CSE 400: Fundamentals of Probability in Computing

Lecture 5: Bayes' Theorem, Random Variables, and Probability Mass Function

Instructor: Dhaval Patel, PhD
Date: January 20, 2026

1. Bayes' Theorem

Weighted Average of Conditional Probabilities

Let A and B be events. We may express A as

$$A = AB \cup AB^c$$

for, in order for an outcome to be in A, it must either be in both A and B, or be in A but not in B.

Since AB and AB^c are mutually exclusive, by **Axiom 3**, we have:

$$\Pr(A) = \Pr(AB) + \Pr(AB^c)$$

Using conditional probability,

$$\Pr(A) = \Pr(A \mid B)\Pr(B) + \Pr(A \mid B^c)[1 - \Pr(B)]$$

Interpretation as stated in the lecture: The probability of event A is a **weighted average** of the conditional probabilities, with weights given by the probabilities of the events on which it is conditioned.

2. Bayes' Theorem — Learning by Example

Example 3.1 (Part 1/2)

An insurance company believes that people can be divided into two classes:

- Accident prone
- Not accident prone

Statistics show:

- An accident-prone person has an accident within one year with probability 0.4
- A non-accident-prone person has an accident within one year with probability 0.2

Assume:

- 30% of the population is accident prone

Question: What is the probability that a new policyholder will have an accident within one year?

Solution

Let:

- A_1: policyholder has an accident within one year
- A: policyholder is accident prone

Conditioning on whether the policyholder is accident prone:

$$\Pr(A_1) = \Pr(A_1 \mid A)\Pr(A) + \Pr(A_1 \mid A^c)\Pr(A^c)$$

Substitute values:

$$\Pr(A_1) = (0.4)(0.3) + (0.2)(0.7)$$

$$\Pr(A_1) = 0.12 + 0.14 = 0.26$$

Example 3.1 (Part 2/2)

Suppose that a new policyholder **has had an accident** within one year. **Question:** What is the probability that he or she is accident prone?

Solution

The desired probability is:

$$\Pr(A \mid A_1) = \frac{\Pr(AA_1)}{\Pr(A_1)}$$

Using conditional probability:

$$\Pr(AA_1) = \Pr(A)\Pr(A_1 \mid A)$$

Thus:

$$\Pr(A \mid A_1) = \frac{(0.3)(0.4)}{0.26}$$

$$= \frac{0.12}{0.26} = \frac{6}{13}$$

3. Bayes' Theorem

Formal Introduction: Law of Total Probability and Bayes Formula

Let $B_1, B_2, \ldots, B_n$ be mutually exclusive events such that:

$$\bigcup_{i=1}^{n} B_i = S$$

Then:

$$\Pr(A) = \sum_{i=1}^{n} \Pr(AB_i)$$

This is known as the **Law of Total Probability** (Formula 3.4).

Using:

$$\Pr(AB_i) = \Pr(B_i \mid A)\Pr(A)$$

we obtain:

$$\Pr(B_i \mid A) = \frac{\Pr(A \mid B_i)\Pr(B_i)}{\sum_{j=1}^{n} \Pr(A \mid B_j)\Pr(B_j)}$$

This is known as the **Bayes Formula** (**Proposition 3.1**).
Where:

- $\Pr(B_i)$ is the **a priori probability**
- $\Pr(B_i \mid A)$ is the **posteriori probability**

4. Bayes Formula — Learning by Example

Example 3.2 (Card Problem)

There are three cards:

- One card: both sides red (RR)
- One card: both sides black (BB)
- One card: one red, one black (RB)

The cards are mixed in a hat. One card is randomly selected and placed on the ground.
Given: the upper side is red **Question:** What is the probability that the other side is black?

Solution

Let:

- RR: card is all red
- BB: card is all black
- RB: card is red–black
- R: upturned side is red

We want:

$$\Pr(RB \mid R)$$

Using Bayes' formula:

$$\Pr(RB \mid R) = \frac{\Pr(R \mid RB)\Pr(RB)}{\Pr(R \mid RR)\Pr(RR) + \Pr(R \mid RB)\Pr(RB) + \Pr(R \mid BB)\Pr(BB)}$$

Substitute values:

$$= \frac{(1/2)(1/3)}{(1)(1/3) + (1/2)(1/3) + (0)(1/3)}$$

$$= \frac{1/6}{1/3 + 1/6} = \frac{1/6}{1/2} = \frac{1}{3}$$

Discussion of Incorrect Reasoning

Guessing 1/2 is incorrect because it assumes equal likelihood between RR and RB.

There are **6 equally likely outcomes**:

$$R_1, R_2, B_1, B_2, R_3, B_3$$

Only outcome R_3 corresponds to a red side whose other side is black.

Thus:

$$\Pr = \frac{1}{3}$$

5. Random Variables

Motivation and Concept

When an experiment is performed, interest is often in a **function of the outcome**, not the outcome itself.

Examples:

- Dice tossing: sum of dice
- Coin flipping: number of heads

Definition: Real-valued functions defined on the sample space are called **Random Variables**.

- Values determined by outcomes
- Probabilities assigned to values

The distribution can be visualized as a bar diagram where:

- x-axis: values
- y-axis: $\Pr[X = a]$

6. Random Variables — Example

Coin Toss Example

Experiment: Tossing 3 fair coins Let Y = number of heads

Possible values: $0, 1, 2, 3$

$$\Pr(Y = 0) = \Pr(t, t, t) = \frac{1}{8}$$

$$\Pr(Y = 1) = \Pr(t, t, h; t, h, t; h, t, t) = \frac{3}{8}$$

$$\Pr(Y = 2) = \Pr(t, h, h; h, t, h; h, h, t) = \frac{3}{8}$$

$$\Pr(Y = 3) = \Pr(h, h, h) = \frac{1}{8}$$

Since Y must take one of these values:

$$\sum \Pr(Y = y) = 1$$

7. Probability Mass Function (PMF)

A random variable that takes at most a **countable number of values** is called **discrete**.

Let X be a discrete random variable with range:

$$R_X = \{x_1, x_2, x_3, \dots\}$$

The function:

$$p(x_k) = \Pr(X = x_k)$$

is called the **Probability Mass Function (PMF)**.

Since X must take one of its possible values:

$$\sum_k p(x_k) = 1$$

8. PMF Examples

Example: Two Independent Tosses of a Fair Coin

(As shown in lecture slides)

Example (Given PMF)

The PMF is:

$$p(i) = c\frac{\lambda^i}{i!}, \quad i = 0, 1, 2, \dots$$

where $\lambda > 0$

Find:

$$\Pr(X = 0), \quad \Pr(X > 2)$$

(As presented in the lecture slides)