Conditional Probability and Bayes' Theorem:
An Extremely Simple, Intuitive Introduction

Saurabh

February 22, 2025

Contents

1 Introduction and Motivation

Imagine you are trying to figure out whether it will rain today. You look at the sky and see dark clouds. How does seeing those clouds change the way you think about the chance of rain? Or, suppose you want to know if an email is spam after noticing it contains a suspicious word. How does this new clue affect the likelihood the email is spam?

These everyday questions can be tackled using two powerful ideas in probability:

1. **Conditional Probability:** How likely something is, given that something else has happened.

2. **Bayes' Theorem:** A rule for updating our beliefs in a systematic way when new evidence appears.

Core Idea

Bayes' Theorem is like a tool that lets you start with what you already believe, see some new evidence, and then figure out a new (and often improved) belief.

2 Basic Concepts in Probability

2.1 Sample Space and Events

- **Sample Space** (Ω): All possible outcomes of an experiment or situation. For a coin toss, the sample space is $\{\text{Heads}, \text{Tails}\}$.

- **Event**: A subset of the sample space. For instance, "getting Heads" is an event that takes up one part of the coin-toss sample space.

2.2 Unconditional Probability

$$P(A) = \text{chance that event } A \text{ happens, without extra info.}$$

For example, the probability of rolling a 6 on a fair die is $P(6) = \frac{1}{6}$.

2.3 Conditional Probability

Conditional probability answers: "What is the probability of A happening, now that I know B happened?" Formally:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \quad \text{provided } P(B) > 0.$$

If you know your friend already ate a slice of pizza, and there was only one slice to begin with, the probability that there is still pizza left is immediately zero!

3 Law of Total Probability

This law helps you split a complicated problem into simpler pieces. Suppose $\{B_1, B_2, \ldots, B_n\}$ is a complete set of disjoint events (they don't overlap and cover all possibilities). Then, the probability of an event A can be found by looking at each B_i separately and then adding up the results:

$$P(A) = \sum_{i=1}^{n} P(A \mid B_i)\, P(B_i).$$

Quick Example: Imagine you have two bags of fruits. Bag 1 has 40% of the fruits, and Bag 2 has 60%. You want to find the probability of picking an apple from the collection as a whole.

$$P(\text{Apple}) = P(\text{Apple} \mid \text{Bag 1}) \times 0.40 \; + \; P(\text{Apple} \mid \text{Bag 2}) \times 0.60.$$

This sums up the probabilities from each bag, weighted by how likely you are to pick from that bag.

Why is This Helpful?

It's easier to break things into parts than to handle everything at once. The law of total probability is like dividing a big task into smaller, more manageable tasks.

4 Bayes' Theorem

Bayes' Theorem is a simple formula that tells you how to update your belief about something after you see new evidence.

4.1 The Formula

$$\underbrace{P(B \mid A)}_{\text{Posterior}} = \frac{\underbrace{P(A \mid B)}_{\text{Likelihood}} \times \underbrace{P(B)}_{\text{Prior}}}{\underbrace{P(A)}_{\text{Evidence}}}.$$

Here's the meaning of each term in plain language:

- **Posterior** ($P(B \mid A)$): Your *new* belief about B after seeing A.
- **Prior** ($P(B)$): Your *original* belief about B before seeing A.
- **Likelihood** ($P(A \mid B)$): How likely you are to see A if B is indeed true.
- **Evidence** ($P(A)$): The overall chance of seeing A, calculated using the law of total probability.

4.2 Why Does This Make Sense?

Think of it this way: you start with a guess (the prior). Then, if your observation (A) is something that would be more or less expected under B, you adjust your guess up or down. This updated belief is your posterior.

5 Relatable Examples

5.1 Example 1: Spam Filtering

- Let S = "Email is Spam."
- Let K = "Email contains a suspicious keyword."

