Week 12: Bayes' rule and random variables

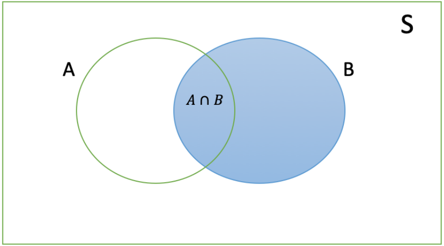
Table of contents

Experiment, sample spaces, events

- An **experiment** is a situation with uncertain outcomes.
- The **sample space** of an experiment is the set Ω of all possible outcomes of the experiment.
- An **event** is a subset of the sample space Ω.

Conditional probability

- Conditional probability: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$



- The sample space for *possible outcomes* is B.
  $P(A|B) =$ probability of event $A \cap B$ in the sample space B.

Independent events

- A and B are **independent** $\Leftrightarrow$ one of the following equations holds

  $P(A \cap B) = P(A)P(B)$, or $P(A|B) = P(A)$, or $P(B|A) = P(B)$.

- A and B are independent means

  "the information that B occurs does not affect the probability that A occurs, and vice versa"

Multiplication rule for conditional probability

Exercise 1. Let A, B, C be events. Show that

(a) $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$.

In particular if A, B are independent, then $P(A \cap B) = P(A)P(B)$.

(b) $P(A \cap B \cap C) = P(A|B \cap C)P(B|C)P(C)$.

If A, B, C are **mutually independent**, then

$P(A \cap B \cap C) = P(A)P(B)P(C)$

Mutually independent

Three events A, B, C are called **mutually independent** if

1. any two events are independent

$$P(A \cap B) = P(A)P(B),\ P(A \cap C) = P(A)P(C),\ P(B \cap C) = P(B)P(C)$$

2. and $P(A \cap B \cap C) = P(A)P(B)P(C)$

Example 1

You have a flight from Amsterdam to Sydney with a stopover in Dubai. The probabilities that a luggage is put on the wrong plane at the different airports are

$$\text{Amsterdam} : 0.05, \ \text{Dubai} : 0.03.$$

What is the probability that your luggage doesn't reach Sydney with you?

Solution

Example 1

You have a flight from Amsterdam to Sydney with a stopover in Dubai. The probabilities that a luggage is put on the wrong plane at the different airports are

$$\text{Amsterdam} : 0.05, \;\; \text{Dubai} : 0.03.$$

What is the probability that your luggage doesn't reach Sydney with you?

Solution

A=event that the luggage is put on the correct plane at Amsterdam

D=event that the luggage is put on the correct plane at Dubai.

You may assume A and D are independent (why?)

Example 1

You have a flight from Amsterdam to Sydney with a stopover in Dubai. The probabilities that a luggage is put on the wrong plane at the different airports are

$$\text{Amsterdam} : 0.05, \ \text{Dubai} : 0.03.$$

What is the probability that your luggage doesn't reach Sydney with you?

Solution

A=event that the luggage is put on the correct plane at Amsterdam

D=event that the luggage is put on the correct plane at Dubai.

You may assume A and D are independent (why?)

$\overline{A \cap D}$ = event that your luggage doesn't reach Sydney is

$$P\left(\overline{A \cap D}\right) = 1 - P(A \cap D)$$

Example 2

You have a flight from Amsterdam to Sydney with stopovers in Dubai and Singapore. The probabilities that a luggage is put on the wrong plane at the different airports are

$$\text{Amsterdam} : 0.05, \ \text{Dubai} : 0.03, \ \text{Singapore} : 0.01.$$

What is the probability that your luggage doesn't reach Sydney with you?

Solution

Example 2

You have a flight from Amsterdam to Sydney with stopovers in Dubai and Singapore. The probabilities that a luggage is put on the wrong plane at the different airports are

$$\text{Amsterdam} : 0.05, \ \text{Dubai} : 0.03, \ \text{Singapore} : 0.01.$$

What is the probability that your luggage doesn't reach Sydney with you?

Solution

A=event that the luggage is put on the correct plane at Amsterdam

D=event that the luggage is put on the correct plane at Dubai.

S= event that the luggage is put on the correct plane at Singapore.

You may assume A, D, S are mutually independent.

