Lecture note 3: Conditional probability

Foundation of Probability Theory/STA 203

Zhuosong ZHANG

Department of Statistics and Data Science, SUSTech

Fall, 2023

Learning objectives

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

- Understanding the concept of conditional probability.
- Be able to calculate it for simple examples and real-world problems.
- Apply multiplication rule to calculate probabilities in more complex situations.
- Learn about the law of total probability and its applications in calculating conditional probabilities.
- Use Bayes' theorem to find conditional probabilities.
- Understand the concept of independence, and know the difference between independent events and disjoint events.
- Learn about the concepts of experiments and independent trials.

Conditional probability

Example 1 (Sex and Sports)

Two psychologists surveyed 478 children in grades 4, 5, and 6 in elementary schools in Michigan. Among other questions, they asked the students whether their primary goal was to get good grades, to be popular, or to be good at sports. Here is a contingency table giving the counts of the students by their goals and sex:

	Grades	Popular	Sports	Total
Boys	117	50	60	**227**
Girls	130	91	30	**251**
Total	**247**	**141**	**90**	**478**

Questions:

- What is the probability that a randomly chosen student is to excel at sports?
- What if we are given the information that the selected student is a girl?

Solution.

- For the first question, by the classical probability model, let Ω be the collection of all the 478 students, E be the students whose goal was to be good at sports. Then,

$$\mathbb{P}(E) = \frac{\#\text{ of students excel at sports}}{\text{Total number of students}} = \frac{90}{478} \approx 0.188.$$

- For the second question, the probability space has changed. We now let Ω' be the collection of all girls in these 478 students, and let E' be the collection of girls whose goal was to good at sports. Then,

$$\mathbb{P}(E') = \frac{\#\text{ of female students excel at sports}}{\text{Total number of female students}} = \frac{30}{251} \approx 0.120. \qquad \blacksquare$$

The probability might change if we are provided with more information!

Conditional probability

- Conditional probability provides us with a way to reason about the outcome of an experiment, based on partial information.
- In an experiment involving two successive rolls of a die, you are told that the sum of the two rolls is 9. How likely is it that the first roll was a 6?
- In a word guessing game, the first letter of the word is a “t”. What is the likelihood that the second letter is an “h”?
- How likely is it that a person has a certain disease given that a medical test was negative?
- A spot shows up on a radar screen. How likely is it to correspond to an aircraft?

Definition 2 (Conditional probability)

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and $E, F \in \mathscr{F}$ be two events such that $\mathbb{P}(F) > 0$. The conditional probability of E given F is defined by

$$\mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Example 3

In the "sex and sports" example, let F be the event that the randomly chosen student is a girl. Then,

$$\mathbb{P}(E|F) = \frac{\frac{\#\text{of female students who are excel at sports}}{\text{Total number of students}}}{\frac{\#\text{ female students}}{\text{Total number of students}}} = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)}.$$

Example 4

In an experiment involving two successive rolls of a die, you are told that the sum of the two rolls is 9. How likely is it that the first roll was a 6?

Solution.

Let E be the event that the first roll was 6, and F be the event that the sum of the two rolls is 9. We have

$$\mathbb{P}(E \cap F) = \mathbb{P}(\{(6,3)\}) = \frac{1}{36},$$

$$\begin{aligned}\mathbb{P}(F) &= \mathbb{P}(\{(3,6),(4,5),(5,4),(6,3)\}) \\ &= \frac{4}{36} = \frac{1}{9}.\end{aligned}$$

Then,

$$\mathbb{P}(E|F) = \frac{1/36}{1/9} = \frac{1}{4}. \quad \blacksquare$$

Conditional probabilities specify a probability law

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and let F be an event such that $\mathbb{P}(F) > 0$. Define the conditional probability map $\widetilde{\mathbb{P}}(\cdot) : \mathscr{F} \to [0, 1]$ as

$$\widetilde{\mathbb{P}}(E) = \mathbb{P}(E|F).$$

Then, $(\Omega, \mathscr{F}, \widetilde{\mathbb{P}})$ is also a probability space.

Remark

Because $(\Omega, \mathscr{F}, \widetilde{\mathbb{P}})$ is a probability space, $\mathbb{P}(\cdot|F)$ satisfies all of the properties of a general probability, for example,

$$\mathbb{P}(\varnothing|F) = 0, \quad \mathbb{P}(E^c|F) = 1 - \mathbb{P}(E|F), \quad \mathbb{P}(E_1|F) \leqslant \mathbb{P}(E_2|F) \text{ if } E_1 \subset E_2$$

However,

$$\mathbb{P}(E|F^c) \neq 1 - \mathbb{P}(E|F) \quad \text{in general.}$$

To show that $\widetilde{\mathbb{P}}$ is a probability measure, we only need to show the three axioms.

