AN INVARIANCE PRINCIPLE FOR THE EMPIRICAL PROCESS WITH RANDOM SAMPLE SIZE

BY M. CSÖRGÖ[1] AND S. CSÖRGÖ

Communicated by David Blackwell, January 19, 1970

Let $C=C[0, 1]$ be the space of continuous functions on $[0, 1]$ with the uniform topology, that is the distance between two points x and y (two functions x and y of $t\in[0, 1]$) is defined by

$$\rho(x, y) = \sup_t |x(t) - y(t)| .$$

Let $\mathcal{B}$ be the σ-field of Borel sets of C.

Let $(\Omega, \mathcal{A}, P)$ be some probability space and W be the Wiener measure on $(C, \mathcal{B})$ with the corresponding Wiener process $\{W_t(\omega):0\leqq t\leqq 1\}$, $\omega\in\Omega$; that is W_t has values in C and is specified by $E(W_t)=0$ and $E(W_sW_t)=s$ if $s\leqq t$. Let W^0 be the Gaussian measure on $(C, \mathcal{B})$ constructed by setting $W_t^0=W_t-tW_1$. Then $W_t^0\in C$, $E(W_t^0)=0$ and $E(W_s^0W_t^0)=s(1-t)$ if $s\leqq t$. Also $W_0^0=W_1^0=0$ with probability 1 and $\{W_t^0:0\leqq t\leqq 1\}$ is called the tied down Wiener process or the Brownian bridge.

Let $S_n=\xi_1+\cdots+\xi_n$, $S_0=0$, $n=1, 2, \cdots$ be the partial sum sequence of random variables $\{\xi_n\}$ defined on $(\Omega, \mathcal{A}, P)$. Define a random element X_n of C by

$$X_n(t, \omega) = W_n(t, \omega) + (nt - [nt])\xi_{[nt]+1}(\omega)/n^{1/2} - tW_n(1, \omega) \tag{1}$$

where $W_n(t, \omega)=S_{[nt]}(\omega)/n^{1/2}$. The following theorem is an immediate consequence of L. Breiman's analysis of §§13.5 and 13.6 in his book [3].

THEOREM B. *Suppose the random variables $\xi_1, \xi_2, \cdots$ are independent and identically distributed with mean zero and variance* 1. *Then the random functions X_n defined by* (1) *satisfy*

$$X_n \xrightarrow{\mathfrak{D}} W^0. \tag{2}$$

Here (2), and also similar relations later on, are interpreted in accordance with (4.5) and (4.7) of Billingsley's book [2], depending on

AMS Subject Classifications. Primary 6030, 6040; Secondary 6270, 6271.

Key Words and Phrases. Invariance principle, weak convergence, randomly selected partial sums, empirical process, random-sample-size Kolmogorov-Smirnov statistics.

[1] Work done while the author was a Canada Council Fellow in 1969–1970 at the University of Vienna.

whether W^0 is construed as a random function or as a measure in the spirit of [2, p. 65]; the meaning is the same for the two interpretations. Since $h(x) = \sup_t |x(t)|$ with $x(t) = w(t) - tw(1)$ is a continuous function on C in the sup-norm metric, (2) implies

$$\sup_t |X_n(t)| \xrightarrow{\mathfrak{D}} \sup_t |W_t^0|,$$

an invariance principle, as statements like this are often called. Similarly,

$$\sup_t X_n(t) \xrightarrow{\mathfrak{D}} \sup_t W_t^0, \qquad \inf_t X_n(t) \xrightarrow{\mathfrak{D}} \inf_t W_t^0.$$

For each n, let ν_n be a positive-integer-valued random variable defined on the same probability space as the ξ_n. Define X_n, a random element of C, as in (1), and Y_n, another random element of C, by

$$Y_n(t, \omega) = X_{\nu_n(\omega)}(t, \omega). \tag{3}$$

THEOREM 1. *Suppose the random variables* $\xi_1, \xi_2, \cdots$ *are independent and identically distributed with mean zero and variance* 1. *If*

$$\nu_n/n \xrightarrow{P} \nu, \tag{4}$$

where ν *is a positive random variable, and*

$$\xi_{[\nu_n t]}(\omega)/(\nu_n(\omega))^{1/2} \xrightarrow{P} 0, \qquad \textit{for every fixed } t, \tag{5}$$

then the random functions Y_n *defined by* (3) *satisfy*

$$Y_n \xrightarrow{\mathfrak{D}} W^0. \tag{6}$$

COROLLARY 1. *Under the same assumptions as in Theorem* 1 (6) *implies*

$$\sup_t |Y_n(t)| \xrightarrow{\mathfrak{D}} \sup_t |W_t^0|,$$

$$\sup_t Y_n(t) \xrightarrow{\mathfrak{D}} \sup_t W_t^0,$$

$$\inf_t Y_n(t) \xrightarrow{\mathfrak{D}} \inf_t W_t^0.$$

REMARK 1. Let D be the space D of Chapter 3 of P. Billingsley's book [2]. Define random elements X_n^*, Y_n^* of D by