- Suppose $P(S) = 0.1$ (10% of emails are spam).
- Suppose $P(K \mid S) = 0.8$ (80% of spam emails contain the keyword).
- Suppose $P(K \mid \neg S) = 0.2$ (20% of non-spam emails also have that keyword).

Step 1: Overall Probability of Keyword
Using the law of total probability:

$$P(K) = P(K \mid S)\,P(S) + P(K \mid \neg S)\,P(\neg S).$$

$$P(K) = (0.8)(0.1) + (0.2)(0.9) = 0.08 + 0.18 = 0.26.$$

Step 2: Update the Belief (Bayes' Theorem)

$$P(S \mid K) = \frac{P(K \mid S)\,P(S)}{P(K)} = \frac{0.8 \times 0.1}{0.26} \approx 0.3077.$$

So even though seeing the keyword makes it more likely that an email is spam, it's only around 30.8% likely—less than half—because non-spam emails sometimes have that keyword too, and most emails aren't spam to begin with.

5.2 Example 2: Do I Need an Umbrella?

- Let $R =$ "It will rain today."
- Let $C =$ "There are dark clouds."
- Suppose $P(R) = 0.2$ (20% chance of rain on a typical day).
- Suppose $P(C \mid R) = 0.9$ (90% of rainy days have obvious dark clouds).
- Suppose $P(C \mid \neg R) = 0.3$ (30% of non-rainy days still have some dark clouds).

First, find $P(C)$:

$$P(C) = P(C \mid R)\,P(R) + P(C \mid \neg R)\,P(\neg R) = (0.9)(0.2) + (0.3)(0.8) = 0.18 + 0.24 = 0.42.$$

Apply Bayes' Theorem:

$$P(R \mid C) = \frac{P(C \mid R)\,P(R)}{P(C)} = \frac{0.9 \times 0.2}{0.42} \approx 0.4286.$$

Seeing dark clouds increases the likelihood of rain from 20% to about 42.9%. That's a big jump, but notice it's still not a guarantee!

6 Practice Problems

Below are some simple and very relatable problems. Try solving them first on your own to practice the concepts of conditional probability and Bayes' theorem. Detailed solutions follow each problem.

6.1 Problem 1: Medical Test

- **Scenario:** A certain disease affects 1% of the population. A medical test can detect the disease accurately 95% of the time if you actually have the disease (true positive), but it also has a 5% chance of incorrectly indicating disease when you do not have it (false positive).
- **Question:** If you take the test and the result is positive, what is the probability you actually have the disease?

Solution

$$\text{Let } D = \text{“Person has the disease”}, \quad \neg D = \text{“Person does not have the disease”}.$$
$$P(D) = 0.01, \quad P(\neg D) = 0.99.$$
$$P(\text{positive} \mid D) = 0.95, \quad P(\text{positive} \mid \neg D) = 0.05.$$

First, compute the overall probability of a positive test:

$$P(\text{positive}) = P(\text{positive} \mid D) \times P(D) + P(\text{positive} \mid \neg D) \times P(\neg D).$$
$$= (0.95)(0.01) + (0.05)(0.99) = 0.0095 + 0.0495 = 0.059.$$

Now use Bayes' theorem to find $P(D \mid \text{positive})$:

$$P(D \mid \text{positive}) = \frac{P(\text{positive} \mid D)\, P(D)}{P(\text{positive})} = \frac{0.95 \times 0.01}{0.059} \approx 0.161.$$

So, even with a "positive" result, there's about a 16.1% chance you have the disease. The low prevalence (only 1% of people have it) heavily affects this result.

6.2 Problem 2: Rain and the Weather App

- **Scenario:** A weather app claims that it's correct 80% of the time about whether it will rain or not. You know that generally, it rains on 30% of days in your area.
- **Question:** If the app says "It will rain tomorrow," what is the probability it really will rain?