(i) Non-negativity: For any $E \in \mathscr{F}$, $\mathbb{P}(E) \geqslant 0$, then

$$\widetilde{\mathbb{P}}(E) = \mathbb{P}(E|F) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)} \geqslant 0.$$

(ii) Normalization:

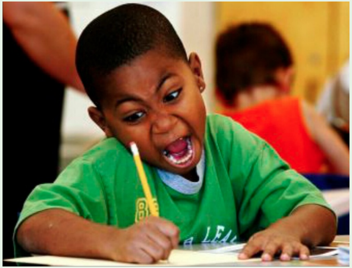
$$\widetilde{\mathbb{P}}(\Omega) = \frac{\mathbb{P}(\Omega \cap F)}{\mathbb{P}(F)} = \frac{\mathbb{P}(F)}{\mathbb{P}(F)} = 1.$$

(iii) Countable additivity: Let $E_1, E_2, \ldots$ be a sequence of mutually exclusive events, then $(E_1 \cap F), (E_2 \cap F), \ldots$ are also mutually exclusive. Then,

$$\widetilde{\mathbb{P}}\left(\bigcup_{i=1}^{\infty} E_i\right) = \frac{1}{\mathbb{P}(F)}\, \mathbb{P}\left(\bigcup_{i=1}^{\infty} (E_i \cap F)\right) = \frac{1}{\mathbb{P}(F)} \sum_{i=1}^{\infty} \mathbb{P}(E_i \cap F) = \sum_{i=1}^{\infty} \widetilde{\mathbb{P}}(E_i).$$

Example 5 (Examination time)

A student is taking a one-hour-time-limit makeup examination. Suppose the probability that the student will finish the exam in less than x hours is $x/2$, for all $0 \leqslant x \leqslant 1$. Then, given that the student is still working after 0.75 hour, what is the conditional probability that the full hour is used?



Solution.

Let $\Omega = [0,1] \cup \{N\}$ be the sample space where $x \in [0,1]$ represents that the student finishes the exam in less than x hours, and N represents that the student didn't finish the exam in one hour; in other words, the full hour is used. Then, we can define the probability measure as follows: for any Borel set $A \subset [0,1]$, $\mathbb{P}(A) = \frac{m(A)}{2}$, where $m(A)$ is the Lebesgue measure of A, and $\mathbb{P}(\{N\}) = 1 - \mathbb{P}([0,1]) = 0.5$.

Let E be the event that the student didn't finish the exam in one hour, and let F be the event that the student is still working after 0.75 hour, then

$$E = \{N\}, \quad F = (0.75, 1] \cup \{N\},$$

and

$$\begin{aligned} \mathbb{P}(E \cap F) &= \mathbb{P}(E) = \mathbb{P}(\{N\}) = 0.5, \\ \mathbb{P}(F) &= \mathbb{P}((0.75, 1]) + \mathbb{P}(\{N\}) \\ &= \frac{0.25}{2} + \frac{1}{2} = 0.625. \end{aligned}$$

Therefore,

$$\begin{aligned} \mathbb{P}(E|F) &= \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(F)} \\ &= \frac{0.5}{0.625} = 0.8. \end{aligned}$$

■

Example

Sometimes the conditional probability can be derived directly by properly choosing a probability space.

Example 6 (Bridge card)

In the card game bridge, the 52 cards are dealt out equally to 4 players—called East, West, North, and South. If North and South have a total of 8 spades among them, what is the probability that East has 3 of the remaining 5 spades?

Solution.

Let Ω be the collections of all possible cards of the East given that 26 cards has been given to the North and South, among which there are 8 spades. Then,

$$N = |\Omega| = \binom{26}{13}.$$

Let E be the event that the East has 3 of the remaining 5 cards, then

$$|E| = \binom{5}{3}\binom{26-5}{13-3} = \binom{5}{3}\binom{21}{10}.$$

Therefore,

$$\mathbb{P}(E) = \frac{\binom{5}{3}\binom{21}{10}}{\binom{26}{13}} \approx 0.339. \qquad \blacksquare$$

Multiplication rule

Multiplication rule

From the definition of conditional probability,

$$\mathbb{P}(E \cap F) = \mathbb{P}(E|F)\,\mathbb{P}(F).$$

In words, the probability that both E and F occur is equal to the probability that F occurs multiplied by the conditional probability of E given that F occurred.

Example

Example 7 (Course taking)

Celine is undecided as to whether to take a French course or a chemistry course. She estimates that her probability of receiving an A grade would be $1/2$ in a French course and $2/3$ in a chemistry course. If Celine decides to base her decision on the flip of a fair coin, what is the probability that she gets an A in chemistry?

