Contents

20%

14Marks

14 Marks

Scanned with CamScanner

December

GTU. Winter 2019

Chap = 1, chap. 2 → 14 Marks
Chap 3, chap 4 → 14 Marks
Chap = 5 → 18 Marks
Chap = 6 → 17 Marks
Chap = 7 → 7 Marks

70 Marks.

from:- D.G. BORAD

-: Shreenathji Engineering Zone :-

DGPatel

Scanned with CamScanner

CHAPTER 3

Basic Statistics

Chapter Outline

3.1 INTRODUCTION

A discrete random variable is described by probability function or probability mass function. Similarly, a continuous random variable is described by its probability density function. Instead of a function, a more compact description can be made by a few parameters, known as statistical measures, that are representative of the distribution. In descriptive statistics, statistical measures are used to summarize a set of observations in order to communicate the information as simply as possible. The observations are described in

(i) a measure of location or central tendency, such as arithmetic mean
(ii) a measure of statistical dispersion like standard deviation
(iii) a measure of the shape of the distribution like skewness or kurtosis
(iv) if more than one variable is measured, a measure of statistical dependence such as correlation coefficient.

Scanned with CamScanner

3.2 MEASURES OF CENTRAL TENDENCY

In statistics, a central tendency or measure of central tendency is a central or typical value of a probability distribution. It is also called a center or location of the distribution. Measures of central tendency are often called averages. An average is a single value which can be taken as a representative of the whole distribution. There are five types of measures of central tendency or averages which are commonly used.

(i) Arithmetic mean or mean or expectation
(ii) Median
(iii) Mode
(iv) Geometric mean
(v) Harmonic mean

1. Mean The mean or average value (μ) of the probability distribution of a discrete random variable X is called as expectation and is denoted by $E(X)$.

$$\mu = E(X) = \sum_{i=1}^{\infty} x_i\, p(x_i) = \sum x\, p(x)$$

where $p(x)$ is the probability mass function of the discrete random variable X.
Expectation of any function $\phi(x)$ of a random variable X is given by

$$E[\phi(x)] = \sum_{i=1}^{\infty} \phi(x_i)\, p(x_i) = \sum \phi(x)\, p(x)$$

Some important results on expectation:

(i) $E(X + k) = E(X) + k$
(ii) $E(aX \pm b) = aE(X) \pm b$
(iii) $E(X + Y) = E(X) + E(Y)$ provided $E(X)$ and $E(Y)$ exists.
(iv) $E(XY) = E(X)\, E(Y)$ if X and Y are two independent random variables.

2. Median The median is the point which divides the entire distribution into two equal parts. If X is a random variable, the value of $X = x$ for which the cumulative distribution function $F(x) = \frac{1}{2}$ is called the median of X. For a discrete random variable X, if there exits no x such that $F(x) = \frac{1}{2}$ then the median M of probability distribution is given by

$$M = \frac{1}{2}(x_k + x_{k+1})$$

where $F(x_k) < \frac{1}{2}$ and $F(x_{k+1}) > \frac{1}{2}$ and x_k and x_{k+1} are two consecutive values of X.

3. Mode The mode is the value of discrete random variable X for which the probability is maximum.

4. Geometric Mean The geometric mean G of a random variable X is defined by $\log G = E(\log X)$. The geometric mean of the probability distribution of a discrete random variable X is given by

$$\log G = \sum_{i=1}^{\infty} (\log x_i) p(x_i) = \sum (\log x)\, p(x)$$

where $p(x)$ is the probability mass function of the discrete random variable X.

5. Harmonic Mean The harmonic mean of a random variable X is defined by $\frac{1}{H} = E\left(\frac{1}{X}\right)$. The harmonic mean of the probability distribution of a discrete random variable X is given by

$$\frac{1}{H} = \sum_{i=1}^{\infty} \frac{1}{x_i} p(x_i) = \sum \frac{1}{x} p(x)$$

where $p(x)$ is the probability mass function of the discrete random variable X.

3.3 MEASURES OF DISPERSION

A measnre of central tendency is a representative value of the random variable. But it is important to know how the values are clustered around or scattered away from the measure of central tendency. The property of the random variable or its distribution by which its values are clustered around or scattering away from the central value is called dispersion. There are three types of measures of dispersion which are commonly used.

(i) Quartile Deviation
(ii) Mean Deviation
(iii) Standard Deviation

1. Quartile Deviation Quartile deviation or semi-inter quartile range of the probability distribution of a discrete random variable X is given by

$$Q = \frac{1}{2}(Q_3 - Q_1)$$

where Q_1 and Q_3 are the first and third quartiles of the distribution respectively.

2. Mean Deviation Mean deviation of the probability distribution of a discrete random variable X is given by

$$\begin{aligned} MD &= E\{|X - \mu|\} \\ &= \sum_{i=1}^{\infty} |x_i - \mu|\, p(x_i) \\ &= \sum |x - \mu|\, p(x) \end{aligned}$$

where $p(x)$ is the probability mass function of the discrete random variable X.

Scanned with CamScanner

3. Standard Deviation *Standard deviation* is the positive square root of the arithmetic mean of the squares of the deviations of the given values from their arithmetic mean. It is denoted by the Greek letter σ.

$$\text{SD} = \sigma = \sqrt{\sum_{i=1}^{\infty} x_i^2\, p(x_i) - \mu^2}$$

$$= \sqrt{E(X^2) - \mu^2}$$

$$= \sqrt{E(X^2) - [E(X)]^2}$$

Variance Variance characterizes the variability in the distributions since two distributions with same mean can still have different dispersion of data about their means. Variance of the probability distribution of a discrete random variable X is given by

$$\begin{aligned}
\text{Var}(X) = \sigma^2 &= E(X-\mu)^2 \\
&= E(X^2 - 2X\mu + \mu^2) \\
&= E(X^2) - E(2X\mu) + E(\mu^2) \\
&= E(X^2) - 2\mu\, E(X) + \mu^2 \qquad [\because E(\text{constant}) = (\text{constant})] \\
&= E(X^2) - 2\mu\mu + \mu^2 \\
&= E(X^2) - \mu^2 \\
&= E(X^2) - [E(X)]^2
\end{aligned}$$

Some important results on variance:

(i) $\text{Var}(k) = 0$

(ii) $\text{Var}(kX) = k^2\, \text{Var}(X)$

(iii) $\text{Var}(X + k) = \text{Var}(X)$

(iv) $\text{Var}(aX + b) = a^2\, \text{Var}(X)$

Example 1

A random variable X has the following distribution:

X	1	2	3	4	5	6
$P(X=x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

Find (i) mean, (ii) variance, and (iii) $P(1 < X < 6)$.

Solution

(i) $\text{Mean} = \mu = \sum x p(x)$

$$= 1\left(\frac{1}{36}\right) + 2\left(\frac{3}{36}\right) + 3\left(\frac{5}{36}\right) + 4\left(\frac{7}{36}\right) + 5\left(\frac{9}{36}\right) + 6\left(\frac{11}{36}\right)$$

$$= \frac{161}{36}$$

$$= 4.47$$

(ii) $\text{Variance} = \sigma^2 = \sum x^2 p(x) - \mu^2$

$$= 1\left(\frac{1}{36}\right) + 4\left(\frac{3}{36}\right) + 9\left(\frac{5}{36}\right) + 16\left(\frac{7}{36}\right) + 25\left(\frac{9}{36}\right) + 36\left(\frac{11}{36}\right) - (4.47)^2$$

$$= \frac{791}{36} - 19.98$$

$$= 1.99$$

(iii) $P(1 < X < 6) = P(X=2) + P(X=3) + P(X=4) + P(X=5)$

$$= \frac{3}{36} + \frac{5}{36} + \frac{7}{36} + \frac{9}{36}$$

$$= \frac{24}{36}$$

$$= 0.67$$

Example 2

The probability distribution of a random variable X is given below. Find (i) E(X), (ii) Var(X), *(iii) E(2X − 3), and (iv)* Var (2X − 3)

X	−2	−1	0	1	2
$P(X=x)$	0.2	0.1	0.3	0.3	0.1

Solution

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

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

$$= 0$$

(ii) $\text{Var}(X) = \sum x^2\, p(x) - [E(X)]^2$

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

$$= 1.6$$

(iii) $E(2X - 3) = 2E(X) - 3$

$$= 2(0) - 3$$

$$= -3$$

(iv) $\text{Var}(2X - 3) = (2)^2\, \text{Var}(X)$

$$= 4(1.6)$$

$$= 6.4$$

Scanned with CamScanner

Example 3

Mean and standard deviation of a random variable X are 5 and 4 respectively. Find $E(X^2)$ and standard deviation of $(5-3X)$.

Solution

$$E(X) = \mu = 5$$
$$SD = \sigma = 4$$
$$\therefore \quad Var(X) = \sigma^2 = 16$$
$$Var(X) = E(X^2) - [E(X)]^2$$
$$16 = E(X^2) - (5)^2$$
$$\therefore \quad E(X^2) = 41$$
$$Var\,(5-3X) = Var\,(5) - (-3)^2\,Var\,(X)$$
$$= 0 + 9(16)$$
$$= 144$$
$$SD\,(5-3X) = \sqrt{Var\,(5-3X)}$$
$$= \sqrt{144}$$
$$= 12$$

Example 4

A machine produces an average of 500 items during the first week of the month and on average of 400 items during the last week of the month, the probability for these being 0.68 and 0.32 respectively. Determine the expected value of the production. **[Summer 2015]**

Solution

Let X be the random variable which denotes the items produced by the machine. The probability distribution is

X	500	400
$P(X=x)$	0.68	0.32

Expected value of the production $E(X) = \sum x\,p(x)$

$$= 500(0.68) + 400(0.32)$$
$$= 468$$

Example 5

The monthly demand for Allwyn watches is known to have the following probability distribution:

Demand (x)	1	2	3	4	5	6	7	8
Probability $p(x)$	0.08	0.12	0.19	0.24	0.16	0.10	0.07	0.04

Find the expected demand for watches. Also, compute the variance.

Solution

$$E(X) = \Sigma\, x\,p(x)$$
$$= 1(0.08) + 2(0.12) + 3(0.19) + 4(0.24) + 5(0.16) + 6(0.10) + 7(0.07) + 8(0.04)$$
$$= 4.06$$
$$Var(X) = E(X^2) - [E(X)]^2$$
$$= \Sigma\, x^2\,p(x) - [E(X)]^2$$
$$= 1(0.08) + 4(0.12) + 9(0.19) + 16(0.24) + 25(0.16) + 36(0.10) + 49(0.07) + 64(0.04) - (4.06)^2$$
$$= 19.7 - 16.48$$
$$= 3.21$$

Example 6

A discrete random variable has the probability mass function given below:

X	−2	−1	0	1	2	3
$P(X=x)$	0.2	k	0.1	$2k$	0.1	$2k$

Find k, mean, and variance.

Solution

Since $P(X=x)$ is a probability mass function,

$$\sum P(X=x) = 1$$
$$0.2 + k + 0.1 + 2k + 0.1 + 2k = 1$$
$$5k + 0.4 = 1$$
$$5k = 0.6$$
$$k = \frac{0.6}{5} = \frac{3}{25}$$

Hence, the probability distribution is

X	−2	−1	0	1	2	3
$P(X=x)$	$\frac{2}{10}$	$\frac{3}{25}$	$\frac{1}{10}$	$\frac{6}{25}$	$\frac{1}{10}$	$\frac{6}{25}$

Scanned with CamScanner

Mean $= E(X) = \sum x\, p(x)$

$$= (-2)\left(\frac{2}{10}\right) + (-1)\left(\frac{3}{25}\right) + 0 + 1\left(\frac{6}{25}\right) + 2\left(\frac{1}{10}\right) + 3\left(\frac{6}{25}\right)$$

$$= \frac{6}{25}$$

Variance $= \mathrm{Var}(X) = E(X^2) - [E(X)]^2$

$$= \sum x^2\, p(x) - [E(X)]^2$$

$$= 4\left(\frac{2}{10}\right) + 1\left(\frac{3}{25}\right) + 0 + 1\left(\frac{6}{25}\right) + 4\left(\frac{1}{10}\right) + 9\left(\frac{6}{25}\right) - \left(\frac{6}{25}\right)^2$$

$$= \frac{73}{250} - \frac{36}{625}$$

$$= \frac{293}{625}$$

Example 7

A random variable X has the following probability function:

x	0	1	2	3	4	5	6	7
$p(x)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

(i) *Determine k.* (ii) *Evaluate* $P(X < 6)$, $P(X \geq 6)$, $P(0 < X < 5)$ and $P(0 \leq X \leq 4)$. (iii) *Determine the distribution function of X.* (iv) *Find the mean.* (v) *Find the variance.*

Solution

(i) Since $p(x)$ is a probability mass function,

$$\sum p(x) = 1$$

$$0 + k + 2k + 2k + 3k + k^2 + 2k^2 + 7k^2 + k = 1$$

$$10k^2 + 9k - 1 = 1$$

$$(10k - 1)(k + 1) = 0$$

$$k = \frac{1}{10} \text{ or } k = -1$$

$$k = \frac{1}{10} = 0.1 \left[\because\ p(x) \geq 0,\ k \neq -1\right]$$

Hence, the probability function is

X	0	1	2	3	4	5	6	7
$P(X = x)$	0	0.1	0.2	0.2	0.3	0.01	0.02	0.17

(ii) $P(X < 6) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$

$$= 0 + 0.1 + 0.2 + 0.2 + 0.3 + 0.01$$

$$= 0.81$$

$P(X \geq 6) = 1 - P(X < 6)$

$$= 1 - 0.81$$

$$= 0.19$$

$P(0 < X < 5) = P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= 0.1 + 0.2 + 0.2 + 0.3$$

$$= 0.8$$

$P(0 \leq X \leq 4) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$

$$= 0 + 0.1 + 0.2 + 0.2 + 0.3$$

$$= 0.8$$

(iii) Distribution function of X

x	$p(x)$	$F(x)$
0	0	0
1	0.1	0.1
2	0.2	0.3
3	0.2	0.5
4	0.3	0.8
5	0.01	0.81
6	0.02	0.83
7	0.17	1

(iv) $\mu = \sum x p(x)$

$$= 0 + 1(0.1) + 2(0.2) + 3(0.2) + 4(0.3) + 5(0.01) + 6(0.02) + 7(0.17)$$

$$= 3.66$$

(v) $\mathrm{Var}(X) = \sigma^2 = \sum x^2\, p(x) - \mu^2$

$$= 0 + 1(0.1) + 4(0.2) + 9(0.2) + 16(0.3) + 25(0.01) + 36(0.02) + 49(0.17) - (3.66)^2$$

$$= 3.4044$$

Scanned with CamScanner

Example 8

A fair dice is tossed. Let the random variable X denote the twice the number appearing on the dice. Write the probability distribution of X. Calculate mean and variance.

Solution

Let X be the random variable which denotes twice the number appearing on the dice.

(i) Probability distribution of X

x	2	4	6	8	10	12
$p(x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

(ii) Mean $= \mu = \sum xp(x)$

$$= 2\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) + 8\left(\frac{1}{6}\right) + 10\left(\frac{1}{6}\right) + 12\left(\frac{1}{6}\right)$$

$$= 7$$

(iii) Variance $= \sigma^2 = \sum x^2 p(x) - \mu^2$

$$= 4\left(\frac{1}{6}\right) + 16\left(\frac{1}{6}\right) + 36\left(\frac{1}{6}\right) + 64\left(\frac{1}{6}\right) + 100\left(\frac{1}{6}\right) + 144\left(\frac{1}{6}\right) - (7)^2$$

$$= 11.67$$

Example 9

Two unbiased dice are thrown at random. Find the probability distribution of the sum of the numbers on them. Also, find mean and variance.

Solution

Let X be the random variable which denotes the sum of the numbers on two unbiased dice. The random variable X can take values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12. The probability distribution is

X	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Mean $= \mu = \sum xp(x)$

$$= 2\left(\frac{1}{36}\right) + 3\left(\frac{2}{36}\right) + 4\left(\frac{3}{36}\right) + 5\left(\frac{4}{36}\right) + 6\left(\frac{5}{36}\right) + 7\left(\frac{6}{36}\right) + 8\left(\frac{5}{36}\right) + 9\left(\frac{4}{36}\right) + 10\left(\frac{3}{36}\right) + 11\left(\frac{2}{36}\right) + 12\left(\frac{1}{36}\right)$$

$$= \frac{252}{36}$$

$$= 7$$

Variance $= \sigma^2 = \sum x^2 p(x) - \mu^2$

$$= 4\left(\frac{1}{36}\right) + 9\left(\frac{2}{36}\right) + 16\left(\frac{3}{36}\right) + 25\left(\frac{4}{36}\right) + 36\left(\frac{5}{36}\right) + 49\left(\frac{6}{36}\right) + 64\left(\frac{5}{36}\right) + 81\left(\frac{4}{36}\right) + 100\left(\frac{3}{36}\right) + 121\left(\frac{2}{36}\right) + 144\left(\frac{1}{36}\right) - (7)^2$$

$$= \frac{1974}{36} - 49$$

$$= 5.83$$

Example 10

A sample of 3 items is selected at random from a box containing 10 items of which 4 are defective. Find the expected number of defective items.

Solution

Let X be the random variable which denotes the defective items.

Total number of items = 10

Number of good items = 6

Number of defective items = 4

$$P(X=0) = P(\text{no defective item}) = \frac{{}^6C_3}{{}^{10}C_3} = \frac{1}{6}$$

$$P(X=1) = P(\text{one defective item}) = \frac{{}^6C_2\ {}^4C_1}{{}^{10}C_3} = \frac{1}{2}$$

$$P(X=2) = P(\text{two defective items}) = \frac{{}^6C_1\ {}^4C_2}{{}^{10}C_3} = \frac{3}{10}$$

$$P(X=3) = P(\text{three defective items}) = \frac{{}^4C_3}{{}^{10}C_3} = \frac{1}{30}$$

Scanned with CamScanner

Hence, the probability distribution is

X	0	1	2	3
$P(X=x)$	$\frac{1}{6}$	$\frac{1}{2}$	$\frac{3}{10}$	$\frac{1}{30}$

Expected number of defective items $= E(X) = \sum x\, p(x)$

$$= 0 + 1\left(\frac{1}{2}\right) + 2\left(\frac{3}{10}\right) + 3\left(\frac{1}{30}\right)$$

$$= 1.2$$

Example 11

A player tosses two fair coins. He wins ₹ 100 if a head appears and ₹ 200 if two heads appear. On the other hand, he loses ₹ 500 if no head appears. Determine the expected value of the game. Is the game favourable to the players?

