



THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

UCS654: Predictive Analytics using Statistics

Topic
Numericals

Exercise-1

Let \mathcal{S} be a sample space with six sample points, s_1 to s_6 . The events identified on \mathcal{S} are the same as above, namely, $A = \{s_1, s_2\}$, $B = \{s_3, s_4, s_5\}$ and $C = \{s_6\}$ with $P(A) = \frac{1}{3}$, $P(B) = \frac{1}{2}$ and $P(C) = \frac{1}{6}$.

Let $Y(\cdot)$ be the transformation,

$$Y(s_i) = \begin{cases} 1, & i = 1, 2 \\ 2, & i = 3, 4, 5 \\ 3, & i = 6 \end{cases}$$

Show that $Y(\cdot)$ is a random variable by finding $F_Y(y)$. Sketch $F_Y(y)$.

Exercise-2

A random variable X has

$$F_X(x) = \begin{cases} 0 & , \quad x < 0 \\ Kx^2 & , \quad 0 \leq x \leq 10 \\ 100K & , \quad x > 10 \end{cases}$$

- (i) Find the constant K
- (ii) Evaluate $P[X \leq 5]$ and $P[5 < X \leq 7]$
- (iii) What is $f_X(X)$ =?

Exercise-2: Solution

$$\text{i) } F_x(\infty) = 100K = 1 \Rightarrow K = \frac{1}{100}.$$

$$\text{ii) } P(x \leq 5) = F_x(5) = \left(\frac{1}{100}\right) \times 25 = 0.25$$

$$P(5 < X \leq 7) = F_x(7) - F_x(5) = 0.24$$

$$f_x(x) = \frac{dF_x(x)}{dx} = \begin{cases} 0 & , \quad x < 0 \\ 0.02x & , \quad 0 \leq x \leq 10 \\ 0 & , \quad x > 10 \end{cases}$$

Exercise-3

Let $Y = \cos \pi X$, where

$$f_X(x) = \begin{cases} 1, & -\frac{1}{2} < x < \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

Let us find $E[Y]$ and σ_Y^2 . _____

Exercise-3:

Let $Y = \cos \pi X$, where

$$f_X(x) = \begin{cases} 1, & -\frac{1}{2} < x < \frac{1}{2} \\ 0, & \text{otherwise} \end{cases}$$

Let us find $E[Y]$ and σ_Y^2 .

Solution

$$E[Y] = \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos(\pi x) dx = \frac{2}{\pi} = 0.0636$$

$$E[Y^2] = \int_{-\frac{1}{2}}^{\frac{1}{2}} \cos^2(\pi x) dx = \frac{1}{2} = 0.5$$

$$\text{Hence } \sigma_Y^2 = \frac{1}{2} - \frac{4}{\pi^2} = 0.96$$



Covariance and Correlation

If X and Y are two random variables and $g(X, Y) = (X - m_X)(Y - m_Y)$

Where, $m_X = E[X]$ and $m_Y = E[Y]$

$$\lambda_{XY} = COV(X, Y) = E[(X - m_X)(Y - m_Y)]$$

Simplify above equations and find the formula for correlation.

Exercise-4

Let Y be the linear combination of the two random variables X_1 and X_2 as given below.

$$Y = k_1 X_1 + k_2 X_2$$

where, k_1 and k_2 are constants.

Let $E[X_1] = m_1$, $E[X_2] = m_2$, σ_1^2 and σ_2^2 are variances of X_1 and X_2 respectively. The correlation coefficient between X_1 and X_2 is ρ_{12} .

Compute variance of Y , $\sigma_Y^2 =$ _____.

Exercise-5

Let X be a random variable which is uniformly distributed over the interval $(0,1)$. Let Y be chosen from interval $(0,X]$ according to the pdf

$$f(y/x) = \begin{cases} 1/x, & 0 < y \leq x \\ 0, & \text{otherwise.} \end{cases}$$

Compute $E[Y^k]$ for a positive integer k .

Answer: $\frac{1}{(k+1)^2}$

Exercise-6

Consider a continuous random variable X , taking values in the range $[0,255]$ and uniformly distributed. A sample of this random variable has been taken in the form of some images (5X5) in an environment with fixed properties such as illumination, environmental noise. etc. The pixel values in each image represents a random value of X .

What will be the expected mean of each observation in the sample?

What will be the expected variance of each observation in the sample?

Exercise-7

Let $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ are the n -numbers of paired observations for two random variables (X, Y) . Assume $Y = aX^2 + b$, where a and b are constants. Find the formula to compute the values of constants a and b by using the fitting based on MSE.

Book for practice

