

Assignment 5

Vaibhav Chhabra
AI20BTECH11022

Download all latex codes from

<https://github.com/vaibhavchhabra25/AI1103-course/blob/main/Assignment-5/main.tex>

1 PROBLEM

(UGC/MATH 2018 (June set-a)-Q.106) Let $\{X_i\}_{i \geq 1}$ be a sequence of i.i.d. random variables with $E(X_i) = 0$ and $V(X_i) = 1$. Which of the following are true?

- 1) $\frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow 0$ in probability
- 2) $\frac{1}{n^{3/4}} \sum_{i=1}^n X_i \rightarrow 0$ in probability
- 3) $\frac{1}{n^{1/2}} \sum_{i=1}^n X_i \rightarrow 0$ in probability
- 4) $\frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow 1$ in probability

2 SOLUTION

Definition 1. (Convergence in probability)

Let X_1, X_2, \dots be an infinite sequence of random variables, and let Y be another random variable. Then the sequence $\{X_n\}$ converges in probability to Y , if for all $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \Pr(|X_n - Y| \geq \epsilon) = 0 \quad (2.0.1)$$

And we write as $n \rightarrow \infty$, $X_n \rightarrow Y$ in probability.

1)

Theorem 2.1. (Strong Law of Large Numbers)

Let X_1, X_2, \dots be a sequence of i.i.d. random variables, each having finite mean $E(X_i)$. Then for any $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} \Pr\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - E(X_i)\right| \geq \epsilon\right) = 0 \quad (2.0.2)$$

Or, $\frac{1}{n} \sum_{i=1}^n X_i$ converges in probability to $E(X_i)$.

Given,

$$E(X_i) = 0 \quad (2.0.3)$$

$$V(X_i) = 1 \quad (2.0.4)$$

Also, we know that,

$$E(X_i^2) = V(X_i) + (E(X_i))^2 \quad (2.0.5)$$

Putting given values, we get,

$$E(X_i^2) = 1 + 0^2 \quad (2.0.6)$$

$$\implies E(X_i^2) = 1 \quad (2.0.7)$$

So, $E(X_i^2)$ is finite.

Let $F_{X_i}(x)$ be the c.d.f. for the random variable X_i . As $\{X_i\}$ is sequence of i.i.d. random variables, it follows the following conditions $\forall x, x_i \in \mathbb{R}$:

$$\text{a) } F_{X_1}(x) = F_{X_2}(x) = \dots = F_{X_n}(x) = F(x)$$

$$\text{b) } F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = F(x_1)F(x_2) \dots F(x_n)$$

Let $Y_i = X_i^2$. Then for $y \geq 0$,

$$F_{Y_i}(y) = \Pr(Y_i \leq y) \quad (2.0.8)$$

$$\implies F_{Y_i}(y) = \Pr(X_i^2 \leq y) \quad (2.0.9)$$

$$\implies F_{Y_i}(y) = \Pr(-\sqrt{y} \leq X_i \leq \sqrt{y}) \quad (2.0.10)$$

$$\implies F_{Y_i}(y) = \Pr(X_i \leq \sqrt{y}) - \Pr(X_i \leq -\sqrt{y}) \quad (2.0.11)$$

$$\implies F_{Y_i}(y) = F_{X_i}(\sqrt{y}) - F_{X_i}(-\sqrt{y}) \quad (2.0.12)$$

Using (a),

$$F_{Y_i}(y) = F(\sqrt{y}) - F(-\sqrt{y}) \quad (2.0.13)$$

So, by (2.0.13),

$$F_{X_1^2}(y) = F_{X_2^2}(y) = \dots = F_{X_n^2}(y) = F_1(y) \quad (2.0.14)$$

where $F_1(y)$ is the c.d.f. of $Y_i = X_i^2$.