Solution

$$\text{Let } R = \text{“It rains”}, \quad \neg R = \text{“It does not rain”}.$$
$$P(R) = 0.30, \quad P(\neg R) = 0.70.$$

Given the app is "correct" 80% of the time:

$$P(\text{App says Rain} \mid R) = 0.80, \quad P(\text{App says Rain} \mid \neg R) = 0.20.$$

First, find the probability the app says it will rain:

$$P(\text{App says Rain}) = P(\text{App says Rain} \mid R) \times P(R) + P(\text{App says Rain} \mid \neg R) \times P(\neg R).$$
$$= (0.80)(0.30) + (0.20)(0.70) = 0.24 + 0.14 = 0.38.$$

Using Bayes' theorem for $P(R \mid \text{App says Rain})$:

$$P(R \mid \text{App says Rain}) = \frac{0.80 \times 0.30}{0.38} = \frac{0.24}{0.38} \approx 0.6316\ (63.16\%).$$

Thus, if the app says it will rain, there is about a 63% chance it actually will.

6.3 Problem 3: Drawing Socks

- **Scenario:** You have 2 drawers of socks:
  1. Drawer A has 4 red and 6 blue socks.
  2. Drawer B has 8 red and 2 blue socks.

  You randomly choose a drawer with equal probability (50%), then draw one sock from it.
- **Question:** If you draw a red sock, what is the probability it came from Drawer B?

Solution

$$P(A) = 0.5, \quad P(B) = 0.5.$$
$$P(\text{Red} \mid A) = \frac{4}{4+6} = 0.4, \quad P(\text{Red} \mid B) = \frac{8}{8+2} = 0.8.$$
$$P(\text{Red}) = P(\text{Red} \mid A) \times P(A) + P(\text{Red} \mid B) \times P(B) = (0.4)(0.5) + (0.8)(0.5) = 0.2 + 0.4 = 0.6.$$

Using Bayes:

$$P(B \mid \text{Red}) = \frac{P(\text{Red} \mid B)\, P(B)}{P(\text{Red})} = \frac{0.8 \times 0.5}{0.6} = \frac{0.4}{0.6} = \frac{4}{6} = 0.6667\ (66.67\%).$$

So, if you draw a red sock, there is a two-thirds chance it came from Drawer B.

6.4 Problem 4: Conditional Probability Puzzle

- **Scenario:** You roll two fair six-sided dice. One die shows a 4. What is the probability that the other die also shows a 4?
- **Hint:** Think carefully about the sample space given the new information "at least one die is a 4."

Solution (Step-by-Step)

1) **Original Sample Space (36 outcomes):** All pairs (d_1, d_2) where d_1 and d_2 can be 1 to 6.

2) **Condition (at least one 4):** We only look at outcomes where either the first or second die (or both) is 4. The possible outcomes are:

$$(4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (1,4), (2,4), (3,4), (5,4), (6,4).$$

   That's 11 possible outcomes (not 12, because $(4,4)$ is a single outcome).

3) **Event of interest (both dice show 4):** That is just $(4,4)$.

4) **Conditional probability:**

$$P(\text{both 4s} \mid \text{at least one 4}) = \frac{1}{11}.$$

Therefore, given that at least one die shows a 4, there is $\frac{1}{11} \approx 9.09\%$ chance the other is also a 4.

6.5 Problem 5: Coin Flip with Partial Information

- **Scenario:** You flip a fair coin twice. Your friend peeks at the results and tells you, "At least one of the flips landed Heads."
- **Question:** What is the probability that both flips are Heads, given this new information?

Solution

Step 1: All possible outcomes of flipping a fair coin twice are:

$$\{(H, H), (H, T), (T, H), (T, T)\}.$$

Step 2: Given "at least one is Heads," we remove (T, T).
So the reduced sample space is:

$$\{(H, H), (H, T), (T, H)\}.$$

Step 3: Probability:

$$P(\text{both Heads} \mid \text{at least one Heads}) = \frac{|\{(H, H)\}|}{|\{(H, H), (H, T), (T, H)\}|} = \frac{1}{3} \approx 33.33\%.$$

7 Advanced (Tough) Problems

Ready for more challenging scenarios? These classic puzzles can be tricky, so read carefully!

