Probability-3 Lecture-11

06 September 2024 11:25

- Respect to A·N·Kolmogorov !!!

(classical) Laws of Large Numbers

$(\Omega, \mathcal{A}, P)$ - probability

$X_1, \ldots, X_n$ iid r.v.s with common finite mean μ.

∴ Law of Large Numbers (LLN):

$$\frac{X_1 + \cdots + X_n}{n} \longrightarrow \mu$$

Weak Law of Large Numbers (WLLN):

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{P} \mu$$

Strong Law of Large Numbers (SLLN):

[easier proof: done by Etemadi]

$$\frac{X_1 + \cdots + X_n}{n} \xrightarrow{a.s} \mu$$

"Real-life problem is just an abstract concept."
— Prof. AG, 6^{th} Sept. '24.

Weak Law of Numbers (WLLN):

$X_1, X_2, \ldots, X_n$ - sequence of i.i.d. r.vs with finite common mean

$(\Longleftrightarrow E|X_i| < \infty \quad \forall i)$

n^{th} partial sum, $S_n := X_1 + X_2 + \cdots + X_n$

Aim: To show: $\frac{S_n}{n} \xrightarrow{P} \mu$

Aim: To show: $\frac{S_n}{n} \xrightarrow{P} \mu$

ie, $\forall \varepsilon > 0$, $P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \longrightarrow 0$ as $n \longrightarrow \infty$

we don't have a distribution given. So, computing this isn't possible.

So, idea: to find an upper bound, & show that, that upper bound $\longrightarrow 0$.

Using Chebyshev's inequality,

$$\therefore P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{E\left|S_n - n\mu\right|}{n\varepsilon}$$

(Δ-inequality)
$$\leq \frac{E\left|X_1 - \mu\right| + E\left|X_2 - \mu\right| + \cdots + E\left|X_n - \mu\right|}{n\varepsilon}$$

$$= \frac{n \cdot E\left(\left|X_1 - \mu\right|\right)}{n\varepsilon} \quad \left[\because X_i\text{'s are i.i.d.}\right]$$

$$= \frac{E\left(\left|X_1 - \mu\right|\right)}{\varepsilon}$$

?? We are stuck !!
∴ this approach leads us nowhere.

Just to get a "feeling" of happiness, we'll assume a stronger assumption, ie, X_i's have finite 2^{nd} moment.

then, applying Chebyshev's inequality using 2^{nd} moments,

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) \leq \frac{E\left(\left|S_n - n\mu\right|^2\right)}{n^2\varepsilon^2}$$

$$= \frac{V(S_n)}{n^2\varepsilon^2}$$

$\mu = E(X_i)$
$\therefore n\mu = E(S_n)$
ie, this is variance of S_n.

$$= \frac{V(S_n)}{n^2 \varepsilon^2}$$

variance of S_n.

$$= \frac{n \cdot V(X_1)}{n^2 \varepsilon^2}$$

[$\because V(S_n) = n \cdot V(X_i)$

for this to be true, only the covariances $= 0$

$$= \frac{V(X_1)}{n \varepsilon^2} \to 0 \text{ as } n \to \infty$$

ie, **Remark:**

At this stage, only pairwise independence of the r.v.s is needed. Total independence isn't needed.]

So, we proved a slightly weaker condition than WLLN.

... back to the hypothesis:

$X_1, \ldots, X_n$ iid with finite common mean

technique: replace original seq. $X_1, \ldots, X_n$ by a new seq: $Y_1, \ldots, Y_n$; X_n's truncated appropriate.

Truncation technique:

For each $n \geqslant 1$, $Y_n := \begin{cases} |X_n| , & |X_n| \leqslant n \\ 0 , & \text{else.} \end{cases}$

Y_n - function of X_n.

So, if X_n's are inpendent, Y_n's are independent too!!

What did we lose? identicality of the distributions.

