The Annals of Statistics
1974 Vol. 2, No. 3, 586–596

WEAK CONVERGENCE OF A TWO-SAMPLE EMPIRICAL PROCESS AND A CHERNOFF–SAVAGE THEOREM FOR ϕ-MIXING SEQUENCES[1]

BY THOMAS R. FEARS[2] AND K. L. MEHRA

Iowa State University; University of Alberta and University of California, Berkeley

Using Pyke–Shorack (*Ann. Math. Statist.* (1968) 755–771) approach, based on weak-convergence properties of empirical processes, a Chernoff–Savage theorem concerning the asymptotic normality of two-sample linear rank statistics is proved for stationary ϕ-mixing sequences $\{X_m\}$ and $\{Y_n\}$ of rv's. This main result (Theorem 4.1) is almost as strong as proved by Pyke and Shorack for sequences of independent rv's. The basic tool employed is the following new result concerning the behavior of empirical process $\{U_m(t): 0 \leqq t \leqq 1\}$ near 0 and 1 under ϕ-mixing: For given $\varepsilon > 0$, the $P[(t(1-t))^{-\frac{1}{2}+\delta}|U_m(t)| \leqq \varepsilon \ \forall 0 \leqq t \leqq \theta]$, $(0 < \delta < \frac{1}{2}, 0 < \theta < \frac{1}{2})$, can be made arbitrarily close to 1 by taking m sufficiently large and θ sufficiently small.

1. Summary and introduction. In recent years there has been considerable interest in studying the behavior of nonparametric statistical procedures under dependence, as evidenced by the recent papers of Serfling (1968), where the asymptotic distribution of the Wilcoxon statistic is studied for strongly-mixing processes and of Gastwirth, Rubin *et al.* (1967 and 1971), where the effects of serial dependence on levels and efficiency of the sign and Wilcoxon statistics are studied. Motivated by this interest, we prove in this paper the Chernoff-Savage Theorem (1958) (see also [5] and [7]) concerning the asymptotic normality of two-sample linear rank statistics for the case when two independent sequences $\{X_i: i = 1, 2, \cdots\}$ and $\{Y_i: i = 1, 2, \cdots\}$ of rv's (defined on a probability space $(\Omega, \mathscr{A}, P)$) satisfy the following conditions: (i) $\{X_i\}$ and $\{Y_i\}$ are strictly stationary processes, (ii) have absolutely continuous finite dimensional distributions (with respect to Lebesgue measure) and (iii) satisfy the ϕ-mixing condition: Let $\mathscr{M}_1^k$ and $\mathscr{M}_{k+n}^\infty$ denote the σ-fields generated by $\{X_i: i \leqq k\}$ and $\{X_i: i \geqq k+n\}$ respectively. Then for all $k \geqq 1$ and each $n \geqq 1$, $E_1 \in \mathscr{M}_1^k$ and $E_2 \in \mathscr{M}_{k+n}^\infty$ together imply

$$|P(E_1 \cap E_2) - P(E_1) \cdot P(E_2)| \leqq \phi_n \cdot P(E_1),$$

where ϕ_n is a non-increasing function of positive integers with $0 \leqq \phi_n \leqq 1$ and $\lim_{n\to\infty} \phi_n = 0$. We assume further ($w$ log) that $\{Y_i\}$ satisfies this condition with

Received June 1972; revised March 1973.

[1] This work was supported partially by National Research Council of Canada Grant A-3061.

[2] Now at National Cancer Research Institute, Bethesda, Maryland.

AMS 1970 *subject classifications*. Primary 60F05, 62E20; Secondary 62G10.

Key words and phrases. Chernoff-Savage Theorem under dependence, weak convergence of empirical processes, two-sample empirical process, weak convergence, ϕ-mixing processess.

the same function ϕ satisfying

$$\sum_{n=1}^{\infty} n^2 \phi_n^{\frac{1}{2}} < \infty . \tag{1.1}$$

Taking the first mX's and the first nY's, set $N = m + n$, $\lambda_N = (m/N)$ and suppose that $\lambda_N \in \Lambda = [\lambda_*, 1 - \lambda_*]$ for some $0 < \lambda_* \leqq \frac{1}{2}$. Define now the linear rank-statistic T_N by

$$T_N = m^{-1} \sum_{k=1}^{N} c_{Nk} R_{Nk} , \tag{1.2}$$

where R_{Nk} denotes the number of X's among $X_1, X_2, \cdots, X_m$ which do not exceed the kth order statistic of the combined sample $X_1, X_2, \cdots, X_m, Y_1, Y_2, \cdots, Y_n$ and c_{Nk}, $1 \leqq k \leqq N$, are a given set of constants (to be appropriately defined later). If we set $c_{Nk}^* = \sum_{j \geqq k} c_{Nj}$ and $Z_{Nk} = R_{Nk} - R_{N(k-1)}$ $(R_{N0} = 0)$ for $1 \leqq k \leqq N$, then as shown in Pyke and Shorack [7], T_N has an alternative representation $T_N = m^{-1} \sum_{k=1}^{N} c_{Nk}^* Z_{nk}$. The asymptotic normality of T_N in this form was studied by Chernoff and Savage (1958) under the assumption that X's and Y's are all mutually independent rv's. Now if $F_m(G_n)$ denote the empirical df and $F(G)$ the common continuous df of $X_1, X_2, \cdots, X_m$ $(Y_1, Y_2, \cdots, Y_n)$, and further

