CHAPTER SIX

Mathematical Expectation

LEARNING OBJECTIVES. Upon completion of this chapter, you should be able to :

1. Understand and interpret the expected value and variance of a random variable.
2. State with proof the various properties of expectation and variance.
3. Explain the concept of covariance.
4. Obtain moments of bivariate probability distributions.
5. Discuss some inequality relationships involving expectation, like Cauchy-Schwartz inequality, Jenson's inequality, etc.
6. Demonstrate the concept of conditional expectation and conditional variance.

CHAPTER OUTLINE

6·1. INTRODUCTION

Many frequently used *r.v.*'s can be both characterized and dealt with effectively for practical purposes by consideration of quantities called their *expectation*. For example, a gambler might be interested in his average winnings at a game, a businessman in his average profits on a product, a physicist in the average charge of a particle, and so on. The *'average'* value of a random phenomenon is also termed as its *mathematical expectation* or *expected value*. In this chapter we will define and study this concept in detail, which will be used extensively in subsequent chapters.

6·2. MATHEMATICAL EXPECTATION OR EXPECTED VALUE OF A RANDOM VARIABLE.

Once we have constructed the probability distribution for a random variable, we often want to compute the mean or expected value of the random variable. The *expected value* of a discrete random variable is a weighted average of all possible values of the random variable, where the weights are the probabilities associated with the corresponding values. The mathematical expression for computing the expected value of a discrete random variable X with probability mass function (*p.m.f.*) $f(x)$ is given below :

$$\boxed{E(X) = \sum_x x f(x), \text{ (for discrete r.v.)}} \quad \text{... (6·1)}$$

The mathematical expression for computing the expected value of a continuous random variable X with probability density function (*p.d.f.*) $f(x)$ is, however, as follows :

$$\boxed{E(X) = \int_{-\infty}^{\infty} x f(x)\, dx, \text{ (for continuous r.v.)}} \quad \text{... (6·1a)}$$

provided the right, hand integral in (6·1*a*) or series in (6·1) is absolutely convergent, *i.e.*, provided

$$\int_{-\infty}^{\infty} | x f(x) |\, dx = \int_{-\infty}^{\infty} | x |\, f(x)\, dx < \infty \quad \text{... (6·2)}$$

or

$$\sum_x | x f(x) | = \sum_x | x |\, f(x) < \infty \quad \text{... (6·2a)}$$

Remarks 1. Since absolute convergence implies ordinary convergence, if (6·2) or (6·2 *a*) holds then the series or integral in (6·1) and (6·1*a*) also exists, *i.e.*, has a finite value and in that case we define $E(X)$ by (6·1) or (6·1*a*). It should be clearly understood that although X has an expectation only if L.H.S. in (6·2) or (6·2*a*) exists, *i.e.*, converges to a finite limit, its value is given by (6·1) or (6·1*a*).

2. $E(X)$ exists if $E|X|$ exists.

3. *Expected value and variance of an Indicator Variable.* Consider the indicator variable : $X = I_A$, so that

$$X = I_A = \begin{cases} 1, \text{ if } A \text{ happens} \\ 0, \text{ if } \bar{A} \text{ happens} \end{cases}$$

Now $E(X) = 1 . P(X = 1) + 0 . P(X = 0) \Rightarrow E(I_A) = 1 . P(I_A = 1) + 0 . P(I_A = 0)$

$\therefore$

$$E(I_A) = P(A)$$

This gives us a very useful tool to find $P(A)$, rather than to evaluate $E(X)$.

Thus

$$P(A) = E(I_A) \quad \text{... (6·2b)}$$

$$E(X^2) = 1^2 \cdot P(X=1) + 0^2 \cdot P(X=0) = P(I_A = 1) = P(A)$$

$\therefore$ $$\text{Var } X = E(X^2) - [E(X)]^2 = P(A) - [P(A)]^2 = P(A)[1 - P(A)] = P(A)\,P(\bar{A}) \quad \text{... (6·2c)}$$

$\Rightarrow$ $$\text{Var}(I_A) = P(A)\,P(\bar{A}) \quad \text{...(6·2d)}$$

4. If the $r.v.$ X takes the values $0\,!, 1\,!, 2\,!, \ldots$ with probability law:

$$P(X = x\,!) = \frac{e^{-1}}{x\,!}\,;\, x = 0, 1, 2, \ldots, \text{ then } \sum_{x=0}^{\infty} x\,!\, P(X = x\,!) = e^{-1} \sum_{x=0}^{\infty} 1,$$

which is a divergent series. In this case $E(X)$ does not exist.

More rigorously, let us consider a random variable X which takes the values

$$x_i = (-1)^{i+1}(i+1)\,;\, i = 1, 2, 3, \ldots$$

with the probability law: $p_i = P(X = x_i) = \dfrac{1}{i(i+1)}$; $i = 1, 2, 3, \ldots$

Here $$\sum_{i=1}^{\infty} x_i P(X = x_i) = \sum_{i=1}^{\infty} (-1)^{i+1}\left(\frac{1}{i}\right) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

Using Leibnitz test for alternating series, the series on right-hand side is conditionally convergent, since the terms alternate in sign are monotonically decreasing and coverge to zero.

By conditional convergence we mean that although $\sum_{i=1}^{\infty} p_i x_i$ converges, $\sum_{i=1}^{\infty} |p_i x_i|$ does not converge. So, rigorously speaking, in the above example $E(X)$ does not exist, although $\sum_{i=1}^{\infty} p_i x_i$ is finite, *viz.*, $\log_e 2$.

As another example, let us consider the $r.v.$ X which takes the values $x_k = \dfrac{(-1)^k \cdot 2^k}{k}$, $(k = 1, 2, 3, \ldots)$, with probabilities $p_k = 2^{-k}$. Here also we get

$$\sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -\left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots\right] = -\log_e 2 \text{ and } \sum_{k=1}^{\infty} |x_k|\, p_k = \sum_{k=1}^{\infty} \frac{1}{k}$$

which is a divergent series. Hence in this case also expectation does not exist.

As an illustration of a continuous $r.v.$, let us consider the $r.v.$ X with $p.d.f.$:

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

which is $p.d.f.$ of standard Cauchy distribution. [*c.f.* Chapter 9].

$$\int_{-\infty}^{\infty} |x|\, f(x)\, dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{|x|}{1+x^2}\, dx = \frac{2}{\pi} \int_0^{\infty} \frac{x}{1+x^2}\, dx = \frac{1}{\pi} \Big| \log(1+x^2) \Big|_0^{\infty} \to \infty$$

($\because$ Integrand is an even function of x.)

Since this integral does not converge to a finite limit, $E(X)$ does not exist.

6·3. EXPECTED VALUE OF FUNCTION OF A RANDOM VARIABLE

Consider a $r.v.$ X with $p.d.f.$ ($p.m.f.$) $f(x)$ and distribution function $F(x)$. If $g(\cdot)$ is a function such that $g(X)$ is a $r.v.$ and $E[g(X)]$ exists (*i.e.*, is defined), then

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx \quad \text{(For continuous } r.v.\text{)} \quad \text{... (6·3)}$$

$$E[g(X)] = \sum_x g(x) f(x) \quad \text{(For discrete } r.v.\text{)} \quad \text{... (6·3a)}$$

By definition, the expectation of $Y = g(X)$ is :

$$E[g(X)] = E(Y) = \int_y y \, dH_Y(y) = \int_y y \, h(y) \, dy \qquad \ldots (6\cdot4)$$

or
$$E(Y) = \sum_y y \, h(y), \qquad \ldots (6\cdot4a)$$

where $H_Y(y)$ is the distribution function of Y and $h(y)$ is *p.d.f.* of Y.

[The proof of equivalence of (6·3) and (6·4) is beyond the scope of the book.]

This result extends into higher dimensions. If X and Y have a joint *p.d.f.*, $f(x, y)$ and $Z = h(x, y)$ is a random variable for some function h and if $E(Z)$ exists, then

$$E(Z) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x, y) f(x, y) \, dx \, dy \qquad \ldots (6\cdot5)$$

or
$$E(Z) = \sum_x \sum_y h(x, y) f(x, y) \qquad \ldots (6\cdot5a)$$

Particular Cases 1. If we take $g(X) = X^r$, r being a positive integer, in (6·3),

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x) \, dx, \qquad \ldots (6\cdot5b)$$

which is defined as μ_r', the rth moment (about origin) of the probability distribution.

Thus $\quad \mu_r'$ *(about origin)* $= E(X^r)$. In particular

μ_1' *(about origin)* $= E(X)$ and μ_2' *(about origin)* $= E(X^2)$

Hence, $\quad$ Mean $= \bar{x} = \mu_1'$ *(about origin)* $= E(X) \qquad \ldots (6\cdot6)$

and $\quad \mu_2 = \mu_2' - \mu_1'^2 = E(X^2) - \{E(X)\}^2 \qquad \ldots (6\cdot6a)$

2. If $g(X) = [X - E(X)]^r = (X - \bar{x})^r$, then from (6·3), we obtain

$$E[X - E(X)]^r = \int_{-\infty}^{\infty} [x - E(X)]^r f(x) \, dx = \int_{-\infty}^{\infty} (x - \bar{x})^r f(x) \, dx, \qquad \ldots (6\cdot7)$$

which is μ_r, the rth moment about mean.

In particular, if $r = 2$, we get

$$\boxed{\mu_2 = E[X - E(X)]^2 = \int_{-\infty}^{\infty} (x - \bar{x})^2 f(x) \, dx} \qquad \ldots (6\cdot8)$$

Formulae (6·6a) and (6·8) give the variance of the probability distribution of a continuous *r.v.* X in terms of expectation.

