CS 107, Probability, Spring 2019
Lecture 42

Michael Poghosyan

AUA

08 May 2019

Content

- The Law of Large Numbers
- The Central Limit Theorem

Limit Theorems

The Law of Large Numbers

Recall that we defined $X_1, X_2, \ldots, X_n$ are IID if

- X_1, ... ,X_n are Identically Distributed, i.e., they have the same Distribution (the same CDFs, say);
- $X_1, \ldots, X_n$ are Independent.

The Law of Large Numbers

Recall that we defined $X_1, X_2, ..., X_n$ are IID if

- X_1, ... ,X_n are Identically Distributed, i.e., they have the same Distribution (the same CDFs, say);
- $X_1, ..., X_n$ are Independent.

Note: Please note that since all X_k-s have the same Distribution, then they have the same Expected values, the same Variances and all other characteristics.

The Law of Large Numbers

Recall that we defined $X_1, X_2, ..., X_n$ are IID if

- X_1, ... ,X_n are Identically Distributed, i.e., they have the same Distribution (the same CDFs, say);
- $X_1, ..., X_n$ are Independent.

Note: Please note that since all X_k-s have the same Distribution, then they have the same Expected values, the same Variances and all other characteristics. So

$$\mathbb{E}(X_1) = \mathbb{E}(X_2) = ... = \mathbb{E}(X_n),$$

$$Var(X_1) = Var(X_2) = ... = Var(X_n).$$

The Question

The Questions we consider here are:

Questions

Assume $X_1, X_2, ..., X_n$ are IID r.v.s.

Q1 What is the Distribution of

$$S_n = X_1 + X_2 + ... + X_n?$$

Q2 What is the Distribution of

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}?$$

The Question

The Questions we consider here are:

Questions

Assume $X_1, X_2, ..., X_n$ are IID r.v.s.

Q1 What is the Distribution of

$$S_n = X_1 + X_2 + ... + X_n?$$

Q2 What is the Distribution of

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n}?$$

As you remember, calculation of the Distribution of the sum $X + Y$ is not an easy job (one needs to calculate Convolutions), so calculation of the exact Distribution of S_n and $\overline{X}_n$ is not an easy job, in general.

The Sum and Sample Mean Distribution

It turns out that for some particular cases, we can exactly describe the Distribution of S_n and $\overline{X}_n$:

The Sum and Sample Mean Distribution

It turns out that for some particular cases, we can exactly describe the Distribution of S_n and $\overline{X}_n$:

- If $X_k \sim Bernoulli(p)$, $k = 1, ..., n$ are independent, then

$$S_n = X_1 + ... + X_n \sim Binom(n, p);$$

The Sum and Sample Mean Distribution

It turns out that for some particular cases, we can exactly describe the Distribution of S_n and $\overline{X}_n$:

- If $X_k \sim Bernoulli(p)$, $k = 1, ..., n$ are independent, then
$$S_n = X_1 + ... + X_n \sim Binom(n, p);$$
- If $X_k \sim Binom(m, p)$, $k = 1, ..., n$, are independent, then
$$S_n = X_1 + ... + X_n \sim Binom(n \cdot m, p).$$

The Sum and Sample Mean Distribution

It turns out that for some particular cases, we can exactly describe the Distribution of S_n and $\overline{X}_n$:

- If $X_k \sim Bernoulli(p)$, $k = 1, ..., n$ are independent, then
$$S_n = X_1 + ... + X_n \sim Binom(n, p);$$
- If $X_k \sim Binom(m, p)$, $k = 1, ..., n$, are independent, then
$$S_n = X_1 + ... + X_n \sim Binom(n \cdot m, p).$$
- If $X_k \sim \mathcal{N}(\mu, \sigma^2)$, $k = 1, ..., n$, are Independent, then
$$S_n = X_1 + ... + X_n \sim \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$$
$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

The Sum and Sample Mean Distribution

- If $X_k \sim Pois(\lambda)$, $k = 1, ..., n$, are independent, then
$$S_n = X_1 + ... + X_n \sim Pois(n \cdot \lambda).$$

