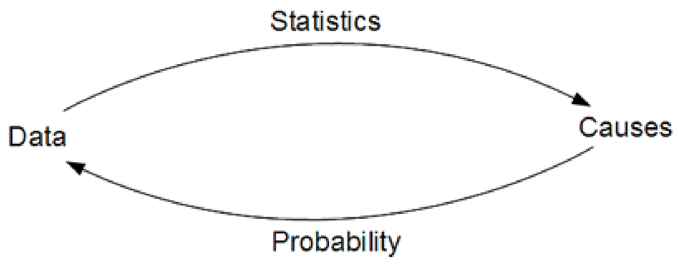
ST101 Unit 2: Probabilities in Data

Contents

Probability

In this unit we will be talking about probability. In a sense, probability is just the opposite of statistics. Put differently, in statistics we are given data and try to infer possible causes, whereas in probability we are given a description of the causes and we try to predict the data.



The reason that we are studying probability rather than statistics is that it will give us a language to describe the relationship between data and the underlying causes.

Flipping coins

Flipping a coin creates data. Each flip of the coin will result in either a ‘head’ or a ‘tail’ result. A fair coin is one that has a 50% chance of coming up heads and a 50% chance of coming up tails. Probability is a method for describing the anticipated outcomes of these coin flips.

Fair Coin

The probability of a coin coming up heads is written using the notation:

P(Heads) =

In a fair coin, the chance of a coin flip coming up heads is 50%. In probability, this is given as a probability of 0.5:

P(Heads) = 0.5

A probability of 1 means that the outcome will always happen. A probability of 0 means that it will never happen. In a fair coin, the probability that a flip will come up tails is:

P(Tails) = 0.5

The sum of the probabilities of all possible outcomes is always 1. So:

P(Heads) + P(Tails) = 1

Loaded Coin

A loaded coin is one that comes up with one outcome much more frequently than the other.

Loaded Coin Quiz

Suppose the probability of heads is 0.75 for a particular coin. What is the probability than the coin-flip will come up tails?

Complementary Outcomes

If we know the probability of an outcome, A, then the probability of the opposite outcome, ¬A ("not A"), is given by:

$$P(\neg A) = 1 - P(A)$$

This is a very basic law of probability.

Two Flips

So what happens when we flip the same, unbiased, coin twice? What is the probability of getting two heads in a row, assuming that P(H) = 0.5?

We can derive the answer to this type of problem using a **truth table**. A truth table enumerates every possible outcome of the experiment. In this case:

Flip-1	Flip-2	Probability
H	H	0.25
H	T	0.25
T	H	0.25
T	T	0.25

In the case of two coin flips, there are four possible outcomes, and because heads and tails are equally likely, each of the four outcomes is equally likely. Since the total probability must equal 1, the probability of each outcome is 0.25.

Another way to consider this is that the probability that we will see a head, followed by another head is the product of the probabilities of the two events:

$$P(H, H) = P(H) \times P(H) = 0.5 \times 0.5 = 0.25$$

So what happens if the coin is loaded?

Well, if the probability of getting a head, P(H), is 0.6, then the probability that we will see a tail, P(T), is going to be 0.4, and the truth table will be:

Flip-1	Flip-2	Probability
H	H	0.6 x 0.6 = 0.36
H	T	0.6 x 0.4 = 0.24
T	H	0.4 x 0.6 = 0.24
T	T	0.4 x 0.4 = 0.16

Notice that the total probability is still 1:

$$0.36 + 0.24 + 0.24 + 0.16 = 1$$

The truth table lists all possible outcomes, so the sum of the probabilities will always be 1.

Two Flips Quiz

Suppose the probability of heads, P(H) = 1. What is the probability of seeing two heads on successive flips, P(H, H)?

One Head

The truth table can get more interesting when we ask different questions. Suppose we flip the coin twice, but what we care about is that *exactly* one of the two flips reveals a head. For a fair coin, where P(H) = 0.5, the probability is:

$$P(\text{Exactly one H}) = 0.5$$

We can see from the truth table that there are exactly two possible outcomes with exactly one head:

Flip-1	Flip-2	Probability
H	H	0.25
H	T	0.25
T	H	0.25
T	T	0.25

The probability of these outcomes is 0.25 + 0.25 = 0.5.

One of Three Quiz

Suppose we take a fair coin where P(H) = 0.5, and we flip is three times. What is the probability that exactly one of those flips will be a head?

One of Three Quiz 2

What about if the coin is loaded with P(H) = 0.6. What is the probability that exactly one flip out of three will be a head?

Even Roll Quiz

Say you have a fair 6-sided die. The probability of each number appearing on any given throw of the die is 1/6.

What is the probability that a throw will be even?

Doubles Quiz

Suppose we throw a fair die twice. What is the probability that we throw the same number on each throw (i.e. a "double")?

Summary

In this section we learned that if we know the probability of an event, P(A) the probability of the opposite event is just 1 – P(A).

