MAT2001 - DA

S.THARUN
19MID0031

STATISTICS FOR ENGINEERS

1) If X and Y have the joint p.d.f $f(x,y) = \frac{1}{3}(x+y)$, $0 \le x \le 1$, $0 \le y \le 2$, then find

(i) $r(x,y)$

(ii) the two lines of regression

(iii) the two regression curves for the means.

(i) $$r_{xy} = \frac{\text{cov}(X,Y)}{\sigma_x \cdot \sigma_y}$$

$$\text{cov}(x,y) = E(x \cdot y) - E(x) \cdot E(y)$$

$$E(x) = \int_{x_1}^{x_2} x \cdot f(x) \cdot dx$$

$$E(y) = \int_{y_1}^{y_2} y \cdot f(y) \cdot dx$$

$$E(x \cdot y) = \int_{x_1}^{x_2}\int_{y_1}^{y_2} xy\, f(x,y) \cdot dx \cdot dy$$

$$\sigma_x = \sqrt{\text{var}(x)} = \sqrt{E(x^2) - [E(x)]^2}$$

$$\sigma_y = \sqrt{\text{var}(y)} = \sqrt{E(y^2) - [E(y)]^2}$$

$$E(x^2) = \int_{x_1}^{x_2} x^2 f(x) \cdot dx$$

$$E(y^2) = \int_{y_1}^{y_2} y^2 \cdot f(y) \cdot dy$$

$$f(x) = \int_{y_1}^{y_2} f(x,y) \cdot dy$$

$$f(y) = \int_{x_1}^{x_2} f(x,y) \cdot dx$$

Scanned with CamScanner

$$f(x) = \int_{y_1}^{y_2} f(x,y)\cdot dy$$

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

$$= \frac{2}{3}(x+1)$$

$$f(y) = \int_{x_1}^{x_2} f(x,y)\cdot dx$$

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

$$= \frac{1}{6}(2y+1)$$

$$E(X) = \int_{x_1}^{x_2} x\cdot f(x)\cdot dx$$

$$= \frac{2}{3}\int_0^1 (x^2+x)\cdot dx = \frac{2}{3}\left[\frac{x^3}{3} + \frac{x^2}{2}\right]_0^1 = \frac{2}{3}\left[\frac{1}{3}+\frac{1}{2}\right]$$

$$= \frac{10}{18} = \frac{5}{9}$$

$$E(Y) = \int_{y_1}^{y_2} y\cdot f(y)\cdot dy$$

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

$$= \frac{1}{6}\left[\frac{16}{3}+2\right] = \frac{16}{18} + \frac{2}{6} = \frac{8}{9} + \frac{2}{6} = \frac{11}{9}$$

$$E(X \cdot Y) = \int_{x_1}^{x_2} \int_{y_1}^{y_2} xy\, f(x,y) \cdot dx\, dy$$

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

$$= \frac{1}{3} \int_0^1 \left[\frac{x^2y^2}{2} + \frac{xy^3}{3}\right]_0^2 \cdot dx \cdot = \frac{1}{3} \int_0^1 \left[2x^2 + \frac{8x}{3}\right] \cdot dx$$

$$= \frac{1}{3} \left[\frac{2x^3}{3} + \frac{8x^2}{6}\right]_0^1 = \frac{1}{3}\left[\frac{2}{3} + \frac{4}{3}\right] = \frac{2}{3}$$

$$E(x^2) = \int_{x_1}^{x_2} x^2 f(x) \cdot dx$$

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

$$= \frac{2}{3}\left[\frac{1}{4} + \frac{1}{3}\right] = \frac{2}{12} + \frac{2}{9} = \frac{6+8}{36} = \frac{14}{36} = \frac{7}{18}$$

$$E(y^2) = \int_{y_1}^{y_2} y^2 f(y) \cdot dy$$

$$= \frac{1}{6}\int_0^2 (2y^3 + y^2) \cdot dy = \frac{1}{6}\left[\frac{2y^4}{4} + \frac{y^3}{3}\right]_0^2 = \frac{1}{6}\left[8 + \frac{8}{3}\right]$$

