YSU ASDS, Statistics, Fall 2019

Lecture 13

Michael Poghosyan

02 Oct 2019

Probability Reminder

Contents

Last Lecture ReCap

- Give the definition of the convergence in the a.s./ Probability / QM / Distributions sense.

Example

Example: Show that if $X_n \sim Binom\left(n, \frac{\lambda}{n}\right)$, then $X_n \xrightarrow{D} Pois(\lambda)$.

Example

Example: Show that if $X_n \sim \mathit{Binom}\left(n, \frac{\lambda}{n}\right)$, then $X_n \xrightarrow{D} \mathit{Pois}(\lambda)$.

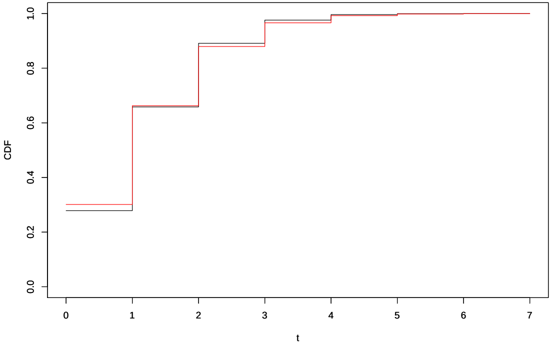
Note: Note that when using $X_n \xrightarrow{D} \mathit{Pois}(\lambda)$ we mean $X_n \xrightarrow{D} X$, where $X \sim \mathit{Pois}(\lambda)$.

Example

Example: Show that if $X_n \sim Binom\left(n, \frac{\lambda}{n}\right)$, then $X_n \xrightarrow{D} Pois(\lambda)$.

Note: Note that when using $X_n \xrightarrow{D} Pois(\lambda)$ we mean $X_n \xrightarrow{D} X$, where $X \sim Pois(\lambda)$.

```
lambda <- 1.2; n <- 10; t <- seq(0,7, 0.1)
plot(t,pbinom(t, size = n, prob = lambda/n), type = "s", ylim = c(0,1), ylab = "CDF")
par(new = T)
plot(t, ppois(t, lambda = lambda), type = "s", col = "red", ylim = c(0,1), ylab = "CDF")
```



Example

Example: Let $X_n \sim Unif[0, n]$. Find the limit in Distributions of $Y_n = \frac{X_n}{n}$.

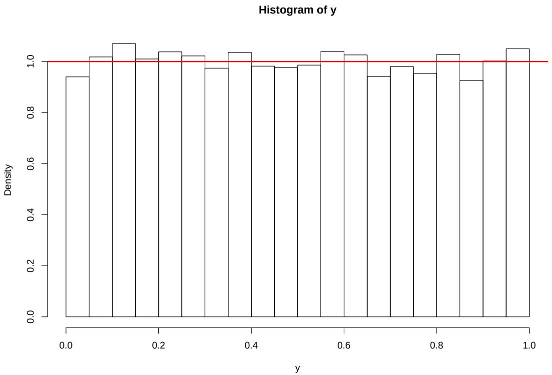
Example

Example: Let $X_n \sim Unif[0, n]$. Find the limit in Distributions of $Y_n = \frac{X_n}{n}$. Let us visually show that $Y_n \xrightarrow{D} Y$, where $Y \sim Unif[0, 1]$.

Example

Example: Let $X_n \sim Unif[0, n]$. Find the limit in Distributions of $Y_n = \frac{X_n}{n}$. Let us visually show that $Y_n \xrightarrow{D} Y$, where $Y \sim Unif[0, 1]$.

```r
n <- 10000 ## We use Y_n
m <- 10000 ## No. of generated numbers
y <- runif(m, min = 0, max = n)/n
hist(y, freq = F)
abline(h = 1, col = "red", lwd = 2)
```



Some Properties

Theorem: Assume $X_n \overset{a.s.(P,qm,D)}{\longrightarrow} X$ and $Y_n \overset{a.s.(P,qm,D)}{\longrightarrow} Y$.

