भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Probability and Stochastic Processes

Convergence of Sequences of Random Variables, Limit Theorems

Karthik P. N.
Assistant Professor, Department of AI
Email: pnkarthik@ai.iith.ac.in

25 November 2024

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Convergence of Sequences of Random Variables

Objective

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables defined w.r.t. $\mathscr{F}$.
Let X be another random variable defined w.r.t. $\mathscr{F}$. We allow X to take $\pm\infty$.

Objective

To define the following forms of convergence.

1. Pointwise convergence; notation: $X_n \xrightarrow{\text{pointwise}} X$.
2. Almost-sure convergence; notation: $X_n \xrightarrow{\text{a.s.}} X$.
3. Mean-squared convergence; notation: $X_n \xrightarrow{\text{m.s.}} X$.
4. Convergence in probability; notation: $X_n \xrightarrow{\text{p}} X$.
5. Convergence in distribution; notation: $X_n \xrightarrow{\text{d}} X$.

Pointwise Convergence

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables defined w.r.t. $\mathscr{F}$.
Let X be another random variable defined w.r.t. $\mathscr{F}$. We allow X to take $\pm\infty$.

Definition (Pointwise Convergence)

We say that the sequence $\{X_n\}_{n=1}^{\infty}$ converges to X pointwise if

$$\forall \omega \in \Omega, \qquad \lim_{n\to\infty} X_n(\omega) = X(\omega).$$

Notation:

$$X_n \xrightarrow{\text{pointwise}} X.$$

Example

$\Omega = [0, 1], \mathscr{F} = \mathscr{B}([0, 1]), \mathbb{P} = \lambda.$
For each $n \in \mathbb{N}$, let

$$X_n(\omega) = \begin{cases} 1, & \omega \in \left[0, \frac{1}{n}\right), \\ 0, & \text{otherwise.} \end{cases}$$

Identify the limit random variable X to which the above sequence converges pointwise.

Convergence in Distribution

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables defined w.r.t. $\mathscr{F}$.
Let X be another random variable defined w.r.t. $\mathscr{F}$. We allow X to take $\pm\infty$.

Definition (Convergence in Distribution)

We say that the sequence $\{X_n\}_{n=1}^{\infty}$ converges to X in distribution if

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \qquad \forall x \in C_{F_X},$$

where C_{F_X} denotes the points of continuity of F_X.

Notation:

$$X_n \xrightarrow{\mathrm{d}} X.$$

Example

Let $X_n = U$ for all $n \in \mathbb{N}$, with $U \sim \mathrm{Unif}([0,1])$.
Let $X = 1 - U$.
Show that does not converge to X pointwise, but $X_n \xrightarrow{\mathrm{d}} X$.

Convergence in Probability

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables defined w.r.t. $\mathscr{F}$.
Let X be another random variable defined w.r.t. $\mathscr{F}$. We allow X to take $\pm\infty$.

Definition (Convergence in Probability)

We say that the sequence $\{X_n\}_{n=1}^{\infty}$ converges to X in probability if

$$\forall \varepsilon > 0, \qquad \lim_{n\to\infty} \mathbb{P}(\{|X_n - X| > \varepsilon\}) = 0.$$

Notation:

$$X_n \xrightarrow{\mathrm{p}} X.$$

Mean-Squared Convergence

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables defined w.r.t. $\mathscr{F}$.
Let X be another random variable defined w.r.t. $\mathscr{F}$. We allow X to take $\pm\infty$.

Definition (Mean-Squared Convergence)

We say that the sequence $\{X_n\}_{n=1}^{\infty}$ converges to X in mean-squared sense if

$$\lim_{n\to\infty} \mathbb{E}[(X_n - X)^2] = 0.$$

Notation:

$$X_n \xrightarrow{\text{m.s.}} X.$$

Almost-Sure Convergence

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.
Let $\{X_n\}_{n=1}^{\infty}$ be a sequence of random variables defined w.r.t. $\mathscr{F}$.
Let X be another random variable defined w.r.t. $\mathscr{F}$. We allow X to take $\pm\infty$.

