STA 131A: Introduction to Probability Theory

Lecture 3: Conditional Probability and Bayes' Rule

Dogyoon Song

Spring 2026, UC Davis

Announcements

Homework 1 is now posted

- Due: Tue, April 7, 11:59 PM PT
- Please follow the submission instructions (e.g., formatting)
- Make sure you can access the homework PDF and know where to submit your solutions
    - Late submissions will not be accepted for any reason and will receive 0 points

If you are unsure about anything related to the course:

1. See the syllabus for information
2. Check Piazza first; someone may already have asked a similar question
3. Feel free to ask on Piazza, in class, or during office hours
    - Reserve email for urgent or private matters only

Agenda

Last time:

- Set theory basics
- Probabilistic models

Today:

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

Recap: Outcomes, events, probability

Outcome: a possible result of an experiment (often denoted ω)

- "Head" or "Tail"
- The outcome 6 on a die roll
- Stock price change of \$5

Event: a subset $A \subseteq \Omega$; the event A occurs if the realized outcome ω lies in A

- $\emptyset, \{6\}, \{1,2,3,4,5,6\}$
- For example, if $A = \{1,2,3\}$, then A occurs if and only if $\omega \in A$, i.e., iff $\omega = 1, 2$, or 3
  - This does **NOT** mean "$\omega = 1$ and $\omega = 2$ and $\omega = 3$"

Probability law: a function P that assigns each event A a number $P(A)$ such that

- $P(A) \geq 0$
- If $A_1, A_2, \ldots$ are pairwise disjoint, then

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

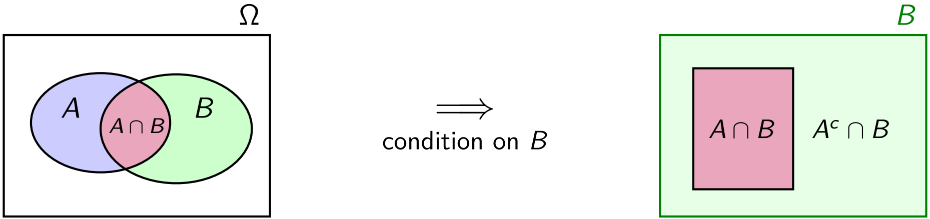
- $P(\Omega) = 1$

Conditional probability: Definition

For an event[1] B with $P(B) > 0$, the **conditional probability** of A given B is

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

- **Example:** $A = \{6\}$, $B = \{2, 4, 6\}$ $\longrightarrow$ $P(A) = 1/6$ vs. $P(A \mid B) = 1/3$
- Conditioning on B means we restrict attention to outcomes in B, and renormalize



Before conditioning

After conditioning on B

[1] In this course, we only condition on events with positive probability

Conditional probability is a probability law

For any fixed B with $P(B) > 0$, the map $P(\cdot \mid B) : A \mapsto P(A \mid B)$ satisfies the axioms

- **Nonnegativity:** $P(A \mid B) \geq 0$ for all events A $\quad \because P(A \cap B) \geq 0, \quad \forall A$
- **Additivity**[2]**:** If $A_1 \cap A_2 = \emptyset$, then

$$
\begin{aligned}
P(A_1 \cup A_2 \mid B) &= \frac{P\big((A_1 \cup A_2) \cap B\big)}{P(B)} && \text{by definition} \\
&= \frac{P\big((A_1 \cap B) \cup (A_2 \cap B)\big)}{P(B)} && \text{by distributivity} \\
&= \frac{P(A_1 \cap B) + P(A_2 \cap B)}{P(B)} && \text{by additivity of } P \\
&= P(A_1 \mid B) + P(A_2 \mid B) && \text{by definition}
\end{aligned}
$$

- **Normalization:** $P(\Omega \mid B) = 1$ $\quad \because P(\Omega \cap B) = P(B)$

$\Longrightarrow$ The map $A \mapsto P(A \mid B)$ is a probability law

[2]Countable additivity follows by the same argument

Conditional probability: A coin-toss example

Example (3 fair coin tosses)

