Lecture 3: Probability Basics (I)

Ye Tian

Department of Statistics, Columbia University
Calculus-based Introduction to Statistics (S1201)

July 7, 2022

Review: data visualization tools

- Continuous (numerical) data:
  - Stem-and-leaf plot
  - Dot plot
  - Histogram
  - Box plot
- Categorical (discrete) data:
  - Bar chart
  - Pie chart (be careful when using it)

Review: shape of data distribution

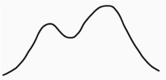
- Modality:
  - Unimodal
  - Bimodal
  - Multimodal
  - Almost uniform (no obvious peaks)
- Skewness:
  - Left-skewed
  - Right-skewed
  - Symmetric

Review: shape of data distribution

- modality

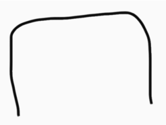
unimodal

bimodal



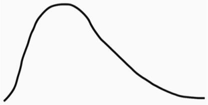
multimodal

uniform



- skewness

right skew



left skew

symmetric

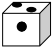
Today's goal

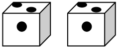
- Understand important concepts that are relative to probability: a random experiment, the sample space, outcomes, events and their relationship
- Know the definition and interpretation of probability and three probability axioms
- Know how to calculate probability in the experiment with equal likely outcomes

Warning: We will see more MATH from today. Be prepared!

Concepts before We Define Probability

Experiments and outcomes

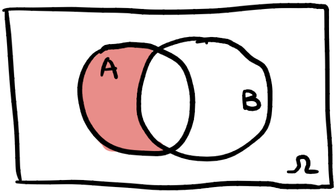
- A **(random) experiment** is any action or process whose outcome is uncertain. E.g. rolling a die 
- An **outcome** is the result of a random experiment. E.g. the die lands with face 4 up
- The **sample space** of an **experiment**, denoted by Ω, is the **set** of all possible outcomes of that experiment. E.g.: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- An **event** is a set (collection) of outcomes. E.g. consider event $A = \{1\}$, event $B = \{3, 5, 6\}$

Exercise: What is the sample space of rolling **two dice**? 

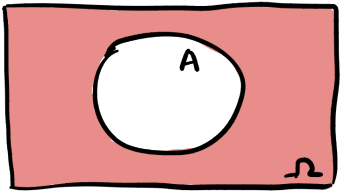
Event (set) operations from set theory

Consider events A and B:

- Set **difference** $A - B$ or $A \backslash B$: outcomes that are in A but not in B
- The **complement** of A: denoted as A^c, which is $\Omega - A$



$A - B$

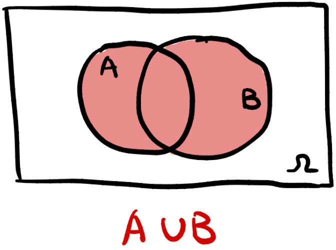


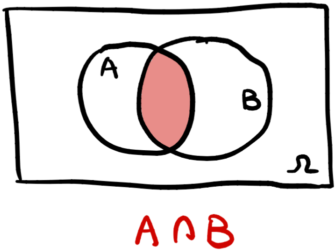
$A^c = \Omega - A$

Event (set) operations from set theory

Consider events A and B:

- The **union** of A and B: denoted as $A \cup B$, which is the event consisting of all outcomes that are in **either** A **or** B
- The **intersection** of A and B: denoted as $A \cap B$, which is the event consisting of all outcomes that are in **both** A and B.

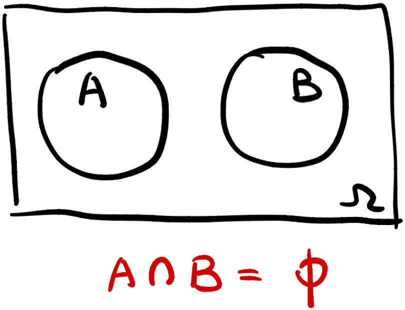




Event (set) operations from set theory

Consider events A and B:

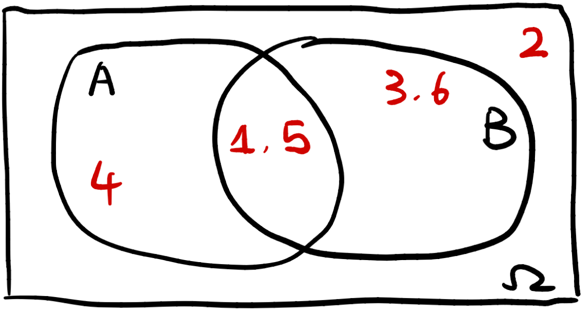
- Let $\emptyset$ denote the **empty set**, which contains no outcomes
- When $A \cap B = \emptyset$, A and B are said to be **mutually exclusive** or **disjoint** events



