



An isometric study of the Lindeberg–Feller central limit theorem via Stein’s method



B. Berckmoes, R. Lowen*, J. Van Casteren

Department of Mathematics and Computer Science, University of Antwerp, Belgium

ARTICLE INFO

Article history:

Received 17 January 2012

Available online 12 April 2013

Submitted by U. Stadtmueller

Keywords:

Approach structure

Distance

Limit

Weak topology

Probability measure

Law

Random variable

Central limit theorem

Lindeberg condition

Kolmogorov metric

Triangular array

ABSTRACT

We use Stein’s method to prove a generalization of the Lindeberg–Feller CLT providing an upper and a lower bound for the superior limit of the Kolmogorov distance between a normally distributed random variable and the rowwise sums of a rowwise independent triangular array of random variables which is asymptotically negligible in the sense of Feller. A natural example shows that the upper bound is of optimal order. The lower bound improves a result by Andrew Barbour and Peter Hall.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction and motivation

One of the most important questions in probability theory reads as follows:

Question 1.1. *Under which conditions is a large sum of independent random variables approximately normally distributed?*

The most common way to address this question formally is within the framework of so-called triangular arrays of random variables, a concept which we now introduce.

By a *standard triangular array* (STA) we mean a triangular array of real square integrable random variables

$$\begin{array}{ccc} \xi_{1,1} & & \\ \xi_{2,1} & \xi_{2,2} & \\ \xi_{3,1} & \xi_{3,2} & \xi_{3,3} \\ \vdots & & \end{array}$$

satisfying the following properties.

(a) $\forall n : \xi_{n,1}, \dots, \xi_{n,n}$ are independent.

* Corresponding author.

E-mail addresses: ben.berckmoes@ua.ac.be (B. Berckmoes), bob.lowen@ua.ac.be (R. Lowen), jan.vancasteren@ua.ac.be (J. Van Casteren).

$$(b) \forall n, k : \mathbb{E} [\xi_{n,k}] = 0.$$

$$(c) \forall n : \sum_{k=1}^n \sigma_{n,k}^2 = 1, \quad \text{where } \sigma_{n,k}^2 = \mathbb{E} [\xi_{n,k}^2].$$

Let $\{\xi_{n,k}\}$ be an STA and ξ a standard normally distributed random variable. [Question 1.1](#) can now be put as follows:

Question 1.2. Under which conditions is the weak limit relation $\sum_{k=1}^n \xi_{n,k} \xrightarrow{w} \xi$ valid?

It turns out that when $\{\xi_{n,k}\}$ satisfies *Feller's negligibility condition* in the sense that

$$\max_{k=1}^n \sigma_{n,k}^2 \rightarrow 0, \quad (1)$$

the Lindeberg–Feller CLT [11] provides us with a completely satisfactory answer to [Question 1.2](#):

Theorem 1.3 (Lindeberg–Feller CLT). Suppose that $\{\xi_{n,k}\}$ satisfies (1). Then the following are equivalent:

$$(a) \sum_{k=1}^n \xi_{n,k} \xrightarrow{w} \xi.$$

$$(b) \forall \epsilon > 0 : \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| \geq \epsilon] \rightarrow 0. \quad (2)$$

(2) is often referred to as Lindeberg's condition.

Throughout the remainder of the paper, ξ and $\{\xi_{n,k}\}$ will be as in [Theorem 1.3](#).

The Kolmogorov distance between random variables η and η' is defined by

$$K(\eta, \eta') = \sup_{x \in \mathbb{R}} |\mathbb{P}[\eta \leq x] - \mathbb{P}[\eta' \leq x]|.$$

In general, K is too strong to metrize weak convergence, but it is well known that if η is continuously distributed, the following are equivalent for any sequence $(\eta_n)_n$:

$$(a) \eta_n \xrightarrow{w} \eta.$$

$$(b) \limsup_{n \rightarrow \infty} K(\eta, \eta_n) = 0.$$

The previous observation reveals that the Lindeberg–Feller CLT in fact gives a necessary and sufficient condition for the number $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ to equal 0. The theorem however does not answer the following question, which is slightly more general than [Question 1.2](#), but nevertheless important from both a theoretical and an applied point of view:

Question 1.4. Under which conditions is the number $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ small?

In this paper we will perform what we call an ‘isometric study’ in which we answer [Question 1.4](#), by providing an upper and a lower bound for the number $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$. These bounds will generalize the Lindeberg–Feller CLT. Stein's method [1,17,18] will turn out to be a powerful and indispensable tool for the elaboration of this program. We would like to point out that related work has been done in [9].

2. An upper bound for $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$

We start with [Lemma 2.1](#), which makes the task of finding an upper bound for $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ considerably more feasible. We let $\mathcal{H}$ stand for the collection of all strictly decreasing functions $h : \mathbb{R} \rightarrow \mathbb{R}$, with a bounded first and second derivative and a bounded and piecewisely continuous third derivative, and for which $\lim_{x \rightarrow -\infty} h(x) = 1$ and $\lim_{x \rightarrow \infty} h(x) = 0$.

Lemma 2.1. If η is continuously distributed, then the formula

$$\limsup_{n \rightarrow \infty} K(\eta, \eta_n) = \sup_{h \in \mathcal{H}} \limsup_{n \rightarrow \infty} |\mathbb{E}[h(\eta) - h(\eta_n)]| \quad (3)$$

is valid for any sequence $(\eta_n)_n$.

Proof. Let $\epsilon > 0$ be arbitrary. The continuity of η allows us to construct points $x_1 < \dots < x_N$ such that for each n

$$K(\eta, \eta_n) \leq \max_{k=1}^N |\mathbb{P}[\eta \leq x_k] - \mathbb{P}[\eta_n \leq x_k]| + \epsilon. \quad (4)$$

But, again invoking the continuity of η , it is also easily seen that for each $x \in \mathbb{R}$ there exist $\delta > 0$ and functions $h, \tilde{h} \in \mathcal{H}$ such that for each n

$$\begin{aligned} |\mathbb{P}[\eta \leq x] - \mathbb{P}[\eta_n \leq x]| &\leq \max\{\mathbb{P}[\eta \leq x - \delta] - \mathbb{P}[\eta_n \leq x], \mathbb{P}[\eta_n \leq x] - \mathbb{P}[\eta \leq x + \delta]\} + \epsilon/2 \\ &\leq \max\{\mathbb{E}[h(\eta) - h(\eta_n)], \mathbb{E}[\tilde{h}(\eta_n) - \tilde{h}(\eta)]\} + \epsilon. \end{aligned} \quad (5)$$

Combining (4) and (5) reveals that there exist functions $h_1, \dots, h_{2N} \in \mathcal{H}$ such that

$$\begin{aligned} \limsup_{n \rightarrow \infty} K(\eta, \eta_n) &\leq \limsup_{n \rightarrow \infty} \max_{k=1}^{2N} |\mathbb{E}[h_k(\eta) - h_k(\eta_n)]| + 2\epsilon \\ &= \max_{k=1}^{2N} \limsup_{n \rightarrow \infty} |\mathbb{E}[h_k(\eta) - h_k(\eta_n)]| + 2\epsilon \\ &\leq \sup_{h \in \mathcal{H}} \limsup_{n \rightarrow \infty} |\mathbb{E}[h(\eta) - h(\eta_n)]| + 2\epsilon. \end{aligned}$$

Hence the left-hand side of (3) is dominated by the right-hand side.