Suppose we toss a fair coin three times and let

$$A = \{\text{more heads than tails}\}, \qquad B = \{\text{1st toss is a head}\}$$

Observe that

$$\begin{aligned} \Omega &= \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}, \\ A &= \{HHH, HHT, HTH, THH\}, \\ B &= \{HHH, HHT, HTH, HTT\} \end{aligned}$$

The event $A \cap B = \{HHH, HHT, HTH\}$

Since all 8 outcomes are equally likely,

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{3/8}{4/8} = \frac{3}{4} \qquad \text{whereas} \qquad P(A) = \frac{4}{8} = \frac{1}{2}$$

Modeling with conditional probability

Example (Fire alarm)

Define two events

$$F = \{\text{there is a fire}\}, \qquad A = \{\text{the alarm goes off}\}$$

The alarm may go off when there is a fire, but it may also produce a false alarm

Thus, there are four possible scenarios:

F	A	Fire and alarm
	A^c	Missed fire
F^c	A	False alarm
	A^c	Nothing happens

Multiplicative rule: $P(A \cap B) = P(A)\, P(B \mid A)$

Conditional probability: Multiplicative rule

Multiplicative rule

Assume $P(A_1 \cap \dots \cap A_k) > 0$ for $k = 1, \dots, n-1$:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1)\,P(A_2 \mid A_1)\,P(A_3 \mid A_1 \cap A_2) \cdots P\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right)$$

This follows from the definition of conditional probability:

$$P\left(\bigcap_{i=1}^{n} A_i\right) = P(A_1) \cdot \frac{P(A_2 \cap A_1)}{P(A_1)} \cdot \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdots \frac{P\left(\bigcap_{i=1}^{n} A_i\right)}{P\left(\bigcap_{i=1}^{n-1} A_i\right)}$$

In words, the probability of a sequence of events is the product of conditional probabilities

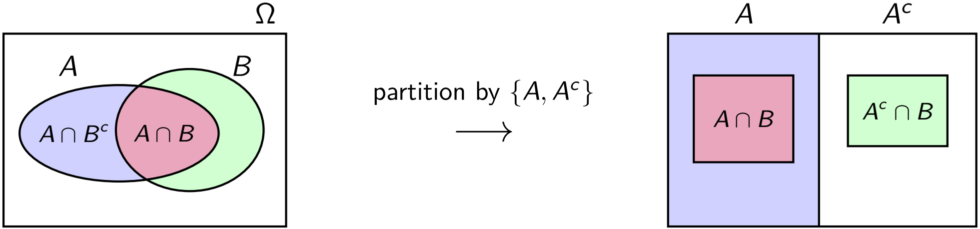
Conditional probability: Law of total probability

Law of total probability

Let $\{A_1, \ldots, A_n\}$ be a partition of Ω such that $P(A_i) > 0$ for all i
Then for any event B,

$$P(B) = \sum_{i=1}^{n} P(A_i \cap B) = \sum_{i=1}^{n} P(A_i) \cdot P(B \mid A_i)$$

For the partition $\{A, A^c\}$, $P(B) = P(A)P(B \mid A) + P(A^c)P(B \mid A^c)$



Law of total probability: Example 1

Example (One match against a random opponent)

You enter a tournament where there are three types of players and let

$$A_i = \text{event of playing with an opponent of type } i, \qquad i \in \{1, 2, 3\},$$
$$B = \text{event of winning}$$

Your chance of meeting each type of players:

$$P(A_1) = 0.6, \qquad P(A_2) = 0.3, \qquad P(A_3) = 0.1$$

Your chance of winning against each type:

$$P(B \mid A_1) = 0.9, \qquad P(B \mid A_2) = 0.7, \qquad P(B \mid A_3) = 0.5$$

Question: What is your chance of winning a game (against random opponents)?

