(P1) For the Mobile Response time problem,

Conditional Probability distributions of Y given X = x

$P_{Y|X}(y|x)$.

y = Response time (nearest second)	x = Number of Bars of Signal Strength: x = 1	x = 2	x = 3
y = 4	0.750	0.400	0.091
y = 3	0.100	0.400	0.091
y = 2	0.100	0.120	0.364
y = 1	0.050	0.080	0.454

Conditional mean of Y given X = 1.

$$\mu_{Y|1} = E(Y|X=1) = \sum_{y \in R_Y} y \cdot P_{Y|X}(y|1)$$

$$= \sum_{y=1}^{4} y \cdot P_{Y|X}(y|1)$$

$$= 1 \cdot P_{Y|X}(1|1) + 2 \cdot P_{Y|X}(2|1) + 3 \cdot P_{Y|X}(3|1) + 4 \cdot P_{Y|X}(4|1)$$

$$= 1 \cdot (0.050) + 2 \cdot (0.100) + 3 \cdot (0.100) + 4 \cdot (0.750)$$

$$= 3.55$$

The conditional mean is interpreted as the expected response time given that one bar of signal is present.

Conditional variance of Y given X = 1 is

$$Var(Y|X=1) = \sum_{y \in R_Y} (y - \mu_{Y|1})^2 P_{Y|X}(y|1)$$

$$= (1-3.55)^2 (0.05) + (2-3.55)^2 0.1 + (3-3.55)^2 (0.1)$$

$$+ (4-3.55)^2 (0.75) = 0.748.$$

(P2). From the previous Problem, the conditional PDF of X given $Y=y$ is

$$f_{X|Y}(x|y) = \frac{f_{X,Y}(x,y)}{f_Y(y)} = \begin{cases} \frac{1}{1-y}, & y \le x \le 1 \\ 0, & \text{otherwise} \end{cases}$$

Find the conditional mean of X given $Y=y$.

$$E(X|Y=y) = \int_{-\infty}^{\infty} x . f_{X|Y}(x|y)\,dx = \int_{x=y}^{x=1} x . \frac{1}{1-y}\,dx$$

$$= \frac{1}{1-y}\left[\frac{x^2}{2}\right]_{x=y}^{x=1}$$

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

Independent r.vs:

Recap: We say two events A & B are independent if $P(A \cap B) = P(A) . P(B)$.

Apply this idea to random variables.

We say r.vs X & Y are independent iff the events $\{X=x\}$ and $\{Y=y\}$ are independent.

Defn: Random variables X & Y are independent

if $P_{X,Y}(x,y) = P_X(x) P_Y(y)$ [Discrete Case]

$f_{X,Y}(x,y) = f_X(x) f_Y(y)$ [Continuous Case].

Remark:

In Problems, to show independence of X & Y, we have to check whether the joint PMF factors into the Product of Marginal pmfs for all pairs x and y.

To prove dependence, we only need to present one counter example, a pair (x, y) with $P_{X,Y}(x,y) \neq P_X(x) \cdot P_Y(y)$.

Example:

A Program consists of two modules. The number of errors, X, in the first module and the number of errors, Y, in the second module have the joint distribution,

$P_{X,Y}(0,0) = P_{X,Y}(0,1) = P_{X,Y}(1,0) = 0.2$, $P_{X,Y}(1,1) = P_{X,Y}(1,2) = P_{X,Y}(1,3) = 0.1$

$P_{X,Y}(0,2) = P_{X,Y}(0,3) = 0.05$.

Find (a). the Marginal distributions of X & Y.
(b) the Probability of no errors in the first module
(C). the distribution of the total number of errors in the program.
Also (d) find out if the errors in the two modules occur independently.

Solution:

It is convenient to organize the joint pmf of X & Y in a table. Adding rowwise and columnwise, we get the Marginal pmfs.

