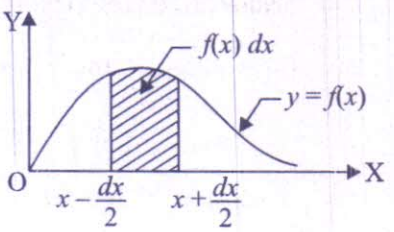
2.4 CONTINUOUS RANDOM VARIABLES

Definition 7 (*Continuous random variable*): A random variable X is said to be continuous if it can assume all possible values between certain limits, *i.e.*, it can take all possible values in a given interval. In other words, a random variable is said to be continuous when its different values cannot be put in one to one correspondence with a set of integers.

Probability density function: Consider the small interval $\left[x-\frac{dx}{2}, x+\frac{dx}{2}\right]$ of x. Let $y = f(x)$ be any continuous function of x such that $f(x)\, dx$ represents the probability that X falls in the infinitesimal interval $\left[x-\frac{dx}{2}, x+\frac{dx}{2}\right]$, symbolically, $P\left(x-\frac{dx}{2}, \le X \le x+\frac{dx}{2}\right) = f(x)\,dx$. Now, $f(x)\,dx$ represents the area bounded by the curve $y = f(x)$, x-axis and the ordinates at the points $x - \frac{dx}{2}$, $x + \frac{dx}{2}$ as shown in the adjacent diagram.



Definition 8 (*Probability density function*): If X is a continuous random variable such that $P\left(x-\frac{dx}{2} \le X \le x+\frac{dx}{2}\right) = f(x)dx$, then $f(x)$ is called the **probability density function (p.d.f.)** of X provided $f(x)$ satisfies the following conditions:

(*i*) $f(x) \geqslant 0, \forall\, x \in R$

(*ii*) $\int_R f(x)\, dx = 1$

where R is the collection of all points in the entire range or spectrum of the random variable X.

(*iii*) $P(E) = \int_E f(x)\, dx$

where E is any well defined event.

Notes: (*i*) The probability for a variate value to lie in the interval of length dx is $f(x)\, dx$ and hence the probability for a variate value to fall in the finite interval $[a, b]$ is

$$P(a \le X \le b) = \int_a^b f(x)\, dx,$$

which represents the area between the curve $y = f(x)$, x-axis and the ordinates at $x = a$ and $x = b$. Further, since total probability is unity, we have $\int_a^b f(x)\, dx = 1$, where $[a, b]$ is the range of the random variable X, *i.e.*, $a \le X \le b$.

(*ii*) $P(X = a) = P(a \le X \le a) = \int_a^a f(x)\, dx = 0$. .

Therefore, probability at a particular point is zero, *i.e.*, it is impossible that a continuous random variable assumes a specific value.

(*iii*) Let $F(x)$ be the distribution function of a continuous random variable X.

Using Property 1 of Art. 2.2 and (*i*), (*ii*), we get

$$P(a < X < b) = P(a \le X < b) = P(a < X \le b) = P(a \le X \le b) = \int_a^b f(x)\, dx = F(b) - F(a).$$

$\Rightarrow$ $f(x)\,dx = dF(x)$. This is known as **probability differential** of X.

$\Rightarrow$ $f(x) = F'(x)$.

(*iv*) **Density curve:** $y = f(x)$ is known as the **probability density curve** or **probability curve.** It gives the graphical representation of the corresponding continuous distribution.

Definition 9 (*Continuous distribution function*): If X is a continuous random variable with probability density function $f(x)$, then the function

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt, \quad -\infty < x < \infty$$

is called the **distribution function (d.f.)** of X.

Notes: (*i*) $0 \le F(x) \le 1, -\infty < x < \infty$.

