The Annals of Mathematical Statistics
1968, Vol. 39, No. 5, 1675–1685

WEAK CONVERGENCE AND A CHERNOFF-SAVAGE THEOREM FOR RANDOM SAMPLE SIZES

BY RONALD PYKE AND GALEN R. SHORACK

University of Washington

1. Summary. A Chernoff and Savage theorem on the asymptotic normality of 2-sample linear rank statistics is here established for random sample sizes. The proof parallels that of Pyke and Shorack (1968), hereafter referred to as PS. A mild restriction on the underlying distributions is needed in the present situation. A result of Pyke (1968) on the weak convergence of the 1-sample empirical process for random sample sizes in the ordinary uniform metric is here extended to other metrics. This extension provides an essential step in the present proof and is also of separate interest. The results extend immediately to c-samples.

2. Introduction and notation. Let $\{X_i : i \geqq 1\}$ and $\{Y_j : j \geqq 1\}$ be independent sequences of independent rv's where the X_i's and Y_j's have continuous df.'s $\underline{F}$ and $\underline{G}$ respectively. Let $\{m_t : t > 0\}$ and $\{n_t : t > 0\}$ be positive integer valued stochastic processes satisfying

$$t^{-1}m_t \rightarrow_p \alpha \quad \text{and} \quad t^{-1}n_t \rightarrow_p \beta \quad \text{as} \quad t \rightarrow \infty \tag{2.1}$$

where $0 < \alpha, \beta < \infty$. Without loss of generality all random elements are defined on the same probability space $(\Omega, \mathcal{A}, P)$. Let

$$N_t = m_t + n_t, \qquad \lambda_t = m_t/N_t, \qquad \lambda_0 = \alpha/(\alpha + \beta). \tag{2.2}$$

Let $\Lambda = [\lambda_*, 1 - \lambda_*]$ where $0 < \lambda_* < \lambda_0 < 1 - \lambda_* < 1$ and let a_λ and b_λ denote the derivatives of $\underline{K}_\lambda = \underline{FH}_\lambda^{-1}$ and $\underline{GH}_\lambda^{-1}$ respectively where

$$\underline{H}_\lambda = \lambda\underline{F} + (1 - \lambda)\underline{G} \tag{2.3}$$

and where we write $\underline{FH}_\lambda^{-1}$ for the composed function $\underline{F}(\underline{H}_\lambda^{-1})$ with inverses defined to be left continuous.

For $t > 0$ let

$$\underline{H}_t = \underline{H}_{\lambda_t}, \qquad H_t = \lambda_t F_t + (1 - \lambda_t)G_t, \qquad \underline{H}_0 = \underline{H}_{\lambda_0}, \tag{2.4}$$

when F_t and G_t are the empirical df's of the first m_t X_i's and the first n_t Y_j's respectively. (Since the choice of subscripts will be consistent throughout, the reader should not be confused by the notation $\underline{H}_\lambda$ and $\underline{H}_t$.) For $t > 0$ let

$$\underline{K}_t = \underline{FH}_t^{-1}, \qquad K_t = \underline{F}H_t^{-1}, \qquad \underline{K}_0 = \underline{FH}_0^{-1}. \tag{2.5}$$

Define stochastic processes $\{L_t(u) : 0 \leqq u \leqq 1\}$ by

$$L_t(u) = N_t^{\frac{1}{2}}[F_t H_t^{-1}(u) - \underline{FH}_t^{-1}(u)], \qquad t > 0. \tag{2.6}$$

Received 19 November 1967.

Let $a_t(a_0)$ and $b_t(b_0)$ denote the derivatives of $\underline{K}_t(\underline{K}_0)$ and $\underline{GH}_t^{-1}(\underline{GH}_0^{-1})$ respectively; which exist a.s. with respect to Lebesgue measure and satisfy $\lambda_t a_t + (1 - \lambda_t)b_t = 1$, $(t \geqq 0)$.

For fixed integers m and n set $N = m + n$ and let

$$T_N = m^{-1}\sum_{1=1}^{N} c_{Ni}^{*} Z_{Ni} \tag{2.7}$$

where the c_{Ni}^{*}'s are arbitrary real constants and where Z_{Ni} equals 1 or 0 depending on whether the ith smallest among $X_1, \cdots, X_m$ and $Y_1, \cdots, Y_n$ is an X or a Y. Define constants c_{Ni} by $c_{Ni}^{*} = c_{Ni} + \cdots + c_{NN}$ and let $\{\nu_N : N \geqq 1\}$ denote a sequence of signed measures on [0, 1] such that ν_N puts measure c_{Ni} at i/N for $i = 1, \cdots, N$ and puts 0 measure elsewhere. For $t > 0$ let

$$T_t = T_{N_t} \quad \text{and} \quad \nu_t = \nu_{N_t}. \tag{2.8}$$

A summation by parts shows that

$$T_t = \sum_{i=1}^{N_t} c_{N_t i} F_t H_t^{-1}(i/N_t) = \int_0^1 F_t H_t^{-1}\, d\nu_t .$$

