# Unit - II (Two Dimensional Random Variables)

* joint distributions
* Marginal and Conditional distributions
* Covariance
* Correlation and Regression
* Transformation of random variables
* Central Limit Theorem (for iid random variables)

**Definition:**

Let S be a sample space associated with a random experiment E. Let X and Y be two random variables defined on S. Then the pair (X, Y) is called a two - dimensional r.v.

**Two Dimensional Discrete r.v**

**Definition:** If the set of all values of (X, Y) is finite (or) countably infinite, then (X, Y) is called a two-dimensional discrete random variable.

**Joint Probability Mass function of discrete random variable**

**Definition:**

Let (X, Y) is a two dimensional discrete r.v. For each possible outcome $(x_i, y_j)$ we associate a number $p(x_i, y_j)$ satisfying the following conditions,

(i) $p(x_i, y_j) \geq 0$, $\forall (x_i, y_j)$

(ii) $\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} p(x_i, y_j) = 1$.

This fn. P is called joint prob. fn. (or) joint prob. mass fn.

## Marginal Probability Function:

Let $(X,Y)$ be a two-dimensional discrete r.v. The marginal probability fn of $X$ is defined by $P_X(x_i) = P(X = x_i) = \sum_{j=1}^{\infty} P(x_i, y_j), i = 1,2,\ldots$

$$= P_{i1} + P_{i2} + P_{i3} + \cdots$$

$$= P_{i.}$$

The set $\{x_i, P_{i.}\}$ is called the marginal probability fn of $X$.

The marginal probability fn of $Y$ is given by $P_Y(y_j) = P(Y = y_j) = \sum_{i=1}^{\infty} p(x_i, y_j), j = 1,2,\ldots$

$$= P_{1j} + P_{2j} + \cdots = P_{.j}$$

The set $\{y_j, P_{.j}\}$ is called the marginal probl fn of $Y$.

## Conditional Probability Function:

Let $(X,Y)$ be a two dimensional discrete r.v, then the conditional probl fn of $X$ given $Y = y_j$ is defined by

$$P(X = x_i | Y = y_j) = \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)} = \frac{P_{ij}}{P_{.j}}$$

Similarly, The conditional probl fn of $Y$ given $X = x_i$ is defined by $P(Y = y_j | X = x_i) = \frac{P_{ij}}{P_{i.}}$

**Note:** If $(X,Y)$ is a two dimensional discrete r.v such that $\boxed{P_{ij} = P_{i.} \times P_{.j}}$, for all $i, j$, then $X$ and $Y$ are said to be independent r.v's.

Problem:

(1) The joint probability mass fn. of X and Y is given by
$P(1,1) = \frac{1}{8}$, $P(1,2) = \frac{1}{4}$, $P(2,1) = 1/8$, $P(2,2) = 1/2$.

(i) Compute the conditional mass fn. of X given Y=2.

(ii) Are X and Y are independent?

(iii) Compute $P(X+Y>2)$

Soln: The joint mass fn. can be shown in the foll. table.

| Y \ X | 1 | 2 | P(y) |
|---|---|---|---|
| 1 | $\frac{1}{8}$ | $\frac{1}{8}$ | $\frac{2}{8}$ |
| 2 | $\frac{1}{4}$ | $\frac{1}{2}$ | $3/4$ |
| P(x) | $3/8$ | $5/8$ | 1 |

∴ The marginal probl. fn. of X is,

| X | 1 | 2 |
|---|---|---|
| P(x) | $3/8$ | $5/8$ |

∴ The marginal probl. fn. of Y is,

| Y | 1 | 2 |
|---|---|---|
| P(y) | $2/8$ | $3/4$ |

(i) The conditional probability fn. of X given Y=2 is,

$$P(X=1 \mid Y=2) = \frac{P(X=1, Y=2)}{P(Y=2)} = \frac{P(1,2)}{P(Y=2)} = \frac{1/4}{3/4}$$

$$= \frac{1}{4} \times \frac{4}{3} = \boxed{\frac{1}{3}}$$

$$P(X=2 \mid Y=2) = \frac{P(X=2, Y=2)}{P(Y=2)} = \frac{P(2,2)}{P(Y=2)} = \frac{1/2}{3/4} = \boxed{2/3}$$

(ii) WKT, X and Y are independent if

$$p(x,y) = p_X(x) \cdot p_Y(y), \ \forall\ x,y.$$

Now, $P_X(1)\, P_Y(1) = \frac{3}{8} \cdot \frac{2}{8} = \frac{3}{32}$

But $P(1,1) = \frac{1}{8} = \frac{4}{32}$

$\therefore P(1,1) \neq P_X(1) \cdot P_Y(1)$.

So X and Y are not independent.

(iii) $P(X+Y > 2) = 1 - P(X+Y \leq 2)$

$= 1 - [P(1,1)] = 1 - \frac{1}{8} = \boxed{7/8}$

## Two Dimensional Continuous random Variables:

**Definition:** If the values of $(X,Y)$ is an uncountably infinite set, (ie) the range space is a region in the xy-plane, then $(X,Y)$ is called two dimensional continuous r.v.

**Joint Probability Density Function of two Dimensional Continuous R.V $(X,Y)$**

If $(X,Y)$ be a two dimensional continuous r.v Then $f(x,y)$ is called the joint p.d.f of $(X,Y)$ provided the foll. conditions are satisfied

(i) $f(x,y) \geq 0, \ \forall\ (x,y)$

(ii) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1$.

**Note:** $P(a_1 \leq X \leq b_1, a_2 \leq Y \leq b_2) = \int_{a_1}^{b_1}\int_{a_2}^{b_2} f(x,y)\, dy\, dx$

## Marginal Distribution Function:

Let $(X,Y)$ be a two dimensional continuous r.v

The Marginal distribution fn of $X$ is given by

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

The Marginal distribution fn of $Y$ is given by

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

**Note:** If $(X,Y)$ is a two dimensional continuous r.v with joint p.d.f $f(x,y)$ such that

$$\boxed{f(x,y) = f_x(x) \cdot f_y(y)}$$

then $X$ and $Y$ are said to be independant r.v.

