

AI 1103 - Challenging Problem 11

T. Rohan
CS20BTECH11064

Download all latex codes from

[https://github.com/rohanthota/
Challenging_problem/main.tex](https://github.com/rohanthota/Challenging_problem/main.tex)

1 PROBLEM

(UGC/MATH 2018 (June set-a)-Q.106) Let $X_{i \geq 1}$ 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 distribution)

A sequence of random variables $Y, Y_1, Y_2 \dots$ converges in distribution to a random variable Y , if

$$\lim_{n \rightarrow \infty} F_{X_n}(a) = F_X(a) \quad \forall a \in \mathbb{R}. \quad (2.0.1)$$

Definition 2. (Convergence in probability)

A sequence of random variables $Y, Y_1, Y_2 \dots$ is said to converge in probability to Y , if

$$\lim_{n \rightarrow \infty} \Pr(|Y_n - Y| > \epsilon) = 0 \quad \forall \epsilon > 0. \quad (2.0.2)$$

Lemma 2.1. If $Y_n \rightarrow Y$ in probability, $Y_n \rightarrow Y$ in distribution.

Lemma 2.2. (Strong Law of Large Numbers)

Let X_1, X_2, \dots, X_n be i.i.d. random variables with expected value $E(X_i) = \mu < \infty$, then,

$$\lim_{n \rightarrow \infty} \Pr\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mu\right| \geq \epsilon\right) = 0 \quad (2.0.3)$$

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

Lemma 2.3. If X_i is a sequence of i.i.d. random variables, satisfying condition

$$F_{X_1}(x) = F_{X_2}(x) = \dots = F_{X_n}(x) = F_X(x) \quad (2.0.4)$$

then,

$$F_{X_1^2}(x) = F_{X_2^2}(x) = \dots = F_{X_n^2}(x) = F_{X^2}(x) \quad (2.0.5)$$

$\forall x \in \mathbb{R}$ where $F_X(x)$ is the c.d.f. of X_i .

Proof. X_i is a sequence of i.i.d. random variables, which means it satisfies the following condition.

$$F_{X_1}(x) = F_{X_2}(x) = \dots = F_{X_n}(x) = F_X(x) \quad (2.0.6)$$

where $F_X(x)$ is the c.d.f. of X_i .

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

$$F_{Y_i}(y) = \Pr(Y_i \leq y) \quad (2.0.7)$$

$$\implies F_{Y_i}(y) = \Pr(X_i^2 \leq y) \quad (2.0.8)$$

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

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

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

Using (2.0.6),

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

From (2.0.12),

$$F_{Y_1}(y) = F_{Y_2}(y) = \dots = F_{Y_n}(y) = F_Y(y) \quad (2.0.13)$$

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

Lemma 2.4. If X_i is a sequence of i.i.d. random variables, satisfying condition

$$F_{X_1, \dots, X_n}(x_1 \dots x_n) = F_X(x_1)F_X(x_2) \dots F_X(x_n) \quad (2.0.14)$$

where $F_X(x)$ is the c.d.f. of X_i , then for $Y_i = X_i^2$

$$F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = F_Y(y_1)F_Y(y_1) \dots F_Y(y_n) \quad (2.0.15)$$

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

Proof. Let $Y_i = X_i^2$. Now, for $y_i \geq 0$, consider

$$F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = \Pr(Y_1 \leq y_1, Y_2 \leq y_2, \dots, Y_n \leq y_n) \quad (2.0.16)$$

$$= \Pr(X_1^2 \leq y_1, X_2^2 \leq y_2, \dots, X_n^2 \leq y_n) \quad (2.0.17)$$

$$= \Pr(-\sqrt{y_1} \leq X_1 \leq \sqrt{y_1}, -\sqrt{y_2} \leq X_2 \leq \sqrt{y_2}, \dots, -\sqrt{y_n} \leq X_n \leq \sqrt{y_n}) \quad (2.0.18)$$

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

$$F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = \Pr(-\sqrt{y_1} \leq X_1 \leq \sqrt{y_1}) \Pr(-\sqrt{y_2} \leq X_2 \leq \sqrt{y_2}) \dots \Pr(-\sqrt{y_n} \leq X_n \leq \sqrt{y_n}) \quad (2.0.19)$$

From (2.0.9) and (2.0.13),

$$F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = F_{Y_1}(y_1)F_{Y_2}(y_2) \dots F_{Y_n}(y_n) \quad (2.0.20)$$

$$= F_Y(y_1)F_Y(y_2) \dots F_Y(y_n) \quad (2.0.21)$$

So,

$$F_{Y_1, Y_2, \dots, Y_n}(y_1, y_2, \dots, y_n) = F_Y(y_1)F_Y(y_1) \dots F_Y(y_n) \quad (2.0.22)$$

□

Lemma 2.5. If X_i is a sequence of i.i.d. random variables, it follows that X_i^2 is also a sequence of i.i.d. random variables.

