INVARIANCE PRINCIPLE FOR SUMS OF INDEPENDENT RANDOM VARIABLES ON A BINARY TREE

By L. G. Gorostiza and A. R. Moncayo

Introduction

Consider a tree and a collection of independent and identically distributed random variables indexed by its nodes. For each realization of the random variables, and each branch of length n of the tree, there is a Donsker path associated to their partial sums on the branch. A stochastic process is defined by choosing at random one of the branches of length n and taking the corresponding Donsker path. We are interested in proving functional central limit theorems for such processes, as $n \to \infty$, that hold for almost all realizations of the random variables.

For the partial sums of length n on certain deterministic trees, and on Galton-Watson trees, central limit theorems holding almost surely have been proved by Stam [10], and by Joffe and Moncayo [4]. Previous and related work has been done by Harris [2], Kharlamov [5], Kolmogorov [6], and Ney [8, 9].

Joffe and Moncayo first tested their ideas on a binary tree [3]. Following them, we also try first on a binary tree, and in this paper we prove an invariance principle for this case. This result of course implies that of [3], and although our approach has some similarities with that work, it is simpler and more general; in particular, we can avoid the detailed use made there of characteristic functions.

Notation, Definitions, Results, and Proofs

For each $n = 1, 2, \cdots$, γ_n denotes the set of all binary sequences (i.e., sequences of zeros and ones) of length n, $\mathcal{B}_n$ is the algebra of all subsets of γ_n, and P_n is the uniform probability measure on $(\gamma_n, \mathcal{B}_n)$. Expectation on the probability space $(\gamma_n, \mathcal{B}_n, P_n)$ is written E_n.

$\gamma = \bigcup_{n=1}^{\infty}\gamma_n$ is a binary tree. Its branches are denoted τ, so $\tau \in \gamma_n$ means that τ is a branch of length n. The particular branch $0 \cdots 0 \in \gamma_n$ is denoted ζ_n. The nodes are named θ, so $\theta \in \tau$ means that θ is a node in the branch τ. The set of nodes at the ends of the branches in γ_n is written ξ_n, so ξ_n is the nth generation of nodes. $\Theta = \bigcup_{n=1}^{\infty} \xi_n$ is the set of all nodes of the tree γ (we exclude the first node).

For $\tau \in \gamma_k$ and $\tau' \in \gamma_n$, with $k \leq n$, the notation $\tau \leq \tau'$ means that the first k nodes of τ' are the nodes of τ (so τ' is an extension of τ). For $\tau \in \gamma_k$ and $k \leq n$, γ_n^{τ} is the set of branches $\{\tau' \in \gamma_n : \tau \leq \tau'\}$ (i.e., the extensions of τ of size n). Clearly, the cardinality of γ_n is 2^n, and for $\tau \in \gamma_k$ and $k \leq n$, γ_n^{τ} has cardinality 2^{n-k}.

Let $\{X(\theta),\ \theta \in \Theta\}$ be independent and identically distributed random variables, with mean zero and variance one, defined on a probability space $(\Omega, \mathfrak{F}, P)$. The notations E and Var refer to expectation and variance on this space. For $\tau \in \gamma_n$, and a natural number k, $1 \leq k \leq n$, $S_k(\tau) = \Sigma_{\theta \in \tau' \in \gamma_k, \tau' \leq \tau} X(\theta)$ is the sum of the random variables on the first k nodes of the branch τ.

On the product probability space $(\Omega \times \gamma_n, \mathfrak{F} \times \mathcal{B}_n, P \times P_n)$ we define a random element Y_n of $C[0, 1]$ as follows: for $\omega \in \Omega, \tau \in \gamma_n, t \in [0, 1]$,

$$Y_n(\omega, \tau; t) = n^{-1/2} S_k(\omega, \tau), \ t = k/n, \ k = 0, 1, \cdots, n,$$
$$\text{linear on } [(k-1)/n, k/n], \ k = 1, \cdots, n.$$

These are the Donsker paths on the branches of γ_n. For fixed $\omega \in \Omega$, the random process $\{Y_n(\omega, \cdot\,; t), t \in [0, 1]\}$ on $(\gamma_n, \mathcal{B}_n, P_n)$ represents a (uniform) random choice of one of the 2^n Donsker paths. In the model of [3], the (random) distribution of the positions of the elements belonging to the nth generation was considered; now we are taking with each element in the nth generation the positions of its whole line of ancestors.