Solution

Let X be the random variable which denotes the number of heads appearing in tosses of two fair coins.

$$S = \{\text{HH, HT, TH, TT}\}$$

$$p(x_1) = P(X=0) = P(\text{no heads}) = \frac{1}{4}$$

$$p(x_2) = P(X=1) = P(\text{one head}) = \frac{2}{4} = \frac{1}{2}$$

$$p(x_3) = P(X=2) = P(\text{two heads}) = \frac{1}{4}$$

Amount to be lost if no head appears $= x_1 = -₹\ 500$

Amount to be won if one head appears $= x_2 = ₹\ 100$

Amount to be won if two heads appear $= x_3 = ₹\ 200$

Expected value of the game $= \mu = \sum x\, p(x)$

$$= x_1\, p(x_1) + x_2\, p(x_2) + x_3\, p(x_3)$$

$$= -500\left(\frac{1}{4}\right) + 100\left(\frac{1}{2}\right) + 200\left(\frac{1}{4}\right)$$

$$= ₹ -25$$

Hence, the game is not favourable to the player.

Example 12

Amit plays a game of tossing a dice. If a number less than 3 appears, he gets ₹ a, otherwise he has to pay ₹ 10. If the game is fair, find a.

Solution

Let X be the random variable which denotes tossing of a dice.

Probability of getting a number less than 3, i.e., 1 or 2 $= p(x_1) = \frac{2}{6} = \frac{1}{3}$

Probability of getting number more than or equal to 3, i.e., 3, 4, 5, or 6 $= p(x_2) = \frac{4}{6} = \frac{2}{3}$

Amount to be received for number less than 3 $= x_1 = ₹\ a$

Amount to be paid for numbers more than or equal to 3 $= x_2 = ₹ -10$

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

$$= x_1\, p(x_1) + x_2\, p(x_2)$$

$$= a\left(\frac{1}{3}\right) + (-10)\left(\frac{2}{3}\right)$$

$$= \frac{a}{3} - \frac{20}{3}$$

For a pair game, $E(x) = 0$.

$$\frac{a}{3} - \frac{20}{3} = 0$$

$$a = 20$$

Example 13

A man draws 2 balls from a bag containing 3 white and 5 black balls. If he is to receive ₹ 14 for every white ball which he draws and ₹ 7 for every black ball, what is his expectation?

Solution

Let X be the random variable which denotes the balls drawn from a bag. 2 balls drawn may be either (i) both white, or (ii) both black, or (iii) one white and one black.

Probability of drawing 2 white balls $= p(x_1) = \frac{^3C_2}{^8C_2} = \frac{3}{28}$

Probability of drawing 2 black balls $= p(x_2) = \frac{^5C_2}{^8C_2} = \frac{10}{28}$

Scanned with CamScanner

Probability of drawing 1 white and 1 black ball = $p(x_3) = \frac{{}^3C_1\,{}^5C_1}{{}^8C_2} = \frac{15}{28}$

Amount to be received for 2 white balls = x_1 = ₹ 14 × 2 = ₹ 28

Amount to be received for 2 black balls = x_2 = ₹ 7 × 2 = ₹ 14

Amount to be received for 1 white and 1 black ball = x_3 = ₹ 14 + ₹ 7 = ₹ 21

$$\text{Expectation} = E(X) = \sum x\,p(x)$$
$$= x_1\,p(x_1) + x_2\,p(x_2) + x_3\,p(x_3)$$
$$= 28\left(\frac{3}{28}\right) + 14\left(\frac{10}{28}\right) + 21\left(\frac{15}{28}\right)$$
$$= ₹\ 19.25$$

Example 14

The probability that there is at least one error in an account statement prepared by A is 0.2 and for B and C, they are 0.25 and 0.4 respectively. A, B, and C prepared 10, 16, and 20 statements respectively. Find the expected number of correct statements in all.

Solution

Let $p(x_1)$, $p(x_2)$ and $p(x_3)$ be the probabilities of the events that there is no error in the account statements prepared by A, B, and C respectively.

$$p(x_1) = 1 - (\text{Probability of at least one error in the account statement prepared by } A)$$
$$= 1 - 0.2$$
$$= 0.8$$

Similarly, $p(x_2) = 1 - 0.25 = 0.75$

$p(x_3) = 1 - 0.4 = 0.6$

Also, $x_1 = 10$, $x_2 = 16$, $x_3 = 20$

$$\text{Expected number of correct statements} = E(X) = \sum x\,p(x)$$
$$= x_1\,p(x_1) + x_2\,p(x_2) + x_3\,p(x_3)$$
$$= 10(0.8) + 16(0.75) + 20(0.6)$$
$$= 32$$

Example 15

A man has the choice of running either a hot-snack stall or an ice-cream stall at a seaside resort during the summer season. If it is a fairly cool summer, he should make ₹ 5000 by running the hot-snack stall, but if the summer is quite hot, he can only expect to make ₹ 1000. On the other hand, if he operates the ice-cream stall, his profit is estimated at ₹ 6500, if the summer is hot, but only ₹ 1000 if it is cool. There is a 40 percent chance of the summer being hot. Should he opt for running the hot-snack stall or the ice-cream stall?

Solution

Let X and Y be the random variables which denote the income from the hot-snack and ice-cream stalls respectively.

Probability of hot summer = p_1 = 40% = 0.4

Probability of cool summer = $p_2 = 1 - p_1 = 1 - 0.4 = 0.6$

$x_1 = 1000$, $x_2 = 5000$, $y_1 = 6500$, $y_2 = 1000$

$$\text{Expected income from hot-snack stall} = E(X)$$
$$= x_1 p_1 + x_2 p_2$$
$$= 1000(0.4) + 5000(0.6)$$
$$= ₹\ 3400$$

$$\text{Expected income from ice-cream stall} = E(Y)$$
$$= y_1 p_1 + y_2 p_2$$
$$= 6500(0.4) + 1000(0.6)$$
$$= ₹\ 3200$$

Hence, he should opt for running the hot-snack stall.

EXERCISE 3.1

1. The probability distribution of a random variable X is given by

X	-2	-1	0	1	2	3
$P(X = x)$	0.1	k	0.2	$2k$	0.3	k

Find k, the mean, and variance. [Ans.: 0.1, 0.8, 2.16]

2. Find the mean and variance of the following distribution:

X	4	5	6	8
$P(X = x)$	0.1	0.3	0.4	0.2

[Ans.: 5.9, 1.49]

Scanned with CamScanner

3. Find the value of k from the following data:

X	0	10	15
$P(X = x)$	$\frac{k-6}{5}$	$\frac{2}{k}$	$\frac{14}{5k}$

Also, find the distribution function and expectation of X.

Ans.: 8,

X	0	10	15
$F(X)$	$\frac{2}{5}$	$\frac{13}{20}$	1

, $\frac{31}{4}$

4. For the following distribution,

X	−3	−2	−1	0	1	2
$P(X = x)$	0.01	0.1	0.2	0.3	0.2	0.15

Find (i) $P(X \geq 1)$, (ii) $P(X < 0)$, (iii) $E(X)$, and (iv) $\text{Var}(X)$

[Ans.: (i) 0.35 (ii) 0.35 (iii) 0.05 (iv) 1.8475]

5. A random variable X has the following probability function:

X	0	1	2	3	4	5	6	7	8
$P(X = x)$	$\frac{k}{45}$	$\frac{k}{15}$	$\frac{k}{9}$	$\frac{k}{5}$	$\frac{2k}{45}$	$\frac{6k}{45}$	$\frac{7k}{45}$	$\frac{8k}{45}$	$\frac{4k}{45}$

Determine (i) k, (ii) mean, (iii) variance, and (iv) SD.

[Ans.: (i) 1 (ii) 0.4622 (iii) 4.9971 (iv) 2.24]

6. A fair coin is tossed until a head or five tails appear. Find (i) discrete probability distribution, and (ii) mean of the distribution.

Ans.: (i)

X	1	2	3	4	5
$P(X = x)$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{16}$

(ii) 1.9

7. Let X denotes the minimum of two numbers that appear when a pair of fair dice is thrown once. Determine (i) probability distribution, (ii) expectation, and (iii) variance.

Ans.: (i)

X	1	2	3	4	5	6
$P(X = x)$	$\frac{11}{36}$	$\frac{9}{36}$	$\frac{7}{36}$	$\frac{5}{36}$	$\frac{3}{36}$	$\frac{1}{36}$

(ii) 2.5278 (iii) 1.9713

8. For the following probability distribution,

X	−3	−2	−1	0	1	2	3
$P(X = x)$	0.001	0.01	0.1	?	0.1	0.01	0.001

Find (i) missing probability, (ii) mean, and (iii) variance.

[Ans.: (i) 0.778 (ii) 0.2 (iii) 0.258]

9. A discrete random variable can take all integer values from 1 to k each with the probability of $\frac{1}{k}$. Show that its mean and variance are $\frac{k+1}{2}$ and $\frac{k^2+1}{2}$ respectively.

10. An urn contains 6 white and 4 black balls; 3 balls are drawn without replacement. What is the expected number of black balls that will be obtained?

[Ans.: $\frac{6}{5}$]

11. A six-faced dice is tossed. If a prime number occurs, Anil wins that number of rupees but if a nonprime number occurs, he loses that number of rupees. Determine whether the game is favourable to the player.

[Ans.: The game is favourable to Anil]

12. A man runs an ice-cream parlour at a holiday resort. If the summer is mild, he can sell 2500 cups of ice cream; if it is hot, he can sell 4000 cups; if it is very hot, he can sell 5000 cups. It is known that for any year, the probability of summer to be mild is $\frac{1}{7}$ and to be hot is $\frac{4}{7}$. A cup of ice cream costs ₹ 2 and is sold for ₹ 3.50. What is his expected profit?

[Ans.: ₹ 6107.14]

13. A player tosses two fair coins. He wins ₹ 1 or ₹ 2 as 1 tail or 1 head appears. On the other hand, he loses ₹ 5 if no head appears. Find the expected gain or loss of the player.

[Ans.: Loss of ₹ 0.25]

Scanned with CamScanner

14. A bag contains 2 white balls and 3 black balls. Four persons A, B, C, D in the order named each draws one ball and does not replace it. The first to draw a white ball receives ₹ 20. Determine their expectations.

[Ans.: ₹ 8, ₹ 6, ₹ 4, ₹ 2]

3.4 MOMENTS

Moment is the arithmetic mean of the various powers of the deviations of items from their assumed mean or actual mean. If the deviations of the items are taken from the arithmetic mean of the distribution, it is known as *central moment*. If the mean of the first power of deviations are taken, the first moment about the mean is obtained and is denoted by μ_1. The mean of the second power of the deviations gives the second moment about the mean and is denoted by μ_2. Similarly, the mean of the cubes of deviations gives third moment about the mean and is denoted by μ_3. The mean of the fourth power of the deviations from the mean gives the fourth moment about the mean and is denoted by μ_4. Thus, the mean of the r^{th} power of deviations gives the r^{th} moment about mean or r^{th} central moment and is denoted by μ_r.

3.4.1 Central Moments or Moments about Actual Mean

The moments about the mean value $\mu = E(X)$ are called central moments and denoted by μ_r

$$\mu_r = E\{(x-\mu)^r\}$$

$$= \sum_{i=1}^{\infty}(x_i-\mu)^r p(x_i)$$

$$= \sum (x-\mu)^r p(x)$$

If frequency distribution is given and $n = \sum f$, then $p(x_i) = \dfrac{\sum f_i}{N}$

$$\mu_r = \frac{\sum_{i=1}^{\infty} f_i(x_i-\mu)^r}{N}$$

3.4.2 Properties of Central Moments

(i) The first moment about the mean is always zero, i.e., $\mu_1 = 0$.

(ii) The second moment about the mean measures variance, i.e.,

$$\mu_2 = \sigma^2 \quad \text{or} \quad \text{SD} = \sigma = \pm\sqrt{\mu_2}$$

(iii) The third moment about the mean measures skewness.

If $\mu_3 > 0$, the distribution is positively skewed.
If $\mu_3 < 0$, the distribution is negatively skewed.
If $\mu_3 = 0$, the distribution is symmetrical.

$$\text{Skewness } \beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

(iv) The fourth moment about the mean measures kurtosis. It gives information on the peakedness or height of the peak of a frequency distribution, i.e., whether it is more peaked or more flat topped than a normal curve.

$$\text{Kurtosis } \beta_2 = \frac{\mu_4}{\mu_2^2}$$

(v) In a symmetric distribution, all odd moments are zero, i.e., $\mu_1 = \mu_3 = \mu_5 = \cdots = \mu_{2r+1} = 0$.

3.4.3 Raw Moments or Moments about Arbitrary Origin

When the actual mean of a distribution is a fraction, it is tedious to calculate central moments. In such cases, moments about an arbitrary origin 'a' is calculated and then these moments are converted into the moments about actual mean. The moments about the arbitrary origin are known as raw moments and are denoted by μ_r'. Thus, μ_1' denotes the first moment about an arbitrary origin, μ_2' denotes the second moment about an arbitrary origin and so on.

$$\mu_r' = E\{(X-a)^r\}$$

$$= \sum_{i=1}^{\infty}(x_i-a)^r p(x_i)$$

$$= \sum (x-a)^r p(x)$$

If frequency distribution is given and $n = \sum f$, then $p(x_i) = \dfrac{\sum f_i}{N}$

$$\mu_r' = \frac{\sum_{i=1}^{\infty} f_i(x_i-a)^r}{N}$$

When $a = 0$, μ_r' is called r^{th} order simple moments.

$$\mu_r' = E\{X^r\}$$

$$= \sum_{i=1}^{\infty} x_i^r p(x_i)$$

Scanned with CamScanner

$$= \sum x^r p(x)$$

$$= \frac{\sum fx^r}{n}$$

3.4.4 Relation between Central Moments and Raw Moments

The moments about the actual mean, i.e., central moments and moments about the arbitrary origin, i.e., raw moments are related with each other by the following equations:

First central moment $\mu_1 = \mu_1' - \mu_1' = 0$

Second central moment $\mu_2 = \mu_2' - (\mu_1')^2$

Third central moment $\mu_3 = \mu_3' - 3\mu_2' \mu_1' + 2(\mu_1')^3$

Fourth central moment $\mu_4 = \mu_4' - 4\mu_3' \mu_1' + 6\mu_2' (\mu_1')^2 - 3(\mu_1')^4$

Similarly, the raw moments can be expressed in terms of central moments.

First raw moment $\mu_1' = \mu - a$

Second raw moment $\mu_2' = \mu_2 + (\mu_1')^2$

Third raw moment $\mu_3' = \mu_3 + 3\mu_2 \mu_1' + (\mu_1')^3$

Fourth raw moment $\mu_4' = \mu_4 + 4\mu_3 \mu_1' + 6\mu_2(\mu_1')^2 + (\mu_1')^4$

Example 1

Calculate the first four moments from the following data:

x	0	1	2	3	4	5	6	7	8
f	5	10	15	20	25	20	15	10	5

Also, calculate the values of β_1 and β_2.

Solution

$$N = \sum f = 125$$

$$\bar{x} = \frac{\sum fx}{N} = \frac{500}{125} = 4$$

x	f	fx	$x-\mu$	$f(x-\mu)$	$f(x-\mu)^2$	$f(x-\mu)^3$	$f(x-\mu)^4$
0	5	0	–4	20	80	–320	1280
1	10	10	–3	–30	90	–270	810
2	15	30	–2	–30	60	–120	240
3	20	60	–1	–20	20	–20	20
4	25	100	0	0	0	0	0
5	20	100	1	20	20	20	20
6	15	90	2	30	60	120	240
7	10	70	3	30	90	270	810
8	5	40	4	20	80	320	1280
	$\sum f = 125$	$\sum fx = 500$		$\sum f(x-\mu) = 0$	$\sum f(x-\mu)^2 = 500$	$\sum f(x-\mu)^3 = 0$	$\sum f(x-\mu)^4 = 4700$

Moments about the actual mean:

$$\mu_1 = \frac{\sum f(x-\mu)}{N} = \frac{0}{125} = 0$$

$$\mu_2 = \frac{\sum f(x-\mu)^2}{N} = \frac{500}{125} = 4$$

$$\mu_3 = \frac{\sum f(x-\mu)^3}{N} = \frac{0}{125} = 0$$

$$\mu_4 = \frac{\sum f(x-\mu)^4}{N} = \frac{4700}{125} = 37.6$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0}{64} = 0$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{37.6}{16} = 2.35$$

Example 2

Calculate the first four moments of the following distribution about the mean:

x	0	1	2	3	4	5	6	7	8
f	1	8	28	56	70	56	28	8	1

Also, evaluate β_1 and β_2.

Scanned with CamScanner

Solution

Let $a = 4$ be the arbitrary origin.

x	f	$x-a$	$f(x-a)$	$f(x-a)^2$	$f(x-a)^3$	$f(x-a)^4$
0	1	−4	−4	16	−64	256
1	8	−3	−24	72	−216	648
2	28	−2	−56	112	−224	448
3	56	−1	−56	56	−56	56
4	70	0	0	0	0	0
5	56	1	56	56	56	56
6	28	2	56	112	224	448
7	8	3	24	72	216	648
8	1	4	4	16	64	256
	$\sum f = 256$		$\sum f(x-a) = 0$	$\sum f(x-a)^2 = 512$	$\sum f(x-a)^3 = 0$	$\sum f(x-a)^4 = 2816$

$$N = \sum f = 256$$

Moments about the arbitrary origin:

$$\mu_1' = \frac{\sum f(x-a)}{N} = \frac{0}{256} = 0$$

$$\mu_2' = \frac{\sum f(x-a)^2}{N} = \frac{512}{256} = 2$$

$$\mu_3' = \frac{\sum f(x-a)^3}{N} = \frac{0}{256} = 0$$

$$\mu_4' = \frac{\sum f(x-a)^4}{N} = \frac{2816}{256} = 11$$

Moments about the actual mean:

$$\mu_1 = 0$$

$$\begin{aligned}\mu_2 &= \mu_2' - (\mu_1')^2 \\ &= 2 - 0 \\ &= 2\end{aligned}$$

$$\begin{aligned}\mu_3 &= \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3 \\ &= 0 - 3(2)(0) + 2(0)^3 \\ &= 0\end{aligned}$$

$$\begin{aligned}\mu_4 &= \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4 \\ &= 11 - 4(0)(0) + 6(2)(0)^2 - 3(0)^4 \\ &= 11\end{aligned}$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{11}{(2)^2} = 2.75$$

Example 3

The first four moments of distribution about $x = 2$ are 1, 2.5, 5.5, and 16. Calculate the four moments about μ.

Solution

$$\mu_1' = 1, \quad \mu_2' = 2.5, \quad \mu_3' = 5.5, \quad \mu_4' = 16$$

Moments about the mean:

$$\mu_1 = 0$$

$$\begin{aligned}\mu_2 &= \mu_2' - (\mu_1')^2 \\ &= 2.5 - (1)^2 \\ &= 1.5\end{aligned}$$

$$\begin{aligned}\mu_3 &= \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3 \\ &= 5.5 - 3(2.5)(1) + 2(1)^3 \\ &= 0\end{aligned}$$

$$\begin{aligned}\mu_4 &= \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4 \\ &= 16 - 4(5.5)(1) + 6(2.5)(1)^2 - 3(1)^4 \\ &= 6\end{aligned}$$

Example 4

The first three moments of a distribution about the value 2 of the variables are 1, 16, and −40. Show that the mean = 3, variance = 15 and $\mu_3 = -86$.

