ONLINE MASTERS IN **DATA SCIENCE**

DSC 215 - PROBABILITY AND STATISTICS FOR DATA SCIENCE

INTRODUCTION TO PROBABILITY

UC San Diego

COMPUTER SCIENCE & ENGINEERING

HALICIOĞLU DATA SCIENCE INSTITUTE

Why Probability

- Probability is the foundation upon which statistics is built.

- Probability is also the foundation of machine learning, artificial intelligence, game theory, and information theory, among other areas related to data science.

- Formalizing probability concepts will help provide a deeper understanding of statistical tools and techniques that we will introduce afterwards.

Defining Probability Formally: the Sample Space

- Suppose you conduct an experiment that can produce a number of possible outcomes.

- This set of all possible outcomes is called the **sample space** of the experiment, and denoted Ω.

- **Examples**:

  - Rolling a die yields $\Omega = \{1,2,3,4,5,6\}$.

  - Flipping a coin yields $\Omega = \{\text{Heads, Tails}\}$.

Defining Probability Formally: the Event Space

- **The set of events, or event space** $\mathscr{F}$ is a set whose elements are themselves sets.
- Specifically every $A \in \mathscr{F}$ is a subset of Ω, i.e., $A \subset \Omega$.
- $\mathscr{F}$ must satisfy certain properties
  - The empty set $\varnothing$ must be in $\mathscr{F}$, i.e.,
    - $\varnothing \in \mathscr{F}.$
  - If $A \in \mathscr{F}$, then it's complement in Ω must be in $\mathscr{F}$, i.e.,
    - $A \in \mathscr{F} \implies A^c \in \mathscr{F}.$
  - If $A_1, A_2, \ldots$ are all in $\mathscr{F}$, then their union must be in $\mathscr{F}$, i.e.,
    - $A_1, A_2, \ldots \in \mathscr{F} \implies A = A_1 \cup A_2 \cup \ldots \in \mathscr{F}$

Defining Probability Formally: the Event Space

- **Example:** For any choice of sample space Ω, the set $\mathscr{F} = \{\emptyset, \Omega\}$ is an acceptable event space.

- **Example:** For any choice of sample space Ω, the power set $\mathscr{F} = P(\Omega)$ (i.e., the set of all subsets of Ω) is an acceptable event space.

- If $\Omega = \{1,2,3\}$, then$P(\Omega) = \big\{\emptyset, \{1\}, \{2\}, \{3\}, \{1,2\}, \{1,3\}, \{2,3\}, \{1,2,3\}\big\}$.

- **Example**: If $\Omega = \{1,2,3\}$, then $\mathscr{F} = \big\{\emptyset, \{1\}, \{2,3\}, \{1,2,3\}\big\}$ is also an acceptable event space.

Defining Probability Formally: Probability Measure

- **A probability measure** is a function $\mathbb{P}: \mathscr{F} \to [0,1]$.

- That is, it is a function that takes elements of $\mathscr{F}$ as input and produces numbers between 0 and 1 as output. In other words, it assigns probabilities to events.

- $\mathbb{P}$ must satisfy the axioms of probability:

  - $\mathbb{P}(A) \geq 0 \quad \forall A \in \mathscr{F}$.

  - $\mathbb{P}(\Omega) = 1$

  - If $A_1, A_2, \ldots$ are disjoint (i.e., $A_i \cap A_j = \varnothing \quad \forall i \neq j$) then
$$\mathbb{P}(A_1 \cup A_2 \cup \ldots) = \sum_i \mathbb{P}(A_i).$$

Defining Probability Formally: Probability Space

- **A probability space** consists of the triple $(\Omega, \mathscr{F}, \mathbb{P})$, in other words it consists of a sample space, an event space, and a probability measure.

- **Example** (rolling a die):

  - The sample space is $\Omega = \{1,2,3,4,5,6\}$.
  - The event space is $\mathscr{F} = P(\Omega) = \{\emptyset, \{1\}, \{2\}, \ldots, \{6\}, \{1,2\}, \{1,3\}, \ldots, \{1,2,3,4,5,6\}\}$
  - If a set $A \in \mathscr{F}$ has i elements, then $\mathbb{P}(A) = \dfrac{i}{6}$. So, for example $\mathbb{P}(\{1,4,6\}) = 3/6 = 0.5$.

Properties of Probability Measures

- Probability measures $\mathbb{P}$ satisfy certain important properties
  - $A \subset B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$.
  - $\mathbb{P}(A \cap B) \leq \min\left(\mathbb{P}(A), \mathbb{P}(B)\right)$.
  - $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$. This is known as the union bound.
  - $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$.
  - If $A_1, A_2, \ldots$ are disjoint sets with $A_1 \cup A_2 \cup \ldots = \Omega$, then $\sum_i \mathbb{P}(A_i) = 1$.
- For example, in our die rolling scenario
  - $\{1\} \subset \{1,2\}$ and $\mathbb{P}(\{1\}) = 1/6$, $\mathbb{P}(\{1,2\}) = 2/6$.
  - $\mathbb{P}(\{1\} \cap \{2,3\}) = 0 \leq \min\left(\mathbb{P}(\{1\}), \mathbb{P}(\{2,3\})\right)$
  - Can you check the other properties on the die rolling example?