The Sum and Sample Mean Distribution

- If $X_k \sim Pois(\lambda)$, $k = 1, ..., n$, are independent, then

$$S_n = X_1 + ... + X_n \sim Pois(n \cdot \lambda).$$

- If $X_k \sim Geom(p)$, $k = 1, ..., n$, are independent, then

$$S_n = X_1 + ... + X_n \sim NegBin(n, p).$$

The Sum and Sample Mean Distribution

- If $X_k \sim Pois(\lambda)$, $k = 1, ..., n$, are independent, then

$$S_n = X_1 + ... + X_n \sim Pois(n \cdot \lambda).$$

- If $X_k \sim Geom(p)$, $k = 1, ..., n$, are independent, then

$$S_n = X_1 + ... + X_n \sim NegBin(n, p).$$

  Here, the Negative Binomial Distribution is the number of failures before the n-th success when doing $Bernoulli(p)$ trials, see `https://en.wikipedia.org/wiki/Negative_binomial_distribution`.

Some Partial Information

Now, going back to the general case.

Some Partial Information

Now, going back to the general case. The general problem of finding the Distribution of S_n and $\overline{X}_n$ is a very hard one, but we can get some partial information about S_n and $\overline{X}_n$:

Some Partial Information

Now, going back to the general case. The general problem of finding the Distribution of S_n and $\overline{X}_n$ is a very hard one, but we can get some partial information about S_n and $\overline{X}_n$:

$$\mathbb{E}(S_n) = \mathbb{E}(X_1 + \ldots + X_n) =$$

Some Partial Information

Now, going back to the general case. The general problem of finding the Distribution of S_n and $\overline{X}_n$ is a very hard one, but we can get some partial information about S_n and $\overline{X}_n$:

$$\mathbb{E}(S_n) = \mathbb{E}(X_1 + ... + X_n) = n \cdot \mathbb{E}(X_1);$$

Some Partial Information

Now, going back to the general case. The general problem of finding the Distribution of S_n and $\overline{X}_n$ is a very hard one, but we can get some partial information about S_n and $\overline{X}_n$:

$$\mathbb{E}(S_n) = \mathbb{E}(X_1 + ... + X_n) = n \cdot \mathbb{E}(X_1);$$

$$Var(S_n) = Var(X_1 + ... + X_n) =$$

Some Partial Information

Now, going back to the general case. The general problem of finding the Distribution of S_n and $\overline{X}_n$ is a very hard one, but we can get some partial information about S_n and $\overline{X}_n$:

$$\mathbb{E}(S_n) = \mathbb{E}(X_1 + ... + X_n) = n \cdot \mathbb{E}(X_1);$$

$$Var(S_n) = Var(X_1 + ... + X_n) = n \cdot Var(X_1);$$

and

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + \ldots + X_n}{n}\right) =$$

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + ... + X_n}{n}\right) = \mathbb{E}(X_1);$$

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + ... + X_n}{n}\right) = \mathbb{E}(X_1);$$

$$Var(\overline{X}_n) = Var\left(\frac{X_1 + ... + X_n}{n}\right) =$$

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + ... + X_n}{n}\right) = \mathbb{E}(X_1);$$

$$Var(\overline{X}_n) = Var\left(\frac{X_1 + ... + X_n}{n}\right) = \frac{Var(X_1)}{n}.$$

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + \ldots + X_n}{n}\right) = \mathbb{E}(X_1);$$

$$Var(\overline{X}_n) = Var\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{Var(X_1)}{n}.$$

Note: Surely you know the well-known saying: *"7 angam chaphir, mek angam ktrir"*.

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + ... + X_n}{n}\right) = \mathbb{E}(X_1);$$

$$Var(\overline{X}_n) = Var\left(\frac{X_1 + ... + X_n}{n}\right) = \frac{Var(X_1)}{n}.$$

Note: Surely you know the well-known saying: *"7 angam chaphir, mek angam ktrir"*. The above relationships are the mathematical statement of this saying $\ddot\smile$

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + \ldots + X_n}{n}\right) = \mathbb{E}(X_1);$$

$$Var(\overline{X}_n) = Var\left(\frac{X_1 + \ldots + X_n}{n}\right) = \frac{Var(X_1)}{n}.$$

Note: Surely you know the well-known saying: *"7 angam chaphir, mek angam ktrir"*. The above relationships are the mathematical statement of this saying $\smile$

Note: These relations are important in the Financial Risk and Investment Theory: they are confirming the importance of the **Diversification**.

