus, $W_n(\underset{\sim}{t})$ is equal to 0 when $t_j=0$ for some $1\le j\le p$. Thus, using a direct multiparameter generalization of Theorem 15.6 of Billingsley (1968, p. 128) [viz., Theorem 3 of Bickel and Wichura (1971)], for the tightness of $\{W_n\}$, it suffices to show that for every pair (B,C) of neighbouring blocks (in E^p) and every $\lambda>0$,

$$P\{\min[|V_n(B)|,|V_n(C)|]>\lambda\} \le K\lambda^{-\gamma}[\mu(B\cup C)]^{\beta}, \tag{2.22}$$

where $K<\infty$, $\gamma>0$ and $\beta>1$. Now, as mentioned earlier, for iid stochastic vectors [or for a *-mixing (stationary) process], $E[V_n^2(B)V_n^2(C)] \le K\mu(B)\mu(C)$, so that (2.22) holds with $\gamma=4$, $\beta=2$. But, for a general ϕ-mixing process, by Lemma 2.2,

$$E[V_n^2(B)V_n^2(C)] \le K_\phi[Q(B)Q(C)+n^{-1}Q(B)+n^{-1}Q(C)],\ K_\phi<\infty, \tag{2.23}$$

where $0\le Q(A) = \mu(A)[1-\mu(A)] \le \mu(A) \le 1$, $\forall\ A\varepsilon E^p$. The last two terms on the rhs of (2.23) relate to a value of $\beta=1$ in (2.22), so that we can not claim $\beta>1$ for every B, $C\varepsilon E^p$. A similar problem arises if one uses the inequality that $E[V_n^2(B)V_n^2(C)] \le \{E[V_n^4(B)]E[V_n^4(C)]\}^{\frac{1}{2}}$ and then uses Lemma 2.1 (for k=1). Thus, presumably, a modified approach is needed.

We may remark that for general ϕ-mixing processes, it has been observed in Sen (1972a) that empirical processes behave quite smoothly for large n. This suggests the following approach. First, using the basic inequality between the modulii of continuity for $C^p[0,1]$ and $D^p[0,1]$ spaces [cf. Billingsley (1968, p. 110) and Neuhaus (1971, p. 1288)] and the fact that $W_n(\underset{\sim}{t})=0$ with probability one when $t_j=0$ for some $1\le j\le p$, it suffices to show that for every $\varepsilon>0$ and $\eta>0$, there exists a $\delta>0$ and an integer n_o, such that

$$P\{\omega_\delta(W_n) > \varepsilon\} < \eta,\ n \ge n_o, \tag{2.24}$$

where for every $0<\delta<1$, $n\ge1$,

$$\omega_\delta(W_n) = \sup_{|\underset{\sim}{t}-\underset{\sim}{s}|<\delta}\{|W_n(\underset{\sim}{t})-W_n(\underset{\sim}{s})|:\ \underset{\sim}{s},\underset{\sim}{t}\varepsilon E^p\}. \tag{2.25}$$

Second, by a direct multiparameter extension of Theorem 8.3 of Billingsley (1968, p. 56), for (2.24), it suffices to show that for every $\underset{\sim}{0}\le\underset{\sim}{b}_o<\underset{\sim}{1}$, $\varepsilon>0$ and $\eta>0$, there exist a δ: $0<\delta<1$ and an integer n_o, such that for $B=\{\underset{\sim}{t}: \underset{\sim}{b}_o \leq \underset{\sim}{t} \leq \underset{\sim}{b}_o + \delta \underset{\sim}{1}\}$ and $n \geq n_o$,

(2.26) $$P\{\sup_{\underset{\sim}{t} \in B} |W_n(\underset{\sim}{t}) - W_n(\underset{\sim}{b}_o)| > \varepsilon\} < \tfrac{1}{2}\eta[\mu(B)+\delta^p],$$

where $\mu(B)$ is defined before (2.21). [Note that $\frac{1}{2}[\mu(A) + \|A\|]$ (where $\|A\|$ is the Euclidean volume of A) is always ≤ 1, $\forall A \in E^p$, and as $\mu(A)$ is bounded from above by any side of A, we have $\delta^p < \mu(B) + \delta^p < \delta + \delta^p \to 0$ as $\delta \to 0$.]

