The Annals of Mathematical Statistics
1968, Vol. 39, No. 3, 755–771

WEAK CONVERGENCE OF A TWO-SAMPLE EMPIRICAL PROCESS AND A NEW APPROACH TO CHERNOFF-SAVAGE THEOREMS[1]

By Ronald Pyke and Galen R. Shorack

University of Washington

0. Summary. An empirical stochastic process for two-sample problems is defined and its weak convergence studied. The results are based upon an identity which relates the two-sample empirical process to the more usual one-sample empirical process. Based on this identity a relatively simple proof of a Chernoff-Savage theorem is obtained. The c-sample analogues of these results are also included.

1. Introduction. Let $X_1, \cdots, X_m, \cdots$, and $Y_1, \cdots, Y_n, \cdots$ be independent sequences of independent random variables (rv) with common continuous distribution functions (df) F and G respectively. Let $N = m + n$ and $\lambda_N = m/N$ and suppose that $0 < \lambda_* \leqq \lambda_N \leqq 1 - \lambda_* < 1$ for some $\lambda_* > 0$. (Here, and throughout, when N is used as a subscript it denotes the pair (m, n).) Let $\Lambda = [\lambda_*, 1 - \lambda_*]$.

When dealing with a single sample $X_1, \cdots, X_m$ most distribution-free tests and statistics are based upon the empirical process $\{U_m(t): 0 \leqq t \leqq 1\}$ defined by

$$U_m(t) = m^{\frac{1}{2}}[F_m F^{-1}(t) - t], \tag{1.1}$$

where F_m is the empirical df of the sample and where for any two df's, F, G we write FG^{-1} for the composed function $FG^{-1}(t) = F(G^{-1}(t))$. (Throughout this paper we take inverse functions to be left continuous; thus $F^{-1}(t) = \inf\{x: F(x) \geqq t\}$.) These 1-sample empirical processes have been studied extensively. Much is known about them including their weak convergence to a separable *tied-down Wiener process* $\{U_0(t): 0 \leqq t \leqq 1\}$; that is, to a Gaussian process with $E[U_0(t)] = 0$ and $E[U_0(s)U_0(t)] = s(1-t)$, $0 \leqq s \leqq t \leqq 1$.

Similarly for the second sample, let G_n denote the empirical df of $Y_1, \cdots, Y_n$ and let $\{V_n(t): 0 \leqq t \leqq 1\}$ be the corresponding empirical process

$$V_n(t) = n^{\frac{1}{2}}[G_n G^{-1}(t) - t]. \tag{1.2}$$

By assumption, the U_m- and V_n-processes are independent. Let $\{V_0(t): 0 \leqq t \leqq 1\}$ denote a second tied-down Wiener process independent of the U_0-process.

Define

$$K_N = FH_N^{-1}, \qquad K_\lambda = FH_\lambda^{-1}, \qquad K = K_{\lambda_N} = FH^{-1}, \tag{1.3}$$

Received 13 February 1967; revised 4 October 1967.

[1] This research was supported in part by the National Science Foundation under G-5719.

where

$$(1.4)\qquad H_N = \lambda_N F_m + (1-\lambda_N)G_n, \qquad H_\lambda = \lambda F + (1-\lambda)G, \qquad H = H_{\lambda_N}.$$

Most non-parametric tests and statistics for the 2-sample problem are based upon the relative ranks of one of the samples. Suppose we let $\{R_{Ni} : 1 \leqq i \leqq N\}$ denote the set of relative ranks of the X-sample; that is, R_{Ni} is the number of observations among $X_1, \cdots, X_m$ which do not exceed the ith order statistic of the combined sample $X_1, \cdots, X_m, Y_1, \cdots, Y_n$. In terms of previous notation one sees that $R_{Ni} = mF_m H_N^{-1}(i/N)$. It is thus natural to propose the 2-sample empirical process $\{L_N(t) : 0 \leqq t \leqq 1\}$ defined by

$$(1.5)\qquad L_N(t) = N^{\frac{1}{2}}[F_m H_N^{-1}(t) - FH^{-1}(t)].$$

We will now show how the L_N-process relates to the linear rank statistics studied by Chernoff and Savage (1958). Define

$$(1.6)\qquad T_N = m^{-1}\sum_{i=1}^N c_{Ni}R_{Ni} = \sum_{i=1}^N c_{Ni}F_m H_N^{-1}(i/N)$$

where $\{c_{Ni} : 1 \leqq i \leqq N\}$ is a given set of constants. If ν_N denotes the signed measure which puts measure c_{Ni} on the point i/N for $1 \leqq i \leqq N$, and puts zero measure elsewhere, then (1.6) may be written as

$$(1.7)\qquad T_N = \int_0^1 F_m H_N^{-1}\,d\nu_N.$$

The class of linear rank statistics described by (1.6) or (1.7) is the same as the one considered by Chernoff and Savage (1958). This can be seen by writing $Z_{Ni} = R_{Ni} - R_{Ni-1}$ for $1 < i \leqq N$ and $Z_{N1} = R_{N1}$. A summation by parts in (1.6) then yields

$$(1.8)\qquad T_N = m^{-1}\sum_{i=1}^N c_{Ni}^* Z_{Ni}$$

where $c_{Ni}^* = c_{Ni} + c_{Ni+1} + \cdots + c_{NN}$. Although the difference between (1.6) and (1.8) is only due to a summation by parts, the relative simplicity of the proofs of Theorems 4.1 and 5.1 given below indicates the usefulness and naturalness of the representation (1.6).

Write $\mu_N = \int_0^1 K\,d\nu_N$ and set

$$(1.9)\qquad T_N^* = N^{\frac{1}{2}}(T_N - \mu_N) = \int_0^1 L_N\,d\nu_N.$$

Let ν be another signed Lebesgue-Stieltjes measure on (0,1) for which $|\nu|([\epsilon, 1-\epsilon]) < \infty$ for all $\epsilon > 0$ where $|\nu| = \nu^+ + \nu^-$ is the total variation of ν. In the statement of Theorems 4.1 and 5.1 conditions are given to insure that $\{\nu_N\}$ converges to ν in such a way as to permit one to substitute ν for ν_N in (1.9).

In Section 2, various results about 1-sample empirical processes are described. In Section 3, the basic identity relating the L_N-process to the 1-sample empirical processes of the two samples is presented. The results of Sections 2 and 3 are used in Sections 4 and 5 to establish in a relatively simple way the weak convergence of the L_N-process relative to various metrics. These results are then used to obtain central limit theorems for linear rank statistics. The extension of all results to the c-sample problem is given in Section 6. Related work is contained in Pyke and Shorack ((1967a) and (1967b)). In the former the analogous limit theorems for random sample sizes are obtained, whereas in the latter questions of asymptotic relative efficiency are considered.

2. The 1-sample empirical process. For $m \geqq 0$ let $\{W_m(t): 0 \leqq t \leqq 1\}$ denote stochastic processes on a probability space $(\Omega, \mathfrak{A}, P)$ whose sample functions are points in some metric space $(\mathfrak{M}, \delta)$.

DEFINITION 2.1. We write $W_m \rightarrow_L W_0$ relative to $(\mathfrak{M}, \delta)$ if $\lim_{m\to\infty} E[\psi(W_m)] = E[\psi(W_0)]$ for all bounded real functionals ψ defined on $\mathfrak{M}$ which are continuous in the δ-metric and are such that $\psi(W_m)$, $m \geqq 0$, are measurable with respect to $\mathfrak{A}$. Such convergence is called *convergence in law* or *weak convergence*. (If the W_m process and W_0-process are measurable with respect to the Borel sets of $(\mathfrak{M}, \delta)$, so that their image laws on $\mathfrak{M}$ are well defined, then the above definition is equivalent to the usual definition of weak convergence as given in Prokhorov (1956) for example.)

Suppose $\mathfrak{M} = D$, the set of all right continuous real valued functions on [0, 1] having only jump discontinuities. In this case two possible metrics are $\delta = \rho$, the uniform metric defined by

$$\rho(f, g) = \sup_{0 \leqq t \leqq 1} |f(t) - g(t)|, \tag{2.1}$$

and $\delta = d$, the Prokhorov metric on D as defined by Prokhorov (1956). Prokhorov showed that (D, d) is a complete separable metric space and that $U_m \rightarrow_L U_0$ relative to (D, d). Actually, since all jumps of the U_m-process equal $m^{-\frac{1}{2}}$, it is possible to show that $U_m \rightarrow_L U_0$ relative to the stronger uniform topology of the non-separable metric space (D, ρ). We will obtain this result as Lemma 2.1. (It should be pointed out that Dudley (1966) gives a definition of weak convergence for non-separable spaces which is more general than Definition 2.1 above in that his use of upper and lower integrals enables him to place a less restrictive assumption of measurability upon the function ψ. Also, the statement "$U_m \rightarrow_L U_0$ on (D, ρ) in the sense of Prokhorov's definition" is false; see Chibisov (1965) for a statement of the measurability difficulties.)

