Math/Stat 394, Winter 2019
F.W. Scholz

Central Limit Theorems and Proofs

The following gives a self-contained treatment of the central limit theorem (CLT). It is based on Lindeberg's (1922) method. To state the CLT which we shall prove, we introduce the following notation. We assume that $X_{n1}, \ldots, X_{nn}$ are independent random variables with means 0 and respective variances $\sigma_{n1}^2, \ldots, \sigma_{nn}^2$ with

$$\sigma_{n1}^2 + \ldots + \sigma_{nn}^2 = \tau_n^2 > 0 \text{ for all } n$$

Denote the sum $X_{n1} + \ldots + X_{nn}$ by S_n and observe that S_n has mean zero and variance τ_n^2, see Fact 8.28 and 8.31 in Anderson et al.

Lindeberg's Central Limit Theorem:
If the Lindeberg condition is satisfied, i.e., if for every $\epsilon > 0$ we have that

$$L_n(\epsilon) = \frac{1}{\tau_n^2} \sum_{i=1}^{n} E\left(X_{ni}^2 I_{\{|X_{ni}| \geq \epsilon\tau_n\}}\right) \longrightarrow 0 \text{ as } n \to \infty ,$$

then for every $a \in R$ we have that

$$P\left(S_n/\tau_n \leq a\right) - \Phi(a) \longrightarrow 0 \text{ as } n \to \infty$$

Proof: Step 1 (convergence of expectations of smooth functions): We will show in Appendix 1 that for certain functions f we have that

$$E\left[f\left(S_n/\tau_n\right)\right] - E\left[f\left(Z\right)\right] \to 0 \text{ as } n \to \infty , \tag{1}$$

where Z denotes a standard normal random variable. If this convergence would hold for any function f and if we then applied it to

$$f_a(x) = I_{(-\infty,a]}(x) = 1 \quad \text{if } x \leq a \text{ and } = 0 \text{ if } x > a,$$

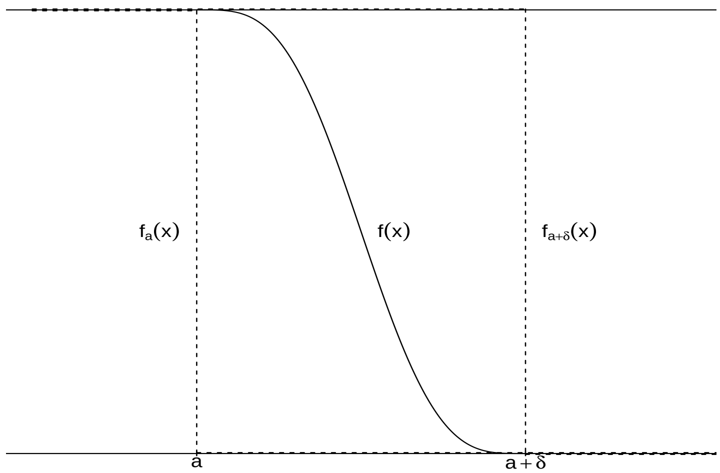
then

$$E\left[f_a\left(S_n/\tau_n\right)\right] - E\left[f_a\left(Z\right)\right] = P\left(S_n/\tau_n \leq a\right) - \Phi(a)$$

and the statement of the CLT would follow. Unfortunately, we cannot directly demonstrate the above convergence (1) for all f, but only for smooth f. Here smooth f means that f is bounded and has three bounded, continuous derivatives as stipulated in Lemma 1 of Appendix 1.

Step 2 (sandwiching a step function between smooth functions): We will approximate $f_a(x) = I_{(-\infty,a]}(x)$, which is a step function with step at $x = a$, by sandwiching it between two smooth functions. In fact, for $\delta > 0$ one easily finds (see Appendix 2 for an explicit example) smooth functions $f(x)$ with $f(x) = 1$ for $x \leq a$, $f(x)$ monotone decreasing from 1 to 0 on $[a, a+\delta]$ and $f(x) = 0$ for $x \geq a + \delta$. Hence we would have