For a given $\varepsilon>0$ and $\delta>0$ (to be chosen later on), select n_o so large that $\delta > n_o^{-\frac{1}{2}} \varepsilon/2p$. Let then (for $n \geq n_o$),

(2.27) $$\underset{\sim}{b}_n(\underset{\sim}{i}) = \underset{\sim}{b}_o + (n^{-\frac{1}{2}} \varepsilon/2p)\underset{\sim}{i}, \quad \underset{\sim}{0} \leq \underset{\sim}{i} \leq \underset{\sim}{m}_n,$$

where $\underset{\sim}{m}_n = m_n \cdot \underset{\sim}{1}$ and

(2.28) $$m_n = [2p\delta n^{\frac{1}{2}}/\varepsilon]+1 \ (\geq 1).$$

Also, let

(2.29) $$B(\underset{\sim}{i},n) = \{\underset{\sim}{t}: \underset{\sim}{b}_n(\underset{\sim}{i}) \leq \underset{\sim}{t} \leq \underset{\sim}{b}_n(\underset{\sim}{i}+\underset{\sim}{1})\}, \quad \underset{\sim}{i} \geq \underset{\sim}{0}.$$

Note that, if $G_{[j]}(t)(=t)$ be the marginal df of Y_{ji}, $1 \leq j \leq p$, then by (2.27), $\forall \underset{\sim}{i} \geq \underset{\sim}{0}$,

$$\sqrt{n}[G(\underset{\sim}{b}_n(\underset{\sim}{i}+\underset{\sim}{1})) - G(\underset{\sim}{b}_n(\underset{\sim}{i}))] \leq \sum_{j=1}^{p} \sqrt{n}[G_{[j]}(i_j+1) - G_{[j]}(i_j)]$$

(2.30) $$= \sum_{j=1}^{p} (\varepsilon/2p) = \varepsilon/2.$$

Hence, on using the fact that for $\underset{\sim}{t} \in B(\underset{\sim}{i},n)$, $G_n(\underset{\sim}{b}_n(\underset{\sim}{i})) \leq G_n(\underset{\sim}{t}) \leq G_n(\underset{\sim}{b}_n(\underset{\sim}{i}+\underset{\sim}{1}))$ and $G(\underset{\sim}{b}_n(\underset{\sim}{i})) \leq G(\underset{\sim}{t}) \leq G(b_n(\underset{\sim}{i}+\underset{\sim}{1}))$, we obtain by (1.4), (2.30) and a few routine steps that

(2.31) $$\sup_{\underset{\sim}{t} \in B(\underset{\sim}{i},n)} |W_n(\underset{\sim}{t}) - W_n(\underset{\sim}{b}_o)|$$

$$\leq \max_{\underset{\sim}{j}=\underset{\sim}{i},\underset{\sim}{i}+\underset{\sim}{1}} |W_n(\underset{\sim}{b}_n(\underset{\sim}{j})) - W_n(\underset{\sim}{b}_o)| + \varepsilon/2, \ \forall \underset{\sim}{i} \geq \underset{\sim}{0}.$$

Consequently,

(2.32) $$\sup_{\underset{\sim}{t}\varepsilon B}|W_n(\underset{\sim}{t})-W_n(\underset{\sim}{b}_o)| \leq \max_{\underset{\sim}{0}\leq\underset{\sim}{i}\leq\underset{\sim}{m}_n} |W_n(\underset{\sim}{b}_n(\underset{\sim}{i}))-W_n(\underset{\sim}{b}_o)|+\varepsilon/2.$$

Thus, it suffices to show that for every $\varepsilon>0$ and $\eta>0$, there exist a δ: $0<\delta<1$ and an integer n_o, such that for $n \geq n_o$ and every $\underset{\sim}{b}_o \varepsilon E^p$,

(2.33) $$P\{\max_{\underset{\sim}{0}\leq\underset{\sim}{i}\leq\underset{\sim}{m}_n} |W_n(\underset{\sim}{b}_n(\underset{\sim}{i}))-W_n(\underset{\sim}{b}_o)| > \tfrac{1}{2}\varepsilon\} < \tfrac{1}{2}\eta[\mu(B)+\delta^p].$$