## Conditional Probability Distribution:

Let $(X,Y)$ be the two dimensional continuous r.v's with joint p.d.f $f(x,y)$. Then the conditional p.d.f of $X$ given by $Y$ is

$$\boxed{f(x|y) = \frac{f(x,y)}{f_y(y)}}$$

Where $f_y(y)$ is the marginal p.d.f of $Y$

The conditional p.d.f of $Y$ given by $X$ is

$$\boxed{f(y|x) = \frac{f(x,y)}{f_x(x)}}$$

Where $f_x(x)$ is the marginal p.d.f of $X$.

Problems:

① If $X$ and $Y$ are random variables having the joint density fn.

$$f(x,y) = \begin{cases} \frac{1}{8}(6-x-y), & 0<x<2,\ 2<y<4 \\ 0, & OW \end{cases}$$

Find (i) $P(X+Y<3)$ (ii) $P(X<1, Y<3)$

(iii) $P(X<1 \mid Y<3)$.

Soln: Given $f(x,y) = \begin{cases} \frac{1}{8}(6-x-y), & 0<x<2,\ 2<y<4 \\ 0, & \text{elsewhere} \end{cases}$

The marginal probl. fn. of $X$ is given by

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

$$= \frac{1}{8}\int_2^4 (6-x-y)\, dy = \frac{1}{8}\left[(6-x)y - \frac{y^2}{2}\right]_2^4$$

$$= \frac{1}{8}\left[(6-x)4 - \frac{4\times 4}{2} - (6-x)2 + 2\right]$$

$$= \frac{1}{8}\left[2(6-x) - 6\right]$$

$$= \frac{6-x-3}{4} = \frac{3-x}{4},$$

$$\therefore f_X(x) = \begin{cases} \frac{3-x}{4}, & 0<x<2 \\ 0, & OW \end{cases}$$

The marginal probl. fn. of $Y$ is given by,

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

$$= \frac{1}{8}\int_0^2 (6-x-y)\, dx$$

$$= \frac{1}{8}\left[(6-y)x - \frac{x^2}{2}\right]_0^2 = \frac{1}{8}\left[(6-y)2 - 2\right]$$

$$= \frac{5-y}{4}$$

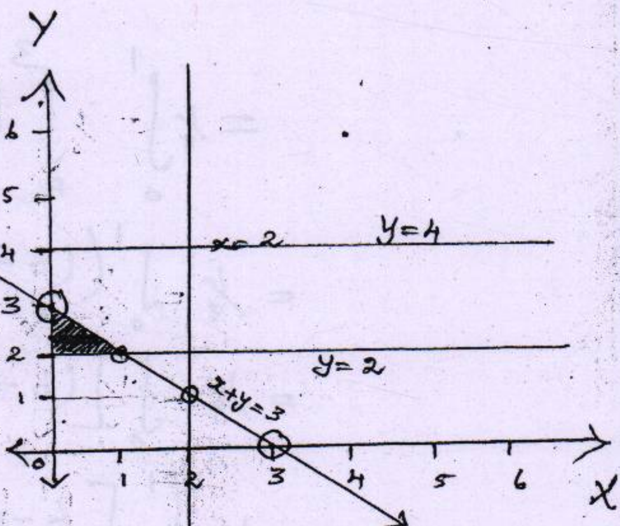
$$\therefore f_Y(y) = \begin{cases} \dfrac{5-y}{4}, & 2 < y < 4 \\ 0, & \text{OW} \end{cases}$$

(i) $P(X+Y<3) = \iint f(x,y)\, dx\, dy$

Here, The limits of $x$ & $y$ are

$x \to 0$ to $3-y$

$y \to 2$ to $3$




$x+y=3$

| | | | | |
|---|---|---|---|---|
| $x$: | 0 | 3 | 1 | 2 |
| $y$: | 3 | 0 | 2 | 1 |

$$\therefore P(X+Y<3) = \frac{1}{8}\int_2^3 \int_0^{3-y} (6-x-y)\, dx\, dy$$

$$= \frac{1}{8}\int_2^3 \left[(6-y)x - \frac{x^2}{2}\right]_0^{3-y} dy$$

$y = 3-x$

$x = 2$ to $2$

$$= \frac{1}{8}\int_2^3 \left[(6-y)(3-y) - \frac{(3-y)^2}{2}\right] dy$$

$$= \frac{1}{8}\int_2^3 (3-y)\left[6-y-\frac{(3-y)}{2}\right] dy$$

$$= \frac{1}{8}\int_2^3 (3-y)\left[\frac{12-2y-3+y}{2}\right] dy$$

$$= \frac{1}{8}\int_2^3 (3-y)\left[\frac{9-y}{2}\right] dy$$

$$= \frac{1}{16}\int_2^3 (27-3y-9y+y^2)\, dy$$

$$= \frac{1}{16}\int_2^3 (27-12y+y^2)\, dy$$

$$= \frac{1}{16}\left[27y - \frac{12y^2}{2} + \frac{y^3}{3}\right]_2^3$$

$$= \frac{1}{16}\left[\frac{19}{3} - 30 + 27\right] = \boxed{\frac{5}{24}}$$

(ii) $P(X<1, Y<3) = \int_{-\infty}^{1}\int_{-\infty}^{3} f(x,y)\, dy\, dx$

$= \frac{1}{8}\int_{0}^{1}\int_{2}^{3} (6-x-y)\, dy\, dx$

$= \frac{1}{8}\int_{0}^{1} \left((6-x)y - \frac{y^2}{2}\right)_{2}^{3} dx$

$= \frac{1}{8}\int_{0}^{1} \left[(6-x)3 - \frac{9}{2} - (6-x)2 + 2\right] dx$

$= \frac{1}{8}\int_{0}^{1} \left[6-x - \frac{5}{2}\right] dx$ $\quad \frac{12-5}{2}$

$= \frac{1}{8}\int_{0}^{1} \left(\frac{7}{2} - x\right) dx = \frac{1}{8}\left[\frac{7}{2}x - \frac{x^2}{2}\right]_{0}^{1}$

$= \frac{1}{8}\left[\frac{7}{2} - \frac{1}{2}\right] = \frac{1}{8} \times \cancel{\frac{6}{2}}^{3} = \boxed{\frac{3}{8}}$

