The Martingale Central Limit Theorem

Steven P. Lalley
University of Chicago

May 27, 2014

1 Lindeberg's Method

One of the most useful generalizations of the central limit theorem is the *martingale central limit theorem* of Paul Lévy. Lévy was in part inspired by Lindeberg's treatment of the central limit theorem for sums of independent – but not necessarily identically distributed – random variables. Lindeberg formulated what, in retrospect, is the right hypothesis, now known as the *Lindeberg condition*,[1] on the summands for the central limit theorem, and in addition he proposed a new approach to proving central limit theorems. The Lindeberg condition plays a central role in the most general form of the martingale central limit theorem, as well, and as Lévy showed, the Lindeberg method of proof can be adapted to martingales. In this section I will show how Lindeberg's method works in the very simplest context, for sums of independent, identically distributed random variables. In section 3, I will show how the method can be generalized to martingales.

Theorem 1. *(Central Limit Theorem) Let $\xi_1, \xi_2, \dots$ be independent, identically distributed random variables with mean zero and variance* 1*. Then for every continuous, bounded function $f : \mathbb{R} \to \mathbb{R}$,*

$$\lim_{n\to\infty} Ef\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i\right) = Ef(Z) \tag{1}$$

where Z is a standard normal random variable.

Proof. It suffices to prove this for C^∞ functions f with compact support, by a standard approximation argument, so I will assume henceforth that f is such a function. Lindeberg's method depends on the fact that the family of normal densities is closed under convolutions, in particular, if X and Y are independent Gaussian random variables then $X + Y$ is also Gaussian. Consequently, if $\zeta_1, \zeta_2, \dots, \zeta_n$ are independent standard normal random variables then

$$Z \overset{\mathcal{D}}{=} \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\zeta_i.$$

Without loss of generality we may assume that the underlying probability space is large enough to support not only the random variables ξ_i but also an independent sequence of i.i.d. standard Gaussian random

[1] Feller, and independently Lévy, later proved that Lindeberg's condition is in some sense *necessary* for the validity of the central limit theorem.

variables ζ_i. The objective is to show that as $n \to \infty$,

$$Ef\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i\right) - Ef\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\zeta_i\right) \longrightarrow 0. \tag{2}$$

For notational ease set

$$\begin{aligned} \xi_i' &= \xi_i/\sqrt{n} \quad \text{and} \\ \zeta_i' &= \zeta_i/\sqrt{n}; \end{aligned}$$

then relation (2) can be re-stated as

$$Ef\left(\sum_{i=1}^{n}\xi_i'\right) - Ef\left(\sum_{i=1}^{n}\zeta_i'\right) \longrightarrow 0.$$

Lindeberg's strategy for proving (2) is to replace the summands ξ_i' in the first expectation by the corresponding Gaussian summands ζ_i', one by one, and to bound at each step the change in the expectation resulting from the replacement of ξ_i' by ζ_i':

$$\left|Ef\left(\sum_{i=1}^{n}\xi_i'\right) - Ef\left(\sum_{i=1}^{n}\zeta_i'\right)\right| \le \sum_{k=1}^{n}\left|Ef\left(\sum_{i=1}^{k}\xi_i' + \sum_{i=k+1}^{n}\zeta_i'\right) - Ef\left(\sum_{i=1}^{k-1}\xi_i' + \sum_{i=k}^{n}\zeta_i'\right)\right| \tag{3}$$

Since the individual terms ξ_i' and ζ_i' account for only a small fraction of the sums, the differences in the value of f can be approximated by using two-term Taylor series approximations. Furthermore, since f has compact support, the derivatives are *uniformly* continuous, and so the remainder terms can be estimated uniformly. In particular, for any $\varepsilon > 0$ there exist $\delta > 0$ and $C < \infty$ such that for any $x, y \in \mathbb{R}$,

$$\begin{aligned} &|f(x+y) - f(x) - f'(x)y - f''(x)y^2/2| \le \varepsilon y^2 \quad \text{if } |y| \le \delta \quad \text{and} \\ &|f(x+y) - f(x) - f'(x)y - f''(x)y^2/2| \le Cy^2 \quad \text{otherwise.} \end{aligned} \tag{4}$$

