2.6

Functions of Random Variables

The previous modules discussed basic properties of events defined in a given sample space and the random variables used to represent those events. The fundamental assumption that was made in those modules is that events can always be defined by **random variables**. However, in many applications, the events are functions of other events. For example, the time until a complex system fails is a function of the time to failure of the individual components that make up the system. This means that the random variable used to represent the time to failure of the complex system is a function of the random variables used to represent the times to failure of the component parts of the system. This module deals with functions of random variables. Because of the complexity involved in computing the c.d.fs and p.d.fs of multiple random variables, the discussion is restricted to functions of at most two random variables.

Functions of One Random Variable: Let X be a r.v. with p.d.f. (or p.m.f.) $f_X(x)$ and c.d.f. $F_X(x)$. Let Y be the new random variable that is a function of X. That is,

$$Y = g(X)$$

Then we are interested in computing p.d.f (or p.m.f.) $f_Y(y)$ and c.d.f. $F_Y(y)$ of Y.

For example, let $Y = X + 5$. Then

$$F_Y(y) = P(Y \leq y) = P[X + 5 \leq y] = P[X \leq y - 5] = F_X(y - 5)$$

Linear Functions: Consider the function $g(X) = aX + b$, where a and b are constants. The c.d.f of Y is given by

$$F_Y(y) = P\{Y \leq y\} = P[aX + b \leq y]$$

$$= P\left[X \leq \frac{y-b}{a}\right] = F_X\left(\frac{y-b}{a}\right)$$

where a is positive .The p.d.f. of Y is given by

$$f_Y(y) = \frac{d}{dy}\left(F_Y(y)\right) = \frac{d}{dy}\left(F_X\left(\frac{y-b}{a}\right)\right) = \left(\frac{d}{du}\left(F_X(u)\right)\right)\left(\frac{du}{dy}\right)$$

where $u = \frac{y-b}{a}$ and $\frac{du}{dy} = \frac{1}{a}$. Thus,

$$f_Y(y) = \left(\frac{1}{a}\right) f_X(u) = \left(\frac{1}{a}\right) f_X\left(\frac{y-b}{a}\right)$$

If $a < 0$, we have,

$$F_Y(y) = P(Y \leq y) = P(aX + b \leq y) = P(aX \leq y - b)$$

$$= P\left(X \geq \frac{y-b}{a}\right) = 1 - \left\{P\left[X \leq \frac{y-b}{a}\right] - P\left[X = \frac{y-b}{a}\right]\right\} \quad (\because a < 0)$$

The change in sign on the second line arises from the fact that a is negative. If X is continuous, $P\left[X = \frac{(y-b)}{a}\right] = 0$. Thus, the c.d.f and p.d.f for the case of negative a are given by

$$F_Y(y) = 1 - P\left[X \leq \frac{y-b}{a}\right]$$

$$= 1 - F_X\left(\frac{y-b}{a}\right)$$

Therefore, $f_Y(y) = \frac{d}{dy}\left(F_Y(y)\right) = -\left(\frac{1}{a}\right) f_X\left(\frac{y-b}{a}\right)$

Therefore, the general p.d.f. of Y is given by

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right)$$

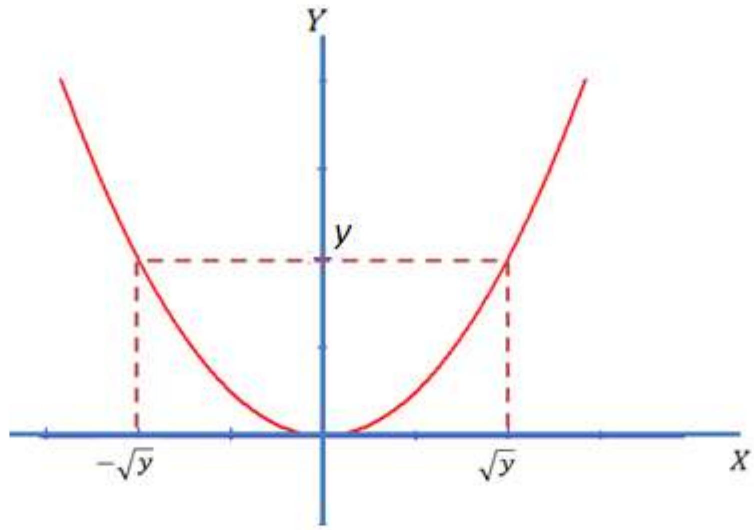
Example 1: Find the p.d.f of Y in terms of the p.d.f of X if $Y = 2X + 7$.

Solution: From the results obtained above,

$$F_Y(y) = F_X\left(\frac{y-7}{2}\right)$$

and $$f_Y(y) = \left(\frac{1}{2}\right) f_X\left(\frac{y-7}{2}\right)$$

Power Functions: Consider the quadratic function $Y = X^2$. The plot of Y against X is shown in the following figure where we see that for one value of Y there are two values of X, namely $\sqrt{Y}$ and $-\sqrt{Y}$.



Thus, the c.d.f of Y is given by

$$F_Y(y) = P[Y \leq y] = P[X^2 \leq y]$$

$$= P\left[|X| \leq \sqrt{y}\right], \quad y > 0$$

$$= P\left[-\sqrt{y} \leq X \leq \sqrt{y}\right]$$

$$= F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right)$$

The p.d.f of Y is given by

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{d}{dy}\left[F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right)\right]$$

Let $u = \sqrt{y} = y^{\frac{1}{2}}$. Thus, $\frac{du}{dy} = \frac{1}{2}y^{-\frac{1}{2}}$ and

$$f_Y(y) = \frac{d}{dy}\left[F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right)\right]$$

$$= \frac{d}{du}\left(F_X(u)\right)\frac{du}{dy} + \frac{d}{du}\left(F_X(-u)\right)\frac{du}{dy}$$

$$= \frac{1}{2} y^{-\frac{1}{2}} \left[\frac{d}{du} \left(F_X(u) \right) + \frac{d}{du} \left(F_X(-u) \right) \right]$$

$$= \frac{1}{2} y^{-\frac{1}{2}} \left[f_X\left(\sqrt{y}\right) + f_X\left(-\sqrt{y}\right) \right]$$

$$= \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} \quad , y > 0$$

If $f_X(x)$ is an even function, then $f_X(x) = f_X(-x)$ and $F_X(-x) = 1 - F_X(x)$. Thus, we have

$$f_Y(y) = \frac{f_X(\sqrt{y}) + f_X(-\sqrt{y})}{2\sqrt{y}} = \frac{2 f_X(\sqrt{y})}{2\sqrt{y}} = \frac{f_X(\sqrt{y})}{\sqrt{y}}$$

Example2: Find the p.d.f of the random variable $Y = X^2$, where X is the standard normal random variable.

Solution: Since the p.d.f. of X is given by $f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, which is an even function, we know that

$$F_Y(y) = F_X\left(\sqrt{y}\right) - F_X\left(-\sqrt{y}\right)$$

$$= 2F_X\left(\sqrt{y}\right) - 1$$

Therefore, if we let $u = \sqrt{y}$, then

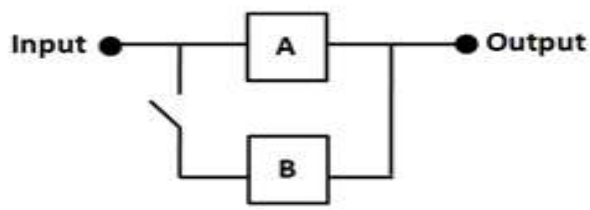
$$f_Y(y) = \frac{dF_Y(y)}{dy} = 2 \frac{dF_X(u)}{du} \frac{1}{2\sqrt{y}}$$

$$= \frac{1}{\sqrt{y}} f_X\left(\sqrt{y}\right)$$

$$= \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \; y > 0$$

Sum of Two Independent Random variables

Consider two independent continuous random variables X and Y. We are interested in computing the c.d.f and p.d.f of their sum $g(X,Y) = S = X + Y$. The random variable S can be used to model the reliability of systems with stand-by connections, as shown in *fig*.1 In such systems, the component A whose time-to-failure is represented by the random variable X is the primary component, and the component B whose time-to-failure is represented by the random variable Y is the backup component that is brought into operation when the primary component fails. Thus, S represents the time until the system fails, which is the sum of the lifetimes of both components.



$fig.1$

Their c.d.f. can be obtained as follows:

$$F_S(s) = P[S \le s] = P[X + Y \le s] = \iint_D f_{XY}(x,y)\,dxdy$$

where D is the set $D = \{(x,y) | x + y \le s\}$, which is the area to the left of the line $s = x + y$ as shown in *fig.* 2.

Thus,

$$F_S(s) = \int_{-\infty}^{\infty}\int_{-\infty}^{s-y} f_{XY}(x,y)\,dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{s-y} f_X(x)\,f_Y(y)\,dxdy$$

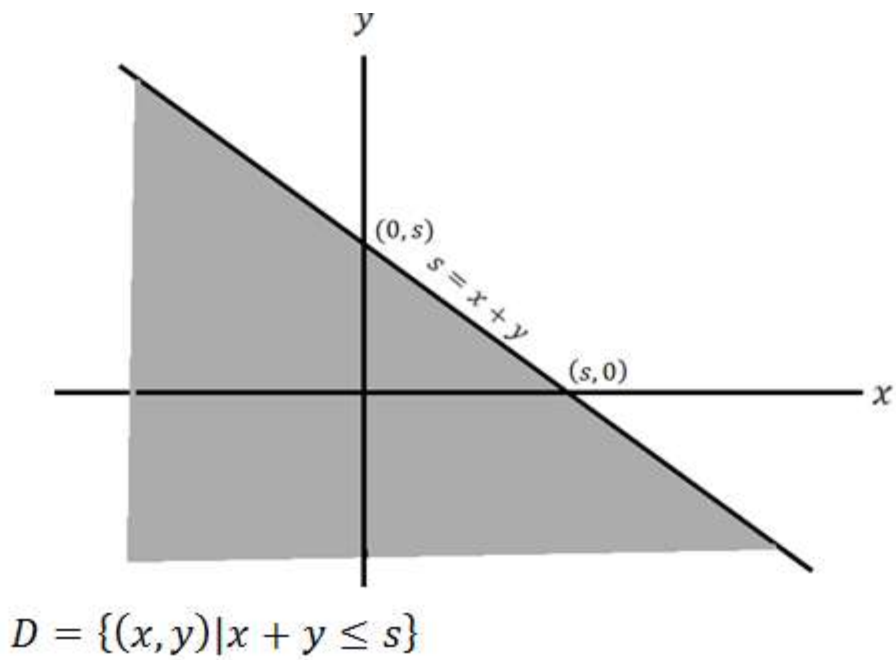
$$= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{s-y} f_X(x)\,dx\right\} f_Y(y)\,dy$$

$$= \int_{-\infty}^{\infty} F_X(s-y)\,f_Y(y)\,dy$$

The p.d.f. of S is obtained by differentiating the c.d.f. , as follows:

$$f_S(s) = \frac{d}{ds} F_S(s) = \frac{d}{ds} \int_{-\infty}^{\infty} F_X(s-y)\, f_Y(y)\, dy$$
$$= \int_{-\infty}^{\infty} \frac{d}{ds} F_X(s-y)\, f_Y(y)\, dy$$
$$= \int_{-\infty}^{\infty} f_X(s-y)\, f_Y(y)\, dy$$

where we have assumed that we can interchange differentiation and integration. The expression on the right-hand side is a well-known result in signal analysis



$fig. 2$

called the **convolution integral**. Thus, we find that the p.d.f of the sum S of two independent random variables X and Y is the convolution of the p.d.fs of the two random variables; that is,

$$f_S(s) = f_X(s) f_Y(s)$$

Example 3: Find the p.d.f. of the sum of X and Y if the two random variables are independent random variables with the common p.d.f.

$$\boldsymbol{f_X(u) = f_Y(u)} = \begin{cases} \boldsymbol{\frac{1}{4}} & \boldsymbol{0 < u < 4} \\ \boldsymbol{0} & \textbf{otherwise} \end{cases}$$

Solution: The limits of integration of the p.d.f of $S = X + Y$ can be computed with the aid of $fig.3$ When $0 \leq s \leq 4$ (see $fig.3\ (a)$ where $f_Y(s-x)$ is shown in dashed lines),

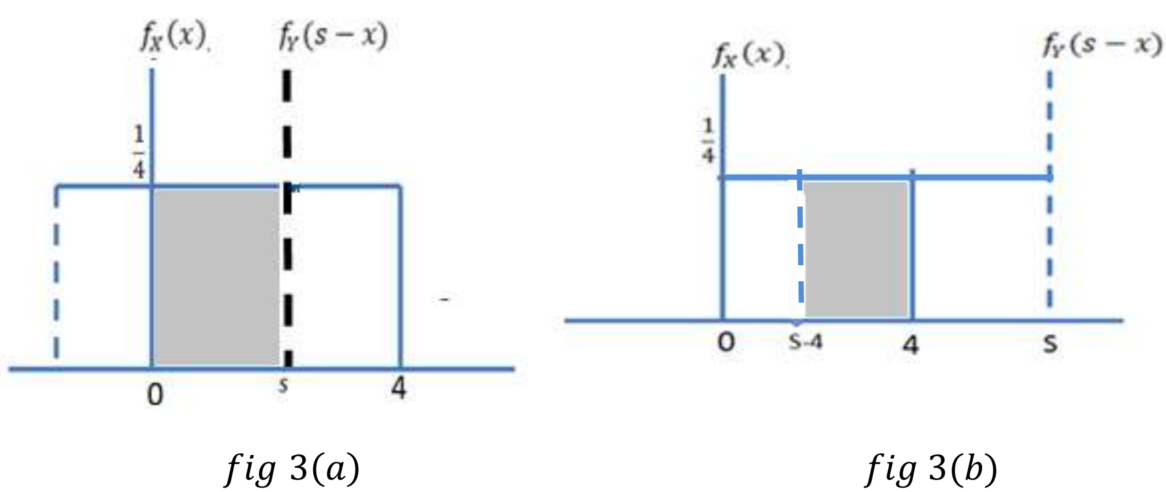
$$f_S(s) = \int_0^s \frac{1}{16} dy = \frac{s}{16}$$

For $4 < s < 8$ (see $fig.3\ (b)$), we obtain

$$f_S(s) = \int_{s-4}^4 dy = \frac{8-s}{16}$$

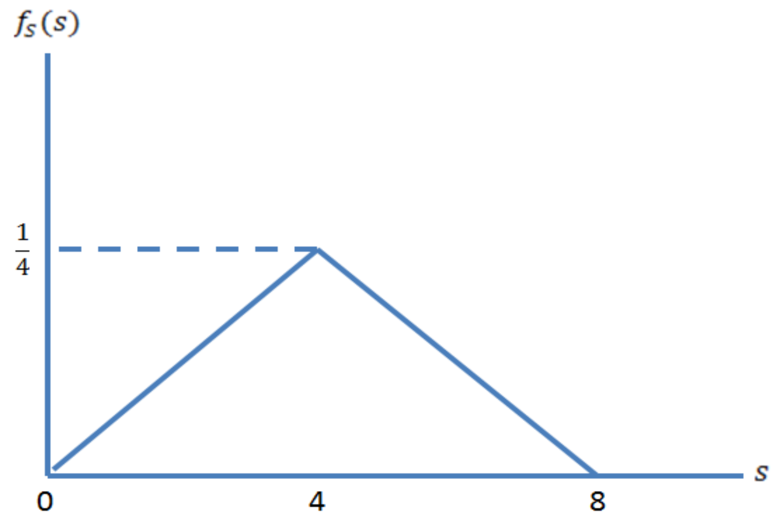
Thus ,

$$f_S(s) = \begin{cases} \frac{s}{16} & , \ 0 \leq s \leq 4 \\ \frac{8-s}{16} & , \ 4 < s < 8 \\ 0 & , \ \text{otherwise} \end{cases}$$



$Fig.3$: Convolution of p.d.fs $(a)\ \ 0 \leq s \leq 4$ and $(b)\ \ \ 4 \leq s \leq 8$

The p.d.f of $S = X + Y$ is illustrated in the following figure.



Example 4: The time X between consecutive snowstorms in winter is a random variable with the p.d.f.

$$\boldsymbol{f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}}$$

Assume it has not snowed up until now. What is the p.d.f. of the time U until the second snowstorm?

Solution: Let X be the random variable that denotes the time until the first snowstorm from the reference time, and let Y be the random variable that denotes the time between the first snowstorm and the second snowstorm. If we assume that the times between snowstorms are independent, then X and Y are independent and identically distributed random variables. That is, the p.d.f of Y is given by

$$f_Y(y) = \begin{cases} \lambda e^{-\lambda y}, & y \geq 0 \\ 0, & \text{otherwise} \end{cases}$$

Thus, $U = X + Y$, and the p.d.f. of U is given by

$$f_U(u) = \int_0^\infty f_X(x) f_Y(u - x) dx$$

Since $f_X(x) = 0$ when $x < 0, f_Y(u - x) = 0$ when $u - x < 0$ (or $x > u$). Thus, the range of interest in the integration is $0 \leq x \leq u$, and we obtain

$$\begin{aligned} f_U(u) &= \int_0^u f_X(x)\, f_Y(u-x)\,dx \\ &= \int_0^u \lambda e^{-\lambda x} \lambda e^{-\lambda(u-x)} dx = \lambda^2 e^{-\lambda u} \int_0^u dx \\ &= \lambda^2 u e^{-\lambda u} \quad u \geq 0 \end{aligned}$$

This is the Erlang – 2 distribution.

Note: A random variable X is said to follow Erlang-k distribution if its p.d.f. is given by

$$f(x) = \begin{cases} \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} & , \quad k = 1,2, \dots ;\ \lambda > 0; x \geq 0 \\ 0 & , \quad x < 0 \end{cases}.$$

Sum of Two Discrete Random Variables

The examples above deal with continuous random variables.
Let $Z = X + Y$, where X and Y are discrete random variables. Then the p.m.f of Z is given by

$$\begin{aligned} p_Z(z) = P[Z = z] = P[X + Y = z] &= \sum_{k \leq z} P[X = k, Y = z - k] \\ &= \sum_{k \leq z} p_{XY}[k, z - k] \end{aligned}$$

If X and Y are independent random variables, then the p.m.f. of Z is the convolution of the p.m.f of X and the p.m.f of Y. That is,

$$p_Z(z) = \sum_{k \leq z} p_{XY}(k, z - k) = \sum_{k \leq z} p_X(k) p_Y(z - k)$$

Sum of Two Independent Binomial Random Variables

Let X and Y be two independent binomial random variables with parameters (n, p) and (m, p), respectively and their sum be Z; that is, $Z = X + Y$. Then the p.m.f of Z is given by

$$\begin{aligned} p_Z(z) &= P[X+Y=z] \\ &= \sum_{k=0}^{n} P[X=k, Y=z-k] = \sum_{k=0}^{n} P[X=k]P[Y=z-k] \\ &= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \binom{m}{z-k} p^{z-k} (1-p)^{m-z+k} \\ &= p^z (1-p)^{n+m-z} \sum_{k=0}^{n} \binom{n}{k}\binom{m}{z-k} \end{aligned}$$

Using the combinatorial identity $\binom{n+m}{z} = \sum_{k=0}^{n} \binom{n}{k}\binom{m}{z-k}$, we obtain

$$p_Z(z) = \binom{n+m}{z} p^z (1-p)^{n+m-z}$$

This result shows that the sum of two independent binomial random variables with parameters (n, p) and (m, p) is a binomial random variable with parameter $(n + m, p)$.

Minimum of Two Independent Random Variables

Consider two independent continuous random variables X and Y. We are interested in a random variable U that is the minimum of X and Y; that is, $U = \min(X, Y)$. The random variable U can be used to represent the reliability of systems with series connections, as shown in $fig. 4$ .Such systems are operational as long as all components are operational. The first component to fail causes the system to fail. Thus, if in the example shown in $fig. 4$, the times-to-failure are represented by the random variables X and Y, then S represents the time until the system fails, which is the minimum of the lifetimes of the two components.

The c.d.f. of U can be obtained as follows:

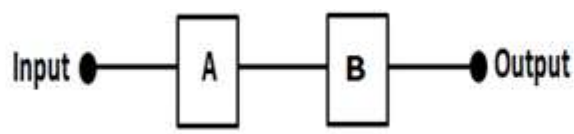
$$F_U(u) = P[U \leq u] = P[\min(X,Y) \leq u] = P[(X \leq u, X \leq Y) \cup (Y \leq u, X > Y)]$$

Since $P[A \cup B] = P[A] + P[B] - P[A \cap B]$, we have that $F_U(u) = F_X(u) + F_Y(u) - F_{XY}(u,u)$. Also, since X and Y are independent, we obtain the c.d.f. and p.d.f. of U as follows:

$$F_U(u) = F_X(u) + F_Y(u) - F_{XY}(u,u) = F_X(u) + F_Y(u) - F_X(u)F_Y(u)$$

$$f_U(u) = \frac{d}{du}F_U(u) = f_X(u) + f_Y(u) - f_X(u)F_Y(u) - F_X(u)f_Y(u)$$

$$= f_X(u)\{1 - F_Y(u)\} + f_Y(u)\{1 - F_X(u)\}$$



$fig.4$

$fig.4$: Series connection modeled by random variable U

Example 5: Assume that $U = \min(X,Y)$, where X and Y are independent random variables with the respective p.d.fs

$$\boldsymbol{f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0}$$

$$\boldsymbol{f_Y(y) = \mu e^{-\mu y} \quad y \geq 0}$$

where $\lambda > 0$ and $\mu > 0$. What is the p.d.f. of U?

Solution: We first obtain the c.d.fs of X and Y, which are as follows:

$$F_X(x) = P[X \leq x] = \int_0^x \lambda e^{-\lambda w}\, dw = 1 - e^{-\lambda x}$$

$$F_Y(y) = P[Y \leq y] = \int_0^y \mu e^{-\mu w}\, dw = 1 - e^{-\mu y}$$

Thus, the p.d.f of U is given by

$$f_U(u) = f_X(u)\{1 - F_Y(u)\} + f_Y(u)\{1 - F_X(u)\}$$

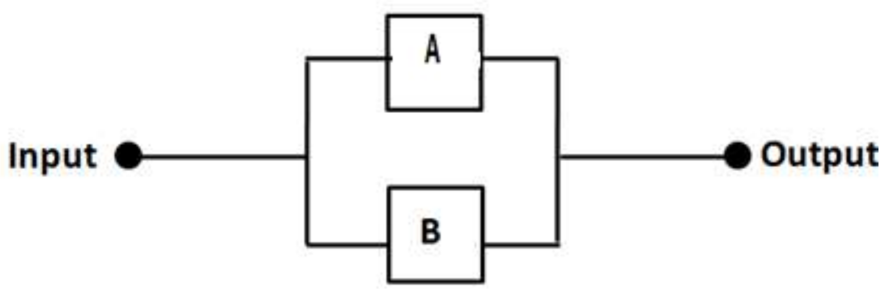
$$= \lambda e^{-\lambda u} e^{-\mu u} + \mu e^{-\mu u} e^{-\lambda u}$$

$$= (\lambda + \mu) e^{-(\lambda+\mu)u} \ , \ u \geq 0$$

This is exponential distribution with mean $\frac{1}{\lambda+\mu}$.

Maximum of Two Independent Random Variables

Consider two independent continuous random variables X and Y. We are interested in the c.d.f. and p.d.f. of the random variable W that is the maximum of the two random variables; that is, $W = \max(X, Y)$. The random variable W can be used to represent the reliability of systems with parallel connections, as shown in



$fig.5$

$fig.5$: Parallel connection modeled by the random variable W

In such systems, we are interested in passing a signal between the two endpoints through either the component labeled A or the component labeled B. Thus, as long as one or both components are operational, the system is operational. This implies that the system is declared to have failed when both paths become unavailable. That is, the reliability of the system depends on the reliability of the last component to fail.

The c.d.f of W can be obtained by noting that if the greater of the two random variables is less than or equal to w, then the smaller random variable must also be less than or equal to w. Thus,

$$F_W(w) = P[W \leq w] = P[\max(X, Y) \leq w] = P[(X \leq w) \cap (Y \leq w)]$$

$$= F_{XY}(w, w)$$

Since X and Y are independent, we obtain the c.d.f and p.d.f of W as follows:

$$F_W(w) = F_{XY}(w, w) = F_X(w)F_Y(w)$$

$$f_W(w) = \frac{d}{dw}F_W(w) = f_X(w)F_Y(w) + F_X(w)f_Y(w)$$

Example 6: Assume that $W = \max(X, Y)$, where X and Y are independent random variables with the respective p.d.fs:

$$\boldsymbol{f_X(x) = \lambda e^{-\lambda x} \quad x \geq 0}$$

$$\boldsymbol{f_Y(y) = \mu e^{-\mu y} \quad y \geq 0}$$

where $\lambda > 0$ and $\mu > 0$. What is the pdf of W.

Solution: We first obtain the c.d.fs of X and Y, which are as follows:

$$F_X(x) = P[X \leq x] = \int_0^x \lambda e^{-\lambda z}\, dz = 1 - e^{-\lambda x}$$

$$F_Y(y) = P[Y \leq y] = \int_0^y \mu e^{-\mu z}\, dz = 1 - e^{-\mu y}$$

Thus, the p.d.f of W is given by

$$f_W(w) = f_X(w)F_Y(w) + F_X(w)f_Y(w)$$

$$= \lambda e^{-\lambda w}(1 - e^{-\mu w}) + \mu e^{-\mu w}\left(1 - e^{-\lambda w}\right)$$

$$= \lambda e^{-\lambda w} + \mu e^{-\mu w} - (\lambda + \mu)e^{-(\lambda+\mu)w} \quad w \geq 0$$

Note that the mean of W is $\frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda+\mu}$.

Two Functions of Two Random Variables

Let X and Y be two random variables with a given joint p.d.f $f_{XY}(x,y)$. Assume that U and W are two functions of X and Y; that is, $U = g(X,Y)$ and $W = h(X,Y)$. Sometimes it is necessary to obtain the joint p.d.f of U and W, $f_{UW}(u,w)$, in terms of the p.d.fs of X and Y.

It can be shown that $f_{UW}(u,w)$ is given by

$$f_{UW}(u,w) = \frac{f_{XY}(x_1,y_1)}{|J(x_1,y_1)|} + \frac{f_{XY}(x_2,y_2)}{|J(x_2,y_2)|} + \cdots + \frac{f_{XY}(x_n,y_n)}{|J(x_n,y_n)|}$$

where $(x_1,y_1), (x_2,y_2), \dots, (x_n,y_n)$ are real solutions of the equations $u = g(x,y)$ and $w = h(x,y)$; and $J(x,y)$ is called the **Jacobian** of the transformation $\{u = g(x,y), w = h(x,y)\}$ and defined by

$$J(x,y) = \begin{vmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \end{vmatrix} = \frac{\partial g}{\partial x} \cdot \frac{\partial h}{\partial y} - \frac{\partial g}{\partial y} \cdot \frac{\partial h}{\partial x}$$

Example 7: Let $\boldsymbol{U = g(X,Y) = X + Y}$ and $\boldsymbol{W = h(X,Y) = X - Y}$. Find $\boldsymbol{f_{UW}(u,w)}$.

Solution: The unique solution to the equations $u = x + y$ and $w = x - y$ is $x = \frac{u+w}{2}$ and $y = \frac{u-w}{2}$. Thus, there is only one set of solutions. Since

$$J(x,y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix} = -2$$

we obtain

$$f_{UW}(u,w) = \frac{f_{XY}(x,y)}{|J(x,y)|} = \frac{1}{|-2|} f_{XY}\left(\frac{u+w}{2}, \frac{u-w}{2}\right) = \frac{1}{2} f_{XY}\left(\frac{u+w}{2}, \frac{u-w}{2}\right)$$

Application of the Transformation Method

Assume that $U = g(X, Y)$, and we are required to find the p.d.f. of U. We can use the above transformation method by defining an auxiliary function $W = X$ or $W = Y$ so we can obtain the joint p.d.f. $f_{UW}(u, w)$ of U and W. Then we obtain the required marginal p.d.f. $f_U(u)$ as follows:

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u, w)dw$$

Example 8: Find the p.d.f. of the random variable $\boldsymbol{U = X + Y}$, where the joint p.d.f. of $\boldsymbol{X}$ and $\boldsymbol{Y}$, $\boldsymbol{f_{XY}(x, y)}$, is given.