Since $U_m \rightarrow_L U_0$ relative to (D, d) and (D, d) is a complete separable metric space it is possible, (see item 3.1.1 of Skorokhod (1956)), to construct processes $\{\tilde{U}_m(t): 0 \leqq t \leqq 1\}$, $m \geqq 0$, with sample functions in D and having the same finite dimensional df's as $\{U_m(t): 0 \leqq t \leqq 1\}$, $m \geqq 0$, but which in addition satisfy $d(\tilde{U}_m, \tilde{U}_0) \rightarrow_{\text{a.s.}} 0$. Let us make an independent construction for the V_n-processes so that

$$d(\tilde{U}_m, \tilde{U}_0) \rightarrow_{\text{a.s.}} 0, \qquad d(\tilde{V}_n, \tilde{V}_0) \rightarrow_{\text{a.s.}} 0 \tag{2.2}$$

where all processes are defined on a single probability space $(\tilde{\Omega}, \tilde{\mathfrak{A}}, \tilde{P})$. This is the probability space we shall work on in what follows. Note that if we set $\tilde{F}_m = m^{\frac{1}{2}}\tilde{U}_m(F) + F$, then $\tilde{F}_m$ is a.s. a df having exactly m discontinuities each of magnitude m^{-1}. (We shall henceforth drop the symbol $\sim$ from the notation.)

Based on the above construction, we shall prove a series of lemmas about the U_m-processes.

WARNING. The results obtained below which involve convergence stronger than convergence in law may apply only to the specially constructed processes. Only the implied convergence in law should be assumed to hold for the original processes unless further checking is done.

LEMMA 2.1. *In view of* (2.2), $\rho(U_m, U_0) \to_{\text{a.s.}} 0$ *as* $m \to \infty$.

PROOF. Since U_0 is continuous with probability one, it suffices to invoke Theorem 1 of Appendix 1 of Prokhorov (1956) which states that d-convergence to a continuous function is equivalent to ρ-convergence. □

For many applications, such as those discussed in Sections 4 and 5 below, the weak convergence of the empirical processes in (D, ρ) is not sufficient. Rather, it is necessary to make use of the specific behavior of these processes near 0 and 1. For the U_0-process, this behavior is best described by the law of the iterated logarithm. Set

$$q_0(t) = \{2t(1 - t) \log [-\log t(1 - t)]\}^{\frac{1}{2}}.$$

Then by the law of the iterated logarithm (cf. Lévy (1948)) it follows that

$$P[\limsup_{t \to a} |U_0(t)/q_0(t)| = 1] = 1, \qquad a = 0, 1. \tag{2.3}$$

To see this, let $\{W(s): s \geqq 0\}$ be Brownian motion with mean zero, $E[W(s)]^2 = s$ and continuous paths a.s. By a transformation of Doob (1949), the finite dimensional df's of $\{U_0(t): 0 \leqq t \leqq 1\}$ agree with those of $\{(1 - t)W(t/(1 - t))$ $0 \leqq t \leqq 1\}$. By applying the usual law of the iterated logarithm to the W-process, (both at $s = 0$ and at $s = +\infty$), and using this transformation, one obtains (2.3).

To study the behavior near 0 and 1 of the U_m-processes we use the following result. This result is equivalent to Lemma 7 of Govindarajulu et al. (1967), but the proof given here is simpler. For a similar result, see the result of Chibisov (1964) quoted as (4.5) in Chibisov (1965).

LEMMA 2.2. *Let q be any non-negative function which is non-decreasing on* $[0, \theta]$ *for* $\theta < 1$. *Then there exists a constant* $c_\theta > 0$ *such that*

$$P[|U_m(t)| \leqq q(t), 0 \leqq t \leqq \theta] \geqq 1 - c_\theta \textstyle\int_0^\theta [q(t)]^{-2}\, dt \tag{2.4}$$

for all $m \geqq 0$. *Moreover,* c_θ *is non-increasing as* $\theta \to 0$.

PROOF. For $m \geqq 1$, let $\{\nu_m(t): 0 \leqq t \leqq \infty\}$ be a separable Poisson process with $E[\nu_m(1)] = m$. Then the left hand side of (2.4) equals $P(A \mid \nu_m(1) = m)$ where $A = [|\nu_m(t) - mt| \leqq m^{\frac{1}{2}}q(t), 0 \leqq t \leqq \theta]$.

Since $\{\nu_m(t) - mt: t \geqq 0\}$ is a martingale, $\{|\nu_m(t) - mt|^2: t \geqq 0\}$ is a submartingale with $E|\nu_m(t) - mt|^2 = mt$. Therefore, by Theorem 5.1 of Birnbaum and Marshall (1961)

$$P(A^c) \leqq \textstyle\int_0^\theta [q(t)]^{-2}\, dt. \tag{2.5}$$

(Since the ν_m-process has independent increments, (2.5) may also be obtained directly from the Hájek-Rényi version of Kolmogorov's inequality; cf. Gnedenko (1962).) Furthermore, since a Poisson process has stationary and independent increments, we may write

$$\begin{aligned}(2.6)\qquad & P(A^c \mid \nu_m(1) = m) \\ &= \textstyle\sum_{k=0}^m P(A^c \cap [\nu_m(\theta) = k] \cap [\nu_m(1) - \nu_m(\theta) = m - k]) \\ &\quad \cdot (P[\nu_m(1) = m])^{-1} \\ &= \textstyle\sum_{k=0}^m P(A^c \cap [\nu_m(\theta) = k]) P[\nu_m(1-\theta) = m - k](P[\nu_m(1) = m])^{-1}.\end{aligned}$$

If Y is a Poisson rv with parameter λ, then $P[Y = k] \leqq P[Y = [\lambda]]$ for every k where $[\lambda]$ is the greatest integer in λ. By applying this to $P[\nu_m(1-\theta) = m - k]$, it then follows by Stirling's approximation that there exists a constant $c_\theta = O((1-\theta)^{-\frac{1}{2}})$ such that $P[\nu_m(1-\theta) = m-k]/P[\nu_m(1) = m] \leqq c_\theta$ for all k. Using this bound, (2.6) gives $P(A^c \mid \nu_m(1) = m) \leqq c_\theta P(A^c)$ which in conjunction with (2.5) completes the proof for $m \geqq 1$.

For $m = 0$, simply apply the Birnbaum-Marshall inequality directly to the transformed W-process described following (2.3). Thus

$$\begin{aligned}P[|U_0(t)| &\leqq q(t), 0 \leqq t \leqq \theta] \\ &= P[(1-t)|W(t/(1-t))| \leqq q(t), 0 \leqq t \leqq \theta] \\ &= P[|W(s)| \leqq (1+s)q(s/(1+s)), 0 \leqq s \leqq \theta^* = \theta/(1-\theta)] \\ &\geqq 1 - \textstyle\int_0^{\theta^*}(1+s)^{-2}[q(s/(1+s))]^{-2}\,ds = 1 - \int_0^\theta [q(t)]^{-2}\,dt\end{aligned}$$

by the Birnbaum-Marshall inequality since $E[W(s)]^2 = s$. (This result does not use (2.2).) □

From Lemma 2.2 we see the possibility of obtaining weak convergence of the U_m-process relative to more general metrics on D. Whenever well defined set

$$(2.7)\qquad \rho_q(f, g) = \rho(f/q, g/q) \qquad \text{and} \qquad d_q(f, g) = d(f/q, g/q).$$

DEFINITION 2.2. Let $\mathbf{Q}'$ denote the class of all non-negative functions defined on [0, 1] which for some $\epsilon > 0$ are bounded away from zero on $(\epsilon, 1-\epsilon)$, are non-decreasing (non-increasing) on $[0, \epsilon]([1-\epsilon, 1])$ and which have square-integrable reciprocals. Let $\mathbf{Q} = \{q \,\varepsilon\, D : q \geqq q' \text{ for some } q' \,\varepsilon\, \mathbf{Q}'\}$.

THEOREM 2.1. *For* $q \,\varepsilon\, \mathbf{Q}$, $\rho_q(U_m, U_0) \to_p 0$ *and* $d_q(U_m, U_0) \to_p 0$ *as* $m \to \infty$.

