2.7

Correlation coefficient and Bivariate Normal Distribution

Meaning of correlation:

In a bivariate distribution we may be interested to find out if there is any **correlation** or **covariance** between the two variables under study. If the change in one variable affects a change in the other variable, the variables are said to be **correlated**. If the two variables deviate in the same direction, $i.e.$, if the increase (or decrease) in one results in a corresponding increase (or decrease) in the other, **correlation** is said to be **positive**. But, if they constantly deviate in the opposite directions, $i.e.$, if increase (or decrease) in one results in corresponding decrease (or increase) in the other, ***correlation*** *is said to be* ***negative***. For example, the correlation between (i) the heights and weights of a group of persons, and (ii) the income and expenditure; is positive and the correlation between (i) price and demand of a commodity and (ii) the volume and pressure of a perfect gas; is negative. *Correlation is said to be* ***perfect*** if the deviation in one variable is followed by a corresponding and proportional deviation in the other.

Karl Pearson's Coefficient of Correlation:

As a measure of intensity or degree of linear relationship between two variables, **Karl Pearson**, a British Biometrician developed a formula called **correlation coefficient**. Correlation coefficient between two variables X and Y, usually denoted by $\boldsymbol{\rho(X,Y)}$ or $\boldsymbol{\rho_{XY}}$, is a numerical measure of linear relationship between them and is defined by

$$\boldsymbol{\rho(X,Y) = \frac{\sigma_{XY}}{\sigma_X.\sigma_Y} = \frac{cov\,(X,Y)}{\sqrt{V(X)}\,\sqrt{V(Y)}}}$$

where $\sigma_{XY} = cov\,(X,Y) = E\big[\big(X - E(X)\big)\,\big(Y - E(y)\big)\big]$,

$\sigma_X{}^2 = V(X) = E\big[\big(X - E(X)\big)\big]^2$ and $\sigma_Y{}^2 = V(Y) = E\big[\big(Y - E(Y)\big)\big]^2$

Note:

1. $\rho(X, Y)$ provides a measure of linear relationship between X and Y. For non – linear relationship, however, it is not suitable.

2. Karl Pearson's correlation coefficient is also called **product – moment correlation coefficient.**

Properties:

1. $-1 \leq \rho(X, Y) \leq 1$. If $\rho = -1$, the **correlation is perfect and negative**. If $\rho = 1$, the **correlation is perfect and positive**.
2. Correlation coefficient is independent of change of origin and scale. That is, if $U = \frac{X-a}{h}$ and $V = \frac{Y-b}{k}$, then $\rho(U, V) = \rho(X, Y)$

Theorem: Two independent variables are uncorrelated.

Proof:

Consider $\sigma_{XY} = cov\ (X, Y) = E\big[\big(X - E(X)\big)\big(Y - E(X)\big)\big]$

$\Longrightarrow\ \sigma_{XY} = E(X, Y) - E(X). E(Y)$ ...... (1)

If X and Y are independent, then

$E(XY) = E(X). E(Y)$ .... (2)

From (1) and (2), if X and Y are independent, then $\rho(X, Y) = 0$

The converse need not be true. That is, uncorrelated variables need not be independent.

Example 1 : Let $X \sim N(0, 1)$ and $Y = X^2$. Then $E(X) = E(X^3) = 0$.

Solution: Consider $cov\ (X, Y) = E(XY) - E(X). E(Y) = E(X^3) - E(X). E(X^2)$

$$= 0 - 0 = 0$$

$\Longrightarrow cov\ (X, Y) = 0$ but X and Y are related by $Y = X^2$.

Thus, uncorrelated variables need not be independent.

Note: The converse is true if the joint distribution of (X, Y) is bivariate normal.