Define $\mu_t = \int_0^1 \underline{K}_t\, d\nu_t$ and set

$$T_t^{*} = N_t^{\frac{1}{2}}(T_t - \mu_t) = \int_0^1 L_t(u)\, du. \tag{2.9}$$

Let $\{U_0(u) : 0 \leqq u \leqq 1\}$ and $\{V_0(u) : 0 \leqq u \leqq 1\}$ be independent tied-down Wiener processes on [0, 1]; and where defined set

$$\begin{aligned} L_0(u) = (1 - \lambda_0)\{&\lambda_0^{-\frac{1}{2}} b_0(u) U_0(\underline{FH}_0^{-1}(u)) \\ &- (1 - \lambda_0)^{-\frac{1}{2}} a_0(u) V_0(\underline{GH}_0^{-1}(u))\}. \end{aligned} \tag{2.10}$$

Let D denote the space of all right continuous functions on [0, 1]. Let ρ denote the uniform metric on D and let d denote Prokhorov's (1956) metric on D. Let weak convergence of processes, denoted by $\rightarrow_L$, be as in Definition 2.1 of PS. Let the class of functions $\underline{Q}$, be as in PS. Thus $\underline{Q} = \{q \varepsilon D : q \geqq q' \text{ for some } q' \varepsilon \underline{Q}'\}$ where $\underline{Q}'$ denotes the class of all nonnegative functions defined on [0, 1] which for some $\epsilon > 0$ are bounded away from zero on $(\epsilon, 1 - \epsilon)$ and are non-decreasing (non-increasing) on $[0, \epsilon]$ ($[1 - \epsilon, 1]$) and which have square integrable reciprocals. When defined let $\rho_q(f, g) = \rho(f/q, g/q)$ and $d_q = d(f/q, g/q)$ for all $f, g \varepsilon D$.

Let ν denote a signed Lebesgue-Stieltjes measure on (0, 1) and let $|\nu|$ denote its total variation measure.

3. Statements of the main theorems. For $t > 0$ define processes $\{U_t(u) : 0 \leqq u \leqq 1\}$ and $\{V_t(u) : 0 \leqq u \leqq 1\}$ by

$$U_t(u) = m_t^{\frac{1}{2}}[F_t \underline{F}^{-1}(u) - u] \quad \text{and} \quad V_t(u) = n_t^{\frac{1}{2}}[G_t \underline{G}^{-1}(u) - u]. \tag{3.1}$$

THEOREM 1. *For any $q \varepsilon \underline{Q}$, $U_t \rightarrow_L U_0$ relative to (D, ρ_q) and to (D, d_q) as $t \rightarrow \infty$.*

ASSUMPTION 1. The functions $\underline{K}_\lambda$ have derivatives a_λ for all $u \varepsilon (0, 1)$ and for some $\lambda' \varepsilon \Lambda$ the function $a_{\lambda'}$ can be continuously extended to [0, 1]. (See Corollary 4.1 of PS for conditions under which this assumption holds.)

THEOREM 2. *Suppose Assumption* 1 *holds,* $\int_0^1 q\, d|\nu| < \infty$ *for some* $q \in \underset{\sim}{Q}$ *and*

$$\int_{N_t^{-1}}^1 L_t\, d(\nu_t - \nu) \to_p 0. \tag{3.2}$$

Then T_t^* *is asymptotically distributed as* $\int_0^1 L_0\, d\nu$; *which is a* $N(0, \sigma_0^2)$ rv *where* σ_0^2 *is finite and*

$$\begin{aligned}\sigma_0^2 = 2(1-\lambda_0)^2\{\lambda_0^{-1}\textstyle\int_0^1\int_0^v b_0(u)b_0(v)\underline{F}\underline{H}_0^{-1}(u)\,[1-\underline{F}\underline{H}_0^{-1}(v)]\,d\nu(u)\,d\nu(v)\\ + (1-\lambda_0)^{-1}\textstyle\int_0^1\int_0^v a_0(u)a_0(v)\underline{G}\underline{H}_0^{-1}(u)[1-\underline{G}\underline{H}_0^{-1}(v)]\,d\nu(u)\,d\nu(v)\}.\end{aligned} \tag{3.3}$$

Also $\rho_q(L_t', L_0) \to_p 0$ *and* $d_q(L_t', L_0) \to_p 0$ *as* $t \to \infty$; *when* L_t' *equals* L_t *for* $u \geqq N_t^{-1}$ *and equals* 0 *otherwise.*

REMARK. Suppose the measure ν is induced by a non-constant function $-J$ which is of bounded variation on $(\epsilon, 1-\epsilon)$ for all $\epsilon > 0$. Then under the hypothesis of Corollary 5.1 of PS we have that (3.2) holds and that

$$T_t^* = N_t^{\frac{1}{2}}[T_t - \textstyle\int_{-\infty}^{\infty} J(\underline{H}_0)\, d\underline{F}] + o_p(1). \tag{3.4}$$

The quantity on the right of (3.4) is a more interesting statistic than is T_t^*, because it is centered by a fixed quantity. (The proof of Proposition 5.1 in PS establishes (3.2) and the fact that $N_t^{\frac{1}{2}}[\mu_t - \int_{-\infty}^{\infty} J(\underline{H}_t)\, d\underline{F}] = o_p(1)$. That $N_t^{\frac{1}{2}}\int_{-\infty}^{\infty}[J(\underline{H}_t) - J(\underline{H}_0)]\, d\underline{F} = o_p(1)$ follows from similar manipulations.)

