Joint Distribution, Joint Moments

04.04.2020

* Recall:

Experiments that produce a collection of RVs.

Suppose if X and Y are RVs, then we defined

joint CDF as $F_{X,Y}(x,y) = P(X \leq x, Y \leq y)$

joint PDF $f_{X,Y}(x,y)$

joint PMF $= P_{X,Y}(x,y)$

* Joint Moments.

* $E[X^j Y^k]$

$j=1,\ k=0 \Rightarrow E[X^1 Y^0] = E[X]$.

$j=0,\ k=1 \Rightarrow E[X^0 Y^1] = E[Y]$

The individual expectations of X and Y respectively

* $E[X^j Y^k]$:

* $E[X^j Y^k]$ $j = k = 1$

$= E[XY]$

for X, Y Discrete

$$E[XY] = \sum_{x} \sum_{y} xy \, p_{X,Y}(x,y)$$

$E[XY]$ = CORRELATION.

* Similar to variance, $Var[X] = E[(X-\mu_X)^2]$,
we have

$$Cov[X,Y] = E[(X-\mu_X)(Y-\mu_Y)]$$

$\mu_X = E[X]$, $\mu_Y = E[Y]$

* $Var[X] \rightarrow$ Spread of the RV X about its mean

* $Cov[XY]$ -

Let X be the marks scored by Students in Math

Let Y be the marks scored in Physics.

Q: Can we say that a student who scores well in math also scores well in physics?

Ex: Let X be the fuel price
Let Y be the vegetable price

Can we say that if the fuel price increases, the veg. cost increases too?

What we look at is how both RVs jointly behave?

A measure of that is $E[XY]$, $E[(X-\mu_x)(Y-\mu_y)]$

Suppose X : Weight in Kgs
Y : Height in Cms.

What is the unit of measurement of $E[XY]$ & $Cov[XY]$?

Suppose X: voltage

$$Var(X) = E[X^2] = \int_x x^2 f_X(x)\,dx.$$
If $E[X] = 0$.

$$Var[X] = \int_x (x-\mu_x)^2 f_X(x)\,dx. \quad \text{Volt}^2.$$

X : voltage
$E[X] \rightarrow$ Average voltage - Volts.

Similarly what is the unit for

$E[XY]$, where X: Weight in Kgs
Y: Height in Cm

$E[XY]$ Has a unit Kg Cm
$\text{Cov}(XY)$ has unit kg Cm.

(i) Suppose I change the units of measurement from kg to pounds & Cm to inches.

P: Weight in Pounds
C: Height in Inches

P & X
C & Y

1 kg = 2.2 pounds
1 cm = 0.4 inches.

$P = (\;)X \leftarrow$ in kgs
$C = (\;)Y. \leftarrow$ in Cms.

$P \to \alpha X$: α Scaling

$C \to \beta Y$.

$$E[PC] = E[\alpha X \beta Y]$$

$$= \alpha \beta E[XY] \leftarrow$$

When we measure in kg cm, correln was some number.

When we measure in pounds inches, Correlation changes.

$$E[PC] = \text{Scaled } E[XY]$$

Cov: $E[(X-\mu_X)(Y-\mu_Y)]$

kg cm $\to$ pounds inches.

$$Cov(XY) = E[(X-\mu_X)(Y-\mu_X)] = E(XY - X\mu_X - Y\mu_Y + \mu_X\mu_Y)$$

Quantity $\underline{\text{changes}}$ on the units of
depends measurements

X : Weight in kgs.
Y : Weight in kgs

$$E[XY] - \mu_Y E[X] - \mu_X E[Y] + \mu_X\mu_Y$$

1 kg = 1000g

$$\Downarrow$$

W $\to$ Weight in grams
Z $\to$ Weight in grams

$$E[XY] - 2\mu_X\mu_Y + \mu_X\mu_Y$$

$$E[XY] = k\, E[WZ].$$

$$\Downarrow$$

( ) 10^6 . $Cov(XY) = E(XY) - \mu_X\mu_Y$

look at measurement indep of units.

$$\frac{Cov(X,Y)}{\sigma_X \; \sigma_Y} = \frac{Cov(X,Y)}{\sqrt{Var X}\sqrt{Var Y}}$$

$$= \frac{kg\,cm}{kg \cdot cm} = \text{Dimensionless quantity.}$$

$\Rightarrow$ Correlation Coeff t $\rightarrow \rho_{X,Y}$

$$-1 \leq \rho_{X,Y} \leq 1$$

$$-1 \leq \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var\,Y}} \leq 1$$

If $\rho_{X,Y} = -1 \Rightarrow$ If X $\uparrow$es, Y $\downarrow$es

X $\downarrow$es, Y $\uparrow$es.

