Lecture 15

2024 Spring

July 4, 2024

Last Lecture

- Convergence in distribution.

$X_1 \cdots X_n$

$\downarrow \quad \downarrow$

$F_1 \quad F_n$

$\lim_{n\to\infty} F_n(x) = F(x)$, for all x. where $F(x)$ is continuous

$F(x) \to$ cdf of X

$X_n \xrightarrow{d} X$

$F(x)$ is the limiting distribution of X_n

This Lecture

- Finish Convergence in distribution.
- Convergence in probability.

(Cont') Convergence in Distribution

Examples: Continuous a sequence of random variables X_n with range $\mathbb{R}$ and cdf

$$F_n(x) = \frac{\exp(nx)}{1+\exp(nx)}, \quad x \in \mathbb{R}.$$

logistic regression
sigmoid function.

Find the limiting distribution.

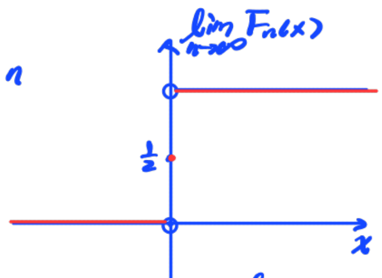
$$F_n(x) = \frac{1}{1+\exp(-nx)}$$

When $x>0$, $\lim_{n\to\infty} F_n(x) = 1$.

$x<0$, $\lim_{n\to\infty} F_n(x) = 0$

$x=0$, $F_n(0) = \frac{1}{2}$ for all n

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0 & x<0 \\ \frac{1}{2} & x=0 \\ 1 & x>0 \end{cases}$$



not a cdf.

$$F(x) = \begin{cases} 0 & x<0 \\ 1 & x\geq 0 \end{cases}$$

is a cdf.

$\Rightarrow$ $F(x)$ is the limiting distribution of X_n.

If $X \sim F(x)$, $Pr(X=0)=1$, we say X is a degenerate r.v.

$X_n \xrightarrow{d} X$ or $X_n \xrightarrow{d} 1$

Here should be 0

↳ converge in distribution to a constant.

Convergence in Probability Convergence in distribution only imposes requirements on the cdfs. It does not directly address the random variable (which is also a function).

If $X=Y$, $Pr(X\le t)=Pr(Y\le t)$ same cdf.

Example: Two random variables may be different but have the same cdf. $X \sim \text{Uniform}(0,1)$ and $Y = 1 - X$.

$X \neq Y$

$F_X(t)=Pr(X\le t)=t,\ t\in(0,1)$

$F_Y(t)=Pr(1-X\le t)=Pr(X\ge 1-t)=t,\ t\in(0,1)$

$\Rightarrow F_X(t)=F_Y(t),$ for $t\in(0,1)$

X is a r.v.

$X(w)$, w is an event. $X(w)\in\mathbb{R}$.

$X: \Omega \mapsto \mathbb{R}$

$F(\cdot)$ is a cdf

$F: \mathbb{R} \mapsto [0,1]$

Convergence in distribution only addresses convergence at the cdf level. The previous example shows that two random variables can be very different while having the same cdf. **Convergence in probability** directly addresses convergence at the random variable level.

Definition: Let $X_1, \ldots, X_n$ be a sequence of random variables, and X be a random variable. We say the sequence $X_1, \ldots, X_n$ converges to X in probability if

$$\lim_{n\to\infty} \Pr(|X_n - X| \geq \epsilon) = 0 \quad \text{or} \quad \lim_{n\to\infty} Pr(|X_n-X|<\varepsilon)=1$$

for any $\epsilon > 0$.

Notation: $X_n \xrightarrow{p} X$ as $n \to \infty$.

Comments:

1. $|X_n-X|$: difference between two rvs. (functions)

   For a given $\varepsilon>0$, $|X_n-X|\ge\varepsilon$ means X_n and X are "not close"

   $|X_n-X|<\varepsilon$. means X_n and X are "close"

   no matter how small ε is, as long as n is large.

   $Pr(X_n$ & X are close$)\to 1$.

2. We call it "in probability" because we are dealing with the limit of the probability function (not cdf).

Theorem: $X_n \xrightarrow{p} X \Rightarrow X_n \xrightarrow{d} X$

$X_n \xrightarrow{d} X \not\Rightarrow X_n \xrightarrow{p} X$

E.g. $X \sim Norm(0,1)$. $Y = -X \sim Norm(0,1)$

let $Z_1 = Y$, $Z_2 = Y$, $\cdots$ $Z_n = Y$. $\Rightarrow Z_n \xrightarrow{d} X$

$Pr(|Z_n - X| < \varepsilon) = Pr(2|X| < \varepsilon) = Pr(-\frac{\varepsilon}{2} < X < \frac{\varepsilon}{2}) \not\to 1$

$\Rightarrow Z_n \not\xrightarrow{p} X$

Convergence to a constant

Definition: Let $X_1, \ldots, X_n$ be a sequence of random variables, and b be a constant. If

$\lim_{n\to\infty} Pr(|X_n - b| < \varepsilon) = 1$. for any $\varepsilon > 0$

Then X_n converge in prob. to the constant b.

$X_n \xrightarrow{p} b$

If. $\lim_{n\to\infty} F_n(x) = \begin{cases} 0 & , x < b \\ 1 & , x > b \end{cases}$

then. X_n converge in distribution to. the constant b.

$X_n \xrightarrow{d} b$.

The limiting dist'n is $F(x) = \begin{cases} 0 & x < b \\ 1 & . x \geq b \end{cases}$

Notation:

$X_n \xrightarrow{p} b$. $X_n \xrightarrow{d} b$.