Example 2: The j.p.m.f of (X, Y) is given below:

$Y \backslash X$	-1	1
0	$\frac{1}{8}$	$\frac{3}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$

Find the correlation coefficient between X and Y

Solution : Computation of marginal p.m.fs

$Y \backslash X$	-1	1	$g(y)$
0	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{4}{8}$
1	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{4}{8}$
$p(x)$	$\frac{3}{8}$	$\frac{5}{8}$	1

We have

$E(X) = \sum x\, p(x) = (-1) \times \frac{3}{8} + 1 \times \frac{5}{8} = -\frac{3}{8} + \frac{5}{8} = \frac{2}{8} = \frac{1}{4},$

$E(X^2) = \sum x^2\, P(x) = (-1)^2 \frac{3}{8} + 1^2 \times \frac{5}{8} = \frac{3}{8} + \frac{5}{8} = 1$, then

$V(X) = E(X^2) - \left(E(X)\right)^2 = 1 - \left(\frac{1}{4}\right)^2 = 1 - \frac{1}{16} = \frac{15}{16}$

Similarly, $E(Y)=\sum y\, g(y)=0\times\frac{4}{8}+1\times\frac{4}{8}=\frac{4}{8}=\frac{1}{2}$

$E(Y^2)=\sum y^2\, g(y)=0^2\times\frac{4}{8}+1^2\times\frac{4}{8}=\frac{4}{8}=\frac{1}{2}$ and

$V(Y)=E(Y^2)-\left(E(Y)\right)^2=\frac{1}{2}-\left(\frac{1}{2}\right)^2=\frac{1}{2}-\frac{1}{4}=\frac{1}{4}$

Further, $E(XY)=0\times(-1)\times\frac{1}{8}+0\times1\times\frac{3}{8}+1\times(-1)\times\frac{2}{8}+1\times1\times\frac{2}{8}=0$

Thus, $cov\ (X,Y)=E(XY)-E(X)E(Y)$

$$0-\frac{1}{4}\times\frac{1}{2}=-\frac{1}{8}$$

$$\therefore \rho\ (X,Y)=\frac{cov\ (X,Y)}{\sqrt{V(X)}\sqrt{V(Y)}}=-\frac{\frac{1}{8}}{\sqrt{\frac{15}{16}\times\frac{1}{4}}}=-\frac{1}{\sqrt{15}}=-0.2582$$

Example 3: Two random variables X and Y have the joint probability density function

$$\boldsymbol{f(x,y)}=\begin{cases}\boldsymbol{2-x-y} & , \quad \boldsymbol{0}<x<1\ ,\boldsymbol{0}<y<1\\ \boldsymbol{0} & , \quad \boldsymbol{otherwise}\end{cases}$$

Find correlation coefficient between X and Y.

Solution: By symmetry in x and y we have $f_1(x)=f_2(y)$, $E(X)=E(Y)$ and $V(X)=V(Y)$

The m.p.d.f X is given by

$$f_1(x)=\int_0^1 f(x,y)\,dy=\int_0^1(2-x-y)\,dy=\frac{3}{2}-x$$

Thus, $f_1(x)=\begin{cases}\frac{3}{2}-x & , \quad if\ \ 0<x<1\\ 0 & , \quad otherwise\end{cases}$

Consider.

$$E(X) = \int_0^1 x f_1(x) dx = \int_0^1 x\left(\frac{3}{2} - x\right) dx = \int_0^1 \left(\frac{3}{2}x - x^2\right) dx = \frac{5}{12}$$

$$E(X^2) = \int_0^1 x f_1(x) = \int_0^1 x^2\left(\frac{3}{2} - x\right) dx = \int_0^1 \left(\frac{3}{2}x^2 - x^3\right) dx = \frac{1}{4}$$

Further,

$$E(XY) = \int_0^1 \int_0^1 xy\, f(x,y) dx\, dy = \int_0^1 \int_0^1 xy\,(2 - x - y) dx\, dy$$

$$= \int_0^1 y\left(\int_0^1 (2x - x^2 - xy) dx\right) dy = \int_0^1 y\left[2.\frac{x^2}{2} - \frac{x^3}{3} - \frac{yx^2}{2}\right]_0^1 dy$$

$$= \int_0^1 y\left(1 - \frac{1}{3} - \frac{y}{2}\right) dy = \int_0^1 y\left(\frac{2}{3} - \frac{y}{2}\right) dy$$