PROOF. For any $0 < a \leqq \epsilon \leqq 1 - \epsilon \leqq b < 1$,

$$\begin{aligned}(2.8)\qquad \rho_q(U_m, U_0) \leqq \sup\nolimits_{0 \leqq t \leqq a, b \leqq t \leqq 1} (|U_m(t)| + |U_0(t)|)/q(t) \\ + \rho(U_m, U_0)/\inf\nolimits_{a \leqq t \leqq b} q(t),\end{aligned}$$

where ϵ is a value associated with q as in Definition 2.2. For each such a, b, the second term on the right converges to zero in probability by Lemma 2.1. Fix $\delta > 0$. Since q^{-2} is integrable, choose $a < \epsilon$ so that $\int_0^a [q(t)]^{-2}\,dt < c_a^{-1}\delta^3$. Then by Lemma 2.2

$$(2.9)\qquad P[\sup\nolimits_{0 \leqq t \leqq a} |U_m(t)/q(t)| \leqq \delta] \geqq 1 - c_a\delta^{-2} \textstyle\int_0^a [q(t)]^{-2}\,dt > 1 - \delta.$$

Since the reversed process defined by $U_m^-(t) = U_m((1-t)-)$ has the same finite dimensional df's as the U_m-process, one may in a similar manner choose $b \geqq 1 - \epsilon$ for which $P[\sup_{b \leqq t \leqq 1} |U_m(t)/q(t)| \leqq \delta] > 1 - \delta$. Since this inequality and (2.9) also hold for $m = 0$, the first term on the right hand side of (2.7) does not exceed 4δ on an event whose probability exceeds $1 - 4\delta$. This completes the proof for ρ_q.

It is straightforward to show that there is a constant c for which $d \leqq c\rho$; Lemma 4.3 of Chibisov (1965) states that $c = 4$ works. Thus, the result for ρ implies the result for d. (For continuous q we could have used Theorem 1 of Appendix 1 of Prokhorov (1956) which states that ρ-convergence to a continuous function is equivalent to d-convergence.) □

This theorem implies that for $q \varepsilon \mathbf{Q}$, $U_m \rightarrow_L U_0$ relative to (D, ρ_q) and relative to (D, d_q). Related results are given by Chibisov (1964), (1965) who proves the weak convergence of the empirical processes relative to d_q when q is a function of degree $\frac{1}{2}$ in neighborhoods of 0 and 1 and satisfies

$$\int_0^1 \exp\{-cq^2(u)/u(1-u)\}[u(1-u)]^{-1}\,du < \infty \quad \text{for all} \quad c > 0.$$

(A function h is of degree r if $h(cu)/h(u)$ converges to c^r as u approaches a specified limit.) In view of this weak convergence and the completeness and separability of (D, d_q), one may apply Skorokhod's result, as was done prior to (2.2), to obtain versions of the empirical processes for which $d_q(U_m, U_0) \rightarrow_{\text{a.s.}} 0$ and $\rho_q(U_m, U_0) \rightarrow_{\text{a.s.}} 0$. It can be shown that the remaining results of this section are valid for these q functions of Chibisov and that therefore the results of the following sections remain valid for this larger class of q functions

The basic identity to be derived as Lemma 3.1 will express L_N as a random linear combination of $U_m(FH_N^{-1})$ and $V_n(GH_N^{-1})$. We therefore need an analogue of Theorem 2.1 for the process $U_m(FH_N^{-1}) = U_m(K_N)$; the analogue will appear as Theorem 2.2.

LEMMA 2.3. *As* $N \rightarrow \infty$, $\rho(HH_N^{-1}, HH^{-1}) \rightarrow_{\text{a.s.}} 0$, *uniformly in all continuous* F *and* G *and all* $\lambda_N \varepsilon [0, 1]$.

PROOF. Since $H_N - H = \lambda_N(F_m - F) + (1 - \lambda_N)(G_n - G)$, the Glivenko-Cantelli theorem implies that $\rho(H_N, H) \rightarrow_{\text{a.s.}} 0$ uniformly in F and G. Also $\rho(HH_N^{-1}, HH^{-1}) \leqq \rho(HH_N^{-1}, H_NH_N^{-1}) + \rho(H_NH_N^{-1}, HH^{-1})$ where $\rho(H_NH_N^{-1}, HH^{-1}) \leqq 1/N$. (This result does not use (2.2).) □

LEMMA 2.4. *If* $q \varepsilon \mathbf{Q}$, *then as* $N \rightarrow \infty$,

$$\rho(U_m(K_N)/q(K_N), U_0(K)/q(K)) \rightarrow_p 0$$

uniformly in all continuous F *and* G *and all* $\lambda_N \varepsilon [0, 1]$.

PROOF. Since $H = \lambda_N F + (1 - \lambda_N)G$, Lemma 2.3 implies that $\rho(K_N, K) \rightarrow_{\text{a.s.}} 0$ and $\rho(GH_N^{-1}, GH^{-1}) \rightarrow_{\text{a.s.}} 0$ uniformly; note that $K_N - K$ and $GH_N^{-1} - GH^{-1}$ are always of the same sign. Therefore, since U_0 is continuous a.s., $\rho(U_0(K_N)/q(K_N), U_0(K)/q(K)) \rightarrow_p 0$, the argument is like that of Theorem 2.1 where a.s. continuity of U_0 is needed to handle the term from the middle of the interval. Thus this lemma follows from Theorem 2.1 and the triangle inequality. □

LEMMA 2.5[2]. *For* $\epsilon > 0$, *there exists* $b > 0$ *such that*

$$P[K_N(t) \leqq 2b\lambda_*^{-1}t \quad \text{for all} \quad t \geqq 1/N] \geqq 1 - \epsilon. \tag{2.10}$$

PROOF. It is well known (cf. [9], Lemma 8) that for every $\epsilon > 0$, there exists $b > 0$ such that $P(A_m) > 1 - \epsilon$ where

$$A_m = [F(t) \leqq bF_m(t) \quad \text{for all} \quad t \quad \text{where } F_m(t) > 0]. \tag{2.11}$$

Therefore, it follows that if $t \geqq 1/N$ and $0 < F_m H_N^{-1}(t)$, then on A_m

$$K_N(t) = FH_N^{-1}(t) \leqq bF_m H_N^{-1}(t) \leqq b\lambda_*^{-1}H_N H_N^{-1}(t) \leqq 2b\lambda_*^{-1}t.$$

On the other hand if $t \geqq 1/N$ but $0 = F_m H_N^{-1}(t) < 1/m$, then on A_m

$$K_N(t) = FH_N^{-1}(t) \leqq FF_m^{-1}(1/m) \leqq bm^{-1} \leqq 2b\lambda_*^{-1}t.$$

Thus on A_m we have $K_N(t) \leqq 2b\lambda_*^{-1}t$ for all $t \geqq 1/N$. □

Let $U_m^*(K_N)$ equal $U_m(K_N)$ for $1/N \leqq t \leqq 1 - 1/N$ and equal 0 otherwise.

THEOREM 2.2. *For* $q \in \mathbf{Q}$

$$\rho_q(U_m^*(K_N), U_0(K)) \to_p 0$$

uniformly in all continuous F *and* G *and all* $\lambda_N \in \Lambda$.

PROOF. As in the proof of Theorem 2.1 the problem reduces to a study of the supremum over intervals $[0, \delta]$ and $[1 - \delta, 1]$ for sufficiently small δ. For over the interior interval $(\delta, 1 - \delta)$, $q(K)$ and, with high probability, $q(K_N)$ are bounded away from zero so that the supremum over $(\delta, 1 - \delta)$ converges in probability to zero by Lemma 2.4. Furthermore, it suffices to consider only the interval $[0, \delta]$ since the interval $[1 - \delta, 1]$ may be treated similarly by considering the reversed process $U_m^-(t) = U_m((1 - t)-)$. Also without loss of generality assume q is non-decreasing.

