Lesson 1.3: Conditional Probability, Multiplication Rule of Probability, and Bayes Theorem

NE Milla, Jr.

2025-01-17

Learning Outcomes

At the end of the lesson, students must be able to

1. Explain the idea of conditional probability,
2. Compute conditional probability,
3. Apply the multiplication rule of probability,
4. Derive and apply the Bayes Theorem.

Introduction

Take two events A and B. For example, let A be the event "Receive a score of 95% on the econometrics exam" and let B be the event "Study econometrics 12 hours a day". We might be interested in the question: *Does B affect the likelihood of A?*

Alternatively, we may be interested in questions such as: *Does attending college affect the likelihood of obtaining a high wage?* Or: *Do tariffs affect the likelihood of price increases?*

These are questions of conditional probability. Abstractly, consider two events A and B. Suppose that we know that B has occurred. Then the only way for A to occur is if the outcome is in the intersection $A \cap B$ So we are asking: "What is the probability that $A \cap B$ occurs, given that B occurs?" The answer is not simply $P(A \cap B)$. Instead we can think of the *new* or *reduced* sample space as B. To do so we normalize all probabilities by $P(B)$. We arrive at the following definition.

Conditional probability

Definition: If $P(B) > 0$, the conditional probability of A given B is

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

The notation $A|B$ is read as "A given B" or "A assuming that B is true". To add clarity we will sometimes refer to $P(A)$ as the *unconditional probability* to distinguish it from $P(A|B)$.

EXAMPLE 1:

Consider the random experiment of rolling a die. What is the probability of recording an even number of dots if it is known in advance that at least 4 dots were recorded?

SOLUTION:

Let A = event of recording at least 4 dots and let B = event of recording even number of dots. Then

$$\begin{aligned} A &= \{4,5,6\} \\ B &= \{2,4,6\}, \text{ and} \\ A \cap B &= \{4,6\} \end{aligned}$$

Hence,

$$\begin{aligned} P(A) &= \frac{3}{6}, \\ P(B) &= \frac{3}{6}, \text{ and} \\ P(A \cap B) &= \frac{2}{6} \end{aligned}$$

Therefore, the probability of recording an even number of dots if it is known in advance that at least 4 dots were recorded is

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{2}{6}}{\frac{3}{6}} = \frac{2}{3}$$

REMARKS:

1. The *unconditional* probability of obtaining an even number of dots is $P(B) = \frac{3}{6}$. This shows that having an additional (prior) information that at least 4 dots are obtained, the probability of an even number of dots increased from 0.5 to 0.667. This is generally true.
2. The *conditioning* event can be viewed as the **reduced** or **restricted** sample space (S^*). That is $S^* = A$. Then, based on this reduced sample space, the probability of recording an even number of dots is $P(B) = \frac{2}{3}$.

EXAMPLE 2:

What is the probability of drawing an odd-numbered card from a deck of playing cards if it is known that the card is red?

SOLUTION:

Let A = event of drawing an odd-numbered card and let B = event of getting a red card. Then

$$\begin{aligned} P(A) &= \frac{16}{52}, \\ P(B) &= \frac{26}{52}, \text{ and} \\ P(A \cap B) &= \frac{8}{52} \end{aligned}$$

Therefore, the probability of drawing an odd-numbered card from a deck of playing cards if it is known that the card is red is given by

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{8}{52}}{\frac{26}{52}} = \frac{8}{26}$$

Alternatively, using the concept of *reduced* sample space, we just focus on the 26 red cards and out of these there are 8 odd-numbered cards. Hence, we get the same answer $\frac{8}{26}$.

REMARKS:

1. Conditional probability satisfies the 3 axioms of probability. That is,

    a. $P(A|B) \geq 0$
    b. $P(A|A) = 1$
    c. Given pairwise mutually exclusive events $A_1, A_2, \cdots,$, then

    $$P\left(\bigcup_{i=1}^{\infty} A_i|B\right) = \sum_{i=1}^{\infty} P(A_i|B)$$

2. Since conditional probability satisfies the three fundamental axioms, all the probability rules we derived earlier have their respective "conditional versions." For example,

    a. $P(A^c|B) = 1 - P(A|B)$ [**complement rule**]
    b. If $A_1 \subset A_2$, then for any set B, $P(A_1|B) \leq P(A_2|B)$ [**monotonicity rule**]
    c. For any given sets $A_1, A_2, \text{and} B$,

    $$P[(A_1 \cup A_2)|B] = P(A_1|B) + P(A_2|B) - P[(A_1 \cap A_2)|B]$$

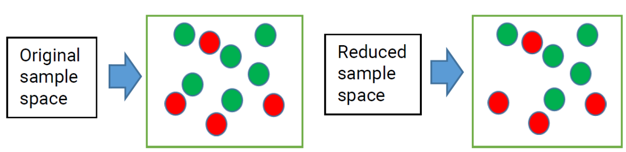
    [**addition rule**]

EXAMPLE 3:

A box contains 6 green balls and 4 red balls. We randomly (and without replacement) draw two balls from the box. What is the probability that the second ball selected is red, given that the first ball selected is green?

SOLUTION:

Consider the following box. The process involves selecting a green ball in the first draw followed by a red ball. Once a green ball has been drawn, we are left with 9 balls (5 green and 4 red).



So that the probability that the second ball selected is red, given that the first ball selected is green is

$$P(R_2|G_1) = \frac{4}{9}$$

The subscripts 1 and 2 indicate the order of the draw (1st and 2nd draws).

EXAMPLE 4:

Refer to the previous example, what is the probability that the balls selected are both red?