Some Properties

Theorem: Assume $X_n \xrightarrow{a.s.(P,qm,D)} X$ and $Y_n \xrightarrow{a.s.(P,qm,D)} Y$. Then

- $X_n + Y_n \xrightarrow{a.s.(P,qm)} X + Y$;

Some Properties

Theorem: Assume $X_n \overset{a.s.(P,qm,D)}{\longrightarrow} X$ and $Y_n \overset{a.s.(P,qm,D)}{\longrightarrow} Y$. Then

- $X_n + Y_n \overset{a.s.(P,qm)}{\longrightarrow} X + Y$;
- $X_n \cdot Y_n \overset{a.s.(P)}{\longrightarrow} X \cdot Y$;

Some Properties

Theorem: Assume $X_n \xrightarrow{a.s.(P,qm,D)} X$ and $Y_n \xrightarrow{a.s.(P,qm,D)} Y$. Then

- $X_n + Y_n \xrightarrow{a.s.(P,qm)} X + Y$;
- $X_n \cdot Y_n \xrightarrow{a.s.(P)} X \cdot Y$;
- If $g \in C(\mathbb{R})$, then $g(X_n) \xrightarrow{a.s.(P,D)} g(X)$

Some Properties

Theorem: Assume $X_n \xrightarrow{a.s.(P,qm,D)} X$ and $Y_n \xrightarrow{a.s.(P,qm,D)} Y$. Then

- $X_n + Y_n \xrightarrow{a.s.(P,qm)} X + Y$;
- $X_n \cdot Y_n \xrightarrow{a.s.(P)} X \cdot Y$;
- If $g \in C(\mathbb{R})$, then $g(X_n) \xrightarrow{a.s.(P,D)} g(X)$

Note: In the general case, if $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} Y$, then not necessarily $X_n + Y_n \xrightarrow{D} X + Y$.

Some Properties

Theorem: Assume $X_n \xrightarrow{a.s.(P,qm,D)} X$ and $Y_n \xrightarrow{a.s.(P,qm,D)} Y$. Then

- $X_n + Y_n \xrightarrow{a.s.(P,qm)} X + Y$;
- $X_n \cdot Y_n \xrightarrow{a.s.(P)} X \cdot Y$;
- If $g \in C(\mathbb{R})$, then $g(X_n) \xrightarrow{a.s.(P,D)} g(X)$

Note: In the general case, if $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} Y$, then not necessarily $X_n + Y_n \xrightarrow{D} X + Y$.

But we have:

Theorem (Slutsky) Assume $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} c$, where $c \in \mathbb{R}$ is a constant. Then

Some Properties

Theorem: Assume $X_n \xrightarrow{a.s.(P,qm,D)} X$ and $Y_n \xrightarrow{a.s.(P,qm,D)} Y$. Then

- $X_n + Y_n \xrightarrow{a.s.(P,qm)} X + Y$;
- $X_n \cdot Y_n \xrightarrow{a.s.(P)} X \cdot Y$;
- If $g \in C(\mathbb{R})$, then $g(X_n) \xrightarrow{a.s.(P,D)} g(X)$

Note: In the general case, if $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} Y$, then not necessarily $X_n + Y_n \xrightarrow{D} X + Y$.

But we have:

Theorem (Slutsky) Assume $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} c$, where $c \in \mathbb{R}$ is a constant. Then

- $X_n + Y_n \xrightarrow{D} X + c$;

Some Properties

Theorem: Assume $X_n \xrightarrow{a.s.(P,qm,D)} X$ and $Y_n \xrightarrow{a.s.(P,qm,D)} Y$. Then

- $X_n + Y_n \xrightarrow{a.s.(P,qm)} X + Y$;
- $X_n \cdot Y_n \xrightarrow{a.s.(P)} X \cdot Y$;
- If $g \in C(\mathbb{R})$, then $g(X_n) \xrightarrow{a.s.(P,D)} g(X)$

Note: In the general case, if $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{D} Y$, then not necessarily $X_n + Y_n \xrightarrow{D} X + Y$.