3. Taking $\quad g(x) = constant = c$, say in (6·3), we get

$$E(c) = \int_{-\infty}^{\infty} c \, f(x) \, dx = c \int_{-\infty}^{\infty} f(x) \, dx = c \qquad \ldots (6\cdot9)$$

$$\therefore \quad E(c) = c \qquad \ldots (6\cdot9a)$$

Remark. The corresponding results for a discrete *r.v.* X can be obtained on replacing integration by summation (Σ) over the given range of the variable X in the formulae (6·5) to (6·9).

In the following sections, we shall establish some more results on 'Expectation' in the form of Theorems, for continuous *r.v.*'s. The corresponding results for discrete *r.v.*'s can be obtained similarly on replacing integration by summation (Σ) over the given range of the variable X and are left as an exercise to the reader.

6·4. PROPERTIES OF EXPECTATION

Property 1. Addition Theorem of Expectation.

If X and Y are random variables, then $E(X + Y) = E(X) + E(Y)$, ... (6·10)

provided all the expectations exist.

Proof. Let X and Y be continuous *r.v.*'s with joint *p.d.f.* $f_{XY}(x, y)$ and marginal *p.d.f.*'s $f_X(x)$ and $f_Y(y)$ respectively. Then by def.,

$$E(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx \quad \text{... (6·11)} \quad \text{and} \quad E(Y) = \int_{-\infty}^{\infty} y f_Y(y)\,dy \quad \text{...(6·12)}$$

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

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x f_{XY}(x, y)\,dx\,dy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f_{XY}(x, y)\,dx\,dy$$

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

$$= \int_{-\infty}^{\infty} x f_X(x)\,dx + \int_{-\infty}^{\infty} y f_Y(y)\,dy$$

$$= E(X) + E(Y) \qquad \text{[On using (6·11) and (6·12)]}$$

The result in (6·10) can be extended to n variables as given below.

Generalisation. *The mathematical expectation of the sum of n random variables is equal to the sum of their expectations, provided all the expectations exist. Symbolically, if* $X_1, X_2, \ldots, X_n$ *are random variables then*

$$E(X_1 + X_2 + \ldots + X_n) = E(X_1) + E(X_2) + \ldots + E(X_n) \quad \text{.. (6·13)}$$

or
$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i), \text{ if all the expectations exist.} \quad \text{... (6·13a)}$$

Proof. Using (6·10), for two *r.v.*'s X_1 and X_2, we get

$$E(X_1 + X_2) = E(X_1) + E(X_2) \quad \Rightarrow \quad \text{(6·13) is true for } n = 2. \quad \text{... (*)}$$

Let us now suppose that (6·13) is true for $n = r$ (say), so that

$$E\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} E(X_i) \quad \text{...(6·14)}$$

$$E\left(\sum_{i=1}^{r+1} X_i\right) = E\left[\sum_{i=1}^{r} X_i + X_{r+1}\right] = E\left(\sum_{i=1}^{r} X_i\right) + E(X_{r+1}) \quad \text{[Using (6·10)]}$$

$$= \sum_{i=1}^{r} E(X_i) + E(X_{r+1}) \quad \text{[Using (6·14)]}$$

$$= \sum_{i=1}^{r+1} E(X_i)$$

Hence if (6·13) is true for $n = r$, it is also true for $n = r + 1$. But we have proved in (*) above that (6·13) is true for $n = 2$. Hence it is true for $n = 2 + 1 = 3$; $n = 3 + 1 = 4$ ...

and so on. Hence by the principle of mathematical induction, (6·13) is true for all positive integral values of n.

Property 2. Multiplication Theorem of Expectation

If X and Y are independent random variables, then $E(XY) = E(X) . E(Y)$ ...(6·15)

Proof. Proceeding as in property 1, we have

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_{XY}(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy\, f_X(x) f_Y(y)\, dx\, dy \qquad \text{[Since X and Y are independent]}$$

$$= \int_{-\infty}^{\infty} x f_X(x)\, dx \int_{-\infty}^{\infty} y f_Y(y)\, dy \qquad \text{[Using (6·11) and (6·12)]}$$

$= E(X) E(Y)$, provided X and Y are independent

Generalisation. *The mathematical expectation of the product of a number of independent random variables is equal to the product of their expectations. Symbolically, if* $X_1, X_2, ..., X_n$ *are n independent r.v.'s, then*

$$\left.\begin{aligned} E(X_1, X_2, \ldots X_n) &= E(X_1) E(X_2) \ldots E(X_n) \\ \text{i.e.,} \quad E\left(\prod_{i=1}^{n} X_i\right) &= \prod_{i=1}^{n} E(X_i), \end{aligned}\right\} \qquad \ldots(6·16)$$

provided all the expectations exist.

Proof. Using (6·15), for two *independent r.v.*'s X_1 and X_2, we get

$$E(X_1 X_2) = E(X_1) E(X_2) \quad \Rightarrow \quad (6·16) \text{ is true for } n = 2. \qquad \ldots(*)$$

Let us now suppose that (6·16) is true for $n = r$, (say) so that

$$E\left(\prod_{i=1}^{r} X_i\right) = \prod_{i=1}^{r} E(X_i) \qquad \ldots(6·17)$$

Thus
$$E\left(\prod_{i=1}^{r+1} X_i\right) = E\left(\prod_{i=1}^{r} X_i \,.\, X^{r+1}\right) = E\left(\prod_{i=1}^{r} X_i\right). E(X_{r+1}) \qquad \text{[Using (6·15)]}$$

$$= \left[\prod_{i=1}^{r} (E X_i)\right] E(X_{r+1}) \qquad \text{[Using (6·17)]}$$

$$= \prod_{i=1}^{r+1} (E X_i)$$

Hence if (6·16) is true for $n = r$, it is also true for $n = r + 1$. Hence using (*), by the principle of mathematical induction we conclude that (6·16) is true for all positive integral values of n.

Property 3. *If X is a random variable and 'a' is constant, then*

(*i*) $E[a\Psi(X)] = a E[\Psi(X)]$ ... (6·18)

(*ii*) $E[\Psi(X) + a] = E[\Psi(X)] + a,$ ... (6·19)

where $\Psi(X)$, *a function of X, is a r.v. and all the expectations exist.*

Proof.

(i) $$E[a\,\Psi(X)] = \int_{-\infty}^{\infty} a\,\Psi(x).f(x)\,dx = a\int_{-\infty}^{\infty} \Psi(x) f(x)\,dx = a\,E[\Psi(X)]$$

(ii) $$E[\Psi(X)+a] = \int_{-\infty}^{\infty} \{\Psi(x)+a\} f(x)\,dx = \int_{-\infty}^{\infty} \Psi(x) f(x)\,dx + a\int_{-\infty}^{\infty} f(x)\,dx$$

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

Cor. (i) If $\Psi(X) = X$, in (6·18) then

$$E(a\,X) = a\,E(X) \quad \text{and} \quad E(X+a) = E(X)+a \qquad \text{... (6·20)}$$

(ii) If $\Psi(X) = 1$ in (6·18) then $E(a) = a$ ... (6·21)

Property 4. *If X is a random variable and a and b are constants, then*

$$E(a\,X+b) = a\,E(X)+b, \qquad \text{... (6·22)}$$

provided all the expectations exist.

Proof. By def., we have

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

Cor. 1. If $b = 0$, then we get $E(a\,X) = a\,.\,E(X)$ ... (6·22*a*)

Cor. 2. Taking $a = 1$, $b = -\overline{X} = -E(X)$, we get $E(X-\overline{X}) = 0$

Remark. If we write $g(X) = a\,X + b$ ... (6·23)

then $g[E(X)] = a\,E(X)+b$ ... (6·23*a*)

Hence from (6·22) and (6·23*a*), $E\{g(X)\} = g\{E(X)\}$ ... (6·24)

Now (6·23) and (6·24) imply *that expectation of a linear function is the same linear function of the expectation*. The result, however, is not true if $g(\cdot)$ is not linear. For instance,

$$E(1/X) \neq \{1/E(X)\} \quad ; \quad E(X^{1/2}) \neq [E(X)]^{1/2}$$
$$E[\log(X)] \neq \log[E(X)] \quad ; \quad E(X^2) \neq [E(X)]^2,$$

since all the functions stated above are non-linear. As an illustration, let us consider a random variable X which assumes only two values + 1 and –1, each with equal probability $\frac{1}{2}$. Then

$$E(X) = 1 \times \frac{1}{2} + (-1) \times \frac{1}{2} = 0 \text{ and } E(X^2) = 1^2 \times \frac{1}{2} + (-1)^2 \times \frac{1}{2} = 1.$$

Thus $$E(X^2) \neq [E(X)]^2$$

For a non-linear function $g(X)$, it is difficult to obtain expressions for $E[g(X)]$ in terms of $g[E(X)]$, say, for $E[\log(X)]$ or $E(X^2)$ in terms of $\log[E(X)]$ or $[E(X)]^2$. However, some results in the form of inequalities between $E[g(X)]$ and $g[E(X)]$ are available, as discussed in later part of the chapter.

Property 5. Expectation of a Linear Combination of Random Variables :

Let $X_1, X_2, \ldots, X_n$ be any n random variables and if $a_1, a_2, \ldots, a_n$ are any n constants, then

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i) \qquad \text{... (6·25)}$$

provided all the expectations exist.

The result is obvious from (6·13) and (6·20).

Property 6. *If* $X \geq 0$ *then* $E(X) \geq 0$.

Proof. If X is a continuous random variable *s.t.* $X \geq 0$, then

$$E(X) = \int_{-\infty}^{\infty} x \cdot p(x)\, dx = \int_{0}^{\infty} x \cdot p(x)\, dx > 0, \qquad [\because \text{ If } X \geq 0, p(x) = 0 \text{ for } x < 0]$$

provided the expectation exists.

Property 7. *If X and Y are two random variables such that* $Y \leq X$*, then* $E(Y) \leq E(X)$*, provided all the expectations exist.*

Proof. Since $Y \leq X$, we have the *r.v.* $Y - X \leq 0 \Rightarrow X - Y \geq 0$

Hence $E(X - Y) \geq 0 \quad \Rightarrow \quad E(X) - E(Y) \geq 0$

$\Rightarrow \quad E(X) \geq E(Y) \quad \Rightarrow \quad E(Y) \leq E(X)$, as desired.