SOLUTION:

We are interested in computing $P(R_1 \cap R_2)$. We can compute this probability using two methods. In method 1, we can use the combination rule as follows:

$$P(R_1 \cap R_2) = \frac{{}_4\mathrm{C}_2 \times {}_6\mathrm{C}_0}{{}_{10}\mathrm{C}_2} = \frac{6}{45} = \frac{2}{15}$$

In the second method, we use the concept of conditional probability.

$$P(R_1 \cap R_2) = P(R_1) \times P(R_2|R_1) = \frac{4}{10} \times \frac{3}{9} = \frac{2}{15}$$

Multiplication Rule of Probability

Suppose A and B are events in the sample space S. Then

$$P(A \cap B) = P(A) \times P(B|A)$$

Or, equivalently,

$$P(A \cap B) = P(B) \times P(A|B)$$

This "rule" follows directly from the definition of conditional probability. All we have to do is apply "cross" multiplication to the conditional probability formula.

The multiplication rule allows us to approach calculating the probability of an intersection "sequentially." First, calculate $P(A_1)$ for the first event. Next, calculate $P(A_2|A_1)$ for the second event (given the first). Next, calculate $P(A_3|A_1 \cap A_2)$ for the third event (given the first two), and so on.

EXAMPLE 5:

You are dealt a hand of 5 cards at random. What is the probability they are all spades?

SOLUTION:

Define the events $A_i, i = 1, 2, 3, 4, 5$. Assuming the cards are randomly drawn from the deck, then

$$\begin{aligned} P(A_1) &= \frac{13}{52} \\ P(A_2|A_1) &= \frac{12}{51} \\ P(A_3|A_1 \cap A_2) &= \frac{11}{50} \\ P(A_4|A_1 \cap A_2 \cap A_3) &= \frac{10}{49} \\ P(A_5|A_1 \cap A_2 \cap A_3 \cap A_4) &= \frac{9}{48} \end{aligned}$$

Therefore, the probability all five cards are spades is

$$\begin{aligned} P\left(\bigcap_{i=1}^{5} A_i\right) &= P(A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) \\ &= P(A_1) \times P(A_2|A_1) \times P(A_3|A_1 \cap A_2) \times P(A_4|A_1 \cap A_2 \cap A_3) \times P(A_5|A_1 \cap A_2 \cap A_3 \cap A_4) \\ &= \frac{13}{52} \times \frac{12}{51} \times \frac{11}{50} \times \frac{10}{49} \times \frac{9}{48} \\ &\approx 0.0005 \end{aligned}$$

Bayes Theorem

Now we are ready to state one of the most useful results in conditional probability: *Bayes' Theorem*. Suppose that we know $P(A|B)$, but we are interested in the probability $P(B|A)$. Using the definition of conditional probability, we have

$$P(B|A)P(A) = P(A|B)P(B)$$

If we divide both sides of the equation by $P(A)$ we obtain

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

which is the famous Bayes' formula. In this formulation, we call $P(B|A)$ as the *posterior* probability and $P(B)$ as the *prior* probability.

Often, in order to find $P(A)$ in Bayes' formula we need to use the law of total probability. We state formally the Bayes' theorem as follows.

Suppose A is an event in S and suppose that $B_1, B_2, B_3, \cdots, B_k$ forms a partition of S. Then

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}$$

EXAMPLE 6:

Three persons A, B and C have applied for a job in a private company. The chance of their selections is in the ratio 1:2:4. The probabilities that A, B and C can introduce changes to improve the profits of the company are 0.8, 0.5 and 0.3, respectively. If the change does not take place, find the probability that it is due to the appointment of C.

SOLUTION:

Let A, B, and C be the event that applicants A, B, and C are appointed or hired, respectively. Further, let D be the event that changes happen to improve profits of the company.

Then, $P(A) = 1/7, P(B) = 2/7, P(C) = 4/7$. Also, $P(D|A) = 0.8, P(D|B) = 0.5, P(D|C) = 0.3$.

From the given probabilities, we have

$$\begin{aligned}
P(D^c|A) &= 1 - P(D|A) = 1 - 0.8 = 0.2 \\
P(D^c|B) &= 1 - P(D|B) = 1 - 0.5 = 0.5 \\
P(D^c|C) &= 1 - P(D|C) = 1 - 0.3 = 0.7
\end{aligned}$$

We calculate $P(C|D^c)$ using the Bayes' Theorem as follows

$$\begin{aligned}
P(C|D^c) &= \frac{P(D^c|C)P(C)}{P(D^c|A)P(A) + P(D^c|B)P(B) + P(D^c|C)P(C)} \\
&= \frac{0.7(4/7)}{0.2(1/7) + 0.5(2/7) + 0.7(4/7)} \\
&= 0.7
\end{aligned}$$

Additional Examples

1. A fair coin is tossed three times. Find the probability of getting three heads, given that there are at least two heads.

2. Three balls are picked at random one by one and without replacement from a box containing four white and eight black balls. Let A = {The first ball is white}, B = {the second ball is white}, and C = {The third ball is white}. Find:

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

3. A fair die is tossed twice. Find the probability that

   a. the sum is 9, given that the sum is greater than 6
   b. the sum is 7, given that the sum is odd

4. Tom and Mary agree to meet at some place at a certain time. The probability that Tom keeps his appointment is 0.7, and that either of them keeps the appointment is 0.9. Given that Tom did not keep the appointment, what is the probability that Mary didn't either.

5. A insurance company has insured 4000 doctors, 8000 teachers and 12000 businessmen. The chances of a doctor, teacher and businessman dying before the age of 58 is 0.01, 0.03 and 0.05, respectively. If one of the insured people dies before 58, find the probability that he is a doctor.