But we have:

Theorem (Slutsky) Assume $X_n \xrightarrow{D} X$ and $Y_n \xrightarrow{P} c$, where $c \in \mathbb{R}$ is a constant. Then

- $X_n + Y_n \xrightarrow{D} X + c$;
- $X_n \cdot Y_n \xrightarrow{D} c \cdot X$.

Relationship between Convergence Types

Theorem: (Convergence Relationship Diagram)

$$\begin{tikzcd} X_n \xrightarrow{a.s.} X \arrow[dr] & & \\ & X_n \xrightarrow{\mathbb{P}} X \arrow[r] & X_n \xrightarrow{D} X \\ X_n \xrightarrow{qm} X \arrow[ur] & & \end{tikzcd}$$

Relationship between Convergence Types

Theorem: (Convergence Relationship Diagram)

$$
\begin{tikzcd}
\boxed{X_n \xrightarrow{a.s.} X} \arrow[dr] & & \\
& \boxed{X_n \xrightarrow{\mathbb{P}} X} \arrow[r] & \boxed{X_n \xrightarrow{D} X} \\
\boxed{X_n \xrightarrow{qm} X} \arrow[ur] & &
\end{tikzcd}
$$

Note: Inverse implications are not always correct.

Relationship between Convergence Types

Theorem: (Convergence Relationship Diagram)

$$\begin{tikzcd}
X_n \xrightarrow{a.s.} X \arrow[dr] & & \\
& X_n \xrightarrow{\mathbb{P}} X \arrow[r] & X_n \xrightarrow{D} X \\
X_n \xrightarrow{qm} X \arrow[ur] & &
\end{tikzcd}$$

Note: Inverse implications are not always correct. But, say, the following holds: If $X_n \xrightarrow{D} X$ and $X \equiv \textit{constant}$, then $X_n \xrightarrow{P} X$ (X_n and X are defined on the same Probability space).

Limit Theorems

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s.

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s. This means that:

- All X_n-s have the same Distribution.

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s. This means that:

- All X_n-s have the same Distribution. In particular, all numerical partial characteristics of X_n coincide.

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s. This means that:

- All X_n-s have the same Distribution. In particular, all numerical partial characteristics of X_n coincide. In particular,

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

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

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s. This means that:

- All X_n-s have the same Distribution. In particular, all numerical partial characteristics of X_n coincide. In particular,

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

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

We will use this many-many

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s. This means that:

- All X_n-s have the same Distribution. In particular, all numerical partial characteristics of X_n coincide. In particular,

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

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

We will use this many-many-many-many-... times.

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s. This means that:

- All X_n-s have the same Distribution. In particular, all numerical partial characteristics of X_n coincide. In particular,

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

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

We will use this many-many-many-many-... times.

- X_n-s are independent.

Sequence of IID r.v.

Assume X_n is a sequence of **Independent, Identically Distributed (IID)** r.v.s. This means that:

- All X_n-s have the same Distribution. In particular, all numerical partial characteristics of X_n coincide. In particular,

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

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

We will use this many-many-many-many-... times.

- X_n-s are independent. Say, in particular,

$$Var(X_1+X_2+...+X_n) = Var(X_1)+Var(X_2)+...+Var(X_n) = n \cdot Var(X_1).$$

Idea of the LLN and CLT

Assume we have a sequence X_n of IID rvs.