We also learned about composite events where the probability is given by:

$$P(A) \times P(A) \times \ldots \times P(A)$$

Now technically, these conditional events imply **independence**. This just means that the outcome of the second coin flip does not depend on the outcome of the first. In the next section we will look at dependence.

Conditional Probability

In real life, things depend on each other. For example, people can be born smart or dumb. For simplicity, let's assume that whether they're born smart or dumb is just nature's equivalent of the flip of a coin.

Now, whether they become a Stanford professor is not entirely independent of their intelligence. In general, becoming a Stanford professor is not very likely. The probability may only be 0.0001, but it also depends on their intelligence. If they are born smart, the probability may be higher.

In the previous section, subsequent events like coin tosses were independent of what had happened before. We are now going to look at some more interesting cases where the outcome of the first event does have an impact on the probability of the outcome of the second.

Cancer Example

Let's suppose that there is a patient who may be suffering from cancer. Let's say that the probability of a person getting this cancer is 0.1:

$$P(\text{Cancer}) = 0.1$$

$$P(\neg\text{Cancer}) = 0.9$$

Now, we don't know whether the person actually has cancer, but there is a blood test that we can give. The outcome of the test may be positive, or it may be negative, but like any good test, it tells us something about the thing we really care about – in this case whether or not the person has cancer.

Let's say that the probability of a positive test when a person has cancer is 0.9:

$$P(\text{Positive} \mid \text{Cancer}) = 0.9$$

and, $P(\text{Negative} \mid \text{Cancer}) = 0.1$

The sum of the possible test outcomes will always e equal to 1.

This is called the **sensitivity** of the test. Now, this notation says that the result of the test depends on whether or not the person has cancer. This is known as a **conditional probability**.

In order to fully specify the test, we also need to specify the probability of a positive test in the case of a person who doesn't have cancer. In this case, we will say that this is 0.2:

$$P(\text{Positive} \mid \neg\, \text{Cancer}) = 0.2$$

$$P(\text{Negative} \mid \neg\, \text{Cancer}) = 0.8$$

This is the **specificity** of the test. We now have all the information we need to derive the truth table:

Cancer	Test	P( )
Y	Positive	0.1 x 0.9 = 0.09
Y	Negative	0.1 x 0.1 = 0.01
N	Positive	0.9 x 0.2 = 0.18
N	Negative	0.9 x 0.8 = 0.72
		$\Sigma = 1.0$

We can now use the truth table to find the probability that we will see a positive tets result

$$P(\text{Positive}) = 0.09 + 0.18 = 0.27$$

Total Probability

Let’s put this into mathematical notation. We were given the probability of having cancer, P(C), from which we were able to derive the probability of not having cancer:

$$P(\neg C) = 1 - P(C)$$

We also had the two conditional probabilities, $P(+ \mid C)$ and $P(+ \mid \neg C)$, and from these we were able to derive the probabilities of a negative test:

$$P(- \mid C) = 1 - P(+ \mid C)$$

$$\text{and } P(- \mid \neg C) = 1 - P(+ \mid \neg C)$$

Then, the probability of a positive test result was:

$$P(+) = P(C).P(+ \mid C) + P(\neg C).P(+ \mid \neg C)$$

This is known as **total probability**. Let’s consider another example.

Two Coins

Image that we have a bag containing two coins. We know that coin1 is fair, and coin2 is loaded, so that:

$$P_1(H) = 0.5 \text{ and } P_1(T) = 0.5$$

$$P_2(H) = 0.9 \text{ and } P_2(T) = 0.1$$

We now pick a coin from the bag. Each coin has an equal probability of being picked from the bag. We flip the coin once.

Two Coins Quiz 1

What is the probability that the coin comes up heads?

Two Coins Quiz 2

Let's say that we now flip the coin twice. What is the probability that we will see a head first followed by a tail?

Two Coins Quiz 3

Now the bag contains two new coins. both are loaded:

$$P(H \mid 1) = 1$$

$$P(H \mid 2) = 0.6$$

The probability of picking coin 1 is still 0.5:

$$P(1) = 0.5$$

What is the probability of flipping the coin twice and seeing two tails?

Bayes Rule

In this section, we introduce what may be the Holy Grail of probabilistic inference. It's called Bayes' Rule. The rule is based on work by Reverent Thomas Bayes who used the principle to infer the existence of God. In doing so, he created a new family of methods that have vastly influenced artificial intelligence and statistics.