For given $\epsilon, \eta > 0$, choose b by Lemma 2.5 to satisfy (2.10). Then use Lemma 2.2 to choose $\alpha > 0$ to satisfy

$$P[|U_m(t)| \leqq \eta q(\lambda_* t/2b), \quad 0 < t < \alpha] > 1 - \epsilon;$$

and then choose $\delta > 0$ so that

$$P[K_N(t) < \alpha, 0 < t \leqq \delta] > 1 - \epsilon$$

for all sufficiently large N. This choice of δ is possible by Lemma 2.3 since $K_N(t) \leqq \lambda_*^{-1}HH_N^{-1}(\delta)$ for $t \leqq \delta$. With these choices of b, α and δ one obtains

$$\begin{aligned} P[\sup_{0 \leqq t \leqq \delta} |U_m^*(K_N(t))|/q(t) \leqq \eta] \\ \geqq P([|U_m(t)| \leqq \eta q(\lambda_* t/2b), \quad 0 < t < \alpha] \cap [K_N(t) < \alpha, 0 < t \leqq \delta] \\ \cap [K_N(t) \leqq 2b\lambda_*^{-1}(t), t \geqq 1/N]) \end{aligned}$$

for N sufficiently large. Since η and ϵ are arbitrary the proof is complete. □

---

[2] The authors are indebted to the referee for Lemma 2.5. The linear bounds on K_N provided by this result allow stronger statements to be made about the behavior of empirical processes in the tails than can be obtained by iterating Lemma 2.2. The main results of this paper are now able to be stated for $q \in Q$ rather than for the slightly more restrictive class of functions satisfying

$$\int_0^1 [t(1 - t)]^{-\frac{1}{2}}[q(t)]^{-1}\, dt < \infty.$$

3. The 2-sample empirical process. Recall that by Lemma 2.1

(3.1) $\rho(U_m, U_0) \to_{\text{a.s.}} 0, \quad \rho(V_n, V_0) \to_{\text{a.s.}} 0 \quad \text{as} \quad m, n \to \infty.$

Our study and application of the 2-sample empirical process $\{L_N(t): 0 \leqq t \leqq 1\}$ defined in (1.5) depends on the following identity which relates it to the 1-sample empirical U_m- and V_n-processes.

LEMMA 3.1. *With probability* 1,

(3.2) $$L_N(t) = (1 - \lambda_N)\{\lambda_N^{-\frac{1}{2}} B_N(t) U_m(FH_N^{-1}(t)) - (1 - \lambda_N)^{-\frac{1}{2}} A_N(t) V_n(GH_N^{-1}(t)\} + \delta_N(t)$$

for all $t \varepsilon (0, 1]$ *where*

(3.3) $$\delta_N(t) = A_N(t) N^{\frac{1}{2}}[H_N H_N^{-1}(t) - t],$$

(3.4) $$A_N(t) = [K(u_t) - K(t)]/(u_t - t), \qquad u_t = HH_N^{-1}(t),$$

and B_N *is defined by*

(3.5) $$\lambda_N A_N(t) + (1 - \lambda_N) B_N(t) = 1.$$

(In (3.2) L_N *is defined by left continuity at any otherwise undefined points.)*

PROOF. The relationship (3.2) follows immediately from (1.3), (1.4) and (1.5) by writing

$$L_N = N^{\frac{1}{2}}[F_m H_N^{-1} - FH^{-1}] = N^{\frac{1}{2}}[m^{-\frac{1}{2}} U_m(FH_N^{-1}) + FH_N^{-1} - FH^{-1}]$$
$$= N^{\frac{1}{2}}[m^{-\frac{1}{2}} U_m(FH_N^{-1}) + A_N(HH_N^{-1} - HH^{-1})]$$

provided $HH_N^{-1} \neq HH^{-1}$. Since with probability one this proviso is satisfied except at a finite number of points, and since

$$N^{\frac{1}{2}}[HH_N^{-1} - HH^{-1}]$$
$$= N^{\frac{1}{2}}[HH_N^{-1} - H_N H_N^{-1} + H_N H_N^{-1} - HH^{-1}]$$
$$= -\lambda_N^{\frac{1}{2}} U_m(FH_N^{-1}) - (1 - \lambda_N)^{\frac{1}{2}} V_n(GH_N^{-1}) + N^{\frac{1}{2}}[H_N H_N^{-1} - HH^{-1}]$$

the desired result follows. If $HH_N^{-1} = HH^{-1}$, it follows from the first line of the proof that the proper expression for L_N is $L_N = \lambda_N^{-\frac{1}{2}} U_m(FH^{-1})$. □

In the representation (3.2), one has for all $t \varepsilon [0, 1]$ that

(3.6) $$|A_N(t)| \leqq \lambda_N^{-1}, \qquad |B_N(t)| \leqq (1 - \lambda_N)^{-1}, \qquad |\delta_N(t)| \leqq \lambda_N^{-1} N^{-\frac{1}{2}}.$$

The first two inequalities follow from (3.5) since $A_N(t)$ and $B_N(t)$ are of the same sign, whereas the third inequality derives from (3.3) and the fact that $0 \leqq H_N H_N^{-1}(t) - t \leqq N^{-1}$.

From the representation (3.2) of the L_N-process, it is easy to see what must be the natural limiting process. We define first of all the $\bar{L}_N$-process by

(3.7) $$\bar{L}_N(t) = (1 - \lambda_N)\{\lambda_N^{-\frac{1}{2}} B_N(t) U_0(FH^{-1}(t)) - (1 - \lambda_N)^{-\frac{1}{2}} A_N(t) V_0(GH^{-1}(t))\}$$

for $0 \leqq t \leqq 1$. In view of Lemma 2.4 (for $q \equiv 1$) and its analogue for the V_n-process, the following limiting result is immediate.

LEMMA 3.2. *As* $N \to \infty$, $\rho(L_N, \bar{L}_N) \to_p 0$.

To obtain the natural limiting process from (3.7) it remains to study the limiting behavior of B_N and A_N. By their definition, A_N and B_N are difference quotients of FH^{-1} and GH^{-1} respectively. In view of (1.4), $\lambda_N FH^{-1}(t) + (1 - \lambda_N)GH^{-1}(t) = t$, so that both FH^{-1} and GH^{-1} are absolutely continuous. Let a_N and b_N denote the derivatives of FH^{-1} and GH^{-1} respectively, which therefore exist a.e. on [0, 1] with respect to Lebesgue measure. In some results, we will let λ_N converge to a fixed λ_0. Let a_0, b_0 denote the derivatives of FH_0^{-1}, GH_0^{-1} where $H_0 = \lambda_0 F + (1 - \lambda_0)G$. Now, wherever defined, set

$$(3.8) \quad L_0(t) = (1 - \lambda_0)^{\frac{1}{2}}\{\lambda_0^{-\frac{1}{2}} b_0(t) U_0(FH_0^{-1}(t)) - (1 - \lambda_0)^{-\frac{1}{2}} a_0(t) V_0(GH_0^{-1}(t))\}$$

and

$$(3.9) \quad L_{0N}(t) = (1 - \lambda_N)^{\frac{1}{2}}\{\lambda_N^{-\frac{1}{2}} b_N(t) U_0(FH^{-1}(t)) - (1 - \lambda_N)^{-\frac{1}{2}} a_N(t) V_0(GH^{-1}(t))\}.$$

These expressions respectively determine the natural limiting processes for L_N when λ_N does or does not converge. In the following sections the convergence of L_N to these limits will be considered.

4. Convergence in the ρ_q and d_q metrics. We now consider the convergence of the L_N-process to the limiting L_0-process when $\lambda_N \to \lambda_0$. Let $q \varepsilon \mathbf{Q}$, let δ_N^* be defined to equal δ_N on $[1/N, 1 - 1/N]$ and 0 elsewhere, and let L_N' equal L_N on $[1/N, 1]$ and 0 elsewhere. By (3.6), $\rho_q(\delta_N^*, 0) = o(1)$. Also

$$\sup_{1-1/N \leqq t < 1} |L_N(t)/q(t)| = o(1)$$

since $|L_N(t)| = N^{\frac{1}{2}} |1 - FH^{-1}(t)| \leqq N^{\frac{1}{2}} \lambda_*^{-1}(1 - t)$ in this interval. Thus, in order to prove that

$$(4.1) \quad \rho_q(L_N', L_0) \to_p 0$$

it will, by (3.2) and the triangle inequality, suffice to prove that

$$\rho(B_N, 0)[\rho_q(U_m^*(K_N), U_0(K)) + \rho_q(U_0(K), U_0(K_0))] + \rho(B_N, b_0)\rho_q(U_0(K_0), 0),$$

and the analogous quantity in A_N, V_n and G, converge in probability to 0. Apply (3.6) to $\rho(B_N, 0)$, Theorem 2.2 to $\rho_q(U_m^*(K_N), U_0(K))$, the proof of Lemma 2.4 to $\rho_q(U_0(K), U_0(K_0))$ and Lemma 2.2 to $\rho_q(U_0(K_0), 0)$ to obtain that (4.1) holds provided

$$(4.2) \quad \rho(B_N, b_0) \to_p 0.$$

We will now make a differentiability assumption on the functions $\{K_\lambda : \lambda \varepsilon \Lambda\}$ that will imply (4.2). This assumption is satisfied in many important cases, as Corollary 4.1 indicates.

