Probability in Computing

- Reminders
  - Submit signed Collaboration and Honesty policy on Gradescope
  - HW1 due Thursday
  - Have fun with Non-transitive dice!

- LECTURE 2
  - Basic concepts in probability
  - Probability Function
  - Axioms of Probability
  - Probability rules
  - Tree diagram method
  - Monty Hall problem
  - Birthday Paradox

Prof. Tiago Januario
Slides by Prof. Sofya Raskhodnikova
• Random experiment: a repeatable procedure

• Outcome: result of the experiment

• Sample space $\Omega$: set of all possible outcomes

• Event $E$: a subset of the sample space

• Probability function $Pr$: assigns a probability $Pr(E)$ to each event $E$
assigns a probability \( \Pr(E) \) to each event \( E \)

- **Experiment:** toss a fair coin

<table>
<thead>
<tr>
<th>( E )</th>
<th>( \emptyset )</th>
<th>{H}</th>
<th>{T}</th>
<th>{H,T}</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Pr(E) )</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Experiment:** roll a die

\( Tiago \) Januario: Probability in Computing
assigns a probability $\Pr(E)$ to each event $E$

- Experiment: toss a fair coin 3 times
What principles did we use to come up with those probability functions?

- **Symmetry**: each outcome of the coin toss (or die roll) is equally likely
- The probability of each outcome is a number between 0 and 1
- **Additivity**: for events with more than one outcome, the probability of the event is the sum of the probabilities of its outcomes

**Note**: some experiments are not symmetric (toss a coin until H)
Axioms of Probability

• Probability function assigns a probability to each event
• A probability function must satisfy the following properties, called axioms of probability
  – Non-negativity: \( \Pr(E) \geq 0 \) for all events \( E \subseteq \Omega \)
  – Additivity: if \( A \) and \( B \) are disjoint events then
    \[ \Pr(A \cup B) = \Pr(A) + \Pr(B) \]
    In particular, for each event \( E \), \( \Pr(E) \) is the sum of probabilities of outcomes in \( E \)
    if \( A_1, A_2, \ldots \) are disjoint events then
    \[ \Pr(A_1 \cup A_2 \cup \ldots) = \Pr(A_1) + \Pr(A_2) + \ldots \]
  – Normalization: \( \Pr(\Omega) = 1 \)
Axioms of Probability

- **Probability function** assigns a probability to each event
- A probability function must satisfy the following properties, called **axioms of probability**
  - **Non-negativity**: \( \Pr(E) \geq 0 \) for all events \( E \subseteq \Omega \)
  - **Additivity**: if \( A \) and \( B \) are disjoint events then
    \[
    \Pr(A \cup B) = \Pr(A) + \Pr(B)
    \]
  - **Normalization**: \( \Pr(\Omega) = 1 \)

\[\text{Do we need to add } \Pr(E) \leq 1?\]
• 50 students showed up to office hours today
• 20 are Red Sox fans (R), 25 are Patriots fans (P)
• Professor Tiago chooses a student uniformly at random.
• Let \( p \) be the probability that she chooses a Red Sox or a Patriot fan.

What is the range of possible values for \( p \)?

A. \( p \geq 0.5 \)
B. \( p \leq 0.4 \)
C. \( 0.4 \leq p \leq 0.9 \)
D. \( 0.5 \leq p \leq 0.9 \)
E. \( 0.4 \leq p \leq 0.5 \)
Probability: Cast of Characters

- Random experiment: a repeatable procedure
- Outcome: result of the experiment
- Sample space $\Omega$: set of all possible outcomes
- Event: a subset of the sample space
- Probability function $Pr$: assigns a probability $Pr(E)$ to each event $E$

Frequent example: uniform probability spaces, where all outcomes have the same probability, $\frac{1}{|\Omega|}$. 
In a standard deck of 52 cards, how many different "full house" poker hands are there? (A full house consists of five cards with three of them having one value and two having a different value; the order of cards doesn't matter, so \{5\heartsuit, 5\diamondsuit, 5\spadesuit, Q\heartsuit, Q\clubsuit\} and \{5\diamondsuit, 5\heartsuit, Q\heartsuit, 5\spadesuit, Q\clubsuit\} are the same full house hand.)

A. \(13 \cdot 6 \cdot 6 \cdot 4\)

B. \(13 \cdot 12 \cdot 6 \cdot 4\)

C. \(13^2 \cdot 11 \cdot 6^2 \cdot 2\)

D. \(\binom{52}{5}\)

E. None of the above
We toss a coin 5 times. The outcome is a string of length 5 of H’s and T’s, where H means “heads” and T means “tails”. How many outcomes are there where the first toss is H or we have exactly 2 T’s?

