भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Stochastic Processes

Interpretation of CLT, What CLT is Not, Local CLT

Karthik P. N.

Assistant Professor, Department of AI

Email: pnkarthik@ai.iith.ac.in

14 February 2025

Weak Law of Large Numbers (WLLN)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{X_n\}_{n=1}^{\infty}$ be defined w.r.t. $\mathscr{F}$.

Theorem (Weak Law of Large Numbers)

Let $\{X_n\}_{n=1}^{\infty}$ be i.i.d. with $\mathbb{E}[|X_1|] < +\infty$. Further, let $\mathbb{E}[X_1] = \mu$. Let

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

Then,

$$\frac{S_n}{n} \xrightarrow{\text{p.}} \mu.$$

More formally, for every $\varepsilon > 0$,

$$\lim_{n \to \infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = 0.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Strong Law of Large Numbers (SLLN)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{X_n\}_{n=1}^{\infty}$ be defined w.r.t. $\mathscr{F}$.

Theorem (Strong Law of Large Numbers)

Let $\{X_n\}_{n=1}^{\infty}$ be i.i.d. with $\mathbb{E}[|X_1|] < +\infty$. Further, let $\mathbb{E}[X_1] = \mu$. Let

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

Then,

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

More formally,

$$\mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Central Limit Theorem (CLT)

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Let $\{X_n\}_{n=1}^{\infty}$ be defined w.r.t. $\mathscr{F}$.

Theorem (Central Limit Theorem)

Let $\{X_n\}_{n=1}^{\infty}$ be i.i.d. with mean $\mathbb{E}[X_1] = \mu \in \mathbb{R}$ and $\mathrm{Var}(X_1) = \sigma^2 < +\infty$. Let $S_n = \sum_{i=1}^{n} X_i$. Then,

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}} = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{\mathrm{d}} X, \qquad X \sim \mathcal{N}(0,1).$$

More formally,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\mathrm{Var}(S_n)}} \leq x\right) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, \mathrm{d}t \qquad \forall x \in \mathbb{R}.$$

Proof of CLT

Let

$$Z_i = \frac{X_i - \mathbb{E}[X_i]}{\sqrt{\mathrm{Var}(X_1)}}, \qquad U_n = \frac{\sum_{i=1}^{n} Z_i}{\sqrt{n}}$$

Proof of CLT

Let

$$Z_i = \frac{X_i - \mathbb{E}[X_i]}{\sqrt{\mathrm{Var}(X_1)}}, \qquad U_n = \frac{\sum_{i=1}^{n} Z_i}{\sqrt{n}}$$

$$C_{Z_i}(s) \quad =$$

Proof of CLT

Let

$$Z_i = \frac{X_i - \mathbb{E}[X_i]}{\sqrt{\mathrm{Var}(X_1)}}, \qquad U_n = \frac{\sum_{i=1}^{n} Z_i}{\sqrt{n}}$$

$$c_{Z_i}(s) \quad = \quad 1 - \frac{s^2}{2} + o(s^2), \qquad s \in \mathbb{R},$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Proof of CLT

Let

$$Z_i = \frac{X_i - \mathbb{E}[X_i]}{\sqrt{\mathrm{Var}(X_1)}}, \qquad U_n = \frac{\sum_{i=1}^{n} Z_i}{\sqrt{n}}$$

$$c_{Z_i}(s) = 1 - \frac{s^2}{2} + o(s^2), \qquad s \in \mathbb{R},$$

$$c_{U_n}(s)$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Proof of CLT

Let

$$Z_i = \frac{X_i - \mathbb{E}[X_i]}{\sqrt{\mathrm{Var}(X_1)}}, \qquad U_n = \frac{\sum_{i=1}^{n} Z_i}{\sqrt{n}}$$

$$\begin{aligned} C_{Z_i}(s) &= 1 - \frac{s^2}{2} + o(s^2), \qquad s \in \mathbb{R}, \\ C_{U_n}(s) &= \left(C_{Z_1}\left(\frac{s}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{s^2}{2n} + o\left(\frac{s^2}{n}\right)\right)^n \end{aligned}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Proof of CLT

