The Annals of Probability
1974, Vol. 2, No. 4, 642-651

ON THE WEAK CONVERGENCE OF EMPIRICAL PROCESSES IN SUP-NORM METRICS

BY NEVILLE E. O'REILLY
Rutgers University

Necessary and sufficient conditions for the weak convergence in stronger than the usual metrics of empirical, inverse empirical and partial sum processes are obtained. The results are related to work of Chibisov and of Pyke and Shorack.

1. Introduction. Let $U_{1n} \leqq U_{2n} \leqq \cdots \leqq U_{nn}$ be random variables obtained by ordering n independent observations from the uniform distribution on [0, 1]. Set $U_{0n} \equiv 0$, $U_{n+1,n} \equiv 1$ and let F_n and F_n^{-1} be defined by

$$F_n(t) = k/n, \ U_{kn} \leqq t < U_{k+1,n}, \qquad k = 0, 1, \cdots, n$$
$$= 1, \qquad t = 1 \tag{1.1}$$

and

$$F_n^{-1}(t) = U_{kn}, \ k/(n+1) \leqq t < (k+1)/(n+1), \qquad k = 0, 1, \cdots, n$$
$$= 1, \qquad t = 1. \tag{1.2}$$

That is, F_n is the empirical distribution function of a sample of n uniform random variables and F_n^{-1} is a version of its inverse. Two sequences of related processes are the empirical process

$$U_n(t) = n^{\frac{1}{2}}(F_n(t) - t), \qquad 0 \leqq t \leqq 1 \tag{1.3}$$

and the inverse empirical process

$$V_n(t) = n^{\frac{1}{2}}(F_n^{-1}(t) - t), \qquad 0 \leqq t \leqq 1. \tag{1.4}$$

These processes have realizations in the space $D[0, 1]$ of real valued functions on [0, 1] having only jump discontinuities.

In studying the asymptotic distribution of statistics based on ordered observations (in particular linear rank statistics and linear combinations of order statistics) various authors, starting with Pyke and Shorack in [7], have made use of the weak convergence of U_n and V_n to a tied down Wiener process in metrics stronger than the usual Prohorov and sup-norm metrics; that is, weak convergence in metrics of the form

$$d_q(x, y) = d(x/q, y/q) \qquad \text{and}$$
$$\rho_q(x, y) = \rho(x/q, y/q) \tag{1.5}$$

where d is the Prohorov metric, ρ is the sup-norm metric and q is a nonnegative function approaching zero at the endpoints of the interval [0, 1].

Received April 30, 1973; revised November 5, 1973.
AMS 1970 *subject classification.* Primary F15, G50, G99, E20.
Key words and phrases. Weak convergence, empirical process, sup-norm metric.

A question naturally arises as to how fast the function $q(t)$ can be allowed to approach zero at the interval endpoints and convergence in the metric ρ_q and d_q continue to hold. Pyke and Shorack [7] give a sufficient condition which essentially is monotonicity of q near the endpoints and convergence of the integral

$$\int_0^1 q^{-2}(t)\,dt\,. \tag{1.6}$$

A finer result was given by Chibisov [2] for the process U_n. Chibisov showed that, roughly speaking, U_n has in the limit as $n \to \infty$ the same sets of upper and lower class functions as the Wiener (and tied-down Wiener) process. He also gave, under the assumption that $q(t)$ and $q(1-t)$ are regularly varying at 0 of order $0 < \alpha \leqq \frac{1}{2}$, a necessary and sufficient condition for the convergence of U_n to a tied-down Wiener process in (D, ρ_q); namely

$$\int_0^1 t^{-1} \exp(-\varepsilon h_i^2(t))\,dt < \infty \qquad \text{for every} \quad \varepsilon > 0\,, \qquad i = 1, 2 \tag{1.7}$$

where $h_1(t) = t^{-\frac{1}{2}}q(t)$ and $h_2(t) = t^{-\frac{1}{2}}q(1-t)$. (Chibisov's proof utilises a criterion, 2.2 on page 148 of [2], for membership in the upper and lower classes of the Wiener process which is given without proof and does not appear to be in the literature; in particular it is not in P. Lévy's monograph [5] to which the reader is referred in [2]).