Let's think about the cancer example from the previous section. Say that there is a specific cancer that occurs in 1% of the population. There is a test for this cancer that has a 90% chance of a positive result if the person has cancer. The specificity of the test is 90%, i.e. there is a 90% chance of a negative test result if the person doesn't have cancer:

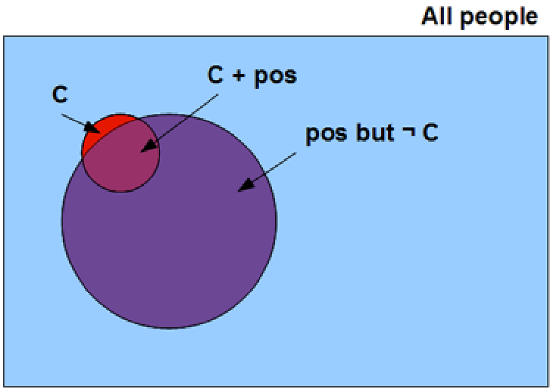
$$P(C) = 0.01$$

$$P(+ \mid C) = 0.9$$

$$P(- \mid \neg C) = 0.9$$

So, here is the question. What is the probability that a person has cancer, given that they have had a positive test?

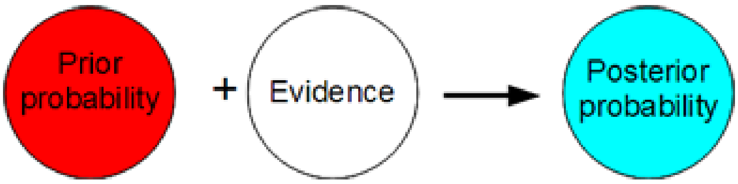
Let's show the figures on a diagram:



Only 1% of the people have this cancer. 99% are cancer-free. The test for this cancer catches 90% of those who have the cancer, which is 90% of the cancer circle. But the test can also give a positive result even when the person doesn't have cancer. In our case a false-positive can occur in 10% of cases – *which is 10% of the total population*. The remaining area represents the case of people who don't have the cancer and get a negative result from the test.

In fact, the area **C + pos** in the diagram above is actually about 8.3% of the total area representing a positive test result. So a positive test has only raised the probability that the person has cancer by a factor of about 8.

So this is the basis of Bayes' Rule. We start with some prior probability before we run the test, and then we get some evidence from the test itself, which leads us to what is known as a posterior probability:



In our example, we have the prior probability, P(C), and we obtain the posterior probabilities as follow. First we calculate what are known as the **joint probabilities**:

$$P(C \mid pos) = P(C) \,.\, P(Pos \mid C)$$

$$P(\neg C \mid pos) = P(\neg C) \,.\, P(Pos \mid \neg C)$$

Given the values in our example we get:

$$P(C \mid pos) = 0.01 \times 0.9 = 0.009$$

$$P(\neg C \mid pos) = 0.99 \times 0.1 = 0.099$$

These values are *non-normalised* – they do not sum to 1. In terms of our diagram above, they are the absolute areas of the regions representing a positive test.

We obtain the posterior probabilities by normalising the joint probabilities. To do this, we divide each of the joint probabilities by the probability of a positive test result:

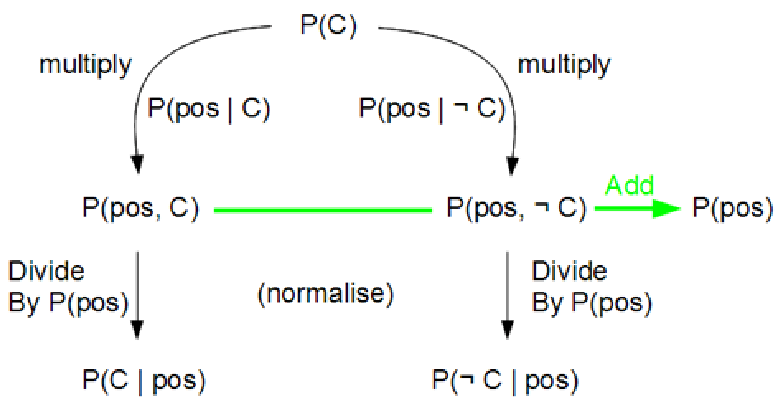
$$P(pos) = P(C \mid pos) + P(\neg C \mid pos)$$

So the posterior probabilities are:

$$P(C \mid pos) = \frac{P(C).P(pos \mid C)}{P(pos)}$$

$$P(\neg C \mid pos) = \frac{P(\neg C).P(pos \mid \neg C)}{P(pos)}$$

We can represent the process of calculating Bayes' Rule in a diagram:



Let's say we get a positive test. We have a prior probability, and a test with a given sensitivity and specificity:

Prior: P(C)

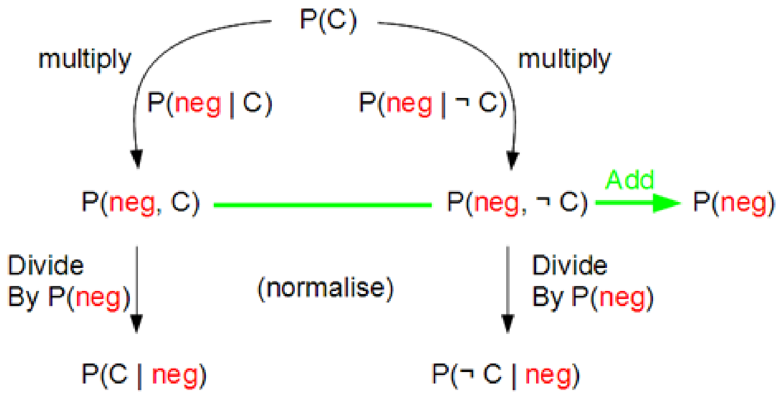
Sensitivity: P(pos | C)

Specificity: P(pos | ¬ C)

We multiply the prior by the sensitivity and by the specificity. This gives us a number that combines the cancer hypothesis with the test result for each of the cases – cancer or non-cancer. We add these numbers (normally, they do not add up to 1), to get the total probability of a positive test.

