Chapter 5

Bayes' Theorem

Topics to Cover

- Bayes' theorem
- Prior and posterior probabilities
- Use of partitions in probabilistic inference

Learning Objectives

- Derive Bayes' theorem from conditional probability.
- Compute posterior probabilities using partitions.
- Interpret Bayes' rule as updating beliefs after observing an event.

5.1 Motivation: reversing conditioning

From Chapter 4, we compute:

$$P(A \mid B)$$

In many situations:

- an event A is observed
- we want to infer which underlying event E_i led to this observation

This motivates reversing conditional probabilities.

5.2 Bayes' Theorem

Starting point. From conditional probability:

$$P(E \cap A) = P(E \mid A)P(A)$$

$$P(E \cap A) = P(A \mid E)P(E)$$

Theorem (Bayes). For $P(A) > 0$:

$$\boxed{P(E \mid A) = \frac{P(A \mid E)P(E)}{P(A)}}$$

5.3 Role of partitions

Let $\{E_1, E_2, \ldots, E_n\}$ be a partition of Ω:

- $E_i \cap E_j = \emptyset \quad (i \neq j)$
- $\bigcup_{i=1}^{n} E_i = \Omega$

Then any event A can be written as:

$$A = \bigcup_{i=1}^{n} (A \cap E_i)$$

Since the terms are disjoint:

$$P(A) = \sum_{i=1}^{n} P(A \cap E_i)$$

Using conditional probability:

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

This is the Law of Total Probability.

5.4 Bayes' rule with partitions

Let $\{E_1, \ldots, E_n\}$ be a partition.
Then:

$$\boxed{P(E_i \mid A) = \frac{P(A \mid E_i)\, P(E_i)}{\sum_{j=1}^{n} P(A \mid E_j)\, P(E_j)}}$$

5.5 Diagram: decomposition using partitions

$$P(A) = \sum_{i} P(A \cap E_i) = \sum_{i} P(A \mid E_i)\, P(E_i)$$

5.6 Example: two possible cases

Suppose:

- exactly one of E_1 or E_2 occurs
- event A is observed

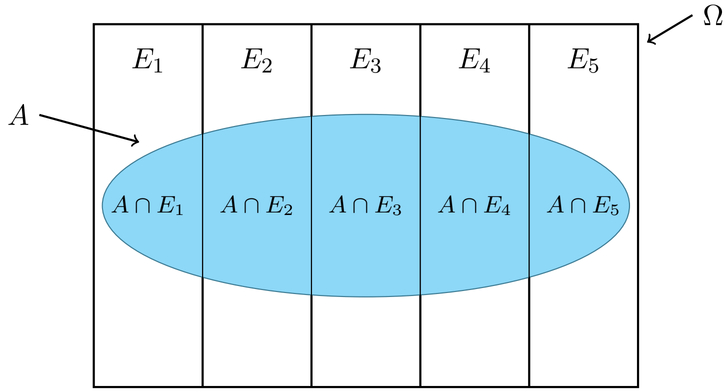

Figure 5.1: Decomposition $A = \bigcup_{i=1}^{n}(A \cap E_i)$ with elliptical event A

Then:

$$P(E_1 \mid A) = \frac{P(A \mid E_1)P(E_1)}{P(A \mid E_1)P(E_1) + P(A \mid E_2)P(E_2)}$$

Interpretation. The observation A updates our belief about whether E_1 or E_2 occurred.

5.7 Example: coin selection

- One of two coins is selected
- E_1: coin 1 chosen, E_2: coin 2 chosen
- A: head observed

Then:

$$P(E_1 \mid A) = \frac{P(A \mid E_1)P(E_1)}{P(A \mid E_1)P(E_1) + P(A \mid E_2)P(E_2)}$$

5.8 Interpretation

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

- Prior: $P(E_i)$
- Likelihood: $P(A \mid E_i)$
- Posterior: $P(E_i \mid A)$

The denominator ensures normalization:

$$\sum_i P(E_i \mid A) = 1$$

5.9 Modeling checklist

1. Identify a partition $\{E_i\}$
2. Specify prior probabilities $P(E_i)$
3. Specify likelihoods $P(A \mid E_i)$
4. Compute:
$$P(A) = \sum_i P(A \mid E_i)\, P(E_i)$$
5. Compute posterior probabilities $P(E_i \mid A)$

KEEN 3Cs

Curiosity

How does observing an event change our beliefs about underlying causes? What happens when prior probabilities are very small?

Connections

Bayes' theorem builds directly on:

- Conditional probability (Chapter 4)
- Law of Total Probability

Creating Value

Construct a model where observations are used to infer which scenario among several possibilities occurred.

Practice

A. Remember and Understand

1. [Remember]State Bayes' theorem.
2. [Understand]Explain prior and posterior probabilities.

B. Apply and Analyze

1. [Apply]Compute $P(E_1 \mid A)$ for two cases.
2. [Analyze]Verify that the denominator equals $P(A)$.

C. Evaluate and Create

1. [Evaluate]Explain how prior probabilities influence the result.
2. [Create]Construct a three-case inference model.