GATE: ST - 32.2023

EE22BTECH11039 - Pandrangi Aditya Sriram*

Question: Let $\{X_n\}_{n\geq1}$ be a sequence of independent and identically distributed random variables each having a mean 4 and variance 9. If $Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ for $n \geq 1$, then $\lim_{n\to\infty} \mathrm{E}\left[\left(\frac{Y_n-4}{\sqrt{n}}\right)^2\right]$ (in integer) equals ________. (GATE ST 2023)

Solution:

1) **Theory:** For all X_i which as i.i.d's, mean $\mu = 4$ and variance $\sigma^2 = 9$,

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{1}$$

The mean of a sum of i.i.d random variables is calculated as

$$\mathrm{E}[Y_n] = \mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \tag{2}$$

$$= \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}[X_i] \tag{3}$$

$$= \frac{1}{n}(n\mu) \tag{4}$$

$$= \mu \tag{5}$$

The variance of a sum of i.i.d random variables is calculated as

$$\mathrm{var}(Y_n) = \mathrm{E}\left[\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^2\right] - \left(\mathrm{E}\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right]\right)^2 \tag{6}$$

$$= \frac{1}{n^2}\left\{\mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] - \left(\mathrm{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2\right\} \tag{7}$$

But

$$\mathrm{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] = \mathrm{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} X_iX_j\right] \tag{8}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}\left[X_iX_j\right] \tag{9}$$

and

$$\left(\mathrm{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2 = \left(\sum_{i=1}^{n} \mathrm{E}[X_i]\right)^2 \tag{10}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \mathrm{E}[X_i]\,\mathrm{E}\left[X_j\right] \tag{11}$$

Putting (9) and (11) in (7), and using the definition of covariance,

$$\mathrm{var}(Y_n) = \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\left(\mathrm{E}\left[X_iX_j\right] - \mathrm{E}[X_i]\,\mathrm{E}\left[X_j\right]\right)\right\} \tag{12}$$

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}\mathrm{cov}\left(X_i, X_j\right)\right\} \tag{13}$$

As all the variables are i.i.d's and are thus uncorrelated,

$$\mathrm{cov}\left(X_i, X_j\right) = \begin{cases} 0 & \text{if } i \neq j \\ \mathrm{var}(X_i) & \text{if } i = j \end{cases} \tag{14}$$

Putting (14) in (13),

$$\mathrm{var}(Y_n) = \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathrm{cov}(X_i, X_i)\right) \tag{15}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n}\mathrm{var}(X_i)\right) \tag{16}$$

$$= \frac{1}{n^2}\left(\sum_{i=1}^{n}\sigma^2\right) \tag{17}$$

$$= \frac{\sigma^2}{n} \tag{18}$$

Consider the term $\left(\frac{Y_n-\mu}{\sqrt{n}}\right)^2$. Calculating its expectation,

$$\mathrm{E}\left[\left(\frac{Y_n-\mu}{\sqrt{n}}\right)^2\right] = \frac{1}{n}\mathrm{E}\left[(Y_n-\mu)^2\right] \tag{19}$$

$$= \frac{1}{n}\mathrm{var}(Y_n) \tag{20}$$

$$= \frac{\sigma^2}{n^2} \tag{21}$$

Substituting $\sigma^2 = 9$ and $\mu = 4$, we get

$$\lim_{n\to\infty} \mathrm{E}\left[\left(\frac{Y_n - 4}{\sqrt{n}}\right)^2\right] = \lim_{n\to\infty} \frac{9}{n^2} = 0 \tag{22}$$

2) **Simulation:** Any distribution with mean $\mu = 4$ and variance $\sigma^2 = 9$ can be used for the variable X_{ij} for all $i, j \in \mathbb{N}$; as shown in the Theory part, the limit is always zero regardless of the distribution. The most straightforward distribution that can be used for X_{ij} is:

$$p_{X_{ij}}(x) = \begin{cases} 0.5 & \text{if } x \in \{1, 7\} \\ 0 & \text{otherwise} \end{cases} \tag{23}$$

This distribution has the following characteristics:

$$\mu = \mathrm{E}\left[X_{ij}\right] = 0.5 \times 1 + 0.5 \times 7 = 4 \tag{24}$$

$$\sigma^2 = \mathrm{E}\left[X_{ij}^2\right] - \left(\mathrm{E}\left[X_{ij}\right]\right)^2 \tag{25}$$

$$= \left(0.5 \times 1^2 + 0.5 \times 7^2\right) - 4^2 \tag{26}$$

$$= 9 \tag{27}$$

A matrix $X_{n\times m}$ is generated for all $i \leq n$ and $j \leq m$. Using this matrix, a set of m values for Y_j is generated as

$$Y_j = \frac{1}{n} \sum_{i=1}^{n} X_{ij} \tag{28}$$

Now, the expression $\frac{(Y_j - 4)^2}{n}$ is calculated for all $j \leq m$ and their expectancy is calculated as follows:

$$\mathrm{E}\left[\left(\frac{Y_n - 4}{\sqrt{n}}\right)^2\right] = \frac{1}{m} \sum_{j=1}^{m} \frac{\left(Y_j - 4\right)^2}{n} \tag{29}$$

To calculate the limit $n \to \infty$, different values of n are taken, and the expected value is calculated (taking a fixed small value of m to reduce computational time) for each case. This output is plotted and is seen to be close to the curve $\frac{9}{n^2}$, as derived in (22). In both cases, we can observe the limit tends towards zero.

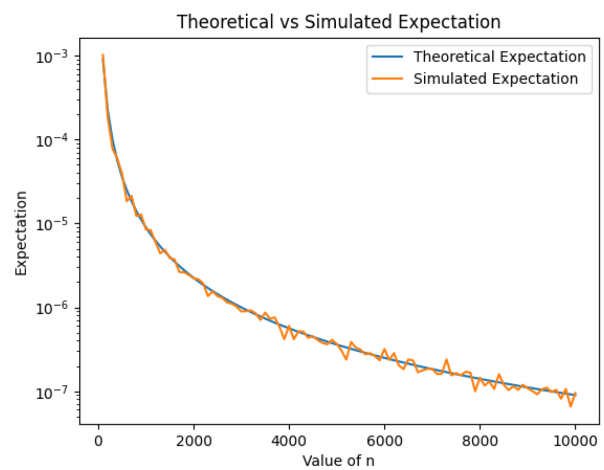

Fig. 2. Expectation vs n