Solution

$$a = 2, \quad \mu_1' = 1, \quad \mu_2' = 16, \quad \mu_3' = 16, \quad \mu_3' = -40$$

Scanned with CamScanner

$$\mu_1' = \mu - a$$
$$1 = \mu - 2$$
$$\therefore \quad \mu = 3$$

Mean = 3

$$\mu_2 = \mu_2' - (\mu_1')^2$$
$$= 16 - (1)^2$$
$$= 15$$

Variance = μ_2 = 15

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$
$$= -40 - 3(16)(1) + 2(1)^3$$
$$= -86$$

EXERCISE 3.2

1. Calculate the first four moments about the mean from the following data:

x	1	2	3	4	5
f	2	3	5	4	1

[Ans.: 0, 1.262, 0.722, 3.795]

2. Calculate the first four moments about the mean and also the value of β_2 from the following table:

x	0	1	2	3	4	5	6	7	8
f	1	8	28	156	170	56	28	8	1

[Ans.: 0, 1.294, 0.642, 0.582, 3.93]

3. The first four moments of a distribution about the value 4 of the variables are 1, 4, 10, and 45. Show that the mean = 5, variance = 3, and μ_3 = 0.
4. The first four central moments of a distribution are 0, 2.5, 0.7, and 18.75. Calculate β_1 and β_2.

[Ans.: 0.031, 3]

5. The values of μ_1, μ_2, μ_3 and μ_4 are 0, 9.2, 3.6, and 1.22 respectively. Find skewness and kurtosis of the distribution.

[Ans.: 0.129, 1.4]

6. The first four moments about the working mean 28.5 of a distribution are 0.294, 7.144, 14.409, and 454.98. Calculate the moments about the mean. Also, evaluate β_1 and β_2.

[Ans.: 28.794, 7.058, 36.151, 408.738, 3.717, 8.205]

3.5 SKEWNESS

Skewness is a measure that refers to the extent of symmetry or asymmetry in a distribution. A distribution is said to be *symmetrical* when its mean, median, and mode are equal, and the frequencies are symmetrically distributed about the mean. A symmetrical distribution when plotted on a graph will give a perfectly bell-shaped curve which is known as a *normal curve* (Fig. 3.1).

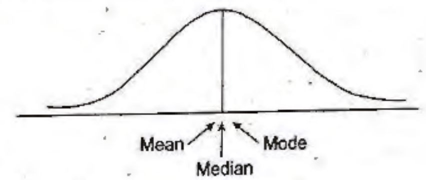
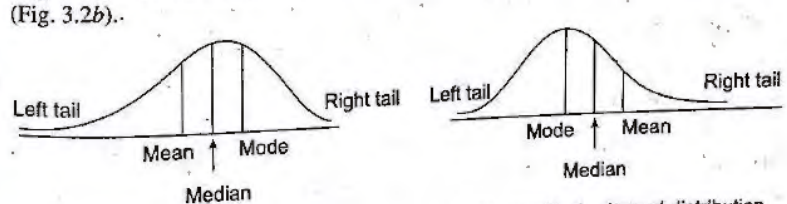


Fig. 3.1

A distribution is said to be asymmetrical or skewed when the mean, median, and mode are not equal, i.e., the mean median, and mode do not coincide. If the curve has a longer tail towards the left, it is said to be a negatively skewed distribution (Fig. 3.2*a*). If the curve has a longer tail towards the right, it is said to be positively skewed (Fig. 3.2*b*)..



(a) Negatively skewed distribution

(b) Positively skewed distribution

Fig. 3.2

Skewness gives an idea of the nature and degree of concentration of observations about the mean.

3.5.1 Measures of Skewness

A measure of skewness gives the extent and direction of skewness of a distribution. These measures can be absolute or relative. The absolute measures are also known as measures of skewness.

Absolute skewness = Mean – Mode

Scanned with CamScanner

If the value of the mean is greater than the mode, the skewness will be positive and if the value of the mean is less than the mode, the skewness will be negative.

The relative measures of skewness is called the *coefficient of skewness*.

3.5.2 Karl Pearson's Coefficient of Skewness

Karl Pearson's coefficient of skewness denoted by S_k, is given by

$$S_k = \frac{\text{Mean} - \text{Mode}}{\text{Standard Deviation}} = \frac{\text{Mean} - \text{Mode}}{\sigma}$$

When the mode is ill-defined and the distribution is moderately skewed, the averages have the following relationship:

$$\text{Mode} = 3\ \text{Median} - 2\ \text{Mean}$$

$$S_k = \frac{\text{Mean} - (3\ \text{Median} - 2\ \text{Mean})}{\text{Standard Deviation}} = \frac{3(\text{Mean} - \text{Median})}{\text{Standard Deviation}} = \frac{3(\text{Mean} - \text{Median})}{\sigma}$$

The coefficient of skewness usually lies between –1 and 1.

For a positively skewed distribution, $S_k > 0$.

For a negatively skewed distribution, $S_k < 0$.

For a symmetrical distribution, $S_k = 0$.

3.6 KURTOSIS

Measures of central tendency, dispersion and skewness of a random variable cannot give a complete idea about the probability distribution. In order to analyse the probability distribution completely, another characteristic, Kurtosis is required. Kurtosis means the convexity of the probability curve of the distribution. It measures the degree of peakedness of distribution and is given by

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4}$$

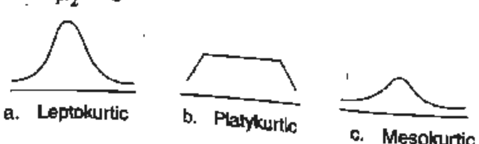


Fig. 3.3

The curves with $\beta_2 > 3$ is called Leptokurtic and those with $\beta_2 < 3$ are called platykurtic. The normal curve for which $\beta_2 = 3$ is called Mesokurtic.

As $\beta_1 = \frac{\mu_3^2}{\mu_2^3}$ and $\beta_2 = \frac{\mu_4}{\mu_2^2}$ determine the shape of the probability curve, these are called Pearson's shape coefficients.

Example 1

From the marks scored by 100 students in Section A and 100 students in Section B of a class, the following measures were obtained:

Section A	$\mu_A = 55$	$\sigma_A = 15.4$	Mode = 58.72
Section B	$\mu_A = 53$	$\sigma_B = 15.4$	Mode = 48.83

Determine which distribution of marks is more skewed.

Solution

$$S_{k_A} = \frac{\text{Mean} - \text{Mode}}{\sigma_A} = \frac{55 - 58.72}{15.4} = -0.24$$

$$S_{k_B} = \frac{\text{Mean} - \text{Mode}}{\sigma_B} = \frac{53 - 48.83}{15.4} = 0.27$$

$$|0.27| > |-0.24|$$

Hence, the distribution of marks of Section B is more skewed.

Example 2

For a group of 10 items, $\sum x = 452$, $\sum x^2 = 24270$, and mode = 43.7. Find Karl Pearson's coefficient of skewness.

Solution

$n = 10$, $\sum x = 452$, $\sum x^2 = 24270$, mode = 43.7

$$\mu = \frac{\Sigma x}{n} = \frac{452}{10} = 45.2$$

$$\sigma = \sqrt{\frac{\Sigma x^2}{n} - \left(\frac{\Sigma x}{n}\right)^2} = \sqrt{\frac{24270}{10} - \left(\frac{452}{10}\right)^2} = 19.59$$

Scanned with CamScanner

$$S_k = \frac{\text{Mean} - \text{Mode}}{\sigma}$$
$$= \frac{45.2 - 43.7}{19.59}$$
$$= 0.077$$

Example 3

In a distribution, the mean = 65, median = 70, coefficient of skewness = –0.6. Find the mode and coefficient of variation.

Solution

$\mu = 65$, Median = 70, $S_k = -0.6$

Mode = 3 Median – 2 Mean = 3(70) – 2(65) = 80

$$S_k = \frac{\text{Mean} - \text{Mode}}{\sigma}$$
$$-0.6 = \frac{65 - 80}{\sigma}$$
$$\therefore \quad \sigma = 25$$
$$\text{CV} = \frac{\sigma}{\bar{x}} \times 100 = \frac{25}{65} \times 100 = 38.64\%$$

Example 4

The following information was obtained from the records of a factory relating to wages:

Arithmetic mean = ₹ 56.8, Median = ₹ 59.5, *Standard deviation* = ₹ 12.4

Give the information about the distribution of wages.

Solution

$\mu = 56.8$, Median = 59.5, $\sigma = 12.4$

$$S_k = \frac{3(\text{Mean} - \text{Median})}{\sigma} = \frac{3(56.8 - 59.5)}{12.4} = -0.65$$
$$\text{Mode} = 3\,\text{Median} - 2\,\text{Mean} = 3(59.5) - 2(56.8) = 64.9$$

Hence, the maximum wages is ₹ 64.9.

There is a negative skewness in wages.

Example 5

For a moderately skewed distribution of retail price for men's shoes, it is found that the mean price is ₹ 20 and the median price is ₹ 17. If the coefficient of variation is 20%, find the Pearson's coefficient of skewness.

Solution

$\mu = 20$, Median = 17, CV = 20%

$$\text{CV} = \frac{\sigma}{\bar{x}} \times 100$$
$$20 = \frac{\sigma}{20} \times 100$$
$$\therefore \quad \sigma = 4$$
$$S_k = \frac{3(\text{Mean} - \text{Median})}{\sigma} = \frac{3(20 - 17)}{4} = 2.25$$

Example 6

Find the mean, SD, quartiles, median and Karl Pearson's coefficient of skewness for the following probability distribution:

$X = x$	1	2	3	4	5	6	7	8
$p(x)$	0.008	0.032	0.142	0.216	0.240	0.206	0.143	0.013

Solution

(i) Mean $= \mu = \sum x\,p(x)$
$$= 1(0.008) + 2(0.032) + 3(0.142) + 4(0.216) + 5(0.240) + 6(0.206) + 7(0.143) + 8(0.013)$$
$$= 4.903$$

(ii) $\text{Var}(X) = \sigma^2 = \sum x^2 p(x) - \mu^2$
$$= 1(0.008) + 4(0.032) + 9(0.142) + 16(0.216) + 25(0.240) + 36(0.206) + 49(0.143) + 64(0.013) - (4.903)^2$$
$$= 2.086$$
$$SD = \sqrt{\text{Var}(X)} = \sqrt{2.086} = 1.444$$

(iii) $F(3) = 0.008 + 0.032 + 0.142 = 0.182 < 0.25$
$F(4) = 0.008 + 0.032 + 0.142 + 0.216 = 0.398 > 0.25$
$$Q_1 = \frac{1}{3}(3 + 4) = 3.5$$
$F(5) = 0.008 + 0.032 + 0.142 + 0.216 + 0.240 = 0.638 < 0.75$
$F(6) = 0.008 + 0.032 + 0.142 + 0.216 + 0.240 + 0.206 = 0.844 > 0.75$

Scanned with CamScanner

$$Q_2 = \frac{1}{2}(4+5) = 4.5$$

$$Q_3 = \frac{1}{2}(5+6) = 5.5$$

(iv) Median $= Q_2 = 4.5$

(v) Pearson's coefficient of skewness

$$S_k = \frac{\text{Mean} - \text{Median}}{\text{SD}} = \frac{4.903 - 4.5}{1.444} = 0.279$$

Example 7

Find the mean, median QD, MD, SD, β_1 and β_2 of the following probability distribution:

$X = x$	0	1	2	3	4	5	6	7	8
$p(x)$	0.004	0.036	0.1	0.232	0.280	0.204	0.112	0.028	0.004

Solution

(i) Mean $= \mu = \Sigma x\, p(x)$

$$= 0 + 1(0.036) + 2\,(0.1) + 3(0.232) + 4(0.280) + 5(0.204) + 6(0.112) + 7(0.028) + 8(0.04)$$

$$= 3.972$$

(ii) Median

$$F(3) = 0.004 + 0.036 + 0.1 + 0.232 = 0.372 < 0.5$$

$$F(4) = 0.004 + 0.036 + 0.1 + 0.232 + 0.280 = 0.652 > 0.5$$

$$\text{Median } M = \frac{1}{2}(3+4) = 3.5$$

(iii) Mode is the value of X for which $P(X = x)$ is maximum.

Mode $= 4$ $\quad [\because P(X = 4) = 0.280$ is maximum probability$]$

(iv) Variance $= \sigma^2 = \sum x^2 p(x) - \mu^2$

$$= 0 + 1\,(0.036) + 4(0.1) + 9(0.232) + 16(0.28) + 25(0.204) + 36(0.112) + 49(0.028) + 64(0.004) - (3.972)^2$$

$$= 1.987$$

$$\text{SD} = \sqrt{\text{Var}(X)} = \sqrt{1.987} = 1.41$$

(v) $F(2) = 0.004 + 0.036 + 0.1 = 0.14 < 0.25$

$F(3) = 0.372 > 0.25$

$$Q_1 = \frac{1}{2}(2+3) = 2.5$$

$F(4) = 0.652 < 0.75$

$F(5) = 0.004 + 0.036 + 0.1 + 0.232 + 0.280 + 0.204 = 0.856 > 0.75$

$$Q_3 = \frac{1}{2}(4+5) = 4.5$$

$$QD = \frac{1}{2}(Q_3 - Q_1) = \frac{1}{2}(4.5 - 2.5) = 1$$

(vi) MD $= \sum |x - \mu|\, p(x)$

$$= 3.972(0.004) + 2.972(0.036) + 1.972(0.1) + 0.972(0.232) + 0.028\,(0.28) + 1.028(0.204) + 2.028(0.112) + 3.028(0.028) + 4.028(0.004)$$

$$= 1.091$$

(vii) $\mu_1' = \mu = 3.972$

$\mu_2' = E(X^2) = \Sigma x^2 p(x)$

$$= 0 + 1(0.036) + 4(0.1) + 9(0.232) + 16(0.280) + 25(0.204) + 36(0.112) + 49(0.028) + 64(0.004)$$

$$= 17.764$$

$\mu_3' = E(X^3) = \Sigma x^3 p(x)$

$$= 0 + 1(0.036) + 8\,(0.1) + 27(0.232) + 64(0.280) + 125(0.204) + 216(0.112) + 343(0.028) + 512(0.004)$$

$$= 86.364$$

$\mu_4' = E(X^4) = \Sigma x^4 p(x)$

$$= 0 + 1(0.036) + 16(0.1) + 81(0.232) + 256(0.280) + 625(0.204) + 1296(0.112) + 2401(0.028) + 4096(0.004)$$

$$= 448.372$$

$\mu_2 = \sigma^2 = 1.987$

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2(\mu_1')^3$$

$$= 86.364 - 3(17.764)\,(3.972) + 2(3.972)^3$$

$$= 0.019$$

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

$$= 448.372 - 4(86.364)\,(3.972) + 6(17.764)\,(3.972)^2 - 3(3.972)^4$$

$$= 11.053$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{(0.019)^2}{(1.987)^3} = 0.00005$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{11.053}{(1.987)^2} = 2.8$$

Scanned with CamScanner

EXERCISE 3.3

1. Karl Pearson's measure of skewness of a distribution is 0.5. Its median and mode are respectively 42 and 36. Find the coefficient of variation.
   [Ans.: 40]

2. From the marks scored by 120 students in Section *A* and 120 students in Section *B* of a class, the following measures are obtained:

Section *A*	$\bar{x} = 46.83$	SD = 14.8	mode = 51.67
Section *B*	$\bar{x} = 47.83$	SD = 14.8	mode = 47.07

   Determine which distribution of marks is more skewed.
   [Ans.: Section *A*]

3. For a moderately skewed data, the arithmetic mean is 200, the coefficient of variation is 8, and Karl Pearson's coefficient of skewness is 0.3. Find the mode and median.
   [Ans.: 195.2, 198.4]

4. Karl Pearson's coefficient of skewness of a distribution is 0.32. Its standard deviation is 6.5 and the mean is 29.6. Find the mode and median for the distribution.
   [Ans.: 27.52, 28.9]

5. The median, mode and coefficient of skewness for a certain distribution are respectively 17.4, 15.3, and 0.35. Find the coefficient of variation.
   [Ans.: 48.78%]

6. In a distribution, mean = 65, median = 70, coefficient of skewness = -6. Find the mode and coefficient of variation.
   [Ans.: 80, 39.78%]

3.7 MEASURES OF STATISTICS FOR CONTINUOUS RANDOM VARIABLES

1. Mean The mean or average value (μ) of the probability distribution of a continuous random variable X is called the expectation and is denoted by $E(X)$.

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

Expectation of any function $\phi(x)$ of a continuous random variable X is given by

$$E[\phi(x)] = \int_{-\infty}^{\infty} \phi(x)\, f(x)\, dx$$

2. Median The median is the point which divides the entire distribution into two equal parts. In case of a continuous distribution, the median is the point which divides the total area into two equal parts. Thus, if a continuous random variable X is defined from a to b and M is the median,

$$\int_a^M f(x)\, dx = \int_M^b f(x)\, dx = \frac{1}{2}$$

By solving any one of this equation, the median is obtained.

3. Mode The mode is value of x for which $f(x)$ is maximum. Mode is given by

$$f'(x) = 0 \text{ and } f''(x) < 0 \quad \text{for} \quad a < x < b$$

4. Geometric Mean The geometric mean of the probability distribution of a continuous random variable X is given by

$$\log G = \int_{-\infty}^{\infty} (\log x) f(x)\, dx$$

5. Harmonic Mean The harmonic mean of the probability distribution of a continuous random variable X is given by

$$\frac{1}{H} = \int_{-\infty}^{\infty} \frac{1}{x} f(x)\, dx$$

6. Quartile Deviation The r^{th} quartile of the probability distribution of a continuous random variable X and denoted by Q_r, is given by

$$\int_{-\infty}^{Q_r} f(x)\, dx = \frac{r}{4},\ r = 1, 2, 3$$

Q_1 and Q_3 are called the first (lower) and the third (upper) quartiles respectively. Q_2 is the median (middle or second quartile). Quartile deviation or semi-interquartile range of the probability distribution of a continuous random variable X is given by

$$Q = \frac{1}{2}(Q_3 - Q_1)$$

7. Mean Deviation Mean deviation of the probability distribution of a continuous random variable X is given by

Scanned with CamScanner

$$\text{MD} = \int_{-\infty}^{\infty} |x-\mu| f(x)\,dx$$

8. Standard Deviation The standard deviation of the probability distribution of a continuons random variable X is given by

$$\text{SD} = \sqrt{\text{Var}(X)} = \sigma$$

9. Variance The variance of the probability distribution of a continuous random variable X is given by

$$\text{Var}(X) = \sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx$$

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

10. Moments Central moments or moments about actual mean of the probability distribution of a continuous random variable X is given by

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

Raw moments or moments about arbitrary origin of the probability distribution of a continuous random variable X is given by

$$\mu_r' = \int_{-\infty}^{\infty} (x-a)^r f(x)\,dx$$

When $a = 0$, μ_r' is called r^{th} order simple moments.