Idea of the LLN and CLT

Assume we have a sequence X_n of IID rvs. We want to study the behavior of either the sum

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

Idea of the LLN and CLT

Assume we have a sequence X_n of IID rvs. We want to study the behavior of either the sum

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

or the average

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

Idea of the LLN and CLT

Assume we have a sequence X_n of IID rvs. We want to study the behavior of either the sum

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

or the average

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

Note: Not an easy task to find the Distribution of S_n or $\overline{X}_n$.

Idea of the LLN and CLT

Assume we have a sequence X_n of IID rvs. We want to study the behavior of either the sum

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

or the average

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

Note: Not an easy task to find the Distribution of S_n or $\overline{X}_n$. Even for $n = 2$.

Idea of the LLN and CLT

Assume we have a sequence X_n of IID rvs. We want to study the behavior of either the sum

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

or the average

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

Note: Not an easy task to find the Distribution of S_n or $\overline{X}_n$. Even for $n = 2$. We need Convolutions!

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) =$$

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) =$$

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

so *the mean of the means is the mean* $\ddot\smile$, and

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

so *the mean of the means is the mean* $\ddot\smile$, and

$$Var(S_n) =$$

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

so *the mean of the means is the mean* $\ddot\smile$, and

$$Var(S_n) = n \cdot Var(X_1), \qquad Var(\overline{X}_n) =$$

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

so *the mean of the means is the mean* $\ddot\smile$, and

$$Var(S_n) = n \cdot Var(X_1), \qquad Var(\overline{X}_n) = \frac{Var(X_1)}{n}.$$

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

so *the mean of the means is the mean* $\ddot\smile$, and

$$Var(S_n) = n \cdot Var(X_1), \qquad Var(\overline{X}_n) = \frac{Var(X_1)}{n}.$$

The last property is the mathematical proof of the effectivness of "7 angam chapir, mek angam ktrir" $\ddot\smile$

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

so *the mean of the means is the mean* $\ddot\smile$, and

$$Var(S_n) = n \cdot Var(X_1), \qquad Var(\overline{X}_n) = \frac{Var(X_1)}{n}.$$

The last property is the mathematical proof of the effectivness of "7 angam chapir, mek angam ktrir" $\ddot\smile$

The interpretation of $\mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1)$ and $Var(\overline{X}_n) = \frac{Var(X_1)}{n}$:

What we know about S_n and $\overline{X}_n$

Some important known facts about S_n and $\overline{X}_n$ in the general case:

$$\mathbb{E}(S_n) = n \cdot \mathbb{E}(X_1), \qquad \mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1);$$

so *the mean of the means is the mean* $\ddot\smile$, and

$$Var(S_n) = n \cdot Var(X_1), \qquad Var(\overline{X}_n) = \frac{Var(X_1)}{n}.$$

The last property is the mathematical proof of the effectivness of "7 angam chapir, mek angam ktrir" $\ddot\smile$

The interpretation of $\mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1)$ and $Var(\overline{X}_n) = \frac{Var(X_1)}{n}$:the values of $\overline{X}_n$ are centered at $\mathbb{E}(X_1)$ and are becoming more and more concentrated around that number as n increases.

Some Particular results about S_n and $\overline{X}_n$

The following are important cases when we can find exactly the Distribution of S_n and/or $\overline{X}_n$:

Some Particular results about S_n and $\overline{X}_n$

The following are important cases when we can find exactly the Distribution of S_n and/or $\overline{X}_n$:

- If $X_k \sim \mathcal{N}(\mu, \sigma^2)$, $k = 1, ..., n$, are Independent, then

$$S_n = X_1 + ... + X_n \sim$$

Some Particular results about S_n and $\overline{X}_n$

The following are important cases when we can find exactly the Distribution of S_n and/or $\overline{X}_n$:

- If $X_k \sim \mathcal{N}(\mu, \sigma^2)$, $k = 1, ..., n$, are Independent, then

$$S_n = X_1 + ... + X_n \sim \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$$

and

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

Some Particular results about S_n and $\overline{X}_n$

The following are important cases when we can find exactly the Distribution of S_n and/or $\overline{X}_n$:

- If $X_k \sim \mathcal{N}(\mu, \sigma^2)$, $k = 1, ..., n$, are Independent, then

$$S_n = X_1 + ... + X_n \sim \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$$

and

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

Some Particular results about S_n and $\overline{X}_n$

The following are important cases when we can find exactly the Distribution of S_n and/or $\overline{X}_n$:

- If $X_k \sim \mathcal{N}(\mu, \sigma^2)$, $k = 1, ..., n$, are Independent, then

$$S_n = X_1 + ... + X_n \sim \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$$

  and

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

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

$$S_n = X_1 + ... + X_n \sim$$

Some Particular results about S_n and $\overline{X}_n$

The following are important cases when we can find exactly the Distribution of S_n and/or $\overline{X}_n$:

- If $X_k \sim \mathcal{N}(\mu, \sigma^2)$, $k = 1, ..., n$, are Independent, then

$$S_n = X_1 + ... + X_n \sim \mathcal{N}(n \cdot \mu, n \cdot \sigma^2)$$

  and

$$\overline{X}_n = \frac{X_1 + X_2 + ... + X_n}{n} \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

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

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

Some Particular results about S_n and $\overline{X}_n$

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

$$S_n = X_1 + ... + X_n \sim$$

Some Particular results about S_n and $\overline{X}_n$

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

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

Some Particular results about 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

$$S_n = X_1 + ... + X_n \sim$$

Some Particular results about 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 Weak LLN

Now, what can be said about S_n and $\overline{X}_n$ in the general case?

The Weak LLN

Now, what can be said about S_n and $\overline{X}_n$ in the general case? LLN and CLT help us in this matter, they describe the *asymptotic* properties of these guys:

The Weak LLN

Now, what can be said about S_n and $\overline{X}_n$ in the general case? LLN and CLT help us in this matter, they describe the *asymptotic* properties of these guys:

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

Now, what can be said about S_n and $\overline{X}_n$ in the general case? LLN and CLT help us in this matter, they describe the *asymptotic* properties of these guys:

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

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

i.e., 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

Now, what can be said about S_n and $\overline{X}_n$ in the general case? LLN and CLT help us in this matter, they describe the *asymptotic* properties of these guys:

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

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

i.e., 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 Strong LLN

The Strong Law of Large Numbers, SLLN, Kolmogorov

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

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

The Strong LLN

The Strong Law of Large Numbers, SLLN, Kolmogorov

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

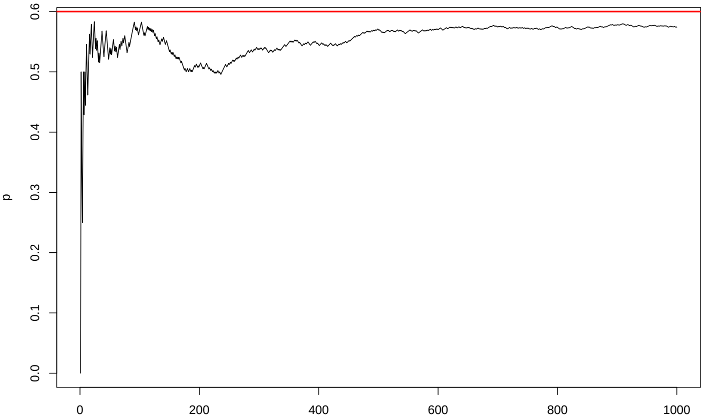
$$\frac{X_1 + X_2 + ... + X_n}{n} \xrightarrow{a.s.} \mathbb{E}(X_1), \qquad n \to +\infty,$$

that is,

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

Visualization of the LLN

```r
n <- 1000; expect <- 0.6
X <- rbinom(n, 1, expect)
S <- cumsum(X); p <- S/(1:n)
plot(p, type = "l")
abline(expect,0, col = "red", lwd = 2)
```



Supplements, LLN

Sometimes we are required to calculate limits of the form:

$$\lim_{n \to +\infty} \frac{g(X_1) + g(X_2) + ... + g(X_n)}{n}$$

in the *Probability* or *a.s.* sense, for some nice function g.