$$H_N = \lambda_N F_m + (1 - \lambda_N) G_n \quad \text{and} \tag{1.3}$$
$$H_\lambda = \lambda F + (1 - \lambda) G \quad \text{and} \quad H_{\lambda_N} = H ,$$

then $R_{Nk} = F_m H_N^{-1}(k/N)$, where $H_N^{-1}(t) = \inf\{x : H_N(x) \geqq t\}$, and (1.2) takes the form $\int_0^1 F_m H_N^{-1} d\nu_N$, with ν_N denoting the signed measure which assigns measure c_{Nk} at the point (k/N), $1 \leqq k \leqq N$. Setting $\mu_N = \int_0^1 FH^{-1} d\nu_N$, Pyke and Shorack (1968) studied the asymptotic distribution of

$$T_N^* = N^{\frac{1}{2}}(T_N - \mu_N) = \int_0^1 L_N(t)\, d\nu_N(t) \tag{1.4}$$

by studying the weak convergence, as $N \to \infty$, of the two-sample empirical process $\{L_N(t) : 0 \leqq t \leqq 1\}$, defined by

$$L_N(t) = [F_m H_N^{-1}(t) - FH^{-1}(t)] , \tag{1.5}$$

to a Gaussian process L_0 (see (3.8) of [7]) relative to various metrics.

The relatively straightforward manner with which using the Pyke–Shorack approach we are able to extend the Chernoff–Savage theorem to the ϕ-mixing case should support their assertion about the "usefulness and naturalness" of the representation (1.2) and their approach. (In this connection the reader is referred to another possible approach for such problems introduced by Hájek (1968).) The work in this paper is also motivated by a possible subsequent use of these results in devising and studying appropriate nonparametric procedures for dependent sequences. The basic result of this paper, namely Theorem 2.1, is given in Section 2. Section 3 proves a weak convergence property of the process L_N needed for the Chernoff–Savage Theorem which is given in Section 4.

2. Preliminary results: one-sample empirical processes. Throughout this paper we shall attempt consistently to follow the notation introduced by Pyke and Shorack. Accordingly, consider all relevant notation introduced in Section 1 to be for the strictly stationary ϕ-mixing sequences $\{X_i\}$ and $\{Y_i\}$, and let $\{U_m(t): 0 \leqq t \leqq 1\}$ and $\{V_n(t): 0 \leqq t \leqq 1\}$ denote the corresponding one sample empirical processes defined by

$$U_m(t) = m^{\frac{1}{2}}[F_m F^{-1}(t) - t] \quad \text{and} \quad V_n(t) = n^{\frac{1}{2}}[G_n G^{-1}(t) - t]\,. \tag{2.1}$$

The results of this section will rely heavily on Theorems 12.2, 15.2 and 22.1 of Billingsley (1968). Lemma 2.1 below is concerned with the behavior of the U_m-processes for ϕ-mixing sequences near 0 and 1. (K_ϕ or K_ϕ' etc. are used as generic constants throughout.)

LEMMA 2.1. *Let $q(t) = K[t(1-t)]^{\frac{1}{2}-\delta}$, $0 \leqq t \leqq 1$, for some δ with $0 < \delta < \frac{1}{2}$ and assume that the ϕ-mixing sequence $\{X_i\}$ satisfies the conditions imposed in Section 1 including (1.1). Then for given $\varepsilon > 0$, there exists a θ $(0 < \theta < \frac{1}{2})$ and an integer m_0, depending only on ε and ϕ (and not on the particular $\{X_i\}$), such that for $m \geqq m_0$*

$$P\left[\sup_{0 \leqq t \leqq \theta} \left|\frac{U_m(t)}{q(t)}\right| \geqq \varepsilon\right] \leqq \varepsilon\,; \tag{2.2}$$

(m_0 and θ also depend on q and consequently on K and δ).

PROOF. Let $g_t(x) = I_{(-\infty, F^{-1}(t)]}(x) - t$ and consider M real points $0 < s_1 < s_2 < \cdots < s_M = \theta < \frac{1}{2}$, with $s_l = (l\theta/M)$, $1 \leqq l \leqq M$. Then for every pair (j, k), $1 < j < k \leqq M$,

$$\begin{aligned} E\left|\frac{g_{s_k}(X_1)}{q(s_{k-1})} - \frac{g_{s_j}(X_1)}{q(s_{j-1})}\right|^2 &= \frac{s_k(1-s_k)}{q^2(s_{k-1})} + \frac{s_j(1-s_j)}{q^2(s_{j-1})} - \frac{2s_j(1-s_k)}{q(s_{j-1})q(s_{k-1})} \\ &= \frac{s_k}{q^2(s_{k-1})} + \frac{s_j}{q^2(s_{j-1})} - \frac{2s_j}{q(s_{k-1})q(s_{j-1})} - \left[\frac{s_k}{q(s_{k-1})} - \frac{s_j}{q(s_{j-1})}\right]^2 \\ &\leqq \frac{1}{q(s_{k-1})}\left[\frac{s_k}{q(s_{k-1})} - \frac{s_j}{q(s_{j-1})}\right] + \frac{s_j}{q(s_{j-1})}\left[\frac{1}{q(s_{j-1})} - \frac{1}{q(s_{k-1})}\right]. \end{aligned} \tag{2.3}$$

