

[Toggle the table of contents](#)

# Bayes' theorem

In probability theory and statistics, **Bayes' theorem** (alternatively **Bayes' law** or **Bayes' rule**), named after Thomas Bayes, describes the probability of an event, based on prior knowledge of conditions that might be related to the event.<sup>[1]</sup> For example, if the risk of developing health problems is known to increase with age, Bayes' theorem allows the risk to an individual of a known age to be assessed more accurately by conditioning it relative to their age, rather than simply assuming that the individual is typical of the population as a whole.

One of the many applications of Bayes' theorem is Bayesian inference, a particular approach to statistical inference. When applied, the probabilities involved in the theorem may have different probability interpretations. With Bayesian probability interpretation, the theorem expresses how a degree of belief, expressed as a probability, should rationally change to account for the availability of related evidence. Bayesian inference is fundamental to Bayesian statistics, being considered by one authority as; "to the theory of probability what Pythagoras's theorem is to geometry."<sup>[2]</sup>

## History

Bayes' theorem is named after the Reverend Thomas Bayes (/beɪz/), also a statistician and philosopher. Bayes used conditional probability to provide an algorithm (his Proposition 9) that uses evidence to calculate limits on an unknown parameter. His work was published in 1763 as *An Essay towards solving a Problem in the Doctrine of Chances*. Bayes studied how to compute a distribution for the probability parameter of a binomial distribution (in modern terminology). On Bayes's death his family transferred his papers to a friend, the minister, philosopher, and mathematician Richard Price.

Over two years, Richard Price significantly edited the unpublished manuscript, before sending it to a friend who read it aloud at the Royal Society on 23 December 1763.<sup>[3]</sup> Price edited<sup>[4]</sup> Bayes's major work "An Essay towards solving a Problem in the Doctrine of Chances" (1763), which appeared in *Philosophical Transactions*,<sup>[5]</sup> and contains Bayes' theorem. Price wrote an introduction to the paper which provides some of the philosophical basis of Bayesian statistics and chose one of the two solutions offered by Bayes. In 1765, Price was elected a Fellow of the Royal Society in recognition of his work on the legacy of Bayes.<sup>[6][7]</sup> On 27 April a letter sent to his friend Benjamin Franklin was read out at the Royal Society, and later published, where Price applies this work to population and computing 'life-annuities'.<sup>[8]</sup>

Independently of Bayes, Pierre-Simon Laplace in 1774, and later in his 1812 *Théorie analytique des probabilités*, used conditional probability to formulate the relation of an updated posterior probability from a prior probability, given evidence. He reproduced and extended Bayes's results in 1774, apparently unaware of Bayes's work.<sup>[note 1][9]</sup> The Bayesian interpretation of probability was developed mainly by Laplace.<sup>[10]</sup>

About 200 years later, Sir Harold Jeffreys put Bayes's algorithm and Laplace's formulation on an axiomatic basis, writing in a 1973 book that Bayes' theorem "is to the theory of probability what the Pythagorean theorem is to geometry".<sup>[11]</sup>

Stephen Stigler used a Bayesian argument to conclude that Bayes' theorem was discovered by Nicholas Saunderson, a blind English mathematician, some time before Bayes;<sup>[12][13]</sup> that interpretation, however, has been disputed.<sup>[14]</sup> Martyn Hooper<sup>[15]</sup> and Sharon McGrayne<sup>[16]</sup> have argued that Richard Price's contribution was substantial:

By modern standards, we should refer to the Bayes–Price rule. Price discovered Bayes's work, recognized its importance, corrected it, contributed to the article, and found a use for it. The modern convention of employing Bayes's name alone is unfair but so entrenched that anything else makes little sense.<sup>[16]</sup>

## Statement of theorem

Bayes' theorem is stated mathematically as the following equation:<sup>[17]</sup>

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

where **A** and **B** are events and  $P(B) \neq 0$ .

- $P(A|B)$  is a conditional probability: the probability of event **A** occurring given that **B** is true. It is also called the posterior probability of **A** given **B**.
- $P(B|A)$  is also a conditional probability: the probability of event **B** occurring given that **A** is true. It can also be interpreted as the likelihood of **A** given a fixed **B** because  $P(B|A) = L(A|B)$ .
- $P(A)$  and  $P(B)$  are the probabilities of observing **A** and **B** respectively without any given conditions; they are known as the prior probability and marginal probability.

## Proof

### For events

Bayes' theorem may be derived from the definition of conditional probability:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ if } P(B) \neq 0,$$

where  $P(A \cap B)$  is the probability of both A and B being true. Similarly,

$$P(B|A) = \frac{P(A \cap B)}{P(A)}, \text{ if } P(A) \neq 0.$$

Solving for  $P(A \cap B)$  and substituting into the above expression for  $P(A|B)$  yields Bayes' theorem:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}, \text{ if } P(B) \neq 0.$$

### For continuous random variables

For two continuous random variables  $X$  and  $Y$ , Bayes' theorem may be analogously derived from the definition of conditional density:

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$$

$$f_{Y|X=x}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

Therefore,

$$f_{X|Y=y}(x) = \frac{f_{Y|X=x}(y)f_X(x)}{f_Y(y)}.$$

## General case

Let  $P_Y^x$  be the conditional distribution of  $Y$  given  $X = x$  and let  $P_X$  be the distribution of  $X$ . The joint distribution is then  $P_{X,Y}(dx, dy) = P_Y^x(dy)P_X(dx)$ . The conditional distribution  $P_X^y$  of  $X$  given  $Y = y$  is then determined by

$$P_X^y(A) = E(1_A(X)|Y = y)$$

Existence and uniqueness of the needed conditional expectation is a consequence of the Radon–Nikodym theorem. This was formulated by Kolmogorov in his famous book from 1933. Kolmogorov underlines the importance of conditional probability by writing "I wish to call attention to ... and especially the theory of conditional probabilities and conditional expectations ..." in the Preface.<sup>[18]</sup> The Bayes theorem determines the posterior distribution from the prior distribution. Bayes' theorem can be generalized to include improper prior distributions such as the uniform distribution on the real line.<sup>[19]</sup> Modern Markov chain Monte Carlo methods have boosted the importance of Bayes' theorem including cases with improper priors.<sup>[20]</sup>

## Examples

### Recreational mathematics

Bayes' rule and computing conditional probabilities provide a solution method for a number of popular puzzles, such as the Three Prisoners problem, the Monty Hall problem, the Two Child problem and the Two Envelopes problem.

