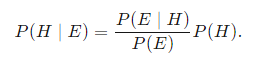
**Bayes' Theorem and Conditional Probability**

[**Andrew Ellinor**](https://brilliant.org/profile/andrew-ujvet5/), [**Christopher Williams**](https://brilliant.org/profile/christopher-pb1ay7/), [**Adam Strandberg**](https://brilliant.org/profile/adam-hcz865/), and

14 others

contributed

**Bayes' theorem** is a formula that describes how to update the probabilities of hypotheses when given evidence. It follows simply from the axioms of [conditional probability](https://brilliant.org/wiki/conditional-probability-distribution/), but can be used to powerfully reason about a wide range of problems involving belief updates.

Given a hypothesis *H* and evidence *E*, Bayes' theorem states that the relationship between the probability of the hypothesis before getting the evidence *P*(*H*) and the probability of the hypothesis after getting the evidence *P*(*H*∣*E*) is

Many modern [machine learning](https://brilliant.org/wiki/machine-learning/) techniques rely on Bayes' theorem. For instance, spam filters use Bayesian updating to determine whether an email is real or spam, given the words in the email. Additionally, many specific techniques in statistics, such as calculating [*p*-values](https://brilliant.org/wiki/statistical-significance/) or [interpreting medical results](https://brilliant.org/wiki/bayesian-theory-in-science-and-math/), are best described in terms of how they contribute to updating hypotheses using Bayes' theorem.

**Contents**

Explaining Counterintuitive Results

Deriving Bayes' Theorem

Visualizing Bayes’ Theorem

Diagnosing Disease

More Examples

## **Explaining Counterintuitive Results**

Probability problems are notorious for yielding surprising and counterintuitive results. One famous example--or a pair of examples--is the following:

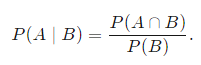
1. A couple has 2 children and the older child is a boy. If the probabilities of having a boy or a girl are both 50%, what's the probability that the couple has two boys?   
   We already know that the older child is a boy. The probability of two boys is equivalent to the probability that the younger child is a boy, which is 50\%50%.
2. A couple has two children, of which at least one is a boy. If the probabilities of having a boy or a girl are both 50\%50%, what is the probability that the couple has two boys?

At first glance, this appears to be asking the same question. We might reason as follows: “We know that one is a boy, so the only question is whether the other one is a boy, and the chances of that being the case are 50\%50%. So again, the answer is 50\%50%.”

This makes perfect sense. It also happens to be incorrect.

## **Deriving Bayes' Theorem**

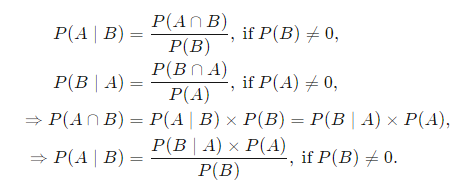
Bayes' theorem centers on relating different [conditional probabilities](https://brilliant.org/wiki/conditional-probability-distribution/). A conditional probability is an expression of how probable one event is given that some other event occurred (a fixed value). For instance, "what is the probability that the sidewalk is wet?" will have a different answer than "what is the probability that the sidewalk is wet given that it rained earlier?"

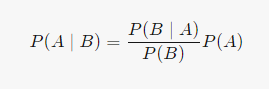
For a [joint probability distribution](https://brilliant.org/wiki/discrete-random-variables-joint-probability/) over events *A* and *B*,  *P*(*A*∩*B*), the conditional probability of *A* given *B* is defined as

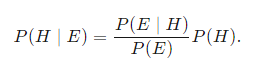
In the sidewalk example, where *A* is "the sidewalk is wet" and *B* is "it rained earlier," this expression reads as "the probability the sidewalk is wet given that it rained earlier is equal to the probability that the sidewalk is wet and it rains over the probability that it rains."

Note that  *P*(*A*∩*B*) is the probability of both *A* and *B* occurring, which is the same as the probability of *A* occurring times the probability that *B* occurs given that *A* occurred: P(A).*P*(*B*∣*A*)×*P*(*A*). Using the same reasoning, *P*(*A*∩*B*) is also the probability that *B* occurs times the probability that *A* occurs given that *B* occurs: *P*(*A*∣*B*)×*P*(*B*). The fact that these two expressions are equal leads to Bayes' Theorem. Expressed mathematically, this is:

Notice that our result for dependent events and for Bayes’ theorem are both valid when the events are independent. In these instances, *P*(*A*∣*B*)=*P*(*A*) and *P*(*B*∣*A*)=*P*(*B*), so the expressions simplify.

Notice that our result for dependent events and for Bayes’ theorem are both valid when the events are independent. In these instances,  *P*(*A*∣*B*)=*P*(*A*) and  *P*(*B*∣*A*)=*P*(*B*), so the expressions simplify.

