An mgf proof of the central limit theorem
STK4011 Autumn 2019

Emil Aas Stoltenberg
Department of Mathematics, University of Oslo
November 21, 2019

Here is a theorem that is often called the Lindeberg–Lévy central limit theorem. The proof I present here is an ε-generalisation of the proof found in Inlow (2010).

Theorem 1. *Let $X_1,\ldots,X_n$ be independent random variables with mean zero and variance $\mathrm{E}\,X_j^2 = \sigma_j^2 < \infty$ for $j = 1,\ldots,n$. Define $B_n^2 = \sum_{j=1}^n \sigma_j^2$ and $Z_n = \sum_{j=1}^n X_j$. We are going to show that*

$$\frac{Z_n}{B_n} \overset{d}{\to} \mathrm{N}(0,1), \tag{1}$$

provided the Lindeberg condition, that for every $\varepsilon > 0$,

$$L_n(\varepsilon) \coloneqq \frac{1}{B_n^2}\sum_{j=1}^n X_j^2 I\{|X_j| \geq \varepsilon B_n\} \to 0, \quad \text{as } n \to \infty, \tag{2}$$

is in force.

Normally, this result is proved using characteristic functions. In this note, we'll prove this theorem using moment generating functions, but without assuming that the moment generating functions of the X_j's exist. Recall that by assuming that the mgf exists, we are in effect assuming that the moments of all orders exist. That, we do not want to assume.

Proof. For every $\varepsilon > 0$, we can write

$$X_j = X_j I\{|X_j| < \varepsilon B_n\} + X_j I\{|X_j| \geq \varepsilon B_n\}.$$

Define $\xi_{n,j}$ by

$$\xi_{n,j} = \mathrm{E}\,X_j I\{|X_j| < \varepsilon B_n\} = -\mathrm{E}\,X_j I\{|X_j| \geq \varepsilon B_n\}.$$

where we use that $\mathrm{E}\,X_j = 0$. Taking plus-minus $\xi_{n,j}$,

$$X_j = (X_j I\{|X_j| < \varepsilon B_n\} - \xi_{n,j}) + (X_j I\{|X_j| \geq \varepsilon B_n\} + \xi_{n,j}) =: V_{n,j} + W_{n,j},$$

by which we define $V_{n,j}$ and $W_{n,j}$, and note that these have $\mathrm{E}\,V_{n,j} = 0$ and $\mathrm{E}\,V_{n,j} = 0$. With this notation,

$$\frac{Z_n}{B_n} = \frac{1}{B_n}\sum_{j=1}^n V_{n,j} + \frac{1}{B_n}\sum_{j=1}^n W_{n,j}.$$

We are now going to show that $\sum_{j=1}^n V_{n,j}/B_n$ converges in distribution to a standard normal, and that $\sum_{j=1}^n W_{n,j}/B_n$ converges in probability to zero. Then (2) follows from the Slutsky–Cramér rules, according to which $V_n \to_d V$ and $W_n \to_p w$ implies $V_n + W_n \to_d V + w$.

Note that $V_{n,j}/B_n$ is a bounded random variable,

$$\begin{aligned}|V_{n,j}/B_n| &= |X_j I\{|X_j| < \varepsilon B_n\} - \xi_{n,j}|/B_n = \big|X_j I\{|X_j| < \varepsilon B_n\} - \mathrm{E}\,X_j I\{|X_j| < \varepsilon B_n\}\big|/B_n \\ &\leq X_j I\{|X_j| < \varepsilon B_n\}\big|/B_n + \big|\mathrm{E}\,X_j I\{|X_j| < \varepsilon B_n\}\big|/B_n = 2\varepsilon.\end{aligned}$$

