UNIT 2 MATHEMATICAL EXPECTATION AND DISCRETE PROBABILITY DISTRIBUTIONS

Marketed by
SIA GROUP

PART-A

SHORT QUESTIONS WITH SOLUTIONS

Q1. What is a Random variable?

Model Paper-III, Q1(c)

Answer :

A random variable is a variable which takes particular value (numerical value) and the value is determined by the result of the random experiment.

Always random variables are denoted by capital letters and the corresponding values by small letters.

For example, if a fair dice is rolled and if 'X' denotes the number obtained then 'X' is called a random variable.

Thus, 'X' can take any one of the particular values as 1, 2, 3, 4, 5 or 6, each with a probability $\frac{1}{6}$.

These values can be tabulated as follows,

X	1	2	3	4	5	6
P(X)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

All the possible outcomes of the random experiment is called as a 'sample space' and is denoted by 'S'.

Q2. What is discrete random variable and a continuous random variable?

Model Paper-II, Q1(c)

Answer :

Discrete Random Variable

If a random variable 'X' takes a finite number of values or countably infinite of values, then it is known as '*Discrete random variable*'.

Example

Number of candidates selected for a clerical post in a bank, number of movies released in a particular month, number of customers served at a e-seva centre etc.

Therefore, a discrete random variable is always a whole number, it cannot be a rational number.

Continuous Random Variable

If a random variable 'X' takes any of the possible values (fractional or integral values) at any point of time is called as continuous random variable.

Continuous random variable is known for its accuracy and is based on measurement.

Example

Height, weight, the time interval between two customers who arrived to the post office, temperature etc.

Q3. What is variance of a random variable?

Answer :

Variance of a probability distribution of a discrete random variable is given by the sum of the products of the squared deviation of expected values and all the individual values of the random variable and their respective probabilities.

Variance is denoted by Var(X) or σ^2.

It is given by,

$$\sigma^2 = E(X^2) - \mu^2$$

If 'X' is a discrete random variable then the variance is given by,

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x) \text{ and,}$$

If 'X' is continuous random variable then the variance is given by,

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty}[(x-\mu)^2] = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)\,dx$$

Q4. What is chebyshev's inequality? Model Paper-I, Q1(c)

Answer :

Chebyshev's inequality provides assurity that for a wide class of probability distribution almost all values are near to mean. According to Chebyshev's inequality if x is a random variable with mean value $\bar{x}$ and variance as σ_x^2, then for any positive integer k,

$$P\left[\left|x-\bar{x}\right| \geq k\sigma_x\right] \leq \frac{1}{k^2}$$

$$P\left[\left|x-\bar{x}\right| < k\sigma_x\right] \geq 1-\frac{1}{k^2}$$

Q5. What is Bernoulli's Distribution? Model Paper-II, Q1(d)

Answer :

Bernoulli's distribution is an experiment that can have either of the two possible outcomes (i.e., success or failure). Its probability depends only on the parameter 'p'.

Consider 'X' be a random variable that can have two outcomes i.e., success and failure.

Here,

$$X\,(\text{Success}) = 1$$

$$X\,(\text{Failure}) = 0$$

The Probability Mass Function (p.m.f) of a random variable x is given below,

$$P(X=x) = \begin{cases} p & ; \text{ for } x = 1 \text{ (success)} \\ & \text{or} \\ q \text{ or } 1-p & ; \text{ for } x = 0 \text{ (failure)} \end{cases}$$

Here,

p = Probability of success

q = Probability of failure.

Q6. What is Binomial Distribution? Model Paper-III, Q1(d)

Answer :

In an experiment consisting of 'n' repeated trials where all the trials are independent and each trial results in only two possible outcomes. The two outcomes are 'success' represented by 'p' and 'failure' represented by 'q' respectively. Therefore, the binomial density function is given by,

$$B(x, n, p) = {}^nC_x p^x (q)^{n-x}$$

Where,

$$q = 1 - p \qquad [\because p + q = 1]$$

Q7. What is Geometric distribution?

Answer :

Geometric distribution is the distribution in which number of trials needed inorder to achieve the first success to determined. This distribution is categorized into two different types based on the following criteria,

(i) Number of Failures

If 'X' is the random variable that denote number of failures needed prior to the occurrence of first success then the probability mass function is given by,

$$P(X=x) = \begin{cases} p.q^x & \text{Where } x = 0,1,2,3,... \\ 0 & \text{Otherwise} \end{cases}$$

(ii) Number of Trials

If 'X' is the random variable that denote number of trials needed prior to the occurrence of first success then the probability mass function is given by,

$$P(X=x) = \begin{cases} p.q^{x-1} & \text{Where } x = 0,1,2,3,... \\ 0 & \text{Otherwise} \end{cases}$$

Q8. What is Poisson distribution?

Answer : Model Paper-I, Q1(d)

Poisson distribution was discovered in the year 1837 by a french mathematician and physicist Simeon Denis Poisson. It is preferred for those events in which the outcomes occur at random instants of time wherein the interest lies only in the number of occurrences.

A random variable 'x' is said to have a Poisson distribution if it assumes only positive values and it is given by,

$$P(x, \lambda) \begin{cases} = \dfrac{e^{-\lambda}\lambda^x}{x!}; & x = 0, 1, 2, ..., n \\ = 0\ ; & \text{Otherwise} \end{cases}$$

Here 'λ' is known as the parameter of distribution and is greater than '0' i.e., $\lambda > 0$. The notation $X \sim P(\lambda)$ is used to denote that X is a Poisson variate with parameter 'λ'.

PART-B

ESSAY QUESTIONS WITH SOLUTIONS

2.1 MATHEMATICAL EXPECTATION

2.1.1 Mean of a Random Variable

Q9. Define the term 'Random Variable'. What are the two types of random variables? Explain.

Model Paper-I, Q4(a)

Answer :

Random Variable

A random variable is a variable which takes particular value (numerical value) and the value is determined by the result of the random experiment.

Always random variables are denoted by capital letters and the corresponding values by small letters.

For example, if a fair dice is rolled and if 'X' denotes the number obtained then 'X' is called a random variable.

Thus, 'X' can take any one of the particular values as 1, 2, 3, 4, 5 or 6, each with a probability $\frac{1}{6}$.

These values can be tabulated as follows,

X	1	2	3	4	5	6
P(X)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

All the possible outcomes of the random experiment is called as a 'sample space' and is denoted by 'S'.

Types of Random Variables

There are two types of random variables,

1. Discrete random variable
2. Continuous random variable.

1. Discrete Random Variable

If a random variable 'X' takes a finite number of values or countably infinite of values, then it is known as '*Discrete random variable*'.

Example

Number of candidates selected for a clerical post in a bank, number of movies released in a particular month, number of customers served at a e-seva centre etc.

Therefore, a discrete random variable is always a whole number, it cannot be a rational number.

2. Continuous Random Variable

If a random variable 'X' takes any of the possible values (fractional or integral values) at any point of time is called as continuous random variable.

Continuous random variable is known for its accuracy and is based on measurement.

Example

Height, weight, the time interval between two customers who arrived to the post office, temperature etc.

Q10. Explain the mean or expectation and variance of a random variable.

Answer :

Expectation of Random Variable

The arithmetic mean of a discrete random variable of the probability distribution is called as 'expectation of random variable' or 'expected value of the random variable'.

Let 'X' be the discrete random variable with the corresponding probability distribution $f(X)$, then the expected value is given by,

$$E(X) = \Sigma X.f(X)$$

$$\mu = \Sigma X.f(X)$$

Expected value is usually represented by a greek letter 'μ'.

In other words, the sum of the product of the random variables and its corresponding probabilities is called as 'expected value of the random variable'.

If a discrete random value 'X' has 'x_1', 'x_2', 'x_3' ..... 'x_n' as all the possible values with corresponding probabilities $f(x_1)$, $f(x_2)$, $f(x_3)$, ... $f(x_n)$, then the expected value is given by,

$$\mu = E(X) = \sum_x x f(x)$$

and if 'X' is continuous random variable then the expected value is given by,

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx$$

Example

Find the expected value of the variable from the given data

X	0	1	2	3
f(X)	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Solution

X	f(X)	Xf(X)
0	$\frac{1}{8}$	0
1	$\frac{3}{8}$	$\frac{3}{8}$
2	$\frac{3}{8}$	$\frac{6}{8}$
3	$\frac{1}{8}$	$\frac{3}{8}$
		$\Sigma X.f(X) = \frac{12}{8}$

Expected value, $E(X)$, $\mu = \Sigma X.f(X)$

Expected value, $\mu = \Sigma X.f(X) = \frac{12}{8}$

$\boxed{\mu = 1.5}$

PROBLEMS

Q11. A lot containing 7 components is sampled by a quality inspector, the lot contains 4 good components and 3 defective components. A sample of 3 is taken by the inspector. Find the expected value of the number of good components in this sample.

Solution :

Let 'X' be the number of good components and $E(X)$ be the expected value.

Here, x can take the values of 0,1,2,3

Given that,

Number of good components = 4

Number of defective components = 3

Probability of $X = 0$

$$P(X=0) = \frac{{}^4C_0\,{}^3C_{3-0}}{{}^7C_3} = \frac{1\times 1}{35} = \frac{1}{35}$$

Probability of $X = 1$

$$P(X=1) = \frac{{}^4C_1\,{}^3C_{3-1}}{{}^7C_3} = \frac{4\times {}^3C_2}{35} = \frac{4\times 3}{35} = \frac{12}{35}$$

Probability of $X = 2$

$$P(X=2) = \frac{{}^4C_2\,{}^3C_{3-2}}{{}^7C_3} = \frac{6\times {}^3C_1}{35} = \frac{6\times 3}{35} = \frac{18}{35}$$

Probability of $X = 3$

$$P(X=3) = \frac{{}^4C_3\,{}^3C_{3-3}}{{}^7C_3} = \frac{4\times {}^3C_0}{{}^7C_3} = \frac{4\times 1}{35} = \frac{4}{35}$$

Expected number of good components

$$E(X) = \sum x_i\, p(x=i)$$

$$= 0[p(x=0)] + 1[p(x=1)] + 2[p(x=2)] + 3\,[p(x=3)]$$

$$= 0\left(\frac{1}{35}\right) + 1\left(\frac{12}{35}\right) + 2\left(\frac{18}{35}\right) + 3\left(\frac{4}{35}\right)$$

$$= 0 + \frac{12}{35} + \frac{36}{35} + \frac{12}{35} = \frac{60}{35} = 1.7$$

$$\therefore \quad E(X) = 1.7$$

Q12. A salesperson for a medical device company has two appointments on a given day. At the first appointment, he believes that he has a 70% chance to make the deal, from which he can earn $1000 commission if successful. On the other hand, he thinks he only has a 40% chance to make the deal at the second appointment, from which, if successful, he can make $1500. What is his expected commission based on his own probability belief? Assume that the appointment results are independent of each other.

Solution :

Given that,

Number of appointments a salesperson has for a medical device company in a day = 2.

Possibility that salesperson has a chance = 70% to make a deal and earn $1000 at first appointment.

Possibility that salesperson has a chance = 40% to make a deal and earn $1500 at second appointment.

So, the salesperson can have 4 possible commission totals i.e., $0, $1000, $1500, $2500.

Now, assume that the appointment results are independent of each other and so the associated probabilities can be calculated as follows,

$$f(\$0) = \left(1-\frac{70}{100}\right)\left(1-\frac{40}{100}\right)$$

$$= (1-0.7)(1-0.4)$$

$$= (0.30)(0.60)$$

$$\boxed{f(\$0) = \$0.18}$$

$$f(\$1000) = \left(\frac{70}{100}\right)\left(1-\frac{40}{100}\right)$$

$$= (0.7)(1-0.4)$$

$$= (0.7)(0.6)$$

$$\boxed{f(\$1000) = \$0.42}$$

$$f(\$1500) = \left(1 - \frac{70}{100}\right)\left(\frac{400}{100}\right)$$

$$= (1 - 0.7)(0.4)$$

$$= (0.3)(0.4)$$

$$\boxed{f(\$1500) = \$0.12}$$

$$f(\$2500) = \left(\frac{70}{100}\right)\left(\frac{40}{100}\right)$$

$$= (0.7)(0.4)$$

$$\boxed{f(\$2500) = \$0.28}$$

The expected commission for the salesperson is calculated as shown below,

$$\because \quad \mu = E(X) = \sum_x x f(x)$$

$$E(X) = (\$0)(f(\$0)) + (\$1000)(f(\$1000) + (\$1500)(f(\$1500) + (\$2500)(f(\$2500))$$

$$E(X) = (\$0)(\$0.18) + (\$1000)(\$0.42) + (\$1500)(\$0.12) + (\$2500)(\$0.28)$$

$$E(X) = \$0 + \$420 + \$180 + \$700$$

$$\boxed{E(X) = \$1300}$$

Hence, the expected commission for the salesperson is \$1300.

Q13. Let X be the random variable that denotes the life in hours of a certain electronic device. The probability density function is

$$f(x) = \begin{cases} \dfrac{20,000}{x^3}, & x > 100, \\ 0, & \text{elsewhere} \end{cases}$$

Find the expected life of this type of device.

Solution :

Given that,

X is a random variable that denotes the life of a certain electronic device in hours.

The probability density function is,

$$f(x) = \begin{cases} \dfrac{20,000}{x^3}, & x > 100 \\ 0, & \text{elsewhere} \end{cases}$$

The mean or expected value of X is

$$\mu = E(X) = \int_{-\infty}^{\infty} x.f(x)\,dx$$

The expected life of a certain electronic device is calculated as follows,

$$E(X) = \int_{100}^{\infty} x.f(x)\,dx$$

$$= \int_{100}^{\infty} x.\frac{20,000}{x^3}\,dx$$

$$= \int_{100}^{\infty} \frac{20,000}{x^2}\,dx$$

$$= 20,000 \int_{100}^{\infty} \frac{1}{x^2}\,dx$$

$$= 20,000 \int_{100}^{\infty} x^{-2}\,dx$$

$$= 20,000 \left[\frac{x^{-2+1}}{-2+1}\right]_{100}^{\infty}$$

$$= 20,000 \left[\frac{x^{-1}}{-1}\right]_{100}^{\infty}$$

$$= 20,000 \left[\frac{-1}{x}\right]_{100}^{\infty}$$

$$= 20,000 \left[\frac{-1}{\infty} + \frac{1}{100}\right]$$

$$= 20,000 \left[0 + \frac{1}{100}\right]$$

$$= \frac{20,000}{100}$$

$$= 200$$

$$\boxed{E(x) = 200}$$

Hence, the expected life of this type of device is 200 hours.