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x)\,dx$$

11. Skewness Skewness of the probability distribution of a continuous random variable X is given by

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3}$$

12. Kurtosis Kurtosis of the probability distribution of a continuous random variable X is given by

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4}$$

Note The formulae of various measures of central tendency, dispersion, skewness and kurtosis of discrete probability distribution can be easily extended to the case of continuous probability distribution by simply replacing $p(x)$ by $f(x)dx$ and the summation by integration over the specified range of the variable X.

Example 1

For the continuous random variable having pdf

$$f(x) = 4x^3, \quad 0 \le x \le 1$$
$$= 0 \quad \textit{otherwise}$$

Find the mean and variance of X.

Solution

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

$$= \int_{-\infty}^{0} x f(x)\,dx + \int_{0}^{1} x f(x)\,dx + \int_{1}^{\infty} x f(x)\,dx$$

$$= 0 + \int_{0}^{1} x(4x^3)\,dx + 0$$

$$= 4\int_{0}^{1} x^4\,dx$$

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

$$= 4\left(\frac{1}{5} - 0\right)$$

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

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

$$= \int_{-\infty}^{0} x^2 f(x)\,dx + \int_{0}^{1} x^2 f(x)\,dx + \int_{1}^{\infty} x^2 f(x)\,dx - \mu^2$$

$$= 0 + \int_{0}^{1} x^2 (4x^3)\,dx + 0 - \left(\frac{4}{5}\right)^2$$

Scanned with CamScanner

$$= 4\int_0^1 x^5\,dx - \frac{16}{25}$$

$$= 4\left|\frac{x^6}{6}\right|_0^1 - \frac{16}{25}$$

$$= \frac{4}{6} - \frac{16}{25}$$

$$= \frac{2}{75}$$

Example 2

For the triangular distribution

$$\begin{aligned} f(x) &= x & 0 < x \le 1 \\ &= 2 - x & 1 \le x \le 2 \\ &= 0 & \textit{otherwise} \end{aligned}$$

Find the mean and variance.

Solution

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

$$= \int_{-\infty}^{0} x f(x)\,dx + \int_0^1 x f(x)\,dx + \int_1^2 x f(x) + \int_2^{\infty} x f(x)\,dx$$

$$= 0 + \int_0^1 x \cdot x\,dx + \int_1^2 x(2-x)\,dx + 0$$

$$= \int_0^1 x^2\,dx + \int_1^2 (2x - x^2)\,dx$$

$$= \left|\frac{x^3}{3}\right|_0^1 + \left|2\frac{x^2}{2} - \frac{x^3}{3}\right|_1^2$$

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

$$= \frac{1}{3} + \frac{4}{3} - \frac{2}{3}$$

$$= 1$$

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

$$= \int_{-\infty}^{0} x^2 f(x)\,dx + \int_0^1 x^2 f(x)\,dx + \int_1^2 x^2 f(x)\,dx + \int_2^{\infty} x^2 f(x)\,dx - \mu^2$$

$$= 0 + \int_0^1 x^2 \cdot x\,dx + \int_1^2 x^2(2-x)\,dx + 0 - 1$$

$$= \int_0^1 x^3\,dx + \int_1^2 (2x^2 - x^3)\,dx - 1$$

$$= \left|\frac{x^4}{4}\right|_0^1 + \left|\frac{2x^3}{3} - \frac{x^4}{4}\right|_1^2 - 1$$

$$= \left(\frac{1}{4} - 0\right) + \left[\left(\frac{16}{3} - \frac{16}{4}\right) - \left(\frac{2}{3} - \frac{1}{4}\right)\right] - 1$$

$$= \frac{7}{6} - 1$$

$$= \frac{1}{6}$$

Example 3

If the probability density function of X is given by

$$f(x) = \begin{cases} \dfrac{x}{2} & 0 < x \le 1 \\ \dfrac{1}{2} & 1 < x \le 2 \\ \dfrac{3-x}{2} & 2 < x < 3 \\ 0 & \textit{otherwise} \end{cases}$$

Find the expected value of $f(x) = x^2 - 5x + 3$.

Solution

$$E[E\phi(x)] = \int_{-\infty}^{\infty} \phi(x) \cdot f(x)\,dx$$

$$E(x^2 - 5x + 3) = \int_{-\infty}^{\infty} (x^2 - 5x + 3)\, f(x)\,dx$$

Scanned with CamScanner

$$= \int_0^1 (x^2-5x+3)\frac{x}{2}\,dx + \int_1^2 (x^2-5x+3)\frac{1}{2}\,dx + \int_2^3 (x^2-5x+3)\left(\frac{3-x}{2}\right)dx$$

$$= \frac{1}{2}\int_0^1 (x^3-5x^2+3x)\,dx + \frac{1}{2}\int_1^2 (x^2-5x+3)\,dx + \frac{1}{2}\int_2^3 (-x^3+8x^2-18x+9)\,dx$$

$$= \frac{1}{2}\left|\frac{x^4}{4}-\frac{5x^3}{3}+\frac{3x^2}{2}\right|_0^1 + \frac{1}{2}\left|\frac{x^3}{3}-\frac{5x^2}{2}+3x\right|_1^2 + \frac{1}{2}\left|-\frac{x^4}{4}+\frac{8x^3}{3}-\frac{18x^2}{2}+9x\right|_2^3$$

$$= \frac{1}{2}\left(\frac{1}{4}-\frac{5}{3}+\frac{3}{2}\right) + \frac{1}{2}\left(\frac{8}{3}-10+6-\frac{1}{3}+\frac{5}{2}-3\right) + \frac{1}{2}\left(-\frac{81}{4}+\frac{216}{3}-\frac{162}{2}+27+\frac{16}{4}-\frac{64}{3}+\frac{72}{2}-18\right)$$

$$= \frac{1}{24}-\frac{13}{12}-\frac{19}{24}$$

$$= -\frac{11}{6}$$

Example 4

A continuous random variable has the probability density function

$$f(x) = kxe^{-\lambda x} \qquad x \ge 0, \lambda > 0$$
$$= 0 \qquad \textit{otherwise}$$

Determine (i) k, (ii) mean, and (iii) variance.

Solution

Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{-\infty}^{0} f(x)\,dx + \int_0^{\infty} f(x)\,dx = 1$$

$$0 + \int_0^{\infty} k\,x e^{-\lambda x}\,dx = 1$$

$$k\int_0^{\infty} x e^{-\lambda x}\,dx = 1$$

$$k\left|x\frac{e^{-\lambda x}}{-\lambda} - 1\frac{e^{-\lambda x}}{\lambda^2}\right|_0^{\infty} = 1$$

$$k\left[(0-0)-\left(0-\frac{1}{\lambda^2}\right)\right] = 1$$

$$k = \lambda^2$$

Hence, $f(x) = \lambda^2 x e^{-\lambda x} \qquad x \ge 0, \lambda = 0$

$= 0 \qquad$ otherwise

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

$$= \int_{-\infty}^{0} x\,f(x)\,dx + \int_0^{\infty} x\,f(x)\,dx$$

$$= 0 + \int_0^{\infty} x\,\lambda^2\,x e^{-\lambda x}\,dx$$

$$= \lambda^2 \int_0^{\infty} x^2 e^{-\lambda x}\,dx$$

$$= \lambda^2\left|x^2\left(\frac{e^{-\lambda x}}{-\lambda}\right) - 2x\left(\frac{e^{-\lambda x}}{\lambda^2}\right) + 2\left(\frac{e^{-\lambda x}}{-\lambda^3}\right)\right|_0^{\infty}$$

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

$$= \frac{2}{\lambda}$$

(iii) Variance $= \sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$

$$= \int_{-\infty}^{0} x^2 f(x)\,dx + \int_0^{\infty} x^2 f(x)\,dx - \mu^2$$

$$= 0 + \int_0^{\infty} x^2\,\lambda^2\,x e^{-\lambda x}\,dx - \left(\frac{2}{\lambda}\right)^2$$

$$= \lambda^2 \int_0^{\infty} x^3 e^{-\lambda x}\,dx - \frac{4}{\lambda^2}$$

Scanned with CamScanner

$$= \lambda^2 \left| x^3\left(\frac{e^{-\lambda x}}{-\lambda x}\right) - 3x^2\left(\frac{e^{-\lambda x}}{\lambda^2}\right) + 6x\left(\frac{e^{-\lambda x}}{-\lambda^3}\right) - 6\left(\frac{e^{-\lambda x}}{\lambda^4}\right) \right|_0^\infty - \frac{4}{\lambda^2}$$

$$= \lambda^2 \left[(0-0+0-0) - \left(0-0+0-\frac{6}{\lambda^4}\right)\right] - \frac{4}{\lambda^2}$$

$$= \frac{6}{\lambda^2} - \frac{4}{\lambda^2}$$

$$= \frac{2}{\lambda^2}$$

Example 5

The probability density $f(x)$ of a continuous random variable is given by $f(x) = k\,e^{-|x|}$, $-\infty < x < \infty$ (i) show that $k = \frac{1}{2}$, and (ii) find the mean and variance of the distribution. (iii) Also, find the probability that the variate lies between 0 and 4.

Solution

(i) Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{-\infty}^{\infty} k\,e^{-|x|}\,dx = 1$$

$$k\int_{-\infty}^{\infty} e^{-|x|}\,dx = 1$$

$$2k\int_{0}^{\infty} e^{-|x|}\,dx = 1 \qquad [\because e^{-|x|} \text{ is an even function}]$$

$$2k\int_{0}^{\infty} e^{-x}\,dx = 1 \qquad [\because |x| = x \quad 0 \le x \le \infty]$$

$$2k\left|-e^{-x}\right|_0^\infty = 1$$

$$-2k\,(0-1) = 1$$

$$k = \frac{1}{2}$$

Hence, $f(x) = \frac{1}{2}e^{-|x|} \qquad -\infty < x < \infty$

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

$$= \frac{1}{2}\int_{-\infty}^{\infty} x\,e^{-|x|}\,dx$$

$$= 0 \qquad [\because \text{ the integrand is an odd function}]$$

(iii) $\text{Var}(X) = \sigma^2 = \int_{-\infty}^{\infty} x^2\,f(x)\,dx - \mu^2$

$$= \frac{1}{2}\int_{-\infty}^{\infty} x^2\,e^{-|x|}\,dx - 0$$

$$= 2\left(\frac{1}{2}\right)\int_{0}^{\infty} x^2\,e^{-|x|}\,dx \qquad [\because \text{ the integrand is an even function}]$$

$$= \int_{0}^{\infty} x^2\,e^{-|x|}\,dx$$

$$= \left| x^2\frac{e^{-x}}{-1} - 2x\frac{e^{-x}}{1} + 2\frac{e^{-x}}{-1} \right|_0^\infty$$

$$= 0 - (-2)$$

$$= 2$$

(iii) Probability that the variate lies between 0 and 4

$$P(0 < X < 4) = \int_0^4 f(x)\,dx$$

$$= \frac{1}{2}\int_0^4 e^{-|x|}\,dx$$

$$= \frac{1}{2}\int_0^4 e^{-x}\,dx \qquad [\because |x| = x \quad 0 < x < 4]$$

$$= -\frac{1}{2}\left|e^{-x}\right|_0^4$$

$$= -\frac{1}{2}(e^{-4} - 1)$$

$$= 0.4908$$

Example 6

The daily consumption of electric power is a random variable X with probability density function

Scanned with CamScanner

$$f(x) = k\,x\,e^{-\frac{x}{3}} \quad x > 0$$
$$= 0 \quad x \le 0$$

Find the value of k, the expectation of X, and the probability that on a given day, the electric consumption is more than the expected value.

Solution

Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$$

$$\int_{-\infty}^{0} f(x)\,\mathrm{d}x + \int_{0}^{\infty} f(x)\,\mathrm{d}x = 1$$

$$0 + \int_{0}^{\infty} k\,x\,e^{-\frac{x}{3}}\,\mathrm{d}x = 1$$

$$k\left| x\left(\frac{e^{-\frac{x}{3}}}{-\frac{1}{3}}\right) - (1)\left(\frac{e^{-\frac{x}{3}}}{\frac{1}{9}}\right)\right|_0^{\infty} = 1$$

$$k[(0-0)-(0-9)] = 1$$
$$9k = 1$$
$$k = \frac{1}{9}$$

Hence, $f(x) = \frac{1}{9} x e^{-\frac{x}{3}} \quad x > 0$

$$= 0 \quad x \le 0$$

$$E(X) = \int_{-\infty}^{\infty} x\,f(x)\,\mathrm{d}x$$
$$= \int_{-\infty}^{0} x\,f(x)\,\mathrm{d}x + \int_{0}^{\infty} x\,f(x)\,\mathrm{d}x$$
$$= 0 + \int_{0}^{\infty} x \cdot \frac{1}{9} x\,e^{-\frac{x}{3}}\,\mathrm{d}x$$
$$= \frac{1}{9}\int_{0}^{\infty} x^2\,e^{-\frac{x}{3}}\,\mathrm{d}x$$

$$= \frac{1}{9}\left| x^2\left(\frac{e^{-\frac{x}{3}}}{-\frac{1}{3}}\right) - 2x\left(\frac{e^{-\frac{x}{3}}}{\frac{1}{9}}\right) + 2\left(\frac{e^{-\frac{x}{3}}}{-\frac{1}{27}}\right)\right|_0^{\infty}$$
$$= \frac{1}{9}(0-0+0+54)$$
$$= 6$$

$$P(X > 6) = \int_{0}^{6} f(x)\,\mathrm{d}x$$
$$= \int_{0}^{6} \frac{1}{9} x\,e^{-\frac{x}{3}}\,\mathrm{d}x$$
$$= \frac{1}{9}\int_{0}^{6} x\,e^{-\frac{x}{3}}\,\mathrm{d}x$$
$$= \frac{1}{9}\left| x\left(\frac{e^{-\frac{x}{3}}}{-\frac{1}{3}}\right) - 1\left(\frac{e^{-\frac{x}{3}}}{\frac{1}{9}}\right)\right|_0^{\infty}$$
$$= \frac{1}{9}\left[(0-0) - \left(-18e^{-2} - 9e^{-2}\right)\right]$$
$$= 3\,e^{-2}$$
$$= 0.406$$

Example 7

Let X be a random variable with E(X) = 10 and Var(X) = 25. Find the positive values of a and b such that Y = aX – b has an expectation of 0 and a variance of 1.

Solution

$$E(Y) = E(aX - b)$$
$$0 = aE(X) - b$$
$$= a(10) - b$$
$$10a - b = 0$$
$$\mathrm{Var}(Y) = \mathrm{Var}(aX - b)$$
$$1 = a^2\ \mathrm{Var}(X)$$
$$= a^2(25)$$

Scanned with CamScanner

$$25a^2 = 1$$
$$a = \frac{1}{5}$$
$$b = 2$$

Example 8

A continuous random variable X is distributed over the interval [0, 1] with pdf $f(x) = ax^2 + bx$, where a, b are constants. If the mean of X is 0.5, find the values of a and b.

Solution

Since $f(x)$ is probability density function,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{\infty} f(x)\,dx = 1$$

$$0 + \int_{0}^{1} (ax^2 + bx)\,dx + 0 = 1$$

$$\left|\frac{ax^3}{3} + \frac{bx^2}{2}\right|_0^1 = 1$$

$$\frac{a}{3} + \frac{b}{2} = 1$$

$$2a + 3b = 6 \qquad \text{...(1)}$$

Also, $\mu = 0.5$

$$\int_{0}^{1} x\, f(x)\,dx = 0.5$$

$$\int_{0}^{1} x\,(ax^2 + bx)\,dx = 0.5$$

$$\int_{0}^{1} (ax^3 + bx^2)\,dx = 0.5$$

$$\left|\frac{ax^4}{4} + \frac{bx^3}{3}\right|_0^1 = 0.5$$

$$\frac{a}{4} + \frac{b}{3} = 0.5$$

$$3a + 4b = 6 \qquad \text{...(2)}$$

Solving Eqs (1) and (2),

$$a = -6, \qquad b = 6$$

Example 9

A continuous random variable X has the pdf defined by $f(x) = A + Bx$, $0 \le x \le 1$. If the mean of the distribution is $\frac{1}{3}$, find A and B.

Solution

Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_{-\infty}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{\infty} f(x)\,dx = 1$$

$$0 + \int_{0}^{1} (A + Bx)\,dx + 0 = 1$$

$$\left|Ax + \frac{Bx^2}{2}\right|_0^1 = 1$$

$$A + \frac{B}{2} = 1 \qquad \text{...(1)}$$

Also,

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

$$\int_{-\infty}^{\infty} x\, f(x)\,dx = \frac{1}{3}$$

$$\int_{-\infty}^{0} x\, f(x)\,dx + \int_{0}^{1} x\, f(x)\,dx = \frac{1}{3}$$

$$0 + \int_{0}^{1} x\,(A + Bx)\,dx = \frac{1}{3}$$

$$\int_{0}^{1} (Ax + Bx^2)\,dx = \frac{1}{3}$$

$$\left|\frac{Ax^2}{2} + \frac{Bx^3}{3}\right|_0^1 = \frac{1}{3}$$

$$\frac{A}{2} + \frac{B}{3} = \frac{1}{3}$$

$$3A + 2B = 2 \qquad \text{...(2)}$$

Scanned with CamScanner

Solving Eqs (1) and (2),

$A = 2, \quad B = -2$

Example 10

A continuous random variable has probability density function $f(x) = 6(x - x^2) \quad 0 \le x \le 1.$

Find the (i) mean, (ii) variance, (iii) median, and (iv) mode.