If Y is a random variable bounded by $K < \infty$, then its moment generating function clearly exists, because $M_n(t) = \mathrm{E}\,e^{tY} \leq \mathrm{E}\,e^{|tY|} \leq e^{|t|K} < \infty$. So, since $|V_{n,j}/B_n| \leq 2\varepsilon$ its moment generating function exists, and is given by

$$M_{n,j}(t) = 1 + \frac{1}{2}\mathrm{E}\,(V_{n,j}/B_n)^2 + r_{n,j}(t) = 1 + \frac{1}{2}\mathrm{E}\,\sigma_{n,j}^2/B_n^2 + r_{n,j}(t)$$

where $\sigma_{n,j}^2 = \mathrm{Var}\,V_{n,j}$ and

$$r_{n,j}(t) = \frac{t^3}{6}\mathrm{E}\,\{(V_{n,j}/B_n)^3\exp(bV_{n,j}/B_n)\},$$

for some b between t and zero. Note that,

$$\sigma_j^2 = \mathrm{Var}\, X_j = \mathrm{Var}\, V_{n,j} + \mathrm{Var}\, W_{n,j} \geq \sigma_{n,j}^2.$$

Moreover, let $|t| \leq 1$ so that $|b| < 1$, then using the bound on $V_{n,j}/B_n$ we found above,

$$\begin{aligned} |r_{n,j}(t)| &\leq \frac{1}{6}\mathrm{E}\,\{|V_{n,j}/B_n|(V_{n,j}/B_n)^2 \exp(|V_{n,j}/B_n|)\} \\ &\leq \frac{\varepsilon e^{2\varepsilon}}{3}\frac{1}{B_n^2}\, E\, V_{n,j}^2 \leq \frac{\varepsilon e^{2\varepsilon}}{3}\frac{\sigma_j^2}{B_n^2}. \end{aligned}$$

Since the $V_{n,1}/B_n, \ldots, V_{n,n}/B_n$ are independent, the mgf of $\sum_{j=1}^n V_{n,j}/B_n$ is $M_n(t) = \prod_{j=1}^n M_{n,j}(t)$, so that

$$M_n(t) = \prod_{i=1}^n \{1 + \sigma_{n,j}^2/B_n + r_{n,j}(t)\} = \prod_{i=1}^n \{1 + z_{n,j}\} \tag{3}$$

where $z_{n,j} = \sigma_{n,j}^2/B_n + r_{n,j}(t)$. From Lemma A.1 on page 1290 in Nils' 1990 Beta paper article, we need that (i) $\sum_{j=1}^n z_{n,j} \to z$; (ii) that $\max_{j \leq n} |z_{n,j}| \to 0$; and (iii) that $\limsup_{n\to\infty} \sum_{j=1}^n |z_{n,j}|$ is bounded. For (i) we are going to show that

$$\sum_{j=1}^n z_{n,j} = \sum_{j=1}^n \frac{\sigma_{n,j}^2}{B_n^2} + \sum_{j=1}^n r_{n,j}(t) \to 1 + 0 = 1, \tag{4}$$

as $n \to \infty$. Looking back at the definitions of $V_{n,j}$ and $W_{n,j}$,

$$\begin{aligned} \sigma_{n,j}^2 &= \mathrm{E}\, V_{n,j}^2 = \mathrm{E}\,(X_j - W_{n,j})^2 = \sigma_j^2 - 2\mathrm{E}\, X_j W_{n,j} + \mathrm{E}\, W_{n,j}^2 \\ &= \sigma_j^2 - 2\mathrm{E}\, X_j^2 I\{|X_j| \geq \varepsilon B_n\} + \mathrm{E}\, W_{n,j}^2, \end{aligned}$$

thus

$$\Big|\sum_{j=1}^n \frac{\sigma_{n,j}^2}{B_n^2} - 1\Big| \leq 2L_n(\varepsilon) + \frac{1}{B_n^2}\sum_{j=1}^n \mathrm{E}\, W_{n,j}^2$$

