Sometimes we already know something about the outcome of an experiment before we compute a probability. This extra information can change the probability. For example, the probability that a randomly chosen card is a king is $\frac{4}{52} = \frac{1}{13}$. But if someone peeks at the card and tells us that it is a face card, then the probability that it is a king is much larger. In this handout we learn how to compute such probabilities.

Example 1. A fair die is rolled. We are told that the result is an even number. What is the probability that the result is a 6?

Solution. Without the extra information, the probability of a 6 would be $\frac{1}{6}$. But now we know that the result is one of 2, 4, or 6. These three outcomes are still equally likely, and only one of them is a 6. So the probability is $\frac{1}{3}$.

The extra information made the sample space smaller: instead of $S = \{1, 2, 3, 4, 5, 6\}$, we only need to look at the even outcomes, $\{2, 4, 6\}$.

Definition. The probability that F occurs when we already know that E has occurred is called the **conditional probability** of F given E. We denote it by $P(F \mid E)$ and read it as "the probability of F given E." If all outcomes are equally likely, then

$$P(F \mid E) = \frac{n(E \text{ and } F)}{n(E)}.$$

In words: the event we know about, E, becomes the new sample space, and we count how many of its outcomes are also in F.

In Example 1, E is "even" and F is "6." So $n(E) = 3$, $n(E \text{ and } F) = 1$, and $P(6 \mid \text{even}) = \frac{1}{3}$.

Example 2. Two fair dice are rolled, and we are told that the sum is 8. Find the probability that

a) one of the dice shows a 3, b) the two dice show the same number.

Solution. Of the 36 equally likely outcomes, five have a sum of 8:

$$(2, 6) \quad (3, 5) \quad (4, 4) \quad (5, 3) \quad (6, 2)$$

These five outcomes form the new sample space.

a) Two of them, $(3, 5)$ and $(5, 3)$, contain a 3. So $P(\text{a 3} \mid \text{sum is 8}) = \frac{2}{5}$.

b) Only $(4, 4)$ shows the same number twice. So $P(\text{doubles} \mid \text{sum is 8}) = \frac{1}{5}$.

© Hidegkuti, 2026

Two-Way Tables

Data about two questions are often collected in a two-way table. Conditional probabilities are easy to read from such a table: the given information tells us which row or column to look at.

Example 3. An instructor recorded, for each of her 120 students, whether the student did the homework regularly and whether the student passed the course.

	passed	failed	total
did homework regularly	63	7	70
did not do homework regularly	20	30	50
total	83	37	120

A student is selected at random. Find the probability that the student

a) passed, b) passed, given that the student did the homework regularly,

c) passed, given that the student did not do the homework regularly,

d) did the homework regularly, given that the student passed.

Solution. a) 83 of the 120 students passed, so $P(\text{passed}) = \dfrac{83}{120} \approx 0.69$.

b) We only look at the first row: 70 students did the homework regularly, and 63 of them passed.

$$P(\text{passed} \mid \text{homework}) = \frac{63}{70} = 0.9$$

c) We only look at the second row: 50 students, and 20 of them passed.

$$P(\text{passed} \mid \text{no homework}) = \frac{20}{50} = 0.4$$

d) Now we only look at the first column: 83 students passed, and 63 of them did the homework regularly.

$$P(\text{homework} \mid \text{passed}) = \frac{63}{83} \approx 0.76$$

Be careful. The answers to parts b) and d) are different. In general, $P(F \mid E)$ and $P(E \mid F)$ are *not* the same: the given event tells us which group we are looking at, and the two groups are different. In part b) we look at the students who did the homework; in part d) we look at the students who passed.

The Formula for Conditional Probability

In Example 3, we can divide the top and the bottom of the fraction by the total number of students, 120:

$$P(\text{passed} \mid \text{homework}) = \frac{63}{70} = \frac{\frac{63}{120}}{\frac{70}{120}} = \frac{P(\text{homework and passed})}{P(\text{homework})}$$

© Hidegkuti, 2026

This gives a formula that works even when we do not know the number of outcomes, only probabilities.