7.1 Problem 6: The Monty Hall Problem

- **Scenario:** You are on a game show. There are 3 doors: behind one door is a prize (car), behind the other two are goats. You pick a door, but you don't open it yet. The host, who knows where the prize is, opens one of the other two doors to show a goat. Then you get a choice: stick with your original door, or switch to the remaining unopened door.
- **Question:** If you want the best chance of winning the car, should you stick or switch? And what is the probability of winning in each strategy?

Solution (Summary)

1) **Prior Probability:** The chance you initially chose the prize door is $1/3$. The chance the prize is behind one of the other two doors is $2/3$.

2) **Host's Action:** The host will always open a door with a goat, which gives you extra information.

3) **Posterior Probability:**

    - If you **stick** with your original door, your chance of winning remains $1/3$.

- If you **switch** to the other unopened door, your chance of winning is 2/3.

In simpler terms, once the host reveals a goat, it's more likely the car is behind the *other* door you did *not* choose initially. Hence, **switching doubles your chances of winning**.

7.2 Problem 7: The Two-Coin Box Puzzle (Bertrand's Box Variant)

- **Scenario:** You have three boxes:
  1. Box 1: Two gold coins (GG)
  2. Box 2: Two silver coins (SS)
  3. Box 3: One gold and one silver (GS)

  You randomly choose one box (each box has a 1/3 chance of being picked). Then you pick *one* coin at random from that box, and you see that it is **gold**.

- **Question:** What is the probability the other coin in the same box is also gold?

Solution (Step-by-Step)

1) **Possible ways to pick a gold coin:**
   - **Box 1 (GG):** It has 2 gold coins. If you pick this box, you *always* draw gold, no matter which coin you pick. Probability of choosing Box 1 is $\frac{1}{3}$, and given Box 1, probability of picking a gold coin is 1.
   - **Box 2 (SS):** No chance of gold here, so this scenario does not contribute.
   - **Box 3 (GS):** It has 1 gold and 1 silver coin. Probability of choosing Box 3 is $\frac{1}{3}$, and given Box 3, probability of picking gold is $\frac{1}{2}$.

2) **Compute $P(\textbf{Gold})$:** Using the law of total probability,

$$P(\text{Gold}) = P(\text{Gold} \mid \text{Box 1})\,P(\text{Box 1}) + P(\text{Gold} \mid \text{Box 2})\,P(\text{Box 2}) + P(\text{Gold} \mid \text{Box 3})\,P(\text{Box 3}).$$

$$= (1) \times \frac{1}{3} + (0) \times \frac{1}{3} + \left(\frac{1}{2}\right) \times \frac{1}{3} = \frac{1}{3} + 0 + \frac{1}{6} = \frac{1}{2}.$$

3) **Probability that the box is GG, given you drew gold:**

$$P(\text{Box 1} \mid \text{Gold}) = \frac{P(\text{Gold} \mid \text{Box 1})\,P(\text{Box 1})}{P(\text{Gold})} = \frac{(1) \times \frac{1}{3}}{\frac{1}{2}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}.$$

4) If you are in Box 1 (GG), the other coin is obviously gold. Hence:

$$P(\text{other coin gold} \mid \text{Gold}) = \frac{2}{3}.$$

So there is a $\frac{2}{3}$ chance you are in the box with two gold coins (since that box is more "likely" to give you a gold coin in the first place), meaning there's a 2-out-of-3 chance the other coin is also gold.

Final Thoughts

These problems illustrate how to apply both *conditional probability* and *Bayes' theorem* in everyday scenarios, from medical tests to simple dice and coin puzzles—and even tricky game shows and coin-box mind-benders. By carefully considering how new information changes the sample space (or updates your prior belief), you gain a powerful framework for analyzing uncertainty in the real world.