Recalling that $\xi_{n,j} = -\mathrm{E}\, X_j I\{|X_j| \geq \varepsilon B_n\}$ and using Jensen's inequality

$$\mathrm{E}\, W_{n,j}^2 = \mathrm{E}\, X_j^2 I\{|X_j| \geq \varepsilon B_n\} + \xi_{n,j}^2 \leq 2\mathrm{E}\, X_j^2 I\{|X_j| \geq \varepsilon B_n\}.$$

from which we see that

$$\frac{1}{B_n^2}\sum_{j=1}^n \mathrm{E}\, W_{n,j}^2 \leq 2L_n(\varepsilon),$$

so this term goes to zero by the Lindeberg condition, and we conclude that $\sum_{j=1}^n \sigma_{n,j}^2/B_n^2 \to 1$. In addition,

$$\Big|\sum_{j=1}^n r_{n,j}(t)\Big| \leq \frac{\varepsilon e^{2\varepsilon}}{3}\frac{\sum_{j=1}^n \sigma_j^2}{B_n^2} \leq \frac{\varepsilon e^{2\varepsilon}}{3}, \tag{5}$$

which, since $\varepsilon > 0$ was arbitrary, can be made arbitrarily small. This takes care of the $\sum_{j=1}^n z_{n,j} \to z$ part of the lemma needed for convergence in (3). Clearly, $\max_{j\leq n} r_{n,j}(t) \to 0$ because $\varepsilon > 0$ is arbitrary. This establishes (4).

For the second condition of the lemma, we need that

$$\frac{\max_{j\leq n} \sigma_j^2}{B_n^2} \to 0, \quad \text{as } n \to \infty. \tag{6}$$

But for any $\varepsilon > 0$

$$\sigma_j^2 = \mathrm{E}\, X_j^2 \leq \varepsilon^2 B_n^2 + \mathrm{E}\, X_j^2 I\{|X_j| \geq \varepsilon B_n\} \leq \varepsilon^2 B_n^2 + B_n^2 L_n(\varepsilon).$$

Divide through by B_n^2 and take the maximum on both sides, and (6) follows because $\varepsilon > 0$ is arbitrary.

That the third condition of the lemma is satisfied follows from (5) and (6). In conclusion, the moment generating function of $\sum_{j=1}^n V_{n,j}/B_n$ tends to that of a standard normal, that is

$$M_n(t) \to e^{-t^2/2}, \quad \text{as } n \to \infty.$$

It remains to show that $\sum_{j=1}^n W_{n,j}/B_n$ converges in probability to zero. Since the $W_{n,1},\dots,W_{n,n}$ are independent, this can be done by an application of Chebyshev's inequality: For any $a>0$,

$$\Pr(|\sum_{j=1}^n W_{n,j}/B_n| \geq a) \leq \frac{1}{a^2B_n^2}\sum_{j=1}^n \operatorname{Var} W_{n,j} \leq \frac{2}{a^2B_n^2}\sum_{j=1}^n \operatorname{E} X_j^2 I\{|X_j|\geq \varepsilon B_n\} = \frac{2}{a}L_n(\varepsilon).$$

where the second to last inequality is Jensen's, and the right hand side tends to zero because the Lindeberg condition holds. □

For completeness, here is the lemma from Nils' Beta process article that was used in the proof.

Lemma 2. *Let $z_{n,j}$ be a sequence of real number so that* (i) $\sum_{j=1}^n z_{n,j}\to z$*;* (ii) $\max_{j\leq n}|z_{n,j}|\to 0$*; and* (iii) $\limsup_n \sum_{j=1}^n |z_{n,j}| \leq M < \infty$*, i.e., that the series is absolutely convergent. Then*

$$\prod_{j=1}^n (1+z_{n,j}) \to \exp(z).$$

Proof. We have that

$$\begin{aligned}\log(1+z) &= \sum_{k=1}^\infty (-1)^{k+1}z^k/k = z - z^2/2 + z^3/3 - z^4/4 + \cdots \\ &= z + z^2\{-1/2 + z/3 - z^2/4 + \cdots\}z + z^2K(z),\end{aligned}$$