(*ii*) $F(-\infty) = \lim_{x \to -\infty} \int_{-\infty}^{x} f(t)\,dt = 0$

and $F(\infty) = \lim_{x \to \infty} \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{\infty} f(t)\,dt = 1.$

Example 1: *Find the constant k so that the function*

$$f(x) = \begin{cases} kx^2, & 0 < x < 3 \\ 0, & otherwise \end{cases}$$

is a probability density funcion (p.d.f.). Find the distribution function and evaluate $P(1 < X < 2)$.

Solution: From the property of p.d.f. and given condition, we get

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

$$\Rightarrow \quad \int_0^3 kx^2\,dx = 1$$

$$\Rightarrow \quad k\left[\frac{x^3}{3}\right]_0^3 = 1$$

$$\Rightarrow \quad 9k = 1$$

$$\therefore \quad k = \frac{1}{9}.$$

Also, $f(x) \ge 0, \forall x$ and $k = \frac{1}{9}$. So, $f(x)$ is a possible p.d.f. for $k = \frac{1}{9}$.

The distribution function of X is given by

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt$$

When $x \le 0$:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt = 0 \qquad (\because f(t) = 0, \forall t \le 0)$$

When $0 < x < 3$:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} 0\,dt + \int_0^x \frac{t^2}{9}\,dt$$

$$= \frac{1}{9}\left[\frac{t^3}{3}\right]_0^x = \frac{x^3}{27}.$$

When $x \geq 3$: $\quad F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt = \int_{-\infty}^{0} 0\,dt + \int_{0}^{3} \frac{t^2}{9}\,dt + \int_{3}^{x} 0\,dt = 1$

Hence, the distribution function of X is:

$$F(x) = \begin{cases} 0 & , \quad x \leq 0 \\ \dfrac{x^3}{27} & , \quad 0 < x < 3 \\ 1 & , \quad x \geq 3 \end{cases}$$

Now,
$$P(1 < X < 2) = \int_1^2 f(x)\,dx = \int_1^2 \frac{x^2}{9}\,dx = \frac{1}{9}\left[\frac{x^3}{3}\right]_1^2$$
$$= \frac{1}{27}(2^3 - 1^3) = \frac{7}{27}.$$

Note: $\quad P(1 < X < 2) = F(2) - F(1) = \dfrac{2^3}{27} - \dfrac{1^3}{27} = \dfrac{7}{27}.$

Example 2: *The time one has to wait for a bus at a bus stand is observed to be a random phenomenon governed by the r.v. X with the following p.d.f.:*

$$\begin{aligned} f(x) &= 0 && : \quad x < 0 \\ &= \frac{1}{9}(x+1) && : \quad 0 \leq x < 1 \\ &= \frac{4}{9}\left(x - \frac{1}{2}\right) && : \quad 1 \leq x < \frac{3}{2} \\ &= \frac{4}{9}\left(\frac{5}{2} - x\right) && : \quad \frac{3}{2} \leq x < 2 \\ &= \frac{1}{9}(4 - x) && : \quad 2 \leq x < 3 \\ &= \frac{1}{9} && : \quad 3 \leq x < 6 \\ &= 0 && : \quad x \geq 6. \end{aligned}$$

Let events A and B are defined as follows:

$A \equiv$ One waits between 0 and 2 min. inclusive

$B \equiv$ One waits between 1 and 3 min. inclusive

Show that (i) $P(B/A) = \dfrac{2}{3}$ (ii) $P(\overline{A}\,\overline{B}) = \dfrac{1}{3}$.

Solution: (*i*) By definition, $P(B/A) = \dfrac{P(AB)}{P(A)}$ ...(1)

Now,

$$P(A) = \int_0^2 f(x)\,dx = \int_0^1 \frac{1}{9}(x+1)\,dx + \int_1^{3/2} \frac{4}{9}\left(x - \frac{1}{2}\right)dx + \int_{3/2}^{2} \frac{4}{9}\left(\frac{5}{2} - x\right)dx$$

$$=\frac{1}{9}\left[\frac{x^2}{2}+x\right]_0^1+\frac{4}{9}\left[\frac{x^2}{2}-\frac{x}{2}\right]_1^{3/2}+\frac{4}{9}\left[\frac{5}{2}x-\frac{x^2}{2}\right]_{3/2}^2=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{1}{2}.$$

