2.3

Mathematical Expectation

The term expectation is used for the process of averaging when a random variable is involved. It is the number used to locate the centre of the probability distribution (p.m.f or p.d.f) of a random variable. A probability distribution is described by certain satisfied measures which are computed using mathematical expectation (or expectation)

Let X be a random variable defined on a sample space S. Let $g(.)$ be a function of X such that $g(X)$ is a random variable. Then the **expected value of $g(X)$** is defined by

$$E(g(X)) = \begin{cases} \sum\limits_{x} g(x)\, p(x) & \text{if } X \text{ is a d.r.v with p.m.f. } p(x) \\ \int\limits_{-\infty}^{\infty} g(x)\, f(x)dx & \text{if } X \text{ is a c.r.v with p.d.f. } f(x) \end{cases} \quad \text{--------(1)}$$

provided these values exist.

Mean and moments:

i. Let $g(X) = X$. Then, by formula (1), **expected value of X** is defined by

$$E(X) = \mu = \begin{cases} \sum\limits_{x} x\, p(x) & \text{if } X \text{ is a d.r.v with p.m.f. } p(x) \\ \int\limits_{-\infty}^{\infty} x\, f(x)dx & \text{if } X \text{ is a c.r.v with p.d.f. } f(x) \end{cases}$$

Then $E(X)$ is called the **mean of the random variable X** and it is denoted by μ.

ii. Let $g(X) = (X - A)^r$ where A is an arbitrary constant and r is a non negative integer. Then the formula (1) gives

$$E(X-A)^r = \mu_r' = \begin{cases} \sum\limits_{x} (x-A)^r\, p(x) & \text{if } X \text{ is a d.r.v with p.m.f. } p(x) \\ \int\limits_{-\infty}^{\infty} (x-A)^r\, f(x)dx & \text{if } X \text{ is a c.r.v with p.d.f. } f(x) \end{cases}$$

The quantity $E(X-A)^r$ is called the $\boldsymbol{r^{th}}$ **moment about $\boldsymbol{A}$** and it is denoted by μ'_r. If $A = 0$, then μ'_r are known as **Raw Moments**.

iii. Let $g(X) = \left(X - E(X)\right)^r = (X-\mu)^r$. Then the formula (1) gives

$$E(X-\mu)^r = \mu_r = \begin{cases} \sum_X (x-\mu)^r \, p(x) & if\ X\ is\, a\, d.r.v\, with\ p.m.f.\ p(x) \\ \int_{-\infty}^{\infty} (x-\mu)^r \, f(x)dx & if\ X\ is\, a\, c.r.v\, with\ p.d.f.\ f(x) \end{cases}$$

The function $E(X-\mu)^r$ is called the $\boldsymbol{r^{th}}$ **central moment of $\boldsymbol{X}$** and it is denoted by $\boldsymbol{\mu_r}$

iv. If $r = 2$, then $\mu_2 = \sigma^2 = E(X-\mu)^2$ and it is known as the **variance of the random variable $\boldsymbol{X}$** and it is denoted by $\boldsymbol{V(X)}$ or $\boldsymbol{\sigma^2}$.

v. Mean (μ) and variance (σ^2) are important statistical measures of a probability distribution.

Example 1: Let $\boldsymbol{X}$ be a d.r.v with the p.m.f. given below:

$\boldsymbol{x}$	$\boldsymbol{-3}$	$\boldsymbol{6}$	$\boldsymbol{9}$
$\boldsymbol{p(x)}$	$\boldsymbol{\frac{1}{6}}$	$\boldsymbol{\frac{1}{2}}$	$\boldsymbol{\frac{1}{3}}$

Find $\boldsymbol{E(X)}$and $\boldsymbol{E(X^2)}$.

Solution:

$$E(X) = \sum_x x\,p(x) - 3 \times \frac{1}{6} + 6 \times \frac{1}{2} + 9 \times \frac{1}{3} = -\frac{1}{2} + 3 + 3 = \frac{11}{2}$$

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

Example 2: Find the expectation of the number on a die when thrown.