Supplements, LLN

Sometimes we are required to calculate limits of the form:

$$\lim_{n \to +\infty} \frac{g(X_1) + g(X_2) + ... + g(X_n)}{n}$$

in the *Probability* or *a.s.* sense, for some nice function g.Clearly, under the condition that $\mathbb{E}(g(X_1))$ and $Var(g(X_1))$ are finite, or $\mathbb{E}(|g(X_1)|) < +\infty$, we will have

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

Supplements, LLN

Sometimes we are required to calculate limits of the form:

$$\lim_{n \to +\infty} \frac{g(X_1) + g(X_2) + ... + g(X_n)}{n}$$

in the *Probability* or *a.s.* sense, for some nice function g.Clearly, under the condition that $\mathbb{E}(g(X_1))$ and $Var(g(X_1))$ are finite, or $\mathbb{E}(|g(X_1)|) < +\infty$, we will have

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

Say, for example,

$$\frac{X_1^2 + X_2^2 + ... + X_n^2}{n} \xrightarrow{\mathbb{P}, a.s.}$$

Supplements, LLN

Sometimes we are required to calculate limits of the form:

$$\lim_{n\to+\infty} \frac{g(X_1)+g(X_2)+...+g(X_n)}{n}$$

in the *Probability* or *a.s.* sense, for some nice function g.Clearly, under the condition that $\mathbb{E}(g(X_1))$ and $Var(g(X_1))$ are finite, or $\mathbb{E}(|g(X_1)|) < +\infty$, we will have

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

Say, for example,

$$\frac{X_1^2+X_2^2+...+X_n^2}{n} \xrightarrow{\mathbb{P},a.s.} \mathbb{E}(X_1^2), \qquad n\to+\infty.$$

CLT

The LLN says that the values of $\overline{X}_n$ are concentrated around $\mathbb{E}(X_1)$.

CLT

The LLN says that the values of $\overline{X}_n$ are concentrated around $\mathbb{E}(X_1)$. But it is not giving us an idea about how the values are distributed around that mean value.

CLT

The LLN says that the values of $\overline{X}_n$ are concentrated around $\mathbb{E}(X_1)$. But it is not giving us an idea about how the values are distributed around that mean value. CLT helps us in this regard.

CLT

The LLN says that the values of $\overline{X}_n$ are concentrated around $\mathbb{E}(X_1)$. But it is not giving us an idea about how the values are distributed around that mean value. CLT helps us in this regard.

To give the general idea of the CLT, let us use the following transform: for a r.v. X, let us denote

$$Standardize(X) = \frac{X - \mathbb{E}(X)}{\sqrt{Var(X)}} = \frac{X - \mathbb{E}(X)}{SD(X)},$$

the Standardization (normalization, scaling) of X.

CLT

The LLN says that the values of $\overline{X}_n$ are concentrated around $\mathbb{E}(X_1)$. But it is not giving us an idea about how the values are distributed around that mean value. CLT helps us in this regard.

To give the general idea of the CLT, let us use the following transform: for a r.v. X, let us denote

$$Standardize(X) = \frac{X - \mathbb{E}(X)}{\sqrt{Var(X)}} = \frac{X - \mathbb{E}(X)}{SD(X)},$$

the Standardization (normalization, scaling) of X. Clearly,

$$\mathbb{E}(Standardize(X)) = 0 \qquad \text{and} \qquad Var(Standardize(X)) = 1.$$