The probability that your luggage doesn't reach Sydney is

$$P\left(\overline{A \cap D \cap S}\right) = 1 - P(A \cap D \cap S)$$

Partition - Definition

- $B_1, \ldots, B_n$ is a **partition** of Ω if
  1. $\cup_{i=1}^{n} B_i = \Omega$ and
  2. $B_1, \ldots, B_n$ are pairwise disjoint

Partition - Definition

- $B_1, \ldots, B_n$ is a **partition** of Ω if
  1. $\cup_{i=1}^{n} B_i = \Omega$ and
  2. $B_1, \ldots, B_n$ are pairwise disjoint
- Examples
  1. $\{1\}, \{2,3\}$ is a partition of $\{1,2,3\}$.
  2. $\{1,2\}, \{2,3\}$ is not a partition of $\{1,2,3\}$.
  3. B and $\bar{B}$ is a partition of Ω.

Law of total probability

Theorem 1

Let P be a probability measure on Ω. Assume that $B_1, \ldots, B_n$ is a partition of Ω. Then for any event A, we have

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$

Corollary of Theorem 1

Corollary 1

Let A and B be events in the sample space Ω. Then

$$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B}).$$

Complement of conditional events

- A **conditional event** is an event of the form $A|B$, read as "A given B".
- The complement of $A|B$ is $\bar{A}|B$.

Complement of conditional events

- A **conditional event** is an event of the form $A|B$, read as "A given B".
- The complement of $A|B$ is $\bar{A}|B$.
- Question: What is the relation between $P(A|B)$ and $P(\bar{A}|B)$?

$$\begin{aligned} P(A|B) + P(\bar{A}|B) &= P(A)? \\ P(A|B) + P(\bar{A}|B) &= P(B)? \\ P(A|B) + P(\bar{A}|B) &= 1? \end{aligned}$$

Probability of the complement of a conditional event

Lemma 2

Let A, B be two events with $P(B) > 0$. Then

$$P(A|B) + P(\bar{A}|B) = 1$$

Probability of the complement of a conditional event

Lemma 2

Let A, B be two events with $P(B) > 0$. Then

$$P(A|B) + P(\bar{A}|B) = 1$$

- Since A and $\bar{A}$ for a partition for Ω

$$P(B) = P(B \cap A) + P(B \cap \bar{A}) = P(A \cap B) + P(\bar{A} \cap B),$$

- Hence

$$\begin{aligned} P(B) &= P(A|B)P(B) + P(\bar{A}|B)P(B) \\ 1 &= P(A|B) + P(\bar{A}|B) \end{aligned}$$

Bayes' rule (simplified version)

Theorem 2

Let A, B be events with $P(A) > 0$, $P(B) > 0$. Then

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

Bayes' rule (simplified version)

Theorem 2

Let A, B be events with $P(A) > 0$, $P(B) > 0$. Then

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

- $P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$

Bayes' rule (simplified version)

Theorem 2

Let A, B be events with $P(A) > 0$, $P(B) > 0$. Then

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

- $P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$
- Since B and $\bar{B}$ form a partition of Ω,

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$$

Bayes' rule (simplified version)

Theorem 2

Let A, B be events with $P(A) > 0$, $P(B) > 0$. Then

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

- $P(B|A) = \dfrac{P(B \cap A)}{P(A)} = \dfrac{P(A \cap B)}{P(A)} = \dfrac{P(A|B)P(B)}{P(A)}$
- Since B and $\bar{B}$ form a partition of Ω,

$$P(A) = P(A \cap B) + P(A \cap \bar{B}) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$$

- $P(B|A) = \dfrac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$

Example 3

1 in $100,000$ people has a rare disease for which there is a fairly accurate diagnostic test. This test is correct 99.0% of the time when given to a person selected at random who has the disease, and correct 99.5% of the time when given to a person selected at random who does not have the disease. Find

(a) The probability that a person who tests positive actually has the disease?

(b) The probability that a person who tests negative does not have the disease?

Solution

A=event that a randomly selected person has the disease.

B=event that a randomly selected person tests positive.

Need to compute $P(A|B)$ and $P(\bar{A}|\bar{B})$.