Solution: We define the auxiliary random variable $W = X$. Then the unique solution to the two equations is $x = w$ and $y = u - w$, and the Jacobian of the transformation is

$$J(x, y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

Since there is only one solution to the equations, we have that

$$f_{UW}(u, w) = \frac{f_{XY}(w, u - w)}{|-1|} = f_{XY}(w, u - w)$$

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u, w)dw = \int_{-\infty}^{\infty} f_{XY}(w, u - w)dw$$

This reduces to the convolution integral. We obtained earlier when X and Y are independent.

Example 9: Find the p.d.f. of the random variable $U = X - Y$, where the joint p.d.f. of X and Y is given.

Solution: We define the auxiliary random variable $W = X$. Then the unique solution to the two equations is $x = w$ and $y = w - u$, and the Jacobian of the transformation is

$$J(x,y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} 1 & -1 \\ 1 & 0 \end{vmatrix} = 1$$

Since there is only one solution to the equations, we have that

$$f_{UW}(u,w) = f_{XY}(w, w-u)$$

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u,w)dw = \int_{-\infty}^{\infty} f_{XY}(w, w-u)dw$$

Example 10: The joint p.d.f of two random variables X and Y is given by $f_{XY}(x,y)$. If we define the random variable $U = XY$, determine the p.d.f of U.

Solution: We define the auxiliary random variable $W = X$. Then the unique solution to the two equations is $x = w$ and $y = \frac{u}{x} = \frac{u}{w}$, and the Jacobian of the transformation is

$$J(x,y) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{vmatrix} = \begin{vmatrix} y & x \\ 1 & 0 \end{vmatrix} = -x = -w$$

Since there is only one solution to the equations, we have that

$$f_{UW}(u,w) = \frac{f_{XY}(x,y)}{|J(x,y)|} = \frac{1}{|w|} f_{XY}\left(w, \frac{u}{w}\right)$$

$$f_U(u) = \int_{-\infty}^{\infty} f_{UW}(u,w)dw = \int_{-\infty}^{\infty} \frac{1}{|w|} f_{XY}\left(w, \frac{u}{w}\right) dw$$

P1:

Obtain the p.d.f. of $Z = X + Y$, where X and Y are two independent random variables with the following p.d.fs:

$$f_X(x) = \begin{cases} \dfrac{1}{b-a} & ; \quad a < x < b \\ 0 & ; \quad otherwise \end{cases}$$

$$f_Y(y) = \begin{cases} \dfrac{1}{d-c} & ; \quad c < y < d \,; \; d - c < b - a \\ 0 & ; \quad otherwise \end{cases}$$

Solution:

The two p.d.fs are shown in the *fig.* 1. To evaluate the limits of integration of the p.d.f. of Z, we consider the following regions represented by the diagram shown in *fig.* 2.

When $z < a + c$, $f_Z(z) = 0$. When $a + c \leq z \leq a + d$ (see *figure 2(i)*, we obtain)

$$f_Z(z) = \frac{1}{(b-a)(d-c)} \int_a^{z-c} dy = \frac{z-c-a}{(b-a)(d-c)}$$

When $a + d \leq z \leq \mathrm{b} + \mathrm{c}$ (see *fig.* $2(ii)$), we obtain)

$$f_Z(z) = \frac{1}{(b-a)(d-c)} \int_{z-d}^{z-c} dy = \frac{1}{b-a}$$

When $b + c \leq z \leq b + d$ (see *fig.* $2(ii)$), we obtain

$$f_Z(z) = \frac{1}{(b-a)(d-c)} \int_{z-d}^{b} dy = \frac{b+d-z}{(b-a)(d-c)}$$

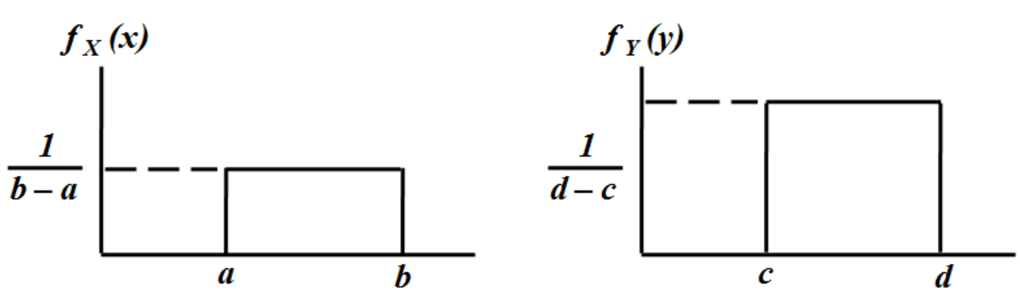


Figure 1: p.d.fs of X and Y

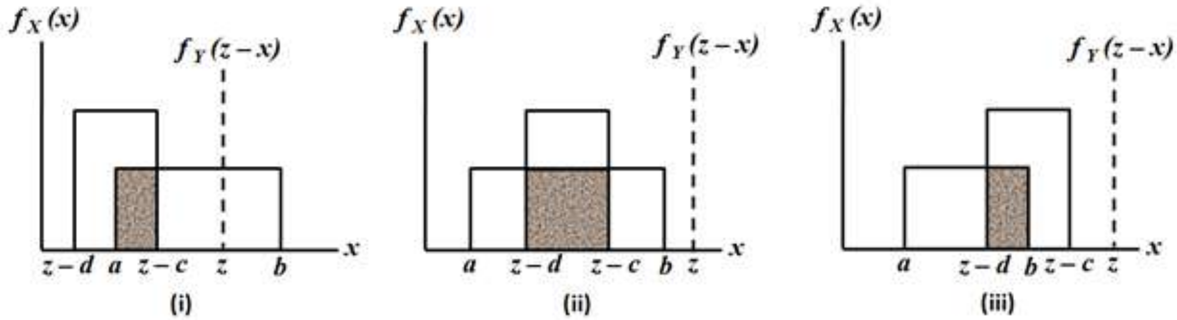
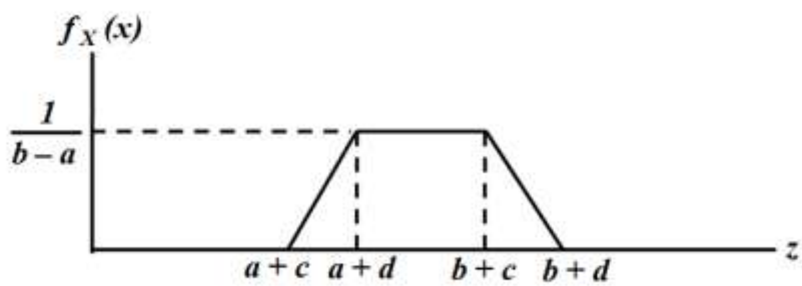


Figure 2: Convolution of the p.d.fs for different values z

Finally, when $z > b + d$, $f_Z(z) = 0$, thus, the p.d.f of Z is given by

$$f_Z(z) = \begin{cases} 0 & , \quad z < a + c \\ \dfrac{z - a - c}{(b-a)(d-c)} & , \quad a + c \le z \le a + d \\ \dfrac{1}{b-a} & , \quad a + d \le z \le b + c \\ \dfrac{b + d - z}{(b-a)(d-c)} & , \quad b + c \le z \le b + d \\ 0 & , \quad z > b + d \end{cases}$$

The p.d.f is graphically illustrated in the following figure which is a **trapezoid**. Note that when $b - a = d - c$, the p.d.f reduces to an **isosceles triangle** centered at $z = \frac{a+c+b+d}{2}$. In the special case when $a = c$ and $b = d$, the isosceles triangle is centered at $z = a + b$.



P2:

Find the p.d.f of U, which is the sum of X and Y that are independent random variables with the following p.d.fs:

$$\boldsymbol{f_X(x) = \lambda e^{-\lambda x} \quad , \quad x \geq 0}$$
$$\boldsymbol{f_Y(y) = \lambda^2 y e^{-\lambda y} \quad , \quad y \geq 0}$$

Solution:

Since X and Y are independent random variables and the p.d.f of U is given by

$$f_U(u) = \int_0^u f_X(x) f_Y(u-x)\, dx$$
$$= \int_0^u \lambda e^{-\lambda x}\, \lambda^2 (u-x) e^{-\lambda(u-x)} dx$$

$$= \lambda^3 e^{-\lambda u} \int_0^u (u-x)\, dx$$

$$= \lambda^3 e^{-\lambda u} \left[ux - \frac{x^2}{2}\right]_0^u = \lambda^3 e^{-\lambda u} \left[u^2 - \frac{u^2}{2}\right]$$

$$= \frac{\lambda^3 u^2 e^{-\lambda u}}{2}$$

$$= \frac{\lambda^3 u^2 e^{-\lambda u}}{2!} \quad u \geq 0$$

This is known as **Erlang – 3 distribution**.

2.6. Functions of Random Variables:

Exercise

1) The p.m.f. of X is given by

x	-2	-1	1	2
$p(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Find the probability distribution of $Y = 4X + 3$.

2) A r.v. X is exponentially distributed with parameter1. Find the p.d.f. of $y = \sqrt{x}$

3) Let X be a c.r.v with p.d.f.

$$f(x) = \begin{cases} \frac{x}{12} & , \quad 1 < x < 5 \\ 0 & , \quad otherwise \end{cases}$$

find the p.d.f. of $y = 2X + 4$

4) If the c.r.v X has the p.d.f.

$$f(x) = \begin{cases} \frac{2}{9}(x+1) & , \quad -1 < x < 2 \\ 0 & , \quad otherwise \end{cases}$$

find the p.d.f. of $y = X^2$

5) Find the $p.d.f$ of $W = X + Y$ where X and Y are independent r.v s with the following p.d.f s:

$$f_X(x) = \lambda e^{-\lambda x}, x \geq 0$$

and $$f_Y(y) = \mu e^{-\mu y}, \ \ y \geq 0, \text{ where } \lambda \neq \mu.$$

6) The j.p.d.f of two random variables X and Y is given by $f_{XY}(xy)$. Find the p.d.f of $V = \frac{X}{Y}$

ANSWERS

1.

Y	-5	-1	7	11
$p(y)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

2. $f(y) = \begin{cases} 2y\, e^{-y^2}, & y \geq 0 \\ 0\quad, & otherwise \end{cases}$

3. $f(y) = \begin{cases} \frac{y-4}{48}\ , & 6 < y < 14 \\ 0\ , & otherwise \end{cases}$

4. $f(y) = \begin{cases} \frac{2}{9\sqrt{y}} & , 0 < y < 1 \\ \frac{1}{9}\left(\frac{\sqrt{y}+1}{\sqrt{y}}\right) & , 1 < y < 4 \\ 0 & , \ otherwise \end{cases}$

5. $f_W(w) = \begin{cases} \frac{\lambda\mu}{\lambda-\mu}\ \left(e^{-\mu w} - e^{-\lambda w}\right), & w \geq 0 \\ 0, & otherwise \end{cases}$

6. $f_V(v) = \int_{-\infty}^{\infty} |w|\, f_{XY}(vw, w) dw$

2.7

Correlation coefficient and Bivariate Normal Distribution

Meaning of correlation:

In a bivariate distribution we may be interested to find out if there is any **correlation** or **covariance** between the two variables under study. If the change in one variable affects a change in the other variable, the variables are said to be **correlated**. If the two variables deviate in the same direction, $i.e.$, if the increase (or decrease) in one results in a corresponding increase (or decrease) in the other, **correlation** is said to be **positive**. But, if they constantly deviate in the opposite directions, $i.e.$, if increase (or decrease) in one results in corresponding decrease (or increase) in the other, ***correlation** is said to be **negative***. For example, the correlation between (i) the heights and weights of a group of persons, and (ii) the income and expenditure; is positive and the correlation between (i) price and demand of a commodity and (ii) the volume and pressure of a perfect gas; is negative. *Correlation is said to be **perfect*** if the deviation in one variable is followed by a corresponding and proportional deviation in the other.

Karl Pearson's Coefficient of Correlation:

As a measure of intensity or degree of linear relationship between two variables, **Karl Pearson**, a British Biometrician developed a formula called **correlation coefficient**. Correlation coefficient between two variables X and Y, usually denoted by $\boldsymbol{\rho(X,Y)}$ or $\boldsymbol{\rho_{XY}}$, is a numerical measure of linear relationship between them and is defined by

$$\boldsymbol{\rho(X,Y) = \frac{\sigma_{XY}}{\sigma_X . \sigma_Y} = \frac{cov\,(X,Y)}{\sqrt{V(X)}\,\sqrt{V(Y)}}}$$

where $\sigma_{XY} = cov\,(X,Y) = E\big[\big(X - E(X)\big)\,\big(Y - E(y)\big)\big]$,

$\sigma_X{}^2 = V(X) = E\big[\big(X - E(X)\big)\big]^2$ and $\sigma_Y{}^2 = V(Y) = E\big[\big(Y - E(Y)\big)\big]^2$

Note:

1. $\rho(X, Y)$ provides a measure of linear relationship between X and Y. For non – linear relationship, however, it is not suitable.

2. Karl Pearson's correlation coefficient is also called **product – moment correlation coefficient.**

Properties:

1. $-1 \leq \rho(X, Y) \leq 1$. If $\rho = -1$, the **correlation is perfect and negative**. If $\rho = 1$, the **correlation is perfect and positive**.
2. Correlation coefficient is independent of change of origin and scale. That is, if $U = \frac{X-a}{h}$ and $V = \frac{Y-b}{k}$, then $\rho(U, V) = \rho(X, Y)$

Theorem: Two independent variables are uncorrelated.

Proof:

Consider $\sigma_{XY} = cov\ (X, Y) = E\big[\big(X - E(X)\big)\big(Y - E(X)\big)\big]$

$\Longrightarrow\ \sigma_{XY} = E(X, Y) - E(X). E(Y)$ ...... (1)

If X and Y are independent, then

$E(XY) = E(X). E(Y)$ .... (2)

From (1) and (2), if X and Y are independent, then $\rho(X, Y) = 0$

The converse need not be true. That is, uncorrelated variables need not be independent.

Example 1 : Let $\boldsymbol{X \sim N(0, 1)}$ and $\boldsymbol{Y = X^2}$. Then $\boldsymbol{E(X) = E(X^3) = 0}$.

Solution: Consider $cov\ (X, Y) = E(XY) - E(X). E(Y) = E(X^3) - E(X). E(X^2)$

$$= 0 - 0 = 0$$

$\Longrightarrow cov\ (X, Y) = 0$ but X and Y are related by $Y = X^2$.

Thus, uncorrelated variables need not be independent.

Note: The converse is true if the joint distribution of (X, Y) is bivariate normal.

Example 2: The j.p.m.f of (X, Y) is given below:

$Y \backslash X$	-1	1
0	$\frac{1}{8}$	$\frac{3}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$

Find the correlation coefficient between X and Y

Solution : Computation of marginal p.m.fs

$Y \backslash X$	-1	1	$g(y)$
0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{4}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{4}{8}$
$p(x)$	$\frac{3}{8}$	$\frac{5}{8}$	1

We have

$E(X) = \sum x\, p(x) = (-1) \times \frac{3}{8} + 1 \times \frac{5}{8} = -\frac{3}{8} + \frac{5}{8} = \frac{2}{8} = \frac{1}{4}$,

$E(X^2) = \sum x^2\, P(x) = (-1)^2 \frac{3}{8} + 1^2 \times \frac{5}{8} = \frac{3}{8} + \frac{5}{8} = 1$, then

$V(X) = E(X^2) - \big(E(X)\big)^2 = 1 - \left(\frac{1}{4}\right)^2 = 1 - \frac{1}{16} = \frac{15}{16}$

Similarly, $E(Y) = \sum y\, g(y) = 0 \times \frac{4}{8} + 1 \times \frac{4}{8} = \frac{4}{8} = \frac{1}{2}$

$E(Y^2) = \sum y^2\, g(y) = 0^2 \times \frac{4}{8} + 1^2 \times \frac{4}{8} = \frac{4}{8} = \frac{1}{2}$ and

$V(Y) = E(Y^2) - \left(E(Y)\right)^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$

Further, $E(XY) = 0 \times (-1) \times \frac{1}{8} + 0 \times 1 \times \frac{3}{8} + 1 \times (-1) \times \frac{2}{8} + 1 \times 1 \times \frac{2}{8} = 0$

Thus, $cov\ (X, Y) = E(XY) - E(X)E(Y)$

$$0 - \frac{1}{4} \times \frac{1}{2} = -\frac{1}{8}$$

$$\therefore \rho\ (X, Y) = \frac{cov\ (X,Y)}{\sqrt{V(X)}\,\sqrt{V(Y)}} = -\frac{\frac{1}{8}}{\sqrt{\frac{15}{16} \times \frac{1}{4}}} = -\frac{1}{\sqrt{15}} = -0.2582$$

Example 3: Two random variables X and Y have the joint probability density function

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

Find correlation coefficient between X and Y.

Solution: By symmetry in x and y we have $f_1(x) = f_2(y)$, $E(X) = E(Y)$ and $V(X) = V(Y)$

The m.p.d.f X is given by

$$f_1(x) = \int_0^1 f(x, y)\, dy = \int_0^1 (2 - x - y)\, dy = \frac{3}{2} - x$$

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

Consider.

$$E(X)=\int_0^1 xf_1(x)dx=\int_0^1 x\left(\frac{3}{2}-x\right)dx=\int_0^1\left(\frac{3}{2}x-x^2\right)dx=\frac{5}{12}$$

$$E(X^2)=\int_0^1 xf_1(x)=\int_0^1 x^2\left(\frac{3}{2}-x\right)dx=\int_0^1\left(\frac{3}{2}x^2—x^3\right)dx=\frac{1}{4}$$

Further,

$$E(XY)=\int_0^1\int_0^1 xy\,f(x,y)dx\,dy\ =\int_0^1\int_0^1 xy\,(2-x-y)dx\,dy$$

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

$$=\int_0^1 y\left(1-\frac{1}{3}-\frac{y}{2}\right)dy=\int_0^1 y\left(\frac{2}{3}-\frac{y}{2}\right)dy$$

$$=\int_0^1\left(\frac{2}{3}y-\frac{y^2}{2}\right)dy=\left[\frac{y^3}{3}-\frac{y^3}{6}\right]_0^1=\frac{1}{3}-\frac{1}{6}=\frac{1}{6}$$

$$\therefore E(XY)=\frac{1}{6}$$

Thus, $V(X)=E(X^2)-\left(E(X)\right)^2=\frac{1}{4}-\left(\frac{5}{12}\right)^2=\frac{1}{4}-\frac{25}{144}=\frac{36-25}{144}=\frac{11}{144}$

and $cov\ (X,Y)=E(XY)-E(X)E(Y)=\frac{1}{6}-\left(\frac{5}{12}\right)^2=\frac{1}{6}-\frac{25}{144}=\frac{24-25}{144}=-\frac{1}{144}$

∴ The correlation coefficient is given by

$$\rho(X,Y)=\frac{cov\ (X,Y)}{\sqrt{V(x)}\sqrt{V(Y)}}=-\frac{\frac{1}{144}}{\sqrt{\frac{11}{144}}\sqrt{\frac{11}{144}}}=-\frac{\frac{1}{144}}{\frac{11}{144}}=-\frac{1}{11}$$

Bivariate Normal Distribution:

The bivaraite normal distribution is a generalization of a normal distribution for a single value.

Let X and Y be two normally correlated variables with correlation coefficient ρ. Let $E(X) = \mu_1,\ V(X) = \sigma_1^2\ , E(Y) = \mu_2$ and $V(Y) = \sigma_2^2$.

Definition: The bivariate continuous random variable (X, Y) is said to follow bivariate normal distribution with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and ρ if its p.d.f. is given by

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right\}\right];$$

$$-\infty < x, y, \mu_1, \mu_2 < \infty,\ \ \sigma_1 > 0, \sigma_2 > 0 \text{ and } -1 < \rho < 1.$$

Notation: $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Read as (X, Y) follows **bivariate normal distribution** with parameters $\mu_1, \mu_2 \sigma_1^2, \sigma_2^2$ and ρ.

Note: The curve $z = f(x, y)$ which is the equation of a surface in three dimensions is called the **Normal correlation surface**.

Marginal p.d.fs of X and Y: The m.p.d.f of X is given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)dy$$

Let $v = \frac{y-\mu_2}{\sigma_2}$, then $y = \mu_2 + \sigma_2 v$ and $dy = \sigma_2 dv$

Therefore,

$$f_1(x) = \frac{\sigma_2}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\, v\left(\frac{x-\mu_1}{\sigma_1}\right) + v^2\right\}\right] dv$$

$$= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] \int_{-\infty}^{\infty} exp\left[-\frac{1}{2(1-\rho^2)}\left\{v - \rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right\}^2\right] dv$$

Let $\frac{1}{\sqrt{1-\rho^2}}\left[v - \rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right] = t$. Then $dv = \sqrt{1-\rho^2}\ dt$

$\therefore\ f_1(x) = \frac{1}{2\pi\sigma_1} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)\right] \int_{-\infty}^{\infty} exp\left(-\frac{t^2}{2}\right) dt$

$= \frac{1}{2\pi\sigma_1} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] . \sqrt{2\pi}$

$\Rightarrow\ f_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]$ for $-\infty < x < \infty$

Similarly, it can be shown that

$f_2(y) = \frac{1}{\sigma_2\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{y-\mu_1}{\sigma_2}\right)^2\right]$ for $-\infty < x < \infty$

Hence $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$.

Note: If $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\ )$, then $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$

Conditional p.d.fs of X and Y

The conditional probability density function (c.p.d.f.) of X for given Y is given by

$$f_{1|2}(x|y) = \frac{f(x,y)}{f_2(y)}$$

$= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2 \left(1-(1-\rho^2)\right)\right\}\right]$

$= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}}\ exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1)^2 - 2\rho\frac{\sigma_1}{\sigma_2}(x-\mu_1)(y-\mu_2) + \frac{\sigma_1^2}{\sigma_2^2}\rho^2(y-\mu_2)^2\right\}\right]$

$= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1) - \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]$

Therefore, $f_{1|2}(x|y) = \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1) - \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]$

which is the univariate normal distribution with mean

$E(X|Y=y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)$ and

$V(X|Y=y) = \sigma_1^2(1-\rho^2)$

Thus, the c.p.d.f of X for fixed Y is given by

$$(X|Y=y) \sim N\left[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2), \sigma_1^2(1-\rho^2)\right]$$

Similarly, the c.p.d.f of Y for fixed $X = x$ is given by

$$f_{2|1}(y|x) = \frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)\sigma_2^2}\left\{(y-\mu_2) - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right\}^2\right], -\infty < y < \infty$$

Thus, the c.p.d.f of Y for fixed X is given by

$$(Y|X=x) \sim N\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2)\right]$$

Example 4: If $(X,Y) \sim BN\ (5, 10, 1, 25, \rho)$ where $\rho > 0$, find ρ when $P(4 < Y < 16|X = 5) = 0.954$

Solution:

Here $\mu_1 = 5, \mu_2 = 10,\ \sigma_1^2 = 1, \sigma_2^2 = 25$. We know that $(Y|X = x) \sim N[\mu, \sigma^2]$

where $\mu = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1)$ and $\sigma^2 = \sigma_2^2(1-\rho^2)$.

Here $\mu = 10 + \rho \times \frac{5}{1}(5-5) = 10$ and $\sigma^2 = 25(1-\rho^2)$

Thus $(Y|X = 5) \sim N[10, 25(1-\rho^2)]$. We want to find ρ so that $P(4 < Y < 16|X = 5) = 0.954$

Let $Z = \frac{Y-\mu}{\sigma} = \frac{Y-10}{5\sqrt{1-\rho^2}} \sim N(0,1) \Rightarrow P\left(\frac{4-10}{\sigma} < Z < \frac{16-10}{\sigma}\right) = 0.954$

$\Rightarrow P\left(-\frac{6}{\sigma} < Z < \frac{6}{\sigma}\right) = 0.954 \Rightarrow P\left(0 < Z < \frac{6}{\sigma}\right) = 0.477$

From standard normal table, we have $\frac{6}{\sigma} = 2 \ \Rightarrow \sigma = 3 \ \Rightarrow \sigma^2 = 9$

$\Rightarrow 25(1-\rho^2) = 9 \Rightarrow 1 - \rho^2 = \frac{9}{25} \Rightarrow \rho^2 = 1 - \frac{9}{25} = \frac{16}{25} \Rightarrow \rho = \frac{4}{5} = 0.8$

Example 5: Find $cor(X,Y)$ for the jointly normal distribution

$$\boldsymbol{f(x,y)=\frac{1}{2\pi\sqrt{3}}exp\left[-\frac{\{(2x-y)^2+2xy\}}{6}\right]}\ ,-\infty<x,y<\infty$$

Solution: Given $(X,Y)\sim BN(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho)$. Then its p.d.f. is given by

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2}-2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2}+\frac{(y-\mu_2)^2}{\sigma_2^2}\right\}\right]--(1)$$

We have

$$f(x,y)=\frac{1}{2\pi\sqrt{3}}exp\left[-\frac{\{(2x-y)^2+2xy\}}{6}\right]\ ,-\infty<x,y<\infty$$

$$i.e,\quad f(x,y)=\frac{1}{2\pi\sqrt{3}}exp\left[-\frac{\{4x^2+y^2-2xy\}}{6}\right]----(\mathbf{2})$$

Comparing (1) and (2), we get $\mu_1=\mu_2=0$. Then (1) becomes

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}exp\left[-\frac{\left\{\frac{x^2}{\sigma_1{}^2}+\frac{y^2}{\sigma_2{}^2}-2\rho\frac{xy}{\sigma_1\sigma_2}\right\}}{2(1-\rho^2)}\right]—(3)$$

Comparing (2) and (3) , we find

$\sigma_1\sigma_2\sqrt{1-\rho^2}=\sqrt{3}\,,\sigma_1{}^2(1-\rho^2)=\frac{3}{4},\sigma_2{}^2(1-\rho^2)=3$ and $\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)}=\frac{1}{3}$

On solving we get $\sigma_1{}^2=1\,,\sigma_2{}^2=4,\rho^2=\frac{1}{4}$

Thus $cor(X,Y)=\rho=\sqrt{\frac{1}{4}}=\pm\frac{1}{2}$

Example 6: Determine the parameters of the bivariate normal distribution

$$\boldsymbol{f(x,y)=c\,exp\left[-\frac{\{16(x-2)^2-12(x-2)(y+3)+9(y+3)^2\}}{216}\right]}$$

Solution: If $(X,Y)\sim BN(\mu_1,\mu_2,\sigma_1{}^2,\sigma_2{}^2,\rho)$, then

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}exp\left[-\frac{\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)^2+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}}{2(1-\rho^2)}\right]$$

Comparing these functions, we get

$$\mu_1=2,\mu_2=-3,c=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},\frac{16}{216}=\frac{1}{2(1-\rho^2){\sigma_1}^2}$$

$$\frac{9}{216}=\frac{1}{2(1-\rho^2){\sigma_2}^2},\frac{12}{216}=\frac{2\rho}{2\sigma_1\sigma_2(1-\rho^2)}$$

$\therefore (1-\rho^2){\sigma_1}^2=\frac{27}{4}, (1-\rho^2){\sigma_2}^2=12, \sigma_1\sigma_2(1-\rho^2)=18\rho$

$\Longrightarrow (1-\rho^2)^2{\sigma_1}^2{\sigma_2}^2=81=(18\rho)^2\Longrightarrow \rho^2=\frac{1}{4},$

Further, $\sigma_1=3$ and $\sigma_2=4$.

