Functions of Random Variables

→ We have discussed specific types of distribution, both discrete and continous. Now we discuss the distributions of functions of Random variables.

- functions of R.V.'s are required statistical hypothesis test, estimation or even statistical graphics, example sum of random variables, linear combinations of R.V., sums of squares of R.V.

#. Transformations of Variables :-

→ We shall first discuss the transformation of discrete random variables.

→ X :→ discrete random variable with probability distribution $f(x)$ and suppose that $Y = u(X)$ defines a one-to-one transformation between the values of X and Y, so that the equation $y = u(x)$ can be uniquely solved for x in terms of y, say, $x = w(y)$. Then the probability distribution of Y is

$$g(y) = f[w(y)].$$

$$g(y) = P[Y=y]$$
$$= P[X=x]$$
$$= P[X = w(y)]$$
$$= f(w(y)).$$

under the transformation the probability is conserved.

Ex:- Let X be a binomial random variable with $X = \{0,1,2,3\}$ and the probability of a success in X as $(2/5)$. Find the probability distribution of the R.v. $Y = X^2$.

Soln:- Since X = no. of success in the given experiment and max. 3-success are possible

$\Rightarrow$ No. of trails, $n = 3$.

$$\Rightarrow f(x) = \begin{cases} {}^3C_x \left(\frac{2}{5}\right)^x \left(\frac{3}{5}\right)^{3-x} & ; x = 0,1,2,3. \\ 0 & ; \text{elsewhere.} \end{cases}$$

Transformation:-

$Y = X^2 \Rightarrow X = \sqrt{Y}$ (we have taken only +ve square root as $x \geq 0$.)

$X \rightarrow \{0,1,2,3\} \Rightarrow Y = \{0,1,4,9\}$

$$\Rightarrow g(y) = f(w(y)) = \begin{cases} {}^3C_{\sqrt{y}} \left(\frac{2}{5}\right)^{\sqrt{y}} \left(\frac{3}{5}\right)^{3-\sqrt{y}} & ; y = 0,1,4,9 \\ 0 & ; \text{elsewhere.} \end{cases}$$

For more than one random variable.

Suppose that X_1 and X_2 are discrete random variables with joint probability distribution $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ define one-to-one transformation between the points (x_1, x_2) and (y_1, y_2) so that the equations,

$$y_1 = u_1(x_1, x_2) \text{ and } y_2 = u_2(x_1, x_2)$$

may be uniquely solved for x_1 and x_2 in terms of y_1 and y_2, say $x_1 = w_1(y_1, y_2)$ and $x_2 = w_2(y_1, y_2)$. Then the joint probability distribution of Y_1 and Y_2 is

$$g(y_1, y_2) = f[w_1(y_1, y_2), w_2(y_1, y_2)].$$

\# Let X_1 and X_2 be discrete random variables with joint prob. distribution,

$$f(x_1, x_2) = \begin{cases} \frac{x_1 x_2}{18} ; & x_1 = 1, 2 ; \; x_2 = 1, 2, 3 \\ 0 ; & \text{elsewhere.} \end{cases}$$

Find the probability distribution of the random variable $Y = X_1 X_2$.

Soln → We shall define the second r.v. in the transformed co-ordinate system to solve for x_1 & x_2 uniquely, as $Y' = X_2$.

→ $x_2 = \{1, 2, 3\} \Rightarrow \cancel{\alpha} \cdot y' = \{1, 2, 3\}$

$y = \{1, 2, 3, 4, 6\}$ as $y = x_1 x_2$ we take all possible combinations of (x_1, x_2)

→ Solving for (x_1, x_2) in terms of $\cancel{(y_1, y_2)}$ (y, y').

$\boxed{x_2 = y'} = w_2(y, y')$ and $x_1 = \frac{y}{x_2} = \frac{y}{y'} = w_1(y, y')$

$$g(y, y') = f(w_1(y, y'), w_2(y, y'))$$

$$= \frac{1}{18} \frac{y}{y'} \times y' = \frac{y}{18}$$

$$\Rightarrow g(y, y') = \begin{cases} \frac{y}{18} ; & y = 1, 2, 3, 4, 6 ; \; y' = 1, 2, 3 \\ 0 ; & \text{elsewhere.} \end{cases}$$

→ For calculating distribution of y we sum $g(y, y')$ over all the values of y'

$$g_1(y) = \sum_{y'} g(y, y') = \sum_{y'} \frac{y}{18}$$

$$= \frac{y}{18}\left[1+2+3\right]$$

$$= \frac{y}{18} \times 6 = \frac{y}{3}$$

#. Let X_1 and X_2 be two independent random variables having Poisson distributions with parameters μ_1 and μ_2, respectively. Find the distribution of the random variably ~~of~~ $Y_1 = X_1 + X_2$.