$$X_n^*(t, \omega) = W_n(t, \omega) - tW_n(1, \omega), \tag{7}$$

$$Y_n^*(t, \omega) = X_{\nu_n(\omega)}^*(t, \omega) \tag{8}$$

with $W_n(t, \omega)$ as in (1). Then Theorem B holds for X_n^* of (7) and, omitting condition (5), Theorem 1 holds for Y_n^* of (8). Also, in defining Y_n of (3) and Y_n^* of (8) it is not essential that the random variables $\{\xi_n\}$ involved should be independent and identically distributed with unit variance. We have stated Theorem 1 here for random elements of C and for independent identically distributed random variables having unit variance only because it is, as will be shown later, directly applicable in this form to prove the random-sample-size Kolmogorov-Smirnov theorems. More general versions of Theorem 1 and detailed proofs of them will appear in [4]. We also note that for Y_n of (3) one postulates (5), for it is not true in general that $\xi_{[nt]}/n^{1/2} \xrightarrow{P} 0$ and (4) imply (5).

For the proof of Theorem 1 we use Theorem B, Theorems 7.7, 8.1, 8.2 of P. Billingsley's book [2] and results of A. Rényi [7] and J. Mogyoródi [5]. First we show that for a single time point s $\{X_n^*(s)\}$ is mixing with the normal distribution function $N(0, s(1-s))$ in the sense of A. Rényi's definition of mixing sequences of events [7] and that it also satisfies the tightness condition of F. J. Anscombe [1]. Then, using Theorem B, Theorem 7.7 of [2] and Theorem 2 of [5], we show that the finite-dimensional distributions of Y_n of (3) converge to those of W^0. Next it is verified that the sequence $\{Y_n\}$ is tight in the sense of Theorem 8.2 of [2] and then Theorem 1 follows from Theorem 8.1 of [2]. Details of this proof will appear in [4].

Let $U_1, \cdots, U_n$ be independent random variables uniformly distributed on $[0, 1]$. The order statistics are defined as follows: $U_1^{(n)}$ is the smallest, and so forth; $U_n^{(n)}$ is the largest. Let

$$F_n(t) = (\text{the number of the } U_i \leq t)/n, \qquad t \in [0, 1].$$

Define the Kolmogorov-Smirnov statistics

$$D_n^+ = n^{1/2} \sup_t (F_n(t) - t) = n^{1/2} \max_{k \leq n} (k/n - U_k^{(n)}),$$

$$D_n^- = n^{1/2} \inf_t (F_n(t) - t) = n^{1/2} \min_{k \leq n} (k/n - U_k^{(n)}),$$

$$D_n = n^{1/2} \sup_t \left| t - F_n(t) \right| = n^{1/2} \max_{k \leq n} \left| U_k^{(n)} - k/n \right|,$$

and the random-sample-size Kolmogorov-Smirnov statistics $\Delta_n^+ = D_{\nu_n}^+$, $\Delta_n^- = D_{\nu_n}^-$, $\Delta_n = D_{\nu_n}$.

THEOREM 2. *Under condition* (4) *of Theorem* 1 *we have*

$$\Delta_n^+ \xrightarrow{\mathfrak{D}} \sup_t W_t^0, \qquad \Delta_n^- \xrightarrow{\mathfrak{D}} \inf_t W_t^0, \qquad \Delta_n \xrightarrow{\mathfrak{D}} \sup_t | W_t^0 |.$$

PROOF OF THEOREM 2. Let $S(n)=\zeta_1+\cdots+\zeta_n$, $n=1, 2, \cdots$ be the partial sum sequence of independent exponential random variables $\{\zeta_n\}$ with mean 1. L. Breiman [3, §13.6] obtained the following representation of D_n

$$\begin{aligned} D_n &\overset{\mathfrak{D}}{=} n^{1/2} \max_{k\leq n} \left| \frac{S(k)}{S(n+1)} - \frac{k}{n} \right| \\ &\overset{\mathfrak{D}}{=} \frac{n}{S(n+1)} \max_{k\leq n} \left| \frac{S(k)-k}{n^{1/2}} - \frac{k}{n}\, \frac{S(n+1)-n}{n^{1/2}} \right|, \end{aligned} \tag{9}$$

with analogous expressions for D_n^+ and D_n^-. Here $\overset{\mathfrak{D}}{=}$ means that the random variables in question have the same distribution. Put $\xi_n=\zeta_n-1$, $S_k=S(k)-k$ and $W_n(t,\omega)=S_{[nt]}(\omega)/n^{1/2}$. Then