$$f_a(x) \leq f(x) \leq f_{a+\delta}(x) \text{ for all } x \in R$$

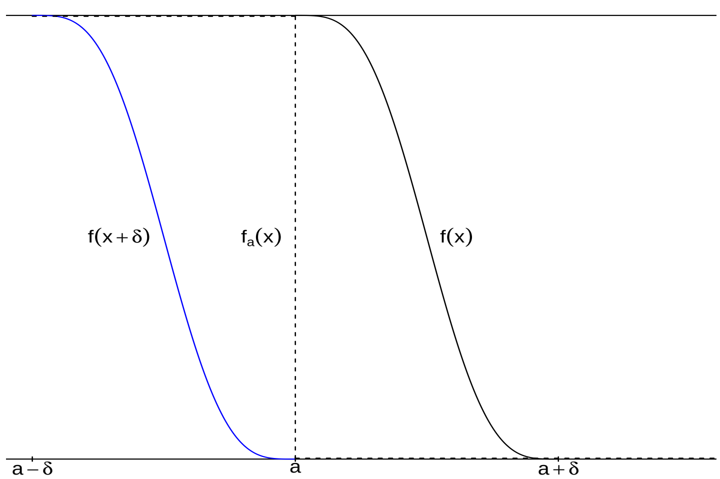


Since $f_{a+\delta}(x+\delta) = f_a(x)$, we also get from the second previous "$\leq$"

$$f_a(x) = f_{a+\delta}(x+\delta) \geq f(x+\delta)$$

Combining these, we can bracket $f_a(x)$ for all $x \in R$ by

$$f(x+\delta) \leq f_a(x) \leq f(x)$$



Step 3 (the approximation argument): The following diagram will clarify the approximation strategy. The inequalities result from the bracketing in the previous step.

$$\begin{array}{ccccc}
E\left[f\left(\frac{S_n}{\tau_n}+\delta\right)\right] & \leq & E\left[f_a\left(\frac{S_n}{\tau_n}\right)\right] & \leq & E\left[f\left(\frac{S_n}{\tau_n}\right)\right] \\
\wr\wr & & & & \wr\wr \\
E\left[f\left(Z+\delta\right)\right] & \leq & E\left[f_a\left(Z\right)\right] & \leq & E\left[f\left(Z\right)\right]
\end{array}$$

$\wr\wr$ means that for any fixed $\delta > 0$ and for fixed f the terms above and below $\wr\wr$ become arbitrarily close, say within $\epsilon/3$ of each other, as $n \to \infty$ (see Step 1). Since $f(Z+\delta) - f(Z) = 0$ for

$Z \notin [a-\delta, a+\delta]$ and $|f(Z+\delta)-f(Z)| \le 1$ for $a-\delta \le Z \le a+\delta$ we have[1]

$$\begin{aligned}|E[f(Z+\delta)]-E[f(Z)]| &= \left|E\left[\left(f(Z+\delta)-f(Z)\right) I_{[a-\delta \le Z \le a+\delta]}\right]\right| \le E\left|\left[I_{[a-\delta \le Z \le a+\delta]}\right]\right| \\ &= P(a-\delta \le Z \le a+\delta) = \Phi(a+\delta)-\Phi(a-\delta) \\ &\le \Phi(\delta)-\Phi(-\delta) \le 2\delta\phi(0) = 2\delta/\sqrt{2\pi}\end{aligned}$$

The latter bound can be made as small as $\epsilon/3$, no matter what f is, by taking $\delta = \epsilon\sqrt{2\pi}/6$. With this choice of δ, and n sufficiently large, the above diagram entails that

$$\left|E\left[f_a\left(\frac{S_n}{\tau_n}\right)\right] - E\left[f_a(Z)\right]\right| \le \frac{\epsilon}{3}+\frac{\epsilon}{3}+\frac{\epsilon}{3} = \epsilon$$

Since this can be done for any $\epsilon > 0$ we have shown that

$$\left|E\left[f_a\left(\frac{S_n}{\tau_n}\right)\right] - E\left[f_a(Z)\right]\right| \to 0 \quad \text{as} \quad n \to \infty \quad \square$$

From Lindeberg's CLT other special versions follow at once. The first, the Lindeberg-Levy CLT, considers independent *identically* distributed random variables with finite variance σ^2. The second, the Liapunov CLT, considers independent, but not necessarily identically distributed random variables with *finite third moments*.