In this paper necessary and sufficient conditions for the weak convergence of V_n and U_n are given in Section 3 and Section 4, respectively. These results are more general than those of Pyke and Shorack and Chibisov's result is stated as a corollary to Theorem 2. Section 2 contains preliminary results and in Section 5 it is indicated how the corresponding results for the partial sum and two sample empirical processes can be similarly derived. Only the metric ρ_q where

$$\rho_q(x, y) = \sup\{|x(t) - y(x)|/q(t) : 0 \leqq t \leqq 1\}$$

is treated; the results for d_q are immediate consequences of those for ρ_q since the limiting process involved is almost everywhere continuous.

2. Preliminary results. Let $W(t)$ denote a standard Wiener process; that is a separable Gaussian process with $E\,W(t) = 0$, $E[W(s)W(t)] = \min(s, t)$. Set $W_0(t) = W(t) - t\,W(1)$ so that $W_0(t)$ is a Wiener process "tied-down" at $t = 1$. The proof of the main results requires that the probability that $W_0(t)/q(t)$ exceed an arbitrary ε near zero (or one) be small and since the local behavior of $W_0(t)$ is the same as that of $W(t)$ we need the following proposition.

PROPOSITION 2.1. *Let $q(t)$ be a continuous nonnegative function on* $[0, 1]$ *which is non-decreasing in a neighborhood of* 0. *Then a necessary and sufficient condition that $\varepsilon q(t)$ be in the upper class of the Wiener process for every $\varepsilon > 0$ is the convergence of the integral*

$$\int_0^1 t^{-1} \exp(-\varepsilon h^2(t))\,dt \tag{2.2}$$

for all $\varepsilon > 0$ and where $h(t) = t^{-\frac{1}{2}}q(t)$.

PROOF. Suppose the integral in (2.2) is finite for every $\varepsilon > 0$. Then for $0 < \lambda < 1$,

$$\exp(-h^2(t)) \leqq (-\log(\lambda))^{-1} \int_{\lambda t}^{t} s^{-1} \exp(-\lambda h^2(s))\, ds$$

where the right side tends to zero with t. Thus $h(t)$ becomes infinite as t approaches zero and for any $\varepsilon > 0$ and all small enough t

$$h(t) \leqq \exp(\varepsilon h^2(t)/2) .$$

Hence for small θ

$$\int_0^\theta t^{-1} h(t) \exp(-\varepsilon h^2(t))\, dt \leqq \int_0^\theta t^{-1} \exp\left(-\frac{\varepsilon}{2}(t)\right) dt < \infty . \tag{2.3}$$

By the well-known Kolmogorov criterion (see e.g., [4] page 33) the finiteness of the integral on the left side of (2.3) is sufficient for $\varepsilon q(t)$ to be in the upper class. (Although the Kolmogorov test is stated in [4] under a monotonicity assumption on $h(t)$ this is not needed for the relevant half of the test).

For the proof in the other direction we first introduce some notation. Let $b_j = \theta\lambda^j$, $L_j = (\theta\lambda^{j+1}, \theta\lambda^j]$, $B_{jx} = \{-q'(b_k) - x \leqq W(t) - W(b_j) \leqq q'(b_k) - x$ for all t in L_k, all $k = 0, 1, \cdots, j-1\}$, $A_{jx} = \{b_j W(t) - tW(b_j) > b_j q'(b_j) - tx$ for some t in $L_j\}$, $j = 0, 1, 2, \cdots$, where for given $\varepsilon > 0$, $0 < \lambda < 1$, $q'(t) = (\lambda\varepsilon/2)^{\frac{1}{2}} q(t)$. Further let $F_j(x)$ denote the distribution function of $W(b_j)$.
Then

$$\begin{aligned}
&P(W(t) > q'(t) \text{ some } 0 < t \leqq \theta) \\
&\quad \geqq P \textstyle\bigcup_{j=0}^\infty \{W(t) > q'(b_j) \text{ some } t \text{ in } L_j\} \\
&\quad \geqq \textstyle\sum_{j=1}^\infty P\{W(t) > q'(b_j) \text{ some } t \text{ in } L_j, \\
&\qquad\qquad W(t) \leqq q'(b_k) \text{ all } t \text{ in } L_k, k = 0, 1, \cdots, j-1\} \\
&\quad \geqq \textstyle\sum_{j=1}^\infty P\{W(t) > q'(b_j) \text{ some } t \text{ in } L_j, \\
&\qquad\qquad -q'(b_k) \leqq W(t) \leqq q'(b_k) \text{ all } t \text{ in } L_k, k = 0, 1, \cdots, j-1\} .
\end{aligned} \tag{2.4}$$