Q14. Suppose that the number of cars X that pass through a car wash between 4:00 P.M. and 5:00 P.M. on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
P(X = x)	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Solution : Model Paper-I, Q4(b)

Let 'X' be a random variable that denotes number of cars that pass through a car wash between 4:00 P.M to 5:00 P.M on any sunny Friday.

The probability distribution for the random variable 'X' is given as,

x	4	5	6	7	8	9
P(X = x)	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Given that,

$g(X) = 2X - 1$ denotes the amount of many paid to the attendant by the manager in dollars.

The expected value of the random variable $g(X)$ is,

$$\mu_g(X) = E[g(X)] = \sum_x g(x) f(x)$$

The expected earnings of the attendant is calculated as shown below,

$$E[2X-1] = \sum_{x=4}^{9}(2x-1)f(x)$$

$$= (2(4)-1)(P(X=4)) + (2(5)-1)(P(X=5)) + (2(6)-1)(P(X=6)) + (2(7)-1)(P(X=7)) + (2(8)-1)(P(X=8)) + (2(9)-1)(P(X=9))$$

$$= 7\left(\frac{1}{12}\right) + 9\left(\frac{1}{12}\right) + 11\left(\frac{1}{4}\right) + 13\left(\frac{1}{4}\right) + 15\left(\frac{1}{6}\right) + 17\left(\frac{1}{6}\right)$$

$$= \frac{7}{12} + \frac{9}{12} + \frac{11}{4} + \frac{13}{4} + \frac{15}{6} + \frac{17}{6}$$

$$= 0.58 + 0.75 + 2.75 + 3.25 + 2.50 + 2.83$$

$$= 12.66$$

Hence, the expected earnings of the attendant for the time period (i.e, 4:00 PM to 5:00 PM) is $12.66.

Q15. Let X be a random variable with density function $f(x) = \begin{cases} \frac{x^2}{3} & -1 < x < 2 \\ 0 & \text{elsewhere} \end{cases}$. Find the expected value of $g(X) = 4X + 3$.

Solution :

Given that,

'X' is a random variable

$$f(x) = \begin{cases} \frac{x^2}{3} & -1 < x < 2 \\ 0 & \text{elsewhere} \end{cases}$$

$$g(x) = 4x + 3$$

$$E[g(x)] = E[4x+3]$$

$$= \int_{-1}^{2} \frac{(4x+3)}{3} x^2\,dx$$

$$= \frac{1}{3}\int_{-1}^{2}(4x^3 + 3x^2)\,dx$$

$$= \frac{1}{3}\left[\int_{-1}^{2} 4x^3\,dx + \int_{-1}^{2} 3x^2\,dx\right]$$

$$= \frac{1}{3}\left[4\left(\frac{x^4}{4}\right)_{-1}^{2} + 3\left(\frac{x^3}{3}\right)_{-1}^{2}\right]$$

$$= \frac{1}{3}\left[[(2)^4 - (-1)^4] + [(2)^3 - (-1)^3]\right]$$

$$= \frac{1}{3}[16 - 1 + 8 + 1]$$

$$= \frac{24}{3}$$

$$= 8$$

Therefore, the expected value of $g(x) = 4x + 3$ is 8.

Q16. Let X and Y be the random variables with joint probability distribution indicated in the table below. Find the expected value of g(X, Y) = XY.

f(x, y)		x = 0	x = 1	x = 2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{4}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Solution :

Given that,

X, Y are the random variables

$g(X, Y) = XY$

The joint probability distribution for X, Y random variables is given as shown below,

f(x, y)		x = 0	x = 1	x = 2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{4}$	$\frac{15}{28}$	$\frac{3}{28}$	1

The expected value of the random variable $g(X, Y)$ is,

$$\mu_g(X, Y) = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

The expected value for the given function is calculated as shown below,

$$E[g(X, Y)] = E(XY)$$

$$= \sum_{x=0}^{2} \sum_{y=0}^{2} xy f(x, y)$$

$$= (0)(0)f(0, 0) + (0)(1)f(0, 1) + (0)(2)f(0, 2) + (1)(0)f(1, 0) + (1)(1)f(1, 1) + (1)(2)f(1, 2) + (2)(0)f(2, 0) + (2)(1)f(2, 1) + (2)(2)f(2, 2)$$

$$= 0 + 0 + 0 + 0 + f(1, 1) + 2f(1, 2) + 0 + 2f(2, 1) + 4f(2, 2)$$

$$= \frac{3}{14} + 2(0) + 2(0) + 4(0)$$

$$= \frac{3}{14}$$

$$\boxed{E(XY) = \frac{3}{14}}$$

$\therefore$ The expected value for $E(XY)$ is $\frac{3}{14}$

Q17. Find E(Y/X) for the density function,

$$f(x, y) = \begin{cases} \dfrac{x(1+3y^2)}{4}, & 0 < x < 2, 0 < y < 1, \\ 0, & \text{elsewhere} \end{cases}$$

Solution :

Given that,

$$f(x, y) = \begin{cases} \dfrac{x(1+3y^2)}{4}, & 0 < x < 2, 0 < y < 1 \\ 0, & \text{elsewhere} \end{cases}$$

The expected value of the random variable g(X, Y) is,

$$u_g(X, Y) = E[g(X, Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx dy$$

Now, $E\left(\frac{Y}{X}\right)$ is calculated as follows,

$$E\left(\frac{Y}{X}\right) = \int_0^1 \int_0^2 y f(x, y)\, dx dy$$

$$= \int_0^1 y\left(\frac{1+3y^2}{2}\right)$$

$$= \frac{1}{2}\int_0^1 (y+3y^3)\, dy$$

$$= \frac{1}{2}\left[\int_0^1 y dy + \int_0^1 3y^3 dy\right]$$

$$= \frac{1}{2}\left[\left[\frac{y^{1+1}}{1+1}\right]_0^1 + 3\left[\frac{y^{3+1}}{3+1}\right]_0^1\right]$$

$$= \frac{1}{2}\left[\left[\frac{y^2}{2}\right]_0^1 + 3\left[\frac{y^4}{4}\right]_0^1\right]$$

$$= \frac{1}{2}\left[\left(\frac{1}{2} - \frac{0}{2}\right) + 3\left(\frac{1}{4} - \frac{0}{4}\right)\right]$$

$$= \frac{1}{2}\left[\left(\frac{1}{2}\right) + 3\left(\frac{1}{4}\right)\right]$$

$$= \frac{1}{2}\left[\left(\frac{1}{2} + \frac{3}{4}\right)\right]$$

$$= \frac{1}{2}\left(\frac{5}{4}\right)$$

$$\boxed{\therefore E\left(\frac{Y}{X}\right) = \frac{5}{8}}$$

2.1.2 Variance and Covariance of Random Variables

Q18. Explain about the variance of a random variable.

Model Paper-II, Q4(a)

Answer :

Variance of a Random Variable

Variance of a probability distribution of a discrete random variable is given by the sum of the products of the squared deviation of expected values and all the individual values of the random variable and their respective probabilities.

Variance is denoted by Var(X) or σ^2.

It is given by,

$$\sigma^2 = E(X^2) - \mu^2$$

If 'X' is discrete random variable then the variance is given by,

$$\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x) \text{ and,}$$

If 'X' is continuous random variable then the variance is given by,

$$\sigma^2 = E[(X-\mu)^2] = \int_{-\infty}^{\infty}\left[(x-\mu)^2\right] = \int_{-\infty}^{\infty}(x-\mu)^2 f(x)\,dx$$

Example

Calculate variance and standard deviation from the given data.

X	1	2	3
f(X)	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

Solution

$\text{Var}(X)(\sigma^2) = E(X^2) - \mu^2$

X	f(X)	X.f(X)	X².f(X)
1	$\frac{1}{4}$	$\frac{1}{4}$	$1^2 \times \frac{1}{4} = \frac{1}{4}$
2	$\frac{2}{4}$	$\frac{4}{4} = 1$	$2^2 \times \frac{2}{4} = 4 \times \frac{2}{4} = 2$
3	$\frac{1}{4}$	$\frac{3}{4}$	$3^2 \times \frac{3}{4} = \frac{9}{4}$
		$E(X) = \Sigma X.f(X) = 2$	$\mu^2 = \Sigma X^2.f(X) = 4.5$

$\sigma^2 = 4.5 - 2^2 = 4.5 - 4 = 0.5$

$\boxed{\sigma^2 = 0.5}$

Standard deviation, $\sigma = \sqrt{\text{Var}(X)}$

$\Rightarrow \sigma^2 = 0.5$

$\Rightarrow \sigma = \sqrt{0.5} = 0.707$

$\boxed{\sigma = 0.707}$

Q19. Define covariance. Prove that the covariance of two random variables X and Y with means μ_X and μ_Y respectively is given by, $\sigma_{XY} = E(XY) - \mu_X\mu_Y$

Answer :

Covariance

The covariance of two random variables is referred as their second order joint moment.

Let X and Y be the random variables with joint probability distribution $f(x, y)$.

The covariance of X and Y is,

$$\sigma_{XY} = E[(X-\mu_X)(Y-\mu_Y)] = \sum_x \sum_y (x-\mu_X)(y-\mu_Y) f(x, y) \text{ only if } X \text{ and } Y \text{ are discrete.}$$

Proof

Consider, $\sigma_{XY} = \sum_x \sum_y (x-\mu_X)(y-\mu_Y) f(x, y)$

$$= \sum_x \sum_y xy f(x, y) - \sum_x \sum_y y\mu_X f(x, y) - \sum_x \sum_y x\mu_Y f(x, y) + \sum_x \sum_y \mu_X\mu_Y f(x, y)$$

Since, the mean of X is $\mu_X = \sum_x x f(x, y)$ and the mean of Y is $\mu_Y = \sum_y y f(x, y)$

Also, $\sum_x \sum_y f(x, y) = 1$ for joint discrete distribution.

$$\sigma_{XY} = \sum_x \sum_y xy f(x, y) - \mu_X \sum_x \sum_y y f(x, y) - \mu_Y \sum_x \sum_y x f(x, y) + \mu_X \mu_Y \sum_x \sum_y f(x, y)$$

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y - \mu_Y\mu_X + \mu_X\mu_Y(1)$$

$$\sigma_{XY} = E(XY) - 2\mu_X\mu_Y + \mu_X\mu_Y$$

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y$$

$$\boxed{\sigma_{XY} = E(XY) - \mu_X\mu_Y}$$

Thus, the covariance of two random variables X, Y is $\sigma_{XY} = E(XY) - \mu_X\mu_Y$

Q20. The variance of a random variable X is,

$$\sigma^2 = E(X^2) - \mu^2.$$

Answer :

Let 'X' be a random variable with probability distribution $f(x)$ and mean 'μ'.

The variance of X is $\sigma^2 = E[(X-\mu)^2] = \sum_x (x-\mu)^2 f(x)$, only if x is discrete.

So, $\sigma^2 = \sum_x (x-\mu)^2 f(x)$

$$\sigma^2 = \sum_x (x^2 - 2\mu x + \mu^2) f(x) \qquad [\because (a-b)^2 = a^2 - 2ab + b^2]$$

$$\sigma^2 = \sum_x x^2 f(x) - \sum_x 2\mu x f(x) + \sum_x \mu^2 f(x)$$

Since, the mean of X, $E(X) = \mu = \sum_x x f(x)$

$$\sigma^2 = \sum_x x^2 f(x) - 2\mu \sum_x x f(x) + \mu^2 \sum_x f(x)$$

$$\sigma^2 = E(X^2) - 2\mu(\mu) + \mu^2 \qquad \left[\because \sum_x f(x) = 1\right]$$

$$\sigma^2 = E(X^2) - \mu^2$$

Thus, the variance of a random variable X is $\sigma^2 = E(X^2) - \mu^2$

PROBLEMS

Q21. Let the random variable X represent the number of automobiles that are used for official business purposes on any given workday. The probability distribution for company A is,

x	1	2	3
f(x)	0.3	0.4	0.3

and that for company B is,

x	0	1	2	3	4
f(x)	0.2	0.1	0.3	0.3	0.1

Show that the variance of the probability distribution for company B is greater than that for company A.

Solution :

Given that,

'X' is the random variable that represents the number of automobiles.

The probability distribution for company A is given as shown below,

x	1	2	3
f(x)	0.3	0.4	0.3

The mean of 'X' is given by,

$$\mu = E(X) = \sum_x x f(x)$$

The mean for company A is calculated as shown below,

$$\mu_A = E(X) = \sum_x x f(x)$$

$$\mu_A = (1)(0.3) + (2)(0.4) + (3)(0.3)$$

$$= 0.3 + 0.8 + 0.9$$

$$\boxed{\mu_A = 2}$$

∴ The mean for company A is 2

The variance of X is given by,

$$\sigma^2 = E[(X - \mu)^2] = \sum_x (x - \mu)^2 f(x)$$

The variance of random variable 'X' for company 'A' is calculated as follows,

$$\sigma_A^2 = \sum_{x=1}^{3} (x - 2)^2 f(x)$$

$$= (1 - 2)^2(0.3) + (2 - 2)^2(0.4) + (3 - 2)^2(0.3)$$

$$= (1)(0.3) + (0)(0.4) + (1)(0.3)$$

$$\boxed{\sigma_A^2 = 0.6}$$

∴ The variance for company A is 0.6

The probability distribution for company B is given as shown below,

x	0	1	2	3	4
f(x)	0.2	0.1	0.3	0.3	0.1

The mean for company B is calculated as shown below,

$$\mu_B = E(X) = \sum_x x f(x)$$

$$= (0)(0.2) + (1)(0.1) + (2)(0.3) + (3)(0.3) + (4)(0.1)$$

$$= 0 + 0.1 + 0.6 + 0.9 + 0.4$$

$$\boxed{\mu_B = 2}$$

∴ The mean for company B is 2

The variance for company B is calculated as follows,

$$\sigma_B^2 = \sum_{x=1}^{4} (x - 2)^2 + f(x)$$

$$= (0 - 2)^2(0.2) + (1 - 2)^2(0.1) + (2 - 2)^2(0.3) + (3 - 2)^2(0.3) + (4 - 2)^2(0.1)$$

$$= (4)(0.2) + (1)(0.1) + (0)(0.3) + (1)(0.3) + (4)(0.1)$$

$$= 0.8 + 0.1 + 0 + 0.3 + 0.4$$

$$\boxed{\sigma_B^2 = 1.6}$$

∴ The variance of company B is 1.6.