Lindeberg-Levy CLT: Let $Y_1, \ldots, Y_n$ be independent and identically distributed random variables with common mean μ and finite positive variance σ^2 and let $T_n = Y_1 + \ldots + Y_n$. Then

$$\text{for all } a \in R \qquad P\left(\frac{T_n - n\mu}{\sqrt{n}\sigma} \le a\right) - \Phi(a) \longrightarrow 0 \quad \text{as} \quad n \to \infty$$

Proof: Let $X_i = Y_i - \mu$, then X_i has mean zero and variance σ^2. Note that $S_n = X_1 + \ldots + X_n = T_n - n\mu$ has variance $\tau_n^2 = n\sigma^2$ so that

$$\frac{T_n - n\mu}{\sqrt{n}\sigma} = \frac{S_n}{\tau_n}$$

The Lindeberg function $L_n(\epsilon)$ becomes

$$L_n(\epsilon) = \frac{1}{n\sigma^2}\sum_{i=1}^{n} E\left[X_i^2 I_{[|X_i| \ge \epsilon\sqrt{n}\sigma]}\right] = \frac{1}{n\sigma^2} nE\left[X_1^2 I_{[|X_1| \ge \epsilon\sqrt{n}\sigma]}\right] \to 0$$

as $n \to \infty$. For example, for a continuous r.v. with mean zero and finite variance σ^2 this last convergence follows by noting[2]

$$\sigma^2 = \int_{(-a,a)} x^2 f(x)dx + \int_{(-a,a)^c} x^2 f(x)dx\ ,$$

where the first summand on the right converges to σ^2 as $a \to \infty$. Thus the second term must go to zero. But the second term is just

$$\int_{(-a,a)^c} x^2 f(x)dx = E\left[X^2 I_{[|X| \ge a]}\right] \quad \square$$

[1] Using $|E(X)| \le E(|X|)$ which follows from $-|x| \le x \le |x|\ \forall x$

[2] We express the following in terms of densities, but it holds generally, i.e., for discrete r.v.'s as well.

Liapunov CLT: Let $Y_1, \ldots, Y_n$ be independent r.v. with means $\mu_1, \ldots, \mu_n$, variances $\sigma_1^2, \ldots, \sigma_n^2$ and finite absolute central moments, i.e., $E(|Y_i - \mu_i|^3) < \infty$. If Liapunov's condition is satisfied, i.e.,

$$\ell_n = \frac{1}{\tau_n^3} \sum_{i=1}^{n} E\left(|Y_i - \mu_i|^3\right) \to 0 \text{ as } n \to \infty ,$$

then

$$P\left(\frac{\sum Y_i - \sum \mu_i}{\sqrt{\sum \sigma_i^2}} \le a\right) - \Phi(a) \to 0 \text{ as } n \to \infty$$

Proof: We will simply show that Liapunov's condition implies Lindeberg's condition. Note that $X_i = Y_i - \mu_i$ has mean zero and variance σ_i^2.

$$\ell_n = \frac{1}{\tau_n^3} \sum E\left(|X_i|^3\right) \ge \frac{1}{\tau_n^3} \sum E\left(|X_i|^3 I_{[|X_i| \ge \epsilon\tau_n]}\right) \ge \frac{\epsilon\tau_n}{\tau_n^3} \sum E\left(|X_i|^2 I_{[|X_i| \ge \epsilon\tau_n]}\right) = \epsilon L_n(\epsilon) \quad \square$$

Appendix 1

Demonstration of Step 1: A crucial part of the proof is the following form of Taylor's formula. You may want to skip the technical proof which is provided only for completeness.