Soln. → Since X_1 and X_2 are independent, we can write

$$f(x_1, x_2) = f(x_1) f(x_2) = \frac{e^{-\mu_1} \mu_1^{x_1}}{x_1!} \frac{e^{-\mu_2} \mu_2^{x_2}}{x_2!}$$

Where $x_1 = 0, 1, 2, ---$, $x_2 = 0, 1, 2, ---$. ($\because$ X_1 & X_2 are Poisson R.v.)

→ Let us define the second R.v. in new ~~sp~~ coordinates as $Y_2 = X_2$.

→ Inverse functions are given by

$$x_1 = \underbrace{y_1 - y_2}_{= w_1(y_1, y_2)} ; \qquad x_2 = y_2 = w_2(y_1, y_2)$$

Since $x_1 \geq 0 \Rightarrow y_1 - y_2 \geq 0 \Rightarrow y_1 \geq y_2$.

$$\Rightarrow y_1 = 0, 1, 2, ---- ; \quad y_2 = 0, 1, 2, -, y_1$$

$$g(y_1, y_2) = f\left(w_1(y_1, y_2),\ w_2(y_1, y_2)\right)$$

$$= \frac{e^{-\mu_1} \mu_1^{y_1 - y_2}}{(y_1 - y_2)!} \frac{e^{-\mu_2} \mu_2^{y_2}}{y_2!}$$

$$= \frac{e^{-(\mu_1 + \mu_2)} \mu_1^{y_1 - y_2} \mu_2^{y_2}}{(y_1 - y_2)!\ y_2!}$$

→ The marginal prob. distribution of y_1 is

$$h(y_1) = \sum_{y_2=0}^{y_1} g(y_1, y_2) = e^{-(\mu_1+\mu_2)} \sum_{y_2=0}^{y_1} \frac{\mu_1^{y_1-y_2} \mu_2^{y_2}}{(y_1 - y_2)!\, y_2!}$$

$$= \frac{e^{-(\mu_1+\mu_2)}}{y_1!} \underbrace{\sum_{y_2=0}^{y_1} \frac{y_1!}{(y_1-y_2)!\, y_2!} \mu_1^{(y_1-y_2)} \mu_2^{y_2}}$$

This sum is the binomial expansion of $(\mu_1+\mu_2)^{y_1}$

$$\Rightarrow \quad h(y_1) = \frac{e^{-(\mu_1+\mu_2)}}{y_1!} (\mu_1+\mu_2)^{y_1}$$

= Poisson's distribution for y_1 with mean $\mu_1 + \mu_2$.

Transformation of Continous random variables:-

Thm:- Suppose that X is a continous random variable with probability distribution $f(x)$. Let $Y = u(X)$ define a one-to-one correspondence b/w the values of X and Y so that the eqn. $y = u(x)$ can be uniquely solved for x in terms of y, say $x = w(y)$. Then the probability distribution of Y

is:- $g(y) = f[w(y)]\,|J|$,

where $J = w'(y)$ ~~and~~ $= \frac{d}{dy}(w(y))$ and is called the Jacobian of the transformation.

The speed of a molecule in a uniform gas at equilibrium is a random variable V whose probability distribution is given by

$$f(v) = \begin{cases} k v^2 \exp(-b v^2) & ; v>0 \\ 0 & ; \text{elsewhere.} \end{cases}$$

where k & b are ~~is~~ ~~an~~ appropriate constants. Find the prob. distribution of the kinetic energy of the molecule W, where $W = mV^2/2$.

Soln. → $W = \frac{mV^2}{2} \Rightarrow V^2 = \frac{2W}{m}$

$\Rightarrow w = u(v)$

$\Rightarrow v = \sqrt{\frac{2w}{m}}$ = ~~tot~~ $u_1(w)$.

→ since $v > 0 \Rightarrow w > 0$.