ASSUMPTION 4.1. The functions K_λ have derivatives a_λ for all $t \varepsilon (0, 1)$, and for some λ', $a_{\lambda'}$ is continuous on $(0, 1)$ and has one-sided limits at 0 and 1.

It is an easy consequence of (1.4) and Assumption 4.1 that the functions GH_λ^{-1} have derivatives b_λ satisfying $\lambda a_\lambda + (1 - \lambda)b_\lambda = 1$ and that $b_{\lambda'}$ is continuous on $(0, 1)$ and has one-sided limits at 0 and 1.

DEFINITION 4.1. A family of functions $\{h_\lambda : \lambda \varepsilon \Lambda\}$ is said to be *uniformly equicontinuous* (uec) if for all $\epsilon > 0$ there exists $\delta_\epsilon > 0$ such that $|h_\lambda(u) - h_\lambda(v)| < \epsilon$ for all $\lambda \varepsilon \Lambda$ and for all u, v in the domain of the functions with $|u - v| < \delta_\epsilon$.

LEMMA 4.1. *Under Assumption* 4.1 $\{a_\lambda : \lambda \varepsilon \Lambda\}$ *and* $\{b_\lambda : \lambda \varepsilon \Lambda\}$ *are two* uec *families of functions on* $[0, 1]$ *and* $a_\lambda(b_\lambda)$ *converges uniformly to* $a_0(b_0)$ *as* $\lambda \to \lambda_0$.

PROOF. Differentiate $K_\lambda = K_{\lambda'}(H_{\lambda'}H_\lambda^{-1})$ to obtain

$$(4.3) \qquad a_\lambda = a_{\lambda'}(H_{\lambda'}H_\lambda^{-1})[(\lambda' - \lambda)a_\lambda + (1 - \lambda')](1 - \lambda)^{-1}.$$

If we set $\alpha_\lambda = a_{\lambda'}(H_{\lambda'}H_\lambda^{-1})$ and let ϕ_λ denote the remaining factors in (4.3) so that $a_\lambda = \alpha_\lambda\phi_\lambda$, then

$$[a_\lambda(u) - a_\lambda(v)][1 - \alpha_\lambda(v)(\lambda' - \lambda)(1 - \lambda)^{-1}] = [\alpha_\lambda(u) - \alpha_\lambda(v)]\phi_\lambda(u).$$

But for $\lambda \varepsilon \Lambda$, then ϕ_λ's are uniformly bounded since $|a_\lambda| \leqq 1$ and $\alpha_\lambda(\lambda' - \lambda)(1 - \lambda)^{-1}$ is uniformly bounded away from 1. Thus for $\lambda \varepsilon \Lambda$, $|a_\lambda(u) - a_\lambda(v)| \leqq C\,|\alpha_\lambda(u) - \alpha_\lambda(v)|$ for some constant C. To complete the proof that $\{a_\lambda : \lambda \varepsilon \Lambda\}$ is uec, apply to the definition of α_λ the fact that $\{H_{\lambda'}H_\lambda^{-1} : \lambda \varepsilon \Lambda\}$ is uec; since for $u \leqq v$ we have $0 \leqq FH_\lambda^{-1}(v) - FH_\lambda^{-1}(u) \leqq \lambda^{-1}(v - u) \leqq \lambda_*^{-1}(v - u)$. □

LEMMA 4.2. *Under Assumption* 4.1 $\rho(A_N, a_0) \to_{\text{a.s.}} 0$ *and* $\rho(B_N, b_0) \to_{\text{a.s.}} 0$ *as* $N \to \infty$.

PROOF. By the definition of A_N in (3.4), the mean value theorem and Lemma 4.1 we have that a_N is continuous and $A_N(t) = a_N(t_N)$ for some t_N between t and $HH_N^{-1}(t)$. Consequently, $\rho(A_N, a_N) = \sup_{0 \leqq t \leqq 1} |a_N(t_N) - a_N(t)|$. Since, by Lemma 2.3, $t_N \to t$ uniformly on $[0, 1]$ as $N \to \infty$, it follows from the uec of the a_N's that $\rho(A_N, a_N) \to_{\text{a.s.}} 0$. Also, since $\lambda_N \to \lambda_0$, $\rho(a_N, a_0) \to 0$ by Lemma 4.1. An application of the triangle inequality completes the proof. □

Due to the left continuity of the L_N-process we introduce D^- for the set of all left continuous functions on [0, 1].

THEOREM 4.1. (a) *Suppose Assumption* 4.1 *holds*, $\lambda_N \to \lambda_0$ *and* $q \varepsilon \mathbf{Q}$. *Then* $\rho_q(L_N', L_0) \to_p 0$ *so that* $L_N' \to_L L_0$ *relative to* (D^-, ρ_q). *The same statement holds for* d_q.

(b) *If in addition the measures* $\{\nu_N : N \geqq 1\}$ *and* ν *of Section* 1 *satisfy*

(i) $\int_{1/N}^1 L_N\, d(\nu_N - \nu) \to_p 0$, *and*,

(ii) $\int_0^1 q\, d|\nu| < \infty$,

then $T_N^* \to_p \int_0^1 L_0\, d\nu$, *a* $N(0, \sigma_0^2)$ rv *where*

$$(4.4) \quad \sigma_0^2 = 2(1 - \lambda_0)^2\{\lambda_0^{-1} \int_0^1 \int_0^v b_0(u)\, b_0(v) FH_0^{-1}(u)[1 - FH_0^{-1}(v)]\, d\nu(u)\, d\nu(v) + (1 - \lambda_0)^{-1} \int_0^1 \int_0^v a_0(u) a_0(v) GH_0^{-1}(u)[1 - GH_0^{-1}(v)]\, d\nu(u)\, d\nu(v)\}.$$

PROOF. (a) Lemma 4.2 shows that (4.1) holds under Assumption 4.1. Convergence in law follows easily by definition. Results for d_q follow by an analogous argument, since $d_q \leqq 4\rho_q$ as remarked in the proof of Theorem 2.1.

(b) Now

$$|T_N^* - \int_0^1 L_0\,d\nu| = |\int_0^1 L_N'\,d(\nu_N - \nu) + \int_0^1 (L_N' - L_0)\,d\nu|$$
$$\leqq |\int_0^1 L_N'\,d(\nu_N - \nu)| + \rho_q(L_N', L_0)\int_0^1 q\,d|\nu| \to_p 0.$$

The fact that the limiting rv has a normal distribution relies on the existence and finiteness of its variance, which in turn follows from (ii) since

$$E[\int_0^1 L_0\,d\nu]^2 \leqq c\int_0^1\int_0^v u(1-v)\,d|\nu|(u)\,d|\nu|(v)$$
$$\leqq c\int_0^1\int_0^v\{[u(1-u)v(1-v)]^{\frac{1}{2}}[q(u)q(v)]^{-1}\}q(u)q(v)\,d|\nu|(u)\,d|\nu|(v)$$

for some constant c. ▯

Since the L_N-process takes jumps only at points i/N and since the L_0-process has continuous paths a.s., the mappings L_N $(N \geqq 0)$ from $(\Omega, \mathfrak{A})$ to (D^-, ρ_q) are measurable; see Billingsley (1967), p. 308. Thus the weak convergence in Theorem 4.1 is valid in the sense of Prokhorov's usual definition.

It is necessary to write L_N' and not L_N in Theorem 4.1 since if the smallest observation is an X, then $L_N(0+) = N^{\frac{1}{2}}m^{-1}$.

REMARK. If λ_N does not converge, then in Theorem 4.1 we may write $\rho_q(L_N', L_{0N}) \to_p 0$ and $T_N^* - \int_0^1 L_{0N}\,d\nu \to_p 0$ where $\int_0^1 L_{0N}\,d\nu$ is $N(0, \sigma_{0N}^2)$ with σ_{0N}^2 given by the obvious modification of (4.4). In this form, the theorem relates most closely to the main theorem of Chernoff and Savage (1958); σ_{0N}^2 is given as equation (4.3) in this reference.

COROLLARY 4.1. *The conclusion of Theorem* 4.1 *is valid if in place of Assumption* 4.1 *we have either* (i) $F = G$ *or* (ii) *there are numbers* $-\infty \leqq \alpha_1, \alpha_2, \gamma_1, \gamma_2 \leqq +\infty$ *for which* $F(\alpha_1) = G(\alpha_1) = 0$, $F(\alpha_2) = G(\alpha_2) = 1$, *the boundary of the set of common zeros of the continuous functions* $f = F'$ *and* $g = G'$ *is finite and* $\gamma_i = \lim_{x\to\alpha_i}[g(x)/f(x)]$ *exists for* $i = 1, 2$.

