The Lindeberg central limit theorem

Jordan Bell

May 29, 2015

1 Convergence in distribution

We denote by $\mathscr{P}(\mathbb{R}^d)$ the collection of Borel probability measures on $\mathbb{R}^d$. Unless we say otherwise, we use the **narrow topology** on $\mathscr{P}(\mathbb{R}^d)$: the coarsest topology such that for each $f \in C_b(\mathbb{R}^d)$, the map

$$\mu \mapsto \int_{\mathbb{R}^d} f d\mu$$

is continuous $\mathscr{P}(\mathbb{R}^d) \to \mathbb{C}$. Because $\mathbb{R}^d$ is a Polish space it follows that $\mathscr{P}(\mathbb{R}^d)$ is a Polish space.[1] (In fact, its topology is induced by the **Prokhorov metric**.[2])

2 Characteristic functions

For $\mu \in \mathscr{P}(\mathbb{R}^d)$, we define its **characteristic function** $\tilde{\mu} : \mathbb{R}^d \to \mathbb{C}$ by

$$\tilde{\mu}(u) = \int_{\mathbb{R}^d} e^{iu\cdot x} d\mu(x).$$

Theorem 1. If $\mu \in \mathscr{P}(\mathbb{R})$ has finite kth moment, $k \geq 0$, then, writing $\phi = \tilde{\mu}$:

1. $\phi \in C^k(\mathbb{R})$.
2. $\phi^{(k)}(v) = (i)^k \int_{\mathbb{R}} x^k e^{ivx} d\mu(x)$.
3. $\phi^{(k)}$ is uniformly continuous.
4. $|\phi^k(v)| \leq \int_{\mathbb{R}} |x|^k d\mu(x)$.

[1] Charalambos D. Aliprantis and Kim C. Border, *Infinite Dimensional Analysis: A Hitchhiker's Guide*, third ed., p. 515, Theorem 15.15; `http://individual.utoronto.ca/jordanbell/notes/narrow.pdf`

[2] Onno van Gaans, *Probability measures on metric spaces*, `http://www.math.leidenuniv.nl/~vangaans/jancol1.pdf`; Bert Fristedt and Lawrence Gray, *A Modern Approach to Probability Theory*, p. 365, Theorem 25.

Proof. For $0 \le l \le k$, define $f_l : \mathbb{R} \to \mathbb{C}$ by

$$f_l(v) = \int_{\mathbb{R}} x^l e^{ivx} d\mu(x).$$

For $h \neq 0$,

$$\left| x^l e^{ivx} \frac{e^{ihx} - 1}{h} \right| \le |x^l \cdot x| = |x|^{l+1},$$

so by the dominated convergence theorem we have for $0 \le l \le k - 1$,

$$\begin{aligned}
\lim_{h \to 0} \frac{f_l(v+h) - f_l(v)}{h} &= \lim_{h \to 0} \int_{\mathbb{R}} x^l e^{ivx} \frac{e^{ihx} - 1}{h} d\mu(x) \\
&= \int_{\mathbb{R}} x^l e^{ivx} \left(\lim_{h \to 0} \frac{e^{ihx} - 1}{h} \right) d\mu(x) \\
&= \int_{\mathbb{R}} i x^{l+1} e^{ivx} d\mu(x).
\end{aligned}$$

That is,

$$f_l' = i f_{l+1}.$$

And, by the dominated convergence, for $\epsilon > 0$ there is some $\delta > 0$ such that if $|w| < \delta$ then

$$\int_{\mathbb{R}} |x|^k |e^{iwx} - 1| d\mu(x) < \epsilon,$$

hence if $|v - u| < \delta$ then

$$\begin{aligned}
|f_k(v) - f_k(u)| &= \left| \int_{\mathbb{R}} x^k e^{iux} (e^{i(v-u)x} - 1) d\mu(x) \right| \\
&\le \int_{\mathbb{R}} |x|^k |e^{i(v-u)x} - 1| d\mu(x) \\
&< \epsilon,
\end{aligned}$$

showing that f_k is uniformly continuous. As well,

$$|f_k(v)| \le \int_{\mathbb{R}} |x|^k d\mu(x)$$

But $\phi = f_0$, i.e. $\phi^{(0)} = f_0$, so

$$\phi^{(1)} = f_0' = i f_1, \quad \phi^{(2)} = (i f_1)' = (i)^2 f_2, \quad \cdots, \quad \phi^{(k)} = (i)^k f_k.$$

