THE LINDEBERG-LÉVY THEOREM FOR MARTINGALES

PATRICK BILLINGSLEY

The central limit theorem of Lindeberg [7] and Lévy [3] states that if \( \{ m_1, m_2, \cdots \} \) 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 \( E\{ m^2 \} \), assuming that \( E\{ m_1 \} = 0 \), which entails no loss of generality. In the following result, the assumption of independence is weakened.

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

\[
E\{ u_n | u_1, \cdots, u_{n-1} \} = 0
\]

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

The condition (1) is exactly the requirement that the partial sums \( \sum_{k=1}^{n-1} u_k \) 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 \( \Omega \) 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 \( \mathcal{F}_n \) is the Borel field generated by \( \{ u_n, u_{n-1}, u_{n-2}, \cdots \} \), then, by (1),

\[
E\{ u_n | \mathcal{F}_{n-1} \} = 0,
\]

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

Let \( \sigma_n^2 = E\{ u_n^2 | \mathcal{F}_{n-1} \} \) 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.
\[ (3) \quad \lim_{n \to \infty} n^{-1} \sum_{k=1}^{n} \sigma_k^2 = \sigma^2 \]

with probability one. Let \( \sigma_n^2 = \sigma_1^2 + \cdots + \sigma_n^2 \), put \( m_t = \min \{ n : s_n^2 \geq t \} \) for \( t > 0 \), let \( c_t \) be that number such that \( 0 < c_t \leq 1 \) and \( s_{m_{t-1}}^2 + c_t \sigma_{m_t}^2 = t \), and, finally, let \( s_t = \sum_{i=0}^{m_t-1} u_{m_{t-1} + i} + c_t u_{m_t} \). 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

\[ (4) \quad \lim_{t \to \infty} P \left( t^{-1/2} z_t \leq x \right) = \Phi(x), \]

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

\[ (5) \quad p \lim_{n \to \infty} n^{-1/2} \left| \sum_{k=1}^{n} u_k - z_{n+t} \right| = 0. \]

To prove (4), define new variables \( \bar{u}_1, \bar{u}_2, \ldots \) by

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

If \( \Lambda \in \mathfrak{F}_{k-1} \) then, since \( \{ m_t > k \} \in \mathfrak{F}_{k-1} \), the variables \( u_k^2 \) and \( \sigma_k^2 \) 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 \( \sigma_k^2 = E \{ u_k^2 | \mathfrak{F}_{k-1} \} \),

\[ (6) \quad \sigma_k^2 = \begin{cases} \sigma_k^2 & \text{if } m_t > k \\ c_t^2 \sigma_k^2 & \text{if } m_t = k \\ 0 & \text{if } m_t < k. \end{cases} \]

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

\[ (7) \quad E \{ \bar{u}_k | \mathfrak{F}_{k-1} \} = 0, \]

with probability one.

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

\[ \eta_n = t^{-1/2} (\bar{u}_1 + \cdots + \bar{u}_n + \sigma_{n+1} \xi_{n+1} + \sigma_{n+2} \xi_{n+2} + \cdots) \]

then \( \eta_n = t^{-1/2} \mathfrak{Z}_t \) for \( n \geq m_t \). Moreover, since \( E \{ \eta_0 | \mathfrak{B} \} = 0 \) and
\[ E\{\eta_0\| (\mathcal{B}) \} = t^{-1} \sum_{k=1}^{\infty} \sigma_k^2 = 1, \eta_0 \text{ has the unit normal distribution. Therefore} \]
\[ |E\{\exp(ist^{-1/2}z_t)\} - \exp(-s^2/2)| \]
\[ \leq \sum_{n=1}^{\infty} \left| E\{\exp(isz_n)\} - E\{\exp(isz_{n-1})\} \right|. \]

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

If \( \gamma' = E\{\gamma\| \xi_n, (\mathcal{B})\} \), then
\[ \gamma' = \exp \left( -(s^2/2t) \sum_{k=n+1}^{\infty} \sigma_k^2 \right), \]
and hence, since \( \sum_{k=n+1}^{\infty} \sigma_k^2 = t - \sum_{k=1}^{n-1} \sigma_k^2 \), \( \gamma' \) is measurable \( \mathcal{G}_{n-1} \). It follows that
\[ |E\{\exp(isz_n)\} - E\{\exp(isz_{n-1})\}| = |E\{\alpha \gamma\}| \]
\[ = |E\{\alpha \gamma'\}| = |E\{\alpha \gamma' E\{\beta\| \mathcal{G}_{n-1}\}\}| \]
\[ \leq E\{|E\{\beta\| \mathcal{G}_{n-1}\}|\}. \]

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
\[ g(w) = \sup_{0 \leq t \leq w} |1 - \exp(iw)| / 2. \]