$$= \frac{32}{18} = \frac{16}{9}$$

$$var(x) = E[x^2] - [E(x)]^2 = \frac{7}{18} - \frac{25}{81} = \frac{13}{162}$$

$$var(y) = E(y^2) - [E(y)]^2 = \frac{16}{9} - \frac{121}{81} = \frac{23}{81}$$

$$\sigma_x = \sqrt{var(x)} = 0.283278861$$

$$\sigma_y = \sqrt{var(y)} = 0.532870169$$

Scanned with CamScanner

$$\operatorname{cov}(x,y) = E(x\cdot y) - E(x)\cdot E(y)$$

$$= \frac{2}{3} - \frac{5}{9}\cdot\frac{11}{9} = \frac{2}{3} - \frac{55}{81} = -\frac{1}{81}$$

$$= -0.012345679$$

$$r_{xy} = \frac{\operatorname{cov}(x,y)}{\sigma_x \cdot \sigma_y} = \frac{-0.012345679}{0.283278861 \times 0.532870169}$$

$$= \frac{-0.012345679}{0.150950854}$$

$$= -0.081786082$$

Scanned with CamScanner

(ii) Regression line y on x

$$y - y_1 = r(x - \bar{x})\frac{\sigma_y}{\sigma_x}$$

$$y - \frac{11}{9} = (-0.081786082)\left(\frac{0.532870169}{0.283278861}\right)\left(x - \frac{5}{9}\right)$$

$$y - \frac{11}{9} = (-0.153846154)\left(x - \frac{5}{9}\right)$$

$$9y - 11 = -1.384615388x + 0.769230771$$

$$\boxed{y = -0.153846154x + 1.307692308}$$

Regression line x on y

$$x - x_1 = r\frac{\sigma_x}{\sigma_y}(y - y_1)$$

$$\left(x - 5/9\right) = (-0.081786082)\left(\frac{0.283278861}{0.532870169}\right)\left(y - \frac{11}{9}\right)$$