Property 8. $|E(X)| \leq E|X|$*, provided the expectations exist.* ... (6·26)

Proof. Since $X \leq |X|$, we have by Property 7, $E(X) \leq E|X|$ ... (*)

Again since $-X \leq |X|$, we have by Property 7, $E(-X) \leq E|X|$

$\therefore \quad -E(X) \leq E|X|$ ... (**)

From (*) and (**), we get the desired result $|E(X)| \leq E|X|$.

Property 9. *If* μ_r' *exists, then* μ_s' *exists for all* $1 \leq s \leq r$.

Mathematically, if $E(X^r)$ *exists, then* $E(X^s)$ *exist for all* $1 \leq s \leq r$*, i.e.,*

$$E(X^r) < \infty \quad \Rightarrow \quad E(X^s) < \infty, \ \forall\ 1 \leq s \leq r \qquad \text{... (6·27)}$$

Proof. $\displaystyle\int_{-\infty}^{\infty} |x|^s\, dF(x) = \int_{-1}^{1} |x|^s\, dF(x) + \int_{|x|>1} |x|^s\, dF(x)$

If $s < r$, then $|x|^s < |x|^r$, for $|x| > 1$.

$$\therefore \quad \int_{-\infty}^{\infty} |x|^s\, dF(x) \leq \int_{-1}^{1} |x|^s\, dF(x) + \int_{|x|>1} |x|^r\, dF(x)$$

$$\leq \int_{-1}^{1} dF(x) + \int_{|x|>1} |x|^r\, dF(x),$$

since for $-1 < x < 1$, $|x|^s < 1$.

$$\therefore \quad \int_{-\infty}^{\infty} |x|^s\, dF(x) \leq 1 + E|X|^r < \infty \qquad [\because E(X^r) \text{ exists}]$$

$$\Rightarrow \qquad E(X^s) \text{ exists}, \ \forall\ 1 \leq s \leq r$$

Remark. The above result states that if the moments of a specified order exist, then all the lower order moments automatically exist. However, the converse is not true, *i.e.*, we may have distributions for which all the moments of a specified order exist but no higher order moment exist. For example, for the *r.v.* with *p.d.f.* :

$$p(x) = \begin{cases} 2/x^3 & ; \ x \geq 1 \\ 0 & ; \ x < 1 \end{cases}$$

we have :

$$E(X) = \int_{1}^{\infty} x\, p(x)\, dx = 2\int_{1}^{\infty} x^{-2}\, dx = \left| \left(\frac{-2}{x} \right) \right|_{1}^{\infty} = 2$$

$$E(X^2) = \int_{1}^{\infty} x^2\, p(x)\, dx = 2\int_{1}^{\infty} \frac{1}{x}\, dx = \infty$$

Thus for the above distribution, 1st order moment (mean) exists but 2nd order moment (variance) does not exist.

As another illustration, consider a *r.v.* X with *p.d.f.* :

$$p(x) = \frac{(r+1)\,a^{r+1}}{(x+a)^{r+2}}\,;\ x \ge 0,\ a > 0$$

$$\mu_r' = E(X^r) = (r+1)\,a^{r+1}\int_0^{\infty}\frac{x^r}{(x+a)^{r+2}}\,dx$$

Put $x = ay$ and using Beta integral : $\displaystyle\int_0^{\infty}\frac{x^{m-1}}{(1+x)^{m+n}} = \beta(m, n)$, we shall get, on simplification :

$$\mu_r' = (r+1)\,a^r\,.\,\beta(r+1, 1) = a^r$$

However, $\displaystyle \mu_{r+1}' = E(X^{r+1}) = (r+1)\,a^{r+1}\int_0^{\infty}\frac{x^{r+1}}{(x+a)^{r+2}}dx \to \infty,$

as the integral is not convergent. Hence in this case only the moments up to *r*th order exist and higher order moments do not exist.

Property 10. *If X and Y are independent random variables, then*

$$E[h(X)\,.\,k(Y)] = E[h(X)]\,E[k(Y)] \qquad \text{... (6·28)}$$

where h(.) is a function of X alone and k(·) is a function of Y alone, provided expectations on both sides exist.

Proof. Let $f_X(x)$ and $g_Y(y)$ be the marginal *p.d.f.*'s of X and Y respectively. Since X and Y are independent, their joint *p.d.f.* $f_{XY}(x, y)$ is given by :

$$f_{XY}(x, y) = f_X(x)\,f_Y(y) \qquad \text{... (*)}$$

By def., for continuous *r.v.*'s

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

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(x)\,k(y)\,f(x)\,g(y)\,dx\,dy \qquad \text{[From (*)]}$$

Since $E[h(X)\,k(Y)]$ exists, the integral on the right-hand side is absolutely convergent and hence by Fuibini's theorem for integrable functions, we can change the order of integration to get

$$E[h(X)\,k(Y)] = \left[\int_{-\infty}^{\infty} h(x) f(x)\,dx\right]\left[\int_{-\infty}^{\infty} k(y)\,g(y)\,dy\right] = E[h(X)]\,.\,E[k(Y)],$$

as desired.

Remark. The result can be proved for discrete random variables X and Y on replacing integration by summation over the given range of X and Y.

6·5. PROPERTIES OF VARIANCE

If X is a random variable, then $V(aX + b) = a^2\,V(X)$, ...(6·29)

where a and b are constants.

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

$$\therefore \quad Y - E(Y) = a\,[X - E(X)]$$

Squaring and taking expectation of both sides, we get

$$E[Y - E(Y)]^2 = a^2 E[X - E(X)]^2$$

$$\Rightarrow \quad V(Y) = a^2 V(X) \quad \text{or} \quad V(aX + b) = a^2 V(X),$$

where $V(X)$ is written for variance of X.

Cor. (*i*) If $b = 0$, then $V(aX) = a^2 V(x)$ ... (6·29*a*)

$\Rightarrow$ Variance is not independent of change of scale.

(*ii*) If $a = 0$, then $V(b) = 0 \Rightarrow$ Variance of a constant is zero. ... (6·29*b*)

(*iii*) If $a = 1$, then $V(X + b) = V(X)$ ... (6·29*c*)

$\Rightarrow$ Variance is independent of change of origin.

6·6. COVARIANCE

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

$$Cov(X,Y) = E[\{X - E(X)\}\{Y - E(Y)\}] \qquad \text{... (6·30)}$$

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

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

$$= E(XY) - E(X) E(Y) \qquad \text{... (6·30a)}$$

If X and Y are independent then $E(XY) = E(X) E(Y)$ and hence in this case

$$Cov(X, Y) = E(X) E(Y) - E(X) E(Y) = 0 \qquad \text{... (6·30b)}$$

Remarks 1.
$$Cov(aX, bY) = E[\{aX - E(aX)\}\{bY - E(bY)\}]$$

$$= E[a\{X - E(X)\}\, b\{Y - E(Y)\}]$$

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

$$= ab\, Cov(X, Y) \qquad \text{... (6·31)}$$

2.
$$Cov(X + a, Y + b) = Cov(X, Y) \qquad \text{... (6·31a)}$$

3.
$$Cov\left(\frac{X - \bar{X}}{\sigma_X}, \frac{Y - \bar{Y}}{\sigma_Y}\right) = \frac{1}{\sigma_X \sigma_Y} Cov(X, Y) \qquad \text{... (6·31b)}$$

4. Similarly, we shall get :

$$Cov(aX + b, cY + d) = ac\, Cov(X, Y) \qquad \text{... (6·31c)}$$

$$Cov(X + Y, Z) = Cov(X, Z) + Cov(Y, Z) \qquad \text{... (6·31d)}$$

$$Cov(aX + bY, cX + dY) = ac\sigma_X^2 + bd\sigma_Y^2 + (ad + bc)\, Cov(X, Y) \qquad \text{... (6·31e)}$$

5. If X and Y are independent, $Cov(X, Y) = 0$. [c.f. (6·30*b*)]. However, the converse is not true. For illustrations see Chapter 10 on Correlation.

6·6·1. Variance of a Linear Combination of Random Variables

Let $X_1, X_2, \ldots, X_n$ be n random variables, then

$$V\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i^2 V(X_i) + 2 \mathop{\sum_{i=1}^{n} \sum_{j=1}^{n}}_{(i<j)} a_i a_j\, Cov(X_i, X_j) \qquad \text{... (6·32)}$$

Proof. Let
$$U = a_1 X_1 + a_2 X_2 + \ldots + a_n X_n$$

so that
$$E(U) = a_1 E(X_1) + a_2 E(X_2) + \ldots + a_n E(X_n)$$

$$\therefore \quad U - E(U) = a_1 \{X_1 - E(X_1)\} + a_2 \{X_2 - E(X_2)\} + \ldots + a_n \{X_n - E(X_n)\}$$

Squaring and taking expectation of both sides, we get

$$E[U - E(U)]^2 = a_1^2 E[X_1 - E(X_1)]^2 + a_2^2 E[X_2 - E(X_2)]^2 + \ldots + a_n^2 E[X_n - E(X_n)]^2$$

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

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

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

Remarks 1. If $a_i = 1\,;\, i = 1, 2, \ldots, n$, then

$$V(X_1 + X_2 + \ldots + X_n) = V(X_1) + V(X_2) + \ldots + V(X_n) + 2 \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i<j}}^{n} \operatorname{Cov}(X_i, X_j) \qquad \ldots (6{\cdot}32a)$$

2. If $X_1, X_2, \ldots X_n$ are independent (pairwise), then $\operatorname{Cov}(X_i, X_j) = 0, (i \neq j)$.
Thus from (6·32) and (6·32*a*), we get

$$\left.\begin{aligned} V(a_1X_1 + a_2X_2 + \ldots + a_n X_n) &= a_1^2\, V(X_1) + a_2^2\, V(X_2) + \ldots + a_n^2\, V(X_n) \\ \text{and} \quad V(X_1 + X_2 + \ldots + X_n) &= V(X_1) + V(X_2) + \ldots + V(X_n), \end{aligned}\right\} \qquad \ldots (6{\cdot}32b)$$

provided $X_1, X_2, \ldots, X_n$ are independent.

