Joint Expectation

Recall:

$$
\mathbb{E}[X] = \int_\Omega xf_X(x) \, dx.
$$

How about the expectation for two variables?

Definition

Let $X$ and $Y$ be two random variables. The **joint expectation** is

$$
\mathbb{E}[XY] = \sum_{y \in \Omega_Y} \sum_{x \in \Omega_X} xy \cdot p_{X,Y}(x, y) \quad (1)
$$

if $X$ and $Y$ are discrete, or

$$
\mathbb{E}[XY] = \int_{y \in \Omega_Y} \int_{x \in \Omega_X} xy \cdot f_{X,Y}(x, y) \, dx \, dy \quad (2)
$$

if $X$ and $Y$ are continuous. Joint expectation is also called **correlation**.
Example 1. Let $X$ be a coin flip, $Y$ be a dice. Find $\mathbb{E}[XY]$.

Solution.

\[
\begin{array}{cccccc}
\hline
& 1 & 2 & 3 & 4 & 5 & 6 \\
\hline
X = 0 & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} \\
X = 1 & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} & \frac{1}{12} \\
\hline
\end{array}
\]

\[
\mathbb{E}[XY] = \sum_{x,y} p_{X,Y}(x, y) = \frac{21}{12}.
\]
Today’s lecture

- What is joint expectation?
- Why define joint expectation as $\mathbb{E}[XY]$?
- Independence
- Covariance
- Correlation
Understanding $\mathbb{E}[XY]$ 

So what is $\mathbb{E}[XY]$?

$$\mathbb{E}[XY] = \sum_{i=1}^{N} \sum_{j=1}^{N} x_i y_j \cdot p_{X,Y}(x_i, y_j),$$

This can actually be written in the matrix form

$$\text{PMF} = \begin{bmatrix}
p_{X,Y}(x_1, y_1) & p_{X,Y}(x_1, y_2) & \cdots & p_{X,Y}(x_1, y_N) \\
p_{X,Y}(x_2, y_1) & p_{X,Y}(x_2, y_2) & \cdots & p_{X,Y}(x_2, y_N) \\
\vdots & \vdots & \ddots & \vdots \\
p_{X,Y}(x_N, y_1) & p_{X,Y}(x_N, y_2) & \cdots & p_{X,Y}(x_N, y_N)
\end{bmatrix}. \quad (3)$$
So, $\mathbb{E}[XY]$ is actually ...

\[
\mathbb{E}[XY] = \begin{bmatrix} x_1 & \cdots & x_N \end{bmatrix} \begin{bmatrix}
\begin{bmatrix}
pxy(x_1, y_1) & \cdots & pxy(x_1, y_N) \\
\vdots & \ddots & \vdots \\
pxy(x_N, y_1) & \cdots & pxy(x_N, y_N)
\end{bmatrix}
\end{bmatrix} y
\]

\[
= x^T P y.
\] (4)

**Why correlation is defined as $\mathbb{E}[XY]$?**

- $\mathbb{E}[XY]$ is a weighted inner product between the states:

\[
\mathbb{E}[XY] = x^T P y.
\] (5)

- $x$ and $y$ are the states of the random variables $X$ and $Y$.
- Inner product measures the similarity between two vectors.
Then, what is $x^T P y$?

**Definition**

Let $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^N$ be two vectors. Define the *cosine angle* $\cos \theta$

$$
\cos \theta = \frac{x^T y}{\|x\|\|y\|},
$$

where $\|x\| = \sqrt{\sum_{i=1}^{N} x_i^2}$ is the norm of the vector $x$, $\|y\| = \sqrt{\sum_{i=1}^{N} y_i^2}$ is the norm of the vector $y$. 

**Distributions**

$p_X(x)$ or $P_X$

$p_{X,Y}(x,y)$ or $P_{X,Y}$

$p_Y(y)$ or $P_Y$
The denominators in cosine angle

\[
\mathbb{E}[X^2] = \sum_{i=1}^{N} x_i x_i \cdot p_X(x_i)
\]

\[
= \left[ \begin{array}{c} x_1 \\ \vdots \\ x_N \end{array} \right]^T \begin{bmatrix} p_X(x_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_X(x_N) \end{bmatrix} \left[ \begin{array}{c} x_1 \\ \vdots \\ x_N \end{array} \right]
= x^T P_X x = \left\| x \right\|_{P_X}^2,
\]

\[
\mathbb{E}[Y^2] = \sum_{j=1}^{N} y_j y_j \cdot p_Y(y_j)
\]

\[
= \left[ \begin{array}{c} y_1 \\ \vdots \\ y_N \end{array} \right]^T \begin{bmatrix} p_Y(y_1) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_Y(y_N) \end{bmatrix} \left[ \begin{array}{c} y_1 \\ \vdots \\ y_N \end{array} \right]
= y^T P_Y y = \left\| y \right\|_{P_Y}^2.
\]
So, the cosine angle is

\[ \cos \theta = \frac{x^T P_{XY} y}{\|x\|_{P_X} \|y\|_{P_Y}}, \tag{7} \]

which is the same as

\[ \cos \theta = \frac{x^T P_{XY} y}{\|x\|_{P_X} \|y\|_{P_Y}} = \frac{\mathbb{E}[XY]}{\sqrt{\mathbb{E}[X^2]} \sqrt{\mathbb{E}[Y^2]}}. \tag{8} \]

What is \( \mathbb{E}[XY] \) then?

- \( \mathbb{E}[XY] \) is the cosine angle between \( x \) and \( y \)
- Measures how similar they are
- Weight by the PMF/PDF
Therefore, ... 

Since the cosine angle is between -1 and 1, we have that

$$-1 \leq \frac{E[XY]}{\sqrt{E[X^2]} \sqrt{E[Y^2]}} \leq 1. \quad (9)$$

If you like to prove this formally, ... 

**Theorem (Cauchy-Schwarz Inequality)**

*For any random variables $X$ and $Y$, it holds that*

$$(E[XY])^2 \leq E[X^2]E[Y^2]. \quad (10)$$

**Proof:** See ebook.
Questions?