(iii) $P(X<1 \mid Y<3) = \frac{P(X<1, Y<3)}{P(Y<3)}$ —— ①

Now,

$P(Y<3) = \int_{-\infty}^{3} f(y)\, dy$

$= \frac{1}{4}\int_{2}^{3} (5-y)\, dy = \frac{1}{4}\left(5y - \frac{y^2}{2}\right)_{2}^{3}$

$= \frac{1}{4}\left[15 - \frac{9}{2} - 10 + \frac{4}{2}\right]$

$= \frac{1}{4}\left[5 - \frac{5}{2}\right] = \frac{1}{4}\left(\frac{10-5}{2}\right) = \frac{1}{4} \times \frac{5}{2}$

$= \boxed{\frac{5}{8}}$

From ①,

$\therefore P(X<1 \mid Y<3) = \frac{3/8}{5/8} = \frac{3}{8} \times \frac{8}{5} = \boxed{\frac{3}{5}}$

8

(2) Given the joint p.d.f. of $(X,Y)$ as

$$f(x,y) = \begin{cases} 8xy, & 0 < x < y < 1 \\ 0, & \text{Otherwise} \end{cases}$$

Find the marginal and conditional probability fnst of X and Y. Are X and Y are independent?

Soln:

Given, $f(x,y) = \begin{cases} 8xy, & 0 < x < y < 1 \\ 0, & \text{OW} \end{cases}$

The Marginal p.d.f of X is

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

$$= \int_x^1 8xy\, dy = 8\cancel{x}\left(x\frac{y^2}{2}\right)_x^1$$

$$= 8x\left[\frac{1}{2} - \frac{x^2}{2}\right]$$

$$= 4x(1-x^2)$$

$$\therefore f_X(x) = \begin{cases} 4x(1-x^2), & 0 < x < 1 \\ 0, & \text{OW} \end{cases}$$

The Marginal p.d.f. of Y is

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

$$= 8\int_0^y xy\, dx = 8y\left(\frac{x^2}{2}\right)_0^y$$

$$= \frac{\overset{4}{\cancel{8}} y^3}{\cancel{2}} = 4y^3.$$

$$\therefore f_Y(y) = \begin{cases} 4y^3, & 0 < y < 1 \\ 0, & \text{OW} \end{cases}$$

The conditional p.d.f of $X$ given $Y=y$ is,

$$f(x \mid Y=y) = \frac{f(x,y)}{f_Y(y)}$$

$$= \frac{8xy}{4y^3} = \frac{2x}{y^2}$$

$$\therefore f(x \mid y) = \begin{cases} \frac{2x}{y^2}, & 0 < x < y, \ 0 < y < 1 \\ 0, & \text{OW} \end{cases}$$

The conditional p.d.f of $Y$ given $X=x$ is,

$$f(y \mid X=x) = \frac{f(x,y)}{f_X(x)} = \frac{8xy}{4x(1-x^2)}$$

$$= \frac{2y}{1-x^2}$$

$$f(y \mid x) = \begin{cases} \frac{2y}{1-x^2}, & x < y < 1, \ 0 < x < 1 \\ 0, & \text{OW} \end{cases}$$

Now, $f_X(x) \cdot f_Y(y) = 4x(1-x^2) \cdot 4y^3$

$$= 16y^3 x (1-x^2)$$

$$\neq f(x,y)$$

$\therefore$ $X$ and $Y$ are not independent

## Covariance :

Let $X$ and $Y$ be two random variables defined on the same sample space. The covariance of $X$ and $Y$ is denoted by $cov(X,Y)$ and is defined by $cov(X,Y) = E[XY] - E[X]E[Y]$.

**Note:**

(i) $Cov(X,X) = Var(X)$

(ii) If $X$ and $Y$ are independent, then

$Cov(X,Y) = 0$.

If $X$ and $Y$ are independent, then

$$E[XY] = E[X]E[Y]$$

$$\Rightarrow cov(X,Y) = 0.$$

(iii) $Cov(X,Y) = cov(Y,X)$

(iv) $Cov(aX, bY) = ab\,cov(X,Y)$, where $a$ & $b$ are constants.

(v) $Cov(X+a, Y+b) = cov(X,Y)$.

(vi) $Cov(aX+b, cY+d) = ac\,cov(X,Y)$

(vii) $Cov(X+Y, Z) = cov(X,Z) + cov(Y,Z)$

## Relation Between Variance and Covariance :

(i) $Var(X+Y) = Var(X) + Var(Y) + 2\,cov(X,Y)$

(ii) $Var(X-Y) = var(X) + Var(Y) - 2\,cov(X,Y)$

(iii) If $X$ and $Y$ are independent, then

$Cov(X,Y) = 0$

$\therefore Var(X \pm Y) = Var(X) + Var(Y)$.

$Var(ax) = a^2 Var\,x$

## Correlation:

Two variables X and Y are said to be correlated if a change in the value of one of the variables causes a change in the value of the other variable.

Though, covariance is a measure of linear relationship between two variates X and Y, it does not tell us the degree of relationship (or) strength of the relationship.

The correlation co-efficient is a measure which gives the degree of relationship between X and Y.

**Definition:**

Let X and Y are two r.v. The co-efficient of correlation between X and Y is denoted by $\rho_{xy}$ (or) $r_{xy}$ and is defined by

$$\rho_{xy} = \frac{\operatorname{cov}(X,Y)}{\sigma_X \cdot \sigma_Y} \quad \text{if } \sigma_X \neq 0,\ \sigma_Y \neq 0.$$

**Note:**

(1) This is called Karl pearson's correlation co-efficient.

(2) If X and Y are independent, then $\operatorname{cov}(X,Y)=0$ $\therefore \rho_{XY}=0$

But the converse is not true (ie) if $\rho_{XY}=0$ We cannot say X and Y are independent.

(3) If $\operatorname{cov}(X,Y)=0 \Rightarrow \rho_{XY}=0 \Rightarrow$ X and Y are uncorrelated.

(4) Correlation co-efficient always lies between $-1$ and $1$ $-1 \le \rho_{xy} \le 1$

(5) Correlation co-efficient is unaffected by change of origin and scale.

Note:

(1) If $\rho_{xy} = 1$, then there is a perfect positive correlation.