Again, $P(AB)=P(1\le X\le 2)=\int_1^2 f(x)\,dx=\int_1^{3/2}\frac{4}{9}\left(x-\frac{1}{2}\right)dx+\int_{3/2}^2\frac{4}{9}\left(\frac{5}{2}-x\right)dx$

$$=\frac{1}{6}+\frac{1}{6}=\frac{1}{3}.$$

Therefore, from (1), we get

$$P(B/A)=\frac{1/3}{1/2}=\frac{2}{3}.$$

(*ii*) $\bar{A}\bar{B}\equiv$ Waiting time is more than 3 min.

$$\therefore\quad P(\bar{A}\bar{B})=P(X>3)=\int_3^\infty f(x)\,dx=\int_3^6\frac{1}{9}dx+\int_6^\infty 0\,dx$$

$$=\frac{1}{9}[x]_3^6=\frac{3}{9}=\frac{1}{3}.$$

Example 3: *The probability density function is given by $f(x)=kx^2$, $0\le x\le 6$; $f(x)=k(12-x)^2$, $6\le x\le 12$; $f(x)=0$, elsewhere.*

(*i*) *Evaluate the constant k.* (*ii*) *Find $P(6\le X\le 9)$.* (*W.B.U.T. 2007*)

Solution: (*i*) The given function $f(x)$ is a possible probability density function if $f(x)\ge 0$, $\forall\, x$ and $\int_{-\infty}^{\infty}f(x)dx=1$.

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

$$\Rightarrow\quad \int_{-\infty}^0 0\,dx+\int_0^6 k\,x^2dx+\int_6^{12}k\,(12-x)^2\,dx+\int_{12}^\infty 0\,dx=1$$

$$\Rightarrow\quad k\left[\frac{x^3}{3}\right]_0^6+k\left[-\frac{1}{3}(12-x)^3\right]_6^{12}=1$$

$$\Rightarrow\quad 72k+72k=1$$

$$\Rightarrow\quad k=\frac{1}{144}$$

$$\therefore\quad f(x)=\begin{cases}\dfrac{x^2}{144}, & \text{for } 0\le x\le 6\\[2mm] \dfrac{(12-x)^2}{144}, & \text{for } 6\le x\le 12\\[2mm] 0, & \text{elsewhere}\end{cases}$$

Obviously, $f(x)\geqslant 0$, $\forall x$. So, $f(x)$ is a possible p.d.f. for $k=\dfrac{1}{144}$.

(*ii*) $$P(6 \le X \le 9) = \int_6^9 f(x)\,dx = \int_6^9 \frac{(12-x)^2}{144}\,dx = \frac{1}{144}\left[-\frac{1}{3}(12-x)^3\right]_6^9$$

$$= \frac{1}{144}(-9+72) = \frac{63}{144} = \frac{7}{16}.$$

2.5 MEAN OR EXPECTATION OF A RANDOM VARIABLE

Definition 10 (*Mean or expectation*): (*i*) If X is a discrete random variable which can assume values x_1, $x_2, \ldots x_n, \ldots$ with respective probability $P(X = x_i) = p(x_i) = p_i$; $i = 1, 2, \ldots$, then its **mean** or **expectation** is denoted and defined as

$$m = E(X) = \sum_{i=1}^{\infty} p_i x_i, \text{ such that } \sum_{i=1}^{\infty} p_i = 1.$$

(*ii*) If X is a continuous random variable with probability density function $f(x)$ then its **mean** or **expectation** is denoted and defined as

$$m = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx, \text{ such that } \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

We state below (without proof) some properties of mean.

Properties of mean (or expectation)

Let X, Y, Z are random variables and a, b are some real constants.