Basic Idea of the CLT

The basic idea of the CLT is the following: if we have a sequence of IID r.v. X_n, and we consider their sum S_n or their average $\overline{X}_n$, then

$$Standardize(S_n) \xrightarrow{D} \mathcal{N}(0,1)$$

and

$$Standardize(\overline{X}_n) \xrightarrow{D} \mathcal{N}(0,1).$$

Basic Idea of the CLT

The basic idea of the CLT is the following: if we have a sequence of IID r.v. X_n, and we consider their sum S_n or their average $\overline{X}_n$, then

$$Standardize(S_n) \xrightarrow{D} \mathcal{N}(0,1)$$

and

$$Standardize(\overline{X}_n) \xrightarrow{D} \mathcal{N}(0,1).$$

Btw, trivially,

$$Standardize(\overline{X}_n) = Standardize(S_n),$$

and these two versions of CLT are the same.

Basic Idea of the CLT

The basic idea of the CLT is the following: if we have a sequence of IID r.v. X_n, and we consider their sum S_n or their average $\overline{X}_n$, then

$$Standardize(S_n) \xrightarrow{D} \mathcal{N}(0,1)$$

and

$$Standardize(\overline{X}_n) \xrightarrow{D} \mathcal{N}(0,1).$$

Btw, trivially,

$$Standardize(\overline{X}_n) = Standardize(S_n),$$

and these two versions of CLT are the same.

So for large n, the Distribution of the $Standardize(S_n)$ or $Standardize(\overline{X}_n)$ is approximately Standard Normal.

Basic Idea of the CLT

The basic idea of the CLT is the following: if we have a sequence of IID r.v. X_n, and we consider their sum S_n or their average $\overline{X}_n$, then

$$Standardize(S_n) \xrightarrow{D} \mathcal{N}(0,1)$$

and

$$Standardize(\overline{X}_n) \xrightarrow{D} \mathcal{N}(0,1).$$

Btw, trivially,

$$Standardize(\overline{X}_n) = Standardize(S_n),$$

and these two versions of CLT are the same.

So for large n, the Distribution of the $Standardize(S_n)$ or $Standardize(\overline{X}_n)$ is approximately Standard Normal. And this **independent of the initial Distribution of X_k**!

Basic Idea of the CLT

The basic idea of the CLT is the following: if we have a sequence of IID r.v. X_n, and we consider their sum S_n or their average $\overline{X}_n$, then

$$Standardize(S_n) \xrightarrow{D} \mathcal{N}(0,1)$$

and

$$Standardize(\overline{X}_n) \xrightarrow{D} \mathcal{N}(0,1).$$

Btw, trivially,

$$Standardize(\overline{X}_n) = Standardize(S_n),$$

and these two versions of CLT are the same.

So for large n, the Distribution of the $Standardize(S_n)$ or $Standardize(\overline{X}_n)$ is approximately Standard Normal. And this **independent of the initial Distribution of** X_k!

Easy and beautiful, isn't it?

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

We Standardize S_n:

$$Standardize(S_n) =$$

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

We Standardize S_n:

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

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

We Standardize S_n:

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

Now we use

$$\mathbb{E}(S_n) =$$

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

We Standardize S_n:

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

Now we use

$$\mathbb{E}(S_n) = n \cdot \mu, \qquad Var(S_n) =$$

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

We Standardize S_n:

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

Now we use

$$\mathbb{E}(S_n) = n \cdot \mu, \qquad Var(S_n) = n \cdot \sigma^2.$$

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

We Standardize S_n:

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

Now we use

$$\mathbb{E}(S_n) = n \cdot \mu, \qquad Var(S_n) = n \cdot \sigma^2.$$

Then,

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

CLT, in the Sums form

Assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

We consider sums

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

We Standardize S_n:

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

Now we use

$$\mathbb{E}(S_n) = n \cdot \mu, \qquad Var(S_n) = n \cdot \sigma^2.$$

Then,

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

The CLT states:

$$\frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma} \xrightarrow{D} \mathcal{N}(0,1).$$

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$.