A. 6  
B. 10  
C. 16  
D. 20  
E. None of the above
We toss a coin 5 times. The outcome is a string of length 5 of H’s and T’s, where H means “heads” and T means “tails”. What is the probability that the first toss is H or we have exactly 2 T’s?
• Probabilistic experiment: drawing a sample of $k$ elements from a set $S$ of $n$ elements.

Ex.: Tossing a coin 5 times. Then $S=\{H,T\}$. We are drawing 5 elements out of 2 with replacement.

• The outcome of a probabilistic experiment is called a sample point.

Ex.: THHTT is an example of a sample point.

• The set of possible outcomes is called the sample space. It is denoted by $\Omega$.

Ex.: $\Omega = \{HHHHH, HHHHT, ..., TTTTT\}$ and $|\Omega| = 32$. 
A **probability space** is a sample space $\Omega$, together with a probability $\Pr[\omega]$ for each sample point $\omega$, such that

- Each probability is between 0 and 1:
  \[ 0 \leq \Pr[\omega] \leq 1 \text{ for all } \omega \in \Omega. \]

- The sum of probabilities of all outcomes is 1:
  \[ \sum_{\omega \in \Omega} \Pr[\omega] = 1. \]

**Frequent example**: **uniform** sampling, that is, all sample points have the same probability, $\frac{1}{|\Omega|}$. 
Events

• Event is a subset of a sample space $\Omega$.
• Probability of event $E$ is the sum of probabilities of all elements of $E$:

$$\Pr[E] = \sum_{\omega \in E} \Pr[\omega].$$
### Probability space has three components

- Sample space $\Omega$
- Family of allowable events $E \subseteq \Omega$
- A probability function $\Pr$ that maps events $E$ to $\mathbb{R}$ such that:
  - $\Pr(E) \in [0,1]$ for any event $E$;
  - $\Pr(\Omega) = 1$;
  - For any finite or countable sequence of pairwise disjoint events $E_1, E_2, \ldots$,
    \[ \Pr\left( \bigcup_{i \geq 1} E_i \right) = \sum_{i \geq 1} \Pr(E_i). \]
How Compute Probabilities

1. Find sample space
2. Define events of interest
3. Determine outcome probabilities
   - For uniform sample spaces: \( \frac{1}{|\Omega|} \) for each outcome
4. Compute event probabilities
   - For uniform sample spaces: \( \Pr(E) = \frac{|E|}{|\Omega|} \)
Example: Dice

Roll two dice. What is the probability that

- the sum is at least 10?
- there is at least one 6?

1. \( \Omega = \{(i, j): 1 \leq i, j \leq 6\} \)

2. \( A = \text{event that sum is at least 10}; \)

\[ B = \text{event that there is at least one 6} \]

3. Uniform and \(|\Omega| = 36\), so all outcomes have probability \( \frac{1}{36} \)

4. \( \Pr(A) = \frac{|A|}{|\Omega|} = \)

\( \Pr(B) = \frac{|B|}{|\Omega|} = \)
Shuffle a deck of 52 cards. What is the probability that all red cards come before all black cards?

1. $\Omega$ is the set of all permutations (orderings) of the deck

2. $E = \text{event that all red cards come before all black cards}$;

3. Uniform and $|\Omega| = 52!$, so all outcomes have probability $\frac{1}{52!}$

4. $\Pr(E) = \frac{|E|}{|\Omega|} =$
Example: Poker Hands

Shuffle a deck of 52 cards and deal a 5-card hand (unordered). What is the probability that it is a flush (all cards have the same suit)?

1. \( \Omega \) is the set of all poker hands; \( |\Omega| = \binom{52}{5} \)
2. \( E \) = event that it is a flush;
   \( |E| = 4 \cdot \binom{13}{5} \)
3. Uniform and \( |\Omega| = \binom{52}{5} \), so all outcomes have probability \( \frac{1}{\binom{52}{5}} \)
4. \( \Pr(E) = \frac{|E|}{|\Omega|} = \)
Example: Fair Coins

Toss a fair coin \( n \) times. What is the probability that you get exactly \( n/2 \) heads?