Solution.

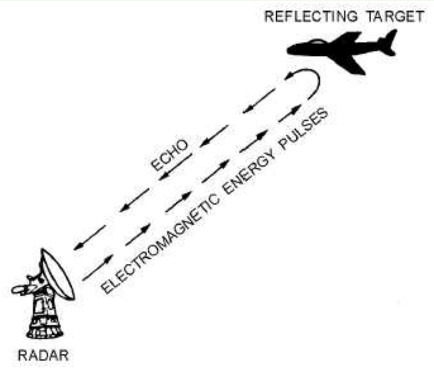
Let C denote the event that Celina takes chemistry and A denote the event that she receives an A in whatever course she takes. Then, the probability that she gets an A in chemistry is

$$\mathbb{P}(A \cap C) = \mathbb{P}(C)\,\mathbb{P}(A|C) = \frac{1}{2} \times \frac{2}{3} = \frac{1}{3}. \qquad \blacksquare$$

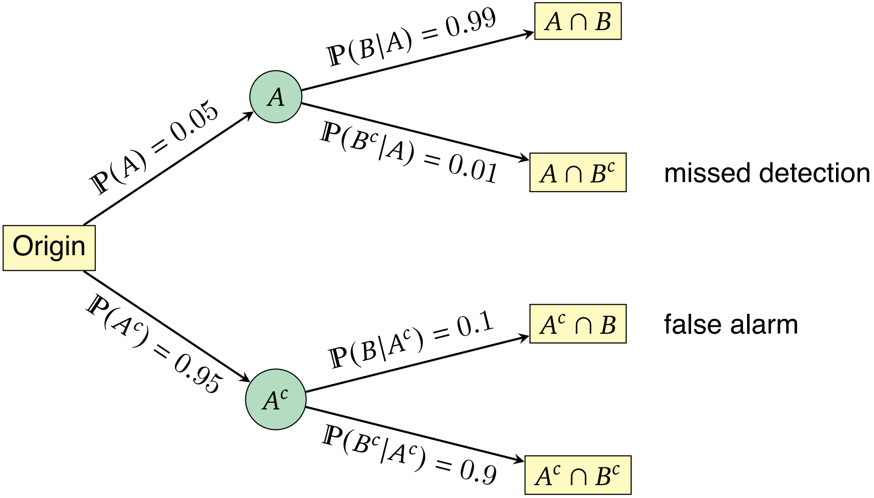
南方科技大學
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 8 (Aircraft detection)

If an aircraft is present in a certain area, a radar detects it and generates an alarm signal with probability 0.99. If an aircraft is not present. The radar generates a (false) alarm, with probability 0.10. We assume that an aircraft is present with probability 0.05. What is the probability of no aircraft presence and a false alarm? What is the probability of aircraft presence and no detection?



Let A be the event that an aircraft is present, and let B be the event that the radar generates an alarm. Then, we have the following graph:



Examples

Example 9 (Drawing balls)

Suppose that an urn contains 8 red balls and 4 white balls. We draw 2 balls from the urn without replacement.

(a) If we assume that at each draw each ball in the urn is equally likely to be chosen, what is the probability that both balls drawn are red?

(b) Now suppose that the balls have different weights, with each red ball having weight r and each white ball having weight w. Suppose that the probability that a given ball in the urn is the next one selected is its weight divided by the sum of the weights of all balls currently in the urn. Now what is the probability that both balls are red?

Solution to (a).

Let R_1 and R_2 denote the events that the first and second balls drawn are red, respectively. Then,

$$\mathbb{P}(R_1) = \frac{8}{12} = \frac{2}{3}.$$

Given that the first ball selected is red, there are 8 remaining red balls and 4 white balls, so

$$\mathbb{P}(R_2|R_1) = \frac{7}{11}.$$

Therefore, by the multiplication rule,

$$\mathbb{P}(R_1 \cap R_2) = \mathbb{P}(R_1)\,\mathbb{P}(R_2|R_1) = \left(\frac{2}{3}\right)\left(\frac{7}{11}\right) = \frac{14}{33}. \quad \blacksquare$$

Solution to (b).