Some Partial Information

$$\mathbb{E}(\overline{X}_n) = \mathbb{E}\left(\frac{X_1 + ... + X_n}{n}\right) = \mathbb{E}(X_1);$$

$$Var(\overline{X}_n) = Var\left(\frac{X_1 + ... + X_n}{n}\right) = \frac{Var(X_1)}{n}.$$

Note: Surely you know the well-known saying: *"7 angam chaphir, mek angam ktrir"*. The above relationships are the mathematical statement of this saying ¨⌣

Note: These relations are important in the Financial Risk and Investment Theory: they are confirming the importance of the **Diversification**.

Folklore: Diversification in one sentence: Do not put all your eggs into one basket!

Intro to LLN and CLT

Now, we want to study more properties of Distributions of $\overline{X}_n$ and/or S_n in the general case.

Intro to LLN and CLT

Now, we want to study more properties of Distributions of $\overline{X}_n$ and/or S_n in the general case.
The two famous Limit Theorems in Probability Theory, The Law of the Large Numbers (LLN) and the Central Limit Theorem (CLT) help us to get an information about the **asymptotic** (i.e., limiting, or, for large n) properties of $\overline{X}_n$ and S_n.

The Weak LLN

The Weak Law of Large Numbers, WLLN

If $X_1, X_2, ..., X_n$ are IID, with finite $\mathbb{E}(X_1)$ and Variance $Var(X_1)$, then

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(X_1), \qquad n \to +\infty,$$

The Weak LLN

The Weak Law of Large Numbers, WLLN

If $X_1, X_2, ..., X_n$ are IID, with finite $\mathbb{E}(X_1)$ and Variance $Var(X_1)$, then

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(X_1), \qquad n \to +\infty,$$

where $\mathbb{P}$ over the arrow sign means the convergence is in Probability:

The Weak LLN

The Weak Law of Large Numbers, WLLN

If $X_1, X_2, ..., X_n$ are IID, with finite $\mathbb{E}(X_1)$ and Variance $Var(X_1)$, then

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(X_1), \qquad n \to +\infty,$$

where $\mathbb{P}$ over the arrow sign means the convergence is in Probability: for any $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mathbb{E}(X_1)\right| \geq \varepsilon\right) \to 0, \qquad n \to +\infty.$$

The Weak LLN

The Weak Law of Large Numbers, WLLN

If $X_1, X_2, ..., X_n$ are IID, with finite $\mathbb{E}(X_1)$ and Variance $Var(X_1)$, then

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(X_1), \qquad n \to +\infty,$$

where $\mathbb{P}$ over the arrow sign means the convergence is in Probability: for any $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\frac{X_1 + X_2 + ... + X_n}{n} - \mathbb{E}(X_1)\right| \geq \varepsilon\right) \to 0, \qquad n \to +\infty.$$

Note: This means that for any $\varepsilon > 0$, the chances that $\overline{X}_n$ is far from $\mathbb{E}(X_1)$ more than ε, is very small, if n is large.

The Weak LLN: The Idea

The idea of the LLN is the following:

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$.

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$.

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ?

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.
And what we need to have to say that $\overline{X}_n$ is close to μ?

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.
And what we need to have to say that $\overline{X}_n$ is close to μ? That the Variance of $\overline{X}_n$ is close to 0!

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.
And what we need to have to say that $\overline{X}_n$ is close to μ? That the Variance of $\overline{X}_n$ is close to 0! And we know that

$$Var(\overline{X}_n) =$$

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.
And what we need to have to say that $\overline{X}_n$ is close to μ? That the Variance of $\overline{X}_n$ is close to 0! And we know that

$$Var(\overline{X}_n) = \frac{Var(X_1)}{n}$$

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.
And what we need to have to say that $\overline{X}_n$ is close to μ? That the Variance of $\overline{X}_n$ is close to 0! And we know that

$$Var(\overline{X}_n) = \frac{Var(X_1)}{n} \to 0, \qquad n \to +\infty.$$

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.
And what we need to have to say that $\overline{X}_n$ is close to μ? That the Variance of $\overline{X}_n$ is close to 0! And we know that

$$Var(\overline{X}_n) = \frac{Var(X_1)}{n} \to 0, \qquad n \to +\infty.$$

Ura!