To obtain a suitable bound on the right side of (2.3), first note that the first term in (2.3) which equals

$$\begin{aligned} \frac{1}{q(s_{k-1})} \sum_{j<l\leqq k}\left[\frac{s_l}{q(s_{l-1})} - \frac{s_{l-1}}{q(s_{l-2})}\right] &\leqq \frac{1}{q(s_{k-1})} \sum_{j<l\leqq k} \frac{(s_l - s_{l-1})}{q(s_{l-2})} \\ &\leqq (\theta/M) \sum_{j<l\leqq k} [1/q^2(s_{l-2})]\,, \end{aligned} \tag{2.4}$$

where we have used the monotonicity of q in $(0, \frac{1}{2})$. For the second term in (2.3), which can be expressed as

$$\frac{s_j}{q(s_{j-1})} \sum_{j<l\leqq k}\left[\frac{1}{q(s_{l-2})} - \frac{1}{q(s_{l-1})}\right], \tag{2.5}$$

note that since $((t+\alpha)/q(t))$ is non-decreasing in t for $0 < \alpha \leqq t < \theta$ and $s_l \leqq 2s_{l-1}$ for $2 \leqq l \leqq M$, we obtain, using the mean value theorem, that for each l, $j < l \leqq k$, there exists a t_0 with $s_{l-2} < t_0 < s_{l-1}$, such that

$$\frac{s_j}{q(s_{j-1})}\left[\frac{1}{q(s_{l-2})} - \frac{1}{q(s_{l-1})}\right] = \frac{s_j}{q(s_{j-1})} \frac{(s_{l-1} - s_{l-2})(\frac{1}{2} - \delta)(1 - 2t_0)}{K[t_0(1 - t_0)]^{\frac{3}{2}-\delta}}$$
$$(2.6) \qquad \leqq \frac{s_{l-1}}{q(s_{l-2})} \frac{(\theta/M)(\frac{1}{2} - \delta)(1 - 2t_0)}{q(t_0)[t_0(1 - t_0)]}$$
$$\leqq (\theta/Mq^2(s_{l-2}))\,.$$

From (2.3) to (2.6) and the inequality $|q^2(s_{l-1})/q^2(s_{l-2})| < 2$ for all $2 < l \leqq M$, we obtain

$$(2.7) \qquad E\left|\frac{g_{s_k}(X_1)}{q(s_{k-1})} - \frac{g_{s_j}(X_1)}{q(s_{j-1})}\right|^2 \leqq \frac{4\theta}{M} \sum_{j<l\leqq k} [1/q^2(s_{l-1})]$$

for each $1 < j < k \leqq M$. Further, also for each $1 < k \leqq M$

$$(2.8) \qquad E\left|\frac{g_{s_k}(X_1)}{q(s_{k-1})}\right|^2 \leqq \left(\frac{\theta}{M}\right)\frac{k}{q^2(s_{k-1})} \leqq \frac{2\theta}{M} \sum_{1<l\leqq k} [1/q^2(s_{l-1})]\,.$$

Consider now for each pair (j, k), $1 < j < k \leqq M$, the sequence $\{\eta_i\}_{i=1}^m$ with

$$(2.9) \qquad \eta_i = \eta_i^*/2\{\sum_{j<l\leqq k} [1/q^2(s_{l-1})]\}^{\frac{1}{2}} \qquad \text{and}$$
$$\eta_i^* = \left[\frac{g_{s_k}(X_i)}{q(s_{k-1})} - \frac{g_{s_j}(X_i)}{q(s_{j-1})}\right].$$

and observe that since $[q^2(s_j)/q^2(s_{j-1})] < 2$ for all $1 < j \leqq M$,

$$(2.10) \qquad |\eta_i^*|^2 \leqq 2\left[\frac{1}{q^2(s_{k-1})} + \frac{1}{q^2(s_{j-1})}\right] \leqq 4 \sum_{j<l\leqq k} [1/q^2(s_{l-1})]\,.$$

From (2.9) and (2.10) it follows that $|\eta_i| \leqq 1$ for $1 \leqq i \leqq m$. We can thus apply Lemma 22.1 of Billingsley (1968) to conclude the existence of a constant K_ϕ such that

$$(2.11) \qquad E|\sum_{i=1}^m \eta_i|^4 \leqq K_\phi[m^2E^2(\eta_1^2) + mE(\eta_1^2)]\,,$$

so that from (2.1), (2.7), (2.9) and (2.11) we obtain for $1 < j < k \leqq M$,

$$(2.12) \qquad E\left|\frac{U_m(s_k)}{q(s_{k-1})} - \frac{U_m(s_j)}{q(s_{j-1})}\right|^4 \leqq K_\phi\left(1 + \frac{M}{m\theta}\right)\left(\frac{\theta}{M}\right)^2 [\sum_{j<l\leqq k} (1/q^2(s_{l-1}))]^2\,.$$

Similarly, inequality (2.8) and the same argument yield

$$(2.13) \qquad E\left|\frac{U_m(s_k)}{q(s_{k-1})}\right|^4 \leqq K_\phi\left(1 + \frac{M}{m\theta}\right)\left(\frac{\theta}{M}\right)^2 [\sum_{1<l\leqq k} (1/q^2(s_{l-1}))]^2\,.$$