Conditioning each term on $W(b_j)$ makes the right-hand side of (2.4) equal to

$$\textstyle\sum_{j=1}^\infty \int_{-\infty}^\infty P(A_{jx}) P(B_{jx}) F_j(dx) , \tag{2.5}$$

which in turn is bounded below by

$$\begin{aligned}
&\textstyle\sum_{j=1}^\infty P\{W_0(t) > b_j^{-\frac{1}{2}} q'(b_j) \text{ for some } 0 < t < 1\} \int_0^\infty P(B_{jx}) F_j(dx) \\
&\quad \geqq \tfrac{1}{2} \textstyle\sum_{j=1}^\infty \exp(-\lambda\varepsilon h^2(b_j)) P\{-q'(t) \leqq W(t) \leqq q'(t) \text{ all } b_j < t \leqq \theta\} \\
&\quad \geqq \tfrac{1}{4} \textstyle\sum_{j=1}^\infty \exp(-\lambda\varepsilon h^2(\theta\lambda^j)) \qquad \text{for small enough } \theta \\
&\quad \geqq (-4\log(\lambda))^{-1} \textstyle\sum_{j=1}^\infty \int_{b_j}^{b_{j-1}'} t^{-1} \exp(-\varepsilon h^2(t))\, dt \\
&\quad = (-4\log(\lambda))^{-1} \int_0^\theta t^{-1} \exp(\varepsilon h^2(t))\, dt .
\end{aligned} \tag{2.6}$$

Since the left side of (2.4) approaches zero with θ and bounds the right side of (2.6) from above, the convergence of the integral in (2.2) is shown. The two lemmas below are needed to obtain bounds on distribution tails.

LEMMA 2.7. *Let $Y_1, Y_2, \cdots, Y_n$ be independent random variables. For $1 \leqq m < n$,* *set* $S_j = Y_1 + Y_2 + \cdots + Y_j$, $T_j = Y_j + Y_{j+1} + \cdots + Y_n$, $m \leqq j \leqq n$ *and* $T_{n+1} = 0$. *Let* $a \leqq 0 \leqq b$ *and* $B_j = \{T_j \geqq a\}$, $\gamma \leqq \min(P(B_j), m \leqq j \leqq n+1)$. *Then*

$$\gamma P\{\max(S_j, m \leqq j \leqq n) \geqq b\} \leqq P\{S_n \geqq a + b\}. \tag{2.8}$$

PROOF. See Freedman [3] on page 12.

LEMMA 2.9. *Let* $Z_{n+1} = Y_1 + Y_2 + \cdots + Y_{n+1}$ *where the* Y_i *are independent random variables each with distribution function*

$$F(y) = 0, \quad y < 0; \qquad = 1 - \exp(-y), \quad y \geqq 0. \tag{2.10}$$

Then for any sequence $\alpha_n \downarrow 0$,

$$P\{|(Z_{n+1}/n) - 1| > \alpha_n\} \leqq 5(n/2\pi)^{\frac{1}{2}} \exp(-n\alpha_n^2/8) = r(n), \quad \textit{say}. \tag{2.11}$$

PROOF. This lemma is due to Rosenkrantz and is proven in Lemma 4 of [9].

In order to avoid measurability problems occasioned by nonseparability of (D, ρ_q) (and overlooked by Chibisov in [2]) we understand "weak convergence" in the sense of Definition 2.1 of [7]. For separable spaces this definition reduces to the Prokhorov definition.

3. Inverse empirical process. For the process V_n we prove the following theorem.

THEOREM 1. *Let* $q(t)$ *be a continuous, nonnegative function on* $[0, 1]$, *bounded away from zero on* $[\gamma, 1-\gamma]$ *for some* $\gamma > 0$, *non-decreasing* (*non-increasing*) *on* $[0, \gamma]$ $([1-\gamma, 1])$. *Then*

$$\int_0^1 t^{-1} \exp(-\varepsilon h_i^2(t))\, dt < \infty, \qquad \text{for all } \varepsilon > 0, \qquad i = 1, 2 \tag{3.1}$$

is both necessary and sufficient for the weak convergence of V_n *to* W_0 *in* ρ_q, *where* $h_1(t) = t^{-\frac{1}{2}}q(t)$ *and* $h_2(t) = t^{-\frac{1}{2}}q(1-t)$.

