AN ASYMPTOTIC REPRESENTATION OF THE SAMPLE DISTRIBUTION FUNCTION

BY DAVID R. BRILLINGER

Communicated by David Blackwell, January 10, 1969

1. Let $X_1, \cdots, X_n$ be independent observations from the uniform distribution on $[0, 1]$. Let $F_n(x)$ = the proportion of the $X_j \leqq x$. We will prove

THEOREM. *There is a random function* $\{G_n(x); 0 \leqq x \leqq 1\}$, *with the same distribution as* $\{F_n(x); 0 \leqq x \leqq 1\}$ *for each* n, *and there is a Brownian motion* W, *such that for the Brownian* $B(x) = n^{-1/2}W(nx)$

$$\sup_{0 \leqq x \leqq 1} \left| n^{1/2}[G_n(x) - x] - [B(x) - xB(1)] \right| = O[n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}] \tag{1}$$

almost surely as $n \to \infty$.

This theorem is of use in the investigation of the asymptotic behavior of functionals of $\{F_n(x); 0 \leqq x \leqq 1\}$, especially functionals dependent on n.

2. We construct $G_n(x)$ as follows; let $Y_1, Y_2, \cdots$ be independent exponential variables with mean 1. Let $S(k) = Y_1 + \cdots + Y_k$, $k = 1, 2, \cdots$ and let $S(0) = 0$. Set

$$G_n(x) = k/n \qquad \text{if } S(k)/S(n+1) \leqq x < S(k+1)/S(n+1).$$

This $\{G_n(x); 0 \leqq x \leqq 1\}$ has the same distribution as $\{F_n(x); 0 \leqq x \leqq 1\}$ for each n. We now record a series of lemmas.

LEMMA 1. *There is a Brownian motion* W *such that*

$$\sup_{1 \leqq k \leqq n} \left| k - S(k) - W(k) \right| = O[n^{1/4}(\log n)^{1/2}(\log \log n)^{1/4}] \tag{2}$$

almost surely as $n \to \infty$.

PROOF. This result is deducible from Theorem 1.5 of Strassen [8].

LEMMA 2. *Almost surely as* $n \to \infty$

$$\sup_{0 \leqq x \leqq 1} \left| S(nG_n(x)) - xS(n+1) \right| = O[n^{1/4}]. \tag{3}$$

PROOF.

$$
\begin{aligned}
&| S(nG_n(x)) - xS(n+1) | \\
&\qquad = | S(k) - xS(n+1) | \qquad \text{if } S(k) \leqq xS(n+1) < S(k+1) \\
&\qquad \leqq S(k+1) - S(k) \qquad \text{if } S(k) \leqq xS(n+1) < S(k+1). \\
&\qquad \leqq \max_{1 \leqq k \leqq n} Y_k
\end{aligned}
$$

and one sees, by elementary calculations, that this last $= O[n^{1/4}]$ almost surely as $n \to \infty$.

LEMMA 3. *Almost surely as* $n \to \infty$

$$
\sup_{0 \leqq x \leqq 1} | nG_n(x) - S(nG_n(x)) - W(nG_n(x)) | = O[n^{1/4}(\log n)^{1/2}(\log \log n)^{1/4}]. \tag{4}
$$

PROOF.

$$
\begin{aligned}
&| nG_n(x) - S(nG_n(x)) - W(nG_n(x)) | \\
&\qquad = | k - S(k) - W(k) | \qquad \text{if } S(k) \leqq xS(n+1) < S(k+1) \\
&\qquad \leqq \sup_{1 \leqq k \leqq n} | k - S(k) - W(k) |
\end{aligned}
$$

and (4) follows from (2).

LEMMA 4. *Almost surely as* $n \to \infty$

$$
\sup_{0 \leqq x \leqq 1} | G_n(x) - x | = O[n^{-1/2}(\log \log n)^{1/2}]. \tag{5}
$$

PROOF. See Theorem 2* in Chung [3].

We next define the Brownian motion B by $B(x) = n^{-1/2}W(nx)$ and then have

LEMMA 5. *Almost surely as* $n \to \infty$

$$
\sup_{0 \leqq x \leqq 1} | B(G_n(x)) - B(x) | = O[n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}]. \tag{6}
$$

PROOF. (6) follows from (5) and Lévy's Hölder condition for Brownian motion (see Itô and McKean [4]) extended to apply to the interval $[0, n]$.

PROOF OF THEOREM. Up to an error term

$$
O[n^{-1/4}(\log n)^{1/2}(\log \log n)^{1/4}],
$$

that is uniform in x, almost surely as $n \to \infty$

$$\begin{aligned} n^{1/2}G_n(x) &= n^{-1/2}S(nG_n(x)) + n^{-1/2}W(nG_n(x)) \qquad \text{from (4)},\\ &= n^{-1/2}xS(n+1) + B(G_n(x)) \qquad \text{from (3)},\\ &= n^{-1/2}x[(n+1) - W(n+1)] + B(x) \qquad \text{from (2) and (6)},\\ &= n^{1/2}x - xB(1) + B(x), \end{aligned}$$

giving (1).

3. We may use the probability integral transformation to deduce a representation of the sample distribution function of observations from any continuous distribution. The results of Rosenkrantz [7] may be adapted to obtain rates of convergence in distribution for certain functionals of $F_n(x)$. The announcement of Kiefer [5] suggests that the error term in (1) may be best possible.

Bickel [1] and Billingsley [2] consider the weak convergence of the process $n^{1/2}[F_n(x)-x]$ to $W(x)-xW(1)$. Pyke and Root [6] let the distribution of Y depend on n and then prove

$$\sup_{0\leqq x\leqq 1} \left| n^{1/2}[G_n(x) - x] - [W(x) - xW(1)] \right| = o(1)$$

almost surely as $n\rightarrow\infty$. I would like to thank Professor Pyke for the remark that B, as constructed above, depends on n.

References

1. P. J. Bickel, *Some contributions to the theory of order statistics*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), vol. I: *Statistics*, Univ. of California Press, Berkeley, Calif., 1967, pp. 575–591.

2. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968.

3. K. L. Chung, *An estimate concerning the Kolmogorov limit distribution*, Trans. Amer. Math. Soc. **67** (1949), 36–50.

4. K. Itô and H. P. McKean, Jr., *Diffusion processes and their sample paths*, Springer-Verlag, Berlin, 1965.

5. J. Kiefer, *The deviations in Skorokhod-Strassen approximation*, Notices Amer. Math. Soc. **15** (1968), 936–937.

6. R. Pyke and D. Root, *An application of stopping times to obtain weak convergence*, Technical Report No. 16, University of Washington, Seattle, Wash., 1968.

7. W. A. Rosenkrantz, *On rates of convergence for the invariance principle*, Trans. Amer. Math. Soc. **129** (1967), 542–552.

8. V. Strassen, *Almost sure behavior of sums of independent random variables and martingales*, Proc. Fifth Berkeley Sympos. Math. Statist. and Probability (Berkeley, Calif., 1965/66), vol. II: *Contributions to probability theory*, Part 1, Univ. of California Press, Berkeley, Calif., 1967, pp. 315–343.

The London School of Economics and Political Science,
Aldwych, England