Thus, $c=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}=\frac{1}{2\pi\times3\times4\sqrt{1-\frac{1}{4}}}=\frac{1}{12\pi\sqrt{3}}$

$\therefore (X,Y)\sim BN\left(2,3,9,16,\frac{1}{2}\right)$

Example 7: If $X\sim N(\mu,\sigma^2)$ and $(Y|x)\sim N(x,\sigma^2)$, show that

$$(X,Y)\sim BN(\mu,\mu,\sigma^2,2\sigma^2,\rho)\ .$$

Solution: We are given that

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],-\infty<x<\infty$$

$$g(y|x)=\frac{1}{\sqrt{2\pi}\sigma}exp\left[-\frac{1}{2}\left(\frac{y-x}{\sigma}\right)^2\right],-\infty<y<\infty$$

$$\therefore h(x,y)=g(y|x)f(x)=\frac{1}{2\pi\sigma^2}exp\left[-\frac{1}{2}\left\{\left(\frac{x-\mu}{\sigma}\right)^2+\left(\frac{y-x}{\sigma}\right)^2\right\}\right]$$

Consider $\left(\frac{y-x}{\sigma}\right)^2=\left(\frac{y-\mu+\mu-x}{\sigma}\right)^2=\left(\frac{y-\mu}{\sigma}\right)^2+\left(\frac{x-\mu}{\sigma}\right)^2-2\left(\frac{x-\mu}{\sigma}\right)\left(\frac{y-\mu}{\sigma}\right)$

Thus, $h(x,y)=\frac{1}{2\pi\sigma^2}exp\left[-\frac{1}{2}\left\{2\left(\frac{x-\mu}{\sigma}\right)^2+\left(\frac{y-\mu}{\sigma}\right)^2-2\left(\frac{x-\mu}{\sigma}\right)\left(\frac{y-\mu}{\sigma}\right)\right\}\right]$

The bivariate normal p.d.f. is given by

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}\right]$$

On comparing $h(x,y)$ with $f(x,y)$, we get

$$\sigma_1\sigma_2\sqrt{1-\rho^2}=\sigma^2 \quad , \quad {\sigma_1}^2(1-\rho^2)=\frac{1}{2}\sigma^2$$

$$\frac{\sigma_1\sigma_2(1-\rho^2)}{\rho}=\sigma^2 \quad , \quad {\sigma_2}^2(1-\rho^2)=\sigma^2 , \quad \mu_1=\mu_2=\mu$$

On solving, we get $\rho^2=\frac{1}{2}$ ${\sigma_2}^2=2\sigma^2$, ${\sigma_1}^2=\sigma^2$.

Thus, $(X,Y)\sim BN\left(\mu,\mu,\sigma^2,2\sigma^2,\frac{1}{\sqrt{2}}\right)$

Example 8: The variables X and Y are connected by the equation $aX+bY+c=0$. Show that the correlation between them is -1 if signs of a and b are same and $+1$ if they are different signs.

Solution: Given $aX+bY+c=0\Longrightarrow aE(X)+bE(Y)+c=0$

$\therefore a[X-E(X)]+b[Y-E(Y)]=0\Longrightarrow[X-E(X)]=-\frac{b}{a}[Y-E(Y)]$

$\therefore cov(X,Y)=E[\{X-E(X)\}\{Y-E(Y)\}]=-\frac{b}{a}E\big(Y-E(Y)\big)^2=-\frac{b}{a}{\sigma_Y}^2$ and

${\sigma_X}^2=E\big(X-E(X)\big)^2=\frac{b^2}{a^2}E\big(Y-E(Y)\big)^2=\frac{b^2}{a^2}{\sigma_Y}^2$

$$\therefore \rho=\frac{cov(X,Y)}{\sigma_X\cdot\sigma_Y}=\frac{-\frac{b}{a}{\sigma_Y}^2}{\sqrt{{\sigma_Y}^2}\sqrt{\frac{b^2}{a^2}{\sigma_Y}^2}}=\frac{-\frac{b}{a}{\sigma_Y}^2}{\left|\frac{b}{a}\right|{\sigma_Y}^2}=\frac{-\frac{b}{a}}{\left|\frac{b}{a}\right|}$$

$$\therefore \rho=\frac{cov(X,Y)}{\sigma_X\cdot\sigma_Y}=\begin{cases}1 , & if\ a\ and\ b\ have\ opposite\ signs\\ -1 , & if\ a\ and\ b\ have\ same\ signs\end{cases}$$

2.7 .Correlation coefficient and Bivariate Normal Distribution

Exercise:

1. Find the correlation coefficient between X and Y for each of the j.p.d.f. $f(x,y)$ of (X,Y) given below:

   (i) $f(x,y) = \begin{cases} \frac{3}{2}(x^2+y^2) & , \quad 0 \le x \le 1\,, \ 0 \le y \le 1 \\ 0 & , \quad otherwise \end{cases}$

   (ii) $f(x,y) = \begin{cases} (x+y) & , \quad 0 \le x\,, y \le 1 \\ 0 & , \quad otherwise \end{cases}$

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

2. If X, Y and Z are uncorrelated r.vs with 0 mean and standard deviations $5, 12$ and 9 repecitvely and $U = X + Y$ and $V = Y + Z$, then find the correlation coefficient between U and V.

3. If X, Y, Z are uncorrelated r.vs having same variance, find the correlation coefficient between $(X+Y)$ and $(Y+Z)$.

4. If the independent r.vs X and Y have variance 36 and 16 respectively, find the correlation coefficient between $(X+Y)$ and $(X-Y)$.

ANSWERS

1. (i) -0.2055 (ii) $-\frac{1}{11}$ (iii) 0.8

2. $\frac{48}{65}$

3. $\frac{1}{2}$

4. $\frac{5}{13}$

Unit – 3

Probability Inequalities and Generating Functions

3.1

Probability Inequalities

Inequalities are useful for bounding quanties that might otherwise be hard to compute. They will also be used in the **theory of convergence** and **limit theorems**.

Chebychev's Inequality

When we want to find the probability of an event described by a random variable, its c.d.f or p.d.f. or p.m.f. is required. If it is not known but its *mean* and *variance* are known, we can use **Chebychev's inequality** to find the **upper bound** or **lower bound** for the probability of the event.

Theorem 1: If X is a random variable with mean μ and variance σ^2, then

$$\boldsymbol{P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}} \qquad \dots\dots\dots\dots..(\mathbf{1})$$

where $\epsilon > 0$

Proof: The proof is given for a continuous random variable. Let X be a continuous r.v. with p.d.f. $f(x)$. Then

$$\sigma^2 = E\big(X - E(X)\big)^2 = E(X-\mu)^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

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

$$\geq \int_{-\infty}^{\mu-\epsilon} (x-\mu)^2 f(x)dx + \int_{\mu+\epsilon}^{\infty} (x-\mu)^2 f(x)dx$$

In the first integral,

$$x \leq \mu - \epsilon \Longrightarrow -x \geq -\mu + \epsilon \Longrightarrow -(x-\mu) \geq \epsilon \Longrightarrow (x-\mu)^2 \geq \epsilon^2$$

In the third integral, $x \geq \mu + \epsilon \Longrightarrow (x-\mu) \geq \varepsilon \Longrightarrow (x-\mu)^2 \geq \epsilon^2$

$$\therefore \sigma^2 \geq \epsilon^2 \left[\int_{-\infty}^{\mu-\epsilon} f(x)dx + \int_{\mu+\epsilon}^{\infty} f(x)dx\right]$$

$$= \epsilon^2[P(X \leq \mu - \epsilon) + P(X \geq \mu + \epsilon)]$$

$$= \epsilon^2 P[\mu - \epsilon \geq X \geq \mu + \epsilon] = \epsilon^2 P[-\epsilon \geq X - \mu \geq \epsilon]$$

$$= \epsilon^2 P[|X-\mu| \geq \epsilon]$$

Thus, $\sigma^2 \geq \epsilon^2 P[|X-\mu| \geq \epsilon] \Longrightarrow P[|X-\mu| \geq \epsilon] \leq \frac{\sigma^2}{\epsilon^2}$

Note: The proof is similar as in the case of d.r.v.X except that integration is replaced by summation.

Alternative forms:

$$P[|X-\mu| < \epsilon] = 1 - P[|X-\mu| \geq \epsilon] \geq 1 - \frac{\sigma^2}{\epsilon^2}$$

$$\Longrightarrow \boldsymbol{P}[|\boldsymbol{X} - \boldsymbol{\mu}| < \epsilon] \geq \mathbf{1} - \frac{\boldsymbol{\sigma^2}}{\boldsymbol{\epsilon^2}} \quad \dots\dots\dots\dots.(2)$$

Let $\epsilon = k\sigma$ for $k > 0$. Then from (1), we have

$$\boldsymbol{P}[|\boldsymbol{X} - \boldsymbol{\mu}| \geq \boldsymbol{k\sigma}] \leq \frac{\mathbf{1}}{\boldsymbol{k^2}} \quad \dots\dots\dots\dots\dots..(3)$$

and from (2), we have

$$\boldsymbol{P}[|\boldsymbol{X} - \boldsymbol{\mu}| < k\sigma] \geq \mathbf{1} - \frac{\mathbf{1}}{\boldsymbol{k^2}} \quad \dots\dots\dots\dots(4)$$

Example 1: If a r.v. X has mean 12 and variance 9 and the probability distribution is unknown, then find $P(6 < X < 18)$.

Solution: Since the probability distribution of X is not known, we can't find the value of the required probability. We can find only a lower bound for probability using Chebychev's inequality. We have, for $\epsilon > 0$.

$$P[|X - \mu| < \epsilon] \geq 1 - \frac{\sigma^2}{\epsilon^2}$$

Given $E(X) = \mu = 12$ and $V(X) = \sigma^2 = 9$.

Then $P[|X - 12| < \epsilon] \geq 1 - \frac{9}{\epsilon^2} \Rightarrow P[-\epsilon < (X - 12) < \epsilon] \geq 1 - \frac{9}{\epsilon^2}$

$$\Rightarrow P[12 - \epsilon < X < 12 + \epsilon] \geq 1 - \frac{9}{\epsilon^2}$$

Let $\epsilon = 6$. Then $P[6 < X < 18] \geq 1 - \frac{9}{36} = 1 - \frac{1}{4} = 0.75$

$$\Rightarrow P[6 < X < 18] \geq 0.75$$

Thus, the probability of X lying between 6 and 18 is alseast 75%.

Example 2: A d.r.v. X takes the values $-1, 0$ and 1 with probabilities $\frac{1}{8}, \frac{3}{4}$ and $\frac{1}{8}$ respectively. Evaluate $P[|X - \mu| \geq 2\sigma]$ and compare it with the upper bound given by Chebychev's inequality.

Solution: We have,

X	-1	0	1
$p(x)$	$\frac{1}{8}$	$\frac{3}{4}$	$\frac{1}{8}$

Then $E(X) = \mu = \sum xp(x) = -1 \times \frac{1}{8} + 0 \times \frac{3}{4} + 1 \times \frac{1}{8} = 0$

and $E(X^2) = \sum x^2 p(x) = 1 \times \frac{1}{8} + 0 \times \frac{3}{4} + 1 \times \frac{1}{8} = \frac{2}{8} = \frac{1}{4}$

Hence $\sigma^2 = V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{1}{4} - 0 = \frac{1}{4}$

Consider $P[|X-\mu| \geq 2\sigma] = P\left[|X| \geq 2 \cdot \frac{1}{2}\right] = P[|X| \geq 1]$

$$= P(X = -1,1)$$

$$= P(X = -1) + P(X = 1)$$

$$= \frac{1}{8} + \frac{1}{8} = \frac{2}{8} = \frac{1}{4} = 0.25$$

$\Longrightarrow P[|X-\mu| \geq 2\sigma] \leq 0.25$

On the other hand, by Chebychev's inequality,

$$P[|X-\mu| \geq 2\sigma] \leq \frac{1}{2^2} = \frac{1}{4}$$

Note that the two values are same.

Example 3: Use Chebychev's inequality to find how many times must a fair coin be tossed in order that the probability that the ratio of the number of heads to the number of tosses will lie between 0.45 and 0.55 will be at least 0.95.

Solution: Let X denote the number of heads obtained when a fair coin is tossed n times. Then $X \sim B(n,p)$. That is $E(X) = np$ and $V(X) = npq$.

Let $Y = \frac{X}{n}$. Then $E(Y) = E\left(\frac{X}{n}\right) = \frac{1}{n}E(X) = \frac{np}{n} = p$

and $V(Y) = V\left(\frac{X}{n}\right) = E\left(\left(\frac{X}{n}\right)^2\right) - \left(E\left(\frac{X}{n}\right)\right)^2 = \frac{1}{n^2}(E(X^2) - \left(E(X)\right)^2$

$$= \frac{1}{n^2}V(X) = \frac{npq}{n^2} = \frac{pq}{n}.$$

Since $p = \frac{1}{2}$ for a fair coin, $E(Y) = \frac{1}{2}$ and $V(Y) = \frac{1}{4n}$

By Chebychev's inequality for Y

$$P\left[\left|Y - \frac{1}{2}\right| < \epsilon\right] \geq 1 - \frac{\frac{1}{4n}}{\epsilon^2} = 1 - \frac{1}{4n\epsilon^2}$$

$$\Longrightarrow P\left[\frac{1}{2}-\epsilon < Y < \frac{1}{2}+\epsilon\right] \geq 1-\frac{1}{4n\epsilon^2}$$

Notice that, if $\epsilon = 0.05$ then $P(0.45 < Y < 0.55) \geq 1-\frac{1}{4n\epsilon^2}$

Now, find n when $\epsilon = 0.05$ and $1-\frac{1}{4n\epsilon^2} = 0.95 \Longrightarrow 1-\frac{1}{n\times 4\times(0.05)^2} = 0.95$

$\Longrightarrow 1-\frac{1}{0.01\times n} = 0.95 \Longrightarrow \frac{1}{0.01\times n} = 0.05 \Longrightarrow n = \frac{1}{0.01\times 0.05} = \frac{1}{0.0005} = \frac{10000}{5} = 2000$

Thus, $n = 2000$

Bienayme – Chebychev's inequality

Theorem 3: Let $g(X)$ be a non-negative function of a r.v. X . Then for every $\boldsymbol{k} > 0$, we have

$$\boldsymbol{P[g(X) \geq k] \leq \frac{E[g(X)]}{k}} \dots\dots\dots\dots.(\mathbf{1})$$

Proof: Here we shall prove the theorem for continuous random variable. The proof can be adapted to the case of discrete random variable on replacing integration by summation over the given range of the variable.

Let S be the set of all X for which $g(X) \geq k$.That is, $S = \{x \mid g(x) \geq k\}$.

Now, $E\left[g(X)\right] = \int_S g(x)f(x)dx,$ where $f(x)$ is the p.d.f. of X

$$\geq k\int_S f(x)dx \quad \left(on\ S, g(x) \geq k\right)$$

$$= kP\left[g(X) \geq k\right]$$

$$\Rightarrow P\left[g(X) \geq k\right] \leq \frac{E\left[g(X)\right]}{k}$$

Note:

1. If $g(X) = \left(X - E(X)\right)^2 = (X-\mu)^2$, then $E\left(g(X)\right) = V(X) = \sigma^2$ and replacing k by $\epsilon^2\sigma^2$ in equation (1), we get

$$\boldsymbol{P[(X-\mu)^2 \geq \epsilon^2\sigma^2] \leq \frac{\sigma^2}{\epsilon^2\sigma^2} = \frac{1}{\epsilon^2}}$$

$$\Rightarrow \boldsymbol{P[|X-\mu| \geq \epsilon\sigma] \leq \frac{1}{\epsilon^2}}$$

   which is **Chebychev's inequality.**

2. If $g(X) = |X|$ in (1), then we get for any $k > 0$,

$$\boldsymbol{P[|X| \geq k] \leq \frac{E(|X|)}{k}}$$

   which is known as **Markov's inequality.**

3. If $g(X) = |X|^r$ in (1), then we get

$$\boldsymbol{P[|X|^r \geq k^r] \leq \frac{E(|X|^r)}{k^r}}$$

   which is known as **generalized Markov's inequality.**

Cauchy – Schwartz Inequality

When the j.p.d.f. of X and Y is known, upper bound for expected value of the product of X and Y can be found by using Cauchy – Schwartz inequality when second moments about origin of X and Y are given ($i.e.$, $E(X^2)$ and $E(Y^2)$ are given).

Theorem 2: For any two random variables X and Y

$$\boldsymbol{\left(E(XY)\right)^2 \leq E(X^2)E(Y^2)}$$

Proof: Consider $E(X - tY)^2 \geq 0$ for any real number t. That is,

$E(X^2 - 2tXY + t^2Y^2) = E(X^2) - 2tE(XY) + t^2E(Y^2) \geq 0$
which is a quadratic expression in t. This expression is always positive only when t

has complex roots. This is possible only when discriminant of the expression is negative. Thus,

$$4\left(E(XY)\right)^2 - 4E(X^2)E(Y^2) \le 0$$

$$\Rightarrow \left(E(XY)\right)^2 \le E(X^2)E(Y^2)$$

Hence the result.

Example 4: The j.p.d.f. of (X, Y) is given by

$$f(x, y) = \frac{x+y}{21} \text{ for } x = 1, 2, 3 \text{ and } y = 1, 2.$$

Verify whether Cauchy-Schwartz inequality.

Solution: The joint and marginal p.m.fs $f_1(x)$ and $f_2(y)$ of X and Y respectively are given in the following table.

$Y \backslash X$	1	2	3	$f_2(y)$
1	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{9}{21}$
2	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{12}{21}$
$f_1(x)$	$\frac{5}{21}$	$\frac{7}{21}$	$\frac{9}{21}$	1

$$E(X) = \sum_{x=1}^{3} x f_1(x) = 1 \times \frac{5}{12} + 2 \times \frac{7}{21} + 3 \times \frac{9}{21} = \frac{46}{21}$$

$$E(X^2) = \sum_{x=1}^{3} x^2 f(x) = 1^2 \times \frac{5}{12} + 2^2 \times \frac{7}{21} + 3^2 \times \frac{9}{21} = \frac{114}{21}$$

$$E(Y) = \sum_{y=1}^{2} y f_2(y) = 1 \times \frac{9}{21} + 2 \times \frac{12}{21} = \frac{33}{21}$$

$$E(Y^2) = \sum_{y=1}^{2} y^2 f_2(y) = 1^2 \times \frac{9}{21} + 2^2 \times \frac{12}{21} = \frac{57}{21}$$

$$E(XY) = \sum_{x=1}^{3}\sum_{y=1}^{2} xyf(x,y)$$

$$= 1 \times 1 \times \frac{2}{21} + 1 \times 2 \times \frac{3}{21} + 1 \times 3 \times \frac{4}{21} + 2 \times 1 \times \frac{3}{21} + 2 \times 2 \times \frac{4}{21} + 2 \times 3 \times \frac{5}{21}$$

$$= \frac{1}{21}(2 + 6 + 12 + 6 + 16 + 30) = \frac{71}{21}$$

Verification of Cauchy-Schwatz inequality:

Here $\left(E(XY)\right)^2 = \left(\frac{71}{21}\right)^2 = 11.755$ and $E(X^2)E(Y^2) = \frac{114}{21} \times \frac{57}{21} = 14.735$

Note that $\left(E(XY)\right)^2 \leq E(X^2)E(Y^2)$.

Example 5: Let X and Y be c.r.vs with j.p.d.f.

$$\boldsymbol{f(x,y) = \begin{cases} x + y, & 0 < x < 1, 0 < y < 1 \\ 0, & otherwise \end{cases}}$$

Verify Cauchy-Schwartz inequality.

Solution: The m.p.d.f. of X is given by

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

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

Since $f(x,y)$ is symmetric in x and y, the m.p.d.f. of Y is given by

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

Now,

$$E(X^2) = \int_0^1 x^2\left(x+\frac{1}{2}\right)dx = \int_0^1 \left(x^3+\frac{x^2}{2}\right)dx = \left[\frac{x^4}{4}+\frac{x^3}{6}\right]_0^1 = \frac{1}{4}+\frac{1}{6} = \frac{10}{24}$$

Similarly, $E(Y^2) = \frac{10}{24}$ .Now,

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyf(x,y)dxdy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy(x+y)dxdy$$

$$= \int_0^1 \left\{\int_0^1 (x^2y+xy^2)dx\right\}dy = \int_0^1 \left(\frac{x^3y}{3}+\frac{x^2y^2}{2}\right)_0^1 dy$$

$$= \frac{1}{6}\int_0^1 (2y+3y^2)dy = \left[\frac{y^2}{6}+\frac{y^3}{6}\right]_0^1 = \frac{2}{6} = \frac{1}{3}$$

Thus $\left(E(XY)\right)^2 = \left(\frac{1}{3}\right)^2 = \frac{1}{9} = 0.111$, and $E(X^2)E(Y^2) = \frac{10}{24}\times\frac{10}{24} = 0.1736$

Hence $\left(E(XY)\right)^2 \leq E(X^2)E(Y^2)$.

P1:

A symmetric die is thrown 600 times. Find the lower bound for the probability of getting 80 to 120 sixes.

Solution:

Let X be the total number of sixes.

Then $X \sim B\left(600, \frac{1}{6}\right)$, $E(X) = np = 600 \times \frac{1}{6} = 100$

and $V(X) = np(1-p) = 600 \times \frac{1}{6} \times \frac{5}{6} = \frac{500}{6}$.

Using Chebychev's inequality, we get

$$P\{|X - E(X)| < k\sigma\} \geq 1 - \frac{1}{k^2} \Rightarrow P\left\{|X - 100| < k\sqrt{\frac{500}{6}}\right\} \geq 1 - \frac{1}{k^2}$$

Therefore, $P\left\{100 - k\sqrt{\frac{500}{6}} < X < 100 + k\sqrt{\frac{500}{6}}\right\} \geq 1 - \frac{1}{k^2}$

Taking $k = \frac{20}{\sqrt{\frac{500}{6}}}$, we get $P(80 < X < 120) \geq 1 - \frac{1}{400 \times \left(\frac{6}{500}\right)} = \frac{19}{24}$

The lower bound for the probability of getting 80 to 120 sixes.

P2:

For geometric distribution $p(x) = 2^{-x}$; $x = 1, 2, 3, \ldots$, prove that Chebychev's inequality gives $P\{|X-2| \le 2\} > \frac{1}{2}$, while the actual probability is $\frac{15}{16}$.

Solution:

$$E(X) = \sum x\,p(x) = \sum_{x=1}^{\infty} \frac{x}{2^x} = \frac{1}{2} + \frac{2}{2^2} + \frac{3}{2^3} + \frac{4}{2^4} + \ldots$$

$$= \frac{1}{2}\left(1 + 2A + 3A^2 + 4A^3 + \ldots\right) = \frac{1}{2}(1-A)^{-2} = 2, \left(A = \frac{1}{2}\right)$$

$$E(X^2) = \sum x^2\,p(x) = \sum_{x=1}^{\infty} \frac{x^2}{2^x} = \frac{1}{2} + \frac{4}{2^2} + \frac{9}{2^3} + \ldots$$

$$= \frac{1}{2}\left\{1 + 4A + 9A^2 + \ldots\right\}, where\, A = \frac{1}{2}$$

$$= \frac{1}{2}(1+A)(1-A)^{-3} = 6$$

$\therefore Var(X) = \sigma^2 = E(X^2) - \{E(X)\}^2 = 6 - 4 = 2 \Longrightarrow \sigma = \sqrt{2}$

Using Chebychev's inequality, we get $P\{|X - E(X)| > k\sigma|\} \le \frac{1}{k^2}$

With $k = \sqrt{2}$, we get $P\{|X-2| > \sqrt{2}\cdot\sqrt{2}\} \le \frac{1}{2} \Longrightarrow P[|X-2| \le 2] > 1 - \frac{1}{2} = \frac{1}{2}$

The actual probability is given by

$P\{|X-2| \le 2\} = P\{0 \le X \le 4\} = P\{X = 1,2,3 \;or\; 4\}$

$$= \frac{1}{2} + \left(\frac{1}{2}\right)^2 + \left(\frac{1}{2}\right)^3 + \left(\frac{1}{2}\right)^4 = \frac{15}{16}$$

P3:

Does there exist a variate X for which

$$\boldsymbol{P[\mu - 2\sigma \leq X \leq \mu + 2\sigma] = 0.6} \quad \text{……….} \ (\mathbf{1})$$

Solution:

We have

$$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma) = P(|X - \mu| \leq 2\sigma)$$

By Chebychev's inequality $P(|X - \mu| \leq 2\sigma) \geq 1 - \frac{1}{2^2} = 0.75$

Since lower bound for the probability is 0.75, there does not exist a r.v. X for which the equation (1) holds.

P4:

(a) For a p.m.f

$$\boldsymbol{p(x)} = \begin{cases} \frac{1}{8} & , \quad \boldsymbol{x = -1} \\ \frac{6}{8} & , \quad \boldsymbol{x = 0} \\ \frac{1}{8} & , \quad \boldsymbol{x = 1} \end{cases}$$

Find $\boldsymbol{P(|X - \mu| \geq 2\sigma)}$.

(b) Compare this result with that obtained by using Chebychev's inequality.

Solution:

(a)

x:	-1	0	1
$p(x)$:	$\frac{1}{8}$	$\frac{6}{8}$	$\frac{1}{8}$

$\therefore \mu = E(X) = \sum x\, p(x) = -1 \times \frac{1}{8} + 1 \times \frac{1}{8} = 0$ and

$$E(X^2) = \sum x^2 p(x) = 1 \times \frac{1}{8} + 1 \times \frac{1}{8} = \frac{1}{4}$$

$$\therefore Var(X) = E(X^2) - \{E(X)\}^2 = \frac{1}{4} \Longrightarrow \sigma = \frac{1}{2}$$

$$P\{|X - \mu| \geq 2\sigma\} = P\{|X| \geq 1\} = 1 - P(|X| < 1)$$

$$= 1 - P(-1 < X < 1) = 1 - P(X = 0) = \frac{1}{4}$$

(b) $P\{|X - \mu| \geq 2\sigma\} \leq \frac{1}{2^2}$ (By Chebychev's inequality: $P\{|X - \mu| \geq k\sigma\} \leq \frac{1}{k^2}$)

In this case, both results are same.

Note: This example shows that, in general, Chebychev's inequality cannot be improved.

3.1 .Probability inequalities

Exercise:

1. The Chebychev's inequality for random variable X is $(-2 < X < \infty) \geq \frac{21}{25}$, find $E(X)$ and $V(X)$.

2. Two unbiased dice are thrown. If X is the sum of the numbers showing up, prove that $P(|X - 7| \geq 3) \leq \frac{35}{54}$. Compare this with the actual probability.

3. If X is the number scored in a throw of a fair die, find the upper bound for $P(|X - \mu| \geq 2.5)$ where $\mu = E(X)$. Also find the actual probability.

4. If X is a r.v. such that $E(X) = 3$ and $E(X^2) = 13$, find the lower bound of $P(-2 \leq X \leq 8)$.

5. A discrete random variable X is specified by $p(-a) = p(a) = \frac{1}{8}$ and $p(0) = \frac{3}{4}$. Compute
   (i) $P(|X| \geq 2\sigma)$ and
   (ii) Chebychev's inequality bound.

