AN INVARIANCE PRINCIPLE FOR THE EMPIRICAL PROCESS WITH RANDOM SAMPLE SIZE

BY M. CSÖRGŐ AND S. CSÖRGŐ

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 $\sigma$-field of Borel sets of $C$. Let $(\Omega, \mathcal{A}, P)$ be some probability space and $W$ be the Wiener measure on $(C, \mathcal{A})$ with the corresponding Wiener process \{\$W_t(\omega) : 0 \leq t \leq 1\$, $\omega \in \Omega$; that is $W_t$ has values in $C$ and is specified by $E(W_t) = 0$ and $E(W_s W_t) = s$ if $s \leq t$. Let $W^0$ be the Gaussian measure on $(C, \mathcal{A})$ constructed by setting $W^0_t = W_t - tW_1$. Then $W_t^0 \in C$, $E(W_0^0) = 0$ and $E(W_s^0 W_t^0) = s(1 - t)$ if $s \leq t$. Also $W_0^0 = W_1^0 = 0$ with probability 1 and \{\$W_t^0 : 0 \leq t \leq 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

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

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

$$(2) \quad X_n \overset{D}{\rightarrow} W^0.$$

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

$$\sup_t \left| X_n(t) \right| \overset{\mathcal{D}}{\longrightarrow} \sup_t \left| W^0_t \right|,$$

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

$$\sup_t X_n(t) \overset{\mathcal{D}}{\longrightarrow} \sup_t W^0_t, \quad \inf_t X_n(t) \overset{\mathcal{D}}{\longrightarrow} \inf_t W^0_t.$$

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

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

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

$$\frac{\nu_n}{n} \overset{P}{\longrightarrow} \nu,$$

where $\nu$ is a positive random variable, and

$$\xi_{\nu_n(t)}(\omega)/\nu_n(\omega) \overset{P}{\rightarrow} 0, \quad \text{for every fixed } t,$$

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

$$Y_n \overset{\mathcal{D}}{\longrightarrow} W^0.$$

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

$$\sup_t \left| Y_n(t) \right| \overset{\mathcal{D}}{\longrightarrow} \sup_t \left| W^0_t \right|,$$

$$\sup_t Y_n(t) \overset{\mathcal{D}}{\longrightarrow} \sup_t W^0_t,$$

$$\inf_t Y_n(t) \overset{\mathcal{D}}{\longrightarrow} \inf_t W^0_t.$$

**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
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 $\xi_{[n t]}/n^{1/2}$ implies (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: $U_{[1]}$ is the smallest, and so forth; $U_{[n]}$ is the largest. Let

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

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 $\Delta_n^+ = D_n^+, \Delta_n^- = D_n^-, \Delta_n = D_n$. 

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
**THEOREM 2.** Under condition (4) of Theorem 1 we have

\[ \Delta_n^+ \overset{\mathcal{D}}{\to} \sup_t W_t^0, \quad \Delta_n^- \overset{\mathcal{D}}{\to} \inf_t W_t^0, \quad \Delta_n \overset{\mathcal{D}}{\to} \sup_t |W_t^0|. \]

**Proof of Theorem 2.** Let \( S(n) = \xi_1 + \cdots + \xi_n, \) \( n = 1, 2, \cdots \) be the partial sum sequence of independent exponential random variables \( \{\xi_n\} \) with mean 1. L. Breiman [3, §13.6] obtained the following representation of \( D_n \)

\[ D_n = n^{1/2} \max_{k \leq n} \left| \frac{S(k)}{S(n+1)} - \frac{k}{n} \right|, \]

(9)

with analogous expressions for \( D_n^+ \) and \( D_n^- \). Here \( \mathcal{D} \) means that the random variables in question have the same distribution. Put \( \xi_n = \xi_n - 1, S_k = S(k) - k \) and \( W_n(t, \omega) = S_{\lfloor nt \rfloor} - S_{\lfloor nt \rfloor} \). Then

\[ D_n = \sup_t \left| X_n^*(t, \omega) \right|, \quad \text{for } n \text{ large}, \]

(10)

\[ D_n = \sup_t \left| X_n(t, \omega) \right|, \quad \text{for } n \text{ large}, \]

where \( X_n^* \) and \( X_n \) are respectively defined in terms of the above \( \xi_n \) 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 \( E(\xi_n) = \sigma^2(\xi_n) = 1 \) and hence \( n/S(n+1) \rightarrow 1 \) and \( \xi_{n+1}/n^{1/2} \rightarrow 0 \), while the second asymptotic representation of (10) is the consequence of \( \xi_{\lfloor nt \rfloor}/n^{1/2} \rightarrow 0 \) uniformly in \( t \). The \( X_n \) of (10) satisfy the conditions of Theorem B and the usual Kolmogorov-Smirnov theorems follow. For \( \Delta_n \) we have (9) with \( n \) replaced by \( \nu_n \) on both sides. Now we show

\[ \Delta_n = \sup_t \left| Y_n^*(t, \omega) \right|, \quad \text{for } n \text{ large}, \]

(11)

\[ \Delta_n = \sup_t \left| Y_n(t, \omega) \right|, \quad \text{for } n \text{ large}, \]

where \( Y_n^* \) and \( Y_n \) are respectively defined in terms of the above \( \xi_n \) and \( W_n \) via (8) and (3); we also have the analogous asymptotic expressions for \( \Delta_n^+ \) and \( \Delta_n^- \). It is true in general that if \( \{Z_n\} \) is a sequence of random variables such that \( Z_n \overset{\text{b.a.}}{\rightarrow} Z \) and \( \{\nu_n\} \) is a sequence of
positive-integer-valued random variables such that \( \nu_n \xrightarrow{P} + \infty \), then
\[ Z_n \xrightarrow{P} 1. \]
Now condition (4) of Theorem 1 implies \( \nu_n \xrightarrow{P} 1 \) and we have
\[ n/S(n+1) \xrightarrow{a.s.} 1. \]
Consequently, \( \nu_n/S(\nu_n+1) \xrightarrow{P} 1 \). Using the fact that the \( \xi_n \) are exponential random variables with mean 1 and that
\[ \nu_n \xrightarrow{P} + \infty, \]
it can be easily shown that \( \xi_{n+1}/\nu_n^{1/2} \) and \( \xi_{[\nu_n]}+1/\nu_n^{1/2} \) 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
\[ Y_n \approx W^0. \]
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
\[ f(t) = 1/t, \quad 1/(1 - t) \quad \text{and} \quad 1/[\mu(1 - t)]^{1/2}, \]
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


McGill University, Montreal, Quebec, Canada

József Attila University, Szeged, Hungary

Mathematisches Institut der Universität, Wien, Austria