Theorem: $X_n \xrightarrow{p} b \Leftrightarrow X_n \xrightarrow{d} b$

Equivalence holds only when b is a constant.

the "$\Leftarrow$" part is not true for a non-degenerate r.v. X.

Proof & Implication: Prove $X_n \xrightarrow{P} b \iff X_n \xrightarrow{d} b$.

① For "$\Rightarrow$", we can use the theorem.

② For "$\Leftarrow$", if $X_n \xrightarrow{d} b$, then $X_n \xrightarrow{P} b$.

$$0 \le Pr(|X_n - b| \ge \varepsilon) = Pr(\{X_n \ge \varepsilon + b\} \cup \{X_n \le b - \varepsilon\})$$

b/c two events have no overlap

$$= Pr(X_n \ge b+\varepsilon) + Pr(X_n \le b-\varepsilon)$$
$$= 1 - Pr(X_n < b+\varepsilon) + \underline{Pr(X_n \le b-\varepsilon)}$$
$$\le 1 - Pr(X_n \le b + \tfrac{\varepsilon}{2}) + F_n(b-\varepsilon)$$
$$= 1 - F_n(b+\tfrac{\varepsilon}{2}) + F_n(b-\varepsilon)$$

$A_1 = \{X_n < b+\varepsilon\}$

$A_2 = \{X_n \le b + \frac{\varepsilon}{2}\}$

For any $a \in A_2$, $a \in A_1$

$\Rightarrow A_2 \subseteq A_1$

$Pr(A_2) \le Pr(A_1)$

$\underline{1 - Pr(A_2) \ge 1 - Pr(A_1)}$

$$0 \le \lim_{n\to\infty} Pr(|X_n - b| \ge \varepsilon) \le \lim_{n\to\infty}\{1 - F_n(b+\tfrac{\varepsilon}{2}) + F_n(b-\varepsilon)\}$$
$$= 1 - \lim_{n\to\infty} F_n(b+\tfrac{\varepsilon}{2}) + \lim_{n\to\infty} F_n(b-\varepsilon)$$
$$= 1 - 1 + 0 = 0$$

$$\Rightarrow \lim_{n\to\infty} Pr(|X_n - b| \ge \varepsilon) = 0 \Rightarrow X_n \xrightarrow{P} b.$$

Examples:

①. $X_1 \cdots X_n \overset{i.i.d.}{\sim} Unif(0,1)$

$X_{(1)} = \min\{X_1, \cdots, X_n\}$, $X_{(n)} = \max\{X_1 \cdots X_n\}$

From previous lecture, $X_{(1)} \xrightarrow{d} 0$ $\quad X_{(n)} \xrightarrow{d} 1$

$\Rightarrow X_{(1)} \xrightarrow{P} 0$, $X_{(n)} \xrightarrow{P} 1$.

②. $X_1 \cdots X_n$ i.i.d. with pdf. $f(x) = e^{-(x-\theta)}$, $x > \theta$.

prove $X_{(1)} = \min\{X_1 \cdots X_n\}$. $X_{(1)} \xrightarrow{P} \theta$

<u>Method 1</u>: (by definition of $\xrightarrow{P}$)

$$Pr(|X_{(1)} - \theta| \ge \varepsilon) = Pr(X_{(1)} \ge \theta+\varepsilon) + \underbrace{Pr(X_{(1)} \le \theta - \varepsilon)}_{=0}$$
$$= Pr(X_{(1)} \ge \theta + \varepsilon)$$
$$= \{Pr(X_1 \ge \theta+\varepsilon)\}^n = e^{-n\varepsilon}$$

$$\lim_{n\to\infty} Pr(|X_{(1)} - \theta| \ge \varepsilon) = \lim_{n\to\infty} e^{-n\varepsilon} = 0$$

$$\Rightarrow X_{(1)} \xrightarrow{P} \theta$$

Examples:

Method 2 (prove $X_{(1)} \xrightarrow{d} \theta$)

$$F_n(x) = Pr(X_{(1)} \le x) = \begin{cases} 0 & x \le \theta \\ 1 - Pr(X_{(1)} > x) & x > \theta \end{cases}$$

$$1 - Pr(X_{(1)} > x) = 1 - Pr(X_1 > x, \cdots, X_n > x)$$
$$= 1 - \{Pr(X_1 > x)\}^n$$
$$= 1 - e^{-n(x-\theta)}$$

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0 & , x \le \theta \\ 1 & , x > \theta \end{cases}$$ is not a cdf.

but its limiting distribution (cdf)

$$F(x) = \begin{cases} 0 & , x < \theta \\ 1 & , x \ge \theta \end{cases}$$

If $X \sim F(x)$
$Pr(X = \theta) = 1$
X is a degenerate r.v.

$\Rightarrow X_{(1)} \xrightarrow{d} \theta \Rightarrow X_{(1)} \xrightarrow{P} \theta$.

Preview

we have " $\xrightarrow{d}$ " " $\xrightarrow{P}$ "

$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. $X_1, \cdots X_n$ are iid.

①. Weak Law of Large Number (WLLN)

$\bar{X}_n \xrightarrow{P} E(X_1)$

②. Central Limit Theorem (CLT)

Some function of $\bar{X}_n \xrightarrow{d} Norm(0,1)$

③. Find out the asymptotic behavior of some functions of X_n

E.g. $X_n \xrightarrow{d} X$. what about $\underline{h(X_n)}$?

$\begin{cases} \text{Slutsky theorem} \\ \text{Delta method.} \end{cases}$