Consequently, for each k,

$$\begin{aligned} &Ef\left(\sum_{i=1}^{k}\xi_i' + \sum_{i=k+1}^{n}\zeta_i'\right) - Ef\left(\sum_{i=1}^{k-1}\xi_i' + \sum_{i=k}^{n}\zeta_i'\right) \\ &\quad = Ef'\left(\sum_{i=1}^{k-1}\xi_i' + \sum_{i=k+1}^{n}\zeta_i'\right)(\xi_k' - \zeta_k') + \frac{1}{2}Ef''\left(\sum_{i=1}^{k-1}\xi_i' + \sum_{i=k+1}^{n}\zeta_i'\right)((\xi_k')^2 - (\zeta_k')^2) + ER_k(A) + ER_k(B) \end{aligned} \tag{5}$$

where

$$\begin{aligned} R_k(A) &\le \varepsilon(\xi_k')^2 + C(\xi_k')^2\mathbf{1}\{|\xi_k'| \ge \delta\} \quad \text{and} \\ R_k(B) &\le \varepsilon(\zeta_k')^2 + C(\zeta_k')^2\mathbf{1}\{|\zeta_k'| \ge \delta\}. \end{aligned}$$

The crucial feature of the expansion (5) is the independence of the individual terms ξ_i' and ζ_i'; this guarantees that the first two expectations on the right side of (5) split (as products of expectations), and since

ξ'_k and ζ'_k have the same mean and variance, it follows that *the first two expectations on the right side are* 0. Consequently, for each k,

$$\begin{aligned}\left|Ef\left(\sum_{i=1}^{k}\xi'_i+\sum_{i=k+1}^{n}\zeta'_i\right)-Ef\left(\sum_{i=1}^{k-1}\xi'_i+\sum_{i=k}^{n}\zeta'_i\right)\right| &\\ \leq E|R_k(A)|+E|R_k(B)| &\\ \leq \varepsilon E(\xi'_k)^2+\varepsilon E(\zeta'_k)^2)+CE(\xi'_k)^2\mathbf{1}\{|\xi'_k|\geq\delta\}+CE(\zeta'_k)^2)\mathbf{1}\{|\zeta'_k|\geq\delta\} &\\ \leq n^{-1}\varepsilon(E(\xi_k)^2+E(\zeta_k)^2)+n^{-1}CE(\xi_k)^2\mathbf{1}\{|\xi_k|\geq\sqrt{n}\delta\}+n^{-1}CE(\zeta_k)^2)\mathbf{1}\{|\zeta_k|\geq\sqrt{n}\delta\} &\\ \leq 2\varepsilon n^{-1}+n^{-1}CE(\xi_k)^2\mathbf{1}\{|\xi_k|\geq\sqrt{n}\delta\}+n^{-1}CE(\zeta_k)^2)\mathbf{1}\{|\zeta_k|\geq\sqrt{n}\delta\}.\end{aligned}$$

Substituting this bound in inequality (3) now yields

$$\left|Ef\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i\right)-Ef\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\zeta_i\right)\right|\leq 2\varepsilon+CE(\xi_1)^2\mathbf{1}\{|\xi_1|\geq\sqrt{n}\delta\}+CE(\zeta_1)^2)\mathbf{1}\{|\zeta_1|\geq\sqrt{n}\delta\}.$$

Since $E\xi_1^2=1<\infty$ and $E\zeta_1^2=1<\infty$, the dominated convergence theorem implies that the last two expectations converge to zero as $n\to\infty$, and so

$$\limsup_{n\to\infty}\left|Ef\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\xi_i\right)-Ef\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\zeta_i\right)\right|\leq 2\varepsilon.$$