For the converse inequality, it suffices to remark that for any $h \in \mathcal{H}$

$$|\mathbb{E}[h(\eta) - h(\eta_n)]| \leq \int_0^1 |\mathbb{P}[\eta \leq h^{-1}t] - \mathbb{P}[\eta_n \leq h^{-1}t]| dt \leq K(\eta, \eta_n). \quad \square$$

With the Lindeberg–Feller CLT in mind, it seems plausible that bounds for $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ should be based on Lindeberg’s condition (2). Hence this leads us to the definition of the number

$$\text{Lin}(\{\xi_{n,k}\}) = \sup_{\epsilon > 0} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2; |\xi_{n,k}| \geq \epsilon] \quad (6)$$

which we call the *Lindeberg index*. It is clear that $\{\xi_{n,k}\}$ satisfies Lindeberg’s condition if and only if $\text{Lin}(\{\xi_{n,k}\}) = 0$. We will provide an example, inspired by one of the problems posed in [11, Chapter XV], which illustrates that the Lindeberg index is not trivial in our context.

Fix $0 < \alpha < 1$, let $\beta = \frac{\alpha}{1-\alpha}$ and put

$$s_n^2 = (1 + \beta)n - \beta \sum_{k=1}^n k^{-1} = n + \beta \sum_{k=1}^n (1 - k^{-1}). \quad (7)$$

Notice that $s_n^2 \rightarrow \infty$. Now consider the STA $\{\eta_{\alpha,n,k}\}$ such that

$$\mathbb{P}[\eta_{\alpha,n,k} = -1/s_n] = \mathbb{P}[\eta_{\alpha,n,k} = 1/s_n] = \frac{1}{2} (1 - \beta k^{-1}) \quad (8)$$

and

$$\mathbb{P}[\eta_{\alpha,n,k} = -\sqrt{k}/s_n] = \mathbb{P}[\eta_{\alpha,n,k} = \sqrt{k}/s_n] = \frac{1}{2} \beta k^{-1}. \quad (9)$$

Proposition 2.2. The STA $\{\eta_{\alpha,n,k}\}$ satisfies Feller’s condition (1) and

$$\text{Lin}(\{\eta_{\alpha,n,k}\}) = \alpha. \quad (10)$$

Proof. For Feller’s negligibility condition we observe

$$\max_{k=1}^n \mathbb{E}[\eta_{\alpha,n,k}^2] = \frac{(1 + \beta) - \beta n^{-1}}{s_n^2} \rightarrow 0.$$

In order to calculate $\text{Lin}(\{\eta_{\alpha,n,k}\})$, we fix $\epsilon > 0$ so that $\epsilon^2(1 + \beta) \leq 1$ and hence $\epsilon^2 s_n^2 \leq n$. Then for n so that $\epsilon s_n > 1$ and $k \leq n$ we have

$$\mathbb{E}[\eta_{\alpha,n,k}^2; |\eta_{\alpha,n,k}| \geq \epsilon] = \begin{cases} \beta/s_n^2 & \text{if } k \geq \epsilon^2 s_n^2 \\ 0 & \text{if } k < \epsilon^2 s_n^2. \end{cases}$$

It follows that

$$\begin{aligned} \sum_{k=1}^n \mathbb{E}[\eta_{\alpha,n,k}^2; |\eta_{\alpha,n,k}| \geq \epsilon] &= \frac{\beta(n - \lceil \epsilon^2 s_n^2 \rceil + 1)}{s_n^2} \\ &= \frac{\beta n}{(1 + \beta)n - \beta \sum_{k=1}^n k^{-1}} - \beta \frac{\lceil \epsilon^2 s_n^2 \rceil - 1}{s_n^2} \\ &= \frac{\beta}{(1 + \beta) - \beta \frac{1}{n} \sum_{k=1}^n k^{-1}} - \beta \frac{\lceil \epsilon^2 s_n^2 \rceil - 1}{s_n^2}, \end{aligned}$$

whence

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E}[\eta_{\alpha,n,k}^2; |\eta_{\alpha,n,k}| \geq \epsilon] &= \lim_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E}[\eta_{\alpha,n,k}^2; |\eta_{\alpha,n,k}| \geq \epsilon] \\ &= \frac{\beta}{1 + \beta} - \beta \epsilon^2 \\ &= \alpha - \beta \epsilon^2, \end{aligned}$$

which proves the proposition. $\square$

We now embark on the search for an upper bound for $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ in terms of $\text{Lin}(\{\xi_{n,k}\})$. We will discuss and compare two different methods, which are both known to lead to a proof of the sufficiency of Lindeberg's condition for normal convergence in the Lindeberg–Feller CLT.

2.1. The classical method

The ‘classical method’ to prove normal convergence appears in many different forms such as e.g. characteristic functions [13] or Gaussian transforms [7], but it always involves estimating an expression of the type

$$\left| \mathbb{E} \left[h(\xi) - h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| \quad (11)$$

based on the following three key observations. Their proofs are elementary and can be found in the literature.

(I) The random variable ξ is infinitely divisible in the sense that for each n

$$\xi = \sum_{k=1}^n \eta_{n,k} \quad (12)$$

where $\{\eta_{n,k}\}$ is the STA consisting of normally distributed random variables with $\mathbb{E}[\eta_{n,k}^2] = \sigma_{n,k}^2$.

(II) Let $\{\eta_{n,k}\}$ be as in (I). Then, for any bounded and continuous function $h : \mathbb{R} \rightarrow \mathbb{R}$,

$$\left| \mathbb{E} \left[h \left(\sum_{k=1}^n \eta_{n,k} \right) - h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| \leq \sum_{k=1}^n \sup_{a \in \mathbb{R}} |\mathbb{E}[h(a + \eta_{n,k}) - h(a + \xi_{n,k})]|. \quad (13)$$

(III) Let $h : \mathbb{R} \rightarrow \mathbb{R}$ have a bounded second derivative and a bounded and piecewise continuous third derivative. Then for any $a, x \in \mathbb{R}$

$$\left| h(a + x) - h(a) - h'(a)x - \frac{1}{2}h''(a)x^2 \right| \leq \min \left\{ \|h''\|_{\infty} x^2, \frac{1}{6} \|h'''\|_{\infty} |x|^3 \right\}. \quad (14)$$

Combining (12)–(14) for our purpose, yields for any $h \in \mathcal{H}$ and $\epsilon > 0$

