AI1103-Assignment-4

Name: Vikhyath Sai Kothamasu
Roll Number: CS20BTECH11056

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Question

Let $\{X_j\}$ be a sequence of independent Bernoulli random variables with $\mathbb{P}(X_j = 1) = \frac{1}{4}$ and let $Y_n = \frac{1}{n} \sum_{j=1}^{n} X_j^2$. Then Y_n converges, in probability, to _____ .

Solution

A sequence of random variables $Y_1, Y_2, Y_3 \ldots$ converges, in probability, to a random variable Y if

$$\lim_{n \to \infty} \Pr(|Y_n - Y| \geq \epsilon) = 0 \quad \forall \epsilon > 0 \qquad (0.0.1)$$

Similarly, a sequence of random variables $Y_1, Y_2, Y_3 \ldots$ converges, in mean square, to a random variable Y if

$$\lim_{n \to \infty} E(|Y_n - Y|^2) = 0 \qquad (0.0.2)$$

A random variable converges, in probability, to a value if it converges, in mean square, to the same particular value by Markov's Inequality.

$$Y_n \xrightarrow{\mu_s} c \Rightarrow Y_n \xrightarrow{p} c \qquad (0.0.3)$$

Proof for (0.0.3): For any $\epsilon > 0$

$$\Pr(|Y_n - Y| \geq \epsilon) = \Pr(|Y_n - Y|^2 \geq \epsilon^2) \qquad (0.0.4)$$

$$\Pr(|Y_n - Y| \geq \epsilon) \leq \frac{E|Y_n - Y|^2}{\epsilon^2}$$
$$\text{(by Markov's Inequality)} \quad (0.0.5)$$

$$\lim_{n \to \infty} E(|Y_n - Y|^2) = 0 \qquad (0.0.6)$$

$$0 \leq \lim_{n \to \infty} \Pr(|Y_n - Y| \geq \epsilon) \leq \frac{0}{\epsilon^2} \qquad (0.0.7)$$

$$\lim_{n \to \infty} \Pr(|Y_n - Y| \geq \epsilon) = 0 \quad \forall \epsilon > 0 \qquad (0.0.8)$$

And for a random variable to converge, in mean square, the required conditions are:

$$\lim_{n \to \infty} E(Y_n) = c \text{ for some constant } c \in \mathbb{R} \qquad (0.0.9)$$

$$\lim_{n \to \infty} \sigma^2 = 0 \qquad (0.0.10)$$

Given in the question that $\{X_j\}$ is a sequence of random variables with

$$\Pr(X_j = 1) = \frac{1}{4} \qquad (0.0.11)$$

$$\Pr(X_j = 0) + \Pr(X_j = 1) = 1 \qquad (0.0.12)$$

$$\Pr(X_j = 0) = 1 - \frac{1}{4} = \frac{3}{4} \qquad (0.0.13)$$

$$X_j \in \{0, 1\} \qquad (0.0.14)$$

Since $0^2 = 0$ and $1^2 = 1$,

$$X_j^2 = X_j \quad \forall j \in \{1, 2, \ldots, n\} \qquad (0.0.15)$$

Thus,

$$Y_n = \frac{1}{n}\sum_{j=1}^{n} X_j^2 \qquad (0.0.16)$$

$$= \frac{1}{n}\sum_{j=1}^{n} X_j \qquad (0.0.17)$$

$$= \frac{X_1 + X_2 + \ldots X_n}{n} \qquad (0.0.18)$$

For $\Pr(Y_n = y)$,

$$\frac{X_1 + X_2 + \ldots X_n}{n} = y \qquad (0.0.19)$$

$$X_1 + X_2 + \ldots X_n = ny \qquad (0.0.20)$$

$$ny \in \{0, 1, 2, \ldots, n-1, n\} \qquad (0.0.21)$$

$$y \in \{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\} \qquad (0.0.22)$$

For equation (0.0.20), the number of possible combinations is

$$= {}^nC_{ny} \qquad (0.0.23)$$

Then,

$$\Pr(Y_n = y) = \sum_{x_1, x_2, \ldots, x_n=0}^{y} \Pr(X_1 = x_1, X_2 = x_2, \ldots$$

$$X_{n-1} = x_{n-1}, X_n = y - x_1 - x_2 - \ldots - x_{n-1}) \qquad (0.0.24)$$

$$\Pr(Y_n = y) = {}^nC_{ny}\left(\frac{1}{4}\right)^{ny}\left(\frac{3}{4}\right)^{n-ny} \qquad (0.0.25)$$

Let us assume

$$k = ny \qquad (0.0.26)$$

$$k \in \{0, 1, 2, \ldots, n-1, n\} \qquad (0.0.27)$$

$$E(Y_n) = \sum_{y=0}^{1} y \times \Pr(Y_n = y) \qquad (0.0.28)$$

$$= \sum_{y=0}^{1} y \times {}^nC_{ny}\left(\frac{1}{4}\right)^{ny}\left(\frac{3}{4}\right)^{n-ny} \qquad (0.0.29)$$

$$= \sum_{k=0}^{n} \frac{k}{n} \times {}^nC_k\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k} \qquad (0.0.30)$$

$$= 0 + \sum_{k=1}^{n} \frac{k}{n} \times \frac{n}{k} \times {}^{n-1}C_{k-1}\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k} \qquad (0.0.31)$$

$$= \sum_{k=1}^{n} {}^{n-1}C_{k-1}\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k} \qquad (0.0.32)$$