3. If $a_1 = 1 = a_2$ and $a_3 = a_4 = \ldots = a_n = 0$, then from (6·32), we get

$$V(X_1 + X_2) = V(X_1) + V(X_2) + 2 \operatorname{Cov}(X_1, X_2)$$

Again if $a_1 = 1,\ a_2 = -1$ and $a_3 = a_4 = \ldots = a_n = 0$, then

$$V(X_1 - X_2) = V(X_1) + V(X_2) - 2 \operatorname{Cov}(X_1, X_2)$$

Thus we have

$$V(X_1 \pm X_2) = V(X_1) + V(X_2) \pm 2 \operatorname{Cov}(X_1, X_2) \qquad \ldots (6{\cdot}32c)$$

If X_1 and X_2 are independent, then $\operatorname{Cov}(X_1, X_2) = 0$ and we get

$$V(X_1 \pm X_2) = V(X_1) + V(X_2) \qquad \ldots (6{\cdot}32d)$$

Example 6·1. *Let X be a random variable with the following probability distribution :*

x	:	-3	6	9
$P(X = x)$	:	$1/6$	$1/2$	$1/3$

Find E (X) and E (X²) and using the laws of expectation, evaluate E (2X + 1)².

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

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

$$\therefore \quad E(2X+1)^2 = E(4X^2 + 4X + 1) = 4E(X^2) + 4E(X) + 1 = 4 \times \frac{93}{2} + 4 \times \frac{11}{2} + 1 = 209.$$

Example 6·2. *(a) Find the expectation of the number on a die when thrown.*

(b) Two unbiased dice are thrown. Find the expected values of the sum of numbers of points on them.

Solution. (*a*) 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, ..., 6 each with equal probability $\frac{1}{6}$ · Hence

$$E(X) = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \ldots + \frac{1}{6} \times 6 = \frac{1}{6}(1 + 2 + 3 + \ldots + \qquad \frac{1}{6} \times \frac{6 \times 7}{2} = \frac{7}{2} \qquad \ldots(*)$$

Remark. This does not mean that in a random throw of a dice, the player will get the number $\frac{7}{2} = 3.5$. In fact, one can never get this (fractional) number in a throw of a dice. Rather, this implies that if the player tosses the dice for a "long" period, then on the average toss he will get $\frac{7}{2} = 3.5$.

(*b*) The probability function of X (the sum of numbers obtained on two dice), is

Value of X : x	2	3	4	5	6	7	....	11	12
Probability	1/36	2/36	3/36	4/36	5/36	6/36	....	2/36	1/36

$$E(X) = \sum_i p_i x_i$$

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

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

Aliter. Let X_i be the number obtained on the ith dice $(i = 1, 2)$ when thrown. Then the sum of the number of points on two dice is given by :

$$S = X_1 + X_2 \Rightarrow E(S) = E(X_1) + E(X_2) = \frac{7}{2} + \frac{7}{2} = 7 \qquad \text{[On using (*)]}$$

Remark. This result can be generalised to the sum of points obtained in a random throw of n dice. Then

$$E(S) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{7}{2} = \frac{7n}{2}.$$

Example 6·3. *In four tosses of a coin, let X be the number of heads. Tabulate the 16 possible outcomes with the corresponding values of X. By simple counting, derive the probability distribution of X and hence calculate the expected value of X.*

Solution. Let H represent a head, T a tail and X, the random variable denoting the number of heads.

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

The random variable X takes the values 0, 1, 2, 3 and 4. Since, from the above table, we find that the number of cases favourable to the coming of 0, 1, 2, 3 and 4 heads are 1, 4, 6, 4 and 1 respectively, we have

$$P(X = 0) = \frac{1}{16},\ P(X = 1) = \frac{4}{16} = \frac{1}{4},\ P(X = 2) = \frac{6}{16} = \frac{3}{8},\ P(X = 3) = \frac{4}{16} = \frac{1}{4},\ P(X = 4) = \frac{1}{16}.$$

The probability distribution of X can be summarized as follows :

x :	0	1	2	3	4
$p(x)$:	$\frac{1}{16}$	$\frac{1}{4}$	$\frac{3}{8}$	$\frac{1}{4}$	$\frac{1}{16}$

$$\therefore \quad E(X) = \sum_{x=0}^{4} x\,p(x) = 1\cdot\frac{1}{4}+2\cdot\frac{3}{8}+3\cdot\frac{1}{4}+4\cdot\frac{1}{16}=\frac{1}{4}+\frac{3}{4}+\frac{3}{4}+\frac{1}{4}=2.$$

Example 6·4. *An urn contains 7 white and 3 red balls. Two balls are drawn together, at random from this urn. Compute the probability that neither of them is white. Find also the probability of getting one white and one red ball. Hence compute the expected number of white balls drawn.*

Solution. Let X denote the number of white balls drawn. The probability distribution of X is obtained as follows :

x :	0	1	2
$p(x)$:	$\frac{^3C_2}{^{10}C_2}=\frac{1}{15}$	$\frac{^7C_1\times {^3C_1}}{^{10}C_2}=\frac{7}{15}$	$\frac{^7C_2}{^{10}C_2}=\frac{7}{15}$

Then expected number of white balls drawn is :

$$E(X) = 0\times\frac{1}{15}+1\times\frac{7}{15}+2\times\frac{7}{15}=\frac{21}{15}.$$

Example 6·5. *A gamester has a disc with a freely revolving needle. The disc is divided into 20 equal sectors by thin lines and the sectors are marked 0, 1, 2, ..., 19. The gamester treats 5 or any multiple of 5 as lucky numbers and zero as a special lucky number. He allows a player to whirl the needle on a charge of 10 paise. When the needdle stops at the lucky number the gamester pays back the player twice the sum charged and at the special lucky number the gamester pays to the player 5 times of the sum charged. Is the game fair ? What is the expectation of the player ?*

Solution.

Event	Favourable	$p(x)$	Player's Gain (x)
Lucky number	5, 10, 15	3/20	$20-10=10\,p$
Special lucky No.	0	1/20	$50-10=40\,p$
Other numbers	1, 2, 3, 4, 6, 7, 8, 9, 11, 12, 13, 14, 16, 17, 18, 19	16/20	$-10\,p$

$$\therefore \quad E(X) = \frac{3}{20}\times 10+\frac{1}{20}\times 40-\frac{16}{20}\times 10=-\frac{9}{2}\neq 0,$$ *i.e.*, the game is *not fair*.

Example 6·6. *A box contains 2^n tickets among which ^nC_i tickets bear the number i ; $i = 0, 1, 2, ..., n$. A group of m tickets is drawn. What is the expectation of the sum of their numbers ?*

Solution. Let X_i ; $i = 1, 2, ..., m$ be the variable representing the number on the ith ticket drawn. Then the sum 'S' of the numbers on the tickets drawn is given by :

$$S = X_1+X_2+\dots+X_m=\sum_{i=1}^{m} X_i, \text{ so that } E(S)=\sum_{i=1}^{m} E(X_i)$$

X_i is a random variable which can take any one of the possible values 0, 1, 2, ..., n with respective probabilities : $^nC_0/2^n$, $^nC_1/2^n$, $^nC_2/2^n$, ..., $^nC_n/2^n$.

$$\therefore \quad E(X_i) = \frac{1}{2^n}\left(1.\,^nC_1+2.\,^nC_2+3.\,^nC_3+\dots+n.\,^nC_n\right)$$

$$= \frac{1}{2^n}\left(1.n+2.\frac{n(n-1)}{2!}+3.\frac{n(n-1)(n-2)}{3!}+\dots+n.1\right)$$

$$= \frac{n}{2^n}\left\{1+(n-1)+\frac{(n-1)(n-2)}{2!}+\dots+1\right\}$$

$$= \frac{n}{2^n} \left({}^{n-1}C_0 + {}^{n-1}C_1 + {}^{n-1}C_2 + \dots + {}^{n-1}C_{n-1} \right) = \frac{n}{2^n} \cdot (1+1)^{n-1} = \frac{n}{2}$$

$$\therefore \quad E(S) = \sum_{i=1}^{m} E(X_i) = \sum_{i=1}^{m} (n/2) = \frac{mn}{2}.$$

Example 6·7. *A coin is tossed until a head appears. What is the expectation of the number of tosses required ?*

Solution. Let X denote the number of tosses required to get the first head. Then X can materialise in the following ways :

Event	x	*Probability*, $p(x)$
H	1	$\frac{1}{2}$
TH	2	$\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$
TTH	3	$\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8}$
⋮	⋮	⋮

$$\therefore \quad E(X) = \sum_{x=1}^{\infty} x\, p(x) = 1 \times \frac{1}{2} + 2 \times \frac{1}{4} + 3 \times \frac{1}{8} + 4 \times \frac{1}{16} + \dots \qquad \dots (*)$$

This is an arithmetic-geometric series with ratio of GP being $r = \frac{1}{2}$.

Let $\quad S = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + \dots$

Then $\quad \frac{1}{2} S = \qquad \frac{1}{4} + 2 \cdot \frac{1}{8} + 3 \cdot \frac{1}{16} + \dots$

$$\therefore \quad \left(1 - \frac{1}{2}\right) S = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots \quad \Rightarrow \quad \frac{1}{2} S = \frac{\frac{1}{2}}{1 - \frac{1}{2}} = 1 \quad \text{or} \quad S = 2.$$

[Since the sum of an infinite G. P. with first term a and common ratio r (< 1) is $\frac{a}{1-r}$.]