$$\begin{aligned} \left| \mathbb{E} \left[h(\xi) - h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| &= \left| \mathbb{E} \left[h \left(\sum_{k=1}^n \eta_{n,k} \right) - h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| \\ &\leq \sum_{k=1}^n \sup_{a \in \mathbb{R}} |\mathbb{E}[h(a + \eta_{n,k}) - h(a + \xi_{n,k})]| \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^n \sup_{a \in \mathbb{R}} \left| \mathbb{E} \left[h(a + \eta_{n,k}) - h(a) - h'(a)0 - \frac{1}{2} h''(a) \sigma_{n,k}^2 \right] \right. \\
&\quad \left. - \mathbb{E} \left[h(a + \xi_{n,k}) - h(a) - h'(a)0 - \frac{1}{2} h''(a) \sigma_{n,k}^2 \right] \right| \\
&= \sum_{k=1}^n \sup_{a \in \mathbb{R}} \left| \mathbb{E} \left[h(a + \eta_{n,k}) - h(a) - h'(a) \eta_{n,k} - \frac{1}{2} h''(a) \eta_{n,k}^2 \right] \right. \\
&\quad \left. - \mathbb{E} \left[h(a + \xi_{n,k}) - h(a) - h'(a) \xi_{n,k} - \frac{1}{2} h''(a) \xi_{n,k}^2 \right] \right| \\
&\leq \sum_{k=1}^n \sup_{a \in \mathbb{R}} \mathbb{E} \left[\left| h(a + \eta_{n,k}) - h(a) - h'(a) \eta_{n,k} - \frac{1}{2} h''(a) \eta_{n,k}^2 \right| \right] \\
&\quad + \sum_{k=1}^n \sup_{a \in \mathbb{R}} \mathbb{E} \left[\left| h(a + \xi_{n,k}) - h(a) - h'(a) \xi_{n,k} - \frac{1}{2} h''(a) \xi_{n,k}^2 \right| \right] \\
&\leq \frac{1}{6} \|h'''\|_{\infty} \sum_{k=1}^n \mathbb{E} [|\eta_{n,k}|^3] + \frac{1}{6} \|h'''\|_{\infty} \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3; |\xi_{n,k}| < \epsilon] \\
&\quad + \|h''\|_{\infty} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| \geq \epsilon] \\
&\leq \frac{1}{6} \|h'''\|_{\infty} \mathbb{E} [|\xi|^3] \left(\sum_{k=1}^n \sigma_{n,k}^3 \right) + \frac{1}{6} \|h'''\|_{\infty} \epsilon \left(\sum_{k=1}^n \sigma_{n,k}^2 \right) \\
&\quad + \|h''\|_{\infty} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| \geq \epsilon] \\
&\leq \frac{1}{6} \|h'''\|_{\infty} \left(\mathbb{E} [|\xi|^3] \max_{k=1}^n \sigma_{n,k} + \epsilon \right) \left(\sum_{k=1}^n \sigma_{n,k}^2 \right) \\
&\quad + \|h''\|_{\infty} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| \geq \epsilon] \\
&= \frac{1}{6} \|h'''\|_{\infty} \left(\mathbb{E} [|\xi|^3] \max_{k=1}^n \sigma_{n,k} + \epsilon \right) + \|h''\|_{\infty} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| \geq \epsilon].
\end{aligned}$$

Because of (1), the previous calculation shows that for any $h \in \mathcal{H}$

$$\limsup_{n \rightarrow \infty} \left| \mathbb{E} \left[h(\xi) - h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| \leq \|h''\|_{\infty} \text{Lin}(\{\xi_{n,k}\}). \quad (15)$$

Recalling (3), we see that (15) proves that Lindeberg's condition is sufficient for the weak limit relation $\sum_{k=1}^n \xi_{n,k} \xrightarrow{w} \xi$ to hold. However, since $\|h''\|_{\infty}$ blows up if we let h run through the collection $\mathcal{H}$, (15) is useless for the derivation of an upper bound for $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$. We conclude that although the classical method suffices to decide when the number $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ is 0, it is not subtle enough to decide when it is small.

2.2. Stein's method

Whereas the classical method provides an upper bound for (11) based on a direct analysis of the function h , Stein's method first transforms (11) to an expression which is easier to analyze. The basics of the method are contained in the following lemma. The proofs can be found in e.g. [1].

Lemma 2.3. Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be measurable and bounded. Put

$$f_h(x) = e^{x^2/2} \int_{-\infty}^x (h(t) - \mathbb{E}[h(\xi)]) e^{-t^2/2} dt. \quad (16)$$

Then for any $x \in \mathbb{R}$

$$\mathbb{E}[h(\xi)] - h(x) = xf_h(x) - f'_h(x). \quad (17)$$

Moreover, if h is absolutely continuous, then

$$\|f''_h\|_\infty \leq 2 \|h'\|_\infty, \quad (18)$$

and if $h_z = 1_{]-\infty, z]}$ for $z \in \mathbb{R}$, then for all $x, y \in \mathbb{R}$

$$|f'_{h_z}(x) - f'_{h_z}(y)| \leq 1. \quad (19)$$

We will apply Lemma 2.3 in order to derive an upper bound for

$$\limsup_{n \rightarrow \infty} K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right).$$

We first need three additional lemmata.

Stein's method was used by Barbour and Hall to derive Berry–Esseen type bounds in [2]. The following lemma is inspired by their paper.

Lemma 2.4. Let $h \in \mathcal{H}$ and put

$$\delta_{n,k} = f_h \left(\sum_{i \neq k} \xi_{n,i} + \xi_{n,k} \right) - f_h \left(\sum_{i \neq k} \xi_{n,i} \right) - \xi_{n,k} f'_h \left(\sum_{i \neq k} \xi_{n,i} \right) \quad (20)$$

and

$$\epsilon_{n,k} = f'_h \left(\sum_{i \neq k} \xi_{n,i} + \xi_{n,k} \right) - f'_h \left(\sum_{i \neq k} \xi_{n,i} \right) - \xi_{n,k} f''_h \left(\sum_{i \neq k} \xi_{n,i} \right). \quad (21)$$

Then

$$\mathbb{E} \left[\left(\sum_{k=1}^n \xi_{n,k} \right) f_h \left(\sum_{k=1}^n \xi_{n,k} \right) - f_h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] = \sum_{k=1}^n \mathbb{E}[\xi_{n,k} \delta_{n,k}] - \sum_{k=1}^n \sigma_{n,k}^2 \mathbb{E}[\epsilon_{n,k}]. \quad (22)$$

Proof. Calculate the right-hand side of (22) and recall that $\xi_{n,k}$ and $\sum_{i \neq k} \xi_{n,i}$ are independent, $\mathbb{E}[\xi_{n,k}] = 0$ and $\sum_{k=1}^n \sigma_{n,k}^2 = 1$. $\square$

The following lemma is a straightforward application of Taylor's theorem.

Lemma 2.5. Let $f : \mathbb{R} \rightarrow \mathbb{R}$ have a bounded derivative and a bounded and piecewise continuous second derivative. Then for any $a, x \in \mathbb{R}$

$$|f(a+x) - f(a) - f'(a)x| \leq \min \left\{ \left(\sup_{x_1, x_2 \in \mathbb{R}} |f'(x_1) - f'(x_2)| \right) |x|, \frac{1}{2} \|f''\|_\infty x^2 \right\}. \quad (23)$$

We finally observe that the favorable inequality (19) is extendable to all functions in the collection $\mathcal{H}$. In the lemma below we use the so-called layer cake representation which says that if ζ is a $[0, 1]$ -valued random variable then

$$\mathbb{E}[\zeta] = \int_0^1 \mathbb{P}[\zeta \geq s] ds.$$

Lemma 2.6. Let $h \in \mathcal{H}$. Then for all $x, y \in \mathbb{R}$

$$|f'_h(x) - f'_h(y)| \leq 1. \quad (24)$$

Proof. From (16) we derive that

$$f'_h(x) = xe^{x^2/2} \int_{-\infty}^x (h(t) - \mathbb{E}[h(\xi)]) e^{-t^2/2} dt + h(x) - \mathbb{E}[h(\xi)]. \quad (25)$$

