

4.3: Conditional Probability

Suppose you want to figure out if you should buy a new car. When you first go and look, you find two cars that you like the most. In your mind they are equal, and so each has a 50% chance that you will pick it. Then you start to look at the reviews of the cars and realize that the first car has had 40% of them needing to be repaired in the first year, while the second car only has 10% of the cars needing to be repaired in the first year. You could use this information to help you decide which car you want to actually purchase. Both cars no longer have a 50% chance of being the car you choose. You could actually calculate the probability you will buy each car, which is a conditional probability. You probably wouldn't do this, but it gives you an example of what a conditional probability is.

Conditional probabilities are probabilities calculated after information is given. This is where you want to find the probability of event A happening after you know that event B has happened. If you know that B has happened, then you don't need to consider the rest of the sample space. You only need the outcomes that make up event B. Event B becomes the new sample space, which is called the **restricted sample space**, R. If you always write a restricted sample space when doing conditional probabilities and use this as your sample space, you will have no trouble with conditional probabilities. The notation for conditional probabilities is $P(A, \text{ given } B) = P(A|B)$. The event following the vertical line is always the restricted sample space.

Example 4.3.1 conditional probabilities

- Suppose you roll two dice. What is the probability of getting a sum of 5, given that the first die is a 2?
- Suppose you roll two dice. What is the probability of getting a sum of 7, given the first die is a 4?
- Suppose you roll two dice. What is the probability of getting the second die a 2, given the sum is a 9?
- Suppose you pick a card from a deck. What is the probability of getting a Spade, given that the card is a Jack?
- Suppose you pick a card from a deck. What is the probability of getting an Ace, given the card is a Queen?

Solution

- a. Since you know that the first die is a 2, then this is your restricted sample space, so

$$R = \{(2,1), (2,2), (2,3), (2,4), (2,5), (2,6)\}$$

Out of this restricted sample space, the way to get a sum of 5 is $\{(2,3)\}$. Thus

$$P(\text{sum of 5} | \text{the first die is a 2}) = \frac{1}{6}$$

- b. Since you know that the first die is a 4, this is your restricted sample space, so

$$R = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$$

Out of this restricted sample space, the way to get a sum of 7 is $\{(4,3)\}$. Thus

$$P(\text{sum of 7} | \text{the first die is a 4}) = \frac{1}{6}$$

- c. Since you know the sum is a 9, this is your restricted sample space, so

$$R = \{(3,6), (4,5), (5,4), (6,3)\}$$

Out of this restricted sample space there is no way to get the second die a 2. Thus

$$P(\text{second die is a 2} | \text{sum is 9}) = 0$$

- d. Since you know that the card is a Jack, this is your restricted sample space, so

$$R = \{JS, JC, JD, JH\}$$

Out of this restricted sample space, the way to get a Spade is $\{JS\}$. Thus

$$P(\text{Spade} | \text{Jack}) = \frac{1}{4}$$

- e. on: Since you know that the card is a Queen, then this is your restricted sample space, so

$$R = \{QS, QC, QD, QH\}$$

Out of this restricted sample space, there is no way to get an Ace, thus

$$P(\text{Ace} | \text{Queen}) = 0$$

If you look at the results of Example 4.3.7 part d and Example 4.3.1 part b, you will notice that you get the same answer. This means that knowing that the first die is a 4 did not change the probability that the sum is a 7. This added knowledge did not help you in any way. It is as if that information was not given at all. However, if you compare Example 4.3.7 part b and Example 4.3.1 part a, you will notice that they are not the same answer. In this case, knowing that the first die is a 2 did change the probability of getting a sum of 5. In the first case, the events sum of 7 and first die is a 4 are called **independent events**. In the second case, the events sum of 5 and first die is a 2 are called **dependent events**.

Events A and B are considered **independent events** if the fact that one event happens does not change the probability of the other event happening. In other words, events A and B are independent if the fact that B has happened does not affect the probability of event A happening and the fact that A has happened does not affect the probability of event B happening. Otherwise, the two events are dependent. In symbols, A and B are independent if

$$P(A|B) = P(A) \text{ or } P(B|A) = P(B)$$

Example 4.3.2 independent events

- Suppose you roll two dice. Are the events “sum of 7” and “first die is a 3” independent?
- Suppose you roll two dice. Are the events “sum of 6” and “first die is a 4” independent?
- Suppose you pick a card from a deck. Are the events “Jack” and “Spade” independent?
- Suppose you pick a card from a deck. Are the events “Heart” and “Red” card independent?
- Suppose you have two children via separate births. Are the events “the first is a boy” and “the second is a girl” independent?
- Suppose you flip a coin 50 times and get a head every time, what is the probability of getting a head on the next flip?