$$x - 5/9 = -0.04347826\,(y - 11/9)$$

$$9x - 5 = -0.391304346y + 0.478260867$$

$$9x = -0.391304346y + 5.478260867$$

$$\boxed{x = -0.04347826y + 0.608695652}$$

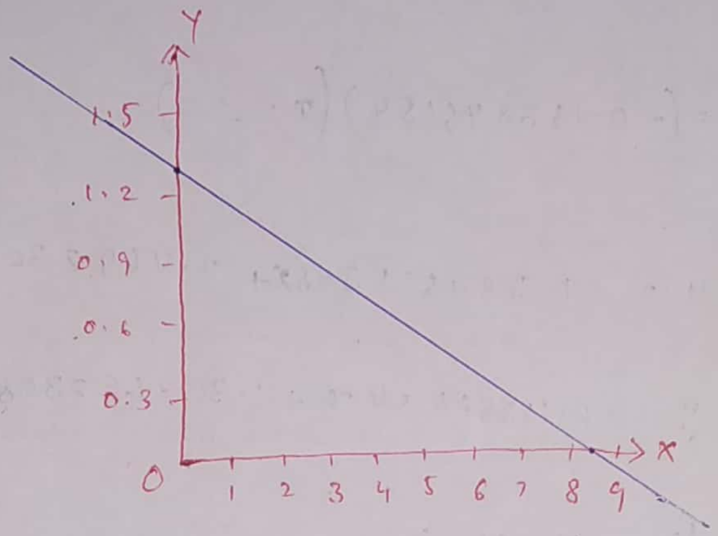
19M1D0031

(iii) y on x

$y = -0.153846154x + 1.307692308$

y-intercept = 1.307692308 ≃ 1.3

x-intercept = 8.499999994 ≃ 8.5

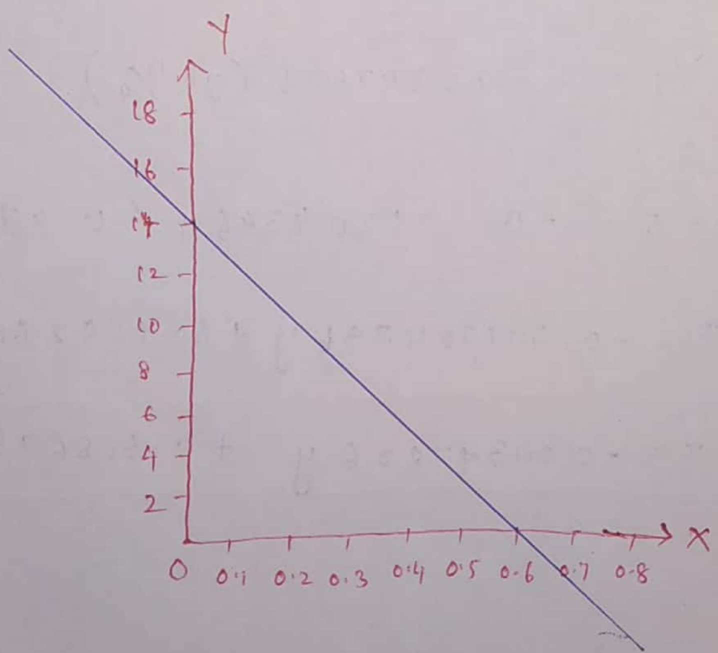


x on y

$x = -0.0434782 6y + 0.608695652$

x-intercept = 0.608695652 ≃ 0.6

y-intercept = 14.00000028 ≃ 14



2) If X, Y and Z are uncorrelated r.v's with zero means and S.D's 5, 13 and 9 respectively and if $U = X+Y$ and $V = Y+Z$, find the correlation co-efficient between U and V.

$E(X) = E(Y) = E(Z) = 0$

$cov(X,Y) = 0$, $cov(Y,Z) = 0$, $cov(Z,X) = 0$

$\sigma_x = 5$; $\sigma_y = 12$; $\sigma_z = 9$

$$\left.\begin{array}{l} U = X+Y \\ V = Y+Z \end{array}\right\} \Rightarrow r(U,V) = ?$$

$$\frac{cov(U,V)}{\sigma_u \cdot \sigma_v} = \frac{E(U \cdot V) - E(U)E(V)}{\sigma_u \sigma_v}$$

$E(U) = E(X+Y) = E(X) + E(Y) = 0$

$E(V) = E(Y+Z) = E(Y) + E(Z) = 0$

$var(X) = E(X^2) - [E(X)]^2 = E(X^2) - 0$

$\sigma_x^2 = E(x^2)$

$\Rightarrow E(x^2) = 25$

$E(y^2) = 144$

$E(z^2) = 81$

$E(UV) = E[(x+y)(y+z)]$

$= E[xy + xz + y^2 + yz]$

$= E[xy] + E(xz) + E(y^2) + E(yz)$

$= 0 + 0 + 144 + 0$

$E(UV) = 144$

Scanned with CamScanner

$var(U) = E(U^2) - [E(u)]^2 = E(u^2)$

$= E[(x+y)^2] = E(x^2+y^2+2xy) = E(x^2) + E(y^2) + E(2xy)$

$= E(x^2) + E(y^2) + 2E(x) \cdot E(y)$

$= 25 + 144 = 169$

$\sigma_u = \sqrt{var(u)} = \sqrt{169} = 13$

$|||^{ly}$

$var(V) = E(V^2) - [E(V)]^2 = E(V^2)$

$= E(y^2) + E(z^2) + 2E(y) \cdot E(z)$

$= 144 + 81 = 225$

$\sigma_v = \sqrt{var(v)} = \sqrt{225} = 15$

$$r_{uv} = \frac{cov(u,v)}{\sigma_u \cdot \sigma_v}$$

$$= \frac{E(uv) - E(u) \cdot E(v)}{\sigma_u \cdot \sigma_v} = \frac{E(uv) - 0}{\sigma_u \cdot \sigma_v}$$

$$= \frac{144}{(13)(15)} = \frac{144}{195} = 0.73846154$$

3) The life length x of an electronic component follows an exponential distribution, there are two processes by which the component may be manufactured. The expected life length of the component is 100h, if process I is used to manufacture while it is 150h, if process II is used. The cost of manufacturing a single component process one is Rs.10, while it is Rs.20 for process II. Moreover if the component lasts less than the guaranteed life of 200h, a loss of Rs.50 is to borne by the manufacturer. which process is advantageous to the manufacturer?