REMARK. In Theorem 2, Assumption 1 may be replaced by the assumption that $\lambda_{N_t} - \lambda_0 = o_p(N_t^{-\frac{1}{2}})$ and K_0 is differentiable a.e. $|\nu|$.

EXAMPLE. (1-*sample tests of symmetry*). Let $\xi_1, \cdots, \xi_N$ be iid with continuous df Ψ. Let $\lambda_0 = 1 - \Psi(0)$ and suppose $0 < \lambda_0 < 1$. It is desired to test whether Ψ is symmetric about the origin. With probability one we may let $0 < X_1 < \cdots < X_{m_N}(0 < Y_1 < \cdots < Y_{n_N})$ be an ordering of those $|\xi|$'s among $\xi_1, \cdots, \xi_N$ for which $\xi > 0(\xi < 0)$. For $x \geqq 0$ let $F(x) = [\Psi(x) - (1-\lambda_0)]/\lambda_0$ and $G(x) = [(1-\lambda_0) - \Psi(-x)]/(1-\lambda_0)$ be the conditional df's of ξ given that ξ is positive and negative respectively. Consider $T_N = m_N^{-1}\sum_{i=1}^N c_{Ni}^* Z_{Ni}$; Wilcoxon, normal scores and other 1-sample tests of symmetry can be represented as $\lambda_N T_N$ where $\lambda_N = m_N/N$. When (3.4) holds

$$N^{\frac{1}{2}}[T_N - \textstyle\int_{-\infty}^{\infty} J(\underline{H}_0)\, d\underline{F}]$$

is asymptotically $N(0, \sigma_0^2)$ with σ_0^2 given by (3.3). See Puri and Sen (1966) for a corresponding result for the more usual statistic $\lambda_N T_N$. Our result also holds if N is a random sample size.

4. Proof of Theorem 1. For $q \equiv 1$ this theorem is proved in Pyke (1968). If one divides by q in a few obvious places one may recopy all of that proof verbatim except for the statement (Pyke's equation (2.10)):

$$M_n \equiv \max_{1 \leqq k \leqq n} (k/n)^{\frac{1}{2}}\rho(U_k/q, 0) = O_p(1) \tag{4.1}$$

where $U_k(u) = k^{\frac{1}{2}}[F_k\underline{F}^{-1}(u) - u]$ for $u \in [0, 1]$ and where F_k is the empirical df of $X_1, \cdots, X_k$. (In proving random sample size theorems one shows that with high probability the random situation does not differ significantly from an appropriate fixed sample size situation; M_n arises when that difference is considered. Hence it is necessary to introduce this notation for the fixed sample size process. The context should insure that the reader will not be confused by the notation U_k, F_k for fixed sample size k and U_t, F_t for random size N_t.) To prove (4.1) with $q \equiv 1$, Pyke used a weak convergence result of Dudley (1966); a result which does not apparently allow for division by $q \varepsilon \bar{Q}$. The general outline of the proof of (4.1) given below is related to that used by Donsker (1952).

As in Pyke (1968), one may write

$$(4.2)\quad P[M_n > y] = P[\sup_{0 \leqq u,t \leqq 1} |Z_n(u, t)|/q(u) > n^{\frac{1}{2}}y \mid V_n(1, 1) = 0]$$

where

$$(4.3)\quad Z_n(u, t) = V_n(u, t) - uV_n(1, t), \qquad V_n(u, t) = N_n(u, t)\text{-}nut$$

and $\{N_n(u, t) : 0 \leqq u, t \leqq 1\}$ is a Poisson process over the unit square with parameter $E[N_n(1, 1)] = n$. Let

$$(4.4)\qquad S = \inf \{t : \sup_{0 \leqq u \leqq 1} |Z_n(u, t)|/q(u) > n^{\frac{1}{2}}y\}$$

where one sets $S = 1$ if this set is empty. Write $S_J = ([SJ] + 1)/J$ where [ ] is the greatest integer function. For $\delta > 0$, define the event

$$(4.5)\qquad A_J = [\sup_{0 \leqq u \leqq 1} |Z_n(u, S) - Z_n(u, S_J)|/q(u) \geqq n^{\frac{1}{2}}\delta]$$

and set

$$(4.6)\quad \pi_{nJ}(y) = P[\max_{1 \leqq j \leqq J} \sup_{0 \leqq u \leqq 1} |Z_n(u, j/J)|/q(u) > n^{\frac{1}{2}}y \mid V_n(1, 1) = 0].$$