Hence, the variance of the probability distribution for company B i.e., $\sigma_B^2 = 1.6$ is greater than company A i.e., $\sigma_A^2 = 0.6$.

Q22. Let the random variable X represent the number of defective parts for a machine when 3 parts are sampled from a production line and tested. The following is the probability distribution of X.

x	0	1	2	3
f(x)	0.51	0.38	0.10	0.01

Calculate σ^2.

Solution :

Given that,

'X' is a random variable that represents the number of defective parts

The probability distribution of X is given as shown below,

x	0	1	2	3
f(x)	0.51	0.38	0.10	0.01

The variance of a random variable X is given by,

$$\sigma^2 = E(X^2) - \mu^2$$

The mean of X is, $\mu = E(X) = \sum_x x\, f(x)$

$$\therefore \quad \mu = (0)(0.51) + (1)(0.38) + (2)(0.10) + (3)(0.01)$$

$$\boxed{\mu = 0.61}$$

$$E(X^2) = \sum_x x^2 f(x)$$

$$= (0)^2(0.51) + (1)^2(0.38) + (2)^2(0.10) + (3)^2(0.01)$$

$$= (0)(0.51) + (1)(0.38) + (4)(0.10) + (9)(0.01)$$

$$= 0 + 0.38 + 0.40 + 0.09$$

$$\boxed{E(x^2) = 0.87}$$

Variance, $\sigma^2 = E(X^2) - \mu^2$

$$= 0.87 - (0.61)^2$$

$$= 0.87 - 0.37$$

$$\boxed{\sigma^2 = 0.50}$$

$\therefore$ The variance $\sigma^2 = 0.50$

Q23. The weekly demand for a drinking-water product, in thousands of liters, from a local chain of efficiency stores is a continuous random variable X having the probability density

$$f(x) = \begin{cases} 2(x-1), & 1<x<2, \\ 0, & \text{elsewhere} \end{cases}$$

Find the mean and variance of X.

Solution :

Given that,

'X' is a continuous random variable that denotes the weekly demand for drinking-water product.

The probability density function is,

$$f(x) = \begin{cases} 2(x-1), & 1<x<2 \\ 0, & \text{elsewhere} \end{cases}$$

The mean of X is,

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)\,dx$$

$$= \int_1^2 x.2(x-1)\,dx$$

$$= 2\int_1^2 x(x-1)\,dx$$

$$= 2\int_1^2 (x^2 - x)\,dx$$

$$= 2\left(\int_1^2 x^2 dx - \int_1^2 x dx\right)$$

$$= 2\left(\left[\frac{x^{2+1}}{2+1}\right]_1^2 - \left[\frac{x^{1+1}}{1+1}\right]_1^2\right)$$

$$= 2\left(\left[\frac{x^3}{3}\right]_1^2 - \left[\frac{x^2}{2}\right]_1^2\right)$$

$$= 2\left[\left(\frac{8}{3}\right) - \left(\frac{1}{3}\right) - \left(\frac{4}{2}\right) + \left(\frac{1}{2}\right)\right]$$

$$= 2\left[\frac{5}{6}\right]$$

$$\boxed{\mu = \frac{5}{3}}$$

$\therefore$ The mean for continuous random variable 'X' is $\frac{5}{3}$

$$E(X^2) = \int_1^2 x^2.2(x-1)\,dx$$

$$= 2\int_1^2 x^2(x-1)\,dx$$

$$= 2\int_1^2 (x^3 - x^2)\,dx$$

$$= 2\left[\int_1^2 x^3 dx - \int_1^2 x^2 .dx\right]$$

$$= 2\left(\left[\frac{x^{3+1}}{3+1}\right]_1^2 - \left[\frac{x^{2+1}}{2+1}\right]_1^2\right)$$

$$= 2\left(\left[\frac{x^4}{4}\right]_1^2 - \left[\frac{x^3}{3}\right]_1^2\right)$$

$$= 2\left[\left(\frac{16}{14}\right) - \left(\frac{1}{4}\right) - \left(\frac{8}{3}\right) + \left(\frac{1}{3}\right)\right]$$

$$= 2\left[\frac{17}{12}\right]$$

$$\boxed{E(X^2) = \frac{17}{16}}$$

Variance, $\sigma^2 = E(X^2) - \mu^2$

$$= \frac{17}{16} - \left(\frac{5}{3}\right)^2$$

$$= \frac{17}{16} - \frac{25}{9}$$

$$\boxed{\sigma^2 = \frac{1}{18}}$$

$\therefore$ The variance for continuous random variable 'X' is $\frac{1}{18}$

Q24. Calculate the variance of g(X) = 2X + 3, where X is a random variable with probability distribution

x	0	1	2	3
f(x)	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

Solution :

Given that,

'X' is a random variable and variance of

$g(X) = 2X + 3$

The probability distribution of X is given as,

x	0	1	2	3
f(x)	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{8}$

The mean of the random variable $g(x)$ is

$$\mu_{g(x)} = E[g(X)] = \sum_x g(x) f(x)$$

$$\therefore \quad \mu_{2X+3} = E(2X+3)$$

$$= \sum_{x=0}^{3} (2x+3) f(x)$$

$$= (2(0)+3)\left(\frac{1}{4}\right) + (2(1)+3)\left(\frac{1}{8}\right) + (2(2)+3\left(\frac{1}{2}\right) + (2(3)+3)\left(\frac{1}{8}\right)$$

$$= (3)\left(\frac{1}{4}\right) + (5)\left(\frac{1}{8}\right) + (7)\left(\frac{1}{2}\right) + (9)\left(\frac{1}{8}\right)$$

$$= \frac{3}{4} + \frac{5}{8} + \frac{7}{2} + \frac{9}{8}$$

$$= 6$$

$$\boxed{\mu_{2X+3} = 6}$$

The variance of the random variable $g(X)$ is,

$$\sigma^2 g(X) = E\{[g(x) - \mu_{g(x)}]^2\} = \sum_x [g(x) - \mu_{g(x)}]^2 f(x)$$

$$\sigma^2_{2X+3} = E\{[(2X+3) - \mu_{(2X+3)}]^2\}$$

$$= E[(2X+3-6)^2]$$

$$= E[(2X-3)^2]$$

$$= E[(2X)^2 - 2(2X)(3) + (3)^2]$$

$$= E[4X^2 - 12X + 9]$$

$$= \sum_{x=0}^{3} (4x^2 - 12x + 9) f(X)$$

$$= (4(0)^2 - 12(0) + 9)(f(0)) + (4(1)^2 - 12(1) + 9)(f(1)) + (4(2)^2 - 12(2) + 9)(f(2)) + (4(3)^2 - 12(3) + 9)(f(3))$$

$$= (9)\left(\frac{1}{4}\right) + (4 - 12 + 9)\left(\frac{1}{8}\right) + (16 - 24 + 9)\left(\frac{1}{2}\right) + (36 - 36 + 9)\left(\frac{1}{8}\right)$$

$$= \frac{9}{4} + \frac{1}{8} + \frac{1}{2} + \frac{9}{8}$$

$$= 4$$

$$\boxed{\sigma^2_{2X+3} = 4}$$

$\therefore$ The variance is 4

Q25. Let X be a random variable with density function $f(x) = \begin{cases} \frac{x^2}{3} & ; -1 < x < 2 \\ 0 & ; \text{elsewhere} \end{cases}$.

Find the variance of the random variable g(X) = 4X + 3.

Model Paper-II, Q4(b)

Solution :

Given that,

'X' is a random variable

$$f(x) = \begin{cases} \frac{x^2}{3} & ; -1 < x < 2 \\ 0 & ; \text{elsewhere} \end{cases}$$

$$g(x) = 4x + 3$$

$$E[g(x)] = E[4x + 3]$$

$$= \int_{-1}^{2} \frac{(4x+3)}{3} x^2 \, dx$$

$$= \frac{1}{3} \int_{-1}^{2} (4x^3 + 3x^2) \, dx$$

$$= \frac{1}{3} \left[\int_{-1}^{2} 4x^3 \, dx + \int_{-1}^{2} 3x^2 \, dx \right]$$

$$= \frac{1}{3} \left[4\left(\frac{x^4}{4}\right)_{-1}^{2} + 3\left(\frac{x^3}{3}\right)_{-1}^{2} \right]$$

$$= \frac{1}{3} [[(2)^4 - (-1)^4] + [(2)^3 - (-1)^3]]$$

$$= \frac{1}{3} [16 - 1 + 8 + 1]$$

$$= \frac{24}{3} = 8$$

The variance of the random variable $g(x)$ is,

$$\sigma^2_{g(x)} = E\{[g(x) - \mu_{g(x)^2}]\} = \int_{-\infty}^{\infty} [g(x) - \mu_{g(x)}]^2 f(x) \, dx$$

$$\sigma^2_{4X+3} = E\{[(4X+3) - \mu_{(4X+3)}]^2\}$$

$$= E[(4X + 3 - 8)^2]$$

$$= E[(4X - 5)^2]$$

$$= \int_{-1}^{2} (4x-5)^2 \frac{x^2}{3} \, dx$$

$$= \frac{1}{3} \int_{-1}^{2} [(4x)^2 - 2(4x)(5) + (5)^2] x^2 \, dx$$

$$= \frac{1}{3} \int_{-1}^{2} (16x^2 - 40x + 25) x^2 \, dx$$

$$= \frac{1}{3} \int_{-1}^{2} (16x^4 - 40x^3 + 25x^2) \, dx$$

$$= \frac{1}{3} \left[\int_{-1}^{2} 16x^4 \, dx - \int_{-1}^{2} 40x^3 \, dx + \int_{-1}^{2} 25x^2 \, dx \right]$$

$$= \frac{1}{3} \left[16 \int_{-1}^{2} x^4 \, dx - 40 \int_{-1}^{2} x^3 \, dx + 25 \int_{-1}^{2} x^2 \, dx \right]$$

$$= \frac{1}{3} \left(16 \left[\frac{x^{4+1}}{4+1} \right]_{-1}^{2} - 40 \left[\frac{x^{3+1}}{3+1} \right]_{-1}^{2} + 25 \left[\frac{x^{2+1}}{2+1} \right]_{-1}^{2} \right)$$

$$= \frac{1}{3}\left(16\left[\frac{x^5}{5}\right]_{-1}^{2} - 40\left[\frac{x^4}{4}\right]_{-1}^{2} + 25\left[\frac{x^3}{3}\right]_{-1}^{2}\right)$$

$$= \frac{1}{3}\left(16\left[\frac{32}{5} - \frac{(-1)}{5}\right] - 40\left[\frac{16}{4} - \frac{1}{4}\right] + 25\left[\frac{8}{3} - \frac{-1}{3}\right]\right)$$

$$= \frac{1}{3}\left(16\left(\frac{33}{5}\right) - 40\left(\frac{15}{4}\right) + 25\left(\frac{9}{3}\right)\right)$$

$$= \frac{1}{3}\left(\frac{528}{5} - \frac{600}{4} + \frac{225}{3}\right)$$

$$= \frac{1}{3}\left(\frac{153}{5}\right)$$

$$= \frac{51}{5}$$

$$\boxed{\sigma^2_{4X+3} = \frac{51}{5}}$$

$\therefore$ The variance is $\frac{51}{5}$

Q26. Let the number of blue refills be X and the number of red refills be Y. Two refills for a ballpoint pen are selected at random from a certain box, and the following is the joint probability distribution.

f(x, y)		x = 0	x = 1	x = 2	h(y)
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
	g(x)	$\frac{5}{14}$	$\frac{15}{28}$	$\frac{3}{28}$	1

Find the covariance of X and Y, and also find the correlation coefficient between X and Y.