$$\lambda_1 = \frac{1}{100} \Rightarrow f_1 = \lambda_1 e^{-\lambda_1 x} = \frac{e^{-x/100}}{100}$$

$$\lambda_2 = \frac{1}{150} \Rightarrow f_2 = \lambda_2 e^{-\lambda_2 x} = \frac{e^{-x/150}}{150}$$

$$\underset{(f_1)}{P_1(x \leq 200)} = \int_0^{200} \frac{e^{-x/100}}{100} \cdot dx$$

$$= \frac{1}{100} \int_0^{200} e^{-x/100} \cdot dx$$

On solving,

$$= \frac{1}{100} (100 - 100e^{-2}) = \frac{86.46}{100} = 0.8646 \text{ --- } (1)$$

Scanned with CamScanner

$$P_2(x<200) \, (f_2) = \int_0^{200} \frac{e^{-x/150}}{150} \cdot dx$$

$$= \frac{1}{150} \int_0^{200} e^{-x/150} \cdot dx$$

On solving,

$$= \frac{1}{150} \left(150 - 150 e^{-4/3}\right)$$

$$= \frac{1}{150} \left(150 - 150 e^{-1.33}\right)$$

$$= \frac{110.46}{150} = 0.7364 \quad \text{—— (2)}$$

Now,

Avg cost for Process 1 $= 10 + 50 \times P_1(x \leq 200)$

$= 10 + 50\,(0.8646)$

$= 10 + 43.23$

$=$ ₹53.23 —— (3)

Avg cost for process 2 $= 20 + 50 \times P(x \leq 200)$

$= 20 + 50 \times 0.7364$

$= 20 + 36.82$

$=$ ₹56.82 — (4)

From (3) and (4)

Process I is advantageous than process II for the manufacturer.

4) If the density function of a continuous r.v X is $f(x) = ce^{-b(x-a)}$, $a \le x$, where a, b, c are constants, show that $b = c = \frac{1}{\sigma}$ and $a = \mu - \sigma$ where $\mu = E(X)$ and $\sigma = var(X)$

$$f(x) = ce^{-b(x-a)}, \quad x \ge a$$

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

$$\Rightarrow \int_{a}^{\infty} ce^{-b(x-a)} \, dx = 1$$

$$\Rightarrow ce^{ab} \int_{a}^{\infty} e^{-bx} \, dx = 1$$

$$\Rightarrow -\frac{ce^{ab}}{b} \left[e^{-bx} \right]_{a}^{\infty} = 1$$

$$\Rightarrow -\frac{ce^{ab}}{b} \left[0 - e^{-ab} \right] = 1$$

$$\Rightarrow -\frac{ce^{ab}}{b} \left[-e^{-ab} \right] = 1$$

$$\Rightarrow \frac{c}{b} = 1$$

$$\Rightarrow \boxed{c = b}$$

$$\text{Mean} = E(x) = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx$$

$$= \int_{a}^{\infty} x\, c\, e^{-(b(x-a))} \cdot dx = \int_{a}^{\infty} x\, b\, e^{-b(x-a)} \cdot dx$$

$$= b e^{ab} \int_{a}^{\infty} x e^{-bx} \cdot dx$$

$u = x$ $\quad dv = e^{-bx} \cdot dx$

$du = dx$ $\quad v = \frac{e^{-bx}}{-b}$

$$= b e^{ab} \left[-\frac{x e^{-bx}}{b} \Big|_{a}^{\infty} + \int_{a}^{\infty} \frac{e^{-bx}}{b} \cdot dx \right]$$

$$= b e^{ab} \left[\left(0 + \frac{a e^{-ba}}{b}\right) + \frac{1}{b} \left[\frac{e^{-bx}}{-b} \right]_{a}^{\infty} \right]$$

$$= e^{ab} \left[a e^{-ab} - \frac{1}{b} (0 - e^{-ab}) \right]$$

$$= e^{ab} \left[a e^{-ab} + \frac{e^{-ab}}{b} \right]$$

$$\boxed{\mu = E(x) = a + \frac{1}{b}}$$

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

$u = x^2$ $\quad dv = e^{-bx} \cdot dx$

$du = 2x \cdot dx$ $\quad v = \frac{e^{-bx}}{-b}$

$$= \int_{a}^{\infty} c x^2 e^{-b(x-a)} \cdot dx$$

$$= c e^{ab} \int_{a}^{\infty} x^2 e^{-bx} \cdot dx$$

$$= ce^{ab}\left[-\frac{x^2 e^{-bx}}{b}\Bigg|_a^{\infty} + \int_a^{\infty} \frac{2xe^{-bx}}{b}\cdot dx\right]$$

