CSE400: Fundamentals of Probability in Computing

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

Swayam Prajapati, AU2440087

Group 17

February 06, 2026

Contents

1 Bayes' Theorem

1.1 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.

As AB and AB^c are mutually exclusive, we have, by Axiom 3,

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

Using conditional probability,

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

Hence, the probability of event A is a weighted average of the 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 believes that people can be divided into two classes: those who are accident prone and those who are not. The company's statistics show that an accident-prone person will have an accident within a fixed one-year period with probability 0.4, whereas this probability is 0.2 for a person who is not accident prone. Assume that 30% of the population is accident prone.

Find: The probability that a new policyholder will have an accident within one year.

Solution:

Let

$$A_1 = \{\text{policyholder has an accident within one year}\}$$
$$A = \{\text{policyholder is accident prone}\}$$

Then,

$$\begin{aligned}\Pr(A_1) &= \Pr(A_1 \mid A)\Pr(A) + \Pr(A_1 \mid A^c)\Pr(A^c)\\ &= (0.4)(0.3) + (0.2)(0.7) = 0.26\end{aligned}$$

Example 3.1 (Part 2)

Suppose that a new policyholder has an accident within one year of purchasing a policy. What is the probability that he or she is accident prone?

Solution:

The desired probability is

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

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

$$= \frac{(0.3)(0.4)}{0.26} = \frac{6}{13}$$

1.3 Formal Introduction: Law of Total Probability

Suppose that $B_1, B_2, \ldots, B_n$ are mutually exclusive events such that

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

By writing

$$A = \bigcup_{i=1}^{n} AB_i$$

and using the fact that the events AB_i are mutually exclusive, we obtain

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

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

1.4 Bayes Formula

Using

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

we get

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

This is known as the **Bayes Formula (Proposition 3.1)**.
Here,

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

1.5 Example 3.2 (Cards)

Three cards are identical except:

- One card: red–red (RR)
- One card: black–black (BB)
- One card: red–black (RB)

One card is randomly selected and placed on the ground. If the upturned side is red, find the probability that the other side is black.

Let

$$R = \{\text{upturned side is red}\}$$

The desired probability is

$$\Pr(RB \mid R) = \frac{\Pr(RB \cap R)}{\Pr(R)}$$

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

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

2 Random Variables

2.1 Motivation and Concept

A random variable is a real-valued function defined on the sample space.

- Dice tossing: focus on the sum of dice
- Coin tossing: focus on the number of heads

Definition: A random variable X on a sample space Ω is a function

$$X : \Omega \to \mathbb{R}$$

We restrict attention to discrete random variables unless stated otherwise.

2.2 Distribution

For a random variable X, let

$$\{\omega \in \Omega : X(\omega) = a\}$$

denote an event, abbreviated as $\{X = a\}$.

The collection of probabilities

$$\Pr(X = a)$$

for all possible values of a is called the **distribution** of X.

These probabilities can be visualized using a bar diagram.

2.3 Example

Example 1: Tossing three fair coins.

Let Y be the number of heads.

$$\Pr(Y=0)=\frac{1}{8},\quad \Pr(Y=1)=\frac{3}{8},\quad \Pr(Y=2)=\frac{3}{8},\quad \Pr(Y=3)=\frac{1}{8}$$

Since Y must take one of the values 0 through 3,

$$1=\sum_{i=0}^{3}\Pr(Y=i)$$

3 Probability Mass Function

3.1 Concept

A random variable that can take on at most a countable number of 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

$$p_X(x)=\Pr(X=x)$$

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

Since X must take one of the values x_k,

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

3.2 Example: Two Independent Tosses of a Fair Coin

Let X be the number of heads obtained.

$$p_X(x)=\begin{cases}\frac{1}{4}, & x=0 \text{ or } 2\\ \frac{1}{2}, & x=1\\ 0, & \text{otherwise}\end{cases}$$

$$\Pr(X>0)=\Pr(X=1)+\Pr(X=2)=\frac{3}{4}$$

3.3 Example

The PMF of a random variable X is given by

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

where $\lambda > 0$.

Since

$$\sum_{i=0}^{\infty} p(i) = 1$$

$$c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = 1$$

and since

$$e^{\lambda} = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}$$

we obtain

$$c = e^{-\lambda}$$