Convergence, LLN and CLT

Hayk Aprikyan, Hayk Tarkhanyan

Convergence of Functions

In calculus, we used to work with *sequences of numbers:*

$$x_1,\ x_2,\ x_3,\ \ldots,\ x_n,\ \ldots$$

and in case the sequence got closer and closer to some number x as n gets larger and larger, we said that the sequence *converges* to x:

$$\lim_{n\to\infty} x_n = x$$

Convergence of Functions

If instead, we had a *sequence of functions:*

$$f_1 \quad , f_2 \quad , f_3 \quad , \ldots, f_n \quad , \ldots$$

and

Convergence of Functions

If instead, we had a *sequence of functions:*

$$f_1(a),\ f_2(a),\ f_3(a),\ \ldots,\ f_n(a),\ \ldots$$

and at some point a, their values got closer to some number b as $n \to \infty$, we say that the sequence $\{f_n\}$ *converges* to b **at the point** a:

$$\lim_{n\to\infty} f_n(a) = b$$

Convergence of Functions

For example, consider the sequence of functions:

$$f_n(x) = \frac{x}{n}$$

Convergence of Functions

For example, consider the sequence of functions:

$$f_n(x) = \frac{x}{n}$$

At point $a = 5$, we have:

$$f_1(5) = 5, \quad f_2(5) = 2.5, \quad f_3(5) = \frac{5}{3}, \quad f_4(5) = 1.25, \quad \ldots$$

Convergence of Functions

For example, consider the sequence of functions:

$$f_n(x) = \frac{x}{n}$$

At point $a = 5$, we have:

$$f_1(5) = 5, \quad f_2(5) = 2.5, \quad f_3(5) = \frac{5}{3}, \quad f_4(5) = 1.25, \quad \ldots$$

and indeed, $\lim_{n \to \infty} f_n(5) = 0$.

For example, consider the sequence of functions:

$$f_n(x) = \frac{x}{n}$$

At point $a = 5$, we have:

$$f_1(5) = 5, \quad f_2(5) = 2.5, \quad f_3(5) = \frac{5}{3}, \quad f_4(5) = 1.25, \quad \ldots$$

and indeed, $\lim_{n\to\infty} f_n(5) = 0$.

Question

What happens at the point $a = 6$?

Convergence of Functions

For example, consider the sequence of functions:

$$f_n(x) = \frac{x}{n}$$

At point $a = 5$, we have:

$$f_1(5) = 5, \quad f_2(5) = 2.5, \quad f_3(5) = \frac{5}{3}, \quad f_4(5) = 1.25, \quad \ldots$$

and indeed, $\lim\limits_{n\to\infty} f_n(5) = 0$.

Question

What happens at the point $a = 6$? At $a = 0$?

Convergence of Functions

For example, consider the sequence of functions:

$$f_n(x) = \frac{x}{n}$$

At point $a = 5$, we have:

$$f_1(5) = 5, \quad f_2(5) = 2.5, \quad f_3(5) = \frac{5}{3}, \quad f_4(5) = 1.25, \quad \dots$$

and indeed, $\lim_{n \to \infty} f_n(5) = 0$.

Question

What happens at the point $a = 6$? At $a = 0$?

As we see, this sequence of functions converges to 0 at **every point** $a \in \mathbb{R}$:

$$\lim_{n \to \infty} f_n(a) = 0 \qquad \text{for all } a \in \mathbb{R}$$

Convergence of Functions

In a more general case, if we denote

- the limit at the point 1 by $g(1)$

Convergence of Functions

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$

Convergence of Functions

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$
- the limit at the point 6.4 by $g(6.4)$

Convergence of Functions

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$
- the limit at the point 6.4 by $g(6.4)$
- ...