Now, by (1.4), (2.21) and (2.27), $\forall\ \underset{\sim}{i} \geq \underset{\sim}{0}$,

(2.34) $$W_n(\underset{\sim}{b}_n(\underset{\sim}{i}+\underset{\sim}{1}))-W_n(\underset{\sim}{b}_o) = \Sigma_{\underset{\sim}{0}\leq\underset{\sim}{j}\leq\underset{\sim}{i}}\ V_n(B(\underset{\sim}{j},n))$$

$$= S_n(\underset{\sim}{i}),\ \text{say},$$

where by Lemma 2.1, under $A_1(\phi)<\infty$,

(2.35) $$E[V_n^4(B(\underset{\sim}{i},n))] \leq K_\phi[\mu^2(B(\underset{\sim}{i},n))+n^{-1}\mu(B(\underset{\sim}{i},n))].$$

Unfortunately, $\mu(B(\underset{\sim}{i},n)$, though bounded from above by $\varepsilon/2\sqrt{n}$, can be arbitrarily close to 0. For example, if $G(\underset{\sim}{t})$ is degenerate on a lower dimensional space, then $\mu(B)$ may be equal to 0 for some $B\varepsilon E^p$. To overcome this difficulty, we define

(2.36) $$\lambda_n(\underset{\sim}{i}) = \max\{\mu(B(\underset{\sim}{i},n)),\ (\varepsilon/2p\sqrt{n})^p\},\ \forall\ \underset{\sim}{i} \geq \underset{\sim}{0}.$$

Now, (2.36) implies that $\lambda_n(\underset{\sim}{i}) \geq (\varepsilon/2p\sqrt{n})^p$ i.e., $(2p/\varepsilon)^2\lambda_n^{2/p}(\underset{\sim}{i}) \geq \frac{1}{n}$. Hence, from (2.35), we have under $A_1(\phi)<\infty$,

(2.37) $$E[V_n^4(B(\underset{\sim}{i},n))] \leq K_\phi\{\lambda_n^2(\underset{\sim}{i})+(2p/\varepsilon)^2[\lambda_n(\underset{\sim}{i})]^{1+2/p}\}.$$

Since for p=1, the proof is considered in Billingsley (1968, p. 198) and Sen (1971), we confine ourselves to $p\geq 2$, so that $1+2/p\leq 2$. Then, from the above, we have, $\forall\ \underset{\sim}{i} \geq \underset{\sim}{0}$, (2.38) $$E[V_n^4(B(\underset{\sim}{i},n))] \le K_{\phi,\varepsilon}[\lambda_n(\underset{\sim}{i})]^\beta, \beta = 1+2/p > 1,$$

where

(2.39) $$K_{\phi,\varepsilon} \le K_\phi\{1 + (2p/\varepsilon)^2\} < \infty, \forall\ \varepsilon>0.$$

Now Bickel and Wichura (1971, Theorem 1) have considered a multiparameter extension of Theorem 12.5 of Billingsley (1968), under the assumption that (2.22) holds for every $(B,C)\varepsilon E^p$. By the same method of induction, it follows, on pretty much the same line, that under (2.38), Theorem 12.2 of Billingsley (1968, p. 94) extends to a multiple array, as in (2.34). Thus, we have for every $\varepsilon>0$,

(2.40) $$P\{\max_{\underset{\sim}{0}\le\underset{\sim}{i}\le\underset{\sim}{m}_n-\underset{\sim}{1}} |S_n(\underset{\sim}{i})| > \tfrac{1}{2}\varepsilon\}$$
$$\le (16\ K^*_{\phi,\varepsilon}/\varepsilon^4)(\sum_{\underset{\sim}{0}\le\underset{\sim}{i}\le\underset{\sim}{m}_n-\underset{\sim}{1}} \lambda_n(\underset{\sim}{i}))^\beta,\ \beta = 1+2/p,$$

where $K^*_{\phi,\varepsilon}(<\infty)$ depends on ε through $K_{\phi,\varepsilon}$ in (2.39). Now, $\lambda_n(\underset{\sim}{i}) \le \mu(B(\underset{\sim}{i},n)) + (\varepsilon/2p\sqrt{n})^p$, $\forall\ \underset{\sim}{i} \ge 0$, so that the rhs of (2.40) is bounded by