Now let $\xi_1 = U_m(s_2)/q(s_1)$, $\xi_i = [U_m(s_{i+1})/q(s_i)] - [U_m(s_i)/q(s_{i-1})]$ for $1 < i < M$ ($\xi_0 = 0$) and use (2.12), (2.13) and Theorem 12.2 of [1] to conclude

$$(2.14) \qquad P[\max_{1\leqq i\leqq M} |U_m(s_{i+1})/q(s_i)| \geqq \varepsilon]$$
$$\leqq \frac{K_\phi}{\varepsilon^4}\left(1 + \frac{M}{m\theta}\right)\left[\left(\frac{\theta}{M}\right) \sum_{l=1}^{M-1} (1/q^2(s_l))\right]^2.$$

Now it can be easily seen (as in Billingsley [1]; see (22.17) page 199) that for

$s_i = (i\theta/M) \leqq t \leqq s_{i+1} = ((i+1)\theta/M)$, $i = 1, 2, \cdots, M-1$,

$$|U_m(t)| \leqq |U_m(s_{i+1})| + m^{\frac{1}{2}}(\theta/M) + |U_m(s_i)| ,$$

so that using the monotonicity of q in $(0, \frac{1}{2})$ we obtain

$$\left|\frac{U_m(t)}{q(t)}\right| \leqq \left|\frac{U_m(s_{i+1})}{q(s_i)}\right| + \left|\frac{U_m(s_i)}{q(s_i)}\right| + \frac{m^{\frac{1}{2}}(\theta/M)}{q(s_i)} ,$$

which implies that

$$\sup_{(\theta/M)\leqq t\leqq\theta}\left|\frac{U_m(t)}{q(t)}\right| \leqq 2\max_{1\leqq i\leqq M}\left|\frac{U_m(s_{i+1})}{q(s_i)}\right| + \frac{|U_m(s_1)|}{q(s_1)} + \frac{1}{KM^{\delta}}(2m\theta/M)^{\frac{1}{2}} . \tag{2.15}$$

For given θ, ε and a sufficiently large m (m will depend on θ and ε), now choose M such that

$$\frac{4\theta m}{\varepsilon} > M > \frac{2\theta m}{\varepsilon} \quad \text{and} \quad \frac{1}{KM^{\delta}} \leqq \frac{\varepsilon^{\frac{1}{2}}}{4} ; \tag{2.16}$$

(for large m, say $m \geqq m_0 = m_0(\theta, \varepsilon)$, (2.16) is certainly possible so that M chosen above depends on θ and ε. Later we shall also choose θ suitably). Now since

$$E\left|\frac{U_m(s_1)}{q(s_1)}\right|^2 \leqq \frac{K_\phi'}{K^2} s_1^{2\delta} \qquad (s_1 = (\theta/M)) \tag{2.17}$$

(see Lemma 3 page 172 of [1]), from (2.14), (2.15) and (2.16) we obtain for the choice of M in (2.16)

$$\begin{aligned} &P\left[\sup_{(\theta/M)\leqq t\leqq\theta}\left|\frac{U_m(t)}{q(t)}\right| \geqq \frac{\varepsilon}{2}\right] \\ &\leqq P\left[2\max_{1\leqq i<M}\left|\frac{U_m(s_{i+1})}{q(s_i)}\right| \geqq \frac{\varepsilon}{8}\right] + P\left[\left|\frac{U_m(s_1)}{q(s_1)}\right| \geqq \frac{\varepsilon}{8}\right] \\ &\leqq \frac{K_\phi}{\varepsilon^5}\left(\int_0^\theta \frac{1}{q^2(t)}\,dt\right)^2 + \frac{K_\phi'}{K^2\varepsilon^2} s_1^{2\delta} ; \end{aligned} \tag{2.18}$$

(K_ϕ, K_ϕ' in (2.18) are generic constants depending on ϕ and q alone). Further from (2.16) again, we have

$$\begin{aligned} &P\left[\sup_{0\leqq t<(\theta/M)}\left|\frac{U_m(t)}{q(t)}\right| < \frac{\varepsilon}{2}\right] \\ &\geqq P\left[\left\{\sup_{0\leqq t<(\theta/M)}\left|\frac{U_m(t)}{q(t)}\right| < \frac{\varepsilon}{2}\right\} \cap \left\{F_m F^{-1}\left(\frac{\theta}{M}\right) = 0\right\}\right] \\ &= P\left[\left\{\sup_{0\leqq t<(\theta/M)}\left|\frac{m^{\frac{1}{2}}t}{q(t)}\right| < \frac{\varepsilon}{2}\right\} \cap \left\{F_m F^{-1}\left(\frac{\theta}{M}\right) = 0\right\}\right] \\ &= P\left[F_m F^{-1}\left(\frac{\theta}{M}\right) = 0\right] = 1 - P\left[\bigcup_{i=1}^m \left\{F(X_i) \leqq \frac{\theta}{M}\right\}\right] \\ &\geqq 1 - (m\theta/M) \geqq 1 - \frac{\varepsilon}{2} . \end{aligned} \tag{2.19}$$

From (2.18) and (2.19), we have for $m \geqq m_0 = m_0(\varepsilon, \theta)$

$$P\left[\sup_{0\leqq t\leqq\theta}\left|\frac{U_m(t)}{q(t)}\right| \geqq \varepsilon\right] \leqq \frac{K_\phi}{\varepsilon^5}\left(\int_0^\theta \frac{1}{q^2(t)}\,dt\right)^2 + \frac{K_\phi'}{K^2\varepsilon^2}\theta^{2\delta} + \frac{\varepsilon}{2}. \tag{2.20}$$

Choosing θ sufficiently small so that the sum of the first two terms in (2.20) does not exceed $\varepsilon/2$, the desired result follows. □

Let $C = C[0, 1]$ denote the space of continuous functions on $[0, 1]$ and $D = D[0, 1]$ denote the space of all functions on $[0, 1]$ that are right continuous, possess limits on the left and are continuous on the left at 1. Let ρ denote the uniform metric on D and d the Skorokhod metric (see Billingsley (1968) page 115) which makes (D, d) a complete separable metric space.

