Tutorial

Ex: Three fair coins are tossed. Let X denote the number of heads on the first two coins and Y denote the number of tails on the last two coins. Find $cov(X,Y)$. Also evaluate correlation coeff^n.

Sol^n: Three fair coins are tossed then sample space is given by $S = \{HHH, HHT, HTH, TTH, THT, TTT, HTT, THH\}$.

X : no. of heads on first two coins

$R_X : \{0, 1, 2\}$

Y : no. of tails on the last two coins

$R_Y : \{0, 1, 2\}$

The joint PMF of (X,Y) is given in the table.

X \ Y	0	1	2	$p_X(x)$
0	0	1/8	1/8	2/8
1	1/8	2/8	1/8	4/8
2	1/8	1/8	0	2/8
$p_Y(y)$	$\frac{2}{8}$	$\frac{4}{8}$	$\frac{2}{8}$	1

$$E(X) = \sum_{x=0}^{2} x\, p_X(x)$$

$$= 0 \cdot p_X(0) + 1\, p_X(1) + 2\, p_X(2)$$

$$= 0 \cdot 2/8 + 1 \cdot 4/8 + 2 \cdot 2/8$$

$$= 1$$

Similarly $\boxed{E(Y) = 1}$

$$E(XY) = \sum_{y=0}^{2}\sum_{x=0}^{2} xy\, p_{X,Y}(x,y)$$

$$= \sum_{y=0}^{2}\left[0 + y\, p_{XY}(1,y) + 2y\, p_{X,Y}(2,y)\right]$$

$$= \sum_{y=0}^{2} y\, p_{X,Y}(1,y) + 2\sum_{y=0}^{2} y\, p_{X,Y}(2,y)$$

$$= p_{X,Y}(1,1) + 2\, p_{X,Y}(1,2) + 2\left[p_{X,Y}(2,1) + 2p_{X,Y}(2,2)\right]$$

$$= \frac{2}{8} + 2\cdot\frac{1}{8} + 2\left[\tfrac{1}{8} + 2\cdot 0\right] = 3/4.$$

$$\therefore Cov(X,Y) = E(XY) - E(X)\cdot E(Y).$$
$$= 3/4 - 1\cdot 1 = -\frac{1}{4} = -0.25$$

$\boxed{Cov(X,Y) = -0.25}$. Try your self for $\rho_{X,Y}$.

<u>Ex</u>: Let joint pdf of X and Y be given by

$$f_{X,Y}(x,y) = \begin{cases} e^{-x}, & 0 < y < x < \infty \\ 0, & \text{otherwise} \end{cases}$$

Determine the correlation coefficient between X and Y.

<u>Solution</u>: Note that correlation coeff^n between X & Y is

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \cdot \sigma_Y}$$

where $Cov(X,Y) = E(XY) - (EX)(EY)$

$\sigma_x \to$ standard deviation of X

$$= \sqrt{V(X)} = \sqrt{EX^2 - (EX)^2}$$

Similarly

$\sigma_y \longrightarrow$ standard deviation of Y

$$= \sqrt{E(Y^2) - (EY)^2}$$

To evaluate $\rho_{X,Y}$ we require marginal pdfs of X and Y both.

Marginal pdf of X is

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy = \int_0^x e^{-x}\,dy$$

$$= x e^{-x}.$$

$$\therefore \boxed{f_X(x) = x e^{-x}, \quad 0 < x < \infty}$$

Similarly $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx = \int_y^{\infty} e^{-x}\,dx$

$$= e^{-y}$$

$$\therefore \boxed{f_Y(y) = e^{-y}, \quad 0 < y < \infty}$$

$$\therefore E(X) = \int_0^{\infty} x \cdot f_X(x)\,dx = \int_0^{\infty} x^2 e^{-x}\,dx$$