$\rho_{X,Y} = 1 \Rightarrow$ If X $\uparrow$es, Y $\uparrow$es

Direct relation

$\rho_{X,Y} = 0 \Rightarrow$ X & Y are on their own

$\rightarrow$ ORTHOGONAL

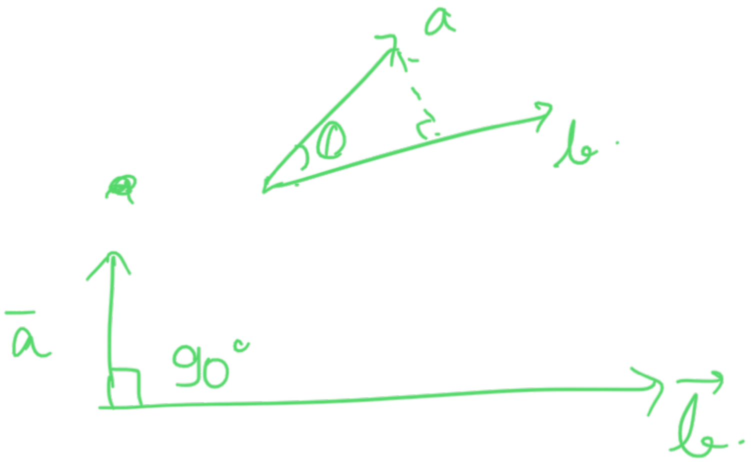
$$-1 \leq \cos\theta \leq 1$$

Suppose a & b are 2 vectors with angle θ between them

$$-1 \leq \frac{a \cdot b}{|a|\,|b|} \leq 1 \qquad (*)$$

$$-1 \leq \frac{Cov(X,Y)}{\sigma_X \sigma_Y} \leq 1 \qquad (**)$$

Cos θ measures Correlation.



Why should $\rho_{X,Y}$ be bounded b/w -1 & 1?

Proof: Let σ_X^2 & σ_Y^2 be Var X & Var(Y) resp.

Let W be a RV defined as

$W = X - aY$, where a is a const!

$$\mathrm{Var}[W] = E\left[W - E[W]\right]$$

$$= E\left[(X - aY)^2\right] - \left(E\left[(X - aY)\right]\right)^2$$

$$\Downarrow$$

$$\mathrm{Var}[W] = E\left[X^2 - 2aXY + a^2Y^2\right] \qquad (\mu_x - a\mu_y)^2$$

$$= E\left[X^2 - 2aXY + a^2Y^2\right] - \left(\mu_x^2 - 2a\mu_x\mu_y + a^2\mu_y^2\right)$$

$$\Rightarrow \left(\underbrace{E[X^2] - \mu_x^2}\right) + a^2 E\left(Y^2 - \mu_y^2\right)$$

$$- 2a\left(\underbrace{E(XY) - \mu_x\mu_y}\right)$$

$$= \mathrm{Var}(X) + a^2\,\mathrm{Var}(Y) - 2a\,\mathrm{Cov}(X,Y).$$

$\mathrm{Var}[W] \geq 0$ for any a.

Since $\mathrm{Var}(W)$ is a squared quantity

$\mathrm{Var}(W) \geq 0$ for any a

$$\Rightarrow \mathrm{Var}(X) + a^2\,\mathrm{Var}(Y) - 2a\,\mathrm{Cov}(X,Y) \geq 0$$

$$\Rightarrow 2a\,\mathrm{Cov}(X,Y) \leq \mathrm{Var}(X) + a^2\,\mathrm{Var}(Y).$$

Choose $a = \dfrac{\sigma_x}{\sigma_y}$

$$2\,\mathrm{Cov}(x,y) \le \frac{1}{a}\left(\mathrm{Var}(x) + a^2\,\mathrm{Var}(y)\right)$$

$$2\,\mathrm{cov}(x,y) \le \frac{\sigma_y}{\sigma_x}\left(\sigma_x^2 + \left(\frac{\sigma_x}{\sigma_y}\right)^2 \sigma_y^2\right)$$

$$\le \frac{\sigma_y \cdot \sigma_x^{\cancel{2}}}{\cancel{\sigma_x}} + \frac{\sigma_x^{\cancel{2}}}{\cancel{\sigma_y^2}\,\cancel{\sigma_x}} \cdot \cancel{\sigma_y^2}\,\sigma_y$$

$$2\,\mathrm{cov}(x,y) \le \underline{2\,\sigma_x\,\sigma_y}$$

$$\Rightarrow \mathrm{cov}(x,y) \le \sigma_x\,\sigma_y.$$

$$\Rightarrow \frac{\mathrm{cov}(x,y)}{\sigma_x\,\sigma_y} \le 1. \quad = \rho_{xy} \le 1. \qquad (1)$$

