$$f_{x_1 x_2 x_3 | x_4 x_5}(x_1, x_2, x_3)$$

$$= \frac{f_{x_1 x_2 x_3 x_4 x_5}(x_1, x_2, x_3, x_4, x_5)}{f_{x_4 x_5}(x_4, x_5)}$$

Independent if and only if,

$$f_{x_1 x_2 \cdots x_p}(x_1, x_2, \cdots, x_p) = f_{x_1}(x_1) f_{x_2}(x_2) \cdots f_{x_p}(x_p).$$

COVARIANCE AND CORRELATION

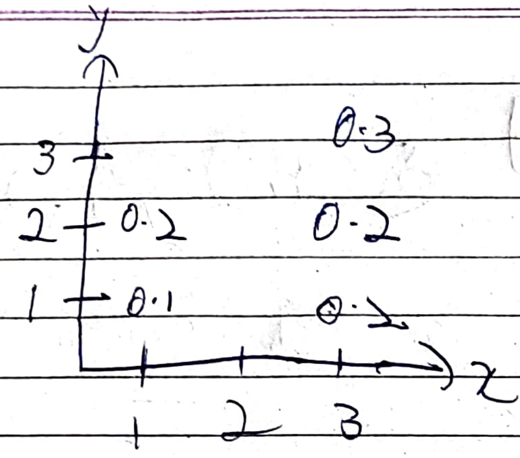
A common measure of the relationship b/w the ~~variables~~ two random variables is covariance. It is usually linear

discrete

$$E[h(X,Y)] = \begin{cases} \sum\sum_R h(x,y) f_{xy}(x,y) \\ \iint_R h(x,y) f_{xy}(x,y)\,dx\,dy \end{cases}$$

X, Y continuous

Q)



Calculate $E[(X-\mu_x)(Y-\mu_y)]$

$E(X) = x f(x)$; so

$E[(X-\mu_x)(Y-\mu_y)] = (x-\mu_x)(y-\mu_y) f_{xy}(x,y)$

$\mu_x = \sum_R x f_{xy}(x,y)$

$= (1\times 0.1 + 1\times 0.2) + (3\times 0.2 + 3\times 0.2 + 3\times 0.3)$

$= 2.4$

$\mu_y = \sum_R y f_{xy}(x,y)$

$= (1\times 0.1 + 1\times 0.2) + (2\times 0.2 + 2\times 0.2) + 3\times 0.3$

$= 2.0$

$E[(X-\mu_x)(Y-\mu_y)] = \sum(x-\mu_x)(y-\mu_x)\, f_{xy}(x,y)$

(for each point in X, Y)

$\Rightarrow (1-2.4)(1-2.0)\times 0.1$
$+(1-2.4)(2-2.0)\times 0.2 + (3-2.4)(1-2.0)$
$\times 0.2 + (3-2.4)(2-2.0)\times 0.2 + (3-2.4)$
$(3-2.0)\times 0.3 = 0.2$

Covariance denoted by cov(X, Y) or σ_{xy} b/w random variable X and Y is,

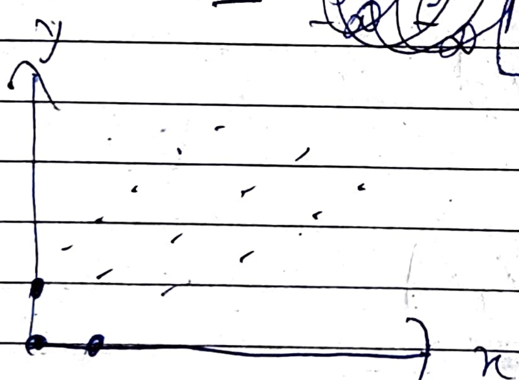
$\sigma_{XY} = E[(X-\mu_x)(Y-\mu_y)] = E(XY)$
$- \mu_x \mu_y$

For continous,

$$E[(Y-\mu_y)(X-\mu_x)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_x)(y-\mu_y)\, f_{xy}(x,y)\, dx\, dy$$

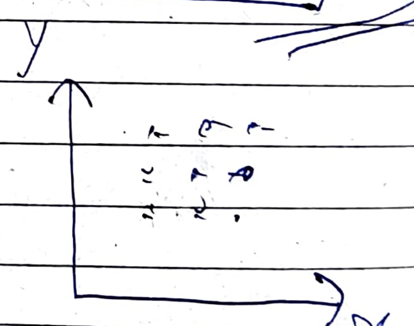
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[xy-\mu_x y - x\mu_y+\mu_x\mu_y]\, dx\, dy$$

