CSE400 – Fundamentals of Probability in Computing

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

Instructor: Dhaval Patel, PhD

Date: January 20, 2026

Outline

- Bayes' Theorem
  - Weighted Average of Conditional Probabilities
  - Learning by Example
  - Formal Introduction: Law of Total Probability and Bayes' Theorem
- Random Variables
  - Motivation and Concept
  - Examples
- Probability Mass Function (PMF)
  - Concept
  - Examples
- Class Participation – Quiz

---

1. Bayes' Theorem

1.1 Weighted Average of Conditional Probabilities

Let (A) and (B) be events.

Event (A) can be expressed as:

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

For an outcome to be in (A), it must either be in both (A) and (B), or be in (A) but not in (B).

The events (AB) and (AB^c) are mutually exclusive. By **Axiom 3**:

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

Using conditional probability:

$$\Pr(AB) = \Pr(A \mid B)\,\Pr(B)$$

$$\Pr(AB^c) = \Pr(A \mid B^c)\,\Pr(B^c)$$

Substituting:

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

Statement:

The probability of event (A) is a weighted average of conditional probabilities, with weights given by the probabilities of the conditioning events.

1.2 Learning by Example

Example 3.1 (Part 1)

An insurance company divides people into two classes:

- Accident prone
- Not accident prone

Statistics:

- An accident-prone person has an accident within a fixed 1-year period with probability (0.4).
- A person who is not accident prone has an accident within the same period with probability (0.2).
- (30%) of the population is accident prone.

Let:

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

Conditioning on whether the policyholder is accident prone:

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

Substituting:

$$\Pr(A_1) = (0.4)(0.3) + (0.2)(0.7) = 0.26$$

Example 3.1 (Part 2)

Suppose a policyholder has an accident within one year.

Find the probability that the policyholder is accident prone.

The desired probability:

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

Using conditional probability:

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

Substituting:

$$\Pr(A \mid A_1) = \frac{(0.3)(0.4)}{0.26} = \frac{6}{13}$$

---

2. Formal Introduction: Law of Total Probability and Bayes' Theorem

2.1 Law of Total Probability

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

2.2 Bayes' Formula

Using:

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

we obtain the **Bayes Formula (Proposition 3.1)**.

Definitions:

- ($\Pr(B_i)$) : *a prioriprobability*
- ($\Pr(B_i \mid A)$) : *a posterioriprobabilityofevent(B_i)given(A)*

2.3 Learning by Example

Example 3.2

Three cards:

- One card with both sides red (RR)
- One card with both sides black (BB)
- One card with one red side and one black side (RB)

One card is randomly selected and placed on the ground.

The upper side is red.

Let:

- (RR, BB, RB): events describing the selected card
- (R): event that the upturned side is red

The required probability:

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

Substituting:

$$\Pr(RB \mid R) = \frac{(1/2)(1/3)}{(1)(1/3) + (1/2)(1/3) + (0)(1/3)} = \frac{1}{3}$$

The experiment consists of six equally likely outcomes:

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

The other side of the upturned red side is black only if the outcome is (R_3).

Thus, the conditional probability is (1/3).

---

3. Random Variables

3.1 Motivation and Concept

When an experiment is performed, interest may lie in a function of the outcome rather than the outcome itself.

Examples:

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

These real-valued functions defined on the sample space are called **random variables**.

- Values are determined by outcomes of an experiment
- Probabilities are assigned to possible values

The distribution of a random variable can be visualized as a bar diagram:

- x-axis: values the random variable can take
- height at value (a): (Pr$[X = a]$)

3.2 Examples

Example 1: Tossing Three Fair Coins

The experiment consists of tossing three fair coins.

Let (Y) denote the 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 the values (0) through (3):

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

---

4. Probability Mass Function (PMF)

4.1 Concept

A random variable that can take at most a countable number of possible values is said to be **discrete**.

Let (X) be a discrete random variable with range:

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

The function:

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

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

Since (X) must take one of the values (x_k) : $\sum_k p(x_k) = 1$

4.2 Example: Two Independent Tosses of a Fair Coin

(As presented in the lecture slides.)

4.3 Example

The probability mass function of a random variable (X) is given by:

$$p(i) = c\lambda^i, \quad i = 0, 1, 2, \ldots$$

where ($\lambda > 0$).

Find:

- ($\Pr(X = 0)$)
- ($\Pr(X > 2)$)

(As stated and shown in the lecture slides.)

---

5. Class Participation

Students were instructed to switch to **Campuswire** for class participation and discussion.