Bayes' rule (general version)

Theorem 3

Let $A_1, A_2, \ldots, A_n$ be a partition of Ω and let A be an event. Assume $P(A) > 0$ and $P(A_i) > 0$ for all i. Then for any $k \in \{1, \ldots, n\}$, we have

$$P(A_k|A) = \frac{P(A|A_k)P(A_k)}{\sum_{i=1}^{n} P(A|A_i)P(A_i)}.$$

Interpretation of Bayes' rule

$$P(A_k|A) = \frac{P(A|A_k)P(A_k)}{\sum_{i=1}^{n} P(A|A_i)P(A_i)}$$

- A_i's are possible causes for the occurrence of A.
- The Bayes' formula computes the probability that A_k caused A, given that A occurred.

Proof of Theorem 2

Writing $P(A_k \cap A)$ in two different ways, we have

$$P(A_k|A)P(A) = P(A|A_k)P(A_k) \Rightarrow P(A_k|A) = \frac{P(A|A_k)P(A_k)}{P(A)}$$

Summary on Bayes' rules

- Partition $\Omega = B \cup \bar{B}$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|\bar{B})P(\bar{B})}$$

- Partition $\Omega = A_1 \cup A_2 \cup A_3$

$$P(A_1|A) = \frac{P(A|A_1)P(A_1)}{P(A|A_1)P(A_1) + P(A|A_2)P(A_2) + P(A|A_3)P(A_3)}$$

- General partition $\Omega = A_1 \cup \cdots \cup A_n$

$$P(A_1|A) = \frac{P(A|A_1)P(A_1)}{P(A|A_1)P(A_1) + \cdots + P(A|A_n)P(A_n)}$$

Example 4

A factory uses 3 machines M_1, M_2, M_3 to produce certain items.

- M_1 produces 50% of the items, of which 3% are defective.
- M_2 produces 30% of the items, of which 4% are defective.
- M_3 produces 20% of the items, of which 5% are defective.

Suppose that a defective item is found. What is the probability that it came from M_2?

$A_1, A_2, A_3 =$ events that a given item comes from M_1, M_2, M_3.
$A =$ event that a given item is defective.

Discussion

- A dice is thrown three times

$$\Omega = \{abc : a, b, c \in \{1, \ldots, 6\}\}$$

- Usually we are not interested in the whole Ω (too complex), but only extract information of interests, for examples,

Discussion

- A dice is thrown three times

$$\Omega = \{abc : a, b, c \in \{1, \ldots, 6\}\}$$

- Usually we are not interested in the whole Ω (too complex), but only extract information of interests, for examples,
  - the sum of all numbers that show up, or
  - the number of sixes, or
  - the number of ones
- Each of these quantities is a random variable.

Random variables

- A **random variable** on the sample space Ω is a function

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

  that is, X assigns a real number to each possible outcome.
- Capital letters $X, Y, Z, \ldots$ denote random variables.
  Small letters $x, y, z, \ldots$ denote **possible values** of X, Y, Z.

Example 5

- $X = \#$ heads in 3 coin tosses.

$$\Omega = \{\text{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}\}.$$

Example 5

- $X = \#$ heads in 3 coin tosses.

$$\Omega = \{\text{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}\}.$$

- X is a function $X : \Omega \to \mathbb{R}$

$$\begin{aligned} X(\text{HHH}) &= 3, \\ X(\text{HHT}) &= X(\text{HTH}) = X(\text{THH}) = 2, \\ X(\text{HTT}) &= X(\text{THT}) = X(\text{TTH}) = 1, \\ X(\text{TTT}) &= 0 \end{aligned}$$

Example 5

- $X = \#$ heads in 3 coin tosses.

$$\Omega = \{\text{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}\}.$$

- X is a function $X : \Omega \to \mathbb{R}$

$$\begin{aligned} X(\text{HHH}) &= 3, \\ X(\text{HHT}) &= X(\text{HTH}) = X(\text{THH}) = 2, \\ X(\text{HTT}) &= X(\text{THT}) = X(\text{TTH}) = 1, \\ X(\text{TTT}) &= 0 \end{aligned}$$

- The set of possible values of X is $\{0, 1, 2, 3\}$.

Example 6

$X = \#$ number of heads in 3 consecutive fair-coin tosses.
Find $P(X = 3), P(X \leq 1)$ and $P(X \neq 2)$.

Example 7

$p \in [0, 1]$. A calibrated coin has chance of landing head is p.