Furthermore, for all $h \in \mathcal{H}$, applying the layer cake representation to the constant random variable ζ with value $h(x)$ on the probability space $([0, 1], \mathcal{B}_{[0,1]}, \lambda)$, with $\mathcal{B}_{[0,1]}$ the Borel σ -field and λ the Lebesgue measure, we get

$$h(x) = \int_0^1 \lambda(\{t \in [0, 1] : \zeta(t) \geq s\}) ds = \int_0^1 \Phi(x, s) ds$$

with

$$\Phi(x, s) = \begin{cases} 0 & \text{if } h(x) \langle s \Leftrightarrow x \rangle h^{-1}(s) \\ 1 & \text{if } h(x) \geq s \Leftrightarrow x \leq h^{-1}(s). \end{cases}$$

Put differently,

$$\Phi(x, s) = 1_{]-\infty, h^{-1}(s)]}(x) = h_{h^{-1}(s)}(x)$$

and hence

$$h(x) = \int_0^1 h_{h^{-1}(s)}(x) ds. \quad (26)$$

Combining (25) and (26) and applying Fubini yields

$$\begin{aligned} f'_h(x) - f'_h(y) &= xe^{x^2/2} \int_{-\infty}^x (h(t) - \mathbb{E}[h(\xi)]) e^{-t^2/2} dt + h(x) - ye^{y^2/2} \int_{-\infty}^y (h(t) - \mathbb{E}[h(\xi)]) e^{-t^2/2} dt - h(y) \\ &= xe^{x^2/2} \int_{-\infty}^x \int_0^1 (h_{h^{-1}(s)}(t) - \mathbb{E}[h_{h^{-1}(s)}(\xi)]) e^{-t^2/2} ds dt + \int_0^1 h_{h^{-1}(s)}(x) ds \\ &\quad - ye^{y^2/2} \int_{-\infty}^y \int_0^1 (h_{h^{-1}(s)}(t) - \mathbb{E}[h_{h^{-1}(s)}(\xi)]) e^{-t^2/2} ds dt - \int_0^1 h_{h^{-1}(s)}(y) ds \\ &= \int_0^1 \left[xe^{x^2/2} \int_{-\infty}^x (h_{h^{-1}(s)}(t) - \mathbb{E}[h_{h^{-1}(s)}(\xi)]) e^{-t^2/2} dt + h_{h^{-1}(s)}(x) \right. \\ &\quad \left. - ye^{y^2/2} \int_{-\infty}^y (h_{h^{-1}(s)}(t) - \mathbb{E}[h_{h^{-1}(s)}(\xi)]) e^{-t^2/2} dt + h_{h^{-1}(s)}(y) \right] ds \\ &= \int_0^1 [f'_{h_{h^{-1}(s)}}(x) - f'_{h_{h^{-1}(s)}}(y)] ds \end{aligned}$$

and the lemma follows from (19). $\square$

Combining (17), (18), (22), (23) and (24) yields for $h \in \mathcal{H}$ and $\epsilon > 0$

$$\begin{aligned} \left| \mathbb{E} \left[h(\xi) - h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| &= \left| \mathbb{E} \left[\left(\sum_{k=1}^n \xi_{n,k} \right) f_h \left(\sum_{k=1}^n \xi_{n,k} \right) - f'_h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| \\ &\leq \sum_{k=1}^n \mathbb{E} [|\xi_{n,k} \delta_{n,k}|] + \sum_{k=1}^n \sigma_{n,k}^2 \mathbb{E} [|\epsilon_{n,k}|] \\ &\leq \frac{1}{2} \|f''_h\|_{\infty} \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^3; |\xi_{n,k}| < \epsilon] \\ &\quad + \left(\sup_{x_1, x_2 \in \mathbb{R}} |f'_h(x_1) - f'_h(x_2)| \right) \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^2; |\xi_{n,k}| \geq \epsilon] \\ &\quad + \left(\sup_{x_1, x_2 \in \mathbb{R}} |f''_h(x_1) - f''_h(x_2)| \right) \sum_{k=1}^n \sigma_{n,k}^2 \mathbb{E} [|\xi_{n,k}|] \\ &\leq \frac{1}{2} \|f''_h\|_{\infty} \epsilon + \sum_{k=1}^n \mathbb{E} [|\xi_{n,k}|^2; |\xi_{n,k}| \geq \epsilon] + \left(\sup_{x_1, x_2 \in \mathbb{R}} |f''_h(x_1) - f''_h(x_2)| \right) \max_{k=1}^n \sigma_{n,k}. \end{aligned}$$

Because of (1), the previous calculation shows that for any $h \in \mathcal{H}$

$$\limsup_{n \rightarrow \infty} \left| \mathbb{E} \left[h(\xi) - h \left(\sum_{k=1}^n \xi_{n,k} \right) \right] \right| \leq \text{Lin}(\{\xi_{n,k}\}). \quad (27)$$

Recalling (3), we see that (27) entails the following beautiful generalization of the sufficiency of Lindeberg's condition in the Lindeberg–Feller CLT.

Theorem 2.7. *The inequality*

$$\limsup_{n \rightarrow \infty} K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right) \leq \text{Lin} \left(\{\xi_{n,k}\} \right) \quad (28)$$

is valid.

3. A lower bound for $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$

In [2] the following theorem is proved.

Theorem 3.1. *There exists a constant $C > 0$, not depending on $\{\xi_{n,k}\}$, such that*

$$\limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} \left[\xi_{n,k}^2 \left(1 - e^{-\frac{1}{4} \xi_{n,k}^2} \right) \right] \leq C \limsup_{n \rightarrow \infty} K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right). \quad (29)$$

Moreover, (29) can be shown to hold with $C \leq 41$.

Let Φ_L be the collection of all non-decreasing functions $\phi : \mathbb{R}^+ \rightarrow [0, 1]$ which are strictly positive on $\mathbb{R}_0^+$ and for which $\lim_{\epsilon \downarrow 0} \phi(\epsilon) = 0$.

Inspired by Theorem 3.1, we define for $\phi \in \Phi_L$ the number

$$\widetilde{\text{Lin}}_{\phi}(\{\xi_{n,k}\}) = \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \phi(|\xi_{n,k}|)] \quad (30)$$

which we call the *relaxed Lindeberg index (with respect to ϕ)*. Furthermore, we put for $\gamma > 0$

$$\phi_{\gamma}(x) = 1 - e^{-\gamma x^2} \quad (31)$$

and

$$\widetilde{\text{Lin}}_{\gamma}(\{\xi_{n,k}\}) = \widetilde{\text{Lin}}_{\phi_{\gamma}}(\{\xi_{n,k}\}). \quad (32)$$

The following proposition collects some basic properties of the relaxed Lindeberg index.

