CHAPTER 7

Mathematical Expectations

7.1 EXPECTATION

Let X be a random variable defined on a probability space $(\Omega, \mathcal{A}, P(\bullet))$, let F be the distribution function of X. Then the expectation of the random variable X is denoted by E(X) and is defined for continuous random variable as

$$E(X) = \int_{-\infty}^{\infty} x\, dF(x) = \int_{-\infty}^{\infty} x f(x)\, dx \quad \text{Provided} \int_{-\infty}^{\infty} f(x) = 1$$

The expectation is said to exist if and only if,

$$\int_{-\infty}^{\infty} |x|\, dF(x) < \infty$$

The range of the integral depends on the range in which the random variable is defined.

In case, we are interested in finding the expectation of a function of the random variable X, $g(x)$ (say), then we have

$$E(g(x)) = \int_{-\infty}^{\infty} g(x)\, dF(x) = \int_{-\infty}^{\infty} g(x) f(x)\, dx \qquad ...(1)$$

provided the expectation exists.

Let X be a discrete random variable, which takes the values $x_1, x_2, ...$ with corresponding probabilities $p_1, p_2, ...$ then the expectation of the random variable X is given by

$$E(X) = \sum_{i=1}^{\infty} x_i p_i \text{ provided } \sum_{i=1}^{\infty} p_i = 1$$

and $\sum_{i=1}^{\infty} |x_i| p_i < \infty$ is satisfied.

7.2 Some Results Based on Expectation

1. If C is a constant then E(C) = C

We have,

$$E(C) = \int_{-\infty}^{\infty} C\,dF(x) = C\int_{-\infty}^{\infty} f(x)\,dx = C \times 1 = C.$$

2. If C is a constant and X is a random variable then,

$$E(CX) = CE(X)$$

We have,

$$E(CX) = \int_{-\infty}^{\infty} Cx\,dF(x) = C\int_{-\infty}^{\infty} x\,dF(x)$$

$$= CE(X)$$

3. If C and a are constants and X is a random variable then,

$$E(CX + a) = CE(X) + a$$

We have,

$$E(CX + a) = \int_{-\infty}^{\infty} (CX + a)\,dF(x)$$

$$= C\int_{-\infty}^{\infty} x\,dF(x) + \int_{-\infty}^{\infty} a\,dF(x)$$

$$= C\int_{-\infty}^{\infty} x\,dF(x) + a\int_{-\infty}^{\infty} dF(x) = CE(X) + a.$$

4. If X and Y be two random variable and if a, b are two constants then,

$$E(aX + bY) = aE(X) + bE(Y)$$

Let $f(x, y)$ be the joint density of (X, Y).

So, we have, $E(aX + bY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (ax + by)\, f(x,y)\,dx\,dy$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ax\, f(x,y) + by\, f(x,y)\,dx\,dy$$

$$= a\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\, f(x,y)\,dx\,dy + b\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\, f(x,y)\,dx\,dy$$

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

$$= a\int_{-\infty}^{\infty} x f_1(x)dx + b\int_{-\infty}^{\infty} y f_2(y)dy$$

$$= aE(X) + bE(Y)$$

Thus, $E(aX + bY) = aE(X) + bE(Y)$

This result can be extended to a number of variable, *i.e.*,

$$E\left(\sum a_i X_i\right) = \sum a_i E(X_i)$$

Note : Here we extend the definition of expectation of single variate to two variables, *i.e.*, if $f(x, y)$ is a function involving x and y such that,

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

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

provided $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |\phi(x, y)|\, dF(x, y) < \infty$ and $f(x, y)$ is the joint pdf of (x, y) such that, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)\, dx\, dy = 1$

5. If X and Y are two random variable and if ϕ_1 (X) and ϕ_2 (Y) are two functions of X and Y respectively, then

$$E[\phi_1(X)\, \phi_2(Y)] = E[\phi_1(X)]\, E[\phi_2(Y)]$$

provided X and Y are independent.

Let $f(x, y)$ be the joint distribution of (X, Y). Now, if X and Y independent then we have,

$$f(x, y) = f_1(x) \cdot f_2(y)$$

Thus, $$E[\phi_1(X)\, \phi_2(Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_1(x)\phi_2(y) f(x, y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \phi_1(x)\phi_2(y) f_1(x)\, f_2(y)\, dx\, dy$$

$$= \int_{-\infty}^{\infty} \phi_1(x)\, f_1(x)\, dx \int_{-\infty}^{\infty} \phi_2(y)\, f_2(y)\, dy$$

$$= E[\phi_1(X)] \cdot E[\phi_2(Y)]$$

Note : (*i*) This can be extended to any number of functions, *i.e.*, $E\left[\prod \phi_i(X_i)\right] = \prod E[\phi_i(X_i)]$ provided X_i's are independent of each other.

(*ii*) In particular, if $\phi_1(X) = X$ and $\phi_2(Y) = Y$ then

We have,

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

6. If $g(x) = g_1(x) + g_2(x)$, then

$$E[g(X)] = E[g_1(X)] + E[g_2(X)]$$

provided the expectation exists.

We have, $$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, dF(x) = \int_{-\infty}^{\infty} [g_1(x) + g_2(x)]\, dF(x)$$

$$= \int_{-\infty}^{\infty} g_1(x)\, dF(x) + \int_{-\infty}^{\infty} g_2(x)\, dF(x)$$

$$= E[g_1(X)] + E[g_2(X)]$$

7. If $g_1(x) \le g_2(x)$ then $E[g_1(X)] \le E[g_2(X)]$ provided the expectation exists.

Now, $$E[g_1(X)] = \int_{-\infty}^{\infty} g_1(x)\, dF(x) \le \int_{-\infty}^{\infty} g_2(x)\, dF(x)$$

$$= E[g_2(X)]$$

Thus, $$E[g_1(X)] \le E[g_2(X)]$$

8. $$|E[g_1(X)]| \le E|g_1(X)|$$

We have, $$|E[g_1(X)]| = \left| \int_{-\infty}^{\infty} g_1(x)\, dF(x) \right| \le \int_{-\infty}^{\infty} |g_1(x)|\, dF(x)$$

$$= E[|g_1(X)|]$$

9. $E\left(\frac{1}{X}\right) \ge \frac{1}{E(X)}$, where X is a random variable.