where $K(z) = -1/2 + z/3 - z^2/4 + \cdots$. Note that whenever $|z|\leq 1/2$,

$$|K(z)| \leq 1/2 + 1/6 + 1/16 + \cdots = 1/2(1 + 1/3 + 1/4 + \cdots) = 1/2.$$

Since $\max_{j\leq n}|z_{n,j}|\to 0$, all the $|z_{n,j}|\leq 1/2$ for n large enough. Now, for n big enough for $|K(z_{n,j})|\leq 1/2$ for all j,

$$\sum_{j=1}^n z_{n,j}^2 |K(z_{n,j})| \leq \max_{j\leq n}|z_{n,j}| \sum_{j=1}^n |z_{n,j}|,$$

where the right hand side tends to zero by assumption (ii) and (iii). Thus $\sum_{j=1}^n \log(1+z_{n,j})\to z$ and the result follows because $\exp(z)$ is a continuous function. □

Applications

Example 1. (Independent zero-ones) Let $X_1,\dots,X_n$ be independent Bernoulli with success probabilities $p_1,\dots,p_n$. Let $B_n = \{\sum_{i=1}^n p_i(1-p_i)\}^{1/2}$, then Z_n given by

$$Z_n = \frac{\sum_{i=1}^n (X_i - p_i)}{B_n} \overset{d}{\to} \mathrm{N}(0,1).$$

The $X_1-p_1,\dots,X_n-p_n$ are mean zero with variance $p_i(1-p_i)$. It suffices to check the Lindeberg condition. Here we can simply use that

$$\begin{aligned}\operatorname{E}(X_i-p_i)^2 I\{|X_i-p_i|\geq \varepsilon B_n\} &= p_i(1-p_i)^2 I\{|1-p_i|\geq \varepsilon B_n\} + (1-p_i)p_i^2 I\{|p_i|\geq \varepsilon B_n\} \\ &\leq p_i(1-p_i)\big(I\{|1-p_i|\geq \varepsilon B_n\} + I\{|p_i|\geq \varepsilon B_n\}\big) \\ &\leq p_i(1-p_i) I\{1\geq \varepsilon B_n\}.\end{aligned}$$

Then for any $\varepsilon>0$,

$$\frac{1}{B_n^2}\sum_{i=1}^n \operatorname{E}(X_i-p_i)^2 I\{|X_i-p_i|\geq \varepsilon B_n\} \leq I\{1\geq \varepsilon B_n\},$$

which tends to zero provided $B_n\to\infty$. If $B_n\to\infty$, then $B_n^2\to\infty$, and

$$B_n^2 = \sum_{i=1}^n p_i(1-p_i) = \sum_{i=1}^n \{p_i - p_i^2\} \to \infty,$$

which may only happen if $\sum_{i=1}^n p_i\to\infty$. In effect, since both $\sum_{i=1}^n p_i$ and $\sum_{i=1}^n(1-p_i)$ are bigger than B_n^2, we'll have that they both tend to infinity. To see that we cannot have Z_n converging to a standard normal if $\sum_{i=1}^n p_i < \infty$, note that by the Borel–Cantelli lemma $\sum_{i=1}^n p_i<\infty$ implies that $P(X_i = 1 \,\text{infinitely often}) = 0$. Consequently, there exists an n_0 such that for all $n\geq n_0$, $X_n = 0$. This entails that for all $n \geq n_0$ we'll have $Z_n = \sum_{i=1}^{n_0}(X_i - p_i)/B_n + \sum_{i=n_0+1}^{n}(-p_i)/B_n$, in which case Z_n obviously does not converge to a standard normal. The same conclusion can be drawn assuming that $\sum_{i=1}^{n}(1-p_i) < \infty$, showing that a necessary condition for $Z_n \to_d \mathrm{N}(0,1)$ is that the p_i are bounded away from zero and one.

