Department of Health Sciences,
1971 University Blvd,
Liberty University,
Lynchburg, VA 24515

@jpPabl0
jpmonteagudo@liberty.edu
@jpmonteagudo28
http://www.jpmonteagudo.com

The Lindeberg– Lévy's 1922 "Classic" Central Limit Theorem

Here I provide a more technical summary of Paul Lévy's special case of the classic central limit theorem as discussed in his 1924 article, "Théorie des erreurs. La loi de Gauss et les lois exceptionnelles" and also mentioned by Hans Fischer (2011) in *A History of the Central Limit Theorem*.

For a sequence of distribution functions (F_k), $k \in \mathbb{R}$ of independent random variables X_k, each with zero expectation and variance 1, let

$$\forall \epsilon > 0 \exists a > 0 \exists k \in \mathbb{N} : \int_{|\xi| \leq a} \xi^2 dF_k(\xi) \geq 1 - \epsilon$$

Let $(m_k)_{k \in \mathbb{N}}$ be a sequence of positive numbers with

$$\frac{\max_{k=1}^{n} m_k^2}{\sum_{k=1}^{n} m_k^2} \to 0, (n \to \infty).$$

Then

$$\lim_{n \to \infty} P\left(\frac{\sum_{k=1}^{n} (m_k X_k)}{\sqrt{\sum_{k=1}^{n} m_k^2}} \leq x \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \, dt$$

The first statement shows that for all positive ϵ– errors– there exists an a and a positive integer k that's greater than zero such that the integral of ξ^2– the second moment or variance of X_k– with respect to its distribution function over a specific range $|\xi| \leq a$ will be greater than or equal to $1 - \epsilon$. In other words, for any small level of error, we can find a specific range (determined by a) and a specific distribution from a sequence (determined by k) such that most of the "density" of that distribution falls within that range. This integration then controls the behavior of the tails of the distribution.

Next, we have a sequence of positive integers m_k used as scaling factors or weights for the random variables X_k. The only admissible m_k weights are those for which, as the number of terms (n) increases, the ratio of the maximum squared scaling factor to the sum of squared scaling factors tends to zero. As you consider an increasingly large number of terms in the sum of scaled independent random variables, no single scaling factor dominates the contribution to the sum. The impact of the largest squared scaling factor becomes negligible compared to the cumulative effect of all squared scaling factors.

Then, as $n-$ the sample size$-$ increases towards infinity, the distribution of the standardized sum $\frac{\sum_{k=1}^{n}(m_k X_k)}{\sqrt{\sum_{k=1}^{n} m_k^2}}$, formed by the weighted sum of the independent random variables X_k, converges to standard normal distribution. This means the probability of this standardized sum being less than a specific value x corresponds to the area less than or equal to x under the standard normal curve.

At the time Lévy presented his formulation of the CLT, Lindeberg submitted a note to the Paris Academy in which he proposed a more general condition than Aleksandr Lyapunov's. Under Lindeberg's condition, we have a random variable X_k with $\mu = 0$ and $\sigma^2 = 1$ and Levy's formulation can be obtained by substituting

$$\forall \epsilon > 0 \exists a > 0 \exists k \in \mathbb{N} : \int_{|\xi| \leq a} \xi^2 dF_k(\xi) \geq 1 - \epsilon$$

by

$$\forall \epsilon > 0 \exists a > 0 \exists k \in \mathbb{N} : \frac{1}{\sigma_k^2} \int_{|\xi| \leq a\sigma_k} \xi^2 \, dF_k(\xi) \geq 1 - \epsilon.$$

Making $\sigma_{k_2} = m_{k_2}$ and $r_{n^2} = \sum_{k=1}^{n} \sigma_k^2$, we get

$$\forall t > 0 \forall \eta > 0 \exists n_0 \forall n \geq n_0 : \frac{1}{r_n^2} \int_{|x| \leq r_n t} x^2 \, dF_k(x) \geq 1 - \eta.$$

Arriving at the Classic Central Limit Theorem

After dealing with characteristic and moment–generating functions, we arrive at the "classic" CLT

$$\sqrt{n}(\bar{X}_k - E[X_k]) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

for every real number z,

$$\lim_{n \to \infty} P\left(\frac{\sqrt{n}(\bar{X}_k - E[X_k])}{\sigma} \leq \frac{z}{\sigma}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{z}{\sigma}} e^{-\frac{t^2}{2}} \, dt.$$

For any sequence of independent, identically distributed (i.i.d) random variables with mean equal zero and finite variance, as n approaches infinity, the random variables $\sqrt{n}(\bar{X}_k - E[X_k])$ converge to a normal distribution.

