Def: a random variable X is a quantity which can take one of a number of value with a set of probabilities

e.g. earnings in a game that depend on a finite # of outcomes

discrete: X takes values $x_1, \dots, x_n$ with probabilities $p_1, \dots, p_n$ $\quad (p_1 + \dots + p_n = 1)$

continuous: X takes any real number as values, according to a probability distribution

$p(x) : \mathbb{R} \to \mathbb{R}_{\geq 0}$

$\left(\int_{-\infty}^{\infty} p(x)\, dx = 1 \right)$

e.g. temperature

$\mathrm{Prob}(X \in [a,b]) = \int_a^b p(x)\, dx$

- advantage: now you can do algebraic manipulations

→ if X is a random variable, so is $X + 7$

→ ——— " - ————, so is X^2

→ if X and Y are random variables, so are $X + Y$, $X \cdot Y$ ……

Def: the expected value (mean) of a random variable X is

$$E[X] = \mu \begin{cases} x_1 p_1 + \dots + x_n p_n & \text{if } X = \text{discrete} \\ \int_{-\infty}^{\infty} p(x) \cdot x \cdot dx & \text{if } X = \text{continuous} \end{cases}$$

Def: the variance of a random variable X is

$$E\left[(X - E[X])^2\right] = \Sigma \begin{cases} (x_1-\mu)^2 p_1 + \dots + (x_n-\mu)^2 p_n & \text{if } X = \text{discrete} \\ \int_{-\infty}^{\infty} p(x) \cdot (x-\mu)^2 dx & \text{if } X = \text{continuous} \end{cases}$$

Def: the covariance of two random variables X and Y is

$$E\left[(X - E[X])(Y - E[Y])\right] = \Sigma_{XY}$$

the variance of X = the covariance of X and X

$\cdot$ X and Y are discrete : let $p_{ij} = \text{Prob}(X = x_i \text{ and } Y = y_j)$

$X \rightsquigarrow x_1, \dots, x_n$

$Y \rightsquigarrow y_1, \dots, y_m$

$$\mu = E[X] = \sum_i x_i \cdot \text{Prob}(X = x_i)$$

$$= \sum_{i,j} x_i \cdot \underbrace{\text{Prob}(X = x_i \text{ and } Y = y_j)}_{p_{ij}}$$

$$= \sum_{i,j} p_{ij}\, x_i$$

similarly, $\nu = E[Y] = \sum_{i,j} p_{ij}\, y_j$

$$\Sigma_{XY} = E\left[(X-\mu)(Y-\nu)\right]$$

$$= \sum_{i,j} p_{ij}\,(x_i - \mu)(y_j - \nu)$$

2 independent coin tosses

$$p_{HH} = p_{HT} = p_{TH} = p_{TT} = \frac{1}{4}$$

X = # of heads of toss 1 $\longrightarrow \mu = 1/2$

Y = # of heads of toss 2 $\longrightarrow \nu = 1/2$

$$\Sigma_{XY} = \frac{1}{4}\cdot\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{4}\cdot\frac{1}{2}\cdot\left(-\frac{1}{2}\right) + \frac{1}{4}\left(-\frac{1}{2}\right)\frac{1}{2} + \frac{1}{4}\cdot\left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) = 0$$

2 dependent coin tosses (assume coins are glued together)

$$p_{HH} = p_{TT} = \frac{1}{2}, \quad p_{TH} = p_{HT} = 0$$

X = # of heads of toss 1 $\longrightarrow \mu = 1/2$

Y = # of heads of toss 2 $\longrightarrow \nu = 1/2$

$$\Sigma_{XY} = 1/2 \cdot 1/2 \cdot 1/2 + 1/2 \cdot \left(-\frac{1}{2}\right)\left(-\frac{1}{2}\right) + 0 \cdot \left(-\frac{1}{2}\right)\cdot\frac{1}{2} + 0 \cdot \frac{1}{2}\left(-\frac{1}{2}\right) = \frac{1}{4}$$

Interpretation : • if $\Sigma_{XY} > 0$, then X and Y are called correlated

• if $\Sigma_{XY} < 0$, ——— " ——— anti-correlated

• if $\Sigma_{XY} = 0$, ——— " ——— uncorrelated

(if X and Y are independent, then they are uncorrelated)

$$\boxed{\begin{aligned} &\mathrm{Prob}(X=a \text{ and } Y=b) \\ &\quad \parallel \\ &\mathrm{Prob}(X=a)\cdot \mathrm{Prob}(Y=b) \end{aligned}}$$

Thm (Cauchy-Schwartz)

$$\boxed{|\Sigma_{XY}| \leq \sqrt{\Sigma_{XX} \cdot \Sigma_{YY}}}$$

equality if $X=Y$ or if $X=-Y$

$\mathrm{Tr}\, K = \Sigma_{XX} + \Sigma_{YY} > 0$

$\det K = \Sigma_{XX}\Sigma_{XY} - \Sigma_{XY}^2 \geq 0$

by C-S

Def : the covariance matrix of two random variables X and Y is defined as

$$K = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$$

is a positive semi-definite symmetric matrix

actually positive definite unless X and Y are perfectly (anti)-correlated

$$K = \sum_{i,j} p_{ij} \begin{bmatrix} (x_i-\mu)^2 & (x_i-\mu)(y_j-\nu) \\ (y_j-\nu)(x_i-\mu) & (y_j-\nu)^2 \end{bmatrix} = \sum_{i,j} p_{ij} \begin{bmatrix} x_i-\mu \\ y_j-\nu \end{bmatrix} \begin{bmatrix} x_i-\mu & y_j-\nu \end{bmatrix}$$

independent coin tosses :

$$K = \frac{1}{4}\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} + \frac{1}{4}\begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \end{bmatrix} + \frac{1}{4}\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + \frac{1}{4}\begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

dependent coin toss :

$$K = \frac{1}{2}\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \end{bmatrix} + 0\begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}\begin{bmatrix} \frac{1}{2} & -\frac{1}{2} \end{bmatrix} + 0\begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & \frac{1}{2} \end{bmatrix} + \frac{1}{2}\begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}\begin{bmatrix} -\frac{1}{2} & -\frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

det = 0.

Erratum: all these entries should be 1/4

K is diagonal if and only if X and Y are uncorrelated

Continuous case: X, Y

$$p(x,Y) : \mathbb{R}^2 \longrightarrow \mathbb{R}_{\geq 0} \quad , \quad \mathrm{Prob}\Big(X \in [a,b] \text{ and } Y \in [c,d]\Big) = \int_a^b \int_c^d p(x,Y)\, dx\, dY$$

$$\Sigma_{XY} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,Y) \cdot (X - \overset{\mu = E[X]}{\mu})(Y - \overset{\nu = E[Y]}{\nu})\, dx\, dY$$

$$K = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,Y) \cdot \underbrace{\begin{bmatrix} X - \mu \\ Y - \nu \end{bmatrix} \begin{bmatrix} X - \mu & Y - \nu \end{bmatrix}}_{\begin{bmatrix} (X-\mu)^2 & (X-\mu)(Y-\nu) \\ (Y-\nu)(X-\mu) & (Y-\nu)^2 \end{bmatrix}} dx\, dY$$