Let

$$Z_i = \frac{X_i - \mathbb{E}[X_i]}{\sqrt{\mathrm{Var}(X_1)}}, \qquad U_n = \frac{\sum_{i=1}^{n} Z_i}{\sqrt{n}}$$

$$\begin{aligned} C_{Z_i}(s) &= 1 - \frac{s^2}{2} + o(s^2), \qquad s \in \mathbb{R}, \\ C_{U_n}(s) &= \left(C_{Z_1}\left(\frac{s}{\sqrt{n}}\right)\right)^n = \left(1 - \frac{s^2}{2n} + o\left(\frac{s^2}{n}\right)\right)^n \end{aligned}$$

$$\begin{aligned} \lim_{n\to\infty} C_{U_n}(s) &= \lim_{n\to\infty} \left(1 - \frac{s^2}{2n} + o\left(\frac{s^2}{n}\right)\right)^n \\ &= e^{\frac{-s^2}{2}} \end{aligned}$$

Some Notes about CLT

- From SLLN, we know that

$$\frac{S_n - \mathbb{E}[S_n]}{n} \xrightarrow{\text{a.s.}} 0.$$

Some Notes about CLT

- From SLLN, we know that

$$\frac{S_n - \mathbb{E}[S_n]}{n} \xrightarrow{\text{a.s.}} 0.$$

- Thus, the distribution of $S_n - \mathbb{E}[S_n]$, when divided by n, is degenerate for large n

Some Notes about CLT

- From SLLN, we know that
$$\frac{S_n - \mathbb{E}[S_n]}{n} \xrightarrow{\text{a.s.}} 0.$$
- Thus, the distribution of $S_n - \mathbb{E}[S_n]$, when divided by n, is degenerate for large n
- CLT: the distribution of $S_n - \mathbb{E}[S_n]$, when divided by $\sqrt{n}$, is non-degenerate for large n

Some Notes about CLT

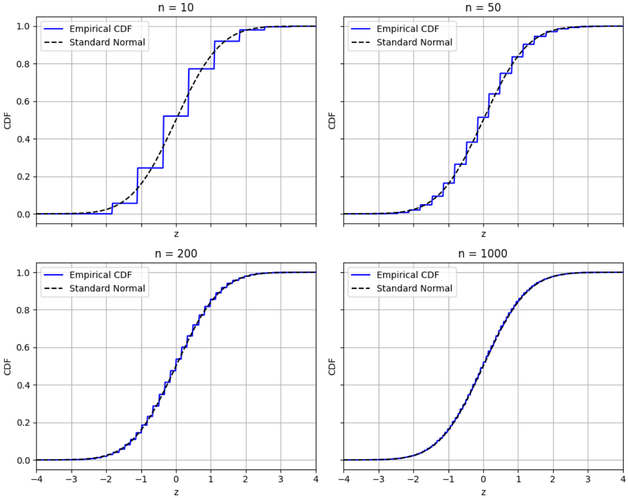
- From SLLN, we know that

$$\frac{S_n - \mathbb{E}[S_n]}{n} \xrightarrow{\text{a.s.}} 0.$$

- Thus, the distribution of $S_n - \mathbb{E}[S_n]$, when divided by n, is degenerate for large n
- CLT: the distribution of $S_n - \mathbb{E}[S_n]$, when divided by $\sqrt{n}$, is non-degenerate for large n
- According to CLT, for large n,

$$\mathbb{P}\left(\frac{S_n - \mathbb{E}[S_n]}{\sigma\sqrt{n}} > t\right) \approx \mathbb{P}(X > t), \qquad X \sim \mathcal{N}(0, 1)$$

Demonstration of CLT



What CLT is Not

- CLT is not a statement about convergence of the PDFs of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ to Gaussian PDF

What CLT is Not

- CLT is not a statement about convergence of the PDFs of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ to Gaussian PDF
- If $X_1, X_2, \ldots$ are discrete random variables, then $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ is a discrete random variable, and hence does not admit any PDF.