Solution :

Given that,

X, Y are the random variables

$g(X, Y) = XY$

The joint probability distribution for X, Y random variables is given as shown below,

f(x, y)		x = 0	x = 1	x = 2	Row Totals
y	0	$\frac{3}{28}$	$\frac{9}{28}$	$\frac{3}{28}$	$\frac{15}{28}$
	1	$\frac{3}{14}$	$\frac{3}{14}$	0	$\frac{3}{7}$
	2	$\frac{1}{28}$	0	0	$\frac{1}{28}$
Column Totals		$\frac{5}{4}$	$\frac{15}{28}$	$\frac{3}{28}$	1

The expected value of the random variable $g(X, Y)$ is,

$$\mu_g(X, Y) = E[g(X, Y)] = \sum_x \sum_y g(x, y) f(x, y)$$

The expected value for the given function is calculated as shown below,

$$E[g(X, Y)] = E(XY)$$

$$= \sum_{x=0}^{2}\sum_{y=0}^{2} xyf(x, y)$$

$$= (0)(0)f(0, 0) + (0)(1)f(0, 1) + (0)(2)f(0, 2) + (1)(0)f(1, 0) + (1)(1)f(1, 1) + (1)(2)f(1, 2) + (2)(0)f(2, 0) + (2)(1)f(2, 1) + (2)(2)f(2, 2)$$

$$= 0 + 0 + 0 + 0\,f(1, 1) + 2f(1, 2) + 0 + 2f(2, 1) + 4f(2, 2)$$

$$= \frac{3}{14} + 2(0) + 2(0) + 4(0)$$

$$= \frac{3}{14}$$

$$\boxed{E(XY) = \frac{3}{14}}$$

$\therefore$ The expected value for $E(XY)$ is $\frac{3}{14}$

The mean for random variable 'X' is given by,

$$\mu_X = \sum_x xg(x)$$

$$\mu_X = \sum_{x=0}^{2} xg(x)$$

$$= (0)\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + (2)\left(\frac{3}{28}\right)$$

$$= 0 + \frac{15}{28} + \frac{6}{28}$$

$$\boxed{\mu_X = \frac{3}{4}}$$

The mean for random variable 'Y' is

$$\mu_Y = \sum_Y yh(y)$$

$$= \sum_{y=0}^{2} yh(y)$$

$$= (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (2)\left(\frac{1}{28}\right)$$

$$= 0 + \frac{3}{7} + \frac{2}{28}$$

$$\boxed{\mu_Y = \frac{1}{2}}$$

The covariance of two random variables X, Y is given by,

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y$$

$$= \frac{3}{14} - \left(\frac{3}{4}\right)\left(\frac{1}{2}\right)$$

$$= \frac{3}{14} - \frac{3}{8}$$

$$\boxed{\sigma_{XY} = -\frac{9}{56}}$$

$$E(X^2) = \sum_x x^2 g(x)$$

$$= \sum_{x=0}^{2} x^2 g(x)$$

$$= (0)^2 (g(0) + (1)^2 (g(1)) + (2)^2 (g(2))$$

$$= (0)^2\left(\frac{5}{14}\right) + (1)\left(\frac{15}{28}\right) + 4\left(\frac{3}{28}\right)$$

$$= 0 + \frac{15}{28} + \frac{12}{28} = \frac{27}{28}$$

$$\boxed{E(X^2) = \frac{27}{28}}$$

Similarly,

$$E(Y^2) = \sum_y y^2 h(y)$$

$$= \sum_{y=0}^{2} y^2 h(y)$$

$$= (0)^2(h(1)) + (1)^2(h(2)) + (2)^2(h(3))$$

$$= (0)\left(\frac{15}{28}\right) + (1)\left(\frac{3}{7}\right) + (4)\left(\frac{1}{28}\right)$$

$$= 0 + \frac{3}{7} + \frac{4}{28} = \frac{4}{7}$$

$$\boxed{E(Y^2) = \frac{4}{7}}$$

The variance of random variable 'X' is,

$$\sigma_X^2 = E(X^2) - \mu_X^2$$

$$= \frac{27}{28} - \left(\frac{3}{4}\right)^2$$

$$= \frac{27}{28} - \frac{9}{16}$$

$$\boxed{\sigma_X^2 = \frac{45}{112}}$$

The variance of random variable 'Y' is,

$$\sigma_Y^2 = E(Y^2) - \mu_Y^2$$

$$= \frac{4}{7} - \left(\frac{1}{2}\right)^2$$

$$= \frac{4}{7} - \frac{1}{4}$$

$$\boxed{\sigma_Y^2 = \frac{9}{28}}$$

The correlation coefficient of X and Y is,

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

$$= \frac{\frac{-9}{56}}{\sqrt{\frac{45}{112}}\sqrt{\frac{9}{28}}}$$

$$= \frac{\frac{-9}{56}}{\frac{9\sqrt{5}}{56}} = \frac{-1}{\sqrt{5}}$$

$$\boxed{\rho_{XY} = \frac{-1}{\sqrt{5}}}$$

$\therefore$ The correlation coefficient between X and Y is $\frac{-1}{\sqrt{5}}$

Q27. The fraction X of male runners and the fraction Y of female runners who compete in marathon races are described by the joint density function.

$$f(x, y) = \begin{cases} 8xy, & 0 \le y \le x \le 1, \\ 0, & \text{elsewhere} \end{cases}$$

Find the covariance of X and Y and also find the correlation coefficient of X and Y?

Solution :

Given that,

X is the fraction of male runners

Y is the fraction of female runners

$$f(x, y) = \begin{cases} 8xy, & 0 \le y \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

The following are the marginal density functions,

$$g(x) = \begin{cases} 4x^3, & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

$$h(y) = \begin{cases} 4y(1-y^2), & 0 \le y \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

Now, $\mu_X = E(X) = \int_0^1 x.4x^3 dx$

$$= \int_0^1 4x^4 dx$$

$$= 4\int_0^1 x^4 dx = 4\left[\frac{x^{4+1}}{4+1}\right]_0^1$$

$$= 4\left[\frac{x^5}{5}\right]_0^1$$

$$= \left[\frac{1}{5} - \frac{0}{5}\right]$$

$$= 4\left[\frac{1}{5}\right] = \frac{4}{5}$$

$$\boxed{\therefore\ \mu_X = \frac{4}{5}}$$

Similarly, $\mu_Y = \int_0^1 y.4y(1-y^2)dy$

$$= \int_0^1 4y^2(1-y^2)dy$$

$$= 4\int_0^1 (y^2 - y^4)dy$$

$$= 4\left[\int_0^1 y^2 dy - \int_0^1 y^4 dy\right]$$

$$= 4\left[\left[\frac{y^{2+1}}{2+1}\right]_0^1 - \left[\frac{y^{4+1}}{4+1}\right]_0^1\right] = 4\left[\left[\frac{y^3}{3}\right]_0^1 - \left[\frac{y^5}{5}\right]_0^1\right]$$

$$= 4\left[\left[\frac{1}{3} - \frac{0}{3}\right] - \left[\frac{1}{5} - \frac{0}{5}\right]\right]$$

$$= 4\left[\left[\frac{1}{3}\right] - \left[\frac{1}{5}\right]\right]$$

$$= 4\left(\frac{2}{15}\right) = \frac{8}{15}$$

$$\boxed{\therefore\ \mu_Y = \frac{8}{15}}$$

The joint density function is,

$$E(XY) = \int_0^1 \int_0^1 x^2y^2\,dx\,dy$$

$$= 8\int_0^1 \int_0^1 x^2y^2\,dx\,dy$$

$$= 8\int_0^1 x^2dx \int_0^1 y^2dy$$

$$= 8\left[\left[\frac{x^{2+1}}{2+1}\right]_0^1 \left[\frac{y^{2+1}}{2+1}\right]_0^1\right]$$

$$= 8\left[\left[\frac{x^3}{3}\right]_0^1 \left[\frac{y^3}{3}\right]_0^1\right]$$

$$= 8\left[\left[\frac{1}{3}-\frac{0}{3}\right]\left[\frac{1}{3}-\frac{0}{3}\right]\right]$$

$$= 8\left(\frac{1}{3}\right)\left(\frac{1}{3}\right)$$

$$E(XY) = \frac{8}{9}$$

The covariance σ_{XY} is,

$$\sigma_{XY} = E(XY) - \mu_X\mu_Y$$

$$= \frac{8}{9} - \left(\frac{4}{5}\right)\left(\frac{8}{15}\right)$$

$$= \frac{8}{9} - \frac{32}{75}$$

$$= \frac{104}{225}$$

$$\boxed{\sigma_{XY} = \frac{104}{225}}$$

$$E(X^2) = \int_0^1 x^2.g(x)\,dx$$

$$= \int_0^1 x^2.4x^3dx$$

$$= \int_0^1 4x^5dx$$

$$= 4\int_0^1 x^5dx$$

$$= 4\left[\frac{x^{5+1}}{5+1}\right]_0^1$$

$$= 4\left[\frac{x^6}{6}\right]_0^1$$

$$= 4\left[\frac{1}{6}-0\right]$$

$$= \frac{4}{6}$$

$$= \frac{2}{3}$$

$$\boxed{E(X^2) = \frac{2}{3}}$$

Similarly, $E(Y^2) = \int_0^1 y^2h(y)\,dy$

$$= \int_0^1 y^2 4y(1-y^2)\,dy$$

$$= 4\int_0^1 y^2y(1-y^2)\,dy$$

$$= 4\int_0^1 y^2(y-y^3)\,dy$$

$$= 4\int_0^1 (y^3-y^5)\,dy$$

$$= 4\left[\int_0^1 y^3dy - \int_0^1 y^5dy\right]$$

$$= 4\left[\left[\frac{y^{3+1}}{3+1}\right]_0^1 - \left[\frac{y^{5+1}}{5+1}\right]_0^1\right]$$

$$= 4\left[\left[\frac{y^4}{4}\right]_0^1 - \left[\frac{y^6}{6}\right]_0^1\right]$$

$$= 4\left[\left[\frac{1}{4}-0\right]-\left[\frac{1}{6}-0\right]\right]$$

$$= 4\left(\frac{1}{4}-\frac{1}{6}\right)$$

$$= 4\left(\frac{1}{12}\right)$$

$$= \frac{1}{3}$$

$$\boxed{E(Y^2) = \frac{1}{3}}$$

The variance of random variable X is,

$$\sigma_X^2 = E(X^2) - \mu_X^2$$

$$= \frac{2}{3} - \left(\frac{4}{5}\right)^2$$

$$= \frac{2}{3} - \frac{16}{25}$$

$$\boxed{\sigma_X^2 = \frac{2}{75}}$$

The variance of random variable 'Y' is,

$$\sigma_Y^2 = E(Y^2) - \mu_Y^2$$

$$= \frac{1}{3} - \left(\frac{8}{15}\right)^2$$

$$= \frac{1}{3} - \frac{64}{225}$$

$$= \frac{11}{225}$$

$$\boxed{\sigma_Y^2 = \frac{11}{225}}$$

The correlation coefficient of X and Y is,

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X\sigma_Y}$$

$$= \frac{\frac{104}{225}}{\sqrt{\frac{2}{75}}\sqrt{\frac{11}{225}}}$$

$$= 12.801$$

$$\boxed{\rho_{XY} = 12.801}$$

2.1.3 Means and Variances of Linear Combinations of Random Variables

Q28. If a and b are constants then,

$E(aX + b) = aE(X) + b$.

Answer :

The expected value for $aX + b$ is given by,

$$E(aX+b) = \int_{-\infty}^{\infty} (ax+b)f(x)dx$$

$$E(aX+b) = \int_{-\infty}^{\infty} axf(x)dx + \int_{-\infty}^{\infty} bf(x)dx$$

$$E(aX+b) = a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx$$

$$E(aX+b) = aE(X) + b(1) \qquad \left[\because E(X) = \int_{-\infty}^{\infty} xf(x) \text{ and } \int_{-\infty}^{\infty} f(x)dx = 1\right]$$

$$E(aX+b) = aE(X) + b$$

$$\boxed{\therefore E(aX+b) = aE(X) + b}$$

Q29. The expected value of the sum or difference of two or more functions of a random variable X is the sum or difference of the expected values of the functions. That is,

$E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$

Answer :

Given that,

X is a random variable

The expected value for the sum or difference of two or more functions with random variable X is,

$$= E[g(X) \pm h(X)]$$

$$= \int_{-\infty}^{\infty} [g(X) \pm h(X)]f(x)dx \qquad \left[\because E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx\right]$$

$$= \int_{-\infty}^{\infty} g(x)f(x)dx \pm \int_{-\infty}^{\infty} h(x)f(x)dx$$

$$= E[g(X)] \pm E[h(X)] \qquad \left[\begin{array}{l}\because E[g(x)] = \int_{-\infty}^{\infty} g(x)f(x)dx \\ \quad E[h(x)] = \int_{-\infty}^{\infty} h(x)f(x)dx\end{array}\right]$$

$$\therefore \quad E[g(X) \pm h(X)] = E[g(X)] \pm E[h(X)]$$

Thus, the expected value of the sum or difference of two or more functions with random variable X is equal to the sum or difference of the expected values of the functions.

Q30. The expected value of the sum or difference of two or more functions of the random variables X and Y is the sum or difference of the expected values of the functions That is,

$E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)]$.

Answer :

Model Paper-III, Q4(a)

Given that,

X, Y are random variables

The expected value for the sum or difference of two or more functions is,

$$= E[g(X, Y) \pm h(X, Y)]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} [g(x, y) \pm h(x, y)]f(x, y)\,dxdy \qquad \left[\because E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f(x, y)\,dxdy\right]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x, y)f(x, y)\,dxdy \pm \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x, y)f(x, y)\,dxdy$$

$$= E[g(X, Y)] \pm E[h(X, Y)] \qquad \left[\begin{aligned}\because E[g(X,Y)] &= \int_{-\infty}^{\infty} g(x, y)f(x, y)\,dxdy \\ E[h(X,Y)] &= \int_{-\infty}^{\infty} h(x, y)f(x, y)\,dxdy\end{aligned}\right]$$

$$\therefore \quad E[g(X, Y) \pm h(X, Y)] = E[g(X, Y)] \pm E[h(X, Y)]$$

Thus, the expected value of the sum or difference of two or more functions with random variables X, Y is equal to the sum or difference of the expected values of the functions.

Hence proved.

Q31. Let X and Y be two independent random variables. Then

E(XY) = E(X)E(Y).