Convergence of Functions

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$
- the limit at the point 6.4 by $g(6.4)$
- $\dots$
- the limit at the point a by $g(a)$

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$
- the limit at the point 6.4 by $g(6.4)$
- ...
- the limit at the point a by $g(a)$

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$
- the limit at the point 6.4 by $g(6.4)$
- $\ldots$
- the limit at the point a by $g(a)$

then we will get a new function $g(x)$,

Convergence of Functions

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$
- the limit at the point 6.4 by $g(6.4)$
- ...
- the limit at the point a by $g(a)$

then we will get a new function $g(x)$, and we can say that

$$f_n(a) \to g(a) \qquad \text{for all } a \in \mathbb{R}$$

Convergence of Functions

In a more general case, if we denote

- the limit at the point 1 by $g(1)$
- the limit at the point 2 by $g(2)$
- the limit at the point 6.4 by $g(6.4)$
- $\ldots$
- the limit at the point a by $g(a)$

then we will get a new function $g(x)$, and we can say that

$$f_n(a) \to g(a) \qquad \text{for all } a \in \mathbb{R}$$

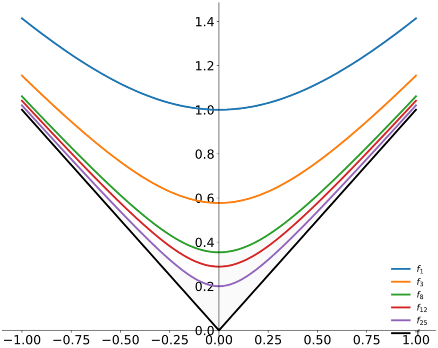
or simply that the sequence of functions $\{f_n\}$ *converges* to the function g.

Convergence of Functions

Definition

A sequence of functions $\{f_n\}$ *converges* to a function g, if for every point a and for every $\varepsilon > 0$, there exists a natural number N such that for all $n \geq N$,

$$|f_n(a) - g(a)| < \varepsilon$$



Random Variable as a Function

What does all this have to do with random variables?

Random Variable as a Function

What does all this have to do with random variables?

Whenever we have a random variable X, we can think of it as a function:

$$\text{for each outcome } \omega \in \Omega \;\mapsto\; \text{some value}$$

Example

Random Variable as a Function

What does all this have to do with random variables?

Whenever we have a random variable X, we can think of it as a function:

$$\text{for each outcome } \omega \in \Omega \;\mapsto\; \text{some value}$$

Example

For example, if X denotes the height of a random person, we can say:

- $\Omega =$ the set of all people in the world
- for each person ω, $X(\omega)$ is the height of that person

Random Variable as a Function

What does all this have to do with random variables?

Whenever we have a random variable X, we can think of it as a function:

$$\text{for each outcome } \omega \in \Omega \;\mapsto\; \text{some value}$$

Example

For example, if X denotes the height of a random person, we can say:

- $\Omega =$ the set of all people in the world
- for each person ω, $X(\omega)$ is the height of that person

i.e.

```
X = lambda ω:  the height of ω
X(Arthur Abraham)
>> 1.75
```

Random Variable as a Function

Example

If X denotes the outcome of tossing a coin, we can say:

- $\Omega = \{\mathtt{H}, \mathtt{T}\}$
- $X(\mathtt{H}) = 1$ and $X(\mathtt{T}) = 0$

Random Variable as a Function

Example

If X denotes the outcome of tossing a coin, we can say:

- $\Omega = \{\mathtt{H}, \mathtt{T}\}$
- $X(\mathtt{H}) = 1$ and $X(\mathtt{T}) = 0$

In some cases, the formula about how X depends on ω can be more complicated, and it is not always easy to keep track of it – but theoretically, X is always a function from Ω to $\mathbb{R}$.

Random Variable as a Function

Example

If X denotes the outcome of tossing a coin, we can say:

- $\Omega = \{\mathtt{H}, \mathtt{T}\}$
- $X(\mathtt{H}) = 1$ and $X(\mathtt{T}) = 0$

In some cases, the formula about how X depends on ω can be more complicated, and it is not always easy to keep track of it – but theoretically, X is always a function from Ω to $\mathbb{R}$.

Question

So can we talk about the convergence of random variables?

Random Variable as a Function

Example

If X denotes the outcome of tossing a coin, we can say:

- $\Omega = \{\mathtt{H}, \mathtt{T}\}$
- $X(\mathtt{H}) = 1$ and $X(\mathtt{T}) = 0$

In some cases, the formula about how X depends on ω can be more complicated, and it is not always easy to keep track of it – but theoretically, X is always a function from Ω to $\mathbb{R}$.