Event (set) operations from set theory

Example: Consider rolling a die. Events $A = \{1, 4, 5\}$, $B = \{1, 3, 5, 6\}$

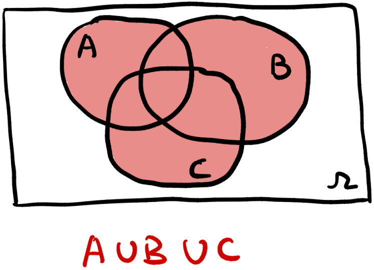
- $A \cup B =?$, $A - B =?$, $A \cap B =?$
- $A^c =?$, $B^c =?$
- $A^c \cap B =?$
- Are A and B mutually exclusive?

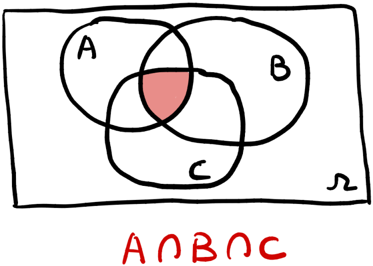


Event (set) operations from set theory: ≥ 3 events

Definitions of union, intersection and mutually exclusive events can be extended to the case of multiple events. Consider events A, B and C:

- The **union** of A, B and C: denoted as $A \cup B \cup C$, which is the event consisting of all outcomes that are in **either** A **or** B **or** C
- The **intersection** of A, B and C: denoted as $A \cap B \cap C$, which is the event consisting of all outcomes that are in **all** A, B **and** C.

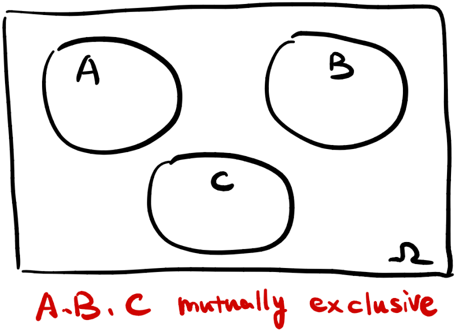




Event (set) operations from set theory: ≥ 3 events

Definitions of union, intersection and mutually exclusive events can be extended to the case of multiple events. Consider events A, B and C:

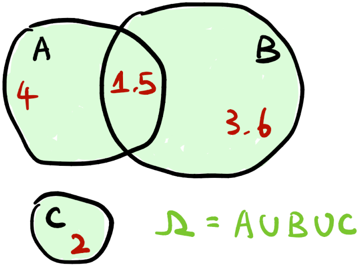
- When no two events of A, B and C have any outcomes in common (i.e. $A \cap B = A \cap C = B \cap C = \emptyset$), they are said to be **mutually exclusive** events (or **(pairwisely) disjoint**)

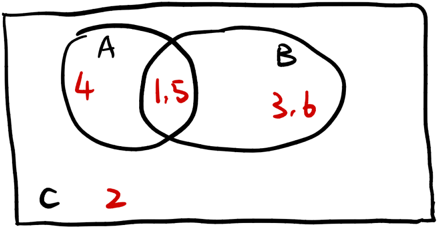


Event (set) operations from set theory: ≥ 3 events

Example: Consider rolling a die. Sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$.
Events $A = \{1, 4, 5\}$, $B = \{1, 3, 5, 6\}$, $C = \{2\}$.

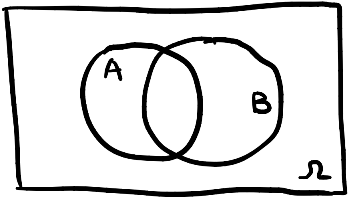
- $A \cup B \cup C =?$
- $(B^c \cap A) \cup C =?$
- $(A \cup B)^c \cap C =?$
- Are A and C mutually exclusive? What about A, B and C?