Answer :

Given that,

X and Y are two independent random variables

The expected value for XY is

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)\,dxdy$$

Consider $f(x, y) = g(x)h(y)$, since X, Y are independent so,

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyg(x)h(y)\,dxdy$$

$$= \int_{-\infty}^{\infty} xg(x)\,dx \int_{-\infty}^{\infty} yh(y)\,dy$$

$$E(XY) = E(X)E(Y) \qquad \left[\begin{aligned}\because E(X) &= \int_{-\infty}^{\infty} xg(x)\,dx \\ E(Y) &= \int_{-\infty}^{\infty} yh(y)\,dy\end{aligned}\right]$$

$$\boxed{\therefore\ E(XY) = E(X)E(Y)}$$

Q32. If X and Y are random variables with joint probability distribution f(x, y) and a, b, and c are constants, then,

$$\sigma^2_{aX+bY+c} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{XY}$$

Answer :

Given that,

X, Y are two random variables

Consider, variance $\sigma^2_{aX+bY+c}$

$$= E\{[(aX + bY + c) - \mu_{aX+bY+c}]^2\} \qquad [\because \sigma^2 = E[(X - \mu)^2]] \qquad \text{...(1)}$$

Now, consider $\mu_{aX+bY+c}$

$$\mu_{aX+bY+c} = E(aX + bY + c)$$

$$aX + bY + c = aE(X) + bE(Y) + c \qquad [\because \mu = E(X)]$$

$$\boxed{\mu_{aX+bY+c} = a\mu_X + b\mu_Y + c}$$

Substitute equation (2) in equation (1), $\qquad [\because \mu_x = E(X), \mu_y = E(Y)] \qquad \ldots (2)$

$$\sigma^2_{aX+bY+c} = E\{[(aX + bY + c) - a\mu_x + b\mu_y + c]^2\}$$

$$\sigma^2_{aX+bY+c} = E\{[a(X - \mu_x) + b(Y - \mu_y)]^2\}$$

$$\sigma^2_{aX+bY+c} = a^2E[(X - \mu_x)^2] + b^2E[(Y - \mu_y)^2] + 2abE[(X - \mu_x)(Y - \mu_y)]$$

$$\sigma^2_{aX+bY+c} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{XY} \qquad \left[\because \sigma^2_X = (X-\mu_X)^2, \sigma^2_Y = (X-\mu_Y)^2, \sigma_{XY} = (X-\mu X)(Y-\mu Y)\right]$$

$$\boxed{\therefore \sigma^2_{aX+bY+c} = a^2\sigma^2_X + b^2\sigma^2_Y + 2ab\sigma_{XY}}$$

Hence proved.

Q33. Consider two independent random variables X and Z with means μ_x and μ_z and variances σ^2_X and σ^2_Z, respectively. Consider a random variable,

Y = X/Z.

Give approximations for E(Y) and Var(Y).

Answer :

Given that,

X, Z are two independent random variables containing means μ_x, μ_y and variances σ^2_X, σ^2_Z

Also, given that random variable

$$Y = \frac{X}{Z}$$

Here, $\frac{dy}{dx} = \frac{d}{dx}\left(\frac{X}{Z}\right)$

$$= \frac{1}{Z}\left(\frac{dX}{dx}\right)$$

$$= \frac{1}{Z}(1)$$

$$\boxed{\frac{dy}{dx} = \frac{1}{Z}}$$

So, $\boxed{\frac{d^2y}{dx^2} = 0}$

$$\frac{dy}{dz} = \frac{d}{dz}\left(\frac{X}{Z}\right)$$

$$= x\frac{d}{dz}(Z^{-1})$$

$$= x(-1)Z^{-1-1}$$

$$= -xZ^{-2}$$

$$\boxed{\frac{dy}{dz} = \frac{-x}{Z^2}}$$

So, $\frac{d^2y}{dz^2} = \frac{d^2}{dz^2}\left(\frac{-x}{Z^2}\right)$

$$= (-x)(-2)Z^{-2-1}$$

$$= 2xZ^{-3}$$

$$= \frac{2x}{Z^3}$$

$$\boxed{\frac{d^2y}{dz^2} = \frac{2x}{Z^3}}$$

Now,

$$E(Y) = \frac{\mu_X}{\mu_Z} + \frac{\mu_X}{\mu_Z^3}\sigma_Z^2$$

$$E(Y) = \frac{\mu_X}{\mu_Z}\left(1 + \frac{\mu_X}{\mu_Z^2}\sigma_Z^2\right)\sigma$$

and $\text{Var}(Y) = \frac{1}{\mu_Z^2}\sigma_X^2 + \frac{\mu_X^2}{\mu_Z^4}\sigma_Z^2$

$$\text{Var}(Y) = \frac{1}{\mu_Z^2}\left(\sigma_X^2 + \frac{\mu_X^2}{\mu_Z^2}\sigma_Z^2\right)$$

PROBLEMS

Q34. Suppose that the number of cars X that pass through a car wash between 4:00 P.M. and 5:00 P.M. on any sunny Friday has the following probability distribution:

x	4	5	6	7	8	9
P(X = x)	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Let f(X) = 2X – 1 represent the amount of money, in dollars, paid to the attendant by the manager. Find the attendant's expected earnings for this particular time period.

Solution :

Let 'X' be a random variable that denotes number of cars that pass through a car wash between 4:00 P.M to 5:00 P.M on any sunny Friday.

The probability distribution for the random variable 'X' is given as,

x	4	5	6	7	8	9
P(X = x)	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{6}$	$\frac{1}{6}$

Given that,

$$f(X) = 2X - 1$$

Now, consider $E(aX + b) = aE(X) + b$

The expected value for $f(X)$ is,

$$E(2X - 1) = 2E(X) - 1$$

The mean of random variable 'X' is given by,

$$\mu = E(X) = \sum_x xf(x)$$

$$\therefore \quad E(X) = \sum_{x=4}^{9} xf(x)$$

$$= (4)(f(4)) + (5)(f(5)) + (6)(f(6)) + 7(f(7)) + 8(f(8)) + 9(f(9))$$

$$= (4)\left(\frac{1}{12}\right) + (5)\left(\frac{1}{12}\right) + (6)\left(\frac{1}{4}\right) + (7)\left(\frac{1}{4}\right) + 8\left(\frac{1}{6}\right) + 9\left(\frac{1}{6}\right)$$

$$= \frac{4}{12} + \frac{5}{12} + \frac{6}{4} + \frac{7}{4} + \frac{8}{6} + \frac{9}{6}$$

$$= \frac{41}{6}$$

$$\boxed{E(X) = \frac{41}{6}}$$

$$E(2X - 1) = 2E(X) - 1$$

$$= 2\left(\frac{41}{6}\right) - 1$$

$$= \frac{41}{3} - 1$$

$$\boxed{E(2X - 1) = 12.67}$$

$\therefore$ The expected earnings of the attendant for the timeperiod (i.e., 4:00 PM to 5:00 PM) is \$12.67.

Q35. Let X be a random variable with density function

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere} \end{cases}$$

Find the expected value of g(X) = 4X + 3.

Solution :

Given that,

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

Consider, $E(aX + b) = \int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx$

The expected value for $g(x)$ is,

$$E(4X + 3) = E(4X) + E(3)$$

$$E(4X + 3) = 4(E(X)) + E(3) \quad \text{... (1)}$$

Now, consider mean $E(X) = \int_{-\infty}^{\infty} xf(x)dx$

$$E(X) = \int_{-1}^{2} x.\frac{x^2}{3}dx$$

$$= \int_{-1}^{2} \frac{x^2}{3}dx$$

$$= \frac{1}{3}\int_{-1}^{2} x^3 dx$$

$$= \left(\frac{1}{3}\right)\left[\frac{x^{3+1}}{3+1}\right]_{-1}^{2}$$

$$= \frac{1}{3}\left(\frac{16}{4} - \frac{1}{4}\right)$$

$$= \frac{1}{3}\left(\frac{15}{4}\right)$$

$$= \frac{5}{4}$$

$$\boxed{E(X) = \frac{5}{4}}$$

Now, substitute $E(X)$ in equation (1).

$$E(4X+3) = 4(E(X)) + E(3)$$
$$= 4\left(\frac{5}{4}\right) + 3 \quad [\because E(3) = 3]$$
$$= 5 + 3$$
$$\boxed{E(4X+3) = 8}$$

$\therefore$ The expected value is 8.

Q36. Let X be a random variable with probability distribution as follows,

x	0	1	2	3
f(x)	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$

Find the expected value of $Y = (X-1)^2$.

Solution :

Given that,

'X' is a random variable

$Y = (X-1)^2$

The probability distribution is given as,

x	0	1	2	3
f(x)	$\frac{1}{3}$	$\frac{1}{2}$	0	$\frac{1}{6}$

We have, $E[(X-1)^2] = E(X^2 - 2X + 1)$
$$= E(X^2) - E(2X) + E(1)$$
$$= E(X^2) - 2E(X) + E(1) \quad \text{... (1)}$$

Here,
$$E(X) = \sum_x x.f(x)$$
$$= \sum_{x=0}^{3} x.f(x)$$
$$= (0)(f(0)) + (1)(f(1)) + (2)(f(2)) + (3)(f(3))$$
$$= (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (2)(0) + (3)\left(\frac{1}{6}\right)$$
$$= 0 + \frac{1}{2} + 0 + \frac{3}{6}$$
$$= 1$$
$$\boxed{E(X) = 1}$$
$$E(X^2) = \sum_x x^2 f(x)$$
$$= \sum_{x=0}^{3} x^2 f(x)$$
$$= (0)^2(f(0)) + (1)^2(f(1)) + (2)^2(f(2)) + (3)^2(f(3))$$
$$= (0)\left(\frac{1}{3}\right) + (1)\left(\frac{1}{2}\right) + (4)(0) + (9)\left(\frac{1}{6}\right)$$
$$= 0 + \frac{1}{2} + 0 + \frac{9}{6}$$
$$= 2$$
$$\boxed{E(X^2) = 2}$$

Now, on substituting the values of $E(X)$, $E(X^2)$ in equation (1) we get,

$$E[(X-1)^2] = 2 - 2(1) + 1$$
$$= 2 - 2 + 1 \quad [\because E(1) = 1]$$
$$\boxed{E[(X-1)^2] = 1}$$

$\therefore$ The expected value is 1

Q37. The weekly demand for a certain drink, in thousands of liters, at a chain of (__) venience stores is a continuous random variable $g(X) = X^2 + X - 2$, where X (___) the density function.

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

Find the expected value of the weekly demand for the drink.

Solution :

Given that,

$$f(x) = \begin{cases} 2(x-1), & 1 < x < 2 \\ 0, & \text{elsewhere} \end{cases}$$

$g(X) = X^2 + X - 2$

The expected value for g(x) is,

$$E(X^2 + X - 2) = E(X^2) + E(X) - E(2) \quad \text{... (1)}$$

Here, Mean $E(X) = \int_{-\infty}^{\infty} xf(x)dx$

$$E(X) = \int_1^2 x.2(x-1)dx$$
$$= 2\int_1^2 x(x-1)dx$$
$$= 2\left[\int_1^2 x^2 dx - \int_1^2 x dx\right]$$
$$= 2\left(\left[\frac{x^{2+1}}{2+1}\right]_1^2 - \left[\frac{x^{1+1}}{1+1}\right]_1^2\right)$$
$$= 2\left(\left[\frac{x^3}{3}\right]_1^2 - \left[\frac{x^2}{2}\right]_1^2\right)$$
$$= 2\left(\left(\frac{8}{3} - \frac{1}{3}\right) - \left(\frac{4}{2} - \frac{1}{2}\right)\right)$$
$$= 2\left(\left(\frac{7}{3}\right) - \left(\frac{3}{2}\right)\right)$$
$$= 2\left(\frac{5}{6}\right)$$
$$E(X) = \frac{5}{3}$$
$$\boxed{\therefore E(X) = \frac{5}{3}}$$

Now, consider

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$= \int_1^2 x^2.2(x-1)\,dx$$

$$= 2\int_1^2 x^2.(x-1)\,dx$$

$$= 2\int_1^2 (x^3 - x^2)\,dx$$

$$= 2\left[\int_1^2 x^3 dx - \int_1^2 x^2 dx\right]$$

$$= 2\left[\frac{x^{3+1}}{3+1}\right]_1^2 - \left[\frac{x^{2+1}}{2+1}\right]_1^2$$

$$= 2\left[\frac{x^4}{4}\right]_1^2 - \left[\frac{x^3}{3}\right]_1^2$$

$$= 2\left(\frac{15}{4} - \frac{7}{3}\right)$$

$$= 2\left(\frac{17}{12}\right)$$

$$E(X^2) = \frac{17}{6}$$

$$\boxed{\therefore E(X^2) = \frac{17}{6}}$$

Now, substitute the values of $E(X)$, $E(X^2)$ in equation (1),

$$E(X^2 + X - 2) = E(X^2) + E(X) - E(2)$$

$$= \frac{17}{6} + \frac{5}{3} - 2 \qquad [\because E(2) = 2]$$

$$= \frac{5}{2}$$

$$\boxed{\therefore E(X^2 + X - 2) = \frac{5}{2}}$$

$\therefore$ The weekly demand for the drink in thousands of liters is,

$$= \frac{5}{2} \times 1000$$

$$= 2500 \text{ liters.}$$

Thus, the expected value of the weekly demand for the drink is 2500 liters.

Q38. It is known that the ratio of gallium to arsenide does not affect the functioning of gallium-arsenide wafers, which are the main components of microchips. Let X denote the ratio of gallium to arsenide and Y denote the functional wafers retrieved during a 1-hour period X and Y are independent random variables with the joint density function.

$$\mathbf{f(x, y) = \begin{cases} \dfrac{x(1+3y^2)}{4}, & 0<x<2, 0<y<1 \\ 0, & \text{elsewhere} \end{cases}}$$

Show that E(XY) = E(X)E(Y).

Model Paper-III, Q4(b)

Solution :

Given that,

X, Y are independent random variables.

Where, 'X' denotes the ratio of gallium to arsenide, and 'Y' denotes the functional wafers retrieved during a 1-hour period.

