Outline

- Basics of probability theory
- Random variable and distributions: Expectation and Variance
Definition of Probability

- **Experiment**: toss a coin twice
- **Sample space**: possible outcomes of an experiment
  - $S = \{HH, HT, TH, TT\}$
- **Event**: a subset of possible outcomes
  - $A=\{HH\}$, $B=\{HT, TH\}$
- **Probability of an event**: an number assigned to an event $\text{Pr}(A)$
  - Axiom 1: $0 \leq \text{Pr}(A) \leq 1$
  - Axiom 2: $\text{Pr}(S) = 1$, $\text{Pr}(\emptyset) = 0$
  - Axiom 3: For two events $A$ and $B$, $\text{Pr}(A \cup B) = \text{Pr}(A) + \text{Pr}(B) - \text{Pr}(A \cap B)$
  - Proposition 1: $\text{Pr}(\neg A) = 1 - \text{Pr}(A)$
  - Proposition 2: For every sequence of disjoint events $\text{Pr}(\bigcup_i A_i) = \sum_i \text{Pr}(A_i)$
Joint Probability

- For events A and B, **joint probability** $\Pr(AB)$ (also shown as $\Pr(A \cap B)$) stands for the probability that both events happen.

- Example: $A=\{HH\}$, $B=\{HT, TH\}$, what is the joint probability $\Pr(AB)$?
  
  Zero
Independence

- Two events $A$ and $B$ are independent in case
  \[ \Pr(AB) = \Pr(A) \Pr(B) \]

- A set of events $\{A_i\}$ is independent in case
  \[ \Pr(\bigcap_i A_i) = \prod_i \Pr(A_i) \]
Independence

- Two events $A$ and $B$ are independent in case
  \[ \Pr(AB) = \Pr(A)\Pr(B) \]
- A set of events $\{A_i\}$ is independent in case
  \[ \Pr(\bigcap_i A_i) = \prod_i \Pr(A_i) \]
- Example: Drug test

<table>
<thead>
<tr>
<th></th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success</td>
<td>200</td>
<td>1800</td>
</tr>
<tr>
<td>Failure</td>
<td>1800</td>
<td>200</td>
</tr>
</tbody>
</table>

A = \{A patient is a Woman\}
B = \{Drug fails\}

Will event A be independent from event B?

\[ \Pr(A) = 0.5, \ \Pr(B) = 0.5, \ \Pr(AB) = 9/20 \]
Independence

Consider the experiment of tossing a coin twice

Example I:
- $A = \{HT, HH\}$, $B = \{HT\}$
- Will event $A$ independent from event $B$?

Example II:
- $A = \{HT\}$, $B = \{TH\}$
- Will event $A$ independent from event $B$?

Disjoint $\neq$ Independence

If $A$ is independent from $B$, $B$ is independent from $C$, will $A$ be independent from $C$?

Not necessarily, say $A = C$
If $A$ and $B$ are events with $\Pr(A) > 0$, the conditional probability of $B$ given $A$ is

$$\Pr(B \mid A) = \frac{\Pr(AB)}{\Pr(A)}$$
If A and B are events with \( \Pr(A) > 0 \), the \textit{conditional probability of B given A} is

\[
\Pr(B \mid A) = \frac{\Pr(AB)}{\Pr(A)}
\]

\textbf{Example: Drug test}

\begin{tabular}{|c|c|c|}
\hline
 & Women & Men \\
\hline
Success & 200 & 1800 \\
\hline
Failure & 1800 & 200 \\
\hline
\end{tabular}

A = \{Patient is a Woman\}

B = \{Drug fails\}

\[\Pr(B\mid A) = ?\]

\[\Pr(A\mid B) = ?\]
Conditioning

- If A and B are events with \( \Pr(A) > 0 \), the *conditional* probability of B given A is

\[
\Pr(B \mid A) = \frac{\Pr(AB)}{\Pr(A)}
\]

- Example: Drug test

A = {Patient is a Woman}
B = {Drug fails}
\[
\Pr(AB) = \frac{18}{20} \quad \Pr(A) = \frac{20}{20}\]

<table>
<thead>
<tr>
<th></th>
<th>Women</th>
<th>Men</th>
</tr>
</thead>
<tbody>
<tr>
<td>Success</td>
<td>200</td>
<td>1800</td>
</tr>
<tr>
<td>Failure</td>
<td>1800</td>
<td>200</td>
</tr>
</tbody>
</table>

- Given A is independent from B, what is the relationship between \( \Pr(A \mid B) \) and \( \Pr(A) \)?

\[
\Pr(A \mid B) = \Pr(A)
\]
Outline

- Basics of probability theory
- Bayes’ rule
- Random variable and probability distribution: Expectation and Variance
Random Variable and Distribution

- A *random variable* $X$ is a numerical outcome of a random experiment.

- The *distribution* of a random variable is the collection of possible outcomes along with their probabilities:
  - **Discrete case:** $\Pr(X = x) = p_\theta(x)$
  - **Continuous case:** $\Pr(a \leq X \leq b) = \int_a^b p_\theta(x)dx$

- The *support* of a discrete distribution is the set of all $x$ for which $\Pr(X=x) > 0$.

- The *joint distribution* of two random variables $X$ and $Y$ is the collection of possible outcomes along with the joint probability $\Pr(X=x, Y=y)$. 
Random Variable: Example

- Let $S$ be the set of all sequences of three rolls of a die. Let $X$ be the sum of the number of dots on the three rolls.
- What are the possible values for $X$?
- $\Pr(X = 3) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{216}$,
- $\Pr(X = 5) =$ ?
Expectation

- A random variable $X \sim \Pr(X=x)$. Then, its expectation is
  \[ E[X] = \sum_x x \Pr(X = x) \]

  - In an empirical sample, $x_1, x_2, \ldots, x_N$,
  \[ E[X] = \frac{1}{N} \sum_{i=1}^{N} x_i \]

- Continuous case: $E[X] = \int_{-\infty}^{\infty} xp_{\theta}(x)dx$

- In the discrete case, expectation is indeed the average of numbers in the support weighted by their probabilities

- Expectation of sum of random variables
  \[ E[X_1 + X_2] = E[X_1] + E[X_2] \]
Expectation: Example

- Let $S$ be the set of all sequence of three rolls of a die. Let $X$ be the sum of the number of dots on the three rolls.

- Exercise: What is $E(X)$?

- Let $S$ be the set of all sequence of three rolls of a die. Let $X$ be the product of the number of dots on the three rolls.

- Exercise: What is $E(X)$?
Variance

The variance of a random variable $X$ is the expectation of $(X-\mu)^2$:

$$
\operatorname{Var}(X) = \mathbb{E}[(X-\mu)^2] \\
= \mathbb{E}[X^2] + \mathbb{E}[\mu^2] - 2\mu \mathbb{E}[X] = \\
= \mathbb{E}[X^2] + \mathbb{E}[\mu]^2 - 2\mu \mu = \\
= \mathbb{E}[X^2] - \mu^2
$$