THE LINDEBERG-LÉVY THEOREM FOR MARTINGALES[1]

PATRICK BILLINGSLEY

The central limit theorem of Lindeberg [7] and Lévy [3] states that if $\{u_1, u_2, \cdots\}$ is an independent, identically distributed sequence of random variables with finite second moments, then the distribution of $n^{-1/2}\sum_{k=1}^{n} u_k$ approaches the normal distribution with mean 0 and variance $E\{u_1^2\}$, assuming that $E\{u_1\}=0$, which entails no loss of generality. In the following result, the assumption of indepence is weakened.

THEOREM. *Let $\{u_1, u_2, \cdots\}$ be a stationary, ergodic stochastic process such that $E\{u_1^2\}$ is finite and*

$$E\{u_n \| u_1, \cdots, u_{n-1}\} = 0 \tag{1}$$

with probability one. Then the distribution of $n^{-1/2}\sum_{k=1}^{n} u_k$ approaches the normal distribution with mean 0 and variance $E\{u_1^2\}$.

The condition (1) is exactly the requirement that the partial sums $\sum_{k=1}^{n} u_k$ form a martingale. The theorem will be proved by sharpening the methods of [1, §9], which in turn are based on work of Lévy; see [4], [5, Chapter 4], and [6, pp. 237 ff]. The debt to Lévy will be clear to anyone familiar with these papers.

In proving the theorem, we may assume that the process is represented in the following way. Let Ω be the cartesian product of a sequence of copies of the real line, indexed by the integers $n=0, \pm 1, \pm 2, \cdots$. Let u_n be the coordinate variables, let $\mathcal{B}$ be the Borel field generated by them, and let P be that probability measure on $\mathcal{B}$ with the finite-dimensional distributions prescribed by the original process. If $\mathfrak{F}_n$ is the Borel field generated by $\{u_n, u_{n-1}, u_{n-2}, \cdots\}$, then, by (1),

$$E\{u_n \| \mathfrak{F}_{n-1}\} = 0, \tag{2}$$

with probability one, for $n=0, \pm 1, \cdots$.

Let $\sigma_n^2 = E\{u_n^2 \| \mathfrak{F}_{n-1}\}$ and let $\sigma^2 = E\{\sigma_n^2\} = E\{u_n^2\}$. If T is the shift operator then, as is easily shown, $\sigma_n^2 = T^n \sigma_0^2$. Since T is ergodic, it follows by the ergodic theorem that

Received by the editors November 3, 1960.

[1] This research was supported in part by Research Grant No. NSF-G10368 from the Division of Mathematical, Physical, and Engineering Sciences of the National Science Foundation.

$$\lim_{n\to\infty} n^{-1} \sum_{k=1}^{n} \sigma_k^2 = \sigma^2 \tag{3}$$

with probability one. Let $s_n^2 = \sigma_1^2 + \cdots + \sigma_n^2$, put $m_t = \min\{n: s_n^2 \geqq t\}$ for $t > 0$, let c_t be that number such that $0 < c_t \leqq 1$ and $s_{m_t-1}^2 + c_t^2\sigma_{m_t}^2 = t$, and, finally, let $z_t = u_1 + \cdots + u_{m_t-1} + c_t u_{m_t}$. It follows from (3) that $\sum_k \sigma_k^2$ diverges with probability one; hence m_t and the other variables are well defined. We will first show that

$$\lim_{t\to\infty} P\{t^{-1/2}z_t \leqq x\} = \Phi(x), \tag{4}$$

where $\Phi(x)$ is the unit normal distribution function. The proof of the theorem will then be completed by showing that

$$\operatorname{p\,lim}_{n\to\infty} n^{-1/2}\left| \sum_{k=1}^{n} u_k - z_{n\sigma^2} \right| = 0. \tag{5}$$

