Main class 2: Defining probability, Conditional Probability

Izzy Grosof

September 19, 2025

1 Takeaways

- An *experiment* is a random scenario.
- An *outcome* ω is one possible way that experiment could turn out.
- The *sample space* Ω is the set of all possible outcomes. Written symbolically, $\omega \in \Omega$. "$\in$" means "in" or "element of".
- An *event* E is a subset of the possible outcomes. Written symbolically, $E \subset \Omega$. "$\subset$" means "subset of".
- We will use the following definition of probability in this class, introduced by Andrey Kolmogorov, called the "three axioms of probability":

  Definition 1. A probability $\mathbb{P}$ is defined to be a function mapping *events* (subsets of sample space $E \subset \Omega$) to *real numbers* that satisfies the following three properties (axioms):

  Non-negativity $\mathbb{P}\{E\} \geq 0$, for any event E.

  Normalization $\mathbb{P}\{\Omega\} = 1$, where Ω is the entire sample space of all possible outcomes.

  Countably-infinite Additivity If $E_1, E_2, E_3 \ldots$ is a finite or countably-infinite sequence of events, where each pair of events is disjoint ($E_i \cap E_j = \emptyset$, $\forall i \neq j$), then

  $$\mathbb{P}\{E_1 \cup E_2 \cup E_3 \cup \ldots\} = \mathbb{P}\{E_1\} + \mathbb{P}\{E_2\} + \mathbb{P}\{E_3\} + \ldots.$$

  $\cup$ means "Union", $\cap$ means "Intersection". "countably infinite" means "infinite like the integers, not like the real numbers".

- Conditional Probability – $\mathbb{P}\{A \mid B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}$
- Law of Total Probability – Adding up cases, using the conditional probabilities in each case.

  $$\mathbb{P}\{B\} = \mathbb{P}\{A \cap B\} + \mathbb{P}\{\overline{A} \cap B\} = \mathbb{P}\{A\}\mathbb{P}\{B \mid A\} + \mathbb{P}\{\overline{A}\}\mathbb{P}\{B \mid \overline{A}\}$$

- Bayes rule – From $\mathbb{P}\{A \mid B\}$ to $\mathbb{P}\{B \mid A\}$.

  $$\mathbb{P}\{B \mid A\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{A\}} = \frac{\mathbb{P}\{A \mid B\}\mathbb{P}\{B\}}{\mathbb{P}\{A\}}$$

- Independence – If $A \perp B$, then $\mathbb{P}\{A \mid B\} = \mathbb{P}\{A\}$.

2 Defining probability

- An *experiment* is a random scenario.
- An *outcome* ω is one possible way that experiment could turn out.
- The *sample space* Ω is the set of all possible outcomes. Written symbolically, $\omega \in \Omega$. "$\in$" means "in" or "element of".
- An *event* E is a subset of the possible outcomes. Written symbolically, $E \subset \Omega$. "$\subset$" means "subset of".
- We will use the following definition of probability in this class, introduced by Andrey Kolmogorov, called the "three axioms of probability":

  Definition 2. A probability $\mathbb{P}$ is defined to be a function mapping *events* (subsets of sample space $E \subset \Omega$) to *real numbers* that satisfies the following three properties (axioms):

  Non-negativity $\mathbb{P}\{E\} \geq 0$, for any event E.

  Normalization $\mathbb{P}\{\Omega\} = 1$, where Ω is the entire sample space of all possible outcomes.

  Countably-infinite Additivity If $E_1, E_2, E_3 \ldots$ is a finite or countably-infinite sequence of events, where each pair of events is disjoint ($E_i \cap E_j = \emptyset$, $\forall i \neq j$), then

  $$\mathbb{P}\{E_1 \cup E_2 \cup E_3 \cup \ldots\} = \mathbb{P}\{E_1\} + \mathbb{P}\{E_2\} + \mathbb{P}\{E_3\} + \ldots.$$

  $\cup$ means "Union", $\cap$ means "Intersection". "countably infinite" means "infinite like the integers, not like the real numbers".

Summary:
Ex: Flip 5 coins. Ex outcome: Sequence of heads and tails, HHTHT Ex event: Exactly 3 heads.
Events either happen or they don't – yes or no.
Events have probabilities.
Goals for today:

- Conditional probability
- Law of total probability
- Bayes' rule
- Independence

3 Conditional Probability – 8 min (10:05 - 10:13)

"Probability of event A, given that event B occurred".

Important – say you're running a manufacturing process, and you previously calculated the chance that a component would fail before next scheduled maintenance, and it was low enough for comfort. But now, you've found some small cracks in the component, and you need to know the conditional probability: Fail soon given little cracks.

"Probability of A given B".

$\mathbb{P}\{A \mid B\}$.

Draw picture: Venn diagram, A and B overlapping, dots all around.

Intersection: $A \cap B$

To calculate:

$$\mathbb{P}\{A \mid B\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{B\}}$$
$$\mathbb{P}\{A \cap B\} = \mathbb{P}\{A \mid B\}\mathbb{P}\{B\}$$

Note: $\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\&B\} = \mathbb{P}\{A \textit{ and } B\}$.

4 Law of total probability – 8 min (10:13 - 10:21)

Last formula was great if you know unconditional probabilities $\mathbb{P}\{A \cap B\}$ and $\mathbb{P}\{B\}$, and you want to know conditional probabilities.

What if we know conditional probabilities, and we want to know unconditional probabilities?