We have, for any random variable X,

$$X = E(X)\left[1 + \frac{X - E(X)}{E(X)}\right]$$

$$= E(X)\,[1 + Z], \text{ where } Z = \frac{X - E(X)}{E(X)} \qquad ....(2)$$

Again, $$E(Z) = \frac{E(X) - E(X)}{E(X)} = 0$$

Now, $X = E(X)\,[1 + Z]$

So, $\dfrac{1}{X} = \dfrac{1}{E(X)[1+Z]}$

$$E\left[\frac{1}{X}\right] = E\left[\frac{1}{E(X)[1+Z]}\right] = \frac{1}{E(X)}\,E\left[\frac{1}{1+Z}\right] \quad ...(3)$$

Now, $\dfrac{1}{1+Z} = \dfrac{1-Z^2+Z^2}{1+Z} = 1 - Z + \dfrac{Z^2}{1+Z}$

So, $E\left[\dfrac{1}{1+Z}\right] = E\left[1 - Z + \dfrac{Z^2}{1+Z}\right]$

$$= 1 - E(Z) + E\left(\frac{Z^2}{1+Z}\right) = 1 + E\left(\frac{Z^2}{1+Z}\right) \quad ...(4)$$

Replacing (4) in (3), we have

$$E\left[\frac{1}{X}\right] = \frac{1}{E(X)}\left\{1 + E\left(\frac{Z^2}{1+Z}\right)\right\}$$

Now Z^2 and $1+Z$ both are positive,

So, $\dfrac{Z^2}{1+Z} \geq 0 \quad \Rightarrow \quad E\left(\dfrac{Z^2}{1+Z}\right) \geq 0$

So, $1 + E\left(\dfrac{Z^2}{1+Z}\right) \geq 1$

Thus, $E\left[\dfrac{1}{X}\right] \geq \dfrac{1}{E(X)}$

7.3 VARIANCE OF A RANDOM VARIABLE

An important characteristic of a random variable is its variance, which is a measure of dispersion of the random variable. The variance of a random variable may be defined as the expected value of the square of the deviation of the random variable from its expected value.

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

In case we are dealing with a function of the random variable, $\phi(X)$, we have

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

$$= E\left[\phi^2(X)+[E(\phi(X))]^2 /-2\phi(X)\, E[\phi(x)]\right]$$

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

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

In particular if $\phi(X) = X$ then we have,

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

We know that variance is always positive,

So, $\quad Var(X) \geq 0$

$\Rightarrow \quad E(X^2) - [E(X)]^2 \geq 0$

$\Rightarrow \quad E(X^2) \geq [E(X)]^2$

7.4 SOME RESULTS BASED ON VARIANCE

1. If $\phi(X)$ be a function of a random variable X, and if a is a constant, then

$$Var[\phi(X) + a] = Var[\phi(X)]$$

From the definition of variance we have,

$$Var[\phi(X) + a] = E[\phi(X) + a - E[\phi(X) + a]]^2$$

$$= E\left[\phi(X)+a-E\left[\phi(X)\right]-a\right]$$

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

$$= Var[\phi(X)]$$

2. If $\phi(X)$ be a function of random variable X, then for a constant b,

$$Var[b\,\phi(X)] = b^2\, Var(\phi(X))$$

We have $Var[b\,\phi(X)] = E[(b\,\phi(X) - b\,E(\phi(X)))^2]$

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

$$= b^2\, E[\phi(X) - E(\phi(X))]^2$$

$$= b^2\, Var[\phi(X)]$$

3. $Var[a\,\phi(X) + b] = a^2\, Var[\phi(X)]$

We have, $Var[a\,\phi(X) + b] = E[(a\,\phi(X) + b) - E(a\,\phi(X) + b)]^2$

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

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

$$= a^2\, Var(X)$$

4. Variance of Linear Combination of random variables.

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

then, $$Var\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2\, Var(X_i) + 2\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j\, Cov(X_i, Y_j)$$

Let, $U = a_1X_2 + a_2X_2 + \ldots + a_n X_n$

$$E(U) = [a_1X_1 + a_2X_2 + \ldots + a_nX_n]$$
$$= E(a_1 X_1) + E(a_2 X_2) + \ldots + E(a_n X_n)$$
$$= a_1 E(X_1) + a_2 E(X_2) + \ldots + a_n E(X_n)$$

So, $U - E(U) = a_1 [X_1 - E(X_1)] + a_1 [X_2 - E(X_2)] + \ldots + a_n [X_n - E(X_n)]$

Squaring both sides and taking expectation we have,

$$E[U - E(U)]^2 = a_1^2 E[X_1 - E(X_1)]^2 + a_2^2 E[X_2 - E(X_2)]^2 + \ldots$$
$$+ a_n^2 E[X_n - E(X_n)]^2 + 2a_1a_2 E[X_1 - E(X_1)][X_2 - E(X_2)] + 2a_1a_3 E[X_1 - E(X_1)]$$
$$[X_3 - E(X_3)] + \ldots + 2a_{n-1}a_n E[X_n - E(X_n)][X_{n-1} - E(X_{n-1})]$$

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

$$= \sum_{j=1}^{n} a_i^2 \text{ Var }(X_i) + 2\sum_{\substack{i=1 \\ i\neq j}}^{n}\sum_{j=1}^{n} a_i a_j \text{ Cov }(X_i, X_j)$$

Note :

(*i*) If $a_i = 1, \forall$ then.

$$\text{Var }(X_1 + X_2 + \ldots + X_n) = \text{Var }(X_1) + \text{Var }(X_2) + \ldots + \text{Var }(X_n) + 2\sum_{\substack{i=1 \\ i\neq j}}^{n}\sum_{j=1}^{n} \text{Cov }(X_i, X_j)$$

(*ii*) If $X_1, X_2, \ldots, X_n$ are pairwise independent, then $\text{Cov }(X_i, X_j) = 0\ \forall i \neq j$.

Thus we have,

$$\text{Var}\left[\sum a_i X_i\right] = a_1^2 \text{ Var }(X_1) + a_2^2 \text{ Var }(X_1) + \ldots + a_n^2 \text{ Var }(X_n)$$
$$= \sum_{j=1}^{n} a_i^2 \text{ Var }(X_i)$$

(*iii*) If $a_1 = 1, a_2 = 1$, and $a_3 = a_4 = \ldots = a_n = 0$

Then, $\text{Var }(X_1 + X_2) = \text{Var }(X_1) + \text{Var }(X_2) + 2 \text{ Cov }(X_1, X_2)$

(*iv*) Again if $a_1 = 1, a_2 = -1, a_3 = a_4 = \ldots = a_n = 0$

Then, $\text{Var }(X_1 - X_2) = \text{Var }(X_1) + \text{Var }(X_2) - 2 \text{ Cov }(X_1, X_2)$

(*v*) If X_1 and X_2 are independent and $a_1 = a_2 = 1$ and $a_3 = a_4 = \ldots = a_n = 0$

hen,

$$\text{Var }(X_1 \pm X_2) = \text{Var }(X_1) + \text{Var }(X_2)$$

Theorem 1.1 : If X and Y are independent random variables, then

$\text{Var}(XY) = \text{Var}(X)\,\text{Var}(Y) + [E(Y)]^2\,\text{Var}(X) + [E(X)]^2\,\text{Var}(Y)$