1. (*i*) If $X = a$, then $E(X) = a$
   (*ii*) If $a \le X \le b$, then $a \le E(X) \le b$.
2. *Transformation property:*
   (*i*) If $Y = aX$, then $E(Y) = a\,E(X)$.
   (*ii*) If $Y = a \pm bX$, then $E(Y) = a \pm b\,E(X)$.
3. If $Z = aX \pm bY$, then $E(Z) = a\,E(X) \pm b\,E(Y)$.
4. $E(XY) = E(X)\,E(Y)$, provided X and Y are independent.
5. If $g(x)$ is a continuous function, then

$$E\{g(X)\} = \sum p_i\, g(x_i), \text{ for a discrete distribution}$$
$$= \int_{-\infty}^{\infty} g(x) f(x)\,dx, \text{ for a continuous distribution}$$

Note: Mean of a random variable X is a measure of location of the density of X.

(*) Insert E-1 →

Example 1: *Let X be a discrete random variable assuming values 1, 2, 3, ... and suppose that E(X) exists. Show that* $E(X) = \sum_{i=1}^{\infty} P(X \ge i)$.

Solution: Here,
$$\begin{aligned} E(X) &= 1.\,P(X=1) + 2.\,P(X=2) + 3.\,P(X=3) + 4.\,P(X=4) + \ldots \\ &= \{P(X=1) + P(X=2) + P(X=3) + P(X=4) + \ldots\} \\ &\quad + \{P(X=2) + P(X=3) + P(X=4) + \ldots\} \\ &\quad + \{P(X=3) + P(X=4) + \ldots\} + \ldots \\ &= P(X \ge 1) + P(X \ge 2) + P(X \ge 3) + \ldots \\ &= \sum_{i=1}^{\infty} P(X \ge i). \end{aligned}$$

Expectation of a linear combination of random variables

Theorem 1: If $X_1, X_2, \cdots, X_n$ be any n random variables, then $E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)$

where a_i's $(i=1,2,\ldots,n)$ are n constants.

Proof: Let us prove this result by using mathematical induction.

For $n=1$: $E(a_1 X_1) = a_1 E(X_1)$.

Therefore, the result is true for $n=1$. Let us assume that the result is true for $n=k$,

i.e., $Y = a_1 X_1 + a_2 X_2 + \cdots + a_k X_k$

$\Rightarrow E(Y) = a_1 E(X_1) + a_2 E(X_2) + \cdots + a_k E(X_n)$ [by assumption]

Then $E(a_1 X_1 + a_2 X_2 + \cdots + a_k X_k + a_{k+1} X_{k+1})$

$= E(Y + a_{k+1} X_{k+1})$

$= E(Y) + a_{k+1} E(X_{k+1})$

$= a_1 E(X_1) + a_2 E(X_2) + \cdots + a_k E(X_k) + a_{k+1} E(X_{k+1})$

Therefore, the result is true for $n=k+1$ and hence it is true for any positive integer n by mathematical induction.

86

Example 2: *Let X be a random variable with the following probability mass function:*

x	-3	6	9
$p(x)$	$1/6$	$1/2$	$1/3$

Find: (i) E(X) *(ii) $E(X^2)$* *(iii) $E(2X + 1)^2$.*

Solution:

(*i*) $E(X) = \sum xp(x) = (-3)\times\frac{1}{6}+6\times\frac{1}{2}+9\times\frac{1}{3}=\frac{11}{2}.$

(*ii*) $E(X^2) = \sum x^2 p(x) = 9\times\frac{1}{6}+36\times\frac{1}{2}+81\times\frac{1}{3}=\frac{93}{2}$

(*iii*) $E(2X+1)^2 = E(4X^2+4X+1) = 4E(X^2)+4E(X)+1$

$$= 4\times\frac{93}{2}+4\times\frac{11}{2}+1 \qquad \text{[By (i) and (ii)]}$$

$$= 186 + 22 + 1 = 209.$$

Example 3: *A continuous random variable X is distributed over the interval [0, 1] with p.d.f. $f(x) = ax^2 + bx$, where a and b are constants. If the mean of X is 0.25, find the values of a and b.*