(2.41) $$(16\ K^*_{\phi,\varepsilon}/\varepsilon^4)[\mu(B_n)+\delta_n^p]^\beta,\ \beta = 1+2/p>1,$$

where $B_n = \{\underset{\sim}{t}: \underset{\sim}{b}_o\le\underset{\sim}{t}\le\underset{\sim}{b}_o+(\varepsilon/2p\sqrt{n})\underset{\sim}{m}_n\}$ and $\delta_n = m_n(\varepsilon/2p\sqrt{n})$. By (2.27), $0\le\mu(B_n)-\mu(B) \le \varepsilon/2\sqrt{n}$ and $\delta\le\delta_n\le\delta+\varepsilon/2p\sqrt{n}$. Thus, using the fact that $\delta_n^p\le\mu(B_n) + \delta_n^p \le \delta_n + \delta_n^p$, we obtain that for every $\eta>0$, there exist a $\delta>0$ and an n_o, such that

(2.42) $$\mu(B_n) + \delta_n^p \le \frac{3}{2}(\mu(B)+\delta^p),\ n \ge n_o,$$

(2.43) $$(16\ K^*_{\phi,\varepsilon}/\varepsilon^4)[\mu(B_n)+\delta_n^p]^{2/p} < \eta/3,\ n \ge n_o,$$

which completes the proof of (2.33). Hence, we have the following.

LEMMA 2.4. *Under* (1.1) *and* $A_1(\phi)<\infty$, $\{W_n\}$ *is tight i.e.*, (2.24) *holds*.

From Lemmas 2.1 and 2.4, we arrive at the main theorem of this section.

THEOREM 1. Under (1.1) and $A_1(\phi)<\infty$, as $n\to\infty$, W_n converges in law (in the Skorokhod J_1-topology on $D^p[0,1]$) to a Gaussian process W for which (2.2) holds.

For later use in Section 3, we consider the following.

LEMMA 2.5. Under (1.1) and $A_1(\phi)<\infty$,

$$\sup_n E\{\sup_{\underset{\sim}{t}\in E^p} [W_n^2(\underset{\sim}{t})]\} < \infty. \tag{2.44}$$

Proof: For a non-negative random variable Y, $E\{YI(Y>k)\} = kP\{Y>k\}+ \int_k^\infty [1-P\{Y>t\}]dt$, for every $k>0$, where $I(n)$ is the indicator function. So, to prove (2.44), it suffices to show that for every $\lambda>1$,

$$\sup_n P\{\sup_{\underset{\sim}{t}\in E^p} |W_n(\underset{\sim}{t})| > \lambda\} \le K\,\lambda^{-\gamma},\ \gamma > 2,\ K < \infty. \tag{2.45}$$

To prove (2.45), we virtually repeat the steps (2.27) through (2.41), where (i) in (2.27)-(2.28), we take both ε and δ equal to 1, (ii) in (2.36), we replace ε by 1 (so that $K_{\phi,1} \le K_\phi(1+4p^2)$, in (2.39)), and in (2.40), $\frac{1}{2}\varepsilon$ by $\frac{1}{2}\lambda$, $\lambda>1$. Since, then $\lambda_n(\underset{\sim}{i}) \le 1/p\sqrt{n}$, the rhs of (2.40) reduces to

$$\begin{aligned}&(16\,K^*_{\phi,1})(1+(1+1/p\sqrt{n})^p)/\lambda^4\\ &\le C_\phi\lambda^{-4},\ C_\phi < \infty,\ \forall\ n.\end{aligned} \tag{2.46}$$

Q.E.D.