### Drug testing

Suppose, a particular test for whether someone has been using cannabis is 90% sensitive, meaning the true positive rate (TPR) = 0.90. Therefore, it leads to 90% true positive results (correct identification of drug use) for cannabis users.

The test is also 80% specific, meaning true negative rate (TNR) = 0.80. Therefore, the test correctly identifies 80% of non-use for non-users, but also generates 20% false positives, or false positive rate (FPR) = 0.20, for non-users.

Assuming 0.05 prevalence, meaning 5% of people use cannabis, what is the probability that a random person who tests positive is really a cannabis user?

The Positive predictive value (PPV) of a test is the proportion of persons who are actually positive out of all those testing positive, and can be calculated from a sample as:

$$\text{PPV} = \text{True positive} / \text{Tested positive}$$

If sensitivity, specificity, and prevalence are known, PPV can be calculated using Bayes theorem. Let  $P(\text{User}|\text{Positive})$  mean "the probability that someone is a cannabis user given that they test positive," which is what is meant by PPV. We can write:



Figure 1: Using a frequency box to show  $P(\text{User}|\text{Positive})$  visually by comparison of shaded areas

$$\begin{aligned} P(\text{User}|\text{Positive}) &= \frac{P(\text{Positive}|\text{User})P(\text{User})}{P(\text{Positive})} \\ &= \frac{P(\text{Positive}|\text{User})P(\text{User})}{P(\text{Positive}|\text{User})P(\text{User}) + P(\text{Positive}|\text{Non-user})P(\text{Non-user})} \\ &= \frac{0.90 \times 0.05}{0.90 \times 0.05 + 0.20 \times 0.95} = \frac{0.045}{0.045 + 0.19} \approx 19\% \end{aligned}$$

The fact that  $P(\text{Positive}) = P(\text{Positive}|\text{User})P(\text{User}) + P(\text{Positive}|\text{Non-user})P(\text{Non-user})$  is a direct application of the Law of Total Probability. In this case, it says that the probability that someone tests positive is the probability that a user tests positive, times the probability of being a user, plus the probability that a non-user tests positive, times the probability of being a non-user. This is true because the classifications user and non-user form a partition of a set, namely the set of people who take the drug test. This combined with the definition of conditional probability results in the above statement.

In other words, even if someone tests positive, the probability that they are a cannabis user is only 19%—this is because in this group, only 5% of people are users, and most positives are false positives coming from the remaining 95%.

If 1,000 people were tested:

- 950 are non-users and 190 of them give false positive (0.20 × 950)
- 50 of them are users and 45 of them give true positive (0.90 × 50)

The 1,000 people thus yields 235 positive tests, of which only 45 are genuine drug users, about 19%. See Figure 1 for an illustration using a frequency box, and note how small the pink area of true positives is compared to the blue area of false positives.

Sensitivity or specificity

The importance of specificity can be seen by showing that even if sensitivity is raised to 100% and specificity remains at 80%, the probability of someone testing positive really being a cannabis user only rises from 19% to 21%, but if the sensitivity is held at 90% and the specificity is increased to 95%, the probability rises to 49%.

Actual	Test Positive	Test Negative	Total	Actual	Test Positive	Test Negative	Total
User	45	5	50	User	50	0	50
Non-user	190	760	950	Non-user	190	760	950
Total	235	765	1000	Total	240	760	1000
90% sensitive, 80% specific, PPV=45/235 ≈ 19%				100% sensitive, 80% specific, PPV=50/240 ≈ 21%			

Actual	Test Positive	Test Negative	Total
User	45	5	50
Non-user	47	903	950
Total	92	908	1000
90% sensitive, 95% specific, PPV=45/92 ≈ 49%			

Cancer rate

Even if 100% of patients with pancreatic cancer have a certain symptom, when someone has the same symptom, it does not mean that this person has a 100% chance of getting pancreatic cancer. Assuming the incidence rate of pancreatic cancer is 1/100000, while 10/99999 healthy individuals have the same symptoms worldwide, the probability of having pancreatic cancer given the symptoms is only 9.1%, and the other 90.9% could be "false positives" (that is, falsely said to have cancer; "positive" is a confusing term when, as here, the test gives bad news).

Based on incidence rate, the following table presents the corresponding numbers per 100,000 people.

Symptom Cancer	Yes	No	Total
Yes	1	0	1
No	10	99989	99999
Total	11	99989	100000

Which can then be used to calculate the probability of having cancer when you have the symptoms:

$$\begin{aligned}
 P(\text{Cancer}|\text{Symptoms}) &= \frac{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer})}{P(\text{Symptoms})} \\
 &= \frac{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer})}{P(\text{Symptoms}|\text{Cancer})P(\text{Cancer}) + P(\text{Symptoms}|\text{Non-Cancer})P(\text{Non-Cancer})} \\
 &= \frac{1 \times 0.00001}{1 \times 0.00001 + (10/99999) \times 0.99999} = \frac{1}{11} \approx 9.1\%
 \end{aligned}$$

## Defective item rate

A factory produces items using three machines—A, B, and C—which account for 20%, 30%, and 50% of its output respectively. Of the items produced by machine A, 5% are defective; similarly, 3% of machine B's items and 1% of machine C's are defective. If a randomly selected item is defective, what is the probability it was produced by machine C?

Condition Machine	Defective	Flawless	Total
A	10	190	200
B	9	291	300
C	5	495	500
Total	24	976	1000

Once again, the answer can be reached without using the formula by applying the conditions to a hypothetical number of cases. For example, if the factory produces 1,000 items, 200 will be produced by Machine A, 300 by Machine B, and 500 by Machine C. Machine A will produce  $5\% \times 200 = 10$  defective items, Machine B  $3\% \times 300 = 9$ , and Machine C  $1\% \times 500 = 5$ , for a total of 24. Thus, the likelihood that a randomly selected defective item was produced by machine C is  $5/24$  ( $\sim 20.83\%$ ).

This problem can also be solved using Bayes' theorem: Let  $X_i$  denote the event that a randomly chosen item was made by the  $i^{\text{th}}$  machine (for  $i = A, B, C$ ). Let  $Y$  denote the event that a randomly chosen item is defective. Then, we are given the following information:

$$P(X_A) = 0.2, \quad P(X_B) = 0.3, \quad P(X_C) = 0.5.$$

If the item was made by the first machine, then the probability that it is defective is 0.05; that is,  $P(Y|X_A) = 0.05$ . Overall, we have

$$P(Y|X_A) = 0.05, \quad P(Y|X_B) = 0.03, \quad P(Y|X_C) = 0.01.$$

To answer the original question, we first find  $P(Y)$ . That can be done in the following way:

$$P(Y) = \sum_i P(Y|X_i)P(X_i) = (0.05)(0.2) + (0.03)(0.3) + (0.01)(0.5) = 0.024.$$

Hence, 2.4% of the total output is defective.