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

We Standardize $\overline{X}_n$:

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

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

We Standardize $\overline{X}_n$:

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

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

We Standardize $\overline{X}_n$:

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

We use

$$\mathbb{E}(\overline{X}_n) =$$

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

We Standardize $\overline{X}_n$:

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

We use

$$\mathbb{E}(\overline{X}_n) = \mu, \qquad Var(\overline{X}_n) =$$

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

We Standardize $\overline{X}_n$:

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

We use

$$\mathbb{E}(\overline{X}_n) = \mu, \qquad Var(\overline{X}_n) = \frac{\sigma^2}{n}.$$

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

We Standardize $\overline{X}_n$:

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

We use

$$\mathbb{E}(\overline{X}_n) = \mu, \qquad Var(\overline{X}_n) = \frac{\sigma^2}{n}.$$

Then,

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

CLT, in the Averages form

Again, assume X_n be a sequence of IID r.v. with finite expectation $\mu = \mathbb{E}(X_i)$ and variance $\sigma^2 = Var(X_i)$. We consider the Averages

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

We Standardize $\overline{X}_n$:

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

We use

$$\mathbb{E}(\overline{X}_n) = \mu, \qquad Var(\overline{X}_n) = \frac{\sigma^2}{n}.$$

Then,

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

The CLT states:

$$\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{D} \mathcal{N}(0,1).$$

Two forms of CLT

Of course, these two forms of the CLT are the same: we have

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

and

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

Two forms of CLT

Of course, these two forms of the CLT are the same: we have

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

and

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

Now,

$$\frac{S_n - n \cdot \mu}{\sqrt{n} \cdot \sigma} = \frac{n \cdot (\frac{S_n}{n} - \mu)}{\sqrt{n} \cdot \sigma} = \frac{\frac{S_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}},$$

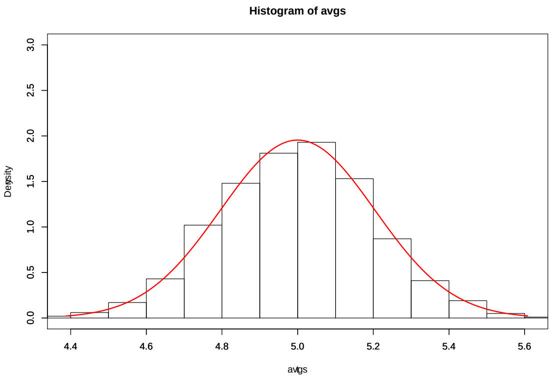
so

$$Standardize(S_n) = Standardize(\overline{X}_n).$$

Hence, the above two versions of CLT are the same, just one is in terms of S_n, the other one is in terms of $\overline{X}_n$.

CLT, Visually

```r
n <- 600 # Sample Size
m <- 1000 # no of Samples
rate <- 0.2
x <- rexp(n*m, rate = rate)
theo.mean <- 1/rate #theoretical mean
theo.sd <- 1/rate #theoretical SD
m <- matrix(x, ncol = m); d <- data.frame(m)
avgs <- sapply(d, mean)
a = theo.mean-3*theo.sd/sqrt(n); b = theo.mean+3*theo.sd/sqrt(n)
hist(avgs, freq = F, xlim = c(a, b), ylim=c(0,3))
par(new = T)
t <- seq(a,b, 0.01)
y <- dnorm(t, mean = theo.mean, sd = theo.sd/sqrt(n))
plot(t,y, type = "l", col="red", lwd = 2, , xlim = c(a,b), ylim=c(0,3))
```



CLT, Visually, v2

```r
n <- 600 # Sample Size
m <- 1000 # no of Samples
rate <- 0.2
x <- rexp(n*m, rate = rate)
m <- matrix(x, ncol = m); d <- data.frame(m)
avgs <- sapply(d, mean)
qqnorm(avgs, ylab = "Averages"); qqline(avgs)
```