PROOF. (i). This follows since Assumption 4.1 is satisfied when $F = G$. (ii). Suppose first that the set of common zeros is empty. Since $a_\lambda(t) = (d/dt)FH_\lambda^{-1}(t) = (\lambda + (1-\lambda)[gH_\lambda^{-1}(t)/fH_\lambda^{-1}(t)])^{-1}$, the existence of γ_1 and γ_2 implies in this case that Assumption 4.1 holds.

In the general case, consider the finite set of points t for which $H_{\lambda_0}^{-1}(t)$ is in the boundary of the set of common zeros; these points are interior to $(0, 1)$. Hence, we may throw away intervals of arbitrarily small $|\nu|$-measure that contain these points. On the complement of these intervals Assumption 4.1, with Λ replaced by a sufficiently small interval about λ_0, holds. Now apply the proof of Lemma 4.2 to each interval in the complement. ▯

5. Convergence in other metric spaces. A study of linear rank statistics leads naturally to a study of integral functionals on stochastic processes; see (1.9). It thus seems natural to consider integral metrics on these processes. Indeed, this will allow us to obtain results under weaker hypotheses than previously assumed. This follows since the B_N's may converge in an integral norm when they fail to converge in a supremum norm.

Whenever well defined let

$$\|f\|_\nu = \int_0^1 |f(t)|\, d\,|\nu|\,(t). \tag{5.1}$$

We shall drop the subscript and write $\|\cdot\|$ in case ν is Lebesgue measure on $(0, 1)$.

LEMMA 5.1. *If* $\int_0^1 q\, d|\nu| < \infty$ *for some* $q \in \mathbf{Q}$ *then*

$$\|U_m^*(K_N) - U_0(K)\|_\nu \to_p 0$$

uniformly in all continuous F *and* G *and all* $\lambda_N \in \Lambda$.

PROOF. This is an immediate consequence of Theorem 2.2 since $\|f\|_\nu \leqq \rho_q(f, 0) \int_0^1 q\, d|\nu|$. □

As in Section 4, one may show that $\|\delta_N^*\|_\nu = o(1)$, and $|\int_{1-1/N}^1 L_N\, d\nu| = o(1)$. Thus by analogy with equations (4.1) and (4.2) we have that in order to establish

$$\|L_N' - L_0\|_\nu \to_p 0 \qquad (\|L_N' - L_{0N}\|_\nu \to_p 0) \tag{5.2}$$

it suffices to find conditions under which

$$\|(B_N - b_0)q\|_\nu \to_p 0 \qquad (\|(B_N - b_N)q\|_\nu \to_p 0) \tag{5.3}$$

for some $q \in \mathbf{Q}$.

Let $C_0 \subset (0, 1)$ be the set on which $K_0 = FH_0^{-1}$ is differentiable. By the remark at the end of Section 3, the Lebesgue measure of C_0 is 1.

LEMMA 5.2. *If* $\lambda_N = \lambda_0 + O(N^{-\frac{1}{2}})$ *and* μ *is a Lebesgue-Stieltjes measure on* $(0, 1)$ *for which* $\mu(C_0^c) = 0$, *then* $A_N \to a_0$ *and* $B_N \to b_0$ *in* $P \times \mu$*-measure as* $N \to \infty$.

PROOF. By (3.5), it suffices to prove the result about A_N. From (1.3) and (3.4) one obtains that $K = FH^{-1} = FH_0^{-1}(H_0H^{-1}) = K_0(H_0H^{-1})$ and that $A_N = A_{N1}A_{N2}$ where

$$A_{N1} = [K_0(H_0H_N^{-1}) - K_0(H_0H^{-1})][H_0H_N^{-1} - H_0H^{-1}]^{-1},$$

$$A_{N2} = [H_0H_N^{-1} - H_0H^{-1}][HH_N^{-1} - HH^{-1}]^{-1}.$$

By (1.4), $H_0 = \lambda_0 F + (1 - \lambda_0)G = H + (\lambda_0 - \lambda_N)(F - G)$, so that

$$H_0H_N^{-1}(t) - H_0H^{-1}(t) = HH_N^{-1}(t) - t + (\lambda_0 - \lambda_N)D_N(t)$$

where $D_N = FH_N^{-1} - FH^{-1} - GH_N^{-1} + GH^{-1}$. Therefore,

$$A_{N2} = 1 + (\lambda_0 - \lambda_N)D_N(d_N - N^{-\frac{1}{2}}W_N)^{-1} \tag{5.4}$$

where $0 \leqq d_N(t) = H_NH_N^{-1}(t) - t \leqq 1/N$ and $W_N = N^{\frac{1}{2}}(H_NH_N^{-1} - HH_N^{-1})$. By definitions (1.1) and (1.2),

$$W_N = \lambda_N^{\frac{1}{2}}U_m(FH_N^{-1}) + (1 - \lambda_N)^{\frac{1}{2}}V_n(GH_N^{-1}).$$

It therefore follows by arguments similar to those used to prove Lemma 2.3 and 2.4 that $\rho(W_N, W_0) \to_p 0$ where $W_0 = \lambda_0^{\frac{1}{2}}U_0(FH_0^{-1}) + (1 - \lambda_0)^{\frac{1}{2}}V_0(GH_0^{-1})$. By Lemma 2.3, $\rho(D_N, 0) \to_{\text{a.s.}} 0$. Let $\{N_k : k \geqq 1\}$ be a subsequence for which

$\rho(W_{N_k}, W_0) \to_{\text{a.s.}} 0$. For each $\omega \in \Omega$ let $C_{0,\omega} = \{t \in [0, 1]: W_0(t) \neq 0\}$. By applying to (5.4) the convergences just mentioned and the hypothesis $\lambda_N - \lambda_0 = O(N^{-\frac{1}{2}})$, it follows that every subsequence has a subsequence, $\{N_k\}$ say, for which

$$(5.5) \qquad P\{w \in \Omega: A_{N_k 2})t) \to 1 \text{ for all } t \in C_{0,\omega}\} = 1.$$

To study the limit of A_{N1}, observe first that for any function p

$$\frac{p(x) - p(y)}{x - y} = \frac{p(x) - p(x_0)}{x - x_0}\,\frac{x - x_0}{x - y} + \frac{p(x_0) - p(y)}{x_0 - y}\,\frac{x_0 - y}{x - z},$$

Thus if p is differentiable at x_0 with derivative p', then $[p(x) - p(y)]/(x - y) \to p'(x_0)$ if $x, y \to x_0$ in such a way that $(x_0 - y)/(x - y)$ remains bounded. As $N \to \infty$ over a subsequence $\{N_k\}$ for which $\rho(W_{N_k}, W_0) \to_{\text{a.s.}} 0$, one obtains that for each $t \in C_{0,\omega}$ and almost all ω,

$$(H_0 H_0^{-1} - H_0 H^{-1})(H_0 H_N^{-1} - H_0 H^{-1})^{-1}$$
$$= (\lambda_N - \lambda_0)(FH^{-1} - GH^{-1})[d_N - N^{-\frac{1}{2}} W_N + (\lambda_0 - \lambda_N) D_N]^{-1}$$

remains bounded. Therefore

$$(5.6) \qquad P\{\omega \in \Omega: A_{N_k 1}(t) \to a_0(t) \text{ for all } t \in C_0 \cap C_{0,\omega}\} = 1.$$

Since $P[W_0(t) = 0] = 0$ for each $t \in (0, 1)$ and since the set of pairs (t, ω) for which $t \in C_{0,\omega}$ is measurable because of the almost sure continuity of the W_0-process, it follows from Fubini's theorem that $\mu(C_{0,\omega}^c) = 0$ for almost all ω. The proof is completed by combining (5.5) and (5.6). □

THEOREM 5.1. (a) *Suppose* K_0 *is differentiable* a.e. $|\nu|$, $\lambda_N = \lambda_0 + O(N^{-\frac{1}{2}})$. *and* $\int_0^1 q\, d|\nu| < \infty$ *for some* $q \in \mathbf{Q}$. *Then* $\|L_N' - L_0\|_\nu \to_p 0$ *so that* $L_N' \to_L L_0$ *relative to* $(D^-, \|\cdot\|_\nu)$.

(b) *Suppose in addition that the measures* $\{\nu_N : N \geqq 1\}$ *of Section* 1 *satisfy*

(i) $\int_{1/N}^1 L_N\, d(\nu_N - \nu) \to_p 0$.

Then $T_N \to_p \int_0^1 L_0\, d\nu$ *which is a* $N(0, \sigma_0^2)$ rv *with* σ_0^2 *given by* (4.4).