Then

$$\begin{aligned}
\pi_{nJ}(y) &\leqq P[M_n > y] = P[S < 1 \mid V_n(1, 1) = 0] \\
&= P(A_J^c \cap [S < 1] \mid V_n(1, 1) = 0) + P(A_J \cap \\
(4.7)\qquad &\qquad [S < 1] \mid V_n(1, 1) = 0) \\
&\leqq \pi_{nJ}(y - \delta) + P(A_J \cap [S < 1] \mid V_n(1, 1) = 0) \\
&= \pi_{nJ}(y - \delta) + P(A_J \cap [S < 1] \cap \\
&\qquad [V_n(1, 1) = 0])/P[V_n(1, 1) = 0].
\end{aligned}$$

Define the events

$$B_{1J} = [\sup_{0 \leqq u \leqq 1} (1 - u)|V_n(u, S_J) - V_n(u, S|/q(u) \geqq n^{\frac{1}{2}}\delta/2]$$

and

$$\begin{aligned}
B_{2J} = [\sup_{0 \leqq u \leqq 1} u|[V_n(1, S_J) &- V_n(u, S_J)] \\
&- [V_n(1, S) - V_n(u, S)]|/q(u) \geqq n^{\frac{1}{2}}\delta/2.]
\end{aligned}$$

Then

$$\begin{aligned}
(4.8)\quad &P(A_J \cap [S < 1] \cap [V_n(1, 1) = 0]) \\
&\qquad \leqq \textstyle\sum_{i=1}^{2} P(B_{iJ} \cap [S < 1] \cap [V_n(1, 1) = 0])
\end{aligned}$$

since $A_J \subset B_{1J} \cup B_{2J}$. Using the independent increments of the Poisson process we have for $s < 1 - 1/J$ and $i = 1, 2$ that

$$\begin{aligned} &P(B_{iJ} \cap [S < 1] \cap [V_n(1, 1) = 0] \mid S = s, V_n(1, S) = v) \\ &\quad = \textstyle\sum_z P(B_{iJ}^s \cap [V_n(1, S_J) - V_n(1, S) = z] \\ &\qquad\qquad \cap [V_n(1, 1) = 0] \mid S = s, V_n(1, S) = v) \\ &\quad = \textstyle\sum_z P(B_{iJ}^s \cap [V_n(1, s_J) - V_n(1, s) = z])P[V_n(1, 1) \\ &\qquad\qquad - V_n(1, s_J) = -v - z] \end{aligned} \tag{4.9}$$

for each v, where $s_J = ([sJ] + 1)/J_0$ and B_{iJ}^s is defined as B_{iJ} only with s and s_J in place of S and S_J. We use the fact that if Z is Poisson with mean λ then there is a constant c such that

$$\max_z P[Z = z] = P[Z = [\lambda]] \leqq c\lambda^{-\frac{1}{2}}; \tag{4.10}$$

the inequality in (4.10) follows from Stirlings approximation. It follows by using (4.10) that for $s < 1 - 1/J$ the term on the right hand side of (4.9) does not exceed

$$\begin{aligned} cJ^{\frac{1}{2}}n^{-\frac{1}{2}}\textstyle\sum_z P(B_{iJ}^s \cap [V_n(1, s_J) - V_n(1, s) = z]) &= cJ^{\frac{1}{2}}n^{-\frac{1}{2}}P(B_{iJ}^s) \\ &\leqq cJ^{\frac{1}{2}}n^{-\frac{1}{2}}(4/n\delta^2)\textstyle\int_0^1 [q(u)]^{-2}n(s_J - s)\,du \\ &\leqq c^*\delta^{-2}J^{-\frac{1}{2}}n^{-\frac{1}{2}} \end{aligned}$$

where c^* denotes a generic constant and the next to the last inequality follows from Theorem 5.1 of Birnbaum and Marshall (1961). Thus taking expectations in (4.9) we get

$$P(B_{iJ} \cap [S < 1 - 1/J] \cap [V_n(1, 1) = 0]) \leqq c^*\delta^{-2}J^{-\frac{1}{2}}n^{-\frac{1}{2}}. \tag{4.11}$$

A separate argument is necessary for the case $1 - 1/J \leqq S < 1$. Define

$$S^* = \inf\{t \geqq 1 - 1/J: \sup_{0 \leqq u \leqq 1} |Z_n(u, t)|/q(u) > n^{\frac{1}{2}}y\}$$

where one sets $S^* = 1$ if this set is empty. Define $S_J{}^*$ and B_{iJ}^* $(i = 1, 2)$ in terms of S^* analogously to their unstarred counterparts. Then, for $i = 1, 2$, one obtains

$$\begin{aligned} &P(B_{ij} \cap [1 - 1/J \leqq S < 1] \cap [V_n(1, 1) = 0]) \\ &\quad = \textstyle\sum_{z,w} P(B_{iJ}^* \cap [S^* = S < 1] \cap [V_n(1, 1) - V_n(1, S^*) = z] \\ &\qquad\qquad \cap [V_n(1, S^*) - V_n(1, 1 - 1/J) = w] \cap [V_n(1, 1 - 1/J) = -z - w]) \\ &\quad \leqq \textstyle\sum_{z,w} P(B_{iJ}^* \cap [V_n(1, 1) - V_n(1, S^*) = z]) \\ &\qquad\qquad \cdot P[V_n(1, S^*) - V_n(1, 1 - 1/J) = w]P[V_n(1, 1 - 1/J) = -z - w] \end{aligned}$$

because of independent increments and the fact that S^* is a stopping time. By (4.10) the last factor is bounded by $cn^{-\frac{1}{2}}(1 - 1/J)^{-\frac{1}{2}} \leqq 2cn^{-\frac{1}{2}}$ for large J.