□

If $\phi \in C^k(\mathbb{R})$, **Taylor's theorem** tells us that for each $x \in \mathbb{R}$,

$$\begin{aligned}
\phi(x) &= \sum_{l=0}^{k-1} \frac{\phi^{(l)}(0)}{l!} x^l + \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} \phi^{(k)}(t) dt \\
&= \sum_{l=0}^{k} \frac{\phi^{(l)}(0)}{l!} x^l + \int_0^x \frac{(x-t)^{k-1}}{(k-1)!} (\phi^{(k)}(t) - \phi^{(k)}(0)) dt \\
&= \sum_{l=0}^{k} \frac{\phi^{(l)}(0)}{l!} x^l + R_k(x),
\end{aligned}$$

and $R_k(x)$ satisfies

$$|R_k(x)| \leq \left(\sup_{0 \leq u \leq 1} |\phi^{(k)}(ux) - \phi^{(k)}(0)| \right) \cdot \frac{|x|^k}{k!}.$$

Define $\theta_k : \mathbb{R} \to \mathbb{C}$ by $\theta_k(0) = 0$ and for $x \neq 0$

$$\theta_k(x) = \frac{k!}{x^k} \cdot R_k(x),$$

with which, for all $x \in \mathbb{R}$,

$$\phi(x) = \sum_{l=0}^{k} \frac{\phi^{(l)}(0)}{l!} x^l + \frac{1}{k!} \theta_k(x) x^k.$$

Because R_k is continuous on $\mathbb{R}$, θ_k is continuous at each $x \neq 0$. Moreover,

$$|\theta_k(x)| \leq \sup_{0 \leq u \leq 1} |\phi^{(k)}(ux) - \phi^{(k)}(0)|,$$

and as $\phi^{(k)}$ is continuous it follows that θ_k is continuous at 0. Thus θ_k is continuous on $\mathbb{R}$.

Lemma 2. If $\mu \in \mathscr{P}(\mathbb{R})$ have finite kth moment, $k \geq 0$, and for $0 \leq l \leq k$,

$$M_l = \int_{\mathbb{R}} x^l d\mu(x),$$

then there is a continuous function $\theta : \mathbb{R} \to \mathbb{C}$ for which

$$\tilde{\mu}(x) = \sum_{l=0}^{k} \frac{(i)^l M_l}{l!} x^l + \frac{1}{k!} \theta(x) x^k.$$

The function θ satisfies

$$|\theta(x)| \leq \sup_{0 \leq u \leq 1} |\tilde{\mu}^{(k)}(ux) - \tilde{\mu}^{(k)}(0)|.$$

Proof. From Theorem 1, $\tilde{\mu} \in C^k(\mathbb{R})$ and

$$\tilde{\mu}^{(l)}(0) = (i)^l \int_{\mathbb{R}} x^l d\mu(x) = (i)^l M_l.$$

Thus from what we worked out above with Taylor's theorem,

$$\tilde{\mu}(x) = \sum_{l=0}^{k} \frac{(i)^l M_l}{l!} x^l + \frac{1}{k!}\theta_k(x) x^k,$$

for which

$$|\theta_k(x)| \leq \sup_{0 \leq u \leq 1} |\tilde{\mu}^{(k)}(ux) - \tilde{\mu}^{(k)}(0)|.$$

□

For $a \in \mathbb{R}$ and $\sigma > 0$, let

$$p(t, a, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-a)^2}{2\sigma^2}\right), \qquad t \in \mathbb{R}.$$

Let γ_{a,σ^2} be the measure on $\mathbb{R}$ whose density with respect to Lebesgue measure is $p(\cdot, a, \sigma^2)$. We call γ_{a,σ^2} a **Gaussian measure**. We calculate that the first moment of γ_{a,σ^2} is a and that its second moment is σ^2. We also calculate that

$$\tilde{\gamma}_{a,\sigma^2}(x) = \exp\left(iax - \frac{1}{2}\sigma^2 x^2\right).$$

Lévy's continuity theorem is the following.[3]

Theorem 3 (Lévy's continuity theorem)**.** Let μ_n be a sequence in $\mathscr{P}(\mathbb{R}^d)$.