Solution

a. To determine if they are independent, you need to see if $P(A|B) = P(A)$. It doesn't matter which event is A or B, so just assign one as A and one as B.

Let A = sum of 7 = {(1,6), (2,5), (3,4), (4,3), (5,2), (6,1)} and B = first die is a 3 = {(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)}
 $P(A|B)$ means that you assume that B has happened. The restricted sample space is B,

$$R = \{(3,1), (3,2), (3,3), (3,4), (3,5), (3,6)\}$$

In this restricted sample space, the way for A to happen is {(3,4)}, so

$$P(A|B) = \frac{1}{6}$$

$$\text{The } P(A) = \frac{6}{36} = \frac{1}{6}$$

$P(A|B) = P(A)$ Thus “sum of 7” and “first die is a 3” are independent events.

b. To determine if they are independent, you need to see if $P(A|B) = P(A)$. It doesn't matter which event is A or B, so just assign one as A and one as B.

Let A = sum of 6 = {(1,5), (2,4), (3,3), (4,2), (5,1)} and B = first die is a 4 = {(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)}, so

$$P(A) = \frac{5}{36}$$

For $P(A|B)$, the restricted sample space is B,

$$R = \{(4,1), (4,2), (4,3), (4,4), (4,5), (4,6)\}$$

In this restricted sample space, the way for A to happen is {(4,2)}, so

$$P(A|B) = \frac{1}{6}$$

In this case, “sum of 6” and “first die is a 4” are dependent since $P(A|B) \neq P(A)$.

c. To determine if they are independent, you need to see if $P(A|B) = P(A)$. It doesn't matter which event is A or B, so just assign one as A and one as B.

Let A = Jack = {JS, JC, JD, JH} and B = Spade {2S, 3S, 4S, 5S, 6S, 7S, 8S, 9S, 10S, JS, QS, KS, AS}

$$P(A) = \frac{4}{52} = \frac{1}{13}$$

For $P(A|B)$, the restricted sample space is B,

$$R = \{2S, 3S, 4S, 5S, 6S, 7S, 8S, 9S, 10S, JS, QS, KS, AS\}$$

In this restricted sample space, the way A happens is {JS}, so

$$P(A|B) = \frac{1}{13}$$

In this case, "Jack" and "Spade" are independent since $P(A|B) = P(A)$.

d. To determine if they are independent, you need to see if $P(A|B) = P(A)$. It doesn't matter which event is A or B, so just assign one as A and one as B.

Let A = Heart = {2H, 3H, 4H, 5H, 6H, 7H, 8H, 9H, 10H, JH, QH, KH, AH} and B = Red card = {2D, 3D, 4D, 5D, 6D, 7D, 8D, 9D, 10D, JD, QD, KD, AD, 2H, 3H, 4H, 5H, 6H, 7H, 8H, 9H, 10H, JH, QH, KH, AH}, so

$$P(A) = \frac{13}{52} = \frac{1}{4}$$

For $P(A|B)$, the restricted sample space is B,

$$R = \{2D, 3D, 4D, 5D, 6D, 7D, 8D, 9D, 10D, JD, QD, KD, AD, 2H, 3H, 4H, 5H, 6H, 7H, 8H, 9H, 10H, JH, QH, KH, AH\}$$

In this restricted sample space, the way A can happen is 13,

$$P(A|B) = \frac{13}{26} = \frac{1}{2}$$

In this case, "Heart" and "Red" card are dependent, since $P(A|B) \neq P(A)$.

e. In this case, you actually don't need to do any calculations. The gender of one child does not affect the gender of the second child, the events are independent.

f. Since one flip of the coin does not affect the next flip (the coin does not remember what it did the time before), the probability of getting a head on the next flip is still one-half.

Multiplication Rule:

Two more useful formulas: If two events are dependent, then $P(A \text{ and } B) = P(A) * P(B|A)$

If two events are independent, then $P(A \text{ and } B) = P(A) * P(B)$

If you solve the first equation for $P(B|A)$, you obtain $P(B|A) = \frac{P(A \text{ and } B)}{P(A)}$, which is a formula to calculate a conditional probability. However, it is easier to find a conditional probability by using the restricted sample space and counting unless the sample space is large.