Conditional Probability Formula. If $P(E) > 0$, then

$$P(F \mid E) = \frac{P(E \text{ and } F)}{P(E)}.$$

Multiplying both sides by $P(E)$ gives the AND rule: $P(E \text{ and } F) = P(E) \cdot P(F \mid E)$.

Example 4. At a college, 60% of the students have a job, and 24% of the students have a job *and* take an online class. A student who has a job is selected at random. Find the probability that this student takes an online class.

Solution. We need $P(\text{online} \mid \text{job})$.

$$P(\text{online} \mid \text{job}) = \frac{P(\text{job and online})}{P(\text{job})} = \frac{0.24}{0.60} = 0.4$$

So 40% of the students who have a job take an online class.

Conditional Probability and Independence

Recall that two events are independent if the occurrence of one does not change the probability of the other. Using our new notation:

Independence. The events E and F are **independent** if $P(F \mid E) = P(F)$. Otherwise, they are **dependent**.

Example 5. A card is drawn from a standard deck of 52 cards.

a) Are the events "the card is red" and "the card is a king" independent?

b) Are the events "the card is a face card" and "the card is a king" independent?

Solution. Without any extra information, $P(\text{king}) = \frac{4}{52} = \frac{1}{13}$.

a) There are 26 red cards, and 2 of them are kings. So $P(\text{king} \mid \text{red}) = \frac{2}{26} = \frac{1}{13}$. This is the same as $P(\text{king})$, so the two events are independent. Knowing the color does not help us guess whether the card is a king.

b) There are 12 face cards, and 4 of them are kings. So $P(\text{king} \mid \text{face card}) = \frac{4}{12} = \frac{1}{3}$. This is much larger than $\frac{1}{13}$, so the two events are dependent.

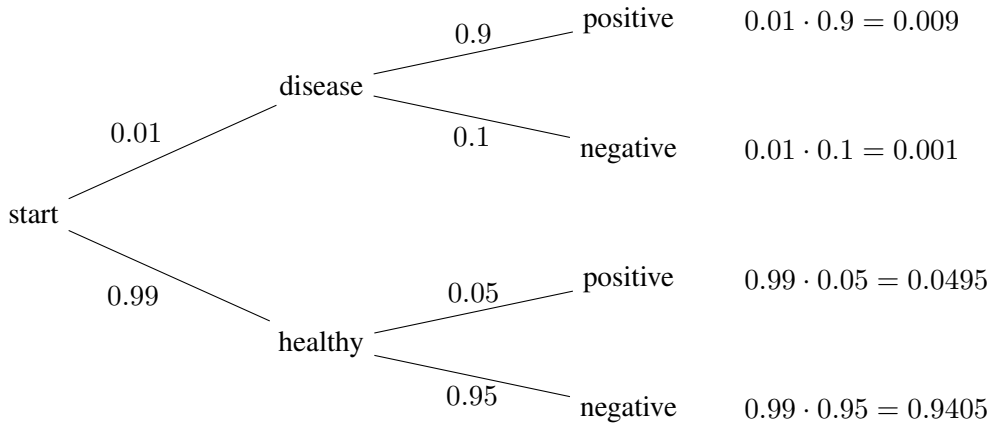
© Hidegkuti, 2026

Medical Tests

One of the most important uses of conditional probability is in understanding medical tests. The answer to the next example surprises most people, including many doctors.

Example 6. A disease affects 1% of the population. A test for the disease is positive for 90% of the people who have the disease. But the test is not perfect: it is also positive for 5% of the people who do not have the disease. A randomly selected person tests positive. What is the probability that this person has the disease?

Solution. We organize the information in a tree diagram. The first branches show whether the person has the disease. The second branches show the test result, given what happened first.