THEOREM 2.1. *Let the function q and the ϕ-mixing sequence $\{X_i\}$ satisfy the conditions of Lemma 2.1. Then, as $m \to \infty$, $(U_m/q) \to_L (U_0/q)$ relative to (D, d), where U_0 is a Gaussian random function on $[0, 1]$ specified by*

$$\begin{aligned} &E\{U_0(t)\} = 0 \qquad \text{and for} \quad 0 \leqq s < t \leqq 1 \\ &E\{U_0(s)U_0(t)\} = s(1-t) + \textstyle\sum_{k=2}^\infty E[g_s(X_1)g_t(X_k)] \\ &\qquad\qquad + \textstyle\sum_{k=2}^\infty E[g_t(X_1)g_s(X_k)], \end{aligned} \tag{2.21}$$

where $g_t(x) = I_{(-\infty, F^{-1}(t)]}(x) - t$. Further $P[(U_0/q) \in C] = P[U_0 \in C] = 1$. (Note that for $0 \leqq s, t \leqq 1$, $E[g_s(X_1)\cdot g_t(X_k)] = F_k(F^{-1}(s), F^{-1}(t)) - st$, where F_k is the df of (X_1, X_k).)

PROOF. First note that for any $0 < a < \frac{1}{2} < b < 1$

$$\begin{aligned} \omega_\eta\left(\frac{U_m}{q}\right) &= \sup_{|s-t|<\eta}\left|\frac{U_m(s)}{q(s)} - \frac{U_m(t)}{q(t)}\right| \\ &\leqq \sup_{t\in\{[0,a-\eta)\cup(b+\eta,1]\}}\sup_{|s-t|<\eta}\left|\frac{U_m(s)}{q(s)} - \frac{U_m(t)}{q(t)}\right| \\ &\quad + \sup_{t\in[a-\eta,b+\eta]}\sup_{|s-t|<\eta}\left|\frac{U_m(s)-U_m(t)}{q(s)} \right.\\ &\quad \left. + U_m(t)\left(\frac{1}{q(s)} - \frac{1}{q(t)}\right)\right| \\ &\leqq 2\sup_{0\leqq t<a}\left|\frac{U_m(t)}{q(t)}\right| + 2\sup_{b<t\leqq 1}\left|\frac{U_m(t)}{q(t)}\right| + \frac{\omega_\eta(U_m)}{A} \\ &\quad + \frac{\omega_\eta(q)\cdot\sup_t|U_m(t)|}{A^2}, \end{aligned} \tag{2.22}$$

where $A = \min(q(a-2\eta), q(b+2\eta))$. By virtue of Lemma 2.1, the first term on the right of (2.22) can be made arbitrarily small with high probability, by choosing a sufficiently small and m sufficiently large. The same can be done with respect to the second term by considering the reverse process $U_m^-(t) = U_m((1-t)^-)$ and using a Lemma 2.1-type result for b sufficiently close to 1 and m appropriately large. The same holds for the third term for sufficiently small η and large m in view of (22.13) of Billingsley [1] and the fact that A remains bounded away from zero as $\eta \to 0$. To conclude a similar assertion for the fourth term note the tightness of $U_m(t)$ implies that $\sup_t |U_m(t)|$ is bounded in probability (Theorems 15.2 and 22.1 of Billingsley (1968)) and that $\lim_{\eta\to 0} \omega_\eta = 0$. Hence from (2.22) it follows that for the process $\{U_m(t)/q(t): 0 \leqq t \leqq 1\}$ condition (ii) of Theorem 15.2 of Billingsley is satisfied. The condition (i) of this theorem is also satisfied for large m, say $m \geqq m_0$, by using Lemma 2.1. This establishes the tightness of the sequence $\{(U_m/q): m \geqq m_0\}$ relative to (D, d). Since $U_m \to_L U_0$ relative to (D, d), the convergence of the appropriate finite dimensional distributions of (U_m/q) to those of (U_0/q) is an obvious conclusion. The proof of $(U_m/q) \to_L (U_0/q)$ follows by Theorem 15.1 and that of $P[(U_0/q) \in C] = 1$ by (2.22) above and Theorem 15.5 of Billingsley (1968). □

Let $\rho_q(f, g) = \rho(f/q, g/q)$ and similarly for d_q.

COROLLARY 2.1. *The conclusion of Theorem* 2.1 *can be strengthened to read*

$$\rho_q(U_m, U_0) \to 0 \qquad \text{and} \qquad d_q(U_m, U_0) \to 0\,,$$

as $m \to \infty$, *where* U_m *and* U_0 *are now processes equivalent* (*in the sense of Skorokhod* [10] *item* 3.1.1), *respectively, to* U_m *and* U_0 *of Theorem* 2.1.