Now all we need to do to obtain the posterior probabilities is to normalise the two numbers by dividing by the total probability, P(pos).

This is our algorithm for Bayes Rule, and we can produce an almost exactly similar diagram for a negative test result as shown below:



Let's work through an example. We start with our prior probability, sensitivity and specificity:

$$P(C) = 0.001$$

$$P(pos \mid C) = 0.9$$

$$P(neg \mid \neg C) = 0.9$$

Cancer Probabilities Quiz

Calculate the probabilities of

- $P(\neg C)$
- $P(neg \mid C)$
- $P(pos \mid \neg C)$

Probability Given Test Quiz

Assume that the test comes back negative. Calculate

- $P(C \mid neg)$ #the combined probability of having cancer given the negative test result
- $P(\neg C \mid neg)$ #the combined probability of being cancer-free given the negative test result

Normaliser Quiz

Calculate the normaliser, P(neg)

Normalising Probability Quiz

What is the posterior probability of cancer, given that we had a negative test result?

What is remarkable about the result is what the posterior probabilities actually mean. Before the test, we had a 1% chance of having cancer. After a negative test result this has gone down by about a factor of 9.

Conversely, before the test there was a 99% chance that we were cancer-free. That number has now gone up to 99.89%, greatly increasing our confidence that we are cancer free.

Let's consider another example. In this case the prior probability, sensitivity and specificity are:

$$P(C) = 0.1$$

$$P(pos \mid C) = 0.9$$

$$P(neg \mid \neg C) = 0.5$$

So the sensitivity is high, but the specificity is much lower.

Disease Test Quiz 1

What are the values of the probabilities:

- $P(\neg C)$
- $P(neg \mid C)$
- $P(pos \mid \neg C)$

Disease Test Quiz 2

What are the values of the probabilities:

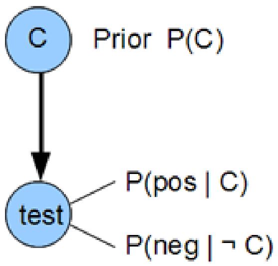
- $P(C, neg)$
- $P(\neg C, neg)$
- $P(neg)$

Disease Test Quiz 3

What are the values of the probabilities:

- P(C | neg)
- P(¬ C | neg)

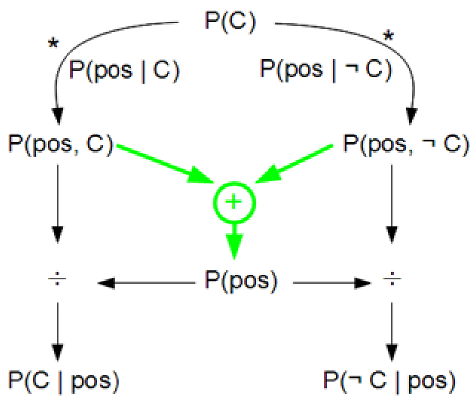
Bayes Rules Summary



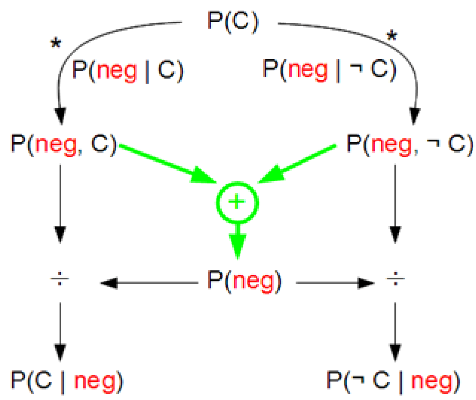
In Bayes Rule, we have a hidden variable that we care about, but can't measure directly. Instead we have a test. We have a prior probability of how often the variable is true. The test characterised by how often it gives a positive result when the variable is true (sensitivity), and how often it gives a negative result when the variable is false (specificity).

Bayes rule then applies the algorithm we saw earlier to calculate the posterior probabilities for the variable given a test outcome:

Positive test:



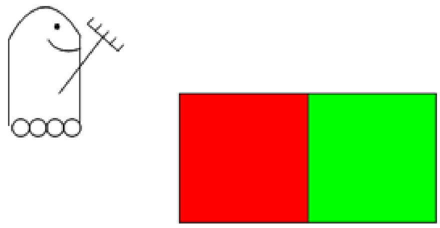
Negative Test:



Robot Sensing

Let's practice using Bayes' Rule with a different example.