$$\begin{aligned} D_n &\overset{\mathfrak{D}}{=} \sup_t | X_n^*(t,\omega) |, \qquad \text{for } n \text{ large}, \\ &\overset{\mathfrak{D}}{=} \sup_t | X_n(t,\omega) |, \qquad \text{for } n \text{ large}, \end{aligned} \tag{10}$$

where X_n^* and X_n are respectively defined in terms of the above ξ_n and W_n via (7) and (1). Analogous asymptotic representations hold for D_n^+ and D_n^-. The first asymptotic representation of (10) for D_n is true because $E(\zeta_n)=\sigma^2(\zeta_n)=1$ and hence $n/S(n+1)\xrightarrow{\text{a.s.}}1$ and $\zeta_{n+1}/n^{1/2}\xrightarrow{P}0$, while the second asymptotic representation of (10) is the consequence of $\xi_{[nt]+1}/n^{1/2}\xrightarrow{P}0$ uniformly in t. The X_n of (10) satisfy the conditions of Theorem B and the usual Kolmogorov-Smirnov theorems follow. For Δ_n we have (9) with n replaced by ν_n on both sides. Now we show

$$\begin{aligned} \Delta_n &\overset{\mathfrak{D}}{=} \sup_t | Y_n^*(t,\omega) |, \qquad \text{for } n \text{ large}, \\ &\overset{\mathfrak{D}}{=} \sup_t | Y_n(t,\omega) |, \qquad \text{for } n \text{ large}, \end{aligned} \tag{11}$$

where Y_n^* and Y_n are respectively defined in terms of the above ξ_n and W_n via (8) and (3); we also have the analogous asymptotic expressions for Δ_n^+ and Δ_n^-. It is true in general that if $\{Z_n\}$ is a sequence of random variables such that $Z_n\xrightarrow{\text{a.s.}}Z$ and $\{\nu_n\}$ is a sequence of positive-integer-valued random variables such that $\nu_n \xrightarrow{P} +\infty$, then $Z_{\nu_n} \xrightarrow{P} Z$. Now condition (4) of Theorem 1 implies $\nu_n \xrightarrow{P} +\infty$ and we have $n/S(n+1) \xrightarrow{\text{a.s.}} 1$. Consequently, $\nu_n/S(\nu_n+1) \xrightarrow{P} 1$. Using the fact that the ζ_n are exponential random variables with mean 1 and that $\nu_n \xrightarrow{P} +\infty$, it can be easily shown that $\zeta_{\nu_n+1}/\nu_n^{1/2}$ and $\xi_{[\nu_n t]+1}/\nu_n^{1/2}$ both converge in probability to zero, the latter one uniformly in t. Hence both asymptotic representations of (11) are true. Also, given condition (4), the Y_n of (11) satisfy the conditions of Theorem 1 and hence $Y_n \xrightarrow{\mathfrak{D}} W^0$. The statements of Theorem 2 now follow from Corollary 1.

REMARK 2. Theorem 2 with $\nu=1$ in (4) was proved by R. Pyke [6] in an interesting and different way, utilizing results about stochastic processes with two-dimensional parameter sets. We should also note that proving appropriate versions of Theorem 1, random-sample-size versions of the Kolmogorov-Smirnov theorems with weight functions like

$$f(t) = 1/t, \qquad 1/(1-t) \quad \text{and} \quad 1/[t(1-t)]^{1/2},$$

which are important in applications, can also be proved in a similar way as well as two or more-sample random-sample-size versions. Statements and proofs for these results will also appear in [4].

REFERENCES

1. F. J. Anscombe, *Large-sample theory of sequential estimation*, Proc. Cambridge Philos. Soc. **48** (1952), 600–607. MR **14**, 487.

2. P. Billingsley, *Convergence of probability measures*, Wiley, New York, 1968. MR **38** #1718.

3. L. Breiman, *Probability*, Addison-Wesley, Reading, Mass., 1968. MR **37** #4841.

4. M. Csörgö and S. Csörgö, *On weak convergence of randomly selected partial sums* (to appear).

5. J. Mogyoródi, *Limit distributions for sequences of random variables with random indices*, Trans. Fourth Prague Conf. on Information Theory, Statistical Decision Functions, Random Processes, (Prague, 1965) Academia, Prague, 1967, pp. 463–470. MR **36** #935.

6. R. Pyke, *The weak convergence of the empirical process with random sample size*, Proc. Cambridge Philos. Soc. **64** (1968), 155–160. MR **36** #3402.

7. A. Rényi, *On mixing sequences of sets*, Acta. Math. Acad. Sci. Hungar. **9** (1958), 215–228. MR **20** #4623.

McGILL UNIVERSITY, MONTREAL, QUEBEC, CANADA

JÓZSEF ATTILA UNIVERSITY, SZEGED, HUNGARY

MATHEMATISCHES INSTITUT DER UNIVERSITAT, WIEN, AUSTRIA