The joint density function is,

$$f(x, y) = \begin{cases} \dfrac{x(1+3y^2)}{4}, & 0<x<2, 0<y<1 \\ 0, & \text{elsewhere} \end{cases}$$

Consider,

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x, y)\,dxdy$$

$$= \int_0^1\int_0^2 xy\frac{x(1+3y^2)}{4}\,dxdy \qquad \text{...(1)}$$

$$= \int_0^1\int_0^2 \frac{x^2y(1+3y^2)}{4}\,dxdy$$

$$= \frac{1}{4}\int_0^1\int_0^2 x^2y(1+3y^2)\,dxdy$$

$$= \frac{1}{4}\int_0^2 x^2dx\int_0^1 y(1+3y^2)\,dy$$

$$= \frac{1}{4}\left(\int_0^2 x^2dx\int_0^1 (y+3y^3)\,dy\right)$$

$$= \frac{1}{4}\left(\int_0^2 x^2dx\left(\int_0^1 ydy + \int_0^1 3y^3dy\right)\right)$$

$$= \frac{1}{4}\left(\left[\frac{x^{2+1}}{2+1}\right]_0^2\left(\left[\frac{y^{1+1}}{1+1}\right]_0^1 + 3\left[\frac{y^{3+1}}{3+1}\right]_0^1\right)\right)$$

$$= \frac{1}{4}\left(\left[\frac{x^3}{3}\right]_0^2\left(\left[\frac{y^2}{2}\right]_0^1 + 3\left[\frac{y^4}{4}\right]_0^1\right)\right)$$

$$= \frac{1}{4}\left(\left(\frac{8}{3} - \frac{0}{3}\right)\left(\left(\frac{1}{2} - \frac{0}{2}\right) + 3\left(\frac{1}{4} - \frac{0}{4}\right)\right)\right)$$

$$= \frac{1}{4}\left(\left(\frac{8}{3}\right)\left(\frac{1}{2} + 3\left(\frac{1}{4}\right)\right)\right)$$

$$= \frac{1}{4}\left(\left(\frac{8}{3}\right)\left(\frac{1}{2} + \frac{3}{4}\right)\right)$$

$$= \frac{1}{4}\left(\left(\frac{8}{3}\right)\left(\frac{5}{4}\right)\right)$$

$$= \frac{1}{4}\left(\frac{10}{3}\right)$$

$$E(XY) = \frac{5}{6}$$

$$\boxed{\therefore E(XY) = \frac{5}{6}}$$

Consider, $E(X) = \int_{-\infty}^{\infty} xg(x)dx$

Now, compare with equation (1) and the obtained $E(X)$ is as shown below,

$$E(X) = \int_0^2 x.\frac{x}{2}dx$$

$$= \frac{1}{2}\int_0^2 x^2 dx$$

$$= \left(\frac{1}{2}\right)\left[\frac{x^{2+1}}{2+1}\right]_0^2$$

$$= \left(\frac{1}{2}\right)\left[\frac{x^3}{3}\right]_0^2$$

$$= \left(\frac{1}{2}\right)\left[\frac{8}{3} - \frac{0}{3}\right]$$

$$= \frac{1}{2} \times \frac{8}{3}$$

$$= \frac{4}{3}$$

$$\boxed{E(X) = \frac{4}{3}}$$

Consider, $E(Y) = \int_{-\infty}^{\infty} yh(y)dy$

Now, compare with equation (1) and the obtained $E(Y)$ is as shown below,

$$E(Y) = \int_0^1 y.\frac{(1+3y^2)}{2}dy$$

$$= \frac{1}{2}\int_0^1 y(1+3y^2)dy$$

$$= \frac{1}{2}\int_0^1 (y+3y^3)dy$$

$$= \frac{1}{2}\left[\int_0^1 y.dy + \int_0^1 3y^3 dy\right]$$

$$= \frac{1}{2}\left[\int_0^1 y.dy + 3\int_0^1 y^3 dy\right]$$

$$= \frac{1}{2}\left(\left[\frac{y^{1+1}}{1+1}\right]_0^1 + 3\left[\frac{y^{3+1}}{3+1}\right]_0^1\right)$$

$$= \frac{1}{2}\left(\left[\frac{y^2}{2}\right]_0^1 + 3\left[\frac{y^4}{4}\right]_0^1\right)$$

$$= \frac{1}{2}\left(\left[\frac{1}{2} - \frac{0}{2}\right] + 3\left[\frac{1}{4} - \frac{0}{4}\right]\right)$$

$$= \frac{1}{2}\left(\frac{1}{2} + 3\left(\frac{1}{4}\right)\right)$$

$$= \frac{1}{2}\left(\frac{1}{2} + \frac{3}{4}\right)$$

$$= \frac{1}{2}\left(\frac{5}{4}\right)$$

$$\boxed{E(Y) = \frac{5}{8}}$$

So, $E(X)E(Y) = \left(\frac{4}{3}\right)\left(\frac{5}{8}\right)$

$$= \left(\frac{5}{6}\right)$$

$$\boxed{\therefore\ E(X)E(Y) = \left(\frac{5}{6}\right)}$$

$$E(X)E(Y) = E(XY)$$

$$\left[\because\ E(XY) = \frac{5}{6}\right]$$

Thus, $E(XY) = E(X)E(Y)$

Hence proved.

Q39. If X and Y are random variables with variances $\sigma_X^2 = 2$, $\sigma_Y^2 = 4$ and covariance $\sigma_{XY} = -2$, find the variance of the random variable $Z = 3X - 4Y + 8$.

Solution :

Given that,

X, Y are random variables whose

Variances are $\sigma_X^2 = 2$ and $\sigma_Y^2 = 4$

Covariance is $\sigma_{XY} = -2$ and

The random variable $Z = 3X - 4Y + 8$

The variance of the random variable Z can be written as,

$$\sigma_Z^2 = \sigma_{3X-4Y+8}^2$$

Consider, $\sigma_{aX+bY+c}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 - 2ab(\sigma_x)(\sigma_y)$

Now, σ_{3X-4Y}^2 $\quad [\because\ \sigma_{aX+c}^2 = a^2\sigma_X^2]$

$$\sigma_Z^2 = 9\sigma_X^2 + 16\sigma_Y^2 - 2(3\sigma_X^2)(4\sigma_Y^2)$$

$$= (9)(2) + (16)(4) - 24(-2)$$

$$[\because\ \sigma_X^2 = 2, \sigma_Y^2 = 4, \sigma_{XY} = -2]$$

$$= 18 + 64 + 48$$

$$\boxed{\sigma_Z^2 = 130}$$

$\therefore$ Variance of z is 130

Q40. Let X and Y denote the amounts of two different types of impurities in a batch of a certain chemical product. Suppose that X and Y are independent random variables with variances $\sigma_X^2 = 2$ and $\sigma_Y^2 = 3$. Find the variance of the random variable $Z = 3X - 2Y + 5$.

Solution :

Given that,

X, Y denotes the amounts of two different types of impurities.

Variances are $\sigma_X^2 = 2$ and $\sigma_Y^2 = 3$

The random variable $Z = 3X - 2Y + 5$

Here, the X, Y are two independent random variables.

The variance of the random variable z can be written as,

$$\sigma_z^2 = \sigma_{3X-2Y+5}^2$$

$$= \sigma_{3X-2Y}^2 \qquad [\because \sigma_{aX+c}^2 = a^2\sigma_X^2]$$

Since, X, Y are independent random variables, so consider $\sigma_{aX-bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2$

$$\sigma_z^2 = 9\sigma_X^2 + 4\sigma_Y^2 \qquad [\because \sigma_X^2 = 2, \sigma_Y^2 = 3]$$

$$= 9(2) + 4(3)$$

$$= 18 + 12$$

$$\sigma_z^2 = 30$$

$$\boxed{\sigma_z^2 = 30}$$

$\therefore$ The variance of z is 30.

2.1.4 Chebyshev's Theorem

Q41. Write short note on Chebyshev's inequality.

Answer :

Chebyshev's inequality provides assurity that for a wide class of probability distribution almost all values are near to mean. According to Chebyshev's inequality if x is a random variable with mean value $\bar{x}$ and variance as σ_x^2, then for any positive integer k,

$$P\left[|x-\bar{x}| \geq k\sigma_x\right] \leq \frac{1}{k^2}$$

$$P\left[|x-\bar{x}| < k\sigma_x\right] \geq 1-\frac{1}{k^2}$$

This inequality has wide application in probability distribution where the mean and variance are defined. For an instance it can be used to prove the weak law of large numbers. This inequality is based on the following two conditions,

(i) The underlying distribution should have a mean.

(ii) The average size of the deviations should be away from this mean.

Q42. State and prove Chebyshev's inequality.

Model Paper-I, Q5(a)

Answer :

Statement

If 'X' is a random variable with mean value $\bar{X}$ and variance as σ_X^2, then for any positive integer K,

$$P\left[|X-\bar{X}| \geq K\sigma_X\right] \leq \frac{1}{K^2}$$

(or)

$$\sigma_X P\left[|X-\bar{X}| \geq K\sigma_X\right] \geq 1-\frac{1}{K^2}$$

Proof

Let us consider that 'X' is a continuous random variable.

From the definition of variance we know that,

$$V(X) = E[X-E(X)]^2$$

$\because$ Here, $V(X) = \sigma_X^2$ and $E(X) = \bar{X}$

$\therefore \quad \sigma_X^2 = E[X - \bar{X}]^2$

$$\sigma_X^2 = \int_{-\infty}^{\infty} (x-\bar{X})^2 f(x)\,dx \quad [\because X \text{ is continuous}]$$

$$= \int_{-\infty}^{\bar{X}-K\sigma_X} (x-\bar{X})^2 f(x)\,dx + \int_{\bar{X}-K\sigma_X}^{\bar{X}+K\sigma_X} (x-\bar{X})^2 f(x)\,dx + \int_{\bar{X}+K\sigma_X}^{\infty} (x-\bar{X})^2 f(x)\,dx \quad \left[\because \int_{\bar{X}-K\sigma_X}^{\bar{X}+K\sigma_X} (x-\bar{X})^2 f(x)\,dx \geq 0\right]$$

$$\Rightarrow \quad \sigma_X^2 \geq \int_{-\infty}^{\bar{X}-K\sigma_X} (x-\bar{X})^2 f(x)\,dx + \int_{\bar{X}+K\sigma_X}^{\infty} (x-\bar{X})^2 f(x)\,dx$$

From the 1st integral we have,

$$X \leq \bar{X} - K\sigma_X$$

$$X - \bar{X} \leq -K\sigma_X$$

$$-(X - \bar{X}) \geq K\sigma_X$$

$$(X - \bar{X})^2 \geq K^2 \sigma_X^2 \quad \text{... (i)}$$

From the 2nd integral we have,

$$X \geq \bar{X} + K\sigma_X$$

$$X - \bar{X} \geq K\sigma_X$$

$$(X - \bar{X})^2 \geq K^2 \sigma_X^2 \quad \text{... (ii)}$$

$\therefore$ From equations (i) and (ii) we can write,

$$\sigma_X^2 \geq \int_{-\infty}^{\bar{X}-K\sigma_X} K^2\sigma_X^2 f(x)\,dx + \int_{\bar{X}+K\sigma_X}^{\infty} K^2\sigma_X^2 f(x)\,dx$$

$$\sigma_X^2 \geq K^2\sigma_X^2 \left[\int_{-\infty}^{\bar{X}-K\sigma_X} f(x)\,dx + \int_{\bar{X}+K\sigma_X}^{\infty} f(x)\,dx\right]$$

$$\sigma_X^2 \geq K^2\sigma_X^2 \left[P(X < \bar{X}-K\sigma_X) + P(X \geq \bar{X}+K\sigma_X)\right]$$

$$\sigma_X^2 \geq K^2\sigma_X^2 \left[P(X-\bar{X} < -K\sigma_X) + P(X-\bar{X} > K\sigma_X)\right]$$

$$\sigma_X^2 \geq K^2\sigma_X^2 \; P\left[|X-\bar{X}| \geq K\sigma x\right]$$

$$\sigma_X^2 \geq K^2\sigma_X^2 \; P\left[|X-\bar{X}| \geq K\sigma_X\right]$$

$$1 \geq \left(\frac{K^2\sigma_x^2}{\sigma_x^2}\right) P\left[|X-\bar{X}| \geq K\sigma_X\right]$$

$$\frac{1}{K^2} \geq P\left[|X-\bar{X}| \geq K\sigma_X\right]$$

$$\therefore \quad P\left[|X-\bar{X}| \geq K\sigma_x\right] \leq \frac{1}{K^2}$$

It can also be written as,

$$P\left[|X-\bar{X}| < K\sigma_X\right] \geq 1 - \frac{1}{K^2}$$

Note

If $K\sigma_x = C$, where 'C' is any positive integer then the chebyshev's inequality will become,

$$P\left[|X - \bar{X}| \geq C\right] \leq \frac{1}{\left(C/\sigma_x\right)^2}$$

$$P\left[|X - \bar{X}| \geq C\right] \leq \frac{\sigma_x^2}{C^2}$$

$$P\left[|X - E(X)| \geq C\right] \leq \frac{V(X)}{C^2}$$

(or)

$$P\left[|X - E(X)| < C\right] \geq 1 - \frac{V(X)}{C^2}$$

PROBLEM

Q43. A random variable X has a mean $\mu = 8$, a variance $\sigma 2 = 9$, and an unknown probability distribution. Find,

(a) $P(-4 < X < 20)$

(b) $P(|X - 8| \geq 6)$.

Solution :

Given that,

X is a random variable,

Mean $\mu = 8$

Variance $\sigma^2 = 9$

(a) $P(-4 < X < 20)$

From chebyshev's Theorem we have,

Consider, $P(\mu - k\sigma < X < \mu + K\sigma) \geq 1 - \frac{1}{k^2}$

$$= P[8 - (4)(3) < X < 8 + (4)(3)] \geq 1 - \frac{1}{(4)^2} \qquad \left[\because \begin{array}{l} \mu = 8 \\ \sigma = 3 \end{array}\right]$$

$$= P[8 - (4)(3) < X < 8 + (4)(3)] \geq 1 - \frac{1}{16}$$

$$P[8 - (4)(3) < X < 8 + (4)(3)] \geq \frac{15}{16}$$

$$\therefore\ P(-4 < X < 20) \geq \frac{15}{16}$$

(b) $P(|X - 8| \geq 6)$

Consider, $P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$

$$= 1 - P(|X - 8| < 6)$$

$$= 1 - P(-6 < X - 8 < 6) \leq \frac{1}{k^2}$$

$$= 1 - P[8 - (2)(3) < X < 8 + (2)(3)] \leq \frac{1}{(2)^2}$$

$$= 1 - P[8 - (2)(3) < X < 8 + (2)(3)] \leq \frac{1}{4}$$

$$\therefore\ P(|X - 8| \geq 6) \leq \frac{1}{4}$$

2.2 DISCRETE PROBABILITY DISTRIBUTIONS

2.2.1 Introduction and Motivation, Binomial Distribution

Q44. Discuss in detail about discrete probability distribution.