The invariance principle is the

THEOREM. *The sequence or processes $\{Y_n(\omega, \cdot\,; t), t \in [0, 1]\}$ converges weakly to a standard Brownian motion process B for P-almost all ω.*

Consequently, for any measurable $f: C[0, 1] \to (-\infty, \infty)$ whose discontinuity set has B-measure zero, $f(Y_n(\omega, \cdot))$ converges weakly to $f(B)$ as $n \to \infty$, for P-almost all ω; and if f is bounded, $E_n f(Y_n(\omega, \cdot)) = 2^{-n} \Sigma_{\tau \in \gamma_n} f(Y_n(\omega, \tau)) \to \int_{C[0,1]} f dB$ as $n \to \infty$, for P-almost all ω. In particular, the functional $f(x) = x(1)$, $x \in C[0, 1)$, yields the central limit theorem of [3].

Proof. The proof is presented in two lemmas. The first lemma gives a sufficient condition for weak convergence of X_n to X on $C[0, 1]$. The second lemma shows that the sufficient condition of the first holds for $X_n = Y_n(\omega)$ and $X = B$, for P-almost all ω, thus completing the proof of the theorem.

A complex-valued function f on $C[0, 1]$ is said to satisfy a Lipschitz condition of order one if there is a constant $M \geq 0$ such that

$$| f(x) - f(y) | \leq M \sup_{0 \leq t \leq 1} | x(t) - y(t) |, \ x, y \in C[0, 1].$$

LEMMA 1. *Let X_n, $n = 1, 2, \cdots$, and X be random elements of $C[0, 1]$ defined respectively on probability spaces $(\Omega_n, \mathfrak{F}_n, P_n)$, $n = 1, 2, \cdots$, and $(\Omega, \mathfrak{F}, P)$. If*

$$\int f(X_n) dP_n \to \int f(X) dP \text{ as } n \to \infty$$

for all bounded complex-valued functions f on $C[0, 1]$ that satisfy a Lipschitz condition of order one, then X_n converges weakly to X as $n \to \infty$.

Proof. Let $f(x) = \exp iuw(x, \delta)$, where $u \in (-\infty, \infty)$ is fixed, and $w(x, \delta) = \sup_{|t-s|<\delta} | x(t) - x(s) |$ is the modulus of continuity of $x \in C[0, 1]$, with fixed $\delta > 0$. Using the inequality $| e^{ia} - 1 | \leq | a |$ for real a, and the inequality $| w(x, \delta) - w(y, \delta) | \leq 2 \sup_{0 \leq t \leq 1} | x(t) - y(t) |$, we obtain $| f(x) - f(y) | \leq 2 | u | \sup_{0 \leq t \leq 1} | x(t) - y(t) |$, and since this holds for all u, we have by hypothesis that $w(X_n, \delta)$ converges weakly to $w(X, \delta)$ as $n \to \infty$; in particular, $P_n[w(X_n, \delta) \geq \epsilon] \to P[w(X, \delta) \geq \epsilon]$ as $n \to \infty$, for $\epsilon > 0$, which implies that $\{X_n\}$ is tight (see [1], Theorem 8.2).

Now let $f(x) = \exp i \Sigma_{j=1}^{m} u_j x(t_j)$, where $t_1, \cdots, t_m \in [0, 1]$ are fixed,

$u_1, \cdots, u_m \in (-\infty, \infty)$ are fixed, and $x \in C[0, 1]$. Again using the inequality $|e^{ia} - 1| \leq |a|$ we find $|f(x) - f(y)| \leq \Sigma_{j=1}^{m} |u_j| \sup_{0 \leq t \leq 1} |x(t) - y(t)|$, and since this holds for all $u_1, \cdots, u_m$, by hypothesis it follows that $(X_n(t_1), \cdots, X_n(t_m))$ converges weakly to $(X(t_1,) \cdots, X(t_m))$ as $n \to \infty$.