Suppose $a = -\dfrac{\sigma_x}{\sigma_y}$ $\qquad$ $\mathrm{Cov}(x,y) = E\left[(X - \mu_x)(Y - \mu_y)\right]$

$$\Rightarrow \mathrm{cov}(x,y) \ge -\sigma_y\,\sigma_x$$

$$\Rightarrow -1 \le \frac{\mathrm{cov}(x,y)}{\sigma_x\,\sigma_y} \qquad (2)$$

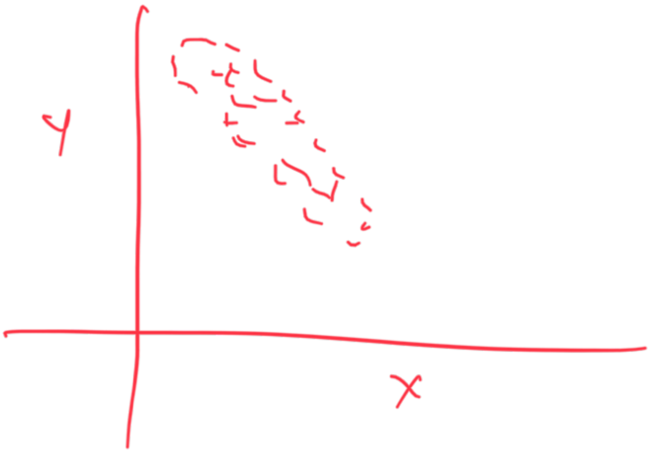
Clubbing (1) & (2) we get

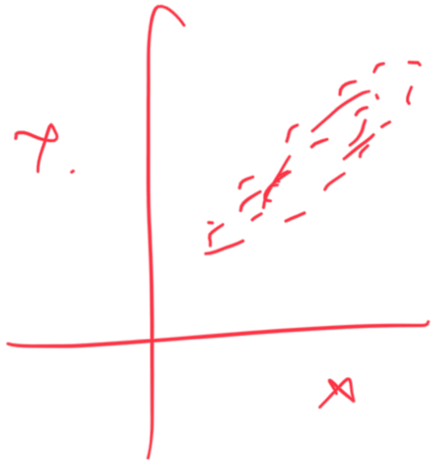
$$-1 \le \rho_{xy} \le 1.$$

If $\rho_{x,y} > 0 \rightarrow$ Positive Correlated

$\rho_{xy} < 0 \rightarrow$ Negative Correlation

$\rho_{xy} = 0 \rightarrow$ Zero Correlation.





<u>NOTE</u>: ρ_{xy} : ONLY MEASURES LINEAR RELATION BETWEEN X & Y.

Ex Cases where $\rho_{x,y} = 0$

X : Mobile No & Y is your Salary

$\rho_{X,Y} > 0$

X : Salary, T is Income tax

$-1 < \rho_{XY} < 0.$

X : Measures the distance of a mobile phone from a base stn

Y : Power of the received signal at the mobile.

* If X and Y are RVs such that $Y = aX + b$

$$\rho_{X,Y} = \begin{cases} -1 & a < 0 \\ 0 & a = 0 \\ 1 & a > 0 \end{cases}$$

* Independent Random Variables:

$$f_{X,Y}(x,y) = f_X(x)\, f_Y(y)$$

$$p_{X,Y}(x,y) = p_X(x)\, p_Y(y)$$

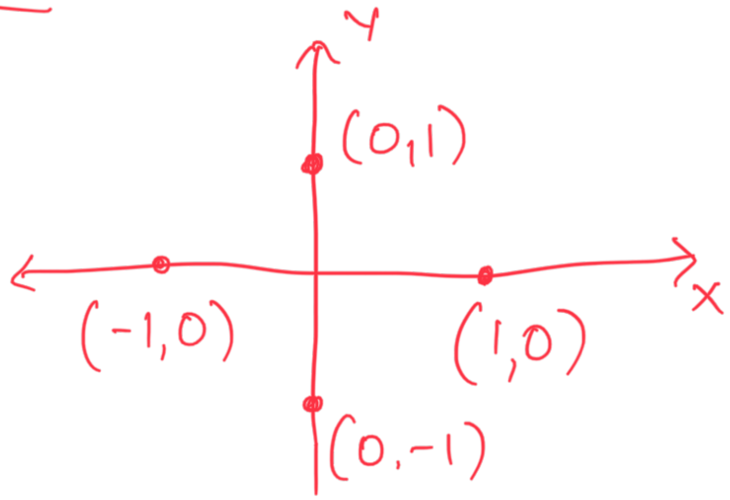
If X & Y are indep, what is $Cov(X,Y)$?

WORK FROM HOME!

If X & Y are indep then $\text{cov}(X,Y) = 0$.

