MA589 Quiz III

Statistical Foundation for Data Science 08:05–08:50 IST

IIT GUWAHATI October 25, 2024

Model Answers of Quiz III

1. (5 points) Let $X_1, X_2, \ldots, X_9$ be independent and identically distributed $N(0, 5)$ random variables. Then find the expected value of $\overline{X}S^2$, where $\overline{X} = \frac{1}{9}\sum_{i=1}^{9} X_i$ and $S^2 = \frac{1}{8}\sum_{i=1}^{9}\left(X_i - \overline{X}\right)^2$.

   Solution: We know that $\overline{X} \sim N(0, \frac{5}{9})$. Thus, $E(\overline{X}) = 0$. Moreover, $\overline{X}$ and S^2 are independent random variables. Therefore,

$$E\left(\overline{X}S^2\right) = E\left(\overline{X}\right) E\left(S^2\right) = 0.$$

2. (5 points) Let (X, Y) be bivariate normal with parameters $\mu_x = 5$, $\sigma_x^2 = 1$, $\mu_y = 10$, $\sigma_y^2 = 25$, and correlation coefficient ρ, where $\rho < 0$. If it is known that the conditional probability of $Y \in (4, 16)$ given $X = 5$ is 0.954, determine the value of ρ. You may use the fact that $\Phi(2) = 0.977$.

   Solution: The conditional distribution of Y given $X = 5$ is $N(10, 25(1 - \rho^2))$. Therefore,

$$\begin{aligned}
& P\left(4 < Y < 16 | X = 5\right) = 0.954 \\
\implies & P\left(\frac{4-10}{5\sqrt{1-\rho^2}} < \frac{Y-10}{5\sqrt{1-\rho^2}} < \frac{16-10}{5\sqrt{1-\rho^2}} \middle| X = 5\right) = 0.954 \\
\implies & P\left(-\frac{6}{5\sqrt{1-\rho^2}} < Z < \frac{6}{5\sqrt{1-\rho^2}}\right) = 0.954, \quad \text{where } Z \sim N(0, 1) \\
\implies & 2\Phi\left(\frac{6}{5\sqrt{1-\rho^2}}\right) - 1 = 0.954 \\
\implies & \Phi\left(\frac{6}{5\sqrt{1-\rho^2}}\right) = 0.977 \\
\implies & \frac{6}{5\sqrt{1-\rho^2}} = 2 \\
\implies & \rho = -0.8, \quad \text{as } \rho < 0 \text{ given.}
\end{aligned}$$

3. (5 points) Let $X_n \sim Bin(n, 0.6)$ for all $n = 1, 2, 3, \ldots$. Then, for all $z \in \mathbb{R}$, show that

$$\lim_{n\to\infty} P\left(\frac{5X_n - 3n}{\sqrt{6n}} \le z\right) = \Phi(z),$$

   where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal distribution.

   Solution: Let $Y_i \overset{i.i.d.}{\sim} Bin(1, 0.6)$ for all $i = 1, 2, 3, \ldots$. Then, $X_n \overset{d}{=} \sum_{i=1}^{n} Y_i$ for all $n = 1, 2, 3, \ldots$. Moreover, $E(Y_i) = 0.6$ and $Var(Y_i) = 0.6 \times 0.4 = 0.24$ for all $i = 1, 2, \ldots$. Now, using CLT, for all $z \in \mathbb{R}$,

$$\begin{aligned}
&\lim_{n\to\infty} P\left(\sqrt{n}\frac{\overline{Y}_n - 0.6}{\sqrt{0.24}} \le z\right) = \Phi(z) \\
\implies &\lim_{n\to\infty} P\left(\sqrt{n}\,\frac{X_n/n - 0.6}{\sqrt{0.24}} \le z\right) = \Phi(z) \\
\implies &\lim_{n\to\infty} P\left(\frac{5X_n - 3n}{\sqrt{6n}} \le z\right) = \Phi(z).
\end{aligned}$$