The Weak LLN: The Idea

The idea of the LLN is the following: let us denote $\mu = \mathbb{E}(X_1)$. We know that the Mean of $\overline{X}_n$ is μ: $\mathbb{E}(\overline{X}_n) = \mu$. Is this enough to say that $\overline{X}_n$ is close to μ? Of course, no.
And what we need to have to say that $\overline{X}_n$ is close to μ? That the Variance of $\overline{X}_n$ is close to 0! And we know that

$$Var(\overline{X}_n) = \frac{Var(X_1)}{n} \to 0, \qquad n \to +\infty.$$

Ura!
The rigorous proof is by using the Chebyshev Inequality: OTB

The Strong LLN

The Strong LLN Says that the above convergence holds also in the Strong Sense, under less restrictive settings:

The Strong Law of Large Numbers, SLLN, Kolmogorov

If $X_1, X_2, ..., X_n$ are IID, with finite $\mathbb{E}(X_1)$, then

$$\frac{X_1 + X_2 + ... + X_n}{n} \xrightarrow{a.s.} \mathbb{E}(X_1), \qquad n \to +\infty,$$

The Strong LLN

The Strong LLN Says that the above convergence holds also in the Strong Sense, under less restrictive settings:

The Strong Law of Large Numbers, SLLN, Kolmogorov

If $X_1, X_2, ..., X_n$ are IID, with finite $\mathbb{E}(X_1)$, then

$$\frac{X_1 + X_2 + ... + X_n}{n} \stackrel{a.s.}{\to} \mathbb{E}(X_1), \qquad n \to +\infty,$$

where *a.s.* over the arrow sign means the convergence is in Almost Sure sense:

The Strong LLN

The Strong LLN Says that the above convergence holds also in the Strong Sense, under less restrictive settings:

The Strong Law of Large Numbers, SLLN, Kolmogorov

If $X_1, X_2, ..., X_n$ are IID, with finite $\mathbb{E}(X_1)$, then

$$\frac{X_1 + X_2 + ... + X_n}{n} \stackrel{a.s.}{\to} \mathbb{E}(X_1), \qquad n \to +\infty,$$

where *a.s.* over the arrow sign means the convergence is in Almost Sure sense:

$$\mathbb{P}\left(\lim_{n\to+\infty} \frac{X_1 + X_2 + ... + X_n}{n} = \mathbb{E}(X_1)\right) = 1.$$

Example:

Example: Assume $X_1, X_2, ..., X_n \sim Unif[-1, 2]$ are IID. Calculate, in the $\mathbb{P}$ and a.s. sense,

$$\lim_{n \to +\infty} \frac{X_1^2 + X_2^2 + ... + X_n^2}{n}.$$

Solution: OTB

Example:

Example: Assume we have a coin, for which the probability of having Heads is $p \in (0, 1)$. We are tossing that coin many times. We calculate the proportion of the Heads for that tosses. What is the limit of that proportion, almost surely, if we repeat tossing infinitely many times?

Solution: OTB

Example:

Question: Why the idea of a Casino works?

Example:

Question: Why the idea of a Casino works? And why the idea of an Insurance works?

Example:

Question: Why the idea of a Casino works? And why the idea of an Insurance works?
Explanation: Sometimes Casinos and Insurance companies pay a lot of money. So why their jobs are profitable?

Example:

Question: Why the idea of a Casino works? And why the idea of an Insurance works?
Explanation: Sometimes Casinos and Insurance companies pay a lot of money. So why their jobs are profitable?
Idea: Because of the LLN!

Example:

Question: Why the idea of a Casino works? And why the idea of an Insurance works?
Explanation: Sometimes Casinos and Insurance companies pay a lot of money. So why their jobs are profitable?
Idea: Because of the LLN! You cannot run away from the fortune, I mean, from Math. $\ddot\smile$.

