Bayes' Theorem Application: Spam Filter Analysis

The Problem: The Word "Win"

You are designing a spam filter. You have analyzed your past email traffic and established the following probabilities:

- **Prior Probability:** 40% of all emails you receive are **Spam** ($P(\text{Spam}) = 0.40$).
- **Likelihood (Spam):** 50% of Spam emails contain the word "Win" ($P(\text{Win} \mid \text{Spam}) = 0.50$).
- **Likelihood (Ham/Safe):** Only 5% of Ham emails contain the word "Win" ($P(\text{Win} \mid \text{Ham}) = 0.05$).

Question: You receive a new email that contains the word "Win". What is the probability that this email is **Spam**?

Method 1: Using the Bayes Formula

$$P(\text{Spam} \mid \text{Win}) = \frac{P(\text{Win} \mid \text{Spam}) \cdot P(\text{Spam})}{P(\text{Win})}$$

Step A: Calculate the Numerator (Spam Part)

Probability that it is Spam *and* has the word "Win":

$$P(\text{Win} \mid \text{Spam}) \cdot P(\text{Spam}) = 0.50 \times 0.40 = \mathbf{0.20}$$

Step B: Calculate the Denominator (Total Probability of "Win")

We must add the probability of "Win" appearing in Spam and in Ham.

$$P(\text{Win}) = [\text{Spam Part}] + [\text{Ham Part}]$$

$$P(\text{Win}) = (0.50 \times 0.40) + (0.05 \times 0.60)$$
$$P(\text{Win}) = 0.20 + 0.03 = \mathbf{0.23}$$

Step C: Final Calculation

$$P(\text{Spam} \mid \text{Win}) = \frac{0.20}{0.23} \approx \mathbf{0.8696}$$

Answer: There is an **86.96%** chance the email is Spam.

Method 2: Using a Contingency Table

Let's imagine a sample bucket of **100 random emails**.

Step A: Fill in the Total Counts (Priors)

- **Spam:** 40% of 100 = **40 emails**.
- **Ham:** 60% of 100 = **60 emails**.

Step B: Apply the Word Counts (Likelihoods)

- **Spam Row:** 50% of 40 have "Win" $\rightarrow$ **20**. (20 do not).
- **Ham Row:** 5% of 60 have "Win" $\rightarrow$ **3**. (57 do not).

Step C: Completed Table

Type	Has "Win"	No "Win"	Total
Spam	**20**	20	40
Ham	**3**	57	60
TOTAL	**23**	77	**100**

Step D: Calculate from the Table

The question asks: *"Given the email contains 'Win'..."* This tells us to look strictly at the **Has "Win"** column.

- Total emails with "Win": **23**
- Spam emails inside that group: **20**

$$P(\text{Spam} \mid \text{Win}) = \frac{\text{Spam w/ Win}}{\text{Total w/ Win}}$$

$$P(\text{Spam} \mid \text{Win}) = \frac{20}{23} \approx \mathbf{0.8696}$$

Conclusion: Both methods confirm that the presence of the word "Win" raises the probability of Spam from 40% to **87%**.

Example 2: The Weather Forecast

The Scenario: You are planning a picnic. You look out the window and see it is **cloudy**. You want to know the probability that it will actually **rain**.

- **Prior Probability:** In your city, it rains on 20% of all days ($P(\text{Rain}) = 0.20$).
- **Likelihood (True Positive):** If it is going to rain, there are clouds 90% of the time ($P(\text{Cloud} \mid \text{Rain}) = 0.90$).
- **Likelihood (False Positive):** If it is **not** going to rain (fair weather), there are still clouds 30% of the time ($P(\text{Cloud} \mid \text{No Rain}) = 0.30$).

Question: Given that it is cloudy, what is the probability it will actually rain?

Method 1: Using the Bayes Formula

$$P(\text{Rain} \mid \text{Cloud}) = \frac{P(\text{Cloud} \mid \text{Rain}) \cdot P(\text{Rain})}{P(\text{Cloud})}$$

Step A: Calculate the Numerator (Rain Part)

$$P(\text{Cloud} \mid \text{Rain}) \cdot P(\text{Rain}) = 0.90 \times 0.20 = \mathbf{0.18}$$

Step B: Calculate the Denominator (Total Cloudy Days)

We combine clouds from rainy days and clouds from dry days.

$$P(\text{Cloud}) = [\text{Rain Part}] + [\text{No Rain Part}]$$

$$P(\text{Cloud}) = (0.18) + (0.30 \times 0.80)$$

$$P(\text{Cloud}) = 0.18 + 0.24 = \mathbf{0.42}$$

Step C: Final Calculation

$$P(\text{Rain} \mid \text{Cloud}) = \frac{0.18}{0.42} \approx \mathbf{0.4286}$$

Answer: There is a **42.9%** chance of rain. (Even though it's cloudy, it is still more likely to stay dry!)

Method 2: Using a Contingency Table

Let's imagine a timeline of **100 days**.

Step A: Fill in the Priors

- **Rainy Days:** 20% of 100 = **20 days**.
- **Dry Days:** 80% of 100 = **80 days**.

Step B: Apply the Forecasts

- **Rain Row:** 90% of 20 are cloudy $\to$ **18**. (2 are clear).
- **Dry Row:** 30% of 80 are cloudy $\to$ **24**. (56 are clear).

Step C: Completed Table

Weather	Cloudy	Clear Sky	Total
Rain	**18**	2	20
No Rain	**24**	56	80
TOTAL	**42**	58	**100**

Step D: Calculate from the Table

The question asks: *"Given it is Cloudy..."* We look strictly at the **Cloudy** column.

- Total Cloudy Days: **42**
- Actual Rainy Days in that group: **18**

$$P(\text{Rain} \mid \text{Cloud}) = \frac{18}{42} \approx \mathbf{0.4286}$$

Conclusion: Both methods show that despite the clouds, the probability of rain is only **42.9%**.