Example 4.3.3 Multiplication rule

- Suppose you pick three cards from a deck, what is the probability that they are all Queens if the cards are not replaced after they are picked?
- Suppose you pick three cards from a deck, what is the probability that they are all Queens if the cards are replaced after they are picked and before the next card is picked?

Solution

a. This sample space is too large to write out, so using the multiplication rule makes sense. Since the cards are not replaced, then the probability will change for the second and third cards. They are dependent events. This means that on the second draw there is one less Queen and one less card, and on the third draw there are two less Queens and 2 less cards.

$$\begin{aligned}
 P(3 \text{ Queens}) &= P(Q \text{ on 1st and } Q \text{ on 2nd and } Q \text{ on 3rd}) \\
 &= P(Q \text{ on 1st}) * P(Q \text{ on 2nd} | Q \text{ on 1st}) * P(Q \text{ on 3rd} | 1st \text{ and } 2nd \text{ } Q) \\
 &= \frac{4}{52} * \frac{3}{51} * \frac{2}{50} \\
 &= \frac{24}{132600}
 \end{aligned}$$

b. Again, the sample space is too large to write out, so using the multiplication rule makes sense. Since the cards are put back, one draw has no affect on the next draw and they are all independent.

$$\begin{aligned}
 P(3 \text{ Queens}) &= P(\text{Queen on 1st and Queen on 2nd and Queen on 3rd}) \\
 &= P(\text{Queen on 1st}) * P(\text{Queen on 2nd}) * P(\text{Queen on 3rd}) \\
 &= \frac{4}{52} * \frac{4}{52} * \frac{4}{52} \\
 &= \left(\frac{4}{52}\right)^3 \\
 &= \frac{64}{140608}
 \end{aligned}$$

Example 4.3.4 application problem

The World Health Organization (WHO) keeps track of how many incidents of leprosy there are in the world. Using the WHO regions and the World Banks income groups, one can ask if an income level and a WHO region are dependent on each other in terms of predicting where the disease is. Data on leprosy cases in different countries was collected for the year 2011 and a summary is presented in Example 4.3.1 ("Leprosy: Number of," 2013).

Table 4.3.1: Number of Leprosy Cases

WHO Region	World Bank Income Group				Row Total
	High Income	Upper Middle Income	Lower Middle Income	Low Income	
Americas	174	36028	615	0	36817
Eastern Mediterranean	54	6	1883	604	2547
Europe	10	0	0	0	10
Western Pacific	26	216	3689	1155	5086
Africa	0	39	1986	15928	17953
South-East Asia	0	0	149896	10236	160132
Column Total	264	36289	158069	27923	222545

- Find the probability that a person with leprosy is from the Americas.
- Find the probability that a person with leprosy is from a high-income country.
- Find the probability that a person with leprosy is from the Americas and a high-income country.
- Find the probability that a person with leprosy is from a high-income country, given they are from the Americas.
- Find the probability that a person with leprosy is from a low-income country.
- Find the probability that a person with leprosy is from Africa.
- Find the probability that a person with leprosy is from Africa and a low-income country.
- Find the probability that a person with leprosy is from Africa, given they are from a low-income country.
- Are the events that a person with leprosy is from "Africa" and "low-income country" independent events? Why or why not?

- j. Are the events that a person with leprosy is from “Americas” and “high-income country” independent events? Why or why not?

Solution

- a. There are 36817 cases of leprosy in the Americas out of 222,545 cases worldwide. So,

$$P(\text{Americas}) = \frac{36817}{222545} \approx 0.165$$

There is about a 16.5% chance that a person with leprosy lives in a country in the Americas.

- b. There are 264 cases of leprosy in high-income countries out of 222,545 cases worldwide. So,

$$P(\text{high-income}) = \frac{264}{222545} \approx 0.0001$$

There is about a 0.1% chance that a person with leprosy lives in a high-income country.

- c. There are 174 cases of leprosy in countries in a high-income country in the Americas out the 222,545 cases worldwide. So,

$$P(\text{Americas and high-income}) = \frac{174}{222545} \approx 0.0008$$

There is about a 0.08% chance that a person with leprosy lives in a high-income country in the Americas.