$P_{X,Y}(x,y)$		$y=0$	1	2	3	$P_X(x)$
x	0	0.20	0.20	0.05	0.05	0.50
	1	0.20	0.10	0.10	0.10	0.50
$P_Y(y)$		0.40	0.30	0.15	0.15	1.00

This solves (a).. (b). $P_X(0) = 0.50$

(c). Let $Z = X+Y$. be the total number of errors

To find the distribution of Z, we first identify the possible values, then find the probability of each value

Range of $Z = \{0, 1, 2, 3, 4\}$

Then $P_Z(0) = P(X+Y=0) = P_{X,Y}(0,0) = 0.20$ $\quad P\{X=0 \cap Y=0\}$

$P_Z(1) = P\{X=0 \cap Y=1\} + P\{X=1 \cap Y=0\}$

$= P_{X,Y}(0,1) + P_{X,Y}(1,0) = 0.20 + 0.20 = 0.40$

$P_Z(2) = P_{X,Y}(0,2) + P_{X,Y}(1,1) = 0.05 + 0.10 = 0.15$

$P_Z(3) = P_{X,Y}(0,3) + P_{X,Y}(1,2) = 0.05 + 0.10 = 0.15$

$P_Z(4) = P_{X,Y}(1,3) = 0.10$

It is a good check to verify that $\sum_{z \in R_Z} P_Z(z) = 1$

(d) To verify independence of X & Y, check if their joint pmf factors into a product of marginal pmfs.

We see that $P_{X,Y}(0,0) = 0.2 = P_X(0) P_Y(0) = (0.5)(0.4)$

Keep checking,... Next $P_{X,Y}(0,1) = 0.2$ whereas $P_X(0) P_Y(1) = (0.5)(0.3) = 0.15$

We found a pair of x and y that violates the formula for independent r.vs.

Therefore, the numbers of errors in two modules are <u>dependent</u>!

Example: Given joint PDF of X, Y is

$$f_{X,Y}(x,y) = \begin{cases} 4xy, & 0 \le x \le 1,\ 0 \le y \le 1 \\ 0, & \text{Otherwise} \end{cases}$$

Are X & Y independent?

First find the Marginal PDFs of X & Y.

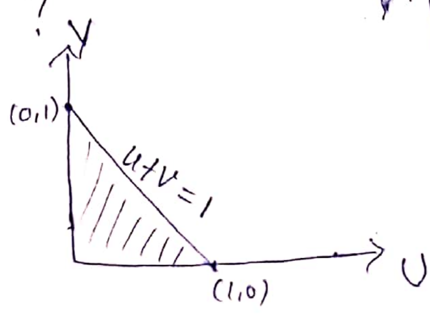
$$f_X(x) = \int_{y=0}^{y=1} f_{X,Y}(x,y)\,dy = \begin{cases} 2x, & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

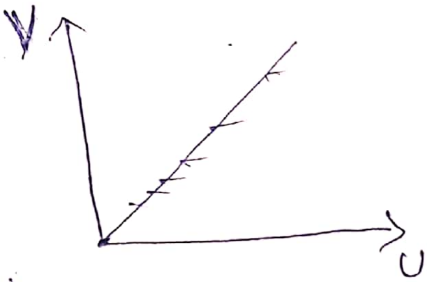
$$f_Y(y) = \int_{x=0}^{x=1} f_{X,Y}(x,y)\,dx = \begin{cases} 2y, & 0 \le x \le 1 \\ 0, & \text{otherwise.} \end{cases}$$

It is easily verified that $f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y)$ for all pairs (x,y). So we conclude that X, Y are independent r.Vs.

(Pi) Given joint PDF of U, V is $f_{U,V}(u,v) = \begin{cases} 24uv, & u \ge 0,\ v \ge 0,\ u+v \le 1 \\ 0, & \text{Otherwise} \end{cases}$

Are U & V independent?





$$f_U(u) = \begin{cases} 12u(1-u)^2, & 0 \le u \le 1 \\ 0, & \text{Otherwise} \end{cases}$$

$$f_V(v) = \begin{cases} 12(v)(1-v)^2, & 0 \le v \le 1 \\ 0, & \text{Otherwise} \end{cases}$$

Clearly U & V are not independent.

Covariance and correlation:

Expectation, variance, and standard deviation characterize the distribution of a single r.v. Now we introduce measures of association of two r.vs.