Solution: Let X be the random variable representing the number on a die when thrown. Then X can take any one of the values 1,2,3,4,5,6 each with equal probability $\frac{1}{6}$. Hence

$$E(X) = \sum_x x\, p(x) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}$$

$$= \frac{1}{6}(1+2+3+4+5+6) = \frac{1}{6}.\frac{6.7}{2} = \frac{7}{2}$$

$$\Longrightarrow \; E(X) = \frac{7}{2}$$

Example 3: Two unbiased dice are thrown. Find the expected values of the sum of numbers of points on them.

Solution: Define X is the sum of the numbers obtained on the two dice and $X = 2,3,4, \dots, 12$ and its probability distribution is given by

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

$$E(X) = \sum x\, p(x) = 2 \times \frac{1}{36} + 3 \times \frac{2}{36} + 4 \times \frac{3}{36} + 5 \times \frac{4}{36} + 6 \times \frac{5}{36} +$$

$$7 \times \frac{6}{36} + 8 \times \frac{5}{36} + 9 \times \frac{4}{36} + 10 \times \frac{3}{36} + 11 \times \frac{2}{36} + 12 \times \frac{1}{36}$$

$$= \frac{1}{36}(2+6+12+20+30+42+40+36+30+22+12) = \frac{252}{36} = 7$$

$$\Longrightarrow \; E(X) = 7$$

Example 4: In four tosses of a coin, let X be the number of heads. Find the mean and variance of X.

Solution: The sample space S consists of $2^4 = 16$ outcomes and the following table gives the outcomes and the value of X for each outcome is

S.No	Out come	X
1	$TTTT$	0
2	$TTTH$	1
3	$TTHT$	1
4	$TTHH$	2
5	$THTT$	1
6	$THTH$	2
7	$THHT$	2
8	$THHH$	3
9	$HTTT$	1
10	$HTTH$	2
11	$HTHT$	2
12	$HTHH$	3
13	$HHTT$	2
14	$HHTH$	3
15	$HHHT$	3
16	$HHHH$	4

$$p(0) = \frac{1}{16}, p(1) = \frac{4}{16}, p(2) = \frac{6}{16}, p(3) = \frac{4}{16}, p(4) = \frac{1}{16}$$

The p.m.f of X is given in the following table:

x	0	1	2	3	4
$p(x)$	$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$

$$\mathrm{E(X)} = \sum_x x\, p(x) = 0 \times \frac{1}{16} + 1 \times \frac{4}{16} + 2 \times \frac{6}{16} + 3 \times \frac{4}{16} + 4 \times \frac{1}{6}$$

$$= \frac{1}{16}(0 + 4 + 12 + 12 + 4) = \frac{32}{16} = 2$$

$$\Rightarrow \mathrm{E(X)} = 2$$

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

$$= (0-2)^2 \times \frac{1}{16} + (1-2)^2 \times \frac{4}{16} + (2-2)^2 \times \frac{6}{16} + (3-2)^2 \times \frac{4}{16} + (4-2)^2 \times \frac{1}{16}$$

$$= 4 \times \frac{1}{16} + 1 \times \frac{4}{16} + 0 \times \frac{6}{16} + 1 \times \frac{4}{16} + 4 \times \frac{1}{16} = \frac{1}{16}(4+4+4+4) = \frac{16}{16} = 1$$

$\Longrightarrow \mathrm{V(X)} = 1$

Example 5: Find the mean and variance of the random variable X, whose p.d.f is given by

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

Solution:

$$E(X) = \int_0^2 x.f(x)dx = \frac{1}{2}\int_0^2 x.dx = \frac{1}{2}\left[\frac{x^2}{2}\right]_0^2 = 1 - 0 = 1$$

$\Longrightarrow$ Mean of the random variable X is 1.