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

$\therefore E(XY) = \frac{1}{6}$

Thus, $V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{1}{4} - \left(\frac{5}{12}\right)^2 = \frac{1}{4} - \frac{25}{144} = \frac{36-25}{144} = \frac{11}{144}$

and $cov\,(X,Y) = E(XY) - E(X)E(Y) = \frac{1}{6} - \left(\frac{5}{12}\right)^2 = \frac{1}{6} - \frac{25}{144} = \frac{24-25}{144} = -\frac{1}{144}$

$\therefore$ The correlation coefficient is given by

$$\rho(X,Y) = \frac{cov\,(X,Y)}{\sqrt{V(x)}\,\sqrt{V(Y)}} = -\frac{\frac{1}{144}}{\sqrt{\frac{11}{144}}\sqrt{\frac{11}{144}}} = -\frac{\frac{1}{144}}{\frac{11}{144}} = -\frac{1}{11}$$

Bivariate Normal Distribution:

The bivaraite normal distribution is a generalization of a normal distribution for a single value.

Let X and Y be two normally correlated variables with correlation coefficient ρ. Let $E(X) = \mu_1,\ V(X) = \sigma_1^2\ , E(Y) = \mu_2$ and $V(Y) = \sigma_2^2$.

Definition: The bivariate continuous random variable (X, Y) is said to follow bivariate normal distribution with parameters $\mu_1, \mu_2, \sigma_1^2, \sigma_2^2$ and ρ if its p.d.f. is given by

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right\}\right];$$

$$-\infty < x, y, \mu_1, \mu_2 < \infty,\ \ \sigma_1 > 0, \sigma_2 > 0 \text{ and } -1 < \rho < 1.$$

Notation: $(X, Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Read as (X, Y) follows **bivariate normal distribution** with parameters $\mu_1, \mu_2\sigma_1^2, \sigma_2^2$ and ρ.

Note: The curve $z = f(x, y)$ which is the equation of a surface in three dimensions is called the **Normal correlation surface**.

Marginal p.d.fs of X and Y: The m.p.d.f of X is given by

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy$$

Let $v = \frac{y-\mu_2}{\sigma_2}$, then $y = \mu_2 + \sigma_2 v$ and $dy = \sigma_2 dv$

Therefore,

$$f_1(x) = \frac{\sigma_2}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\infty} exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\, v\left(\frac{x-\mu_1}{\sigma_1}\right) + v^2\right\}\right] dv$$

$$= \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] \int_{-\infty}^{\infty} exp\left[-\frac{1}{2(1-\rho^2)}\left\{v - \rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right\}^2\right] dv$$

Let $\frac{1}{\sqrt{1-\rho^2}}\left[v - \rho\left(\frac{x-\mu_1}{\sigma_1}\right)\right] = t$. Then $dv = \sqrt{1-\rho^2}\ dt$

$\therefore\ f_1(x) = \frac{1}{2\pi\sigma_1} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)\right] \int_{-\infty}^{\infty} exp\left(-\frac{t^2}{2}\right) dt$

$= \frac{1}{2\pi\sigma_1} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] . \sqrt{2\pi}$

$\Rightarrow\ f_1(x) = \frac{1}{\sigma_1\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right]$ for $-\infty < x < \infty$

Similarly, it can be shown that

$f_2(y) = \frac{1}{\sigma_2\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{y-\mu_1}{\sigma_2}\right)^2\right]$ for $-\infty < x < \infty$

Hence $X{\sim}N(\mu_1, \sigma_1^2)$ and $Y{\sim}N(\mu_2, \sigma_2^2)$.