Note that the Lindeberg condition implies that $B_n \to \infty$. Assume that $B_n \uparrow B_{\max} < \infty$, and think of $\varepsilon = \varepsilon'/B_{\max}$ for some $\varepsilon' > 0$ if necessary (i.e. if the X_i's are bounded rv's), then

$$L_n(\varepsilon) \geq \varepsilon^2 n \Pr(|X_i| \geq \varepsilon B_n) \geq \varepsilon^2 n \Pr(|X_i| \geq \varepsilon B_{\max}),$$

which diverges when $n \to \infty$.

Example 2. (Simple linear regression) Consider the simple linear regression model

$$Y_i = \beta x_i + \zeta_i, \quad i = 1, \ldots, n$$

where $x_1, \ldots, x_n$ are fixed and known covariates and the $\zeta_1, \ldots, \zeta_n$ are independent random variables with mean zero and finite second moments $\mathrm{E}\,\zeta_i^2 = \sigma^2$. We are interested in conditions for convergence in distribution of $\widehat{\beta}_n - \beta$, properly normalised, where $\widehat{\beta}_n$ is the least squares estimator

$$\widehat{\beta}_n = \frac{\sum_{i=1}^n x_i Y_i}{\sum_{i=1}^n x_i^2}.$$

Define the random variables

$$\xi_i = x_i \zeta_i, \quad \text{for } i = 1, \ldots, n,$$

and we can write

$$\{\sum_{i=1}^n x_i^2\}(\hat{\beta}_n - \beta) = \sum_{i=1}^n \xi_{n,i}.$$

The $\xi_1, \ldots, \xi_n$ are independent, mean zero, with variance

$$\mathrm{Var}\,\xi_i = x_i^2 \sigma^2,$$

Define B_n^2 by

$$B_n^2 = \sum_{i=1}^n \mathrm{Var}\,\xi_{n,i} = \sigma^2 \sum_{i=1}^n x_i^2,$$

and recall that $\delta_n^2 = \max_{i \leq n} x_i^2 / \{\sum_{i=1}^n x_i^2\}$ with $\delta_n \to 0$ by assumption. Check the Lindeberg condition. For any $\varepsilon > 0$,

$$\begin{aligned}\frac{1}{B_n^2}\sum_{i=1}^n \mathrm{E}\,\xi_{n,i}^2 I\{|\xi_{n,i}| \geq \varepsilon B_n\} &= \frac{1}{B_n^2}\sum_{i=1}^n x_i^2 \mathrm{E}\,\zeta_i I\{|\zeta_i| \geq \varepsilon B_n/|x_i|\} \\ &\leq \frac{1}{B_n^2}\sum_{i=1}^n x_i^2 \mathrm{E}\,\zeta_i^2 I\{|\zeta_i| \geq \varepsilon\sigma/\delta_n\} = (1/\sigma^2)\mathrm{E}\,\zeta_1^2 I\{|\zeta_1| \geq \varepsilon\sigma/\delta_n\},\end{aligned}$$

because the $\zeta_1, \ldots, \zeta_n$ are i.i.d. the expectation part goes outside the sum. As $\delta_n \to 0$, the right hand side tends to zero, and the Lindeberg condition is satisfied. From the central limit theorem we had proved in lecture 14. November, this means that

$$(\sum_{i=1}^n x_i^2)^{1/2}(\hat{\beta}_n - \beta) = \frac{\sigma}{B_n}\sum_{i=1}^n \xi_i \overset{d}{\to} \mathrm{N}(0, \sigma^2).$$

This means that for n large enough,

$$\hat{\beta}_n - \beta \approx_d \mathrm{N}(0, \frac{\sigma^2}{\sum_{i=1}^n x_i^2}),$$

which we recognise as the exact distribution of $\hat{\beta}_n - \beta$ when the $\zeta_1, \ldots, \zeta_n$ are independent $\mathrm{N}(0, \sigma^2)$.