Lemma 1 (Taylor's Formula): Let f be a bounded function defined on R with three bounded continuous derivatives $f^{(0)} = f$, $f^{(1)}$, $f^{(2)}$ and $f^{(3)}$, i.e., with

$$\sup_{x \in R} \left| f^{(i)}(x) \right| = M_f^{(i)} < \infty \text{ for } i = 0, 1, 2, 3.$$

Then for all $h \in R$

$$g(h) = \sup_{x \in R} \left| f(x+h) - f(x) - f^{(1)}(x)h - \frac{1}{2} f^{(2)}(x)h^2 \right| \le K_f \min\left(h^2, |h|^3\right)$$

Proof: From the usual form of Taylor's theorem we have for any h

$$f(x+h) - f(x) - f^{(1)}(x)h - \frac{1}{2} f^{(2)}(x)h^2 = \frac{h^3}{6} f^{(3)}(\tilde{x})$$

where $\tilde{x}$ lies between x and $x + h$. Hence for all $h \in R$

$$\left| f(x+h) - f(x) - f^{(1)}(x)h - \frac{1}{2} f^{(2)}(x)h^2 \right| \le \frac{M_f^{(3)}}{6} |h|^3 \tag{2}$$

On the other hand by simple application of the triangle inequality ($|a + b + \ldots| \le |a| + |b| + \ldots$) we also have

$$\begin{aligned} |f(x+h) - f(x) - f^{(1)}(x)h &- \frac{1}{2} f^{(2)}(x)h^2| \\ &\le M_f^{(0)} + M_f^{(0)} + M_f^{(1)}|h| + \frac{1}{2} M_f^{(2)} h^2 \\ &= h^2 \left(\frac{2M_f^{(0)}}{h^2} + \frac{M_f^{(1)}}{|h|} + \frac{1}{2} M_f^{(2)} \right) \end{aligned}$$

which for large h, say $|h| \geq b_f$, and large K', say $K' = \frac{1}{2}M_f^{(2)} + 1$, can be bounded by $K'h^2$. For $|h| < b_f$ we have from the inequality (2)

$$\left| f(x+h) - f(x) - f^{(1)}(x)h - \frac{1}{2}f^{(2)}(x)h^2 \right| \leq \frac{M_f^{(3)}}{6}|h|^3 \leq b_f \frac{M_f^{(3)}}{6} h^2$$

Taking

$$K_f = \max\left(b_f \frac{M_f^{(3)}}{6}, K', \frac{M_f^{(3)}}{6} \right)$$

we have for $h \in R$

$$\left| f(x+h) - f(x) - f^{(1)}(x)h - \frac{1}{2}f^{(2)}(x)h^2 \right| \leq K_f h^2 \text{ and } \leq K_f |h|^3 \, ,$$

hence $\leq \min(K_f h^2, K_f |h|^3)$ $\square$.

From Lemma 1 we obtain the following Corollary.

Corollary 1: Under the conditions of Lemma 1 we have

$$f(x+h_1) - f(x+h_2) = [f^{(1)}(x)(h_1 - h_2) + \frac{1}{2}f^{(2)}(x)(h_1^2 - h_2^2)] + r$$

with

$$|r| \leq g(h_1) + g(h_2) \leq K_f \left(\min(h_1^2, |h_1|^3) + \min(h_2^2, |h_2|^3) \right)$$

Proof:

$$\begin{aligned}
|r| &= \left| f(x+h_1) - f(x+h_2) - [f^{(1)}(x)(h_1 - h_2) + \frac{1}{2}f^{(2)}(x)(h_1^2 - h_2^2)] \right| \\
&= \Bigg| f(x+h_1) - f(x) - f^{(1)}(x)h_1 - \frac{1}{2}f^{(2)}(x)h_1^2 \\
&\qquad\qquad - \left[f(x+h_2) - f(x) - f^{(1)}(x)h_2 - \frac{1}{2}f^{(2)}(x)h_2^2 \right] \Bigg| \\
&\leq g(h_1) + g(h_2)
\end{aligned}$$

by the triangle inequality $\square$.

To demonstrate the convergence in (1) consider independent normal random variables $V_{n1}, \ldots, V_{nn}$ with mean zero and respective variances $\sigma_{n1}^2, \ldots, \sigma_{nn}^2$, so that[3]

$$Z = \frac{V_{n1} + \ldots + V_{nn}}{\tau_n}$$

is normal with mean zero and variance one, i.e., standard normal. These V_{ni}'s are also taken to be independent of the X_{nj}'s.

[3]The sum of independent normal r.v.'s is a normal r.v., see Fact 7.9 in Anderson et al.