Solution: Since $f(x)$ is a p.d.f., we have

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

$$\Rightarrow \quad \int_0^1 (ax^2+bx)\,dx = 1$$

$$\Rightarrow \quad \left[a\frac{x^3}{3}+b\frac{x^2}{2}\right]_0^1 = 1$$

$$\Rightarrow \quad 2a+3b = 6 \qquad \ldots(1)$$

Given, $E(X) = 0.25$

$$\Rightarrow \quad \int_{-\infty}^{\infty} xf(x)\,dx = 0.25$$

$$\Rightarrow \quad \int_0^1 x(ax^2+bx)\,dx = \frac{1}{4}$$

$$\Rightarrow \quad \left[\frac{ax^4}{4}+\frac{bx^3}{3}\right]_0^1 = \frac{1}{4}$$

$$\Rightarrow \quad 3a+4b = 3 \qquad \ldots(2)$$

Solving (1) and (2), we get

$$a = -15,\ b = 12.$$

Example 4: *A random variable X has the density function $f(x) = x,\ 0 \le x \le 1;\ f(x) = \frac{1}{2},\ 1 < x \le 2$. Find the mean of X.* *(W.B.U.T. 2011)*

Solution:

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_0^1 x.xdx + \int_1^2 x.\frac{1}{2}dx$$

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

$$= \frac{1}{3} + \frac{1}{2}\left(2 - \frac{1}{2}\right) = \frac{1}{3} + \frac{3}{4} = \frac{13}{12},$$

this is the required mean of X.

2.6 VARIANCE AND STANDARD DEVIATION

Definition 11 (*Variance*): (*i*) If X is a discrete random variable which can assume values $x_1, x_2, \ldots x_n, \ldots$ with respective probability $P(X = x_i) = p(x_i) = p_i$; $i = 1, 2, \ldots$, and mean m, then its **variance** is denoted and defined as

$$\text{Var}(X) = E\{(X - m)^2\} = E[\{X - E(X)\}^2] = \sum_{i=1}^{\infty}(x_i - m)^2 p_i.$$

(*ii*) If X is a continuous random variable with probability density function $f(x)$ and mean m, then its **variance** is denoted and defined as

$$\text{Var}(X) = E\{(X - m)^2\} = E[\{X - E(X)\}^2] = \int_{-\infty}^{\infty}(x - m)^2 f(x)dx.$$

Definition 12 [*Standard Deviation* (*S.D.*)]: The **standard deviation (S.D.)** of a random variable X is denoted by $\sigma(X)$ or σ_x or simply σ and is defined as the positive square root of Var (X)

$$\therefore \qquad \sigma = +\sqrt{\text{Var}(X)} \quad \Rightarrow \quad \sigma^2 = \text{Var}(X).$$

Note: Variance of a random variable X is a measure of the spread or dispersion of the density of X. Also Var $(X) = 0$ implies that the whole mass is concentrated about the mean.

Properties of variance: *Let X is a random variable and a, b are some real constants.*

(*i*) $Var(X) = E(X^2) - \{E(X)\}^2 = E(X^2) - m^2$, *where* $m = E(X)$.

(*ii*) $Var(aX + b) = a^2\ Var(X)$

(*iii*) $Var(a) = 0$

(*iv*) $Var(X) = E\{X(X - 1)\} - m(m - 1)$, *where* $m = E(X)$.

Proof:

(*i*)

$$\begin{aligned}\text{L.H.S.} &= \text{Var}(X) = E\{(X - m)^2\} = E(X^2 - 2mX + m^2)\\ &= E(X^2) - E(2mX) + E(m^2)\\ &= E(X^2) - 2m\,E(X) + m^2\\ &= E(X^2) - 2m \,.\, m + m^2 = E(X^2) - m^2 = \text{R.H.S.}\end{aligned}$$

88

(ii) Now, $E(aX + b) = a\,E(X) + b$ [see Property 2 (ii), Art. 2.5]

$\therefore$ L.H.S. $= \text{Var}\,(aX + b) = E[\{aX + b - E\,(aX + b)\}^2]$

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

$= a^2\,\text{Var}\,(X) =$ R.H.S.