Answer :

Discrete probability distribution can describe the behavior of a random variable in the form of a histogram (or) table (or) a formula. However, the observations obtained from the various statistical experiments exhibit similar behavior. Since the discrete random variables are related to these experiments, they can be defined by the same probability distribution and also represented by using a single formula. Therefore, only a limited set of probability distributions are required to demonstrate all the related discrete random variables.

The following are some of the discrete probability distributions,

(i) **Binomial Distribution**

For answer refer Unit-II, Q46.

(ii) **Geometric Distribution**

For answer refer Unit-II, Q54.

(iii) **Negative Binomial Distribution**

For answer refer Unit-II, Q53.

(iv) **Poisson Distribution**

For answer refer Unit-II, Q58.

Q45. Discuss in detail about Bernoulli's distribution.

Answer : Model Paper-II, Q5(a)

Bernoulli's Distribution

Bernoulli's distribution is an experiment that can have either of the two possible outcomes (i.e., success or failure). Its probability depends only on the parameter 'p'.

Consider 'X' be a random variable that can have two outcomes i.e., success and failure.

Here,

X(Success) = 1

X(Failure) = 0

The Probability Mass Function (p.m.f) of a random variable x is given below,

$$P(X = x) = \begin{Bmatrix} p & ; \text{ for } x = 1 \text{ (success)} \\ & \text{or} \\ q \text{ or } 1-p & ; \text{ for } x = 0 \text{ (failure)} \end{Bmatrix}$$

Here,

p = Probability of success

q = Probability of failure

Examples

1. Result of a student in an examination can be pass or fail.
2. Result of tossing a coin can be head or tail.

Properties of Bernoulli's Distribution

The Bernoulli distribution should possess the properties mentioned below,

(i) The experiment should contain repeated number of trials.

(ii) The outcome of every trial is either success or failure.

(iii) The probability of success i.e., 'p', remains constant for successive trials.

(iv) The repeated trials are considered to be independent in nature.

Q46. Explain in detail about Binomial distribution. Model Paper-III, Q5(a)

Answer :

Binomial Distribution

In an experiment consisting of 'n' repeated trials where all the trials are independent and each trial results in only two possible outcomes. The two outcomes are 'success' represented by 'p' and 'failure' represented by 'q' respectively. Therefore, the binomial density function is given by,

$$B(x, n, p) = {}^nC_x p^x (q)^{n-x}$$

Where,

$$q = 1 - p \qquad [\because p + q = 1]$$

Properties

- Mean of binomial distribution = np
- Variance of binomial distribution = npq
- Mode of binomial distribution : Mode is the value of X corresponding to the highest value of $P(X)$ in the distribution.

Example

Binomial distribution is used to identify the following,

1. The number of defective items in a set of 100 items.
2. The number of girls present in a family of 10 children.

Applications and Uses of Binomial Distribution

The applications and uses of binomial distribution are as follows,

1. Binomial distribution is applicable in case of repeated trials such as,
   - Number of telephone calls received at a telephone exchange during a particular period of time.
   - Students appearing in the examinations.
   - The number of births taken place in a nursing home.
2. It is used in survey sampling.
3. It is used in industry related problems.
4. It is used in quality control charts.
5. It is used in counting the defective products in a company.

PROBLEMS

Q47. The probability that a certain kind of component will survive a shock test is 3/4. Find the probability that exactly 2 of the next 4 components tested survive.

Solution :

Given that,

The probability that a specific component will survive a shock test is, $p = \frac{3}{4}$

Here, the number of components that will be tested are, $n = 4$

The probability distribution of the binomial random variable X is given by,

$$b(x; n,p) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2....n.$$

The probability that exactly 2 components will survive is,

$$b\left(2;4,\frac{3}{4}\right) = \binom{4}{2}\left(\frac{3}{4}\right)^2\left(1-\frac{3}{4}\right)^{4-2} \qquad [\because q = 1-p]$$

$$= \binom{4}{2}\left(\frac{3}{4}\right)^2\left(\frac{1}{4}\right)^2$$

$$= \left(\frac{4!}{2!2!}\right)\frac{(3)^2(1)^2}{(4)^4}$$

$$= \frac{24}{4}\times\frac{9}{256}$$

$$= \frac{27}{128}$$

$\therefore$ The probability that exactly two components will survive a shock test is $\frac{27}{128}$

Q48. The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that (a) at least 10 survive, (b) from 3 to 8 survive, and (c) exactly 5 survive?

Solution : Model Paper-I, Q5(b)

Given that,

The probability that a patient recovers from a rare blood disease is, $p = 0.4$.

The number of people contracted to this disease are, $n = 15$

Let, X be the number of people who survive from this disease.

(a) Atleast 10 Survive

The probability that atleast 10 people will survive is,

$$P(X \geq 10) = 1 - P(X < 10)$$

$$= 1 - \sum_{x=0}^{9} b(x; n, p)$$

$$= 1 - \sum_{x=0}^{9} b(x; 15, 0.4)$$

$$= 1 - 0.9662$$

$$= 0.03$$

$$\boxed{\therefore P(X \geq 10) = 0.03}$$

(b) From 3 to 8 Survive

The probability that, 3 to 8 people survive is,

$$P(3 \leq X \leq 8) = \sum_{x=3}^{8} b(x; 15, 0.4)$$

$$= \sum_{x=0}^{8} b(x; 15, 0.4) - \sum_{x=0}^{2} b(x; 15, 0.4)$$

$$= 0.9050 - 0.0271 \qquad [\because \text{From binomial probability sums table}]$$

$$= 0.88$$

$$\boxed{\therefore P(3 \leq X \leq 8) = 0.88}$$

(c) **Exactly 5 will Survive**

The probability that exactly 5 people will survive is,

$P(X = 5) = b(5; 15, 0.4)$

$= \sum_{x=0}^{5} b(x; 15, 0.4) - \sum_{x=0}^{4} b(x; 15, 0.4)$

$= 0.4032 - 0.2173$ [∵ From binomial probability sums table]

$= 0.19$

$\boxed{\therefore P(X = 5) = 0.19}$

Q49. A large chain retailer purchases a certain kind of electronic device from a manufacturer. The manufacturer indicates that the defective rate of the device is 3%.

(a) The inspector randomly picks 20 items from a shipment. What is the probability that there will be at least one defective item among these 20?

(b) Suppose that the retailer receives 10 shipments in a month and the inspector randomly tests 20 devices per shipment. What is the probability that there will be exactly 3 shipments each containing at least one defective device among the 20 that are selected and tested from the shipment?

Solution :

(a) Let, X be the number of defective devices among 20.

The probability that there will be atleast one defective item is,

$P(X \geq 1) = 1 - P(X = 0)$

$= 1 - b(0; 20, 0.03)$

$= 1 - (0.03)^0(1 - 0.03)^{20-0}$

$= 1 - (1)(0.97)20$

$= 1 - 0.54$

$= 0.46$

$\boxed{\therefore P(X \geq 1) = 0.46}$

(b) Let, Y be the number of shipments containing atleast one defective item. Here, the probability that there will be atleast one defective item among 20 is, $p = 0.46$

The probability that there will be exactly 3 shipments is,

$P(Y = 3) = \binom{10}{3}(0.46)^3(1 - 0.46)^{10-3}$

$= (120)(0.46)^3(0.54)^7$

$= 120(0.0013)$

$= 0.16$

$\boxed{\therefore P(Y = 3) = 0.16}$

Q50. It is conjectured that an impurity exists in 30% of all drinking wells in a certain rural community. In order to gain some insight into the true extent of the problem, it is determined that some testing is necessary. It is too expensive to test all of the wells in the area, so 10 are randomly selected for testing.

(a) Using the binomial distribution, what is the probability that exactly 3 wells have the impurity, assuming that the conjecture is correct?

(b) What is the probability that more than 3 wells are impure?

Solution :

Given that,

The number of drinking wells that are tested randomly are, $n = 10$

The probability of the existence of impurities in drinking water is, $p = 30\% = \frac{30}{100} = 0.3$

The probability distribution of the binomial random variable X is given by,

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, x = 0, 1, 2, \ldots n$$

(a) The probability that exactly 3 drinking wells are impure is,

$$b(3; 10, 0.3) = \sum_{x=0}^{3} b(x; 10, 0.3) - \sum_{x=0}^{2} b(x; 10, 0.3)$$

$= 0.6496 - 0.3828$ [$\because$ From the binomial probability sums table]

$= 0.27$

$\boxed{b(3; 10, 0.3) = 0.27}$

(b) The probability that more than 3 drinking wells are impure is,

$$P(X > 3) = 1 - P(X \le 3)$$

$$= 1 - \sum_{x=0}^{3} b(x; 10, 0.3)$$

$= 1 - 0.6496$ [$\because$ From the binomial probability sums table]

$= 0.35$

$\boxed{P(X > 3) = 0.35}$

Q51. The probability that a patient recovers from a rare blood diseases is 0.4. If 15 people are known to have contracted this disease. Find the mean and variance of binomial random variable and then use chebyshev's theorem to interpret the interval $\mu \pm 2\sigma$.

Solution :

Given that,

The probability that a patient recovers from a rare blood disease is, $p = 0.4$

The number of people contracted to this disease are, $n = 15$

Here, the mean of binomial random variable is given by,

$\mu = np$

$= (15)(0.4)$

$\boxed{\text{Mean}, \mu = 6}$

The variance of binomial random variable is given by,

$\sigma^2 = npq$

$= (15)(0.4)(1 - 0.4)$

$= (15)(0.4)(0.6)$

$= 3.6$

$\boxed{\text{Variance}, \sigma^2 = 3.6}$

By using Chebyshev's theorem to interpret the interval $\mu \pm 2\sigma$

$$\mu \pm 2\sigma = 6 \pm 2(1.90)$$

Here,

$$\mu - 2\sigma = 6 - 2(1.90)$$
$$= 6 - 3.8 \text{ and}$$
$$\mu + 2\sigma = 6 + 2(1.90)$$
$$= 6 + 3.8$$
$$= 9.80.$$

Thus, the Chebyshev's theorem states that the number of people who will survive among all 15 patients will have a probability of atleast $\frac{3}{4}$ that fall within the interval from 2.2 to 9.8 i.e. inclusively from 2 to 9.

Q52. It is conjectured that an impurity exists in 30% of all drinking wells in a certain rural community. In order to gain some insight into the true extent of the problem, it is determined that some testing is necessary. It is too expensive to test all of the wells in the area, so 10 are randomly selected for testing. The notion that 30% of the wells are impure is merely a conjecture put forth by the area water board. Suppose 10 wells are randomly selected and 6 are found to contain the impurity. What does this imply about the conjecture? Use a probability statement.

Solution :

Given that,

A conjecture states that the 30% of the drinking wells are impure. The probability that 6 wells are found to be impure from 10 wells that are selected randomly is,

$$P(X \geq 6) = \sum_{x=0}^{10} b(x; 10, 0.3) - \sum_{x=0}^{5} b(x; 10, 0.3)$$

$$= 1 - 0.9527 \qquad [\because \text{ From binomial probability sums table}]$$

$$= 0.05$$

$$\boxed{P(X \geq 6) = 0.05}$$

Hence, there is 5% chance that more than 6 wells will be impure when 30% of all wells are impure.

Thus, the given conjecture is false and the problem of impure wells is considered to be much severe.

2.2.2 Geometric Distributions

Q53. Explain in detail about Negative Binomial Distribution.

Answer :

Negative Binomial Distribution

Negative binomial distribution can be defined as the binomial distribution except that it can perform any number of trials for fixed number of successes.

According to the succession of bernoulli trials, consider the following,

$P(X = n)$ = Probability of 'n' trials that vary to give 'r' successes $(n > r)$.

Assume that $P(X = n)$ be the probability that $(r - 1)$ successes can be obtained in $(n - 1)$ trial and r^{th} succcess in the final n^{th} trial.

So,

$$P(X = n) = {}^{n-1}C_{r-1}\, p^{r-1} q^{(n-1)-(r-1)}.p$$
$$= {}^{n-1}C_{r-1}\, p^{r-1} q^{n-1-r+1}.p$$
$$= {}^{n-1}C_{r-1}\, p^{r} q^{n-r} \quad \text{where } n = r, r+1, r+2, r+3, \ldots$$

$\therefore$ The negative probability distribution for 'n' number of trial needed for give 'r' successes is given by,

$$P(X = n) = {}^{n-1}C_{r-1}\, p^{r} q^{n-r}$$

Alternative Method

The probability that there are total of 'x' failures which precedes the r^{th} success in $(x + r)$ trials is given by,

$$P(x) = {}^{-r}C_x \, p^r (-q)^x$$

Here,

$x = 0, 1, 2, 3, \ldots$ and $r \geq 0$.

The conditions that must be satisfied for negative binomial distribution are as follows,

(i) It should have sequence of independent trials.

(ii) It should give either success or failure.

(iii) It should have constant probability of success from one trial to another.

(iv) It should continue the process until the required number of successes are obtained.

Q54. Discuss in brief about Geometric distribution.

Answer :

Geometric Distribution

Geometric distribution is the distribution in which number of trials needed inorder to achieve the first success to determined. This distribution is categorized into two different types based on the following criteria,

(i) **Number of Failures**

If 'X' is the random variable that denote number of failures needed prior to the occurrence of first success then the probability mass function is given by,

$$P(X = x) = \begin{cases} p.q^x & \text{Where } x = 0,1,2,3,\ldots \\ 0 & \text{Otherwise} \end{cases}$$

(ii) **Number of Trials**

If 'X' is the random variable that denote number of trials needed prior to the occurrence of first success then the probability mass function is given by,

$$P(X = x) = \begin{cases} p.q^{x-1} & \text{Where } x = 0,1,2,3,\ldots \\ 0 & \text{Otherwise} \end{cases}$$

Examples

1. Iterative tossing of a coin until the occurrence of first head.
2. Digital transmission of specified amount of bits until the occurrence of first error.

