CSE 400: Fundamentals of Probability in Computing
Lecture 5: Bayes' Theorem, Random Variables, and Probability Mass Function
Group No.: S1 G17
Date: January 20, 2026

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.

Since AB and AB^c are mutually exclusive, 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 1 (Example 3.1, Part 1)**.** *An insurance company divides people into two classes: accident-prone and not accident-prone. The probability that an accident-prone person has an accident within one year is* 0.4*, while for a non-accident-prone person it is* 0.2*. If* 30% *of the population is accident-prone, find the probability that a new policyholder has an accident within one year.*

Solution. Let A_1 denote the event that the policyholder has an accident within one year, and let A denote the event that the policyholder is accident-prone. Then,

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

Substituting values,

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

Example 2 (Example 3.1, Part 2)**.** *Suppose that a new policyholder has an accident within one year. What is the probability that the policyholder is accident-prone?*

Solution. The desired probability is

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

Substituting values,

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

1.3 Formal Introduction: Law of Total Probability and Bayes Formula

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

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

Then, exactly one of the events $B_1, B_2, \ldots, B_n$ must occur. 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).

Proposition 1 (Bayes Formula)**.** *For events A and $B_1, B_2, \ldots, B_n$ as above,*

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

Here, $\Pr(B_i)$ is the *a priori* probability, and $\Pr(B_i \mid A)$ is the *posteriori* probability of B_i given A.

1.4 Learning by Example

Example 3 (Example 3.2)**.** *Three cards are identical in form: one has both sides red (RR), one has both sides black (BB), and one has one side red and one side black (RB). A card is randomly selected and placed on the ground. If the upper side is red, find the probability that the other side is black.*

Solution. Let RR, BB, and RB denote the events that the chosen card is all red, all black, or red-black, respectively. Let R denote the event that the upturned side is red. Then,

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

Substituting values,

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

2 Random Variables

2.1 Motivation and Concept

When an experiment is performed, interest often lies in a function of the outcome rather than the outcome itself. These real-valued functions defined on the sample space are called **random variables**.

Definition 1. *A random variable X on a sample space Ω is a function*

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

that assigns a real number $X(\omega)$ to each sample point $\omega \in \Omega$.

Unless otherwise stated, attention is restricted to discrete random variables, which take values in a finite or countably infinite subset of $\mathbb{R}$.

2.2 Distribution of a Random Variable

Let a be any number in the range of a random variable X. The event

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

is denoted by $X = a$. The probability $\Pr(X = a)$ is defined as the probability of this event. The collection of probabilities $\Pr(X = a)$ for all possible values of a is called the **distribution** of X.

2.3 Example

Example 4. *Suppose an experiment consists of tossing three fair coins. Let Y denote the number of heads that appear.*

Then Y takes values $0, 1, 2, 3$ with probabilities

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

Since Y must take one of these values,

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

3 Probability Mass Function

3.1 Concept

Definition 2. *A random variable that can take on 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

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

is called the **probability mass function (PMF)** of X. Since X must take one of the values in R_X,

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

3.2 Example: Two Independent Tosses of a Fair Coin

Let X denote the number of heads obtained when a fair coin is tossed twice. Then the PMF of X is

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

Hence,

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

3.3 Example

The probability mass function 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

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

we have

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

Using

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

it follows that

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

Therefore,

$$\Pr(X = 0) = p(0) = e^{-\lambda},$$

and

$$\Pr(X > 2) = 1 - [p(0) + p(1) + p(2)].$$