(2) If $\rho_{xy} = -1$, then there is a perfect negative correlation.

(3) If $\rho_{xy} = 0$, then the variables are uncorrelated.

Problems:

(1) The joint p.m.f of X and Y is given below:

| y \ x | -1 | 1 |
|---|---|---|
| 0 | 1/8 | 3/8 |
| 1 | 2/8 | 2/8 |

Find the correlation co-efficient of (X,Y).

Soln:

| y \ x | -1 | 1 | P(y) |
|---|---|---|---|
| 0 | 1/8 | 3/8 | 4/8 |
| 1 | 2/8 | 2/8 | 4/8 |
| P(x) | 3/8 | 5/8 | 1 |

The marginal density fn of $X$ is

| $x$ : | $-1$ | $1$ |
|---|---|---|
| $p(x)$ : | $3/8$ | $5/8$ |

The marginal density fn of $Y$ is

| $y$ : | $0$ | $1$ |
|---|---|---|
| $p(y)$ : | $4/8 = 1/2$ | $4/8$ |

$$E[X] = \sum x\, p(x) = (-1) \times 3/8 + 1 \times 5/8 = \boxed{1/4}$$

$$E[X^2] = \sum x^2\, p(x) = (-1)^2 \times 3/8 + 1^2 \times 5/8 = \boxed{1}$$

$$E[Y] = \sum y\, p(y) = 0 \times \tfrac{1}{2} + 1 \times 1/2 = \boxed{1/2}$$

$$E[Y^2] = \sum y^2\, p(y) = 0^2 \times 1/2 + 1^2 \times 1/2 = \boxed{1/2}$$

$$E[XY] = \sum \sum xy\, p(x,y)$$

$$= (0) \times (-1) \times 1/8 + 0 \times 1 \times 3/8 + 1 \times (-1) \times 2/8 + 1 \times 1 \times 2/8 = 0 + 0 - 2/8 + 2/8 = \boxed{0}$$

$$\sigma_x^2 = E[X^2] - \{E[X]\}^2 = 1 - (1/4)^2 = \frac{15}{16}$$

$$\boxed{\sigma_x = \frac{\sqrt{15}}{4}}$$

$$\sigma_y^2 = E[Y^2] - \{E[Y]\}^2 = \tfrac{1}{2} - (1/2)^2 = 1/4$$

$$\boxed{\sigma_y = 1/2}$$

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

$$= 0 - (1/4)(1/2) = -1/8$$

$$\therefore r_{xy} = \frac{\mathrm{Cov}(X,Y)}{\sigma_x \sigma_y} = \frac{-1/8}{\left(\frac{\sqrt{15}}{4}\right)(1/2)} = \frac{-1}{\sqrt{15}} = \boxed{-0.258}$$

(2) Suppose that the two D.r.v's $(X, Y)$ has the joint p.d.f

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

Obtain the correlation co-efficient between X and Y

Soln: Given $f(x,y) = \begin{cases} x+y, & 0<x<1,\ 0<y<1 \\ 0, & OW \end{cases}$

The Marginal p.d.f of $X$ is given by,

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 (x+y)\,dy$$

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

$$\therefore f_X(x) = \begin{cases} x+\frac{1}{2}, & 0<x<1 \\ 0, & OW \end{cases}$$

The Marginal p.d.f of $Y$ is given by,

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

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

$$\therefore f_Y(y) = \begin{cases} y+\frac{1}{2}, & 0<y<1 \\ 0, & OW \end{cases}$$

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

$$= \boxed{\frac{7}{12}}$$

$$E[x^2] = \int_0^1 x^2 f(x)\,dx = \int_0^1 x^2 (x + 1/2)\,dx = \int_0^1 \left(x^3 + \frac{x^2}{2}\right) dx$$

$$= \left(\frac{x^4}{4} + \frac{x^3}{6}\right)_0^1 = \left(\frac{1}{4} + \frac{1}{6}\right) = \frac{3+2}{12} = \boxed{5/12}$$

$$E[Y] = \int_0^1 y f(y)\,dy = \int_0^1 y\,(y + 1/2)\,dy = \boxed{7/12}$$

$$E[Y^2] = \int_0^1 y^2 f(y)\,dy = \int_0^1 y^2 (y + 1/2)\,dy = \boxed{5/12}$$

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

$$= \int_0^1 \int_0^1 xy\,(x+y)\,dx\,dy$$

$$= \int_0^1 y\,(x^2 + xy)\,dx\,dy$$

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

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

$$= \frac{2}{6} = \boxed{1/3}$$

$$\sigma_x^2 = E[x^2] - [E[x]]^2 = 5/12 - \left(\frac{7}{12}\right)^2 = \frac{11}{144}$$

$$\boxed{\sigma_x = \frac{\sqrt{11}}{12}}$$

$$\sigma_y^2 = E[Y^2] - [E[Y]]^2 = 5/12 - (7/12)^2 = \frac{11}{144}$$

$$\boxed{\sigma_Y = \frac{\sqrt{4}}{12}}$$

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

$$= 1/3 - (7/12 \times 7/12) = \frac{-1}{144}$$

$$\therefore r_{xy} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{-1/144}{\frac{\sqrt{4}}{12} \times \frac{\sqrt{4}}{12}} = \frac{-1}{11} = \boxed{-0.0909}$$

(3) If $X, Y, Z$ are uncorrelated r.v's with zero means and standard deviations, 5, 12, 9 respectively and $U = X+Y$, $V = Y+Z$ find $r_{UV}$.

Soln:

Given $X, Y, Z$ are uncorrelated r.v.

$$\therefore \rho_{XY} = 0, \rho_{XZ} = 0, \rho_{YZ} = 0$$

$$\Rightarrow cov(X,Y) = 0, \; cov(X,Z) = 0, \; Cov(Y,Z) = 0 \quad (1)$$

Also given, $\sigma_X^2 = 25$, $\sigma_Y^2 = 144$, $\sigma_Y^2 = 81$ (2)

and $U = X+Y$, $V = Y+Z$.