1. If $\mu \in \mathscr{P}(\mathbb{R}^d)$ and $\mu_n \to \mu$, then for each $\tilde{\mu}_n$ converges to $\tilde{\mu}$ pointwise.
2. If there is some function $\phi : \mathbb{R}^d \to \mathbb{C}$ to which $\tilde{\mu}_n$ converges pointwise and ϕ is continuous at 0, then there is some $\mu \in \mathscr{P}(\mathbb{R}^d)$ such that $\phi = \tilde{\mu}$ and such that $\mu_n \to \mu$.

3 The Lindeberg condition, the Lyapunov condition, the Feller condition, and asymptotic negligibility

Let $(\Omega, \mathscr{F}, P)$ be a probability and let X_n, $n \geq 1$, be independent L^2 random variables. We specify when we impose other hypotheses on them; in particular,

[3] `http://individual.utoronto.ca/jordanbell/notes/martingaleCLT.pdf`, p. 19, Theorem 15.

we specfify if we suppose them to be identically distributed or to belong to L^p for $p > 2$.

For a random variable X, write

$$\sigma(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{E(|X - E(X)|^2)}.$$

Write

$$\sigma_n = \sigma(X_n),$$

and, using that the X_n are independent,

$$s_n = \sigma\left(\sum_{j=1}^n X_j\right) = \left(\sum_{j=1}^n \sigma_j^2\right)^{1/2}$$

and

$$\eta_n = E(X_n).$$

For $n \geq 1$ and $\epsilon > 0$, define

$$\begin{aligned} L_n(\epsilon) &= \frac{1}{s_n^2} \sum_{j=1}^n E((X_j - \eta_j)^2 | |X_j - \eta_j| \geq \epsilon s_n) \\ &= \frac{1}{s_n^2} \sum_{j=1}^n \int_{|x-\eta_j| \geq \epsilon s_n} (x - \eta_j)^2 d(X_{j*}P)(x). \end{aligned}$$

We say that the sequence X_n **satisfies the Lindeberg condition** if for each $\epsilon > 0$,

$$\lim_{n \to \infty} L_n(\epsilon) = 0.$$

For example, if the sequence X_n is identically distributed, then $s_n^2 = n\sigma_1^2$, so

$$\begin{aligned} L_n(\epsilon) &= \frac{1}{n\sigma_1^2} \sum_{j=1}^n \int_{|x-\eta_1| \geq \epsilon n^{1/2}\sigma_1} (x - \eta_1)^2 d(X_{1*}P)(x) \\ &= \frac{1}{\sigma_1^2} \int_{|x-\eta_1| \geq \epsilon n^{1/2}\sigma_1} (x - \eta_1)^2 d(X_{1*}P). \end{aligned}$$

But if μ is a Borel probability measure on $\mathbb{R}$ and $f \in L^1(\mu)$ and K_n is a sequence of compact sets that exhaust $\mathbb{R}$, then[4]

$$\int_{\mathbb{R} \setminus K_n} |f| d\mu \to 0, \qquad n \to \infty.$$

Hence $L_n(\epsilon) \to 0$ as $n \to \infty$, showing that X_n satisfies the Lindeberg condition.

[4] V. I. Bogachev, *Measure Theory*, volume I, p. 125, Proposition 2.6.2.

We say that the sequence X_n **satisfies the Lyapunov condition** if there is some $\delta > 0$ such that the X_n are $L^{2+\delta}$ and

$$\lim_{n\to\infty} \frac{1}{s_n^{2+\delta}} \sum_{j=1}^n E(|X_j - \eta_j|^{2+\delta}) = 0.$$

In this case, for $\epsilon > 0$, then $|x - \eta| \geq \epsilon s_n$ implies $|x - \eta|^{2+\delta} \geq |x - \eta|^2 (\epsilon s_n)^\delta$ and hence

$$\begin{aligned}
L_n(\epsilon) &\leq \frac{1}{s_n^2} \sum_{j=1}^n \int_{|x-\eta_j|\geq \epsilon s_n} \frac{|x-\eta_j|^{2+\delta}}{(\epsilon s_n)^\delta} d(X_{j*}P)(x) \\
&= \frac{1}{\epsilon s_n^{2+\delta}} \sum_{j=1}^n \int_{|x-\eta_j|\geq \epsilon s_n} |x-\eta_j|^{2+\delta} d(X_{j*}P)(x) \\
&= \frac{1}{\epsilon s_n^{2+\delta}} \sum_{j=1}^n \int_{|X_j-\eta_j|\geq \epsilon s_n} |X_j-\eta_j|^{2+\delta} dP \\
&\leq \frac{1}{\epsilon s_n^{2+\delta}} \sum_{j=1}^n E(|X_j - \eta_j|^{2+\delta}) \\
&\to 0.
\end{aligned}$$

This is true for each $\epsilon > 0$, showing that if X_n satisfies the Lyapunov condition then it satisfies the Lindeberg condition.