Actually, it follows that under (1.1) and $A_1(\phi)$,

$$\sup_n E\{\sup_{\underset{\sim}{t}\in E^p} |W_n(\underset{\sim}{t})|^{2+\delta}\} < \infty, \tag{2.47}$$

for every $0\le\delta<1$, and in general, if $A_k(\phi)<\infty$ for some $k\ge1$, then

$$\sup_n E\{\sup_{\underset{\sim}{t}\in E^p} |W_n(\underset{\sim}{t})|^{2k+\delta}\} < \infty,\ \forall\ ,\ 0 \le \delta < 1. \tag{2.48}$$

3. Weak convergence of $\{W_{N_\nu}\}$. For two real valued functions $Z(\underset{\sim}{t})$ and $Z^*(\underset{\sim}{t})$, defined on E^p, we let

$$(3.1)\qquad \rho(Z,Z^*) = \sup\{|Z(\underset{\sim}{t})-Z^*(\underset{\sim}{t})|: \underset{\sim}{t}\varepsilon E^p\}.$$

Then

$$(3.2)\qquad W_n^* = \sup_{\underset{\sim}{t}\varepsilon E^p} |W_n(\underset{\sim}{t})| = \rho(W_n,0).$$

LEMMA 3.1. Under (1.1) and $A_1(\phi)<\infty$, for every positive ε, there exists a $K_\varepsilon(<\infty)$, such that for every n,

$$(3.3)\qquad P\{\max_{1\le k\le n} (k/n)^{\frac{1}{2}}\rho(W_k,0) > K_\varepsilon\} < \varepsilon.$$

Proof. Without any loss of generality, we take $K_\varepsilon>1$, as otherwise, a smaller value of K_ε can always be replaced by a value greater than one. As by definition of W_n, with probability one, $\rho(W_n,0) \le n^{\frac{1}{2}}$, for every n, we need to prove (3.3) only for $n > K_\varepsilon^2(>1)$. Let then

$$(3.4)\qquad n_s = s[n^{\frac{1}{2}}],\ s=0,1,\dots,m_n = [n^{\frac{1}{2}}]+1,$$

and note that by (1.4), (3.2) and (3.4),

$$(3.5)\qquad \max_{n_{s-1}<k\le n_s} |k^{\frac{1}{2}}W_k^* - n_s^{\frac{1}{2}}W_{n_s}^*| \le (n_s-n_{s-1}) \le n^{\frac{1}{2}},\ \forall\, s\ge 1,$$

so that

$$(3.6)\qquad \max_{1\le k\le n} (k/n)^{\frac{1}{2}}W_k^* \le \max_{1\le s\le m_n} (n^{-1}n_s)^{\frac{1}{2}}W_{n_s}^* +1.$$

Thus, it suffices to show that for every $0<\varepsilon<1$, there exists a positive $K_\varepsilon^*(<\infty)$, such that

$$(3.7)\qquad P\{\max_{1\le s\le m_n} (n^{-1}n_s)^{\frac{1}{2}}W_{n_s}^* > K_\varepsilon^*\} < \varepsilon.$$

Let us then define $S_{ns} = \xi^*_{n1}+\ldots+\xi^*_{ns}$, $s\geq 1$, where

$$\xi^*_{ns} = n^{-\frac{1}{2}}(n_s^{\frac{1}{2}}W^*_{n_s} - n_{s-1}^{\frac{1}{2}}W^*_{n_{s-1}}),\ s\geq 1. \tag{3.8}$$

Now, by (1.4), (3.2), (3.4) and (3.8), for every ℓ: $1\leq\ell\leq m_n-s$, $s\geq 1$,

$$\begin{aligned} |\xi^*_{n_s}+\ldots+\xi^*_{n_{s+\ell}}| &= |n^{-\frac{1}{2}}(n_{s+\ell}^{\frac{1}{2}}W^*_{n_{s+\ell}} - n_s^{\frac{1}{2}}W^*_{n_s})| \\ &= n^{-\frac{1}{2}}\left|\sup_{\underset{\sim}{t}\in E^p}\left|\sum_{i=1}^{n_{s+\ell}}[c(\underset{\sim}{t}-\underset{\sim}{Y}_i)-G(\underset{\sim}{t})]\right| - \sup_{\underset{\sim}{t}\in E^p}\left|\sum_{i=1}^{n_s}[c(\underset{\sim}{t}-\underset{\sim}{Y}_i)-G(\underset{\sim}{t})]\right|\right| \\ &\leq n^{-\frac{1}{2}}[\sup_{\underset{\sim}{t}\in E^p}|\sum_{i=n_s+1}^{n_{s+\ell}}[c(\underset{\sim}{t}-\underset{\sim}{Y}_i)-G(\underset{\sim}{t})]|] \\ &= ([n_{s+\ell}-n_s]/n)^{\frac{1}{2}}W^*_{n_{s+\ell}-n_s}, \end{aligned} \tag{3.9}$$