To Find

$$r_{UV} = \frac{cov(U,V)}{\sigma_U \sigma_V}.$$

Now, $Cov(U,V) = cov(X+Y, Y+Z)$

$$= Cov(X+Y, Y) + Cov(X+Y, Z)$$

$$= cov(X,Y) + cov(Y,Y) + cov(X,Z) + cov(Y,Z)$$

$$= 0 + Cov(Y,Y) + 0 + 0 \text{ (from (1))}$$

$$= cov(Y,Y) = Var(Y) = 144$$

$$\sigma_U^2 = Var(U) = Var(X+Y)$$

$$= Var(X) + Var(Y) + 2cov(X,Y)$$

$$= 25 + 144 + 2 \times 0$$

$$= 169$$

$$\boxed{\therefore \sigma_U = 13}$$

$$\sigma_V^2 = Var(V) = Var(Y+Z)$$
$$= Var(Y) + Var(Z) + 2\,cov(Y,Z)$$
$$= 144 + 81 + 2\times 0 = 225$$

$$\therefore \sigma_V = 15$$
$$\therefore \rho_{UV} = \frac{cov(U,V)}{\sigma_U \sigma_V} = \frac{144}{13\times 15} = \frac{48}{65} = \boxed{0.74}$$

## Regression:

Regression is a mathematical measure of the average relationship between two (or) more variables interms of the original limits of the data.

**Lines of Regression:**

The line of regression of $Y$ on $X$ is given by
$$y-\bar{y} = r\frac{\sigma_y}{\sigma_x}(x-\bar{x})$$
$$\Rightarrow y-\bar{y} = b_{yx}(x-\bar{x})$$

The line of regression of $X$ on $Y$ is given by,
$$x-\bar{x} = r\frac{\sigma_x}{\sigma_y}(y-\bar{y})$$
$$\Rightarrow x-\bar{x} = b_{xy}(y-\bar{y})$$

**Note:** Both the lines of regression passes through $(\bar{x},\bar{y})$

**Regression co-efficients:**

Regression co-efficient of $Y$ on $X$ is
$$b_{yx} = r\frac{\sigma_y}{\sigma_x}.$$

Regression co-efficient of X on Y is

$$b_{xy} = r\frac{\sigma_x}{\sigma_y}$$

Correlation co-efficient: $r = \pm\sqrt{b_{xy}\cdot b_{yx}}$.

Angle Between the Regression Lines: If $\theta$ is angle between 2 regression lines of 2 variables X & Y then,

$$\tan\theta = \left(\frac{1-r^2}{r}\right)\left(\frac{\sigma_x\cdot\sigma_y}{\sigma_x^2+\sigma_y^2}\right)$$

Problems:

(1) From the following data, find

(i) the two regression eqn. (ii) the co-efficient of correlation between the marks in Economics and Statistics, (iii) the most likely marks in statistics when marks in Economics are 30.

| Marks in Economics x | 25 | 28 | 35 | 32 | 31 | 36 | 29 | 38 | 34 | 32 |
|---|---|---|---|---|---|---|---|---|---|---|
| Marks in Statistics y | 43 | 46 | 49 | 41 | 36 | 32 | 31 | 30 | 33 | 39 |

Soln:

| X | Y | $X^2$ | $Y^2$ | XY |
|---|---|---|---|---|
| 25 | 43 | 625 | 1849 | 1075 |
| 28 | 46 | 784 | 2116 | 1288 |
| 35 | 49 | 1225 | 2401 | 1715 |
| 32 | 41 | 1024 | 1681 | 1312 |
| 31 | 36 | 961 | 1296 | 1116 |
| 36 | 32 | 1296 | 1024 | 1152 |
| 29 | 31 | 841 | 961 | 899 |
| 38 | 30 | 1444 | 900 | 1140 |
| 34 | 33 | 1156 | 1089 | 1122 |
| 32 | 39 | 1024 | 1521 | 1248 |
| 320 | 380 | 10380 | 14838 | 12067 |

Here $n = 10$

$$E[x] = \bar{x} = \frac{\Sigma x}{n} = \frac{320}{10} = 32$$

$$E[y] = \bar{y} = \frac{\Sigma y}{n} = \frac{380}{10} = 38$$

$$E[x^2] = \frac{\Sigma x^2}{n} = \frac{10380}{10} = 1038$$

$$E[y^2] = \frac{\Sigma y^2}{n} = \frac{14838}{10} = 1483.8$$

$$E[xy] = \frac{\Sigma xy}{n} = \frac{12067}{10} = 1206.7$$

$$Cov(x,y) = E[xy] - E[x] \cdot E[y]$$

$$= 1206.7 - (32)(38)$$

$$= \boxed{-9.3}$$

$$\sigma_x^2 = Var(x) = E[x^2] - (E[x])^2$$

$$= 14$$

$$\sigma_y^2 = Var(y) = E[y^2] - (E[y])^2$$

$$= 39.8$$

$$r = \frac{cov(x,y)}{\sigma_x \sigma_y}$$

Now,

$$r\frac{\sigma_y}{\sigma_x} = \frac{cov(x,y)}{\sigma_x^2}$$

$$\boxed{b_{yx}} = \frac{cov(x,y)}{\sigma_x^2} = \frac{-9.3}{14} = -0.6643$$

$$b_{xy} = \frac{cov(x,y)}{\sigma_y^2} = \frac{-9.3}{39.8} = -0.2337$$

Eqn:- of the line of regression of $x$ on $y$ is given by $x - \bar{x} = b_{xy}(y - \bar{y})$

$$\Rightarrow x - 32 = (-0.2337)(y - 38)$$

$$\Rightarrow x + 0.2337y - 40.8806 = 0$$

$$\Rightarrow \boxed{x = -0.2337y + 40.8806}$$

Eqn:- of the line of regression of $y$ on $x$ is

$$y - \bar{y} = b_{yx}(x - \bar{x})$$

$$\Rightarrow y - 38 = (-0.6643)(x - 32)$$

$$\Rightarrow \boxed{y = -0.6643x + 59.2576.}$$

Co-efficient of correlation:

$$r = \pm\sqrt{b_{xy} \cdot b_{yx}}$$

$$r = \pm\ 0.394$$

To find the most likely marks in statistics($y$) when marks in Economics ($x$) are 30,

$$y = -0.6643x + 59.2576$$

$$= -0.6643(30) + 59.2576$$

$$= 39.3286$$

$$\boxed{y \simeq 39}$$

(2) The two lines of regression are

$$8x - 10y + 66 = 0 \quad \text{---} (1)$$

$$40x - 18y - 214 = 0 \quad \text{---} (2)$$

The variance of $x$ is 9. Find (i) mean values of $x$ and $y$ (ii) correlation co-efficient between $x$ and $y$.