PROOF. We may use Theorem 2.1 and the above referred theorem of Skorokhod to replace (U_m/q) and (U_0/q) by "equivalent" processes ξ_m and ξ_0, respectively, (on a space $(\tilde\Omega, \tilde{\mathscr{A}}, \tilde P)$) such that $d(\xi_m, \xi_0) \to_{\text{a.s.}} 0$. Setting $\tilde U_m = q\xi_m$ $(m \geqq 0)$ and ignoring $\sim$ for convenience we obtain the d-convergence above. The ρ-convergence a.s. follows since $P[(U_0/q) \in C] = 1$, the ρ and d convergence being then equivalent. □

3. Weak convergence of the L_N-process. We now turn to the convergence of the process L_N. Let U_0 be the Gaussian process defined by (2.21) and V_0 an independent process given by (2.21) with X's (F) replaced by Y's (G). Given Corollary 2.1 for the equivalent process U_m and V_n $(m, n \geqq 0)$, the remaining arguments below run parallel to those of [7] and [8]; and so we shall simply sketch the proofs, providing details only when they seem necessary. Let

$$(3.1) \qquad K_N = FH_N^{-1}\,, \qquad K_\lambda = FH_\lambda^{-1}\,, \qquad K = K_{\lambda_N} = FH^{-1} \qquad \text{and}$$
$$K_0 = FH_0^{-1} \qquad \text{where} \quad H_0 = H_{\lambda_0}\,.$$

Further, let a_λ, a_N and a_0 denote the a.e. (w.r. to Lebesgue measure) derivatives of K_λ, K and K_0 and b_λ, b_N and b_0 denote the a.e. derivatives of GH_λ^{-1}, GH^{-1} and GH_0^{-1} respectively. Define $\{L_0(t); 0 \leqq t \leqq 1\}$ by

$$(3.2) \qquad L_0 = (1 - \lambda_0)\{\lambda_0^{-\frac12} b_0 U_0(FH_0^{-1}) - (1 - \lambda_0)^{-\frac12} a_0 V_0(GH_0^{-1})\}\,;$$

and define L_N' by letting it equal L_N on $[1/N, 1]$ and equal 0 on $[0, 1/N]$. Now assume that

(3.3) $\quad a_\lambda$ exist for all $t \in (0, 1)$ and for some λ', $a_{\lambda'}$ is continuous on $(0, 1)$ and has left (right) limit at one (zero).

(Condition (3.3) implies that the same holds for b_λ also; see also Corollary 4.1 of [7] for conditions under which (3.3) holds.) For the equivalent processes we can now state

THEOREM 3.1. *Let* $0 < \tau < \frac{1}{2}$ *and* $q(t) = K[t(1-t)]^{\frac{1}{2}-\tau}$ *for* $t \in [0, 1]$. (a) *If* (3.3) *holds and* $\lambda_N \to \lambda_0$, *as* $N \to \infty$, *then* $\rho_q(L_N', L_0) \to_{\text{a.s.}} 0$; *further* (b) *if* a_0 *exists* a.e. ν, *for some Lebesgue–Stieltjes signed measure* ν *on* (0, 1) *with* $\int_0^1 q\, d|\nu| < \infty$, *and* $\lambda_N = \lambda_0 + O(N^{-\frac{1}{2}})$, *then* $\|L_N' - L_0\|_\nu = \int |L_N' - L_0|\, d|\nu| \to 0$, *as* $N \to \infty$. (The processes L_N and L_0 here are the "equivalent" processes.)

PROOF. The proof is analogous to those of Theorems 4.1 (a) and 5.1 (a) of [7]. To see this, first note that because of the absolute continuity assumption for finite dimensional distributions of $\{X_i\}$ and $\{Y_i\}$, the a.s. representation (3.2) of [7] for the L_N-process (defined by (1.5)) remains true. Accordingly, if we show that the conclusion of Theorem 2.2 of [7] remains true in the present situation, the rest of the arguments of [7] would apply verbatim.

First, to show that Lemma 2.3 of [7] remains valid, we need only remark that with probability one F_m converges uniformly to F for ergodic, and therefore for ϕ-mixing, stationary processes by Tucker (1959). Lemma 2.5 of [7] also requires a minor change. Its statement should now read: for given $\varepsilon, \tau > 0$ $(\tau < \frac{1}{2})$, there exist a $\beta > 0$ and an N_0 such that for $N \geqq N_0$,

$$P\left[K_N(t) \leqq \beta t^{1-\tau} \text{ for } t \geqq \frac{1}{N}\right] > 1 - \varepsilon\,. \tag{3.4}$$

To prove (3.4), note that since $K(t) \leqq \lambda_*^{-1}t$ for all $t \in [0, 1]$ and $\rho(K_N, K) \to_{\text{a.s.}} 0$, the problem reduces to the study of the intervals $[0, \theta]$ and $[1-\theta, 1]$ for sufficiently small θ. Clearly we need to consider only the interval $[0, \theta]$. We now choose θ according to Lemma 2.1 such that

$$P[A_m] \geqq 1 - \varepsilon\,, \qquad \text{where} \quad A_m = \{U_m(t) \leqq q'(t) \ \text{ for } \ 0 \leqq t \leqq \theta\} \tag{3.5}$$

with $q'(t) = [t(1-t)]^{\frac{1}{2}-\delta}$ and $\delta = \tau/2(1-\tau)$. Now if $F_m H_N^{-1}(t) > 0$, then $(1/m) \leqq F_m H_N^{-1}(t) \leqq 2\lambda_*^{-1}t$, so that on A_m,

$$\begin{aligned} K_N(t) &= F_m H_N^{-1}(t) - m^{-\frac{1}{2}} U_m(K_N(t)) \\ &\leqq F_m H_N^{-1}(t) + (F_m H_N^{-1}(t))^{\frac{1}{2}} q'(K_N(t)) \\ &\leqq (2\lambda_*^{-1}t) + (2\lambda_*^{-1}t)^{\frac{1}{2}}(K_N(t))^{\frac{1}{2}-\delta}\,. \end{aligned} \tag{3.6}$$

