Laws of Large numbers

Consider a sequence of random variables $\{X_n\}$ with a common mean μ. It is common practice to determine μ on the basis of the sample mean defined by the relation

$$\frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{N} X_i$$

where $S_n = \sum_{i=1}^{N} X_i$. We assert that $\frac{S_n}{n} \to \mu$ as $n \to \infty$. Note that μ is a deterministic constant whereas $\frac{S_n}{n}$ is a function of n random variables. The laws of large numbers are the basis of such assertions.

More generally, suppose $\{X_n\}$ is a sequence of random variables with $\mu_i = EX_i,\ i = 1, 2, ..., n$.

Then,

$$E\frac{S_n}{n} = \frac{1}{n}\sum_{i=1}^{n} EX_i$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mu_i$$

The sequence $\{X_n\}$ is said to obey the *strong law of large numbers* if

$$\frac{S_n}{n} \xrightarrow{a.s.} \frac{1}{n}\sum_{i=1}^{n} \mu_i \ .$$

Similarly, $\{X_n\}$ is said to obey the *weak law of large numbers* if

$$\frac{S_n}{n} \xrightarrow{P.} \frac{1}{n}\sum_{i=1}^{n} \mu_i \ .$$

We consider weak law for a more general case and the strong law for the special case when $\{X_n\}$ is a sequence of iid random variables.

Theorem 1 Weak law of large numbers(WLLN): Suppose $\{X_n\}$ is a sequence of random variables defined on a probability space $(S, \mathbb{F}, P)$ with finite mean $\mu_i = EX_i,\ i = 1, 2, ..., n$ and finite second moments. If

$$\lim_{n\to\infty} \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1, j\neq i}^{n} cov(X_i, X_j) = 0,$$

then $\frac{S_n}{n} \xrightarrow{P.} \frac{1}{n}\sum_{i=1}^{n} \mu_i$.

Proof: We have

$$E\left(\frac{S_n}{n}-\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)^2 = E\left(\frac{1}{n}\sum_{i=1}^{n}X_i-\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)^2$$

$$=\frac{1}{n^2}E\left(\sum_{i=1}^{n}(X_i-\mu_i)\right)^2$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}E(X_i-\mu_i)^2+\frac{1}{n^2}\sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{j=1,j\neq i}}^{\mathrm{n}}E(X_i-\mu_i)(X_j-\mu_j)$$

$$=\frac{1}{n^2}\sum_{i=1}^{n}\sigma_{i_i}^2+\frac{1}{n^2}\sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{j=1,j\neq i}}^{\mathrm{n}}cov(X_i,X_j)$$

$$\therefore \lim_{n\to\infty}E\left(\frac{S_n}{n}-\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)^2=\lim_{n\to\infty}\left(\frac{1}{n^2}\sum_{i=1}^{n}\sigma_{i_i}^2+\frac{1}{n^2}\sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{j=1,j\neq i}}^{\mathrm{n}}cov(X_i,X_j)\right)$$

Now $\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sigma_{i_i}^2=0$, as each $\sigma_{i_i}^2$ is finite. Also,

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{j=1,j\neq i}}^{\mathrm{n}}cov(X_i,X_j)=0$$

$$\lim_{n\to\infty}E\left(\frac{S_n}{n}-\frac{1}{n}\sum_{i=1}^{n}\mu_i\right)^2=0$$

$$\Rightarrow \frac{S_n}{n}\xrightarrow{m.s..}\frac{1}{n}\sum_{i=1}^{n}\mu_i$$

$$\Rightarrow \frac{S_n}{n}\xrightarrow{P.}\frac{1}{n}\sum_{i=1}^{n}\mu_i$$

Special Case of the WLLN

(a) Suppose $\{X_n\}$ is a sequence of independent and identically distributed random variables defined on a probability space $(S,\mathbb{F},P)$

Then we have

$EX_i=\text{constant}=\mu(say)$

$\text{var}(X_i)=\text{constant}=\sigma^2(say)$ *and*

$\text{cov}(X_i,X_j)=0$

$$\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sigma_{i_i}^2=\lim_{n\to\infty}\frac{1}{n^2}\sum_{i=1}^{n}\sigma^2=0$$

$$\therefore \frac{S_n}{n}\xrightarrow{P.}\mu$$

(b) Suppose $\{X_n\}$ is a sequence of independent random variables defined on a probability space $(S,\mathbb{F},P)$ with the mean $\mu_i=EX_i,\ i=1,2,\ldots,n$ and finite second moments .