To prove (4), define new variables $\tilde{u}_1, \tilde{u}_2, \cdots$ by

$$\tilde{u}_k = \begin{cases} u_k & \text{if } m_t > k \\ c_t u_k & \text{if } m_t = k \\ 0 & \text{if } m_t < k. \end{cases}$$

If $\Lambda \in \mathfrak{F}_{k-1}$ then, since $\{m_t > k\} \in \mathfrak{F}_{k-1}$, the variables u_k^2 and σ_k^2 have the same integrals over $\Lambda\{m_t > k\}$, by (2). Similarly, since c_t^2, multiplied by the indicator function of $\{m_t = k\}$, is measurable $\mathfrak{F}_{k-1}$, it follows that u_k^2 and $c_t^2\sigma_k^2$ have the same integrals over $\Lambda\{m_t = k\}$. Therefore, if $\tilde{\sigma}_k^2 = E\{\tilde{u}_k^2 \| \mathfrak{F}_{k-1}\}$,

$$\tilde{\sigma}_k^2 = \begin{cases} \tilde{\sigma}_k^2 & \text{if } m_t > k \\ c_t^2\tilde{\sigma}_k^2 & \text{if } m_t = k \\ 0 & \text{if } m_t > k \end{cases} \tag{6}$$

except on a set of measure zero. Similar arguments show that

$$E\{\tilde{u}_k \| \mathfrak{F}_{k-1}\} = 0, \tag{7}$$

with probability one.

Adjoin to the space random variables $\xi_1, \xi_2, \cdots$, each normally distributed with mean 0 and variance 1, which are independent of each other and of the Borel field $\mathfrak{B}$. If

$$\eta_n = t^{-1/2}(\tilde{u}_1 + \cdots + \tilde{u}_n + \tilde{\sigma}_{n+1}\xi_{n+1} + \tilde{\sigma}_{n+2}\xi_{n+2} + \cdots)$$

then $\eta_n = t^{-1/2}z_t$ for $n \geqq m_t$. Moreover, since $E\{\eta_0 \| \mathfrak{B}\} = 0$ and

$E\{\eta_0^2 \| \mathfrak{B}\} = t^{-1}\sum_{k=1}^{\infty} \bar{\sigma}_k^2 = 1$, η_0 has the unit normal distribution. Therefore

$$(8)\qquad \begin{aligned} &| E\{\exp(ist^{-1/2}z_t)\} - \exp(-s^2/2) | \\ &\qquad\qquad \leq \sum_{n=1}^{\infty} | E\{\exp(is\eta_n)\} - E\{\exp(is\eta_{n-1})\} | . \end{aligned}$$

Write

$$\begin{aligned} \alpha &= \exp\left(ist^{-1/2} \sum_{k=1}^{n-1} \bar{u}_k \right), \\ \beta &= \exp(ist^{-1/2}\bar{u}_n) - \exp(ist^{-1/2}\bar{\sigma}_n\xi_n), \\ \gamma &= \exp\left(ist^{-1/2} \sum_{k=n+1}^{\infty} \bar{\sigma}_k\xi_k \right). \end{aligned}$$

If $\gamma' = E\{\gamma \| \xi_n, \mathfrak{B}\}$, then

$$\gamma' = \exp\left(-(s^2/2t) \sum_{k=n+1}^{\infty} \bar{\sigma}_k^2 \right),$$

and hence, since $\sum_{k=n+1}^{\infty} \bar{\sigma}_k^2 = t - \sum_{k=1}^{n} \bar{\sigma}_k^2$, γ' is measurable $\mathfrak{F}_{n-1}$. It follows that

$$\begin{aligned} | E\{\exp(is\eta_n)\} - E\{\exp(is\eta_{n-1})\} | &= | E\{\alpha\beta\gamma\} | \\ = | E\{\alpha\beta\gamma'\} | &= | E\{\alpha\gamma' E\{\beta \| \mathfrak{F}_{n-1}\}\} | \\ \leq E\{ | E\{\beta \| \mathfrak{F}_{n-1}\} | \}. \end{aligned}$$