PROOF. Represent the empirical process V_n in terms of a "partial sum" process S_n along the lines of Breiman [1] page 286. That is, let $Y_1, Y_2, \cdots$ be independent random variables with distribution given by (2.10). Set $Z_i = Y_1 + Y_2 + \cdots + Y_i$, $i \geqq 1$, $Z_0 = 0$. The process S_n is defined by

$$\begin{aligned} S_n(t) &= (Z_k - k)(n+1)^{-\frac{1}{2}}, \; k/(n+1) \\ &\leqq t < (k+1)/(n+1), \qquad k = 0, 1, \cdots, n \\ &= (Z_{n+1} - n - 1)(n+1)^{-\frac{1}{2}}, \qquad t = 1. \end{aligned}$$

Then

$$V_n \cong \frac{(n^2+n)^{\frac{1}{2}}}{Z_{n+1}} T_n + \frac{n^{\frac{1}{2}}(n+1)}{Z_{n+1}} (e_n(t) - t) \tag{3.2}$$

where

$$T_n = S_n(t) - tS_n(1) \tag{3.3}$$

and $e_n(t) = [(n+1)t]/(n+1)$ ($[x]$ denotes largest integer $\leqq x$) and "$\cong$" means the two processes have the same distribution.

Choose a version $\bar{S}_n$ of S_n ($\bar{S}_n$ has same distribution as S_n) defined on a common probability space with a Wiener process $\bar{W}$ such that $\rho(\bar{S}_n, \bar{W}) \to 0$ almost surely as $n \to \infty$. Let $\bar{T}_n$ be defined in terms of $\bar{S}_n$ by (3.3) and $\bar{V}_n$ be defined in terms of $\bar{T}_n$ by (3.2). Then $\rho(\bar{V}_n, \bar{W}_0) \to 0$ as $n \to \infty$ where $\bar{W}_0 = \bar{W}(t) - t\,\bar{W}(1)$. Suppose (3.1) holds. In order to show the weak convergence of V_n to W_0 in (D, ρ_q) it is sufficient to show that $\rho_q(\bar{V}_n, \bar{W}_0) \to 0$ in probability as $n \to \infty$. This in turn will follow from

$$\lim_{\theta\to 0} \limsup_{n\to\infty} P\{|V_n(t)| > \varepsilon q(t) \text{ some } 0 < t \leqq \theta\} = 0 \quad \text{for all } \varepsilon > 0 \tag{3.4}$$

since the corresponding result for $1 - \theta \leqq t < 1$ is obtained by replacing $q(t)$ by $q(1-t)$ in (3.4). Let $\alpha_n = (8 \log n/n)^{\frac{1}{2}}$ and apply Lemma 2.9 to get

$$\begin{aligned}
&P\{|V_n(t)| > \varepsilon q(t), \text{ some } t \text{ in } (0, \theta]\} \\
&\quad \leqq P\left\{\frac{(n^2+n)^{\frac{1}{2}}}{Z_{n+1}} |T_n| > \frac{\varepsilon}{2} q(t),\right. \\
&\qquad\qquad \left.\text{some } t \text{ in } (0, \theta], |Z_{n+1}/n - 1| \leqq \alpha_n\right\} \\
&\qquad + P\left\{\frac{n^{\frac{1}{2}}(n+1)}{Z_{n+1}} |e_n(t) - t| > \frac{\varepsilon}{2} q(t),\right. \\
&\qquad\qquad \left.\text{some } t \text{ in } (0, \theta], |Z_{n+1}/n - 1/ \leqq \alpha_n\right\} + O(n^{-\frac{1}{2}}) \\
&\quad \leqq P\left\{|T_n| > \frac{\varepsilon}{3} q(t), \text{ some } t \leqq \theta\right\} + O(n^{-\frac{1}{2}}), \qquad n \text{ large} \\
&\quad = P_3 + O(n^{-\frac{1}{2}}), \quad \text{say.}
\end{aligned} \tag{3.5}$$

Now bound P_3 from above;

$$P_3 \leqq P_4(\varepsilon/6) + P_5(\varepsilon/6) \tag{3.6}$$

where

$$P_4(\varepsilon) = P\{|S_n(t)| > \varepsilon q(t), \text{ some } t \text{ in } (0, \theta]\}$$

and

$$\begin{aligned}
P_5(\varepsilon) &= P\{t|S_n(1)| > \varepsilon q(t), \text{ some } t \text{ in } (0, \theta]\}. \\
&\leqq P\{|S_n(1)| > \varepsilon q(\theta)/\theta\} \\
&\to_n P\{|Z| > \varepsilon q(\theta)/\theta\}, \qquad \text{where } Z \sim N(0, 1) \\
&\to_\theta 0.
\end{aligned} \tag{3.7}$$