(iii) L.H.S. $= \text{Var}\,(a) = E[\{a - E(a)\}^2] = E[(a - a)^2] = 0 =$ R.H.S.

(iv) L.H.S. $= \text{Var}\,(X) = E\{(X - m)^2\} = E(X^2 - 2mX + m^2)$

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

$= E(X^2 - X) + E(X) - E(2mX) + E\,(m^2)$

$= E\{X(X - 1)\} + m - 2mE(X) + m^2$

$= E\{X(X - 1)\} + m - 2m\,.\,m + m^2$

$= E\{X(X - 1)\} - m(m - 1)$

$=$ R.H.S.

Definition 13 [*Coefficient of Variation (C.V.)*]: The ratio of the standard deviation to the mean is known as the **coefficient of variation (C.V.).** Note that this measure is **unit free** and is often expressed as a percentage.

$$\therefore \text{ Coefficient of variation } = \frac{\text{S.D.}}{\text{Mean}} \times 100.$$

(*) Insert E-2 →

Example 1: *Given the following probability distribution of X:*

x	-3	-2	-1	0	1	2	3
$p(x)$	0.05	0.10	0.30	0	0.35	0.10	0.10

Compute (i) $E(X)$ (ii) $E(2X \pm 3)$ (iii) $E(X^2)$ (iv) Var (X) (v) Var $(2X \pm 1)$.

Solution: Observe that $\sum p(x) = 1$ and $p(x) \geqslant 0, \forall x$.

(i) $E(X) = \sum xp(x) = (-3)\,(0.05) + (-2)\,(0.10) + (-1)\,(0.30) + 0$
$+ 1\,(0.35) + 2(0.10) + 3\,(0.10) = 0.20$

(ii) $E(2X \pm 3) = 2\,E(X) \pm 3 = 2(0.20) \pm 3 = 0.40 \pm 3$
$= 3.40, -2.60$

(iii) $E(X^2) = \sum x^2p(x) = (-3)^2\,(0.05) + (-2)^2\,(0.10)$
$+ (-1)^2\,(0.30) + 0 + 1^2\,(0.35) + 2^2\,(0.10) + 3^2\,(0.10)$
$= 0.45 + 0.4 + 0.3 + 0.35 + 0.4 + 0.9$
$= 2.8$

(iv) $\text{Var}\,(X) = E(X^2) - \{E(X)\}^2 = 2.8 - (0.20)^2 = 2.76.$

(v) $\text{Var}\,(2X \pm 1) = 2^2\,\text{Var}\,(X)$ [$\because$ Var $(aX \pm b) = a^2$ Var (X)]
$= 4 \times 2.76 = 11.04.$

Example 2: *A continuous random variable X has the following p.d.f.:*

$$f(x) = \begin{cases} \frac{1}{4}(x+2) & , \quad -1 < x < 1 \\ 0 & , \quad \text{elsewhere} \end{cases}$$

Find the mean and variance of X

Variance of a linear combination of random variables

If $X_1, X_2, \ldots, X_n$ be any n random variables, then $Var\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j Cov(X_i, X_j)$

Proof: Let $U = a_1 X_1 + a_2 X_2 + \cdots + a_n X_K$.