Answers:

1. $E(X) = 3$ and $V(X) = 4$
2. Actual probability $= \frac{1}{3}$
3. Upper bound $= 0.47$ and actual probability $= 0$
4. $\frac{21}{25}$
5. (i) $\frac{1}{4}$ (ii) $\frac{1}{4}$

Probability & Statistics
HW 7 - Chebyshev & Empirical & Z score

Name Key

1. The average IQ of the students in one calculus class is 110, with a standard deviation of 5; the average IQ of students in another class is 106, with a standard deviation of 4. A student has an IQ of 112, in which class is he ranked higher?

$$z_1 = \frac{112-110}{5} \qquad z_2 = \frac{112-106}{4}$$

$$z_1 = 0.4 \qquad z_2 = 1.5 \quad \rightarrow \quad \boxed{\text{Ranked higher in 2nd Class.}}$$

2. The average price of the wagon at Car Dealer A is $25,000, with a standard deviation of $4000. The average price at Car Dealer B is $20,000, with a standard deviation of $2000. If I spent $23,000 at dealer A and my sister paid $18,000 at dealer B, which was a better deal?

$$z_A = \frac{23{,}000 - 25{,}000}{4{,}000} \qquad z_B = \frac{18{,}000 - 20{,}000}{2\,000}$$

$$z_A = -0.5 \qquad z_B = -1 \quad \rightarrow \quad \boxed{\text{Dealer B was the better deal.}}$$

3. The average score on an English final examination was 85, with a standard deviation of 5; the average score on a history final exam was 110, with a standard deviation of 8. I made an 80 on the English test and a 100 on the history test. Which test was better?

$$z_E = \frac{80-85}{5} \qquad z_H = \frac{100-110}{8}$$

$$z_E = -1 \qquad z_H = -1.25 \quad \rightarrow \quad \boxed{\text{English}}$$

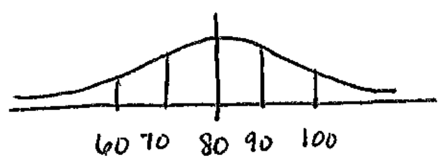
4. The average age of the accountants at Three Rivers Corp. is 26, with a standard deviation of 6; the average salary of the accountants is $31,000 with a standard deviation of $4000. I'm applying at both places and I'm 31 and want to make $35,000, in which category will I be higher?

$$z_{Age} = \frac{31-26}{6} \qquad z_{salary} = \frac{35000-31000}{4000}$$

$$z_A = 0.83 \qquad z_S = 1 \quad \rightarrow \quad \boxed{\text{Salary}}$$

5. Using Chebyshev's theorem, solve the following problems for a distribution with a mean of 80 and a standard deviation of 10

a. At least what percentage of values will fall between 60 and 100?



$$K = 2 \rightarrow \left(1 - \frac{1}{2^2}\right)(100) = \boxed{75\%}$$

b. At least what percentage of values will fall between 65 and 95?

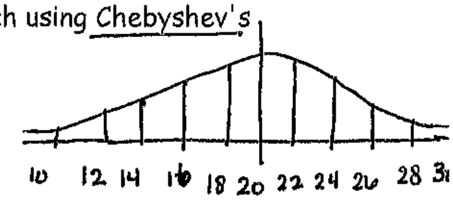
$$K = 1.5 \rightarrow \left(1 - \frac{1}{1.5^2}\right)(100) = \boxed{56\%}$$

$$z = \frac{95-80}{10} \qquad z = \frac{65-80}{10}$$

$$z = 1.5 \qquad z = -1.5$$

Thus, $K = 1.5$

6. The mean of a distribution is 20 and the standard deviation is 2. Answer each using Chebyshev's theorem.



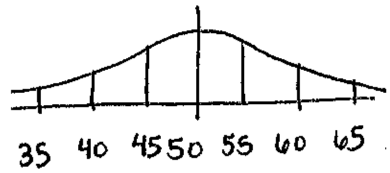
a. At least what percentage of the values will fall between 10 and 30?

$K = 5$ from Graph or $z = \frac{30-20}{2}$, $z = 5$ $\Big\}$ $\left[1 - \frac{1}{(5)^2}\right](100) = \boxed{96\%}$

b. At least what percentage of the values will fall between 12 and 28?

$K = 4$ or $z = \frac{28-20}{2}$, $z = 4$ $\Big\}$ $\left[1 - \frac{1}{(4)^2}\right](100) = \boxed{93.75\%}$

7. The mean of a distribution is 50 and the standard deviation is 5. Answer each using The Empirical Rule.



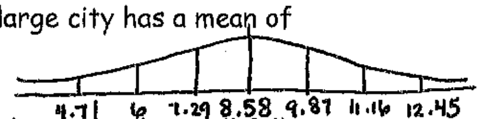
a. At least what percentage of the values will fall between 45 and 55?

1 st dev $\rightarrow$ $\boxed{68\%}$ OR $z = \frac{55-50}{5}$, $z = 1$

b. At least what percentage of the values will fall between 35 and 65?

3 st dev $\rightarrow$ $\boxed{99.7\%}$

8. A sample of the hourly wages of employees who work in restaurants in a large city has a mean of \$8.58 and a standard deviation of \$1.29.



a. Using Chebyshev's theorem, find the range in which at least 75% of the data values will fall.

$.75 = 1 - \frac{1}{K^2}$

$-\frac{1}{K^2} = -0.25$ $\rightarrow$ $-0.25K^2 = -1$, $K^2 = 4$, $K = 2$

$\therefore$ $8.58 \pm 2(1.29)$

$\boxed{\$6 \text{ to } \$11.16}$

b. Using the Empirical Rule, find the range in which at least 95% of the data values will fall.

95% $\rightarrow$ 2 st dev   Thus, range is $\boxed{\$6 \text{ to } \$11.16}$

9. A sample of the labor costs per hour to assemble a certain product has a mean of \$15.72 and a standard deviation of \$2.15. Using Chebyshev's theorem, find the values in which at least 88.89% of the data will lie.

$.8889 = 1 - \frac{1}{K^2}$

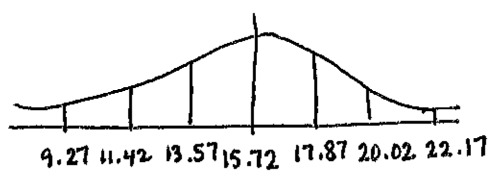
$-0.1111 = -\frac{1}{K^2}$

$0.1111K^2 = 1$

$K^2 = 9 \rightarrow K = 3$

$\therefore$ $15.72 \pm 3(2.15)$

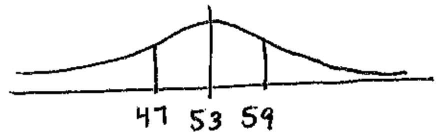
$\boxed{\$9.27 \text{ to } \$22.17}$



10. The average score on a special test of knowledge of wood refinishing has a mean of 53 and a standard deviation of 6. Using the Empirical Rule, find the values in which at least 68% of the data will lie.

68% $\rightarrow$ 1st dev $\therefore$ $53 \pm 1(6)$

$\boxed{47 \text{ to } 59}$



3.2

Moment Generating Function

Certain derivations presented in modules 2.4, 2.5 and 2.6 have been somewhat heavy on algebra. For example, determining the mean and variance of the **Binomial distribution** turned out to be fairly tiresome. Another example of hard work was determining the set of probabilities associated with a sum , $P(X+Y=t)$. Many of these tasks are greatly simplified by using **probability generating functions.**

Moment Generating Function: The moment generating function (m.g.f) of a random variable X is dentoed by $M_X(t)$ and it is defined as

$$\boldsymbol{M_X(t) = E(e^{tX})}$$

$$\therefore\ M_X(t) = E(e^{tX}) = E\left[1 + \frac{tX}{1!} + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots\right]$$

$$= E(1) + \frac{t}{1!}E(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \cdots + \frac{t^r}{r!}E(X^r) + \cdots + \infty$$

$$\therefore\ M_X(t) = 1 + \frac{t}{1!}\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \cdots + \frac{t^r}{r!}\mu_r' + \cdots + \infty \qquad \ldots\ldots \quad (1)$$

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu_r'$$

which gives the m.g.f in terms of moments.

Therefore the coeffceint of $\frac{t^r}{r!}$ in $M_X(t)$ is μ_r' , where $r = 1, 2, 3, \ldots$ and $\mu_r' = E(X^r)$, moment about origin.

The m.g.f of X about mean $\mu = \mu_1' = E(X)$ is defined as

$$M_{X-\mu}(t) = E\left[e^{t(X-\mu)}\right] = E\left[1 + \frac{t}{1!}(X-\mu) + \frac{t^2}{2!}(X-\mu)^2 + \frac{t^3}{3!}(X-\mu)^3 + \cdots\right]$$

$$= 1 + \frac{t}{1!}E(X-\mu) + \frac{t^2}{2!}E(X-\mu)^2 + \frac{t^3}{3!}E(X-\mu)^3 + \cdots$$

$$= 1 + \frac{t}{1!}\mu_1 + \frac{t^2}{2!}\mu_2 + \frac{t^3}{3!}\mu_3 + \cdots$$

where $E(x-\mu)^r = \mu_r$ is known as the r^{th} central moment for $r = 1, 2, \ldots$

Note that $\mu_1 = E(X-\mu) = E(X) - \mu = \mu - \mu = 0$

Since $M_X(t)$ generates moments, it is called **moment generating function**.

If X is a discrete random variable with p.m.f. $p(x)$ then

$$M_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx} p(x)$$

If X is a continuous random variable with p.d.f. $f(x)$, then

$$M_X(t) = E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

Moments Using Moment Generating Function:

Differentiating equation (1) with respect to t and then putting $t = 0$, gives

$$\mu_1' = \left[\frac{d}{dt} M_X(t)\right]_{t=0}$$

$$\mu_2' = \left[\frac{d^2}{dt^2} M_X(t)\right]_{t=0}$$

In general,

$$\mu_r' = \left[\frac{d^r}{dt^r} M_X(t)\right]_{t=0}, \; r = 1, 2, 3, \ldots$$

Note: Moment generating function $M_X(t)$ is used to calculate the higher moments.

Theorems on Moment Generating Function:

Theorem 1: $\boldsymbol{M_{aX}(t) = M_X(at)}$, where $\boldsymbol{a}$ is a constant.

Proof: By definition $M_{aX}(t) = E(e^{taX}) = E(e^{atX}) = M_X(at)$

Therefore, $M_{aX}(t) = M_X(at)$

Theorem 2: The moment generating function of the sum of n independent random variables is equal to the product of their respective moment generating functions, $\boldsymbol{i.e.}$, $\boldsymbol{M_{X_1+X_2+X_3+\cdots+X_n}(t) = M_{X_1}(t)M_{X_2}(t)M_{X_3}(t)\dots M_{X_n}(t)}$

Proof: By definition,

$$M_{X_1+X_2+X_3+\cdots+X_n}(t) = \left[E^{t(X_1+X_2+X_3+\cdots+X_n)}\right]$$

$$= E(e^{tX_1})E(e^{tX_2})E(e^{tX_3})\dots E(e^{tX_n})$$

(Since $X_1, X_2, \dots, X_n$ are independent).

Therefore, $M_{X_1+X_2+X_3+\cdots+X_n}(t) = M_{X_1}(t)M_{X_2}(r)\dots M_{X_n}(t)$
Hence the proof.

Uniqueness Theorem of Moment Generating Function:

The m.g.f. of a distribution, if exists, uniquely determines the distribution. This implies that corresponding to a given probability distribution, there is only one m.g.f (provided it exists) and corresponding to a given m.g.f, there is only one probability distribution. Hence $M_X(t) = M_Y(t) \Longrightarrow X$ and Y are identically distributed.

Effect of Change of Origin and Scale on Moment Generating Function:

Let a random variable X be transformed to a new variable U by changing both the origin and scale in X as $= \frac{X-a}{h}$, where a and h are constants.

The m.g.f of U (about origin) is given by

$$M_U(t) = E(e^{tU}) = E\left[e^{t\left(\frac{X-a}{h}\right)}\right] = E\left(e^{\left(\frac{tX}{h}-\frac{ta}{h}\right)}\right) = e^{-\frac{at}{h}}E\left(e^{\left(\frac{tX}{h}\right)}\right)$$

$$\therefore\ \boldsymbol{M_{\frac{X-a}{h}}(t) = e^{-\frac{at}{h}}M_X\left(\frac{t}{h}\right)}$$

Note: If $Y = aX + b$, then $M_Y(t) = e^{bt}M_X(at)$

Example 1: If X represents the outcome when a fair die is tossed, find the m.g.f. of X and hence, find $E(X)$ and $Var(X)$.

Solution: When a fair die is tossed

$$P(X = x) = \frac{1}{6}, \quad x = 1, 2, 3, 4, 5, 6$$

$$\therefore \quad M_X(t) = \sum_{X=1}^{6} e^{tx} P(X = x) = \frac{1}{6} \sum_{x=1}^{6} e^{tx}$$

$$= \frac{1}{6}(e^t + e^{2t} + e^{3t} + e^{4t} + e^{5t} + e^{6t})$$

$$E(X) = \left[\frac{d}{dt} M_X(t)\right]_{t=0} = \frac{1}{6}[e^t + 2e^{2t} + 3e^{3t} + 4e^{4t} + 5e^{5t} + 6e^{6t}]_{t=0}$$

$$= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2}$$

$\therefore$ Mean $= E(X) = \frac{7}{2}$

Now, $E[X^2]_{t=0} = \left\{\frac{d^2}{dt^2}[M_X(t)]\right\}_{t=0}$

$$= \frac{1}{6}[e^t + 4e^{2t} + 9e^{3t} + 16e^{4t} + 25e^{5t} + 36e^{6t}]_{t=0}$$

$$= \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = \frac{91}{6}$$

$$V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

Example 2: Find the m.g.f. of the random variable X whose probability function $P(X = x) = \frac{1}{2^x}, x = 1, 2, 3, \ldots$ and hence find its mean.

Solution: By definition,

$$M_X(t) = E\left(e^{tX}\right) = \sum_{x=0}^{\infty} e^{tx} P(X = x) = \sum_{x=1}^{\infty} e^{tX}\left(\frac{1}{2^x}\right) = \sum_{x=1}^{\infty}\left(\frac{e^t}{2}\right)^x$$

$$=\left[\frac{e^t}{2}+\left(\frac{e^t}{2}\right)^2+\left(\frac{e^t}{2}\right)^3+\cdots\right]$$

$$=\frac{e^t}{2}\left[1+\frac{e^t}{2}+\left(\frac{e^t}{2}\right)^2+\cdots\right]=\frac{e^t}{2}\left(1-\frac{e^t}{2}\right)^{-1}$$

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

Therefore, $M_X(t)=\frac{e^t}{2-e^t}$

$$\mu_1'=\left[\frac{d}{dt}M_X(t)\right]_{t=0}=\left[\frac{d}{dt}\left(\frac{e^t}{2-e^t}\right)\right]_{t=0}=\left[\frac{(2-e^t)e^t-e^t(-e^t)}{(2-e^t)^2}\right]_{t=0}=\frac{(2-1)1+1}{(2-1)^2}=2$$

Thus, $E(X)$ =mean= 2

Example 3: If the moments of a random variable X are defined by $E(X^r)=0.6, r=1,2,\ldots$ .Show that $P(X=0)=0.4$, $P(X=1)=0.6$, and $P(X\geq 2)=0$.

Solution: We know that

$$M_X(t)=\sum_{r=0}^{\infty}\frac{t^r}{r!}\mu_r'$$

where $\mu_r'=E(X^r)=0.6$

$$\therefore\ M_X(t)=1+\sum_{r=1}^{\infty}\frac{t^r}{r!}\mu_r'=1+\sum_{r=1}^{\infty}\frac{t^r}{r!}(0.6)=1+(0.6)\left(\frac{t}{1!}+\frac{t^2}{2!}+\frac{t^3}{3!}+\cdots\right)$$

$$=1+(0.6)(e^t-1)=1-0.6+0.6e^t=0.4+0.6e^t \qquad \ldots\ldots.(1)$$

But by definition,

$$M_X(t)=E\left(e^{tx}\right)=\sum_{r=0}^{\infty}e^{tx}P(X=x)$$

$M_X(t)=P(X=0)+e^tP(X=1)+e^{2t}P(X=2)+e^{3t}P(X=3)+$...+... ...... (2)

From equations (1) and (2), we have

$$0.4 + 0.6e^t = P(X=0) + e^t P(X=1) + e^{2t} P(X=2) + e^{3t} P(X=3) + \cdots$$

Equating the coefficients of like terms on both sides,

$$P(X=0) = 0.4,\ P(X=1) = 0.6$$

$$P(X=2) = P(X=3) = \cdots = 0 \quad \Longrightarrow P(X \geq 2) = 0$$

Example 4: Find the m.g.f. of a random variable whose moments are $\boldsymbol{\mu_r = (r+1)!\ 2^r}$.

Solution: By definition, we have $M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\ \mu_r'$

$$= \sum_{r=0}^{\infty} \frac{t^r}{r!}\ (r+1)!\,2^r = \sum_{r=0}^{\infty} (r+1)(2t)^r$$

$$= 1 + 2(2t) + 3(2t)^2 + \cdots = (1-2t)^{-2} = \frac{1}{(1-2t)^2}$$

$$\therefore \quad M_X(t) = \frac{1}{(1-2t)^2}$$

Example 5: If $\boldsymbol{X \sim B(n,p)}$, find the m.g.f of $\boldsymbol{X}$ and hence find its mean and variance.

Solution: Since $X \sim B(n,p)$, its p.m.f. is given by

$p(x) = \binom{n}{x} p^x q^{n-x},\ x = 0, 1, 2, \dots, n$ and $q = 1 - p$.

Then the m.g.f. of X is given by

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{n} e^{tx} p(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} \left(pe^t\right)^x q^{n-x}$$

$$\Longrightarrow M_X(t) = (q + pe^t)^n$$

$\frac{d}{dt} M_X(t) = n(q + pe^t)^{n-1} pe^t \Longrightarrow$Mean$= \mu_1' = \left[\frac{dM_X(t)}{dt}\right]_{t=0} = np$

Next, $\frac{d^2}{dt^2}\left(M_X(t)\right) = np[(n-1)(q+pe^t)^{n-2}pe^{2t} + (q+pe^t)^{n-1}e^t]$

$\Longrightarrow\ \mu_2' = \left[\frac{d^2}{dt^2}\left(M_X(t)\right)\right]_{t=0} = np[(n-1)p+1] = np[np-p+1] = np(np+q)$

$\Longrightarrow\ \mu_2' = n^2p^2 + npq$

Now, variance $= \sigma^2 = \mu_2' - (\mu_1')^2 = n^2p^2 + npq - n^2p^2 = npq$

Thus, $\mu = np$ and $\sigma^2 = npq$

Note that $\sigma^2 = npq = \mu q$ where $(0 < q < 1)$ .Thus, $\mu > \sigma^2$.

Note: *For binomial distribution, mean is always greater than variance.*

Example 6 : If $X \sim P(\lambda)$, find its m.g.f. and hence find its mean and variance.

Solution: Since $X \sim P(\lambda)$, then its p.m.f. is given by

$$p(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{\lambda!}\ , x = 0, 1, \dots \text{ and } \lambda > 0$$

The m.g.f. of X is given by

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tx} p(x) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^t\right)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

$\Longrightarrow\ M_X(t) = e^{\lambda(e^t-1)}$

Since $\frac{d}{dt}\left(M_X(t)\right) = e^{\lambda(e^t-1)}\lambda e^t$; Mean$= \mu = \mu' = \left[\frac{d}{dt}\left(M_X(t)\right)\right]_{t=0} = \lambda$.

Now, $\frac{d^2}{dt^2}\left(M_X(t)\right) = \lambda\left[e^{\lambda(e^t-1)}e^t + e^{\lambda(e^t-1)}\lambda e^{2t}\right]$

Then $\mu_2' = \left[\frac{d^2 M_X(t)}{dt^2}\right]_{t=0} = \lambda(1+\lambda) = \lambda + \lambda^2$

Thus, variance $= \sigma^2 = \mu_2' - (\mu_1')^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$

Therefore, $\mu = \sigma^2 = \lambda$

Note: *Mean and variance are same for Poisson distribution.*

Example 7: If $X \sim NB(r, p)$, find its m.g.f. and hence find its mean and variance.

Solution: Since $X \sim NB(r, p)$, its p.m.f. is given by

$$p(x) = \binom{-r}{x} p^r (-q)^x, \qquad x = 0, 1, 2, \dots$$

The m.g.f of X is given by

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \binom{-r}{x} p^r (-q)^x$$

$$= \sum_{x=0}^{\infty} \binom{-r}{x} p^r \left(-qe^t\right)^x = p^r \sum_{x=0}^{\infty} \binom{-r}{x} \left(-qe^t\right)^x$$

$\Rightarrow \quad M_X(t) = p^r (1 - qe^t)^{-r}.$

Now, $\frac{d}{dt}\left(M_X(t)\right) = p^r(-r)(1 - qe^t)^{-(r+1)}(-qe^t) = qrp^r(1 - qe^t)^{-(r+1)}e^t$

Mean$= \mu = \mu_1' = \left[\frac{d}{dt}\left(M_X(t)\right)\right]_{t=0} = qrp^r(1-q)^{-(r+1)} = qrp^r(p)^{-(r+1)} = \frac{rq}{p}$

Further, $\frac{d^2}{dt^2}\left(M_X(t)\right) = rqp^r \frac{d}{dt}\left\{(1 - qe^t)^{-(r+1)}e^t\right\}$

$$= rqp^r\left\{-(r+1)(1 - qe^t)^{-(r+2)}(-qe^{2t}) + (1 - qe^t)^{-(r+1)}e^t\right\}$$

Then $\mu_2' = \left[\frac{d^2}{dt^2}\left(M_X(t)\right)\right]_{t=0} = rqp^r\left[(r+1)qp^{-(r+2)} + p^{-(r+1)}\right]$

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

$\Rightarrow \quad \mu_2' = \dfrac{r^2q^2}{p^2} + \dfrac{rq}{p^2}$

Hence, variance $= \sigma^2 = \mu_2' - (\mu_1')^2 = \frac{r^2q^2}{p^2} + \frac{rq}{p^2} - \frac{r^2q^2}{p^2} \Rightarrow \quad \sigma_2^2 = \frac{rq}{p^2}$

Example 8:Let X be a random variable with p.d.f.

$$f(x) = \begin{cases} \frac{1}{3} e^{-\frac{x}{3}} & , \quad x > 0 \\ 0 & , \quad otherwise \end{cases}$$

Find

(i) $P(X > 3)$

(ii) **M.g.f. of X**

(iii) $E(X)$ **and** $Var(X)$

Solution:

(i) $$P(X > 3) = \int_3^\infty f(x)\, dx = \int_3^\infty \frac{1}{3} e^{-\frac{x}{3}} dx = \frac{1}{3}\left[\frac{e^{-\frac{x}{3}}}{-\frac{1}{3}}\right]_3^\infty = -(0 - e^{-1}) = e^{-1} = \frac{1}{e}$$

(ii) $$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} f(x) dx$$

$$= \int_0^\infty e^{tx} \frac{1}{3} e^{-\frac{x}{3}} dx = \frac{1}{3}\int_0^\infty e^{\left(t-\frac{1}{3}\right)x} dx = \frac{1}{3}\int_0^\infty e^{-\left(\frac{1}{3}-t\right)x} dx = \frac{1}{3}\left[\frac{e^{-\left(\frac{1}{3}-t\right)x}}{-\left(\frac{1}{3}-t\right)}\right]_0^\infty$$

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

$$M_X(t) = \frac{1}{1-3t} = (1-3t)^{-1}$$

$$\frac{d}{dt}[M_X(t)] = -(1-3t)^{-2}(-3) = 3(1-3t)^{-2}$$

(iii) $$E(X) = Mean = \left[\frac{d}{dt} M_X(t)\right]_{t=0} = 3$$

$$\frac{d^2}{dt^2}[M_X(t)] = -6(1-3t)^{-3}(-3) = 18(1-3t)^{-3}$$

$$E(X^2) = \left[\frac{d^2}{dt^2} M_X(t)\right]_{t=0} = 18$$

$$Var(X) = E(X^2) - [E(X)]^2 = 18 - 9 = 9$$

Example 9: Let X be a discrete random variable with p.d.f.

$$\boldsymbol{p(x) = \begin{cases} \dfrac{1}{x(x+1)} & , \quad x = 1, 2 \ldots \\ 0 & , \quad otherwise \end{cases}}$$

Show that $E(X)$ does not exist even though m.g.f. exist.

Solution:

$$E(X) = \sum_{x=1}^{\infty} x\, p(x) = \sum_{x=1}^{\infty} \frac{1}{x+1} = \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots = \sum_{x=1}^{\infty} \frac{1}{x} - 1$$

But $\boldsymbol{\sum_{x=1}^{\infty} \frac{1}{x}}$ is a divergent series.

Therefore, $E(x)$ does not exist and hence, no moment exists.

Now, m.g.f. of X is given by

$$M_X(t) = \sum_{x=1}^{\infty} p(x) e^{tx} = \sum_{x=1}^{\infty} \frac{e^{tx}}{x(x+1)}$$

Substituting $z = e^t$,

$$M_X(t) = \sum_{x=1}^{\infty} \frac{z^x}{x(x+1)} = \frac{z}{1.2} + \frac{z^2}{2.3} + \frac{z^3}{3.4} + \ldots$$

$$= z\left(1 - \frac{1}{2}\right) + z^2\left(\frac{1}{2} - \frac{1}{3}\right) + z^3\left(\frac{1}{3} - \frac{1}{4}\right) + \cdots$$

$$= \left(z + \frac{z^2}{2} + \frac{z^3}{3} + \cdots\right) - \frac{z}{2} - \frac{z^2}{3} - \frac{z^4}{4} \ldots$$

$$= -\log(1 - z) - \frac{1}{z}\left(\frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \cdots\right)$$

$$= -\,log(1-z) + 1 + \frac{1}{z}\,log(1-z)\,, |z| < 1$$

$$= 1 + \left(\frac{1}{z} - 1\right) log(1-z)\,,\ \ |z| < 1$$

$$M_X(t) = \begin{cases} 1 + (e^{-\,t} - 1)\,log(1-e^t) & , \qquad\qquad t < 0 \\ 1 & , \qquad for \quad t = 0 \end{cases}$$

and $M_X(t)$ does not exist for $t > 0$.

3.2

Moment Generating Function

Certain derivations presented in modules 2.4, 2.5 and 2.6 have been somewhat heavy on algebra. For example, determining the mean and variance of the **Binomial distribution** turned out to be fairly tiresome. Another example of hard work was determining the set of probabilities associated with a sum , $P(X + Y = t)$. Many of these tasks are greatly simplified by using **probability generating functions.**

Moment Generating Function: The moment generating function (m.g.f) of a random variable X is dentoed by $M_X(t)$ and it is defined as

$$\boldsymbol{M_X(t) = E(e^{tX})}$$

$$\therefore\ M_X(t) = E(e^{tX}) = E\left[1 + \frac{tX}{1!} + \frac{(tX)^2}{2!} + \frac{(tX)^3}{3!} + \cdots\right]$$

$$= E(1) + \frac{t}{1!}E(X) + \frac{t^2}{2!}E(X^2) + \frac{t^3}{3!}E(X^3) + \cdots + \frac{t^r}{r!}E(X^r) + \cdots + \infty$$

$$\therefore\ M_X(t) = 1 + \frac{t}{1!}\mu_1' + \frac{t^2}{2!}\mu_2' + \frac{t^3}{3!}\mu_3' + \cdots + \frac{t^r}{r!}\mu_r' + \cdots + \infty \qquad \ldots\ldots \quad (1)$$

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu_r'$$

which gives the m.g.f in terms of moments.

Therefore the coeffceint of $\frac{t^r}{r!}$ in $M_X(t)$ is μ_r' , where $r = 1, 2, 3, \ldots$ and $\mu_r' = E(X^r)$, moment about origin.