For example, if X_n are identically distributed and $L^{2+\delta}$, then

$$\frac{1}{s_n^{2+\delta}} \sum_{j=1}^n E(|X_j - \eta_j|^{2+\delta}) = \frac{1}{n^{\delta/2}\sigma_1^{2+\delta}} E(|X_j - \eta_j|^{2+\delta}) \to 0,$$

showing that X_n satisfies the Lyapunov condition.

Another example: Suppose that the sequence X_n is bounded by M almost surely and that $s_n \to \infty$. $|X_n| \leq M$ almost surely implies that

$$|\eta_n| = |E(X_n)| \leq E(|X_n|) \leq E(M) = M.$$

Therefore $|X_n - \eta_n| \leq |X_n| + |\eta_n| \leq 2M$ almost surely. Let $\delta > 0$. Then, as $s_n^2 = n\sigma_1^2$,

$$\begin{aligned}
\frac{1}{s_n^{2+\delta}} \sum_{j=1}^n E(|X_j - \eta_j|^{2+\delta}) &\leq \frac{1}{s_n^{2+\delta}} \sum_{j=1}^N E(|X_j - \eta_j|^2)(2M)^\delta \\
&= \frac{(2M)^\delta}{s_n^\delta} \\
&\to 0,
\end{aligned}$$

showing that X_n satisfies the Lyapunov condition.

We say that a sequence of random variables X_n satisfies the **Feller condition** when

$$\lim_{n\to\infty} \max_{1\leq j\leq n} \frac{\sigma_j}{s_n} = 0,$$

where $\sigma_j = \sigma(X_j) = \sqrt{\operatorname{Var}(X_j)}$ and

$$s_n = \left(\sum_{j=1}^n \sigma_j^2 \right)^{1/2}.$$

We prove that if a sequence satisfies the Lindeberg condition then it satisfies the Feller condition.[5]

Lemma 4. If a sequence of random variables X_n satisfies the Lindeberg condition, then it satisfies the Feller condition.

Proof. Let $\epsilon > 0$¿ For $n \geq 1$ and $1 \leq k \leq n$, we calculate

$$\begin{aligned}
\sigma_k^2 &= \int_{\mathbb{R}} (x-\eta_k)^2 d(X_{k*}P)(x) \\
&= \int_{|x-\eta_k|<\epsilon s_n} (x-\eta_k)^2 d(X_{k*}P)(x) + \int_{|x-\eta_k|\geq\epsilon s_n} (x-\eta_k)^2 d(X_{k*}P)(x) \\
&\leq (\epsilon s_n)^2 + \sum_{j=1}^n \int_{|x-\eta_j|\geq\epsilon s_n} (x-\eta_j)^2 d(X_{j*}P)(x) \\
&= \epsilon^2 s_n^2 + s_n^2 L_n(\epsilon).
\end{aligned}$$

Hence

$$\max_{1\leq k\leq n} \left(\frac{\sigma_k}{s_n} \right)^2 \leq \epsilon^2 + L_n(\epsilon),$$

and so, because the X_n satisfy the Lindeberg condition,

$$\limsup_{n\to\infty} \max_{1\leq k\leq n} \left(\frac{\sigma_k}{s_n} \right)^2 \leq \epsilon^2.$$

This is true for all $\epsilon > 0$, which yields

$$\lim_{n\to\infty} \max_{1\leq k\leq n} \left(\frac{\sigma_k}{s_n} \right)^2 = 0,$$

namely, that the X_n satisfy the Feller condition. □

We do not use the following idea of an asymptotically negligible family of random variables elsewhere, and merely take this as an excsuse to write out

[5] Heinz Bauer, *Probability Theory*, p. 235, Lemma 28.2.

what it means. A family of random variables $X_{n,j}$, $n \geq 1$, $1 \leq j \leq k_n$, is called **asymptotically negligible**[6] if for each $\epsilon > 0$,

$$\lim_{n\to\infty} \max_{1\leq j\leq k_n} P(|X_{n,j}| \geq \epsilon) = 0.$$

A sequence of random variables X_n converging in probability to 0 is equivalent to it being asymptotically negligible, with $k_n = 1$ for each n.