Proof : From the formula for variance we have,

$$\text{Var}(XY) = E[(XY)^2] - [E(XY)]^2$$

$$= E[X^2Y^2] - [E(X)\,E(Y)]^2 \qquad [\because X, Y \text{ are independent}]$$

$$= E(X^2)\,E(Y^2) - [E(X)]^2\,[E(Y)]^2 \qquad [\because X^2 \text{ and } Y^2 \text{ are also independent}]$$

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

$$= \text{Var}(X)\,\text{Var}(Y) + [E(Y)]^2\,\text{Var}(X) + [E(X)]^2\,\text{Var}(Y).$$

7.5 SOME FORMULAE RELATED TO CONTINUOUS PROBABILITY DISTRIBUTIONS

(i) $P[c \le X \le d] = \int_c^d f(x)\,dx$

(ii) The mean of the random variable μ_1', is given by

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

(iii) The rth raw moment is given by

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

(iv) The rth central moment is given by

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

(v) The geometric mean G is given by

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

(vi) The harmonic mean, H is given by

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

(*vii*) Mean deviation about μ is given by

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

(*viii*) The median m is given by

$$\int_{-\infty}^{m} f(x)\,dx = \int_{m}^{\infty} f(x)\,dx = \frac{1}{2}$$

(*ix*) The lower quartile Q_1 and the upper quartile Q_3 is given by

$$\int_{-\infty}^{Q_1} f(x)\,dx = \frac{1}{4} \text{ and } \int_{Q_3}^{\infty} f(x)\,dx = \frac{1}{4}$$

(*x*) The mode of a distribution $f(x)$ is obtained from the solution of the equation

$f'(x) = 0$ provided $f''(x) < 0$ for the value of x obtained from $f'(x) = 0$.

Illustration 7.1 : At a party n people put their hats in the center of a room where the hats are mixed together. Each person than randomly selects one. We are interested to know the mean and variance of X the number that select their own hat.

Solution : Let $X = X_1 + X_2 + \ldots + X_n$

where, $X_i = 1$, if the i^{th} person selects his/her own hat.

$= 0$ otherwise

So, $$P[X_i = 1] = \frac{1}{n}$$

Thus, $$E[X_i] = 1.P[X_i = 1] + 0 \,.\, P[X_i = 0] = \frac{1}{n}$$

$$\text{Var}(X_i) = E\left[X_i^2\right] - \left[E(X_i)\right]^2$$

Also, $$E\left[X_i^2\right] = 1^2 \,.\, P[X_i = 1] + 0^2 \,.\, P[X_i = 1] = \frac{1}{n}$$

Thus, $$V(X_i) = E\left(X_i^2\right) - \left[E(X_i)\right]^2 = \frac{1}{n} - \left(\frac{1}{n}\right)^2 = \frac{n-1}{n^2}$$

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

Now, $X_iX_j = 1$, if the i^{th} and j^{th} person both selects their own hats

$= 0$, otherwise.

and thus, $$E(X_iX_j) = P[X_i = 1, X_j = 1]$$
$$= P[X_i = 1] \,.\, P[X_j = 1 | X_i = 1]$$
$$= \frac{1}{n} \times \frac{1}{n-1} = \frac{1}{n(n-1)}$$

Thus, $$\text{Cov}(X_i, X_j) = \frac{1}{n(n-1)} - \left(\frac{1}{n}\right)^2 = \frac{n-(n-1)}{n^2(n-1)} = \frac{1}{n^2(n-1)}$$

Now, $$E(X) = \sum E(X_i) = \sum \frac{1}{n} = n \times \frac{1}{n} = 1$$

$$\text{Var}(X) = \sum \text{Var}(X_i) + 2\sum_{i \neq j} \text{Cov}(X_i, X_j)$$

$$= n.\frac{n-1}{n^2} + 2\frac{1}{n^2(n-1)}\,{}^{n}C_2$$

$$= \frac{n-1}{n} + \frac{1}{n} = \frac{n}{n} = 1$$

Illustration 7.2 : An urn contains a white and b black balls, c balls are drawn out of it without replacement. Determine the mathematical expectation of the number of white balls drawn.

Solution : Let us define a random variable X_i such that,

$$X_i = 1 \text{ if the } i^{th} \text{ ball is white,}$$
$$= 0 \text{ if the } i^{th} \text{ ball is black}$$

with, $i = 1, 2, \ldots, c$.

Now, $$P[X_i = 1] = \frac{a}{a+b},\; P[X_i = 0] = \frac{b}{a+b}$$

$$E[X_i] = 1.\frac{a}{a+b} + 0.\frac{b}{a+b} = \frac{a}{a+b}$$

The number of r white balls among the c drawn is given by,

$$r = X_1 + X_2 + \ldots + X_c.$$

So, $$E[r] = E[X_1] + E[X_2] + \ldots + E[X_c] = \frac{ac}{a+b}$$

Illustration 7.3 : Suppose r balls are drawn one at a time without replacement, from a bag containing n white and m black balls. Find the expected value and the variance of the number of black balls drawn.

Solution : Let the variable X_k be defined in the following manner :

$$X_k = 1 \text{ if the } k\text{th ball is black}$$
$$= 0 \text{ if the } k\text{th ball is white}$$

$$P[X_k = 1] = \frac{m}{m+n},\; P[X_k = 0] = \frac{n}{m+n}$$

$\therefore$ $$E[X_k] = \frac{m}{m+n} = E\left[X_k^2\right]$$

$$\text{Var}[X_k] = E\left[X_k^2\right] - [E(X_k)]^2 = \frac{m}{m+n} - \frac{m^2}{(m+n)^2}$$

$$= \frac{mn}{(m+n)^2}$$

Now, if $j \neq k$ then $X_jX_k = 1$, if the jth and kth balls drawn are black, otherwise $X_jX_k = 0$.

So, $$P[X_jX_k = 1] = \frac{m(m-1)}{(m+n)(m+n-1)}$$

Thus, $$E[X_jX_k] = \frac{m(m-1)}{(m+n)(m+n-1)}$$

So, $$\text{Cov}\,(X_jX_k) = E\,[X_jX_k] - E[X_j]\,E[X_k]$$

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

$$= \frac{mn}{(m+n)^2(m+n-1)}$$

$\therefore$ $$E[S_r] = E[X_1 + X_2 + \ldots + X_r]$$
$$= E[X_1] + E[X_2] + \ldots + E[X_r]$$
$$= \frac{mr}{m+n}$$

$$\text{Var}\,[S_r] = \sum \text{Var}(X_k) + 2\sum_{j \neq k} \text{Cov}(X_j, X_k)$$

$$= r.\,\frac{mn}{(m+n)^2} - r(r-1)\,\frac{mn}{(m+n)^2(m+n-1)}$$

$$= \frac{mnr\,(m+n-r)}{(m+n)^2(m+n-1)}$$

Illustration 7.4 : A man with n keys wants to open door and tries the keys independently and at random. Find the expectation and variane of the number of trials (*a*) if unsuccessful keys are not eliminated from further selection (*b*) if they are.