Definition (Almost-Sure Convergence)

We say that the sequence $\{X_n\}_{n=1}^{\infty}$ converges to X almost surely if

$$\mathbb{P}\left(\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$$

Notation:

$$X_n \xrightarrow{\text{a.s.}} X.$$

Question:

How do we know that $\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} \in \mathscr{F}$?

Recap of Limits

Recall the following definition of limits of sequences of real numbers:
Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of real numbers.

Limit of a Sequence of Real Numbers

We say that $x \in \mathbb{R} \cup \{\pm\infty\}$ is the limit of the sequence $\{x_n\}_{n=1}^{\infty}$ if

$$\forall \varepsilon > 0, \quad \exists N_\varepsilon \in \mathbb{N} \quad \text{such that} \quad |x_n - x| < \varepsilon \quad \forall n \geq N_\varepsilon.$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Recap of Limits

Recall the following definition of limits of sequences of real numbers:
Let $\{x_n\}_{n=1}^{\infty}$ be a sequence of real numbers.

Limit of a Sequence of Real Numbers

We say that $x \in \mathbb{R} \cup \{\pm\infty\}$ is the limit of the sequence $\{x_n\}_{n=1}^{\infty}$ if

$$\forall \varepsilon > 0, \quad \exists N_\varepsilon \in \mathbb{N} \quad \text{such that} \quad |x_n - x| < \varepsilon \quad \forall n \geq N_\varepsilon.$$

Equivalently,

$$\forall q \in \mathbb{Q},\ q > 0, \quad \exists N_q \in \mathbb{N} \quad \text{such that} \quad |x_n - x| < q \quad \forall n \geq N_q.$$

Limits of Sequences of Random Variables

Limit of a Sequence of Random Variables

Given $\omega \in \Omega$, we say that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ if

$$\forall q \in \mathbb{Q},\ q > 0, \quad \exists N_q(\omega) \in \mathbb{N} \quad \text{such that} \quad |X_n(\omega) - X(\omega)| < q \quad \forall n \geq N_q(\omega).$$

Limits of Sequences of Random Variables

Limit of a Sequence of Random Variables

Given $\omega \in \Omega$, we say that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ if

$$\forall q \in \mathbb{Q},\ q > 0, \quad \exists N_q(\omega) \in \mathbb{N} \quad \text{such that} \quad |X_n(\omega) - X(\omega)| < q \quad \forall n \geq N_q(\omega).$$

$$|X_n(\omega) - X(\omega)| < q \quad \forall n \geq N_q(\omega) \iff \omega \in \bigcap_{n=N_q(\omega)}^{\infty} \{\omega' \in \Omega : |X_n(\omega') - X(\omega')| < q\}$$

Limits of Sequences of Random Variables

Limit of a Sequence of Random Variables

Given $\omega \in \Omega$, we say that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ if

$$\forall q \in \mathbb{Q},\ q > 0, \quad \exists N_q(\omega) \in \mathbb{N} \quad \text{such that} \quad |X_n(\omega) - X(\omega)| < q \quad \forall n \geq N_q(\omega).$$

$$|X_n(\omega) - X(\omega)| < q \quad \forall n \geq N_q(\omega) \iff \omega \in \bigcap_{n=N_q(\omega)}^{\infty} \{\omega' \in \Omega : |X_n(\omega') - X(\omega')| < q\}$$

$$\exists N_q \in \mathbb{N} \qquad \iff \qquad \omega \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{\omega' \in \Omega : |X_n(\omega') - X(\omega')| < q\}$$

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Limits of Sequences of Random Variables