Hence, substituting in (*), we have $\qquad E(X) = 2.$

Example 6·8. *What is the expectation of the number of failures preceding the first success in an infinite series of independent trials with constant probability p of success in each trial ?*

Solution. Let the random variable X denote the number of failures preceding the first success. Then X can take the values $0, 1, 2, \dots, \infty$. We have

$$P(X = x) = p(x) = P(x \text{ failures precede the first success}) = q^x p,$$

where $q = 1 - p$, is the probability of failure in a trial. Then by *def.*,

$$E(X) = \sum_{x=0}^{\infty} x\, p(x) = \sum_{x=0}^{\infty} x\, q^x p = pq \sum_{x=1}^{\infty} x\, q^{x-1} = pq\,(1 + 2q + 3q^2 + 4q^3 + \dots) \qquad \dots (*)$$

Now $\quad 1 + 2q + 3q^2 + 4q^3 + \dots$ is an infinite arithmetic-geometric series.

Let $\qquad S = 1 + 2q + 3q^2 + 4q^3 + \dots$

$\qquad\quad qS = \qquad q + 2q^2 + 3q^3 + \dots$

$$\therefore \quad (1-q)\,S = 1 + q + q^2 + q^3 + \dots = \frac{1}{1-q} \quad \Rightarrow \quad S = \frac{1}{(1-q)^2}$$

$$\therefore \quad 1 + 2q + 3q^2 + 4q^3 \dots = \frac{1}{(1-q)^2}. \quad \text{Hence} \quad E(X) = \frac{pq}{(1-q)^2} = \frac{pq}{p^2} = \frac{q}{p}. \quad \text{[From (*)]}$$

Example 6·9. *A box contains 'a' white and 'b' black balls. 'c' balls are drawn at random. Find the expected value of the number of white balls drawn.*

Solution. Let a variable X_i, associated with ith draw, be defined as follows :

$$X_i = \begin{cases} 1, \text{ if } i\text{th ball drawn is white} \\ 0, \text{ if } i\text{th ball drawn is black} \end{cases}$$

Then the number 'S' of the white balls among 'c' balls drawn is given by :

$$S = X_1 + X_2 + \ldots + X_c = \sum_{i=1}^{c} X_i \quad \Rightarrow \quad E(S) = \sum_{i=1}^{c} E(X_i) \qquad \ldots(*)$$

Now $\quad P(X_i = 1) = P$ (of drawing a white ball) $= \dfrac{a}{a+b}$

and $\quad P(X_i = 0) = P$ (of drawing a black ball) $= \dfrac{b}{a+b}$

$\therefore \quad E(X_i) = 1 . P(X_i = 1) + 0 . P(X_i = 0) = \dfrac{a}{a+b}$

Hence $\quad E(S) = \displaystyle\sum_{i=1}^{c} \left(\frac{a}{a+b}\right) = \frac{ca}{a+b}$. $\qquad$ [From (*)]

Example 6·10. *Let the r.v. X have the distribution :*

$$P(X = 0) = P(X = 2) = p\,;\, P(X = 1) = 1 - 2p, \text{ for } 0 \le p \le \tfrac{1}{2}.$$

For what p is the Var (X) a maximum ?

Solution. Here the r.v. X takes the values 0, 1, and 2 with respective probabilities $p, 1 - 2p$ and $p, 0 \le p \le \frac{1}{2}$. Thus

$$E(X) = 0 \times p + 1 \times (1 - 2p) + 2 \times p = 1,\ E(X^2) = 0 \times p + 1^2 \times (1 - 2p) + 2^2 \times p = 1 + 2p$$

$$\therefore \quad \text{Var}(X) = E(X^2) - \{E(X)\}^2 = 2p\,;\ 0 \le p \le \tfrac{1}{2}$$

Obviously, for $0 \le p \le \frac{1}{2}$, Var (X) is maximum when $p = \frac{1}{2}$, and $[\text{Var}\ (X)]_{max} = 2 \times \frac{1}{2} = 1$.

Example 6·11. *(Random Walk Problem). Starting from the origin, unit steps are taken to the right with probability p and to the left with probability q (= 1 – p). Assuming independent movements, find the mean and variance of the distance moved from origin after n steps .*

Solution. Let us associate a variable X_i with the ith step defined as follows :

$$X_i = \begin{cases} +1, \text{ if the } i\text{th step is towards the right,} \\ -1, \text{ if the } i\text{th step is towards the left.} \end{cases}$$

Then $S = X_1 + X_2 + \ldots + X_n = \Sigma X_i$, represents the random distance moved from origin after n steps.

$$E(X_i) = 1 \times p + (-1) \times q = p - q \quad \text{and} \quad E(X_i^2) = 1^2 \times p + (-1)^2 \times q = p + q = 1$$

$$\therefore \quad \text{Var}(X_i) = E(X_i^2) - \{E(X_i)\}^2 = (q + p)^2 - (p - q)^2 = 4pq$$

$$\text{Hence } E(S_n) = \sum_{i=1}^{n} E(X_i) = n(p - q) \quad \text{and} \quad V(S_n) = \sum_{i=1}^{n} V(X_i) = 4npq$$

[$\because$ Movements of steps are independent.]

Example 6·12. *Let X be a r.v. with mean μ and variance σ^2. Show that $E(X-b)^2$, as a function of b, is minimised when $b = \mu$.*

Solution.
$$E(X-b)^2 = E[(X-\mu)+(\mu-b)]^2$$
$$= E(X-\mu)^2 + (\mu-b)^2 + 2(\mu-b)E(X-\mu) = \text{Var}(X) + (\mu-b)^2 \quad [\because E(X-\mu)=0]$$

$$\Rightarrow \quad E(X-b)^2 \geq Var(X), \qquad (*)$$

since $(\mu-b)^2$, being the square of a real quantity is always non-negative.

The sign of equality holds in (*) iff $(\mu-b)^2 = 0 \Rightarrow \mu = b$.

Hence $E(X-b)^2$ is minimised when $\mu = b$ and its minimum value is $E(X-\mu)^2 = \sigma_X^2$.

Remark. This result states that the sum of squares of deviations is minimum when taken about mean.

Example 6·13. *In a sequence of Bernoulli trials, let X be the length of the run of either successes or failures starting with the first trial. Find E (X) and V (X).*

Solution. Let 'p' denote the probability of success. Then $q = 1-p$, is the probability of failure. X = 1 means that we can have any of the possibilities SF and FS with respective probabilities pq and qp.

$$\therefore \quad P(X=1) = P(SF) + P(FS) = pq + qp = 2pq$$

Similarly $\quad P(X=2) = P(SSF) + P(FFS) = p^2q + q^2p$

In general $\quad P(X=r) = P(SSS...SF) + P(FFF...FS) = p^r.q + q^r.p$

$$\therefore \quad E(X) = \sum_{r=1}^{\infty} r\,P(X=r) = \sum_{r=1}^{\infty} r(p^r.q + q^r.p) = pq\left(\sum_{r=1}^{\infty} r.p^{r-1} + \sum_{r=1}^{\infty} r.q^{r-1}\right)$$

$$= pq\{(1+2p+3p^2+...) + (1+2q+3q^2+...)\} = pq\{(1-p)^{-2} + (1-q)^{-2}\}$$

$$= pq(q^{-2}+p^{-2}) = pq\left(\frac{1}{q^2}+\frac{1}{p^2}\right) = \frac{p}{q}+\frac{q}{p}$$

(See Remark 1 to Example 6·15)

$$V(X) = E(X^2) - \{E(X)\}^2 = E\{X(X-1)\} + E(X) - \{E(X)\}^2 \qquad ...(*)$$

$$\text{Now } E\{X(X-1)\} = \sum_{r=2}^{\infty} r(r-1)P(X=r) = \sum_{r=2}^{\infty} r(r-1)(p^rq+q^rp)$$

$$= \sum_{r=2}^{\infty} r(r-1)p^rq + \sum_{r=2}^{\infty} r(r-1)q^rp$$

$$= p^2q\sum_{r=2}^{\infty} r(r-1)p^{r-2} + q^2p\sum_{r=2}^{\infty} r(r-1)q^{r-2}$$

$$= 2p^2q\sum_{r=2}^{\infty}\frac{r(r-1)}{2}p^{r-2} + 2q^2p\sum_{r=2}^{\infty}\frac{r(r-1)}{2}q^{r-2}$$

$$= 2p^2q(1-p)^{-3} + 2q^2p(1-q)^{-3} = 2\left(\frac{p^2}{q^2}+\frac{q^2}{p^2}\right)$$

$$\therefore \quad V(X) = 2\left(\frac{p^2}{q^2}+\frac{q^2}{p^2}\right) + \left(\frac{p}{q}+\frac{q}{p}\right) - \left(\frac{p}{q}+\frac{q}{p}\right)^2 = \left(\frac{p}{q}-\frac{q}{p}\right)^2 + \left(\frac{p}{q}+\frac{q}{p}\right) \quad ...[\text{From } (*)]$$

Example 6·14. *A deck of n numbered cards is thoroughly shuffled and the cards are inserted into n numbered cells one by one. If the card number 'i' falls in the cell 'i', we count it as a match, otherwise not. Find the mean and variance of total number of such matches.*

Solution. Let us associate a *r.v.*, X_i with the *i*th draw defined as follows :

$$X_i = \begin{cases} 1, \text{ if the } i\text{th card dealt has the number '}i\text{' on it} \\ 0, \text{ otherwise} \end{cases}$$

The total number of matches '*S*' is given by

$$S = X_1 + X_2 + ... + X_n = \sum_{i=1}^{n} X_i \quad \Rightarrow \quad E(S) = \sum_{i=1}^{n} E(X_i)$$