Proposition 3.2. *For any $\phi \in \Phi_L$ we have*

$$\widetilde{\text{Lin}}_{\phi}(\{\xi_{n,k}\}) \leq \text{Lin}(\{\xi_{n,k}\}) \quad (33)$$

and, in addition,

$$\{\xi_{n,k}\} \text{ satisfies Lindeberg's condition } \Leftrightarrow \widetilde{\text{Lin}}_{\phi}(\{\xi_{n,k}\}) = 0. \quad (34)$$

Furthermore, for any sequence $(\phi_n)_n$ in Φ_L we have

$$\phi_n \uparrow 1_{]0, \infty[} \Rightarrow \widetilde{\text{Lin}}_{\phi_n}(\{\xi_{n,k}\}) \uparrow \text{Lin}(\{\xi_{n,k}\}). \quad (35)$$

In particular,

$$\gamma \uparrow \infty \Rightarrow \widetilde{\text{Lin}}_{\gamma}(\{\xi_{n,k}\}) \uparrow \text{Lin}(\{\xi_{n,k}\}). \quad (36)$$

Finally, for any $\gamma > 0$,

$$\widetilde{\text{Lin}}_{2\gamma}(\{\xi_{n,k}\}) \leq 2\widetilde{\text{Lin}}_{\gamma}(\{\xi_{n,k}\}). \quad (37)$$

Proof. (33) follows from

$$\begin{aligned} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \phi(|\xi_{n,k}|)] &= \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \phi(|\xi_{n,k}|); |\xi_{n,k}| \geq \epsilon] + \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \phi(|\xi_{n,k}|); |\xi_{n,k}| < \epsilon] \\ &\leq \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2; |\xi_{n,k}| \geq \epsilon] + \phi(\epsilon) \end{aligned}$$

and the fact that $\lim_{\epsilon \downarrow 0} \phi(\epsilon) = 0$.

Notice that (33) entails that $\widetilde{\text{Lin}}_\phi(\{\xi_{n,k}\}) = 0$ if $\{\xi_{n,k}\}$ satisfies Lindeberg's condition. For the converse implication, observe that, for any $\epsilon > 0$,

$$\phi(\epsilon) \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2; |\xi_{n,k}| > \epsilon] \leq \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 \phi(|\xi_{n,k}|)]$$

and $\phi(\epsilon) > 0$. This proves (34).

In order to prove (35), we choose for $\epsilon > 0$ a natural number n_0 such that

$$\phi_{n_0}(\epsilon) \geq 1 - \epsilon.$$

Then

$$\limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2; |\xi_{n,k}| > \epsilon] \leq \limsup_{n \rightarrow \infty} \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 \phi_{n_0}(|\xi_{n,k}|)] + \epsilon$$

and (35) follows.

Finally, (37) follows from the observation that

$$\sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 \phi_{2\gamma}(|\xi_{n,k}|)] = \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 \phi_\gamma(|\xi_{n,k}|)] + \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 \phi_\gamma(|\xi_{n,k}|) e^{-\gamma \xi_{n,k}^2}]. \quad \square$$

Proposition 3.3 provides an explicit formula for $\widetilde{\text{Lin}}_\gamma(\{\eta_{\alpha,n,k}\})$, where $\{\eta_{\alpha,n,k}\}$ is the STA determined by (2), (7), (8) and (9). Recall that $\{\eta_{\alpha,n,k}\}$ satisfies Feller's condition (1) and that $\text{Lin}(\{\eta_{\alpha,n,k}\}) = \alpha$ (Proposition 2.2).

Proposition 3.3. *The formula*

$$\widetilde{\text{Lin}}_\gamma(\{\eta_{\alpha,n,k}\}) = \alpha \left(1 - \frac{1 - e^{-\gamma(1-\alpha)}}{\gamma(1-\alpha)} \right) \quad (38)$$

is valid.

Proof. An easy calculation leads to

$$\sum_{k=1}^n \mathbb{E}[\eta_{\alpha,n,k}^2 \phi_\gamma(|\eta_{\alpha,n,k}|)] = \frac{n}{s_n^2} \left(1 - e^{-\frac{\gamma}{s_n^2}} \right) - \beta \left(1 - e^{-\frac{\gamma}{s_n^2}} \right) \left(\frac{1}{s_n^2} \sum_{k=1}^n k^{-1} \right) + \beta \frac{n}{s_n^2} - \beta \left(\frac{1}{s_n^2} \sum_{k=1}^n e^{-\frac{\gamma k}{s_n^2}} \right).$$

Arguing analogously as in the proof of Proposition 2.2, we see that

$$\frac{n}{s_n^2} \rightarrow \frac{1}{1+\beta} \quad \text{and} \quad \frac{1}{s_n^2} \sum_{k=1}^n k^{-1} \rightarrow 0.$$

Also,

$$\frac{1}{s_n^2} \sum_{k=1}^n e^{-\frac{\gamma k}{s_n^2}} = e^{-\frac{\gamma}{s_n^2}} \frac{1 - e^{-\frac{n\gamma}{s_n^2}}}{s_n^2 \left(1 - e^{-\frac{\gamma}{s_n^2}} \right)} \rightarrow \frac{1 - e^{-\frac{\gamma}{1+\beta}}}{\gamma}.$$

Recalling that $\alpha = \frac{\beta}{1+\beta}$, the previous limit relations yield the desired result. $\square$

In terms of the relaxed Lindeberg index, Theorem 3.1 is turned into the following.

Theorem 3.4. *There exists a constant $C_{1/4} > 0$, not depending on $\{\xi_{n,k}\}$, such that*

$$\widetilde{\text{Lin}}_{1/4}(\{\xi_{n,k}\}) \leq C_{1/4} \limsup_{n \rightarrow \infty} K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right). \quad (39)$$

Moreover, (39) can be shown to hold with $C_{1/4} \leq 41$.

Theorem 3.4 generalizes the necessity of Lindeberg's condition in the Lindeberg–Feller CLT and therefore constitutes a counterpart for Theorem 2.7. It also has the following.

Corollary 3.5. *For each $\gamma > 0$ there exists a constant $C_\gamma > 0$, not depending on $\{\xi_{n,k}\}$, such that*

$$\widetilde{\text{Lin}}_\gamma(\{\xi_{n,k}\}) \leq C_\gamma \limsup_{n \rightarrow \infty} K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right). \quad (40)$$

Proof. Combine (37) and (39). $\square$

In the same spirit we have obtained Theorem 3.6, which is stronger than Theorem 3.4. It provides the sharpest lower bound for $\limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right)$ we have produced so far. We defer the proof to Appendix A.

Theorem 3.6. *There exists a constant $C_{1/2} > 0$, not depending on $\{\xi_{n,k}\}$, such that*

$$\widetilde{\text{Lin}}_{1/2}(\{\xi_{n,k}\}) \leq C_{1/2} \limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right). \quad (41)$$

Moreover, (41) can be shown to hold with $C_{1/2} \leq 30.3$.

4. Conclusion

Theorems 2.7 and 3.6 together yield Theorem 4.1, our main result, which answers Question 1.4 by providing an upper and a lower bound for the number $\limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right)$ which constitute a generalization of the Lindeberg–Feller CLT.

Theorem 4.1. *There exists a constant $\widetilde{C}_{1/2} > 0$, not depending on $\{\xi_{n,k}\}$, such that*

$$\widetilde{C}_{1/2} \widetilde{\text{Lin}}_{1/2}(\{\xi_{n,k}\}) \leq \limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right) \leq \text{Lin}(\{\xi_{n,k}\}). \quad (42)$$

Moreover, we can take $\widetilde{C}_{1/2} \geq 0.033$.

Applying (10), (38) and (42) to the STA $\{\eta_{\alpha,n,k}\}$, determined by (2), (7), (8) and (9), we obtain the following.

Theorem 4.2. *There exists a constant $\widetilde{C}_{1/2} > 0$ such that for each $\alpha > 0$*

$$\widetilde{C}_{1/2} \alpha \left(1 - \frac{1 - e^{-\frac{1}{2}(1-\alpha)}}{\frac{1}{2}(1-\alpha)}\right) \leq \limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \eta_{\alpha,n,k}\right) \leq \alpha. \quad (43)$$

Moreover, we can take $\widetilde{C}_{1/2} \geq 0.033$.