Finally, since $\varepsilon>0$ is arbitrary, the convergence (2) must hold. □

2 The Lindeberg Condition

Lindeberg's second insight was that a similar pairing of non-Gaussian, mean-zero random variables ξ_i with Gaussian, mean-zero random variables ζ_i of the same variance could be carried out even when the random variables ξ_i are not identically distributed, because sums of independent Gaussian random variables are still Gaussian, even if the summands have different variances. When variances are matched, so that $E\xi_i^2=E\zeta_i^2$, most of the proof given above carries through directly: in particular, the first two terms in the Taylor series would (after taking expectations) cancel, leaving only the remainder terms $ER_k(A)$ and $ER_k(B)$. Thus, the real issue in generalizing the central limit theorem is to formulate a hypothesis that will guarantee that the sum of the expectations $ER_k(A)$ and $ER_k(B)$ will be small.

Triangular Arrays: A *triangular array* is a doubly-indexed family $\{\xi_{n,i}\}_{n\geq1,1\leq i\leq m(n)}$ of random variables.

Lindeberg's Conditions: A triangular array $\{\xi_{n,i}\}_{n\geq1,1\leq i\leq m(n)}$ of independent random variables satisfies Lindeberg's conditions if

(A) $E\xi_{n,i}=0$ for all n,i.
(B) $\sum_{i\leq m(n)}E\xi_{n,i}^2=1$.
(C) For every $\delta>0$,

$$\lim_{n\to\infty}\sum_{i=1}^{m(n)}E\xi_{n,i}^2\mathbf{1}\{|\xi_{n,i}|\geq\delta\}=0. \qquad (6)$$

Theorem 2. *(Lindeberg's Central Limit Theorem) If $\{\xi_{n,i}\}$ is a triangular array that satisfies Lindeberg's conditions, then as $n \to \infty$*

$$\sum_{i=1}^{m(n)} \xi_{n,i} \xrightarrow{\mathscr{D}} \text{Normal}\,(0,1). \tag{7}$$

The proof is very nearly identical to Lindeberg's proof of the central limit theorem. As an exercise, you should fill in the details.

3 Martingale Central Limit Theorem

Independence is used in the proof of the central limit theorem (and of Lindeberg's generalization to triangular arrays of independent random variables) only in the evaluation of the first two expectations on the right side of equation (5). These involve only the first two conditional moments of the random variables ξ_i given the $\sigma-$algebra generated by $\xi_1, \xi_2, \ldots, \xi_{i-1}$. Lévy realized that independence is more than is needed for this purpose: in fact, only the martingale property is needed.

Assume now that each row of the triangular array $\{\xi_{n,i}\}_{i \le m(n)}$ is a martingale difference sequence, that is, for each row n there is a filtration $\{\mathscr{F}_{n,i}\}_{0 \le i \le m(n)}$ such that the sequence $\{\xi_{n,i}\}_{i \le m(n)}$ is adapted to the filtration and

$$E(\xi_{n,i} \,|\, \mathscr{F}_{n,i-1}) = 0. \tag{8}$$

Write

$$S_{n,k} = \sum_{i=1}^{k} \xi_{n,i} \quad \text{and} \quad V_{n,k}^2 = \sum_{i=1}^{k} E(\xi_{n,i}^2 \,|\, \mathscr{F}_{n,i-1}). \tag{9}$$

Theorem 3. *(P. Lévy) Assume in addition to* (8) *that the sum of the conditional variances in each row is* 1*, that is, $V_{n,m(n)}^2 = 1$, and assume that the triangular array $\{\xi_{n,i}\}_{i \le m(n)}$ satisfies the* Lindeberg condition*, that is, for every $\delta > 0$,*

$$\lim_{n\to\infty} \sum_{i=1}^{m(n)} E\xi_{n,i}^2 \mathbf{1}\{|\xi_{n,i}| \ge \delta\} = 0. \tag{10}$$

Then as $n \to \infty$,

$$S_{n,m(n)} \xrightarrow{\mathscr{D}} \text{Normal}(0,1). \tag{11}$$

REMARK. There are many variants of this theorem. In one of the more useful of these, The Lindeberg condition is replaced by the hypothesis

$$\lim_{n\to\infty} \sum_{i=1}^{m(n)} E(\xi_{n,i}^2 \mathbf{1}\{|\xi_{n,i}| \ge \delta\} \,|\, \mathscr{F}_{n,i-1}) = 0 \quad \text{in probability.} \tag{12}$$

See the book by HEYDE & HALL for a more detailed discussion. In many applications the rather strong hypothesis that $V_{n,m(n)} = 1$ is not satisfied. For this reason, the following variant of Lévy's theorem (which we will not prove) is often cited.