Relaxed assumptions in the Central Limit Theorem

The classic central limit theorem states that given (i.i.d) random variables X_i with expectation μ and variance $\sigma^2 < \infty$; the distribution of X_i will converge to a standard normal distribution; **however**, we can weaken these conditions to create variations of the CLT.

Non-identically distributed random variables

Let X_n be any sequence of independent random variables with $\mu = 0$ and $\sigma^2 < \infty$, with variance

$$S_n^2 = \sum_i^n \sigma_n^2$$

Lyapunov's CLT

The larger the δ, the faster the tails of the distribution must decay.

If for some $\delta > 0$, Lyapunov's condition

$$\lim_{n\to\infty} \frac{1}{S_n^{2+\delta}} \sum_{i=1}^n E[|X_n - \mu_i|^{2+\delta}] = 0$$

is satisfied, then a sum of $\frac{X_n - \mu_n}{s_n}$ converges to a standard normal random variable as n goes to infinity.

$$\frac{1}{S_n} \sum_{i=1}^n (X_n - \mu_n) \xrightarrow{d} \mathcal{N}(0,1)$$

Lindeberg's CLT

Lindeberg's condition also requires that X_n be independent, random variable with finite variance and that for some $\epsilon > 0$,

$$\lim_{n\to\infty} \frac{1}{S_n^2} \sum_{i=1}^n E[(X_n - \mu_n)^2 \cdot 1_{\{|X_n - \mu_n| > \epsilon_{S_n}\}}] = 0.$$

$1_{\{\ldots\}}$ represents an indicator function

If the previous condition is satisfied, then

$$\frac{1}{S_n} \sum_i^n (X_n - \mu_n) \xrightarrow{d} \mathcal{N}(0,1)$$

Infinite Variance– The Generalized CLT

Under the Generalized CLT (GCLT), we require that X_n be i.i.d but don't impose a constraint of finite variance.

A distribution is said to have infinite variance when the upper bound of its expectation is unknown

For random variables X_n with infinite variance, we would need sequences of constants $a_n > 0$ and $b_n > 0$ such that

$$\frac{\sum_{i=1}^{n}(X_n) - b_n}{a_n} \xrightarrow{d} Z,$$

with Z being a non-degenerate random variable for some $0 < \alpha \leq 2$. Therefore, if the sums of the i.i.d random variables converge in distribution to Z, then Z must be a stable distribution. When $a = 2$, X_n converges to a normally distributed random variable.

References

Amir, Ariel. 2019. "An elementary renormalization-group approach to the Generalized Central Limit Theorem and Extreme Value Distributions." *ArXiv.org*, https://doi.org/10.1088/1742-5468/ab5b8c.

Bolthausen, Erwin, and Mario V Wüthrich. 2013. "Bernoulli's Law of Large Numbers."

Fischer, Hans. 2011. *History of the Central Limit Theorem: From Classical to Modern Probability Theory.* http://www.springer.com/series/4142.

Kwak, Sang Gyu, and Jong Hae Kim. 2017. "Central limit theorem: The cornerstone of modern statistics." *Korean Journal of Anesthesiology* 70 (2): 144–156. https://doi.org/10.4097/kjae.2017.70.2.144.

Lévy, Paul. 1924. "Théorie des erreurs.La loi de Gauss et les lois exceptionnelles." *Bulletin de la S.M.F.* 52:49–85.

Loeve, M. 1950. *Fundamental Limit Theorems of Probability Theory.*