Exercise 1. Generalise the example above to the regression model $Y_i = \beta_0+\beta_1 x_i+\zeta_i$, where $\zeta_1,\ldots,\zeta_n$ are i.i.d. with $\mathrm{E}\,\zeta_i = 0$ and $\mathrm{Var}\,\zeta_i^2 = \sigma^2 < \infty$. Do this by defining $\xi_{n,i} = \zeta_i(x_i - \bar{x}_n)$, where $\bar{x}_n = n^{-1}\sum_{i=1}^n x_i$, and so on. You'll then discover why the Linderberg–Lévy theorem is typically stated in terms of triangular arrays

$$\begin{matrix} \xi_{1,1} & & \\ \xi_{2,1} & \xi_{2,2} & \\ \xi_{3,1} & \xi_{3,2} & \xi_{3,3} \\ \vdots & & \ddots \end{matrix}$$

See, for example, Ferguson (1996) for such a statement of the theorem.

1. Liapunov, skewness, and more

The skewness $\mathrm{E}\,Z^3$ of a standard normal random variable is zero. The Lindeberg–Lévy theorem is often stated in terms of the Liapunov condition: For some $\delta > 0$,

$$\frac{1}{B_n^{2+\delta}}\sum_{i=1}^n \mathrm{E}\,|X_i|^{2+\delta} \to 0, \tag{7}$$

when $n \to \infty$. Let $\varepsilon > 0$, then

$$\mathrm{E}\,|X_i|^{2+\delta} \geq \mathrm{E}\,|X_i|^{2+\delta} I\{|X_i| \geq \varepsilon B_n\} \geq \varepsilon^\delta B_n^\delta \mathrm{E}\,|X_i|^2 I\{|X_i| \geq \varepsilon B_n\},$$

so that

$$\frac{1}{B_n^2}\sum_{i=1}^n \mathrm{E}\,X_i^2 I\{|X_i| \geq \varepsilon B_n\} \leq \frac{1}{\varepsilon^\delta}\frac{1}{B_n^{2+\delta}}\sum_{i=1}^n \mathrm{E}\,|X_i|^{2+\delta},$$

and the Liapunov condition is seen to imply the Lindeberg condition. Since $\mathrm{E}\,X_i^3 \leq \mathrm{E}\,|X_i|^3$, this also shows that of the skewness of the sequence $X_1,\ldots,X_n$ tends to 0, the Lindeberg condition is satisfied.

Example 3. Here is an interesting skewness-related example from Li et al. (2014). Suppose $X_1,\ldots,X_n$ are i.i.d. mean zero random variables with variance σ^2, skewness $\gamma = \mathrm{E}\,(X_i/\sigma)^3$ and kurtosis $\kappa = \mathrm{E}\,(X_i/\sigma)^4$. Recall that the kurtosis is always bigger than one (use Jensen's inequality), and that for the normal distribution the kurtosis is 3. The estimator

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^n X_i^2,$$

is unbiased, and

$$\sqrt{n}(\widehat{\sigma}_n^2 - \sigma^2) \overset{d}{\to} \mathrm{N}(0, \sigma^4(\kappa-1)),$$

provided the kurtosis is finite. Let

$$\widehat{\beta}_n = \frac{\sum_{i=1}^n X_i^3}{\sum_{i=1}^n X_i^2},$$

be an estimator of the regression of the X_i^2 on X_i, that is

$$\beta = \mathrm{Cov}(X_i^2, X_i)/\mathrm{Var}(X_i) = \mathrm{E}\,X_i^3/\mathrm{E}\,X_i^2 = \sigma\gamma.$$