Choose $0 < \lambda < 1$ and let $N = N(\lambda, \theta)$ be the largest integer less than or equal to $-\log[(n+1)\theta]/\log(\lambda)$. Set $k_j = (n+1)\theta\lambda^j$, $j = 0, 1, \cdots, N$ and let $\bar{k}_j$ be the largest integer less than or equal to k_j, $j = 0, 1, \cdots, N$. Then $P_4(\varepsilon)$ is bounded above by

$$\textstyle\sum_{j=0}^{N} P\{|S_k| > \varepsilon(n+1)^{\frac{1}{2}} q(\theta\lambda^{j+1}) \text{ for some } k_{j+1} \leqq k \leqq k_j\} \tag{3.8}$$

where $S_k = Z_k - k$. Apply Lemma 2.7 and Chebyshev's inequality to get the jth term of (3.8) bounded above by

$$\tfrac{4}{3}P\{|S_{\bar{k}_j}| > \varepsilon(n+1)^{\frac{1}{2}}q(\theta\lambda^{j+1}) - 2(k_j - k_{j+1})^{\frac{1}{2}}\}, \tag{3.9}$$

which in turn for $\theta < \theta^*(\lambda)$ is bounded by

$$\tfrac{4}{3}P\{|S_{\bar{k}_j}| \geqq \bar{k}_j^{\frac{1}{2}}\bar{f}_k\} \tag{3.10}$$

where $\bar{f}_j = \varepsilon/2(k_{j+1}/\bar{k}_j)^{\frac{1}{2}}\, h(\theta\lambda^{j+1})$. Now apply the Berry–Esséen type bound of Nagaev [6] to get

$$P_4(\varepsilon) \leqq \tfrac{8}{3}\sum_{j=0}^{N}[1 - \Phi(f_j)] + L\sum_{j=0}^{N}\bar{k}_j^{-\frac{1}{2}}\bar{f}_j^{-3}, \tag{3.11}$$

where $f_j = \varepsilon/2(\theta\lambda^j)^{-\frac{1}{2}}q(\theta\lambda^{j+1})$, L is a constant which depends only on $E(Y_1^3)$ and Φ is the standard normal distribution function. The first term on the right side of (3.11) can be bounded above by a quantity which is independent of n and goes to 0 as $\theta \to 0$; for

$$\begin{aligned}\sum_{j=0}^{N}[1 - \Phi(f_j)] &\leqq K\sum_{j=0}^{\infty}\exp(-f_j^2/\lambda)\\ &\leqq -K(\log(\lambda))^{-1}\sum_{j=0}^{\infty}\int_{\theta\lambda^{j+2}}^{\theta\lambda^{j+1}} t^{-1}\exp((-\varepsilon\lambda)^2h^2(t)/2)\,dt\\ &= -K(\log(\lambda))^{-1}\int_0^{\theta\lambda} t^{-1}\exp((-\varepsilon\lambda)^2h^2(t)/2)\,dt\\ &\to 0 \quad \text{as} \quad \theta \to 0.\end{aligned} \tag{3.12}$$

The second term on the right side of (3.11) is bounded as follows:

$$\begin{aligned}\sum_{j=0}^{N}\bar{k}_j^{-\frac{1}{2}}\bar{f}_j^{-3} &\leqq \sum_{j=0}^{N}k_j^{-\frac{1}{2}}f_j^{-3}\\ &\leqq (\lambda\varepsilon^2)^{-\frac{3}{2}}\sum_{j=0}^{N-1}k_j^{-\frac{1}{2}}(h(\theta\lambda^{j+1}))^{-3}\\ &\quad + (\lambda\varepsilon^2)^{-\frac{3}{2}}(h(\theta\lambda^{N+1}))^{-3}, \qquad \text{since} \quad n\theta\lambda^N \geqq 1\\ &\leqq C_1(n+1)^{-\frac{1}{2}}b(\theta)\int_{(n+1)^{-1}}^{\theta}x^{-\frac{3}{2}}\,dx + C_2 b((n+1)^{-1})\end{aligned} \tag{3.13}$$

where $b(t) = \max(h^{-3}(S): 0 < S \leqq t)$. Performing the integration yields

$$\sum_{j=0}^{N}\bar{k}_j^{-\frac{1}{2}}f_j^{-3} \leqq C_1 b(\theta) + C_2 b((n+1)^{-1}) \tag{3.14}$$

so that from (3.14), (3.12) and (3.11) it follows that

$$\lim_{\theta\to 0}\limsup_{n\to\infty}P_4(\varepsilon) = 0 \qquad \text{for all} \quad \varepsilon > 0. \tag{3.15}$$

The sufficiency of the condition (3.1) follows from (3.4), (3.5), (3.6), (3.7) and (3.15).