Two paths end with a positive test. So $P(\text{positive}) = 0.009 + 0.0495 = 0.0585$, and

$$P(\text{disease} \mid \text{positive}) = \frac{P(\text{disease and positive})}{P(\text{positive})} = \frac{0.009}{0.0585} \approx 0.154.$$

So a person who tests positive has only about a 15% chance of having the disease!

To see why, imagine $10\,000$ people. (We use $10\,000$ rather than $1\,000$ so that every count is a whole number.) About 1% of them, 100 people, have the disease, and 90% of those, 90 people, test positive. The other $9\,900$ people are healthy, but 5% of them, 495 people, also test positive.

	positive	negative	total
disease	90	10	100
no disease	495	9 405	9 900
total	585	9 415	10 000

We only look at the first column: out of the 585 positive tests, only 90 belong to people who have the disease:

$$P(\text{disease} \mid \text{positive}) = \frac{90}{585} \approx 0.154$$

Because the disease is rare, the healthy people with a false positive result outnumber the sick people. This is why a positive result is usually checked with a second test.

© Hidegkuti, 2026

Bayes' Theorem

In Example 6, we knew $P(\text{positive} \mid \text{disease}) = 0.9$, but the question asked for the reverse, $P(\text{disease} \mid \text{positive})$. Reversing a conditional probability like this is so common that the method has a name. It is called Bayes' theorem, after Thomas Bayes, an English minister of the 1700s.

Bayes' Theorem. For any two events A and B,

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(A) \cdot P(B \mid A) + P(\text{not } A) \cdot P(B \mid \text{not } A)}.$$

The top is the probability of the path "A, then B" in the tree diagram. The bottom adds up all the paths that end with B, so it is $P(B)$.

In Example 6, A is "disease" and B is "positive," and Bayes' theorem gives exactly the computation we did there:

$$P(\text{disease} \mid \text{positive}) = \frac{0.01 \cdot 0.9}{0.01 \cdot 0.9 + 0.99 \cdot 0.05} = \frac{0.009}{0.0585} \approx 0.154$$

Example 7. A factory has two machines that make light bulbs. Machine A makes 60% of the bulbs, and machine B makes the other 40%. Of the bulbs made by machine A, 2% are defective. Of the bulbs made by machine B, 5% are defective. A randomly selected bulb is defective. What is the probability that it was made by machine A?

Solution. We know $P(\text{defective} \mid \text{A}) = 0.02$, and we need $P(\text{A} \mid \text{defective})$. Here "not A" means machine B, so $P(\text{not A}) = 0.4$ and $P(\text{defective} \mid \text{not A}) = 0.05$. By Bayes' theorem,

$$P(\text{A} \mid \text{defective}) = \frac{0.6 \cdot 0.02}{0.6 \cdot 0.02 + 0.4 \cdot 0.05} = \frac{0.012}{0.012 + 0.02} = \frac{0.012}{0.032} = 0.375.$$

Even though machine A makes most of the bulbs, a defective bulb more likely came from machine B, because machine B makes defective bulbs much more often.

© Hidegkuti, 2026



Practice Problems

1. A fair die is rolled. Find the probability that
   a) the result is 2, given that it is even,
   b) the result is greater than 3, given that it is odd,
   c) the result is even, given that it is greater than 3.

2. A card is drawn from a standard deck of 52 cards. Find
   a) $P(\text{queen} \mid \text{face card})$ b) $P(\text{heart} \mid \text{red})$ c) $P(\text{face card} \mid \text{heart})$ d) $P(\text{red} \mid \text{ace})$

3. Two fair dice are rolled. Find the probability that
   a) the sum is 7, given that the first die shows a 4,
   b) the two dice show the same number, given that the sum is 6,
   c) the sum is 10 or more, given that at least one of the dice shows a 6.