Solution

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

$$= \int_{-\infty}^{0} x\, f(x)\, dx + \int_{0}^{1} x\, f(x)\, dx + \int_{1}^{\infty} x\, f(x)\, dx$$

$$= 0 + \int_{0}^{1} x\, 6(x - x^2)\, dx + 0$$

$$= 6\int_{0}^{1} (x^2 - x^3)\, dx$$

$$= 6\left|\frac{x^3}{3} - \frac{x^4}{4}\right|_0^1$$

$$= 6\left(\frac{1}{3} - \frac{1}{4}\right)$$

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

(ii) $\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x)\, dx - \mu^2$

$$= \int_{-\infty}^{0} x^2 f(x)\, dx + \int_{0}^{1} x^2\, f(x)\, dx + \int_{1}^{\infty} x^2 f(x)\, dx - \mu^2$$

$$= 0 + \int_0^1 x^2\, 6(x - x^2)\, dx + 0 - \frac{1}{4}$$

$$= 6\int_0^1 (x^3 - x^4)\, dx - \frac{1}{4}$$

$$= 6\left|\frac{x^4}{4} - \frac{x^5}{5}\right|_0^1 - \frac{1}{4}$$

$$= 6\left(\frac{1}{4} - \frac{1}{5}\right) - \frac{1}{4}$$

$$= \frac{6}{20} - \frac{1}{4}$$

$$= \frac{1}{20}$$

(iii) $$\int_{a}^{M} f(x)\, dx = \int_{M}^{b} f(x)\, dx = \frac{1}{2}$$

$$\int_{0}^{M} 6(x - x^2)\, dx = \frac{1}{2}$$

$$6\left|\frac{x^2}{2} - \frac{x^3}{3}\right|_0^M = \frac{1}{2}$$

$$6\left(\frac{M^2}{2} - \frac{M^3}{3}\right) = \frac{1}{2}$$

$$3M^2 - 2M^3 = \frac{1}{2}$$

$$4M^3 - 6M^2 + 1 = 0$$

$$(2M - 1)(2M^2 - 2M - 1) = 0$$

$$M = \frac{1}{2} \quad \text{or} \quad M = \frac{1 \pm \sqrt{3}}{2}$$

$$M = \frac{1}{2} \text{ lies in } (0, 1)$$

Hence, median $M = \frac{1}{2}$

(iv) Mode is the value of x for which $f(x)$ is maximum. For $f(x)$ to be maximum, $f'(x) = 0$ and $f''(x) < 0$.

$$f'(x) = 0$$

$$6(1 - 2x) = 0$$

$$x = \frac{1}{2}$$

$$f''(x) = -12x$$

At $x = \frac{1}{2}$, $f''(x) = -12 < 0$

Hence, $f(x)$ is maximum at $x = \frac{1}{2}$.

Mode $= \frac{1}{2}$

Scanned with CamScanner

Example 11

The probability density function of a random variable X is

$$f(x)=\frac{1}{2}\sin x \quad 0\le x\le \pi$$
$$=0 \quad \textit{otherwise}$$

Find the mean, mode, and median of the distribution and also, find the probability between 0 and $\frac{\pi}{2}$.

Solution

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

$$=\int_{-\infty}^{0} x f(x)\,dx+\int_{0}^{\pi} x f(x)\,dx+\int_{\pi}^{\infty} x f(x)\,dx$$
$$=0+\int_0^{\pi} x\left(\frac{1}{2}\sin x\right)dx+0$$
$$=\frac{1}{2}\int_0^{\pi} x\sin x\,dx$$
$$=\frac{1}{2}\left|-x\cos x+\sin x\right|_0^{\pi}$$
$$=\frac{\pi}{2}$$

(ii) Mode is the value of x for which $f(x)$ is maximum. For $f(x)$ to be maximum, $f'(x)=0$ and $f''(x)<0$.

$$f'(x)=0$$
$$\cos x=0$$
$$x=\frac{\pi}{2}$$
$$f''(x)=-\frac{1}{2}\sin x$$

At $x=\frac{\pi}{2}, f''(x)=-\frac{1}{2}<0$

Hence, $f(x)$ is maximum of $x=\frac{\pi}{2}$,

Mode $=\frac{\pi}{2}$

(iii) $\int_a^M f(x)\,dx=\int_M^b f(x)\,dx=\frac{1}{2}$

$$\int_0^M \frac{1}{2}\sin x\,dx=\int_M^{\pi}\frac{1}{2}\sin x\,dx=\frac{1}{2}$$
$$\int_0^M \frac{1}{2}\sin x\,dx=\frac{1}{2}$$
$$-\frac{1}{2}\left|\cos x\right|_0^M=\frac{1}{2}$$
$$-\frac{1}{2}(\cos M-1)=\frac{1}{2}$$
$$1-\cos M=0$$
$$\cos M=0$$
$$M=\frac{\pi}{2}$$

Hence, median $M=\frac{\pi}{2}$

(iv) $P\left(0<X<\frac{\pi}{2}\right)=\int_0^{\frac{\pi}{2}} f(x)\,dx$

$$=\int_0^{\frac{\pi}{2}}\frac{1}{2}\sin x\,dx$$
$$=-\frac{1}{2}\left|\cos x\right|_0^{\frac{\pi}{2}}$$
$$=-\frac{1}{2}(0-1)$$
$$=\frac{1}{2}$$

Example 12

The cumulative distribution function of a continuous random variable X is $F(x)=1-e^{-2x} \quad x\ge 0$

$$=0 \quad x<0$$

Find the (i) the probability density function, (ii) mean, and (iii) variance.

Scanned with CamScanner

Solution

(i) $f(x)=\dfrac{d}{dx}F(x)$

$$f(x)=\frac{1}{2}e^{-2x} \qquad x\ge 0$$
$$=0 \qquad x<0$$

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

$$=\int_{-\infty}^{0} x\, f(x)\,dx+\int_{0}^{\infty} x\, f(x)\,dx$$
$$=0+\int_{0}^{\infty} x\cdot\frac{1}{2}e^{-2x}\,dx$$
$$=\frac{1}{2}\int_{0}^{\infty} x\, e^{-2x}\,dx$$
$$=\frac{1}{2}\left|x\left(\frac{e^{-2x}}{-2}\right)-1\left(\frac{e^{-2x}}{4}\right)\right|_{0}^{\infty}$$
$$=\frac{1}{2}\left[(0-0)-\left(0-\frac{1}{4}\right)\right]$$
$$=\frac{1}{8}$$

(iii) $\mathrm{Var}(X)=\int_{-\infty}^{\infty} x^2 f(x)\,dx-\mu^2$

$$=\int_{-\infty}^{0} x^2 f(x)\,dx+\int_{0}^{\infty} x^2 f(x)\,dx-\mu^2$$
$$=0+\int_{0}^{\infty} x^2\cdot\frac{1}{2}e^{-2x}\,dx-\left(\frac{1}{8}\right)^2$$
$$=\frac{1}{2}\int_{0}^{\infty} x^2 e^{-2x}dx-\frac{1}{64}$$
$$=\frac{1}{2}\left|x^2\left(\frac{e^{-2x}}{-2}\right)-2x\left(\frac{e^{-2x}}{4}\right)+2\left(\frac{e^{-2x}}{-8}\right)\right|_{0}^{\infty}-\frac{1}{64}$$
$$=\frac{1}{2}\left[(0-0-0)-\left(0-0-\frac{1}{4}\right)\right]-\frac{1}{64}$$
$$=\frac{1}{8}-\frac{1}{64}$$
$$=\frac{7}{64}$$

Example 13

A continuous random variable X has the distribution function

$$F(x)=0 \qquad x\le 1$$
$$=k(x-1)^4 \qquad 1<x\le 3$$
$$=1 \qquad x>3$$

Determine (i) $f(x)$, (ii) k, and (iii) mean.

Solution

(i) $f(x)=\dfrac{d}{dx}F(x)$

$$f(x)=0 \qquad x\le 1$$
$$=4k(x-1)^3 \qquad 1<x\le 3$$
$$=0 \qquad x>3$$

(ii) Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\,dx=1$$
$$\int_{-\infty}^{1} f(x)dx+\int_{1}^{3} f(x)dx+\int_{3}^{\infty} f(x)dx=1$$
$$0+\int_{1}^{3} 4k(x-1)^3\,dx+0=1$$
$$4k\left|\frac{(x-1)^4}{4}\right|_{1}^{3}=1$$
$$k(16-0)=1$$
$$k=\frac{1}{16}$$

Hence, $f(x)=0 \qquad x\le 1$

$$=\frac{1}{4}(x-1)^3 \qquad 1<x\le 3$$
$$=0 \qquad x>3$$

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

$$=\int_{-\infty}^{1} x\, f(x)\,dx+\int_{1}^{3} x\, f(x)\,dx+\int_{3}^{\infty} x\, f(x)\,dx$$
$$=0+\int_{1}^{3} x\cdot\frac{1}{4}(x-1)^3\,dx+0$$
$$=\frac{1}{4}\int_{1}^{3} x(x-1)^3\,dx$$

Scanned with CamScanner

$$= \frac{1}{4}\int_0^2 (t+1)\, t^3\, dt \qquad \begin{bmatrix} \text{Putting } x-1=t \\ \text{When } x=1, t=0 \\ \text{When } x=3, t=2 \end{bmatrix}$$

$$= \frac{1}{4}\int_0^2 (t^4+t^3)\, dt$$

$$= \frac{1}{4}\left|\frac{t^5}{5}+\frac{t^4}{4}\right|_0^2$$

$$= \frac{1}{4}\left[\left(\frac{2^5}{5}+\frac{2^4}{4}\right)-(0)\right]$$

$$= 2.6$$

Example 14

If the density function of a random variable X is given by

$$f(x) = kx\,(1-x), \quad 0 \le x \le 1,$$

find (i) AM, (ii) HM, (iii) Median, (iv) Mode, (v) SD, (vi) MD about the mean.

Solution

(i) Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$$

$$\int_0^1 kx(1-x)\mathrm{d}x = 1$$

$$k\int_0^1 (x-x^2)\,\mathrm{d}x = 1$$

$$k\left|\frac{x^2}{2}-\frac{x^3}{3}\right|_0^1 = 1$$

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

$$k = 6$$

Hence, $\quad f(x) = 6x(1-x), \quad 0 \le x \le 1$

(ii) $\text{AM} = \mu = E(x) = \int_{-\infty}^{\infty} xf(x)\,\mathrm{d}x$

$$= \int_0^1 x \cdot 6x(1-x)\mathrm{d}x$$

$$= 6\int_0^1 (x^2-x^3)\mathrm{d}x$$

$$= 6\left|\frac{x^3}{3}-\frac{x^4}{4}\right|_0^1$$

$$= 6\left(\frac{1}{3}-\frac{1}{4}\right)$$

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

(iii) $\dfrac{1}{H} = \int_{-\infty}^{\infty} \dfrac{1}{x} f(x)\mathrm{d}x$

$$= \int_0^1 \frac{1}{x}\cdot 6x(1-x)\mathrm{d}x$$

$$= 6\int_0^1 (1-x)\mathrm{d}x$$

$$= 6\left|x-\frac{x^2}{2}\right|_0^1$$

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

$$= 3$$

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

(iv) $\displaystyle\int_0^M f(x)\,\mathrm{d}x = \frac{1}{2}$

$$\int_0^M 6x(1-x)\,\mathrm{d}x = \frac{1}{2}$$

$$6\int_0^M (x-x^2)\mathrm{d}x = \frac{1}{2}$$

$$6\left|\frac{x^2}{2}-\frac{x^3}{3}\right|_0^M = \frac{1}{2}$$

Scanned with CamScanner

$$6\left(\frac{M^2}{2}-\frac{M^3}{3}\right)=\frac{1}{2}$$

$$3M^2-2M^3=\frac{1}{2}$$

$$6M^2-4M^3=1$$

$$4M^3-6M^2+1=0$$

$$M=\frac{1}{2},\frac{1}{2}\pm\frac{\sqrt{3}}{2}$$

The values $M=\frac{1}{2}\pm\frac{\sqrt{3}}{2}$ lie outside (0, 1).

Hence, $M=\frac{1}{2}$

(v) Mode is the value of x for which $f(x)$ is maximum. For $f(x)$ to be maximum, $f'(x)=0$ and $f''(x)<0$.

$$f'(x)=0$$

$$6-12x=0$$

$$x=\frac{1}{2}$$

$$f''(x)=-12<0$$

Hence, $f(x)$ is maximum at $x=\frac{1}{2}$

Mode $=\frac{1}{2}$

As the mean, median and mode are equal, the distribution is symmetrical.

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

$$=\int_0^1 x^2\cdot 6x(1-x)\,dx$$

$$=6\int_0^1 (x^3-x^4)\,dx$$

$$=6\left|\frac{x^4}{4}-\frac{x^5}{5}\right|_0^1$$

$$=6\left(\frac{1}{4}-\frac{1}{5}\right)$$

$$=\frac{3}{10}$$

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

$$=\frac{3}{10}-\left(\frac{1}{2}\right)^2$$

$$=\frac{1}{20}$$

$$SD=\sqrt{\text{Var}(X)}=\sqrt{\frac{1}{20}}=\frac{1}{2\sqrt{5}}$$

(vii) Mean deviation about the mean

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

$$=\int_0^1\left|x-\frac{1}{2}\right|6x(1-x)\,dx$$

$$=\int_0^{\frac{1}{2}}\left(\frac{1}{2}-x\right)6x(1-x)\,dx+\int_{\frac{1}{2}}^1\left(x-\frac{1}{2}\right)6x(1-x)\,dx$$

$$=\int_0^{\frac{1}{2}}(3x-9x^2+6x^3)\,dx+\int_{\frac{1}{2}}^1(-3x+9x^2-6x^3)\,dx$$

$$=\left|\frac{3x^2}{2}-3x^3+\frac{3x^4}{2}\right|_0^{\frac{1}{2}}+\left|-\frac{3x^2}{2}+3x^3-\frac{3x^4}{2}\right|_{\frac{1}{2}}^1$$

$$=\left(\frac{3}{8}-\frac{3}{8}+\frac{3}{32}\right)+\left(-\frac{3}{2}+3-\frac{3}{2}\right)-\left(-\frac{3}{8}+\frac{3}{8}-\frac{3}{32}\right)$$

$$=\frac{3}{16}$$

Example 15

Prove that geometric mean G of the distribution

$$f(x)=6(2-x)(x-1),\ 1\le x\le 2$$

is given by $6\log(16G)=19$.

Scanned with CamScanner

Solution

$$\log G = \int_{-\infty}^{\infty} (\log x) f(x)\,dx$$

$$= \int_{1}^{2} (\log x)\, 6(2-x)(x-1)\,dx$$

$$= -6\int_{1}^{2} (x^2-3x+2)\log x\,dx$$

$$= -6\left[\left|\left(\frac{x^3}{3}-\frac{3x^2}{2}+2x\right)\log x\right|_1^2 - \int_1^2\left(\frac{x^3}{3}-\frac{3x^2}{2}+2x\right)\frac{1}{x}\,dx\right]$$

$$= -6\left[\left(\frac{8}{3}-6+4\right)\log 2 - \int_1^2\left(\frac{x^2}{3}-\frac{3x}{2}+2\right)dx\right]$$

$$= -6\left[\frac{2}{3}\log 2 - \left|\frac{x^3}{9}-\frac{3x^2}{4}+2x\right|_1^2\right]$$

$$= -6\left[\frac{2}{3}\log 2 - \left(\frac{8}{9}-3+4\right)+\left(\frac{1}{9}-\frac{3}{4}+2\right)\right]$$

$$= -6\left[\frac{2}{3}\log 2 - \frac{17}{9}+\frac{49}{36}\right]$$

$$= -4\log 2 + \frac{19}{6}$$

$$\log G + 4\log 2 = \frac{19}{6}$$

$$\log(G\times 2^4) = \frac{19}{6}$$

$$\log(16G) = \frac{19}{6}$$

Example 16

The probability distribution of a random variable X is

$$f(x) = k\sin\frac{\pi}{5}x, \quad 0 \le x \le 5$$

Determine the constant k and obtain the median and quartiles of the distribution.

Solution

Since $f(x)$ is a probability distribution,

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

$$\int_0^5 k\sin\frac{\pi}{5}x\,dx = 1$$

$$k\left|\frac{-\cos\frac{\pi}{5}x}{\frac{\pi}{5}}\right|_0^5 = 1$$

$$\frac{5k}{\pi}(-\cos\pi+\cos 0) = 1$$

$$\frac{5k}{\pi}[-(-1)+1] = 1$$

$$\frac{10k}{\pi} = 1$$

$$k = \frac{\pi}{10}$$

Hence, $$f(x) = \frac{\pi}{10}\sin\frac{\pi}{5}x, \quad 0 \le x \le 5$$

The r^{th} quartile Q_r is given by

$$\int_{-\infty}^{Q_r} f(x)\,dx = \frac{r}{4}, \quad r = 1, 2, 3$$

$$\int_0^{Q_r} \frac{\pi}{10}\sin\frac{\pi}{5}x\,dx = \frac{r}{4}$$

$$\frac{\pi}{10}\left|\frac{-\cos\frac{\pi}{5}x}{\frac{\pi}{5}}\right|_0^{Q_r} = \frac{r}{4}$$

$$\frac{1}{2}\left(-\cos\frac{\pi}{5}Q_r+\cos 0\right) = \frac{r}{4}$$

$$-\cos\frac{\pi}{5}Q_r + 1 = \frac{r}{2}$$

Scanned with CamScanner

$$\cos\frac{\pi}{5}Q_r = 1-\frac{r}{2}$$

$$\frac{\pi}{5}Q_r = \cos^{-1}\left(1-\frac{r}{2}\right)$$

$$Q_r = \frac{5}{\pi}\cos^{-1}\left(1-\frac{r}{2}\right)$$

$$Q_1 = \frac{5}{\pi}\cos^{-1}\left(1-\frac{1}{2}\right) = \frac{5}{\pi}\cos^{-1}\left(\frac{1}{2}\right) = \frac{5}{\pi}\left(\frac{\pi}{3}\right) = \frac{5}{3}$$

$$Q_2 = \frac{5}{\pi}\cos^{-1}(1-1) = \frac{5}{\pi}\cos^{-1}(0) = \frac{5}{\pi}\left(\frac{\pi}{2}\right) = \frac{5}{2}$$

$$Q_3 = \frac{5}{\pi}\cos^{-1}\left(1-\frac{3}{2}\right) = \frac{5}{\pi}\cos^{-1}\left(-\frac{1}{2}\right) = \frac{5}{\pi}\left(\frac{2\pi}{3}\right) = \frac{10}{3}$$

$$\text{Median} = Q_2 = \frac{5}{2}$$

Example 17

Find the median, mode and quartile deviation of continuous random variable X, given that its density functions is

$$f(x) = \frac{k}{1+x^2}, \quad -\infty < x < \infty.$$

Solution

(i) Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = 1$$

$$\int_{-\infty}^{\infty} \frac{k}{1+x^2}\mathrm{d}x = 1$$

$$2k\int_{0}^{\infty} \frac{1}{1+x^2}\mathrm{d}x = 1 \quad \left[\because \int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx, \text{ if } f(x) \text{ is even function}\right]$$

$$2k\left|\tan^{-1}x\right|_0^{\infty} = 1$$

$$2k(\tan^{-1}\infty - \tan^{-1}0) = 1$$

$$2k\left(\frac{\pi}{2}\right) = 1$$

$$k = \frac{1}{\pi}$$

Hence, $$f(x) = \frac{1}{\pi(1+x^2)}, -\infty < x < \infty$$

(ii) The r^{th} quartile Q_r is given by

$$\int_{-\infty}^{Q_r} f(x)\mathrm{d}x = \frac{r}{4}, \quad r = 1, 2, 3$$

$$\int_{-\infty}^{Q_r} \frac{1}{\pi(1+x^2)}\mathrm{d}x = \frac{r}{4}$$

$$\frac{1}{\pi}\left|\tan^{-1}x\right|_{-\infty}^{Q_r} = \frac{r}{4}$$

$$\frac{1}{\pi}\left[\tan^{-1}Q_r - \tan^{-1}(-\infty)\right] = \frac{r}{4}$$

$$\frac{1}{\pi}\left[\tan^{-1}Q_r - \left(-\frac{\pi}{2}\right)\right] = \frac{r}{4}$$

$$\tan^{-1}Q_r = \frac{\pi}{4}(r-2)$$

$$Q_r = \tan\left\{\frac{\pi}{4}(r-2)\right\}$$

$$Q_1 = \tan\left(-\frac{\pi}{4}\right) = -1$$

$$Q_2 = \tan(0) = 0$$

$$Q_3 = \tan\left(\frac{\pi}{4}\right) = 1$$

$$QD = \frac{1}{2}(Q_3 - Q_1)$$

$$= \frac{1}{2}[1-(-1)]$$

$$= 1$$

(iii) Median $Q_2 = 0$

(iv) Mode is the value of x for which $f(x)$ is maximum. For $f(x)$ to be maximum $f'(x) = 0$ and $f''(x) < 0$.