Answer: $P(B) = \sum_{i=1}^{3} P(A_i)P(B \mid A_i) = 0.6 \times 0.9 + 0.3 \times 0.7 + 0.1 \times 0.5 = 0.80$

Law of total probability: Example 2

Example (The Monty Hall problem)

A prize is equally likely to be found behind any of three closed doors

1. You point to one of the doors
2. A friend, who knows where the prize is, opens one of the other two doors and always reveals no prize
3. You may either stay with your original choice or switch to the other unopened door
4. You win the prize if it lies behind your final choice

Question: Should you stay or switch?

Answer: You should switch

Let $C = \{\text{your initial choice is correct}\}$. Then $P(C) = 1/3$, so $P(C^c) = 2/3$

- If C occurs, then switching loses
- If C^c occurs, the friend is forced to open the only remaining door with no prize, so switching wins

$$P(\text{win by staying}) = P(C) = \frac{1}{3}, \qquad P(\text{win by switching}) = P(C^c) = \frac{2}{3}$$

Bayes' rule

Often, we know $P(B \mid A)$, but want $P(A \mid B)$

- A: a hypothesis/model/state of the world
- B: observed data or evidence
- Example: $A =$ having a disease, $\quad B =$ positive screening result

After observing B, we update $P(A)$ to $P(A \mid B)$

- *Prior:* $P(A)$ and $P(A^c)$
- *Likelihood:* $P(B \mid A)$ and $P(B \mid A^c)$
- *Posterior:* $P(A \mid B)$ and $P(A^c \mid B)$

Bayes' rule states that

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)} \qquad \text{where} \qquad P(B) = P(B \mid A)P(A) + P(B \mid A^c)P(A^c)$$

- *Posterior* = *Likelihood* × *Prior* / *Evidence*

Bayes' rule: Example 1

Example (Disease screening)

Let A = disease and B = positive screening result

- Suppose $P(A) = 0.01$
- $P(B \mid A) = 0.8$ (true positive)
- $P(B \mid A^c) = 0.1$ (false positive; the test is positive despite the non-existence of disease)

Q: If the test is positive, how likely is it that the person actually has the disease?

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B \mid A)\, P(A) + P(B \mid A^c)\, P(A^c)} = \frac{0.8 \times 0.01}{0.8 \times 0.01 + 0.1 \times 0.99} \approx 0.075$$

- A positive result raises the probability from 1% to about 7.5%, but it is still far from certain
- When the disease is rare, even a fairly accurate test can produce many false positives

Bayes' rule: Example 2

Example (Fair or double-headed?)

Let A = the coin is double-headed and A^c = the coin is fair

Let B = the first toss is Head

- Suppose $P(A) = 0.5$
- $P(B \mid A) = 1$ (the double-headed (biased) coin always lands Head)
- $P(B \mid A^c) = 0.5$ (the fair coin lands Head with probability $1/2$)

Q: After observing one Head, how likely is it that the coin is double-headed?

$$
\begin{aligned}
P(A \mid B) &= \frac{P(B \mid A)P(A)}{P(B \mid A)P(A) + P(B \mid A^c)P(A^c)} \\
&= \frac{1 \times 0.5}{1 \times 0.5 + 0.5 \times 0.5} = \frac{2}{3} \approx 0.667.
\end{aligned}
$$

- One Head raises the probability from 50% to 66.7%
- Data increase the probability of the model that better explains the observation

Wrap-up

Conditional probability

- $P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$: restrict attention to B and renormalize
- Multiplicative rule: $P(A \cap B) = P(B)P(A \mid B) = P(A)P(B \mid A)$
- Law of total probability: If $\{A_i\}$ is a partition, then $P(B) = \sum_i P(A_i)\, P(B \mid A_i)$

Bayes' rule

- Bayes' rule reverses conditioning:

$$P(A \mid B) = \frac{P(B \mid A)\, P(A)}{P(B)}$$

- Posterior = likelihood × prior / evidence

Suggested reading: [BT08, Ch. 1.3 & 1.4]

References

Dimitri Bertsekas and John N Tsitsiklis.
Introduction to probability, volume 1.
Athena Scientific, 2nd edition, 2008.