PROOF. (a) By Lemma 5.2 with $\mu = |\nu|$, every subsequence of B_N contains a further subsequence converging to b_0 a.e. $P \times |\nu|$. Thus by the dominated convergence theorem we obtain $\|(B_N - b_0)q\|_{|\nu|} \to_{\text{a.s.}} 0$. Convergence in law follows easily by definition.

(b) Now

$$(5.7) \qquad |T_N^* - \textstyle\int_0^1 L_0\, d\nu| \leqq |\int_0^1 L_N'\, d(\nu_N - \nu)| + |\int_0^1 (L_N' - L_0)\, d\nu|$$

which equals $o_p(1)$ by (i) and since we have established (5.3) for $|\nu|$. □

Note that Theorem 4.1(a) has a stronger conclusion than Theorem 5.1(a) since the ρ_q-topology on D^- contains the $\|\cdot\|_\nu$-topology on D^- when $\int_0^1 q\, d\,|\,\nu\,| < \infty$. However, the functional $\int_0^1 \cdot\, d\nu$ is $\|\cdot\|_\nu$-continuous a.s. on D^- with respect to the measure of the L_0-process, so that it was possible to obtain part (b) of Theorem 5.1 in analogy with Theorem 4.1(b).

We next discuss a relationship between Theorem 5.1(b) and Theorem 1 of Chernoff and Savage (1958). Define a function J_N on [0, 1] by letting $J_N(t) = c_{Ni}^*$ for $(i-1)/N < t \leqq i/N$ for $i = 1, \cdots, N$ and $J_N(0) = J_N(0+)$ so that $T_N = \int_0^1 J_N \, dF_m H_N^{-1}$ by (1.7). Let $-J$ denote a non-constant function of bounded variation on $(\epsilon, 1-\epsilon)$ for all $\epsilon > 0$ which induces the Lebesgue-Stieltjes measure ν. Chernoff and Savage consider the statistic $N^{\frac{1}{2}}[T_N - \int_{-\infty}^{\infty} J(H)\, dF] = T_N^* + \gamma_N$ where $\gamma_N = N^{\frac{1}{2}} \int_0^1 [J_N(H) - J(H)]\, dF$; this equality is obtained by summing μ_N by parts.

Proposition 5.1. *Suppose*

(1) $N^{\frac{1}{2}} \int_0^1 |J_N(t) - J_N(t-)| \, dH_N H_N^{-1}(t) = o(1)$,

(2) $J_N(0) = o(N^{\frac{1}{2}})$ *and* $J_N(1) = o(N^{\frac{1}{2}})$,

(3) $|J(t)| \leqq K[t(1-t)]^{-\frac{1}{2}+\delta}$ *for some constants* $K, \delta > 0$ *and*

(4) $N^{\frac{1}{2}} \int_0^1 |J_N(t) - J(t)| \, dt = o(1)$.

Then

(a) $\gamma_N = o(1)$,

(b) $\int_{1/N}^1 L_N \, d(\nu_N - \nu) = o_p(1)$ *uniformly in all continuous* F *and* G *and all* $\lambda_N \varepsilon \Lambda$, *and*

(c) *there exists* $q \varepsilon \mathbf{Q}$ *such that* $\int_0^1 q \, d|\nu| < \infty$.

Proof. (a) $\gamma_N \leqq \lambda_*^{-1} N^{\frac{1}{2}} \int_0^1 |J_N(H) - J(H)| \, dH = o(1)$ by (1).

(b) $I_N \equiv \int_{1/N}^1 L_N \, d(\nu_N - \nu) = -\int_{1/N}^{1-1/N} L_N \, d(J_N - J) + o(1)$ by (3) and the bound on L_N in $[1 - 1/N, 1]$. Thus $I_N = -\int_{1/N}^{1-1/N} L_N(t+) \, d(J_N(t) - J(t)) + o(1)$ using the integration by parts formula on page 419 of Hewitt and Stromberg in conjunction with (2), (3), (4) and the fact that $|L_N(t) - L_N(t+)| = O(N^{\frac{1}{2}})$ for all t. A second integration by parts used in conjunction with (2), (3) and the bound on L_N at $1/N$ and $1 - 1/N$ shows that

$$\begin{aligned} I_N &= \int_{1/N}^{1-1/N} [J_N(t) - J(t-)] \, dL_N(t) + o(1) \\ &\leqq \lambda_*^{-1} N^{\frac{1}{2}} [\int_0^1 |J_N(t) - J(t-)| \, dH_N H_N^{-1}(t) + \int_0^1 |J_N(t) - J(t)| \, dt] \\ &= o(1) \end{aligned}$$

where the last equality follows from (1) and (4).

(c) Let $q(t) = [t(1-t)]^{\frac{1}{2}-\delta/2}$ and use integration by parts. □

Corollary 5.1. *Proposition* 5.1 *remains true if* (4) *is replaced by*

(4′) $J = J_d + J_c$ *where* J_d *is a saltus function taking only a finite number of jumps and where* J_c *has a continuous derivative* J_c' *on intervals* $(0, a_1), (a_1, a_2), \cdots, (a_s, 1)$ *which satisfies*

$$|J_c'(t)| \leqq K[t(1-t)]^{-\frac{3}{2}+\delta} \quad \textit{for} \quad t \neq \textit{any } a_i.$$

Proof. We give the proof for $s = 0$. Let $t_N = (1 + [Nt])/N$ where $[\cdot]$ is the greatest integer function. Now by (2) and (3)

$$\begin{aligned} I_N \equiv N^{\frac{1}{2}} \int_0^1 |J_N - J| \, dt &= N^{\frac{1}{2}} \int_{2/N}^{1-2/N} |J_N - J| \, dt + o(1) \\ &\leqq N^{\frac{1}{2}} \int_0^1 |J_N(t) - J(t_N)| \, dt + N^{\frac{1}{2}} \int_0^1 |J_d(t_N) - J_d(t)| \, dt \\ &\qquad + N^{\frac{1}{2}} \int_{2/N}^{1-2/N} |J_c(t_N) - J_c(t)| \, dt + o(1) \\ &= N^{\frac{1}{2}} \int_{2/N}^{1-2/N} |J_c(t_N) - J_c(t)| \, dt + o(1) \end{aligned}$$

where the last equality follows from (1) and the (4′) condition on J_d. But

$$N^{\frac{1}{2}} \int_{\frac{1}{2}}^{1-2/N} |J_c(t_N) - J_c(t)|\, dt \leqq N^{\frac{1}{2}} \int_{\frac{1}{2}}^{1-1/N} N^{-1} K[t(1-t)]^{-\frac{3}{2}+\delta}\, dt = o(1)$$

since (4′) makes possible an application of the mean value theorem to show $|J_c(t_N) - J_c(t)| \leqq N^{-1}K[(t + 1/N)(1 - t - 1/N)]^{-\frac{3}{2}+\delta}$ when $t \leqq t_N \leqq t + 1/N$; and where the factor $t + 1/N$ accounts for the change in the limits of integration. The integral from $2/N$ to $\frac{1}{2}$ is similar. Thus $I_N = o(1)$, so that (4) holds. If $s \neq 0$, then separate off sufficiently small intervals about each a_i and repeat the above proof. □

Theorem 2 of Chernoff and Savage shows that (1), (2), (3) and (4′) hold when $J = \Psi^{-1}$ where Ψ is a df satisfying

$$|(d^{(i)}/dt^{(i)})\Psi^{-1}(t)| \leqq K[t(1-t)]^{-i-\frac{1}{2}+\delta} \quad \text{for} \quad i = 0, 1, 2$$

and where c^*_{Ni} is the expectation of the ith smallest order statistic in a sample of size N from a population having df Ψ. This shows that Theorem 5.1(b) is broadly applicable. Statements that condition (4) in Theorem 1 of Chernoff and Savage, which puts bounds on J, J' and J'', need hold only a.e. appear frequently in the literature; though the validity of that statement is suspect (see Govindarajulu, et al. (1967)). Proposition 5.1 and its corollary are weak enough to cover all the cases where Chernoff and Savages' (4) failed at a finite number of points; and this is weak enough to cover all cases of known interest.

Remark. Let $q \varepsilon \mathbf{Q}$ be fixed. Consider a class $\mathbf{J}$ of sequences $(J, J_1, J_2, \cdots)$ for which $\int_0^1 q\, d|J|$ is uniformly bounded and for which the hypotheses of Proposition 5.1 are uniform in $\mathbf{J}$. Then the conclusion of Theorem 5.1(b) is uniform in $\mathbf{J}$. To see this, note that the term $\|L_N' - L_0\|_\nu$ in (5.7) is $o_p(1)$ uniformly in $\mathbf{J}$ by the uniform boundedness of $\int_0^1 q\, d|J|$ in $\mathbf{J}$. The term $|\int_{1/N}^1 L_N\, d(\nu_N - \nu)|$ in (5.7) is $o_p(1)$ uniformly in $\mathbf{J}$ by the proof of Proposition 5.1.