We are given that  $Y$  has occurred, and we want to calculate the conditional probability of  $X_C$ . By Bayes' theorem,

$$P(X_C|Y) = \frac{P(Y|X_C)P(X_C)}{P(Y)} = \frac{0.01 \cdot 0.50}{0.024} = \frac{5}{24}$$

Given that the item is defective, the probability that it was made by machine C is  $5/24$ . Although machine C produces half of the total output, it produces a much smaller fraction of the defective items. Hence the knowledge that the item selected was defective enables us to replace the prior probability  $P(X_C) = 1/2$  by the smaller posterior probability  $P(X_C | Y) = 5/24$ .

## Interpretations

The interpretation of Bayes' rule depends on the interpretation of probability ascribed to the terms. The two predominant interpretations are described below. Figure 2 shows a geometric visualization.

### Bayesian interpretation

In the Bayesian (or epistemological) interpretation, probability measures a "degree of belief". Bayes' theorem links the degree of belief in a proposition before and after accounting for evidence. For example, suppose it is believed with 50% certainty that a coin is twice as likely to land heads than tails. If the coin is flipped a number of times and the outcomes observed, that degree of belief will probably rise or fall, but might even remain the same, depending on the results. For proposition  $A$  and evidence  $B$ ,

- $P(A)$ , the *prior*, is the initial degree of belief in  $A$ .
- $P(A|B)$ , the *posterior*, is the degree of belief after incorporating news that  $B$  is true.
- the quotient  $\frac{P(B|A)}{P(B)}$  represents the support  $B$  provides for  $A$ .

For more on the application of Bayes' theorem under the Bayesian interpretation of probability, see Bayesian inference.

### Frequentist interpretation

In the frequentist interpretation, probability measures a "proportion of outcomes". For example, suppose an experiment is performed many times.  $P(A)$  is the proportion of outcomes with property  $A$  (the prior) and  $P(B)$  is the proportion with property  $B$ .  $P(B|A)$  is the proportion of outcomes with property  $B$  *out of* outcomes with property  $A$ , and  $P(A|B)$  is the proportion of those with  $A$  *out of* those with  $B$  (the posterior).

The role of Bayes' theorem is best visualized with tree diagrams such as Figure 3. The two diagrams partition the same outcomes by  $A$  and  $B$  in opposite orders, to obtain the inverse probabilities. Bayes' theorem links the different partitionings.

### Example

An entomologist spots what might, due to the pattern on its back, be a rare subspecies of beetle. A full 98% of the members of the rare subspecies have the pattern, so  $P(\text{Pattern} | \text{Rare}) = 98\%$ . Only 5% of members of the common subspecies have the pattern. The rare subspecies is 0.1% of the total population. How likely is the beetle having the pattern to be rare: what is  $P(\text{Rare} | \text{Pattern})$ ?

From the extended form of Bayes' theorem (since any beetle is either rare or common),

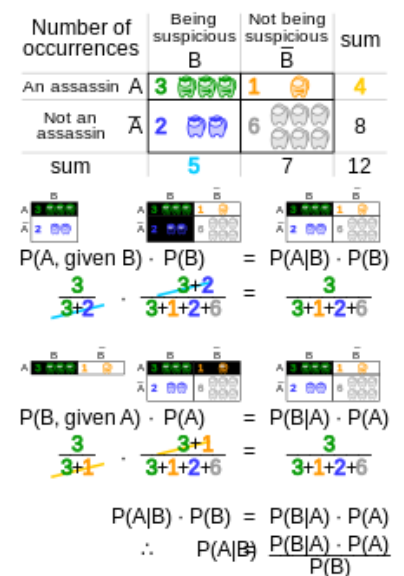


Figure 2: A geometric visualisation of Bayes' theorem

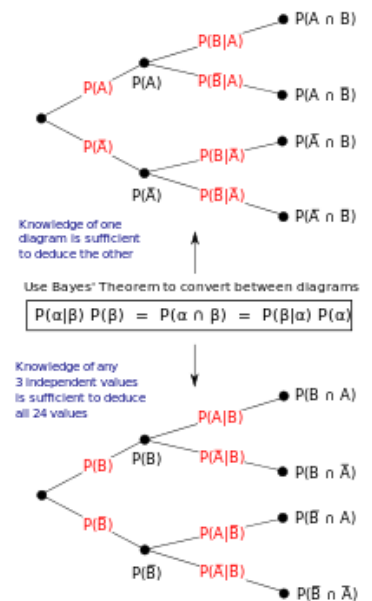


Figure 3: Illustration of frequentist interpretation with tree diagrams

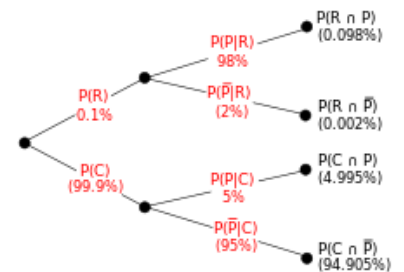


Figure 4: Tree diagram illustrating the beetle example.  $R$ ,  $C$ ,  $P$  and  $\bar{P}$  are the events rare, common, pattern and no pattern. Percentages in parentheses are calculated. Three independent values are given, so it is possible to calculate the inverse tree.

$$\begin{aligned}
 P(\text{Rare}|\text{Pattern}) &= \frac{P(\text{Pattern}|\text{Rare})P(\text{Rare})}{P(\text{Pattern})} \\
 &= \frac{P(\text{Pattern}|\text{Rare})P(\text{Rare})}{P(\text{Pattern}|\text{Rare})P(\text{Rare}) + P(\text{Pattern}|\text{Common})P(\text{Common})} \\
 &= \frac{0.98 \times 0.001}{0.98 \times 0.001 + 0.05 \times 0.999} \\
 &\approx 1.9\%
 \end{aligned}$$

## Forms

### Events

## Simple form

For events  $A$  and  $B$ , provided that  $P(B) \neq 0$ ,

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

In many applications, for instance in Bayesian inference, the event  $B$  is fixed in the discussion, and we wish to consider the impact of its having been observed on our belief in various possible events  $A$ . In such a situation the denominator of the last expression, the probability of the given evidence  $B$ , is fixed; what we want to vary is  $A$ . Bayes' theorem then shows that the posterior probabilities are proportional to the numerator, so the last equation becomes:

$$P(A|B) \propto P(A) \cdot P(B|A).$$

In words, the posterior is proportional to the prior times the likelihood.<sup>[21]</sup>