Theorem 4. *(Martingale Central Limit Theorem) Assume in addition to* (8) *that* (12) *holds, and that $V_{n,m(n)} \xrightarrow{P} 1$ as $n \to \infty$. Then*

$$S_{n,m(n)} \xrightarrow{\mathscr{D}} \text{Normal}(0,1). \tag{13}$$

Proof. As in the proof of the central limit theorem, it suffices to prove that for every C^∞ function f,

$$\lim_{n\to\infty} Ef(S_{n,m(n)}) = Ef(Z) \tag{14}$$

where Z is standard normal. The strategy will be the same as in Lindeberg's proof of the central limit theorem: we will match the martingale differences $\xi_{n,i}$ with independent, mean-zero, normal random variables $\zeta_{n,i}$ in such a way that the conditional variances (given $\mathscr{F}_{n,i-1}$) agree. Assume that the probability space is large enough to support all of the random variables $\xi_{n,i}$ and in addition independent standard normal random variables $Z_{n,i}$. Set

$$\zeta_{n,i} = \sigma_{n,i} Z_{n,i} \quad \text{where} \quad \sigma^2_{n,i} = E(\xi^2_{n,i} \,|\, \mathscr{F}_{n,i-1}).$$

Then

$$\left| Ef\left(\sum_{i=1}^{m(n)} \xi_{n,i}\right) - Ef\left(\sum_{i=1}^{m(n)} \zeta_{n,i}\right)\right| \le \sum_{k=1}^{m(n)} \left| Ef\left(\sum_{i=1}^{k} \xi_{n,i} + \sum_{i=k+1}^{m(n)} \zeta_{n,i}\right) - Ef\left(\sum_{i=1}^{k-1} \xi_{n,i} + \sum_{i=k}^{m(n)} \zeta_{n,i}\right)\right|. \tag{15}$$

To bound the terms on the right side of (14), we will use Taylor's theorem (4) in much the same way as earlier:

$$\begin{aligned} f&\left(\sum_{i=1}^{k} \xi_{n,i} + \sum_{i=k+1}^{m(n)} \zeta_{n,i}\right) - f\left(\sum_{i=1}^{k-1} \xi_{n,i} + \sum_{i=k}^{m(n)} \zeta_{n,i}\right) \\ &= f'\left(\sum_{i=1}^{k-1} \xi_{n,i} + \sum_{i=k+1}^{m(n)} \zeta_{n,i}\right)(\xi_k - \zeta_k) + \frac{1}{2} f''\left(\sum_{i=1}^{k-1} \xi_{n,i} + \sum_{i=k+1}^{m(n)} \zeta_{n,i}\right)(\xi_k^2 - \zeta_k^2) \\ &\qquad + R_{n,k}(A) + R_{n,k}(B) \end{aligned} \tag{16}$$

where

$$\begin{aligned} R_{n,k}(A) &\le \varepsilon(\xi_{n,k})^2 + C(\xi_{n,k})^2 \mathbf{1}\{|\xi_{n,k}| \ge \delta\} \quad \text{and} \\ R_{n,k}(B) &\le \varepsilon(\zeta_{n,k})^2 + C(\zeta_{n,k})^2 \mathbf{1}\{|\zeta_{n,k}| \ge \delta\}. \end{aligned}$$

