Lecture 1: Events and their Probabilities

PhD Mathematics II: Probability

Juan Llavador Peralt

IIES

November 3, 2024

Outline for Today

Basic Measure Theoretic Probability

Conditional Probability

Independence

Example

- Roll two fair dice
- What is the probability that the sum of the two dice is 10?
- What are all the possible outcomes?
- Intuitively the probability of a 10 is $3/36 = 1/12$

Example

- Roll two fair dice
- What is the probability that the sum of the two dice is 10?
- What are all the possible outcomes?

$$\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)
\end{array}$$

- Intuitively the probability of a 10 is $3/36 = 1/12$

Example

- Roll two fair dice
- What is the probability that the sum of the two dice is 10?
- What are all the possible outcomes?

$$
\begin{array}{cccccc}
(1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\
(2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\
(3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\
(4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\
(5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\
(6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6)
\end{array}
$$

- Intuitively the probability of a 10 is $3/36 = 1/12$

Experiments, Sample Spaces, and Events

Experiment:

- Repeatable procedure with a well-defined set of possible outcomes

Sample Space:

- Set of all possible outcomes of an experiment
- Denoted by Ω
- $\omega \in \Omega$ is a generic outcome

Event:

- Subset of the sample space $A \subseteq \Omega$
- The complement of A is $A^c = \{\omega \in \Omega : \omega \notin A\}$

σ-algebra

- Let Ω be a non-empty set
- $\mathcal{F}$ is a collection of subsets of Ω
- $\mathcal{F}$ is a σ-**algebra** if:
  1. $\emptyset \in \mathcal{F}$
  2. $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
  3. $A_1, A_2, \dots \in \mathcal{F} \Rightarrow \cup_{i=1}^{\infty} A_i \in \mathcal{F}$
- We write $\sigma(A)$ to denote the smallest σ-algebra containing A

$$\sigma(A) = \{\emptyset, A, A^c, \Omega\}$$

- And $\mathcal{P}(\Omega)$ to denote the collection of all subsets of Ω (the power set)

Measurable Space and Measure

- A **measurable space** is a pair $(\Omega, \mathcal{F})$ where Ω is a set and $\mathcal{F}$ is a σ-algebra on Ω
- A **measure** is a function $\mu : \mathcal{F} \to [0, \infty]$ such that:
  1. $\mu(A) \geq 0$ for all $A \in \mathcal{F}$
  2. $\mu(\emptyset) = 0$
  3. If $A_1, A_2, \dots \in \mathcal{F}$ are pairwise disjoint, then

$$\mu\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

- Example: the counting measure $\mu(A) = |A|$ for $|\Omega| < \infty$

Probability Measure and Space

- A **probability measure** is a measure $\mathbb{P}$ on a measurable space $(\Omega, \mathcal{F})$ such that:
  1. $\mathbb{P} : \mathcal{F} \to [0, 1]$
  2. $\mathbb{P}(\Omega) = 1$
- The triplet $(\Omega, \mathcal{F}, \mathbb{P})$ is a **probability space**
- Example: $|\Omega| < \infty$ and $\mathbb{P}(A) := \frac{|A|}{|\Omega|}$
- Example: Indicator function is a probability measure:

$$\mathbb{P}(A) := \mathbf{1}_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

Kolmogorov's Probability Axioms

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space where $\mathbb{P}$ is a probability measure that satisfies:

1. **Non-negativity**: $\mathbb{P}(A) \geq 0$ for all $A \in \mathcal{F}$
2. **Normalization**: $\mathbb{P}(\Omega) = 1$
3. **Countable Additivity**: If $A_1, A_2, \dots \in \mathcal{F}$ are pairwise disjoint, then

$$\mathbb{P}\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Basic Properties

1. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
2. If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$
3. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

We can extend (3) to n events:

$$\mathbb{P}\left(\cup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) - \cdots + (-1)^{n+1} \mathbb{P}(A_1 \cap \cdots \cap A_n)$$

Boole's Inequality

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space
- Let $A_1, A_2, \dots \in \mathcal{F}$ be a countable sequence of events
- Then $\mathbb{P}\left(\cup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

Sketch of Proof:

- Define $B_1 = A_1$, $B_2 = A_2 \setminus B_1$, $B_3 = A_3 \setminus (B_1 \cup B_2)$, $\dots$, $B_n = A_n \setminus \cup_{i=1}^{n-1} B_i$
- Then $B_1, B_2, \dots$ are pairwise disjoint
- Furthermore $\cup_{i=1}^{\infty} A_i = \cup_{i=1}^{\infty} B_i$ (why?) and $B_i \subseteq A_i$
- Then $\mathbb{P}\left(\cup_{i=1}^{\infty} A_i\right) = \mathbb{P}\left(\cup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$