Clearly, $\hat{\beta}_n \to_p \beta = \sigma\gamma$, and if the skewness is zero, then $\beta = 0$. Consider the estimator

$$\widehat{\sigma}^2_{\mathrm{skew},n} = \frac{1}{n}\sum_{i=1}^n X_i^2 - \widehat{\beta}_n \frac{1}{n}\sum_{i=1}^n X_i = \frac{1}{n}\sum_{i=1}^n (X_i^2 - \widehat{\beta}_n X_i).$$

This estimator is consistent for σ^2. Note that

$$\frac{1}{n}\sum_{i=1}^n \mathrm{Cov}(X_i^2, \widehat{\beta}_n X_i) = \frac{1}{n}\sum_{i=1}^n \mathrm{Cov}(X_i^2 - \widehat{\beta}_n X_i + \widehat{\beta}_n X_i, \widehat{\beta}_n X_i) = 0 + \frac{1}{n}\sum_{i=1}^n \mathrm{Var}\,\widehat{\beta}_n X_i,$$

since $\widehat{\beta}_n$ is the least squares solution. Moreover,

$$\frac{1}{n}\sum_{i=1}^n \mathrm{Var}\,\widehat{\beta}_n X_i = \beta^2\sigma^2 + \frac{1}{n}\sum_{i=1}^n \{\mathrm{E}\,(\widehat{\beta}_n - \beta)^2 X_i^2 + \mathrm{E}\,(\widehat{\beta}_n - \beta)\beta X_i^2\},$$

which converges to $\beta^2\sigma^2 = \sigma^4\gamma^2$. [xx Some care is needed to show this. Include it xx]. The variance of $\widehat{\sigma}^2_{\mathrm{skew},n}$ is therefore,

$$\begin{aligned}\widehat{\sigma}^2_{\mathrm{skew},n} &= \frac{1}{n^2}\sum_{i=1}^n \big\{\mathrm{Var}\,X_i^2 + \mathrm{Var}\,\widehat{\beta}_n X_i - 2\mathrm{Cov}(X_i^2, \widehat{\beta}_n X_i)\big\} \\ &= \frac{\sigma^4(\kappa-1)}{n} - \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}\,\widehat{\beta}_n X_i \to \frac{\sigma^4(\kappa - 1 - \gamma^2)}{n},\end{aligned}$$

as $n \to \infty$, and

$$\sqrt{n}(\widehat{\sigma}^2_{\mathrm{skew},n} - \sigma^2) \overset{d}{\to} \mathrm{N}\{0, \sigma^4(\kappa - 1 - \gamma^2)\}.$$

This shows that if the X_i's are skewed, then $\widehat{\sigma}^2_{\mathrm{skew},n}$ improves on $\widehat{\sigma}^2_n$ in terms of asymptotic efficiency.

Exercise 2. The example above needs the dominated convergence theorem for the situation where we only have convergence in probability. Prove that if $Z_n \overset{p}{\to} Z$, and $|Z| \leq Y$ for some Y with $\mathrm{E}\,Y < \infty$, the $\mathrm{E}\,Z_n \to \mathrm{E}\,Z$.

2. Mixing, middle ground, etc.

Suppose $X_1, X_2, \ldots$ is a sequence of mean zero random variables with variance $\mathrm{E}\,X_i^2 = \sigma^2$, and

$$\mathrm{Cov}\,(X_j, X_l) = \sigma^2\rho^{|j-l|},$$

for some correlation coefficient $\rho \in (-1, 1)$. To gain insight, suppose that the mgf of these X_i's exists, and that they satisfy the Liapunov condition (7) for $\delta = 1$. With

$$Z_n = n^{-1/2}\sum_{i=1}^n X_i,$$

we then have that

$$M_{Z_n}(t) = 1 + \frac{1}{2}t^2\mathrm{E}\,Z_n^2 + \frac{t^3}{6}\mathrm{E}\,\{Z_n^3\exp(bZ_n)\},$$

for some b between zero and t. The Liapunov condition ensures that the last term on the right vanishes, and

$$n\,\mathrm{E}\,Z_n^2 = n\sigma^2 + \frac{\sigma^2\rho(n-1)}{1-\rho} - \frac{\sigma^2\rho^2(1-\rho^{n-1})}{(1-\rho)^2}.$$