First rewrite the approximation difference in the following telescoped fashion (to simplify notation we write X_i, V_i and σ_i for X_{ni}, V_{ni} and σ_{ni}, respectively)

$$\begin{aligned}
E\left[f\left(\frac{S_n}{\tau_n}\right)\right]-E\left[f\left(Z\right)\right] &= E\left[f\left(\frac{X_1+\ldots+X_n}{\tau_n}\right)\right]-E\left[f\left(\frac{X_1+\ldots+X_{n-1}+V_n}{\tau_n}\right)\right]\\
&+ E\left[f\left(\frac{X_1+\ldots+X_{n-1}+V_n}{\tau_n}\right)\right]-E\left[f\left(\frac{X_1+\ldots+X_{n-2}+V_{n-1}+V_n}{\tau_n}\right)\right]\\
&+ \qquad\qquad\qquad\qquad \ldots\ldots\ldots\\
&+ \qquad\qquad\qquad\qquad \ldots\ldots\ldots\\
&+ E\left[f\left(\frac{X_1+V_2+\ldots+V_n}{\tau_n}\right)\right]-E\left[f\left(\frac{V_1+\ldots+V_n}{\tau_n}\right)\right]
\end{aligned}$$

Now let $Y_i = X_1+\ldots+X_i+V_{i+2}+\ldots+V_n$ for $i=1,\ldots,n-2$ and $Y_0 = V_2+\ldots+V_n$ and $Y_{n-1} = X_1+\ldots+X_{n-1}$. With this notation and employing the two term Taylor expansion of Corollary 1 one can express a typical difference term in the above telescoped sum as

$$E\left[f\left(\frac{Y_i+X_{i+1}}{\tau_n}\right)\right]-E\left[f\left(\frac{Y_i+V_{i+1}}{\tau_n}\right)\right]=E\left[\left(\frac{X_{i+1}-V_{i+1}}{\tau_n}\right)f^{(1)}\left(\frac{Y_i}{\tau_n}\right)+\frac{X_{i+1}^2-V_{i+1}^2}{2\tau_n^2}f^{(2)}\left(\frac{Y_i}{\tau_n}\right)+R_{n,i+1}\right]$$

where

$$|R_{n,i+1}|\le g\left(\frac{X_{i+1}}{\tau_n}\right)+g\left(\frac{V_{i+1}}{\tau_n}\right)\le K_f\left[\min\left(\frac{X_{i+1}^2}{\tau_n^2},\frac{|X_{i+1}|^3}{\tau_n^3}\right)+\min\left(\frac{V_{i+1}^2}{\tau_n^2},\frac{|V_{i+1}|^3}{\tau_n^3}\right)\right] \qquad (3)$$

Noting the independence of Y_i and (X_{i+1},V_{i+1}) and $E(X_{i+1})=E(V_{i+1})=0$ we find[4]

$$E\left[\left(\frac{X_{i+1}-V_{i+1}}{\tau_n}\right)f^{(1)}\left(\frac{Y_i}{\tau_n}\right)\right]=E\left[\frac{X_{i+1}-V_{i+1}}{\tau_n}\right]E\left[f^{(1)}\left(\frac{Y_i}{\tau_n}\right)\right]=0$$

Similarly,

$$E\left[\frac{X_{i+1}^2-V_{i+1}^2}{2\tau_n^2}f^{(2)}\left(\frac{Y_i}{\tau_n}\right)\right]=E\left[\frac{X_{i+1}^2-V_{i+1}^2}{2\tau_n^2}\right]E\left[f^{(2)}\left(\frac{Y_i}{\tau_n}\right)\right]=\frac{\sigma_{i+1}^2-\sigma_{i+1}^2}{2\tau_n^2}E\left[f^{(2)}\left(\frac{Y_i}{\tau_n}\right)\right]=0$$

Thus

$$\left|E\left[f\left(\frac{Y_i+X_{i+1}}{\tau_n}\right)\right]-E\left[f\left(\frac{Y_i+V_{i+1}}{\tau_n}\right)\right]\right|\le E\left|R_{n,i+1}\right|$$