The m.g.f of X about mean $\mu = \mu_1' = E(X)$ is defined as

$$M_{X-\mu}(t) = E\left[e^{t(X-\mu)}\right] = E\left[1 + \frac{t}{1!}(X-\mu) + \frac{t^2}{2!}(X-\mu)^2 + \frac{t^3}{3!}(X-\mu)^3 + \cdots\right]$$

$$= 1 + \frac{t}{1!}E(X-\mu) + \frac{t^2}{2!}E(X-\mu)^2 + \frac{t^3}{3!}E(X-\mu)^3 + \cdots$$

$$= 1 + \frac{t}{1!}\mu_1 + \frac{t^2}{2!}\mu_2 + \frac{t^3}{3!}\mu_3 + \cdots$$

where $E(x-\mu)^r = \mu_r$ is known as the r^{th} central moment for $r = 1, 2, \ldots$

Note that $\mu_1 = E(X-\mu) = E(X) - \mu = \mu - \mu = 0$

Since $M_X(t)$ generates moments, it is called **moment generating function**.

If X is a discrete random variable with p.m.f. $p(x)$ then

$$M_X(t) = E\left(e^{tX}\right) = \sum_x e^{tx} p(x)$$

If X is a continuous random variable with p.d.f. $f(x)$, then

$$M_X(t) = E\left(e^{tX}\right) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

Moments Using Moment Generating Function:

Differentiating equation (1) with respect to t and then putting $t = 0$, gives

$$\mu_1' = \left[\frac{d}{dt} M_X(t)\right]_{t=0}$$

$$\mu_2' = \left[\frac{d^2}{dt^2} M_X(t)\right]_{t=0}$$

In general,

$$\mu_r' = \left[\frac{d^r}{dt^r} M_X(t)\right]_{t=0}, \; r = 1, 2, 3, \ldots$$

Note: Moment generating function $M_X(t)$ is used to calculate the higher moments.

Theorems on Moment Generating Function:

Theorem 1: $\boldsymbol{M_{aX}(t) = M_X(at)}$, **where $\boldsymbol{a}$ is a constant.**

Proof: By definition $M_{aX}(t) = E(e^{taX}) = E(e^{atX}) = M_X(at)$

Therefore, $M_{aX}(t) = M_X(at)$

Theorem 2: The moment generating function of the sum of n independent random variables is equal to the product of their respective moment generating functions, $\boldsymbol{i.e.}$, $\boldsymbol{M_{X_1+X_2+X_3+\cdots+X_n}(t) = M_{X_1}(t)M_{X_2}(t)M_{X_3}(t) \dots M_{X_n}(t)}$

Proof: By definition,

$$M_{X_1+X_2+X_3+\cdots+X_n}(t) = \left[E^{t(X_1+X_2+X_3+\cdots+X_n)}\right]$$

$$= E(e^{tX_1})E(e^{tX_2})E(e^{tX_3}) \dots E(e^{tX_n})$$

(Since $X_1, X_2, \dots, X_n$ are independent).

Therefore, $M_{X_1+X_2+X_3+\cdots+X_n}(t) = M_{X_1}(t)M_{X_2}(r) \dots M_{X_n}(t)$
Hence the proof.

Uniqueness Theorem of Moment Generating Function:

The m.g.f. of a distribution, if exists, uniquely determines the distribution. This implies that corresponding to a given probability distribution, there is only one m.g.f (provided it exists) and corresponding to a given m.g.f, there is only one probability distribution. Hence $M_X(t) = M_Y(t) \Longrightarrow X$ and Y are identically distributed.

Effect of Change of Origin and Scale on Moment Generating Function:

Let a random variable X be transformed to a new variable U by changing both the origin and scale in X as $= \frac{X-a}{h}$, where a and h are constants.

The m.g.f of U (about origin) is given by

$$M_U(t) = E(e^{tU}) = E\left[e^{t\left(\frac{X-a}{h}\right)}\right] = E\left(e^{\left(\frac{tX}{h}-\frac{ta}{h}\right)}\right) = e^{-\frac{at}{h}}E\left(e^{\left(\frac{tX}{h}\right)}\right)$$

$$\therefore\ \boldsymbol{M_{\frac{X-a}{h}}(t) = e^{-\frac{at}{h}}M_X\left(\frac{t}{h}\right)}$$

Note: If $Y = aX + b$, then $M_Y(t) = e^{bt}M_X(at)$

Example 1: If X represents the outcome when a fair die is tossed, find the m.g.f. of X and hence, find $E(X)$ and $Var(X)$.

Solution: When a fair die is tossed

$$P(X = x) = \frac{1}{6}, \quad x = 1, 2, 3, 4, 5, 6$$

$$\therefore \quad M_X(t) = \sum_{X=1}^{6} e^{tx} P(X = x) = \frac{1}{6} \sum_{x=1}^{6} e^{tx}$$

$$= \frac{1}{6}(e^t + e^{2t} + e^{3t} + e^{4t} + e^{5t} + e^{6t})$$

$$E(X) = \left[\frac{d}{dt} M_X(t)\right]_{t=0} = \frac{1}{6}[e^t + 2e^{2t} + 3e^{3t} + 4e^{4t} + 5e^{5t} + 6e^{6t}]_{t=0}$$

$$= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = \frac{7}{2}$$

$\therefore$ Mean $= E(X) = \frac{7}{2}$

Now, $E[X^2]_{t=0} = \left\{\frac{d^2}{dt^2}[M_X(t)]\right\}_{t=0}$

$$= \frac{1}{6}[e^t + 4e^{2t} + 9e^{3t} + 16e^{4t} + 25e^{5t} + 36e^{6t}]_{t=0}$$

$$= \frac{1}{6}(1 + 4 + 9 + 16 + 25 + 36) = \frac{91}{6}$$

$$V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}.$$

Example 2: Find the m.g.f. of the random variable X whose probability function $P(X = x) = \frac{1}{2^x}, x = 1, 2, 3, \ldots$ and hence find its mean.

Solution: By definition,

$$M_X(t) = E\left(e^{tX}\right) = \sum_{x=0}^{\infty} e^{tx} P(X = x) = \sum_{x=1}^{\infty} e^{tX}\left(\frac{1}{2^x}\right) = \sum_{x=1}^{\infty}\left(\frac{e^t}{2}\right)^x$$

$$= \left[\frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + \left(\frac{e^t}{2}\right)^3 + \cdots\right]$$

$$= \frac{e^t}{2}\left[1 + \frac{e^t}{2} + \left(\frac{e^t}{2}\right)^2 + \cdots\right] = \frac{e^t}{2}\left(1 - \frac{e^t}{2}\right)^{-1}$$

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

Therefore, $M_X(t) = \frac{e^t}{2-e^t}$

$$\mu_1' = \left[\frac{d}{dt}M_X(t)\right]_{t=0} = \left[\frac{d}{dt}\left(\frac{e^t}{2 - e^t}\right)\right]_{t=0} = \left[\frac{(2 - e^t)e^t - e^t(-e^t)}{(2 - e^t)^2}\right]_{t=0} = \frac{(2-1)1 + 1}{(2-1)^2} = 2$$

Thus, $E(X)$ =mean= 2

Example 3: If the moments of a random variable X are defined by $E(X^r) = 0.6, r = 1, 2, \ldots$ .Show that $P(X = 0) = 0.4$, $P(X = 1) = 0.6$, and $P(X \geq 2) = 0$.

Solution: We know that

$$M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu_r'$$

where $\mu_r' = E(X^r) = 0.6$

$$\therefore\ M_X(t) = 1 + \sum_{r=1}^{\infty} \frac{t^r}{r!}\mu_r' = 1 + \sum_{r=1}^{\infty} \frac{t^r}{r!}(0.6) = 1 + (0.6)\left(\frac{t}{1!} + \frac{t^2}{2!} + \frac{t^3}{3!} + \cdots\right)$$

$$= 1 + (0.6)(e^t - 1) = 1 - 0.6 + 0.6e^t = 0.4 + 0.6e^t \qquad \ldots\ldots(1)$$

But by definition,

$$M_X(t) = E(e^{tx}) = \sum_{r=0}^{\infty} e^{tx}P(X = x)$$

$M_X(t) = P(X = 0) + e^tP(X = 1) + e^{2t}P(X = 2) + e^{3t}P(X = 3) + \ldots + \ldots$ …… (2)

From equations (1) and (2), we have

$$0.4 + 0.6e^t = P(X = 0) + e^t P(X = 1) + e^{2t} P(X = 2) + e^{3t} P(X = 3) + \cdots$$

Equating the coefficients of like terms on both sides,

$$P(X = 0) = 0.4, \; P(X = 1) = 0.6$$

$$P(X = 2) = P(X = 3) = \cdots = 0 \quad \Rightarrow P(X \geq 2) = 0$$

Example 4: Find the m.g.f. of a random variable whose moments are $\boldsymbol{\mu_r = (r + 1)!\; 2^r}$.

Solution: By definition, we have $M_X(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!} \mu_r'$

$$= \sum_{r=0}^{\infty} \frac{t^r}{r!} (r+1)!\,2^r = \sum_{r=0}^{\infty} (r+1)(2t)^r$$

$$= 1 + 2(2t) + 3(2t)^2 + \cdots = (1 - 2t)^{-2} = \frac{1}{(1-2t)^2}$$

$\therefore \quad M_X(t) = \frac{1}{(1-2t)^2}$

Example 5: If $\boldsymbol{X \sim B(n, p)}$, find the m.g.f of $\boldsymbol{X}$ and hence find its mean and variance.

Solution: Since $X \sim B(n, p)$, its p.m.f. is given by

$p(x) = \binom{n}{x} p^x q^{n-x}, \; x = 0, 1, 2, \ldots, n$ and $q = 1 - p$.

Then the m.g.f. of X is given by

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{n} e^{tx} p(x) = \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} \left(pe^t\right)^x q^{n-x}$$

$\Longrightarrow \; M_X(t) = (q + pe^t)^n$

$\frac{d}{dt} M_X(t) = n(q + pe^t)^{n-1} pe^t \Rightarrow$Mean$= \mu_1' = \left[\frac{dM_X(t)}{dt}\right]_{t=0} = np$

Next, $\frac{d^2}{dt^2}\left(M_X(t)\right) = np[(n-1)(q+pe^t)^{n-2}pe^{2t} + (q+pe^t)^{n-1}e^t]$

$\implies \mu_2' = \left[\frac{d^2}{dt^2}\left(M_X(t)\right)\right]_{t=0} = np[(n-1)p+1] = np[np-p+1] = np(np+q)$

$\implies \mu_2' = n^2p^2 + npq$

Now, variance $= \sigma^2 = \mu_2' - (\mu_1')^2 = n^2p^2 + npq - n^2p^2 = npq$

Thus, $\mu = np$ and $\sigma^2 = npq$

Note that $\sigma^2 = npq = \mu q$ where $(0 < q < 1)$ .Thus, $\mu > \sigma^2$.

Note: *For binomial distribution, mean is always greater than variance.*

Example 6 : If $X \sim P(\lambda)$, find its m.g.f. and hence find its mean and variance.

Solution: Since $X \sim P(\lambda)$, then its p.m.f. is given by

$$p(x) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{\lambda!}, x = 0, 1, \ldots \text{ and } \lambda > 0$$

The m.g.f. of X is given by

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tx} p(x) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^t\right)^x}{x!} = e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

$\implies M_X(t) = e^{\lambda(e^t-1)}$

Since $\frac{d}{dt}\left(M_X(t)\right) = e^{\lambda(e^t-1)}\lambda e^t$; Mean$= \mu = \mu' = \left[\frac{d}{dt}\left(M_X(t)\right)\right]_{t=0} = \lambda$.

Now, $\frac{d^2}{dt^2}\left(M_X(t)\right) = \lambda\left[e^{\lambda(e^t-1)}e^t + e^{\lambda(e^t-1)}\lambda e^{2t}\right]$

Then $\mu_2' = \left[\frac{d^2 M_X(t)}{dt^2}\right]_{t=0} = \lambda(1+\lambda) = \lambda + \lambda^2$

Thus, variance $= \sigma^2 = \mu_2' - (\mu_1')^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$

Therefore, $\mu = \sigma^2 = \lambda$

Note: *Mean and variance are same for Poisson distribution.*

Example 7: If $X \sim NB(r, p)$, find its m.g.f. and hence find its mean and variance.

Solution*:* Since $X \sim NB(r, p)$, its p.m.f. is given by

$$p(x) = \binom{-r}{x} p^r (-q)^x, \qquad x = 0, 1, 2, \ldots$$

The m.g.f of X is given by

$$M_X(t) = \sum_{x=0}^{\infty} e^{tx} \binom{-r}{x} p^r (-q)^x$$

$$= \sum_{x=0}^{\infty} \binom{-r}{x} p^r \left(-qe^t\right)^x = p^r \sum_{x=0}^{\infty} \binom{-r}{x} \left(-qe^t\right)^x$$

$\Longrightarrow \quad M_X(t) = p^r (1 - qe^t)^{-r}.$

Now, $\frac{d}{dt}\left(M_X(t)\right) = p^r(-r)(1 - qe^t)^{-(r+1)}(-qe^t) = qrp^r(1 - qe^t)^{-(r+1)}e^t$

Mean$= \mu = \mu_1' = \left[\frac{d}{dt}\left(M_X(t)\right)\right]_{t=0} = qrp^r(1 - q)^{-(r+1)} = qrp^r(p)^{-(r+1)} = \frac{rq}{p}$

Further, $\frac{d^2}{dt^2}\left(M_X(t)\right) = rqp^r \frac{d}{dt}\left\{(1 - qe^t)^{-(r+1)}e^t\right\}$

$$= rqp^r\left\{-(r+1)(1 - qe^t)^{-(r+2)}(-qe^{2t}) + (1 - qe^t)^{-(r+1)}e^t\right\}$$

Then $\mu_2' = \left[\frac{d^2}{dt^2}\left(M_X(t)\right)\right]_{t=0} = rqp^r\left[(r+1)qp^{-(r+2)} + p^{-(r+1)}\right]$

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

$$\Longrightarrow \quad \mu_2' = \frac{r^2q^2}{p^2} + \frac{rq}{p^2}$$

Hence, variance $= \sigma^2 = \mu_2' - (\mu_1')^2 = \frac{r^2q^2}{p^2} + \frac{rq}{p^2} - \frac{r^2q^2}{p^2} \Longrightarrow \quad \sigma_2^2 = \frac{rq}{p^2}$

Example 8:Let X be a random variable with p.d.f.

$$f(x) = \begin{cases} \frac{1}{3} e^{-\frac{x}{3}} & , \quad x > 0 \\ 0 & , \quad otherwise \end{cases}$$

Find

- **(i)** $P(X > 3)$
- **(ii)** **M.g.f. of X**
- **(iii)** $E(X)$ **and** $Var(X)$

Solution:

(i) $$P(X>3) = \int_3^\infty f(x)\,dx = \int_3^\infty \frac{1}{3} e^{-\frac{x}{3}} dx = \frac{1}{3}\left[\frac{e^{-\frac{x}{3}}}{-\frac{1}{3}}\right]_3^\infty = -(0 - e^{-1}) = e^{-1} = \frac{1}{e}$$

(ii) $$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} f(x) dx$$

$$= \int_0^\infty e^{tx} \frac{1}{3} e^{-\frac{x}{3}} dx = \frac{1}{3}\int_0^\infty e^{\left(t-\frac{1}{3}\right)x} dx = \frac{1}{3}\int_0^\infty e^{-\left(\frac{1}{3}-t\right)x} dx = \frac{1}{3}\left[\frac{e^{-\left(\frac{1}{3}-t\right)x}}{-\left(\frac{1}{3}-t\right)}\right]_0^\infty$$

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

$$M_X(t) = \frac{1}{1-3t} = (1-3t)^{-1}$$

$$\frac{d}{dt}[M_X(t)] = -(1-3t)^{-2}(-3) = 3(1-3t)^{-2}$$

(iii) $$E(X) = Mean = \left[\frac{d}{dt} M_X(t)\right]_{t=0} = 3$$

$$\frac{d^2}{dt^2}[M_X(t)] = -6(1-3t)^{-3}(-3) = 18(1-3t)^{-3}$$

$$E(X^2) = \left[\frac{d^2}{dt^2} M_X(t)\right]_{t=0} = 18$$

$$Var(X) = E(X^2) - [E(X)]^2 = 18 - 9 = 9$$

Example 9: Let X be a discrete random variable with p.d.f.

$$\boldsymbol{p(x) = \begin{cases} \dfrac{1}{x(x+1)} & , \quad x = 1, 2 \ldots \\ 0 & , \quad otherwise \end{cases}}$$

Show that $E(X)$ does not exist even though m.g.f. exist.

Solution:

$$E(X) = \sum_{x=1}^{\infty} x\,p(x) = \sum_{x=1}^{\infty} \frac{1}{x+1} = \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots = \sum_{x=1}^{\infty} \frac{1}{x} - 1$$

But $\boldsymbol{\sum_{x=1}^{\infty} \frac{1}{x}}$ is a divergent series.

Therefore, $E(x)$ does not exist and hence, no moment exists.

Now, m.g.f. of X is given by

$$M_X(t) = \sum_{x=1}^{\infty} p(x) e^{tx} = \sum_{x=1}^{\infty} \frac{e^{tx}}{x(x+1)}$$

Substituting $z = e^t$,

$$M_X(t) = \sum_{x=1}^{\infty} \frac{z^x}{x(x+1)} = \frac{z}{1.2} + \frac{z^2}{2.3} + \frac{z^3}{3.4} + \ldots$$

$$= z\left(1 - \frac{1}{2}\right) + z^2\left(\frac{1}{2} - \frac{1}{3}\right) + z^3\left(\frac{1}{3} - \frac{1}{4}\right) + \cdots$$

$$= \left(z + \frac{z^2}{2} + \frac{z^3}{3} + \cdots\right) - \frac{z}{2} - \frac{z^2}{3} - \frac{z^4}{4} \ldots$$

$$= -\log(1 - z) - \frac{1}{z}\left(\frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \cdots\right)$$

$$= -\,log(1-z) + 1 + \frac{1}{z} log(1-z)\,, |z| < 1$$

$$= 1 + \left(\frac{1}{z} - 1\right) log(1-z)\,,\ |z| < 1$$

$$M_X(t) = \begin{cases} 1 + (e^{-t} - 1)\, log(1 - e^{t}) & , \qquad t < 0 \\ 1 & , \quad for \quad t = 0 \end{cases}$$

and $M_X(t)$ does not exist for $t > 0$.

P1:

Find the m.g.f. of uniform distribution $U[a, b]$ and hence obtain the mean and variance of the distribution.

P2:

Find the m.g.f. of Normal $N(\mu, \sigma^2)$ distribution and hence find its mean and variance.

Solution:

Since $X \sim N(\mu, \sigma^2)$, its p.d.f. is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

The m.g.f. of X is given by

$$M_X(t) = E\left[e^{tX}\right] = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tx}\, exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$$

Let $z = \frac{x-\mu}{\sigma}$. Thus $x = \mu + \sigma z$ and $dx = \sigma . dz$

Thus, $M_X(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} exp\left(t(\mu+\sigma z)\right) exp\left(-\frac{z^2}{2}\right)dz$

$$= \frac{1}{\sqrt{2\pi}} e^{t\mu}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}\left(z^2 - 2t\sigma z\right)\right]dz$$

$$= \frac{1}{\sqrt{2\pi}} e^{t\mu}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}\left\{(z-\sigma t)^2 - \sigma^2 t^2\right\}\right]dz$$

$$= \frac{1}{\sqrt{2\pi}} e^{t\mu + \frac{1}{2}\sigma^2 t^2}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(z-\sigma t)^2\right]dz$$

Let $u = z - \sigma t \implies du = dz$

$$= e^{t\mu + \frac{1}{2}\sigma^2 t^2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}u^2} du$$

$$= e^{t\mu + \frac{1}{2}\sigma^2 t^2}$$

$\Longrightarrow\ M_X(t) = e^{t\mu+\frac{1}{2}\sigma^2 t^2} \Longrightarrow\ M_X'(t) = e^{t\mu+\frac{1}{2}\sigma^2 t^2}(\mu + \sigma^2 t)$

$\Longrightarrow\ \mu$ =Mean= $\mu_1' = M_X'(t)|_{t=0} = \mu$

Further, $M_X''(t) = e^{t\mu+\frac{1}{2}\sigma^2 t^2}(\mu + \sigma^2 t)^2 + e^{t\mu+\frac{1}{2}\sigma^2 t^2}(\sigma^2)$

$\Longrightarrow\ \mu_2' = M_X''(t)|_{t=0} = \mu^2 + \sigma^2 \Longrightarrow\ \mu_2' = \mu^2 + \sigma^2$

The variance is given by

Variance $= \mu_2' - (\mu_1')^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$

P3:

Find the m.g.f. of geometric $G(p)$ distribution and hence obtain its mean and variance.

Solution:

If $X \sim G(p)$, its p.m.f is given by $p(x) = q^x p$ for $x = 0, 1, 2, \dots, \quad 0 < p < 1$.

Then the m.g.f. is given by

$$M_X(t) = E\left[e^{tX}\right] = \sum_{x=0}^{\infty} e^{tx} q^x p = \sum_{x=0}^{\infty} \left(qe^t\right)^x = \frac{p}{1 - qe^t}$$

$$\Longrightarrow \; M_X(t) = \frac{p}{1 - qe^t}$$

Then $\mu_1' = M_X'(t)|_{t=0} = pq(1-q)^{-2} = \frac{q}{p}$ and $\mu_2' = M_X''(t)|_{t=0} = \frac{q}{p} + \frac{2q^2}{p^2}$

and $\mu_2 = \mu_2' - (\mu_1')^2 = \frac{q}{p} + \frac{2q^2}{p^2} - \left(\frac{q}{p}\right)^2 = \frac{q}{p} + \frac{q^2}{p^2} = \frac{qp+q^2}{p^2} = \frac{q(p+q)}{p^2} = \frac{q}{p^2}$

Thus, Mean$= \mu = \mu_1' = \frac{q}{p}$ and variance $= \sigma^2 = \frac{q}{p^2}$

P4:

Find the m.g.f. of exponential $E(\lambda)$ distribution and hence find its mean and variance.

Solution:

Since $X \sim E(\lambda)$, its p.d.f. is given by $f(x) = \lambda e^{-\lambda x}$ for $x > 0$ and $\lambda > 0$

The m.g.f. of X is given by

$$M_X(t) = E\left[e^{tX}\right] = \lambda \int_0^\infty e^{tx} e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda - t)x} dx$$

$$= \frac{\lambda}{\lambda - t} = \left(1 - \frac{t}{\lambda}\right)^{-1} = \sum_{r=0}^{\infty} \left(\frac{t}{\lambda}\right)^r \ for\ \lambda > t$$

$$\Rightarrow M_X(t) = \sum_{r=0}^{\infty} \left(\frac{t}{\lambda}\right)^r \ for\ \lambda > t$$

But $\mu_r' =$ coefficient of $\frac{t^r}{r!}$ in $M_X(t) = \frac{r!}{\lambda^r}$ for $r = 1, 2, \ldots$

Thus, $\mu_1' = \frac{1}{\lambda}$ and $\mu_2' = \frac{2}{\lambda^2}$ and hence $\mu_2 = \mu_2' - (\mu_1')^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$

Thus, $\mu = \frac{1}{\lambda}$ and $\sigma^2 = \frac{1}{\lambda^2}$

3.2. Moment Generating Function

Exercise:

1. Find the m.g.f of a.r.v. whose moments are given by $\mu'_r = (r+1)!\,2^r$

2. If $M_X(t) = \frac{3}{3-t}$, find standard deviation of X

3. Find the m.g.f. of a.r.v X whose p.d.f is given by

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

4. Find the m.g.f and hence find the mean and variance of a.r.v. X whose p.d.f. is given by

   i. $$f(x) = \begin{cases} 2e^{-2x} \ , & x \ge 0 \\ 0 \ , & otherwise \end{cases}$$

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

   iii. $$f(x) \begin{cases} \frac{1}{3}e^{-\frac{x}{3}} \ , & x \ge 0 \\ 0 \ , & otherwise \end{cases}$$

   iv. $$f(x) \begin{cases} \frac{1}{k} \ , & 0 < x < k \\ 0 \ , & otherwise \end{cases}$$

   v. $$f(x) = \lambda e^{-\lambda(x-a)} \ , x \ge a$$

Answers:

1. $M_X(t) = \frac{1}{(1-2t)^r}$

2. $\frac{1}{3}$

3. $\frac{1}{2t^2}(1 + 2t\, e^{2t} - e^{2t})$

4.

   i. $M_X(t) = \frac{2}{2-t}\ ,\ \mu = \frac{1}{2}\ ,\ \sigma^2 = \frac{1}{4}$

   ii. $M_X(t) = \begin{cases} \frac{e^{2t}-e^{-t}}{3t} & , \quad t \neq 0 \\ 1 & , \quad t = 0 \end{cases}$

   iii. $M_X(t) = (1-3t)^{-1},\ \mu = 3\ ,\ \sigma^2 = 9$

   iv. $M_X(t) = \frac{e^{tk}-1}{kt}\ ,\mu = \frac{k}{2}\ ,\ \sigma^2 = \frac{k^2}{\sqrt{2}}$

   v. $M_X(t) = \frac{\lambda e^{at}}{\lambda - t}\ ,\mu = \frac{9\lambda+1}{\lambda},\sigma^2 = \frac{1}{\lambda^2}$

3.3

Characteristic Function

In some cases m.g.f. does not exist. For example, consider the p.m.f. given by

$$p(x) = \begin{cases} \frac{6}{\pi^2 x^2}, & x = 1, 2, 3, \dots \\ 0, & otherwise \end{cases}$$

Its m.g.f. is given by

$$M_X(t) = \sum_{x=1}^{\infty} e^{tx} p(x) = \frac{6}{\pi^2} \sum_{x=1}^{\infty} \frac{e^{tx}}{X^2},$$

which is divergent. Thus, $M_X(t)$ does not exist. A more serviceable function that the m.g.f. is the **characteristic function.**

Characteristic function: The characteristic function of a.r.v. X is defined by

$$\phi_X(t) = E\left[e^{itX}\right] = \begin{cases} \int e^{itx} f(x) dx \;\; if\; X \;\; is\; a\; c.r.v.with\, p.d.f\; f(x) \\ \sum_x e^{itx}\; p(x) \;\; if\; X \;\; is\; a\; d.r.v.\; with\;\; p.m.f\; p(x) \end{cases}$$

where $i = \sqrt{-1}$, the imaginary number.

Note:

1. $|\emptyset_X(t)| = \left|E\left(e^{itX}\right)\right| \leq E\left(\left|e^{itX}\right|\right) = E\left(\sqrt{cos^2 tX + sin^2 tX}\right) = E(1) = 1$

   Since $|\emptyset_X(t)| \leq 1$, $\emptyset_X(t)$ always exists for any **probability distribution**.

2. $\emptyset_X(t) = E\left[e^{itX}\right] = E\left[1 + (it)X + \frac{(it)^2}{2!}X^2 + \frac{(it)^3}{3!}X^3 + \cdots\right]$

$$= 1 + (it)E(X) + \frac{(it)^2}{2!}E(X^2) + \frac{(it)^3}{3!}E(X^3) + \cdots$$

$$= 1 + (it)\mu_1' + \frac{(it)^2}{2!}\mu_2' + \frac{(it)^3}{3!}\mu_3' + \cdots$$

where $\mu'_r = E(X^r) = r^{th}$ moment abut origin for $r = 1, 2, \ldots$

3. If $\emptyset_X(t)$ is given, then the r^{th} moment about origin is given b

$$\boldsymbol{\mu'_r = coefficient\ of\ \frac{(it)^r}{r!}\ \text{in}\ \emptyset_X(t).}$$

Properties:

1. $\boldsymbol{\emptyset_X(0) = 1}$

Proof: $\emptyset_X(t) = E\left[e^{itX}\right] = E(1)$ when $t = 0$

$$= 1$$

Thus, $\emptyset_X(0) = 1$

2. $\boldsymbol{|\emptyset_X(t)| \leq 1}$ **for all real** $\boldsymbol{t}$**.**

Proof: $|\emptyset_X(t)| = \left|E\left(e^{itX}\right)\right| \leq E\left(\left|e^{itX}\right|\right) = E\left(\sqrt{cos^2tX + sin^2tX}\right) = E(1) = 1$

$\Longrightarrow |\emptyset_X(t)| \leq 1$ for all real t

3. $\boldsymbol{\emptyset_X(t)}$ **continuous function of** $\boldsymbol{t}$ **in** $(-\infty, \infty)$.