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

$$\int_0^2 (x-1)^2 .f(x)dx = \frac{1}{2}\int_0^2 (x-1)^2 .dx = \frac{1}{2}\left[\frac{(x-1)^3}{3}\right]_0^2 = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$$

$\Longrightarrow$ Variance of the random variable X is $\frac{1}{3}$

Example 6: Find the mean of the random variable X whose p.d.f. is given by

$$f(x) = \begin{cases} e^{-x}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

Solution:

$$E(X) = \int_0^\infty x\, f(x) dx = \frac{1}{2}\int_0^2 x e^{-x} dx = \left[-xe^{-x}\right]_0^\infty + \int_0^\infty e^{-x} dx = 0 + \left[-e^{-x}\right]_0^\infty = 0+1=1$$

$\Longrightarrow E(X) = 1$

Theorems on Mathematical Expectation:

The following theorems are proved by assuming that the random variables are continuous. If the random variables are discrete, the proof remains the same except replacing integration by summation.

Theorem 1: If X is a random variable and a and b are constants then

$$\boldsymbol{E(aX + b) = aE(X) + b}.$$

Proof: Let X be a c.r.v with p.d.f. $f(x)$. Then

$E(aX + b) =$

$$\int_{-\infty}^{\infty} (ax+b)\, f(x).dx = a\int_{-\infty}^{\infty} x\, f(x)\, dx + b\int_{-\infty}^{\infty} f(x).dx = a\,E(X) + b \qquad \left(\because \int_{-\infty}^{\infty} f(x).dx = 1\right)$$

Corollary 1: If $b = 0$, then $E(aX) = aE(X)$

Corollary 2: If $X = 1$ and $b = 0$, then $E(a) = a$

Theorem 2: Addition Theorem of mathematical expectation.

If X and Y are random variables, then $E(X + Y) = E(X) + E(Y)$ provided all the expectations exist.

Proof: Let X and Y be continuous random variables with j.p.d.f. $f(x, y)$ and m.p.d.fs be $f_1(x)$ and $f_2(y)$ respectively. Then by definition,

$E(X) = \int_{-\infty}^{\infty} x\, f_1(x).dx$ and $E(Y) = \int_{-\infty}^{\infty} y.f_2(y).dy$

Now, $E(X + Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\, f(x,y).dx.dy$

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

$$= \int_{-\infty}^{\infty} x \left[\int_{-\infty}^{\infty} f(x,y)\, dy \right] dx + \int_{-\infty}^{\infty} y \left[\int_{-\infty}^{\infty} f(x,y)\, dx \right] dy$$

$$= \int_{-\infty}^{\infty} x\, f_1(x)\, dx + \int_{-\infty}^{\infty} y\, f_2(y)\, dy = E(X) + E(Y)$$

$$\therefore E(X+Y) = E(X) + E(y)$$

Generalization: If $X_1, X_2, X_3 \dots X_n$ are random variables, then $E(X_1 + X_2 + X_3 + \cdots + X_n) = E(X_1) + E(X_2) + E(X_3) \dots + E(X_n)$ provided all the expectations exist.

Theorem 3: Multiplication Theorem of mathematical Expectations

If X and Y are independent random variables, then $E(X\,Y) = E(X)\, E(Y)$.

Proof: Let X and Y be continuous random variables with j.p.d.f. $f(x, y)$ and m.p.d.fs be $f_1(x)$ and $f_2(y)$ respectively. Then by definition,

$$E(X) = \int_{-\infty}^{\infty} x\, f_1(x)\, dx \text{ and } E(Y) = \int_{-\infty}^{\infty} y\, f_2(y)\, dy$$

Now, $E(X.Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (xy)\, f(x,y)\, dx\, dy$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x,y)\, f_1(x)\, f_2(y)\, dx\, dy \qquad (\because X \text{ and } Y \text{ are independent})$$

$$= \left(\int_{-\infty}^{\infty} x\, f_1(x)\, dx \right) \left(\int_{-\infty}^{\infty} y\, f_2(y)\, dy \right) = E(X)\, E(Y)$$

Generalization: If $X_1, X_2, X_3 \dots X_n$ are independent random variables, then $E(X_1 X_2 X_3 \dots . X_n) = E(X_1) E(X_2) E(X_3) \dots E(X_n)$.