Question

So can we talk about the convergence of random variables?

As it turns out, the answer is yes.

Convergence of Random Variables

Everyday Anahit randomly chooses a number from the interval $[0, 1]$ and multiplies the number by itself n times, where n is the number of days since she started doing this.

Example

- If on the first day, she chooses 0.5, then $X_1 = 0.5^1 = 0.5$

Convergence of Random Variables

Everyday Anahit randomly chooses a number from the interval $[0, 1]$ and multiplies the number by itself n times, where n is the number of days since she started doing this.

Example

- If on the first day, she chooses 0.5, then $X_1 = 0.5^1 = 0.5$
- If on the second day, she chooses 0.8, then $X_2 = 0.8^2 = 0.64$

Convergence of Random Variables

Everyday Anahit randomly chooses a number from the interval $[0, 1]$ and multiplies the number by itself n times, where n is the number of days since she started doing this.

Example

- If on the first day, she chooses 0.5, then $X_1 = 0.5^1 = 0.5$
- If on the second day, she chooses 0.8, then $X_2 = 0.8^2 = 0.64$
- If on the third day, she chooses 0.9, then $X_3 = 0.9^3 = 0.729$

Convergence of Random Variables

Everyday Anahit randomly chooses a number from the interval $[0, 1]$ and multiplies the number by itself n times, where n is the number of days since she started doing this.

Example

- If on the first day, she chooses 0.5, then $X_1 = 0.5^1 = 0.5$
- If on the second day, she chooses 0.8, then $X_2 = 0.8^2 = 0.64$
- If on the third day, she chooses 0.9, then $X_3 = 0.9^3 = 0.729$

Convergence of Random Variables

Everyday Anahit randomly chooses a number from the interval $[0, 1]$ and multiplies the number by itself n times, where n is the number of days since she started doing this.

Example

- If on the first day, she chooses 0.5, then $X_1 = 0.5^1 = 0.5$
- If on the second day, she chooses 0.8, then $X_2 = 0.8^2 = 0.64$
- If on the third day, she chooses 0.9, then $X_3 = 0.9^3 = 0.729$

In other words, on the n-th day, she chooses a random number Y_n from $[0, 1]$ and defines $X_n = Y_n^n$.

Convergence of Random Variables

Everyday Anahit randomly chooses a number from the interval $[0, 1]$ and multiplies the number by itself n times, where n is the number of days since she started doing this.

Example

- If on the first day, she chooses 0.5, then $X_1 = 0.5^1 = 0.5$
- If on the second day, she chooses 0.8, then $X_2 = 0.8^2 = 0.64$
- If on the third day, she chooses 0.9, then $X_3 = 0.9^3 = 0.729$

In other words, on the n-th day, she chooses a random number Y_n from $[0, 1]$ and defines $X_n = Y_n^n$.

Question

What happens to X_n as n gets larger and larger? Does it converge to some number?

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.)*.

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*

Convergence of Random Variables

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.).*

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*

Convergence of Random Variables

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.)*.

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*
- The middle case: $X_n \xrightarrow{p} X$ *in probability*

Convergence of Random Variables

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.).*

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*
- The middle case: $X_n \xrightarrow{p} X$ *in probability*

Convergence of Random Variables

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.)*.

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*
- The middle case: $X_n \xrightarrow{p} X$ *in probability*
- The weakest case: $X_n \xrightarrow{d} X$ *in distribution*

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.)*.

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*
  when $X_n \to X$ a.s., i.e. for almost any ω, the values of X_n get closer and closer to the value of X
- The middle case: $X_n \xrightarrow{p} X$ *in probability*
- The weakest case: $X_n \xrightarrow{d} X$ *in distribution*

Convergence of Random Variables

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.)*.

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*
  when $X_n \to X$ a.s., i.e. for almost any ω, the values of X_n get closer and closer to the value of X
- The middle case: $X_n \xrightarrow{p} X$ *in probability*
  when for any ε,
  $$\mathbb{P}[|X - X_n| > \varepsilon] < \varepsilon$$
  after some point n
- The weakest case: $X_n \xrightarrow{d} X$ *in distribution*

Convergence of Random Variables

Definition

If something happens with probability 1 (i.e. it can also not happen, but this event has probability zero) we say that it happens *almost surely (a.s.)*.