[$\because$ truncation levels are different]

ie, Y_n's are not identically distributed.

$$Y_n \overset{d}{=} |X_1| \cdot \mathbb{1}_{|X_1| \leqslant n}$$

d ← in distributed

Recall our aim: $\frac{S_n}{n} \xrightarrow{P} \mu$

$$\Longleftrightarrow \frac{S_n}{n} - \mu \xrightarrow{P} 0$$

$$\Longleftrightarrow \frac{S_n - n\mu}{n} \xrightarrow{P} 0$$

$$\Longleftrightarrow \frac{S_n - E(S_n)}{n} \xrightarrow{P} 0. \quad \text{—— } 1$$

Here, let $T_n := Y_1 + Y_2 + \cdots + Y_n$

we will prove:

$$\frac{T_n - E(T_n)}{n} \xrightarrow{P} 0 \quad \text{—— } 2$$

firstly, why doing this suffices? ie, to show: $2 \Rightarrow 1$

Observation - 1:

$$P(Y_n \neq X_n) = P(|X_n| > n)$$
$$= P(|X_1| > n) \quad \left[\because X_1, \ldots, X_n \text{ are iids.}\right]$$

$\because E|X_n| < \infty \quad \forall n,$

$$\therefore \sum_n P(Y_n \neq X_n) = \sum_n P(|X_n| > n)$$
$$= \sum_n P(|X_1| > n) < \infty$$

$\Rightarrow$ By Borel-Cantelli lemma,

$$P(Y_n \neq X_n \text{ for infinitely many } n) = 0$$

$$\Longleftrightarrow P(Y_n = X_n \text{ for all but finitely many } n) = 1$$

ie, $\exists\, n_0$ sufficiently large, s.t $\forall\, n > n_0$, this holds.

$$\Rightarrow P\left(\frac{T_n}{n} - \frac{S_n}{n} \to 0\right) = 1$$

$$\Rightarrow \frac{T_n}{n} - \frac{S_n}{n} \xrightarrow{a.s} 0$$

[Consider 2 real seq^n: a_n & b_n s.t, $\exists\, n_0$ large s.t $\forall\, n > n_0$, $a_n = b_n$. Then $\frac{n}{\cdot}$ ... $\frac{n}{\cdot}$,

$$\Rightarrow \frac{T_n}{n} - \frac{S_n}{n} \xrightarrow{a.s} 0$$

$n > n_0$, $a_n = b_n$. Then $\dfrac{\sum_1^n a_k}{n} = \dfrac{\sum_1^n b_k}{n}$

$$\Rightarrow \frac{T_n}{n} - \frac{S_n}{n} \xrightarrow{p} 0$$

this + ② $\Rightarrow \dfrac{S_n}{n} - \dfrac{E(T_n)}{n} \xrightarrow{p} 0$

Observation 2 :

$$\frac{E(T_n)}{n} = \frac{1}{n} \cdot \sum_{k=1}^{n} E(Y_k)$$

$$= \frac{1}{n} \cdot \sum_{k=1}^{n} E\left(X_k \cdot 1_{|X_k| \le k}\right)$$

$$= \frac{1}{n} \cdot \sum_{k=1}^{n} E\left(X_1 \cdot 1_{|X_1| \le k}\right)$$

Now, $E\left(X_1 \cdot 1_{|X_1| \le n}\right) \longrightarrow \mu = \dfrac{E(S_n)}{n}$

$\because X_1 \cdot 1_{|X_1| \le n} \le X_1$, & $E|X_1| = \mu < \infty$

$\therefore$ By DCT, this follows.

Cesaro mean of that

$$\therefore \frac{1}{n} \sum_{k=1}^{n} E\left(X_1 \cdot 1_{|X_1| \le k}\right) \longrightarrow \frac{E(S_n)}{n}$$

$$\therefore \frac{E(T_n)}{n} \longrightarrow \frac{E(S_n)}{n}.$$

$$\Leftrightarrow \frac{E(T_n)}{n} - \frac{E(S_n)}{n} \longrightarrow 0.$$

Note that, this has all moments finite!!