If, on the other hand, V_n converges weakly to W_0 in (D, ρ_q), choose versions so that $\rho_q(V_n, W_0) \to 0$ almost surely as $n \to \infty$ for these versions. Then for any $\varepsilon > 0$

$$\begin{aligned}\rho_q(V_n, W_0) &\geqq \sup\{(V_n(t) - W_0(t))/q(t),\ t < \varepsilon^2/(n+1)\}\\ &\geqq \sup\{(-(nt)^{\frac{1}{2}}t^{\frac{1}{2}} - W_0(t))/q(t),\ t < \varepsilon^2/(n+1)\}\\ &\geqq \sup\{(-\varepsilon t^{\frac{1}{2}} - W_0(t))/q(t),\ t < \varepsilon^2/(n+1)\}.\end{aligned} \tag{3.16}$$

For large enough n we have almost surely that

$$-W_0(t) < \varepsilon t^{\frac{1}{2}}(1 + h_1(t)) \qquad \text{for all} \quad t < \varepsilon^2/(n+1). \tag{3.17}$$

This implies that $t^{\frac{1}{2}}\ (1 + h_1(t))$ is in the upper class of the Wiener process for every $\varepsilon > 0$. By Proposition 2.1

$$\int_0^1 t^{-1} \exp(-\varepsilon(1 + h_1)^2 \, dt < \infty$$

and by Lemma 1 of [2], $1 + h_1(t)$ and hence $h_1(t)$ approaches infinity as $t \to 0$. For large enough n then

$$(3.18) \qquad P\{|W_0(t)| > \varepsilon q(t),\ t < (n+1)^{-1}\} \leqq P\{\sup(|V_n(t) - W_0(t)|/q(t) : t < (n+1)^{-1}) > \varepsilon/2\}\,.$$

Thus $\varepsilon q(t)$ is in the upper class of the Wiener process for every $\varepsilon > 0$ and Proposition 2.1 shows that (3.1) holds for $i = 1$. The case $i = 2$ follows similarly by mapping $t \to 1 - t$ and replacing $q(t)$ by $q(1 - t)$.

REMARK 1. The inverse empirical distribution function is often defined by

$$F_n^{-1}(t) = \inf(s : F_n(s) \geqq t)\,.$$

If this version of F_n^{-1} is used in (1.4), Theorem 1 continues to hold provided V_n modified by being set to zero for t in $[0, n^{-1}]$ and t in $[1 - n^{-1}, 1]$.

REMARK 2. Let Q_1 denote the class of all functions q^* in D which are bounded below by a function q with the properties assumed in Theorem 1. Then V_n converges weakly to W_0 in (D, ρ_{q^*}) and (D, d_{q^*}) for all q^* in Q_1.

4. The empirical process. Theorem 1 applies to the empirical process U_n as well as to the process V_n. This is stated and proved in Theorem 2 below. Since in the proof of Theorem 2, U_n is replaced by a Poisson process rather than a partial sum process we need an analogue of Lemma 2.7 for the Poisson process.

LEMMA 4.1. *Let $X_1(t)$ denote a Poisson process with parameter* 1 *and all paths constant except for upward jumps of height* 1. *Let a, b and c be nonnegative constants with $c > 2(b - a)^{\frac{1}{2}}$. Then*

$$(4.2) \qquad P(|X_1(t) - t| \geqq c, \ \textit{some}\ \ a < t \leqq b) \leqq \tfrac{4}{3} P(|X_1(b) - b| \geqq c - 2(b - a)^{\frac{1}{2}})\,.$$

PROOF. The proof mimics that of Lemma 16 of [3] page 18. Namely, let $0 < c' < c$ and let

$$A_n = \{|X_1(j2^{-n}) - j2^{-n}| \geqq c' \ \text{ for some } \ j = [a2^n] + 1, \cdots, [b2^n]\}\,.$$