$$= 2$$

$$E(X^2) = \int_0^\infty x^2 f_X(x)\,dx = \int_0^\infty x^3 e^{-x}\,dx$$

$$= 6.$$

$$\therefore V(X) = E(X^2) - (EX)^2 = 6 - 4 = 2.$$

$$\therefore \sigma_X = \sqrt{2}.$$

Next, $E(Y) = \int_0^\infty y e^{-y}\,dy = 1$

$$E(Y^2) = \int_0^\infty y^2 e^{-y}\,dy = 2.$$

$$V(Y) = 2 - 1 = 1.$$

$$\sigma_Y = 1.$$

Finally we compute

$$E(XY) = \int_{x=0}^{\infty}\int_{y=0}^{x} xy\, f_{X,Y}(x,y)\,dy\,dx$$

$$= \int_0^\infty x \int_0^x y e^{-x}\,dy\,dx$$

$$= \int_0^\infty x e^{-x}\left(\frac{y^2}{2}\right)_0^x dx = \frac{1}{2}\int_0^\infty x^3 e^{-x}\,dx$$

$$= 3.$$

$$\therefore Cov(X,Y) = E(XY) - E(X)E(Y) = 3 - 2\cdot 1 = 1.$$

$$\therefore \rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{1}{\sqrt{2}\cdot 1} = \frac{1}{\sqrt{2}}.$$

Ex: Let $(X,Y) \sim BVN(0,0,1,1,0.5)$. Find the expected value of $e^{\frac{XY}{2}}$.

Solution: By Definition $E\left(e^{XY/2}\right) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{\frac{xy}{2}} f_{X,Y}(x,y)\,dx\,dy$

where $f_{X,Y}(x,y) \sim BVN(0,0,1,1,0.5)$.

Try to simplify this double integral to get the answer.

However we apply the iterated expectation technique to obtain the result.

Thus we have

$$E\left\{e^{\frac{XY}{2}}\right\} = E\left[E\left\{e^{\frac{XY}{2}} \,\middle|\, X=x\right\}\right] \quad \text{——— (1)}$$

~~The~~ The inner expectation is evaluated with respect to the conditional distribution of $Y|X=x$.

$$\boxed{\begin{array}{l}\text{Recall if } (X,Y) \sim BVN(\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho) \text{ then} \\ Y|X=x \sim N\left(\mu_2+\rho\sigma_2\left(\frac{x-\mu_1}{\sigma_1}\right),\ \sigma_2^2(1-\rho^2)\right)\end{array}}$$

In our case $(X,Y) \sim BVN(0,0,1,1,0.5)$

It is easily verified that

$$Y \mid X=x \sim N\left(\frac{x}{2}, \frac{3}{4}\right)$$

So

$$E\left[e^{\frac{XY}{2}} \,\middle|\, X=x\right] = E\left(e^{\frac{x}{2}\cdot\frac{x}{2} + \frac{1}{2}\frac{x^2}{4}\cdot\frac{3}{4}}\right)$$

$$= E\left(e^{x^2/4 + \frac{3x^2}{32}}\right) = E\left(e^{\frac{11}{32}x^2}\right) \quad \text{—} (2)$$

$\therefore$ From Equation (1) we have (after using (2))

$$E\left(e^{\frac{XY}{2}}\right) = E\left[E\left(e^{\frac{XY}{2}} \,\middle|\, X=x\right)\right]$$

$$= E\left(e^{\frac{11}{32}x^2}\right)$$

For the given problem

$X \sim N(0,1)$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{11}{32}x^2} \cdot e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{5}{32}x^2} dx$$

$$= \sqrt{\frac{16}{5}}.$$

Ex: In a statistical study, the heights of husbands and their wives are measured and found to have a bivariate normal distribution. Let X_1 denote the height of a randomly selected husband and X_2 be the heights of his wife with $\mu_1 = 68$ inches, $\mu_2 = 64''$, $\sigma_1 = 4''$, $\sigma_2 = 3.6''$, $\rho = 0.25$. With these data

(i) find the expected height of a man whose wife is $61''$ tall.

(ii) what is the prob. of wife being taller than her husband if the husband is of average height

(iii) What is the probability of the wife being taller than the third quartile of all the wives heights, if her husbands height is at the third quartiles of all the husband heights.