Therefore, after summing over z and w one obtains

$$\begin{aligned} P(B_{ij} \cap [1 - 1/J \leqq S < 1] \cap [V_n(1, 1) = 0]) & \\ &\leqq 2cn^{-\frac{1}{2}}P(B_{iJ}^*) \\ &\leqq c^* n^{-\frac{1}{2}}\delta^{-2}E(\textstyle\int_0^1 [q(u)]^{-2}(1 - S^*)\, du) \\ &\leqq c^*\delta^{-2}n^{-\frac{1}{2}}J^{-1} \end{aligned}$$

by means of the Birnbaum-Marshall inequality and the fact that $1 - S^* \leqq 1/J$ by definition.

Combining this last inequality with (4.11) implies that (4.8) never exceeds $c^*\delta^{-2}(Jn)^{-\frac{1}{2}}$. Upon substitution of this bound into (4.7) and application of (4.10) to the term $P[V_n(1, 1) = 0]$, it follows that

$$\pi_{nJ}(y) \leqq P(M_n > y) \leqq \pi_{nJ}(y - \delta) + c^*\delta^{-2}J^{-\frac{1}{2}} \tag{4.12}$$

for all $\delta > 0$.

We now study the limit of $\pi_{nJ}(y)$ as $n \to \infty$. To do this, consider J fixed and define

$$X_{nj}(u) = n^{-\frac{1}{2}}[Z_n(u, r_j) - Z_n(u, r_{j-1})] \tag{4.13}$$

where $r_j = j/J$ and consider the J independent processes $\{X_{nj}(u) : 0 \leqq u \leqq 1\}$, $(1 \leqq j \leqq J)$. By definition we may rewrite (4.13) as

$$\begin{aligned} X_{nj}(u) &= n^{-\frac{1}{2}}\{N_n(u, r_j) - N_n(u, r_{j-1}) - u[N_n(1, r_j) - N_n(1, r_{j-1})]\} \\ &= (N_{nj}/n)^{\frac{1}{2}}N_{nj}^{\frac{1}{2}}\, [N_{nj}(u)/N_{nj} - u] \end{aligned} \tag{4.14}$$

where we have set $N_{nj}(u) = N_n(u, r_j) - N_n(u, r_{j-1})$ and $N_{nj} = N_{nj}(1)$. Under the condition $[V_n(1, 1) = 0]$, (equivalently $[N_n(1, 1) = n]$), the random vector $(N_{n1}, N_{n2}, \cdots, N_{nJ})$ has the multinomial distribution of sample size n and with cell probabilities $p_j = J^{-1}$, $1 \leqq j \leqq J$. Under the condition $[N_n(1, 1) = n, N_{nj} = m_j\,; 1 \leqq j \leqq J]$, the X_{nj}-processes are conditionally independent while each X_{nj}-process has conditionally the same probability structure as $(m_j/n)^{\frac{1}{2}}W_{m_j}$ where W_K is the empirical process

$$W_K(u) = K^{\frac{1}{2}}[F_K(u) - u], \qquad 0 \leqq u \leqq 1, \tag{4.15}$$

based on a sample of K uniform-$(0, 1)$ rv's. By Theorem 2.1 of PS it is possible to construct W_K-processes for which it is true that $\rho_q(W_K, W_0) \to_{\text{a.s.}} 0$ where W_0 is a tied-down Wiener process and $q \varepsilon \underline{Q}$. Therefore, by the above mentioned conditional independence, it is possible to construct a sequence of multinomial random vectors $\{(N'_{n1} \cdots, N'_{nJ}) : n \geqq 1\}$ with equal cell probabilities $p_j = J^{-1}$ and J independent sequences of empirical processes

$$\{W_K^{(1)}, \cdots, W_K^{(J)}) : K \geqq 1\}$$

for which

$$\max_{1 \leqq j \leqq J} \rho_q((N'_{nj}/n)^{\frac{1}{2}} W^{(j)}_{N'_{nj}}, \quad J^{-\frac{1}{2}} W_0^{(j)}) \rightarrow_p 0$$

where $W_0^{(1)}, \cdots, W_0^{(J)}$ are independent tied-down Wiener processes. In terms of the X_{nj}-processes this implies in particular that

$$\begin{aligned} \pi_{nJ}(y) &= P[\max_{1 \leqq j \leqq J} \sup_{0 \leqq u \leqq 1} |X_{n1}(u) + \cdots \\ &\quad + X_{nj}(u)|/q(u) > y \mid V_n(1, 1) = 0] \\ &\rightarrow P[\max_{1 \leqq j \leqq J} \sup_{0 \leqq u \leqq 1} |X_{01}(u) + \cdots + X_{0j}(u)|/q(u) > y] \end{aligned} \tag{4.16}$$