Consider a robot living in a world that has exactly two places. There is a red place, R, and a green place, G:



Initially, the robot has no idea of it's location, so the prior probabilities are:

$$P(R) = P(G) = 0.5$$

The robot has sensors that allow it to 'see' its environment, but these sensors are somewhat unreliable:

$$P(\text{see R} \mid \text{in R}) = 0.8$$

$$p(\text{see G} \mid \text{in G}) = 0.8$$

Robot Sensing Quiz 1

Suppose the robot sees red. What is the posterior probability that the robot is in the red cell? What is the posterior probability that it is in the green cell?

Robot Sensing Quiz 2

Suppose the prior probabilities are now:

$$P(R) = 0$$

$$P(G) = 1$$

Once again the robot sees red. What is the posterior probability that the robot is in the red cell? What is the posterior probability that it is in the green cell?

Robot Sensing Quiz 3

Given the probabilities:

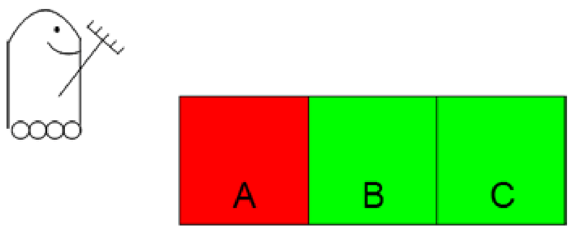
$$P(R) = P(G) = 0.5$$

$$P(\text{see } R \mid \text{in } R) = 0.8$$

$$p(\text{see } G \mid \text{in } G) = 0.5$$

Now the robot sees red. Calculate the posterior probability that the robot is in the red cell, and the posterior probability that it is in the green cell.

Let’s make things a little more complicated. Suppose that there are now three places in the robot’s world, one red and two green. For simplicity, we will label these A, B and C:



So the hidden variable now has three states. We will assume that each place has the same prior probability:

$$P(A) = P(B) = P(C) = 1/3$$

The robot sees red, and we know that:

$$P(R \mid \mathbf{A}) = 0.9$$

$$P(G \mid \mathbf{B}) = 0.9$$

$$P(G \mid \mathbf{C}) = 0.9$$

We can solve for the posterior probabilities exactly as before.

$$P(A, R) = P(A) \text{ x } P(R \mid A) = 1/3 \text{ x } 0.9 = 0.3$$

$$P(B, R) = P(B) \text{ x } P(R \mid B) = 1/3 \text{ x } 0.1 = 0.0333$$

$$P(C, R) = P(C) \text{ x } P(R \mid C) = 1/3 \text{ x } 0.1 = 0.0333$$

So, the normaliser is

$$P(R) = 0.3 + 0.0333 + 0.0333 = 0.3667$$

Which gives the posterior probabilities:

$$P(A \mid R) = 0.82$$

$$P(B \mid R) = 0.09$$

$$P(C \mid R) = 0.09$$

Generalising

The last example showed that there may be more than two states of the hidden variable that we are interested in. There may be 3, 4, 5 or any other number. We can solve these cases using exactly the same methods, but we have to keep track of more values.

In fact, there can be more than just two outcomes of the test. For example, the robot may see red, green or blue. This means that our measurement probabilities will be more elaborate, but the actual method for calculating the posterior probabilities will remain the same.

We can now deal with very large problems, that have many possible hidden causes, by applying Bayes' Rule to determine the posterior probabilities.

Sebastian at Home Quiz

Sebastian has a problem. He travels a lot. It has got so bad that he sometimes wakes up in bed not knowing what country he is in. He is only at home 40% of the time

$$P(\text{away}) = 0.6$$

$$P(\text{home}) = 0.4$$

In the summer, Sebastian lives in California. Typically, it doesn't rain in California in the summer. Whereas, there is a much higher chance of rain in many of the countries that Sebastian travels to:

$$P(\text{rain} \mid \text{home}) = 0.01$$

$$P(\text{rain} \mid \text{away}) = 0.3$$

Let's say that he wakes up and hears it raining. What is the probability that he is at home in California?

Programming Bayes' Rule (optional)

This section is optional.

To print a number in Python, we just use the print statement thus:

```
print 0.3
```

will cause 0.3 to be printed by the Python interpreter.

A **function** or procedure in Python takes some inputs and produces outputs. A function lets us use the same code to operate on different data by passing that data as the input to the function. We define a Python function as follows:

```
def <Name>(<Parameters>):
   <Block>
```

For example, the following simple function, **f**, just returns whatever parameter **p** it is given:

```
def f(p):
   return p

print f(0.3)
```

The print statement just prints the output of the function, given the input 0.3. This has exactly the same effect as before, but in this case, the print statement isn't printing 0.3 directly, but rather it is printing the output of the function .

