Conditional Probability and Bayes' Theorem

Lesson 04

Digital Finance

Outline

Motivation

New information changes probabilities

- Probability of rain tomorrow: 30%
- Given dark clouds now: probability increases!

Finance examples:

- Probability of default given economic recession
- Probability stock rises given positive earnings
- Probability of fraud given suspicious transactions

Conditional probability captures how knowledge affects uncertainty.

Definition of Conditional Probability

Intuition: Given that B happened, what's the chance A also happened?
Visual idea: Zoom into the "B world" – what fraction of B is also A?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\text{overlap of A and B}}{\text{everything in B}}$$

Example: 100 students: 30 study math, 20 study both math and physics.

- P(physics | math) = 20/30 = 67% of math students also study physics

Read as: "Probability of A given B" or "P of A, given B"

The vertical bar "|" means "given that" or "conditional on."

Multiplication Rule

Rearranging the definition:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A)$$

Chain rule (multiple events):

$$P(A \cap B \cap C) = P(A) \cdot P(B|A) \cdot P(C|A \cap B)$$

Example: Drawing cards without replacement

- $P(\text{two aces}) = P(\text{1st ace}) \cdot P(\text{2nd ace}|\text{1st ace})$
- $= \frac{4}{52} \cdot \frac{3}{51} = \frac{1}{221}$

The multiplication rule breaks joint probabilities into conditionals.

Statistical Independence

Events A and B are independent if:

$$P(A|B) = P(A) \quad \text{or equivalently} \quad P(A \cap B) = P(A) \cdot P(B)$$

Meaning: Knowing B occurred doesn't change probability of A **Examples:**

- Two coin flips: outcomes are independent
- Stock returns on non-adjacent days: approximately independent
- Drawing with replacement: independent

Warning: Independence $\neq$ mutually exclusive!

Mutually exclusive events are highly dependent (one rules out the other).

Conditional Independence

A **and** B **conditionally independent given** C**:**

$$P(A \cap B|C) = P(A|C) \cdot P(B|C)$$

Key insight:

- Events can be dependent marginally
- But independent once we condition on something

Example:

- Two students' test scores are dependent
- But independent given they studied the same amount

Conditional independence is fundamental in probabilistic modeling.

Bayes' Theorem

The big idea: How do we update our beliefs when we get new evidence?

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

In plain English:

- **Prior** $P(A)$: What we believed *before* seeing evidence
- **Likelihood** $P(B|A)$: How likely is this evidence if A is true?
- **Posterior** $P(A|B)$: What we believe *after* seeing evidence

Formula in words: New belief = (Old belief × How well evidence fits) / (Normalizer)

Prior + Evidence → Posterior. That's Bayesian updating!

Bayes' Theorem with Total Probability

Extended form using law of total probability:

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)}$$

General form (partition $A_1, \ldots, A_n$):

$$P(A_i|B) = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{j=1}^{n} P(B|A_j) \cdot P(A_j)}$$

The denominator ensures probabilities sum to 1.

Example: Medical Test

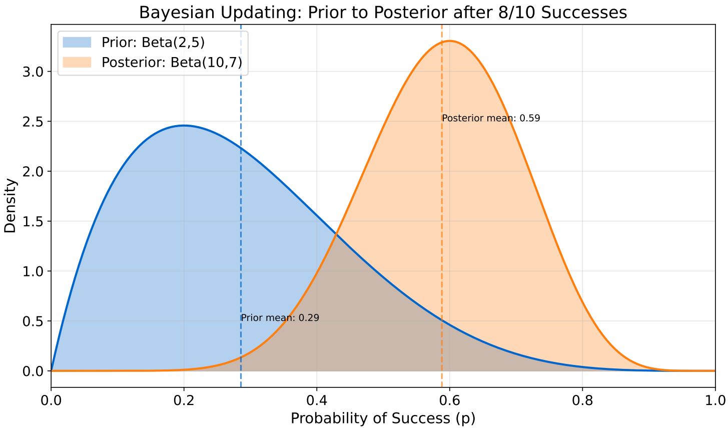
Setup: Disease affects 1%. Test: 95% detects disease, 10% false alarm.
Think about 1,000 people:

- 10 have disease → 9.5 test positive (true positives)
- 990 are healthy → 99 test positive (false positives)
- Total positives: about 109 people

If you test positive: $\mathrm{P}(\text{disease}) = 9.5 \,/\, 109 \approx 8.8\%$
Why so low? The false positives (99) vastly outnumber true positives (9.5) because healthy people (990) far outnumber sick people (10).

Base rate matters! Rare diseases mean most positive tests are false alarms.