Solution : Suppose the man succeeds to open his door at the rth trial. In other words he fails in the $(r-1)$ trials and succeeds at the rth trial.

(*a*) Probability of success at the first trial $= \frac{1}{n}$.

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

Probability of success at the 3rd trial $= \frac{n-1}{n} \times \frac{n-1}{n} \times \frac{1}{n} = \left(\frac{n-1}{n}\right)^2 \cdot \frac{1}{n}$.

In this case there will be an infinite number of trials as unsuccessful keys are not eliminated.

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

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

$$= \frac{1}{n}\left[1 - \frac{n-1}{n}\right]^{-2}$$

$$= \frac{1}{n} \times n^2 = n.$$

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

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

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

$$= \frac{1}{n}.\frac{2n-1}{n}.n^3 - n^2$$

$$= 2n^2 - n - n^2$$

$$= n(n-1)$$

(*b*) Probability of success at the first trial $= \frac{1}{n}$.

Probability of failure $= 1 - \frac{1}{n} = \frac{n-1}{n}$.

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

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

So, probability of success at the nth trial $= \frac{1}{n}$.

Thus, $E(X) = 1.\frac{1}{n} + 2.\frac{1}{n} + \ldots + n.\frac{1}{n} = \frac{n+1}{2}$

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

$$= \left[1^2.\frac{1}{n} + 2^2.\frac{1}{n} + \ldots + n^2.\frac{1}{n}\right] - \left[\frac{n+1}{2}\right]^2$$

$$= \frac{1}{n}\frac{n(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4}$$

$$= \frac{n^2-1}{12}$$

Illustration 7.5 : Obtain the rth central moment in terms of survival function.

Solution : We know that, the survival function is given by, $S(x) = 1 - F(x)$.

We have,
$$E[X^r] = \int_0^\infty x^r f(x)\,dx = \int_0^\infty x^r\, dF(x)$$

$$= -\int_0^\infty x^r\, d[1-F(x)]$$

$$= -x^r\,[1-F(x)]\big]_0^\infty + \int_0^\infty r x^{r-1}[1-F(x)]\,dx$$

$$= \int_0^\infty r x^{r-1}\, S(x)\,dx$$

Illustration 7.6 : If X is a random variable with mean μ and variance σ^2, then show that,

$$F(x-\mu) \le \frac{1}{1+\left(\frac{x-\mu}{\sigma}\right)} \quad \text{if } x < \mu$$

$$\ge \frac{1}{1+\frac{1}{\left(\frac{x-\mu}{\sigma}\right)^2}} \quad \text{if } x \ge \mu$$

Solution : Let Y be a random variable with mean 0 and variance σ^2. Now let us consider

$$\int_{-\infty}^{\infty} (y-x)\,d\{F(x)\} = \int_{-\infty}^{\infty} y\,d\{F(y)\} - \int_{-\infty}^{\infty} x\,d\{F(y)\}$$

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

$$= -x, \qquad \text{.....(1)}$$

Now, $\int_{-\infty}^{\infty} (y-x)\,d\{F(y)\} = \int_{-\infty}^{x} (y-x)\,d\{F(y)\} + \int_{x}^{\infty} (y-x)\,d\{F(y)\}$

Let $y \ge x$

So, the first integral is negative and the second integral is positive and LHS is also +ve.

So, we have,

+ ve quantity = (– ve quantity) + (+ ve quantity)

Thus,

LHS ≤ 2nd integral

$$\int_{-\infty}^{\infty} (y-x)\, d\{F(y)\} \le \int_{x}^{\infty} (y-x)\, dF(y)$$

$$\Rightarrow \quad -x \le \int_{x}^{\infty} (y-x)\, dF(y) \quad \text{using (1)}$$

$$\Rightarrow \quad x^2 \le \left[\int_{x}^{\infty} (y-x)\, dF(y)\right]^2 \quad \text{using (1)} \qquad ....(2)$$

Now, by Cauchy – Schwartz's inequality,

$$\int_{x}^{\infty} (y-x)^2\, d\{F(y)\} \times \int_{x}^{\infty} (1)^2\, d\{F(y)\} \ge \left[\int_{x}^{\infty} (y-x)\, d\{F(y)\}\right]^2$$

$$\Rightarrow \int_{x}^{\infty} (y-x)^2\, d\{F(y)\} \times [F(y)]_x^{\infty} \ge \left[\int_{x}^{\infty} (y-x)\, d\{F(y)\}\right]^2$$

$$\Rightarrow \int_{x}^{\infty} (y-x)^2\, d\{F(y)\}\, [1-F(x)] \ge \left[\int_{x}^{\infty} (y-x)\, d\{F(y)\}\right]^2$$

$$\Rightarrow \int_{x}^{\infty} (y-x)^2\, d\{F(y)\}[1-F(x)] \ge x^2 \qquad \text{[using (2)]}$$

$$\text{So,} \int_{x}^{\infty} (y-x)^2\, d\{F(y)\} \ge \frac{x^2}{[1-F(x)]}$$

$$\Rightarrow \int_{-\infty}^{\infty} (y-x)^2\, d\{F(y)\} \ge \frac{x^2}{[1-F(x)]}$$

$$\Rightarrow \int_{-\infty}^{\infty} (x^2 + y^2 - 2xy)\, d\{F(y)\} \ge \frac{x^2}{[1-F(x)]}$$

$$\Rightarrow \int_{-\infty}^{\infty} x^2\, d\{F(y)\} + \int_{-\infty}^{\infty} y^2\, d\{F(y)\} - 2x \int_{-\infty}^{\infty} y\, d\{F(y)\} \geq \frac{x^2}{[1-F(x)]}$$

$$\Rightarrow \int_{-\infty}^{\infty} y^2 f(y)\, dy + x^2 - 2x \int_{-\infty}^{\infty} y f(y)\, dy \geq \frac{x^2}{[1-F(x)]}$$

$$\Rightarrow \sigma^2 + x^2 \geq \frac{x^2}{[1-F(x)]}$$

$$\Rightarrow 1 - F(x) \geq \frac{x^2}{\sigma^2 + x^2}$$

$$\Rightarrow F(x) \geq 1 - \frac{x^2}{\sigma^2 + x^2} = \frac{\sigma^2 + x^2 - x^2}{\sigma^2 + x^2} = \frac{\sigma^2}{\sigma^2 + x^2}$$

So, $$F(x) \geq \frac{\sigma^2}{\sigma^2 + x^2}$$

Changing the origin at μ we have,

$$F(x-\mu) \geq \frac{\sigma^2}{(x-\mu)^2 + \sigma^2}$$

So, $$F(x-\mu) \geq \frac{1}{1 + \left(\frac{x-\mu}{\sigma}\right)^2} \text{ for } x \leq \mu$$

Similarly, it can be shown that,

$$F(x-\mu) \leq \frac{1}{1 + \frac{1}{\left(\frac{x-\mu}{\sigma}\right)^2}} \text{ for } x \leq \mu.$$