$$= \boxed{E(XY) - \mu_x\mu_y}$$



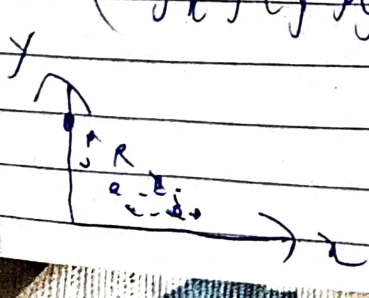
positive covariance

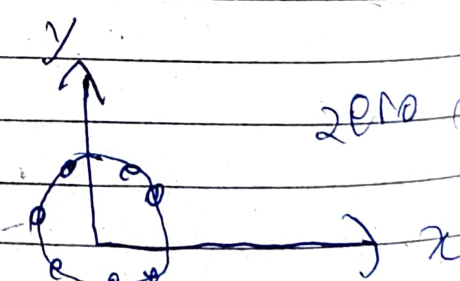
$(x-\mu_x)(y-\mu_y) > 0$



zero covariance

negative covariance





zero covariance

Q) In transmission 4 bits, where X is acceptable and Y is suspect bits.
What is type of covariance ?

$X + Y \leq 4$

If $X = 4$, $Y = 0$,

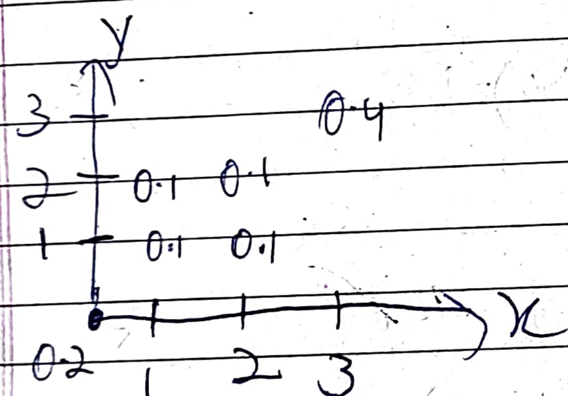
hence, negative covariance.

CORRELATION b/w random variable X and Y is denoted as ρ_{XY}.

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

Since $\sigma_x > 0$, $\sigma_y > 0$, the sign of correlation (+, -ve or zero) depends on covariance.

Also, $-1 \leq \rho_{XY} \leq 1$ +

Q) 

Calculate σ_{XY} and ρ_{XY}

$$E(XY) = \sum xy \, f_{xy}(x,y)$$
$$= 0 \times 0 \times 0.2 + 1 \times 1 \times 0.1 + \ldots \; 3 \times 3 \times 0.4 = 4.5$$

$$E(X) = \sum x f_{xy}(x,y)$$
$$= 0 \times 0.2 + 1 \times 0.2 + 2 \times 0.2 + 3 \times 0.4 = 1.8$$

$$V(X) = (0-1.8)^2 \times 0.2 + (1-1.8)^2 \times 0.2$$
$$+ (2-1.8)^2 \times 0.2 + (3-1.8)^2 \times 0.4$$
$$= \underline{1.36}$$

Marginal Distribution of Y is same as for X, $E(Y) = 1.8$ and $V(Y) = 1.36$,

Consequently,

$$\sigma_{XY} = E(XY) - E(X)(EY)$$
$$(\mu_x)(\mu_y)$$

$$= 4.5 - (1.8)(1.8) = 1.26$$

$$\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{1.26}{\sqrt{1.36}\sqrt{1.36}} = \underline{0.926}$$

If X and Y are independent random variables,

$$\sigma_{XY} = \rho_{XY} = 0$$

BIVARIATE NORMAL DISTRIBUTION

PDF is

$$f_{xy}(x, y; \sigma_x, \sigma_y, \mu_x, \mu_y, \rho)$$

$$= \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(x-\mu_x)^2}{\sigma_x^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right]\right\}$$

for $-\infty < x, y < \infty$, with parameters $\sigma_x, \sigma_y > 0$, $-\infty < \mu_x < \infty$, $-\infty < \mu_y < \infty$, and $-1 < \rho < 1$

It also integrates to 1. Also, the bivariate normal PDF is positive over the entire plane of real numbers

The Bivariate Normal PDF,

$$f_{xy}(x,y) = \frac{1}{\sqrt{2\pi}} e^{-0.5(x^2+y^2)}$$

is a special case with $\sigma_x = 1$, $\sigma_y = 1$, $\mu_x = 0$, $\mu_y = 0$ and $\rho = 0$

If X and Y have a bivariate normal distribution with joint PDF $f_{xy}(x, y; \sigma_x, \sigma_y, \mu_x, \mu_y, \rho)$, the marginal probability distributions of X and Y are normal with means μ_x and μ_y and standard deviation σ_x and σ_y, respectively.
The correlation b/w X and Y is ρ. so, if $\rho = 0$, X and Y are independent.

LINEAR COMBINATIONS

Given random variables $X_1, X_2, \dots X_p$ and constants $c_1, c_2, \dots c_p$

$$Y = c_1 X_1 + c_2 X_2 + c_p X_p$$

is a linear combination of $X_1, X_2, \dots X_p$