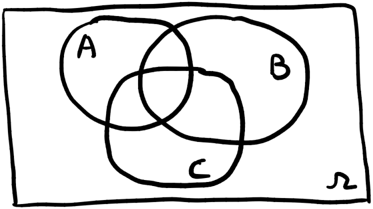




Event (set) operation properties

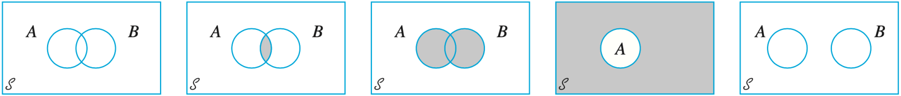
- When calculating union, intersection and complement involved with ≥ 3 events (sets), these properties may be helpful:
  - Commutative property: $A \cup B = B \cup A$, $A \cap B = B \cap A$
  - Associative property: $(A \cup B) \cup C = A \cup (B \cup C)$
  - Distributive property:
    - $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
    - $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
  - de Morgan rule:
    - $(A_1 \cup ... \cup A_n)^c = A_1^c \cap ... \cap A_n^c$
    - $(A_1 \cap ... \cap A_n)^c = A_1^c \cup ... \cup A_n^c$





Venn diagrams

A pictorial representation of events and manipulations with events is obtained by using **Venn diagrams**.



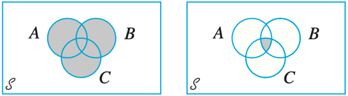
(a) Venn diagram of events A and B

(b) Shaded region is $A \cap B$

(c) Shaded region is $A \cup B$

(d) Shaded region is A'

(e) Mutually exclusive events



(f) Shaded region is $A \cup B \cup C$

(g) Shaded region is $A \cap B \cap C$

Probability, Axioms and Interpretations

Probability

Given an experiment and a sample space Ω, we would like to know the likelihood that an event can occur, which is defined as **probability**.

- The objective of probability is to assign to each event A a number $\mathbb{P}(A)$, called the probability of the event A
  From this perspective: $\mathbb{P}(\cdot)$: $A \in$ a set of all events $\rightarrow$ a number $\mathbb{P}(A)$ (a mapping/function which maps an event to a number)
- According to your intuition/common sense, if we roll a fair die, what's the probability of getting 2 face up (i.e. $\mathbb{P}(\{2\})$)? Can it be $1/5$?
- Given an experiment, the probability is a **determined** function. It is an intrinsic function related to the random experiment and it cannot be arbitrary!

Axioms of probability

Given any experiment and a sample space Ω:

- $0 \leq \mathbb{P}(A) \leq 1$ for any event $A \subseteq \Omega$
- $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$
- If A_1, A_2, A_3, ..., A_n are disjoint (mutually exclusive) events, then

$$\mathbb{P}(A_1 \cup A_2 \cup \ldots \cup A_n) = \sum_{i=1}^{n} \mathbb{P}(A_i).$$

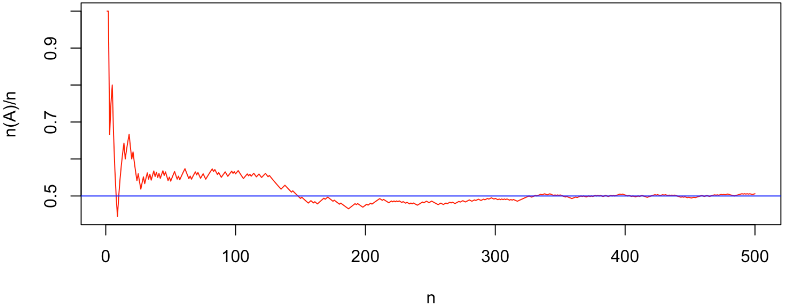
  Here n can be $+\infty$ (an infinite collection).

Any probability **must** satisfy three axioms above. Do you think these axioms make sense? Can you explain why?

Interpretations of probability

Now we are tossing a fair coin and denote event A as "the heads face up". We know $\mathbb{P}(A) = 0.5$.

Now we repeat this game for n times and denote $n(A)$ as the number of replications on which A occurs. We call $n(A)/n$ as **relative frequency**.



The relative frequency of an event will converge to its probability when $n \to +\infty$! In other words, the probability of an event can be seen as its **long-run relative frequency**.

Probability Calculation

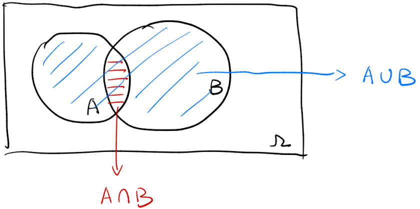
Properties of probability (I)

Suppose A, B and C are some events.

- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
  $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B)$
- $\mathbb{P}(A \cup B \cup C) =$
  $\mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)$
  $\mathbb{P}(A \cap B \cap C) =$
  $\mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cup B) - \mathbb{P}(A \cup C) - \mathbb{P}(B \cup C) + \mathbb{P}(A \cup B \cup C)$

Proof:

- A and A^c are mutually exclusive; $A \cup A^c = \Omega$; Then by Axioms 2 and 3
- Venn diagram:



Properties of probability (II)

Suppose A and B are some events.