From which we see that

$$Z_n \overset{d}{\to} \mathrm{N}\big(0, \sigma^2\frac{1+\rho}{1-\rho}\big),$$

as $n \to \infty$. This result holds under weaker conditions than those employed here, by way of characteristic functions or the techniques used to prove Theorem 1. The Liapunov condition with $\delta = 1$ is probably also stronger than needed.

Example 4. (Middle ground asymptotics) Suppose that $0 = t_{n,0} < t_{n,1} < \cdots < t_{n,n} = \tau_n$ are equidistantly spaced observation times at which we see i.i.d. $X_{t_{n,0}}, \ldots, X_{t_{n,n}}$ with mean zero, variance σ^2, and covariance

$$\mathrm{Cov}(X_{t_{n,i}}, X_{t_{n,j}}) = \sigma^2\rho^{|t_{n,j} - t_{n,i}|},$$

where $\rho = \rho(a) = \exp(-a)$ for some positive parameter a. Set $\Delta_n = \tau_n/n$ and assume that $t_{n,j} = j\Delta_n$. Let now,

$$Z_n = \sum_{i=0}^n X_i.$$

Under the same assumptions as above, which can be weakened, we see that

$$\mathrm{E}\,Z_n^2 = \sigma^2\big\{(n+1) + \frac{n\rho(a)^{\Delta_n}}{1-\rho(a)^{\Delta_n}} - \frac{\rho(a)^{2\Delta_n}\{1-\rho(a)^{n\Delta_n}\}}{\{1-\rho(a)^{\Delta_n}\}^2}\big\}.$$

Suppose that for some fixed $\tau > 0$ and $0 < \alpha < 1$,

$$\tau_n = n^\alpha\tau,$$

where we, for concreteness, take $\alpha = 1/2$. Then $\Delta_n = \tau/n^{1/2}$, and consequently

$$\frac{\Delta_n}{n}\,\mathrm{E}\,Z_n^2 = \sigma^2\Big\{\Delta_n\frac{n+1}{n} + \frac{\Delta_n\rho(a)^{\Delta_n}}{1-\rho(a)^{\Delta_n}} - \frac{\Delta_n}{n}\frac{\rho(a)^{2\Delta_n}\{1-\rho(a)^{n\Delta_n}\}}{\{1-\rho(a)^{\Delta_n}\}^2}\Big\}.$$

The first term tends to zero, and the second term to σ^2/a. Since $\sqrt{n} = \tau/\Delta_n$, the third term is

$$\frac{\Delta_n^3}{\tau^2}\frac{\rho(a)^{2\Delta_n}\{1-\rho(a)^{\sqrt{n}\tau}\}}{\{1-\rho(a)^{\Delta_n}\}^2},$$

which vanishes. In conclusion,

$$Z_n = (\Delta_n/n)^{1/2}\sum_{i=0}^{n} X_i \xrightarrow{d} \mathrm{N}(0,\sigma^2/a).$$

References

Ferguson, T. S. (1996). *A course in large sample theory.* Chapman & Hall/CRC.

Hjort, N. L. (1990). Nonparametric Bayes estimators based on Beta processes in models for life history data. *The Annals of Statistics*, 18:1259–1294.

Inlow, M. (2010). A moment generating function proof of the Lindeberg–Lévy central limit theorem. *The American Statistician*, 64:228–230.

Li, Y., Mykland, P. A., Renault, E., Zhang, L., and Zheng, X. (2014). Realized volatility when sampling times are possibly endogenous. *Econometric Theory*, 30:580–605.