- d. In this case you know that the person is in the Americas. You don't need to consider people from Eastern Mediterranean, Europe, Western Pacific, Africa, and South-east Asia. You only need to look at the row with Americas at the start. In that row, look to see how many leprosy cases there are from a high-income country. There are 174 countries out of the 36,817 leprosy cases in the Americas. So,

$$P(\text{high-income} | \text{Americas}) = \frac{174}{36817} \approx 0.0047$$

There is 0.47% chance that a person with leprosy is from a high-income country given that they are from the Americas.

- e. There are 27,923 cases of leprosy in low-income countries out of the 222,545 leprosy cases worldwide. So,

$$P(\text{low-income}) = \frac{27923}{222545} \approx 0.125$$

There is a 12.5% chance that a person with leprosy is from a low-income country.

- f. There are 17,953 cases of leprosy in Africa out of 222,545 leprosy cases worldwide. So,

$$P(\text{Africa}) = \frac{17953}{222545} \approx 0.081$$

There is an 8.1% chance that a person with leprosy is from Africa.

- g. There are 15,928 cases of leprosy in low-income countries in Africa out of all the 222,545 leprosy cases worldwide. So,

$$P(\text{Africa and low-income}) = \frac{15928}{222545} \approx 0.072$$

There is a 7.2% chance that a person with leprosy is from a low-income country in Africa.

- h. In this case you know that the person with leprosy is from low-income country. You don't need to include the high income, upper-middle income, and lowermiddle income country. You only need to consider the column headed by lowincome. In that column, there are 15,928 cases of leprosy in Africa out of the 27,923 cases of leprosy in low-income countries. So,

$$P(\text{Africa} | \text{low-income}) = \frac{15928}{27923} \approx 0.570$$

There is a 57.0% chance that a person with leprosy is from Africa, given that they are from a low-income country.

- i. In order for these events to be independent, either $P(\text{Africa} | \text{low-income}) = P(\text{Africa})$ or $P(\text{low-income} | \text{Africa}) = P(\text{low-income})$ have to be true. Part (h) showed $P(\text{Africa} | \text{low-income}) \approx 0.570$ and part (f) showed $P(\text{Africa}) \approx 0.081$. Since these are not equal, then these two events are dependent.

- j. In order for these events to be independent, either $P(\text{Americas} | \text{high-income}) = P(\text{Americas})$ or $P(\text{high-income} | \text{Americas}) = P(\text{high-income})$ have to be true. Part (d) showed

$P(\text{high-income} \mid \text{Americas}) \approx 0.0047$ and part (b) showed $P(\text{high-income}) \approx 0.001$. Since these are not equal, then these two events are dependent.

A big deal has been made about the difference between dependent and independent events while calculating the probability of *and* compound events. You must multiply the probability of the first event with the conditional probability of the second event.

Why do you care? You need to calculate probabilities when you are performing sampling, as you will learn later. But here is a simplification that can make the calculations a lot easier: when the sample size is very small compared to the population size, you can assume that the conditional probabilities just don't change very much over the sample.

For example, consider acceptance sampling. Suppose there is a big population of parts delivered to you factory, say 12,000 parts. Suppose there are 85 defective parts in the population. You decide to randomly select ten parts, and reject the shipment. What is the probability of rejecting the shipment?

There are many different ways you could reject the shipment. For example, maybe the first three parts are good, one is bad, and the rest are good. Or all ten parts could be bad, or maybe the first five. So many ways to reject! But there is only **one** way that you'd accept the shipment: if **all ten** parts are good. That would happen if the first part is good, **and** the second part is good, **and** the third part is good, and so on. Since the probability of the second part being good is (slightly) dependent on whether the first part was good, technically you should take this into consideration when you calculate the probability that all ten are good.

The probability of getting the first sampled part good is $\frac{12000 - 85}{12000} = \frac{11915}{12000}$. So the probability that all ten being good is $\frac{11915}{12000} * \frac{11914}{11999} * \frac{11913}{11998} * \dots * \frac{11906}{11991} \approx 93.1357\%$. If instead you assume that the probability doesn't change much, you get $\left(\frac{11915}{12000}\right)^{10} \approx 93.1382\%$. So as you can see, there is not much difference. So here is the rule: if the sample is very small compared to the size of the population, then you can assume that the probabilities are independent, even though they aren't technically. By the way, the probability of rejecting the shipment is $1 - 0.9314 = 0.0686 = 6.86\%$