$\therefore E(U) = a_1 E(X_1) + a_2 E(X_2) + \cdots + a_n E(X_n)$

$\Rightarrow U - E(U) = a_1\{X_1 - E(X_1)\} + a_2\{X_2 - E(X_2)\} + \cdots + a_n\{X_n - E(X_n)\}$

$$\Rightarrow E\{U - E(U)\}^2 = a_1^2\{X_1 - E(X_1)\}^2 + a_2^2\{X_2 - E(X_2)\}^2 + \cdots + a_n^2\{X_n - E(X_n)\}^2 + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j E[\{X_i - E(X_i)\}\{X_j - E(X_j)\}]$$

$$\Rightarrow Var(U) = a_1^2 Var(X_1) + a_2^2 Var(X_2) + \cdots + a_n^2 Var(X_n) + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j Cov(X_i, X_j)$$

$$\Rightarrow Var\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var(X_i) + 2\sum_{i=1}^{n}\sum_{\substack{j=1 \\ i<j}}^{n} a_i a_j Cov(X_i, X_j).$$

Note: If $X_1, X_2, \ldots, X_n$ are mutually independent, then

$$\begin{aligned} Cov(X_i, X_j) &= E[\{X_i - E(X_i)\}\{X_j - E(X_j)\}] \\ &= E\{X_i - E(X_i)\} E\{X_j - E(X_j)\} \\ &= \{E(X_i) - E(X_i)\}\{E(X_j) - E(X_j)\} \\ &= 0. \end{aligned}$$

Therefore, in this situation:

$$Var\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var(X_i).$$

Solution: Observe that $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-1}^{1}\frac{1}{4}(x+2)\,dx = \left[\frac{1}{4}\left(\frac{x^2}{2}+2x\right)\right]_{-1}^{1} = 1.$

Also, $f(x) \geq 0, \quad \forall\, x.$

So, $f(x)$ is a possible probability density function.

Now,
$$\text{mean} = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-1}^{1} x\cdot\frac{1}{4}(x+2)\,dx$$
$$= \frac{1}{4}\left[\frac{x^3}{3}+2\frac{x^2}{2}\right]_{-1}^{1} = \frac{1}{4}\left(\frac{1}{3}+\frac{1}{3}\right) = \frac{1}{6}.$$
$$\text{Var}(X) = \int_{-\infty}^{\infty}\{x - E(X)\}^2 f(x)\,dx$$
$$= \int_{-1}^{1}\left(x-\frac{1}{6}\right)^2 \frac{1}{4}(x+2)\,dx$$
$$= \frac{1}{4}\int_{-1}^{1}\left(x^2-\frac{x}{3}+\frac{1}{36}\right)(x+2)\,dx$$
$$= \frac{1}{4}\int_{-1}^{1}\left(x^3+\frac{5x^2}{3}-\frac{23}{36}x+\frac{1}{18}\right)dx$$
$$= \frac{1}{4}\left[\frac{x^4}{4}+\frac{5x^3}{9}-\frac{23}{72}x^2+\frac{x}{18}\right]_{-1}^{1}$$
$$= \frac{1}{4}\left(\frac{10}{9}+\frac{2}{18}\right) = \frac{11}{36}.$$

Note:
$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) = \int_{-1}^{1} x^2\cdot\frac{1}{4}(x+2)\,dx = \frac{1}{4}\left[\frac{x^4}{4}+\frac{2}{3}x^3\right]_{-1}^{1}.$$
$$= \frac{1}{4}\left(\frac{2}{3}+\frac{2}{3}\right) = \frac{1}{3}.$$

$\therefore$
$$\text{Var}(X) = E(X^2) - \{E(X)\}^2 = \frac{1}{3}-\left(\frac{1}{6}\right)^2 = \frac{11}{36}.$$

*Insert E-3 – E-5 →

MISCELLANEOUS EXAMPLES

Example 1: *If*
$$F(x) = 0, \quad -\infty < x < 0$$
$$= \frac{1}{5}, \quad 0 \leq x < 1$$
$$= \frac{3}{5}, \quad 1 \leq x < 3$$
$$= 1, \quad 3 \leq x < \infty,$$

2.7 CORRELATION COEFFICIENT

1. If X and Y are two random variables, then covariance between them is denoted and defined as

$$\begin{aligned} \text{Cov}(X,Y) &= E[\{X-E(X)\}\{Y-E(Y)\}] \\ &= E\{XY - YE(X) - XE(Y) + E(X)E(Y)\} \\ &= E(XY) - E(Y)E(X) - E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

In particular if X and Y are two independent random variables, then $E(XY) = E(X)E(Y)$ so that $\text{Cov}(X,Y) = 0$.