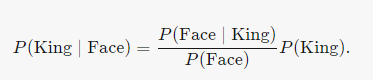
**Bayes' Theorem**

While this is an equation that applies to any probability distribution over events *A* and *B*, it has a particularly nice interpretation in the case where A*A* represents a hypothesis *H* and *B* represents some observed evidence *E*. In this case, the formula can be written as

This relates the probability of the hypothesis before getting the evidence *P*(*H*), to the probability of the hypothesis after getting the evidence, *P*(*H*∣*E*). For this reason, *P*(*H*) is called the **prior probability**, while  *P*(*H*∣*E*) is called the **posterior probability**. The factor that relates the two, ​, is called the **likelihood ratio**. Using these terms, Bayes' theorem can be rephrased as "the posterior probability equals the prior probability times the likelihood ratio."

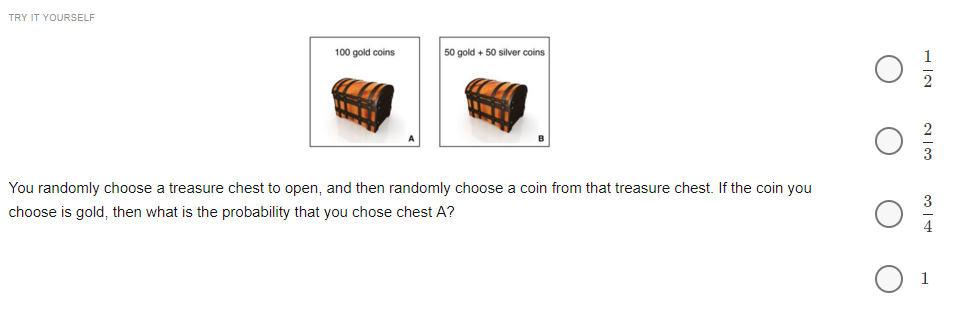
Example:

If a single card is drawn from a standard deck of playing cards, the probability that the card is a king is 4/52, since there are 4 kings in a standard deck of 52 cards. Rewording this, if King is the event "this card is a king," the prior probability P(King) 

If evidence is provided (for instance, someone looks at the card) that the single card is a face card, then the posterior probability *P*(King∣Face) can be calculated using Bayes' theorem:

Since every King is also a face card, *P*(Face∣King)=1. Since there are 3 face cards in each suit (Jack, Queen, King) , the probability of a face card is *P*(Face)=3/13​. Combining these gives a likelihood ratio of 

Using Bayes' theorem gives 



Bayes' theorem clarifies the two-children problem from the first section:

### **1. A couple has two children, the older of which is a boy. What is the probability that they have two boys?**

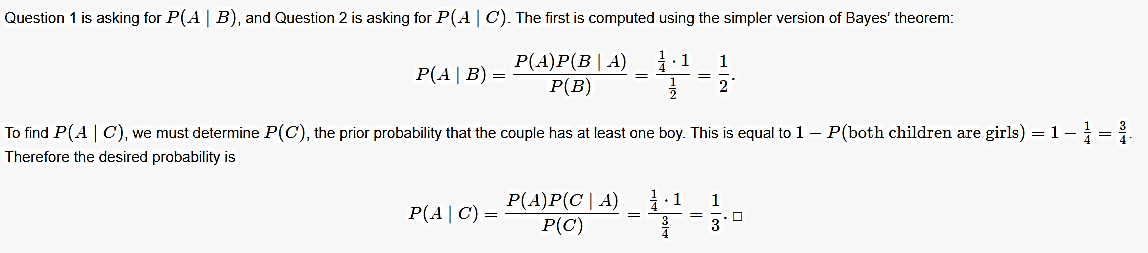
### **2. A couple has two children, one of which is a boy. What is the probability that they have two boys?**

Define three events, *A*, *B*, and *C*, as follows:

A = both children are boys

B = the older child is a boy

C = one of their children is a boy

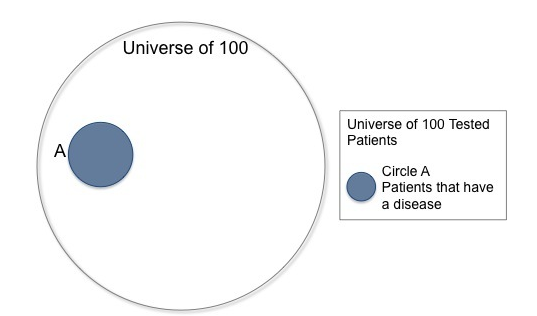


## **Visualizing Bayes’ Theorem**

Venn diagrams are particularly useful for visualizing Bayes' theorem, since both the diagrams and the theorem are about looking at the intersections of different spaces of events.