→ $J = \frac{d}{dw} u_1(w) = \sqrt{\frac{2}{m}} \times \frac{1}{2} w^{-1/2} = \sqrt{\frac{1}{2m}}\, w^{-1/2}$

$g(w) = f(u_1(w)) = k\left(\frac{2w}{m}\right) \exp\left[-b \times \frac{2w}{m}\right]$

$$g(w) = \begin{cases} \frac{2k}{m} w \exp\left[-\frac{2b}{m} w\right] & ; w>0 \\ 0 & ; \text{elsewhere.} \end{cases}$$

Scanned with CamScanner

~~Thm~~

\# Suppose that X_1 and X_2 are continuous random variables with joint probability distribution $f(x_1, x_2)$. Let $Y_1 = u_1(X_1, X_2)$ and $Y_2 = u_2(X_1, X_2)$ define a one-to-one transformation b/w the points (x_1, x_2) and (y_1, y_2) so that the equations $y_1 = u_1(x_1, x_2)$ and $y_2 = u_2(x_1, x_2)$ may be uniquely solved for x_1 and x_2 in terms of y_1 and y_2, say $x_1 = w_1(y_1, y_2)$ and $x_2 = w_2(y_1, y_2)$. Then the joint probability distribution of Y_1 and Y_2 is

$$g(y_1, y_2) = f[w_1(y_1, y_2), w_2(y_1, y_2)] \, |J|,$$

where the Jacobian is the 2×2 determinant,

$$J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$$

and $\frac{\partial x_1}{\partial y_1}$ is the partial derivative of x_1 w.r.t. y_1 and other partial derivatives are defined in a similar manner.

\# Let X_1 and X_2 be two continuous random variables with joint probability distribution,

$$f(x_1, x_2) = \begin{cases} 4x_1 x_2 & ; \ 0 < x_1 < 1, \ 0 < x_2 < 1 \\ 0 & ; \ \text{elsewhere.} \end{cases}$$

Find the joint probability distribution of $Y_1 = X_1^2$ and $Y_2 = X_1 X_2$.

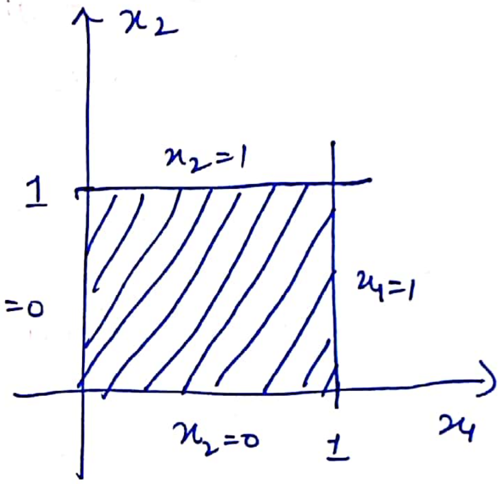
Soln:- The inverse soln. of $y_1 = x_1^2$ and $y_2 = \cancel{x_1 x_2}\ x_1 x_2$ are $x_1 = \sqrt{y_1}$, $x_2 = y_2/\sqrt{y_1}$, and the Jacobian is

$$J = \begin{vmatrix} \frac{1}{2\sqrt{y_1}} & 0 \\ \frac{-y_2}{2y_1^{3/2}} & \frac{1}{\sqrt{y_1}} \end{vmatrix} = \frac{1}{2y_1}$$

$\rightarrow$ Next we find the domain of y_1 and y_2.

- $0 < x_1 < 1$ and $0 < x_2 < 1$

  $\Rightarrow$ The region in the (x_1, x_2) is the interior of square bounded by $x_1 = 0$, $x_2 = 0$, $x_1 = 1$, $x_2 = 1$



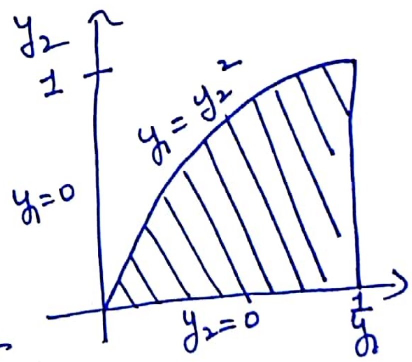
- These boundaries are transformed to

$x_1 = 0 \Rightarrow y_1 = 0$

$x_2 = 0 \Rightarrow y_2 = 0.$

$x_1 = 1 \Rightarrow y_1 = 1$

$x_2 = 1 \Rightarrow \frac{y_2}{\sqrt{y_1}} = 1 \Rightarrow y_1 = y_2^2$



$0 < x_1 < 1 \Rightarrow 0 < \sqrt{y_1} < 1 \Rightarrow \boxed{0 < y_1 < 1}$

$0 < x_2 < 1 \Rightarrow 0 < \frac{y_2}{\sqrt{y_1}} < 1 \Rightarrow \boxed{y_2 > 0}$ and