Limit of a Sequence of Random Variables

Given $\omega \in \Omega$, we say that $\lim_{n\to\infty} X_n(\omega) = X(\omega)$ if

$$\forall q \in \mathbb{Q},\ q > 0, \quad \exists N_q(\omega) \in \mathbb{N} \quad \text{such that} \quad |X_n(\omega) - X(\omega)| < q \quad \forall n \geq N_q(\omega).$$

$$|X_n(\omega) - X(\omega)| < q \quad \forall n \geq N_q(\omega) \iff \omega \in \bigcap_{n=N_q(\omega)}^{\infty} \{\omega' \in \Omega : |X_n(\omega') - X(\omega')| < q\}$$

$$\exists N_q \in \mathbb{N} \quad \iff \quad \omega \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{\omega' \in \Omega : |X_n(\omega') - X(\omega')| < q\}$$

$$\forall q \in \mathbb{Q}, q > 0 \quad \iff \quad \omega \in \bigcap_{\substack{q\in\mathbb{Q}:\\ q>0}} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{\omega' \in \Omega : |X_n(\omega') - X(\omega')| < q\}$$

Limits of Sequences of Random Variables

In summary, we have

$$\left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\} \quad = \quad \bigcap_{\substack{q\in\mathbb{Q}:\\ q>0}} \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} \{\omega' \in \Omega : |X_n(\omega') - X(\omega')| < q\}.$$

A Figure to Remember

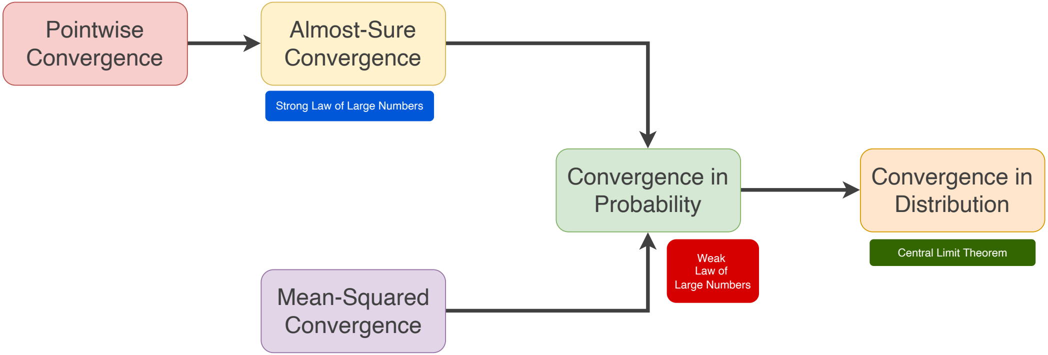

Figure: Implications of various forms of convergence.

Examples

- $\Omega = [0, 1], \mathscr{F} = \mathscr{B}([0, 1]), \mathbb{P} = \lambda.$
  For each $n \in \mathbb{N}$, let
  $$X_n(\omega) = \begin{cases} 1, & \omega \in \left[0, \frac{1}{n}\right), \\ 0, & \text{otherwise.} \end{cases}$$
  Identify the limits RV X and forms of convergence.

Examples

- $\Omega = [0,1], \mathscr{F} = \mathscr{B}([0,1]), \mathbb{P} = \lambda.$
  $X_1 = \mathbf{1}_{[0,1]}$
  $X_2 = \mathbf{1}_{[0,\frac{1}{2}]}, \quad X_3 = \mathbf{1}_{[\frac{1}{2},1]}$
  $X_4 = \mathbf{1}_{[0,\frac{1}{4}]}, \quad X_5 = \mathbf{1}_{[\frac{1}{4},\frac{1}{2}]}, \quad X_6 = \mathbf{1}_{[\frac{1}{2},\frac{3}{4}]}, \quad X_7 = \mathbf{1}_{[\frac{3}{4},1]}$
  and so on.
  Identify the limit and forms of convergence.

Examples

- $X_n = \mathcal{N}\left(0, \frac{1}{n}\right)$ for each $n \in \mathbb{N}$.
  Identify the limit and the forms of convergence.

Examples

- $\mathbb{P}(\{X_n = 1\}) = \frac{1}{n^2} = 1 - \mathbb{P}(\{X_n = 0\})$ for all $n \in \mathbb{N}$. Identify the limit and the forms of convergence.

Limit Theorems

Limit Theorems

Objective

To put down the formal statements of

1. Laws of large numbers:
   1.1 Weak law of large numbers (an instance of convergence in probability).
   1.2 Strong law of large numbers (an instance of almost-sure convergence).
2. Central limit theorem (an instance of convergence in distribution).