Illustration 7.7 : Let X be a random variable with pdf

$$f(x) = C\,(1 - x^2),\ 0 < x < 1$$

Then find

(*i*) C (*ii*) E(X) (*iii*) Median of X (*iv*) $P\left[\frac{1}{2} < X < \frac{3}{4}\right]$

Solution : (*i*) Since $f(x)$ is a pdf so we have,

$$\int_0^1 f(x)dx = 1 \quad \Rightarrow \int_0^1 C(1-x^2)dx = 1$$

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

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

$\Rightarrow$

$$C.\frac{2}{3}=1 \quad \Rightarrow C=\frac{3}{2}$$

$\Rightarrow$

Thus, $$f(x)=\frac{3}{2}(1-x^2),\ 0<x<1.$$

(ii) $$E(X)=\int_0^1 x\,f(x)\,dx=\int_0^1 x\,\frac{3}{2}(1-x^2)\,dx$$

$$=\frac{3}{2}\int_0^1 [x-x^3]\,dx=\frac{3}{2}\left[\frac{x^2}{2}-\frac{x^4}{4}\right]_0^1$$

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

(iii) Let m be the median of the distribution.

So, $$\int_0^m f(x)\,dx=\frac{1}{2}\Rightarrow\int_0^m \frac{3}{2}(1-x^2)\,dx=\frac{1}{2}$$

$$\Rightarrow \quad 3\int_0^m (1-x^2)\,dx=1\Rightarrow\left[x-\frac{x^3}{3}\right]_0^m=\frac{1}{3}$$

$$\Rightarrow \quad m-\frac{m^3}{3}=\frac{1}{3}$$

$$\Rightarrow \quad 3m-m^3=1$$

$$\Rightarrow \quad m^3-3m+1=0$$

Since the equation cannot be solved directly so we can use any method of numerical analysis and find an approximate root of the equation lying between [0, 1]. We find $m = 0.34375$ (approx).

(iv) $$P\left[\frac{1}{2}<X<\frac{3}{4}\right]=\int_{1/2}^{3/4} f(x)\,dx$$

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

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

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

$$= \frac{3}{2}\left[\frac{3}{4}-\frac{9}{64}-\frac{1}{2}+\frac{1}{24}\right]$$

$$= \frac{3}{2}\,[0.75 - 0.1406 - 0.5 + 0.0417]$$

$$= 0.22665$$

Illustration 7.8 : If the possible values of a variate X are 0, 1, 2, 3, ... then

$$E(X) = \sum P(X > n),\ n = 0, 1, 2, ...$$

Solution : We assume that the series $\sum P(X > n)$ is convergent.

Here we have, $E(X) = \sum_{n=0}^{\infty} n.P_n[X = n]$

Now, using $P[X = k] = p_k$, we observe that,

$$P[X > 0] = P[X = 1] + P[X = 2] + ...$$
$$= p_1 + p_2 + p_3 + ...$$
$$P[X > 1] = P[X = 2] + P[X = 3] + ...$$
$$= p_2 + p_3 + ...$$
$$P[X > 2] = P[X = 3] + P[X = 4] + ...$$
$$= p_3 + p_4 + ...$$

Adding all these terms, we have,

$$\sum P[X > n] = P[X > 0] + P[X > 1] + P[X > 2] ....$$
$$= p_1 + 2p_2 + 3p_3 + 4p_4 + ...$$
$$= \sum_{n=0}^{\infty} n\, p_n = E[X]$$

So, $$E[X] = \sum P[X > n]$$

Illustration 7.9 : With the help of an example show that, E(XY) = E(X).E(Y) even though X and Y are not independent.

Solution : Let us consider a case where Y is dependent on X, *i.e.*, $Y = X^2$.

Also, X follows the density function,

$$f(x) = \frac{1}{2};\ -1 \le x \le 1$$

So, $$E(Y) = E(X^2) = \int_{-1}^{1} x^2 f(x)dx = \int_{-1}^{1} \frac{x^2}{2} dx$$

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

So, $$E(XY) = \int_{-1}^{1} \frac{x^3}{2} dx = \frac{1}{2}\left[\frac{x^4}{4}\right]_{-1}^{1}$$

$$= \frac{1}{2}\left[\frac{1}{4} - \frac{1}{4}\right] = 0$$

and $$E(X) = \frac{1}{2}\int_{-1}^{1} x\,dx = \frac{1}{2}\left[\frac{x^2}{2}\right]_{-1}^{1} = 0$$

Thus, $$E(XY) = 0$$

$$E(X)\,E(Y) = \frac{1}{3} \times 0 = 0$$

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

But $Y = X^2$ implying Y is dependent on X.

Illustration 7.10 : Calculate the expectation of the following functions.

(a) $f(x) = \frac{1}{2\sqrt{x}}, 0 < x < 1$

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

(c) $f(x) = 1 - |1 - x|, 0 \le x \le 2$

(d) $f(x) = \frac{(n+m-1)!}{(n-1)!\,(m-1)!} x^{n-1} (1-x)^{m-1}, 0 < x < 1$

Solution : (a) Given $f(x) = \frac{1}{2\sqrt{x}}, \quad 0 < x < 1$

$$E(X) = \int_0^1 x \frac{1}{2\sqrt{x}} dx$$

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

$$= \frac{1}{2} \cdot \left.\frac{x^{\frac{1}{2}+1}}{\frac{1}{2}+1}\right|_0^1 = \frac{1}{3}\cdot\left[x^{3/2}\right]_0^1$$

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

(*b*) Given $f(x) = \frac{1}{2}x^2 e^{-x}$ $\quad 0 < x < \infty$

$$E(X) = \int_0^\infty x.\frac{1}{2}x^2 e^{-x}\,dx$$

$$= \frac{1}{2}\int_0^\infty x^3 e^{-x}\,dx$$

$$= \frac{1}{2}\left(x^3.\int e^{-x}\,dx - \int 3x^2.\left[\int e^{-x}\,dx\right]dx\right)$$

$$= \frac{1}{2}\left[x^3.\frac{e^{-x}}{-1} - \int 3x^2\left[\frac{e^{-x}}{-1}\right]dx\right]_0^\infty$$

$$= \frac{1}{2}\left[-x^3e^{-x} - 3x^2e^{-x} + 6\int xe^{-x}\,dx\right]_0^\infty$$

$$= \frac{1}{2}\left[-x^3e^{-x} - 3x^2e^{-x} + 6\left[x.\frac{e^{-x}}{-1} - \int\frac{e^{-x}}{-1}dx\right]\right]_0^\infty$$

$$= \frac{1}{2}\left[-x^3e^{-x} - 3x^2e^{-x} - 6xe^{-x} - 6.e^{-x}\right]_0^\infty$$