Note: If $(X, Y){\sim}BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho\ )$, then $X{\sim}N(\mu_1, \sigma_1^2)$ and $Y{\sim}N(\mu_2, \sigma_2^2)$

Conditional p.d.fs of X and Y

The conditional probability density function (c.p.d.f.) of X for given Y is given by

$$f_{1|2}(x|y) = \frac{f(x,y)}{f_2(y)}$$

$= \frac{1}{\sigma_1\sqrt{2\pi}\,\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\left(1-(1-\rho^2)\right)\right\}\right]$

$= \frac{1}{\sigma_1\sqrt{2\pi}\,\sqrt{1-\rho^2}}\ exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1)^2 - 2\rho\frac{\sigma_1}{\sigma_2}(x-\mu_1)(y-\mu_2) + \frac{\sigma_1^2}{\sigma_2^2}\rho^2(y-\mu_2)^2\right\}\right]$

$= \frac{1}{\sigma_1\sqrt{2\pi}\,\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1) - \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]$

Therefore, $f_{1|2}(x|y) = \frac{1}{\sigma_1\sqrt{2\pi}\,\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)\sigma_1^2}\left\{(x-\mu_1) - \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)\right\}^2\right]$

which is the univariate normal distribution with mean

$E(X|Y = y) = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2)$ and

$V(X|Y = y) = \sigma_1^2(1-\rho^2)$

Thus, the c.p.d.f of X for fixed Y is given by

$$(X|Y=y) \sim N\left[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(y-\mu_2), \sigma_1^2(1-\rho^2)\right]$$

Similarly, the c.p.d.f of Y for fixed $X = x$ is given by

$$f_{2|1}(y|x) = \frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-\rho^2}} exp\left[\frac{-1}{2(1-\rho^2)\sigma_2^2}\left\{(y-\mu_2) - \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right\}^2\right], -\infty < y < \infty$$

Thus, the c.p.d.f of Y for fixed X is given by

$$(Y|X=x) \sim N\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2(1-\rho^2)\right]$$

Example 4: If $(\boldsymbol{X},\boldsymbol{Y}) \sim \boldsymbol{BN}\ (\mathbf{5},\mathbf{10},\mathbf{1},\mathbf{25},\boldsymbol{\rho})$ where $\boldsymbol{\rho} > 0$, find $\boldsymbol{\rho}$ when $\boldsymbol{P}(\mathbf{4} < Y < 16|X = 5) = \mathbf{0.954}$

Solution:

Here $\mu_1 = 5, \mu_2 = 10$, $\sigma_1^2 = 1, \sigma_2^2 = 25$. We know that $(Y|X=x) \sim N[\mu, \sigma^2]$

where $\mu = \mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)$ and $\sigma^2 = \sigma_2^2(1-\rho^2)$.

Here $\mu = 10 + \rho \times \frac{5}{1}(5-5) = 10$ and $\sigma^2 = 25(1-\rho^2)$

Thus $(Y|X=5) \sim N[10, 25(1-\rho^2)]$. We want to find ρ so that
$P(4 < Y < 16|X = 5) = 0.954$

Let $Z = \frac{Y-\mu}{\sigma} = \frac{Y-10}{5\sqrt{1-\rho^2}} \sim N(0,1) \Rightarrow P\left(\frac{4-10}{\sigma} < Z < \frac{16-10}{\sigma}\right) = 0.954$

$\Rightarrow P\left(-\frac{6}{\sigma} < Z < \frac{6}{\sigma}\right) = 0.954 \Rightarrow P\left(0 < Z < \frac{6}{\sigma}\right) = 0.477$

From standard normal table, we have $\frac{6}{\sigma} = 2 \;\; \Rightarrow \sigma = 3 \;\; \Rightarrow \sigma^2 = 9$

$$\Rightarrow 25(1-\rho^2) = 9 \Rightarrow 1-\rho^2 = \frac{9}{25} \Rightarrow \rho^2 = 1 - \frac{9}{25} = \frac{16}{25} \Rightarrow \rho = \frac{4}{5} = 0.8$$