If events  $A_1, A_2, \dots$ , are mutually exclusive and exhaustive, i.e., one of them is certain to occur but no two can occur together, we can determine the proportionality constant by using the fact that their probabilities must add up to one. For instance, for a given event  $A$ , the event  $A$  itself and its complement  $\neg A$  are exclusive and exhaustive. Denoting the constant of proportionality by  $c$  we have

$$P(A|B) = c \cdot P(A) \cdot P(B|A) \text{ and } P(\neg A|B) = c \cdot P(\neg A) \cdot P(B|\neg A).$$

Adding these two formulas we deduce that

$$1 = c \cdot (P(B|A) \cdot P(A) + P(B|\neg A) \cdot P(\neg A)),$$

or

$$c = \frac{1}{P(B|A) \cdot P(A) + P(B|\neg A) \cdot P(\neg A)} = \frac{1}{P(B)}.$$

## Alternative form

Another form of Bayes' theorem for two competing statements or hypotheses is:

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

For an epistemological interpretation:

For proposition  $A$  and evidence or background  $B$ ,<sup>[22]</sup>

Contingency table

Background Proposition	B	$\neg B$ (not B)	Total
A	$P(B A) \cdot P(A)$ = $P(A B) \cdot P(B)$	$P(\neg B A) \cdot P(A)$ = $P(A \neg B) \cdot P(\neg B)$	$P(A)$
$\neg A$ (not A)	$P(B \neg A) \cdot P(\neg A)$ = $P(\neg A B) \cdot P(B)$	$P(\neg B \neg A) \cdot P(\neg A)$ = $P(\neg A \neg B) \cdot P(\neg B)$	$P(\neg A) = 1 - P(A)$
Total	$P(B)$	$P(\neg B) = 1 - P(B)$	1

- $P(A)$  is the prior probability, the initial degree of belief in  $A$ .
- $P(\neg A)$  is the corresponding initial degree of belief in *not-A*, that  $A$  is false, where  $P(\neg A) = 1 - P(A)$
- $P(B|A)$  is the conditional probability or likelihood, the degree of belief in  $B$  given that proposition  $A$  is true.
- $P(B|\neg A)$  is the conditional probability or likelihood, the degree of belief in  $B$  given that proposition  $A$  is false.
- $P(A|B)$  is the posterior probability, the probability of  $A$  after taking into account  $B$ .

## Extended form

Often, for some partition  $\{A_j\}$  of the sample space, the event space is given in terms of  $P(A_j)$  and  $P(B | A_j)$ . It is then useful to compute  $P(B)$  using the law of total probability:



$$P(B) = \sum_j P(B \cap A_j),$$

Or (using the multiplication rule for conditional probability),<sup>[23]</sup>

$$P(B) = \sum_j P(B|A_j)P(A_j),$$

$$\Rightarrow P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_j P(B|A_j)P(A_j)}.$$

In the special case where  $A$  is a binary variable:

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

## Random variables

Consider a sample space  $\Omega$  generated by two random variables  $X$  and  $Y$ . In principle, Bayes' theorem applies to the events  $A = \{X = x\}$  and  $B = \{Y = y\}$ .

$$P(X=x|Y=y) = \frac{P(Y=y|X=x)P(X=x)}{P(Y=y)}$$

However, terms become 0 at points where either variable has finite probability density. To remain useful, Bayes' theorem must be formulated in terms of the relevant densities (see Derivation).

### Simple form

If  $X$  is continuous and  $Y$  is discrete,

$$f_{X|Y=y}(x) = \frac{P(Y=y|X=x)f_X(x)}{P(Y=y)}$$

where each  $f$  is a density function.

If  $X$  is discrete and  $Y$  is continuous,

$$P(X=x|Y=y) = \frac{f_{Y|X=x}(y)P(X=x)}{f_Y(y)}.$$

If both  $X$  and  $Y$  are continuous,

$$f_{X|Y=y}(x) = \frac{f_{Y|X=x}(y)f_X(x)}{f_Y(y)}.$$

### Extended form

A continuous event space is often conceptualized in terms of the numerator terms. It is then useful to eliminate the denominator using the law of total probability. For  $f_Y(y)$ , this becomes an integral:

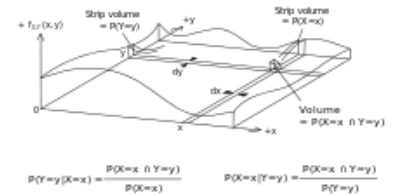


Figure 5: Bayes' theorem applied to an event space generated by continuous random variables  $X$  and  $Y$ . There exists an instance of Bayes' theorem for each point in the domain. In practice, these instances might be parametrized by writing the specified probability densities as a function of  $x$  and  $y$ .

$$f_Y(y) = \int_{-\infty}^{\infty} f_{Y|X=\xi}(y) f_X(\xi) d\xi.$$

## Bayes' rule in odds form

Bayes' theorem in odds form is:

$$O(A_1 : A_2|B) = O(A_1 : A_2) \cdot \Lambda(A_1 : A_2|B)$$

where

$$\Lambda(A_1 : A_2|B) = \frac{P(B|A_1)}{P(B|A_2)}$$

is called the Bayes factor or likelihood ratio. The odds between two events is simply the ratio of the probabilities of the two events. Thus

$$O(A_1 : A_2) = \frac{P(A_1)}{P(A_2)},$$

$$O(A_1 : A_2|B) = \frac{P(A_1|B)}{P(A_2|B)},$$

Thus, the rule says that the posterior odds are the prior odds times the Bayes factor, or in other words, the posterior is proportional to the prior times the likelihood.

In the special case that  $A_1 = A$  and  $A_2 = \neg A$ , one writes  $O(A) = O(A : \neg A) = P(A)/(1 - P(A))$ , and uses a similar abbreviation for the Bayes factor and for the conditional odds. The odds on  $A$  is by definition the odds for and against  $A$ . Bayes' rule can then be written in the abbreviated form

$$O(A|B) = O(A) \cdot \Lambda(A|B),$$

or, in words, the posterior odds on  $A$  equals the prior odds on  $A$  times the likelihood ratio for  $A$  given information  $B$ . In short, **posterior odds equals prior odds times likelihood ratio**.