Our next objective is to show that the expectations of the first two terms on the right side of equation (15) both vanish. In Lindeberg's proof of the central limit theorem the corresponding step was trivial, because the random variables were all independent, but here we must take care in sorting out the dependence of the various terms in the expectations. By construction, $\zeta_{n,i} = \sigma_{n,i} Z_{n,i}$ where the random variables $Z_{n,i}$ are standard normal and independent of $\mathscr{F}_{n,m(n)}$. Consequently, the conditional distribution of the random variable

$$Z'_{n,k} := \frac{\sum_{i=k+1}^{m(n)} \zeta_{n,i}}{\sqrt{1 - V^2_{n,k}}}$$

given $\mathscr{F}_{n,m(n)}$ is the standard normal distribution, and so in particular $Z'_{n,k}$ is *independent* of $\mathscr{F}_{n,m(n)}$. Moreover, since the random variables $Z_{n,i}$ are conditionally independent given $\mathscr{F}_{n,m(n)}$, the random variable $Z'_{n,k}$ is conditionally independent of $\zeta_{n,k} = \sigma_{n,k} Z_{n,k}$. Hence, for any bounded, Borel measurable function g and any nonnegative Borel measurable h,

$$E\left(g\left(\sum_{i=1}^{k-1} \xi_{n,i} + \sum_{i=k+1}^{m(n)} \zeta_{n,i}\right) h(\zeta_{n,k}) \,\middle|\, \mathscr{F}_{n,m(n)}\right) = \iint g\left(\sum_{i=1}^{k-1} \xi_{n,i} + \sqrt{1 - V^2_{n,k}}\, z_1\right) h(\sigma_{n,k} z_2)\, \varphi(z_1)\varphi(z_2)\, dz_1\, dz_2,$$

where $\varphi(z)$ is the standard normal probability density function. Applying this with $h = 1$ and $g = f'$, and using the fact that $V_{n,k}^2$ is measurable with respect to $\mathscr{F}_{n,k-1}$, gives

$$\begin{aligned} Ef'\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sum_{i=k+1}^{m(n)}\zeta_{n,i}\right)\xi_k &= E\int_{-\infty}^{\infty} f'\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sqrt{1-V_{n,k}^2}\,z\right)\xi_k\,\varphi(z)\,dz \\ &= \int_{-\infty}^{\infty} E(f'\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sqrt{1-V_{n,k}^2}\,z\right)E(\xi_k\,|,\mathscr{F}_{n,k-1})\,\varphi(z)\,dz \\ &= 0, \end{aligned}$$

and then with $h_+(x) = x \vee 0$ and $h_- = -x \vee 0$,

$$Ef'\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sum_{i=k+1}^{m(n)}\zeta_{n,i}\right)\zeta_k = E\iint f'\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sqrt{1-V_{n,k}^2}\,z_1\right)z_2\,\varphi(z_1)\varphi(z_2)\,dz_1\,dz_2 = 0.$$

Similar calculations (exercise: fill in the details) with $g = f''$ show that

$$Ef''\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sum_{i=k+1}^{m(n)}\zeta_{n,i}\right)(\xi_k^2 - \zeta_k^2) = E\int f''\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sqrt{1-V_{n,k}^2}\,z_1\right)\sigma_{n,k}^2\left(1 - \int z_2^2\varphi(z_2)\,dz_2\right) = 0.$$

Thus, equation (15) simplifies to

$$Ef\left(\sum_{i=1}^{k}\xi_{n,i} + \sum_{i=k+1}^{m(n)}\zeta_{n,i}\right) - Ef\left(\sum_{i=1}^{k-1}\xi_{n,i} + \sum_{i=k}^{m(n)}\zeta_{n,i}\right) = ER_{n,k}(A) + ER_{n,k}(B),$$

and so by (14),

$$\left|Ef\left(\sum_{i=1}^{m(n)}\xi_{n,i}\right) - Ef\left(\sum_{i=1}^{m(n)}\zeta_{n,i}\right)\right| \le \sum_{k=1}^{m(n)}|ER_{n,k}(A) + ER_{n,k}(B)|.$$

The Lindeberg condition (10) implies that for any $\delta > 0$ the limsup of the right side is $\le 2\delta$. Since $\delta > 0$ is arbitrary, the result (13) follows. □