For example, suppose that $X_{n,j}$ are L^2 random variables each with $E(X_{n,j}) = 0$ and that they satisfy

$$\lim_{n\to\infty} \max_{1\leq j\leq k_n} \mathrm{Var}(X_{n,j}) = 0.$$

For $\epsilon > 0$, by Chebyshev's inequality,

$$P(|X_{n,j}| \geq \epsilon) \leq \frac{1}{\epsilon^2} E(|X_{n,j}|^2) = \frac{1}{\epsilon^2}\mathrm{Var}(X_{n,j}),$$

whence

$$\lim_{n\to\infty} \max_{1\leq j\leq k_n} P(|X_{n,j}| \geq \epsilon) \leq \limsup_{n\to\infty} \max_{1\leq j\leq k_n} \frac{1}{\epsilon^2}\mathrm{Var}(X_{n,j}) = 0,$$

and so the random variables $X_{n,j}$ are asymptotically negligible.

Another example: Suppose that random variables $X_{n,j}$ are identically distributed, with $\mu = {X_{n,j}}_* P$. For $\epsilon > 0$,

$$P\left(\left|\frac{X_{n,j}}{n}\right| \geq \epsilon\right) = P(|X_{n,j}| \geq n\epsilon) = \mu(A_n),$$

where $A_n = \{x \in \mathbb{R} : |x| \geq n\epsilon\}$. As $A_n \downarrow \emptyset$, $\lim_{n\to\infty} \mu(A_n) = 0$. Hence the random variables $\frac{X_{n,j}}{n}$ are asymptotic negligible.

The following is a statement about the characteristic functions of an asymptotically negligible family of random variables.[7]

Lemma 5. Suppose that a family $X_{n,j}$, $n \geq 1$, $1 \leq j \leq k_n$, of random variables is asymptotically negligible, and write $\mu_{n,j} = {X_{n,j}}_* P$ and $\phi_{n,j} = \tilde{\mu}_{n,j}$. For each $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \max_{1\leq j\leq k_n} |\phi_{n,j}(x) - 1| = 0.$$

Proof. For any real t, $|e^{it} - 1| \leq |t|$. For $x \in \mathbb{R}$, $\epsilon > 0$, $n \geq 1$, and $1 \leq j \leq k_n$,

$$\begin{aligned}
|\phi_{n,j}(x) - 1| &= \left|\int_{\mathbb{R}} (e^{ixy} - 1) d\mu_{n,j}(y)\right| \\
&\leq \int_{|y|<\epsilon} |e^{ixy} - 1| d\mu_{n,j}(y) + \int_{|y|\geq\epsilon} |e^{ixy} - 1| d\mu_{n,j}(y) \\
&\leq \int_{|y|<\epsilon} |xy| d\mu_{n,j}(y) + \int_{|y|\geq\epsilon} 2 d\mu_{n,j}(y) \\
&\leq \epsilon|x| + 2P(|X_{n,j}| \geq \epsilon).
\end{aligned}$$

[6]Heinz Bauer, *Probability Theory*, p. 225, §27.2.

[7]Heinz Bauer, *Probability Theory*, p. 227, Lemma 27.3.

Hence

$$\max_{1\le j\le k_n} |\phi_{n,j}(x)-1| \le \epsilon|x| + 2\max_{1\le j\le k_n} P(|X_{n,j}|\ge\epsilon).$$

Using that the family $X_{n,j}$ is asymptotically negligible,

$$\limsup_{n\to\infty}\max_{1\le j\le k_n}|\phi_{n,j}(x)-1|\le 2\epsilon|x|.$$

But this is true for all $\epsilon>0$, so

$$\limsup_{n\to\infty}\max_{1\le j\le k_n}|\phi_{n,j}(x)-1| = 0,$$

proving the claim. □

4 The Lindeberg central limit theorem

We now prove the **Lindeberg central limit theorem**.[8]