- If A and B are mutually exclusive, then $\mathbb{P}(A \cap B) = 0$
- If $A \subseteq B$, then $\mathbb{P}(A) \leq \mathbb{P}(B)$

Proof:

- $A \cap B = \emptyset$; Then by Axiom 2
- $B = (A \cap B) \cup (A^c \cap B) = A \cup (A^c \cap B)$; A and $A^c \cap B$ are mutually exclusive; Then by Axiom 3. (Why does "=" hold?)

Exercise

Suppose A and B are some events, and $\mathbb{P}(A) = 0.2$, $\mathbb{P}(B) = 0.9$.

- Is it possible for A and B to be mutually exclusive?
- Give the range for all possible values of $\mathbb{P}(A \cup B)$ and $\mathbb{P}(A \cap B)$
- Can you illustrate the most "extreme" scenarios (when $\mathbb{P}(A \cup B)$ and $\mathbb{P}(A \cap B)$ achieve maximum or minimum) by Venn diagram?

Proof:

- No. If that's true, then $1 \geq \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) = 1.1$!
- $B \subseteq A \cup B$, therefore $0.9 = \mathbb{P}(B) \leq \mathbb{P}(A \cup B) \leq 1$. Hence $\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) = 1.1 - \mathbb{P}(A \cup B) \in [0.1, 0.2]$ Check: give examples corresponding to the endpoints to verify that these numbers are indeed achievable.

Equally likely outcomes

In many experiments consisting of N outcomes, it is reasonable to assign equal probabilities to all N simple events.

Example:

- tossing a fair coin ($\Omega = \{\mathsf{H}, \mathsf{T}\}, N = 2$) or fair die ($\Omega = \{1, 2, \ldots, 6\}$, $N = 6$)
- selecting one or several cards from a well-shuffled deck of 52 ($N = 52$)

Denote event $A_i = \{i\text{-th outcome}\}$, since $A_1, \ldots, A_N$ are mutually exclusive, $1 = \mathbb{P}(\Omega) = \mathbb{P}(\cup_{i=1}^{N} A_i) = \sum_{i=1}^{N} \mathbb{P}(A_i)$, implying that $\mathbb{P}(A_i) = 1/N$ for $i = 1, \ldots, N$

Consequence: Denote the number of outcomes that event A contains as $N(A)$. Then for any event A,

$$\mathbb{P}(A) = \frac{N(A)}{N}.$$

Equally likely outcomes

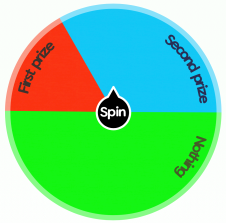
Theorem: In an experiment consisting of N outcomes **with equal probability**, denote the number of outcomes that event A contains as $N(A)$. Then for any event A,

$$\mathbb{P}(A) = \frac{N(A)}{N}.$$

Remark:

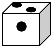
- This result can only be used in the experiment where each outcome has **equal** probability

  Counter-example: Tossing an unfair coin/die, spinning the wheel on the right

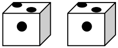


- For such experiments, calculating probability = counting the number of outcomes included in the event!

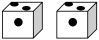
Example 1

We are rolling a fair die . $\Omega = \{1, \ldots, 6\}$.

- What's the probability that we get 2 face up?
  $\mathbb{P}(\{2\}) = 1/N = 1/6$
- What's the probability that we get 2 or 5 face up?
  $\mathbb{P}(\{2, 5\}) = 2/N = 2/6 = 1/3$

Now consider two fair dice .

Example 2

Now consider two fair dice . $\Omega = \{(i,j) : 1 \le i,j \le 6\}$,
$N = |\Omega| = 6 \times 6 = 36$

Die 2

Die 1	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

- What's the probability that one die is 4 and the other is 1?
  $A = \{(4,1),(1,4)\}$, $\mathbb{P}(A) = 2/N = 2/36 = 1/18$
- What's the probability that the sum of two dice is more than 9?
  $B = \{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\}$
  $\mathbb{P}(B) = 6/N = 6/36 = 1/6$

Reading list (optional)

- "Probability and Statistics for Engineering and the Sciences" (9th edition):
  - Chapter 2.1 and 2.2
- "OpenIntro statistics" (4th edition, free online, download `[here]`):
  - Chapter 3.1.1-3.1.4

Many thanks to

- Yang Feng
- Joyce Robbins
- Chengliang Tang
- Owen Ward
- Wenda Zhou
- And all my teachers in the past 25 years