Using the triangle inequality on the above telescoped sum we have

$$\begin{aligned}
\left|E\left[f\left(\frac{S_n}{\tau_n}\right)\right]-E\left[f\left(Z\right)\right]\right| &\le \sum_{i=0}^{n-1}E\left|R_{n,i+1}\right|\le\sum_{i=0}^{n-1}E\left[g\left(\frac{X_{i+1}}{\tau_n}\right)\right]+\sum_{i=0}^{n-1}E\left[g\left(\frac{V_{i+1}}{\tau_n}\right)\right]\\
&= \sum_{i=1}^{n}E\left[g\left(\frac{X_i}{\tau_n}\right)\right]+\sum_{i=1}^{n}E\left[g\left(\frac{V_i}{\tau_n}\right)\right]\ ,
\end{aligned}$$

changing the running index from $i=0,\ldots,n-1$ to $i=1,\ldots,n$ in the last step. We will show that both sums on the right can be made arbitrarily small by letting $n\to\infty$. First consider

$$E\left[g\left(\frac{X_i}{\tau_n}\right)\right]=E\left[g\left(\frac{X_i}{\tau_n}\right)I_{\{|X_i|\le\epsilon\tau_n\}}\right]+E\left[g\left(\frac{X_i}{\tau_n}\right)I_{\{|X_i|>\epsilon\tau_n\}}\right]$$

[4]For independent X and Y we have $E(XY)=E(X)E(Y)$ for finite expectations, Anderson et al., Fact 8.10.

and invoke from (3) the bound $K_f|X_i/\tau_n|^3$ for the first term and the bound $K_f|X_i/\tau_n|^2$ for the second term to get

$$\begin{aligned}
E\left[g\left(\frac{X_i}{\tau_n}\right)\right] &\le K_f E\left[\frac{|X_i|^3}{\tau_n^3} I_{\{|X_i|\le\epsilon\tau_n\}}\right] + K_f E\left[\frac{|X_i|^2}{\tau_n^2} I_{\{|X_i|>\epsilon\tau_n\}}\right] \\
&\le \epsilon\, K_f E\left[\frac{|X_i|^2}{\tau_n^2} I_{\{|X_i|\le\epsilon\tau_n\}}\right] + K_f E\left[\frac{|X_i|^2}{\tau_n^2} I_{\{|X_i|>\epsilon\tau_n\}}\right] \\
&\le \epsilon\, K_f E\left[\frac{|X_i|^2}{\tau_n^2}\right] + K_f E\left[\frac{|X_i|^2}{\tau_n^2} I_{\{|X_i|>\epsilon\tau_n\}}\right] = \epsilon\, K_f \frac{\sigma_i^2}{\tau_n^2} + K_f \frac{1}{\tau_n^2} E\left[X_i^2 I_{\{|X_i|>\epsilon\tau_n\}}\right]
\end{aligned}$$

Summing these we get

$$\begin{aligned}
\sum_{i=1}^n E\left[g\left(\frac{X_i}{\tau_n}\right)\right] &\le K_f \sum_{i=1}^n \epsilon \frac{\sigma_i^2}{\tau_n^2} + K_f \frac{1}{\tau_n^2} \sum_{i=1}^n E\left[X_i^2 I_{\{|X_i|>\epsilon\tau_n\}}\right] \\
&= K_f \epsilon + K_f \frac{1}{\tau_n^2} \sum_{i=1}^n E\left[X_i^2 I_{\{|X_i|>\epsilon\tau_n\}}\right] \\
&= \epsilon\, K_f + K_f L_n(\epsilon) \to \epsilon\, K_f
\end{aligned}$$

as $n \to \infty$ by the assumption of the Lindeberg condition.
Similarly one gets

$$\sum_{i=1}^n E\left[g\left(\frac{V_i}{\tau_n}\right)\right] \le \epsilon\, K_f + K_f \frac{1}{\tau_n^2} \sum_{i=1}^n E\left[V_i^2 I_{\{|V_i|>\epsilon\tau_n\}}\right] \to \epsilon\, K_f \ ,$$

provided we can show that the Lindeberg condition also holds for the V_i's. These two convergences to $K_f\epsilon$ show that the approximation error can be made arbitrarily small as $n \to \infty$, since $\epsilon > 0$ can be any small number, as long as f and hence K_f is kept fixed.
To show the Lindeberg condition for the V_i's, we first note that the Lindeberg condition for the X_i's entails that

$$\max\{\sigma_1/\tau_n, \ldots, \sigma_n/\tau_n\} \to 0 \ \text{ as } \ n \to \infty \tag{4}$$