From the lower bound for $\limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \eta_{\alpha,n,k}\right)$ in (43) we conclude that the upper bound in (42) is of optimal order in the sense that the following holds.

Theorem 4.3. *There does not exist a constant $C > 0$, not depending on $\{\xi_{n,k}\}$, such that*

$$\limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right) \leq C [\text{Lin}(\{\xi_{n,k}\})]^{1+p} \quad (44)$$

with $p > 0$.

We end this paper with the following open question.

Question 4.4. *Does there exist a constant $C > 0$, not depending on $\{\xi_{n,k}\}$, such that*

$$\text{Lin}(\{\xi_{n,k}\}) \leq C \limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right)? \quad (45)$$

A positive answer to Question 4.4 is equivalent with the existence of a constant $C > 0$, not depending on $\{\xi_{n,k}\}$, such that for each $\gamma > 0$

$$\widetilde{\text{Lin}}_{\gamma}(\{\xi_{n,k}\}) \leq C \limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right).$$

Acknowledgment

Ben Berckmoes is Ph.D. fellow at the Fund for Scientific Research of Flanders (FWO).

Appendix A. Proof of Theorem 3.6

We assume without loss of generality that ξ and $\{\xi_{n,k}\}$ are independent. Furthermore, we let $\{\widetilde{\xi}_{n,k}\}$ be a copy of $\{\xi_{n,k}\}$ not depending on ξ and $\{\xi_{n,k}\}$.

Put

$$\psi_{1/2}(x) = 1 - \int_0^1 e^{-\frac{1}{2}(1-s^2)x^2} ds. \quad (46)$$

The following lemma reveals that $\psi_{1/2}$ is closely related to $\phi_{1/2}$.

Lemma 1. *The inequalities*

$$\psi_{1/2} \leq \phi_{1/2} \leq \frac{3}{2}\psi_{1/2} \quad (47)$$

are valid.

Proof. The first inequality is obvious. For the second inequality, we observe that

$$\begin{aligned} \phi_{1/2}(x) - \psi_{1/2}(x) &= e^{-\frac{1}{2}x^2} \int_0^1 \left(e^{\frac{1}{2}s^2x^2} - e^0 \right) ds \\ &= \frac{1}{2} e^{-\frac{1}{2}x^2} \int_0^1 \int_0^1 s^2 x^2 e^{\frac{1}{2}ts^2x^2} dt ds \\ &\leq \frac{1}{2} e^{-\frac{1}{2}x^2} \int_0^1 s^2 x^2 e^{\frac{1}{2}s^2x^2} ds \\ &= \frac{1}{2} e^{-\frac{1}{2}x^2} \int_0^1 s de^{\frac{1}{2}s^2x^2} \\ &= \frac{1}{2} e^{-\frac{1}{2}x^2} \left(e^{\frac{1}{2}x^2} - \int_0^1 e^{\frac{1}{2}s^2x^2} ds \right) \\ &= \frac{1}{2} \psi_{1/2}(x) \end{aligned}$$

and we are done. $\square$

It follows from (47) that $\psi_{1/2}$ belongs to Φ_L . Also, we have the following.

Corollary 2.

$$\widetilde{\text{Lin}}_{\psi_{1/2}}(\{\xi_{n,k}\}) \leq \widetilde{\text{Lin}}_{1/2}(\{\xi_{n,k}\}) \leq \frac{3}{2} \widetilde{\text{Lin}}_{\psi_{1/2}}(\{\xi_{n,k}\}). \quad (48)$$

Lemma 3. *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be bounded, measurable and antisymmetric and put*

$$g(x) = xf(x). \quad (49)$$

Then, for $a \in \mathbb{R}$,

$$\mathbb{E} \left[a^2 \int_0^1 g(\xi + sa) e^{-\frac{1}{2}(1-s^2)a^2} ds \right] = \mathbb{E} [af(\xi + a)]. \quad (50)$$

Proof. First suppose that f is continuously differentiable. Then

$$\begin{aligned} \mathbb{E} \left[a^2 \int_0^1 g(\xi + sa) e^{-\frac{1}{2}(1-s^2)a^2} ds \right] &= \mathbb{E} \left[a^2 \int_0^1 \xi f(\xi + sa) e^{-\frac{1}{2}(1-s^2)a^2} ds \right] + \mathbb{E} \left[a^2 \int_0^1 sa f(\xi + sa) e^{-\frac{1}{2}(1-s^2)a^2} ds \right] \\ &= \mathbb{E} \left[a^2 \int_0^1 f'(\xi + sa) e^{-\frac{1}{2}(1-s^2)a^2} ds \right] + \mathbb{E} \left[a \int_0^1 f(\xi + sa) d \left[e^{\frac{1}{2}s^2a^2} \right] e^{-\frac{1}{2}a^2} \right] \\ &= \mathbb{E} \left[a^2 \int_0^1 f'(\xi + sa) e^{-\frac{1}{2}(1-s^2)a^2} ds \right] \\ &\quad + \left(\mathbb{E} [af(\xi + a)] - \mathbb{E} \left[a^2 \int_0^1 f'(\xi + sa) e^{-\frac{1}{2}(1-s^2)a^2} ds \right] \right) \\ &= \mathbb{E} [af(\xi + a)], \end{aligned}$$

where the second equality follows from the fact that, for $h : \mathbb{R} \rightarrow \mathbb{R}$ continuously differentiable,

$$\mathbb{E}[\xi h(\xi)] = \mathbb{E}[h'(\xi)], \quad (51)$$

which is seen by performing an integration by parts. Now some standard approximation procedures allow us to drop the condition that f be continuously differentiable. $\square$

Lemma 4. Let μ be a symmetric probability distribution on the real line with Fourier transform

$$\widehat{\mu}(t) = \int_{-\infty}^{\infty} e^{-itx} d\mu(x). \quad (52)$$

Then, for $a \in \mathbb{R}$,

$$\mathbb{E}[\widehat{\mu}(\xi + a)] = \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \cos(ax) d\mu(x). \quad (53)$$

Proof. This is standard. $\square$

Lemma 5. Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be bounded with a piecewise continuous derivative. Then, for random variables η, η' ,

$$|\mathbb{E}[h(\eta) - h(\eta')]| \leq \left(\int_{-\infty}^{\infty} |h'(x)| dx \right) K(\eta, \eta'). \quad (54)$$

Proof. Perform an integration by parts on the left side of (54). $\square$

Theorem 6. Let μ be a symmetric probability distribution on the real line, put

$$f(x) = \begin{cases} \frac{1 - \widehat{\mu}(x)}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases} \quad (55)$$

and suppose that f has a bounded second derivative and a bounded and piecewise continuous third derivative. Then

$$\begin{aligned} & \left(\int_{-\infty}^{\infty} \left[1 - \left(1 + \frac{1}{2}x^2 \right) e^{-\frac{1}{2}x^2} \right] d\mu(x) \right) \widetilde{\text{Lin}}_{\psi_{1/2}}(\{\xi_{n,k}\}) \\ & \leq \left(\int_{-\infty}^{\infty} |(\widehat{\mu})'(x)| dx + \int_{-\infty}^{\infty} |f''(x)| dx \right) \limsup_{n \rightarrow \infty} K\left(\xi, \sum_{k=1}^n \xi_{n,k}\right). \end{aligned} \quad (56)$$