4. In a survey, 250 adults were asked whether they have a dog and whether they have a cat.

	has a dog	no dog	total
has a cat	30	45	75
no cat	80	95	175
total	110	140	250

   One of these adults is selected at random. Find the probability that the person
   a) has a dog, b) has a dog, given that the person has a cat,
   c) has a cat, given that the person has a dog,
   d) has no cat, given that the person has no dog.
   e) Are the events "has a dog" and "has a cat" independent? Explain.

5. A fair coin is tossed three times. Find the probability of getting
   a) three heads, given that there is at least one head,
   b) exactly two heads, given that the first toss is heads.

6. A family has two children. Assume that each child is equally likely to be a boy or a girl. Find the probability that both children are girls, given that
   a) at least one of the children is a girl, b) the older child is a girl.

7. Suppose that $P(A) = 0.5$, $P(B) = 0.4$, and $P(A \text{ and } B) = 0.2$.
   a) Find $P(A \mid B)$. b) Find $P(B \mid A)$. c) Are A and B independent? Explain.

© Hidegkuti, 2026

8. A disease affects 2% of the population. A test is positive for 95% of the people who have the disease, and for 10% of the people who do not have it. Imagine $10\,000$ people.

   a) Complete the table.

	positive	negative	total
disease			
no disease			
total			10 000

   b) Find the probability that a person who tests positive has the disease.

   c) Find the probability that a person who tests negative does not have the disease.

9. A bag contains 6 red and 4 green marbles. Two marbles are drawn without replacement. Find the probability that

   a) the second marble is red, b) the first marble is red, given that the second marble is red.

10. An email filter knows that 30% of all emails are spam. The word "free" appears in 40% of the spam emails and in 5% of the other emails. An email contains the word "free." What is the probability that it is spam? Round your answer to three decimal places.

© Hidegkuti, 2026



Answers

1. a) $\frac{1}{3}$ b) $\frac{1}{3}$ c) $\frac{2}{3}$

2. a) $\frac{4}{12} = \frac{1}{3}$ b) $\frac{13}{26} = \frac{1}{2}$ c) $\frac{3}{13}$ d) $\frac{2}{4} = \frac{1}{2}$

3. a) $\frac{1}{6}$ b) $\frac{1}{5}$ c) $\frac{5}{11}$

4. a) $\frac{110}{250} = \frac{11}{25} = 0.44$ b) $\frac{30}{75} = \frac{2}{5} = 0.4$ c) $\frac{30}{110} = \frac{3}{11}$ d) $\frac{95}{140} = \frac{19}{28}$
   e) Dependent, because $P(\text{dog} \mid \text{cat}) = 0.4$ is not equal to $P(\text{dog}) = 0.44$.

5. a) $\frac{1}{7}$ b) $\frac{2}{4} = \frac{1}{2}$

6. a) $\frac{1}{3}$ b) $\frac{1}{2}$

7. a) $\frac{0.2}{0.4} = 0.5$ b) $\frac{0.2}{0.5} = 0.4$ c) Independent, because $P(A \mid B) = 0.5 = P(A)$.

8. a)

	positive	negative	total
disease	190	10	200
no disease	980	8 820	9 800
total	1 170	8 830	10 000

   b) $\frac{190}{1\,170} \approx 0.162$ c) $\frac{8\,820}{8\,830} \approx 0.999$

9. a) $\frac{6}{10} \cdot \frac{5}{9} + \frac{4}{10} \cdot \frac{6}{9} = \frac{30}{90} + \frac{24}{90} = \frac{54}{90} = \frac{3}{5}$ b) $\frac{30/90}{54/90} = \frac{30}{54} = \frac{5}{9}$

10. $\frac{0.3 \cdot 0.4}{0.3 \cdot 0.4 + 0.7 \cdot 0.05} = \frac{0.12}{0.155} \approx 0.774$

For more documents like this, visit our page at https://teaching.martahidegkuti.com and click on Lecture Notes. E-mail questions or comments to mhidegkuti@ccc.edu.