1. \( \Omega \) is the set of all sequences of H’s and T’s of length \( n \)

2. \( E = \) set of all sequences with exactly \( n/2 \) heads.
   \[
   |E| = \binom{n}{n/2} \text{ if } n \text{ is even}
   \]

3. Uniform and \( |\Omega| = 2^n \), so all outcomes have probability \( \frac{1}{2^n} \)

4. \( \Pr(E) = \frac{|E|}{|\Omega|} = \)

\[
\text{Stirling’s approximation:}
\]
\[
n! \approx \sqrt{2\pi n} \left(\frac{n}{e}\right)^n
\]
**Example: Balls and Bins**

Throw 20 labeled balls into 10 labeled bins. Each ball is equally likely to land in any bin, regardless of what happens to other balls.

What is the probability that bin 1 is empty?

1. \( \Omega = \{(b_1, \ldots, b_{20}): 1 \leq b_i \leq 10 \text{ for each } i = 1, \ldots, 20\} \)  
   \(|\Omega| = 10^{20}\)

2. \( A = \text{set of sequences as above, but } 2 \leq b_i \leq 10 \)

3. Uniform and \( |\Omega| = 10^{20} \), so all outcomes have probability \( \frac{1}{10^{20}} \)

4. \( \Pr(E) = \frac{|A|}{|\Omega|} = \)
Axioms of Probability

• Probability function assigns a probability to each event
• A probability function must satisfy the following properties, called axioms of probability
  – Non-negativity: \( \Pr(E) \geq 0 \) for all events \( E \subseteq \Omega \)
  – Additivity: if \( A \) and \( B \) are disjoint events then
    \[
    \Pr(A \cup B) = \Pr(A) + \Pr(B)
    \]
  – Normalization: \( \Pr(\Omega) = 1 \)
We can derive many useful properties from the axioms

- **Complement rule:** \( \Pr(\overline{A}) = 1 - \Pr(A) \)

  **Proof:**
  
  \[
  \Pr(A) + \Pr(\overline{A}) = \Pr(A \cup \overline{A}) \\
  = \Pr(\Omega) \\
  = 1 
  \]

  - by additivity
  - by definition of complement
  - by normalization

\[A \text{ and } \overline{A} \text{ are disjoint}\]
We can derive many useful properties from the axioms

- **Difference rule:** \( \Pr(A/B) = \Pr(A) - \Pr(A \cap B) \)

**Proof:**

\[
\begin{align*}
(A/B) \cup (A \cap B) &= A \\
\Pr(A/B) + \Pr(A \cap B) &= \Pr(A)
\end{align*}
\]

by additivity

\( A/B \) and \( A \cap B \) are disjoint

Tiago Januario: Probability in Computing
We can derive many useful properties from the axioms

- **Monotonicity rule:** If \( A \subseteq B \) then \( \Pr(A) \leq \Pr(B) \)

**Proof:** \( A \cup (B/A) = B \)

\[
\Pr(A) + \Pr(B/A) = \Pr(B) \\
\Pr(A) = \Pr(B) - \Pr(B/A) \\
\leq \Pr(B)
\]

*by additivity*

*by non-negativity*

\( A \) and \( B \setminus A \) are disjoint
We can derive many useful properties from the axioms

- **Inclusion-Exclusion Principle:**
  \[ \Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \]

- **Union Bound:** \( \Pr(A \cup B) \leq \Pr(A) + \Pr(B) \)
• For a bill to come before the US president, it must be passed by both the House and the Senate.
• Suppose that 40% of bills pass the House, 30% the Senate, and 50% pass at least one of the two.

What is the probability the next bill will come before the president?

A. 0.2
B. 0.4
C. 0.5
D. 0.7
E. None of the above
A family has 2 children. Assume that each child is equally likely to be a boy or a girl. Which of the following is more likely?

A. They have 2 boys.
B. They have 2 girls.
C. They have 2 kids of different gender.
D. All the three possibilities above are equally likely.
Rank the likelihood of the following alternatives:

1) Linda is active in the feminist movement
2) Linda is a bank teller
3) Linda is a bank teller and active in the feminist movement

``Linda is 31, single, outspoken, and very bright. She majored in philosophy in college. As a student, she was deeply concerned with racial discrimination and other social issues, and participated in anti-nuclear demonstrations.``
Conjunction Fallacy

Rank the likelihood of the following alternatives:
1) Linda is active in the feminist movement
2) Linda is a bank teller
3) Linda is a bank teller and active in the feminist movement

Between 80 and 95 percent of the subjects ranked: (1) > (3) > (2)

Rank the likelihood of the following alternatives:
1) Linda is active in the feminist movement
2) Linda is a bank teller
3) Linda is a bank teller and active in the feminist movement
Conjunction Fallacy

Rank the likelihood of the following alternatives:

1) Linda is active in the feminist movement
2) Linda is a bank teller
3) Linda is a bank teller and active in the feminist movement