where $\tilde{W}^*_{k-q}$ is the Kolmogorov supremum [cf. (3.2)] for the empirical process based on $\underset{\sim}{Y}_{k+1},\ldots,\underset{\sim}{Y}_q$, $q>k$. Now, on $\tilde{W}^*_{k-q}$, we may use (2.47), so that by (3.9), when $A_1(\phi)<\infty$,

$$\begin{aligned} E|S_{ns+\ell}-S_{ns}|^{2+\delta} &\leq ([n_{s+\ell}-n_s]/n)^{1+\delta/2}E[\tilde{W}^*_{n_{s+\ell}-n_s}]^{2+\delta} \\ &\leq C(\ell/\sqrt{n})^{1+\delta/2},\ C<\infty,\ \forall\ \ell=0,1,\ldots,m_n-s,\ s\geq 1. \end{aligned} \tag{3.10}$$

Therefore, if we use Theorem 12.2 of Billingsley (1968, p. 94), we have by (3.4) and (3.10),

$$(3.11)\qquad P\{\max_{1\le s\le m_n} (n^{-1}n_s)^{\frac{1}{2}} W^*_{n_s} > K^*_\varepsilon\}$$

$$= P\{\max_{1\le s\le m_n} |S_{ns}| > K^*_\varepsilon\}$$

$$\le (K^*_\varepsilon)^{-2-\delta} \cdot K(\textstyle\sum_{s=1}^{m_n} n^{-\frac{1}{2}})^{1+\delta/2}$$

$$\le K(K^*_\varepsilon)^{-2-\delta}(1+n^{-\frac{1}{2}})^{1+\delta/2}$$

$$< \varepsilon, \text{ by proper choice of } K^*_\varepsilon,$$

as $K(<\infty)$ depends on $\delta(>0)$ only. Q.E.D.

As a direct consequence of Lemma 2.5 (or the weak convergence of $\{W_n\}$ to W), it follows that for every $\varepsilon>0$, there exists a positive $K_\varepsilon(<\infty)$, such that

$$(3.12)\qquad P\{\rho(W_n,0) > K_\varepsilon\} < \varepsilon.$$

From (3.3) and (3.12), and proceeding as in the proof of Theorem 2.1 of Pyke (1968), we arrive at the following.

LEMMA 3.2. *(Uniform continuity in probability). For every* $\varepsilon>0$ *and* $\eta>0$, *there exist a* $\delta(>0)$ *and an integer* n_o, *such that for* $n\ge n_o$,

$$(3.13)\qquad P\{\max_{k:\ |k-n|<\delta n} \rho(W_k,W_n) > \varepsilon\} < \eta.$$

We now consider the main theorem of this section.

THEOREM 2. *Under* (1.1), (1.6) *and* $A_1(\phi)<\infty$, $\{W_{N_\nu}\}$ *converges in law* (as $\nu\to\infty$) *in the Skorokhod* J_1*-topology on* $D^p[0,1]$ *to a Gaussian process* W *for which* (2.2) *holds*.