Proof. Put

$$g(x) = xf(x) \quad (57)$$

and

$$\beta_{n,k} = \int_0^1 [\widehat{\mu}(\xi) - \widehat{\mu}(\xi + s\xi_{n,k})] e^{-\frac{1}{2}(1-s^2)\xi_{n,k}^2} ds, \quad (58)$$

$$\gamma_n = g(\xi) - g\left(\sum_{k=1}^n \xi_{n,k}\right), \quad (59)$$

$$\delta_{n,k} = \int_0^1 \left[f'\left(s\xi_{n,k} + \sum_{i \neq k} \xi_{n,i} + \widetilde{\xi}_{n,k}\right) - f'(s\xi_{n,k} + \xi) \right] ds, \quad (60)$$

$$\epsilon_{n,k} = \int_0^1 \left[f'\left(\sum_{i \neq k} \xi_{n,i} + s\xi_{n,k}\right) - f'\left(\sum_{i \neq k} \xi_{n,i} + s\xi_{n,k} + \widetilde{\xi}_{n,k}\right) + \widetilde{\xi}_{n,k} f''\left(\sum_{i \neq k} \xi_{n,i} + s\xi_{n,k}\right) \right] ds. \quad (61)$$

Then

$$\sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 (1 - \widehat{\mu}(\xi)) \psi_{1/2}(|\xi_{n,k}|)] = \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 \beta_{n,k}] + \mathbb{E}[\gamma_n] + \sum_{k=1}^n \mathbb{E}[\xi_{n,k}^2 (\delta_{n,k} + \epsilon_{n,k})] \quad (62)$$

which is seen by calculating the right side applying independence, the fact that $\mathbb{E}[\widetilde{\xi}_{n,k}] = \mathbb{E}[\xi_{n,k}] = 0$ and (50).

It follows from (53) and the inequality $1 - \cos(a) \leq \frac{1}{2}a^2$ that

$$\begin{aligned} \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \beta_{n,k}] &= \int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2} \sum_{k=1}^n \mathbb{E} \left[\xi_{n,k}^2 \int_0^1 [1 - \cos(s\xi_{n,k}x)] e^{-\frac{1}{2}(1-s^2)\xi_{n,k}^2} ds \right] d\mu(x) \\ &\leq \left(\int_{-\infty}^{\infty} \frac{1}{2} x^2 e^{-\frac{1}{2}x^2} d\mu(x) \right) \sum_{k=1}^n \mathbb{E} \left[\xi_{n,k}^2 \int_0^1 s^2 \xi_{n,k}^2 e^{-\frac{1}{2}(1-s^2)\xi_{n,k}^2} ds \right] \\ &= \left(\int_{-\infty}^{\infty} \frac{1}{2} x^2 e^{-\frac{1}{2}x^2} d\mu(x) \right) \sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \psi_{1/2}(|\xi_{n,k}|)]. \end{aligned} \quad (63)$$

From (54) we learn that

$$\mathbb{E} [\gamma_n] \leq \left(\int_{-\infty}^{\infty} |(\widehat{\mu})'(x)| dx \right) K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right) \quad (64)$$

and

$$\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \delta_{n,k}] \leq \left(\int_{-\infty}^{\infty} |f''(x)| dx \right) K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right). \quad (65)$$

Finally, (23) reveals that

$$\sum_{k=1}^n \mathbb{E} [\xi_{n,k}^2 \epsilon_{n,k}] \leq \frac{1}{2} \|f'''\|_{\infty} \sum_{k=1}^n \sigma_{n,k}^4. \quad (66)$$

Now the lemma follows from (1). $\square$

Corollary 7. *There exists a constant $C_{\psi_{1/2}}$, not depending on $\{\xi_{n,k}\}$, such that*

$$\widetilde{\text{Lin}}_{\psi_{1/2}}(\{\xi_{n,k}\}) \leq C_{\psi_{1/2}} \limsup_{n \rightarrow \infty} K \left(\xi, \sum_{k=1}^n \xi_{n,k} \right). \quad (67)$$

Moreover, (67) can be shown to hold with $C_{\psi_{1/2}} \leq 20.2$.

Proof. Let μ_{σ} be normal with mean 0 and variance σ^2 and put

$$f_{\sigma}(x) = \frac{1 - \widehat{\mu}_{\sigma}(x)}{x}. \quad (68)$$

Then

$$\int_{-\infty}^{\infty} \left[1 - \left(1 + \frac{1}{2}x^2 \right) e^{-\frac{1}{2}x^2} \right] d\mu_{\sigma}(x) = 1 - \left(1 + \frac{1}{2} \frac{\sigma^2}{1 + \sigma^2} \right) \frac{1}{\sqrt{1 + \sigma^2}} \quad (69)$$

and

$$\int_{-\infty}^{\infty} |(\widehat{\mu}_{\sigma})'(x)| dx = 2. \quad (70)$$

Furthermore,

$$\int_{-\infty}^{\infty} |f''_{\sigma}(x)| dx = \sigma^2 - 4f'_{\sigma}(R_{\sigma}) \quad (71)$$

with R_{σ} the strictly positive zero of f''_{σ} .

If $\sigma = 1.7$, then $1.4912 < R_{\sigma} < 1.4914$ and applying (56) in this case reveals that we can take $C_{\psi_{1/2}} \approx 20.19$. $\square$

Remark. The authors have attempted to sharpen the upper bound for $C_{\psi_{1/2}}$ by applying Theorem 6 to non-normal probability distributions, but have not succeeded so far.

Proof (Theorem 3.6). Combine (48) and (67). $\square$

Appendix B. The theoretical framework behind the calculations

Since there are many metrics which metrize the weak topology, the reader may well ask why we have singled out the Kolmogorov metric K in our estimations and why we would be interested in assessing $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ the way we did. The deeper fundamental reason for this is to be found in approach theory [4,5,15].

Whereas convergence is handled in the weak topology, notions of approximate convergence can only be handled in a structure that allows for numbers, such as e.g. a metric. In the theory of approximation in metric or normed spaces such notions exist. Indeed, asymptotic radius and asymptotic center of a sequence are well-known [3,8,10,12,14]. Approach theory makes these concepts available in a much wider setting than merely metric spaces.

An approach structure on a set can be thought of as a generalized metric structure in the sense that it is also given by a distance (which we usually denote by δ), but not between pairs of points but between points and sets. As in the case of metric spaces, the natural mappings between approach spaces are contractions. Another analogy with metrics is that a distance δ has a canonical underlying topology which we sometimes refer to as the topology distancized by δ . Since we do not want to go into details and technicalities here, we refer the interested reader to [15].

Why do we then nevertheless end up with a metric, and why the Kolmogorov metric? This is not an arbitrary and ad-hoc choice, it is dictated by the theory, just like weak convergence is dictated by the weak topology. Consider the following setup.

Let $\mathcal{F}$ (respectively $\mathcal{F}_c$) stand for the collection of probability distributions (respectively continuous probability distributions) on the real line and convolution is denoted by $\star$. Now one of the main principles of Bergström's direct convolution method [6,16] is that weak convergence in $\mathcal{F}$ of a sequence $(\eta_n)_n$ to η is equivalent with uniform convergence of the sequence $(\eta_n \star \zeta)_n$ to $\eta \star \zeta$ for every continuous ζ . In other words, if we let $\mathcal{T}_w$ stand for the topology of weak convergence on $\mathcal{F}$ and $\mathcal{T}_K$ for the topology of uniform convergence (i.e. generated by the Kolmogorov metric) on $\mathcal{F}_c$, then $\mathcal{T}_w$ is the weakest topology on $\mathcal{F}$ making all mappings

$$(\mathcal{F} \rightarrow (\mathcal{F}_c, \mathcal{T}_K) : \eta \mapsto \eta \star \zeta)_{\zeta \in \mathcal{F}_c} \quad (72)$$

continuous. Fortunately, we can take the weakest topology with the above property, but this is something we cannot do with metrics, and this is where approach theory comes to the rescue.