Let R_1 and R_2 be as defined in (a). Then,

$$\mathbb{P}(R_1) = \frac{8r}{8r + 4w}.$$

Given that the first ball is red, then the urn contains 7 red and 4 white balls. Then,

$$\mathbb{P}(R_2|R_1) = \frac{7r}{7r + 4w}.$$

Therefore, by the multiplication rule,

$$\mathbb{P}(R_1 \cap R_2) = \mathbb{P}(R_1)\,\mathbb{P}(R_2|R_1) = \left(\frac{8r}{8r + 4w}\right)\left(\frac{7r}{7r + 4r}\right). \qquad \blacksquare$$

General multiplication rule

Proposition 10 (The multiplication rule)

Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and let $E_1, E_2, \dots \in \mathscr{F}$ be a sequence of events. Then,

$$\mathbb{P}(E_1 \cap E_2 \cap \cdots \cap E_n) \\ = \mathbb{P}(E_1)\,\mathbb{P}(E_2|E_1)\,\mathbb{P}(E_3|E_1 \cap E_2)\dots\mathbb{P}(E_n|E_1 \cap E_2 \cap \cdots \cap E_{n-1}).$$

Proof.

The proof follows from a recursive argument. ■

Examples

Example 11 (Playing cards)

An ordinary deck of 52 playing cards is randomly divided into 4 piles of 13 cards each. Compute the probability that each pile has exactly 1 ace.

Solution.

Define the events

$$
\begin{aligned}
E_1 &= \{\text{the A}\spadesuit \text{ is in any one of the piles}\},\\
E_2 &= \{\text{the A}\heartsuit \text{ and A}\spadesuit \text{ are in different piles}\},\\
E_3 &= \{\text{the A}\heartsuit\text{, A}\diamondsuit \text{ and A}\spadesuit \text{ are in different piles}\},\\
E_4 &= \{\text{all of the aces are in different piles}\},
\end{aligned}
$$

then, by the multiplication rule, the desired probability is

$$\mathbb{P}(E_1 \cap E_2 \cap E_3 \cap E_4) = \mathbb{P}(E_1)\,\mathbb{P}(E_2|E_1)\,\mathbb{P}(E_3|E_1 \cap E_2)\,\mathbb{P}(E_4|E_1 \cap E_2 \cap E_3).$$

Now, $\mathbb{P}(E_1) = 1$. Also, as there are 39 slots in piles other than the pile containing A♠, while there are 51 slots overall, it follows that

$$\mathbb{P}(E_2|E_1) = \frac{39}{51}, \quad \mathbb{P}(E_3|E_1 \cap E_2) = \frac{26}{50}, \quad \text{and} \quad \mathbb{P}(E_4|E_1 \cap E_2 \cap E_3) = \frac{13}{49}.$$

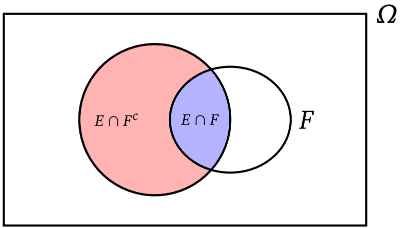
Therefore, $\mathbb{P}(E_1 \cap E_2 \cap E_3 \cap E_4) = \left(\frac{39}{51}\right)\left(\frac{26}{50}\right)\left(\frac{13}{49}\right) \approx 0.105$. ■

Total probability theorem

Introduction

Let E and F be events, and we may express E as

$$E = (E \cap F) \cup (E \cap F^c).$$



As $E \cap F$ and $E \cap F^c$ are mutually exclusive, then

$$\begin{aligned} \mathbb{P}(E) &= \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^c) \\ &= \mathbb{P}(E|F)\,\mathbb{P}(F) + \mathbb{P}(E|F^c)\,\mathbb{P}(F^c). \end{aligned}$$

Partition of Ω

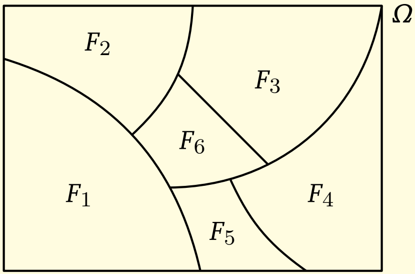
南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Definition 12 (Partition)

We say $F_1, F_2, \ldots$ form a partition of Ω if

(a) they are mutually exclusive, and

(b) $\bigcup_{i=1}^{\infty} F_i = \Omega$.

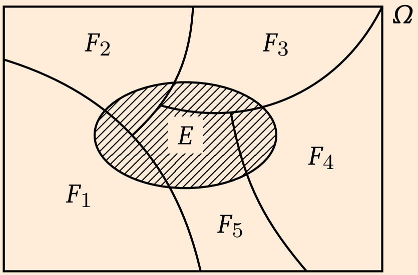


Total probability theorem

Theorem 13 (Law of total probability (Total probability theorem))

