Probability Theory, pt. I

ECON 3640–001

Marcio Santetti

Spring 2022

Motivation

Housekeeping

Notes based on `Blitzstein & Hwang (2019)`, ch. 1

- sections 1.1—1.3, and 1.6—1.7

Motivation

Check out this video:

`Teach statistics before calculus!, by Arthur Benjamin`

Motivation

A quick **quiz**:

Select the statement (a, b, or c) that you agree with most strongly. **There are no wrong answers here!**

How do you interpret the following: "*When flipping a coin, the probability of flipping Heads is 0.5*"?

- (a) If I flip this coin over and over and over and over, roughly 50% of the flips will be Heads.
- (b) If I flip this coin, Heads / Tails are equally plausible.
- (c) Both of the above make sense.

Motivation

A quick **quiz**:

Select the statement (*a*, *b*, or *c*) that you agree with most strongly. **There are no wrong answers here!**

An election is coming up and a pollster claims that "*candidate A has a 0.9 probability of winning.*" How do you interpret this probability?

- (*a*) If we observe the election over and over, candidate A will win roughly 90% of the time.
- (*b*) Candidate A is much more likely to win than to lose.
- (*c*) The pollster's calculation is wrong. Candidate A will either win or lose, thus their probability of winning can only be 0 or 1.

Motivation

A quick **quiz**:

Select the statement (*a* or *b*) that you agree with most strongly. **There are no wrong answers here!**

In a survey with 10 students, they are asked the following 2 questions:

- Will you vote for a Republican in 2024? 10 out of 10 say "yes."
- Each student is given a sample of Pepsi and Coke. 10 out of 10 correctly identify which is which.

In light of these experiments, what do you conclude?

- (*a*) You're more confident that students can distinguish between Coke and Pepsi than that students plan to vote for a Republican candidate.
- (*b*) he evidence in favor of students' intention to vote a Republican is just as strong as the evidence in favor of students' ability to distinguish between Coke and Pepsi.

Motivation

If your answers were (a), (a), and (b), you are fairly **frequentist**.

If your answers were (b), (b), and (a), you are a **Bayesian**.

If your answers to the first two questions were (c) and (c), you have **no side!**

Studying probability

Studying probability

You have probably either used or heard the words *luck*, *coincidence*, *odds*, *chance* in your life to reflect some **guess** or **prediction** about future events.

These can also be understood as **probability statements**.

- Perhaps you don't usually **quantify** these guesses!

Whether you translate such predictions into numbers or not, in the next lectures we will study probability as a **logical path** to measure *uncertainty* and *randomness* through *theoretical principles*.

Studying probability

Probability is the **logic of uncertainty**.

It quantifies uncertainty in a way through which we may

- formalize such uncertainty;
- make informed decisions about uncertain events;
- make inferences about noisy processes and population parameters;
- better understand our surroundings.

Basics of set theory

Basics of set theory

The mathematical framework for probability is built around **sets**.

Suppose we run an *experiment*, and its result is one out of a set of possible *outcomes* (*events*).

Before one can quantify the inherent *randomness* in this experiment, one must understand its possible outcomes.

Consider some definitions:

- A **sample space** (S) is the collection of all possible outcomes of an experiment;
- An **event** is an element or collection of elements in the sample space, typically denoted by capital letters (e.g., A, B, C).
  - Thus an event is a *subset* of the sample space S: $A \subseteq S$.

Basics of set theory

Some **special events**:

- Empty (null) set ($\varnothing$): contains **no** outcomes.
- The complement of an event (A^C): contains **all** outcomes that are not in event A.
- Intersection between events ($A \cap B$): contains all outcomes that are in **both** A and B.
  - A and B are **disjoint** if $A \cap B = \varnothing$.
- Union between events ($A \cup B$): contains all outcomes that are in A **or** B **or** both.
- De Morgan's laws:
  - $(A \cup B)^C = A^C \cap B^C$
  - $(A \cap B)^C = A^C \cup B^C$

Many shades of probability

Many shades of probability

Now it is time to study the **uncertainty** associated with the possible outcomes of an experiment.

To this purpose, we will need **probability**.

Informally, one may define probability as *a measure of uncertainty, ranging between 0 (impossible outcome) and 1 (almost certain event).*

But there are more **formal** ways of defining it:

- The "*naive*" definition of probability;
- The "*frequency*" definition of probability;
- The "*Bayesian*" definition of probability.

Many shades of probability

The **earliest** definition of the probability of an event was to **count** the number of *ways* the event could happen and divide by the *total* number of possible outcomes for the experiment.

$$P^{naive}(A) = \frac{\text{number of ways A can happen}}{\text{total number of outcomes in } S}$$

The naive definition is very **restrictive**.

- It requires S to be *finite*, with equal symmetry across events/outcomes.
- Events must be *equally likely*!

Many shades of probability

Given the **limitations** of the naive definition of probability, we will move on to more general and effective concepts.

> A *probability space* consists of a sample space S and a *probability function* P which takes an event $A \subseteq S$ as input and returns $P(A)$, a real number between 0 and 1, as output.

The function $P(\cdot)$ must satisfy the following **two** axioms:

- $P(\varnothing) = 0$;
- $P(S) = 1$

If events $A_1, A_2, A_3, \ldots$ are **disjoint**, then

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

Many shades of probability

Any function $P(\cdot)$ that satisfies the two axioms is considered a **valid** probability function.

However, the axioms don't tell us how probability should be **interpreted**.

There are different **schools of thought** regarding such interpretation.

The **frequentist** view of probability is that it represents a long-run frequency over a large number of repetitions of an experiment: if we say a coin has probability 1/2 of Heads, that means the coin would land Heads 50% of the time if we tossed it over and over and over.

The **Bayesian** view of probability is that it represents a degree of belief about the event in question, so we can assign probabilities to hypotheses like "candidate A will win the election" or "the defendant is guilty" even if it is not possible to repeat the same election or the same crime over and over again.

Next time: Properties of probability; conditional probability