If we replace the uniform topology $\mathcal{T}_K$ in (72) by its generating metric K , then we end up with the mappings

$$(\mathcal{F} \rightarrow (\mathcal{F}_c, K) : \eta \mapsto \eta \star \zeta)_{\zeta \in \mathcal{F}_c}. \quad (73)$$

We are not able to construct a weakest metric on $\mathcal{F}$, metrizing the weak topology and making all mappings in (73) contractive, it simply does not exist. However, approach spaces allow for so much flexibility that we are able to construct a weakest distance on $\mathcal{F}$, distancizing the weak topology and making all mappings in (73) contractive. We are interested in the approach structure determined by the latter distance. We call it the continuity approach structure and we denote it by δ_c . So we introduced δ_c in exactly the same way as $\mathcal{T}_w$.

It can be shown that the actual distance δ_c is given by

$$\delta_c(\eta, \mathcal{D}) = \sup_{\mathcal{F}_0} \inf_{\psi \in \mathcal{D}} \sup_{\zeta \in \mathcal{F}_0} K(\eta \star \zeta, \psi \star \zeta) \quad (74)$$

where $\eta \in \mathcal{F}$ and $\mathcal{D} \subset \mathcal{F}$ and where the first supremum runs over all finite subsets $\mathcal{F}_0$ of $\mathcal{F}_c$.

The nature of approach structures is such that they allow for many topological-like considerations. Thus in a topological space we can speak of convergent sequences and in an approach space we can speak of the limit operator of a sequence which in our case is given by

$$\lambda_c(\eta_n \rightarrow \eta) = \sup_{\zeta \in \mathcal{F}_c} \limsup_{n \rightarrow \infty} K(\eta \star \zeta, \eta_n \star \zeta). \quad (75)$$

Notice that $(\eta_n)_n$ converges weakly to η if and only if $\lambda_c(\eta_n \rightarrow \eta) = 0$.

The above δ_c is not a metric, it is a genuine approach structure and λ_c is not an asymptotic radius. So why then do we nevertheless end up with just the number $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$? This is because of the peculiarity of the present set-up.

We can prove the following results, where $j(\eta)$ stands for the supremum of the discontinuity jumps of the distribution η and where δ_K stands for the distance generated by the metric K in the usual way, i.e.

$$\delta_K(\eta, \mathcal{D}) = \inf_{\psi \in \mathcal{D}} K(\eta, \psi). \quad (76)$$

Theorem 1. For $\eta \in \mathcal{F}$ and $\mathcal{D} \subset \mathcal{F}$

$$\delta_c(\eta, \mathcal{D}) \leq \delta_K(\eta, \mathcal{D}) \leq \delta_c(\eta, \mathcal{D}) + j(\eta). \quad (77)$$

Corollary 2. For $\eta \in \mathcal{F}_c$ and $\mathcal{D} \subset \mathcal{F}$

$$\delta_c(\eta, \mathcal{D}) = \delta_K(\eta, \mathcal{D}). \quad (78)$$

Corollary 3. For $\eta \in \mathcal{F}_c$ and $(\eta_n)_n$ in $\mathcal{F}$

$$\lambda_c(\eta_n \rightarrow \eta) = \limsup_{n \rightarrow \infty} K(\eta, \eta_n). \quad (79)$$

The number $\limsup_{n \rightarrow \infty} K(\xi, \sum_{k=1}^n \xi_{n,k})$ is therefore not ad-hoc, it is nothing else but the limit-operator of the sequence $(\sum_{k=1}^n \xi_{n,k})_n$ evaluated in ξ for a canonical approach structure on the set of all probability distributions on the real line, and only in our particular case, where we are dealing with a continuous limit distribution, there is equality between the limit-operator for the continuity approach structure δ_c and the asymptotic radius for the Kolmogorov metric K .

References

- [1] A.D. Barbour, L.H.Y. Chen, An Introduction to Stein's Method, Singapore University Press, World Scientific Publishing Co. Pte. Ltd., Singapore, Hackensack, NJ, 2005.
- [2] A.D. Barbour, P. Hall, Stein's method and the Berry–Esseen theorem, Austral. J. Statist. 26 (1984) 8–15.
- [3] Y. Benyamini, Asymptotic centers and best approximation of operators into $C(K)$, Constr. Approx. 1 (1985) 217–229.
- [4] B. Berckmoes, R. Lowen, J. Van Casteren, Approach theory meets probability theory, Topology Appl. 158 (7) (2011) 836–852.
- [5] B. Berckmoes, R. Lowen, J. Van Casteren, Distances on probability measures and random variables, J. Math. Anal. Appl. 374 (2) (2011) 412–428.
- [6] H. Bergström, On the central limit theorem in the case of not equally distributed random variables, Skand. Aktuarietidskr 32 (1949) 37–62.
- [7] H. Bergström, Weak Convergence of Measures, in: Probability and Mathematical Statistics, Academic Press, Inc., Harcourt Brace Jovanovich, Publishers, New York, London, 1982.
- [8] S.C. Bose, Weak convergence to the fixed point of an asymptotically nonexpansive map, Proc. Amer. Math. Soc. 68 (3) (1978) 305–308.
- [9] L.H.Y. Chen, Q.-M. Shao, A non-uniform Berry–Esseen bound via Stein's method, Probab. Theory Related Fields 120 (2) (2001) 236–254.
- [10] M. Edelstein, The construction of an asymptotic center with a fixed-point property, Bull. Amer. Math. Soc. 78 (1972) 206–208.
- [11] W. Feller, An Introduction to Probability Theory and its Applications, Vol. II, second ed., John Wiley & Sons, Inc., New York, London, Sydney, 1971.
- [12] M.R. Haddadi, H. Mazaheri, G.M.H. Labbaf, Relation between fixed point and asymptotical center of nonexpansive maps, Fixed Point Theory Appl. (2011).
- [13] O. Kallenberg, Foundations of Modern Probability, second ed., in: Probability and its Applications (New York), Springer-Verlag, New York, 2002.
- [14] E. Lami Dozo, Asymptotic centers in particular spaces, in: Lecture Notes in Mathematics, vol. 886, 1981, pp. 199–207.
- [15] R. Lowen, Approach Spaces: The Missing Link in the Topology-Uniformity-Metric Triad, in: Oxford Mathematical Monographs, Oxford University Press, 1997.
- [16] L. Råde, A conversation with Harald Bergström, Statist. Sci. 12 (1) (1997) 53–60.
- [17] C. Stein, A bound for the error in the normal approximation to the distribution of a sum of dependent random variables, in: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), in: Probability Theory, vol. II, Univ. California Press, Berkeley, Calif, 1972, pp. 583–602.
- [18] C. Stein, Approximate Computation of Expectations, in: Institute of Mathematical Statistics Lecture Notes-Monograph Series, vol. 7, Institute of Mathematical Statistics, Hayward, CA, 1986.