Continuity of Probability Measure

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

- Let $A_1, A_2, \ldots$ be a sequence of events such that $A_1 \subseteq A_2 \subseteq \ldots$
- This is called an **increasing** sequence of events, define $\lim_{n\to\infty} A_n = \cup_{i=1}^{\infty} A_i$
- Then $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mathbb{P}(A_n)$

Similarly,

- Let $A_1, A_2, \ldots$ be a sequence of events such that $A_1 \supseteq A_2 \supseteq \ldots$
- This is called a **decreasing** sequence of events, define $\lim_{n\to\infty} A_n = \cap_{i=1}^{\infty} A_i$
- Then $\mathbb{P}(\cap_{i=1}^{\infty} A_i) = \lim_{n\to\infty} \mathbb{P}(A_n)$

Proof Sketch

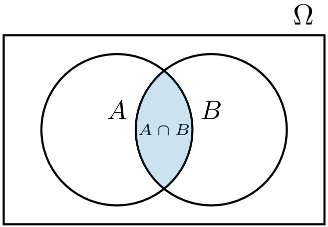
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $A_1, A_2, \dots$ be an increasing sequence of events

- Define $B_1 = A_1$, $B_2 = A_2 \setminus A_1$, $B_3 = A_3 \setminus A_2$, ..., $B_n = A_n \setminus A_{n-1}$
- Then $B_1, B_2, \dots$ are pairwise disjoint
- And $A_n = \cup_{i=1}^n B_i$
- By countable additivity,

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{P}(B_i) = \mathbb{P}\left(\cup_{i=1}^{\infty} B_i\right) = \mathbb{P}\left(\cup_{i=1}^{\infty} A_i\right)$$

Conditional Probability

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space
- Let $A, B \in \mathcal{F}$ be events such that $\mathbb{P}(B) > 0$
- Question: what is the probability that A occurs given that B has occurred?



- The **conditional probability** of A given B is defined as

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

Example

- Let $|\Omega| < \infty$ and $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$
- Then
$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{|A \cap B|/|\Omega|}{|B|/|\Omega|} = \frac{|A \cap B|}{|B|}$$
- In fact, one can show that $\mathbb{P}(\cdot|B) : \mathcal{F} \to [0,1]$ is a probability measure

Application: Simultaenous Events

- Suppose we care about the probability that both A and B occur: $\mathbb{P}(A \cap B)$
- We can use the definition of conditional probability to write

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$$

- Generalizes:

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A|B \cap C)\mathbb{P}(B \cap C) = \mathbb{P}(A|B \cap C)\mathbb{P}(B|C)\mathbb{P}(C)$$

- And so on to n events (see next slide)

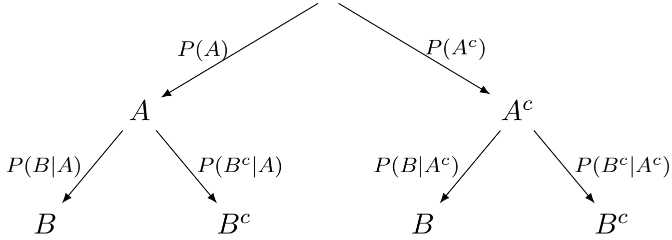
The Multiplication Law

- Let $A_1, A_2, \ldots, A_n \in \mathcal{F}$ be events
- Define $A_0 = \Omega$
- Then

$$\mathbb{P}(\cap_{i=1}^{n} A_i) = \prod_{i=1}^{n} \mathbb{P}(A_i | \cap_{j=0}^{i-1} A_j)$$

Preamble to Bayes' Theorem and Law of Total Probability

- Let $A, B \in \mathcal{F}$



- $\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(B|A) + \mathbb{P}(A^c)\mathbb{P}(B|A^c)$
- $\mathbb{P}(B|A) = \mathbb{P}(B)\frac{\mathbb{P}(A|B)}{\mathbb{P}(A)}$

Partitions (of the Sample Space)

- Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$
- The set $\{A_1, A_2, \ldots, A_n\} \in \mathcal{F}$ is a **partition** of Ω if the following conditions hold:

  1. (**Exhaustive**) $\cup_{i=1}^{n} A_i = \Omega$
  2. (**Mutual Exclusivity**) $A_i \cap A_j = \emptyset$ for all $i \neq j$
  3. (**Non-zero probability**) $\mathbb{P}(A_i) > 0$ for all i

- Note that 3. is not strictly necessary, but convenient

Law of Total Probability

- Let $\{A_1, A_2, \ldots, A_n\}$ be a partition of Ω
- Then for any event $B \in \mathcal{F}$, we have

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i)$$

Proof of the Law of Total Probability

- Let $\{A_1, A_2, \ldots, A_n\}$ be a partition of Ω

$$\begin{aligned}\mathbb{P}(B) &= \mathbb{P}(B \cap \Omega)\\ &= \mathbb{P}(B \cap \cup_{i=1}^n A_i)\\ &= \mathbb{P}(\cup_{i=1}^n B \cap A_i)\\ &= \sum_{i=1}^n \mathbb{P}(B \cap A_i)\\ &= \sum_{i=1}^n \mathbb{P}(B|A_i)\mathbb{P}(A_i)\end{aligned}$$

- the thrid equality follows from the distributive property of the intersection over the union:

$$B \cap \cup_{i=1}^n A_i = \cup_{i=1}^n B \cap A_i$$

Bayes' Theorem

- Let $A, B \in \mathcal{F}$ be events such that $\mathbb{P}(A), \mathbb{P}(B) > 0$
- Let $\{A_1, A_2, \ldots, A_n\}$ be a partition of Ω
- Then:

1. $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)}$
2. $\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A)+\mathbb{P}(B|A^c)\mathbb{P}(A^c)}$
3. $\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_i)}{\sum_{i=1}^{n}\mathbb{P}(B|A_i)\mathbb{P}(A_i)}$

▸ Monty Hall Example

Independence

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space
- Let $A, B \in \mathcal{F}$ be events
- A and B are **independent** if and only if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

- Sometimes written as $A \perp B$
- Intuitively, if A is independent of B:

$$\mathbb{P}(A|B) = \mathbb{P}(A)$$

or

$$\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \mathbb{P}(A)$$

Pairwise and Mutual Independence

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space
- Let $A_1, A_2, \ldots, A_n \in \mathcal{F}$ be events
- What does it mean to say that $\{A_1, A_2, \ldots, A_n\}$ are independent?
- The events are **pairwise independent** if and only if for all $i \neq j$

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

- The events are **mutually independent** if and only if for all $I \subseteq \{1, 2, \ldots, n\}$

$$\mathbb{P}(\cap_{i \in I} A_i) = \prod_{i \in I} \mathbb{P}(A_i)$$

Monty Hall Problem

- You are a contestant on a game show hosted by Monty Hall
- There are three (ex-ante identical) doors
- Behind one of the doors is a car, behind the other two are goats
- Monty asks you to choose a door
- Then, Monty opens another door revealing a goat
- You are given the option to switch doors
- Should you switch?

Monty Hall Problem

- Suppose the three doors are numbered 1, 2, and 3
- Let us define the following events:
  - B_1, B_2, B_3: the car is behind door 1, 2, or 3
  - M_1, M_2, M_3: Monty opens door 1, 2, or 3
- Suppose (WLOG) you choose door 1 and Monty opens door 2
- Doors 1 and 3 are now the only possibilities
- You should switch if $\mathbb{P}(B_3|M_2) > \mathbb{P}(B_1|M_2)$

Monty Hall Problem

- Let's use Bayes' Theorem to calculate $\mathbb{P}(B_1|M_2)$

$$\mathbb{P}(B_1|M_2) = \frac{\mathbb{P}(M_2|B_1)}{\mathbb{P}(M_2|B_1)\mathbb{P}(B_1) + \mathbb{P}(M_2|B_2)\mathbb{P}(B_2) + \mathbb{P}(M_2|B_3)\mathbb{P}(B_3)}\mathbb{P}(B_1)$$

- The doors are ex-ante identical: $\mathbb{P}(B_1) = \mathbb{P}(B_2) = \mathbb{P}(B_3) = 1/3$
- Monty will never open the door with the car: $\mathbb{P}(M_2|B_1) = 1/2$, $\mathbb{P}(M_2|B_2) = 0$, $\mathbb{P}(M_2|B_3) = 1$
- Therefore, $\mathbb{P}(B_3|M_2) = 2/3$
- $\mathbb{P}(B_1|M_2) = 1 - \mathbb{P}(B_3|M_2) = 1/3$
- You should switch!