Theorem 4: Mathematical expectation of a linear combination of random variables.

Let $\boldsymbol{X_1, X_2, X_3 \dots X_n}$ be any $\boldsymbol{n}$ random variables and $\boldsymbol{a_1\ , a_2, a_3, \dots . a_n}$ be any $\boldsymbol{n}$ constants. Then

$$\boldsymbol{E(a_1X_1 + a_2X_2 + a_3X_3 + \cdots + a_nX_n) = a_1E(X_1) + a_2E(X_2) + a_3E(X_3) + \cdots + a_nE(X_n)}$$

provided all the expectations exist.

The proof follows using Theorem 1 and generalization of Theorem 2.

Theorem 5: $\boldsymbol{V(X) = E(X^2) - \left(E(X)\right)^2}$

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

$$= E\left[X^2 - 2XE(X) + \left(E(X)\right)^2\right]$$

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

$= \ E(X^2) - 2E(X)E(X) + \left(E(X)\right)^2 \ \because E(X)$ is a constant and $E\big(E(X)\big) = E(X)$

$\Longrightarrow V(X) = E(X^2) - 2\left(E(X)\right)^2 + \left(E(X)\right)^2$

$\Longrightarrow V(X) = E(X^2) - \left(E(X)\right)^2$

Note: The formula is simple to use instead of $\boldsymbol{E\left(X - E(X)\right)^2}$.

Theorem 6: If $\boldsymbol{X}$ is a random variable, and $\boldsymbol{a}$ and $\boldsymbol{b}$ are constants, then $\boldsymbol{V(ax + b) = a^2\ V(X)}$.

Proof: Let $Y = aX + b$. Then $E(Y) = E(aX + b) = aE(X) + b$ and

$Y - E(Y) = a(X - E(X))$

$\Longrightarrow E\left(Y - E(Y)\right)^2 = a^2E\left(X - E(X)\right)^2$

$\Longrightarrow V(Y) = a^2V(X) \Longrightarrow V(aX + b) = a^2V(X)$

Corollary 1: If $a = 0$, then $V(b) = 0$ $i.e.$, variance of a constant is zero.

Corollary 2: If $b = 0$, then $V(aX) = a^2V(X)$

Covariance: If X and Y are two random variables, then the **covariance** between them is defined by

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))]$$

$$= E[XY - XE(Y) - YE(X) + E(X)E(Y)]$$

$$= E(XY) - E\left(\left(XE(Y)\right) - E(YE(X)\right) + E(E(X)E(Y))$$

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

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

Note:

1. If X and Y are independent , then $\text{Cov}(X, Y) = 0$
2. $\text{Cov}(aX, bY) = ab\text{Cov}(X, Y)$ where a and b are constants.
3. $\text{Cov}(X + a, Y + b) = \text{Cov}(X, Y)$.

Theorem 7: Variance of a linear combination of random variables.

Let $X_1, X_2, X_3, \dots, X_n$ be any n random variables and $a_1, a_2, a_3, \dots, a_n$ are n constants, then

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 V(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov(X_i, X_j)$$

Proof:

Let $U = \sum_{i=1}^{n} a_i X_i$, then $E(U) = \sum_{i=1}^{n} a_i E(X_i)$ and $U - E(U) = \sum_{i=1}^{n} a_i \left(X - E(X_i)\right)$

$$\Rightarrow \left(U - E(U)\right)^2 =$$

$$\sum_{i=1}^{n} a_i^2 \left(X_i - E(X_i)\right)^2 + 2\sum_{\substack{i=1 \\ i<j}}^{n}\sum_{j=1}^{n} a_i a_j \left(X_i - E(X_i)\right)\left(X_j - E(X_j)\right)$$

$$\Rightarrow E\left(U - E(U)\right)^2 =$$

$$\sum_{i=1}^{n} a_i^2 E\left(X_i - E(X_i)\right)^2 + 2\sum_{\substack{i=1 \\ i<j}}^{n}\sum_{j=1}^{n} a_i a_j E\left[\left(X_i - E(X_i)\right)\left(X_j - E(X_j)\right)\right]$$

$$\Rightarrow V\left(\sum_{i=1}^{n} a_i X_i\right) = V(U) = \sum_{i=1}^{n} a_i^2 V(X_i) + 2\sum_{\substack{i=1 \\ i<j}}^{n}\sum_{j=1}^{n} a_i a_j \operatorname{cov}\left(X_i, X_j\right)$$

Note:

1. **If $X_1, X_2, X_3, \dots, X_n$ are independent , then** $V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 V(X_i)$
2. **If $a_1 = a_2 = 1$ and $a_3 = \cdots . = a_n = 0$, then**
   $\boldsymbol{V(X_1 + X_2) = V(X_1) + V(X_2) + 2Cov(X_1, X_2)}$

3. **If $a_1 = 1, a_2 = -1$ and $a_3 = \cdots . = a_n = 0$, then**
   $\boldsymbol{V(X_1 - X_2) = V(X_1) + V(X_2) - 2Cov(X_1, X_2)}$

4. **If X_1 and X_2 are independent , then $V(X_1 \pm X_2) = V(X_1) + V(X_2)$**

Example 7: The j.p.d.f. of X and Y is given by

$$f(x,y) = \begin{cases} 2 - x - y & ,0 < x < 1, 0 < y < 1 \\ 0 & , \quad otherwise \end{cases}$$

Find

i. **m.p.d.fs of X and Y**
ii. **c.p.d.fs of X and Y**
iii. **$V(X)$ and $V(Y)$**

iv. **Covariance between X and Y**

Solutions:

i. $f_1(x) = \int_0^1 f(x,y)dy = \int_0^1 (2-x-y)dy = \left[2y - xy - \frac{y^2}{2}\right]_0^1 = 2 - x - \frac{1}{2} = \frac{3}{2} - x$

$$f_1(x) = \begin{cases} \frac{3}{2} - x & ,0 < x < 1 \\ 0 & ,otherwise \end{cases}$$

Similarly $f_2(y) = \begin{cases} \frac{3}{2} - y & ,0 < y < 1 \\ 0 & ,otherwise \end{cases}$

ii. $f_{1|2}(x|y) = \frac{f(x,y)}{f_2(y)} = \frac{2-x-y}{\frac{3}{2}-y}, 0 < x, y < 1$

and $f_{2|1}(y|x) = \frac{f(x,y)}{f_1(x)} = \frac{2-x-y}{\frac{3}{2}-x},\ 0 < x, y < 1$

iii. $E(X) = \int_0^1 x f_1(x)dx = \int_0^1 x\left(\frac{3}{2} - x\right)dx = \frac{5}{12}$ and

$E(X^2) = \int_0^1 x^2 f_1(x)dx = \int_0^1 x^2\left(\frac{3}{2} - x\right)dx = \frac{1}{4}$

Thus $V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{1}{4} - \frac{25}{144} = \frac{11}{144}$

Similarly $V(Y) = \frac{11}{144}$

iv. $E(XY) = \int_0^1\int_0^1 xy f(x,y)dxdy = \int_0^1\int_0^1 xy(2-x-y)dxdy = \frac{1}{6}$ (verify!)

$$\therefore Cov(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{6} - \frac{5}{12}.\frac{5}{12} = -\frac{1}{144}$$