Now $\quad E(X_i) = 1.P(X_i = 1) + 0.P(X_i = 0) = P(X_i = 1) = \frac{1}{n}$

Hence $\quad E(S) = \sum_{i=1}^{n} \left(\frac{1}{n}\right) = n \cdot \frac{1}{n} = 1$

$$V(S) = V(X_1 + X_2 + ... + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{\substack{i=1 \\ i<j}}^{n} \sum_{j=1}^{n} \text{Cov}(X_i, X_j) \quad ...(1)$$

Now

$$V(X_i) = E(X_i^2) - \{E(X_i)\}^2 = 1^2 . P(X_i = 1) + 0^2 . P(X_i = 0) - \left(\frac{1}{n}\right)^2 = \frac{1}{n} - \frac{1}{n^2} = \frac{n-1}{n^2} \quad ...(2)$$

$$\text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i) E(X_j) \quad ...(3)$$

$$E(X_i X_j) = 1 . P(X_i X_j = 1) + 0 . P(X_i X_j = 0) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)},$$

since $X_i X_j = 1$ if and only if both card numbers *i* and *j* are in their respective matching places and there are $(n-2)$! arrangements of the remaining cards that correspond to this event. Substituting in (3), we get

$$\therefore \quad \text{Cov}(X_i, X_j) = \frac{1}{n(n-1)} - \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n^2(n-1)} \quad ...(4)$$

Substituting from (2) and (4) in (1), we have

$$V(S) = \sum_{i=1}^{n} \left(\frac{n-1}{n^2}\right) + 2 \sum_{\substack{i=1 \\ i<j}}^{n} \sum_{j=1}^{n} \left\{\frac{1}{n^2(n-1)}\right\} = n\left(\frac{n-1}{n^2}\right) + 2 . {}^nC_2 \frac{1}{n^2(n-1)} = \frac{n-1}{n} + \frac{1}{n} = 1.$$

Example 6·15. *If t is any positive real number, show that the function defined by*

$$p(x) = e^{-t}(1 - e^{-t})^{x-1} \quad ...(*)$$

can represent a probability function of a random variable X assuming the values 1, 2, 3, ... Find the E (X) and Var (X) of the distribution.

Solution. We have $e^t > 1, \forall t > 0 \Rightarrow e^{-t} < 1$ or $1 - e^{-t} > 0$

Also $\quad e^{-t} = \frac{1}{e^t} > 0, \forall t > 0$

Hence $\quad p(x) = e^{-t}(1 - e^{-t})^{x-1} \geq 0, \forall t > 0, x = 1, 2, 3, ...$

Also $\sum_{x=1}^{\infty} p(x) = e^{-t} \sum_{x=1}^{\infty} (1-e^{-t})^{x-1} = e^{-t} \sum_{x=1}^{\infty} a^{x-1}$, $(a = 1 - e^{-t})$

$$= e^{-t}(1 + a + a^2 + a^3 + \ldots) = e^{-t} \cdot \frac{1}{(1-a)}$$

$$= e^{-t}[1-(1-e^{-t})]^{-1} = e^{-t} . e^{t} = 1$$

Hence $p(x)$ defined in (*) represents the probability function of a $r.v.$ X.

$$E(X) = \Sigma x . p(x) = e^{-t} \sum_{x=1}^{\infty} x(1-e^{-t})^{x-1} = e^{-t} \sum_{x=1}^{\infty} x . a^{x-1}, \ (a = 1 - e^{-t})$$

$$= e^{-t}(1 + 2a + 3a^2 + 4a^3 + \ldots) = e^{-t}(1-a)^{-2} \qquad \text{(See Remark 1)}$$

$$= e^{-t}(e^{-t})^{-2} = e^{t}$$

$$E(X^2) = \Sigma x^2 p(x) = e^{-t} \sum_{x=1}^{\infty} x^2 . a^{x-1} = e^{-t}(1 + 4a + 9a^2 + 16a^3 + \ldots)$$

$$= e^{-t}(1+a)(1-a)^{-3} = e^{-t}(2-e^{-t})e^{3t} \qquad \text{(See Remark 2)}$$

Hence $\text{Var}(X) = E(X^2) - [E(X)]^2 = e^{-t}(2-e^{-t})e^{3t} - e^{2t}$

$$= e^{2t}[(2-e^{-t})-1] = e^{2t}(1-e^{-t}) = e^{t}(e^{t}-1).$$

***Remarks* 1.** Consider

$$S = 1 + 2a + 3a^2 + 4a^3 + \ldots \text{ (Arithmetic-geometric series)}$$

$$\Rightarrow \quad aS = a + 2a^2 + 3a^3 + \ldots$$

$$\Rightarrow \quad (1-a)S = 1 + a + a^2 + a^3 + \ldots = \frac{1}{(1-a)} \quad \text{or} \quad S = (1-a)^{-2}$$

$$\therefore \quad \sum_{x=1}^{\infty} xa^{x-1} = 1 + 2a + 3a^2 + 4a^3 + \ldots = (1-a)^{-2} \qquad \ldots (*)$$

2. Consider

$$S = 1 + 2^2 . a + 3^2 . a^2 + 4^2 . a^3 + 5^2 . a^4 \ldots$$

$$\Rightarrow \quad S = 1 + 4a + 9a^2 + 16a^3 + 25a^4 + \ldots$$

$$-3a\,S = -3a - 12a^2 - 27a^3 - 48a^4 - \ldots$$

$$+3a^2 S = +3a^2 + 12a^3 + 27a^4 + \ldots$$

$$-a^3 S = -a^3 - 4a^4 - \ldots$$

Adding the above equations, we get

$$(1-a)^3 S = 1 + a \quad \Rightarrow \quad S = (1+a)(1-a)^{-3} \qquad \ldots (**)$$

$$\therefore \quad \sum_{x=1}^{\infty} x^2 a^{x-1} = 1 + 4a + 9a^2 + 16a^3 + \ldots = (1+a)(1-a)^{-3}$$

The results (*) and (**) are quite useful for numerical problems and should be committed to memory.

Example 6·16. *A man with n keys wants to open his door and tries the keys independently and at random. Find the mean and variance of the number of trials required to open the door, (i) if unsuccessful keys are not eliminated from further selection, and (ii) if they are.*

Solution. (*i*) Suppose the man gets the first success at xth trial, *i.e.*, he is unable to open the door in the first $(x-1)$ trials. If unsuccessful keys are not eliminated then X is a random variable which can take the values 1, 2, 3, ... ∞.

Probability of success at the first trial $= \dfrac{1}{n}$

$\therefore$ Probability of failure at the first trial $= 1 - \dfrac{1}{n}$

If unsuccessful keys are not eliminated then the probability of success and consequently of failure is constant for each trial.

Hence $p(x)$ = Probability of 1st success at the x th trial $= \left(1 - \dfrac{1}{n}\right)^{x-1} \cdot \dfrac{1}{n}$.

Thus

$$E(X) = \sum_{x=1}^{\infty} x\,p(x) = \sum_{x=1}^{\infty} x\left(1 - \frac{1}{n}\right)^{x-1} \cdot \frac{1}{n} = \frac{1}{n} \sum_{x=1}^{\infty} x A^{x-1}, \text{ where } A = 1 - \frac{1}{n}.$$

$$E(X) = \frac{1}{n}(1 + 2A + 3A^2 + 4A^3 + \ldots) = \frac{1}{n}(1 - A)^{-2} \qquad \text{[See (*), Example (6·15)]}$$

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

$$E(X^2) = \sum_{x=1}^{\infty} x^2 . p(x) = \sum_{x=1}^{\infty} x^2 \left(1 - \frac{1}{n}\right)^{x-1} \cdot \frac{1}{n}$$

$$= \frac{1}{n} \sum_{x=1}^{\infty} x^2 A^{x-1} = \frac{1}{n}(1 + 2^2 . A + 3^2 . A^2 + 4^2 . A^3 + \ldots)$$

$$= \frac{1}{n}(1 + A)(1 - A)^{-3} \qquad \text{[See (**), Example (6·15)]}$$

$$= \frac{1}{n}\left\{1 + \left(1 - \frac{1}{n}\right)\right\}\left\{1 - \left(1 - \frac{1}{n}\right)\right\}^{-3} = (2n - 1)\,n$$

Hence $V(X) = E(X^2) - \{E(X)\}^2 = (2n - 1)\,n - n^2 = n^2 - n = n\,(n - 1)$

(*ii*) If unsuccessful keys are eliminated from further selection, then the random variable X will take the values from 1 to n. In this case, we have

Probability of success at the first trial $= \dfrac{1}{n}$

Probability of success at the 2nd trial $= \dfrac{1}{n-1}$

Probability of success at the 3rd trial $= \dfrac{1}{n-2}$

and so on.

Hence probability of 1st success at the 2nd trial $= \left(1 - \dfrac{1}{n}\right)\dfrac{1}{n-1} = \dfrac{1}{n}$

Probability of first success at the third trial $= \left(1 - \dfrac{1}{n}\right)\left(1 - \dfrac{1}{n-1}\right) \cdot \dfrac{1}{n-2} = \dfrac{1}{n}$,

and so on. In general, we have

$p(x)$ = Probability of first success at the xth trial $= \dfrac{1}{n}$

$$\therefore \quad E(X) = \sum_{x=1}^{n} x\,p(x) = \frac{1}{n} \sum_{x=1}^{n} x = \frac{n+1}{2}$$

$$E(X^2) = \sum_{x=1}^{n} x^2\,p(x) = \frac{1}{n} \sum_{x=1}^{n} x^2 = \frac{(n+1)(2n+1)}{6}$$

Hence $\quad V(X) = E(X^2) - \{E(X)\}^2 = \dfrac{(n+1)(2n+1)}{6} - \left(\dfrac{n+1}{2}\right)^2$

$$= \frac{n+1}{12}\{2(2n+1) - 3(n+1)\} = \frac{n^2-1}{12}.$$

Example 6·17. *In a lottery m tickets are drawn at a time out of n tickets numbered 1 to n. Find the expectation and the variance of the sum S of the numbers in the tickets drawn.*

Solution. Let X_i denote the score on the ith ticket drawn.

