

Group No.: S1 G17

Lecture: 5

Date: January 20, 2026

1 Bayes' Theorem

1.1 Weighted Average of Conditional Probabilities

Let A and B be events in a probability space. The event A can be expressed as

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

since any outcome in A must either occur together with B or with B^c .

The events AB and AB^c are mutually exclusive. By Axiom 3 of probability,

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

Using the definition of conditional probability,

$$\Pr(AB) = \Pr(A | B) \Pr(B), \quad \Pr(AB^c) = \Pr(A | B^c) \Pr(B^c).$$

Hence,

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

Thus, the probability of an event can be written as a weighted average of conditional probabilities.

1.2 Learning by Example

Example 3.1 (Part 1)

Let

- A_1 : event that a policyholder has an accident within one year,
- A : event that the policyholder is accident-prone.

Given:

$$\Pr(A_1 | A) = 0.4, \quad \Pr(A_1 | A^c) = 0.2, \quad \Pr(A) = 0.3.$$

Using conditioning,

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

Therefore,

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

Example 3.1 (Part 2)

Given that the policyholder has an accident within a year, the probability that the policyholder is accident-prone is

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

1.3 Law of Total Probability

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

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

Then for any event A ,

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

where the events $A \cap B_i$ are mutually exclusive.

Hence,

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

This result is known as the **Law of Total Probability**.

1.4 Bayes Formula

Using $\Pr(A \cap B_i) = \Pr(B_i \mid A) \Pr(A)$ in the law of total probability, 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).

Here,

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

1.5 Example 3.2 (Card Problem)

Three cards are given: RR , BB , and RB . One card is chosen at random and placed on the ground.

Let

RR, BB, RB = events that the chosen card is of the corresponding type,

and let R be the event that the upturned side is red.

The desired probability is

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

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 Definition

A **random variable** is a real-valued function defined on the sample space Ω :

$$X : \Omega \rightarrow \mathbb{R},$$

assigning a real number $X(\omega)$ to each outcome $\omega \in \Omega$.

In this lecture, 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

For a random variable X and a value a in its range,

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

is an event, 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

Let Y be the number of heads obtained when three fair coins are tossed.

The possible values of Y are $\{0, 1, 2, 3\}$ 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,

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

3 Probability Mass Function

3.1 Definition

A random variable that can take 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

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

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

Since X must take one of its possible values,

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

3.2 Example: Two Independent Tosses of a Fair Coin

Let

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\},$$

and define X as the number of heads obtained.

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 PMF of a random variable X is given by

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

where $\lambda > 0$.

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

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

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

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

Thus,

$$\Pr(X = 0) = p(0) = e^{-\lambda}, \quad \Pr(X > 2) = 1 - \sum_{i=0}^2 e^{-\lambda} \frac{\lambda^i}{i!}.$$

End of Lecture 5 Scribe