Scanned with CamScanner

$$f'(x)=0$$

$$-\frac{2x}{\pi(1+x^2)^2}=0$$

$$x=0$$

$$f''(x)=-\frac{2}{\pi}\left[\frac{(1+x^2)^2-x\cdot 2(1+x^2)\cdot 2x}{(1+x^2)^4}\right]$$

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

$$f''(0)=-\frac{2}{\pi}<0$$

Hence, $f(x)$ is maximum at $x = 0$.
Mode = 0

Example 18

Find the mean, variance and the coefficients β_1, β_2 of the distribution $f(x) = kx^2e^{-x}$, $0 < x < \infty$

Solution

Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x=1$$

$$\int_{0}^{\infty} kx^2e^{-x}\mathrm{d}x=1$$

$$k\left|x^2(-e^{-x})-2x\,e^{-x}+2(-e^{-x})\right|_0^{\infty}=1$$

$$k(2e^0)=1$$

$$k=\frac{1}{2}$$

Hence,
$$f(x)=\frac{1}{2}x^2e^{-x}, 0<x<\infty$$

$$\mu_r'=\int_{-\infty}^{\infty} x^r f(x)\mathrm{d}x$$

$$=\int_0^{\infty} x^r\frac{1}{2}x^2e^{-x}\mathrm{d}x$$

$$=\frac{1}{2}\int_0^{\infty} e^{-x}x^{r+2}\mathrm{d}x$$

$$=\frac{1}{2}\overline{\left|r+3\right.}$$

$$=\frac{1}{2}(r+2)!$$

$$\mu_1'=\frac{1}{2}(3!)=3$$

$$\mu_2'=\frac{1}{2}(4!)=12$$

$$\mu_3'=\frac{1}{2}(5!)=60$$

$$\mu_4'=\frac{1}{2}(6!)=360$$

$$\mu_2=\mu_2'-(\mu_1')^2=12-(3)^2=3$$

$$\mu_3=\mu_3'-3\mu_2'\mu_1'+2(\mu_1')^3=60-3(12)(3)+2(3)^3=6$$

$$\mu_4=\mu_4'-4\mu_3'\mu_1'+6\mu_2'(\mu_1')^2-3(\mu_1')^4$$

$$=360-4(6)(3)+6(12)(3)^2-3(3)^4$$

$$=45$$

$$\text{Mean}=\mu_1'=3$$

$$\text{Variance}=\mu_2=3$$

$$\beta_1=\frac{\mu_3^2}{\mu_2^3}=\frac{(6)^2}{(3)^3}=\frac{4}{3}$$

$$\beta_2=\frac{\mu_4}{\mu_2^2}=\frac{45}{(3)^2}=5$$

Example 19

The probability density function of a random variable X is given by $f(x) = kx(2-x)$, $0 \le x \le 2$. Find mean, variance β_1 and β_2.

Solution

Since $f(x)$ is a probability density function,

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x=1$$

Scanned with CamScanner

$$\int_0^2 k\,x(2-x)\,dx = 1$$

$$k\int_0^2 (2x-x^2)\,dx = 1$$

$$k\left|x^2-\frac{x^3}{3}\right|_0^2 = 1$$

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

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

Hence, $f(x) = \frac{3}{4}x(2-x),\ 0 \le x \le 2$

$$\mu'_r = \int_{-\infty}^{\infty} x^r f(x)\,dx$$

$$= \int_0^2 x^r \frac{3}{4}x(2-x)\,dx$$

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

$$= \frac{2(2^{r+1})}{(r+2)(r+3)}$$

$$\mu'_1 = \frac{3(2^2)}{(3)(4)} = 1$$

$$\mu'_2 = \frac{3(2^3)}{(4)(5)} = \frac{6}{5}$$

$$\mu'_3 = \frac{3(2^4)}{(5)(6)} = \frac{8}{5}$$

$$\mu'_4 = \frac{3(2^5)}{(6)(7)} = \frac{16}{7}$$

$$\mu_2 = \mu'_2 - (\mu'_1)^2 = \frac{6}{5} - 1 = \frac{1}{5}$$

$$\mu_3 = \mu'_3 - 3\mu'_2\mu'_1 + 2(\mu'_1)^3 = \frac{8}{5} - 3\left(\frac{6}{5}\right)(1) + 2 = 0$$

$$\mu_4 = \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2(\mu'_1)^2 - 3(\mu'_1)^4$$

$$= \frac{16}{7} - 4\left(\frac{8}{5}\right)(1) + 6\left(\frac{6}{5}\right)(1)^2 - 3(1)^4$$

$$= \frac{3}{35}$$

$$\text{Mean} = \mu'_1 = 1$$

$$\text{Variance} = \mu_2 = \frac{1}{5}$$

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{\frac{3}{35}}{\left(\frac{1}{5}\right)^2} = \frac{15}{7}$$

Example 20

Show that for the symmetrical distribution

$$f(x) = \frac{2a}{\pi}\left(\frac{1}{a^2+x^2}\right), \quad -a \le x \le a$$

$$\mu_2 = \frac{a^2(4-\pi)}{\pi} \quad \textit{and} \quad \mu_4 = a^4\left(1-\frac{8}{3\pi}\right)$$

Solution

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-a}^{a} \frac{2a}{\pi}\left(\frac{1}{a^2+x^2}\right)dx$$

$$= \frac{2a}{\pi}\left|\frac{1}{a}\tan^{-1}\frac{x}{a}\right|_{-a}^{a}$$

$$= \frac{2}{\pi}\left|\tan^{-1}\frac{x}{a}\right|_{-a}^{a}$$

$$= \frac{2}{\pi}\left[\tan^{-1}(1) - \tan^{-1}(-1)\right]$$

$$= 1$$

Hence, $f(x)$ represents a probability density function.

Scanned with CamScanner

$$\mu_1' = \int_{-\infty}^{\infty} x\, f(x)\,dx$$

$$= \int_{-a}^{a} x \frac{2a}{\pi}\left(\frac{1}{a^2+x^2}\right)dx$$

$$= \frac{2a}{\pi}\int_{-a}^{a}\frac{x}{a^2+x^2}dx$$

$$= \frac{2a}{\pi}\left|\frac{1}{2}\log(a^2+x^2)\right|_{-a}^{a}$$

$$= 0 \quad [\because \text{integrand is an odd function of } x]$$

$$\mu_2' = \int_{-\infty}^{\infty} x^2 f(x)\,dx$$

$$= \int_{-a}^{a} x^2 \frac{2a}{\pi}\left(\frac{1}{a^2+x^2}\right)dx$$

$$= \frac{2a}{\pi}\int_{-a}^{a}\frac{x^2}{a^2+x^2}dx$$

$$= \frac{4a}{\pi}\int_{0}^{a}\frac{x^2+a^2-a^2}{a^2+x^2}dx$$

$$= \frac{4a}{\pi}\int_{0}^{a}\left(1-\frac{a^2}{a^2+x^2}\right)dx$$

$$= \frac{4a}{\pi}\left|x - a\tan^{-1}\frac{x}{a}\right|_0^a$$

$$= \frac{4a}{\pi}(a - a\tan^{-1}1)$$

$$= \frac{4a}{\pi}\left(a - a\frac{\pi}{4}\right)$$

$$= \frac{a^2(4-\pi)}{\pi}$$

$$\mu_2 = \mu_2' - (\mu_1')^2 = \frac{a^2(4-\pi)}{\pi} - 0 = \frac{a^2(4-\pi)}{\pi}$$

$$\mu_4 = \mu_4' \qquad (\because \mu_1' = 0)$$

$$\mu_4 = \int_{-\infty}^{\infty} x^4 \cdot f(x)\,dx$$

$$= \int_{-a}^{a} x^4 \cdot \frac{2a}{\pi}\left(\frac{1}{a^2+x^2}\right)dx$$

$$= \frac{2a}{\pi}\int_{-a}^{a}\frac{x^4}{a^2+x^2}dx$$

$$= \frac{4a}{\pi}\int_{0}^{a}\left(x^2 - a^2 + \frac{a^4}{a^2+x^2}\right)dx$$

$$= \frac{4a}{\pi}\left|\frac{1}{3}x^3 - a^2x + a^3\tan^{-1}\frac{x}{a}\right|_0^a$$

$$= \frac{4a}{\pi}\left(\frac{a^3}{3} - a^3 + a^3\tan^{-1}1\right)$$

$$= \frac{4a}{\pi}\left(\frac{a^3}{3} - a^3 + a^3\frac{\pi}{4}\right)$$

$$= a^4\left(1 - \frac{8}{3\pi}\right)$$

EXERCISE 3.4

1. If the probability density function is given by

$$f(x) = kx^2(1-x^3) \qquad 0 \le x \le 1$$
$$= 0 \qquad \text{otherwise}$$

Find (i) k, (ii) $P\left(0 < X < \frac{1}{2}\right)$, (iii) $\bar{X}$, and (iv) σ^2.

$$\left[\textbf{Ans.: } \text{(i) } 6 \text{ (ii) } \frac{15}{64} \text{ (iii) } \frac{9}{14} \text{ (iv) } \frac{9}{245}\right]$$

2. If the probability density function of a random variable is given by

$$f(x) = kx \qquad 0 \le x \le 2$$
$$= 2k \qquad 2 \le x \le 4$$
$$= 6k - kx \qquad 4 \le x \le 6$$

Find (i) k, (ii) $P(1 \le X \le 3)$, and (iii) $\bar{X}$.

$$\left[\textbf{Ans.: } \text{(i) } \frac{1}{2} \text{ (ii) } \frac{1}{3} \text{ (iii) } \frac{383}{36}\right]$$

Scanned with CamScanner

3. If the probability density of a random variable is given by

$$f(x) = k\,x\,e^{-\frac{x}{3}} \quad x > 0$$
$$= 0 \quad x \le 0$$

Find (i) k, (ii) $\overline{X}$, and (iii) σ^2.

$\left[\text{Ans.: (i) } \frac{1}{9} \text{ (ii) } 6 \text{ (iii) } 18\right]$

4. A continuous random variable has the probability density function

$$f(x) = 2e^{-2x} \quad x > 0$$
$$= 0 \quad x \le 0$$

Find (i) $E(X)$, (ii) $E(\overline{X})$, (iii) Var (X), and (iv) SD of X.

$\left[\text{Ans.: (i) } \frac{1}{2} \text{ (ii) } \frac{1}{2} \text{ (iii) } \frac{1}{4} \text{ (iv) } \frac{1}{2}\right]$

5. A random variable X has the pdf

$$f(x) = \frac{k}{1+x^2}, -\infty < x < \infty$$

Determine (i) k, (ii) $P(X \ge 0)$, (iii) mean, and (iv) variance.

$\left[\text{Ans.: (i) } \frac{1}{\pi} \text{ (ii) } \frac{1}{2} \text{ (iii) } 0 \text{ (iv) does not exist}\right]$

6. The distribution function of a continuous random variable X is given by $F(x) = 1-(1+x)e^{-x}, x \ge 0$. Find (i) pdf, (ii) mean, and (iii) variance.

$\left[\text{Ans.: (i) } f(x) = x\,e^{-x}, x \ge 0 \text{ (ii) } 2 \text{ (iii) } 2\right]$

7. If $f(x)$ is the probability density function of a continuous random variable, find k, mean, and variance.

$$f(x) = kx^2 \quad 0 \le x \le 1$$
$$= (2-x)^2 \quad 1 \le x \le 2$$

$\left[\text{Ans.: } 2, \frac{11}{12}, 0.626\right]$

8. A continuous random variable X has the probability density function given by

$$f(x) = 2ax + b \quad 0 \le x \le 2$$
$$= 0 \quad \text{otherwise}$$

If the mean of the distribution is 3, find the constants a and b.

$\left[\text{Ans.: } \frac{3}{2}, -\frac{5}{2}\right]$

9. If X is a continuous random variable with probability density function given by

$$f(x) = k(x-x^3) \quad 0 \le x \le 1$$
$$= 0 \quad \text{otherwise}$$

Find (i) k, (ii) mean, (iii) variance, and (iv) median.

$\left[\text{Ans.: (i) } \frac{1}{2} \text{ (ii) } 0.06 \text{ (iii) } 0.04 \text{ (iv) } 2\right]$

10. The probability density function of a random variable is given by

$$f(x) = 0 \quad x < 2$$
$$= \frac{2x+3}{18} \quad 2 \le x \le 4$$
$$= 0 \quad x > 4$$

Find the mean and variance.

$\left[\text{Ans.: (i) } \frac{83}{27}, 0.33\right]$

11. A continuous random variable X has the probability density function

$$f(x) = x^3 \quad 0 \le x \le 1$$
$$= (2-x)^3 \quad 1 \le x \le 2$$
$$= 0 \quad \text{otherwise}$$

Find $P(0.5 \le X \le 1.5)$ and mean of the distribution.

$\left[\text{Ans.: } \frac{15}{32}, \frac{1}{2}\right]$

12. The probability density function of a continuous random variable X is given by

$$f(x) = kx\,(2-x) \quad 0 \le x \le 2$$

Find k, mean, and variance.

$\left[\text{Ans.: } \frac{3}{4}, 1, \frac{1}{5}\right]$

13. If the density function of a continuous random variable X is given by $f(x) = \lambda e^{-\lambda(x-a)}, a \le x < \infty$, show that $\beta_1 = 4$ and $\beta_2 = 9$.

Scanned with CamScanner

14. If the continuous random variable has the density function $f(x)=\dfrac{kx}{(1+x)^3}, x \geq 0$, find the value of k, median and mode.

$$\left[\text{Ans.: } 2, 1+\sqrt{2}, \frac{1}{2}\right]$$

15. The density function of a continuous random variable X is given by $f(x)=\dfrac{3}{4}x(2-x), 0 \leq x \leq 2$. Find the mean, median, mode, harmonic mean, MD about mean and SD.

$$\left[\text{Ans.: } 1, 1, 1, \frac{2}{3}, \frac{3}{8}, \frac{1}{\sqrt{5}}\right]$$

16. The density function of a continuous random variable X is given by

$$f(x)\begin{cases} \dfrac{1}{16}(3+x^2) & -3 \leq x \leq -1 \\ \dfrac{1}{16}(6-2x^2) & -1 \leq x \leq 1 \\ \dfrac{1}{16}(3-x^2) & 1 \leq x \leq 3 \end{cases}$$

Find the mean, SD and MB about the mean.

$$\left[\text{Ans.: } 0, 1, \frac{13}{16}\right]$$

3.8 EXPECTED VALUES OF TWO DIMENSIONAL RANDOM VARIABLES

If (X, Y) is a two dimensional discrete random variable with joint probability mass function $P(x_i, y_j) = p_{ij}$, then the mathematical expectation of a function $g(x, y)$ is given by

$$E[g(X,Y)] = \sum_{j=1}^{\infty}\sum_{i=1}^{\infty} g(x_i, y_j)p_{ij}$$

$$= \sum_{x}\sum_{y} g(x,y)f(x),y)$$

If (X, Y) is a two dimensional continuous random variable with joint probability density function $f(x, y)$, then the mathematical expectation of a function $g(x, y)$ is given by

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

3.8.1 Properties of Expected Values of Two Dimensional Random Variables

(i) If X and Y are random variables, then $E(X+Y) = E(X) + E(Y)$ provided all the expectations exist.

(ii) If X and Y are independent random variables then $E(XY) = E(X) \cdot E(Y)$.

3.8.2 Conditional Expectation and Conditional Variance

If (X,Y) is a two dimensional discrete random variable with joint probability mass function p_{ij} then the conditional expectations of $g(X, Y)$ is given by

$$E\{g(X,Y)/Y=y_j\} = \sum_{i=1}^{\infty} g(x_i,y_j)P(X=x_i/Y=y_j)$$

$$= \sum_{i=1}^{\infty} \frac{g(x_i,y_j)P(X=x_i, Y=y_j)}{P(Y=y_j)}$$

$$= \sum_{i=1}^{\infty} g(x_i,y_j)\frac{p_{ij}}{p_{*j}}$$

In particular, the conditional expectation of a discrete random variable X given $Y = y_j$ is given by

$$E(X/Y=y_j) = \sum_{i=1}^{\infty} x_i P(X=x_i/Y=y_j)$$

The conditional variance of X given $Y = y_j$ is given by

$$\text{Var}(X/Y=y_j) = E\left[\{X-E(X/Y=y_j)\}^2/Y=y_j\right]$$

If (X, Y) is a two-dimensional continuous random variable with joint probability density function $f(x, y)$, then the conditional expectation of $g(X, Y)$ is given by

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

$$= \int_{-\infty}^{\infty} \frac{g(x,y)f(x,y)dx}{f_Y(y)}$$

In particular, the conditional expectation of X given $Y = y$ is given by

$$E(X/Y=y) = \frac{\int_{-\infty}^{\infty} xf(x,y)dx}{f_Y(y)}$$

Similarly, $$E(Y/X=x) = \frac{\int_{-\infty}^{\infty} yf(x,y)dx}{f_X(x)}$$

Scanned with CamScanner

The conditional variance of X is given by

$$\text{Var}(X/Y=y) = E[\{X - E(X/Y=y)\}^2 / Y=y]$$

Similarly, $$\text{Var}(Y/X=x) = E[\{Y - E(Y/X=x)\}^2 / X=x]$$

3.8.3 Properties of Conditional Expectation

(i) If x and y are independent random variables, then

$$E(Y/X) = E(Y)$$

and $$E(X/Y) = E(X)$$

(ii) $E(XY) = E[X.\, E(Y/X)]$

(iii) $E(X^2Y^2) = E(X^2.\, E(Y^2/X)]$

Example 1

Given a pair of discrete random variable X and Y whose joint probability distribution is given by

Y \ X	2	4
1	0.1	0.15
2	0.2	0.3
3	0.1	0.15

Find the expected value of the function $g(X, Y)$ given that $g(X, Y) = 2X + Y$.