Setting $K_N(t) = z$ and $(2\lambda_*^{-1}t)^{\frac{1}{2}} = u$ in (3.6), we obtain $z \leqq u^2 + u(z^{\frac{1}{2}-\delta})$ which implies, by completing the square on the right and some manipulation, $z^{\frac{1}{2}-\delta}[(4z^{2\delta} + 1)^{\frac{1}{2}} - 1] \leqq 2u$. Since $1 + 4a \geqq (1+a)^2$ for $0 \leqq a \leqq 2$, this yields $z^{\frac{1}{2}+\delta} \leqq 2u$, so that

$$K_N(t) \leqq (2u)^{2/(1+2\delta)} = \beta t^{1-\tau}\,, \tag{3.7}$$

with $\beta = (8/\lambda_*)^{1-\tau} > 1$. Now if $t \geqq 1/N$, but $0 = F_m H_N^{-1}(t) < 1/m$ then on A_m, $K_N = m^{-\frac{1}{2}}|U_m(K_N)| \leqq m^{-\frac{1}{2}}(K_N)^{\frac{1}{2}-\delta}$ which also yields (3.7); thus (3.4) follows from

(3.5) and (3.7). The proof of Theorem 2.2 of [7] also needs some explanation. Let the first inequality now read: By Lemma 2.1 choose $\alpha > 0$ such that for sufficiently large N ($\geqq N_0$)

$$P[|U_m(t)| \leqq \varepsilon q_1((t/\beta)^{1-\tau}) \text{ for all } 0 < t < \alpha] > 1 - \varepsilon\,, \tag{3.8}$$

where β is the same as in (3.4) and $q_1(t) = [t(1-t)]^{\frac{1}{2}}$. This is certainly possible since there exists a constant K^* such that $q_1(t^{1-\tau}/\beta) \geqq K^*[t(1-t)]^{(1-\tau)/2}$. Further, we can find an $\eta > 0$ such that for sufficiently large N

$$P[K_N(t) < \alpha \quad \text{for all } 0 < t \leqq \eta] > 1 - \varepsilon\,. \tag{3.9}$$

Since $q_1(t^{(1-\tau)^2}) \leqq [t(1-t)]^{\frac{1}{2}-\tau}$, the conclusion $\Pr - \lim_{N\to\infty} \rho_q(U_m^*(K_N), U_0(K)) = 0$ of Theorem 2.2 of [7] is obtained, as in [7], from (3.4), (3.8), (3.9) and the fact that $\sup_{0\leqq t\leqq\theta} |U_0(t)/q(t)| = o(1)$, as $\theta \to 0$; the last fact follows from Lemma 2.1, Corollary 2.1 and the inequality

$$|U_0(t)/q(t)| \leqq |U_m(t)/q(t)| + \rho_q(U_m, U_0)\,.$$

The proof is complete. □

The conclusion of Corollary 2.1, and therefore of Theorem 3.1, remains true if q is replaced by any function q^* which, for some $0 < \varepsilon < \frac{1}{2}$, is bounded below by a $q = K[t(1-t)]^{\frac{1}{2}-\delta}$ ($K > 0$, $0 < \delta < \frac{1}{2}$) on $[0, \varepsilon)$ and $(1-\varepsilon, 1]$ and bounded away from zero on $[\varepsilon, 1-\varepsilon]$. However, for proving Chernoff–Savage type theorems, the function q considered above suffices.

4. Chernoff–Savage theorems. We now turn to the asymptotic normality of T_N and conclude this paper by stating a Chernoff–Savage theorem on the lines of Theorem 1 of Pyke and Shorack (1969) and subsequent remarks. Let $\mu = \int_0^1 J(t)\, dFH^{-1}(t)$ where $-J$ denotes a non-constant function of bounded variation on $(\varepsilon, 1-\varepsilon)$ for all $\varepsilon > 0$ and which induces the Lebesgue–Stieltjes measure ν on (0, 1). We need the following conditions:

(C1) $|J(t)| \leqq K[t(1-t)]^{\frac{1}{2}-\delta}$ for some K, $\delta > 0$,
(C2) $N^{-\frac{1}{2}} \sum_{i=1}^N |c_{Ni}^* - J[\min((i/N), 1-1/N)]| \leqq \delta_N$, with $\delta_N = o(1)$,
(C3) $N^{\frac{1}{2}}|\lambda_N - \lambda_0| = O(1)$ and (C3′),
(C3′) FH^{-1} is differentiable a.e. $|\nu|$ for $N \geqq$ some N_0.