Example 5: Find $cor(X,Y)$ for the jointly normal distribution

$$\boldsymbol{f(x,y) = \frac{1}{2\pi\sqrt{3}} exp\left[-\frac{\{(2x-y)^2+2xy\}}{6}\right]}, -\infty < x, y < \infty$$

Solution: Given $(X,Y) \sim BN(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Then its p.d.f. is given by

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{(x-\mu_1)^2}{\sigma_1^2} - 2\rho\frac{(x-\mu_1)(y-\mu_2)}{\sigma_1\sigma_2} + \frac{(y-\mu_2)^2}{\sigma_2^2}\right\}\right] --(1)$$

We have

$$f(x,y) = \frac{1}{2\pi\sqrt{3}} exp\left[-\frac{\{(2x-y)^2+2xy\}}{6}\right], -\infty < x, y < \infty$$

$$i.e, \quad f(x,y) = \frac{1}{2\pi\sqrt{3}} exp\left[-\frac{\{4x^2+y^2-2xy\}}{6}\right] ----(\mathbf{2})$$

Comparing (1) and (2), we get $\mu_1 = \mu_2 = 0$. Then (1) becomes

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} exp\left[-\frac{\left\{\frac{x^2}{\sigma_1{}^2} + \frac{y^2}{\sigma_2{}^2} - 2\rho\frac{xy}{\sigma_1\sigma_2}\right\}}{2(1-\rho^2)}\right] \text{—} (3)$$

Comparing (2) and (3), we find

$\sigma_1\sigma_2\sqrt{1-\rho^2} = \sqrt{3}$, $\sigma_1{}^2(1-\rho^2) = \frac{3}{4}$, $\sigma_2{}^2(1-\rho^2) = 3$ and $\frac{\rho}{\sigma_1\sigma_2(1-\rho^2)} = \frac{1}{3}$

On solving we get $\sigma_1{}^2 = 1$, $\sigma_2{}^2 = 4$, $\rho^2 = \frac{1}{4}$

Thus $cor(X,Y) = \rho = \sqrt{\frac{1}{4}} = \pm\frac{1}{2}$

Example 6: Determine the parameters of the bivariate normal distribution

$$\boldsymbol{f(x,y) = c\, exp\left[-\frac{\{16(x-2)^2 - 12(x-2)(y+3) + 9(y+3)^2\}}{216}\right]}$$

Solution: If $(X,Y) \sim BN(\mu_1, \mu_2, \sigma_1{}^2, \sigma_2{}^2, \rho)$, then

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}exp\left[-\frac{\left\{\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)^2+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}}{2(1-\rho^2)}\right]$$

Comparing these functions, we get

$$\mu_1=2,\mu_2=-3,c=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},\frac{16}{216}=\frac{1}{2(1-\rho^2){\sigma_1}^2}$$

$$\frac{9}{216}=\frac{1}{2(1-\rho^2){\sigma_2}^2},\frac{12}{216}=\frac{2\rho}{2\sigma_1\sigma_2(1-\rho^2)}$$

$\therefore (1-\rho^2){\sigma_1}^2=\frac{27}{4}, (1-\rho^2){\sigma_2}^2=12, \sigma_1\sigma_2(1-\rho^2)=18\rho$

$\Longrightarrow (1-\rho^2)^2{\sigma_1}^2{\sigma_2}^2=81=(18\rho)^2 \Longrightarrow \rho^2=\frac{1}{4},$

Further, $\sigma_1=3$ and $\sigma_2=4$.