Let $F_1, F_2, \dots$ be mutually exclusive events that form a partition of the sample space: $\Omega = \bigcup_{i=1}^{\infty} F_i$. Assume that $\mathbb{P}(E|F_i) = 0$ if $\mathbb{P}(F_i) = 0$. Then,

$$\mathbb{P}(E) = \sum_{i=1}^{\infty} \mathbb{P}(F_i)\,\mathbb{P}(E|F_i).$$



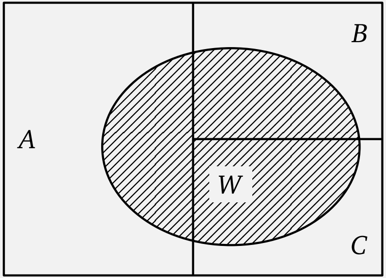
南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 14 (Tennnis tournament)

Suppose there are three types of players in a tennis tournament: A, B, and C. 50% of the contestants in the tournament are A players, 25% are B players, and 25% are C players. Your chance of beating the contestants depends on the class of the player, as follows: 0.3 against an A player 0.4 against a B player 0.5 against a C player. If you play a match in this tournament, what is the probability of your winning the match?

Solution.

Let A be the event that you will playing a match with an A class player, and let B and C be defined similarly. Then, $A \cup B \cup C = \Omega$ and they are also mutually exclusive. Let W be the event that you win the match.



By the law of total probability,

$$
\begin{aligned}
\mathbb{P}(W) &= \mathbb{P}(A)\,\mathbb{P}(W|A) + \mathbb{P}(B)\,\mathbb{P}(W|B) + \mathbb{P}(C)\,\mathbb{P}(W|C) \\
&= (0.3)(0.5) + (0.4)(0.25) + (0.5)(0.25) \\
&= 0.375.
\end{aligned}
$$

Question

Suppose that we have known that you have win the match, what is the conditional probability that you were playing with an A class player?

In general, if we know $\mathbb{P}(W|A)$, how can we reverse the conditioning to get the conditional probability $\mathbb{P}(A|W)$?

Example

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 15 (False positive or false negative)

Let Covid denote the event that a randomly chosen person actually having Covid-19. Let P denote the event of testing positive, and N that of testing negative. Suppose that we know that $\mathbb{P}(P|\text{Covid}^c) = 0.01$ and $\mathbb{P}(N|\text{Covid}) = 0.001$. Moreover, $\mathbb{P}(\text{Covid}) = 0.00005$.

We are now interested in the probability that a person had Covid given that he tested negative $\mathbb{P}(\text{Covid}|N)$?

Solution.

Note that

$$\mathbb{P}(N|\text{Covid}^c) = 1 - \mathbb{P}(P|\text{Covid}^c) = 1 - 0.01 = 0.99,$$

and

$$\mathbb{P}(\text{Covid}^c) = 1 - \mathbb{P}(\text{Covid}) = 1 - 0.00005 = 0.99995.$$

By the law of total probability,

$$\begin{aligned}\mathbb{P}(N) &= \mathbb{P}(N|\text{Covid})\,\mathbb{P}(\text{Covid}) + \mathbb{P}(N|\text{Covid}^c)\,\mathbb{P}(\text{Covid}^c)\\ &= (0.001)(0.00005) + (0.99)(0.99995) = 0.98995055.\end{aligned}$$

Then,

$$\mathbb{P}(\text{Covid}|N) = \frac{\mathbb{P}(N|\text{Covid})\,\mathbb{P}(\text{Covid})}{\mathbb{P}(N)} = \frac{(0.001)(0.00005)}{0.98995055} = 5 \times 10^{-8}. \qquad \blacksquare$$

Bayes' formula

Beyes' formula

Theorem 16

Let $E_1, E_2, \ldots$ be a partition of the sample sample. Then, for any event F such that $\mathbb{P}(F) > 0$, we have

$$\begin{aligned}\mathbb{P}(E_i|F) &= \frac{\mathbb{P}(E_i)\,\mathbb{P}(F|E_i)}{\mathbb{P}(F)} \\ &= \frac{\mathbb{P}(E_i|F)\,\mathbb{P}(F)}{\sum_{j=1}^{\infty}\mathbb{P}(E_j|F)\,\mathbb{P}(F)}.\end{aligned}$$

Remark

The result is a formula known as Bayes' Rule, after the Reverend Thomas Bayes (托马斯·贝叶斯神父, 1702?—1761).

Example 17 (Aircraft and radar, revisited)

What is the probability that the aircraft is present given that the radar generated an alarm?

Solution.