The conditions that are to be satisfied for geometric distribution are as follows,

(i) It should have iterative trials until the occurrence of first success.

(ii) It should have two possible outcomes.

(iii) It should have independent repeated trials.

PROBLEMS

Q55. In an NBA (National Basketball Association) championship series, the team that wins four games out of seven is the winner. Suppose that teams A and B face each other in the championship games and that team A has probability 0.55 of winning a game over team B.

(a) What is the probability that team A will win the series in 6 games?

(b) What is the probability that team A will win the series?

(c) If teams A and B were facing each other in a regional playoff series, which is decided by winning three out of five games, what is the probability that team A would win the series?

Solution : Model Paper-II, Q5(b)

Given that,

The probability of team A winning a game over team B is, $P = 0.55$

The probability distribution of the random variable X, the number of the trial on the k^{th} success occurs, is given by,

$$b^*(x; k, p) = \binom{x-1}{k-1} p^k q^{x-k}, \text{ Where } x = k, k+1, k+2, \ldots.$$

(a) The Probability that the Team A will win the Series in 6 Games

$b^*(6; 4, 0.55)$

Here, $x = 6$ games

$k = 4$ wins

$p = 0.55$

$$b^*(6\,;4, 0.05) = \binom{6-1}{4-1}(0.55)^4(1-0.55)^{6-4}$$

$$= \binom{5}{3}(0.55)^4(0.45)^2$$

$$= (10)(0.0915)(0.2025)$$

$\boxed{b^*(6; 4, 0.05) = 0.1853}$

(b) The Probability that Team A will win the Championship Series

$b^*(4; 4, 0.55) + b^*(5; 4, 0.55) + b^*(6; 4, 0.55) + b^*(7; 4, 0.55)$

$$= \binom{4-1}{4-1}(0.55)^4(1-0.55)^{4-4} + \binom{5-1}{4-1}(0.55)^4(1-0.55)^{5-4} + \binom{6-1}{4-1}(0.55)^4(1-0.55)^{6-4} + \binom{7-1}{4-1}(0.55)^4(1-0.55)^{7-4}$$

$$= \binom{3}{3}(0.55)^4(0.45)^0 + \binom{4}{3}(0.55)^4(0.45)^1 + \binom{5}{3}(0.55)^4(0.45)^2 + \binom{6}{3}(0.55)^4(0.45)^3$$

$= 0.0915 + 0.1647 + 0.1853 + 0.1667$

$= 0.6082$

(c) The Probability that Team A wins the Playoff

$= b^*(3; 3, 0.55) + b^*(4; 3, 0.55) + b^*(5; 3, 0.55)$

$$= \binom{3-1}{3-1}(0.55)^3(1-0.55)^{3-3} + \binom{4-1}{3-1}(0.55)^3(1-0.55)^{4-3} + \binom{5-1}{3-1}(0.55)^3(1-0.55)^{5-3}$$

$$= \binom{2}{2}(0.55)^3(0.45)^0 + \binom{3}{2}(0.55)^3(0.45)^1 + \binom{4}{2}(0.55)^3(0.45)^2$$

$= 0.1663 + 0.2246 + 0.2021$

$= 0.593$

Q56. For a certain manufacturing process, it is known that, on the average, 1 in every 100 items is defective. What is the probability that the fifth item inspected is the first defective item found?

Solution :

Given that,

The probability for defective items is,

$P = \frac{1}{100} = 0.01$

The geometric distribution is,

$g(x; p) = pq^{x-1}, x = 1,2,3,...$

Here, $x = 5$ and $p = 0.01$

The probability that the fifth item inspected will be the first defective item is,

$$P(X=5) = g(5; 0.01) = (0.01)(1-0.01)^{5-1}$$

$$= (0.01)(0.99)^4$$

$$= 0.00960$$

$\boxed{g(5; 0.01) = 0.00960}$

Q57. At a "busy time," a telephone exchange is very near capacity, so callers have difficulty placing their calls. It may be of interest to know the number of attempts necessary in order to make a connection. Suppose that we let p = 0.05 be the probability of a connection during a busy time. We are interested in knowing the probability that 5 attempts are necessary for a successful call.

Solution :

Given that,

The probability of a connection during a busy time is, $p = 0.05$.

The probability of 5 attempts for a successful call is,

$$P(X = 5) = g(5, 0.05)$$

$$= (0.05)(1 - 0.05)^{5-1}$$

$$= (0.05)(0.95)^4$$

$$= (0.05)(0.81)$$

$$= 0.04$$

$$\boxed{\therefore\ P(X = 5) = 0.04}$$

2.2.3 Poisson Distribution

Q58. Explain in detail about Poisson distribution.

Answer :

Poisson Distribution

Poisson distribution was discovered in the year 1837 by a french mathematician and physicist Simeon Denis Poisson. It is preferred for those events in which the outcomes occur at random instants of time wherein the interest lies only in the number of occurrences.

A random variable 'x' is said to have a Poisson distribution if it assumes only positive values and it is given by,

$$P(x, \lambda) \begin{cases} = \dfrac{e^{-\lambda}\lambda^x}{x!}; & x = 0, 1, 2, \ldots, n \\ = 0; & \text{Otherwise} \end{cases}$$

Here 'λ' is known as the parameter of distribution and is greater than '0' i.e., $\lambda > 0$. The notation $X \sim P(\lambda)$ is used to denote that X is a Poisson variate with parameter 'λ'.

Examples

Poisson distribution is used to identify the following,

1. The number of printing mistakes on each page of a book.
2. The number of deaths from rare diseases like heart attack, cancer, etc.

Applications and Uses of Poisson Distribution

The application and uses of poisson distributions are as follows,

(i) It is mostly used in quality control chart.

(ii) It is used in queueing theory.

(iii) It is used in astronomy i.e., photons appearing at telescope.

(iv) It is used in management, finance and insurance.

(v) It is used in earthquake seismology and radioactivity.

(vi) It is used in telecommunication i.e., telephone calls arriving in system.

(vii) It is used in Internet traffic.

Properties

The properties of poisson distribution are,

1. The occurrence of the events is independent i.e., the occurrence of an event in a time interval has no effect on the occurrence of the second event in the same or any other interval.
2. Theoretically, an infinite number of occurrences of the event must be possible in the interval.
3. The probability of single occurrence of the event in a given interval is directly proportional to the length of the interval.
4. In any extremely small portion of the interval, the probability of two or more occurrences of the event is negligible.

Like binomial distribution, poisson distribution also satisfies the two essential properties i.e.,

(i) $f(x) \geq 0$ and

(ii) $\Sigma f(x) = 1$.

Q59. Derive Poisson distribution as limiting case of negative binomial distribution.

Answer : Model Paper-III, Q5(b)

Poisson Distribution as Limiting Case of Negative Binomial Distribution

Poisson distribution as limiting case of negative binomial distribution under the following conditions,

(i) The constant probability of success for every individual trial is indefinitely small

i.e., $p \to 0$

(ii) The number of trials are indefinitely large

i.e., $r \to \infty$

(iii) The Poisson variate λ is finite.

i.e., $\lambda = rp \Rightarrow p = \dfrac{\lambda}{r}$

The probability mass function of negative binomial distribution is given by,

$$P(X=x) = \binom{x+r-1}{x} p^r q^x$$

$$= \binom{x+r-1}{x} \left(\frac{1}{q}\right)^r \left(\frac{p}{q}\right)^x$$

$$= \binom{x+r-1}{x} \left(\frac{1}{1+p}\right)^r \left(\frac{p}{1+p}\right)^x \qquad [\because q = 1+p]$$

$$= \binom{x+r-1}{x} \left(\frac{1}{1+\frac{\lambda}{r}}\right)^r \left(\frac{\frac{\lambda}{r}}{1+\frac{\lambda}{r}}\right)^x \qquad \left[\text{From condition (iii)}, P = \frac{\lambda}{r}\right]$$

$$= \binom{x+r-1}{x} \left(1+\frac{\lambda}{r}\right)^{-r} \left(\frac{\lambda}{r}\right)^x \left(1+\frac{\lambda}{r}\right)^{-x}$$

$$= \frac{(x+r-1)(x+r-2)\ldots(r+1)r}{x!} \left(1+\frac{\lambda}{r}\right)^{-r} \left(1+\frac{\lambda}{r}\right)^{-x} \left(\frac{\lambda^x}{r^x}\right)$$

Applying the limit $r \to \infty$ (i.e., condition (ii))

$$= \lim_{r\to\infty} \left[\frac{\left(1+\frac{x-1}{r}\right)\left(1+\frac{x-2}{r}\right)\ldots\left(1+\frac{1}{r}\right)r^x}{x!} \left(1+\frac{\lambda}{r}\right)^{-r} \left(1+\frac{\lambda}{r}\right)^{-x} \frac{\lambda^x}{r^x}\right]$$

$$= \lim_{r\to\infty} \left[\frac{\left(1+\frac{x-1}{r}\right)\left(1+\frac{x-2}{r}\right)\ldots\left(1+\frac{1}{r}\right)}{x!} \left(1+\frac{\lambda}{r}\right)^{-r} \left(1+\frac{\lambda}{r}\right)^{-x} \lambda^x\right]$$

$$= \frac{\lambda^x}{x!}\left[\left(1+\frac{x-1}{r}\right)\left(1+\frac{x-2}{r}\right)\ldots\left(1+\frac{1}{r}\right)\left(1+\frac{\lambda}{r}\right)^{-r}\left(1+\frac{\lambda}{r}\right)^{-x}\right]$$

$$= \frac{\lambda^x}{x!}\left[\lim_{r\to\infty}\left(1+\frac{\lambda}{r}\right)^{-r} \lim_{r\to\infty}\left(1+\frac{\lambda}{r}\right)^{-x}\right]$$

$$= \frac{\lambda^x}{x!}[e^{-\lambda}]$$

$$= \frac{e^{-\lambda}\lambda^x}{x!}$$

Since, the probability mass function of poisson distribution is obtained, it can be said that the poisson is a limiting distribution of negative binomial distribution.

PROBLEMS

Q60. During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

Solution :

Given that,

Let 'X' be the Poisson random variable

The average number of radioactive particles passing through a counter in 1 millisecond is given by,

$\lambda t = 4$

The probability distribution of the Poisson random variable is,

$$P(x;\lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, x = 0, 1, 2, \ldots$$

The probability of 6 particles entering the counter in 1 millisecond is,

$$P(6;4) = \frac{e^{-4}4^6}{6!}$$

$$= \sum_{x=0}^{6} P(x;4) - \sum_{x=0}^{5} P(x;4)$$

$$= 0.8893 - 0.7851$$ [$\because$ From poisson probability sum table]

$$= 0.1042$$

$$\boxed{P(6;4) = 0.1042}$$

Q61. Ten is the average number of oil tankers arriving each day at a certain port. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

Solution :

Given that,

The average number of oil tankers arriving each day at a certain port is, $\lambda t = 10$

Let 'X' be the number of oil tankers arriving each day.

The probability distribution of the Poisson random variable 'X' is,

$$P(x; \lambda t) = \frac{e^{-\lambda t}(\lambda t)^x}{x!}, x = 0, 1, 2, ...$$

The probability of atmost 15 oil tankers arriving each day is,

$$P(X > 15) = 1 - P(X \leq 15)$$

$$= 1 - \sum_{x=0}^{15} P(x;10)$$

$$= 1 - 0.9513$$ [$\because$ From poisson probability sum table]

$$= 0.05$$

$$\boxed{\therefore\ P(X > 15) = 0.05}$$

Q62. In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.

(a) What is the probability that in any given period of 400 days there will be an accident on one day?

(b) What is the probability that there are at most three days with an accident?

Solution :

Given that,

The probability of an accident on any given day is, $p = 0.005$

Here, $n = 400, p = 0.005$

Let 'X' be a binomial random variable.

(a) The Probability of an Accident on One Day in 400 Days

$\mu = np = (400)(0.005)$

$$\boxed{\mu = 2}$$

$$P(X=1)=\frac{e^{-2}(2)^x}{x!}$$
$$=\frac{e^{-2}(2)^1}{1!}$$
$$=\frac{0.14\times 2}{1}$$
$$=0.27$$

$$\boxed{\therefore P(X=1)=0.27}$$

(b) The Probability of an Accident for atmost Three Days

$$P(X\leq 3)=\sum_{x=0}^{3}\frac{e^{-2}2^x}{x!}$$
$$=e^{-2}\left[\sum_{x=0}^{3}\frac{2^x}{x!}\right]$$
$$=e^{-2}\left[\frac{2^0}{0!}+\frac{2^1}{1!}+\frac{2^2}{2!}+\frac{2^3}{3!}\right]$$
$$=e^{-2}\left[1+2+\frac{4}{2}+\frac{8}{6}\right]$$
$$=e^{-2}\left[\frac{3+6+6+4}{3}\right]$$
$$=e^{-2}\left[\frac{19}{3}\right]$$
$$=0.14[6.33]$$
$$=0.8862$$

$$\boxed{\therefore P(X\leq 3)=0.8862}$$

Q63. In a manufacturing process where glass products are made, defects or bubbles occur, occasionally rendering the piece undesirable for marketing. It is known that, on average, 1 in every 1000 of these items produced has one or more bubbles. What is the probability that a random sample of 8000 will yield fewer than 7 items possessing bubbles?

Solution :

Given that,

Random sample, $n=8000$

The average number of items that produces 1 or more bubbles is, $p=\frac{1}{1000}=0.001$

By using $\mu=np$

$$\mu=(8000)(0.001)$$

$$\boxed{\mu=8}$$

The probability of random sample yielding fewer than 7 items possessing bubbles is,

$$P(X<7)=\sum_{x=0}^{6}b(x;n,p)$$
$$=\sum_{x=0}^{6}b(x;8000,0.001)$$
$$\simeq\sum_{x=0}^{6}P(x;8)$$
$$=0.3134 \qquad [\because \text{From poisson probability sums table}]$$

$$\therefore P(X<7)=0.3134$$