as $n \rightarrow \infty$, where $(X_{01}, \cdots, X_{0J})$ is a vector of independent Wiener processes with X_{0j} equivalent to $J^{-\frac{1}{2}} W_0^{(j)}$. If we let $\pi_J(y)$ denote the limit in (4.16) we have that $\pi_J(y)$ is well defined and $\pi_J(y) \rightarrow 0$ as $y \rightarrow \infty$. It therefore follows by taking limits as $n \rightarrow \infty$ in (4.12), that

$$\begin{aligned} \pi_J(y) &\leqq \liminf_n P[M_n > y] \\ &\leqq \limsup_n P[M_n > y] \leqq \pi_J(y - \delta) + c^*/\delta^2 J^{\frac{1}{2}}. \end{aligned} \tag{4.17}$$

This suffices to establish (4.1) and hence Theorem 1 since the right hand side of (4.17) can be made arbitrarily small by choosing y and J sufficiently large. □

It is possible however to establish a stronger result than that contained in the above proof concerning the convergence in law of the sequence $\{M_n\}$. To do this we first introduce a 2-dimensional parameter Wiener process,

$$\{X(u, s) : 0 \leqq u, s \leqq 1\},$$

defined by saying that (i) $X(\cdot, s)$ for each s is a tied-down Wiener process with mean zero and $E[X(u, s)X(v, s)] = su(1 - v), 0 \leqq u \leqq v \leqq 1$, and (ii) $X(\cdot, r)$ and $X(\cdot, s) - X(\cdot, r)$ are independent for $0 \leqq r \leqq s \leqq 1$. Thus X is a Gaussian process with mean zero and covariance $E[X(u, r)X(v, s)] = \min(r, s)u(1 - v)$ if $u \leqq v$. Equivalently the X-process may be described as having the same finite-dimensional df's as a Brownian motion on the unit square which is tied-down along the line $u = 1$. Assume, without loss of generality, (Lévy (1954), p. 73), that with probability one, X is continuous. With this definition one may rewrite the right hand side of (4.16) as

$$\pi_J(y) = P[\max_{1 \leqq j \leqq J} \sup_{0 \leqq u \leqq 1} |X(u, j/J)|/q(u) > y]. \tag{4.18}$$

THEOREM 3. *As* $n \rightarrow \infty$,

$$M_n \rightarrow_L M \equiv \sup_{0 \leqq u, s \leqq 1} |X(u, s)|/q(u). \tag{4.19}$$

PROOF. The proof is outlined as follows. If one defines $Y(x, s) = (1 + x)X(x/(1 + x), s)$, it follows as in Doob (1949) that Y is a Brownian motion over $[0, \infty) \times [0, 1]$. By Lévy's Holder condition (Lévy (1954), p. 73) and the fact that $q(cu)[u \log(1/u)]^{-\frac{1}{2}} \rightarrow \infty$ as $u \rightarrow 0$ for any $c > 0$, it follows that

$$(4.20) \quad \inf\{x > 0 : |Y(x, s)| > yq(x/(1 + x)) \quad \text{for all} \quad 0 \leqq s \leqq 1\} > 0$$

with probability one. In view of the relationship between X and Y, the a.s. positivity of this stopping time suffices to show that

$$(4.21) \qquad \lim_{J\to\infty} \pi_J(y) = P[M > y].$$

(It should be remarked that if s in (4.20) is restricted to the values j/J, $(1 \leqq j \leqq J)$, the positivity is a consequence of the law of the iterated logarithm for ordinary 1-dimensional Brownian motion. This however would not be enough to establish (4.21).) The proof is now completed by letting first $J \to \infty$ and then $\delta \to 0$ in (4.17). $\square$

5. Proof of Theorem 2. The proof of this theorem corresponds closely to the proof of Theorem 4.1 of PS. We ask the reader to reread that proof (Sections 2, 3 and 4 of PS) at this time; here we indicate only the changes that need to be made. In PS we took the weak convergence of the empirical process as our starting point; in Lemma 1 below we establish the corresponding weak convergence result. Define

$$\begin{aligned}
W_t(u) &= \lambda_t & \text{and} \quad W_0(t) &= \lambda_0 & &\text{for} \quad 0 \leqq u < \tfrac{1}{5}\\
&= 1 - \lambda_t & &= 1 - \lambda_0 & &\text{for} \quad \tfrac{1}{5} \leqq u < \tfrac{2}{5}\\
&= N_t/t & &= \alpha + \beta & &\text{for} \quad \tfrac{2}{5} \leqq u < \tfrac{3}{5}\\
&= U_t(5u - 3) & &= U_0(5u - 3) & &\text{for} \quad \tfrac{3}{5} \leqq u \leqq \tfrac{4}{5}\\
&= V_t(5u - 4) & &= V_0(5u - 4) & &\text{for} \quad \tfrac{4}{5} \leqq u \leqq 1.
\end{aligned}$$

(We wish to replace U_t, V_t, m_t and n_t by random quantities which have the same finite dimensional distributions but which satisfy additional requirements. Introducing the W_t-processes merely provides a convenient way to show that this is possible.)