Weak Law of Large Numbers

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Theorem (Weak Law of Large Numbers)

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. RVs with common, finite expectation $\mathbb{E}[X_1] = \mu$. Let

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

Then,

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

i.e.,

$$\forall \varepsilon > 0, \qquad \lim_{n\to\infty} \mathbb{P}\left(\left\{\omega \in \Omega : \left|\frac{S_n}{n} - \mu\right| > \varepsilon\right\}\right) = 0.$$

Strong Law of Large Numbers

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Theorem (Strong Law of Large Numbers)

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. RVs with common, finite expectation $\mathbb{E}[X_1] = \mu$. Let

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

Then,

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

i.e.,

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

Central Limit Theorem

Fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$.

Theorem (Central Limit Theorem)

Let $X_1, X_2, \ldots$ be a sequence of i.i.d. RVs with common, finite expectation $\mathbb{E}[X_1] = \mu$ and common, finite variance σ^2. Let

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

Then,

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{\mathrm{d}} \mathcal{N}(0, 1),$$

i.e.,

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

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.

Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.

Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x) \, \mathrm{d}x$.

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.
Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.
Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x)\, \mathrm{d}x$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.
Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.
Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x)\, \mathrm{d}x$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Example: $X \sim \mathcal{N}(0, 1)$,

$$g(x) = 10\, e^{-5(x-100)^4}, \qquad x \in \mathbb{R}.$$

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.
Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.
Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x)\, \mathrm{d}x$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Example: $X \sim \mathcal{N}(0, 1)$,

$$g(x) = 10\, e^{-5(x-100)^4}, \qquad x \in \mathbb{R}.$$

- g attains maximum at $x = 100$
- X takes values around 100 with very small probability.

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.

Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.

Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x)\, \mathrm{d}x$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.
Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.
Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x)\, \mathrm{d}x$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Solution:

- Sample $Y_1, Y_2, \ldots \overset{\text{i.i.d.}}{\sim} f_Y$,
- Choose f_Y to have same support as f_X and so that it is simple to sample from on a computer
- Let

$$J_n = \frac{1}{n} \sum_{r=1}^{n} \frac{g(Y_r) f_X(Y_r)}{f_Y(Y_r)}, \qquad n \in \mathbb{N}.$$

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.
Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.
Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x) \, \mathrm{d}x$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.
Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.
Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x)\, dx$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Solution:

- According to SLLN,

$$J_n \xrightarrow{\text{a.s.}}$$

Importance Sampling

Suppose X is a continuous random variable that has PDF f_X.
Let $g : \mathbb{R} \to \mathbb{R}$ be a given function.
Suppose that our interest is to evaluate the value of $\mathbb{E}[g(X)] = \int g(x) f_X(x)\, \mathrm{d}x$.

What If

What if f_X is too difficult to sample from, or $g(X)$ has a very large variance?

Solution:

- According to SLLN,

$$J_n \xrightarrow{\text{a.s.}} \mathbb{E}\left[\frac{g(Y_1) f_X(Y_1)}{f_Y(Y_1)}\right] = \int g(x) f_X(x)\, \mathrm{d}x.$$

Importance Sampling

Going back to the example:

Example: $X \sim \mathcal{N}(0, 1)$,

$$g(x) = 10\, e^{-5(x-100)^4}, \qquad x \in \mathbb{R}.$$

We want to compute $\mathbb{E}[g(X)]$.

Importance Sampling

Going back to the example:

Example: $X \sim \mathcal{N}(0,1)$,

$$g(x) = 10\, e^{-5(x-100)^4}, \qquad x \in \mathbb{R}.$$

We want to compute $\mathbb{E}[g(X)]$.

Solution:

- Choose $f_Y = \mathcal{N}(100, 1)$
- Sample $Y_1, \ldots, Y_n \sim f_Y$
- Set

$$J_n = \frac{1}{n} \sum_{i=1}^{n} \frac{g(Y_n)\, f_X(Y_n)}{f_Y(y_n)}.$$

- SLLN guarantees that $J_n \xrightarrow{\text{a.s.}} \mathbb{E}[g(X)]$