Covariance and Correlation

Definition: Given a probability space $(\Omega, \Pr)$ and two random variables $X: \Omega \to \mathbb{R}$ and $Y: \Omega \to \mathbb{R}$, the covariance of X and Y is

$$Cov(X,Y) = \sum_{w \in \Omega} \Pr(w)(X(w) - E[X])(Y(w) - E[Y]) = E[(X - E[X])(Y - E[Y])]$$

Proposition: $Cov(X,Y) = \sum_{x \in Range(X)} \sum_{y \in Range(Y)} \Pr(X=x, Y=y)(x - E[X])(y - E[Y])$

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

Proof:
$$\begin{aligned} Cov(X,Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[XY] - E[E[Y]X] - E[E[X]Y] + E[X]E[Y] \\ &= E[XY] - E[Y]E[X] - E[X]E[Y] + E[X]E[Y] \\ &= E[XY] - E[X]E[Y] \end{aligned}$$

Observation: If X and Y are independent random variables then $Cov(X,Y) = 0$

Definition: Given two random variables X and Y the correlation between X and Y is $\rho_{X,Y} = \dfrac{Cov(X,Y)}{\sigma_X \sigma_Y}$

$\sigma_X = \sqrt{Var(X)}$

$\sigma_Y = \sqrt{Var(Y)}$

Covariance and Correlation Example

Example: Let's say we flip a fair coin 3 times.
If $X = \#$ of heads and $Y = \begin{cases} 1 & \text{if we have at least 2 heads} \\ 0 & \text{otherwise} \end{cases}$,

what is the correlation between X and Y?

Probability	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
X	0	1	2	3
Y	0	0	1	1

$E[X] = \frac{3}{2}$, $E[Y] = \frac{1}{2}$

$E[XY] = \frac{3}{8} \cdot 2 \cdot 1 + \frac{1}{8} \cdot 3 \cdot 1 = \frac{6+3}{8} = \frac{9}{8}$

$$Cov(X,Y) = E[XY] - E[X]E[Y] = \frac{9}{8} - \frac{3}{2} \cdot \frac{1}{2} = \frac{9}{8} - \frac{6}{8} = \frac{3}{8}$$

$Var(X) = \frac{3}{4}$ $\quad \sigma_X = \sqrt{Var(X)} = \frac{\sqrt{3}}{2}$

$Var(Y) = \frac{1}{4}$ $\quad \sigma_Y = \sqrt{Var(Y)} = \frac{1}{2}$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\frac{3}{8}}{\frac{\sqrt{3}}{2} \cdot \frac{1}{2}} = \frac{\sqrt{3}}{2}$$

Geometric Interpretation of Covariance and Correlation

Idea: Given a probability space $(\Omega, \Pr)$ and random variables $X:\Omega\to\mathbb{R}$ and $Y:\Omega\to\mathbb{R}$, let $\vec{A}$ and $\vec{B}$ be vectors indexed by the atomic events $\omega\in\Omega$ with coordinates

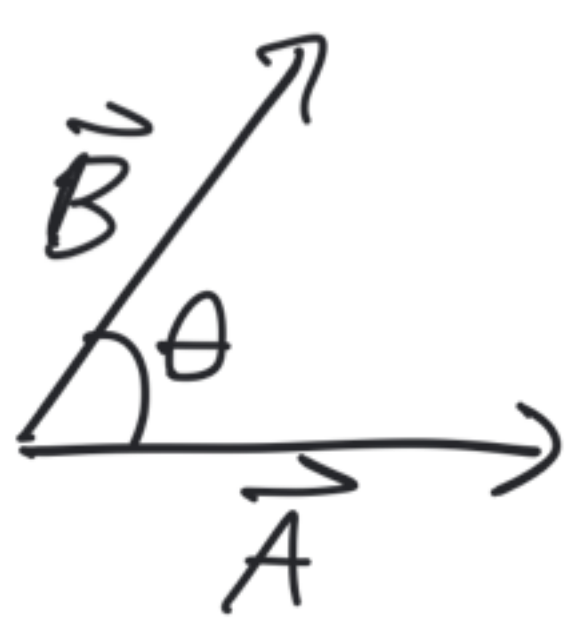
$$\vec{A}_\omega = \sqrt{\Pr(\omega)}\,(X(\omega)-E[X])$$

$$\vec{B}_\omega = \sqrt{\Pr(\omega)}\,(Y(\omega)-E[Y])$$

$$Cov(X,Y) = \vec{A}\cdot\vec{B}$$

$$Var(X) = \vec{A}\cdot\vec{A} = \|\vec{A}\|^2 \qquad \sigma_X = \|\vec{A}\|$$

$$Var(Y) = \vec{B}\cdot\vec{B} = \|\vec{B}\|^2 \qquad \sigma_Y = \|\vec{B}\|$$



$$\vec{A}\cdot\vec{B} = \|\vec{A}\|\cdot\|\vec{B}\|\cdot\cos\theta$$

$$\rho_{X,Y} = \frac{Cov(X,Y)}{\sigma_X\sigma_Y} = \frac{\cancel{\|\vec{A}\|\cdot\|\vec{B}\|}\cdot\cos\theta}{\cancel{\|\vec{A}\|\cdot\|\vec{B}\|}} = \cos\theta$$