Now, for $y_i \geq 0$, consider

$$\begin{aligned} &F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) \\ &= \Pr(Y_1 \leq y_1, Y_2 \leq y_2, \dots, Y_n \leq y_n) \end{aligned} \quad (2.0.15)$$

$$= \Pr(X_1^2 \leq y_1, X_2^2 \leq y_2, \dots, X_n^2 \leq y_n) \quad (2.0.16)$$

$$= \Pr(-\sqrt{y_1} \leq X_1 \leq \sqrt{y_1}, -\sqrt{y_2} \leq X_2 \leq \sqrt{y_2}, \dots, -\sqrt{y_n} \leq X_n \leq \sqrt{y_n}) \quad (2.0.17)$$

Since X_1, X_2, \dots, X_n are independent,

$$\begin{aligned} F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) \\ = \Pr(-\sqrt{y_1} \leq X_1 \leq \sqrt{y_1}) \Pr(-\sqrt{y_2} \leq X_2 \leq \sqrt{y_2}) \\ \dots \Pr(-\sqrt{y_n} \leq X_n \leq \sqrt{y_n}) \end{aligned} \quad (2.0.18)$$

Using (2.0.10) and (2.0.14),

$$F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = F_{Y_1}(y_1) F_{Y_2}(y_2) \dots F_{Y_n}(y_n) \quad (2.0.19)$$

$$= F_1(y_1) F_1(y_2) \dots F_1(y_n) \quad (2.0.20)$$

So,

$$F_{X_1^2, X_2^2, \dots, X_n^2}(y_1, y_2, \dots, y_n) = F_1(y_1) F_1(y_2) \dots F_1(y_n) \quad (2.0.21)$$

By (2.0.14) and (2.0.21), $\{X_i^2\}$ must also be a sequence of i.i.d. random variables.

So, we can apply S.L.L.N. to this sequence.

Then, $\frac{1}{n} \sum_{i=1}^n X_i^2$ converges in probability to $E(X_i^2)$.

Or, $\frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow 1$ in probability.

Thus, option 1 is wrong.

2) Now, we define

$$Y_n = \frac{1}{n^{3/4}} \sum_{i=1}^n X_i \quad (2.0.22)$$

Then,

$$E(Y_n) = \frac{1}{n^{3/4}} E\left(\sum_{i=1}^n X_i\right) \quad (2.0.23)$$

$$\Rightarrow E(Y_n) = \frac{1}{n^{3/4}} (E(X_1) + E(X_2) + \dots + E(X_n)) \quad (2.0.24)$$

Using (2.0.3)

$$E(Y_n) = \frac{1}{n^{3/4}} (0) = 0 \quad (2.0.25)$$

Now,

$$V(Y_n) = V\left(\frac{1}{n^{3/4}} \sum_{i=1}^n X_i\right) \quad (2.0.26)$$

$$\Rightarrow V(Y_n) = \frac{1}{n^{3/2}} V(X_1 + X_2 + \dots + X_n) \quad (2.0.27)$$

As X_1, X_2, \dots, X_n are independent of each other,

$$V(Y_n) = \frac{1}{n^{3/2}} (V(X_1) + V(X_2) + \dots + V(X_n)) \quad (2.0.28)$$

Using (2.0.4)

$$V(Y_n) = \frac{1}{n^{3/2}} (1 + 1 + \dots + 1) = \frac{1}{n^{3/2}} \times n = \frac{1}{n^{1/2}} \quad (2.0.29)$$

Now for any $\epsilon > 0$, consider the probability

$$\Pr(|Y_n - 0| \geq \epsilon) = \Pr(|Y_n - E(Y_n)| \geq \epsilon) \quad (2.0.30)$$

Applying Chebyshev's inequality here, we get,

$$\Pr(|Y_n - 0| \geq \epsilon) \leq \frac{V(Y_n)}{\epsilon^2} = \frac{1}{n^{1/2} \epsilon^2} \quad (2.0.31)$$

So,

$$\lim_{n \rightarrow \infty} \Pr(|Y_n - 0| \geq \epsilon) \leq \lim_{n \rightarrow \infty} \frac{1}{n^{1/2} \epsilon^2} = 0 \quad (2.0.32)$$

$$\Rightarrow \lim_{n \rightarrow \infty} \Pr\left(\left|\frac{1}{n^{3/4}} \sum_{i=1}^n X_i - 0\right| \geq \epsilon\right) = 0 \quad (2.0.33)$$

So, $\frac{1}{n^{3/4}} \sum_{i=1}^n X_i \rightarrow 0$ in probability.

Thus, option 2 is correct.

3)

4) As proved for option (1), $\frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow 1$ in probability. So option 4 is correct.

So the answer must be options 2,4.