LEMMA 1. *$W_t \to_L W_0$ relative to (D, d).*

PROOF. By Theorem 15.1 of Billingsly (1967) it suffices to show that (i) the family of distributions Π_W of the W_t-processes is tight and (ii) the finite dimensional distributions of the W_t-processes converge to those of the W_0-process. We now prove (i). From Pyke (1968) we know that $U_t \to_L U_0$ and $V_t \to_L V_0$ relative to (D, d). Thus the families of distributions Π_U and Π_V of the U_t- and V_t-processes respectively are tight by Billingsley's Theorem 6.2. The definition of tightness of a family Π of probability distributions P on (D, d) implies that for all $\epsilon > 0$ there exists a compact subset K^ϵ of D such that $P(K_\epsilon) > 1 - \epsilon$ for all $P \in \Pi$. Let K_U^ϵ and K_V^ϵ be the guaranteed compact sets for Π_U and Π_V respectively. Since compactness is equivalent to sequential compactness in the metric space (D, d), it is easy to find a compact set K_W^ϵ related to K_U^ϵ and K_V^ϵ such that $P(K_W^\epsilon) > 1 - 3\epsilon$ for all $P \in \Pi_W$. We next prove (ii). Let $u_1, \cdots,$

u_k be any finite subset of $[\frac{3}{5}, 1]$. We will show that $(W_t(u_1), \cdots, W_t(u_k))$ is asymptotically distributed as $(W_0(u_1), \cdots, W_0(u_k))$; the extension to the case of $u_1, \cdots, u_k$ in $[0, 1]$ being trivial. Since the $W_0(u_i)$'s are jointly normal, it suffices to show that the distribution of all linear combinations of the $W_t(u_i)$'s converge to the distribution of the same linear combination of $W_0(u_i)$'s; and this latter fact follows from Theorem 1 of Miller (1961). Alternatively, a direct proof is possible since the linear combination may be written as

$$[\sum_1^{m_t} W_i/m_t^{\frac{1}{2}} + \sum_1^{n_t} Z_j/n_t^{\frac{1}{2}}]$$

where the W_i's are iid and the Z_j's are iid. The result then follows from the proof of the 1-dimensional theorem of Anscombe (1952); Kolmogorov's inequality is used separately on each of two error terms. □

As in equation (2.2) of PS we use item 3.1.1 of Skorokhod (1956) to construct processes $\tilde{W}_t$ and $\tilde{W}_0$ on a probability space $(\tilde{\Omega}, \tilde{\mathfrak{A}}, \tilde{P})$ that have the same finite dimensional distributions as do the processes W_t and W_0 on $(\Omega, \mathfrak{A}, P)$ and which further satisfy $\rho(\tilde{W}_t, \tilde{W}_0) \rightarrow_{\text{a.s.}} 0$ as $t \rightarrow \infty$. We now define for $u \varepsilon [0, 1]$ and $t > 0$

$$\tilde{N}_t = t\tilde{W}_t(\tfrac{2}{5}), \qquad \tilde{m}_t = \tilde{N}_t\tilde{W}_t(0), \qquad \tilde{n}_t = \tilde{N}_t\tilde{W}_t(\tfrac{1}{5}), \qquad \tilde{\lambda}_t = \tilde{m}_t/\tilde{N}_t,$$

$$\tilde{U}_t(u) = \tilde{W}_t((u+3)/5), \qquad \tilde{U}_0(u) = \tilde{W}_0((u+3)/5),$$

$$\tilde{V}_t(u) = \tilde{W}_t(u+4)/5), \qquad \tilde{V}_0(u) = \tilde{W}_0((u+4)/5).$$

Note that

$$\tilde{m}_t/t \rightarrow_{\text{a.s.}} \alpha, \qquad \tilde{n}_t/t \rightarrow_{\text{a.s.}} \beta, \qquad \tilde{\lambda}_t \rightarrow_{\text{a.s.}} \lambda_0$$

and that almost surely $\tilde{U}_t(\tilde{V}_t)$ has $\tilde{m}_t(\tilde{n}_t)$ jumps of size $\tilde{m}_t^{-\frac{1}{2}}(\tilde{n}_t^{-\frac{1}{2}})$ and is otherwise continuous (Lemma 1 was needed to be able to apply Skorokhod's result.)

From now on we suppress the symbol $\sim$ in our notation though everything that follows refers to these specially constructed processes. As the proofs of the lemmas and theorems of PS are now reread, subscripts m, n, N should be replaced by t, $N \rightarrow \infty$ by $t \rightarrow \infty$ and the numbers m, n, N should be subscripted by t. Theorem 2.1 of PS and hence Lemma 2.2 also, has been proved as Theorem 1 of this paper. The proof of Lemma 2.3 of PS carries over as it is, since $N_t \rightarrow_{\text{a.s.}} \infty$ as $t \rightarrow \infty$ (provided we consider the version of Lemma 2.3 that uses (2.2) of PS; otherwise we get only $\rightarrow_p$.)