Solution

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

$$= \sum_x \sum_y (2x+y) f(x,y)$$

$$= \{2(2)+1\}0.1 + \{2(4)+1\}0.15$$
$$+ \{2(2)+2\}0.2 + \{2(4)+2\}0.3$$
$$+ \{2(2)+3\}0.1 + \{2(4)+3\}0.15$$
$$= 8.4$$

Example 2

Let X and Y be two random variables each taking values –1, 0 and 1 and having the joint probability distribution as given below:

Y \ X	–1	0	1	Total p(y)
–1	0	0.1	0.1	0.2
0	0.2	0.2	0.2	0.6
1	0	0.1	0.1	0.2
Total p(x)	0.2	0.4	0.4	1.0

(i) Show that X and Y have different expectation.

(ii) Find E(XY)

(iii) Find Var(X) and Var(Y).

(iv) Given that Y = 0, what is the conditional probability distribution of X?

Solution

(i) $$E(X) = \sum xp(x)$$
$$= -1(0.2) + 0(0.4) + 1(0.4)$$
$$= 0.2$$

$$E(Y) = \sum yp(y)$$
$$= -1(0.2) + 0(0.6) + 1(0.2)$$
$$= 0$$

$$E(X) \neq E(Y)$$

(ii) $$E(XY) = \sum x_i y_j p_{ij}$$
$$= (-1)(-1)(0) + (0)(-1)(0.1) + (1)(-1)(0.1)$$
$$+ (-1)(0)0.2 + (0)(0)(0.2) + (1)(0)(0.2)$$
$$+ (-1)(1)(0) + (0)(1)(0.1) + (1)(1)(0.1)$$
$$= 0$$

(iii) $$E(X^2) = \sum x^2 p(x)$$
$$= (-1)^2(0.2) + 0(0.4) + (1)^2(0.4)$$
$$= 0.6$$

$$\text{Var}(X) = E(X^2) - \{E(X)\}^2$$
$$= 0.6 - (0.2)^2$$
$$= 0.56$$

$$E(Y^2) = \sum y^2 p(y)$$
$$= (-1)^2(0.2) + 0(0.6) + (1)^2(0.2)$$
$$= 0.4$$

Scanned with CamScanner

$$\mathrm{Var}(Y) = E(Y^2) - \{E(Y)\}^2$$
$$= 0.4 - 0$$
$$= 0.4$$

(iv) $$P(X=-1/Y=0) = \frac{P(X=-1, Y=0)}{P(Y=0)}$$
$$= \frac{0.2}{0.6} = \frac{1}{3}$$
$$P(X=0/Y=0) = \frac{P(X=0, Y=0)}{P(Y=0)}$$
$$= \frac{0.2}{0.6} = \frac{1}{3}$$
$$P(X=1/Y=0) = \frac{P(X=1, Y=0)}{P(Y=0)}$$
$$= \frac{0.2}{0.6} = \frac{1}{3}$$

Example 3

If the joint pdf of (X,Y) is given by

$$f(x,y) = \begin{cases} \dfrac{16y}{x^3} & x > 2, 0 < y < 1 \\ 0 & \text{elsewhere} \end{cases}$$

then find E(X,Y).

Solution

$$E(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)\,dx\,dy$$
$$= \int_0^1\int_2^{\infty} xy\left(\frac{16y}{x^3}\right)dx\,dy$$
$$= 16\int_0^1\int_2^{\infty}\left(\frac{y^2}{x^2}\right)dx\,dy$$
$$= 16\int_0^1 y^2\left|-\frac{1}{x}\right|_2^{\infty} dy$$
$$= 16\int_0^1 \frac{1}{2}y^2 dy$$

$$= 8\left|\frac{y^3}{3}\right|_0^1$$
$$= \frac{8}{3}(1-0)$$
$$= \frac{8}{3}$$

Example 4

The joint PDF of (X,Y) is given by

$$f(x,y) = 24xy, \quad x > 0, y > 0, x + y \le 1$$
$$= 0, \quad \text{elsewhere}$$

Find the conditional mean and variance of Y, given X.

Solution

The region of integration is ΔOAB.

In ΔOAB, along vertical strip PQ, limits of y: $y = 0$ to $y = 1 - x$ and x varies from $x = 0$ to $x = 1$.

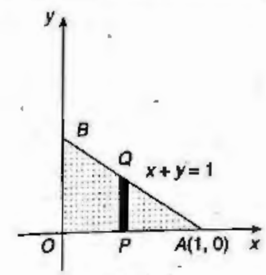


Fig. 3.4

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)dy$$
$$= \int_0^{1-x} 24xy\,dy$$
$$= 24x\left|\frac{y^2}{2}\right|_0^{1-x}$$
$$= 12x(1-x)^2, \quad 0 \le x \le 1$$
$$f(y/x) = \frac{f(x,y)}{f_X(x)}$$
$$= \frac{24xy}{12x(1-x)^2}$$
$$= \frac{2y}{(1-x)^2}$$
$$E(Y/X=x) = \int_{-\infty}^{\infty} yf(y/x)dy$$
$$= \int_0^{1-x} y\cdot\frac{2y}{(1-x)^2}dy$$

Scanned with CamScanner

$$= \frac{2}{(1-x)^2}\left|\frac{y^3}{3}\right|_0^{1-x}$$

$$= \frac{2}{3}(1-x)$$

$$E(Y^2 / x) = \int_{-\infty}^{\infty} y^2 f(y/x)\mathrm{d}y$$

$$= \int_0^{1-x} y^2 \frac{2y}{(1-x)^2}\mathrm{d}y$$

$$\mathrm{Var}(Y^2/x) = (E(Y^2/x) - \{E(Y/x)\}^2$$

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

$$= \frac{1}{2}(1-x)^2 - \frac{4}{9}(1-x)^2$$

$$= \frac{1}{18}(1-x)^2$$

$$= \frac{1}{96}\int_1^5 y^2 \left|\frac{x^3}{3}\right|_0^4 \mathrm{d}y$$

$$= \frac{1}{288}\int_1^5 y^2 \left|x^3\right|_0^4 \mathrm{d}y$$

$$= \frac{1}{288}\int_1^5 64y^2\mathrm{d}y$$

$$= \frac{2}{9}\left|\frac{y^3}{3}\right|_1^5$$

$$= \frac{2}{27}(125-1)$$

$$= \frac{248}{27}$$

(iv) $E(2X+3Y) = 2E(X)+3E(Y)$

$$= 2\left(\frac{8}{3}\right) + 3\left(\frac{31}{9}\right)$$

$$= \frac{47}{3}$$

(v) $E(X^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^2 f(x,y)\mathrm{d}x\,\mathrm{d}y$

$$= \int_1^5\int_0^4 x^2\left(\frac{xy}{96}\right)\mathrm{d}x\,\mathrm{d}y$$

$$= \frac{1}{96}\int_1^5 y\left|\frac{x^4}{4}\right|_0^4 \mathrm{d}y$$

$$= \frac{1}{384}\int_1^5 256y\,\mathrm{d}y$$

$$= \frac{2}{3}\left|\frac{y^2}{2}\right|_1^5$$

$$= \frac{1}{3}(25-1)$$

$$= 8$$

$$\mathrm{Var}(X) = E(X^2) - \{E(X)\}^2$$

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

$$= \frac{8}{9}$$

(vi) $E(Y^2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y^2 f(x,y)\mathrm{d}x\mathrm{d}y$

$$= \int_1^5\int_0^4 y^2\left(\frac{xy}{96}\right)\mathrm{d}x\,\mathrm{d}y$$

$$= \frac{1}{96}\int_1^5 y^3\left|\frac{x^2}{2}\right|_0^4 \mathrm{d}y$$

$$= \frac{1}{192}\int_1^5 16y^3\,\mathrm{d}y$$

$$= \frac{1}{12}\left|\frac{y^4}{4}\right|_1^5$$

$$= \frac{1}{48}(625-1)$$

$$= 13$$

Scanned with CamScanner

$$\text{Var}(Y) = E(Y^2) - \{E(Y)\}^2$$
$$= 13 - \left(\frac{31}{9}\right)^2$$
$$= \frac{92}{81}$$

(vii) $\text{Cov}(X,Y) = E(XY) - E(X)E(Y)$
$$= \frac{248}{27} - \left(\frac{8}{3}\right)\left(\frac{31}{9}\right)$$
$$= 0$$

Example 5

Two random variables X and Y have the following joint probability density function:

$$f(x, y) = \begin{cases} 2 - x - y, & 0 \le x \le 1, 0 \le y \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Find (i) Marginal probability density function of X and Y.
(ii) Conditional density functions
(iii) Var(X) and Var(Y)
(iv) Covariance between X and Y

Solution

(i) $f_X(x) = \int_{-\infty}^{\infty} f(x, y)\,dy$
$$= \int_0^1 (2 - x - y)\,dy$$
$$= \left|2y - xy - \frac{y^2}{2}\right|_0^1$$
$$= \left(2 - x - \frac{1}{2}\right)$$
$$= \frac{3}{2} - x$$

$$\therefore \quad f_X(x) = \begin{cases} \frac{3}{2} - x, & 0 < x < 1 \\ 0, & \text{otherwise} \end{cases}$$

Similarly, $f_Y(y) = \begin{cases} \frac{3}{2} - y, & 0 < y < 1 \\ 0, & \text{otherwise} \end{cases}$

(ii) $f_{X/Y}(x/y) = \frac{f(x,y)}{f_Y(y)}$
$$= \frac{(2 - x - y)}{\left(\frac{3}{2} - y\right)}, \quad 0 < (x, y) < 1$$

$$f_{Y/X}(y/x) = \frac{f(x,y)}{f_X(x)}$$
$$= \frac{(2 - x - y)}{\left(\frac{3}{2} - x\right)}$$

(iii) $E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx$
$$= \int_0^1 x\left(\frac{3}{2} - x\right)dx$$
$$= \left|\frac{3x^2}{4} - \frac{x^3}{3}\right|_0^1$$
$$= \frac{3}{4} - \frac{1}{3}$$
$$= \frac{5}{12}$$

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y)\,dy$$
$$= \int_0^1 y\left(\frac{3}{2} - y\right)dy$$
$$= \left|\frac{3y^2}{4} - \frac{y^3}{3}\right|_0^1$$
$$= \frac{3}{4} - \frac{1}{3}$$
$$= \frac{5}{12}$$

Scanned with CamScanner

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

$$= \int_0^1 x^2\left(\frac{3}{2} - x\right)dx$$

$$= \left|\frac{x^3}{2} - \frac{x^4}{4}\right|_0^1$$

$$= \frac{1}{2} - \frac{1}{4}$$

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

$$\text{Var}(X) = E(X^2) = [E(X)]^2$$

$$= \frac{1}{4} - \left(\frac{5}{12}\right)^2$$

$$= \frac{11}{144}$$

Similarly, $\text{Var}(Y) = \frac{11}{144}$

Example 6

If the joint pdf of (X, Y) is given by

$$f(x, y) = 24y(1 - x), 0 \le y \le x \le 1,$$

then find E(XY).

Solution

The region of integration is ΔOAB. In ΔOAB, along horizontal strip $P'Q'$, Limits of x: $x = y$ to $x = 1$ and y varies from $y = 0$ to $y = 1$.

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f(x, y)\,dx\,dy$$

$$= \int_0^1\int_y^1 xy \cdot 24y(1 - x)\,dx\,dy$$

$$= 24\int_0^1\int_y^1 xy^2(1 - x)\,dx\,dy$$

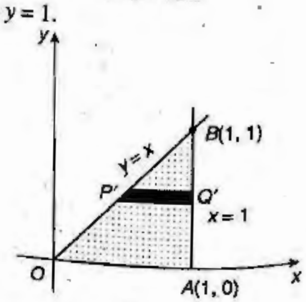


Fig. 3.5

$$= 24\int_0^1 y^2\left|\frac{x^2}{2} - \frac{x^3}{3}\right|_y^1 dy$$

$$= 24\int_0^1 y^2\left(\frac{1}{2} - \frac{1}{3} - \frac{y^2}{2} + \frac{y^3}{3}\right)dy$$

$$= 24\int_0^1 y^2\left(\frac{1}{6} - \frac{y^2}{2} + \frac{y^3}{3}\right)dy$$

$$= 24\int_0^1\left(\frac{y^2}{6} - \frac{y^4}{2} + \frac{y^5}{3}\right)dy$$

$$= 24\left|\frac{y^3}{18} - \frac{y^5}{10} + \frac{y^6}{18}\right|_0^1$$

$$= 14\left(\frac{1}{18} - \frac{1}{10} + \frac{1}{18}\right)$$

$$= \frac{4}{15}$$

Example 7

Two random variables have joint pdf

$$f(x, y) = \begin{cases} \frac{xy}{96}, & 0 < x < 4, \quad 1 < y < 5 \\ 0\ , & \text{elsewhere} \end{cases}$$

Find (i) $E(X)$ (ii) $E(Y)$ (iii) $E(XY)$ (iv) $E(2X + 3Y)$ (v) $\text{Var}(X)$ (vi) $\text{Var}(Y)$ (vii) $\text{Cov}(X,Y)$

Solution

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

$$= \int_1^5\int_0^4 x\left(\frac{xy}{96}\right)dx\,dy$$

$$= \frac{1}{96}\int_1^5 y\left|\frac{x^3}{3}\right|_0^4 dy$$

$$= \frac{1}{96}\int_1^5 y\left(\frac{64}{3}\right)dy$$

Scanned with CamScanner

$$=\frac{2}{9}\left|\frac{y^2}{2}\right|_1^5$$

$$=\frac{2}{9}\left(\frac{25}{2}-\frac{1}{2}\right)$$

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

(ii) $$E(Y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f(x,y)\,dx\,dy$$

$$=\int_1^5\int_0^4 y\left(\frac{xy}{96}\right)dx\,dy$$

$$=\frac{1}{96}\int_1^5 y^2\left|\frac{x^2}{2}\right|_0^4 dy$$

$$=\frac{1}{96}\int_1^5 8y^2\,dy$$

$$=\frac{1}{12}\left|\frac{y^3}{3}\right|_1^5$$

$$=\frac{1}{36}(125-1)$$

$$=\frac{31}{9}$$

(iii) $$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,f(x,y)\,dx\,dy$$

$$=\int_1^5\int_0^4 xy\left(\frac{xy}{96}\right)dx\,dy$$

$$=\frac{1}{96}\int_1^5\int_0^4 x^2y^2\,dx\,dy$$

(iv) $$E(XY)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\,f(x,y)\,dx\,dy$$

$$=\int_0^1\int_0^1 xy(2-x-y)\,dx\,dy$$

$$=\int_0^1 y\left|x^2-\frac{x^3}{3}-\frac{x^2y}{2}\right|_0^1 dy$$

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

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

$$=\left|\frac{y^2}{3}-\frac{y^3}{6}\right|_0^1$$

$$=\frac{1}{3}-\frac{1}{6}$$

$$=\frac{1}{6}$$

(v) $$\mathrm{Cov}(X,Y)=E(XY)-E(X)E(Y)$$

$$=\frac{1}{6}-\left(\frac{5}{12}\right)\left(\frac{5}{12}\right)$$

$$=-\frac{1}{144}$$

Example 8

Let $f(x,y)=8xy,\quad 0<x<y<1$

$\qquad\qquad = 0\quad,\quad$ elsewhere

Find (i) $E\,(Y/X=x)$ *(ii)* $E(XY/X=x)$ (iii) $\mathrm{Var}(Y/X=x)$.

Solution

The region of integration is ΔOAB. In ΔOAB, along vertical strip PQ, limits of y: $y=x$ to $y=1$ and x varies from $x=0$ to $x=1$.

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)\,dy$$

$$=\int_x^1 8xy\,dy$$

$$=8x\left|\frac{y^2}{2}\right|_x^1$$

$$=4x(1-x^2)\quad 0<x<1$$

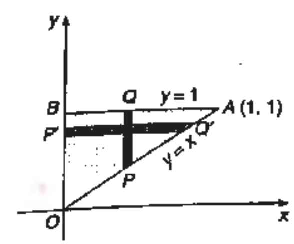


Fig. 3.6

In ΔOAB, along horizontal strip $P'Q'$,

Scanned with CamScanner

Limits of x : $x=0$ to $x=y$ and y varies from $y=0$ to $y=1$

$$f_Y(y)=\int_{-\infty}^{\infty} f(x,y)\,dx$$
$$=\int_0^y 8xy\,dx$$
$$=8y\left|\frac{x^2}{2}\right|_0^y$$
$$=4y^3, \qquad 0<y<1$$

$$f_{X/Y}(x/y)=\frac{f(x,y)}{f_Y(y)}$$
$$=\frac{8xy}{4y^3}$$
$$=\frac{2x}{y^2}$$

$$f_{Y/X}(y/x)=\frac{f(x,y)}{f_X(x)}$$
$$=\frac{8xy}{4x(1-x^2)}$$
$$=\frac{2y}{1-x^2}$$

(i) $E(Y/X=x)=\int_{-\infty}^{\infty} y\,f_{Y/X}(y/x)\,dy$

$$=\int_x^1 y\left(\frac{2y}{1-x^2}\right)dy$$
$$=\frac{2}{1-x^2}\left|\frac{y^3}{3}\right|_x^1$$
$$=\frac{2}{3}\left(\frac{1-x^3}{1-x^2}\right)$$
$$=\frac{2}{3}\left(\frac{1+x+x^2}{1+x}\right)$$

(ii) $E(XY/X=x)=x\,E(Y/X=x)$

$$=\frac{2}{3}\frac{x(1+x+x^2)}{(1+x)}$$

(iii) $E(Y^2/X=x)=\int_{-\infty}^{\infty} y^2 f_{Y/X}(y/x)\,dy$

$$=\int_x^1 y^2\left(\frac{2y}{1-x^2}\right)dy$$
$$=\frac{2}{1-x^2}\left|\frac{y^4}{4}\right|_x^1$$
$$=\frac{1}{2}\left(\frac{1-x^4}{1-x^2}\right)$$

$$\mathrm{Var}(Y/X=x)=E(Y^2/X=x)-\{E(Y/X=x)\}^2$$
$$=\frac{1+x^2}{2}-\left[\frac{2}{3}\left(\frac{1+x+x^2}{1+x}\right)\right]^2$$
$$=\frac{1+x^2}{2}-\frac{4}{9}\frac{(1+x+x^2)^2}{(1+x)^2}$$

EXERCISE 3.5

1. If the pdf of (X, Y) is given by
$$f(x, y)=2-x-y,\ 0\le x\le y\le 1$$
Find $E(X)$ and $E(Y)$. $\left[\text{Ans.: } \frac{5}{12}, \frac{5}{12}\right]$

2. If $f(x,y)=\begin{cases}\frac{1}{\pi}, & 0<x^2+y^2<1\\ 0, & x^2+y^2>1\end{cases}$ [Ans.: 0]

Find the covariance of X, Y.