$$= ce^{ab}\left[\left(0+\frac{a^2 e^{-ab}}{b}\right) + \frac{2}{b}\int_a^{\infty} xe^{-bx}\cdot dx\right]$$

$$= \frac{be^{ab}}{b}\left[a^2 e^{-ab} + 2\left[-\frac{xe^{-bx}}{b}\Bigg|_a^{\infty} + \int_a^{\infty}\frac{e^{-bx}}{b}\cdot dx\right]\right] \quad \because [c=b]$$

$$= e^{ab}\left[a^2e^{-ab} + 2\left[\frac{ae^{-ab}}{b} + \left[\frac{e^{-bx}}{-b^2}\right]_a^{\infty}\right]\right]$$

$$= e^{ab}\left[a^2e^{-ab} + \frac{2ae^{-ab}}{b} + (+2)\frac{e^{-ab}}{b^2}\right]$$

$$\boxed{E(x^2) = a^2 + \frac{2a}{b} + \frac{2}{b^2}}$$

$$var(x) = E(x^2) - [E(x)]^2$$

$$= a^2 + \frac{2a}{b} + \frac{2}{b^2} - \left(a+\frac{1}{b}\right)^2$$

$$= a^2 + \frac{2a}{b} + \frac{2}{b^2} - \left(a^2 + \frac{1}{b^2} + \frac{2a}{b}\right)$$

$$\boxed{var(x) = \frac{1}{b^2}}$$

Scanned with CamScanner

Standard Deviation, $\sigma = \sqrt{var(x)}$

$= \sqrt{1/b^2}$

$\sigma = 1/b = 1/c$ [$\because$ $b = c$

$\Rightarrow \boxed{b = c = \frac{1}{\sigma}}$

$\mu = a + \frac{1}{b} = a + \sigma$

$\Rightarrow \boxed{a = \mu - \sigma}$

5) A study of prices of rice at chennai and vellore gave the following data

	Chennai	Vellore
Mean	19.5	17.75
S.D	1.75	2.5

Also the co-efficient of correlation between the two is 0.8. Estimate the most likely price of rice

(i) at chennai corresponding to the price of 18 at vellore and

(ii) at vellore corresponding to the price of 17 at chennai.

(Chennai) (Vellore)

$\bar{x} = 19.5$ $\bar{y} = 17.75$

$\sigma_x = 1.75$ $\sigma_y = 2.5$ $r_{xy} = 0.8$

Lines of Regression of y on x :-

$$y - \bar{y} = r \frac{\sigma_y}{\sigma_x} (x - \bar{x})$$

$$y - 17.75 = 0.8 \left(\frac{2.5}{1.75} \right) (x - 19.5)$$

$$y - 17.75 = 1.1428x - 22.2846$$

$$y = 1.1428x - 4.5346$$

Lines of Regression of x on y :-

$$x - \bar{x} = r \frac{\sigma_x}{\sigma_y} (y - \bar{y})$$

$$x - 19.5 = 0.8 \left(\frac{1.75}{2.5} \right) (y - \bar{y})$$

$$x - 19.5 = 0.8 \left(\frac{1.75}{2.5} \right) (y - 17.75)$$

$$x - 19.5 = 0.564(y - 17.75)$$

$$x = 0.564y + 9.56$$

(i) $y = 18 \Rightarrow x = ?$

$$x = 0.564\,y + 9.56$$

$$= (0.564)(18) + 9.56$$

$$= 10.08 + 9.56 = 19.64$$

Most likely price at chennai corresponding to the price of 18 at vellore is 19.64

(ii) $x = 17 \Rightarrow y = ?$

$$y = 1.1428\,x - 4.5346$$

$$= 1.1428(17) - 4.5346$$

$$= 19.4276 - 4.5346$$

$$y = 14.893$$

Most likely price of rice at vellore corresponding to the price of 17 at chennai is 14.893