THEOREM 4.1. *Let* $\{X_i\}$ *and* $\{Y_i\}$ *be two independent strictly stationary* ϕ*-mixing sequences of* rv's *satisfying conditions* (ii) *and* (1.1) *of Section* 1. *If* (C1), (C2) *and either* (C3) *or* (3.3) *hold, then as* $N \to \infty$

$$\tilde{T}_N = N^{\frac{1}{2}}(T_N - \mu) \to_P \int_0^1 L_0\, d\nu\,, \tag{4.1}$$

a $N(0, \sigma_0^2)$ rv *with finite variance given by*

$$\sigma_0^2 = \int_0^1 \int_0^1 \operatorname{Cov}(L_0(u), L_0(v))\, d\nu(u)\, d\nu(v)\,.$$

PROOF. The result follows from Theorem 1 of [8], Proposition 5.1 of [7] and Theorem 3.1 above, provided σ_0^2 is shown to be finite. For this the inequality

(K_1 below is some constant)

$$
\begin{aligned}
&|\textstyle\int_0^1 \int_0^t E[U_0(FH_0^{-1}(t))U_0(FH_0^{-1}(s))]\,d\nu(s)\,d\nu(t)| \\
&\quad \leqq K_1 \textstyle\int_0^1 \int_0^t [s(1-t) + 2[s(1-s)t(1-t)]^{\frac{1}{2}} \sum_{k=1}^{\infty} \phi_k^{\frac{1}{2}}]\,d|\nu|(s)\,d|\nu|(t) \\
(4.2) &\quad \leqq K_1[1 + 2\textstyle\sum_{k=1}^{\infty} \phi_k^{\frac{1}{2}}] \int_0^1 \int_0^t \{(s(1-s)t(1-t))^{\frac{1}{2}}/q(t)q(s)\} \\
&\qquad \times q(t)q(s)\,d|\nu|(s)\,d|\nu|(t) \\
&\quad < \infty
\end{aligned}
$$

is sufficient, the first inequality in (4.2) following from $FH_0^{-1}(t) \leqq \lambda_*^{-1}t$, $1 - FH_0^{-1}(t) \leqq \lambda_*^{-1}(1-t)$ and the inequality $E|g_s(X_1)g_t(X_k)| \leqq 2\phi_{k-1}^{\frac{1}{2}}[s(1-s)t(1-t)]^{\frac{1}{2}}$ (see (20.23) of [1]). □

If (C3′) replaces (C3) in above or if (3.3) is replaced by "the a_N's of Lemma 4.2 of [7] form a uniformly equicontinuous family of functions on [0, 1]", then (4.1) becomes $\tilde{T}_n - \int_0^1 L_{0N}\,d\nu \to_P 0$, as $N \to \infty$; (see [7] for a definition of L_{0N} and the outline of proof).

For studying the asymptotic relative efficiency of the statistic T_N relative to another statistic in a given testing or estimation problem involving ϕ-mixing dependence, one would need a Chernoff–Savage theorem for appropriate "contiguous" sequences of the type $\{X_{Ni}\}$ and $\{Y_{Ni}\}$ which also depend on N. (For example, the case when the ϕ-mixing sequences $\{X_i\}$ and $\{Y_i\}$ are replaced by $\{X_i + \tau_N\}$ and $\{Y_i + \eta_N\}$ with constants $\tau_N = N^{-\frac{1}{2}}\tau$ and $\eta_N = N^{-\frac{1}{2}}\eta$, leaving the dependence structure of the two sequences unchanged.) While in simple special cases, it is not difficult to deduce results of Theorem 4.1 type, the question of "uniform" convergence in distribution over sets of ϕ-mixing sequences $\{X_i\}$ and $\{Y_i\}$ needs further investigation. In this connection, it is worth noting that under conditions (C1), (C2) and (C3) the convergence (4.1) above does hold "uniformly" in λ_N, c_{Ni}^*'s and J in the sense of Theorem 1 of [8]. Finally, the extension of above results to the c-sample case and to the case of a ϕ-mixing "type" dependence between $\{X_i\}$ and $\{Y_i\}$ can be accomplished in an obvious manner.

REFERENCES

[1] BILLINGSLEY, PATRICK (1968). *Weak Convergence of Probability Measures*. Wiley, New York.

[2] CHERNOFF, H. and SAVAGE, I. R. (1958). Asymptotic normality and efficiency of certain nonparametric test statistics. *Ann. Math. Statist.* **29** 972–994.

[3] GASTWIRTH, J. L., RUBIN, H., and WOLF, S. S. (1967). The effect of autoregressive dependence on a nonparametric test. *IEEE Trans. Information Theory* **IT-13** 311–313.

[4] GASTWIRTH, J. L. and RUBIN, H. (1971). Effect of dependence on the level of some one-sample tests. *J. Amer. Statist. Assoc.* **66** 816–820.

[5] GOVINDARAJULU, Z., LECAM, L. and RAGHAVACHARI, M. (1967). Generalizations of theorems of Chernoff–Savage on asymptotic normality of nonparametric statistics. *Proc. Fifth Berkeley Symp. Math. Statist. Prob.* 609–638. Univ. of California Press.

[6] HÁJEK, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* **39** 325–346.

[7] PYKE, R. and SHORACK, G. (1968). Weak convergence of two-sample empirical process and a new approach to Chernoff–Savage theorems. *Ann. Math. Statist.* **39** 755–771.

[8] PYKE, R. and SHORACK, G. (1969). A note on Chernoff-Savage theorems. *Ann. Math. Statist.* **40** 1116-1119.

[9] SERFLING, R. J. (1968). The Wilcoxon two-sample statistics on strongly mixing processes. *Ann. Math. Statist.* **39** 1202-1209.

[10] SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes. *Theor. Probability Appl.* (translation in *SIAM*) **1** 261-290.

[11] TUCKER, H. G. (1959). A generalization of Glivenko-Cantelli theorem. *Ann. Math. Statist.* **30** 828-830.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA, CANADA