For example, if a medical test has a sensitivity of 90% and a specificity of 91%, then the positive Bayes factor is  $\Lambda_+ = P(\text{True Positive})/P(\text{False Positive}) = 90\%/(100\% - 91\%) = 10$ . Now, if the prevalence of this disease is 9.09%, and if we take that as the prior probability, then the prior odds is about 1:10. So after receiving a positive test result, the posterior odds of actually having the disease becomes 1:1, which means that the posterior probability of having the disease is 50%. If a second test is performed in serial testing, and that also turns out to be positive, then the posterior odds of actually having the disease becomes 10:1, which means a posterior probability of about 90.91%. The negative Bayes factor can be calculated to be  $91\%/(100\% - 90\%) = 9.1$ , so if the second test turns out to be negative, then the posterior odds of actually having the disease is 1:9.1, which means a posterior probability of about 9.9%.

The example above can also be understood with more solid numbers: Assume the patient taking the test is from a group of 1000 people, where 91 of them actually have the disease (prevalence of 9.1%). If all these 1000 people take the medical test, 82 of those with the disease will get a true positive result (sensitivity of 90.1%), 9 of those with the disease will get a false negative result (false negative rate of 9.9%), 827 of those without the disease will get a true negative result (specificity of 91.0%), and 82 of those without the disease will get a false positive result (false positive rate of 9.0%). Before taking any test, the patient's odds for having the disease is 91:909. After receiving a positive result, the patient's odds for having the disease is

$$\frac{91}{909} \times \frac{90.1\%}{9.0\%} = \frac{91 \times 90.1\%}{909 \times 9.0\%} = 1 : 1$$

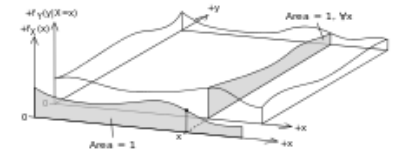


Figure 6: A way to conceptualize event spaces generated by continuous random variables  $X$  and  $Y$

which is consistent with the fact that there are 82 true positives and 82 false positives in the group of 1000 people.

## Correspondence to other mathematical frameworks

### Propositional logic

Using  $P(\neg B|A) = 1 - P(B|A)$  twice, one may use Bayes' theorem to also express  $P(\neg B|\neg A)$  in terms of  $P(A|B)$  and without negations:

$$P(\neg B|\neg A) = 1 - (1 - P(A|B)) \frac{P(B)}{P(\neg A)},$$

when  $P(\neg A) = 1 - P(A) \neq 0$ . From this we can read off the inference

$$P(A|B) = 1 \implies P(\neg B|\neg A) = 1.$$

In words: If certainly  $B$  implies  $A$ , we infer that certainly  $\neg A$  implies  $\neg B$ . Where  $P(B) \neq 0$ , the two implications being certain are equivalent statements. In the probability formulas, the conditional probability  $P(A|B)$  generalizes the logical implication  $B \implies A$ , where now beyond assigning true or false, we assign probability values to statements. The assertion of  $B \implies A$  is captured by certainty of the conditional, the assertion of  $P(A|B) = 1$ . Relating the directions of implication, Bayes' theorem represents a generalization of the contraposition law, which in classical propositional logic can be expressed as:

$$(B \implies A) \iff (\neg A \implies \neg B).$$

In this relation between implications, the positions of  $A$  resp.  $B$  get flipped.

The corresponding formula in terms of probability calculus is Bayes' theorem, which in its expanded form involving the prior probability/base rate  $a$  of only  $A$ , is expressed as:<sup>[24]</sup>

$$P(A|B) = P(B|A) \frac{a(A)}{P(B|A) a(A) + P(B|\neg A) a(\neg A)}.$$

### Subjective logic

Bayes' theorem represents a special case of deriving inverted conditional opinions in subjective logic expressed as:

$$(\omega_{A|B}^S, \omega_{A|\neg B}^S) = (\omega_{B|A}^S, \omega_{B|\neg A}^S) \tilde{\phi} a_A,$$

where  $\tilde{\phi}$  denotes the operator for inverting conditional opinions. The argument  $(\omega_{B|A}^S, \omega_{B|\neg A}^S)$  denotes a pair of binomial conditional opinions given by source  $S$ , and the argument  $a_A$  denotes the prior probability (aka. the base rate) of  $A$ . The pair of derivative inverted conditional opinions is denoted  $(\omega_{A|B}^S, \omega_{A|\neg B}^S)$ . The conditional opinion  $\omega_{A|B}^S$  generalizes the probabilistic conditional  $P(A|B)$ , i.e. in addition to assigning a probability the source  $S$  can assign any subjective opinion to the conditional statement  $(A|B)$ . A binomial subjective opinion  $\omega_A^S$  is the belief in the truth of statement  $A$  with degrees of epistemic uncertainty, as expressed by source  $S$ . Every subjective opinion has a corresponding projected probability  $P(\omega_A^S)$ . The application of Bayes' theorem to projected probabilities of opinions is a homomorphism, meaning that Bayes' theorem can be expressed in terms of projected probabilities of opinions:

$$P(\omega_{A|B}^S) = \frac{P(\omega_{B|A}^S) a(A)}{P(\omega_{B|A}^S) a(A) + P(\omega_{B|\neg A}^S) a(\neg A)}.$$

Hence, the subjective Bayes' theorem represents a generalization of Bayes' theorem.<sup>[25]</sup>

## Generalizations

### Bayes theorem for 3 events

A version of Bayes' theorem for 3 events<sup>[26]</sup> results from the addition of a third event *C* on which all probabilities are conditioned:

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

#### Derivation

Using the chain rule

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

And, on the other hand

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

The desired result is obtained by identifying both expressions and solving for *P(A|B ∩ C)*.

## Use in genetics

In genetics, Bayes' theorem can be used to calculate the probability of an individual having a specific genotype. Many people seek to approximate their chances of being affected by a genetic disease or their likelihood of being a carrier for a recessive gene of interest. A Bayesian analysis can be done based on family history or genetic testing, in order to predict whether an individual will develop a disease or pass one on to their children. Genetic testing and prediction is a common practice among couples who plan to have children but are concerned that they may both be carriers for a disease, especially within communities with low genetic variance.<sup>[27]</sup>

The first step in Bayesian analysis for genetics is to propose mutually exclusive hypotheses: for a specific allele, an individual either is or is not a carrier. Next, four probabilities are calculated: Prior Probability (the likelihood of each hypothesis considering information such as family history or predictions based on Mendelian Inheritance), Conditional Probability (of a certain outcome), Joint Probability (product of the first two), and Posterior Probability (a weighted product calculated by dividing the Joint Probability for each hypothesis by the sum of both joint probabilities). This type of analysis can be done based purely on family history of a condition or in concert with genetic testing.

