Mean and Coveriance

For r.v.s X and Y with j.d.f. $f(x,y)$ and $g : \mathbb{R}^2 \to \mathbb{R}$ be any continuous function

$$E(g(x,y)) = \sum_{y} \sum_{x} g(x,y)\, f(x,y)$$

The covariance, denoted by $cov(X,Y)$ is defined by

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

where $\mu_X = E(X) = \sum_{x} x\, p(x)$, $\mu_Y = E(Y) = \sum_{y} y\, p(y)$

marginal distribution function of X

Theorem For random variable X and Y

$$\sigma_{XY} = E(XY) - E(X)E(Y)$$

proof Let $f(x,y)$ be the joint distribution of X and Y.

Then
$$\sigma_{XY} = E(XY) - E(X)E(Y)$$
$$= \sum_x \sum_y (x-\mu_X)(y-\mu_Y) f(x,y)$$
$$= \sum_x \sum_y (xy - \mu_Y x - \mu_X y + \mu_X \mu_Y) f(x,y)$$
$$= \sum_x \sum_y xy f(x,y) - \mu_Y \sum_x \sum_y x f(x,y)$$
$$- \mu_X \sum_x \sum_y y f(x,y) + \mu_X \mu_Y \sum_x \sum_y f(x,y)$$
$$= E(XY) - \mu_Y \sum_x x \underbrace{p(x)}_{\text{marginal distribution function of } X} - \mu_X \sum_y y \underbrace{p(y)}_{\text{m.d.f of } y} + \mu_X \mu_Y$$
$$= E(XY) - \mu_Y \mu_X - \mu_X \mu_Y + \mu_{XY}$$
$$= E(XY) - \mu_X \mu_Y$$
$$= E(XY) - E(X)E(Y)$$

For two random variables X and Y

- Correlation co-efficient $\rho_{XY} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$ $\boxed{-1 \le \rho_{XY} \le 1}$

$\sigma_X \sigma_Y \rightarrow$ Standard deviations

equivalently

- For discrete r.v. X and Y

$$\rho_{XY} = \frac{n(\sum x_i y_i) - (\sum x_i)(\sum y_i)}{\sqrt{\left(n\sum x_i^2 - (\sum x_i)^2\right)\left(n\sum y_i^2 - (\sum y_i)^2\right)}}$$

Def Two r.v.s X and Y are said to be uncorrelated if $\rho_{XY} = 0$

Steps to compute covariance and correlation coefficient

1) marginal distributions of X and Y
2) Expected value of XY

Example X, Y are r.v.s with j.d.f. given by

X\Y	-1	0	1	total
-1	0	0.1	0.1	0.2
0	0.2	0.2	0.2	0.6
1	0	0.1	0.1	0.2
total	0.2	0.4	0.4	

Sol^n Step 1 marginal distribution of X and Y

x	-1	0	1
$p(x)$	0.2	0.6	0.2

x	-1	0	1
$p(y)$	0.2	0.4	0.4

Step 2

$$E(X) = \sum x\,p(x)$$
$$= (-1)(0.2) + (0)(0.6) + (1)(0.2)$$
$$= 0$$

$$E(Y) = \sum y\,p(y)$$
$$= (-1)(0.2) + (0)(0.4) + (1)(0.4)$$
$$= 0.2$$

$$E(XY) = \sum\sum xy\,p(x,y)$$
$$= (-1)(-1)(0) + (-1)(0)(0.1) + \cdots + (1)(1)(0.1)$$
$$= 0$$

thus $cov(X,Y) = E(XY) - E(X)E(Y)$
$$= 0 - 0\cdot(0.2)$$
$$= 0$$

and $\rho = \frac{cov(X,Y)}{\ } = 0$

$$\rho_{XY} = \frac{\quad}{\sigma_X \sigma_Y} = 0$$

Hence X and Y are uncorrelated.