AI 1103 - Challenging Problem 11

T. Rohan

CS20BTECH11064

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) $\frac{1}{n} \sum_{i=1}^n X_i^2 \to 0$ in probability

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

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

4) $\frac{1}{n} \sum_{i=1}^n X_i^2 \to 1$ in probability

2 Solution

CONVERGENCE IN DISTRIBUTION: A sequence of random variables $Y, Y_1, Y_2 \ldots$ converges in distribution to a random variable Y, if

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

CONVERGENCE IN PROBABILITY: A sequence of random variables $Y, Y_1, Y_2 \ldots$ is said to converge in probability to Y, if

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

THEOREM: If $Y_n \to Y$ in probability, $Y_n \to Y$ in distribution.

STRONG LAW OF LARGE NUMBERS: Let X_1, $X_2, \ldots X_n$ be i.i.d. random variables with expected value $E(X_i) = \mu < \infty$, then,

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

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

CENTRAL LIMIT THEOREM(CLT):
Let X_1, X_2, $\ldots X_n$ 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

$$Z_n = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \quad (2.0.4)$$

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

$$\lim_{n\to\infty} \Pr(Z_n \leq a) = \Phi(a) \ \forall a \in \mathbb{R}. \quad (2.0.5)$$

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

1) 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 $\forall x, x_i \in \mathbb{R}$:

a) $F_{X_1}(x) = F_{X_2}(x) = \ldots = F_{X_n}(x) = F_X(x)$

b) $F_{X_1,\ldots X_n}(x_1 \ldots x_n) = F_X(x_1)F_X(x_2)\ldots F_X(x_n)$

where $F_X(x)$ is the c.d.f. of X_i.
Let $Y_i = X_i^2$. Then for $y \geq 0$,

$$F_{Y_i}(y) = \Pr(Y_i \leq y) \quad (2.0.6)$$

$$\implies F_{Y_i}(y) = \Pr\left(X_i^2 \leq y\right) \quad (2.0.7)$$

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

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

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

Using condition (a),

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

From (2.0.11),

$$F_{X_1^2}(y) = F_{X_2^2}(y) = \ldots = F_{X_n^2}(y) = F_Y(y) \quad (2.0.12)$$

where $F_Y(y)$ is the c.d.f. of $Y_i = X_i^2$.
Now, for $y_i \geq 0$, consider

$$F_{Y_1,Y_2,\ldots,Y_n}(y_1, y_2, \ldots, y_n)$$
$$= \Pr(Y_1 \leq y_1, Y_2 \leq y_2, \ldots, Y_n \leq y_n) \quad (2.0.13)$$

$$= \Pr\left(X_1^2 \le y_1, X_2^2 \le y_2, \ldots, X_n^2 \le y_n\right) \quad (2.0.14)$$

$$= \Pr\left(-\sqrt{y_1} \le X_1 \le \sqrt{y_1}, -\sqrt{y_2} \le X_2 \le \sqrt{y_2},\right.$$
$$\left. \ldots, -\sqrt{y_n} \le X_n \le \sqrt{y_n}\right) \quad (2.0.15)$$

Since $X_1, X_2, \ldots, X_n$ are independent,

$$F_{Y_1,Y_2,\ldots,Y_n}(y_1, y_2, \ldots, y_n) =$$
$$\Pr\left(-\sqrt{y_1} \le X_1 \le \sqrt{y_1}\right)\Pr\left(-\sqrt{y_2} \le X_2 \le \sqrt{y_2}\right)$$
$$\ldots \Pr\left(-\sqrt{y_n} \le X_n \le \sqrt{y_n}\right)$$
$$(2.0.16)$$

From (2.0.8) and (2.0.12),

$$F_{Y_1,Y_2,\ldots,Y_n}(y_1, y_2, \ldots, y_n)$$
$$= F_{Y_1}(y_1)F_{Y_2}(y_2)\ldots F_{Y_n}(y_n) \quad (2.0.17)$$
$$= F_Y(y_1)F_Y(y_2)\ldots F_Y(y_n) \quad (2.0.18)$$

So,

$$F_{X_1^2,X_2^2,\ldots,X_n^2}(y_1, y_2, \ldots, y_n)$$
$$= F_Y(y_1)F_Y(y_1)\ldots F_Y(y_n) \quad (2.0.19)$$

By (2.0.12) and (2.0.19), X_i^2 must also be a sequence of i.i.d. random variables.
We know,

$$E(X_i^2) = V(X_i) + (E(X_i))^2 \quad (2.0.20)$$

Putting given values, we get,

$$E(X_i^2) = 1 \quad (2.0.21)$$

From S.L.L.N, $\frac{1}{n}\sum_{i=1}^n X_i^2$ converges in probability to $E(X_i^2) = 1$.
Therefore, option 1 is incorrect.

2) Now, we define

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

Then,

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

$$\implies E(Y_n) = \frac{1}{n^{3/4}}\left(E(X_1) + E(X_2)\ldots E(X_n)\right) \quad (2.0.24)$$

Since $E(X_i) = 0$

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

Now,

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

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

As $X_1, X_2, \cdots X_n$ are independent of each other,

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

Since $V(X_i) = 1$

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

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

$$\Pr\left(|Y_n - 0| \ge \epsilon\right) = \Pr\left(|Y_n - E(Y_n)| \ge \epsilon\right) \quad (2.0.30)$$

Applying Chebyschev's inequality here, we get,

$$\Pr\left(|Y_n - 0| \ge \epsilon\right) \le \frac{V(Y_n)}{\epsilon^2} = \frac{1}{n^{1/2}\epsilon^2} \quad (2.0.31)$$

So,

$$\lim_{n\to\infty}\Pr\left(|Y_n - 0| \ge \epsilon\right) \le \lim_{n\to\infty}\frac{1}{n^{1/2}\epsilon^2} = 0 \quad (2.0.32)$$

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

So, $\frac{1}{n^{3/4}}\sum_{i=1}^n X_i \to 0$ in probability.
Thus, option 2 is correct.

3) The option states that $\frac{1}{n^{1/2}}\sum_{i=1}^n X_i \to 0$ in probability. This statement implies that $\frac{1}{n^{1/2}}\sum_{i=1}^n X_i \to 0$ in distribution, from the theorem.
Writing the random variable Z_n from CLT, for the i.i.d random variables,

$$Z_n = \frac{X_1 + X_2 + \ldots + X_n - n\mu}{\sqrt{n}\sigma} \quad (2.0.34)$$

$$= \frac{X_1 + X_2 + \ldots + X_n}{\sqrt{n}} \quad (2.0.35)$$

$$= \frac{1}{n^{1/2}} \sum_{i=1}^{n} X_i \qquad (2.0.36)$$

Since $\mu = 0$ and $\sigma = 1$.

According to Central limit theorem,

$$Z_n \to Z, \text{where, } Z \sim N(0, 1) \qquad (2.0.37)$$

which is not what the option states. Therefore, option 3 is incorrect.

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

Therefore, options 2 and 4 are correct.