### Using pedigree to calculate probabilities

Hypothesis	Hypothesis 1: Patient is a carrier	Hypothesis 2: Patient is not a carrier
Prior Probability	1/2	1/2
Conditional Probability that all four offspring will be unaffected	(1/2) · (1/2) · (1/2) · (1/2) = 1/16	About 1
Joint Probability	(1/2) · (1/16) = 1/32	(1/2) · 1 = 1/2
Posterior Probability	(1/32) / (1/32 + 1/2) = 1/17	(1/2) / (1/32 + 1/2) = 16/17

Example of a Bayesian analysis table for a female individual's risk for a disease based on the knowledge that the disease is present in her siblings but not in her parents or any of her four children. Based solely on the status of the subject's siblings and parents, she is equally likely to be a carrier as to be a non-carrier (this likelihood is denoted by

the Prior Hypothesis). However, the probability that the subject's four sons would all be unaffected is  $1/16$  ( $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2}$ ) if she is a carrier, about 1 if she is a non-carrier (this is the Conditional Probability). The Joint Probability reconciles these two predictions by multiplying them together. The last line (the Posterior Probability) is calculated by dividing the Joint Probability for each hypothesis by the sum of both joint probabilities.<sup>[28]</sup>

Using genetic test results

Parental genetic testing can detect around 90% of known disease alleles in parents that can lead to carrier or affected status in their child. Cystic fibrosis is a heritable disease caused by an autosomal recessive mutation on the CFTR gene,<sup>[29]</sup> located on the q arm of chromosome 7.<sup>[30]</sup>

Bayesian analysis of a female patient with a family history of cystic fibrosis (CF), who has tested negative for CF, demonstrating how this method was used to determine her risk of having a child born with CF:

Because the patient is unaffected, she is either homozygous for the wild-type allele, or heterozygous. To establish prior probabilities, a Punnett square is used, based on the knowledge that neither parent was affected by the disease but both could have been carriers:

Mother  Father	W  Homozygous for the wild-type allele (a non-carrier)	M  Heterozygous (a CF carrier)
	W  Homozygous for the wild-type allele (a non-carrier)	MW
M  Heterozygous (a CF carrier)	MW	MM  (affected by cystic fibrosis)

Given that the patient is unaffected, there are only three possibilities. Within these three, there are two scenarios in which the patient carries the mutant allele. Thus the prior probabilities are  $\frac{2}{3}$  and  $\frac{1}{3}$ .

Next, the patient undergoes genetic testing and tests negative for cystic fibrosis. This test has a 90% detection rate, so the conditional probabilities of a negative test are  $\frac{1}{10}$  and 1. Finally, the joint and posterior probabilities are calculated as before.

Hypothesis	Hypothesis 1: Patient is a carrier	Hypothesis 2: Patient is not a carrier
Prior Probability	$\frac{2}{3}$	$\frac{1}{3}$
Conditional Probability of a negative test	$\frac{1}{10}$	1
Joint Probability	$\frac{1}{15}$	$\frac{1}{3}$
Posterior Probability	$\frac{1}{6}$	$\frac{5}{6}$

After carrying out the same analysis on the patient's male partner (with a negative test result), the chances of their child being affected is equal to the product of the parents' respective posterior probabilities for being carriers times the chances that two carriers will produce an affected offspring ( $\frac{1}{4}$ ).

Genetic testing done in parallel with other risk factor identification

Bayesian analysis can be done using phenotypic information associated with a genetic condition, and when combined with genetic testing this analysis becomes much more complicated. Cystic fibrosis, for example, can be identified in a fetus through an ultrasound looking for an echogenic bowel, meaning one that appears brighter than normal on a scan. This is not a foolproof test, as an echogenic bowel can be present in a perfectly healthy fetus. Parental genetic

testing is very influential in this case, where a phenotypic facet can be overly influential in probability calculation. In the case of a fetus with an echogenic bowel, with a mother who has been tested and is known to be a CF carrier, the posterior probability that the fetus actually has the disease is very high (0.64). However, once the father has tested negative for CF, the posterior probability drops significantly (to 0.16).<sup>[28]</sup>

Risk factor calculation is a powerful tool in genetic counseling and reproductive planning, but it cannot be treated as the only important factor to consider. As above, incomplete testing can yield falsely high probability of carrier status, and testing can be financially inaccessible or unfeasible when a parent is not present.

## See also

- Bayesian epistemology
- Inductive probability
- Quantum Bayesianism
- Why Most Published Research Findings Are False*, a 2005 essay in *metascience* by John Ioannidis
- Regular conditional probability



## Notes

1. Laplace refined Bayes's theorem over a period of decades:

- Laplace announced his independent discovery of Bayes' theorem in: Laplace (1774) "Mémoire sur la probabilité des causes par les événements," "Mémoires de l'Académie royale des Sciences de MI (Savants étrangers)," 4: 621–656. Reprinted in: Laplace, "Oeuvres complètes" (Paris, France: Gauthier-Villars et fils, 1841), vol. 8, pp. 27–65. Available on-line at: Gallica (<http://gallica.bnf.fr/ark:/12148/bpt6k77596b/f32.image>). Bayes' theorem appears on p. 29.
- Laplace presented a refinement of Bayes' theorem in: Laplace (read: 1783 / published: 1785) "Mémoire sur les approximations des formules qui sont fonctions de très grands nombres," "Mémoires de l'Académie royale des Sciences de Paris," 423–467. Reprinted in: Laplace, "Oeuvres complètes" (Paris, France: Gauthier-Villars et fils, 1844), vol. 10, pp. 295–338. Available on-line at: Gallica (<http://gallica.bnf.fr/ark:/12148/bpt6k775981/f218.image.langEN>). Bayes' theorem is stated on page 301.
- See also: Laplace, "Essai philosophique sur les probabilités" (Paris, France: Mme. Ve. Courcier [Madame veuve (i.e., widow) Courcier], 1814), page 10 (<https://books.google.com/books?id=rDUJAAAAIAAJ&pg=PA10>). English translation: Pierre Simon, Marquis de Laplace with F. W. Truscott and F. L. Emory, trans., "A Philosophical Essay on Probabilities" (New York, New York: John Wiley & Sons, 1902), p. 15 (<https://google.com/books?id=WxoPAAAAIAAJ&pg=PA15#v=onepage>).