Example:

Question: Why the idea of a Casino works? And why the idea of an Insurance works?
Explanation: Sometimes Casinos and Insurance companies pay a lot of money. So why their jobs are profitable?
Idea: Because of the LLN! You cannot run away from the fortune, I mean, from Math. $\ddot\smile$. In particular, from LLN.

Example:

Question: Why the idea of a Casino works? And why the idea of an Insurance works?
Explanation: Sometimes Casinos and Insurance companies pay a lot of money. So why their jobs are profitable?
Idea: Because of the LLN! You cannot run away from the fortune, I mean, from Math. $\ddot\smile$. In particular, from LLN. **In the long run, Casinos and Insurance Companies will win**!
OTB

Example:

Question: Why the idea of a Casino works? And why the idea of an Insurance works?
Explanation: Sometimes Casinos and Insurance companies pay a lot of money. So why their jobs are profitable?
Idea: Because of the LLN! You cannot run away from the fortune, I mean, from Math. $\ddot\smile$. In particular, from LLN. **In the long run, Casinos and Insurance Companies will win**!
OTB
Similar Idea: This is a quote by Keynes, one of the most important economists of 20th century, see `https://en.wikipedia.org/wiki/John_Maynard_Keynes`: **In the long run, we are all dead** $\ddot\smile$

The Central Limit Theorem

CLT gives more info about the Distribution of S_n and $\overline{X}_n$:

The Central Limit Theorem, CLT

Assume $X_1, X_2, ..., X_n$ are IID with finite Expectation $\mu = \mathbb{E}(X_1)$ and Variance $\sigma^2 = Var(X_1)$.

The Central Limit Theorem

CLT gives more info about the Distribution of S_n and $\overline{X}_n$:

The Central Limit Theorem, CLT

Assume $X_1, X_2, ..., X_n$ are IID with finite Expectation $\mu = \mathbb{E}(X_1)$ and Variance $\sigma^2 = Var(X_1)$. We Standardize S_n (or $\overline{X}_n$):

$$Z_n =$$

The Central Limit Theorem

CLT gives more info about the Distribution of S_n and $\overline{X}_n$:

The Central Limit Theorem, CLT

Assume $X_1, X_2, ..., X_n$ are IID with finite Expectation $\mu = \mathbb{E}(X_1)$ and Variance $\sigma^2 = Var(X_1)$. We Standardize S_n (or $\overline{X}_n$):

$$Z_n = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{Var(S_n)}} =$$

The Central Limit Theorem

CLT gives more info about the Distribution of S_n and $\overline{X}_n$:

The Central Limit Theorem, CLT

Assume $X_1, X_2, ..., X_n$ are IID with finite Expectation $\mu = \mathbb{E}(X_1)$ and Variance $\sigma^2 = Var(X_1)$. We Standardize S_n (or $\overline{X}_n$):

$$Z_n = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma}$$

$$\Bigg(Z_n =$$

The Central Limit Theorem

CLT gives more info about the Distribution of S_n and $\overline{X}_n$:

The Central Limit Theorem, CLT

Assume $X_1, X_2, ..., X_n$ are IID with finite Expectation $\mu = \mathbb{E}(X_1)$ and Variance $\sigma^2 = Var(X_1)$. We Standardize S_n (or $\overline{X}_n$):

$$Z_n = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma}$$

$$\left(Z_n = \frac{\overline{X}_n - \mathbb{E}(\overline{X}_n)}{\sqrt{Var(\overline{X}_n)}} = \right.$$

The Central Limit Theorem

CLT gives more info about the Distribution of S_n and $\overline{X}_n$:

The Central Limit Theorem, CLT

Assume $X_1, X_2, \dots, X_n$ are IID with finite Expectation $\mu = \mathbb{E}(X_1)$ and Variance $\sigma^2 = Var(X_1)$. We Standardize S_n (or $\overline{X}_n$):

$$Z_n = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma}$$

$$\left(Z_n = \frac{\overline{X}_n - \mathbb{E}(\overline{X}_n)}{\sqrt{Var(\overline{X}_n)}} = \frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \right).$$