Since the finite-dimensional distributions of X_n converge weakly to those of X, and $\{X_n\}$ is tight, the lemma is proved (see [1], Theorem 8.1).

LEMMA 2. *For all bounded complex-valued functions f on $C[0, 1]$ that satisfy a Lipschitz condition of order one,*

$$2^{-n}\Sigma_{\tau \in \gamma_n} f(Y_n(\omega, \tau)) \to \int_{C[0,1]} f dB \text{ as } n \to \infty$$

for P-almost all ω.

Proof. For $0 \leq k \leq n$, define a random element $Y_{n,k}$ of $C[0, 1]$ on $(\Omega \times \gamma_n, \mathfrak{F} \times \mathcal{B}_n, P \times P_n)$ as follows: for $\omega \in \Omega, \tau \in \gamma_n, t \in [0, 1]$,

$$Y_{n,k}(\omega, \tau; t) = \begin{cases} 0, \ 0 \leq t \leq k/n, \\ n^{-1/2}(S_j(\omega, \tau) - S_k(\omega, \tau)), \ t = j/n, j = k, \cdots, n, \\ \text{linear on } [(j-1)/n, j/n], j = k+1, \cdots, n. \end{cases}$$

For f satisfying the hypothesis of the lemma, let

$$\varphi_n(\omega) = 2^{-n}\Sigma_{\tau \in \gamma_n} f(Y_n(\omega, \tau)), \ \omega \in \Omega,$$

$$\varphi_{n,k}(\omega) = 2^{-n}\Sigma_{\tau \in \gamma_n} f(Y_{n,k}(\omega, \tau)), \ \omega \in \Omega.$$

We will show that

1. $P[\varphi_n - \varphi_{n,k_n} \to 0 \text{ as } n \to \infty] = 1$, with $k_n \to \infty$, $k_n n^{-1/2} \to 0$ as $n \to \infty$.
2. $E\varphi_{n,k_n} \to \int_{C[0,1]} f dB$ as $n \to \infty$, with $k_n \to \infty$, $k_n n^{-1/2} \to 0$ as $n \to \infty$.
3. $P[\varphi_{n,k_n} - E\varphi_{n,k_n} \to 0 \text{ as } n \to \infty] = 1$, with k_n = integer part of $n^{1/3}$.

These three steps clearly prove the lemma.

Proof of 1.
$$\begin{aligned} |\varphi_n - \varphi_{n,k}| &\leq 2^{-n}\Sigma_{\tau \in \gamma_n} |f(Y_n(\tau)) - f(Y_{n,k}(\tau))| \\ &\leq M2^{-n}\Sigma_{\tau \in \gamma_n} \sup_{0 \leq t \leq 1} |Y_n(\tau, t) - Y_{n,k}(\tau, t)| \\ &\leq M2^{-n}\Sigma_{\tau \in \gamma_n} \sup_{1 \leq j \leq k} n^{-1/2} |S_j(\tau)| \\ &\leq M2^{-n}n^{-1/2}\Sigma_{\tau \in \gamma_n}\Sigma_{\theta \in \tau' \in \gamma_{k}, \tau' \leq \tau} |X(\theta)| \\ &= M2^{-n}n^{-1/2}\Sigma_{j=1}^{k} 2^{n-j}\Sigma_{\theta \in \xi_j} |X(\theta)| \\ &= Mn^{-1/2}\Sigma_{j=1}^{k} Z_j, \end{aligned}$$

where $Z_j = 2^{-j}\Sigma_{\theta \in \xi_j} |X(\theta)|, j = 1, \cdots, k$, are independent, with $EZ_j = \mu$ and $\text{Var } Z_j = 2^{-j}\sigma^2$, where $\mu = E|X(0)|$ and $\sigma^2 = \text{Var}|X(0)|$.