6. The c-sample problem. Let $X_{j1}, \cdots, X_{jn_j}$, $j = 1, \cdots, c$, be independent random samples from populations having continuous df's F_j. Let $N = n_1 + \cdots + n_c$; when used as a subscript N will denote the c-tuple $(n_1, \cdots n_c)$. Let $\boldsymbol{\lambda}_N = (\lambda_{N1}, \cdots, \lambda_{Nc})$ where $\lambda_{Nj} = n_j/N$. Suppose $\lim_{N\to\infty} \lambda_{Nj} = \lambda_{0j}$ exists and $\lambda_* < \lambda_{jN} < 1 - \lambda_*$ for some $\lambda_* > 0$. Let $\boldsymbol{\lambda}_0 = (\lambda_{01}, \cdots, \lambda_{0c})$. Let F_{Nj} denote the empirical df of the jth sample. Then $H_N = \sum_{j=1}^c \lambda_{Nj} F_{Nj}$ is the empirical df of the combined sample. Let $H_0 = \sum_{j=1}^c \lambda_{0j} F_j$, $H_{\boldsymbol{\lambda}} = \sum_{j=1}^c \lambda_j F_j$ where $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_c)$ and $H = H_{\boldsymbol{\lambda}_N}$. For $j = 1, \cdots, c$ define stochastic processes $\{L_{Nj}(t) : 0 \leqq t \leqq 1\}$ and $\{L_{0j}(t) : 0 \leqq t \leqq 1\}$ by

$$L_{Nj}(t) = N^{\frac{1}{2}}[F_{Nj} H_N^{-1}(t) - F_j H^{-1}(t)] \tag{6.1}$$

and

$$L_{0j}(t) = -a_{0j}(t) \textstyle\sum_{i\neq j} \lambda_{0i}^{\frac{1}{2}} U_{0i}(F_i H_0^{-1}(t)) + [\sum_{i\neq j} (\lambda_{0i}/\lambda_{0j}^{\frac{1}{2}}) a_{0i}(t)] U_{0j}(F_j H_0^{-1}(t)) \tag{6.2}$$

where the processes $\{U_{0j}(t) : 0 \leqq t \leqq 1\}$, $j = 1, \cdots, c$, are independent tied-down Wiener processes.

We consider the c-sample analog of Theorem 5.1(b). Let $Z_{Ni}^{(j)}$ equal 1 or 0 depending on whether or not the ith smallest observation from the combined sample of size N is from the jth sample. Consider the linear rank statistics

$$(6.3) \qquad T_{Nj} = n_j^{-1} \sum_{i=1}^{N} C_{Ni} R_{Ni}^{(j)} = n_j^{-1} \sum_{i=1}^{N} C_{Ni}^{*} Z_{Ni}^{(j)}$$

where $R_{Ni}^{(j)} = Z_{N1}^{(j)} + \cdots + Z_{Ni}^{(j)}$, $\{C_{Ni} : 1 \leqq i \leqq N\}$ is a given set of constants and $C_{Ni}^{*} = C_{Ni} + \cdots + C_{NN}$. Let ν_N denote the signed measure which puts measure C_{Ni} on the point i/N for $1 \leqq i \leqq N$ and puts zero measure elsewhere. Then, for $j = 1, \cdots, c$,

$$(6.4) \qquad T_{Nj} = \int_0^1 F_{Nj} H_{N.}^{-1} \, d\nu_N .$$

For $j = 1, \cdots, c$ let $\mu_{Nj} = \int_0^1 F_j H^{-1} \, d\nu_N$ and set $T_{Nj}^{*} = N^{\frac{1}{2}}(T_{Nj} - \mu_{Nj})$. Let ν be a signed Lebesgue-Stieltjes measure on $(0, 1)$ for which $|\nu|([\epsilon, 1 - \epsilon]) < \infty$ for all $\epsilon > 0$.

THEOREM 6.1. *Suppose* $\int_0^1 q \, d|\nu| < \infty$ *for some* $q \in \mathbf{Q}$, *and for* $j = 1, \cdots, c$ *we have* $\lambda_{Nj} = \lambda_{0j} + O(N^{-\frac{1}{2}})$ *and* $K_{0j} = F_j H_0^{-1}$ *is differentiable* a.e. ν. *If*

(i) $N^{\frac{1}{2}} \int_{1/N}^1 L_{Nj} \, d(\nu_N - \nu) \to_p 0$ *for* $j = 1, \cdots, c$,

then $(T_{N1}^{*}, \cdots, T_{Nc}^{*}) \to (\int_0^1 L_{01} \, d\nu, \cdots, \int_0^1 L_{0c} \, d\nu)$, *a* $N(0, \Sigma_0)$ *random vector where* Σ_0 *is straightforwardly determined from* (6.2).

PROOF. Guided by (6.2) and (3.2) represent each L_{Nj} as the sum of c empirical processes (say U_{jn_j}) multiplied by random coefficients (say A_{Nj}) plus a negligible term (say δ_{Nj}). The proof is completed by applying the lemmas of Sections 2, 4, and 5 in a fashion analogous to their use in Theorem 5.1 to obtain that for each j, $T_{Nj} \to_p \int_0^1 L_{0j} \, d\nu$. □

Let D^c denote the c-fold product $D \times \cdots \times D$. On this space consider the metrics $d^c(f, g) = \sum_{j=1}^{c} d(f_j, g_j)$ and $\rho^c(f, g) = \sum_{j=1}^{c} \rho(f_j, g_j)$ for all $f = (f_1, \cdots, f_c)$, $g = (g_1, \cdots, g_c)$ in D^c. Clearly the analogue of Theorem 5.1(a) holds on D^c under the hypotheses of Theorem 6.1. Moreover, this result clearly holds for any metric $\rho^c(f, g)$ for which $\rho(f_{jN}, g_{jN}) \to 0$ as $N \to \infty$ for $j = 1, \cdots, c$ implies $\rho^c(f_N, g_N) \to 0$ as $N \to \infty$; and similarly for d^c.

It is also clear that c-sample results analogous to Theorem 4.1 hold.

REFERENCES

[1] BILLINGSLEY, PATRICK (1967). *Weak convergence of probability measures*. Lecture notes to be published by J. Wiley and Sons, New York.

[2] BIRNBAUM, Z. W. and MARSHALL, A. W. (1961). Some multivariate Chebyshev inequalities with extensions to continuous parameter processes. *Ann. Math. Statist.* **32** 687–703.

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

[4] CHIBISOV, D. M. (1964). Some theorems on the limiting behavior of empirical distribution functions. *Trudy Matem. Inst. im V. A. Steklova* **71** 104–112. (In Russian)

[5] CHIBISOV, D. M. (1965). An investigation of the asymptotic power of tests of fit. *Theor. Prob. and Appl.* (Translated by *SIAM*) **10** 421–437.

[6] DOOB, J. L. (1949). Heuristic approach to the Kolmogorov-Smirnov limit theorems. *Ann. Math. Statist.* **20** 393–403.

[7] DUDLEY, R. M. (1966). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.* **10** 109–126.
[8] GNEDENKO, B. V. (1962). *The Theory of Probability*. Chelsea, New York.
[9] GOVINDARAJULU, Z., LECAM, L. and RAGHAVACHARI, M. (1967). Generalizations of theorems of Chernoff-Savage on asymptotic normality of nonparametric test statistics. *Proc. of Fifth Berkeley Symp. on Math. Statisti. and Prob.* 609–638. Univ. of California Press.
[10] HEWITT, EDWIN and STROMBERG, KARL. (1965). *Real and Abstract Analysis*. Springer-Verlag, New York.
[11] LÉVY, PAUL. (1948). *Processus Stochastiques et Mouvement Brownien*. Gauthier-Villars, Paris.
[12] PROKHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theor. Prob. and Appl.* (translated by *SIAM*) **1** 157–224.
[13] PYKE, RONALD and SHORACK, GALEN R. (1967). Weak convergence and a Chernoff-Savage theorem for random sample sizes. Technical Report No. 14 Math. Dept., University of Washington.
[14] SHORACK, GALEN R. (1967). On asymptotic efficiency. Unpublished University of Washington Technical Report.
[15] SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes. *Theor. Prob. and Appl.* (translated by *SIAM*) **1** 261–290.