$$= \frac{1}{2}\,[0 + 6 \times 1]$$

$$= \frac{6}{2} = 3.$$

(*c*) Given $f(x) = 1 - |1 - x| = 0 \le x \le 2$

$\therefore$

$$E(X) = \int_0^2 xf(x)dx$$

$$= \int_0^2 x.[1-|1-x|]dx = \int_0^2 x.[1-|1-x|]dx$$

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

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

(d) Given $f(x) = \dfrac{\lfloor n+m-1}{\lfloor n-1 \lfloor m-1} x^{n-1}(1-x)^{m-1} \qquad 0 < x < 1$

$$\therefore \quad E(X) = \int_0^1 x.\frac{\lfloor n+m-1}{\lfloor n-1 \lfloor m-1} x^{n-1}(1-x)^{m-1}\,dx$$

$$= \frac{\lfloor n+m-1}{\lfloor n-1 \lfloor m-1} \int_0^1 x^n (1-x)^{m-1}\,dx$$

$$= \frac{\lfloor n+m-1}{\lfloor n-1 \lfloor m-1} \beta(n+1, m)$$

$$= \frac{\lfloor n+m-1}{\lfloor n-1 \lfloor m-1} . \frac{\Gamma n+1 \Gamma m}{\Gamma n+m+1}$$

$$= \frac{\lfloor n+m-1}{\lfloor n-1 \lfloor m-1} . \frac{\lfloor n \lfloor m-1}{\lfloor n+m}$$

$$= \frac{n}{n+m}$$

Illustration 7.11 : Let the p.m.f. of a variate X be given by,

$$P(X = n) = \frac{6}{\pi^2 n^2},\ n = 1, -2, 3, -4, \ldots..$$

Does E(X) exist? Explain

Solution : Given that the p.m.f. is

$$P(X = n) = \frac{6}{\pi^2 n^2},\ n = 1, -2, 3, -4, \ldots.$$

$$\therefore \quad E(X) = \sum_x x.p(x)$$

$$= \sum_{x=n} (x = n).P(x = n)$$

$$= \sum_{(x=n)} (x = n) \times \frac{6}{\pi^2 n^2}$$

$$= \frac{6}{\pi^2}\left\{1.\frac{1}{1^2} + (-2).\frac{1}{(-2)^2} + 3.\frac{1}{3^2} + (-4).\frac{1}{(-4)^2} + \ldots.\right\}$$

$$= \frac{6}{\pi^2}\left\{1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+....\right\}$$

$$= \frac{6}{\pi^2}\log_e 2$$

$\therefore$ E(X) exists

Illustration 7.12 : Let variate X have the distribution,

$$P(X = 0) = P(X = 2) = p;$$

$$P(X = 1) = 1 - 2p, \quad \text{for } 0 \le p \le \frac{1}{2}$$

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

Solution : Here the random variables X has the following probability distribution —

X	:	0	1	2
P (x)	:	p	$1 - 2p$	p

$$\therefore \quad E(X) = \sum_x x \,.\, p(x) = 0 \times p + 1 \times (1 - 2p) + 2p = 1$$

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

$$= 1 + 2p$$

$$\therefore \quad V(X) = E(X^2) - \{E(X)\}^2$$

$$= 1 + 2p - 1$$

$$= 2p$$

Here maximum value of p is $\frac{1}{2}$, therefore V(X) is maximum at $p = \frac{1}{2}$

$$\max\{V(X)\} = 2 \times \frac{1}{2}$$

$$= 1$$

Illustration 7.13 : A pack of cards, n in number are numbered serially and put into a random order such that each of the arrangements have equal probabilities. A random variable X_k is defined such that, $X_k = 1$ if the kth card is in its natural position otherwise it is zero. Compute $E(X_k)$, Var (X_k) and Cov (X_h, X_k). Also if $X = X_1 + ... + X_n$ denote the number of matches then show that E(X) = 1 and Var (X) = 1.

Solution : We have,

$X_k = 1$ if the kth card is in the kth position, with probability $\frac{1}{n}$.

$= 0$ if the kth card is in any other position with probability $\frac{n-1}{n}$.

(i) Find the expectation of X

(ii) Find the median of X

(iii) Find the mode of X.

Solution : We have, for the p.d.f $f(x)$ as,

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

$$= \frac{d}{dx}\left[\frac{1-\cos x}{2}\right]$$

$$= \left[0 - \frac{1}{2}.(-\sin x)\right]$$

$$= \begin{cases} \frac{\sin x}{2}, & 0 \le x < \pi \\ 0, & \text{otherwise} \end{cases}$$

$$\therefore \quad E(X) = \int_0^\pi x \frac{\sin x}{2}\, dx$$

$$= \frac{1}{2}\left[x \int \sin x\, dx - \int\left\{\frac{d}{dx}x \int \sin x\, dx\right\}dx\right]_0^\pi$$

$$= \frac{1}{2}\left[x(-\cos x) - \int 1 \times (-\cos x)\, dx\right]_0^\pi$$

$$= \frac{1}{2}\left[-\pi \cos \pi + \sin \pi + 0 - \sin 0\right]$$

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

Again, $$\int_0^M f(x)\, dx = \frac{1}{2}$$

$$\Rightarrow \quad \int_0^M \frac{\sin x}{2}\, dx = \frac{1}{2}$$

$$\Rightarrow \quad \frac{1}{2}\int_0^M \sin x\, dx = \frac{1}{2}$$

$$\Rightarrow \quad -\cos x\Big]_0^M = 1$$

$$\Rightarrow \quad -\cos M + \cos 0 = 1 \qquad [\because \cos 0 = 1]$$

$$\Rightarrow \quad -\cos M = 0$$

$$\Rightarrow \quad \cos M = 0 = \cos \pi/2$$

$$\Rightarrow \quad M = \pi/2$$

∴ The median of the distribution is at $\pi/2$

(iii) If x is the modal value, then we have

$$f'(x) = 0$$

and $f''(x) < 0$, provided x lies in $[a, b]$

Thus, we have, $f''(x) = 0$

$$\frac{d}{dx}\left(\frac{\sin x}{2}\right) = 0$$

$$\Rightarrow \quad \frac{1}{2}\cos x = 0$$

$$\Rightarrow \quad \cos x = 0$$

$$\Rightarrow \quad x = \pi/2$$

$$\Rightarrow \quad f''(x) < 0$$

Also,

We have, $\frac{d}{dx}\left[\frac{d}{dx}f(x)\right]$

$$= \frac{d}{dx}\left(\frac{\cos x}{2}\right)$$

$$= \frac{1}{2}(-\sin x)$$

which is less than zero at $x = \pi/2$

∴ Modal value of the distribution is at $\pi/2$.

Illustration 7.15 : The distribution function of a variate X is as follows :

$$F(x) = 0 \text{ if } x < 0$$

$$= \frac{x}{8} \qquad \text{if } 0 \le x < 2$$

$$= \frac{x^2}{16} \qquad \text{if } 2 \le x < 4$$

$$= 1 \qquad \text{if } x \ge 4$$

Find E(X) and Var (X).