Theorem 6 (Lindeberg central limit theorem)**.** If X_n is a sequence of independent L^2 random variables that satisfy the Lindeberg condition, then

$$S_{n*}P \to \gamma_1,$$

where

$$S_n = \frac{1}{s_n}\sum_{j=1}^n (X_j-\eta_j) = \frac{\sum_{j=1}^n (X_j - E(X_j))}{\sigma(X_1+\cdots+X_n)}.$$

Proof. The sequence $Y_n = X_n - E(X_n)$ are independent L^2 random variables that satisfy the Lindeberg condition and $\sigma(Y_n)=\sigma(X_n)$. Proving the claim for the sequence Y_n will prove the claim for the sequence X_n, and thus it suffices to prove the claim when $E(X_n)=0$, i.e. $\eta_n=0$.

For $n\ge 1$ and $1\le j\le n$, let

$$\mu_{n,j} = \left(\frac{X_j}{s_n}\right)_* P \qquad \text{and} \qquad \tau_{n,j} = \frac{\sigma_j}{s_n}.$$

The first moment of $\mu_{n,j}$ is

$$\int_{\mathbb{R}} x d\left(\left(\frac{X_j}{s_n}\right)_* P\right)(x) = \int_\Omega \frac{X_j}{s_n} dP = \frac{1}{s_n}E(X_j) = 0,$$

and the second moment of $\mu_{n,j}$ is

$$\int_{\mathbb{R}} x^2 d\left(\left(\frac{X_j}{s_n}\right)_* P\right)(x) = \int_\Omega \left(\frac{X_j}{s_n}\right)^2 dP = \frac{1}{s_n^2}E(X_j^2) = \frac{\sigma_j^2}{s_n^2} = \tau_{n,j}^2,$$

[8]Heinz Bauer, *Probability Theory*, p. 235, Theorem 28.3.

for which

$$\sum_{j=1}^{n} \tau_{n,j}^2 = \frac{1}{s_n^2} \sum_{j=1}^{n} \sigma_j^2 = 1.$$

For $\mu \in \mathscr{P}(\mathbb{R})$ with first moment $\int_{\mathbb{R}} x d\mu(x) = 0$ and second moment $\int_{\mathbb{R}} x^2 d\mu(x) = \sigma^2 < \infty$, Lemma 2 tells us that

$$\tilde{\mu}(x) = M_0 + iM_1 x - \frac{M_2}{2} x^2 + \frac{1}{2}\theta_2(x) x^2 = 1 - \frac{\sigma^2}{2} x^2 + \frac{1}{2}\theta(x) x^2,$$

with

$$|\theta(x)| \leq \sup_{0 \leq u \leq 1} |\tilde{\mu}''(ux) - \tilde{\mu}''(0)|.$$

But by Lemma 1,

$$\tilde{\mu}''(ux) = -\int_{\mathbb{R}} y^2 e^{iuxy} d\mu(y),$$

so

$$\begin{aligned} |\theta(x)| &\leq \sup_{0 \leq u \leq 1} \left| \int_{\mathbb{R}} y^2(-e^{iuxy} + 1) d\mu(y) \right| \\ &\leq \sup_{0 \leq u \leq 1} \int_{\mathbb{R}} y^2 |e^{iuxy} - 1| d\mu(y). \end{aligned}$$

For $0 \leq u \leq 1$, $|e^{iuxy} - 1| \leq |uxy| \leq |xy|$, so for $x \in \mathbb{R}$ and $\epsilon > 0$, with $\delta = \min\left\{\epsilon, \frac{\epsilon}{|x|}\right\}$, when $|y| < \delta$ and $0 \leq u \leq 1$ we have $|e^{iuxy} - 1| < \epsilon$. Thus

$$\begin{aligned} |\theta(x)| &\leq \sup_{0 \leq u \leq 1} \int_{|y|<\delta} y^2 |e^{iuxy} - 1| d\mu(y) + \sup_{0 \leq u \leq 1} \int_{|y| \geq \delta} y^2 |e^{iuxy} - 1| d\mu(y) \\ &\leq \epsilon \int_{|y|<\delta} y^2 d\mu(y) + 2 \int_{|y| \geq \delta} y^2 d\mu(y) \\ &\leq \epsilon\sigma^2 + 2 \int_{|y| \geq \delta} y^2 d\mu(y). \end{aligned}$$