Recall that

$$\mathbb{P}(A) = 0.05, \quad \mathbb{P}(B|A) = 0.99, \quad \mathbb{P}(B|A^c) = 0.1.$$

Applying Bayes' formula, we obtain

$$\begin{aligned}
\mathbb{P}(A|B) &= \frac{\mathbb{P}(A)\,\mathbb{P}(B|A)}{\mathbb{P}(A)\,\mathbb{P}(B|A) + \mathbb{P}(A^c)\,\mathbb{P}(B|A^c)} \\
&= \frac{(0.05)(0.99)}{(0.05)(0.99) + (0.95)(0.1)} \approx 0.3426.
\end{aligned}$$

■

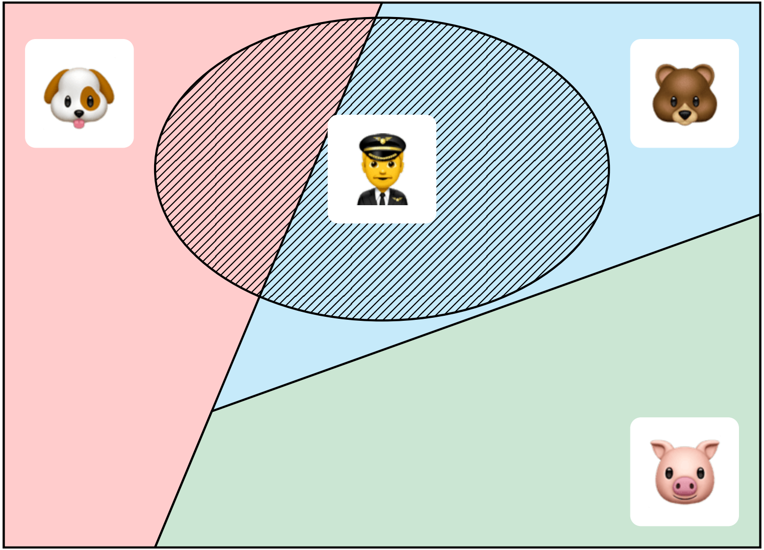
The Three Prisoners problem

Example 18

Once upon a time, there were three prisoners 🐶, 🐻, and 🐷. One day, 🤴 decided to pardon two of them and sentence the last one. One of the prisoners, 🐶, heard the news and wanted to ask a friendly 👮 about his situation. 👮 was honest, and he was allowed to tell 🐶 that 🐻 would be pardoned or that 🐷 would be pardoned, but he could not tell 🐶 whether he would be pardoned.

What is the conditional probability that 🐶 will be pardoned given that the 👮 tells him 🐻 will be pardoned?

从前，有三个囚犯 🐶，🐻，🐷。有一天，国王 🤴 决定赦免其中两个人并判处第三个人。其中一名囚犯 🐶，听到这个消息，想向一位 👮 询问他的情况。👮 是老实人，他可以说哪个会被赦免，但他不准说明谁不会被赦免。假设告诉他告诉了 🐶，囚犯 🐻 将要被赦免，则 🐶 被赦免的条件概率是多少？



Solution.

Let A, B, C denote the event that 🐶、🐻、🐷 will be sentenced, respectively, and let D denote that 👨‍✈️ tells 🐶 that 🐻 will be pardoned. Then,

$$\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{3},$$

and

$$\mathbb{P}(D|A) = \frac{1}{2}, \quad \mathbb{P}(D|B) = 0, \quad \mathbb{P}(D|C) = 1.$$

Therefore, by the Bayes' theorem,

$$\begin{aligned}\mathbb{P}(A|D) &= \frac{\mathbb{P}(D|A)\,\mathbb{P}(A)}{\mathbb{P}(D)} \\ &= \frac{(1/2)(1/3)}{(1/2)(1/3) + (0)(1/3) + (1)(1/3)} = \frac{1}{3}.\end{aligned}$$

Therefore, the conditional probability is $1/3$. ■

Independent events

Independence

- We've said informally that what we mean by independence is that the outcome of one event does not influence the probability of the other.
- With the help of conditional probabilities, we can understand the independence as

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

- No matter whether A happens, the probability of B does not change.
- However, $\mathbb{P}(B|A)$ implies $\mathbb{P}(A) > 0$, which is a strict condition.

Definition 19 (Independence)

Events A and B are said to be mutually independent if

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B).$$

Example 20

A card is selected at random from an ordinary deck of 52 playing cards. If E is the event that the selected card is an A and F is the event that it is a ♠, whether E and F are independent? Why or why not?

Solution.

The answer is "Yes". This follows because $\mathbb{P}(EF) = 1/52$, whereas $\mathbb{P}(E) = 4/52$ and $\mathbb{P}(F) = 13/52$. ∎

Properties of independence

SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Proposition 21

If E and F are independent, then so are E and F^c.

Proof.