$$= \frac{1}{4}\sum_{k=1}^{n} {}^{n-1}C_{k-1}\left(\frac{1}{4}\right)^{k-1}\left(\frac{3}{4}\right)^{(n-1)-(k-1)} \qquad (0.0.33)$$

$$= \frac{1}{4}\sum_{j=0}^{n-1} {}^{n-1}C_j\left(\frac{1}{4}\right)^{j}\left(\frac{3}{4}\right)^{(n-1)-j} \qquad (0.0.34)$$

$$= \frac{1}{4}\left(\frac{1}{4} + \frac{3}{4}\right)^{n-1} = \frac{1}{4} \qquad (0.0.35)$$

$$\lim_{n\to\infty} E(Y_n) = \lim_{n\to\infty} \frac{1}{4} \qquad (0.0.36)$$

$$= \frac{1}{4} \qquad (0.0.37)$$

$$\sigma^2 = \sum_{y=0}^{1} (y - E(Y_n))^2 \times \Pr(Y_n = y) \qquad (0.0.38)$$

$$= \sum_{y=0}^{1} (y - \frac{1}{4})^2 \times {}^nC_{ny}\left(\frac{1}{4}\right)^{ny}\left(\frac{3}{4}\right)^{n-ny} \qquad (0.0.39)$$

$$= \sum_{k=0}^{n} (\frac{k}{n} - \frac{1}{4})^2 \times {}^nC_k\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k} \qquad (0.0.40)$$

$$\sigma^2 = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \times {}^nC_k\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k}$$

$$- \frac{1}{2}\sum_{k=0}^{n} \frac{k}{n} \times {}^nC_k\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k}$$

$$+ \frac{1}{16}\sum_{k=0}^{n} {}^nC_k\left(\frac{1}{4}\right)^{k}\left(\frac{3}{4}\right)^{n-k} \qquad (0.0.41)$$

$$\sigma^2 = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \times {}^nC_k \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} - \frac{1}{2} \times \frac{1}{4} + \frac{1}{16}$$

$$(0.0.42)$$

$$\sigma^2 = 0 + \left(\frac{1}{n}\right)^2 \times {}^nC_1 \left(\frac{1}{4}\right)^1 \left(\frac{3}{4}\right)^{n-1}$$
$$+ \sum_{k=2}^{n} \left(\frac{k}{n}\right)^2 \times {}^nC_k \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} - \frac{1}{8} \quad (0.0.43)$$

$$\sigma^2 = \frac{1}{4n} \left(\frac{3}{4}\right)^{n-1} - \frac{1}{8}$$
$$+ \sum_{k=2}^{n} \left(\frac{k}{n}\right)^2 \times \frac{n(n-1)}{k(k-1)} {}^{n-2}C_{k-2} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} \quad (0.0.44)$$

Fig. 0: The CDF distribution of Y_n for n=10

$$\lim_{n\to\infty} \sigma^2 = \lim_{n\to\infty} \left(\frac{1}{4n} \left(\frac{3}{4}\right)^{n-1} - \frac{1}{8} \right)$$
$$+ \lim_{n\to\infty} \left(\sum_{k=2}^{n} \left(\frac{k}{n}\right) \times \frac{(n-1)}{(k-1)} {}^{n-2}C_{k-2} \left(\frac{1}{4}\right)^k \left(\frac{3}{4}\right)^{n-k} \right)$$
$$(0.0.45)$$

$$\lim_{n\to\infty} \sigma^2 = 0 - \frac{1}{8} + \frac{1}{8} \quad (0.0.46)$$
$$= 0 \quad (0.0.47)$$

To summarise,

$$\lim_{n\to\infty} E(Y_n) = \frac{1}{4} \quad (0.0.48)$$
$$\lim_{n\to\infty} \sigma^2 = 0 \quad (0.0.49)$$
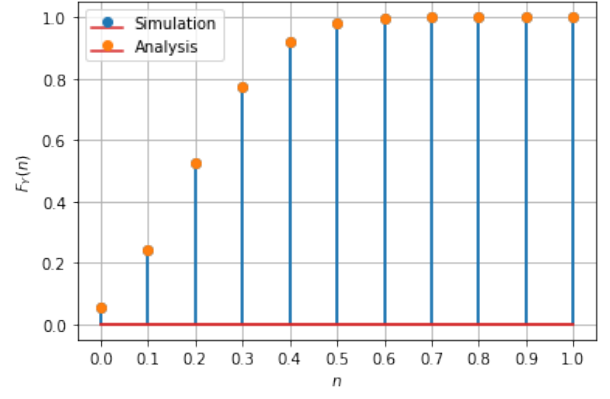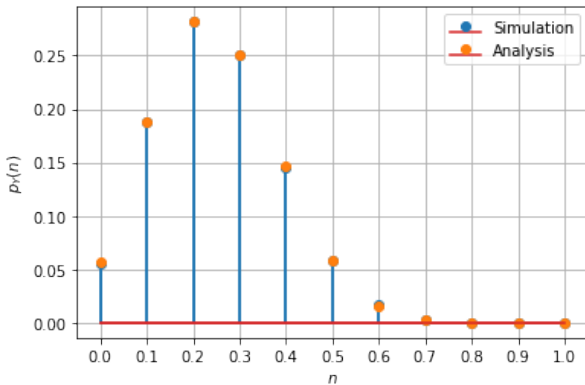
Thus, Y_n converges, in probability, to 0.25.

Fig. 0: The PMF distribution of Y_n for n=10