Solution : The p.d.f, $f(x) = \frac{d}{dx} F(x)$

$$f(x) = 0 \text{ if } x < 0$$

$$= \frac{1}{8} \qquad \text{if } 0 \le x < 2$$

$$= \frac{x}{8} \qquad \text{if } 2 \le x < 4$$

$$= 0 \qquad \text{if } x \ge 4$$

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

$$= \int_0^2 x.\frac{1}{8}\,dx + \int_2^4 x.\frac{x}{8}\,dx$$

$$= \frac{1}{8}\left(\frac{x^2}{2}\right)\Bigg|_0^2 + \frac{1}{8}.\left(\frac{x^3}{3}\right)\Bigg|_2^4$$

$$= \frac{1}{8}\times\frac{4}{2}+\frac{1}{8}\left(\frac{64}{3}-\frac{8}{3}\right)$$

$$= \frac{1}{4}+\frac{1}{24}\times 56$$

$$= \frac{1}{4}+\frac{7}{3}$$

$$= \frac{31}{12}$$

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

$$= \int_0^2 x^2.\frac{1}{8}dx + \int_2^4 x^2.\frac{x}{8}\,dx$$

$$= \frac{1}{8}\left(\frac{x^3}{3}\right)\Bigg|_0^2 + \frac{1}{8}.\left(\frac{x^4}{4}\right)\Bigg|_2^4$$

$$= \frac{1}{24}\times 8+\frac{1}{32}\,(256-16)$$

$$= \frac{1}{3}+\frac{1}{32}\times 240 = \frac{1}{3}+\frac{15}{2}=\frac{47}{6}$$

Now, $V(X) = E(X^2) - E^2(X)$

$$= \frac{47}{6}-\left(\frac{31}{12}\right)^2$$

$$= \frac{167}{144}$$

Illustration 7.16 : A jar has n chips numbered 1, 2, ..., n. A person draws a chip, returns it, draws another, returns it and so on, until a chip is drawn that has been drawn before. Let X be the number of drawings. Find E(X).

Solution : Here let us define D_r as the rth draw. Now X, the number of drawings, as defined in the question will be greater than one, *i.e.*,

$$P(X > 1) = 1$$

$$P(X > 2) = P\,[\text{Distinct numbers on } D_1 \text{ and } D_2]$$

$$= \frac{n}{n}\cdot\frac{n-1}{n} = 1-\frac{1}{n}$$

$P(X > 3) = P[\text{Distinct numbers on } D_1, D_2 \text{ and } D_3]$

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

$$P(X > k) = \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-1}{n}\right), k = 1, 2, 3, \ldots$$

So,
$$P[X = k] = P[X > k-1] - P[X = k]$$

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

$$= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-2}{n}\right)\left[1-1+\frac{k-1}{n}\right]$$

$$= \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{k-2}{n}\right)\left[\frac{k-1}{n}\right], k = 2, 3, \ldots$$

For non-negative integer valued variates,

$$E(X) = \sum_{k=0}^{\infty} P(X > k)$$

$$= P(X > 0) + P(X > 1) + P(X > 2) + \ldots$$

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

Illustration 7.17 : Show that, $E(X) = \int_0^{\infty} [1-F(x)]dx$.

Solution : We have,

$$E(X) = \int_0^{\infty} x f(x)dx = \int_0^{\infty} x\,dF(x)$$

$$= \lim_{\eta\to\infty} \int_0^{\eta} x\,dF(x) \qquad \ldots(1)$$

Now, $\int_0^{\eta} x\,dF(x) = xF(x)\big|_0^{\eta} - \int_0^{\eta} F(x)dx$

$$= \eta F(\eta) - \int_0^{\eta} F(x)dx + \eta - \eta$$

$$= \eta F(\eta) + \int_0^{\eta} dx - \int_0^{\eta} F(x)\, dx - \eta$$

$$= -\eta[1-F(\eta)] + \int_0^{\eta} [1-F(x)]\, dx \qquad \text{...(2)}$$

Now, $\lim\limits_{\eta\to\infty} \eta[1-F(\eta)] = \lim\limits_{\eta\to\infty} \eta \int_0^{\eta} d[F(x)] \le \lim\limits_{\eta\to\infty} \eta \int_{\eta}^{\infty} x\, dF(x) = 0$

$\Rightarrow$ $\lim\limits_{\eta\to\infty} \eta[1-F(\eta)] \le 0 \qquad \text{...(3)}$

Now, $\eta \ge 0$ and $[1-F(\eta)] \ge 0$

So, $\eta[1-F(\eta)] \ge 0$

$\Rightarrow$ $\lim\limits_{\eta\to\infty} \eta[1-F(\eta)] \ge 0$

So, comparing (3) and (4), we have

$$\mathop{\mathrm{Lt}}_{\eta\to\infty} \eta[1-F(\eta)] = 0$$

So, using (1), (2) and (3), we have

$$E(X) = \lim_{\eta\to\infty} \int_0^{\eta} [1-F(x)]dx = \int_0^{\infty} [1-F(x)]dx$$

Illustration 7.18 : Show that

$$E(X) = \int_0^{\infty} [1-F(x)+F(-x)]dx$$

Solution : $E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{-\infty}^{0} x f(x) dx + \int_0^{\infty} x f(x) dx \qquad \text{...(1)}$

$$= I_1 + I_2 \qquad \text{...(2)}$$

Now, $\int_0^{\infty} x f(x) dx = \int_0^{\infty} [1-F(x)]dx = I_2$

Also, $I_1 = \int_{-\infty}^{0} x f(x) dx = \int_{-\infty}^{0} x\, dF(x)$

Let, $x = -z$

$dx = -dz$

When, $x = -\infty, \quad z = \infty$

$x = 0, \qquad z = 0$

$$I_1 = \int_{\infty}^{0} -z\,dF(-z) = \int_{0}^{\infty} z\,dF(-z)$$

$$= zF(-z)\Big|_0^{\infty} + \int_0^{\infty} F(-z)\,dz = \int_0^{\infty} F(-z)\,dz = \int_0^{\infty} F(-x)\,dx \quad ....(3)$$