By Taylor's theorem it follows that if w is real then

$$\exp(iw) = 1 + iw - w^2/2 + \theta w^2 g(|w|),$$

where $|\theta| \leq 1$ and

$$g(w) = \sup_{0 \leq v \leq w} | 1 - \exp(iv) | /2.$$

Note that $g(0) = 0$ and that on $[0, \infty)$, g is continuous, nondecreasing, and bounded by 1. Applying this formula to each of the two terms of β, and using (6) and (7), we obtain

$$\begin{aligned} E\{\beta \| \mathfrak{F}_{n-1}\} &= E\{\theta s^2 t^{-1} \bar{u}_n^2 g(|s| t^{-1/2} |\bar{u}_n|) \| \mathfrak{F}_{n-1}\} \\ &\qquad + E\{\theta s^2 t^{-1} \bar{\sigma}_n^2 \xi_n^2 g(|s| t^{-1/2} \bar{\sigma}_n |\xi_n|) \| \mathfrak{F}_{n-1}\}. \end{aligned}$$

If $h(w) = E\{\xi_1^2 g(w|\xi_1|)\}$ for $w \geq 0$, then h has the same properties g does, and

$$\left| E\{\beta \| \mathfrak{F}_{n-1}\} \right| \leqq s^2 t^{-1} E\{\tilde{u}_n^2 g(|s| t^{-1/2} |\tilde{u}_n|) \| \mathfrak{F}_{n-1}\} + s^2 t^{-1} \tilde{\sigma}_n^2 h(|s| t^{-1/2} \tilde{\sigma}_n).$$

Therefore, by (8),

$$\begin{aligned} (9) \quad & \left| E\{\exp(ist^{-1/2} z_t)\} - \exp(-s^2/2) \right| \\ & \leqq E\left\{ s^2 t^{-1} \sum_{k=1}^{m_t} \left[E\{\tilde{u}_k^2 g(|s| t^{-1/2} |\tilde{u}_k|) \| \mathfrak{F}_{k-1}\} + \tilde{\sigma}_k^2 h(|s| t^{-1/2} \tilde{\sigma}_k) \right] \right\}. \end{aligned}$$

Since g and h are bounded by 1, and since $\sum_{k=1}^{m_t} \tilde{\sigma}_k^2 = t$, the integrand on the right in this expression is bounded by $2s^2$. If we show that the integrand goes to 0 with probability one as t goes to infinity, then it will follow by the dominated convergence theorem that the right-hand member of (9) goes to 0, which will complete the proof of (4).

If $\epsilon > 0$ then $|s| t^{-1/2} < \epsilon$ for all sufficiently large t, and, since g is nondecreasing, we have

$$\begin{aligned} (10) \quad & \limsup_{t \to \infty} t^{-1} \sum_{k=1}^{m_t} E\{\tilde{u}_k^2 g(|s| t^{-1/2} |\tilde{u}_k|) \| \mathfrak{F}_{k-1}\} \\ & \leqq \limsup_{t \to \infty} t^{-1} \sum_{k=1}^{m_t} E\{u_k^2 g(\epsilon |u_k|) \| \mathfrak{F}_{k-1}\}. \end{aligned}$$

It follows from the ergodic theorem that

$$(11) \quad \lim_{n \to \infty} n^{-1} \sum_{k=1}^{n} E\{u_k^2 g(\epsilon |u_k|) \| \mathfrak{F}_{k-1}\} = E\{u_1^2 g(\epsilon |u_1|)\},$$

with probability one. A standard argument of the renewal type applied to (3) shows that