Proof: For $h \neq 0$,

$$|\emptyset_X(t+h) - \emptyset_X(t)| = \left|E\left(e^{i(t+h)X}\right) - E\left(e^{itX}\right)\right| = \left|E\left(e^{i(t+h)X} - e^{itX}\right)\right|$$

$$= \left|E\left\{e^{itX}\left(e^{ihX} - 1\right)\right\}\right| \leq E\left(\left|e^{itX}\right|\left|e^{ihX} - 1\right|\right) = E\left(\left|e^{ihX} - 1\right|\right) \longrightarrow 0 \text{ as } h \longrightarrow 0$$

Thus $\lim\limits_{h \to 0} |\emptyset_X(t+h) - \emptyset_X(t)| = 0$

$\Longrightarrow \lim\limits_{h \to 0} \emptyset_X(t+h) = \emptyset_X(t)$

$\Longrightarrow \emptyset_X(t)$ is a continuous function of t in $(-\infty, \infty)$.

4. $\boldsymbol{\emptyset_X(-t) = \overline{\emptyset_X(t)}}$ **,** $\boldsymbol{i.e}$**,** $\boldsymbol{\emptyset_X(-t)}$ **is the complex conjugate of** $\boldsymbol{\emptyset_X(t)}$**.**

Proof: Here $\overline{\emptyset_X(t)} = \overline{E[e^{itX}]} = E[\cos tX - i \sin tX]$

$\Longrightarrow \emptyset_X(-t) = E[\cos(-tX) + i\sin(-tX)] = E[\cos tX - i \sin tX] = \overline{\emptyset_X(t)}$

Thus, $\emptyset_X(-t) = \overline{\emptyset_X(t)}$

5. If the p.d.f. is even $\boldsymbol{i.e.}$**,** $\boldsymbol{f(-x) = f(x)}$**, then the characteristic function is real valued and even function of** $\boldsymbol{t}$**.**

Proof: We know that , $\phi_X(t) = E\left(e^{itX}\right) = \int_{-\infty}^{\infty} e^{itX} f(x)dx$

Let $x = -y \implies dx = -dy$. Then

$$\phi_X(t) = \int_{\infty}^{-\infty} e^{-ity} f(-y)(-dy) = \int_{-\infty}^{\infty} e^{-ity} f(y)dy \qquad (\because\ f \text{ is an even function})$$

$$= E\left[e^{-itX}\right] = \emptyset_X(-t)$$

Thus, $\emptyset_X(t) = \emptyset_X(-t)$

$\Longrightarrow \emptyset_X(t)$ is an even function of t.

Further, $\overline{\emptyset_X(t)} = \emptyset_X(-t)$ (by property 4)

$= \emptyset_X(t)$ (Since $\emptyset_X(t)$ is even function)

Thus, $\emptyset_X(t)$ is real.

6. **If X is a r.v. with characteristic function $\emptyset_X(t)$ and $\mu_r' = E(X^r)$ exists, then**

$$\boldsymbol{\mu_r' = (-i)^r \frac{d^r}{dt^r}\left(\emptyset_X(t)\right)\Big|_{t=0}}$$

Proof:

$$\frac{d^r}{dt^r}\left(\emptyset_X(t)\right) = \frac{d^r}{dt^r}\left(E\left(e^{itX}\right)\right) = i^r E\left[X^r e^{itX}\right] = i^r E(X^r)$$

Now, $\frac{d^r}{dt^r}\left(\emptyset_X(t)\right)\Big|_{t=0} = i^r E(X^r)$ and $\mu_r' = E(X^r) = \frac{1}{i^r}\frac{d^r}{dt^r}\left(\emptyset_X(t)\right)\Big|_{t=0}$.

Thus, $\mu_r' = (-i)^r \frac{d^r}{dt^r}\left(\emptyset_X(t)\right)\Big|_{t=0}$

7. **Effect of change of origin and scale .**

Let $U = \frac{X-a}{h}$ where a and h are constants.

Then $\emptyset_U(t) = E\left[e^{itU}\right] = E\left[e^{it\left(\frac{X-a}{h}\right)}\right] = e^{-\frac{ita}{h}} E\left[e^{i\left(\frac{t}{h}\right)X}\right]$

$$\Longrightarrow \emptyset_U(t) = e^{-\frac{ita}{h}} \emptyset_X\left(\frac{t}{h}\right)$$

8. **If $X_1, X_2, \dots, X_n$ are independent, then**

$$\boldsymbol{\emptyset_{X_1+\cdots+X_n}(t) = \emptyset_{X_1}(t)\,\emptyset_{X_2}(t) \dots \emptyset_{X_n}(t)}$$

Proof:

$$\emptyset_{X_1+\cdots+X_n}(t) = E\left[e^{it(X_1+X_2+,\dots+X_n)}\right] = E\left[e^{itX_1}.e^{itX_2}\dots e^{itX_n}\right]$$

$$(\because\ X_1, X_2, \dots, X_n \textit{ are independent})$$

$$= E\left[e^{itX_1}\right].E\left[e^{itX_2}\right]\dots E\left[e^{itX_n}\right]$$

$$= \emptyset_{X_1}(t).\emptyset_{X_2}(t)\dots\emptyset_{X_n}(t)$$

$$\Longrightarrow \emptyset_{X_1+\cdots+X_n}(t) = \emptyset_{X_1}(t).\emptyset_{X_2}(t)\dots\emptyset_{X_n}(t)$$

Note: Converse need not be true.

Uniqueness Theorem for Characteristic Functions:

The characteristic function uniquely determines the distribution. That is,
A necessary and sufficient condition for two distributions with p.d.fs $\boldsymbol{f_1(.)}$ and $\boldsymbol{f_2(.)}$ to be identical is that their characteristic function $\emptyset_1(\boldsymbol{t})$ and $\emptyset_2(\boldsymbol{t})$ are identical.

Example 1: If $\boldsymbol{X \sim B(n,p)}$, find its characteristic function and hence obtain its mean and variance.

Solution: Since $X \sim B(n,p)$, its p.m.f. is given by

$$p(x) = \binom{n}{x} p^x q^{n-x} \text{ for } x = 0, 1, 2, \dots, n$$

The characteristic function of X is given by

$$\phi_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{n} e^{itx} p(x) = \sum_{x=0}^{n} e^{itx} \binom{n}{x} p^x q^{n-x} = \sum_{x=0}^{n} \binom{n}{x} \left(pe^{it}\right)^x q^{n-x}$$

$\Rightarrow \ \emptyset_X(t) = \left(q + pe^{it}\right)^n$ and $\frac{d}{dt}\left(\emptyset_X(t)\right) = npi\left(q + pe^{it}\right)^{n-1} e^{it}$

The mean of X is given by

$$\mu = E(X) = \mu' = (-i)\frac{d}{dt}\left(\emptyset_X(t)\right)\Big|_{t=0} = (-i)\left[npi\left(q + pe^{it}\right)^{n-1} e^{it}\right]_{t=0}$$

$$= (-i)\, npi = np$$

Now, $\frac{d^2}{dt^2}\left(\emptyset_X(t)\right) = (npi)\frac{d}{dt}\left[\left(q + pe^{it}\right)^{n-1} e^{it}\right]$

$$= (npi)\left[(n-1)\left(q + pe^{it}\right)^{n-2} pie^{2it} + \left(q + pe^{it}\right)^{n-1} ie^{it}\right]$$

Thus, $\mu_2' = (-i)^2 \frac{d^2}{dt^2}\left(\emptyset_X(t)\right)\Big|_{t=0} = (-i)^2 (npi)[(n-1)pi + i]$

$$= (np)[np - p + 1] = np(np + q) = n^2p^2 + npq$$

$\Rightarrow \ \mu_2' = n^2p^2 + npq$

Therefore, the variance of X is given by

$$\sigma^2 = V(X) = \mu_2' - (\mu_1')^2 = n^2p^2 + npq - n^2p^2$$
$$\Longrightarrow\ \sigma^2 = npq.$$

Example 2: If $X \sim P(\lambda)$, find the characteristic function X and hence obtain its mean and variance.

Solution: Since $X \sim P(\lambda)$, its p.m.f. is given by

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}\ , x = 0, 1, ,2\ \ldots$$

The characteristic function of X is given by

$$\phi_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} p(x) = \sum_{x=0}^{\infty} e^{itx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^{it}\right)^x}{x!} = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda\left(e^{it}-1\right)}$$

$$\Rightarrow \boldsymbol{j_X(t) = e^{\lambda\left(e^{it}-1\right)}}$$

Now, $\frac{d}{dt}\big(\emptyset_X(t)\big) = e^{\lambda(e^{it}-1)}\lambda i e^{it}$

Thus, the mean is given by

$$\mu = \mu_1' = E(X) = (-i)\frac{d}{dt}(\phi_X(t))\Big|_{t=0} = (-i)(\lambda i) = \lambda$$

$$\Longrightarrow\ \mu = \lambda$$

Now, $\frac{d^2}{dt^2}\big(\emptyset_X(t)\big) = (\lambda i)\frac{d}{dt}\left[e^{\lambda(e^{it}-1)}e^{it}\right]$

$$= (\lambda i)\left[e^{\lambda(e^{it}-1)}\lambda i e^{2it} + e^{\lambda(e^{it}-1)} i e^{it}\right]$$

Thus, μ_2' is given by

$$\mu_2' = (-i)^2 \frac{d^2}{dt^2}\left(\phi_X(t)\right)\Big|_{t=0} = (-i)^2(\lambda i)(\lambda i + i) = (-i)^2(i^2)\lambda(\lambda+1)$$

$$= \lambda(\lambda+1) = \lambda^2 + \lambda$$

Hence, the variance is given by $\sigma^2 = \mu_2' - (\mu_1')^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \Rightarrow \ \sigma^2 = \lambda$

Example 3: If $X \sim N(\mu, \sigma^2)$, find the characteristic function of X and hence obtain its mean and variance.

Solution: Since $X \sim N(\mu, \sigma^2)$, its p.d.f. is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] , -\infty < x, \mu < \infty, \sigma > 0$$

The characteristic function of X is given by

$$\phi_X(t) = E\left[e^{itX}\right] = \int_{-\infty}^{\infty} e^{itx} f(x)\,dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]dx$$

Let $\frac{x-\mu}{\sigma} = z \Rightarrow x = \mu + \sigma z \Rightarrow dx = \sigma\, dz$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]dx = \int_{-\infty}^{\infty} e^{it(\mu+\sigma z)} e^{-\frac{1}{2}z^2}\, dz$$

$$= \frac{e^{it\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} exp\left[-\frac{1}{2}\left(z^2 - 2i\sigma zt\right)\right]dz$$

$$= \frac{e^{it\mu}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} exp\left[-\frac{1}{2}\left(z^2 - 2i\sigma zt + i^2\sigma^2 t^2 - i^2\sigma^2 t^2\right)\right]dz$$

$$= \frac{e^{it\mu - \frac{\sigma^2 t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} exp\left[-\frac{1}{2}(z - i\sigma t)^2\right]dz$$

Let $z - i\sigma t = u \Longrightarrow dz = du$

$$= e^{it\mu - \frac{\sigma^2 t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du$$

$$= e^{it\mu - \frac{\sigma^2 t^2}{2}} \qquad \left(\because \quad \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{u^2}{2}} du = 1 \right)$$

$\Longrightarrow \ \emptyset_X(t) = e^{itu - \frac{\sigma^2 t^2}{2}}$

Now, $\frac{d}{dt}\big(\emptyset_X(t)\big) = e^{it\mu - \frac{\sigma^2 t^2}{2}}(i\mu - \sigma^2 t)$

Then $\mu_1' = (-i)\frac{d}{dt}\big(\emptyset_X(t)\big)\Big|_{t=0} = (-i)\ (i\mu) = \mu$

Thus, Mean $= E(X) = \mu$.

Now, $\frac{d^2}{dt^2}\big(\emptyset_X(t)\big) = e^{it\mu - \frac{1}{2}\sigma^2 t^2}(i\mu - \sigma^2 t)^2 + e^{it\mu - \frac{1}{2}t^2\sigma^2}(-\sigma^2)$

Thus, $\mu_2' = (-i)^2\ \frac{d^2}{dt^2}\big(\emptyset_X(t)\big)\Big|_{t=0} = (-i)^2\ [i^2\mu^2 - \sigma^2] = (-1)(-\mu^2 - \sigma^2)$

$\Longrightarrow \ \mu_2' = \mu^2 + \sigma^2$

Hence, the variance is given by

$V(X) = \mu_2' - (\mu_1')^2 = \mu^2 + \sigma^2 - \mu^2 = \sigma^2$

$\Longrightarrow$ Variance $= V(X) = \sigma^2$

Finding p.m.f. (or p.d.f.) when characteristic function is known.

If X is a d.r.v. with characteristic function $\emptyset_X(t)$, then $\emptyset_X(t) = \sum P(X = j)e^{itj}$.

First write the characteristic function in this form and then identify the $P(X = j)$ which is the p.m.f. of the d.r.v. X.

Example 4: Find the p.m.f. of the d.r.v. X whose characteristic function is given by $\boldsymbol{\emptyset_X(t) = \left(q + pe^{it}\right)^n}$.

Solution: We have, $\emptyset_X(t) = \left(q + pe^{it}\right)^n$ and

$$\emptyset_X(t) = \left(q + pe^{it}\right)^n = \sum_{j=0}^{n} \binom{n}{j} \left(pe^{it}\right)^j q^{n-j}$$

$$= \sum_{j=0}^{n} \binom{n}{j} p^j q^{n-j} e^{itj} = \sum_{j=0}^{n} P(X = j) e^{itj}$$

$$= E\left[e^{itX}\right] \text{ where } P(X = j) = \binom{n}{j} p^j q^{n-j}$$

Thus p.m.f. is $p(j) = \binom{n}{j} p^j q^{n-j}$ for $j = 0,1,2, \dots, n$.

Example 5: Find the p.m.f. of a d.r.v. X whose characteristic function is given by

$$\boldsymbol{\emptyset_X(t) = e^{\lambda(e^{it}-1)}}.$$

Solution: We have, $\emptyset_X(t) = e^{\lambda(e^{it}-1)}$

$$\emptyset_X(t) = e^{\lambda(e^{it}-1)} = e^{-\lambda} e^{\lambda e^{it}} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^{it}\right)^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} e^{itx}$$

$$= \sum_{x=0}^{\infty} P(X = x) e^{itx} = E\left[e^{itX}\right]$$

where $P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}$ for $x = 0,1,2, \dots$, which is Poisson distribution with parameter λ.

Theorem 1: If X is a continuous random variable with characteristic function $\emptyset_X(t)$, then its p.d.f. is given by

$$\boldsymbol{f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \emptyset_X(t) dt}$$

Example 6: Find the p.d.f corresponding to the characteristic function

$$\boldsymbol{\emptyset_X(t) = e^{it\mu - \frac{1}{2}t^2\sigma^2}.}$$

Solution:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\,itx} e^{it\mu - \frac{1}{2}t^2\sigma^2} dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{1}{2}[t^2\sigma^2 - 2it(x-\mu)]} dt$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}\left\{t\sigma - i\left(\frac{x-\mu}{\sigma}\right)\right\}^2 + \left(\frac{x-\mu}{\sigma}\right)^2\right] dt$$

$$= \frac{1}{2\pi} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\int_{-\infty}^{\infty} exp\left[-\frac{1}{2}\left\{t\sigma - i\left(\frac{x-\mu}{\sigma}\right)\right\}^2\right] dt$$

Let $t\sigma - i\left(\frac{x-\mu}{\sigma}\right) = u \Rightarrow dt = \frac{du}{\sigma}$.

$$= \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} exp\left(-\frac{u^2}{2}\right) dt$$

$$= \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

Therefore, $f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], -\infty < x < \infty$

$\Longrightarrow X \sim N(\mu, \sigma^2)$

Example 7: Find the p.d.f. corresponding to the characteristic function defined by

$$\emptyset(\boldsymbol{t}) = \begin{cases} \boldsymbol{1 - |t|} & , \quad \boldsymbol{|t| \le 1} \\ \boldsymbol{0} & , \quad \boldsymbol{|t|} > 1 \end{cases}$$

Solution: The p.d.f. of $f(x)$ is given by

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\emptyset(t)dt = \frac{1}{2\pi}\int_{-1}^{1} e^{-itx}\emptyset(t)dt$$

$$= \frac{1}{2\pi}\int_{-1}^{0} e^{-itx}(1+t)dt + \frac{1}{2\pi}\int_{0}^{1} e^{-itx}(1-t)dt$$

($\because$ for $-1 < t < 0, |t| = -t$ and for $0 < t < 1, |t| = t$)

Now,

$$\int_{-1}^{0} e^{-itx}(1+t)dt = \left[\frac{e^{-itx}}{-ix}(1+t)\right]_{-1}^{0} + \frac{1}{ix}\int_{-1}^{0} e^{-itx}dt$$

$$= -\frac{1}{ix} + \frac{1}{ix}\left[\frac{e^{-itx}}{-ix}\right]_{-1}^{0}$$

$$= -\frac{1}{ix} + \frac{1}{(ix)^2}\left(e^{ix} - 1\right)$$

Similarly,

$$\int_{0}^{1} e^{-itx}(1-t)dt = \frac{1}{ix} + \frac{1}{(ix)^2}\left(e^{-ix} - 1\right)$$

$$\therefore f(x) = \frac{1}{2\pi}\left[\frac{1}{(ix)^2}\left\{e^{ix} - 1 + e^{-ix} - 1\right\}\right] = \frac{1}{\pi x^2}\left(1 - \frac{e^{ix} + e^{-ix}}{2}\right)$$

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

P1:

If $X \sim NB(r, p)$, find its characteristic function and hence obtain its mean and variance.

Solution:

Since $X \sim NB(r, p)$, its p.m.f. is given by

$$p(x) = \binom{-r}{x} p^r (-q)^x \, , x = 0,1,2, \ldots$$

The characteristic function of X is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} p(x) \ = \sum_{x=0}^{\infty} e^{itx} \binom{-r}{x} p^r (-q)^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} \left(-qe^{it}\right)^x = p^r \left(1 - qe^{it}\right)^{-r}$$

Thus, $\emptyset_X(t) = p^r \left(1 - qe^{it}\right)^{-r}$

Now,

$$\frac{d}{dt}\left(\emptyset_X(t)\right) = p^r(-r)\left(1 - qe^{it}\right)^{-(r+1)}(-q)ie^{it} = ip^r qre^{it}\left(1 - qe^{it}\right)^{-(r+1)}$$

Thus, mean $= \mu = \mu_1{}' = (-i)\frac{d}{dt}\left(\emptyset_X(t)\right)\Big|_{t=0}$

$$= (-i)ip^r qr(1 - q)^{-(r+1)} = p^r qr(p)^{-(r+1)} = \frac{qr}{p}$$

$$\Rightarrow \mu = \frac{rq}{p}$$

Now, $\frac{d^2}{dt^2}\left(\emptyset_X(t)\right) = p^r rqi \frac{d}{dt}\left[\left(1 - qe^{it}\right)^{-(r+1)} e^{it}\right]$

$$= p^r rqi\left[-(r+1)\left(1 - qe^{it}\right)^{-(r+2)}(-qi)e^{2it} + \left(1 - qe^{it}\right)^{-(r+1)} ie^{it}\right]$$

Thus, $\mu_2{}' = (-i)^2 \frac{d^2}{dt^2}\left(\emptyset_X(t)\right)\Big|_{t=0}$

$$= (-i)^2 p^r rqi\left[(r+1)qip^{-(r+2)} + p^{-(r+1)}i\right]$$

$$= (-i)^2 p^r rqi^2 p^{-(r+2)}[(r+1)q + p]$$

$$= \frac{rq}{p^2}(rq + q + p) = \frac{rq}{p^2}(rq + 1)$$

$$\Longrightarrow \mu_2{}' = \frac{r^2q^2}{p^2} + \frac{rq}{p^2}$$

Thus, the variance is given by

$$\sigma^2 = \mu_2{}' - (\mu_1{}')^2 = \frac{r^2q^2}{p^2} + \frac{rq}{p^2} - \frac{r^2q^2}{p^2}$$

$$\Longrightarrow \sigma^2 = \frac{rq}{p^2}$$

P2:

Find the characteristic function of uniform $\boldsymbol{U[a,b]}$ distribution and hence obtain the mean and variance of the distribution.

Solution:

Since $X \sim U(a,b)$, its p.d.f, is given by

$$f(x) = \frac{1}{b-a}\ , a < x < b$$

The characteristic function is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \int_a^b e^{itx} f(x)dx = \int_a^b e^{itx}\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b e^{itx}dx$$

$$= \frac{1}{b-a}\left[\frac{e^{itx}}{it}\right]_a^b = \frac{1}{(b-a)it}\left(e^{itb} - e^{ita}\right)$$

$$\Longrightarrow \emptyset_X(t) = \frac{e^{itb} - e^{ita}}{it(b-a)}$$

$$\Longrightarrow t\emptyset_X(t) = \tfrac{1}{i(b-a)}\left(e^{itb} - e^{ita}\right) \qquad \dots \dots .. \qquad (1)$$

Differentiating both sides of (1) w.r.t. t, we get

$$t\emptyset'_X(t) + \emptyset_X(t) = \tfrac{1}{i(b-a)}\left[(ib)e^{itb} - (ia)e^{ita}\right] \quad \dots \dots \qquad (2)$$

Again differentiating both sides of (2) w.r.t. t, we get

$$t\emptyset''_X(t) + \emptyset'_X(t) + \emptyset'_X(t) = \tfrac{1}{i(b-a)}\left[(ib)^2 e^{itb} - (ia)^2 e^{ita}\right]$$

$$\Longrightarrow t\emptyset''_X(t) + 2\emptyset'_X(t) = \frac{1}{i(b-a)}\left[(ib)^2 e^{itb} - (ia)^2 e^{ita}\right]$$

In general,

$$t\frac{d^{k+1}}{dt^{k+1}}\left(\emptyset_X(t)\right) + (k+1)\frac{d^k}{dt^k}\left(\emptyset_X(t)\right) = \frac{1}{i(b-a)}\left[(ib)^{k+1}e^{itb} - (ia)^{k+1}e^{ita}\right]$$

If $t = 0$, then

$$(k+1)\frac{d^k}{dt^k}\left(\emptyset_X(t)\right)\Big|_{t=0} = \frac{1}{i(b-a)}[(ib)^{k+1} - (ia)^{k+1}]$$

$$\Rightarrow \frac{d^k}{dt^k}\left(\emptyset_X(t)\right)\Big|_{t=0} = \frac{1}{k+1}\frac{1}{i(b-a)}[(ib)^{k+1} - (ia)^{k+1}] \;\ldots\ldots\ldots\ldots\ldots\; (3)$$

From (3), the mean is given by

$$\mu = \text{mean} = {\mu_1}' = (-i)\frac{d}{dt}\left(\emptyset_X(t)\right)\Big|_{t=0}$$

$$= (-i)\frac{1}{2i(b-a)}[(ib)^2 - (ia)^2] = \frac{1}{2(b-a)}(b^2 - a^2)$$

$$= \frac{1}{2(b-a)}(b-a)(b+a) = \frac{b+a}{2}$$

$$\Rightarrow \mu = \frac{b+a}{2}$$

Again from (3),

$${\mu_2}' = (-i)^2\frac{d^2}{dt^2}\left(\emptyset_X(t)\right)\Big|_{t=0} = (-i)^2\frac{1}{3i(b-a)}((ib)^3 - (ia)^3) =$$

$$= \frac{1}{3(b-a)}(b-a)(b^2 + ab + a^2)$$

$$\Rightarrow {\mu_2}' = \frac{b^2 + ab + a^2}{3}$$

Thus, the variance is given by

$$\sigma^2 = {\mu_2}' - ({\mu_1}')^2 = \frac{b^2+ab+a^2}{3} - \frac{(b+a)^2}{4} = \frac{4(b^2+ab+a^2) - 3(b^2+2ab+a^2)}{12} = \frac{(b-a)^2}{12}$$

$$\Rightarrow \sigma^2 = \frac{(b-a)^2}{12}$$

P3:

Find the characteristic function of exponential $E(\lambda)$ distribution and hence find its mean and variance.

Solution:

Since $X \sim E(\lambda)$, its p.d.f. is given by

$$f(x) = \lambda e^{-\lambda x}, x > 0, \lambda > 0$$

The characteristic function of X is given by

$$\emptyset_X(t) = E[e^{itX}] = \int_0^\infty e^{itx} f(x)dx = \lambda \int_0^\infty e^{itx} e^{-\lambda x} dx = \lambda \int_0^\infty e^{-(\lambda - it)x} dx$$

$$= \lambda \left[\frac{e^{-(\lambda - it)x}}{-(\lambda - it)}\right]_0^\infty = \frac{\lambda}{\lambda - it} = \lambda(\lambda - it)^{-1} = \left(1 - \frac{it}{\lambda}\right)^{-1}$$

$$= \sum_{r=0}^{\infty} \left(\frac{it}{\lambda}\right)^r = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \frac{r!}{\lambda^r}$$

$$\Longrightarrow \emptyset_X(t) = \sum_{r=0}^{\infty} \frac{(it)^r}{r!} \frac{r!}{\lambda^r}$$

Thus, $\mu_r' = coef.$ of $\frac{(it)^r}{r!}$ in $\emptyset_X(t) = \frac{r!}{\lambda^r}$.

Hence, the mean is given by $\mu = \mu_1' = \frac{1}{\lambda}$

Further, $\mu_2' = \frac{2!}{\lambda^2} = \frac{2}{\lambda^2}$ and variance is given by

$$\sigma^2 = \mu_2' - (\mu_1')^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Thus, $\sigma^2 = \frac{1}{\lambda^2}$

P4:

Find the p.m.f. of r.v. whose characteristic function is given by

$$\boldsymbol{\phi_X(t) = \frac{p}{1-qe^{it}}}$$

Solution:

Given $\emptyset_X(t) = \frac{p}{1-qe^{it}} = p\left(1 - qe^{it}\right)^{-1}$

$$= p\sum_{x=0}^{\infty}\left(qe^{it}\right)^x = \sum_{x=0}^{\infty} pq^x e^{itx}$$

$$= \sum_{x=0}^{\infty} P(X = x)e^{itx} = E\left[e^{itX}\right]$$

where $P(X = x) = pq^x$, $x = 0,1,2, \ldots$ which is geometric distribution.

3.3. Characteristic Function

Exercise

1. Find the characteristic function of a r.v. X whose moments are given by $\mu_r{}' = (r+1)!\,2^r$.

2. If $\emptyset_X(t) = \frac{3}{3-it}$, then find standard deviation of X.

3. Find the characteristic function of a r.v.X, whose p.d.f. is given by
$$f(x) = \begin{cases} \frac{x}{2} & , \quad 0 \le x \le 2 \\ 0 & , \quad otherwise \end{cases}$$

4. Find the characteristic function and hence obtain the mean and variance of a r.v.X, whose p.d.f. is given by
$$f(x) = \begin{cases} 2e^{-2x} & , \quad x \ge 0 \\ 0 & , \quad otherwise \end{cases}$$

5. Find the characteristic function of a r.v. X whose p.d.f. is given by
$$f(x) = \begin{cases} \frac{1}{3} & , \quad -1 < x < 2 \\ 0 & , \quad otherwise \end{cases}.$$

6. Find the characteristic function and hence find the mean and variance of a r.v.X whose p.d.f. is given by
$$f(x) = \begin{cases} \frac{1}{3} e^{-\frac{x}{3}} & , \quad x \ge 0 \\ 0 & , \quad otherwise \end{cases}$$

7. A r.v.X has the p.d.f. $f(x) = \begin{cases} \frac{1}{k} & , \quad 0 < x < k \\ 0 & , \quad otherwise \end{cases}$.

   Find the characteristic function, mean and variance of X.

