

The Central Limit Theorem for Triangular Arrays

By Yong Kab Choi & Sung Min Hong
 Gyeongsang National University, Jinju, Korea

The purpose of this note is to show that Lindeberg Theorem holds for triangular arrays and that Lindeberg Theorem represents the special case of the Central Limit Theorem for triangular arrays.

Let X_1, X_2, \dots be independent random variables with c.d.f.s $F_1(x), F_2(x), \dots$ such that $E(X_k) = m_k$ finite, and $V(X_k) = \sigma_k^2 < \infty$, and let $x_n = \sum_{k=1}^n X_k$, $\xi_n = \sum_{k=1}^n m_k$, $\tau_n^2 = \sum_{k=1}^n \sigma_k^2$.

We say that the Lindeberg Condition is satisfied if for each fixed $\varepsilon > 0$,

$$(1) \quad \frac{1}{\tau_n^2} \sum_{k=1}^n \int_{\{|x-m_k| \geq \varepsilon \tau_n\}} (x-m_k)^2 dF_k(x) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Theorem 1. (Lindeberg Theorem)

If the Lindeberg Condition (1) holds, the distribution of $\frac{1}{\tau_n}(x_n - \xi_n)$ converges to the standard normal distribution.

Definition. By a triangular array is meant a double sequence of random variables X_{kn} ($k=1, 2, \dots, n; n=1, 2, \dots$) such that the random variables X_{1n}, \dots, X_{nn} of the n -th row are mutually independent.

For each n let X_{1n}, X_{2n}, \dots be independent random variables with c.d.f.s $F_{1n}(x), F_{2n}(x), \dots$ and its characteristic functions $\varphi_1(t), \varphi_2(t), \dots$ such that $E(X_{kn}) = m_{kn}$ finite, and $V(X_{kn}) = \sigma_{kn}^2 < \infty$, and let

$$T_n = \sum_{k=1}^n X_{kn}, \quad \eta_n = \sum_{k=1}^n m_{kn}, \quad C_n^2 = \sum_{k=1}^n \sigma_{kn}^2.$$

For each fixed $\varepsilon > 0$, let

$$(2) \quad \frac{1}{C_n^2} \sum_{k=1}^n \int_{\{|x-m_{kn}| \geq \varepsilon C_n\}} (x-m_{kn})^2 dF_{kn}(x) \rightarrow 0 \quad \text{as } n \rightarrow \infty.$$

Now we shall show that Lindeberg Theorem holds for triangular arrays.

Theorem 2. If (2) holds, then the distribution of $\frac{1}{C_n}(T_n - \eta_n)$ converges to the standard normal distribution.

Proof. We may assume without loss of generality that all $m_{kn} = 0$. For, otherwise we would replace X_{kn} by $X_{kn} - m_{kn}$, and this involves merely a change of notation.

Let $\varphi_k\left(\frac{t}{C_n}\right)$ be the characteristic function of $\frac{X_{kn}}{C_n}$, then we have to show that

$$(3) \quad \varphi_1\left(\frac{t}{C_n}\right) \cdots \varphi_n\left(\frac{t}{C_n}\right) \rightarrow e^{-t^2/2}.$$

For this purpose we investigate that (3) is equivalent to

$$\sum_{k=1}^n [\varphi_k\left(\frac{t}{C_n}\right) - 1] + \frac{1}{2} t^2 \rightarrow 0.$$

In fact, for any complex numbers such that $|a_k| \leq 1$ and $|b_k| \leq 1$ we have

$$|a_1 \cdots a_n - b_1 \cdots b_n| \leq \sum_{k=1}^n |a_k - b_k|.$$

$e^{\varphi_k \left(\frac{t}{C_n}\right) - 1}$ is the characteristic function of a compound poisson distribution, and for any $\delta > 0$ if $|z|$ is sufficiently small then $|e^z - 1 - z| < \delta |z|$.

From the above inequalities and the Taylor expansion for large n .

$$\begin{aligned} & \left| e^{\sum_{k=1}^n \left[\varphi_k \left(\frac{t}{C_n} \right) - 1 \right]} - \varphi_1 \left(\frac{t}{C_n} \right) \cdots \varphi_n \left(\frac{t}{C_n} \right) \right| \\ & \leq \sum_{k=1}^n \left| e^{\varphi_k \left(\frac{t}{C_n} \right) - 1} - \varphi_k \left(\frac{t}{C_n} \right) \right| \leq \delta \sum_{k=1}^n \left| \varphi_k \left(\frac{t}{C_n} \right) - 1 \right| \\ & \leq \delta \sum_{k=1}^n \frac{t^2 \sigma_{kn}^2}{2 C_n^2} < \delta t^2. \end{aligned}$$

Since δ is arbitrary the first term converges to zero and hence is equivalent. From the Taylor expansion.

$$\left| e^{ix \frac{t}{C_n}} - 1 - ix \frac{t}{C_n} + \frac{x^2 t^2}{2! C_n^2} \right| \leq \left| \frac{x^3 t^3}{3! C_n^3} \right| < \left| \frac{xt}{C_n} \right|^3$$

and the last term is dominated by $\varepsilon x^2 t^3 / C_n^2$ for $|x| < \varepsilon C_n$, and for $x^2 t^2 / C_n^2$ by $|x| > \varepsilon C_n$.

Thus we obtain

$$\begin{aligned} & \sum_{k=1}^n \left[\varphi_k \left(\frac{t}{C_n} \right) - 1 \right] + \frac{1}{2} t^2 = \sum_{k=1}^n \int_{-\infty}^{\infty} \left[e^{ix \frac{t}{C_n}} - 1 - ix \frac{t}{C_n} + \frac{x^2 t^2}{2! C_n^2} \right] \times dF_{kn}(x) \\ & \quad < \sum_{k=1}^n \int_{-\infty}^{\infty} \left| \frac{xt}{C_n} \right|^3 dF_{kn}(x) \\ & \quad = \sum_{k=1}^n \int_{|x| \leq \varepsilon C_n} \frac{\varepsilon t^3 x^2}{C_n^2} dF_{kn}(x) + \sum_{k=1}^n \int_{|x| > \varepsilon C_n} \frac{x^2 t^2}{C_n^2} dF_{kn}(x) = \varepsilon t^3. \end{aligned}$$

From (2) the last equality holds and since ε can be chosen arbitrarily small the first term converges to zero as $n \rightarrow \infty$. Consequently we have induced to the fact that (3) is true, our proof is complete.

Remark. If we set $X_{kn} = \frac{X_k - m_k}{\tau_n}$ then $T_n = \frac{z_n - \hat{\xi}_n}{\tau_n}$.

From this, we note that (2) reduces to (1) and Theorem 1 represents the special case of Theorem 2, and this Theorem 2 will be called the Central Limit Theorem for Triangular Arrays.

References

- [1] W. Feller, *An Introduction to Probability Theory and Its Application II* (1971), John Wiley & Sons, Inc., New York.
- [2] S.S. Wilks, *Mathematical Statistics*, John Wiley & Sons, Inc., New York (1962).
- [3] R.B. Ash, *Real Analysis and Probability*, Academic Press, Inc., New York (1972).