The Central Limit Theorem

CLT gives more info about the Distribution of S_n and $\overline{X}_n$:

The Central Limit Theorem, CLT

Assume $X_1, X_2, \dots, X_n$ are IID with finite Expectation $\mu = \mathbb{E}(X_1)$ and Variance $\sigma^2 = Var(X_1)$. We Standardize S_n (or $\overline{X}_n$):

$$Z_n = \frac{S_n - \mathbb{E}(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma}$$

$$\left(Z_n = \frac{\overline{X}_n - \mathbb{E}(\overline{X}_n)}{\sqrt{Var(\overline{X}_n)}} = \frac{\overline{X}_n - \mu}{\sigma / \sqrt{n}} \right).$$

Then, for any subset $A \subset \mathbb{R}$,

$$\mathbb{P}(Z_n \in A) \to \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

CLT

Explanation: If we standardize S_n or $\overline{X}_n$, we obtain a r.v. Z_n, the Distribution of which tends to the Standard Normal Distribution $\mathcal{N}(0, 1)$.

Explanation: If we standardize S_n or $\overline{X}_n$, we obtain a r.v. Z_n, the Distribution of which tends to the Standard Normal Distribution $\mathcal{N}(0,1)$.
Mathematically, under the conditions of CLT, if $Z \sim \mathcal{N}(0,1)$, then, for any $a < b$,

$$\mathbb{P}(a \leq Z_n \leq b) \to \mathbb{P}(a \leq Z \leq b),$$

Explanation: If we standardize S_n or $\overline{X}_n$, we obtain a r.v. Z_n, the Distribution of which tends to the Standard Normal Distribution $\mathcal{N}(0, 1)$.
Mathematically, under the conditions of CLT, if $Z \sim \mathcal{N}(0, 1)$, then, for any $a < b$,

$$\mathbb{P}(a \leq Z_n \leq b) \to \mathbb{P}(a \leq Z \leq b),$$

i.e.,

$$\mathbb{P}(a \leq Z_n \leq b) \to \frac{1}{\sqrt{2\pi}} \cdot \int_a^b e^{-\frac{x^2}{2}} dx.$$

CLT

Explanation: If we standardize S_n or $\overline{X}_n$, we obtain a r.v. Z_n, the Distribution of which tends to the Standard Normal Distribution $\mathcal{N}(0,1)$.
Mathematically, under the conditions of CLT, if $Z \sim \mathcal{N}(0,1)$, then, for any $a < b$,

$$\mathbb{P}(a \leq Z_n \leq b) \to \mathbb{P}(a \leq Z \leq b),$$

i.e.,

$$\mathbb{P}(a \leq Z_n \leq b) \to \frac{1}{\sqrt{2\pi}} \cdot \int_a^b e^{-\frac{x^2}{2}} dx.$$

If we will denote by $\Phi(x)$ the CDF of $\mathcal{N}(0,1)$ Random Variable, then

$$\mathbb{P}(a \leq Z_n \leq b) \to \Phi(b) - \Phi(a),$$

CLT

Explanation: If we standardize S_n or $\overline{X}_n$, we obtain a r.v. Z_n, the Distribution of which tends to the Standard Normal Distribution $\mathcal{N}(0,1)$.
Mathematically, under the conditions of CLT, if $Z \sim \mathcal{N}(0,1)$, then, for any $a < b$,

$$\mathbb{P}(a \le Z_n \le b) \to \mathbb{P}(a \le Z \le b),$$

i.e.,

$$\mathbb{P}(a \le Z_n \le b) \to \frac{1}{\sqrt{2\pi}} \cdot \int_a^b e^{-\frac{x^2}{2}}\,dx.$$

If we will denote by $\Phi(x)$ the CDF of $\mathcal{N}(0,1)$ Random Variable, then

$$\mathbb{P}(a \le Z_n \le b) \to \Phi(b) - \Phi(a),$$

or, sometimes we write, for large n,

$$\mathbb{P}(a \le Z_n \le b) \approx \Phi(b) - \Phi(a).$$

Usage of the CLT

Assume, for IID r.v.s $X_1, \dots, X_n$, we want to calculate, approximately,

$$\mathbb{P}(a \le X_1 + X_2 + \dots + X_n \le b).$$

Usage of the CLT

Assume, for IID r.v.s $X_1, ..., X_n$, we want to calculate, approximately,

$$\mathbb{P}(a \leq X_1 + X_2 + ... + X_n \leq b).$$

Then we do the following steps: let $S_n = X_1 + ... + X_n$.