Bayesian Updating



Data shifts our beliefs from prior to posterior.

Law of Total Probability

If $A_1, \dots, A_n$ partition the sample space:

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

Simple case (two events):

$$P(B) = P(B|A) \cdot P(A) + P(B|A^c) \cdot P(A^c)$$

Use when:

- Don't know $P(B)$ directly
- But know $P(B|A_i)$ for each scenario

Break a complex probability into simpler conditional pieces.

Example: Portfolio Default

Portfolio has 3 risk categories (illustrative values):

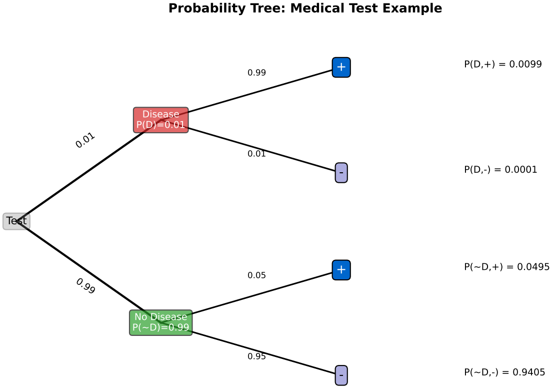
- Low risk (60%): 1% default rate
- Medium risk (30%): 5% default rate
- High risk (10%): 15% default rate

Overall default probability?

$$P(\text{default}) = 0.01 \times 0.60 + 0.05 \times 0.30 + 0.15 \times 0.10$$
$$= 0.006 + 0.015 + 0.015 = 0.036 = 3.6\%$$

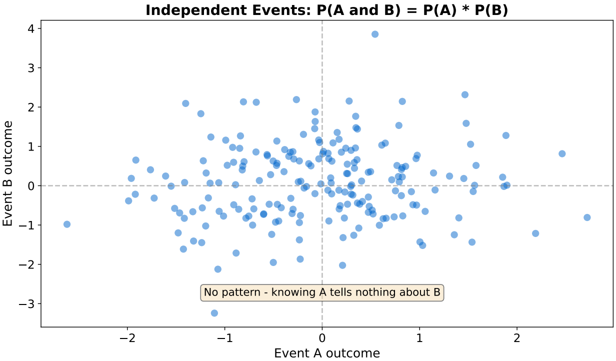
Weighted average of conditional probabilities.

Probability Trees



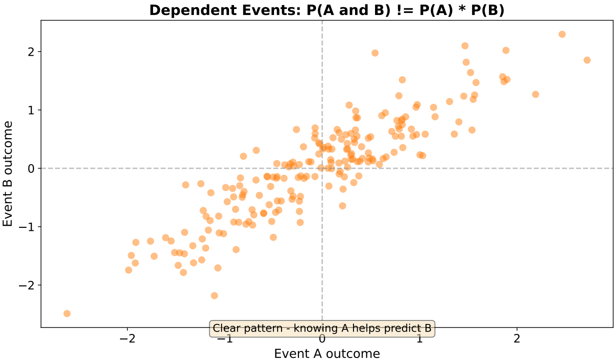
Trees visualize conditional probabilities.

Independent Events



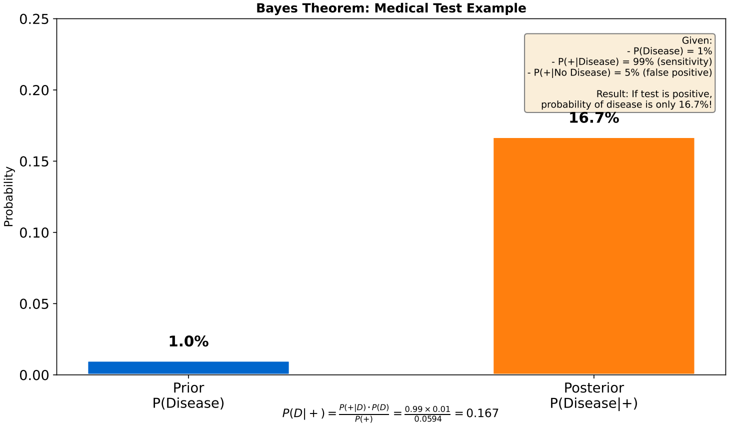
Independent: P(A and B) = P(A) * P(B). Knowing A gives no info about B.

Dependent Events



Dependent: P(A and B) != P(A) * P(B). Knowing A helps predict B.