Let $x \in \mathbb{R}$ and $\epsilon > 0$, and take $\delta = \min\left\{\epsilon, \frac{\epsilon}{|x|}\right\}$. On the one hand, for $n \geq 1$ and $1 \leq j \leq n$, because the first moment of $\mu_{n,j}$ is 0 and its second moment is $\tau_{n,j}^2$,

$$\tilde{\mu}_{n,j}(x) = 1 - \frac{\tau_{n,j}^2}{2} x^2 + \frac{1}{2}\theta_{n,j}(x) x^2,$$

with, from the above,

$$|\theta_{n,j}(x)| \leq \epsilon\tau_{n,j}^2 + 2 \int_{|y| \geq \delta} y^2 d\mu_{n,j}(y).$$

On the other hand, the first moment of the Gaussian measure $\gamma_{0,\tau_{n,j}^2}$ is 0 and its second moment is $\tau_{n,j}^2$. Its characteristic function is

$$\tilde{\gamma}_{0,\tau_{n,j}^2}(x) = \exp\left(-\frac{\tau_{n,j}^2}{2} x^2\right) = 1 - \frac{\tau_{n,j}^2}{2} x^2 + \frac{1}{2}\psi_{n,j}(x) x^2,$$

with, from the above,

$$|\psi_{n,j}(x)| \leq \epsilon\tau_{n,j}^2 + 2\int_{|y|\geq\delta} y^2 d\gamma_{0,\tau_{n,j}^2}(x).$$

In particular, for all $x \in \mathbb{R}$,

$$\tilde{\mu}_{n,j}(x) - \tilde{\gamma}_{0,\tau_{n,j}^2}(x) = \frac{x^2}{2}\left(\theta_{n,j}(x) - \psi_{n,j}(x)\right).$$

For $k \geq 1$ and for $a_l, b_l \in \mathbb{C}$, $1 \leq l \leq k$,

$$\prod_{l=1}^k a_l - \prod_{l=1}^k b_l = \sum_{l=1}^k b_1 \cdots b_{l-1}(a_l - b_l)a_{l+1}\cdots a_k.$$

If further $|a_l| \leq 1$, $|b_l| \leq 1$, then

$$\left|\prod_{l=1}^k a_l - \prod_{l=1}^k b_l\right| \leq \sum_{l=1}^k |a_l - b_l|. \tag{1}$$

Because the X_n are independent, the distribution of

$$S_n = \sum_{j=1}^n \frac{X_j}{s_n}$$

is the convolution of the distributions of the summands:

$$\mu_{n,1} * \cdots * \mu_{n,n},$$

whose characteristic function is

$$\phi_n = \prod_{j=1}^n \tilde{\mu}_{n,j},$$

since the characteristic function of a convolution of measures is the product of the characteristic functions of the measures. Using $\sum_{j=1}^n \tau_{n,j}^2 = 1$ and (1), for $x \in \mathbb{R}$ we have

$$\begin{aligned}
|\phi_n(x) - e^{-\frac{x^2}{2}}| &= \left|\prod_{j=1}^n \tilde{\mu}_{n,j}(x) - \prod_{j=1}^n e^{-\frac{1}{2}\tau_{n,j}^2 x^2}\right| \\
&\leq \sum_{j=1}^n \left|\tilde{\mu}_{n,j}(x) - e^{-\frac{1}{2}\tau_{n,j}^2 x^2}\right| \\
&= \sum_{j=1}^n \left|\tilde{\mu}_{n,j}(x) - \tilde{\gamma}_{0,\tau_{n,j}^2}(x)\right| \\
&= \frac{x^2}{2}\sum_{j=1}^n |\theta_{n,j}(x) - \psi_{n,j}(x)|.
\end{aligned}$$