Putting (2) and (3) in (1), we have,

$$E(X) = \int_0^{\infty} [1 - F(x)]\,dx + \int_0^{\infty} F(-x)\,dx$$

$$= \int_0^{\infty} [1 - F(x) + F(-x)]\,dx$$

Illustration 7.19 : If X is a non-negative integer valued variate, prove that

$$\sum_{k=1}^{\infty} k\,P[X > k] = \frac{1}{2}\,[E(X^2) - E(X)]$$

Solution : $\displaystyle\sum_{k=0}^{\infty} k\,P[X > k] = \sum_{k=1}^{\infty} k \sum_{j=k+1}^{\infty} P(X=j)$

$$= \sum_{j\ge 1} P(X=j) \sum_{k=0}^{j-1} k$$

$$= \sum_{j\ge 1}^{\infty} \frac{1}{2}\, j(j-1)\,P(X=j)$$

$$= \frac{1}{2}\left[\sum P(X=j)\,j^2 - \sum P(X=j)\,j\right]$$

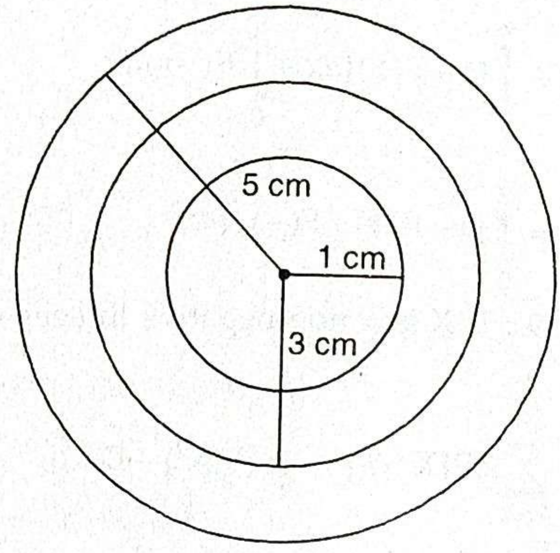
$$= \frac{1}{2}\left[E(X^2) - E(X)\right]$$

Illustration 7.20 : Concentric circles of radius 1 and 3 cm are drawn on a circle of radius 5 cm. A man receives 10, 5 or 3 points if he hits the target inside the smaller circle, inside the middle annular region or inside the outer annlar region respectively. Suppose the man hits the target with probability $\frac{1}{2}$ and then is just as likely to hit one point of the target as the other. Find the expected number E of points he scores each time he fires.

Solution : Let X be a random variable representing the points that the man obtains on firing.

$$\text{So, } P[X = 10] = \frac{1}{2}\frac{\text{area for 10 points}}{\text{area of the target}} = \frac{1}{2}\frac{\pi.1^2}{\pi.5^2} = \frac{1}{50}$$

$$P[X = 5] = \frac{1}{2}\frac{\text{area for 5 points}}{\text{area of the target}} = \frac{1}{2}\frac{\pi.3^2 - \pi.1^2}{\pi.5^2} = \frac{8}{30}$$



$$P[X = 3] = \frac{1}{2}\frac{\text{area for 3 points}}{\text{area of the target}} = \frac{1}{2}\frac{\pi.5^2 - \pi.3^2}{\pi.5^2} = \frac{16}{50}$$

$$P[X = 0] = \frac{1}{2}$$

Thus,
$$E[X] = 10.\frac{1}{50} + 5.\frac{8}{50} + 3.\frac{16}{50} + 0.\frac{1}{2} = \frac{98}{50} = 1.96$$

Illuzstration 7.21 : In a lottery m tickets are drawn at a time out of n tickets numbered from 1 to n. Find the variance of the sum S of the numbers on the tickets drawn.

Solution : Let $X_1, X_2, \ldots, X_n$ be the variables representing the numbers of first, second, ..., nth tickets. Therefore,

$$E[X_i] = \frac{1}{n} + 2.\frac{1}{n} + \ldots + n.\frac{1}{n} = \frac{n+1}{2}$$

In a similar way we find,

$$E\left[X_i^2\right] = \frac{1^2 + 2^2 + \ldots + n^2}{n} = \frac{(n+1)(2n+1)}{6}$$

Thus,
$$V(X_i) = E\left[X_i^2\right] - [E[X_i]]^2 = \frac{n^2 - 1}{12}$$

Now, $$\text{Cov}[X_i, X_j] = E\left[X_i - \frac{n+1}{2}\right]\left[X_j - \frac{n+1}{2}\right]$$

$$= -\sum \frac{1}{n(n-1)} E\left(X_j - \frac{n+1}{2}\right)^2 = -\frac{n+1}{12}$$

Thus, $$V(S) = V(X_1 + ... + X_m)$$

$$= \sum_{i=1}^{m} \text{Var}(X_i) + 2\sum_{i \neq j} \text{Cov}(X_i, X_j)$$

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

Illustration 7.22 : A box contains 2^n tickets among which $\binom{n}{r}$ bear the number r, $(r = 0, 1, 2, ..., n)$. A group of m tickets is drawn at random from the box, and if the random variable X denotes the sum of the numbers on the tickets drawn, show that

$$E(X) = \frac{mn}{2}; \qquad \text{Var}(X) = \frac{mn}{4}\left[1 - \frac{m-1}{2^n - 1}\right]$$

Solution : Let X_i be the random variable denoting the number on the i^{th} ticket drawn on the i^{th} occassion,

Then, $$X = \sum_{i=1}^{n} X_i$$

But $$E(X_i) = \sum_{t=0}^{n} t\,\frac{{}^nC_t}{2^n} = \frac{n}{2}$$

$$E\left[X_i^2\right] = \sum_{i=0}^{n} \frac{t^2\binom{n}{t}\left(\binom{n}{t}-1\right)}{2^n\left(2^n-1\right)} + \sum_{t \neq s} \frac{ts\binom{n}{t}\binom{n}{s}}{2^n\left(2^n-1\right)}; \text{ where } \binom{n}{t} = {}^nC_t$$

$$= \sum_{t=0}^{n} \frac{t^2\binom{n}{t}\left\{\binom{n}{t}-1\right\}}{2^n\left(2^n-1\right)} + \sum_{t=0}^{n} \frac{t\binom{n}{t}\left\{\sum_{s=0}^{n} s\binom{n}{s} - t\binom{n}{t}\right\}}{2^n\left(2^n-1\right)}$$

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

Now, $\text{Cov}(X_i, X_j) = E[X_i X_j] - E(X_i)E(X_j)$

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

$$E(X) = \frac{mn}{2}$$

$$Var(X) = \frac{mn}{4}\left[1 - \frac{m-1}{2^n - 1}\right]$$

Illustration 7.23 : The probability of obtaining a 6 with a biased die is p, $0 < p < 1$. Three players A, B and C roll this die in order, starting with A. The first to throw a 6 with the die wins. Find the probability of winning for A, B and C.

If X is a random variable represents the number of the throw at which the game end then, evaluate E(X) and Var (X).

Solution : Probability that A will win $= p + q^3p + q^6p + ... = \frac{p}{1-q^3}$

Probability that B will win $= qp + q^4 p + q^7 p + ...$ (as B is the second to start with)

$$= \frac{pq}{1-q^3}$$

Probability that C will win $= \frac{pq^2}{1-q^3}$

Also, $\quad P[X = r] = q^{r-1} p, r = 1, 2, ...$

So, $$E[X] = \sum_{r=1}^{\infty} r.P[X = r]$$

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

$$E[X^2] = \sum_{r=1}^{\infty} r^2.P[X = r]$$

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

So, $\quad V(X) = q/p^2.$