Apply Lemma 2.7 along with Chebyshev's inequality to get

$$P(A_n) \leqq \tfrac{4}{3} P\{|X_1([b2^n]2^{-n}) - [b2^n]2^{-n}| \geqq c' - 2(b - a)^{\frac{1}{2}}\} = \tfrac{4}{3} P(B_n)\,,$$

where $B_n = \{|X_1([b]^n]2^{-n}) - [b2^n]2^{-n}| \geqq c' - 2(b - a)^{\frac{1}{2}}\}$. Now A_n increases to A, say, and apart from a set of probability zero,

$$\{|X_1(t) - t| \geqq c, \ \text{ some } \ a < t \leqq b\} \subset A$$

so that

$$P\{|X_1(t) - t| \geqq c, \text{ some } a < t \leqq b\}$$
$$\leqq P(A) \leqq \tfrac{4}{3} \limsup_{n\to\infty} P(B_n) \leqq \tfrac{4}{3} P(|X_1(b) - b| \geqq c' - 2(b-a)^{\frac{1}{2}}) .$$

Let c' increase to c.

THEOREM 2. *Let $q(t)$ be as in the statement of Theorem* 1 *and let U_n be as in* (1.4). *Then* (3.1) *is both a necessary and sufficient condition for the weak convergence of U_n to W_0 in (D, ρ_q).*

PROOF. Suppose (3.1) holds. Let $\bar{U}_n$ and $\bar{W}_0$ be versions of U_n and W_0 defined on a common probability space such that $\rho(\bar{U}_n, \bar{W}_0) \to 0$ almost surely as $n \to \infty$. It is sufficient to show that $\rho_q(\bar{U}_n, W_0) \to 0$ in probability as $n \to \infty$, which in turn will follow from

$$\lim_{\theta\downarrow 0} \limsup_{n\to\infty} P\{U_n(t) > \varepsilon q(t), \text{ some } 0 < t < \theta\} = 0 \quad \text{for every } \varepsilon > 0 . \tag{4.3}$$

The arguments of Chibisov in [2] (Lemmas 3 and 4) show that in place of (4.3) it is sufficient to establish

$$\lim_{\theta\downarrow 0} \limsup_{n\to\infty} P\{|X_n(t) - nt| > \varepsilon n^{\frac{1}{2}} q(t), \text{ some } n^{-1} \leqq t < \theta\} = 0 \tag{4.4}$$

where $X_n(t)$ denotes a Poisson process with parameter n. As in Theorem 1 choose $0 < \lambda < 1$ and let $N = N(\theta, \lambda)$ be the largest integer less than or equal to $-\log(n\theta)/\log(\lambda)$ and $k_j = n\theta\lambda^j, j = 0, 1, \cdots, N$. Now bound the probability in (4.4) by

$$\textstyle\sum_{j=0}^N P\{|X_n(t) - nt| > \varepsilon n^{\frac{1}{2}} q(\theta\lambda^{j+1}), \text{ some } \theta\lambda^{j+1} < t \leqq \theta\lambda^j\} . \tag{4.5}$$

Replace nt by t so that (4.5) becomes

$$\textstyle\sum_{j=0}^N P\{|X_1(t) - t| > \varepsilon n^{\frac{1}{2}} q(\theta\lambda^{j+1}), \text{ some } k_{j+1} < t \leqq k_j\} . \tag{4.6}$$

Apply Lemma (4.1) to majorise (4.6) by

$$\tfrac{4}{3} \textstyle\sum_{j=0}^N P\{|X_1(k_j) - k_j| > \varepsilon n^{\frac{1}{2}} q(\theta\lambda^{j+1}) - 2/k_j(1-\lambda)\} , \tag{4.7}$$

which in turn for small enough θ is bounded above by

$$\tfrac{4}{3} \textstyle\sum_{j=0}^N P\left\{|X_1(k_j) - k_j| > \frac{\varepsilon}{2} k_j^{\frac{1}{2}} h(\theta\lambda^{j+1})\right\} . \tag{4.8}$$

Apply the bound (8) of [2] page 151 to bound (4.8) by

$$\tfrac{8}{3} \textstyle\sum_{j=0}^N \exp[-\varepsilon^2(1-\lambda)^2 h^2(\theta\lambda^{j+1})/8] + \tfrac{8}{3} \textstyle\sum_{j=0}^N \exp[-c_\lambda \varepsilon n^{\frac{1}{2}} q(\theta\lambda^{j+1})/\lambda 4] \tag{4.9}$$

where c_λ is positive constant depending on λ.

