TOPIC 2.7- Correlation

1)

Find the coefficient of correlation between x and y from the following data:

x	65	66	67	67	68	69	70	72
y	67	68	65	68	72	72	69	71

Correlation coefficient $r(x,y) = \dfrac{\text{cov}(x,y)}{\sigma_x \sigma_y}$

where $\text{cov}(x,y) = \dfrac{\sum xy}{n} - \bar{x}\bar{y}$

$\sigma_x = \sqrt{\dfrac{1}{n}\sum x^2 - (\bar{x})^2}$ and $\sigma_y = \sqrt{\dfrac{1}{n}\sum y^2 - (\bar{y})^2}$

x	y	$U = x - 68$	$V = y - 68$	UV	U^2	V^2
65	67	−3	−1	3	9	1
66	68	−2	0	0	4	0
67	65	−1	−3	3	1	9
67	68	−1	0	0	1	0
68	72	0	4	0	0	16
69	72	1	4	4	1	16
70	69	2	1	2	4	1
72	71	4	3	12	16	9
		0	8	24	36	52

$$\bar{U} = \frac{\sum U}{n} = 0 \qquad \bar{V} = \frac{\sum V}{n} = \frac{8}{8} = 1$$

$$\text{Cov}(U,V) = \frac{\sum UV}{n} - \bar{U}\bar{V} = \frac{24}{8} = 3$$

$$\sigma_x = \sqrt{\frac{\sum U^2}{n} - \bar{U}^2} = \sqrt{\frac{36}{8}} = 2.1213$$

$$\sigma_y = \sqrt{\frac{\sum V^2}{n} - \bar{V}^2} = \sqrt{\frac{52}{8} - 1} = \sqrt{\frac{44}{8}}$$

$$= 2.3452$$

$$r(x,y) = r(U,V) = \frac{3}{2.1213 \times 2.3452}$$

$$= \underline{\underline{0.603}}$$

2. Two random variables X and Y have the joint pdf given by $f(x,y) = \begin{cases} K(1-x^2y) ; & 0 \le x, y \le 1 \\ 0 , & \text{otherwise} \end{cases}$

i) Find K (ii) obtain the marginal probability density unctions of X and Y (iii) Also find the correlation oefficient between X and Y .

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

$\Rightarrow \int_0^1\int_0^1 K(1-x^2y)\,dx\,dy = 1$

$\Rightarrow K\int_0^1 \left[x - \frac{x^3y}{3}\right]_0^1 dy = 1$

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

$\Rightarrow K\left[1-\frac{1}{6}\right] = 1 \Rightarrow K\left[\frac{5}{6}\right] = 1$

$$\boxed{K = \frac{6}{5}}$$

(ii) Marginal density function of X

$f(x) = \int_{-\infty}^{\infty} f(x,y)\,dy = \int_0^1 \frac{6}{5}(1-x^2y)\,dy$

$= \frac{6}{5}\left[y - \frac{x^2y^2}{2}\right]_0^1 = \frac{6}{5}\left[1-\frac{x^2}{2}\right]$

$= \frac{3}{5}(2-x^2), \quad 0 \le x \le 1$

Marginal density function of y

$$f(y) = \int_{-\infty}^{\infty} f(x,y)\,dx = \int_0^1 \frac{6}{5}(1-x^2y)\,dx$$

$$= \frac{6}{5}\left[x - \frac{x^3y}{3}\right]_0^1 = \frac{6}{5}\left[1-\frac{y}{3}\right]$$

$$= \frac{2}{5}(3-y), \quad 0 \le y \le 1$$

(iii) $$E[X] = \int_{-\infty}^{\infty} x\,f(x)\,dx = \int_0^1 x\,\frac{3}{5}(2-x^2)\,dx$$

$$= \frac{3}{5}\int_0^1 (2x - x^3)\,dx = \frac{3}{5}\left[\frac{2x^2}{2} - \frac{x^4}{4}\right]_0^1$$

$$= \frac{3}{5}\left[1-\frac{1}{4}\right] = \frac{9}{20}$$

$$E[X^2] = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 x^2\,\frac{3}{5}(2-x^2)\,dx$$

$$= \frac{3}{5}\int_0^1 [2x^2 - x^4]\,dx = \frac{3}{5}\left[\frac{2x^3}{3} - \frac{x^5}{5}\right]_0^1$$

$$= \frac{3}{5}\left[\frac{2}{3}-\frac{1}{5}\right] = \frac{3}{5}\left[\frac{10-3}{15}\right] = \frac{7}{25}$$

$$E[Y] = \int_{-\infty}^{\infty} y\, f(y)\, dy = \int_0^1 y\, \frac{2}{5}(3-y)\, dy$$

$$= \frac{2}{5}\int_0^1 (3y - y^2)\, dy = \frac{2}{5}\left[\frac{3y^2}{2} - \frac{y^3}{3}\right]_0^1$$

$$= \frac{2}{5}\left[\frac{3}{2} - \frac{1}{3}\right] = \frac{2}{5}\left[\frac{9-2}{6}\right] = \frac{7}{15}$$

$$E[Y^2] = \int_{-\infty}^{\infty} y^2 f(y)\, dy = \int_0^1 y^2\, \frac{2}{5}(3-y)\, dy$$

$$= \frac{2}{5}\int_0^1 (3y^2 - y^3)\, dy = \frac{2}{5}\left[\frac{3y^3}{3} - \frac{y^4}{4}\right]_0^1$$

$$= \frac{2}{5}\left[1 - \frac{1}{4}\right] = \frac{3}{10}$$

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

$$= \int_0^1\int_0^1 xy\, \frac{6}{5}(1 - x^2 y)\, dx\, dy$$

$$= \frac{6}{5}\int_0^1\int_0^1 \left[xy - x^3 y^2\right] dx\, dy$$

$$= \frac{6}{5}\int_0^1 \left[\frac{x^2 y}{2} - \frac{x^4 y^2}{4}\right]_0^1 dy$$

$$= \frac{6}{5}\int_0^1 \left[\frac{y}{2} - \frac{y^2}{4}\right] dy = \frac{6}{5}\left[\frac{y^2}{4} - \frac{y^3}{12}\right]_0^1$$

$$= \frac{6}{5}\left[\frac{1}{4} - \frac{1}{12}\right] = \frac{6}{5}\left[\frac{2}{12}\right] = \frac{1}{5}$$

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

$$= \frac{1}{5} - \frac{9}{20}\,\frac{7}{15} = \frac{1}{5}$$

$$= -\frac{1}{100}$$

$$Var(x) = \frac{7}{25} - \left(\frac{9}{20}\right)^2 = \frac{7}{25} - \frac{81}{400}$$

$$= \frac{112-81}{400} = \frac{31}{400}$$

$$Var(y) = \frac{3}{10} - \left(\frac{7}{15}\right)^2 = \frac{3}{10} - \frac{49}{225}$$

$$= \frac{135-98}{450} = \frac{37}{450}$$

$$r(x,y) = \frac{cov(x,y)}{\sigma_x \sigma_y} = \frac{-1/100}{\sqrt{\frac{31}{400}}\sqrt{\frac{37}{450}}}$$

$$= -0.1253$$