Probability Theory, pt. II

ECON 3640–001

Marcio Santetti

Spring 2022

Motivation

Housekeeping

Notes based on `Blitzstein & Hwang (2019)`, ch. 2

- sections 2.1 & 2.2.

Motivation

Last time, we started our study of probability, which is the **logic of uncertainty**.

Despite its different interpretations (e.g., frequentists *vs.* Bayesians), the **theoretical foundations** for both approaches are the same.

Now, we use our knowledge of set theory to get to some crucial **properties** of probability.

Properties of probability

Properties of probability

From the **axioms** studied last time

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

We can derive the following **properties of probability**:

1. **Complement rule**: $P(A^C) = 1 - P(A)$

2. **Subset rule**: If $A \subseteq B$, then $P(A) \leq P(B)$

3. **Inclusion-exclusion rule**: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Properties of probability

Suppose a sample of voters[1] was asked whether they voted for a **Democrat** (D) or **other** candidate (ND) in the last election.

Then, each interviewee was shown a **fake news** article. 81% of them believed in the headline.

Out of those who voted for the **Democrat** candidate, 30% believed in the headline.

Overall, 48% of the sample vote for the **Democrat** candidate.

Convert the above information into *probability statements* and apply the studied *properties*.

1: Example inspired by `this article`.

Conditional probability

Conditional probability

Probability is a language for expressing our **degrees of belief** or uncertainties about events.

As **new evidence** (i.e., information) is incorporated, this may affect our uncertainty about specific events.

Q: This way, how should we **update** our beliefs in light of the evidence we observe?

A: Through **conditional probabilities**.

Conditional probability

Conditional probability shows how to *incorporate* evidence into our understanding of the world in a **logical**, *coherent* manner.

In fact, one can say that **all** probabilities are conditional to some degree, since there is always **background knowledge** built into every probability.

Suppose we **have not** turned on the TV yet and our team is playing. Our assessment about the team winning can be stated as $P(W)$.

We then **turn the TV on** and our team is *losing*.

Presumably, our belief that the team will win should *decrease*.

We may denote this new probability by $P(W|TV)$.

Going from $P(W)$ to $P(W|TV)$ means that we are **conditioning on** TV.

Conditional probability

> "*Conditioning is the soul of statistics.*" (Blitzstein & Hwang, 2019, p. 42)

If A and B are events, and $P(B) > 0$, then the **conditional probability** of A given B, denoted by $P(A|B)$, is defined as

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

where A is the event whose uncertainty we wish to update, and B is the observed evidence (information) we bring in.

Before we see any new data/information, $P(A)$ is our **prior probability** of A.

Consequently, $P(A|B)$ is the **posterior probability** of A.

Conditional probability

Important warning: $P(A|B) \neq P(B|A)$!

Confusing these two quantities is called the "*prosecutor's fallacy*."

Q: How would a *frequentist* interpret $P(A|B)$?

Q: How would a *Bayesian* interpret $P(A|B)$?

Independence

Independence

Events A and B are **independent** if

$$P(A|B) = P(A)$$

This implies that *any knowledge* of B does not help/affect/inform one's belief about A.

Also, from the **conditional probability formula**, independence implies

$$P(A|B) = P(A)P(B)$$

Next time: Law of total probability; Bayes' Theorem