$y_2 < \sqrt{y_1}$ or $\boxed{y_1 > y_2^2}$

$$g(y_1,y_2) = f\left(w_1(y_1,y_2),\, w_2(y_1,y_2)\right)|J|$$

$$= \cancel{4}^{\,2}\,\cancel{\sqrt{y_1}}\,\frac{y_2}{\cancel{\sqrt{y_1}}} \times \frac{1}{\cancel{2}\,y_1} = \frac{2y_2}{y_1}$$

$$\Rightarrow\ g(y_1,y_2) = \begin{cases} \frac{2y_2}{y_1} & ;\ y_2^2 < y_1 < 1,\ 0 < y_2 < 1 \\ 0 & ;\ \text{elsewhere.} \end{cases}$$

\# Rather than the above transformation if we have $Y_1 = X_1^2$ and $Y_2 = X_1 + X_2$. and $0 < x_1 < 1$ and $0 < x_2 < 1$, let us find the transformations and the region for (y_1, y_2)

$$x_1 = \sqrt{y_1} \quad \text{and} \quad x_2 = y_2 - x_1 = y_2 - \sqrt{y_1}$$

→ boundaries transform to :-

$x_1 = 0 \Rightarrow \boxed{y_1 = 0}$

$x_2 = 0 \Rightarrow y_2 = \sqrt{y_1} \Rightarrow \boxed{y_1 = y_2^2}$

$x_1 = 1 \Rightarrow \boxed{y_1 = 1}$

$x_2 = 1 \Rightarrow y_2 - \sqrt{y_1} = 1 \Rightarrow \boxed{y_1 = (y_2 - 1)^2}$

→ Since $x_1 + x_2 > 0 \Rightarrow y_1 + y_2 > 0$.