This follows from

$$\begin{aligned}
\frac{\sigma_i^2}{\tau_n^2} &= \frac{1}{\tau_n^2} E\left[X_i^2 I_{\{|X_i|\le\epsilon\tau_n\}}\right] + \frac{1}{\tau_n^2} E\left[X_i^2 I_{\{|X_i|>\epsilon\tau_n\}}\right] \le \frac{1}{\tau_n^2}\tau_n^2\epsilon^2 E\left[I_{\{|X_i|\le\epsilon\tau_n\}}\right] + \frac{1}{\tau_n^2} E\left[X_i^2 I_{\{|X_i|>\epsilon\tau_n\}}\right] \\
&\le \epsilon^2 + \frac{1}{\tau_n^2} E\left[X_i^2 I_{\{|X_i|>\epsilon\tau_n\}}\right] \le \epsilon^2 + L_n(\epsilon)
\end{aligned}$$

From the Lindeberg condition it follows that the second term vanishes as $n \to \infty$ and hence $\sigma_i^2/\tau_n^2 \le 2\epsilon^2$ as $n \to \infty$ for any $\epsilon > 0$. This shows (4).
The Lindeberg condition for the V_i's is seen as follows, using $|V_i/(\tau_n\epsilon)| > 1$ in the first $\le$:

$$\begin{aligned}
\frac{1}{\tau_n^2} \sum_{i=1}^n E\left[V_i^2 I_{\{|V_i|>\epsilon\tau_n\}}\right] &\le \frac{1}{\epsilon\tau_n^3} \sum_{i=1}^n E\left[|V_i|^3\right] = \frac{1}{\epsilon\tau_n^3} \sum_{i=1}^n \sigma_i^3 E\left[|Z|^3\right] \\
&\le \frac{1}{\epsilon} \max\{\sigma_1/\tau_n, \ldots, \sigma_n/\tau_n\} \frac{\sum_{i=1}^n \sigma_i^2}{\tau_n^2} E\left[|Z|^3\right] \\
&= \frac{1}{\epsilon} \max\{\sigma_1/\tau_n, \ldots, \sigma_n/\tau_n\} E\left[|Z|^3\right] \to 0
\end{aligned}$$

as $n \to \infty$ $\square$.

Appendix 2

Here we will exhibit an explicit function $f(x)$ which has three continuous and bounded derivatives and which is 1 for $x \leq 0$, decreases from 1 to 0 on the interval $[0,1]$ and is 0 for $x \geq 1$. The form of this function taken from Thomasian (1969). By taking

$$h(x) = f\left(\frac{x-a}{\delta}\right)$$

we immediately get a corresponding smooth function which descends from 1 to 0 over the interval $[a, a+\delta]$ instead.
The function $f(x)$ is given as follows for $x \in [0,1]$:

$$f(x) = 1 - 140\left(\frac{1}{4}x^4 - \frac{3}{5}x^5 + \frac{1}{2}x^6 - \frac{1}{7}x^7\right)$$

Then

$$f'(x) = -140x^3(1-x)^3 \;\text{ for }\; x \in [0,1]\ ,$$

hence f is monotone decreasing on $[0,1]$. Further, $f(1) = 0$ and $f(0) = 1$. To check smoothness we have to verify that the first three derivatives join continuously at 0 and 1, i.e., are 0. Clearly, $f'(0) = f'(1) = 0$. Next,

$$f''(x) = -420x^2(1-x)^2(1-2x) \;\text{ with }\; f''(0) = f''(1) = 0$$

and

$$f'''(x) = -420\left(2x(1-x)^2(1-2x) - 2(1-x)x^2(1-2x) - 2x^2(1-x)^2\right)$$

with $f'''(0) = f'''(1) = 0$.

References

Anderson, D.F. Seppäläinen, T. and Valkó, B. (2018), *Introduction to Probability*, Cambridge University Press.

Lindeberg, J.W. (1922). "Eine neue Herleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung." *Mathemat. Z.*, **15**, 211-225.

Thomasian, A. (1969). *The Structure of Probability Theory with Applications.* Mc Graw Hill, New York. (pp. 483-493)