Soln:-

(i) Since both the lines of regression passes through the mean values $\bar{x}$ & $\bar{y}$, the point $(\bar{x}, \bar{y})$ must satisfies the two eqn:-

regression lines,

$$\therefore 8\bar{x} - 10\bar{y} = -66 \quad \text{--- (3)}$$

$$40\bar{x} - 18\bar{y} = 214 \quad \text{--- (4)}$$

Solving (3) & (4) we get, $\bar{x} = 13$, $\bar{y} = 17$.

Let us suppose that the eqn- (1) is the line of regression of $y$ on $x$ & (2) is the line of regression of $x$ on $y$.

$\therefore$ (1) $\Rightarrow$ $y = \frac{8}{10}x + \frac{66}{10}$

$$\Rightarrow b_{yx} = 8/10.$$

(2) $\Rightarrow$ $40x = 18y + 214$

$$\Rightarrow x = \frac{18}{40}y + \frac{214}{40}$$

$$\Rightarrow b_{xy} = \frac{18}{40}$$

$\therefore \gamma = \pm\sqrt{b_{xy} \cdot b_{yx}} = \pm 0.6$

Since both the regression co-efficients are positive $\gamma$ must be +ve

$$\boxed{\therefore \gamma = 0.6}$$

Transformation of Random Variables:

Definition: If $(X, Y)$ is a two D. r.v with joint p.d.f $f(x,y)$ and if $z = g(x,y)$ & $w = h(x,y)$ are two other r.v's then the joint p.d.f of $(Z, W)$ is given by

$$f(z,w) = |J|\, f(x,y), \text{ Where } |J| = \begin{vmatrix} \frac{\partial x}{\partial z} & \frac{\partial x}{\partial w} \\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w} \end{vmatrix}$$

(1) Let $(X,Y)$ be a two-dimensional non-negative continuous r.v having the joint density

$$f(x,y) = \begin{cases} 4xy\, e^{-(x^2+y^2)}, & x \geq 0,\ y \geq 0 \\ 0, & \text{else where} \end{cases}$$

Find the density fn of $U = \sqrt{X^2+Y^2}$

Soln:

Let $u = \sqrt{x^2+y^2}$ & $v = y$

$$u^2 = x^2+y^2$$

$$x^2 = u^2 - y^2$$

$$x^2 = u^2 - v^2$$

$$x = \sqrt{u^2-v^2}, \quad \& \quad y = v$$

$$\frac{\partial x}{\partial u} = \frac{1}{\cancel{2}}(u^2-v^2)^{\frac{1}{2}-1} \cdot \cancel{2}u \ ; \quad \frac{\partial y}{\partial u} = 0$$

$$= \frac{u}{\sqrt{u^2-v^2}}$$

$$\frac{\partial x}{\partial v} = \frac{1}{\cancel{2}}(u^2-v^2)^{\frac{1}{2}-1} \cdot (-\cancel{2}v) \ ; \quad \frac{\partial y}{\partial v} = 1$$

$$= \frac{-v}{\sqrt{u^2-v^2}}$$

$$\therefore |J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{u}{\sqrt{u^2-v^2}} & \frac{-v}{\sqrt{u^2-v^2}} \\ 0 & 1 \end{vmatrix}$$

$$= \frac{u}{\sqrt{u^2-v^2}}$$

The joint p.d.f of $(U,V)$ is

$$f(u,v) = f(x,y) \cdot |J|$$

$$\therefore f(u,v) = 4xy\, e^{-(x^2+y^2)} \cdot \frac{u}{\sqrt{u^2-v^2}}$$

$$= 4 \times \cancel{\sqrt{u^2-v^2}} \cdot v \cdot e^{-u^2} \cdot \frac{u}{\cancel{\sqrt{u^2-v^2}}}$$

$$= 4uv\, e^{-u^2}.$$

Range Space:

Given, $x \geq 0 \Rightarrow \sqrt{u^2-v^2} \geq 0 \Rightarrow u^2 - v^2 \geq 0$

$\Rightarrow u^2 \geq v^2 \Rightarrow u \geq v$

$\Rightarrow v \leq u$.

$y \geq 0 \Rightarrow v \geq 0$

$\therefore$ The joint p.d.f of $(U, V)$ is

$$f(u,v) = \begin{cases} 4uv\, e^{-u^2}, & u \geq 0,\ 0 \leq v \leq u. \\ 0, & \text{Otherwise} \end{cases}$$

The marginal p.d.f of $U$ is

$$f(u) = \int_{-\infty}^{\infty} f(u,v)\, dv$$

$$= 4\int_0^u uv\, e^{-u^2}\, dv$$

$$= \cancel{4}^{2}\, u\, e^{-u^2} \left(\frac{v^2}{\cancel{2}}\right)_0^u = 2u^3 e^{-u^2}.$$

$\therefore$ The marginal p.d.f of U is

$$f(u) = \begin{cases} 2u^3 e^{-u^2}, & u \geq 0 \\ 0, & \text{Otherwise} \end{cases}$$

(2) If X and Y are independent r.v's with p.d.f $e^{-x}$, $x \geq 0$; $e^{-y}$, $y \geq 0$ respectively. Find the density function of $U = \frac{X}{X+Y}$ and $V = X+Y$. Are U & V are independent?

