

Spy or Not? — A Bayes' Rule Game

Quick Setup (before class)

1. **We need 1–3 volunteers** to be suspects.
2. **Run it:** suspect draws a role token privately (in R) → draws a clue card (in R) → reads clue aloud → class computes $P(\text{Spy} \mid \text{Clue})$ in small groups → reveal true role and compare.

Setup

Before class, secretly assign each “suspect” one of two roles:

- **Loyal** — occurs with prior probability $P(\text{Loyal}) = 0.7$
- **Spy** — occurs with prior probability $P(\text{Spy}) = 0.3$

Some clues are more likely to come from a Loyal player; others are more likely to come from a Spy. For example:

Clue	$P(\text{Clue} \mid \text{Loyal})$	$P(\text{Clue} \mid \text{Spy})$
“I was helping the team the whole time.”	0.6	0.2
“I don’t really remember what happened.”	0.3	0.5
“Trust me, it wasn’t me.”	0.1	0.3

The Task

A suspect reveals their clue. Without knowing their true role, we compute:

$$P(\text{Spy} \mid \text{Clue}) = \frac{P(\text{Clue} \mid \text{Spy}) \cdot P(\text{Spy})}{P(\text{Clue} \mid \text{Spy}) \cdot P(\text{Spy}) + P(\text{Clue} \mid \text{Loyal}) \cdot P(\text{Loyal})}$$

Write down their posterior belief *before* the reveal, then the suspect’s true role is announced. Did the group’s belief move in the right direction?

Round2

- **Round 2:** Give a *second* clue for the same suspect. Use their **first posterior as the new prior**, the update $P(\text{Spy} \mid \text{Clue1}, \text{Clue2})$.

Debrief Questions

- Which clue moved your belief the most? Why? (*Connects to likelihood ratios — a clue is “informative” when $P(\text{Clue} \mid \text{Spy})$ and $P(\text{Clue} \mid \text{Loyal})$ are far apart.*)
- What happens to your posterior if a clue is equally likely under both roles?
- How did the *prior* (70/30 split) affect your final answer compared to a 50/50 starting point?