Contents

Joint Probability Mass Function **1**

Covariance **2**

Correlation **2**

Dice Example 3

Dependence and Correlation 3

Correlation and Causation 4

Conditional Expectation **4**

Joint Probability Mass Function

Suppose we have two discrete random variables X and Y on same sample space S

- $P(X = x \text{ and } Y = y)$ is called their joint probability mass function
- Let's go back to sample space S. Remember RV X is really a function mapping from S to a real value, i.e. should really be written $X(\omega)$. Ditto Y.
- Let $E_x = \{\omega \in S : X(\omega) = x\}$ be set of outcomes for which $X = x$
- Let $E_y = \{\omega \in S : Y(\omega) = y\}$ be set of outcomes for which $Y = y$
- $P(X = x) = P(E_x), P(Y = y) = P(E_y)$
- Probability of both is $P(E_x \cap E_y)$ and $P(X = x \text{ and } Y = y) = P(E_x \cap E_y)$

Example: operating system loyalty. Person buys one computers, then another. $X = 1$ if first computer runs windows, else 0. $Y = 1$ is second computer runs windows, else 0.

- Join probability mass function:

	x=0	x=1	P(Y=y)
y=0	0.2	0.3	0.5
y=1	0.1	0.4	0.5
P(X=x)	0.3	0.7	1

- $P(X = 0 \text{ and } Y = 0) = 0.2, P(X = 0 \text{ and } Y = 1) = 0.3$, etc.

Covariance

Say X and Y are random variables with expected values μx and μy. The **covariance** of X and Y is defined as: $Cov(X,Y) = E[(X = \mu_x)(Y = \mu_y)]$

Equavalently:

- $Cov(X,Y) = E[XY - X\mu_y - Y\mu_x + \mu_x\mu_y]$
- $= E[XY] - E[X]\mu_y - E[Y]\mu_x + \mu_x\mu_y$
- $= E[XY] = \mu_x\mu_y - \mu_y\mu_x + \mu_x\mu_y$
- $= E[XY] - \mu_x\mu_y = E[XY] - E[X]E[Y]$

$Cov(X,X) = Var(X)$

Recall when X and Y are independent then $E[XY] = E[X]E[Y]$, so $Cov(X,Y) = 0$. But $Cov(X,Y) = 0$ does **not** imply that X and Y are independent.

Correlation

The **correlation** between X and Y is defined at $Corr(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$

- Also use $\rho_{x,y}$ instead of $Corr(X,Y)$, similarly to the way we λ_x as shorthand for expected value $E[X]$ and σ_x for standard deviation $\sqrt{Var(X)}$ (so $\sigma_x^2 = Var(X)$)
- Sometimes also called the **Pearson correlation coefficient**

Correlation variabes between -1 and 1.

If $X = Y$ then $corr(X,Y) = 1$. If $X = -Y$ then $corr(X,Y) = -1$.

The correlation is another example of a summary statistic. It indicates the strenght of a linear relationship between X and Y. Great case is needed though as it can easily be misleading.

- Correlation says *nothing* about the slope of the line (other than its sign)
- When relationship between X and Y is not roughly linear, correlation coefficient tells us almost nothing

Dice Example

Consider rolling a 6-sided die

- Indicator variable $X = 1$ if roll is $1, 2, 3, 4$
- Indicator variable $Y = 1$ is roll is $3, 4, 5, 6$

What is $Cov(X, Y)$?

- $E[X] = \frac{2}{3}, E[Y] = \frac{2}{3}$
- if $X = 0$ then $Y = 1$ and if $Y = 0$ then $X = 1$

$E[XY] = \sum_x \sum_y xyP(X = x \text{ and } Y = y) = 0 \times 0 \times 0 + 0 \times 1 \times \frac{1}{3} + 1 \times 0 \times \frac{1}{3} + 1 \times 1 \times \frac{1}{3} = \frac{1}{3}$

- $Cov(X, Y) = E[XY] - E[X]E[Y] = \frac{1}{3} - \frac{4}{9} = -\frac{1}{9}$
- Now $P(X = 1) = \frac{2}{3}$ and $P(X = 1 \mid Y = 1) = \frac{1}{2}$
  - So observing $Y = 1$ makes $X = 1$ less likely

Dependence and Correlation

Recall when X and Y are indepedent then $E[XY] = E[X]E[Y]$, so $corr(X, Y) = 0$. But $corr(X, Y) = 0$ does *not* imply that X and Y are indepdent.

Example: X and Y are random variables with joint PMF:

	$x = -1$	$x = 0$	$x = 1$	P(Y=y)
$y = 0$	$\frac{1}{3}$	0	$\frac{1}{3}$	$\frac{2}{3}$
$y = 1$	0	$\frac{1}{3}$	0	$\frac{1}{3}$
$P(X = x)$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1

X takes values $\{-1, 0, 1\}$ with equal probability and $Y = \begin{cases} 1 & X = 0 \\ 0 & \text{if } X \neq 0 \end{cases}$

- $E[X] = -1 \times \frac{1}{3} + 0 \times \frac{1}{3} + 1 \times \frac{1}{3} = 0, E[Y] = 0 \times \frac{2}{3} + 1 \times \frac{1}{3} = \frac{1}{3}$
- Since $XY = 0$ then $E[XY] = 0$
- $Cov(X, Y) = E[XY] - E[X]E[Y] = 0 - 0 = 0$
- Byt X and Y are clearly indepdent

Correlation and Causation

Correlation does not imply causation.

Conditional Expectation

X and Y are jointly distributed discrete random variables.

- Recall conditional PMF of X given $Y = y$ is $P(X = x \mid Y = y) = \frac{P(X=x \text{ and } Y=y)}{P(Y=y)}$
- Define conditional expectation of X given $Y = y$ as $E[X \mid Y = y] = \sum_x xP(X = x \mid Y = y)$
- This is not the same as the expectation $E[X]$
  - E.g. its one thing to ask what the average height of a person in Ireland is and another to ask this once we know that they are male

Roll two six sided dice. X is the value of the sum, Y is the outcome of the first die roll.

- $E[X \mid Y = 6] = \sum_x xP(X = x \mid Y = y) = \frac{1}{6}(7 + 8 + 9 + 10 + 11 + 12) = \frac{57}{6} = 9.5$
- Makes sense: $6 + E[\text{value of second die roll}] = 6 + 3.5$

Linearity:

- $E[\sum_i Y_i \mid X = x] = \sum_i E[Y_i \mid X = x]$
- Proof is the same as for unconditional expectation

Marginalisation:

- $E[X] = \sum_y E[X \mid Y = y]P(Y = y)$