Proof. From Lemma 2.3 and Lemma 2.4, X_i^2 is also a sequence of i.i.d. random variables. □

Lemma 2.6. If X_1, X_2, \dots, X_n are independent random variables,

$$V\left(\sum_{i=1}^n X_i\right) = V(X_1) + V(X_2) + \dots + V(X_n) \quad (2.0.23)$$

Lemma 2.7. (Chebyshev's Inequality)

Let the random variable X have a finite mean μ and a finite variance σ^2 . For every $\epsilon > 0$,

$$\Pr(|X - \mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2} \quad (2.0.24)$$

Lemma 2.8. (Central Limit Theorem)

Let X_1, X_2, \dots, X_n be i.i.d. random variables with expected value $E(X_i) = \mu < \infty$ and $0 < V(X_i) = \sigma^2 < \infty$. Then the random variable

$$Z_n = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}\sigma} \quad (2.0.25)$$

converges in distribution to the standard normal random variable as n goes to infinity, that is

$$\lim_{n \rightarrow \infty} \Pr(Z_n \leq a) = \Phi(a) \quad \forall a \in \mathbb{R}. \quad (2.0.26)$$

where $\Phi(a)$ is the standard normal CDF.

1) From Lemma 2.5, $\{X_i^2\}$ is a sequence of i.i.d. random variables. Now, we know,

$$E(X_i^2) = V(X_i) + (E(X_i))^2 \quad (2.0.27)$$

Putting given values, we get,

$$E(X_i^2) = 1 \quad (2.0.28)$$

From Lemma 2.2,

$$\frac{1}{n} \sum_{i=1}^n X_i^2 \rightarrow 1 \quad (2.0.29)$$

Therefore, option 1 is incorrect.

2)

Lemma 2.9. If $\{X_i\}$ is a sequence of i.i.d. random variables with $E(X_i) = 0$ and $V(X_i) = 1$, then the random variable $Y_n = \frac{1}{n^{3/4}} \sum_{i=1}^n X_i$ has $E(Y_n) = 0$ and $V(Y_n) = \frac{1}{n^{1/2}}$

Proof.

$$E(Y_n) = \frac{1}{n^{3/4}} E\left(\sum_{i=1}^n X_i\right) = 0 \quad (2.0.30)$$

Since $E(X_i) = 0$. Now,

$$V(Y_n) = V\left(\frac{1}{n^{3/4}} \sum_{i=1}^n X_i\right) \quad (2.0.31)$$

$$V(Y_n) = \frac{1}{n^{3/2}} V\left(\sum_{i=1}^n X_i\right) \quad (2.0.32)$$

From Lemma 2.6, and since $V(X_i) = 1$,

$$\Rightarrow V(Y_n) = \frac{1}{n^{3/2}} \times n = \frac{1}{n^{1/2}} \quad (2.0.33)$$

□

Now, from Lemma 2.7 and from Lemma 2.9

$$\lim_{n \rightarrow \infty} \Pr(|Y_n - 0| \geq \epsilon) \leq \lim_{n \rightarrow \infty} \frac{1}{n^{1/2} \epsilon^2} (= 0) \quad (2.0.34)$$

$$\Rightarrow \lim_{n \rightarrow \infty} \Pr\left(\left|\frac{1}{n^{3/4}} \sum_{i=1}^n X_i - 0\right| \geq \epsilon\right) = 0 \quad (2.0.35)$$

From Definition 2,

$$\frac{1}{n^{3/4}} \sum_{i=1}^n X_i \rightarrow 0 \quad (2.0.36)$$

Thus, option 2 is correct.

- 3) Writing the random variable Z_n from Lemma 2.8 for $\{X_i\}$ where $\mu = 0$ and $\sigma = 1$,

$$Z_n = \frac{1}{n^{1/2}} \sum_{i=1}^n X_i \quad (2.0.37)$$

From Lemma 2.8

$$\frac{1}{n^{1/2}} \sum_{i=1}^n X_i \rightarrow N(0, 1) \quad (2.0.38)$$

whereas the option states that

$$\frac{1}{n^{1/2}} \sum_{i=1}^n X_i \rightarrow 0 \quad (2.0.39)$$

from Lemma 2.1.

Therefore, option 3 is incorrect.

- 4) From (2.0.29), option 4 is correct.

Therefore, options 2 and 4 are correct.