Then we have

X_i and X_j are independent

$\therefore \operatorname{cov}(X_i, X_j) = 0$

Again

each $\sigma_{i_i}^2$ is finite

$\therefore \lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} \sigma_{i_i}^2 = 0$

$\therefore \frac{S_n}{n} \xrightarrow{P.} \frac{1}{n} \sum_{i=1}^{n} \mu_i$

(c) Suppose $\{X_n\}$ is a sequence of uncorrelated random variables with the mean $\mu_i = EX_i$, $i = 1, 2, \ldots, n$ and finite second moments defined on a probability space $(S, \mathbb{F}, P)$.

Then we have $\operatorname{cov}(X_i, X_j) = 0$ by definition.

We can proceed as in case (b) to establish the result

$\frac{S_n}{n} \xrightarrow{P.} \frac{1}{n} \sum_{i=1}^{n} \mu_i$

The strong law of large number is based on the concept of almost sure convergence and stated in the following theorem:

Interpretation of relative frequency definition of probability

The relative frequency definition of probability can be interpreted using the *weak law of large number*. Suppose an experiment is repeated n times and a particular event A occurs n_A times. During these repetitions, then,

$$\frac{n_A}{n} \xrightarrow{P} P(A)$$

which is the interpretation .

To prove the above result consider a sequence of random variables $\{I_{A_n}\}$ given by

$$I_{A_n} = 1 \quad \text{if A occurs in } n^{th} \text{ trial}$$
$$= 0 \quad \text{otherwise}$$

Then,

$$EI_{A_n} = 1 \times P(A) + 0 \times (1 - P(A))$$
$$= P(A)$$

Now,

$$\frac{1}{n}\sum_{i=1}^{n} I_{A_i} = \frac{\text{Number of occurences of } A}{n}$$
$$= \frac{n_A}{n}$$

So, using the *weak law of large number*,

$$\frac{n_A}{n} \xrightarrow{P} P(A)$$

Strong Law of Large Numbers

One of the important applications of the a.s. convergence is the *Strong Law of Large Numbers*. The Kolmogorov's strong law of large numbers is stated in the following theorem.

Theorm 1: Suppose $\{X_n\}$ is a sequence of iid random variables defined on a probability space $(S, \mathbb{F}, P)$ with common mean μ and $S_n = \sum_{i=1}^{n} X_i$. Then

$$\frac{S_n}{n} \xrightarrow{a.s} \mu$$

Though the theorem is general, we will prove the following restricted version:

Theorem 2: Suppose $\{X_n\}_{n=1}^{\infty}$ is a sequence of iid random variables defined on a probability space $(S, \mathbb{F}, P)$ with common mean μ and finite fourth central moment ($E(X_n - EX_n)^4 < \infty$). Then

$$\frac{S_n}{n} \xrightarrow{a.s} \mu$$

Proof: We have to prove that

$$P\left(\limsup_{n\to\infty}\left\{s\left|\left|\frac{S_n}{n} - \mu\right| \geq \frac{1}{m}\right.\right\}\right) = 0 \qquad \forall \text{ positive integer } m.$$

Let us examine the fourth moment

$$E\left(\frac{S_n}{n} - \mu\right)^4$$

$$= E\left(\frac{\sum_{i=1}^{n}(X_i - \mu)}{n}\right)^4$$

$$= E\left(\frac{\sum_{i=1}^{n} Y_i}{n}\right)^4$$

where $Y_i = X_i - \mu$

In the expansion of $\left(\sum_{i=1}^{n} Y_i\right)^4$, there will be terms of the form $Y_i^4, Y_i^3Y_j, Y_i^2Y_j^2, Y_iY_jY_k^2, Y_iY_jY_kY_l$.