2. Correlation coefficient between two random variables X and Y usually denoted by $\rho(X,Y)$ [or, $r(X,Y)$] is a numerical measure of linear relationship between them and is defined as

$$\rho(X,Y) = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}, \text{ where } \sigma_X = +\sqrt{\text{Var}(X)} \text{ and } \sigma_Y = +\sqrt{\text{Var}(Y)}.$$

3. $-1 \le \rho(X,Y) \le 1$.

Proof: Let $X^* = \dfrac{X-E(X)}{\sigma_X}$ and $Y^* = \dfrac{Y-E(Y)}{\sigma_Y}$.

Then $E(X^{*2}) = E(Y^{*2}) = 1$ and $\rho(X,Y) = E(X^* Y^*)$.

Now, $0 \le (X^* \pm Y^*)^2 = X^{*2} + Y^{*2} \pm 2X^*Y^*$.

Taking expectations, we get

$$0 \le 2\{1 \pm \rho(X,Y)\} \Leftrightarrow -1 \le \rho(X,Y) \le 1.$$

2.8 MOMENTS

Let X be a random variable and 'a' be a given real number.

$\alpha_K = K^{th}$ order moment about 'a' where K is a positive integer

$= E\{(X-a)^K\}$, if it exists.

In particular: When $a = 0$.

$\alpha_K = K^{th}$ order moment about origin

$= E(X^K)$, where K is a positive integer.

For $K=1$: $\alpha_1 = E(X) = \text{Mean} = m$ (say).

Define: $\mu_K = K^{th}$ order moment about mean $(= m)$

$= E\{(X-m)^K\}$

$= K^{th}$ central moment

$\therefore \mu_1 = E(X-m) = E(X) - m = m - m = 0 =$ first central moment

$\mu_2 = E\{(X-m)^2\} = Var(X) =$ second central moment

Relation between α_K and μ_K

Now, $(X-m)^K = \sum_{r=0}^{K} (-1)^r \, {}^K C_r \, X^{K-r} m^r$

$$\Rightarrow E\{(X-m)^K\} = \sum_{r=0}^{K} (-1)^r \, {}^K C_r \, E(X^{K-r}) m^r$$

$$\Rightarrow \boxed{\mu_K = \sum_{r=0}^{K} (-1)^r \, {}^K C_r \, \alpha_{K-r} m^r.}$$

$\alpha_0 = E(X^0) = 1, \quad \alpha_1 = E(X) = m, \quad \alpha_K = E(X^k).$

K=2: $\mu_2 = \sum_{r=0}^{2} (-1)^r \left({}^2C_r\right) \alpha_{2-r}\, m^r$

$= \alpha_2 - 2\alpha_1 m + \alpha_0 m^2$

$= \alpha_2 - 2m^2 + m^2 = \alpha_2 - m^2.$

K=3: $\mu_3 = \sum_{r=0}^{3} (-1)^r \left({}^3C_r\right) \alpha_{3-r}\, m^r$

$= \alpha_3 - 3\alpha_2 m + 3\alpha_1 m^2 - \alpha_0 m^3$

$= \alpha_3 - 3\alpha_2 m + 2m^3.$

Similarly, for K=4: $\mu_4 = \alpha_4 - 4\alpha_3 m + 6\alpha_2 m^2 - 3m^4.$

Conversely:

$\alpha_2 = \mu_2 + m^2$

$\alpha_3 = \mu_3 + 3\mu_2 m + m^3$

$\alpha_4 = \mu_4 + 4\mu_3 m + 6\mu_2 m^2 + m^4$