## References

- Joyce, James (2003), "Bayes' Theorem" (<https://plato.stanford.edu/archives/spr2019/entries/bayes-theorem/>), in Zalta, Edward N. (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 ed.), Metaphysics Research Lab, Stanford University, retrieved 2020-01-17
- Jeffreys, Sir Harold (1973). *Scientific Inference* (<https://www.worldcat.org/oclc/764571529>). Cambridge: At the University Press. OCLC 764571529 (<https://www.worldcat.org/oclc/764571529>).
- Frame, Paul (2015). *Liberty's Apostle* (<https://www.uwp.co.uk/book/libertys-apostle-richard-price-his-life-and-times/>). Wales: University of Wales Press. p. 44. ISBN 978-1783162161. Retrieved 23 February 2021.
- Allen, Richard (1999). *David Hartley on Human Nature* (<https://books.google.com/books?id=NCu6HhGIAB8C&pg=PA243>). SUNY Press. pp. 243–244. ISBN 978-0791494516. Retrieved 16 June 2013.
- Bayes, Thomas & Price, Richard (1763). "An Essay towards solving a Problem in the Doctrine of Chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S." (<https://doi.org/10.1098%2Frstl.1763.0053>) *Philosophical Transactions of the Royal Society of London*. **53**: 370–418. doi:10.1098/rstl.1763.0053 (<https://doi.org/10.1098%2Frstl.1763.0053>).
- Holland, pp. 46–7.
- Price, Richard (1991). *Price: Political Writings* (<https://books.google.com/books?id=xdH-gjy2vzUC&pg=PR23>). Cambridge University Press. p. xxiii. ISBN 978-0521409698. Retrieved 16 June 2013.
- Mitchell 1911, p. 314.

9. Daston, Lorraine (1988). *Classical Probability in the Enlightenment* (<https://books.google.com/books?id=oq8XNbKyUewC&pg=PA268>). Princeton Univ Press. p. 268. ISBN 0691084971.
10. Stigler, Stephen M. (1986). "Inverse Probability" (<https://books.google.com/books?id=M7yvkERHIIMC&pg=PA99>). *The History of Statistics: The Measurement of Uncertainty Before 1900*. Harvard University Press. pp. 99–138. ISBN 978-0674403413.
11. Jeffreys, Harold (1973). *Scientific Inference* (<https://archive.org/details/scientificinfere0000jeff>) (3rd ed.). Cambridge University Press. p. 31 (<https://archive.org/details/scientificinfere0000jeff/page/31>). ISBN 978-0521180788.
12. Stigler, Stephen M. (1983). "Who Discovered Bayes' Theorem?". *The American Statistician*. **37** (4): 290–296. doi:10.1080/00031305.1983.10483122 (<https://doi.org/10.1080%2F00031305.1983.10483122>).
13. de Vaux, Richard; Velleman, Paul; Bock, David (2016). *Stats, Data and Models* (4th ed.). Pearson. pp. 380–381. ISBN 978-0321986498.
14. Edwards, A. W. F. (1986). "Is the Reference in Hartley (1749) to Bayesian Inference?". *The American Statistician*. **40** (2): 109–110. doi:10.1080/00031305.1986.10475370 (<https://doi.org/10.1080%2F00031305.1986.10475370>).
15. Hooper, Martyn (2013). "Richard Price, Bayes' theorem, and God" (<https://doi.org/10.1111%2Fj.1740-9713.2013.00638.x>). *Significance*. **10** (1): 36–39. doi:10.1111/j.1740-9713.2013.00638.x (<https://doi.org/10.1111%2Fj.1740-9713.2013.00638.x>). S2CID 153704746 (<https://api.semanticscholar.org/CorpusID:153704746>).
16. McGrayne, S. B. (2011). *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines & Emerged Triumphant from Two Centuries of Controversy* (<https://archive.org/details/theorythatwouldn0000mcgr>). Yale University Press. ISBN 978-0300188226.
17. Stuart, A.; Ord, K. (1994), *Kendall's Advanced Theory of Statistics: Volume I – Distribution Theory*, Edward Arnold, §8.7
18. Kolmogorov, A.N. (1933) [1956]. *Foundations of the Theory of Probability*. Chelsea Publishing Company.
19. Taraldsen, Gunnar; Tufto, Jarle; Lindqvist, Bo H. (2021-07-24). "Improper priors and improper posteriors" (<https://doi.org/10.1111%2Fsjos.12550>). *Scandinavian Journal of Statistics*. **49** (3): 969–991. doi:10.1111/sjos.12550 (<https://doi.org/10.1111%2Fsjos.12550>). ISSN 0303-6898 (<https://www.worldcat.org/issn/0303-6898>). S2CID 237736986 (<https://api.semanticscholar.org/CorpusID:237736986>).
20. Robert, Christian P.; Casella, George (2004). *Monte Carlo Statistical Methods* (<http://worldcat.org/oclc/1159112760>). Springer. ISBN 978-1475741452. OCLC 1159112760 (<https://www.worldcat.org/oclc/1159112760>).
21. Lee, Peter M. (2012). "Chapter 1" (<http://www-users.york.ac.uk/~pml1/bayes/book.htm>). *Bayesian Statistics*. Wiley. ISBN 978-1-1183-3257-3.
22. "Bayes' Theorem: Introduction" (<https://web.archive.org/web/20040821012342/http://www.trinity.edu/cbrown/bayesweb/>). *Trinity University*. Archived from the original (<http://www.trinity.edu/cbrown/bayesweb/>) on 21 August 2004. Retrieved 5 August 2014.
23. "Bayes Theorem - Formula, Statement, Proof I Bayes Rule" (<https://www.cuemath.com/data/bayes-theorem/>). *Cuemath*. Retrieved 2023-10-20.
24. Audun Jøsang, 2016, *Subjective Logic; A formalism for Reasoning Under Uncertainty*. Springer, Cham, ISBN 978-3-319-42337-1
25. Audun Jøsang, 2016, *Generalising Bayes' Theorem in Subjective Logic* (<http://folk.uio.no/josang/papers/Josang2016-MFI.pdf>). IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI 2016), Baden-Baden, September 2016
26. Koller, D.; Friedman, N. (2009). *Probabilistic Graphical Models* (<https://web.archive.org/web/20140427083249/http://pgm.stanford.edu/>). Massachusetts: MIT Press. p. 1208. ISBN 978-0-262-01319-2. Archived from the original (<http://pgm.stanford.edu/>) on 2014-04-27.
27. Kraft, Stephanie A; Duenas, Devan; Wilfond, Benjamin S; Goddard, Katrina AB (24 September 2018). "The evolving landscape of expanded carrier screening: challenges and opportunities" (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6752283>). *Genetics in Medicine*. **21** (4): 790–797. doi:10.1038/s41436-018-0273-4 (<https://doi.org/10.1038/s41436-018-0273-4>). PMC 6752283 (<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6752283>). PMID 30245516 (<https://pubmed.ncbi.nlm.nih.gov/30245516>).
28. Ogino, Shuji; Wilson, Robert B; Gold, Bert; Hawley, Pamela; Grody, Wayne W (October 2004). "Bayesian analysis for cystic fibrosis risks in prenatal and carrier screening" (<https://doi.org/10.1097%2F01.GIM.0000139511.83336.8F>). *Genetics in Medicine*. **6** (5): 439–449. doi:10.1097/01.GIM.0000139511.83336.8F (<https://doi.org/10.1097%2F01.GIM.0000139511.83336.8F>). PMID 15371910 (<https://pubmed.ncbi.nlm.nih.gov/15371910>).
29. "Types of CFTR Mutations". Cystic Fibrosis Foundation, [www.cff.org/What-is-CF/Genetics/Types-of-CFTR-Mutations/](http://www.cff.org/What-is-CF/Genetics/Types-of-CFTR-Mutations/).