Let's say that we are going to provide the probability of an event (P = 0.3) to our function, and we want the function to return the probability of the inverse event. To do this we just modify the function as follows:

```
def f(p):
   return 1 - p

print f(0.3)
```

This will print the result 0.7. If we change the input value, the interpreter will print a different output.

Two Flips Quiz

Suppose we have a coin with P(H) = p. Write a function that returns the probability of seeing heads twice, i.e. P(H, H)

Three Flips Quiz

Let's say that we now flip the coin three times. Write a function that returns the probability of seeing heads exactly once.

Flip Two Coins Quiz

We now have two coins. Coin 1 has a probability P(H) = p1, and coin 2 has a probability P(H) = p2. Our function will now need to take two arguments as inputs:

```
def f(p1, p2):
```

Write a function that returns the probability of that both coins come up heads.

Flip One of Two Quiz

We have two coins. Coin 1 has a probability P(H) = p1, and coin 2 has a probability P(H) = p2. We pick one coin from a bag. The probability that we pick coin 1 is P_0, and the probability that we pick coin 2 is 1 - P_0:

$$P(C_1) = P_0$$

$$P(C_2) = 1 - P_0$$

What is the probability that we get heads when we flip the coin?

Write a function with three input arguments, that calculates the probability that we get heads when we flip the coin.

Cancer Example Quiz

Let's return to our cancer example. We have our prior probability, P_0, the probability of a positive test, given cancer, P_1, and there's the probability of a negative test for not-cancer which we'll call P_2.

P© = P_0

$$P(pos \mid C) = P_1$$

$$P(neg \mid \neg C) = P_2$$

What is the formula to calculate probability of a positive test P(pos)?

Calculate Total Quiz

Write a function to calculate the probability of a positive test.

Program Bayes' Rule Quiz

What is the formula to calculate the posterior probability of having cancer following a positive test (using the variable defined above)?

Write a function to calculate the posterior probability of having cancer following a positive test.

Program Bayes' Rule Quiz 2

Write a function to calculate the posterior probability of having cancer following a *negative* test, P(C | neg).

Correlation vs. Causation

In the last unit, we described Simpson's Paradox and showed how it was surprisingly easy to draw false conclusions from data. In this section we will try to give you an insight into a common mistake that is made when interpreting statistical data as a result of confusing **correlation** with **causation**. Newspaper articles frequently confuse correlation with causation.

We will show an example where data is correlated, and show why it is tempting to confuse correlation with causation.

Suppose that you are sick. In fact, you are so sick that you fear that you may die. Fortunately, you're not sick enough so that you can't apply the lessons of this class in order to make a rational decision about whether to go to the hospital.

You consult the data, and find that at your local hospital 40 people were hospitalised, and 4 of these people died. You also find that the majority of the population of your town, 8000 people, didn't visit the hospital, and 20 of these people died at home.

So, 10% of those who were admitted to hospital died, and 0.25% of those who were at home died. This means that the chances of dying in hospital are 40 times greater than the chances of dying at home. There is therefore a ***correlation*** between whether or not you die, and whether or not you are in hospital.

But this doesn't mean that hospitals cause sick people to die.

The statement:

> “The chances of dying in hospital are 40 times greater than dying at home”

shows that there is a ***correlation*** between whether or not you die, and whether or not you are in hospital. Whereas, the statement:

> “Being in a hospital increases your probability of dying by a factor of 40”

is a causal statement. It says that being in hospital causes you to die, not just that being in hospital coincides with the fact that you die. People frequently get this wrong. They observe a correlation, but they suggest that the correlation is causal.

To understand why this could be wrong, let’s look a little deeper into our example.

Considering Health

Let’s say that of the people in the hospital, 36 were sick, and 4 of these died. Four of the people in the hospital were actually healthy, and they all survived.

Of the people who were at home, 40 were actually sick, and 20 of these people died. The remaining 7960 were healthy, but 20 of these people also died (perhaps due to accidents etc.).

These statistics are consistent with our earlier statistics. We have just added another variable - whether a person is sick or healthy.

The percentages of people who died are tabulated below:

	In Hospital	Died	
Sick	36	4	11.11%
Healthy	4	0	0%

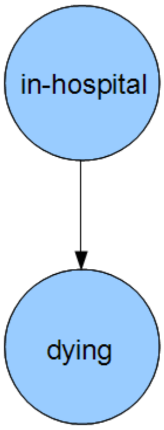
	At Home	Died	
Sick	40	20	50%
Healthy	7960	20	0.2513%

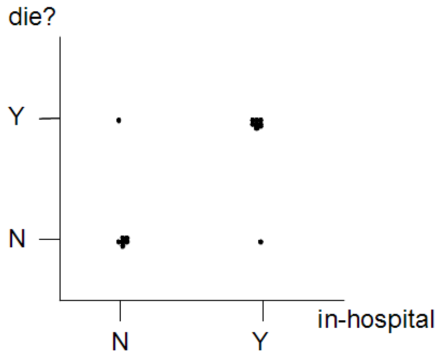
Now, if you are sick, your chances of dying at home are 50% compared with about 11% in the hospital, so you should really make your way to the hospital.