However if $\text{cov}(X,Y) = 0$, X & Y need not be indep!

Example:



There are 4 points as above and each of the points is equiprobable with probability 1/4

Joint PMF.

	$Y=-1$	$Y=0$	$Y=1$		$p_X(i)$
$X=-1$	0	1/4	0	$\Rightarrow$	1/4
$X=0$	1/4	0	1/4	$\Rightarrow$	1/2
$X=1$	0	1/4	0	$\Rightarrow$	1/4
	$\Downarrow$	$\Downarrow$	$\Downarrow$		
$p_Y(j) =$	1/4	1/2	1/4		

$$\sum_{i=-1}^{1} P_x(i) = 1 \;; \quad \sum_{j=-1} P_y(j) = 1$$

$$E[XY] = \sum_{x=-1}^{1}\sum_{y=-1}^{1} xy \; p_{x,y}(x,y) = 0$$

But are X and Y indep ?

$$P_{x,y}(x,y) = p_x(x) \; P_y(y)$$

$$=$$

$$p_{xy}(1,0) = 1/4 \quad \text{from the table above}$$

$$p_x(1) \; P_y(0) = \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{8}$$

$$\Rightarrow \quad p_{xy}(1,0) \neq P_x(1) \, P_y(0)$$

$\Rightarrow$ X & Y are not indep.

It is also clear from the table that whenever X = 1, Y is surely 0.

$\Rightarrow$ All indep RVs are Uncorrelated & all Uncorrelated RVs are not indep.

Ex: 2. Let X be DRV taking values

-1, 0 & 1 with equal probability &

let $Y = X^2$

$E[XY]$

	(i)		$p(X=i)$		$Y = X^2$		$p(Y=j)$
X:	-1	$\Rightarrow$	$1/3$	Y:	1	$=$	$1/3 + 1/3 = 2/3$
	0	$=$	$1/3$		0		$1/3$.
	1	$=$	$1/3$.				

$$E[XY] = \sum_{-1}^{1} \sum_{0}^{1} xy\, P_{X,Y}(x,y)$$

$$\Rightarrow \cancel{(-1)(0)\, P_{X,Y}(-1,0)} + (-1)(1)\, P_{XY}(-1,1)$$
$$+ \cancel{0(0)\, P_{XY}(0,0)} + \cancel{(0,1)\, P_{XY}(0,1)}$$
$$+ \cancel{1(0)\, P_{XY}(1,0)} + (1)(1)\, P_{XY}(1,1)$$

$$= -P_{XY}(-1,1) + P_X(1,1)$$

$$\Rightarrow -2/3 + 2/3 = 0$$

$E[XY] = 0 \Rightarrow$ Tempting to say they are indep.

$\Downarrow$

Uncorrelated

However $Y = X^2$

$\Rightarrow Y$ is dep on X

$\therefore$ Uncorrelated RVs are not necessarily indep.

Aside: Linearity of Expectation operator.

X & Y are DRVs.

$$E[X+Y] = E[X] + E[Y].$$

$$E[X+Y] = \sum_x \sum_y (x+y)\, p_{X,Y}(x,y)$$

$$= \sum_x \sum_y x\, p_{X,Y}(x,y) + \sum_x \sum_y y\, p_{X,Y}(x,y)$$

$$\Rightarrow \sum_x x \underbrace{\boxed{\left(\sum_y p_{X,Y}(x,y)\right)}}_{\text{Marg. } X} + \sum_y y \underbrace{\boxed{\left(\sum_x p_{X,Y}(x,y)\right)}}_{\text{Marg. } Y}$$

$$\Rightarrow \sum_x x\, p_X(x) + \sum_y y\, p_Y(y)$$

$$= E[X] + E[Y].$$

Jointly Gaussian / Bivariate Gaussian:
(2 RVs)

$$f_{X,Y}(x,y) =$$

$$\left[\left(\frac{(x-\mu)^2}{\ldots} + \frac{(y-\mu)^2}{\ldots} - \frac{2\rho\,(x-\mu)(y-\mu)}{\ldots}\right) \right]$$

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

If X & Y are Uncorrelated, $\rho_{XY} = 0$

$$\Rightarrow \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2}}{2} \right)}$$

$$\Rightarrow \frac{1}{\sqrt{2\pi}\sigma_x \sqrt{2\pi}\sigma_y} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}} \cdot e^{-\frac{(y-\mu_y)^2}{2\sigma_y^2}}.$$

$$\Rightarrow f_X(x) f_Y(y)$$

ONLY IN THE CASE OF UNCORRELATED GAUSSIAN RV $\Rightarrow$ INDEPENDENCE

Indep. Gaussian $\Rightarrow$ Uncorrelated

Uncorrelated $\Rightarrow$ Indep only with Jointly Gaussian RVs.