3. Joint pdf of X and Y is given by
$$f(x, y)=3(x+y) \qquad 0\le x\le 1,\ 0\le y\le 1$$
Find $E(Y/X=x)$ and Cov(X,Y). $\left[\text{Ans.: } \frac{(1-x)(x+2)}{3(1+x)}, -\frac{13}{320}\right]$

4. Let $f_{XY}(x, y)=e^{-(x+y)}$ $\quad 0\le x<\infty,\ 0<y<\infty$.
Find Cov(X, Y). [Ans.: 0]

Scanned with CamScanner

5. If the joint pdf of (X, Y) is given by

$$f(x, y) = 2, \quad 0 \le x < y \le 1,$$

find the conditional mean and conditional variance of X given that $Y = y$.

$$\left[\textbf{Ans.: } \frac{y}{2}, \frac{y^2}{12}\right]$$

6. If the joint pdf of (X, Y) is given by

$$f(x, y) = 21\, x^2 y^3, \; 0 \le x < y \le 1$$

find the conditional mean and conditional variance of X, given that $Y = y$, $0 < y < 1$.

$$\left[\textbf{Ans.: } \frac{3y}{4}, \frac{3y^2}{80}\right]$$

7. If the joint pdf of (X, Y) is given by

$$f(x, y) = 3xy\,(x + y), \; 0 < x \le y \le 1,$$

verify that $E\{E(Y/X)\} = E(Y) = \dfrac{17}{24}$.

3.9 BOUNDS ON PROBABILITIES

If the probability distribution of a random variable is known $E(X)$ and $\text{Var}(X)$ can be computed. Conversely, if $E(X)$ and $\text{Var}(X)$ are known, probability distribution of X cannot be constructed and quantities such as $P\{|X - E(X)| \le k\}$ can not be evaluated. Several approximation techniques have been developed to yield upper and /or lower bounds to such probabilities. The most important of such techniques is Chebyshev's inequality.

3.10 CHEBYSHEV'S INEQUALITY

If X is a random variable with mean μ and variance σ^2, then for any positive umber k,

$$P\{|X - \mu| \ge k\sigma\} \le \frac{1}{k^2}$$

or

$$P\{|X - \mu| < k\sigma\} \ge 1 - \frac{1}{k^2}$$

Proof

Let X be a continuous random variable.

$$\sigma^2 = E[X - E(X)]^2$$

$$= E[X - \mu]^2 \qquad [\because \mu = E(X)]$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx \quad \text{where } f(x) \text{ is pdf of } X.$$

$$= \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu-k\sigma}^{\mu+k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx$$

$$\ge \int_{-\infty}^{\mu-k\sigma} (x-\mu)^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} (x-\mu)^2 f(x)\,dx \qquad \text{...(1)}$$

We know that $x \le \mu - k\sigma$ and $x \ge \mu + k\sigma$

$\therefore \quad |x - \mu| \ge k\sigma$

Substituting in Eq. (1),

$$\sigma^2 \ge \int_{-\infty}^{\mu-k\sigma} k^2\sigma^2 f(x)\,dx + \int_{\mu+k\sigma}^{\infty} k^2\sigma^2 f(x)\,dx$$

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

$$= k^2\sigma^2 [P(X \le \mu - k\sigma) + P(X \ge \mu + k\sigma)]$$

$$= k^2\sigma^2 [P(X - \mu \le -k\sigma) + P(X - \mu \ge k\sigma)]$$

$$= k^2\sigma^2 P\{|X - \mu| \ge k\sigma\}$$

$$P\{|X - \mu| \ge k\sigma\} \le \frac{1}{k^2}$$

$$\because P\{|X - \mu| \ge k\sigma\} + P\{|X - \mu| < k\sigma\} = 1$$

$$P\{|X - \mu| < k\sigma\} = 1 - P\{|X - \mu| \ge k\sigma\}$$

$$\ge 1 - \frac{1}{k^2}$$

Note

1. If $k\sigma = c > 0$

$$P\{|X - \mu| \ge c\} \le \frac{\sigma^2}{c^2}$$

and

$$P\{|X - \mu| < c\} \ge 1 - \frac{\sigma^2}{c^2}$$

2. To find the lower bound of probabilities following form of Chebyshev's inequality is used:

$$P\{|X - \mu| < k\sigma\} \ge 1 - \frac{1}{k^2}$$

Scanned with CamScanner

or $$P\{|X-\mu|<c\}\geq 1-\frac{\sigma^2}{c^2}$$

3. To find the upper bound of probabilities following form of Chebyshev's inequality is used;

$$P\{|X-\mu|\geq k\sigma\}\leq \frac{1}{k^2}$$

or $$P\{|X-\mu|\geq c\}\leq \frac{\sigma^2}{c^2}$$

Example 1

A random variable X has a mean $\mu = 12$ and a variance $\sigma^2 = 9$ and unknown probability distribution. Find $P(6 < X < 18)$.

Solution

$$\mu=12,\quad \sigma^2=9$$
$$\sigma=3$$

By Chebyshev's inequality,

$$P\{|X-\mu|<k\sigma\}\geq 1-\frac{1}{k^2}$$
$$P\{-k\sigma<X-\mu<k\sigma\}\geq 1-\frac{1}{k^2}$$
$$P\{\mu-k\sigma<X<\mu+k\sigma\}\geq 1-\frac{1}{k^2}$$
$$P\{12-3k<X<12+3k\}\geq 1-\frac{1}{k^2}$$

Comparing with $P(6 < X < 18)$,

$$12-3k=6$$
$$12+3k=18$$
$$\therefore \quad k=2$$
$$P\{6<X<18\}\geq 1-\frac{1}{4}$$
$$P\{6<X<18\}\geq \frac{3}{4}$$

Example 2

A random variable X has a mean 10 and a variance 4 and unknown probability distribution. Find the value of c such that $P\{|X-10|\geq c\}\leq 0.04$.

Solution

$$\mu=10,\ \sigma^2=4$$
$$\sigma=2$$

By Chebyshev's inequality,

$$P\{|X-\mu|\geq k\sigma\}\leq \frac{1}{k^2}$$

Comparing with $P\{|X-10|\geq c\}\leq 0.04$,

$$\frac{1}{k^2}=0.04$$
$$k=5$$

and
$$k\sigma=c$$
$$c=5(2)=10$$

Example 3

A random variable X has pdf $f(x)=e^{-x}, x\geq 0$. Use Chebyshev's inequality to show that $P\{|X-1|>2\}\leq \frac{1}{4}$ and also, show that the actual probability is given by e^{-3}.

Solution

$$f(x)=e^{-x}$$

The random variable X follows exponential distribution with parameter $\lambda = 1$.

$$E(X)=\mu=\frac{1}{\lambda}=1$$
$$\text{Var}(X)=\sigma^2=\frac{1}{\lambda^2}=1$$

By Chebyshev's inequality,

$$P\{|X-\mu|>k\sigma\}\leq \frac{1}{k^2}$$

Scanned with CamScanner

Comparing with $P\{|X-\mu|>2\}$,

$$k\sigma = 2$$
$$k(1) = 2$$
$$k = 2$$

$$\therefore \quad P\{|X-1|>2\} \le \frac{1}{4}$$

The actual probability is given by

$$P\{|X-1|>2\} = 1 - P\{|X-1| \le 2\}$$
$$= 1 - P\{-1 < X \le 3\}$$
$$= 1 - P\{0 < X \le 3\}$$
$$= 1 - \int_0^3 e^{-x}\,dx$$
$$= 1 - \left|e^{-x}\right|_0^3$$
$$= 1 - e^{-3}$$

Example 4

A random variable X is exponentially distributed with parameter 1. Use Chebyshev's inequality to show that $P\{-1 \le X \le 3\} \ge \frac{3}{4}$. Find the actual probability also.

Solution

For an exponential distribution with parameter $\lambda = 1$,

$$E(X) = \mu = \frac{1}{\lambda} = 1$$
$$\text{Var}(X) = \sigma^2 = \frac{1}{\lambda^2} = 1$$
$$\sigma = 1$$

By Chebyshev's inequality,

$$P\{|X-\mu| < k\sigma\} \ge 1 - \frac{1}{k^2}$$
$$P\{-k\sigma < X-\mu < k\sigma\} \ge 1 - \frac{1}{k^2}$$

$$P\{\mu - k\sigma < X < \mu + k\sigma\} \ge 1 - \frac{1}{k^2}$$
$$P\{1-k < X < 1+k\} \ge 1 - \frac{1}{k^2}$$

Comparing with $P\{-1 \le X \le 3\} \ge \frac{3}{4}$,

$$1 - k = -1$$
$$k = 2$$

$$\therefore \quad P\{-1 \le X \le 3\} \ge 1 - \frac{1}{4}$$
$$\ge \frac{3}{4}$$

The actual probability is given by

$$P\{-1 \le X \le 3\} = P\{0 \le X \le 3\} \quad [\because x > 0 \text{ for exponential distribution}]$$
$$= \int_0^3 f(x)\,dx$$
$$= \int_0^3 e^{-x}\,dx$$
$$= \left|-e^{-x}\right|_0^3$$
$$= -e^{-3} + e^0$$
$$= 1 - e^{-3}$$
$$= 0.9502$$

Example 5

A fair dice is tossed 120 times. Use Chebyshev's inequality to find a lower bound for the probability of getting 80 to 120 sixes.

Solution

Let X be the random variable which denotes number of sixes obtained when a fair dice is tossed by 720 times.

$$n = 720$$

Probability of getting 6 in single toss

$$p = \frac{1}{6}$$
$$q = 1 - p = 1 - \frac{1}{6} = \frac{5}{6}$$

X follows a binomial distribution.

$$\mu = np = (720)\left(\frac{1}{6}\right) = 120$$

$$\sigma^2 = npq = (720)\left(\frac{1}{6}\right)\left(\frac{5}{6}\right) = 100$$

$$\sigma = 10$$

By Chebyshev's inequality,

$$P\{|X-\mu| < k\sigma\} \ge 1-\frac{1}{k^2}$$

$$P\{-k\sigma < X-\mu < k\sigma\} \ge 1-\frac{1}{k^2}$$

$$P\{\mu-k\sigma < X < \mu+k\sigma\} \ge 1-\frac{1}{k^2}$$

$$P\{120-10k < X < 120+10k\} \ge 1-\frac{1}{k^2}$$

Comparing with $P\{80 < X < 120\}$,

$$120-10k = 80$$

$$k = 4$$

$$P\{80 < X < 120\} \ge 1-\frac{1}{4^2}$$

$$P\{80 < X < 120\} \ge \frac{15}{16}$$

Hence, the lower bound for probability $= \frac{15}{16}$

Example 6

Two dice are thrown once. If X is the sum of the numbers sharing up, prove that $P\{|X-7| \ge 3\} \le \frac{35}{34}$. Compare this value with the exact probability.

Solution

Let X_1 and X_2 be the random variables which denote the outcomes of first and second dice.

$$E(X_1) = E(X_2) = \frac{1}{6}(1+2+3+4+5+6) = \frac{7}{2}$$

$$E(X) = E(X_1)+E(X_2) = \mu = \frac{7}{2}+\frac{7}{2} = 7$$

$$E(X_1^2) = E(X_2^2) = \frac{1}{6}(1^2+2^2+3^2+4^2+5^2+6^2) = \frac{91}{6}$$

$$\text{Var}(X_1) = \text{Var}(X_2) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

$$\text{Var}(X) = \text{Var}(X_1+X_2) = (1)^2\text{Var}(X_1)+(1)^2\text{Var}(X_2)$$

$$\sigma^2 = \frac{35}{12}+\frac{35}{12} = \frac{35}{6}$$

$$\sigma = \sqrt{\frac{35}{6}}$$

By Chebyshev's inequality,

$$P\{|X-\mu| \ge k\sigma\} \le \frac{1}{k^2}$$

Comparing with $P\{|X-7| \ge 3\}$,

$$\mu = 7$$

$$k\sigma = 3$$

$$k\sqrt{\frac{35}{6}} = 3$$

$$k = 3\sqrt{\frac{6}{35}}$$

$$\therefore \quad P\{|X-7| \ge 3\} \le \frac{1}{\left(3\sqrt{\frac{6}{35}}\right)^2}$$

$$\le \frac{35}{54}$$

Actual probability is given by

$$P\{|X-7| \ge 3\} = P\{X = 1,2,3,4,10,11,12\}$$

$$= \frac{1}{36}+\frac{2}{36}+\frac{3}{36}+\frac{4}{36}+\frac{3}{36}+\frac{2}{36}+\frac{1}{36}$$

$$= \frac{4}{9}$$

Example 7

Use Chebyshev's inequality to find how many times a fair coin must be tossed in order that probability that the ratio of the number of heads

Scanned with CamScanner

to the number of tosses will the between 0.45 and 0.55 will be at least 0.95.

Solution

Let X be the random variable which denotes the number of heads obtained when a fair coin is tossed n times.

$$p = q = \frac{1}{2}$$

X follows a binomial distribution.

$$\text{Mean} = np \quad \text{and} \quad \text{Var}(X) = npq$$

$$\text{Mean of required ratio } \frac{x}{n} = E\left(\frac{1}{n}X\right) = \frac{1}{n}E(X)$$

$$= \frac{1}{n}np = p = \frac{1}{2}$$

$$\therefore \quad \mu = \frac{1}{2}$$

$$\text{Var}\left(\frac{X}{n}\right) = \left(\frac{1}{n}\right)^2 \text{Var}(X) = \frac{1}{n^2}npq = \frac{pq}{n}$$

$$\sigma = \sqrt{\frac{pq}{n}} = \sqrt{\frac{\frac{1}{2}\cdot\frac{1}{2}}{n}} = \frac{1}{2\sqrt{n}}$$

By Chebyshev's inequality,

$$P\left\{\left|\frac{X}{n} - \mu\right| < k\sigma\right\} \ge 1 - \frac{1}{k^2}$$

$$P\left\{-k\sigma < \frac{X}{n} - \mu < k\sigma\right\} \ge 1 - \frac{1}{k^2}$$

$$P\left\{\mu - k\sigma < \frac{X}{n} < \mu + k\sigma\right\} \ge 1 - \frac{1}{k^2}$$

But $P\left\{0.45 < \frac{X}{n} < 0.55\right\} \ge 0.95$

$$1 - \frac{1}{k^2} = 0.95$$

$$\frac{1}{k^2} = 0.05$$

$$k = \sqrt{20}$$

$$\mu - k\sigma = 0.45$$

$$0.5 - \left(\frac{1}{2\sqrt{n}}\right) = 0.45$$

$$n = 2000$$

Hence, the fair coin must be tossed 2000 times.

Example 8

If X is the number on a dice when it is thrown, prove that $P\{|X - \mu| \ge 2.5\} \le 0.47$, where μ is the mean.

Solution

Let x be the random variable which denotes the number on a dice. The probability function is

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

$$E(X) = \mu = \sum xp(x)$$

$$= 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right)$$

$$= \frac{7}{2}$$

$$\text{Var}(X) = \sigma^2 = \sum x^2 p(x) - \mu^2$$

$$= 1\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 9\left(\frac{1}{6}\right) + 16\left(\frac{1}{6}\right) + 25\left(\frac{1}{6}\right) + 36\left(\frac{1}{6}\right) - \left(\frac{7}{2}\right)^2$$

$$= 2.9167$$

$$\sigma = 1.707$$

By Chebyshev's inequality,

$$P\{|X - \mu| > k\sigma\} < \frac{1}{k^2}$$

Comparing with $P\{|X - \mu| > 2.5\}$,

$$k\sigma = 2.5$$

$$k(1.707) = 2.5$$

$$k = 1.46$$

Scanned with CamScanner

$$\therefore P\{|X-\mu|>2.5\}<\frac{1}{(1.46)^2}$$

$$P\{\{X-\mu\}>2.5\}<0.47$$

Example 9

The number of planes landing at an airport in a 30 minutes interval obeys the Poisson law with mean 25. Use Chebyshev's inequality to find the least chance that the number of planes landing within a given 30 minutes interval will be between 15 and 25.

Solution

Let x be a random variable which denotes the number of planes landing at an airport.

For Poisson distribution,

$$E(X)=\mu=25$$

$$\text{Var}(X)=\sigma^2=\mu=25$$

$$\sigma=5$$

By Chebyshev's inequality,

$$P\{|X-\mu|<k\sigma\}\geq 1-\frac{1}{k^2}$$

$$P\{-k\sigma<X-\mu<k\sigma\}\geq 1-\frac{1}{k^2}$$

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

$$P\{25-5k<X<25+5k\}\geq 1-\frac{1}{k^2}$$

Comparing with $P\{15<X<25\}$,

$$25-5k=15 \text{ and } 25+5k=25$$

$$k=2$$

$$\therefore P\{15<X<25\}\geq 1-\frac{1}{(2)^2}$$

$$\geq\frac{3}{4}$$

EXERCISE 3.6

1. A discrete random variable takes the values –1, 0, 1 with probability $\frac{1}{8},\frac{3}{4},\frac{1}{8}$ respectively. Find $P\{|X-\mu|\geq 25\}$.

   $\left[\text{Ans.: }\frac{1}{4}\right]$

2. Use Chebyshev's inequality to prove that $P\{X=\mu\}=1$ if $\text{Var}(X)=0$.

3. If X is a random variable with $E(X)=3$ and $E(X^2)=13$, find the lower bound for $P(-2<X<8)$ using Chebyshev's inequality.

   $\left[\text{Ans.: }\frac{21}{25}\right]$

4. Can we find a random variable for which $P\{\mu-2\sigma<X<\mu+2\sigma\}=0.6$?

   [Ans.: No]

5. If X denotes the sum of the numbers obtained when 2 dice are drawn, obtain an upper bound for $P\{|X-7|\geq 4\}$. Compare with actual probability.

   $\left[\text{Ans.: }\frac{35}{96},\frac{1}{6}\right]$

6. A fair dice is tossed 720 times. Use Chebyshev's inequality to find a lower bound for getting 100 to 140 sixes.

   $\left[\text{Ans.: }\frac{3}{4}\right]$

7. A pair of dice is rolled 900 times and X denotes the number of times a total of 9 occurs. Find $P(80\leq X\leq 120)$ using Chebyshev's inequality.

   $\left[\text{Ans.: }\frac{2}{9}\right]$

8. A discrete random variable X can assume the values x = 1, 2, 3, ... with probability 2^{-x}. Show that $P\{|X-2|\geq 2\}\leq\frac{1}{2}$, while the actual probability is $\frac{1}{8}$.

9. A random variable X has the pmf $P(X=1)=\frac{1}{18}$, $P(X=2)=\frac{16}{18}$, $P(X=3)=\frac{1}{18}$. Show that there is a value of c such that $P\{|X-\mu|\geq c\}=\frac{\sigma^2}{c^2}$,

Scanned with CamScanner

so that, in general, the bound given by Chebyshev's inequality can not be improved.

10. Using Chebyshev's inequality find how many times a fair coin must be tossed in order that the probability of the ratio of number of heads to the number of tosses will lie between 0.4 and 0.6 will be at least 0.9.

[Ans.: 250]

11. Suppose that number of articles produced in a factory during a week is a random variable with mean 500 and variance 100. What can be said about the probability that a week's production will lie between 400 and 600.

[Ans.: At least 0.99]

Scanned with CamScanner