Sol:-

Since X and Y are independent,

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

$$f(x,y) = \begin{cases} e^{-(x+y)}, & x, y \geq 0 \\ 0, & \text{OW} \end{cases}$$

Also, given, $u = \frac{x}{x+y}$, $v = x+y$.

$$u = \frac{x}{v}, \quad y = v - x$$

$$\Rightarrow x = uv, \quad y = v - uv.$$

$$\frac{\partial x}{\partial u} = v, \quad \frac{\partial y}{\partial u} = -v$$

$$\frac{\partial x}{\partial v} = u, \quad \frac{\partial y}{\partial v} = 1 - u$$

$$\therefore |J| = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} v & u \\ -v & 1-u \end{vmatrix} = v - uv + uv = \boxed{v}.$$

The joint p.d.f of $(U,V)$ is given by,

$$f(u,v) = f(x,y)\,|J|$$

$$= e^{-(x+y)} \cdot v$$

$$= v e^{-v}$$

<u>Range Space:</u>

If $x \geq 0 \Rightarrow uv \geq 0$

$y \geq 0 \Rightarrow v(1-u) \geq 0$

$\therefore$ either $u \geq 0$, $v \geq 0$ and $1-u \geq 0 \Rightarrow 1 \geq u$

(ie) $0 \leq u \leq 1$ & $v \geq 0$

(or) $u \leq 0$, $v \leq 0$, $1-u \leq 0$ (ie) $u \leq 0$, $u > 1$ $\Rightarrow$ which is absurd.

$\therefore$ The range space of $(U,V)$ is given by

$0 \leq u \leq 1$ & $v \geq 0$.

$$\therefore f(u,v) = \begin{cases} v e^{-v}, & 0 \leq u \leq 1,\ v \geq 0 \\ 0, & \text{Otherwise} \end{cases}$$

The Marginal p.d.f of $U$ is,

$$f(u) = \int_{-\infty}^{\infty} f(u,v)\, dv$$

$$= \int_0^{\infty} v e^{-v}\, dv$$

$$= \left[ v \frac{e^{-v}}{-1} - \frac{e^{-v}}{(-1)^2} \right]_0^{\infty} = 1.$$

$\therefore$ U is uniformly distributed in $(0,1)$.

The p.d.f of V is given by,

$$f(v) = \int_{-\infty}^{\infty} f(u,v)\, du$$

$$= \int_0^1 v e^{-v}\, du = v e^{-v} (u)_0^1$$

$$= v e^{-v}$$

$$\therefore f(v) = \begin{cases} v e^{-v}, & v > 0 \\ 0, & \text{ow} \end{cases}$$

Now, $f(u) \cdot f(v) = v e^{-v}$
$= f(u,v)$

$\therefore$ U & V are independent.

**Central Limit Theorem:** (Lindberg - Levy's Form)

Let the r.v's $X_1, X_2, \ldots, X_n$ are independant & identically distributed r.v's with mean $\mu$ and variance $\sigma^2$, $\forall\, i = 1, 2, \ldots, n$.

Then $S_n = X_1 + X_2 + \ldots + X_n$ follows a normal distribution with mean $n\mu$ and variance $n\sigma^2$

(ie) $S_n \sim N(n\mu, \sqrt{n}\,\sigma)$

**Proof:**

To prove as $n \to \infty$

m.g.f of $S_n \to$ m.g.f of Normal variate

(ie) To prove that as $n \to \infty$

$$\text{m.g.f of } Z = \frac{S_n - n\mu}{\sqrt{n}\,\sigma} \to \text{m.g.f of standard normal variate.}$$

$$\text{(ie)} \lim_{n \to \infty} M_Z(t) = e^{\frac{t^2}{2}} \qquad \frac{\bar{X}-\mu}{\sigma}$$

Now,

$$Z = \frac{S_n - n\mu}{\sqrt{n}\,\sigma} = \frac{x_1 + x_2 + \cdots + x_n - n\mu}{\sqrt{n}\,\sigma}$$

$$= \frac{x_1 - \mu + x_2 - \mu + \cdots + x_n - \mu}{\sqrt{n}\,\sigma} = \frac{\sum_{i=1}^{n} x_i - \mu}{\sqrt{n}\,\sigma}$$

$$M_Z(t) = M_{\frac{\sum_{i=1}^{n} x_i - \mu}{\sqrt{n}\,\sigma}}(t) = M_{\sum_{i=1}^{n} x_i - \mu}\left(\frac{t}{\sqrt{n}\,\sigma}\right)$$

$$\left(\because M_{x/a}(t) = M_x(t/a)\right)$$

$$= M_{x_1 - \mu + x_2 - \mu + \cdots + x_n - \mu}\left(\frac{t}{\sqrt{n}\,\sigma}\right)$$

$$= M_{x_1-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right) \cdot M_{x_2-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right) \cdot M_{x_3-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right) \cdots M_{x_n-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right)$$

$$\left(\because M_{x+y}(t) = M_x(t) \cdot M_y(t)\right)$$

$$= \left[M_{x_1-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n \quad (\because \text{m.g.f is same } \forall i) \qquad (1)$$

$$M_X(t) = E(e^{tX}) = 1 + \frac{t}{1!}\mu_1' + \frac{t^2}{2!}\mu_2'$$

$$M_{x_1-\mu}(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu_r' \quad \text{where } E[(x_1-\mu)^r] = \mu_r'$$

$$= \mu_0' + \frac{t}{1!}\mu_1' + \frac{t^2}{2!}\mu_2' + O(t^3)$$

$$= 1 + \frac{t^2}{2!}\sigma^2 + O(t^3) \quad (\because \mu_0' = 1, \mu_1' = 0, \mu_2' = \sigma^2).$$

$$\therefore M_{x_1-\mu}\left(\frac{t}{\sqrt{n}\sigma}\right) = 1 + \frac{t^2}{2!\,n\sigma^2} \times \sigma^2 + O(t^3)$$

$$(x+a)^n = x^n + nc_1 x^{n-1} a + nc_2 x^{n-2} a^2 + \dots + a^n$$

$$= 1 + \frac{t^2}{2n} + O(t^3)$$

From ①, $M_Z(t) = \left[1 + t^2/_{2n} + O(t^3)\right]^n$

$$= \left(1 + t^2/_{2n}\right)^n$$

$$\underset{n\to\infty}{Lt} M_Z(t) = e^{t^2/2} \quad \left(\because \lim_{n\to\infty}\left(1+x/_n\right)^n = e^x\right)$$

$\therefore$ As $n \to \infty$, $S_n$ follows normal dis/- with mean $n\mu$ & variance $n\sigma^2$.

<u>Corollary:</u> If $\bar{X} = \dfrac{x_1 + x_2 + \dots + x_n}{n}$ then $\bar{x}$ follows a normal dis/- with mean $\mu$ and variance $\sigma^2/_n$.