Bayes Theorem: Medical Testing



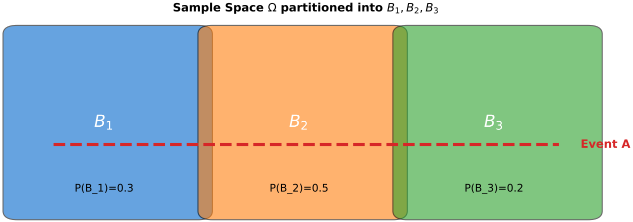
Given:
- P(Disease) = 1%
- P(+|Disease) = 99% (sensitivity)
- P(+|No Disease) = 5% (false positive)

Result: If test is positive, probability of disease is only 16.7%!

$$P(D|+) = \frac{P(+|D) \cdot P(D)}{P(+)} = \frac{0.99 \times 0.01}{0.0594} = 0.167$$

Even accurate tests can have high false positive rates.

Law of Total Probability



$$\boxed{P(A) = P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + P(A|B_3)P(B_3)}$$

Law of Total Probability

Partition the sample space to compute overall probability.

Summary

Conditional probability:

- $P(A|B) = P(A \cap B)/P(B)$
- Updates probabilities with new information

Independence:

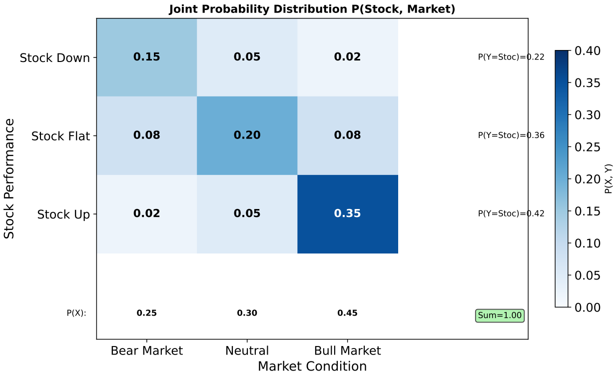
- $P(A \cap B) = P(A) \cdot P(B)$
- Knowledge of one doesn't affect the other

Bayes' theorem:

- Prior × Likelihood → Posterior
- Foundation of Bayesian inference

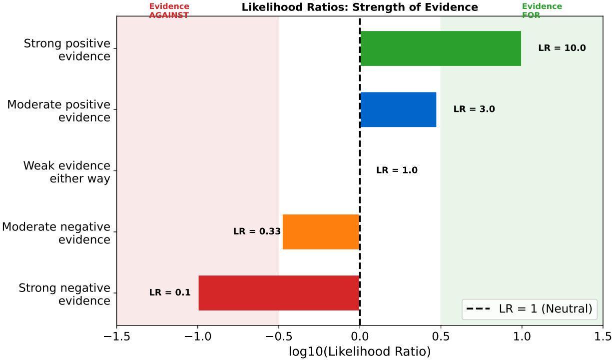
Next lesson: Discrete Random Variables

Joint Probability Heatmap



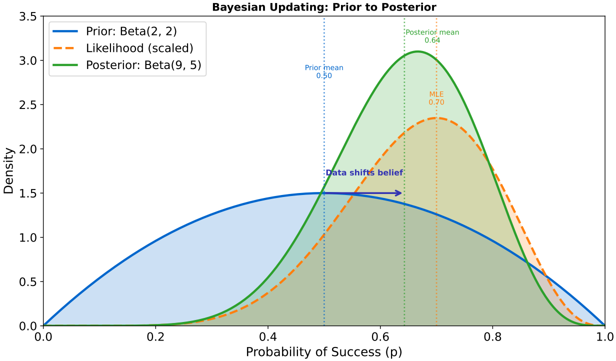
Joint probabilities show how two events occur together.

Likelihood Ratios



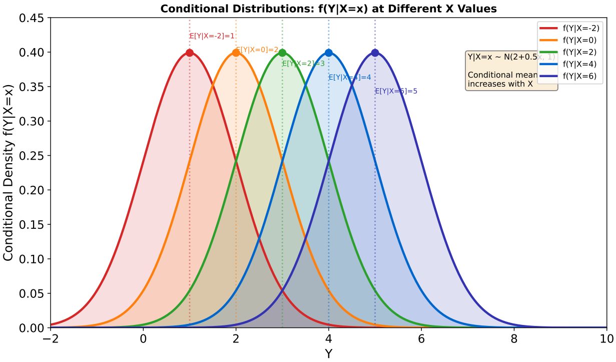
Likelihood ratio = P(evidence|hypothesis) / P(evidence|not hypothesis). LR $>$ 1 supports hypothesis.

Prior vs Posterior Distributions



Evidence shifts our beliefs from prior to posterior.

Conditional Density Slices



Conditional distributions show how Y varies given different values of X.