30. "CFTR Gene – Genetics Home Reference". U.S. National Library of Medicine, National Institutes of Health, [ghr.nlm.nih.gov/gene/CFTR#location](http://ghr.nlm.nih.gov/gene/CFTR#location).

## Bibliography

- This article incorporates text from a publication now in the public domain: Mitchell, John Malcolm (1911). "Price, Richard". In Chisholm, Hugh (ed.). *Encyclopædia Britannica*. Vol. 22 (11th ed.). Cambridge University Press. pp. 314–315.

## Further reading

- Grunau, Hans-Christoph (24 January 2014). "Preface Issue 3/4-2013" (<https://doi.org/10.1365%2Fs13291-013-0077-z>). *Jahresbericht der Deutschen Mathematiker-Vereinigung*. **115** (3–4): 127–128. doi:10.1365/s13291-013-0077-z (<https://doi.org/10.1365%2Fs13291-013-0077-z>).
- Gelman, A, Carlin, JB, Stern, HS, and Rubin, DB (2003), "Bayesian Data Analysis", Second Edition, CRC Press.
- Grinstead, CM and Snell, JL (1997), "Introduction to Probability (2nd edition)," American Mathematical Society (free pdf available) [1] ([https://chance.dartmouth.edu/teaching\\_aids/books\\_articles/probability\\_book/book.html](https://chance.dartmouth.edu/teaching_aids/books_articles/probability_book/book.html)).
- "Bayes formula" ([https://www.encyclopediaofmath.org/index.php?title=Bayes\\_formula](https://www.encyclopediaofmath.org/index.php?title=Bayes_formula)), *Encyclopedia of Mathematics*, EMS Press, 2001 [1994]
- McGrayne, SB (2011). *The Theory That Would Not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines & Emerged Triumphant from Two Centuries of Controversy* (<https://archive.org/details/theorythatwouldn0000mcgr>). Yale University Press. ISBN 978-0-300-18822-6.
- Laplace, Pierre Simon (1986). "Memoir on the Probability of the Causes of Events" (<https://doi.org/10.1214%2Fs0%2F1177013621>). *Statistical Science*. **1** (3): 364–378. doi:10.1214/ss/1177013621 (<https://doi.org/10.1214%2Fs0%2F1177013621>). JSTOR 2245476 (<https://www.jstor.org/stable/2245476>).
- Lee, Peter M (2012), *Bayesian Statistics: An Introduction*, 4th edition. Wiley. ISBN 978-1-118-33257-3.
- Puga JL, Krzywinski M, Altman N (31 March 2015). "Bayes' theorem" (<https://doi.org/10.1038%2Fnmeth.3335>). *Nature Methods*. **12** (4): 277–278. doi:10.1038/nmeth.3335 (<https://doi.org/10.1038%2Fnmeth.3335>). PMID 26005726 (<https://pubmed.ncbi.nlm.nih.gov/26005726>).
- Rosenthal, Jeffrey S (2005), "Struck by Lightning: The Curious World of Probabilities". HarperCollins. (Granta, 2008. ISBN 9781862079960).
- Stigler, Stephen M. (August 1986). "Laplace's 1774 Memoir on Inverse Probability" (<https://doi.org/10.1214%2Fs0%2F1177013620>). *Statistical Science*. **1** (3): 359–363. doi:10.1214/ss/1177013620 (<https://doi.org/10.1214%2Fs0%2F1177013620>).
- Stone, JV (2013), download chapter 1 of "Bayes' Rule: A Tutorial Introduction to Bayesian Analysis" ([http://jim-stone.staff.shef.ac.uk/BookBayes2012/HTML\\_BayesRulev5EbookHTMLFiles/ops/xhtml/ch01BayesJVSone.html](http://jim-stone.staff.shef.ac.uk/BookBayes2012/HTML_BayesRulev5EbookHTMLFiles/ops/xhtml/ch01BayesJVSone.html)), Sebtel Press, England.
- Bayesian Reasoning for Intelligent People* (<http://santafe.edu/~simon/br.pdf>), An introduction and tutorial to the use of Bayes' theorem in statistics and cognitive science.
- Morris, Dan (2016), Read first 6 chapters for free of "Bayes' Theorem Examples: A Visual Introduction For Beginners" (<https://web.archive.org/web/20190213034233/https://www.bayestheorem.net/>) Blue Windmill ISBN 978-1549761744. A short tutorial on how to understand problem scenarios and find P(B), P(A), and P(B|A).

## External links

- Visual explanation of Bayes using trees (<https://www.youtube.com/watch?v=Zxm4Xxvzohk>) on YouTube
- Bayes' frequentist interpretation explained visually (<https://www.youtube.com/watch?v=D8VZqxcu0I0>) on YouTube
- Earliest Known Uses of Some of the Words of Mathematics (B) (<http://jeff560.tripod.com/b.html>). Contains origins of "Bayesian", "Bayes' Theorem", "Bayes Estimate/Risk/Solution", "Empirical Bayes", and "Bayes Factor".
- A tutorial on probability and Bayes' theorem devised for Oxford University psychology students (<http://www.celiagreene.com/charlesmccreery/statistics/bayestutorial.pdf>)
- An Intuitive Explanation of Bayes' Theorem by Eliezer S. Yudkowsky (<http://yudkowsky.net/rational/bayes>)
- Bayesian Clinical Diagnostic Model (<https://kennis-research.shinyapps.io/Bayes-App/>)

Retrieved from "[https://en.wikipedia.org/w/index.php?title=Bayes%27\\_theorem&oldid=1181882054](https://en.wikipedia.org/w/index.php?title=Bayes%27_theorem&oldid=1181882054)"



■