Hence $|\varphi_n - \varphi_{n,k}| \leq Mn^{-1/2}\Sigma_{j=1}^{k}(Z_j - \mu) + M\mu k n^{-1/2}$, and since $\Sigma_{j=1}^{\infty} \text{Var } Z_j < \infty$ implies that $\Sigma_{j=1}^{\infty}(Z_j - \mu)$ converges almost surely (see [7], p. 236), step 1 is proved.

Proof of 2. $E\varphi_{n,k} = E\varphi_n + E(\varphi_{n,k} - \varphi_n)$, where clearly $E\varphi_n = Ef(Y_n(\zeta_n))$.

Since φ_n and $\varphi_{n,k}$ are uniformly bounded, because f is bounded, then by step 1 and the Lebesgue dominated convergence theorem, $E(\varphi_{n,k_n} - \varphi_n) \to 0$ as $n \to \infty$. Now by Donsker's theorem (see [1], Theorem 10.1), $Y_n(\zeta_n)$ converges weakly to B as $n \to \infty$, and therefore $E\varphi_n \to \int_{C[0,1]} f dB$ as $n \to \infty$, and step 2 is proved.

Proof of 3. Let $A_{n,k} = \varphi_{n,k} - E\varphi_{n,k}$, and $T_\tau = \Sigma_{\tau' \in \gamma_n^\tau} f(Y_{n,k}(\tau'))$, $\tau \in \gamma_k$. T_τ, $\tau \in \gamma_k$, are independent and identically distributed, and

$$A_{n,k} = 2^{-n} \Sigma_{\tau \in \gamma_k} (T_\tau - ET_\tau);$$

hence $E \mid A_{n,k} \mid^2 = 2^{-2n+k} E \mid T_{\zeta_k} - ET_{\zeta_k} \mid^2 \leq 2^{-2n+k} E \mid T_{\zeta_k} \mid^2$, and since f is bounded, say by K, $E \mid T_{\zeta_k} \mid^2 \leq K^2 2^{2(n-k)}$, so that $E \mid A_{n,k} \mid^2 \leq K^2 2^{-k}$.

Since $2^{-n^{1/3}} < n^{-2}$ for large n, by Chebyshev's inequality and the Borel-Cantelli lemma we have, for k_n = integer part of $n^{1/3}$, $P[\mid A_{n,k_n} \mid > \epsilon$ infinitely often$] = 0$ for all $\epsilon > 0$, which finishes the proof.

Centro de Investigación del IPN
Universidad Autónoma Metropolitana (Iztapalapa)

References

[1] P. Billingsley, Convergence of Probability Measures, Wiley, (1968).
[2] T. E. Harris, The Theory of Branching Processes, Springer-Verlag, (1963).
[3] A. Joffe and A. R. Moncayo, *On sums of independent random variables defined on a binary tree*, Bol. Soc. Mat. Mexicana., **18**, N° 1, (1973), 50–4.
[4] ———, *Random variables, trees, and branching random walks*, Advances in Math., **10**, N° 3, (1973), 401–16.
[5] B. P. Kharlamov, *The number of generations in a branching process with an arbitrary set of particle types*, Theor. Probability Appl., **14** (1969), 432–49.
[6] A. N. Kolmogorov, *Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung*, Dokl. Akad. Nauk SSSR, **31** (1941), 99–101.
[7] M. Loève, Probability Theory, Third Edition, Van Nostrand, (1963).
[8] P. E. Ney, *The limit distribution of a binary cascade process*, J. Math. Anal. Appl., **10** (1965), 30–6.
[9] ———, *The convergence of a random distribution function associated with brancing processes*, J. Math. Anal. Appl., **12** (1965), 316–27.
[10] A. J. Stam, *On a conjecture by Harris*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete, **5** (1966), 202–06.