Defn: Covariance - summarizes interrelation of two r vs X, Y

The covariance of two r.vs X, Y is

$$\sigma_{XY} = Cov[X,Y] = E((X-\mu_X)(Y-\mu_Y))$$

$$= \begin{cases} \sum_{x \in R_X} \sum_{y \in R_Y} (x-\mu_x)(y-\mu_y) P_{X,Y}(x,y) & \left(\begin{array}{c} X,Y \text{ are} \\ \text{Discrete} \end{array}\right) \\ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_x)(y-\mu_y) f_{X,Y}(x,y)\,dx\,dy & (\text{Continuous}) \end{cases}$$

$$\sigma_{XY} = Cov[X,Y] = E(XY) - E(X)E(Y) = E(XY) - \mu_X \mu_Y.$$

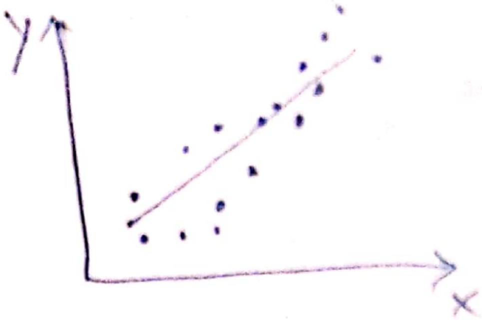
If $Cov[X,Y] > 0$, then positive deviations $(X-\mu_X)$ are more likely to be multiplied by ~~negative~~ Positive $(Y-\mu_Y)$, and negative $(X-\mu_X)$ are more likely to be multiplied by negative $(Y-\mu_Y)$.

In short, Large X imply Large Y, and small X imply small Y.

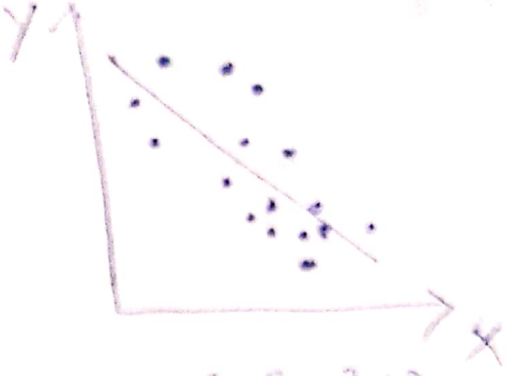
These r vs are positively correlated.

Conversely, $Cov[X,Y] < 0$, means that large X generally correspond to small Y. and small X correspond to large Y. These variables are negatively correlated.

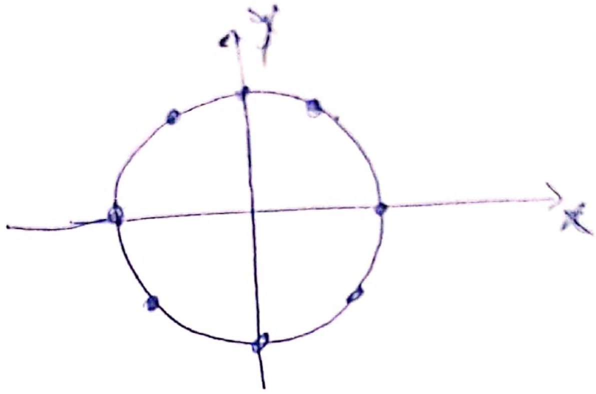
If $Cov[X,Y] = 0$, we say that X & Y are uncorrelated.



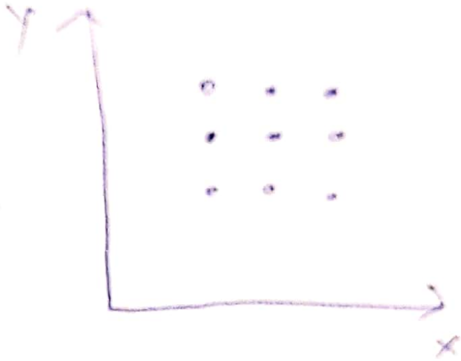
$Cov[X,Y] > 0$



$Cov[X,Y] < 0$



$Cov[X,Y] = 0$



$Cov[X,Y] = 0$.