(ie) $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$.

<u>Problems:</u>

① The burning time of a certain type of lamp is an exponential r.v. with mean 30 hrs. What is the probl- that 144 of these lamps will provided a total of <u>more than</u> 4500 hrs of burning time?

<u>Soln:</u> Gn/- $\mu = E[x] = 30$, $\sigma^2 = Var(x) = 30^2 = 900$

($\because$ In exponential dis/- mean $= \lambda$, var $= \lambda^2$)

$n = 144$.

$S_n = x_1 + x_2 + \dots + x_n$.

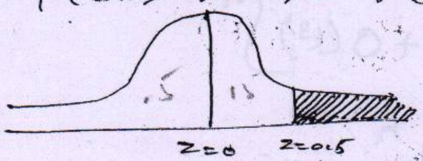
By Using central limit thm/-

$$S_n \sim N(n\mu, \sqrt{n}\sigma)$$

$$S_n \sim N(4320, 360)$$

When $S_n = 4500$, $Z = \frac{S_n - 4320}{360} = \frac{1}{2} = 0.5$

$\therefore P(S_n > 4500) = P(Z > 0.5) = 0.5 - P(0 < Z < 0.5)$




$= 0.5 - 0.1915$

$= \boxed{0.3085}$

(2) The lifetime of a certain brand of an electric bulb may be considered as a R.V with mean 1200 h and s.d 250h. Find the probt, using CLT, that the average life time of 60 bulbs exceeds 1250h.

Soln: Let $X_1, X_2, \ldots, X_{60}$ be the life time of the bulbs.

$\mu = E[X_i] = 1200h$, $i = 1, 2, \ldots, 60$

S.D $\sigma(X_i) = 250h$.

Let $\bar{X}$ denote the average life time of 60 bulbs.

By CLT, $\bar{X}$ follows $N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

To Find

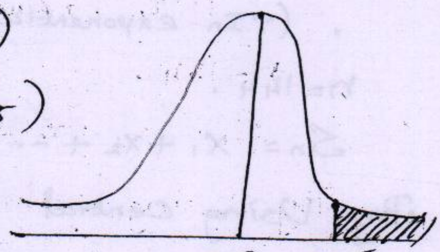
$$P(\bar{X} > 1250) = P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{1250 - \mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$= P\left(Z > \frac{1250 - 1200}{\frac{250}{\sqrt{60}}}\right)$$

$$= P(Z > 1.55)$$

$$= 0.5 - P(0 < Z < 1.55)$$

$$= 0.5 - 0.4394$$




$$= \boxed{0.0606}$$

(3) A dis/- with unknown mean $\mu$ and variance equal to 1.5 Use CLT to find how large a sample should be taken from the dis/- in order that the probl- will be atleast 0.95 that the sample mean $\bar{x}$ will be with in 0.5 of the population mean $\mu$.

Soln:

Let 'n' be the size of the sample.

Gn/- $E[x_i] = \mu$ and $Var(x_i) = 1.5$

Let $\bar{x}$ denote the sample mean

By Corollary to the CLT, $\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

$$(ie)\ \bar{x} \sim N\left(\mu, \frac{\sqrt{1.5}}{\sqrt{n}}\right)$$

We have to find 'n' such that

$$P(|\bar{x}-\mu| < 0.5) \geq 0.95$$

$$P(-0.5 < \bar{x}-\mu < 0.5) \geq 0.95$$

$$\Rightarrow P\left(\frac{-0.5\sqrt{n}}{\sqrt{1.5}} < \frac{\bar{x}-\mu}{\frac{\sqrt{1.5}}{\sqrt{n}}} < \frac{0.5\sqrt{n}}{\sqrt{1.5}}\right) \geq 0.95$$

$$\Rightarrow P(-0.4082\sqrt{n} < Z < 0.4082\sqrt{n}) \geq 0.95$$

$$\Rightarrow 2P(0 < Z < 0.4082\sqrt{n}) \geq 0.95$$

$$\Rightarrow P(0 < Z < 0.4082\sqrt{n}) \geq 0.475$$

From the table of areas under normal curve,

$$P(0 < Z < 1.96) = 0.475$$

$\therefore$ Least value of n is given by $0.4082\sqrt{n} = 1.96$

$$\Rightarrow n = \boxed{23}$$

$\therefore$ The size of the sample must be atleast 23.

(4) A coin is tossed 300 times. What is the probability that heads will appear more than 140 times and less than 150 times Using Central limit Theorem.

Soln: Let $p$ be the prob of getting head in a single trail

$\therefore p = \frac{1}{2};\ q = \frac{1}{2}$ (n is very large as $n \to \infty$, B.D follows an N.D)

Here $n = 300$, (Note that we are applying the B.D).

Now, mean $= np = 150$

Variance $= npq = 75$

To find $P(140 < x < 150) = P(-1.15 < z < 0) = P(0 < z < 1.15)$
$= 0.3749.$

~~(5) The life time of a certain brand of a tube light may be considered as a r.v with mean 1200h and S.D 250 h. Find the probability, using CLT, that the average life time of 60 lights exceeds 1250 h.~~

(5) A random sample of size 100 is taken from a population whose mean is 60 & variance is 400. Using CLT, with what probability can we assert that the mean of the sample will not differ from $\mu = 60$ by more than 4.

Soln $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right) \sim N(60, 2)$

To find $P(|\bar{x} - \mu| \leq 4)$

$$P(|\bar{x} - \mu| \leq 4) = P(-4 \leq \bar{x} - \mu \leq 4)$$
$$= P(-4 \leq \bar{x} - 60 \leq 4)$$
$$= P\left(\frac{-4}{2} \leq \frac{\bar{x} - 60}{2} \leq 4/2\right) = P(-2 \leq z \leq 2)$$
$$= 2P(0 \leq z \leq 2) = 2(0.4773)$$
$$= 0.9546.$$

HW If $x_1, x_2, \ldots x_n$ are poisson variates with parameter $\lambda = 2$, use the CLT, to estimate $P(120 \leq S_n \leq 160)$ where $S_n = x_1 + x_2 + \cdots + x_n$ and $n = 75$.

Ans 0.7866.