Lemma 2.5 of PS can not be recopied until we show that the random sample size analog of equation (2.11) of PS is true; thus we must show for every $\epsilon > 0$ there exists $b > 0$ such that $P(A_t) > 1 - \epsilon$ where

$$(5.1) \qquad A_t = [F(u) \leqq bF_t(u) \quad \text{for all} \quad u \quad \text{where} \quad F_t(u) > 0].$$

Let $B_t = [|m_t/t - \alpha| \leqq \delta]$, $r_t = [t(\alpha - \delta)]$, $s_t = [t(\alpha + \delta)]$ and

$$C_t = [\min\{X_i : 1 \leqq i \leqq r_t\} = \min\{X_i : 1 \leqq i \leqq s_t\}].$$

Then

$$\begin{aligned}
P(A_t) &\geqq P(A_t \cap B_t \cap C_t) \\
&\geqq P([F(u) \leqq bF_t(u) \text{ for all } (m_t, u) \quad \text{such that} \quad r_t \leqq m_t \leqq s_t \quad \text{and} \\
&\qquad F_t(u) > 0] \cap C_t) \\
&\geqq P([(r_t/s_t)bF_{r_t}(u) \geqq F(u) \quad \text{for all } u \text{ such that} \quad F_{r_t}(u) > 0] \cap C_t) \\
&\geqq 1 - \epsilon/2 - P(C_t^c),
\end{aligned}$$

for sufficiently large b, this last inequality following directly from (2.11) of PS. Letting $\xi_t = \min\{X_i : 1 \leqq i \leqq r_t\}$ we have

$$\begin{aligned}
P(C_t) &= E((1 - \xi_t)^{s_t - r_t}) \\
&\rightarrow E(e^{-Z2\delta/(\alpha-\delta)}) = (\alpha - \delta)/(\alpha + \delta)
\end{aligned}$$

where Z is an exponential rv with mean one since $r_t\xi_t \rightarrow_L Z$. Thus for small δ, $P(C_t) \geqq 1 - \epsilon/2$ for t sufficiently large. We may now recopy the rest of the proof of Lemma 2.5 of PS provided we replace A_t by $A_t \cap D_t$ where $D_t = [\lambda_t^{-1} \leqq \lambda_*^{-1}]$ and t is so large that $P(D_t) \geqq 1 - \epsilon$.

Likewise the rest of Section 2 and all of Sections 3 and 4 of PS carry over with only trivial modifications. □

The second remark in Section 3 of this paper follows by rereading Section 5 of PS. (The statement $\lambda_{N_t} - \lambda_0 = o_p(N_t^{-\frac{1}{2}})$ is stronger than is needed for Lemma 5.2; however, the statement $\lambda_{N_t} - \lambda_0 = O_p(N_t^{-\frac{1}{2}})$ is not strong enough.)

NOTE ADDED IN PROOF. An alternate proof of Theorem 1 is possible by deriving a suitable 2-dimensional generalization of the Birnbaum-Marshall inequality which may then be applied directly to bound $P(A_J)$ appropriately. Such an inequality is derivable by paralleling the proof of Wichura's multidimensional Kolmogorov inequality; (Lemma 2.1 of Wichura [14]).

REFERENCES

[1] ANSCOMBE, F. J. (1952). Large sample theory of sequential estimation. *Proc. Cambridge Phil. Soc.* **48** 600–607.

[2] BILLINGSLEY, P. (1967). *Weak convergence of probability measures.* Lecture notes to be published by Wiley, New York.

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

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

[5] DONSKER, M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **23** 277–281.

[6] DUDLEY, R. M. (1966). Weak convergence of probabilities on nonseparable metric spaces and empirical measures on Euclidean spaces. *Illinois J. Math.* **10** 109–126.

[7] LÉVY, PAUL (1954). *Le Mouvement Brownien.* Mem. Sci. Math. **126** Gauthier-Villars, Paris.

[8] MILLER, R. (1961). Asymptotic multivariate occupation time distributions for semi-Markov processes. Stanford Univ. Technical Report No. 70, Dept. of Statistics.

[9] PROKHOROV, YU. V. (1956) Convergence of random processes and limit theorems in probability theory. *Theor. Prob. Appl.* **1** 157–214.

[10] PURI, M. L. and SEN, P. K. (1966). On the asymptotic normality of one sample Chernoff-Savage test statistics. New York Univ. Technical report IMM350, Courant Institute of Mathematical Sciences.

[11] PYKE, R. (1968). The weak convergence of the empirical process with random sample size. *Proc. Cambridge Phil. Soc.* **64** 155–160.

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

[13] SKOROKHOD, A. V. (1956). Limit theorems for stochastic processes *Theor. Prob. Appl.* **1** 261–290.

[14] WICHURA, M. (1968). On the weak convergence of non-Borel probabilities on a metric space. Unpublished Ph.D. dissertation, Columbia Univ.