We distinguish between three possible ways how a sequence of random variables $\{X_n\}$ can converge to a random variable X:

- The strongest case: $X_n \to X$ *almost surely*
  when $X_n \to X$ a.s., i.e. for almost any ω, the values of X_n get closer and closer to the value of X
- The middle case: $X_n \xrightarrow{p} X$ *in probability*
  when for any ε,

  $$\mathbb{P}[|X - X_n| > \varepsilon] < \varepsilon$$

  after some point n
- The weakest case: $X_n \xrightarrow{d} X$ *in distribution*
  when just the PMF/PDFs of X_n converge to the PMF/PDF of X – but the values of X_n do not necessarily get closer to the value of X.

Convergence of Random Variables

Theorem

Almost sure $\Rightarrow$ in probability $\Rightarrow$ in distribution

Theorem

Almost sure $\Rightarrow$ in probability $\Rightarrow$ in distribution

Example

Let X_n be the random variable defined as follows:

- we toss a coin n times
- if we get all heads, then $X_n = 1$
- otherwise (i.e. at least one tail), $X_n = 0$

Convergence of Random Variables

Theorem

Almost sure $\Rightarrow$ in probability $\Rightarrow$ in distribution

Example

Let X_n be the random variable defined as follows:

- we toss a coin n times
- if we get all heads, then $X_n = 1$
- otherwise (i.e. at least one tail), $X_n = 0$

What can we say about the convergence of X_n?

Law of Large Numbers

Finally, we are ready to state two important theorems about the convergence of random variables:

Finally, we are ready to state two important theorems about the convergence of random variables:

Law of Large Numbers (LLN)

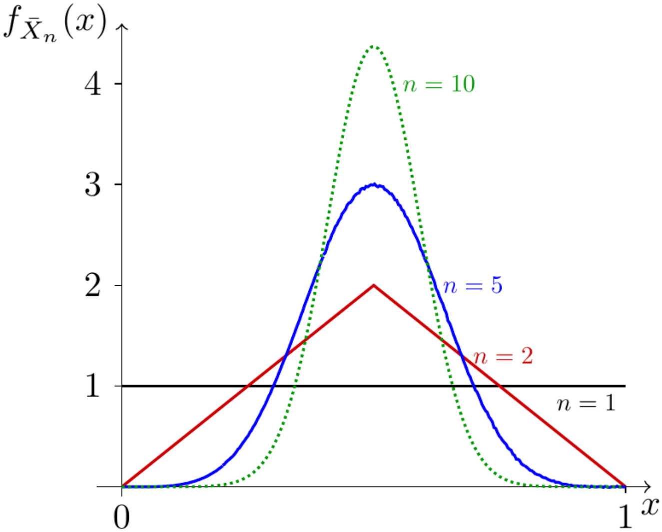
Let $X_1, X_2, \ldots$ be a sequence of independent and identically distributed (i.i.d.) random variables with the same expected value:

$$\mathbb{E}[X_k] = m$$

and with finite variance (again, same). Then their average

$$\frac{X_1 + X_2 + \cdots + X_n}{n} \xrightarrow{p} m \qquad \text{as } n \to \infty$$

Law of Large Numbers



Central Limit Theorem

Ok, so eventually, the average of i.i.d. random variables converges to their expected value. But what about the distribution of the average?

Central Limit Theorem

Ok, so eventually, the average of i.i.d. random variables converges to their expected value. But what about the distribution of the average?

Central Limit Theorem (CLT)

Let $X_1, X_2, \dots$ be a sequence of independent and identically distributed (i.i.d.) random variables with the same expected value:

$$\mathbb{E}[X_k] = m$$

and with finite variance

$$\mathrm{Var}[X_k] = \sigma^2$$

Then their normalized average

$$\sum_{k=1}^{n} \frac{X_k - m}{\sigma\sqrt{n}} \xrightarrow{d} \mathrm{N}(0,1) \qquad \text{as } n \to \infty$$