What CLT is Not

- CLT is not a statement about convergence of the PDFs of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ to Gaussian PDF
- If $X_1, X_2, \ldots$ are discrete random variables, then $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ is a discrete random variable, and hence does not admit any PDF.
- Even if $X_1, X_2, \ldots$ are continuous random variables, and $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ admits a PDF, CLT does not make any claim about the convergence of these PDFs to the Gaussian PDF

Example

Suppose that $X_1, X_2, \cdots \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}([-\sqrt{3},\ +\sqrt{3}])$

- $\mathbb{E}[X_1] =$

Example

Suppose that $X_1, X_2, \cdots \overset{\text{i.i.d.}}{\sim} \text{Unif}([-\sqrt{3},\ +\sqrt{3}])$

- $\mathbb{E}[X_1] = 0$
- $\text{Var}(X_1) =$

Example

Suppose that $X_1, X_2, \cdots \overset{\text{i.i.d.}}{\sim} \text{Unif}([-\sqrt{3},\ +\sqrt{3}])$

- $\mathbb{E}[X_1] = 0$
- $\text{Var}(X_1) = 1$
- $C_{X_1}(s) =$

Example

Suppose that $X_1, X_2, \cdots \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}([-\sqrt{3},\ +\sqrt{3}])$

- $\mathbb{E}[X_1] = 0$
- $\mathrm{Var}(X_1) = 1$
- $C_{X_1}(s) = \dfrac{\sin(s\sqrt{3})}{s\sqrt{3}}, \quad s \in \mathbb{R}.$

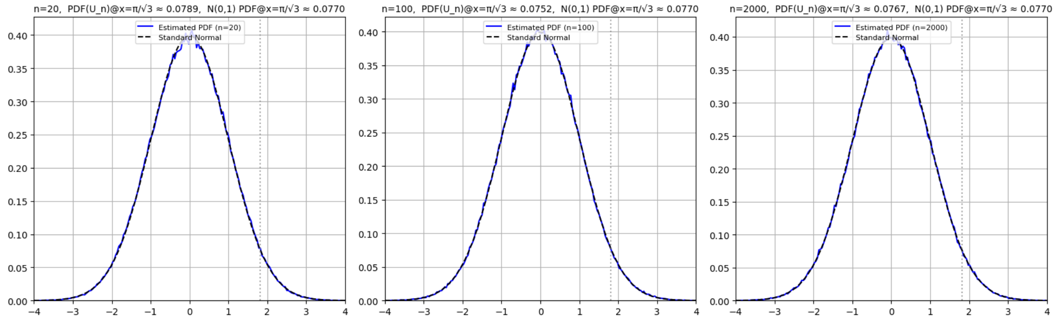
Example

Suppose that $X_1, X_2, \cdots \overset{\text{i.i.d.}}{\sim} \mathrm{Unif}([-\sqrt{3},\ +\sqrt{3}])$

- $\mathbb{E}[X_1] = 0$
- $\mathrm{Var}(X_1) = 1$
- $C_{X_1}(s) = \dfrac{\sin(s\sqrt{3})}{s\sqrt{3}}, \quad s \in \mathbb{R}.$
- As per CLT,

$$C_{\frac{S_n}{\sqrt{n}}}(s) \xrightarrow{n \to \infty} e^{-\frac{s^2}{2}} \quad \forall s \in \mathbb{R}.$$

No Convergence of PDFs for Previous Example



Local CLT

- In many practical examples where $X_1, X_2, ...$ are continuous, one may observe convergence of PDFs of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ to the Gaussian PDF, but this is NOT to be interpreted as a consequence of the CLT
  This may be a consequence of some stronger property playing in hindsight

Local CLT

- In many practical examples where $X_1, X_2, \ldots$ are continuous, one may observe convergence of PDFs of $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ to the Gaussian PDF, but this is NOT to be interpreted as a consequence of the CLT
  This may be a consequence of some stronger property playing in hindsight

Theorem (Local Central Limit Theorem)

Suppose that $X_1, X_2, \cdots \overset{\text{i.i.d.}}{\sim} f_X$. W.l.o.g., let $\mathbb{E}[X_1] = 0$ and $\mathrm{Var}(X_1) = 1$.
Suppose that there exists $r \in \mathbb{N}$ such that

$$\int_{-\infty}^{\infty} |C_{X_1}(s)|^r \, \mathrm{d}s < +\infty.$$

Then,

$$f_{\frac{S_n}{\sqrt{n}}}(x) \xrightarrow{n\to\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad \forall x \in \mathbb{R}.$$