Proof. When in (1.6), $\xi=c$ (a constant >0), with probability one, the proof of the theorem follows readily from Theorem 1 and (3.13). When ξ is a positive random variable, we introduce a sequence $\{k_n\}$ of positive integers, such that

(3.14) $$k_n \to \infty \text{ but } n^{-\frac{1}{2}} k_n \to 0 \text{ as } n\to\infty.$$

Let then

(3.15) $$W_n'(\underset{\sim}{t}) = n^{-\frac{1}{2}} \sum_{i=k_n}^{n} [c(\underset{\sim}{t}-\underset{\sim}{Y}_i)-G(\underset{\sim}{t})],\ \underset{\sim}{t}\varepsilon E^p.$$

Then, by (1.4) and (3.14),

(3.16) $$\rho(W_n,W_n') \le k_n\ n^{-\frac{1}{2}} \to 0 \text{ as } n\to\infty.$$

Since $k_n\to\infty$ as $n\to\infty$ and ξ is defined on the same probability space $(\Omega,\mathcal{A},P)$ as of the $\{X_i\}$, W_n' (and hence, W_n) is a mixing sequence in the sense of Rényi (1958). Consequently, by Lemma 3 of Blum, Hanson and Rosenblatt (1963), we obtain from Lemma 3.2, the following: if $A\varepsilon\mathcal{A}$, then for every $\varepsilon>0$ and $\eta>0$, there exists a $\delta>0$, such that as $n\to\infty$,

(3.17) $$P\{\max_{k:\ |k-n|<\delta n} \rho(W_k,W_n) > \varepsilon|A\} < \eta.$$

Similarly, by Lemma 2.4 and the Blum et al. lemma,

(3.18) $$\lim_{n\to\infty} P\{\omega_\delta(W_n) > \varepsilon|A\} < \eta,\ \forall A\varepsilon\mathcal{A}$$

The last two inequalities insures that the proof of the theorem for independent vectors, considered in Sen (1972b), holds for general ϕ-mixing processes, satisfying (1.1) and $A_1(\phi)<\infty$. For brevity, the details are therefore omitted.

<u>REMARK</u>. Kiefer (1961) has shown that for iid stochastic vectors, for every $\varepsilon>0$, there exists a positive $c_p(\varepsilon)$, $(<\infty)$, such that for p_1,

(3.19) $$P\{\rho(W_n,0) > \lambda\} \le c_p(\varepsilon)\exp\{-(2-\varepsilon)\lambda^2\},$$

for every $\lambda>0$. For ϕ-mixing processes, such a strong result is not known. However, (2.48) shows that if $A_k(\phi)<\infty$, for some $k\geq 1$, then $P\{\rho(W_n,0)>\lambda\}$ is of the order $\lambda^{-2k-\delta}$, $\delta>0$, for every $\lambda>0$.

REFERENCES

[1] BICKEL, P. J., AND WICHURA, M. J. (1971). Convergence criteria for multi-parameter stochastic processes and some applications. Ann. Math. Statist. 42, 1656-1670.

[2] BILLINGSLEY, P. (1968). Convergence of Probability Measures. John Wiley, New York.

[3] BLUM, J. R., HANSON, D. L., AND ROSENBLATT, J. I. (1963). On the central limit theorem for the sum of a random number of independent random variables. Zeit. für Wahrsch. Verw. Geb. 1, 389-393.

[4] KIEFER, J. (1961). On large deviations of the empirics D.F. of vector chance variables and a law of iterated logarithm. Pacific Jour. Math. 11, 649-660.

[5] NEUHAUS, G. (1971). On weak convergence of stochastic processes with multi-dimensional time parameter. Ann. Math. Statist. 42, 1285-1295.

[6] PYKE, R. (1968). The weak convergence of empirical process with random sample sizes. Proc. Cambridge Phil. Soc. 64, 155-160.

[7] RÉNYI, A. (1958). On mixing sequences of sets. Acta Math. Acad. Sci. Hungar. 9, 215-228.

[8] SEN, P. K. (1971). A note on weak convergence of empirical processes for sequences of ϕ-mixing random variables. Ann. Math. Statist. 42, 2131-2133.

[9] SEN, P. K. (1972a). On the Bahadur representation of sample quantiles for sequences of ϕ-mixing random variables. Jour. Mult. Anl. 2, 77-95.

[10] SEN, P. K. (1972b). On weak convergence of empirical processes for random number of independent stochastic vectors. Proc. Cambridge Phil. Soc. (in press).

[11] SEN, P. K. (1973). Limiting behaviour of regular functionals of empirical distributions for stationary *-mixing processes. Zeit. für Wahrsch. Verw Geb. (in press).