Covariance is a measure of Linear relationship between the r.vs. If the relationship between the r.vs is nonlinear, then covariance is zero. [See points on the circle, There is identifiable relationship between the variables. Still the covariance is zero.]

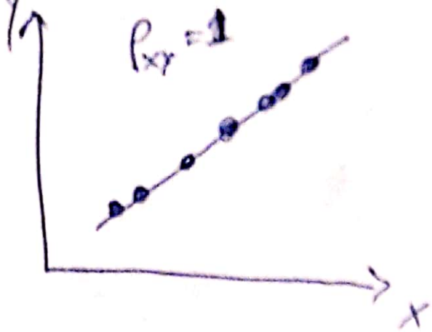
Defn: Correlation (Coefficient) between r.vs X & Y is defined as

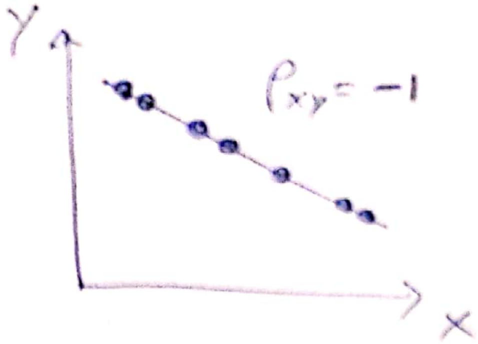
$$\rho_{XY} = \frac{Cov[X,Y]}{\sigma_X \sigma_Y} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$$

[It is free of the units of X and Y.]

- Correlation coefficient is a rescaled, normalized covariance.
- ρ_{XY} satisfies $\boxed{-1 \leq \rho_{XY} \leq 1}$
- where $|\rho_{XY}| = 1$ is possible only when all values of X & Y lie on a straight line (see Fig in the next page)

Scanned with CamScanner





Perfect correlation: $\rho_{xy} = \pm 1$

- Further, values of ρ near 1 indicate strong positive correlation.
- Values near (-1) show strong negative correlation
- Values near 0 show weak correlation or no correlation.

Result:

(i) $Cov[X, Y] = E(XY) - E(X)E(Y)$

(ii) $Var[X+Y] = Var(X) + Var[Y] + 2\,cov[X, Y]$

(iii) $Cov[X, X] = Var(X)$

Theorem: For independent r.vs X & Y.

a) $E(g(X)h(Y)) = E(g(X))E(h(Y))$

b) $E(XY) = E(X)E(Y)$

c) $Cov[X, Y] = 0$

d) $Var[X+Y] = Var[X] + Var[Y]$

e) $E(X \mid Y=y) = E(X)$ for all $y \in R_Y$.

Proof:

a) Since X, Y are independent r.vs $\Rightarrow P_{X,Y}(x,y) = P_X(x)\,P_Y(y)$

$$\Rightarrow E(g(X)h(Y)) = \sum_{x \in R_X} \sum_{y \in R_Y} g(x)\,h(y)\,P_X(x)\,P_Y(y)$$

$$= \sum_{x \in R_X} g(x)\,P_X(x) \cdot \sum_{y \in R_Y} h(y)\,P_Y(y) = E(g(X))\,E(h(Y))$$

(b) Let $g(x) = x$ & $h(y) = y$ in (a). We get the result (b)

(C). From (b), $E(XY) = E(X)E(Y)$

$$\text{COV}[X,Y] = E(XY) - E(X)E(Y) = 0.$$

(d) Follows from result (c) & $\text{Var}[X+Y] = \text{Var}[X] + \text{Var}[Y] + 2\text{Cov}[X,Y]$.

(e) Since $P_{X|Y}(x|y) = P_X(x) \Rightarrow E(X|Y=y) = \sum_{x \in R_X} x \cdot P_X(x) = E(X)$

Remark:

X, Y are independent r.Vs $\Longrightarrow$ $\text{Cov}[X,Y] = 0$

$\not\Longleftarrow$

In general, the converse is NOT true.

~~Therefore~~,