We note that $EY_i^3Y_j = EY_iY_jY_k^2 = EY_iY_jY_kY_l = 0$ as Y_i is of zero mean and the sequence is independent. Therefore, the term of the form EY_i^4 and $EY_i^2EY_j^2$ contribute to the fourth central moment. There are n terms of the form EY_i^4 and ${}^nC_2 \times {}^4C_2 = \frac{n(n-1)}{2} \times 6 = 3n(n-1)$ terms of the form $EY_i^2EY_j^2$.

$$\therefore E\left(\frac{\sum_{i=1}^{n} Y_i}{n}\right)^4 = \frac{1}{n^4}\left[nEY_i^4 + {}^nC_2 \times {}^4C_2\left(EY_i\right)^2\right]$$

$$= \frac{EY_i^4}{n^3} + \frac{3n(n-1)}{n^4}\left(EY_i^2\right)^2$$

$$\le \frac{K}{n^3} + \frac{3}{n^2}K$$

(Assume $EY_i^4 = K < \infty$ and $(EY_i^2)^2 \le EY_i^4 = K$)

$$\therefore E\left(\frac{\sum_{i=1}^{n}\left(X_i - \mu\right)}{n}\right)^4 \le \frac{K}{n^3} + \frac{3K}{n^2}$$

Now $P\left(\left\{s \middle| \left|\frac{1}{n}\sum_{i=1}^{n} X_i(s) - \mu\right| \ge \frac{1}{m}\right\}\right)$

$$= P\left(\left\{s \middle| \left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i(s) - \mu\right)\right)^4 \ge \frac{1}{m^4}\right\}\right)$$

$$\le \frac{E\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \mu\right)\right)^4}{\frac{1}{m^4}} \quad \text{(By Markov inequality)}$$

$$\le \frac{m^4 K}{n^3} + \frac{3m^4}{n^2}K$$

$$\therefore \sum_{n=1}^{\infty} P\left(\left\{s \middle| \left|\frac{1}{n}\sum_{i=1}^{n} X_i(s) - \mu\right| \ge \frac{1}{m}\right\}\right)$$

$$\le \sum_{n=1}^{\infty} \frac{m^4 K}{n^3} + \frac{3m^4}{n^2}K$$

$$< \infty$$

Hence according to the Borel Cantelli Lemma,

$$P\left(\limsup_{n\to\infty}\left\{s \middle| \left|\sum_{i=1}^{\infty}\left(X_i(s) - \mu\right)\right| \ge \frac{1}{m}\right\}\right) = 0$$

Thus, $\frac{S_n}{n} \xrightarrow{a.s} \mu$

Central Limit Theorem

Suppose $\{X_n\}$ is a sequence of independent and identically distributed random variables each with mean μ and variance σ^2 and $S_n = \sum_{i=1}^{n} X_i$ By the weak law of large numbers, $\frac{S_n}{n} \xrightarrow{P} \mu$. Note that the convergence in probability implies the convergence in distribution. Therefore,

$$\frac{S_n}{n} \xrightarrow{d} \mu$$

From the WLLN, we may conclude that for large *n*, $S_n \simeq n\mu$. The *central limit theorem* (CLT) gives the asymptotic distribution of the difference $S_n - n\mu$. The CLT is stated in terms of the standardized average $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$. Clearly, $\mathrm{E}Z_n = 0$ and $\mathrm{Var}(Z_n) = 1$

There are several special cases of the CLT. Here we state and prove the celebrated Lindeberg – Levy central limit theorem.

Theorem: Suppose $\{X_n\}$ is a sequence of i.i.d. random variables with mean μ and variance $\sigma^2 < \infty$. Let $S_n = \sum_{i=1}^{n} X_i$ and $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$. Then $Z_n \xrightarrow{d} Z \sim N(0,1)$ in the sense that

$$\lim_{n\to\infty} F_{Z_n}(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du$$

Proof: We shall prove the theorem using the moment generating function (MGF) of Z_n and the continuity theorem of convergence.