→ $0 < x_1 < 1 \Rightarrow 0 < y_1 < 1$

→ $0 < x_2 < 1 \Rightarrow 0 < y_2 - \sqrt{y_1} < 1$

$\Rightarrow y_2 - \sqrt{y_1} > 0 \Rightarrow \boxed{y_1 < y_2^2}$

$y_2 < 1 + \sqrt{y_1}$

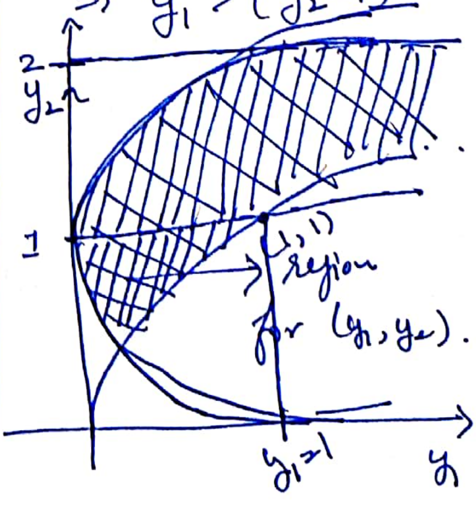
$0 < y_1 < 1$

$\Rightarrow y_2 < 2$

$y_2 - \sqrt{y_1} < 1$

$\Rightarrow \sqrt{y_1} > (y_2 - 1)$

$\Rightarrow y_1 > (y_2 - 1)^2.$

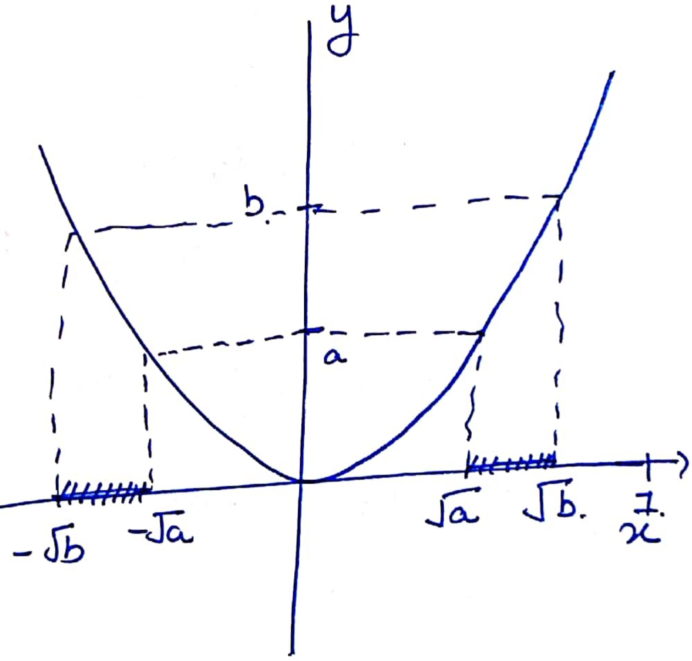


What happens when the transformation $Y = u(X)$; X is ~~is~~ a continous variable, is not a one-to-one transformation?

- Consider $f(x)$ be the distribution fn. of ~~X~~ defined over $-1 < x < 1$ and a transformation $Y = X^2$, then the inverse transformation is

$$x = -\sqrt{y} \quad ; \quad -1 < x < 0$$

$$x = \sqrt{y} \quad ; \quad 0 < x < 1$$



$$P[a < Y < b] = P[-\sqrt{b} < X < -\sqrt{a}] + P[\sqrt{a} < X < \sqrt{b}]$$

$$= \int_{-\sqrt{b}}^{-\sqrt{a}} f(x)\,dx + \int_{\sqrt{a}}^{\sqrt{b}} f(x)\,dx.$$

→ Changing variable in each integral

$$= \int_{b}^{a} f(-\sqrt{y})\, J_1\, dy$$

$$+ \int_{a}^{b} f(\sqrt{y})\, J_2\, dy$$

$$= -\int_{a}^{b} f(-\sqrt{y})\, J_1\, dy + \int_{a}^{b} f(\sqrt{y})\, J_2\, dy.$$

$$J_1 = \frac{d}{dy}(-\sqrt{y_0}) = \frac{-1}{2\sqrt{y}} = -|J_1|$$

$$J_2 = \frac{d}{dy}(\sqrt{y}) = \frac{1}{2\sqrt{y}} = |J_2|$$

$$\Rightarrow P[a < Y < b] = \int_a^b \left(f(-\sqrt{y})\,|J_1| + f(\sqrt{y})\,|J_2| \right) dy$$

If $g(y)$ is the distribution function of random variable y, we get

$$\Rightarrow P[a < Y < b] = \int_a^b g(y)\,dy.$$

→ Comparing both we get

$$g(y) = f(-\sqrt{y})\,|J_1| + f(\sqrt{y})\,|J_2|$$

$$\Rightarrow \quad g(y) = \begin{cases} \dfrac{f(-\sqrt{y}) + f(\sqrt{y})}{2\sqrt{y}} & ; 0 < y < 1 \\ 0 & ; \text{elsewhere.} \end{cases}$$

#. X is a continous random variable with prob. distribution $f(x)$. Let $Y = u(X)$ define a transformation b/w the values of X and Y that is not one-to-one. If the interval over which X is defined can be partitioned in k mutually disjoint sets such that each of the inverse functions

$x_1 = w_1(y)$, $x_2 = w_2(y)$, ------, $x_k = w_k(y)$

If $y = u(x)$ defines a one-to-one correspondence, ~~then~~ then the probability distribution of Y is

$$g(y) = \sum_{i=1}^{k} f[w_i(y)]\,|J_i|,$$

where J_i is $w_i'(y)$; $i = 1, 2, ---, k$.

Moments and Moment-Generating functions:-

→ The r^{th} moment about any point 'a' of the random variable X with distribution function $f(x)$ is given by,

$$E[(X-a)^r] = \begin{cases} \sum_x (x-a)^r f(x) \;; & \text{if } X \text{ is discrete,} \\ \int_{-\infty}^{\infty} (x-a)^r f(x)\,dx \;; & \text{if } X \text{ is continous.} \end{cases}$$

→ If $a =$ origin, then the r^{th} moment about the origin of the random variable X is given by.

$$\mu_r' = E(X^r) = \begin{cases} \sum_x x^r f(x) \;; & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x^r f(x)\,dx \;; & \text{if } X \text{ is continous.} \end{cases}$$

→ The moment-generating function of the random variable X is given by $E(e^{tX})$ and is denoted by $M_X(t)$. Hence,

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_x e^{tx} f(x), & \text{if } X \text{ is discrete,} \\ \int_{-\infty}^{\infty} e^{tx} f(x)\,dx, & \text{if } X \text{ is continous.} \end{cases}$$

→ Let X be a random variable with moment-generating function $M_X(t)$. Then

$$\left.\frac{d^r M_X(t)}{dt^r}\right|_{t=0} = \mu_r'.$$

Ex:- Find the moment-generating function of the binomial random variable X and then use it to verify that $\mu = np$ and $\sigma^2 = npq$.