Homework

Exercise 4.3.1

- Are owning a refrigerator and owning a car independent events? Why or why not?
- Are owning a computer or tablet and paying for Internet service independent events? Why or why not?
- Are passing your statistics class and passing your biology class independent events? Why or why not?
- Are owning a bike and owning a car independent events? Why or why not?
- An experiment is picking a card from a fair deck.
 - What is the probability of picking a Jack given that the card is a face card?
 - What is the probability of picking a heart given that the card is a three?
 - What is the probability of picking a red card given that the card is an ace?
 - Are the events Jack and face card independent events? Why or why not?
 - Are the events red card and ace independent events? Why or why not?
- An experiment is rolling two dice.
 - What is the probability that the sum is 6 given that the first die is a 5?
 - What is the probability that the first die is a 3 given that the sum is 11?
 - What is the probability that the sum is 7 given that the first die is a 2?
 - Are the two events sum of 6 and first die is a 5 independent events? Why or why not?
 - Are the two events sum of 7 and first die is a 2 independent events? Why or why not?
- You flip a coin four times. What is the probability that all four of them are heads?
- You flip a coin six times. What is the probability that all six of them are heads?
- You pick three cards from a deck with replacing the card each time before picking the next card. What is the probability that all three cards are kings?
- You pick three cards from a deck without replacing a card before picking the next card. What is the probability that all three cards are kings?

11. The number of people who survived the Titanic based on class and sex is in Example 4.3.2 ("Encyclopedia Titanica," 2013). Suppose a person is picked at random from the survivors.

Class	Sex		Total
	Female	Male	
1st	134	59	193
2nd	94	25	119
3rd	80	58	138
Total	308	142	450

Table 4.3.2: *Surviving the Titanic*

- What is the probability that a survivor was female?
 - What is the probability that a survivor was in the 1st class?
 - What is the probability that a survivor was a female given that the person was in 1st class?
 - What is the probability that a survivor was a female and in the 1st class?
 - What is the probability that a survivor was a female or in the 1st class?
 - Are the events survivor is a female and survivor is in 1st class mutually exclusive? Why or why not?
 - Are the events survivor is a female and survivor is in 1st class independent? Why or why not?
12. Researchers watched groups of dolphins off the coast of Ireland in 1998 to determine what activities the dolphins partake in at certain times of the day ("Activities of dolphin," 2013). The numbers in Example 4.3.3 represent the number of groups of dolphins that were partaking in an activity at certain times of days.

Activity	Period				Total
	Morning	Noon	Afternoon	Evening	
Travel	6	6	14	13	39
Feed	28	4	0	56	88
Social	38	5	9	10	62
Total	72	15	23	79	189

Table 4.3.3: *Dolphin Activity*

- What is the probability that a dolphin group is partaking in travel?
- What is the probability that a dolphin group is around in the morning?
- What is the probability that a dolphin group is partaking in travel given that it is morning?
- What is the probability that a dolphin group is around in the morning given that it is partaking in socializing?
- What is the probability that a dolphin group is around in the afternoon given that it is partaking in feeding?
- What is the probability that a dolphin group is around in the afternoon and is partaking in feeding?
- What is the probability that a dolphin group is around in the afternoon or is partaking in feeding?
- Are the events dolphin group around in the afternoon and dolphin group feeding mutually exclusive events? Why or why not?
- Are the events dolphin group around in the morning and dolphin group partaking in travel independent events? Why or why not?

Answer

- Independent, see solutions
- Dependent, see solutions
- a. $P(\text{Jack/face card}) = 0.333$, b. $P(\text{heart/card a 3}) = 0.25$, c. $P(\text{red card/ace}) = 0.50$, d. not independent, see solutions, e. independent, see solutions

7. 0.0625

9. 4.55×10^{-4}

11. a. $P(\text{female}) = 0.684$, b. $P(\text{1st class}) = 0.429$, c. $P(\text{female/1st class}) = 0.694$, d. $P(\text{female and 1st class}) = 0.298$, e. $P(\text{female or 1st class}) = 0.816$, f. No, see solutions, g. Dependent, see solutions

This page titled [4.3: Conditional Probability](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Kathryn Kozak](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.

- **4.3: Conditional Probability** by [Kathryn Kozak](#) is licensed [CC BY-SA 4.0](#). Original source: <https://s3-us-west-2.amazonaws.com/oerfiles/statsusingtech2.pdf>.