Suppose the MGF of each RV in the sequence $\{Z_n\}_{n=1}^{\infty}$ and the RV Z exist near $s = 0$. According to the continuity theorem of convergence $\lim_{n\to\infty} F_{Z_n}(z) = F_Z(z)$ if and only if $\lim_{n\to\infty} M_{Z_n}(s) = M_Z(s)$ and $M_Z(s)$ is continuous at $s = 0$.

Here

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

$$\therefore M_{Z_n}(s) = \mathrm{E}e^{Z_n s}$$

$$= Ee^{\sum_{i=1}^{n}\frac{(X_i-\mu)s}{\sigma\sqrt{n}}}$$

$$= \prod_{i=1}^{n} E\,\frac{(X_i-\mu)s}{\sigma\sqrt{n}}$$

$$= \prod_{i=1}^{n} M_{Y_i}\left(\frac{s}{\sqrt{n}}\right)$$

$$= \left(M_{Y_i}\left(\frac{s}{\sqrt{n}}\right)\right)^n$$

The Taylor series expansion of $M_{Y_i}(s)$ is given by

$$M_{Y_i}(s) = \sum_{k=0}^{\infty}\frac{EY_i^k s^k}{k!}$$

Note that $EY_i = 0$ and $EY_i^2 = 1$. Therefore $M_{Y_i}(s)$ near $s = 0$ can be expressed as

$$M_{Y_i}(s) = 1 + \frac{s^2}{2} + o(s^2)$$

$$\therefore M_{Z_n}(s) = \left(M_{Y_i}(\frac{s}{\sqrt{n}})\right)^n$$

$$= \left(1 + \frac{s^2}{2n} + o(\frac{s^2}{n})\right)^n$$

$$\lim_{n\to\infty} M_{Z_n}(s) = \lim_{n\to\infty}\left(1 + \frac{s^2}{2n} + o(\frac{s^2}{n})\right)^n$$

$$= e^{\frac{s^2}{2}}$$

which is the MGF of $X \sim N(0,1)$

Hence, moment generating function converges.

Applying continuity theorem ,

$F_{X_n}(x) \to F_X(x)$

The CLT is true under more general conditions. The i.i.d. part of the Lindeberg – Levy theorem need not be satisfied. We state two of these conditions without proof:

- **Liapounov theorem:** Suppose $\{X_n\}_{n=1}^{\infty}$ is a sequence of independent random variables with mean $\mu_n = EX_n$ and variance $\sigma_n^2 = E(X_n - \mu_n)^2$ and $S_n = \sum_{i=1}^{n} X_i$.

  Clearly $\mu_{S_n} = \sum_{i=1}^{n}\mu_i$ and $\sigma_{S_n}^2 = \sum_{i=1}^{n}\sigma_i^2$. If for some $\delta > 0$,

$$\lim_{n\to\infty} \frac{\sum_{k=1}^{n} E(X_k-\mu_k)^{2+\delta}}{(\sigma_{S_n})^{2+\delta}} = 0,$$

then $\dfrac{S_n-\mu_{S_n}}{\sigma_{S_n}} \xrightarrow{d} Z \sim N(0,1)$

- A corollary of the Liapounov theorem is as follows:

  Suppose each of X_ns are uniformly bounded, that is, $|X_n| < C, \forall n$ and $\lim_{n\to\infty} \sigma_{S_n}^2 \to \infty$. Then it is easy to verify that the conditions of the Liapounov theorem are satisfied. Thus $\dfrac{S_n-\mu_{S_n}}{\sigma_{S_n}} \xrightarrow{d} Z \sim N(0,1)$.

Remark

- The CLT states that the distribution function $F_{Z_n}(z)$ converges to a Gaussian distribution function. The theorem does not say that the pdf $f_{Z_n}(z)$ is a Gaussian pdf in the limit. For example, suppose each X_i has a Bernoulli distribution. Then the pdf of Z_n consists of impulses and can never approach the Gaussian pdf.

- The Cauchy distribution does not meet the conditions for the central limit theorem to hold. As we have noted earlier, this distribution does not have a finite mean or a variance.