Assume that E and F are independent. Since $E = (E \cap F) \cup (E \cap F^c)$ and $E \cap F$ and $E \cap F^c$ are obviously mutually exclusive, we have

$$\mathbb{P}(E) = \mathbb{P}(E \cap F) + \mathbb{P}(E \cap F^c)$$

or, equivalently,

$$\mathbb{P}(E \cap F^c) = \mathbb{P}(E) - \mathbb{P}(E \cap F) = \mathbb{P}(E)[1 - \mathbb{P}(F)] = \mathbb{P}(E)\,\mathbb{P}(F^c)$$

and the result is proved. ■

Example

Suppose now that E is independent of F and is also independent of G. Is E then necessarily independent of $F \cap G$? The answer, somewhat surprisingly, is no, as the following example demonstrates.

Example 22 (Counterexample)

Two fair dice are thrown. Let E denote the event that the sum of the dice is 7. Let F denote the event that the first die equals 4 and G denote the event that the second die equals 3.

(i) Whether E and F are independent?

(ii) Whether E and $F \cap G$ are independent?

Solution to Question (i).

The sample space is given by $\Omega = \{(i,j) : 1 \leqslant i,j \leqslant 6\}$. We have $E = \{(1,6),(2,5),\ldots,(6,1)\}$, then

$$\mathbb{P}(E) = \frac{|E|}{|\Omega|} = \frac{1}{6}.$$

Moreover, $F = \{(4,1),(4,2)\ldots,(4,6)\}$, then

$$\mathbb{P}(F) = \frac{|F|}{|\Omega|} = \frac{1}{6}.$$

On the other hand, $E \cap F = \{(4,2)\}$, then

$$\mathbb{P}(E \cap F) = \frac{1}{36}.$$

Then, it follows that $\mathbb{P}(E \cap F) = \mathbb{P}(E) \cdot \mathbb{P}(F)$ which implies that E and F are independent. ∎

General independence

Definition 23 (Independence for three events)

Three events E, F, and G are said to be independent if:

$$\begin{aligned}
\mathbb{P}(E \cap F \cap G) &= \mathbb{P}(E)\,\mathbb{P}(F)\,\mathbb{P}(G) \\
\mathbb{P}(E \cap F) &= \mathbb{P}(E)\,\mathbb{P}(F) \\
\mathbb{P}(E \cap G) &= \mathbb{P}(E)\,\mathbb{P}(G) \\
\mathbb{P}(F \cap G) &= \mathbb{P}(F)\,\mathbb{P}(G)
\end{aligned}$$

General independence

Definition 24

Of course, we may also extend the definition of independence to more than three events. The events E_1, E_2, ..., E_n are said to be independent if, for every subset $E'_1, E'_2, \ldots, E'_r, r \leqslant n$, of these events:

$$\mathbb{P}(E'_1 E'_2 \cdots E'_r) = \mathbb{P}(E'_1)\,\mathbb{P}(E'_2) \cdots \mathbb{P}(E'_r)$$

Finally, we define an infinite set of events to be independent if every finite subset of those events is independent.

Definition 25

- Sometimes, a probability experiment under consideration consists of performing a sequence of subexperiments .
- More formally, we say that the subexperiments are independent if $E_1, E_2, \ldots$ is necessarily an independent sequence of events whenever E_i is an event whose occurrence is completely determined by the outcome of the ith subexperiment.
- If each subexperiment has the same set of possible outcomes, then the subexperiments are often called trials .

南方科技大学 SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Example 26

An infinite sequence of independent trials is to be performed. Each trial results in a success with probability p and a failure with probability $1-p$. What is the probability that:

(a) at least 1 success occurs in the first n trials?

(b) exactly k successes occur in the first n trials?

$$\underbrace{\text{S}\quad \text{F}\quad \text{F}\quad \text{S}\quad \text{S}\quad \text{F}\quad \text{S}\quad \cdots\quad \text{F}\quad \text{F}}_{n \text{ trials including successes and failures}}$$

Solution to (a).

In order to determine the probability of at least 1 success in the first n trials, it is easiest to compute first the probability of the complementary event: that of no success in the first n trials.

$$\underbrace{\text{F} \quad \text{F} \quad \text{F} \quad \text{F} \quad \text{F} \quad \text{F} \quad \text{F} \quad \cdots \quad \text{F} \quad \text{F}}_{\text{all failures}}$$

Let E_i denote the event of a failure on the ith trial, then $\mathbb{P}(E_i) = 1 - p$ and the probability of no successes is

$$\mathbb{P}(E_1 \cap \cdots \cap E_n) = \prod_{i=1}^{n} \mathbb{P}(E_i) = (1 - p)^n.$$

Therefore, the probability of at least one success is

$$\mathbb{P}(E_1^c \cup \cdots \cup E_n^c) = 1 - (1 - p)^n. \qquad \blacksquare$$

Solution to (b).

Consider any particular sequence of the first n outcomes containing k successes and $n-k$ failures.