Correlation

So why does the hospital example lead us to draw such a wrong conclusion?

We looked at two variables, being in hospital and the chance of dying, and we rightfully observed that these two things are correlated. If we had a scatter-plot with two categories – whether a person was in hospital, & whether or not that person died - we would see increased occurrence of data points as shown below:



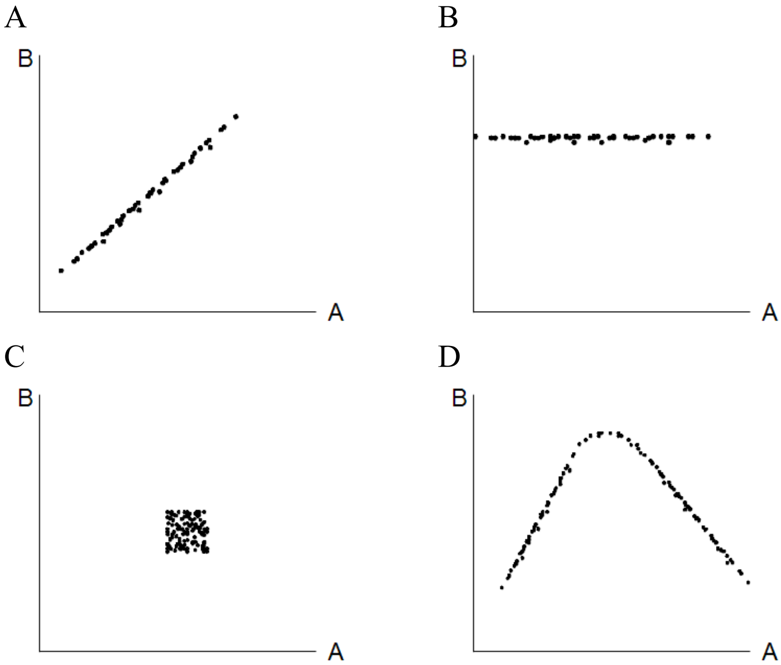


This shows that the data correlates.

So what is correlation? Well, in any plot, data is correlated if knowledge about on variable tells us something about the other.

Correlation Quiz

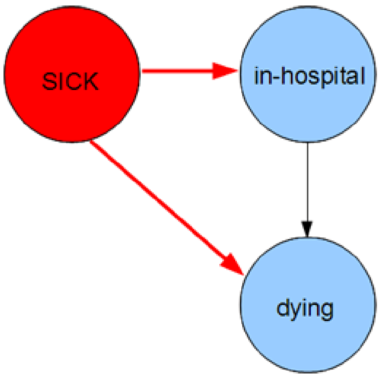
Are the following data pots correlated?



Causation Structure

In the example above, there is clearly a correlation between whether a person is in hospital, and whether or not they die. But we initially left out an important variable: whether or not a person was sick.

In fact, it was the sickness that caused people to die. If we add arcs of causation to our diagram, we find that sickness causes death, and that sickness also causes people to go into hospital:



In fact, in our example, once a person knew that they were sick, being in a hospital *negatively correlated* with them dying. That is, being in a hospital made it less likely that they would die, given that they were sick.

In statistics, we call this a confounding variable. It can be very tempting to just omit this from your data, but if you do, you might find correlations that have absolutely nothing to do with causation.

Fire Correlation

Suppose we study a number of fires. We recorded the number of fire-fighters and the surface area (size) of the fire.

# Fire-fighters	# Size fire
10	100
40	400
200	2000
70	700

Clearly, the number of fire-fighters is correlated with the size of the fire. But fire-fighters don't cause the fires! Getting rid of all the fire-fighters will not get rid of all the fires. This is actually a case of **reverse causation**. The size of the fire determines the number of fire-fighters that will be sent to deal with it.

However, it is impossible to know this just from the data. We only know that this is a case of reverse causation because we already know something about fire and fire-fighters.

Assignment

Check out old news articles in newspapers, or online, and find some that takes data which shows a correlation, and from the data suggests causation, or tells you what to do based on that data.

You will find that the news is full of such abuses of statistics.

Answers

Loaded Coin Quiz

P(Tails) = 1 – P(Heads) = 0.25

Two Flips Quiz

P(H, H) = 1

One of Three Quiz

0.375

Flip-1	Flip-2	Flip-3	Probability
H	H	H	0.125
H	H	T	0.125
H	T	H	0.125
H	T	T	**0.125**
T	H	H	0.125
T	**H**	T	**0.125**
T	T	**H**	**0.125**
T	T	T	0.125

One of Three Quiz 2

P(H) = 0.6
P(T) = 0.4

Flip-1	Flip-2	Flip-3	Probability
H	H	H	0.216
H	H	T	0.144
H	T	H	0.144
H	T	T	**0.096**
T	H	H	0.144
T	**H**	T	**0.096**
T	T	**H**	**0.096**
T	T	T	0.064