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 \( \beta \), and using (6) and (7), we obtain
\[ E\{\beta\| \mathcal{G}_{n-1}\} = E\{\delta s \ t^{-1/2} \bar{u}_n g(|s| t^{-1/2} |\bar{u}_n|)\| \mathcal{G}_{n-1}\} \]
\[ + E\{\delta s \ t^{-1/2} \bar{\sigma}_n \xi_n g(|s| t^{-1/2} |\bar{\sigma}_n| |\xi_n|)\| \mathcal{G}_{n-1}\}. \]

If \( h(w) = E\{\xi^2 g(w | \xi_1)| \) for \( w \geq 0 \), then \( h \) has the same properties \( g \) does, and
Therefore, by (8),
\[
\left| E\{\exp(\text{ist}^{-1/2}\zeta_t)\} - \exp(-s^2/2)\right| \\
\leq E\left\{2^{2-1} \sum_{k=1}^{m_1} \left[ E\left\{\frac{2}{\sigma_k^2} \left| \sum_{i=1}^{k-1} \sigma_i \right| \right| \mathcal{F}_{k-1}\right] + \sigma_k^2 h\left| \frac{s}{\tau_i} \right| \right| \mathcal{F}_{k-1}\right\}. 
\]
Since $g$ and $h$ are bounded by 1, and since $\sum_{k=1}^{m_1} \sigma_k^2 = t$, 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 $|s| t^{-1/2} \leq \epsilon$ for all sufficiently large $t$, and, since $g$ is nondecreasing, we have
\[
\limsup_{t \to \infty} t^{-1} \sum_{k=1}^{m_1} E\left\{\frac{2}{\sigma_k^2} \left| \sum_{i=1}^{k-1} \sigma_i \right| \right| \mathcal{F}_{k-1}\right\} \\
= \limsup_{t \to \infty} t^{-1} \sum_{k=1}^{m_1} E\left\{\frac{2}{\sigma_k^2} \left| u_k \right| \right| \mathcal{F}_{k-1}\right\}. 
\]
It follows from the ergodic theorem that
\[
\lim_{n \to \infty} n^{-1} \sum_{k=1}^{m_1} E\left\{\frac{2}{\sigma_k^2} \left| u_k \right| \right| \mathcal{F}_{k-1}\right\} = E\left\{\frac{2}{\sigma_i^2} \left| u_i \right| \right| \mathcal{F}_{i}\right\}, 
\]
with probability one. A standard argument of the renewal type applied to (3) shows that
\[
\lim_{t \to \infty} m_t/t = \sigma^{-2} 
\]
with probability one. (If $\lambda > 1$ and $k_t = [\lambda \sigma^{-2}]$, then $k_t \geq \lambda^{1/2} \sigma^{-2} t$ for large $t$. But (3) implies that for large $t$, $s^2_t/k_t \geq \sigma^2 \lambda^{-1/2}$, and hence $s^2_t \leq k_t \leq \lambda \sigma^{-2}$; thus $\limsup_t m_t/t \leq \lambda \sigma^{-2}$. A similar argument for $\lambda < 1$ yields (12).) Now (11) and (12) imply that the left-hand member of (10) does not exceed $\sigma^{-2} E\left\{\frac{2}{\sigma_i^2} \left| u_i \right| \right| \mathcal{F}_{i}\right\}$. Since this bound goes to 0 with $\epsilon$ by the dominated convergence theorem, the left-hand member of (10) is 0, with probability one. A similar argument with $\bar{u}_k$ and $g$ replaced by $\bar{\sigma}_k$ and $h$, shows that
\[
\lim_{n \to \infty} t^{-1} \sum_{k=1}^{m_t} \frac{2}{\sigma_k^2} \left| s \right| t^{-1/2} \sigma_k = 0, 
\]
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 \( n \geq n_0 \) then
\[
P\left\{ \left| \frac{m_{n_0} \sigma}{n \sigma^2} - \sigma^{-2} \right| > \epsilon \right\} < \epsilon,
\]
which is possible by (12). If \( n \geq n_0 \) then
\[
P\left\{ \left| \sum_{k=1}^{n} u_k - z_{n_0}^2 \right| > \epsilon \right\} \leq \epsilon + P\left\{ \max_{a \leq i \leq b} \left| \sum_{k=a}^{i} u_k \right| \geq \epsilon n^{1/2}/2 \right\},
\]
where \( a = n - \lfloor \epsilon^2 n \sigma^2 \rfloor \) and \( b = n + \lfloor \epsilon^2 n \sigma^2 \rfloor \). By Kolmogorov's inequality for martingales [2],
\[
P\left\{ \max_{a \leq i \leq b} \left| \sum_{k=a}^{i} u_k \right| \geq \epsilon n^{1/2}/2 \right\} \leq 8 \sigma^2 \cdot
\]
Thus
\[
P\left\{ \left| \sum_{k=1}^{n} u_k - z_{n_0}^2 \right| > \epsilon \right\} \leq (1 + 8 \sigma^2) \epsilon
\]
if \( n \geq n_0 \), 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


University of Chicago