8. Find the characteristic function and hence find the mean and variance of a r.v.X whose p.d.f. is given by $f(x) = \lambda e^{-\lambda(x-a)}$, $x \ge a$.

9. Find the p.m.f. whose characteristic function is given by $\emptyset_X(t) = \frac{e^{it}}{2-e^{it}}$.

Answers:

1. $\emptyset_X(t) = \frac{1}{(1-2it)^2}$
2. $\frac{1}{3}$
3. $\frac{1}{2t^2}\left(e^{2it} - 2ite^{2it} - 1\right)$
4. $\emptyset_X(t) = \frac{2}{2-it}$, $\mu = \frac{1}{2}$, $\sigma^2 = \frac{1}{4}$
5. $\emptyset_X(t) = \begin{cases} \frac{e^{2it}-e^{it}}{3it} & , \quad t \neq 0 \\ 1 & , \quad t = 0 \end{cases}$
6. $(1 - 3it)^{-1}$, $\mu = 3, \sigma^2 = 9$
7. $\emptyset_X(t) = \frac{e^{itk}-1}{kit}$, $\mu = \frac{k}{2}, \sigma^2 = \frac{k^2}{12}$
8. $\emptyset_X(t) = \frac{\lambda e^{ait}}{\lambda - it}$, $\mu = \frac{a\lambda+1}{\lambda}, \sigma^2 = \frac{1}{\lambda^2}$
9. $p(x) = \left(\frac{1}{2}\right)^k$, $k = 1,2,3, \ldots$

3.4

Cumulant Generating Function

Just as the moment generating function (m.g.f.) $M_X(t)$ or characteristic function (ch.f.) $\emptyset_X(t)$ of a r.v.X *generates its moments*, the logarithm of $M_X(t)$ or $\emptyset_X(t)$ generates a sequence of numbers called the **Cumulants of $\boldsymbol{X}$**. Cumulants are of interest for the following two reasons.

1. Moments in terms of cumulants can be obtained easily when compared to obtaining them from m.g.f. or ch.f.
2. j^{th} cumulant of a sum of independent r.vs is simply the sum of the j^{th} cumulants of the summand.

Since the ch.f. exists for every r.v(the m.g.f. need not exist for some r.vs), the cumulant generating function (c.g.f.) is defined as the logarithm of the ch.f.

Cumulant generating function: Let X be a r.v. with characteristic function $\emptyset_X(t) = E\left[e^{itX}\right]$. The cumulant generating function (c.g.f.) of X is defined by

$$\boldsymbol{K_X(t) = ln\left(\emptyset_X(t)\right)} \quad \text{... (1)}$$

for all t in some open interval about 0 in $\boldsymbol{R}$, provided the RHS can be expanded as a convergent series in powers of t.

Thus,

$$K_X(t) = k_1(it) + k_2\frac{(it)^2}{2!} + \cdots + k_r\frac{(it)^r}{r!} \quad \text{....(2)}$$

Note that, $k_j =$ coef of $\frac{(it)^j}{j!}$ in $K_X(t)$ and it is called the $\boldsymbol{j^{th}}$ **Cumulant** of X

We have,

$$\emptyset_X(t) = 1 + \mu_1{}'(it) + \mu_2{}'\frac{(it)^2}{2!} + \cdots + \mu_r{}'\frac{(it)^r}{r!} \quad \text{.... (3)}$$

From (1), (2) and (3), we have

$k_1(it) + k_2\frac{(it)^2}{2!} + \cdots = ln[1 + \mu_1'(it) + \cdots + \cdots]$

$= \left(\mu_1'(it) + \mu_2'\frac{(it)^2}{2!} + \cdots\right) - \frac{1}{2}\left(\mu_1'(it) + \mu_2'\frac{(it)^2}{2!} + \cdots\right)^2 + \frac{1}{3}\left(\mu_1'(it) + \mu_2'\frac{(it)^2}{2!} + \cdots\right)^3 - \cdots$

$(\because ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots)$

Comparing the coefficients of like powers of t, we get the relationship between moments and cumulants. Hence, we have

$k_1 = \mu_1' =$ Mean $= \mu$ and $k_2 = \mu_2' - (\mu_1')^2 =$ variance $= \sigma^2$.
Thus, $\mu = k_1$ and $\sigma^2 = k_2$.

Note:

1. From (2), $K_X(t)$ can be written as

$$K_X(t) = \sum_{j=1}^{\infty} k_j \frac{(it)^j}{j!}$$

   Thus j^{th} cumulant $= k_j =$ coef. of $\frac{(it)^j}{j!}$ in $K_X(t)$.

2. From (2), j^{th} cumulant is obtained as

$$k_j = (-i)^j \left.\frac{d^j K_X(t)}{dt^j}\right|_{t=0}$$

Example 1: If $X \sim B(n, p)$, then obtain the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim B(n, p)$, its p.m.f. is given by

$$p(x) = \binom{n}{x} p^x q^{n-x}, x = 0,1,2, \dots, n$$

Then the characteristic function of X is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx}\binom{n}{x}p^x q^{n-x} = \sum_{x=0}^{\infty}\binom{n}{x}\left(pe^{it}\right)^x q^{n-x}$$

$$= \left(q + pe^{it}\right)^n$$

$$\Rightarrow \emptyset_X(t) = \left(q + pe^{it}\right)^n$$

Thus, the c.g.f. of X is given by

$$K_X(t) = ln\left(\emptyset_X(t)\right) = ln\left[q + pe^{it}\right]^n$$

$$\Rightarrow K_X(t) = n\, ln\left(q + pe^{it}\right)$$

$$= n\, ln\left[q + p\left(1 + (it) + \frac{(it)^2}{2!} + \cdots\right)\right]$$

$$= n\, ln\left[1 + (it)p + \frac{(it)^2}{2!}p + \cdots\right]$$

$$\Rightarrow K_X(t) = n\left[\left\{(it)p + \frac{(it)^2}{2!}p + \cdots\right\} - \frac{1}{2}\left\{(it)p + \frac{(it)^2}{2!}p + \cdots\right\}^2 + \cdots\right]$$

$$\Rightarrow K_X(t) = (it)(np) + \frac{(it)^2}{2!}(np - np^2) + \cdots$$

$\therefore k_1 =$ coef. of $(it) = np$ and $k_2 =$ coef. of $\frac{(it)^2}{2!} = np - np^2 = np(1-p) = npq$

Thus mean and variance are given by $\mu = k_1 = np$ and $\sigma^2 = k_2 = npq$

Example 2: If $X\sim$ Poisson $P(\lambda)$, then find the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim P(\lambda)$, , its p.m.f. is given by

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}\ , x = 0,1,2,\ldots$$

The characteristic function of X is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} p(x) = \sum_{x=0}^{\infty} e^{itx} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^{it}\right)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda(e^{it}-1)}$$

Thus, $\emptyset_X(t) = e^{\lambda(e^{it}-1)}$

The c.g.f. of X is given by

$$K_X(t) = ln\left(\emptyset_X(t)\right) = ln\left[e^{\lambda(e^{it}-1)}\right] = \lambda\left(e^{it} - 1\right)$$

$$= \lambda\left[1 + (it) + \frac{(it)^2}{2!} + \frac{(it)^3}{3!} + \cdots - 1\right]$$

$$\Longrightarrow K_X(t) = (it)\lambda + \frac{(it)^2}{2!}\lambda + \cdots$$

Thus, $k_1 =$ coef. of (it) in $K_X(t) = \lambda$ and $k_2 =$ coef. of $\frac{(it)^2}{2!}$ in $K_X(t) = \lambda$

Hence, mean $=$ variance $= \lambda$

Example 3: If $X \sim NB(r, p)$, then find the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim NB(r, p)$, its p.m.f is given by

$$p(x) = \binom{-r}{x} p^r (-q)^x, x = 0,1,2, \dots$$

The characteristic function of X is given by

$$\emptyset_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} p(x) = \sum_{x=0}^{\infty} e^{itx} \binom{-r}{x} p^r (-q)^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} \left(-qe^{it}\right)^x = p^r \left(1 - qe^{it}\right)^{-r}$$

$$\Longrightarrow \emptyset_X(t) = p^r \left(1 - qe^{it}\right)^{-r}$$

The c.g.f. is given by

$$K_X(t) = ln\left(\emptyset_X(t)\right) = ln\left(p^r\left(1 - qe^{it}\right)^{-r}\right)$$

$$= r\,ln\,p - r\,ln\left(1 - qe^{it}\right)$$

Now, $\frac{d}{dt}\left(K_X(t)\right) = (-r)\frac{-iqe^{it}}{1-qe^{it}} = \frac{irqe^{it}}{1-qe^{it}}$

$$\therefore k_1 = (-i)\frac{d}{dt}\left(K_X(t)\right)\Big|_{t=0} = (-i)\frac{(irq)}{1-q} = \frac{rq}{p}$$

And $\frac{d^2}{dt^2}\left(K_X(t)\right) = irq\frac{d}{dt}\left[\frac{e^{it}}{1-qe^{it}}\right] = irq\left[\frac{(1-qe^{it})ie^{it}+e^{it}qie^{it}}{\left(1-qe^{it}\right)^2}\right]$

$$\therefore k_2 = (-i)^2\frac{d^2}{dt^2}\left(K_X(t)\right)\Big|_{t=0} = (-i)^2(irq)(i)\left[\frac{p+q}{p^2}\right]$$

$$\Longrightarrow k_2 = \frac{rq}{p^2}$$

Thus mean and variance are given by

$\mu = k_1 = \frac{rq}{p}$ and $\sigma^2 = k_2 = \frac{rq}{p^2}$ respectively.

Example 4: If $X \sim N(\mu, \sigma^2)$, then obtain the c.g.f. of X and hence find its mean and variance.

Solution: If $X \sim N(\mu, \sigma^2)$, then its characteristic function can be shown that

$$\emptyset_X(t) = e^{it\mu - \frac{1}{2}\sigma^2 t^2} \qquad \text{(recall!)}$$

Hence, the c.g.f. is given by

$$K_X(t) = ln\left[e^{it\mu - \frac{1}{2}\sigma^2 t^2}\right] = it\mu - \frac{1}{2}\sigma^2 t^2$$

$$\Longrightarrow K_X(t) = (it)\mu + \frac{(it)^2}{2!}\sigma^2$$

$\therefore k_1 =$ coef. of (it) in $K_X(t) = \mu$ and $k_2 =$ coef. of $\frac{(it)^2}{2!}$ in $K_X(t) = \sigma^2$

Thus, mean $= \mu$ and variance $= \sigma^2$.

Example 5: If $X \sim E(\lambda)$, then obtain the c.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim E(\lambda)$, its ch.f. can be shown that

$$\emptyset_X(t) = \frac{\lambda}{\lambda - it} = \left(1 - \frac{it}{\lambda}\right)^{-1} \qquad \text{(recall!)}$$

The c.g.f. of X is given by

$$K_X(t) = ln(\emptyset_X(t) = (-1)\, ln\left(1 - \frac{it}{\lambda}\right) = (-1)\left[-\frac{it}{\lambda} - \frac{1}{2}\left(\frac{it}{\lambda}\right)^2 - \cdots\right]$$

$$\Longrightarrow K_X(t) = \left[(it)\frac{1}{\lambda} + \frac{(it)^2}{2!}\frac{1}{\lambda^2} + \cdots\right]$$

Thus, $k_1 =$ coef. of (it) in $K_X(t) = \frac{1}{\lambda}$ and $k_2 =$ coef. of $\frac{(it)^2}{2!}$ in $K_X(t) = \frac{1}{\lambda^2}$.
Thus, the mean and variance are given by $\mu = \frac{1}{\lambda}$ and $\sigma^2 = \frac{1}{\lambda^2}$ respectively.

Properties of Cumulants: Here we develop some useful properties of cumulants. Let $\boldsymbol{k_n(X)}$ be the $\boldsymbol{n}^{\text{th}}$ **cumulant** of a r.v. X.

Theorem 1: $\boldsymbol{k_n(cX) = c^n k_n(X)}$ for some real constant $\boldsymbol{c}$.

Proof: Consider $\emptyset_{cX}(t) = E\left[e^{itcX}\right] = E\left[e^{i(tc)X}\right] = \emptyset_X(tc)$

$$\Longrightarrow \emptyset_{cX}(t) = \emptyset_X(ct)$$

$$\Longrightarrow ln\,\emptyset_{cX}(t) = ln\,\emptyset_X(ct)$$

$$\Longrightarrow K_{cX}(t) = K_X(ct)$$

Then, $\frac{d^n}{dt^n}\left(K_{cX}(t)\right)\Big|_{t=0} = \frac{d^n}{dt^n}\left(K_X(ct)\right)\Big|_{t=0} = \frac{d^n}{ds^n}\left(K_X(s)\right)\Big|_{s=0} c^n$,where $ct = s$

Therefore, $(-i)^n \frac{d^n}{dt^n}\left(K_{cX}(t)\right)\Big|_{t=0} = (-i)^n c^n \frac{d^n}{dt^n}\left(K_X(t)\right)\Big|_{t=0}$

$\Longrightarrow k_n(cX) = c^n k_n(X)$

Theorem 2: $\boldsymbol{k_n(X+b)} = \begin{cases} \boldsymbol{k_n(X)+b} & , \boldsymbol{if} \quad \boldsymbol{n=1} \\ \boldsymbol{k_n(X)} & , \boldsymbol{if} \quad \boldsymbol{n} > 1 \end{cases}$

Proof: $\emptyset_{X+b}(t) = E\left[e^{it(X+b)}\right] = e^{itb}E\left[e^{itX}\right] = e^{itb}\emptyset_X(t)$

$\Longrightarrow \emptyset_{X+b}(t) = e^{itb}\emptyset_X(t)$

$\Longrightarrow ln[\emptyset_{X+b}(t)] = itb + ln[\emptyset_X(t)]$

$\Longrightarrow K_{X+b}(t) = itb + K_X(t)$

$$\Longrightarrow (-i)^n \frac{d^n}{dt^n}\left(K_{X+b}(t)\right)\Big|_{t=0} = (-i)^n \frac{d^n}{dt^n}\left((itb)\right)\Big|_{t=0} + (-i)^n \frac{d^n}{dt^n}\left(K_X(t)\right)\Big|_{t=0}$$

If $n = 1$, then $K_n(X+b) = b + K_n(X)$.

If $n > 1$, $K_n(X+b) = K_n(X)$.

Theorem3: If $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent random variables and $\boldsymbol{S = X + Y}$, then

$$\boldsymbol{k_n(S) = k_n(X) + k_n(Y).}$$

Proof: Since X and Y are independent,

$$\emptyset_S(t) = \emptyset_X(t)\emptyset_Y(t)$$

$\Longrightarrow ln[\emptyset_S(t)] = ln[\emptyset_X(t)] + ln[\emptyset_Y(t)] \Longrightarrow K_S(t) = K_X(t) + K_Y(t)$

$$\Longrightarrow (-i)^n \frac{d^n}{dt^n}\left(K_S(t)\right)\Big|_{t=0} = (-i)^n \frac{d^n}{dt^n}\left(K_X(t)\right)\Big|_{t=0} + (-i)^n \frac{d^n}{dt^n}\left(K_Y(t)\right)\Big|_{t=0}$$

$\Longrightarrow k_n(S) = k_n(X) + k_n(Y)$

Generalization: If $\boldsymbol{S = X_1 + \cdots + X_m}$ where $\boldsymbol{X_1, X_2, \ldots, X_m}$ are independent random variables, then

$$\boldsymbol{k_n(S) = k_n(X_1) + k_n(X_2) + \cdots + k_n(X_m)}$$

Theorem 4: Let $\mu_j' = E(X^j)$ be the j^{th} moment of X about zero for $j = 1, 2, 3, \dots, n$ where $\mu_0' = 1$. Let $k_1, k_2, \dots, k_n$ be the n cumulants of X. Then

$$\mu_{r+1}' = \sum_{j=0}^{r} \binom{r}{j} \mu_j' k_{(r+1-j)} \qquad \text{.... (1)}$$

for $r = 0, 1, \dots, n-1$.

Proof: For $j = 0, 1, 2, \dots, n$, we have

$$\mu_j' = \frac{d^j}{dt^j}\left(\emptyset_X(t)\right)\Big|_{t=0} \text{ and } k_j = (-i)^j \frac{d^j}{dt^j}\left(K_X(t)\right)\Big|_{t=0}$$

where $\emptyset_X(t) = E\left[e^{itX}\right]$ and $K_X(t) = ln[\emptyset(t)]$ or equivalently, $\emptyset_X(t) = e^{K_X(t)}$. Differentiating this last identify w.r.t. t gives

$$\emptyset_X'(t) = e^{K_X(t)} K_X'(t) \qquad \text{...(2)}$$

and evaluating this at $t = 0$ gives $i\mu_1' = ik_1 \Longrightarrow \mu_1' = k_1$ holds for $r = 0$.

Differentiating (2) for r times, it gives

$$\emptyset_X{}^{(r+1)}(t) = \sum_{j=0}^{r} \binom{r}{j} \emptyset_X{}^{(j)}(t) K_X{}^{(r+1-j)}(t)$$

(Use Leibnitz theorem for the n^{th} derivative of the product of two functions)

and evaluating this at $t = 0$ gives

$$\mu_{r+1}' = \sum_{j=0}^{r} \binom{r}{j} \mu_j' k_{(r+1-j)} \quad \text{for } r = 0, 1, \dots, n-1$$

Note: Taking $r = 0,1,2,3$ in (1) produces

$$\left.\begin{aligned} \mu_1' &= k_1 \\ \mu_2' &= k_2 + \mu_1' k_1 \\ \mu_3' &= k_3 + 2\mu_1' k_2 + \mu_2' k_1 \\ \mu_4' &= k_4 + 3\mu_1' k_3 + 3\mu_2' k_2 + \mu_3' k_1 \end{aligned}\right\} \qquad \text{.....(3)}$$

These recursive formulae can be used to calculate the (μ')s efficiently from ks and vice versa.

Let $\mu_j = E\left[\left(X - E(X)\right)^j\right] = E\left[(X - \mu_1')^j\right]$ for $j = 1,2, \dots$ are unknown as **central moments.**

Then formulae (3) simplify to

$\mu_2 = k_2, \mu_3 = k_3,\ \mu_4 = k_4 + 3{k_2}^2$and $k_2 = \mu_2\ \ , k_3 = \mu_3\ , k_4 = \mu_4 - 3{\mu_2}^2$

Note: Mean $= \boldsymbol{\mu} = \boldsymbol{k_1}$ **and variance** $= \boldsymbol{\mu_2} = \boldsymbol{\sigma^2} = \boldsymbol{k_2}$**.**

P1:

Let X be a.r.v with p.d.f

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

Find the $\boldsymbol{c.g.f}$.of X and hence find its mean and variance.

Solution:

The characteristic function of X is given by

$$\phi_X(t) = E\left[e^{itX}\right] = \int_{-a}^{a} e^{itx} f(x)dx = \frac{1}{2a}\int_{-a}^{a} e^{itx}dx$$

$$= \frac{1}{2a}\left[\frac{e^{itx}}{it}\right]_{-a}^{a} = \frac{1}{2ait}\left[e^{ita} - e^{-ita}\right]$$

$$= \frac{1}{2ait}\left[\left(1 + ita + \frac{(ita)^2}{2!} + \frac{(ita)^3}{3!} + \cdots\right) - \left(1 - ita + \frac{(ita)^2}{2!} - \frac{(ita)^3}{3!} + \cdots\right)\right]$$

$$= \frac{1}{2ait}\left[2ita + 2\frac{(ita)^3}{3!} + 2\frac{(ita)^5}{5!} + \cdots\right]$$

$$\Rightarrow \emptyset_X(t) = \left[1 + \frac{(it)^2}{2!}\frac{a^2}{3} + \frac{(it)^4}{4!}\frac{a^4}{5} + \cdots\right]$$

The c.g.f. of X is given by

$$K_X(t) = ln[\emptyset_X(t)] = ln\left[1 + \frac{(it)^2}{2!}\frac{a^2}{3} + \frac{(it)^4}{4!}\frac{a^4}{5} + \cdots\right]$$

$$\Rightarrow K_X(t) = \left(\frac{(it)^2}{2!}\frac{a^2}{3} + \frac{(it)^4}{4!}\frac{a^4}{5} + \cdots\right) - \frac{1}{2}\left[\frac{(it)^2}{2!}\frac{a^3}{3} + \frac{(it)^4}{4!}\frac{a^4}{5} + \cdots\right]^2 + \cdots$$

$\Rightarrow k_1$ =coeff. of (it) in $K_X(t) = 0$ and k_2 =coeff.of $\frac{(it)^2}{2!}$ in $K_X(t) = \frac{a^3}{3}$

Thus, the mean and variance are given by $\mu = 0$ and $\sigma^2 = \frac{a^3}{3}$

P2:

Let X be r.v. with p.d.f

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

Find the c.g.f. of and hence find its mean and variance.

Solution:

The characteristic function of X is given by

$$\phi_X(t) = E\left[e^{itX}\right] = \int_{-\infty}^{\infty} e^{itx} f(x)\,dx = \frac{1}{2}\int_{-\infty}^{\infty} e^{itx} e^{-|x|}\,dx$$

$$= \frac{1}{2}\left[\int_{-\infty}^{\infty} (\cos tx + i \sin tx) e^{-|x|}\,dx\right] = \frac{1}{2} 2\int_0^{\infty} (\cos t\,x) e^{-|x|}\,dx$$

(Since the integrals in the first and second integrals are even and odd functions of x repsectively).

$$= \int_0^{\infty} e^{-x} \cos tx\,dx$$

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

(On integration by parts)

$$\Rightarrow \quad \emptyset_X(t) = \frac{1}{t^2} - \frac{1}{t^2}\,\emptyset_X(t) \Rightarrow \quad \emptyset_X(t) = \frac{1}{1+t^2}$$

The c.g.f. is given by $K_X(t) = ln\big(\emptyset_X(t)\big) = ln\left[\frac{1}{1+t^2}\right] = -\,ln(1+t^2)$

$$= -\left(t^2 - \frac{t^4}{2} + \frac{t^6}{3} - \cdots\right) = -t^2 + \frac{t^4}{2} - \frac{t^6}{3} + \cdots$$

$$\Rightarrow K_X(t) = \frac{(it)^2}{2!}2 + \cdots$$

$\Rightarrow k_1 = coeff.of\,(it)$ in $K_X(t) = 0$ and $k_2 = coeff.of\ \frac{(it)^2}{2!}$ in $K_X(t) = 2$

Thus, the mean and variance are given by $\mu = 0$ and $\sigma^2 = 2$

P3:

Let X be a.r.v. with p.m.f given by

$$p(x) = \begin{cases} q^x p & , \quad x = 0, 1, 2, \dots, \; 0 < p < 1, \; q = 1 - p \\ 0 & , \quad otherwise \end{cases}$$

Find the c.g.f. and hence find its mean and variance.

Solution:

The characteristic function of X is given by

$$\phi_X(t) = E\left[e^{itX}\right] = \sum_{x=0}^{\infty} e^{itx} p(x) = \sum_{x=0}^{\infty} e^{itx} q^x p$$

$$= p \sum_{x=0}^{\infty} \left(qe^{it}\right)^x = p\left(1 - qe^{it}\right)^{-1}$$

$$\Longrightarrow \; \emptyset_X(t) = p\left(1 - qe^{it}\right)^{-1}$$

The c.g.f. is given by $K_X(t) = ln[\emptyset_X(t)] = ln\, p - ln\left(1 - qe^{it}\right)$

Now, $\frac{d}{dt}\left(K_X(t)\right) = \frac{qie^{it}}{1-qe^{it}} = (iq)\left[\frac{e^{it}}{1-qe^{it}}\right]$

$$\therefore \; k_1 = (-i)\frac{d}{dt}\left(K_X(t)\right)\Big|_{t=0} = (-i)(iq)\frac{1}{p} = \frac{q}{p}$$

Further, $\frac{d^2}{dt^2}\left(K_X(t)\right) = (iq)\left[\frac{(1-qe^{it})ie^{it} + e^{2it}\, iq}{\left(1-qe^{it}\right)^2}\right]$ and

$$k_2 = (-i)^2 \frac{d^2}{dt^2}\left(K_X(t)\right)\Big|_{t=0} = (-i)^2 (iq)\frac{(ip + iq)}{p^2} \Longrightarrow k_2 = \frac{q}{p^2}$$

Thus, the mean and variance are given by $\mu = \frac{q}{p}$ and $\sigma^2 = \frac{q}{p^2}$ respectively.

P4:

Let X be a r.v. with p.d.f.

$$\boldsymbol{f(x) = \frac{1}{2\lambda} exp\left[\frac{-|x-\mu|}{\lambda}\right],} -\infty < x, \mu < \infty, \quad \sigma > 0$$

Find the c.g.f. and hence obtain its mean and variance.

Solution:

It can be shown that the characteristic function is given by

$$\emptyset_X(t) = \frac{e^{it\mu}}{1+\lambda^2 t^2} \qquad \text{(do it!)}$$

The c.g.f. is given by $K_X(t) = ln[\emptyset_X(t)] = ln\left[\frac{e^{it\mu}}{1+\lambda^2 t^2}\right]$

$$= it\mu - ln(1 + \lambda^2 t^2)$$

$$= it\mu - (\lambda^2 t^2 + \cdots)$$

$$\Rightarrow K_X(t) = it\mu + \frac{(it)^2}{2!} 2\lambda^2 + \cdots$$

$$\Rightarrow k_1 = coef.of(it) = \mu \text{ and } k_2 = coef.of \frac{(it)^2}{2!} = 2\lambda^2$$

Thus, the mean and variance are given by μ and $2\lambda^2$ respectively.

3.4. Cumulant Generating Function

Exercise

1. Let $U = \frac{X-a}{h}$ where a and h are real constants. Obtain the cumulants of U interms of cumulants of X.

2. For a distribution, the cumulants are given by
$$k_r = n(r-1)!\,, n > 0$$
Find the ch.f.

3. The p.m.f of X is given by
$$p(x) = q^{r-1}p\,, r = 1, 2, 3, \ldots$$
Find the c.g.f. and hence find its mean and variance.

4. The p.d.f. of X is given by
$$f(x) = \frac{e^{-x}}{(1+e^{-x})^2}, \qquad -\infty < x < \infty$$
Find the c.g.f. and hence find its mean and variance.

Answers

1. $k_1(U) = \frac{k_1(X) - a}{h}$ and $k_r(U) = \frac{k_r(X)}{h^r}$ for $r = 2, 3 \ldots$
2. $\emptyset_X(t) = (1 - it)^{-n}$
3. $\mu = \frac{1}{p}$ and $\sigma^2 = \frac{1+q}{p^2}$
4. $\mu = 0$ and $\sigma^2 = \frac{\pi^2}{3}$

3.5

Probability Generating Function

Let X be a non-negative integer valued random variable with p.m.f. $p(x) = P(X = x)$. Then the **probability generating function** (p.g.f.) of X is defined by

$$\boldsymbol{G_X(t) = E[t^X] = \sum_{x=0}^{\infty} t^x p(x)}$$

where $-1 \leq t \leq 1$ is a dummy variable.