Now, finally,

$$P\left(\left|\frac{T_n - E(T_n)}{n}\right| > \varepsilon\right) = P\left(|T_n - E(T_n)| > n\varepsilon\right)$$

$$\le \frac{E|T_n - E(T_n)|^2}{n^2 \varepsilon^2}$$

(Chebyshev's inequality, and

$$\left(\begin{array}{c}\text{chebyshev's inequality,}\\ \text{wrt } 2^{nd} \text{ moment}\end{array}\right) \leq \frac{E|T_n - E(T_n)|}{n^2\varepsilon^2}$$

$$= \frac{V(T_n)}{n^2\varepsilon^2}$$

$$= \frac{\sum_{k=1}^{n} V(Y_k)}{n^2\varepsilon^2}$$

$$\leq \frac{1}{n^2\varepsilon^2} \cdot \sum_{k=1}^{n} E(Y_k^2) \quad \text{—} (*)$$

$$\leq \frac{1}{n^2\varepsilon^2} \cdot \sum_{k=1}^{n} k \cdot E|Y_k|$$

$$\leq \frac{1}{n^2\varepsilon^2} \cdot \sum_{k=1}^{n} k \cdot \underbrace{E|X_1|}_{\text{common bound for all } Y_k}$$

Again, we're stuck!! both numerator & denominator have orders of n^2.

$\therefore$ (2) $\Rightarrow$ (1)

Choose a sequence $\{a_n\}$ of +ve real nos, s.t $a_n \nearrow \infty$, but $\frac{a_n}{n} \longrightarrow 0$

(i.e, $a_n \nearrow \infty$ "slower" than $n \to \infty$. eg, $a_n = \log(n)$, $a_n = \left(\frac{2}{3}\right)^n$, etc.)

back to $(*)$:

$$P\left(\left|\frac{T_n - ET_n}{n}\right| > \varepsilon\right) \leq \frac{1}{n^2\varepsilon^2} \sum_{k=1}^{n} E(Y_k^2)$$

$$= \frac{1}{n^2\varepsilon^2} \cdot \sum_{k=1}^{n} \left(E X_1^2 \cdot 1_{|X_1| \leq k}\right)$$

$$= \frac{1}{n^2\varepsilon^2} \cdot \sum^{n} E\left(X_1^2 \cdot 1_{|X_1| < a}\right) \quad +$$

$$= \frac{1}{n^2\varepsilon^2}\cdot\sum_{k=1}^{n} E\left(X_1^2\cdot 1_{|X_1|\le a_k}\right) \quad + \quad \frac{1}{n^2\cdot\varepsilon^2}\cdot\sum_{k=1}^{n} E\left(X_1^2\cdot 1_{a_k<|X_1|\le k}\right)$$

1st term

2nd term.

1st term:

$$\frac{1}{n^2\varepsilon^2}\cdot\sum_{k=1}^{n} E\left(X_1^2\cdot 1_{|X_1|\le a_k}\right) = \frac{1}{n^2\varepsilon^2}\sum_{k=1}^{n} E\Big(X_1\cdot \underbrace{X_1\cdot 1_{|X_1|\le a_k}}_{\le a_k}\Big)$$

$$\le \frac{1}{n^2\varepsilon^2}\cdot\sum_{k=1}^{n} a_k\cdot E\left(|X_1|\right)$$

$$\longrightarrow 0. \checkmark$$

2nd term:

$$|X_1|\cdot\underbrace{|X_1|\cdot 1_{a_k<|X_1|\le k}}_{\le k\cdot 1_{a_k<|X_1|}\le n\cdot 1_{a_k<|X_1|}}$$

$$\frac{1}{n^2\varepsilon^2}\cdot\sum_{k=1}^{n} E\left(|X_1|^2\cdot 1_{a_k<|X_1|\le k}\right)$$

$$\le \frac{1}{n^{\cancel{2}}\varepsilon^2}\cdot\sum_{k=1}^{n}\cancel{n}\cdot E\left(|X_1|\cdot 1_{a_k<|X_1|}\right)$$

$$= \frac{1}{n\varepsilon^2}\cdot\sum_{k=1}^{n} E\left(|X_1|\cdot 1_{a_k<|X_1|}\right)$$

$$= \frac{1}{\varepsilon^2}\cdot\frac{\sum_{k=1}^{n} E\left(|X_1|\cdot 1_{|X_1|>a_k}\right)}{n} \longrightarrow 0.$$

Cesaro mean

$$\left[\because\ a_k\to 0\right]$$

(analysis-1 result. Suppose $x_n\to x$ $\Longleftrightarrow$ $\frac{x_1+\cdots+x_n}{n}\to x$)

$\blacksquare$