Soln:-

$$f(x) = \cancel{\binom{n}{x}}\ {}^nC_x\, p^x q^{n-x} \quad ; \quad \cancel{x \geq 0}\ \cancel{x=0,1,2,\ldots} \quad x = 0, 1, 2, \ldots n$$

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

$$= \underbrace{\sum_{x=0}^{n} \binom{n}{x} (pe^t)^x q^{n-x}}_{\text{binomial exp}}$$

$$\boxed{M_X(t) = (pe^t + q)^n}$$

$$\mu_1' = \left.\frac{d}{dt}(M_X(t))\right|_{t=0} = \left.\frac{d}{dt}(pe^t+q)^n\right|_{t=0}.$$

$$= \left. n(pe^t+q)^{n-1} pe^t \right|_{t=0} = n(p+q)^{n-1} p$$

$$\Rightarrow \mu_1' = np \quad (\because p+q=1).$$

$$\mu_2' = \left.\frac{d^2}{dt^2}(M_X(t))\right|_{t=0} = np\left[e^t(n-1)(pe^t+q)^{n-2} pe^t + (pe^t+q)^{n-1} e^t\right] \text{ at } t=0$$

$$= np\left[p(n-1)(p+q)^{n-2} + 1\right]$$

$$\Rightarrow \mu_2' = np[(n-1)p+1]$$

hence, mean $= E(x) = \mu_1' = np$

Variance $= E(x^2) - (E(x))^2$

$$= \mu_2' - \mu^2 = \cancel{n}\,np[np+1-p] - n^2p^2$$

$$= np[\cancel{np}+1-p-\cancel{np}]$$

$$= np[1-p] = npq.$$

→ • <u>Uniqueness Theorem</u>:-

Let X and Y be two random variables with moment generating functions $M_x(t)$ and $M_y(t)$, respectively. If $M_x(t) = M_y(t)$ for all values of t, then X and Y have the same probability distribution.

~~This result is useful~~ in

• ① $M_{x+a}(t) = e^{at}M_x(t)$.

<u>Proof</u>:- $M_{x+a}(t) = E[e^{t(x+a)}] = E[e^{ta}e^{tx}]$

$$= e^{at}E[e^{tx}] = e^{at}M_x(t)$$

• ② If $X_1, X_2, --, X_n$ are independent random variables with moment-generating functions $M_{x_1}(t), ---, M_{x_n}(t)$, respectively, and $Y = X_1 + X_2 + — + X_n$, then

$$M_y(t) = M_{x_1}(t) --- M_{x_n}(t)$$

☆ These results are useful in finding the distribution of linear combinations of independent random variables.

Ex:- If the moment-generating function of poissons' random variable (X) with mean μ is $M_X(t) = e^{\mu_1(e^t-1)}$, then find the distribution of the random variable $Y = X_1 + X_2$ where X_1 & X_2 are independent Poisson's random variables with mean μ_1 & μ_2.

Soln.

$$\left.\begin{aligned} M_{X_1}(t) &= e^{\mu_1(e^t-1)} \\ M_{X_2}(t) &= e^{\mu_2(e^t-1)} \end{aligned}\right\} \rightarrow \text{using the given statement.}$$

~~$M_Y(X_1+$~~

$$M_Y(t) = M_{X_1}(t) M_{X_2}(t) \quad (\text{using the result } 2)$$

$$= e^{\mu_1(e^t-1)} \cdot e^{\mu_2(e^t-1)}$$

$$= e^{(\mu_1+\mu_2)(e^t-1)}.$$

R.H.S. is the moment-generating function of a random variable having Poisson's distribution with the parameters $\mu_1+\mu_2$. Using uniqueness thm. Y is the Poisson's random variable with mean $\mu_1+\mu_2$.