Therefore, for $x \in \mathbb{R}$, $\epsilon > 0$, and $\delta = \min\left\{\epsilon, \frac{\epsilon}{|x|}\right\}$,

$$\begin{aligned}
&|\phi_n(x) - e^{-\frac{x^2}{2}}| \\
\leq &\frac{x^2}{2}\sum_{j=1}^{n}\left(\epsilon\tau_{n,j}^2 + 2\int_{|y|\geq\delta} y^2 d\mu_{n,j}(y) + \epsilon\tau_{n,j}^2 + 2\int_{|y|\geq\delta} y^2 d\gamma_{0,\tau_{n,j}^2}(y)\right) \\
= &\epsilon x^2 + x^2\sum_{j=1}^{n}\int_{|y|\geq\delta} y^2 d\mu_{n,j}(y) + x^2\sum_{j=1}^{n}\int_{|y|\geq\delta} y^2 d\gamma_{0,\tau_{n,j}^2}(y).
\end{aligned}$$

We calculate

$$\begin{aligned}
L_n(\delta) &= \frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|y|\geq\delta s_n} y^2 d(X_{j*}P)(y) \\
&= \frac{1}{s_n^2}\sum_{j=1}^{n}\int_{|X_j|\geq\delta s_n} X_j^2 dP \\
&= \sum_{j=1}^{n}\int_{\left|\frac{X_j}{s_n}\right|\geq\delta}\left(\frac{X_j}{s_n}\right)^2 dP \\
&= \sum_{j=1}^{n}\int_{|y|\geq\delta} y^2 d\left(\left(\frac{X_j}{s_n}\right)_* P\right)(y) \\
&= \sum_{j=1}^{n}\int_{|y|\geq\delta} y^2 d\mu_{n,j}(y).
\end{aligned}$$

Hence, the fact that the X_n satisfy the Lindeberg condition yields

$$\limsup_{n\to\infty}|\phi_n(x) - e^{-\frac{x^2}{2}}| \leq \epsilon x^2 + x^2\limsup_{n\to\infty}\sum_{j=1}^{n}\int_{|y|\geq\delta} y^2 d\gamma_{0,\tau_{n,j}^2}(y). \qquad (2)$$

Write

$$\alpha_n = \max_{1\leq j\leq n}\tau_{n,j} = \max_{1\leq j\leq n}\frac{\sigma_j}{s_n}.$$

We calculate

$$\begin{aligned}
\sum_{j=1}^{n}\int_{|y|\geq\delta} y^2 d\gamma_{0,\tau_{n,j}^2}(y) &= \sum_{j=1}^{n}\int_{|y|\geq\delta} y^2\frac{1}{\tau_{n,j}\sqrt{2\pi}}\exp\left(-\frac{y^2}{2\tau_{n,j}^2}\right)dy \\
&= \sum_{j=1}^{n}\tau_{n,j}^2\int_{|u|\geq\delta/\tau_{n,j}} u^2 d\gamma_{0,1}(u) \\
&\leq \sum_{j=1}^{n}\tau_{n,j}^2\int_{|u|\geq\delta/\alpha_n} u^2 d\gamma_{0,1}(u) \\
&= \int_{|u|\geq\delta/\alpha_n} u^2 d\gamma_{0,1}(u).
\end{aligned}$$

Because the sequence X_n satisfies the Lindeberg condition, by Lemma 4 it satisfies the Feller condition, which means that $\alpha_n \to 0$ as $n \to \infty$. Because $\alpha_n \to 0$ as $n \to \infty$, $\delta/\alpha_n \to \infty$ as $n \to \infty$, hence

$$\int_{|u| \geq \delta/\alpha_n} u^2 d\gamma_{0,1}(u) \to 0$$

as $n \to \infty$. Thus we get

$$\sum_{j=1}^{n} \int_{|y| \geq \delta} y^2 d\gamma_{0,\tau_{n,j}^2}(y) \to 0$$

as $n \to \infty$. Using this with (2) yields

$$\limsup_{n \to \infty} |\phi_n(x) - e^{-\frac{x^2}{2}}| \leq \epsilon x^2.$$

This is true for all $\epsilon > 0$, so

$$\lim_{n \to \infty} |\phi_n(x) - e^{-\frac{x^2}{2}}| = 0,$$

namely, ϕ_n (the characteristic function of $S_{n*}P$) converges pointwise to $e^{-\frac{x^2}{2}}$. Moreover, $e^{-\frac{x^2}{2}}$ is indeed continuous at 0, and $e^{-\frac{x^2}{2}} = \tilde{\gamma}_{0,1}(x)$. Therefore, Lévy's continuity theorem (Theorem 3) tells us that $S_{n*}P$ converges narrowly to $\gamma_{0,1}$, which is the claim. $\square$