The same argument used in (3.15) of Theorem 1 bounds the first term in (4.9) by a quantity independent of n which goes to zero as $\theta \to 0$. Consider the second sum in (4.9). Let $b(\theta) = \min(h(t): 0 < t \leqq \theta)$. Then

$$(4.10) \qquad \sum_{j=0}^{N} \exp[-c_\lambda \varepsilon n^{\frac{1}{2}} q(\theta\lambda^{j+1})/\lambda 4] \leqq \sum_{j=0}^{N} \exp[-c_\lambda \varepsilon (n\theta\lambda^{N+1})^{\frac{1}{2}} b(\theta)\lambda^{-N+j}/\lambda 4] .$$

Use the fact that $n\theta\lambda^{N+1} \geqq \lambda$ and write the terms on the right side of (4.10) in reverse order so that (4.10) is majorised by

$$(4.11) \qquad \sum_{j=0}^{\infty} \exp[-\varepsilon c_\lambda' b(\theta)\lambda^{-j}] \to 0 \qquad \text{as } \theta \to 0 .$$

Sufficiency is thus proved. Necessity follows by an argument similar to that of Theorem 1, except that $(n+1)^{-1}$ is replaced by U_{1n}, the first order statistic. The deviation of U_{1n} from $(n+1)^{-1}$ is then bounded by a Chebyshev type inequality. The details are omitted.

Theorem 2 is more general than Theorem 3 of [2], since for any function $q(t)$ regularly varying of order $0 < \alpha \leqq \frac{1}{2}$ at $t = 0$, it is easily shown there exists $q'(t)$, monotone increasing near zero, such that $q(t) \sim q'(t)$ as $t \to 0$. By Theorem 2, U_n converges weakly to W_0 in (D, ρ_q) (and hence (D, d_q)) for all functions q in the class Q_1.

5. Partial sum processes. The proof of (3.18) Theorem 1 shows that if $Y_1, Y_2, \cdots$, are a sequence of independent, identically distributed random variables with $E(Y_1) = 0$, $E(Y_1^2) = 1$ and $E|Y_1|^3 < \infty$ then for the partial sum process

$$S_n(t) = Z_k n^{-\frac{1}{2}} , \qquad k/n \leqq t < (k+1)/n$$
$$S_n(1) = Z_n n^{-\frac{1}{2}}$$

where $Z_0 = 0$ and $Z_k = \sum_{j=1}^{k} Y_j$, one may state the following theorem.

THEOREM 3. *A necessary and sufficient condition for the weak convergence of $S_n(t)$ to a standard Wiener process $W(t)$ in (D, ρ_q) for a continuous, nonnegative function $q(t)$, bounded away from zero on $(\gamma, 1)$ and non-decreasing on $(0, \gamma)$ for some $\varepsilon > 0$ is the convergence for every $\varepsilon > 0$ of the integral*

$$(5.1) \qquad \int_0^1 t^{-1} \exp(-\varepsilon q^2(t)/t)\, dt .$$

The third moment assumption can be dropped and the theorem continues to hold. A truncation argument along the lines of Robbins and Siegmund [8] in Lemma 5 is needed for the proof which is omitted here.

The arguments of Pyke and Shorack, using Theorem 1 in place of Theorem 2.1 of [7], establish weak convergence of the two sample empirical process L_N (modified at 0) to W_0 in (D, ρ_q) for q in Q_1, a larger class of functions than their class Q.

REFERENCES

[1] BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading.

[2] CHIBISOV, D. M. (1964). Some theorems on the limiting behavior of empirical distribution functions. *Selected Translations in Math. Statist. and Probability* **6** 147-156.

[3] FREEDMAN, D. (1970). *Brownian motion and diffusion*. Holden-Day, San Francisco.

[4] Itô, K. and McKean, H. (1965). *Diffusion processes and their sample paths*. Academic Press, New York.

[5] Lévy, P. (1954). Le mouvement brownien. *Mémor. Sci. Math.* **126** Gauthier-Villars, Paris.

[6] Nagaev, S. V. (1965). Some limit theorems for large deviations. *Theor. Probability Appl.* **10** 214–235.

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

[8] Robbins, H. and Siegmund, D. (1970). Boundary crossing probabilities for the Wiener process and partial sums. *Ann. Math. Statist.* **41** 1410–1429.

[9] Rosenkrantz, W. A. (1969). A rate of convergence for the von Mises statistic. *Trans. Amer. Math. Soc.* **139** 329–337.

Statistics Center
Rutgers University
New Brunswick, New Jersey 08903