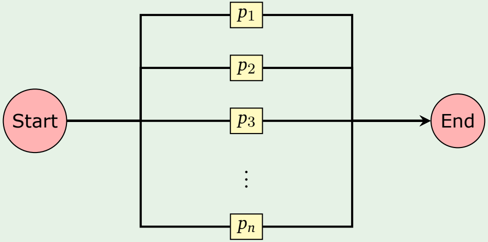
$$\underbrace{\text{S}\quad \text{F}\quad \text{S}\quad \text{F}\quad \text{S}\quad \text{S}\quad \text{F}\quad \cdots\quad \text{F}\quad \text{S}}_{n \text{ trials including } k \text{ successes and } n-k \text{ failures}}$$

Each particular sequence of these events occur with probability $p^k(1-p)^k$.
Since there are $\binom{n}{k}$ such sequences, the desired probability in (b) is

$$\mathbb{P}\{\text{exact } k \text{ successes}\} = \binom{n}{k} p^k (1-p)^k. \qquad \blacksquare$$

Example 27

A system composed of n separate components is said to be a parallel system if it functions when at least one of the components functions. For such a system, if component i, which is independent of the other components, functions with probability $p_i, i = 1, \ldots, n$, what is the probability that the system functions?



Solution.

Let A_i denote the event that component i functions. Then,

$$\begin{aligned}
\mathbb{P}\{\text{system functions}\} &= \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \\
&= 1 - \mathbb{P}\left(\bigcap_{i=1}^{n} A_i^c\right) \\
&= 1 - \prod_{i=1}^{n} \mathbb{P}(A_i^c) \\
&= 1 - \prod_{i=1}^{n} (1 - p_i).
\end{aligned}$$

■

Example 28

Independent trials consisting of rolling a pair of fair dice are performed. What is the probability that an outcome of 5 appears before an outcome of 7 when the outcome of a roll is the sum of the dice?

Solution.

Let E_n denote the event that no 5 or 7 appears on the first $n-1$ trials, and a 5 appears on the nth trial, then the desired probability is

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} E_n\right) = \sum_{n=1}^{\infty} \mathbb{P}(E_n).$$

Since

$$\mathbb{P}\{\text{5 on any trial}\} = \frac{1}{9}, \quad \mathbb{P}\{\text{7 on any trial}\} = \frac{1}{6},$$

we obtain $\mathbb{P}(E_n) = \frac{1}{9}\left(1 - \frac{1}{9} - \frac{1}{6}\right)^{n-1} = \frac{1}{9}\left(\frac{13}{18}\right)^{n-1}$. Thus,

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} E_n\right) = \frac{1}{9}\sum_{n=1}^{\infty}\left(\frac{13}{18}\right)^{n-1} = \frac{2}{5}. \qquad \blacksquare$$

Disjoint events and independent events

- Disjoint events and independent events are two different concepts in probability theory.
- Disjoint events are events that have no outcomes in common, meaning that they cannot occur simultaneously.
- On the other hand, independent events are events where the occurrence of one event does not affect the probability of the other event occurring.

Disjoint events and independent events

Example 29

Suppose we are rolling a fair six-sided die. Let A be the event that we get an even number (⚁, ⚃, or ⚅), and let B be the event that we get a number greater than 4 (⚄ or ⚅).

A and B are not disjoint because there is one outcome (⚅) that satisfies both events. However, A and B are independent because the probability of getting an even number is 1/2 and the probability of getting a number greater than 4 is 1/3, but the probability of getting an even number and a number greater than 4 is $(1/2)(1/3) = 1/6$, which is the probability of getting a 6.

Summary of Conditional Probability

- Conditional probability measures the probability of an event given that another event has occurred.
- The conditional probability of event A given event B is denoted as $P(A|B)$ and is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \text{if } P(B) \neq 0$$

- The multiplication rule states that for any two events A and B, the probability of the intersection of A and B can be calculated as:

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) = \mathbb{P}(B|A) \cdot \mathbb{P}(A)$$

Summary of Conditional Probability

- The Law of Total Probability states that for any partition $B_1, B_2, \ldots, B_n$ of the sample space, the probability of an event A can be calculated as:

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

- Two events A and B are independent if and only if $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$.

- Bayes' Theorem allows us to calculate the conditional probability of one event given another event in terms of the reverse conditional probability:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A) \cdot \mathbb{P}(A)}{\mathbb{P}(B)}$$

Further reading

[1] Sheldon M. Ross (谢尔登·M. 罗斯).

A first course in probability (概率论基础教程): Chapter 3.

10th edition (原书第十版), 机械工业出版社

[2] Sheldon M. Ross (谢尔登·M. 罗斯).

Introduction to Probability Models (概率模型导论): Chapter 1.

12th edition (原书第十二版), 人民邮电出版社