Then $\quad S = X_1 + X_2 + \dots + X_m = \sum_{i=1}^{m} X_i$, is the total score on the m tickets drawn.

$$\therefore \quad E(S) = \sum_{i=1}^{m} E(X_i)$$

Now each X_i is a random variable which assumes the values 1, 2, 3, ... , n each with equal probability $1/n$.

$$\therefore \quad E(X_i) = \frac{1}{n}(1+2+3+\dots+n) = \frac{(n+1)}{2}.$$

$$\text{Hence,} \quad E(S) = \sum_{i=1}^{m}\left(\frac{n+1}{2}\right) = \frac{m(n+1)}{2}$$

$$V(S) = V(X_1 + X_2 + \dots + X_m) = \sum_{i=1}^{m} V(X_i) + 2\sum_{\substack{i,j\\ i<j}}\sum \text{Cov}(X_i, X_j)$$

$$E(X_i^2) = \frac{1}{n}(1^2 + 2^2 + 3^2 + \dots + n^2) = \frac{1}{n}\cdot\frac{n(n+1)(2n+1)}{6} = \frac{(n+1)(2n+1)}{6}$$

$$\therefore \quad V(X_i) = E(X_i^2) - \{E(X_i)\}^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}$$

Also, $\quad \text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$

To find $E(X_i X_j)$, we note that the variables X_i and X_j can take the values as shown below :

X_i	X_j
1	2, 3, ..., n
2	1, 3, ..., n
⋮	⋮
n	1, 2, ..., $(n-1)$

In other words, the variable $X_i X_j$ can take $n(n-1)$ possible values and $P(X_i = l \cap X_j = k) = \dfrac{1}{n(n-1)}$, $k \neq l$. Hence

$$E(X_i X_j) = \frac{1}{n(n-1)}\begin{bmatrix} 1\cdot2 + 1\cdot3 + \dots\dots + \dots\dots 1.n \\ + 2\cdot1 + 2\cdot3 + \dots\dots + \dots\dots 2.n \\ + \dots\dots\dots\dots\dots\dots\dots\dots \\ \vdots \quad \vdots \quad\quad \vdots \\ + n\cdot1 + n\cdot2 + \dots\dots + n.(n-1) \end{bmatrix}$$

$$= \frac{1}{n(n-1)} \begin{bmatrix} 1(1+2+3+\ldots\ldots+n) - 1^2 \\ +2(1+2+3+\ldots\ldots+n) - 2^2 \\ +\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ +n(1+2+\ldots\overline{n-1}+n) - n^2 \end{bmatrix}$$

$$= \frac{1}{n(n-1)} \left\{ (1+2+3+\ldots n)^2 - (1^2+2^2+\ldots+n^2) \right\}$$

$$= \frac{1}{n(n-1)} \left[\left\{ \frac{n(n+1)}{2} \right\}^2 - \frac{n(n+1)(2n+1)}{6} \right] = \frac{(n+1)(3n^2-n-2)}{12(n-1)}$$

$\therefore$ $\text{Cov}(X_i, X_j) = \dfrac{(n+1)(3n^2-n-2)}{12(n-1)} - \left(\dfrac{n+1}{2}\right)^2 = \dfrac{(n+1)}{12(n-1)}\left[3n^2-n-2-3(n^2-1)\right] = -\dfrac{(n+1)}{12}$

Hence $V(S) = \sum\limits_{i=1}^{m} \left(\dfrac{n^2-1}{12}\right) + 2 \sum\sum\limits_{i<j=1}^{m} \left\{-\dfrac{(n+1)}{12}\right\} = \dfrac{m(n^2-1)}{12} + 2 \cdot \dfrac{m(m-1)}{2!} \left\{\dfrac{-(n+1)}{12}\right\},$

[Since there are ${}^mC_2$ covariance terms in $\text{Cov}(X_i, X_j)$]

$\therefore$ $V(S) = \dfrac{m(n+1)}{12}\left[(n-1)-(m-1)\right] = \dfrac{m(n+1)(n-m)}{12}.$

Example 6·18. *A die is thrown $(n+2)$ times. After each throw a '+' is recorded for 4, 5, or 6 and '–' for 1, 2 or 3, the signs forming an ordered sequence. To each, except the first and the last sign, is attached a characteristic random variable which takes the value 1 if both the neighbouring signs differ from the one between them and 0 otherwise. If $X_1, X_2, \ldots, X_n$ are characteristic random variables, find the mean and variance of $X = \sum\limits_{i=1}^{n} X_i$.*

Solution. $X = \sum\limits_{i=1}^{n} X_i \Rightarrow E(X) = \sum\limits_{i=1}^{n} E(X_i)$ ...(i)

Now $E(X_i) = 1.P(X_i = 1) + 0.P(X_i = 0) = P(X_i = 1)$ ...(ii)

For $X_i = 1$, there are the following two mutually exclusive possibilities :

(i) – + –, (ii) + – +

and since the probability of each sign is $\frac{1}{2}$, we have by addition probability theorem:

$P(X_i = 1) = P(i) + P(ii) = \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^3 = \frac{1}{4}$ ...(iii)

$\therefore$ $E(X_i) = \frac{1}{4}$ [From (ii)]. Hence $E(X) = \sum\limits_{i=1}^{n}\left(\frac{1}{4}\right) = \frac{n}{4}$, [From (i)]

$V(X) = V(X_1 + X_2 + \ldots + X_n) = \sum\limits_{i=1}^{n} V(X_i) + 2\sum\limits_{i=1}^{n}\sum\limits_{\substack{j=1 \\ i<j}}^{n} \text{Cov}(X_i, X_j)$ ...(*)

Now $E(X_i^2) = 1^2.P(X_i = 1) + 0^2.P(X_i = 0) = P(X_i = 1) = \frac{1}{4}$ [From (iii)]

$\therefore$ $V(X_i) = E(X_i^2) - \{E(X_i)\}^2 = \frac{1}{4} - \frac{1}{16} = \frac{3}{16}$

$E(X_i X_j) = 1.P(X_i = 1 \cap X_j = 1) + 0.P(X_i = 0 \cap X_j = 0)$
$+ 0.P(X_i = 1 \cap X_j = 0) + 0.P(X_i = 0 \cap X_j = 1)$

$= P(X_i = 1 \cap X_j = 1)$

Since there are following two mutually exclusive possibilities for the event :

$(X_i = 1 \cap X_j = 1)$, (i) $- \ + \ - \ +$ (ii) $+ \ - \ + \ -$, we have

$$P(X_i = 1 \cap X_j = 1) = P(i) + P(ii) = \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^4 = \frac{1}{8}$$

$$\therefore \quad \text{Cov}(X_i, X_j) = E(X_i X_j) - E(X_i)\, E(X_j) = \frac{1}{8} - \frac{1}{4} \times \frac{1}{4} = \frac{1}{16}$$

Hence
$$V(X) = \sum_{i=1}^{n} \frac{3}{16} + 2 \sum_{\substack{i=1 \\ i<j}}^{n} \sum_{j=1}^{n} \text{Cov}(X_i, X_j) \qquad \text{[From (*)]}$$

$$= \frac{3n}{16} + 2\{\text{Cov}(X_1, X_2) + \text{Cov}(X_2, X_3) + \ldots + \text{Cov}(X_{n-1}, X_n)\}$$

$$= \frac{3n}{16} + 2(n-1) \cdot \frac{1}{16} = \frac{5n-2}{16}$$

6·7. SOME INEQUALITIES INVOLVING EXPECTATION

Cauchy-Schwartz Inequality. *If X and Y are random variables taking real values, then*

$$[E(XY)]^2 \leq E(X^2) \cdot E(Y^2) \qquad \ldots (6\cdot33)$$

Proof. Let us consider a real valued function of the real variable t, defined by

$$Z(t) = E(X + tY)^2$$

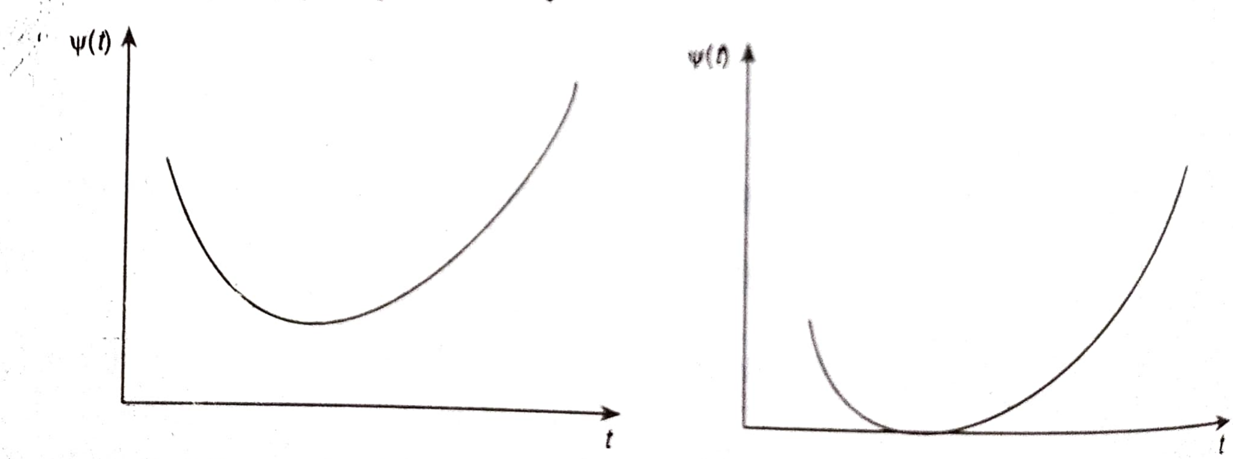
which is always non-negative, since $(X + tY)^2 \geq 0$, for all real X, Y and t.