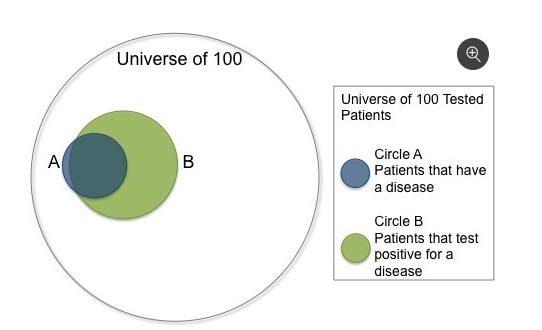
A disease is present in 5 out of 100 people, and a test that is 90% accurate (meaning that the test produces the correct result in 90% of cases) is administered to 100 people. If one person in the group tests positive, what is the probability that this one person has the disease?

The intuitive answer is that the one person is 90% likely to have the disease. But we can visualize this to show that it’s not accurate. First, draw the total population and the 5 people who have the disease:



The circle A represents 5 out 100, or 5% of the larger universe of 100 people.

Next, overlay a circle to represent the people who get a positive result on the test. We know that 90% of those with the disease will get a positive result, so need to cover 90% of circle A, but we also know that 10% of the population who does not have the disease will get a positive result, so we need to cover 10% of the non-disease carrying population (the total universe of 100 less circle A).



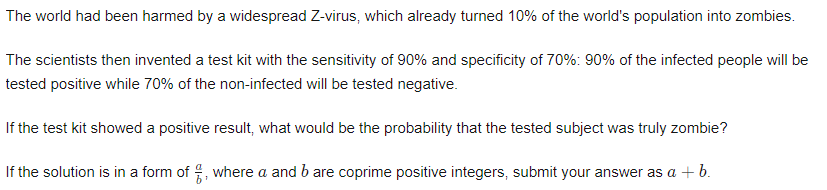
Circle B is covering a substantial portion of the total population. It actually covers more area than the total portion of the population with the disease. This is because 14 out of the total population of 100 (90% of the 5 people with the disease + 10% of the 95 people without the disease) will receive a positive result. Even though this is a test with 90% accuracy, this visualization shows that any one patient who tests positive (Circle B) for the disease only has a 32.14% (4.5 in 14) chance of actually having the disease.

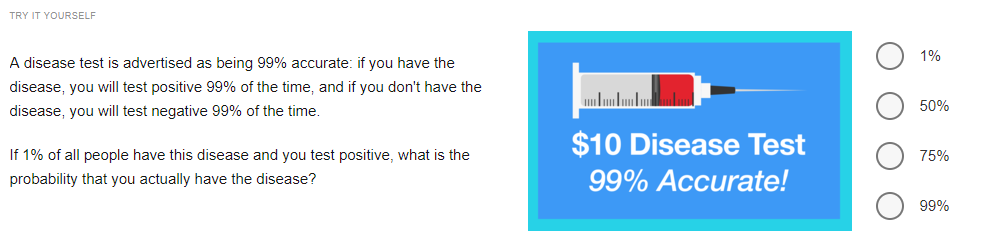
## **Diagnosing Disease**

*Main article:*[*Bayesian theory in science and math*](https://brilliant.org/wiki/bayesian-theory-in-science-and-math/)

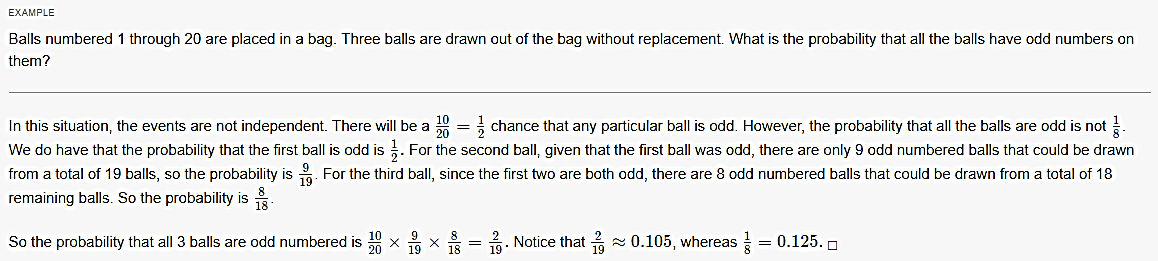
Bayes’ theorem can show the likelihood of getting false positives in scientific studies. An in-depth look at this can be found in [Bayesian theory in science and math](https://brilliant.org/wiki/bayesian-theory-in-science-and-math/).