Usage of the CLT

Assume, for IID r.v.s $X_1, ..., X_n$, we want to calculate, approximately,

$$\mathbb{P}(a \leq X_1 + X_2 + ... + X_n \leq b).$$

Then we do the following steps: let $S_n = X_1 + ... + X_n$.

- We calculate $\mu = \mathbb{E}(X_1)$ and $\sigma = \sqrt{Var(X_1)}$;

Usage of the CLT

Assume, for IID r.v.s $X_1, \dots, X_n$, we want to calculate, approximately,

$$\mathbb{P}(a \le X_1 + X_2 + \dots + X_n \le b).$$

Then we do the following steps: let $S_n = X_1 + \dots + X_n$.

- We calculate $\mu = \mathbb{E}(X_1)$ and $\sigma = \sqrt{Var(X_1)}$;
- We Standardize the above sum:

$$\mathbb{P}(a \le S_n \le b) = \mathbb{P}\left(\frac{a - n \cdot \mu}{\sigma \cdot \sqrt{n}} \le \frac{S_n - n \cdot \mu}{\sigma \cdot \sqrt{n}} \le \frac{b - n \cdot \mu}{\sigma \cdot \sqrt{n}}\right);$$

Usage of the CLT

Assume, for IID r.v.s $X_1, ..., X_n$, we want to calculate, approximately,

$$\mathbb{P}(a \leq X_1 + X_2 + ... + X_n \leq b).$$

Then we do the following steps: let $S_n = X_1 + ... + X_n$.

- We calculate $\mu = \mathbb{E}(X_1)$ and $\sigma = \sqrt{Var(X_1)}$;
- We Standardize the above sum:

$$\mathbb{P}(a \leq S_n \leq b) = \mathbb{P}\left(\frac{a - n\cdot\mu}{\sigma\cdot\sqrt{n}} \leq \frac{S_n - n\cdot\mu}{\sigma\cdot\sqrt{n}} \leq \frac{b - n\cdot\mu}{\sigma\cdot\sqrt{n}}\right);$$

- Use CLT to write that the last Probability is approximately equal to

$$\mathbb{P}(a \leq S_n \leq b) \approx \Phi\left(\frac{b - n\cdot\mu}{\sigma\cdot\sqrt{n}}\right) - \Phi\left(\frac{a - n\cdot\mu}{\sigma\cdot\sqrt{n}}\right).$$

Usage of the CLT

In a similar way, for IID r.v.s $X_1, ..., X_n$, we can calculate, approximately,

$$\mathbb{P}(X_1 + X_2 + ... + X_n \leq b) \qquad \text{or} \qquad \mathbb{P}(X_1 + X_2 + ... + X_n \geq a),$$

or

$$\mathbb{P}(a \leq \frac{X_1 + X_2 + ... + X_n}{n} \leq b),$$

or

$$\mathbb{P}(\frac{X_1 + X_2 + ... + X_n}{n} \leq b), \qquad \mathbb{P}(\frac{X_1 + X_2 + ... + X_n}{n} \geq a).$$

Usage of the CLT

In a similar way, for IID r.v.s $X_1, \dots, X_n$, we can calculate, approximately,

$$\mathbb{P}(X_1 + X_2 + \dots + X_n \leq b) \qquad \text{or} \qquad \mathbb{P}(X_1 + X_2 + \dots + X_n \geq a),$$

or

$$\mathbb{P}(a \leq \frac{X_1 + X_2 + \dots + X_n}{n} \leq b),$$

or

$$\mathbb{P}(\frac{X_1 + X_2 + \dots + X_n}{n} \leq b), \qquad \mathbb{P}(\frac{X_1 + X_2 + \dots + X_n}{n} \geq a).$$

Just use Standardization and the CLT!