Thus $\qquad Z(t) = E(X + tY)^2 \geq 0 \ \forall \ t.$

$\Rightarrow \qquad Z(t) = E[X^2 + 2tXY + t^2Y^2]$

$$= E(X^2) + 2t\, E(XY) + t^2 E(Y^2) \geq 0, \text{ for all } t. \qquad (*)$$

Obviously, $Z(t)$ is a quadratic expression in 't'.



We know that the quadratic expression of the form :

$\psi(t) = At^2 + Bt + C \geq 0$ for all t, implies that the graph of the function $\psi(t)$ either touches the t-axis at only one point or not at all, as exhibited in the above diagrams.

This is equivalent to saying that the discriminant of the function $\psi(t)$, *viz.*, $B^2 - 4AC \leq 0$, since the condition $B^2 - 4AC > 0$ implies that the function $\psi(t)$ has two distinct real roots which is a contradiction to the fact that $\psi(t)$ meets the t-axis either at only one point or not at all. Using this result, we get from (*),

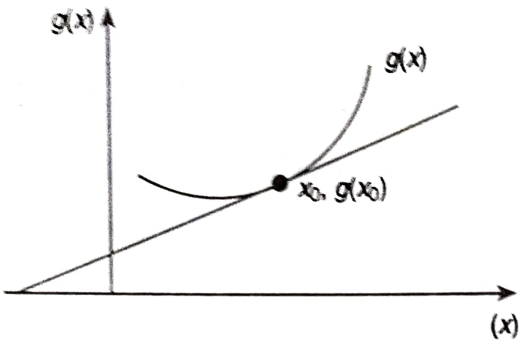
$$4 \cdot \{E(XY)\}^2 - 4 \cdot E(X^2) \cdot E(Y^2) \leq 0 \quad \Rightarrow \quad \{E(XY)\}^2 \leq E(X^2) \cdot E(Y^2)$$

Thus $E(X) \in I$. Now $E(X)$ can be either a left end or a right end point (if end points exist) of I or an interior point of I.

Suppose I has a left end point 'a', *i.e.*, $X \geq a$ and $E(X) = a$. Then $X - a \geq 0$ a.s. and $E(X - a) = 0$.

$$\text{Thus} \quad P(X = a) = 1 \quad \text{or} \quad P\{(X - a) = 0)\} = 1.$$

$$\therefore \quad E\{g(X)\} = E\{g(a)\} \qquad [\because g(x) = g(a) \text{ a.s.}]$$
$$= g(a) \qquad [\because g(a) \text{ is a constant}]$$
$$= g\{E(X)\}.$$



The result can be established similarly if I has a right end point 'b' and $E(X) = b$.

Thus we are now required to establish (6·36) when $E(X) = x_o$, is an interior point of I.

Let $ax + b$ pass through the point $\{x_o, g(x_o)\}$ and let it be below g.

[c.f. Remark 6 above].

$$\therefore \quad E\{g(X)\} \geq E(aX + b) = aE(X) + b$$
$$= ax_o + b = g(x_o) = g\{E(X)\}$$

Hence
$$E\{g(X)\} \geq g\{E(X)\}.$$

Continuous Concave Function (*Definition*). *A continuous function g is concave on an interval I if $(-g)$ is convex.*

Corollary to Jenson's Inequality. *If g is a continuous and concave function on the interval I and X is a r.v. whose values are in I with probability 1, then*

$$E\{g(X)\} \leq g\{E(X)\} \qquad \ldots (6·37)$$

provided the expectations exist.

Remarks 1. Equality holds in (6·36) and corollary (6·37) if the only if
$$P\{g(X) = aX + b\} = 1, \text{ for some } a \text{ and } b.$$

2. Jenson's inequality extends to random vectors. If I is a convex set in n dimensional Euclidean space, *i.e.*, the interval I in (6·36) is transferred to convex set, g is continuous on I, (6·34) holds whenever X_1 and X_2 are any arbitrary vectors in I. The condition $g''(x) \geq 0$ for x interior to I implies $\left(\dfrac{\partial^2 g(x)}{\partial x_i \partial x_j}\right) = M(x)$, (say), is non-negative definite for all x interior to I.

3. If $E(X^2)$ exists, then
$$E(X^2) \geq \{E(X)\}^2, \qquad \ldots (6·38)$$
since $g(X) = X^2$, is convex function of X as $g''(X) = 2 > 0$.

4. If $X > 0$, *i.e.*, X assumes only positive values and $E(X)$ and $E\left(\dfrac{1}{X}\right)$ exist then
$$E\left(\frac{1}{X}\right) \geq \frac{1}{E(X)} \qquad \ldots (6·38a)$$
because $g(X) = \dfrac{1}{X}$, is a convex function of X since $g''(X) = \dfrac{2}{X^3} > 0$ for $X > 0$.

5. If $X > 0$, *a.s.* then
$$E(X^{1/2}) \leq \{E(X)\}^{1/2} \qquad \ldots (6·38b)$$
since $g(X) = X^{1/2}$, $X > 0$ is a concave function as $g''(X) = -\frac{1}{4}X^{-3/2} < 0$, for $X > 0$.

6. If $X > 0$, *a.s.* then
$$E\{\log(X)\} \leq \log\{E(X)\}, \qquad \ldots (6·38c)$$
provided the expectations exist, because $\log X$ is a concave function of X.

Another Useful Inequality. *Let f and g be monotone functions on some subset of the real line and X be a r.v. whose range is in the subset almost surely (a.s.). If the expectations exist, then*

$$E[f(X)\,g(X)] \geq E[f(X)].\,E[g(X)] \qquad \ldots (6\cdot 39)$$

or

$$E[f(X)\,g(X)] \leq E[f(X)].\,E[g(X)] \qquad \ldots (6\cdot 39a)$$

according as f and g are monotone in the same or in the opposite directions.

Proof. Let us consider the case when both the functions f and g are monotone in the same direction. Let x and y lie in the domain of f and g respectively.

If f and g are both monotonically increasing, then

$$y \geq x \quad \Rightarrow \quad f(y) \geq f(x) \text{ and } g(y) \geq g(x)$$

$$\Rightarrow \quad f(y) - f(x) \geq 0 \quad \text{and} \quad g(y) - g(x) \geq 0$$

$$\Rightarrow \quad [f(y) - f(x)].\,[g(y) - g(x)] \geq 0 \qquad \ldots (*)$$

If f and g are both monotonically decreasing then for $y \geq x$, we have

$$f(y) \leq f(x) \quad \text{and} \quad g(y) \leq g(x)$$

$$\Rightarrow \quad f(y) - f(x) \leq 0 \quad \text{and} \quad g(y) - g(x) \leq 0$$

$$\therefore \quad [f(y) - f(x)].\,[g(y) - g(x)] \geq 0 \qquad \ldots (**)$$

Hence if f and g are both monotonic in the same direction, then from (*) and (**), we get the same result, *viz.*, $[f(y) - f(x)].\,[g(y) - g(x)] \geq 0$.

Let us now consider independently and identically distributed (*i.i.d.*) random variables X and Y. Then from above, we get

$$E\,[\{f(Y) - f(X)\}\,\{g(Y) - g(X)\}] \geq 0.$$

$$\Rightarrow \quad E\,[f(Y)\cdot g(Y)] - E\,[f(Y).\,g(X)] - E\,[f(X)\,g(Y)] + E\,[f(X).\,g(X)] \geq 0 \qquad \ldots (6\cdot 40)$$

Since X and Y are *i.i.d. r.v.'s*, we have

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

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

($\because$ X and Y are independent) ($\because$ X and Y are identical)

and $\qquad E[f(X)\,g(Y)] = E\,[f(X)].\,E\,[g(Y)] = E\,[f(X)].\,E\,[g(X)]$

Substituting in (6·40), we get

$$2E\,[f(X).\,g(X)] - 2\,E\,[f(X)].\,E\,[g(X)] \geq 0 \quad \Rightarrow \quad E\,[f(X)\cdot g(X)] \geq E\,[f(X)].\,E\,[g(X)],$$

which establishes the result in (6·39).

Similarly, (6·39*a*) can be established, if f and g are monotonic in opposite directions, *i.e.*, if f is monotonically increasing (decreasing) and g is monotonically decreasing (increasing). *The proof is left as an exercise to the reader.*

Remarks 1. If X is a r. *v.* which takes only non-negative values, *i.e.*, if $X \geq 0$ a.s. then for $\alpha > 0$, $\beta > 0$, $f(X) = X^{\alpha}$ and $g(X) = X^{\beta}$ are monotonic in the same direction. Hence if the expectations exist,

$$E\,(X^{\alpha}.\,X^{\beta}) \geq E\,(X^{\alpha}).\,E\,(X^{\beta}) \quad \Rightarrow \quad E\,(X^{\alpha+\beta}) \geq E\,(X^{\alpha})\,E\,(X^{\beta})\,;\ \alpha > 0,\ \beta > 0 \qquad \ldots (6\cdot 41)$$

In particular, taking $\alpha = \beta = 1$, $E\,(X^2) \geq \{E\,(X)\}^2$, a result already obtained in (6·38).

2. If $X \geq 0$, a.s. and $E\,(X^{\alpha})$ and $E\,(X^{-1})$ exist, then for $\alpha > 0$, we get from (6·39*a*)

$$E\,(X^{\alpha}.\,X^{-1}) \leq E\,(X^{\alpha}).\,E\,(X^{-1}) \quad \Rightarrow \quad E\,(X^{\alpha})\,E\left(\frac{1}{X}\right) \geq E\,(X^{\alpha-1})\,;\ \alpha > 0. \qquad \ldots (6\cdot 42)$$

In particular with $\alpha = 1$, $E\,(X)\,E\left(\frac{1}{X}\right) \geq 1$, a result already obtained in (6·38*a*).