Thus, $c=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}=\frac{1}{2\pi\times3\times4\sqrt{1-\frac{1}{4}}}=\frac{1}{12\pi\sqrt{3}}$

$\therefore (X,Y)\sim BN\left(2,3,9,16,\frac{1}{2}\right)$

Example 7: If $X\sim N(\mu,\sigma^2)$ and $(Y|x)\sim N(x,\sigma^2)$, show that

$$(X,Y)\sim BN(\mu,\mu,\sigma^2,2\sigma^2,\rho)\ .$$

Solution: We are given that

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right],-\infty<x<\infty$$

$$g(y|x)=\frac{1}{\sqrt{2\pi}\sigma}exp\left[-\frac{1}{2}\left(\frac{y-x}{\sigma}\right)^2\right],-\infty<y<\infty$$

$$\therefore h(x,y)=g(y|x)f(x)=\frac{1}{2\pi\sigma^2}exp\left[-\frac{1}{2}\left\{\left(\frac{x-\mu}{\sigma}\right)^2+\left(\frac{y-x}{\sigma}\right)^2\right\}\right]$$

Consider $\left(\frac{y-x}{\sigma}\right)^2=\left(\frac{y-\mu+\mu-x}{\sigma}\right)^2=\left(\frac{y-\mu}{\sigma}\right)^2+\left(\frac{x-\mu}{\sigma}\right)^2-2\left(\frac{x-\mu}{\sigma}\right)\left(\frac{y-\mu}{\sigma}\right)$

Thus, $h(x,y) = \frac{1}{2\pi\sigma^2} exp\left[-\frac{1}{2}\left\{2\left(\frac{x-\mu}{\sigma}\right)^2 + \left(\frac{y-\mu}{\sigma}\right)^2 - 2\left(\frac{x-\mu}{\sigma}\right)\left(\frac{y-\mu}{\sigma}\right)\right\}\right]$

The bivariate normal p.d.f. is given by

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right\}\right]$$

On comparing $h(x,y)$ with $f(x,y)$, we get

$$\sigma_1\sigma_2\sqrt{1-\rho^2} = \sigma^2 \quad , \quad {\sigma_1}^2(1-\rho^2) = \frac{1}{2}\sigma^2$$

$$\frac{\sigma_1\sigma_2(1-\rho^2)}{\rho} = \sigma^2 \quad , \quad {\sigma_2}^2(1-\rho^2) = \sigma^2 \,, \quad \mu_1 = \mu_2 = \mu$$

On solving, we get $\rho^2 = \frac{1}{2}$ ${\sigma_2}^2 = 2\sigma^2$, ${\sigma_1}^2 = \sigma^2$.

Thus, $(X,Y) \sim BN\left(\mu, \mu, \sigma^2, 2\sigma^2, \frac{1}{\sqrt{2}}\right)$

Example 8: The variables X and Y are connected by the equation $aX + bY + c = 0$. Show that the correlation between them is -1 if signs of a and b are same and $+1$ if they are different signs.

Solution: Given $aX + bY + c = 0 \Longrightarrow aE(X) + bE(Y) + c = 0$

$\therefore a[X - E(X)] + b[Y - E(Y)] = 0 \Longrightarrow [X - E(X)] = -\frac{b}{a}[Y - E(Y)]$

$\therefore cov(X,Y) = E[\{X - E(X)\}\{Y - E(Y)\}] = -\frac{b}{a}E\big(Y - E(Y)\big)^2 = -\frac{b}{a}{\sigma_Y}^2$ and

${\sigma_X}^2 = E\big(X - E(X)\big)^2 = \frac{b^2}{a^2}E\big(Y - E(Y)\big)^2 = \frac{b^2}{a^2}{\sigma_Y}^2$

$$\therefore \rho = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y} = \frac{-\frac{b}{a}{\sigma_Y}^2}{\sqrt{{\sigma_Y}^2}\sqrt{\frac{b^2}{a^2}{\sigma_Y}^2}} = \frac{-\frac{b}{a}{\sigma_Y}^2}{\left|\frac{b}{a}\right|{\sigma_Y}^2} = \frac{-\frac{b}{a}}{\left|\frac{b}{a}\right|}$$

$$\therefore \rho = \frac{cov(X,Y)}{\sigma_X \cdot \sigma_Y} = \begin{cases} 1 \,, & if\ a\ and\ b\ have\ opposite\ signs \\ -1 \,, & if\ a\ and\ b\ have\ same\ signs \end{cases}$$