Same diagram.

$$\mathbb{P}\{B\} = \mathbb{P}\{A \cap B\} + \mathbb{P}\{\overline{A} \cap B\} = \mathbb{P}\{A\}\mathbb{P}\{B \mid A\} + \mathbb{P}\{\overline{A}\}\mathbb{P}\{B \mid \overline{A}\}$$

$\overline{A}$ is "not A", A doesn't happen.

Can also do this with partitions with more than two elements. Say that of $A_1, A_2, \ldots A_n$, exactly one happens.

$$\mathbb{P}\{B\} = \sum_{i=1}^{n} \mathbb{P}\{B \mid A_i\}\mathbb{P}\{A_i\}$$

Also works with countably-infinite partitions. Doesn't work with uncountably-infinite partitions (Kolmogorov!).

5 Bayes' Rule – 15 min (10:21 - 10:36)

Say we know a conditional probability in one direction, but we want to know it in the other direction.

Say we know $\mathbb{P}\{A \mid B\}$, and we want to know $\mathbb{P}\{B \mid A\}$.

$$\mathbb{P}\{B \mid A\} = \frac{\mathbb{P}\{A \cap B\}}{\mathbb{P}\{A\}} = \frac{\mathbb{P}\{A \mid B\}\mathbb{P}\{B\}}{\mathbb{P}\{A\}}$$

Just applied the basic formula twice. We can also use the law of total probability to calculate $\mathbb{P}\{A\}$ if we don't know it.

5.1 Breast cancer example

Gerd Gigerenzer, 2006

Was giving statistics workshops to 1,000 gynecologists. People didn't know this – less got it right than if they'd guessed uniformly at random. This matters.

Suppose you're doing a breast cancer screening. A 50-year-old woman comes in for her routine screening. She tests positive. She's alarmed! She wants to know: What's the chance she has breast cancer? Should she immediately look into treatment, or get more tests, or what?

Two events: *Cancer* and *Positive test*.

She wants to know $\mathbb{P}\{Cancer \mid Positive\ test\}$.

We look up some data (was accurate in the 90s):

In the population, without saying anything about test results, patients like this one have cancer 1% of the time. $\mathbb{P}\{Cancer\} = 0.01$

If a patient like this one has cancer, they'll test positive 90% of the time. $\mathbb{P}\{Test\ positive \mid Cancer\} = 0.9$.

If a patient like this one doesn't have cancer, they'll test positive 10% of the time. $\mathbb{P}\{Test\ positive \mid No\ cancer\} = 0.1$.

We want to know two things:

1. What's the chance that a patient like this one will test positive? $\mathbb{P}\{Test\ positive\}$?

2. Given that this patient tested positive, what's the chance that this patient has breast cancer? $\mathbb{P}\{Cancer \mid Test\ positive\}$?

Options:

a. About 90%

b. About 80%

c. About 10%

d. About 1%

A1. About 10%.

A2. About 10%.

What should we do now?

Get more tests. More likely than general population to have cancer, but still unlikely in absolute terms. "Decision making under uncertainty".

6 Independence – 9 min (10:36-10:45)

Last part, we had two events that interacted a lot.

$\mathbb{P}\{Cancer \mid Test\ Positive\} \gg \mathbb{P}\{Cancer\}$.

But sometimes, we want models where things don't interact.

Flipping coin twice.

$\mathbb{P}\{Second\ flip\ heads \mid First\ flip\ heads\} = \mathbb{P}\{Second\ flip\ heads\}$

Independent events!

$$\mathbb{P}\{A \mid B\} = \mathbb{P}\{A\}$$
$$\mathbb{P}\{A \cap B\} = \mathbb{P}\{A\}\mathbb{P}\{B\}$$

Note: goes both ways. $A \perp B$.

If you know $\mathbb{P}\{A\}$ and $\mathbb{P}\{B\}$, you don't know $\mathbb{P}\{A \cap B\}$. But if you know $\mathbb{P}\{A\}, \mathbb{P}\{B\}$, and you know that $A \perp B$, now you know $\mathbb{P}\{A \cap B\}$.

6.1 Mutually independent

Mutually independent: For any subset of events, probability of all equals product of individual probabilities.

If we say that A_1, A_2, A_3 are mutually independent, we're saying

$$\mathbb{P}\{A_1 \cap A_2\} = \mathbb{P}\{A_1\}\mathbb{P}\{A_2\}$$
$$\mathbb{P}\{A_1 \cap A_3\} = \mathbb{P}\{A_1\}\mathbb{P}\{A_3\}$$
$$\mathbb{P}\{A_2 \cap A_3\} = \mathbb{P}\{A_2\}\mathbb{P}\{A_3\}$$
$$\mathbb{P}\{A_1 \cap A_2 \cap A_3\} = \mathbb{P}\{A_1\}\mathbb{P}\{A_2\}\mathbb{P}\{A_3\}$$

Exponentially many relationships. Lots and lots of information to work with.

This was what was going on when flipping coins forever: Modeled all flips as mutually independent.

7 Wrap-up – 5 min (10:45 - 10:50)

- Conditional Probability – $\mathbb{P}\{A \mid B\}$
- Law of Total Probability – Adding up cases, using the conditional probabilities in each case.
- Bayes rule – From $\mathbb{P}\{A \mid B\}$ to $\mathbb{P}\{B \mid A\}$.
- Independence – If $A \perp B$, then $\mathbb{P}\{A \mid B\} = \mathbb{P}\{A\}$.

Reading: 2.1 through 2.6 of IPC, and quiz.