# 2. Law of Large Numbers and Central Limit Theorem

## 1. Law of Large Numbers (LLN)

Let $\{\xi_i\}$ be a sequence of random variables such that $\mathbf{E}\,\xi_i = 0$. Then under general regularity conditions

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \overset{a.s. \text{or}\, p}{\longrightarrow} 0$$

which is called *law of large numbers* (LLN). It is often referred to as *strong law of large numbers* (SLLN) or *weak law of large numbers* (WLLN), depending upon whether the mode of convergence is a.s. or in probability.

### Remarks

(a) If $\mathbf{E}\,\xi_i = \mu$, LLN can be applied to $\{\zeta_i\}$ with $\zeta_i = \xi_i - \mu$. This then yields

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \overset{a.s. \text{or}\, p}{\longrightarrow} \mu$$

For more general case with $\mathbf{E}\,\xi_i = \mu_i$, we have

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i - \frac{1}{n}\sum_{i=1}^{n}\mu_i \overset{a.s. \text{or}\, p}{\longrightarrow} 0$$

and if we assume further that $\frac{1}{n}\sum_{i=1}^{n}\mu_i \to \mu$, then $\frac{1}{n}\sum_{i=1}^{n}\xi_i \overset{a.s. \text{or}\, p}{\longrightarrow} \mu$.

(b) LLN implies that the randomness in the sample mean disappears, as the sample size goes to infinity. With more data, we know more about the sample space, and in the limit our information becomes complete.

(c) The more dependent are the underlying random variables, the less marginal information we get from an additional observation. Therefore, LLN is more difficult to hold for a dependent sequence of random variables. As an extreme case, LLN obviously does not hold for a random constant $\{\xi_i\}$ with $\xi_i = \xi$ for all $i$, since $\frac{1}{n}\sum_{i=1}^{n}\xi_i = \xi$ for all $n$.

We now prove some simple LLN's for i.i.d. random variables.

**Theorem 1** *Let* $\{\xi_i\}$ *be a sequence of* i.i.d. *random variables with* $\mathbf{E}\,\xi_i = 0$.

(a) *If* $\mathbf{E}\,\xi_i^2 < \infty$, *then*

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \quad \to_p \quad 0$$

(b) *If* $\mathbf{E}\,\xi_i^4 < \infty$, *then*

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \quad \to_{a.s.} \quad 0$$

**Proof** The WLLN follows easily from

$$\mathbf{E}\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)^2 = \frac{1}{n}\mathbf{E}\,\xi_i^2 \to 0$$

since $\xrightarrow{\mathcal{L}^2}$ implies $\to_p$.

For the SLLN, notice that

$$\begin{aligned}
\mathbf{E}\left(\sum_{i=1}^{n}\xi_i\right)^4 &= \sum_{i=1}^{n}\mathbf{E}\,\xi_i^4 + 6\sum_{i\neq j}\mathbf{E}\,\xi_i^2\mathbf{E}\,\xi_j^2 \\
&= n\mathbf{E}\,\xi_i^4 + 3n(n-1)\mathbf{E}\,\xi_i^2\mathbf{E}\,\xi_j^2 \\
&= O(n^2)
\end{aligned}$$

which implies

$$\mathbf{E}\sum_{n=1}^{\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)^4 < \infty$$

It follows that

$$\sum_{n=1}^{\infty}\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)^4 < \infty \ \text{ a.s.}$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i \quad \to_{a.s.} \quad 0$$

as was to be shown. ■

## Remarks

(a) The conditions in Theorem 1 are not necessary. They are conditions strong enough to make proofs easy.

(b) As an example, we consider the experiment of an infinite toss of a coin, and define a sequence of Bernoulli random variables $\{\xi_n\}$ by $\xi_i = 1$ if the $i$-th toss is a head. The LLN implies in this specific setting that the sample proportion converges in probability or a.s. to a number, which is perceived as the true probability of getting a head.

## 2. Central Limit Theorem (CLT)

Let $\{\xi_{ni}\}$ be a sequence of random variables such that

$$\mathbf{E}\,\xi_{ni} = 0 \quad \text{and} \quad \sum_{i=1}^{n} \mathbf{E}\,\xi_{ni}^2 = 1$$

Under regularity conditions, we have

$$\sum_i \xi_{ni} \ \to_d \ \mathbf{N}(0,1)$$

which is referred to as *central limit theorem* (CLT). It is known that CLT holds under the Liapounov condition

$$\sum_{i=1}^{n} \mathbf{E}\,|\xi_{ni}|^3 \ \to \ 0$$

or under the Lindeberg condition

$$\sum_{i=1}^{n} \mathbf{E}\,\xi_{ni}^2\,\mathrm{I}\{|\xi_{ni}| > \varepsilon\} \ \to \ 0$$

for any $\varepsilon > 0$.

**Remark** The Liapounov condition is easier to check, but stronger than the Lindeberg condition, as can be easily seen from

$$\sum_{i=1}^{n} \mathbf{E}\,\xi_{ni}^2\,\mathrm{I}\{|\xi_{ni}| > \varepsilon\} \ \le \ \frac{\sum_{i=1}^{n} \mathbf{E}\,|\xi_{ni}|^3}{\varepsilon}$$

We have the following CLT for an i.i.d. sequence of random variables.

**Corollary 2** *Let* $\{X_i\}$ *be* i.i.d. *sequence of random variables with mean zero and variance* $\sigma^2$*. Then*

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_i \quad \to_d \quad \mathbf{N}(0,\sigma^2)$$

**Proof** Define

$$\xi_{ni} = \frac{X_i}{\sigma\sqrt{n}}$$

It is easy to see $\sum_{i=1}^{n} \mathbf{E}\, \xi_{ni}^2 = 1$. Moreover, the Lindeberg condition is satisfied, because

$$\begin{aligned} \sum_{i=1}^{n} \mathbf{E}\, \xi_{ni}^2 \,\mathrm{I}\{|\xi_{ni}| > \varepsilon\} &= \frac{1}{\sigma^2 n} \sum_{i=1}^{n} \mathbf{E}\, X_i^2 \,\mathrm{I}\{|X_i| > \varepsilon\sigma\sqrt{n}\} \\ &= \frac{1}{\sigma^2} \mathbf{E}\, X_i^2 \,\mathrm{I}\{|X_i| > \varepsilon\sigma\sqrt{n}\} \\ &\to 0 \end{aligned}$$

for any $\varepsilon > 0$, by the dominated convergence. Therefore,

$$\sum_{i=1}^{n} \xi_{ni} = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{n} X_i \quad \to_d \quad \mathbf{N}(0,1)$$

from which the stated result follows directly. ∎