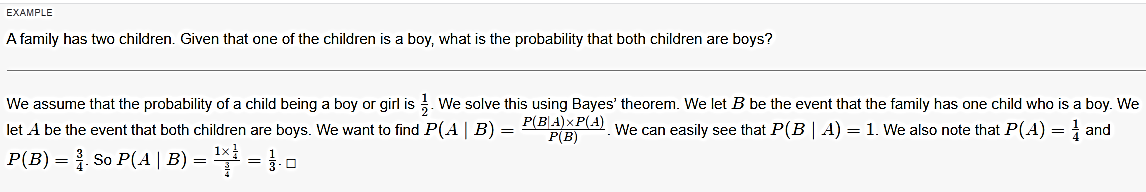
Many medical diagnostic tests are said to be *X*% accurate, for instance 99% accurate, referring specifically to the probability that the test result is correct given your condition (or lack thereof). This is not the same as the posterior probability of having the disease given the result of the test. To see this in action, consider the following problem.

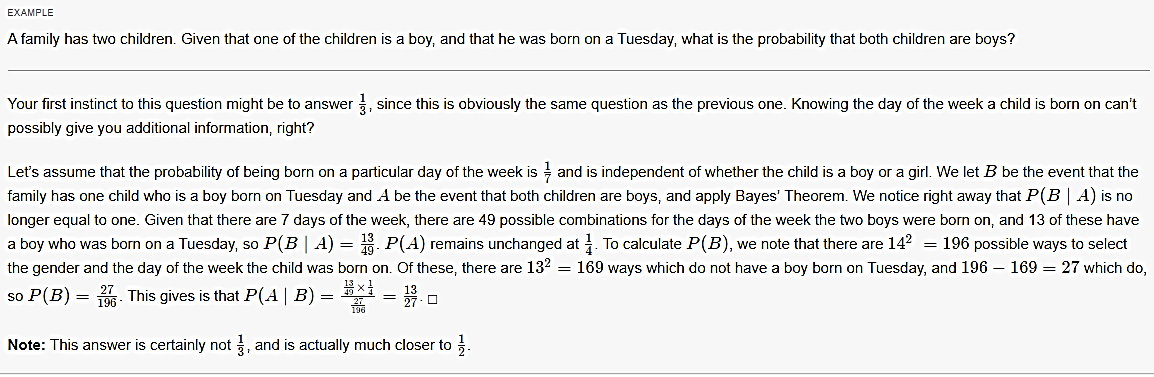


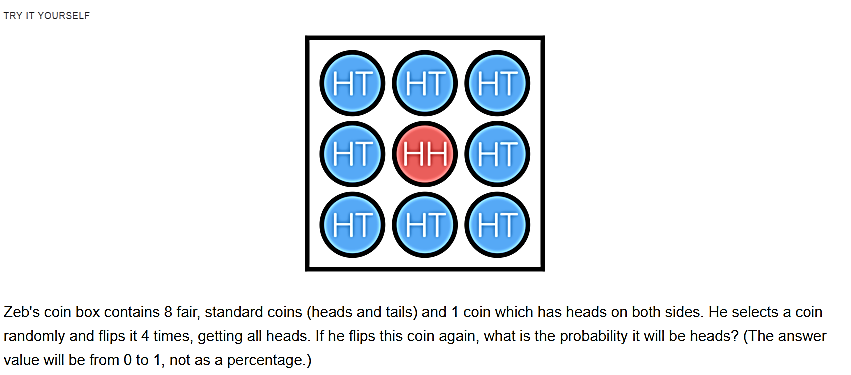


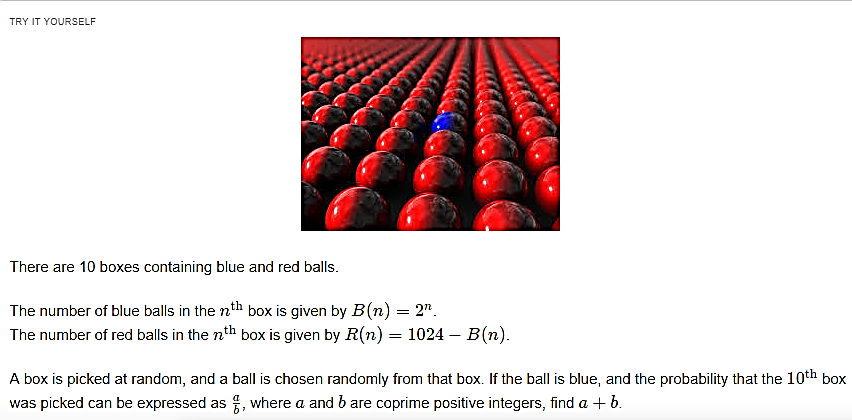
More Examples











Taken from <https://brilliant.org/wiki/bayes-theorem/>