Soln: $(X_1, X_2) \sim BVN(68, 64, 16, (3.6)^2, 0.25)$

(i) Here we want

$$E(X_1 \mid X_2 = 61).$$

$$X_1 \mid X_2 = 61 \sim N\left(68 + 0.25 \times 4 \cdot \left(\frac{61-64}{3.6}\right), 16\cdot(1-0.25^2)\right)$$

Then $E(X_1 | X_2 = 61) = 68 + .25 \times 4 \left(\frac{61-64}{3.6}\right)$

$= 67.17.$

(ii) In this part we want the probability

$P(X_2 > 68 | X_1 = 68).$

To obtain this prob we need conditional distn of $X_2 | X_1 = 68.$

$X_2 | X_1 = 68 \sim N\left(64 + 0.25 \times 3.6 \left(\frac{68-68}{4}\right), (3.6)^2(1-0.25^2)\right)$

$\sim N(64, 12.15)$

$\therefore P(X_2 > 68 | X_1 = 68)$

$= P\left(\frac{X_2 - 64}{\sqrt{12.15}} > \frac{68-64}{\sqrt{12.15}}\right)$

$= P\left(Z > \frac{4}{3.49}\right) = P(Z > 1.15)$

$= \cancel{1 - P(Z \le 1.15)}\ 1 - P(Z \le 1.15)$

$= 1 - \Phi(1.15) = 1 - (0.5 + .3749)$

$= 1 - 0.8749 = 0.126$

(iii) Note that if $Z \sim N(0,1)$ then third quartile $z_{0.75}$ is given by

$$P(Z < z_{0.75}) = 0.75$$

From normal table $z_{0.75} = 0.6745$ —— (*)

$\therefore$ The third quartiles of all the wives heights is

$$P(X_2 < x_{0.75}) = 0.75$$

$$\Rightarrow P\left(\frac{X_2 - E(X_2)}{\sqrt{V(X_2)}} < \frac{x_{0.75} - \mu_2}{\sigma_2}\right) = 0.75$$

$$\Rightarrow P\left(Z < \frac{x_{0.75} - 64}{3.6}\right) = 0.75 \quad \text{—— (**)}$$

From (*) & (**) we must have

$$\frac{x_{0.75} - 64}{3.6} = 0.6745$$

$$\Rightarrow x_{0.75} = 64 + 0.6745 \times 3.6 = 66.428$$

Similarly third quartiles of all husbands heights is

$$x^*_{0.75} = 68 + 0.6745 \times 4 \simeq 70.698$$

Thus we want the probability

$$P(X_2 > 66.428 \mid X_1 = 70.698).$$

Thus to compute this probability we now require the conditional prob. distribution of $X_2 | X_1 = 70.698$

for the given problem

$$X_2 | X_1 = 70.698 \sim N\left(64 + 0.25 \times 3.6 \times \left(\frac{70.698 - 68}{4}\right), (3.6)^2 (1 - 0.25^2)\right)$$

$$\sim N(64.607, (3.4857)^2).$$

$$\therefore P(X_2 > 66.428 | X_1 = 70.698)$$

$$= P\left(\frac{X_2 - 64.607}{3.4857} > \frac{66.428 - 64.607}{3.4857}\right)$$

$$= P(Z > 0.53) = 1 - P(Z \le 0.53)$$

$$= 1 - \Phi(.53) = 1 - (0.5 + 0.2019)$$

$$\simeq 0.321.$$

Ex: Let (X,Y) be jointly distributed as

$$f_{X,Y}(x,y) = \frac{x^2+y^2}{4\pi} e^{-\frac{x^2+y^2}{2}}, \quad -\infty<x<\infty,\ -\infty<y<\infty.$$

Check if X and Y are independent and correlated.

Ex: Let $f_{X,Y}(x,y) = 2-x-y$, $0\le x\le 1, 0\le y\le 1$. Find marginal and conditional pdfs. Further evaluate ρ_{XY}.

Ex: Let $(X,Y) \sim BVN(5, 10, 1, 25, \rho)$.

(i) Let $\rho > 0$ then find if when $P(4<Y<16 \mid X=5) = 0.95$

(ii) If $\rho = 0$ then find $P(X+Y \le 16)$.

Ex: The life of a tube (in hrs) X and filament diameter (in inches) Y are jointly distributed as $BVN(2000, 0.10, 2500, 0.01, 0.87)$. If a filament diameter is 0.098 what is the prob. that tube will last at least 1950 hrs. What is the average expected life of a tube when diameter is 0.098. Find the corresponding conditional variance.