P(exactly 3 Heads) = 0.288

Even Roll Quiz

Three possible outcomes are even (2, 4, 6), so the probability that a throw is even is: 3 x 1/6 = 0.5

Doubles Quiz

The truth table has 36 possible outcomes. Each outcome in the truth table will have a probability of 1/36. Six of these outcomes are "doubles", so the probability of a double is:

6 x 1/36 = 1/6 =0.16667

Two Coins Quiz 1

Pick	Flip	P( )
1	H	0.25
1	T	0.25
2	H	0.45
2	T	0.05

So, P(H) = 0.25 + 0.45 = 0.7

Two Coins Quiz 2

Pick	Flip1	Flip2	P()
1	H	H	0.5 x 0.5 x 0.5 = 0.125
1	H	T	0.5 x 0.5 x 0.5 = 0.125
1	T	H	0.5 x 0.5 x 0.5 = 0.125
1	T	T	0.5 x 0.5 x 0.5 = 0.125
2	H	H	0.5 x 0.9 x 0.9 = 0.405
2	H	T	0.5 x 0.9 x 0.1 = 0.045
2	T	H	0.5 x 0.1 x 0.9 = 0.045
2	T	T	0.5 x 0.1 x 0.1 = 0.005

P(H, T) = 0.125 + 0.045 = 0.17

Two Coins Quiz 3

Pick	Flip1	Flip2	P()
1	H	H	0
1	H	T	0
1	T	H	0
1	T	T	0
2	H	H	0.5 x 0.6 x 0.6 = 0.18
2	H	T	0.5 x 0.6 x 0.4 = 0.12
2	T	H	0.5 x 0.4 x 0.6 = 0.12
2	T	T	0.5 x 0.4 x 0.4 = 0.08

P(T, T) = 0.08

Cancer Probabilities Quiz

- $P(\neg C) = 0.99$
- $P(neg \mid C) = 0.1$
- $P(pos \mid \neg C) = 0.1$

Probability Given Test Quiz

Assume that the test comes back negative. Calculate

- $P(C \mid neg) = 0.01 \times 0.1 = 0.001$
- $P(\neg C \mid neg) = 0.99 \times 0.9 = 0.891$

Normaliser Quiz

$p(neg) = 0.892$

Normalising Probability Quiz

$P(C \mid neg) = 0.0011$
$P(\neg C \mid neg) = 0.9989$

Disease Test Quiz 1

- $P(\neg C) = 0.9$
- $P(neg \mid C) = 0.1$
- $P(pos \mid \neg C) = 0.5$

Disease Test Quiz 2

- $P(C, neg) = 0.1 \times 0.1 = 0.01$
- $P(\neg C, neg) = 0.9 \times 0.5 = 0.45$
- $P(neg) = 0.46$

Disease Test Quiz 3

- $P(C \mid neg) = 0.0217$
- $P(\neg C \mid neg) = 0.9783$

Robot Sensing Quiz 1

P(at R | see R) = 0.8
P(at G | see R) = 0.2

Robot Sensing Quiz 2

P(at R | see R) = 0
P(at G | see R) = 1

Robot Sensing Quiz 3

P(at R | see R) = 0.615
P(at G | see R) = 0.385

Sebastian at Home Quiz

P(rain) = P(home) x P(rain | home) + P(away) x P(rain | away)
P(rain) = 0.4 x 0.01 + 0.6 x 0.3 = 0.184

P(home | rain) = P(home) x P(rain | home) / P(rain)
P(home | rain) = 0.4 x 0.01 / 0.184) = **0.0217**

Two Flips Quiz

```
def f(p):
   return (p * p)
```

Three Flips Quiz

```
def f(p):
   return 3 * p * (1-p) * (1-p)
```

Flip Two Coins Quiz

```
def f(p1,p2):
   return p1 * p2
```

Flip One of Two Quiz

$$P(H) = (P_0 \times P1) + ((1 - P_0) \times P2)$$

```
def f(p0,p1,p2):
    return (p0 * p1) + ((1 - p0) * p2)
```

Cancer Example Quiz

$$P(pos) = (P_1 \times P_0) + ((1 - P_2) \times (1 - P_0))$$

Calculate Total Quiz

```
def f(p0,p1,p2):
    return (p0 * p1) + ((1 - p0) * (1 - p2))
```

Program Bayes' Rule Quiz

$$P(C \mid pos) = P_0 \times P_1 / ((P_0 \times P_1) + ((1 - P_0) \times (1 - P_2))$$

```
def f(p0,p1,p2):
    return ((p0 * p1)/((p0 * p1) + ((1 - p0) * (1 - p2))))
```

Program Bayes' Rule Quiz 2

```
def f(p0,p1,p2):
    return (p0 * (1 - p1))/((p0 * (1 - p1)) + ((1 - p0) * p2))
```

Correlation Quiz

A. Yes
B. No
C. No
D. Yes