$X = \#$ tosses until a head comes up.

Given $n \in \mathbb{Z}^+$. Find $P(X = n)$ and $P(X \leq n)$.

Discrete random variables

A random variable is **discrete** if it takes on only *countably many values*, that is, the set of possible values of X is countable.

- Countable means there is an order to list out everything. Examples
  1. Any finite subset $S = \{a_1, a_2, \ldots, a_n\}$ of $\mathbb{R}$ is countable.
  2. $\mathbb{N} = \{0, 1, 2, \ldots\}$ is countable.
     $\mathbb{Z}^+ = \{1, 2, 3, \ldots\}$ is countable.

Discrete random variables

A random variable is **discrete** if it takes on only *countably many values*, that is, the set of possible values of X is countable.

- Countable means there is an order to list out everything. Examples
  1. Any finite subset $S = \{a_1, a_2, \ldots, a_n\}$ of $\mathbb{R}$ is countable.
  2. $\mathbb{N} = \{0, 1, 2, \ldots\}$ is countable.
     $\mathbb{Z}^+ = \{1, 2, 3, \ldots\}$ is countable.
  3. $\mathbb{R}$ is not countable.
     $[0, 1]$ is not countable.
  4. We only focus on discrete random variables in this course

Examples of discrete random variables

- X = number of heads in 3 coin tosses
  The set of possible values of X is $\{0, 1, 2, 3\} \Rightarrow$ countable.

Examples of discrete random variables

- X = number of heads in 3 coin tosses
  The set of possible values of X is $\{0, 1, 2, 3\} \Rightarrow$ countable.
- X = number of coin tosses until a head comes up
  The set of possible values of X is $\mathbb{Z}^+ = \{1, 2, 3, \dots\} \Rightarrow$ countable.

Probability mass function (PMF)

The **probability mass function (PMF)** of a discrete random variable X is a function $p : \mathbb{R} \to [0, 1]$ defined by

$$p(x) = P(X = x)$$

Example 8

- $X = \#$ heads in 3 independent fair-coin tosses.
  The set of possible values of X is $\{0, 1, 2, 3\}$ and

Example 8

- $X = \#$ heads in 3 independent fair-coin tosses.
  The set of possible values of X is $\{0, 1, 2, 3\}$ and

$$p(0) = P(X = 0) = P(\{\text{TTT}\}) = 1/8$$
$$p(1) = P(X = 1) = P(\{\text{HTT,THT,TTH}\}) = 3/8$$
$$p(2) = P(X = 2) = P(\{\text{HHT,HTH,THH}\}) = 3/8$$
$$p(3) = P(X = 3) = P(\{\text{HHH}\}) = 1/8$$

Example 8

- $X = \#$ heads in 3 independent fair-coin tosses.
  The set of possible values of X is $\{0, 1, 2, 3\}$ and
  $$p(0) = P(X = 0) = P(\{\text{TTT}\}) = 1/8$$
  $$p(1) = P(X = 1) = P(\{\text{HTT,THT,TTH}\}) = 3/8$$
  $$p(2) = P(X = 2) = P(\{\text{HHT,HTH,THH}\}) = 3/8$$
  $$p(3) = P(X = 3) = P(\{\text{HHH}\}) = 1/8$$
- Note that
  $$p(0) + p(1) + p(2) + p(3) = 1.$$

Properties of PMF

Lemma 3

If $X : \Omega \to R$ is a discrete random variable with PMF $p(x)$. Then

$$\sum_{\text{all x}} p(x) = 1$$

Cumulative distribution function (CDF)

- The **cumulative distribution function (CDF)** of a random variable $X : \Omega \to \mathbb{R}$ is a function $F : \mathbb{R} \to [0, 1]$ defined by

$$F(x) = P(X \leq x).$$

Cumulative distribution function (CDF)

- The **cumulative distribution function (CDF)** of a random variable $X : \Omega \to \mathbb{R}$ is a function $F : \mathbb{R} \to [0, 1]$ defined by

$$F(x) = P(X \leq x).$$

- F is a *nondecreasing* function, that is,

$$F(a) \leq F(b) \text{ whenever } a \leq b.$$

Example 9

$X = \#$ heads in 3 independent fair-coin tosses.

Find $p(x)$ and $F(x)$ for all possible values x of X.