Advantages:

1. It is easy to compute.
2. Moments and some probabilities can be obtained easily.
3. The p.m.f. can be obtained easily from p.g.f.
4. It is easy to handle with sum of independent r.vs.

Effect of linear transformation of p.g.f:

Theorem 1: Let X be a discrete random variable with p.g.f. $G_X(t)$. Let $Y = a + bX$ where a and b are real constants. Then $G_Y(t) = t^a G_X(t^b)$

Proof: By the definition of probability generating function, we have, $G_X(t) = E[t^X]$. Then

$$G_Y(t) = E\left[t^{(a+bX)}\right] = E[t^a t^{bX}] = t^a E[(t^b)^X] = t^a G_X(t^b)$$

$$\Longrightarrow G_Y(t) = t^a G_X(t^b)$$

Theorem 2: Additive Property: If X and Y are independent random variables, then for constants a, b, we have

$$\boldsymbol{G_{(aX+bY)}(t) = G_X(t^a) + G_Y(t^b)}$$

Proof: $G_{aX+bY}(t) = E[t^{aX+bY}]$ (by P.g.f.)

$$= E[(t^a)^X (t^b)^Y]$$

$$= E[(t^a)^X]E[(t^b)^Y] \quad (\because X\&Y \text{ are indpendent.})$$

$$= G_X(t^a)G_Y(t^b)$$

Thus, $G_{aX+bY}(t) = G_X(t^a)G_Y(t^b)$

Note: In particular, if $a = b = 1$, then $\boldsymbol{G_{X+Y}(t) = G_X(t)G_Y(t)}$

Generalization: If $X_1, X_2, \dots, X_n$ are independent random variables, then

$$\boldsymbol{G_{(X_1+\cdots+X_n)}(t) = G_{X_1}(t)G_{X_2}(t) \dots G_{X_n}(t)}$$

Relationship between p.g.f. and m.g.f :

The p.g.f. and m.g.f. of a random variable X are defined by $G_X(t) = E[t^X]$ and $M_X(t) = E[e^{tX}]$ respectively.

Now, $M_X(t) = E[e^{tX}] = E[(e^t)^X] = G_X(e^t)$

$$\Longrightarrow \boldsymbol{M_X(t) = G_X(e^t)}$$

Further, $G_X(t) = E[t^X] = E\left[e^{ln(t^X)}\right] = E[e^{X\,ln\,t}] = M_X(ln\,t)$

$$\Longrightarrow \boldsymbol{G_X(t) = M_X(ln\,t)}$$

Theorem 3: p.m.f. from p.g.f : Let $G_X(t)$ be the p.g.f. of a discrete r.v. X that can take the values 0,1,2,…. Then the p.m.f. of X is given by

$$\boldsymbol{p(x) = P(X = x) = \frac{1}{x!} G_X^{(x)}(t)\Big|_{t=0}}$$

Proof: By definition, we have

$$G_X(t) = E(t^X) = \sum_{x=0}^{\infty} t^x p(x)$$

$$= P(X=0)t^0 + P(X=1)t^1 + P(X=2)t^2 + \cdots + P(X=x)t^x + \cdots$$

It can be observed that the coefficient of t^x in $G_X(t)$ is $P(X=x)$. To obtain coefficient of t^x, differentiate $G_X(t)$, x times and substitute $t=0$. Thus,

$$G_X^{(x)}(t) = x(x-1)(x-2)\ldots 2\cdot 1\cdot P(X=x) + (x+1)(x) + \cdots 2\cdot 1\cdot t\cdot P(X=x+1) + \cdots$$

When $t=0$, all terms after the first vanish. Thus,

$$\boldsymbol{P(X=x) = \frac{1}{x!} G_X^{(x)}(t)\Big|_{t=0} = \frac{1}{x!} G_X^{(x)}(0)}$$

Computation of moments using p.g.f:

In the derivation of moments, we use *Taylor's expansion*:

Suppose $f(x)$ has derivatives of all orders at $x=a$. The Taylor's expansion of $f(x)$ at the point $x=a$ is givne by

$$f(x) = f(a) + f'(a)(x-a) + \tfrac{f''(a)}{2!}(x-a)^2 + \cdots + \tfrac{f^n(a)}{n!}(x-a)^n + \cdots$$

$$\Rightarrow f(x) = \sum_{i=0}^{\infty} \frac{f^i(a)}{i!}(x-a)^i$$

The Taylor's expansion of $f(t) = t^X$ about $t=1$ is given by

$$t^X = 1 + X(t-1) + X(X-1)\frac{(t-1)^2}{2!} + X(X-1)(X-2)\frac{(t-1)^3}{3!} + \cdots$$

$$\Rightarrow G_X(t) = E[t^X]$$

$$= 1 + (t-1)E(X) + \tfrac{(t-1)^2}{2!}E[X(X-1)] + \tfrac{(t-1)^3}{3!}E[X(X-1)(X-2)] + \cdots$$

Differentiating (1) w.r.t., t r times and setting $t=1$, we get

$$G_X^{(r)}(t)\Big|_{t=1} = E[X(X-1)\ldots(X-r+1)]$$

$$\Longrightarrow \boldsymbol{E[X(X-1)\ldots(X-r+1)] = G_X^{(r)}(1)} \qquad \ldots(2)$$

which is known as $\boldsymbol{r}^{\text{th}}$ **factorial moment of $\boldsymbol{X}$**. Using these, we can find the moments about origin as follows:

If $r = 1$ in (2), we have

$$\boldsymbol{\mu_1' = E(X) = G_X^{(1)}(1)}$$

If $r = 2$ in (2), we have

$$E[X(X-1)] = E[X^2 - X] = E(X^2) - E(X) = G_X^{(2)}(1)$$

$$\Longrightarrow E(X^2) = G_X^{(2)}(1) + E(X) = G_X^{(2)}(1) + G_X^{(1)}(1)$$

Thus, the second moment about origin is given by

$$\boldsymbol{\mu_2' = E(X^2) = G_X^{(2)}(1) + G_X^{(1)}(1)}$$

Similarly, we can find any moment about origin.

Computation of mean and variance using p.g.f:

Theorem 4: If the r.v. X has p.g.f. $\boldsymbol{G_X(t)}$, then the mean and variance of $\boldsymbol{X}$ are given by

$$\boldsymbol{\mu = E(X) = G_X^{(1)}(1)} \text{ and}$$

$$\boldsymbol{\sigma^2 = V(X) = G_X^{(2)}(1) + G_X^{(1)}(1) - \left[G_X^{(1)}(1)\right]^2}$$

respectively.

Proof: From the above, we have

$$\mu_1' = G_X^{(1)}(1),\ \mu_2' = G_X^{(2)}(1) + G_X^{(1)}(1)$$

Thus , the mean $\mu = \mu_1' = G_X^{(1)}(1)$ and variance $\sigma^2 = \mu_2' - (\mu_1')^2$

$$\Rightarrow \sigma^2 = G_X^{(2)}(1) + G_X^{(1)}(1) - \left(G_X^{(1)}(1)\right)^2$$

Convolution formula:

Theorem 5: If X and Y are independent integer-valued random variables with $\boldsymbol{P(X=x)=p_1(x)}$ and $\boldsymbol{P(Y=y)=p_2(y), x=0,1,2,\ldots}$ and $\boldsymbol{y=0,1,2,\ldots}$, then

$$\boldsymbol{P(X+Y=z)=p(z)=\sum_{x=0}^{z} p_1(x)p_2(z-x)}$$

Proof: We have ,

$$G_X(t) = \sum_{x=0}^{\infty} t^x p_1(x) \quad and \quad G_Y(t) = \sum_{y=0}^{\infty} t^y p_2(y)$$

Now, $\quad G_{X+Y}(t) = G_X(t)G_Y(t) \quad$ (Since X and Y are independt)

$$= \left(\sum_{x=0}^{\infty} t^x p_1(x)\right)\left(\sum_{y=0}^{\infty} t^y p_2(y)\right)$$

$$\Rightarrow G_{X+Y}(t) = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} p_1(x)p_2(y)t^{x+y} \qquad \ldots(1)$$

Let $Z = X + Y$. Then

$$G_Z(t) = E[t^Z] = \sum_{z=0}^{\infty} t^z p(z) \qquad \ldots(2)$$

From (1) and(2), we have

$$\sum_{z=0}^{\infty} t^z p(z) = \sum_{x=0}^{\infty}\sum_{y=0}^{\infty} p_1(x)p_2(y)t^{x+y}$$

$$\Rightarrow \sum_{z=0}^{\infty} t^z p(z) = \sum_{z=0}^{\infty} \left(\sum_{x=0}^{z} p_1(x) p_2(z-x) \right) t^z$$

$$\Rightarrow p(z) = \sum_{x=0}^{z} p_1(x) p_2(z-x) \text{ , for } z = 0,1,2, \dots$$

Example 1: If $X \sim B(n,p)$, then find the p.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim B(n,p)$, its p.m.f. is given by

$$p(x) = \binom{n}{x} p^x q^{n-x} \text{ , } x = 0,1, \dots, n, \qquad 0 < p < 1, q = 1-p$$

The p.g.f. of X is given by

$$G_X(t) = E[t^X] = \sum_{x=0}^{n} t^x p(x) = \sum_{x=0}^{n} t^x \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \binom{n}{x} (tp)^x q^{n-x} = (q+tp)^n$$

$$\Rightarrow G_X(t) = (q+tp)^n$$

Differentiating both sides w.r.t., t we get

$$G_X^{(1)}(t) = n(q+tp)^{n-1} p$$

$\Rightarrow \mu = \text{mean} = \mu_1{}' = G_X^{(1)}(1) = np$ and variance is given by

$$\sigma^2 = G_X^{(2)}(1) + G_X^{(1)}(1) - \left[G_X^{(1)}(1) \right]^2$$

But $G_X^{(2)}(t) = np(n-1)(q+tp)^{n-1} p$

$$\Rightarrow G_X^{(2)}(t) = n(n-1)p^2 = n^2p^2 - np^2$$

Therefore, $\sigma^2 = n^2p^2 - np^2 + np - n^2p^2 = np(1-p) \Rightarrow \sigma^2 = npq$

Example 2: If $X \sim P(\lambda)$, then find the p.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim P(\lambda)$, its p.m.f. is given by

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!}, \ x = 0,1,\dots \text{ and } \lambda > 0$$

The p.g.f. of X is given by

$$G_X(t) = E[t^X] = \sum_{x=0}^{\infty} t^x p(x) = \sum_{x=0}^{\infty} t^x \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^{\infty}\frac{(t\lambda)^x}{x!} = e^{-\lambda}e^{t\lambda} = e^{\lambda(t-1)}$$

$\Longrightarrow G_X(t) = e^{\lambda(t-1)}$

Differentiating both sides w.r.t. t, we get

$G_X^{(1)}(t) = e^{\lambda(t-1)}\lambda$ and $G_X^{(2)}(t) = e^{\lambda(t-1)}\lambda^2$

Thus, $G_X^{(1)}(1) = \lambda$ and $G_X^{(2)}(1) = \lambda^2$

Hence, the mean and variance are given by $\mu = G_X^{(1)}(1) = \lambda$
and $\sigma^2 = G_X^{(2)}(1) + G_X^{(1)}(1) - \left[G_X^{(1)}(1)\right]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$ respectively.

Example 3: If $X \sim NB(r, p)$, then find the p.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim P(\lambda)$, its p.m.f. is given by

$$p(x) = \binom{-r}{x} p^x(-q)^x, x = 0,1,2,\dots$$

The p.g.f. of X is given by

$$G_X(t) = E[t^X] = \sum_{x=0}^{\infty} t^x p(x) = \sum_{x=0}^{\infty} t^x \binom{-r}{x} p^r(-q)^x$$

$$= p^r \sum_{x=0}^{\infty} \binom{-r}{x} (-tq)^x = p^r(1 - tq)^{-r}$$

$\Longrightarrow G_X(t) = p^r(1-tq)^{-r}$

$\Longrightarrow G_X^{(1)}(t) = p^r(-r)(1-tq)^{-(r+1)}(-q) = rqp^r(1-tq)^{-(r+1)}$

$\Longrightarrow G_X^{(2)}(t) = rqp^r\big(-(r+1)\big)(1-tq)^{-(r+2)}(-q) = r(r+1)q^2p^r(1-tq)^{-(r+2)}$

Thus, $G_X^{(1)}(t) = rqp^r p^{-(r+1)} = \frac{rq}{p}$ and

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

$$\Longrightarrow G_X^{(2)}(t) = \frac{r^2q^2}{p^2} + \frac{rq^2}{p^2}$$

Thus, $\mu =$ mean $= G_X^{(1)}(1) = \frac{rq}{p}$ and

$\sigma^2 =$ variance $= G_X^{(2)}(1) + G_X^{(1)}(1) - \left[G_X^{(1)}(1)\right]^2$

$$= \frac{r^2q^2}{p^2} + \frac{rq^2}{p^2} + \frac{rq}{p} - \frac{r^2q^2}{p^2} = \frac{rq}{p^2}(q+p) \Longrightarrow \sigma^2 = \frac{rq}{p^2}$$

Example4: If $X \sim G(p)$, then find the p.g.f. of X and hence obtain its mean and variance.

Solution: Since $X \sim G(p)$, its p.m.f. is given by

$$p(x) = q^x p, \qquad x = 0,1,2,\ldots$$

The p.g.f. of X is given by

$$G_X(t) = E[t^X] = \sum_{x=0}^{\infty} t^x p(x) = \sum_{x=0}^{\infty} t^x q^x p = p\sum_{x=0}^{\infty}(tq)^x = \frac{p}{1-tq}$$

$\Longrightarrow G_X(t) = p(1-tq)^{-1}$

$G_X^{(1)}(t) = p(-1)(1-tq)^{-2}(-q) = pq(1-tq)^{-2}$

$\Longrightarrow G_X^{(1)}(1) = \frac{pq}{p^2} = \frac{q}{p}$

Now, $G_X^{(2)}(t) = pq(-2)(1 - tq)^{-3}(-q) = 2pq^2(1 - tq)^{-3}$

$\Longrightarrow G_X^{(2)}(1) = \frac{2pq^2}{p^3} = \frac{2q^2}{p^2}$

Hence, the mean μ and variance σ^2 of X are given by:

$\mu = G_X^{(1)}(1) = \frac{q}{p}$ and

$\sigma^2 = G_X^{(2)}(1) + G_X^{(1)}(1) - \left[G_X^{(1)}(1)\right]^2$

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

Example 5:The j.p.m.f. of (X, Y) is given in the following table. Prove or diprove

$G_{x+y}(t) = G_X(t)G_y(t)$ iff X and Y are independent.

X \ Y	0	1	2	Total
0	$\frac{1}{9}$	$\frac{2}{9}$	**0**	$\frac{1}{3}$
1	**0**	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{3}$
2	$\frac{2}{9}$	**0**	$\frac{1}{9}$	$\frac{1}{3}$
Total	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	**1**

Solution: Since $P(X = 1, Y = 2) = \frac{2}{9} \neq P(X = 1)P(Y = 2) = \frac{1}{3}\frac{1}{3} = \frac{1}{9}$, it follows that X and Y are not independent.

Now, $G_X(t) = G_Y(t) = \frac{1}{3}(1 + t + t^2)$

Let $Z = X + Y$. Then $Z = 0, 1, 2, 3, 4$. Let $p_i = P(Z = i), i = 0,1,2,3,4$.

$p_0 = P(Z = 0) = P(X + Y = 0) = P(X = 0, Y = 0) = \frac{1}{9}$

$p_1 = P(Z = 1) = P(X + Y = 1) = P(X = 0, Y = 1) + P(X = 1, Y = 0) = \dfrac{2}{9} + 0 = \dfrac{2}{9}$

$p_2 = P(Z = 2) = P(X + Y = 2) = P(X = 0, Y = 2) + P(X = 1, Y = 1) + P(X = 2, Y = 0)$

$$= 0 + \frac{1}{9} + \frac{2}{9} = \frac{3}{9}$$

$p_3 = P(Z = 3) = P(X + Y = 3) = P(X = 1, Y = 2) + P(X = 2, Y = 1) = \frac{2}{9} + 0 = \frac{2}{9}$

$p_4 = P(Z = 4) = P(X + Y = 4) = P(X = 2, Y = 2) = \frac{1}{9}$

The p.d.f. of $Z = X + Y$ is given by

$G_{X+Y}(t) = \frac{1}{9} + \frac{2}{9}t + \frac{3}{9}t^2 + \frac{2}{9}t^3 + \frac{1}{9}t^4$

$\Longrightarrow\ G_{X+Y}(t) = \frac{1}{9}(1 + 2t + 3t^2 + 2t^3 + t^4) = \left[\frac{1}{3}(1 + t + t^2)\right]^2$

$\Longrightarrow\ G_{X+Y}(t) = G_X(t)G_Y(t)$ but X and Y are not independent. Thus the statement is disproved.

Example 6: Can $G_X(t) = \frac{2}{1+t}$ be the p.d.f. of .r.v. X? Give reasons.

Solution: We have $G_X(1) = \frac{2}{1+1} = \frac{2}{2} = 1$

Further, $G_X(t) = \frac{2}{1+t} = 2(1 + t)^{-1} = 2(1 - t + t^2 - t^3 + \cdots)$

$$\Longrightarrow G_X(t) = 2\sum_{x=0}^{\infty}(-1)^x t^x$$

Thus, $p(x) = P(X = x) = coef.of\ t^x$ in $G_X(t) = 2(-1)^x$

$\Longrightarrow p(x) = 2(-1)^x\ , x = 0, 1, 2, \ldots$

Note that it takes negative values also. Hence, $G_X(t)$ is not a p.g.f.

Example 7 : A fair die is thrown n times. Let S be the total number of points. Show that $P(S = n + 5) = \binom{n+4}{5}\left(\frac{1}{6}\right)^n$.

Solution: The p. g. f. of a single throw is given by:

$$G_X(t) = \sum_{x=1}^{6} t^x p(x) = \sum_{x=1}^{6} \frac{t^x}{6}$$

$$= \frac{1}{6}(t + t^2 + \cdots + t^6) = \frac{t}{6}(1 + t + \cdots + t^5) = \frac{t}{6}\frac{(1-t^6)}{1-t}$$

$$\Longrightarrow\ G_X(t) = \frac{t}{6}(1 - t^6)\,(1 - t)^{-1}$$

Since the n throws are identical and independent,

$$G_S(t) = [G_S(t)]^n = \frac{t^n\left(1-t^6\right)^n(1-t)^{-n}}{6^n}$$

$$= \frac{t^n}{6^n}\sum_{j=0}^{n}\binom{n}{j}\left(-t^6\right)^j\sum_{k=0}^{\infty}\binom{n+k-1}{k}t^k$$

$$\Rightarrow G_S\left(t\right) = \frac{1}{6^n}\sum_{j=0}^{n}\sum_{k=0}^{\infty}(-1)^j\binom{n}{j}\binom{n+k-1}{k}\,t^{k+6j+n}$$

$$= \sum_{k=0}^{\infty} P\left(S = k + 6\,j + n\right)\,t^{k+6\,j+n}$$

where,

$$P(S = k + 6j + n) = \frac{1}{6^n} \sum_{j=0}^{n} (-1)^j \binom{n}{j} \binom{n+k-1}{k}$$

Now , $P(S = n + 5) = P(S = k + 6j + n)$ with $j = 0$ and $k = 5$

$$= \frac{1}{6^n} (-1)^0 \binom{n}{0} \binom{n+5-1}{5} = \frac{1}{6^n} \binom{n+4}{5}$$

$$\Longrightarrow P(S = n + 5) = \frac{1}{6^n} \binom{n+4}{5}$$

P1:

Urns U_1 and U_2 have the following distribution

U_1

x	0	1	2	3
$p_1(x)$	0.4	0.2	0.1	0.3

U_2

y	1	2	3
$p_1(y)$	0.3	0.6	0.1

A ball is drawn from each urn and the numbers X, Y appearing are added. Find the p.m.f of $X + Y$ by using p.g.f.

Solution:

Here X and Y are independent. The p.g.fs of X and Y are given by

$$G_X(t) = \sum_{x=0}^{3} t^x p_1(x) = 0.4 + 0.2t + 0.1t^2 + 0.3t^3$$

and

$$G_Y(t) = \sum_{y=1}^{3} t^y p_2(y) = 0.3t + 0.6t^2 + 0.1t^3$$

respectively. Since X and Y are independent

$$G_{X+Y}(t) = G_X(t)G_Y(t) = (0.4 + 0.2t + 0.1t^2 + 0.3t^3)(0.3t + 0.6t^2 + 0.1t^3)$$

$$= 0.12t + 0.30t^2 + 0.19t^3 + 0.17t^4 + 0.19t^5 + 0.03t^6$$

Thus the p.m.f. of $X + Y$ is given by

r	1	2	3	4	5	6
$P(X + Y = r)$	0.12	0.30	0.19	0.17	0.19	0.03

P2:

Lottery tickets bear numbers from 000000 to 999999. Find the probability that a ticket bears a number whose sum of the first three digits equals the sum of the last three digits.

Solution:

Let the ticket bear the number $Y_1Y_2Y_3Y_4Y_5Y_6$. Obviously, all Ys are identically and independently distributed with

$$p(x) = P(Y_j = x) = \frac{1}{10}, \qquad x = 0, 1, 2, \ldots, 9$$

The p.g.f. of each Y is given by

$$G_Y(t) = \frac{1}{10}(1 + t + \cdots + t^9)$$

$$\Longrightarrow \quad G_Y(t) = \frac{1}{10}\frac{(1-t^{10})}{(1-t)}$$

Let $X_1 = Y_1 + Y_2 + Y_3$ and $X_2 = Y_4 + Y_5 + Y_6$.

Then $G_{X_1}(t) = G_{Y_1}(t)G_{Y_2}(t)G_{Y_3}(t)$

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

Similarly $G_{X_2}(t) = \left(\frac{1}{10}\right)^3 (1 - t^{10})^3(1 - t)^{-3}$

Let $G_{X_1}(t) = G_{X_2}(t) = G(t)$

Then $P(X_1 = r) = coef.of\ t^r$ in $G(t)$

and $P(X_2 = r) = coef.of\ t^{-r}$ in $G(t^{-1})$

$\therefore p = P(X_1 = X_2) = coef.of\ t^{\circ}$ in $G(t)G(t^{-1})$

$= coef.of\ t^{\circ}$ in $\left(\frac{1}{10}\right)^3 (1-t^{10})^3(1-t)^{-3}.\left(\frac{1}{10}\right)^3 \left(1-\frac{1}{t^{10}}\right)^3 \left(1-\frac{1}{t}\right)^{-3}$

$\Longrightarrow p = coef.of\ \ t^{\circ}$ in $\left(\frac{1}{10}\right)^3 (1-t^{10})^6.t^{-27}(1-t)^{-6}$

Now, $(1-t^{10})^6 = 1 - \binom{6}{1} t^{10} + \binom{6}{2} t^{20} - \cdots$

and $(1-t)^{-6} = 1 + \binom{6}{5} t + \binom{7}{5} t^2 + \cdots \binom{5+r}{5} t^r + \cdots$

$\therefore\ \ p = \left(\frac{1}{10}\right)^6 \left[\binom{32}{5} - \binom{6}{1}\binom{22}{5} + \binom{6}{2}\binom{12}{5}\right]$

P3:

Find the p.g.f. of $P(X > n)$.

Solution:

The p.g.f. of X is defined by

$G(t) = \sum_{x=0}^{\infty} p(x) t^x$,where $p(x)$ is the p.m.f. of X.

Let $f_n = P(X > n)$ and $p(n) = P(X = n)$

Consider $P(X = n + 1) = P(X > n) - P(X > n + 1)$

$$\Rightarrow \quad p(n + 1) = f_n - f_{n+1}$$

$$\therefore \quad t^n f_n - \left(\frac{1}{t}\right) t^{n+1} f_{n+1} = \frac{1}{t} p(n + 1)$$

$$\Rightarrow \quad \sum_{n=0}^{\infty} t^n f_n - \frac{1}{t} \sum_{n=0}^{\infty} t^{n+1} f_{n+1} = \frac{1}{t} \sum_{n=0}^{\infty} t^{n+1} p(n+1) \qquad \dots (1)$$

Let $H(t) = \sum_{n=0}^{\infty} t^n f_n$ $\qquad \dots (2)$

Now, $\sum_{n=0}^{\infty} t^{n+1} p(n+1) = tp(1) + t^2 p(2) + t^3 p(3) + \cdots$

$$\Rightarrow \sum_{n=0}^{\infty} t^{n+1} p(n+1) = G(t) - p(0) \qquad \dots (3)$$

Note that $\sum_{n=0}^{\infty} t^{n+1} f_{n+1} = H(t) - f_0$ $\qquad \dots (4)$

But $f_0 = P(X > 0) = 1 - P(X = 0) = 1 - p(0)$ $\qquad \dots (5)$

From (4) and (5), we have

$$\sum_{n=0}^{\infty} t^{n+1} f(n+1) = H(t) - 1 + p(0) \quad \dots (6)$$

On substituting (2), (3) and (6) in (1), we get

$$H(t) - \frac{1}{t}[H(t) - 1 + p(0)] = \frac{1}{t}[G(t) - p(0)]$$

$$tH(t) - H(t) + 1 - p(0) = G(t) - p(0)$$

$$\Longrightarrow (t-1)H(t) + 1 = G(t)$$

$$\Longrightarrow H(t) = \frac{G(t)-1}{t-1} \Longrightarrow H(t) = \frac{1-G(t)}{1-t}$$

P4:

Find the p.g.f. of $\boldsymbol{P}(\boldsymbol{X} < n)$.

Solution:

From the relation $P(X < n+1) - P(X < n) = P(X = n)$,

we get $\emptyset_{n+1} - \emptyset_n = p(n)$ where $\emptyset_n = P(X < n)$

$$\therefore \quad \frac{1}{t}\left(\sum_{n=0}^{\infty} t^{n+1}\phi_{n+1}\right) - \left(\sum_{n=0}^{\infty} t^n \phi_n\right) = \left(\sum_{n=0}^{\infty} t^n p(n)\right)$$

or $t^{-1}[H(t) - \emptyset_0] - H(t) = G(t)$

Since, $\emptyset_0 = P(X < 0) = 0$, the above result is

$$t^{-1}H(t) - H(t) = G(t) \Rightarrow (1-t)H(t) = tG(t) \Rightarrow H(t) = \frac{t.G(t)}{1-t}$$

3.5. Probability Generating Function

Exercise

1. Let X be a $\mathrm{r.v.}$ with the following p.m.f.

x	1	2	3	4
$p(x)$	$\frac{1}{3}$	$\frac{1}{6}$	$\frac{1}{8}$	$\frac{3}{8}$

   Find the p.g.f.

2. If $p(x) = P(X = x) = \frac{1}{2^{x-2}}$ for $x = 3, 4, 5, \dots$. Find the p.g.f. of X

3. Let X be a $r.v.$ with p.g.f. $G_X(t)$. Find the p.g.f. of

   (i) $X + 1$
   (ii) $2X$

4. If $G(t)$ is the p.g.f. of X, then find the p.g.f. of $\frac{X-a}{b}$ where a and b are real constants.

5. Let $G_X(t) = \frac{\left(\frac{1}{3}t+\frac{2}{3}\right)^4}{t}$. What is the range of X? What is its p.m.f.?

Answers

1. $G_X(t) = \frac{1}{3}t + \frac{1}{6}t^2 + \frac{1}{8}t^3 + \frac{3}{8}t^4$

2. $G_X(t) = \frac{t^3}{2-t}$

3. $(i) \quad tG_X(t)$

   $(ii) \quad G_X(t^2)$

4. $t^{-\frac{1}{b}} G\left(t^{\frac{1}{b}}\right)$

5.

x	-1	0	1	2	3
$p(x)$	$\frac{16}{81}$	$\frac{32}{81}$	$\frac{24}{81}$	$\frac{8}{81}$	$\frac{1}{81}$