Basics of Covariance

Suppose X is a RV w/ prob meas $\mathbb{P}$ (will skip the formalities of this for now.)

Def

$$\mathbb{E}X := \int_\Omega X(\omega)\, d\mathbb{P}(\omega)$$

Def Given RVs X & Y with finite first moment,

$$\text{Cov}(X,Y) := \mathbb{E}\left[(X-\mathbb{E}X)(Y-\mathbb{E}Y)\right]$$

Def (correlation coeff)

$$\rho := \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

The set of RVs w/ $\mathbb{E}|X|^2 < \infty$ is the L_2 space over functions from $\Omega \to \mathbb{R}$. ~~Cov(X,Y) is a valid inner product on the space. The cosine of the angle θ between X and Y~~. For U, V in this space

$$\langle U, V\rangle := \mathbb{E}UV$$

is a valid inner product. The cosine of the angle θ between U, V $\underbrace{}_{\text{if}}$ $\|U\|_2, \|V\|_2 \neq 0$ given by

$$\cos\theta = \frac{\langle U, V\rangle}{\|U\|_2 \|V\|_2}.$$

If we let $U = X - \mathbb{E}X$, $V = Y - \mathbb{E}Y$ ($\mathbb{E}X, \mathbb{E}Y < \infty$ if $X, Y \in L_2$ by Monotonicity of L_p on probability spaces), then $\text{Cov}(X,Y) = \langle U, V\rangle$ and the corr coeff is given by

$$\rho := \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}$$

Note that ρ is the angle between the centered versions of X & Y

Side note: A useful formula: $\operatorname{Var} X = \mathbb{E}[X^2] - (\mathbb{E}X)^2$.

Example: Suppose X, Y have joint distribution

	x_1	x_2
y_2	a	b
y_1	c	d

$$\mu_X = \mathbb{E}X = x_1(a+c) + x_2(1-(a+c))$$

$$\mu_Y = y_1(c+d) + y_2(1-(c+d))$$

$$\begin{aligned} \operatorname{Var} X &= \mathbb{E}[X^2] - (\mathbb{E}X)^2 \\ &= x_1^2(a+c) + x_2^2(1-(a+c)) - \mu_X \end{aligned}$$

$$\operatorname{Var} Y = y_1^2(c+d) + y_2^2(1-(c+d)) - \mu_Y$$

$$\begin{aligned} \operatorname{Cov}(X,Y) &= \mathbb{E}[(X-\mu_X)(Y-\mu_Y)] \\ &= a(x_1 - \mu_X)(y_2 - \mu_Y) \\ &\quad b(x_2 - \mu_X)(y_2 - \mu_Y) \\ &\quad c(x_1 - \mu_X)(y_1 - \mu_Y) \\ &\quad d(x_2 - \mu_Y)(y_1 - \mu_Y) \end{aligned}$$

Run different values of these on computer. Include $c=b=0$, & $a=d=0$.
Compute ρ.

<u>Def</u> RVs X & Y are said to be independent if $\forall x, y$,

$$\mathbb{P}(X \le x \text{ and } Y \le y) = \mathbb{P}(X \le x)\mathbb{P}(Y \le y)$$

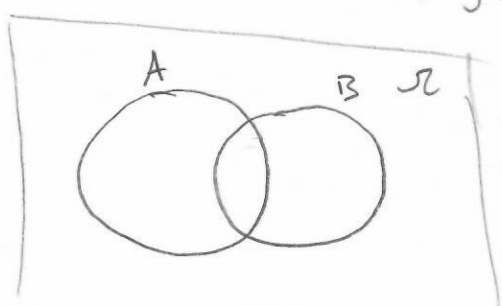
Intuition: We would like to capture the idea that knowing Y provides us no information about X. Bayes thm tells us

$$\mathbb{P}(\underbrace{X \le x}_{A}, \underbrace{Y \le y}_{B}) = \mathbb{P}(A \wedge B) = \mathbb{P}(A|B)\mathbb{P}(A)$$

events

If $\mathbb{P}(A|B) = \mathbb{P}(A)$ then we get $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$

Also, regarding Bayes thm, note that the intuition for this just comes down to measuring the size of sets.



Roughly, If event B has occured, then when we condition on B, we rescale things so B is our universe, so $P(\cdot|B)$ is a probability measure & $\mathbb{P}(B|B) = 1$. If we have $\mathbb{P}(A|B)$, then we get back to normal scaling by multiplying by $\mathbb{P}(B)$.

This can be explained better, but its the gist. You can make this intuition clear & precise w/ a simple example like this. The intuition generalizes, but is most easily formalized using the standard def of indep.

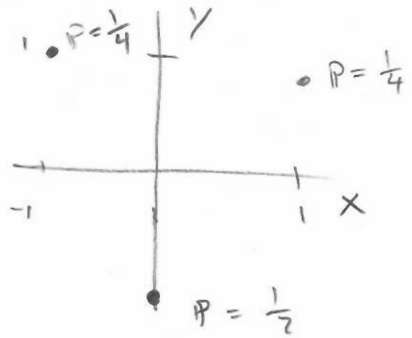
Ex Let $X \sim \mathcal{U}(-1, 1)$. $Y = X^2$

Verify: $\text{Cor}(X, Y) = 0$

But clearly, $\mathbb{P}(Y \le y) \ne \mathbb{P}(Y \le y \mid X \le x)$ since $X \le x \Rightarrow Y \le x^2$

TODO: Show visualization on computer

Ex Let X, Y have joint distribution



Ex

Let X & Y have uniform joint distribution with support where all corners are equidistant from 0.