$$(12) \quad \lim_{t \to \infty} m_t / t = \sigma^{-2}$$

with probability one. (If $\lambda > 1$ and $k_t = [t\lambda\sigma^{-2}]$, then $k_t \geqq \lambda^{1/2}\sigma^{-2} t$ for large t. But (3) implies that for large t, $s_{k_t}^2 / k_t \geqq \sigma^2 \lambda^{-1/2}$, and hence $s_{k_t}^2 \geqq t$, or $m_t \leqq k_t \leqq t\lambda\sigma^{-2}$; thus $\limsup_t m_t/t \leqq \lambda\sigma^{-2}$. A similar argument for $\lambda < 1$ yields (12).) Now (11) and (12) imply that the left-hand member of (10) does not exceed $\sigma^{-2} E\{u_1^2 g(\epsilon |u_1|)\}$. Since this bound goes to 0 with ϵ by the dominated convergence theorem, the left-hand member of (10) is 0, with probability one. A similar argument with $\tilde{u}_k$ and g replaced by $\tilde{\sigma}_k$ and h, shows that

$$\lim_{n \to \infty} t^{-1} \sum_{k=1}^{m_t} \tilde{\sigma}_k^2 h(|s| t^{-1/2} \tilde{\sigma}_k) = 0,$$

with probability one. Thus the integrand on the right in (9) goes to 0, with probability one, which completes the proof of (4).

It remains to prove (5). Given $\epsilon > 0$, choose n_0 so that if $n \geqq n_0$ then

$$P\{ \mid m_{n\sigma^2}/n\sigma^2 - \sigma^{-2} \mid > \epsilon^3 \} < \epsilon,$$

which is possible by (12). If $n \geqq n_0$ then

$$P\left\{ n^{-1/2} \left| \sum_{k=1}^{n} u_k - z_{n\sigma^2} \right| > \epsilon \right\} \leqq \epsilon + P\left\{ \max_{a \leqq l \leqq b} \left| \sum_{k=a}^{l} u_k \right| \geqq \epsilon n^{1/2}/2 \right\},$$

where $a = n - [\epsilon^3 n\sigma^2]$ and $b = n + [\epsilon^3 n\sigma^2]$. By Kolmogorov's inequality for martingales [2],

$$P\left\{ \max_{a \leqq l \leqq b} \left| \sum_{k=a}^{l} u_k \right| \geqq \epsilon n^{1/2}/2 \right\} \leqq (4/\epsilon^2 n) \sum_{k=a}^{b} E\{u_k^2\} \leqq 8\epsilon\sigma^2.$$

Thus

$$P\left\{ n^{-1/2} \left| \sum_{k=1}^{n} u_k - z_{n\sigma^2} \right| > \epsilon \right\} \leqq (1 + 8\sigma^2)\epsilon$$

if $n \geqq n_0$, which establishes (5) and completes the proof of the theorem.

If $u_n = f(z_n)$, where $\{z_n\}$ is a Markov process satisfying the regularity condition described in [1, part (i) of Condition 1.2], and if $E\{f(z_1)^2\}$ is finite when z_1 has the stationary distribution, then, as one can show by the arguments of [1], $n^{-1/2} \sum_{k=1}^{n} u_k$ is asymptotically normal, even if the distribution of z_1 is not the stationary one.

References

1. Patrick Billingsley, *Statistical inference for Markov processes*, Institute of Mathematical Statistics-University of Chicago Statistical Research Monographs, University of Chicago Press, 1961.

2. J. L. Doob, *Stochastic processes*, New York, John Wiley and Sons, 1953.

3. Paul Lévy, *Calcul des probabilités*, Paris, Gauthier-Villars, 1925.

4. ———, *Propriétés asymptotique des sommes de variables enchaînées*, Bull. Sci. Math. vol. 59 (1935) pp. 84–96, 109–128.

5. ———, *Propriétés asymptotiques des sommes de variables aléatoires independantes ou enchaînées*, J. Math. Pures Appl. Ser. 8, vol. 14 (1935) pp. 347–402.

6. ———, *Théorie de l'addition des variables aléatoires*, Paris, Gauthier-Villars, 1937.

7. J. W. Lindeberg, *Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung*, Math. Z. vol. 15 (1922) pp. 211–225.

University of Chicago