



WIKIPEDIA  
The Free Encyclopedia

Main page  
Contents  
Featured content  
Current events  
Random article  
Donate to Wikipedia  
Wikipedia store

Interaction

Help  
About Wikipedia  
Community portal  
Recent changes  
Contact page

Tools

What links here  
Related changes  
Upload file  
Special pages  
Permanent link  
Page information  
Wikidata item  
Cite this page

Print/export

Create a book  
Download as PDF  
Printable version

Languages

العربية  
বাংলা  
Deutsch  
Español  
Français  
Italiano  
Nederlands  
日本語  
Русский

4 more

Edit links

Not logged in Talk Contributions Create account Log in

Article Talk

Read Edit View history

Search Wikipedia

# Bayesian statistics

From Wikipedia, the free encyclopedia



This article **needs additional citations for verification**. Please help [improve this article](#) by [adding citations to reliable sources](#). Unsourced material may be challenged and removed.

Find sources: "Bayesian statistics" – news · newspapers · books · scholar · JSTOR (May 2016) (*Learn how and when to remove this template message*)

**Bayesian statistics** is a theory in the field of [statistics](#) based on the [Bayesian interpretation of probability](#) where [probability](#) expresses a *degree of belief* in an [event](#). The degree of belief may be based on prior knowledge about the event, such as the results of previous experiments, or on personal beliefs about the event. This differs from a number of other [interpretations of probability](#), such as the [frequentist](#) interpretation that views probability as the [limit](#) of the relative frequency of an event after a large number of trials.<sup>[1]</sup>

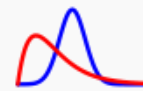
Bayesian statistical methods use [Bayes' theorem](#) to compute and update probabilities after obtaining new data. Bayes' theorem describes the [conditional probability](#) of an event based on data as well as prior information or beliefs about the event or conditions related to the event. For example, in [Bayesian inference](#), Bayes' theorem can be used to estimate the parameters of a [probability distribution](#) or [statistical model](#). Since Bayesian statistics treats probability as a degree of belief, Bayes' theorem can directly assign a probability distribution that quantifies the belief to the parameter or set of parameters.<sup>[1]</sup>

Bayesian statistics was named after [Thomas Bayes](#), who formulated a specific case of Bayes' theorem in [his paper](#) published in 1763. In several papers spanning from the late-1700s to the early-1800s, [Pierre-Simon Laplace](#) developed the Bayesian interpretation of probability. Laplace used methods that would now be considered as Bayesian methods to solve a number of statistical problems. Many Bayesian methods were developed by later authors, but the term was not commonly used to describe such methods until the 1950s. During much of the 20th century, Bayesian methods were unfavorable with many statisticians due to philosophical and practical considerations. Many Bayesian methods required a lot of computation to complete, and most methods that were widely used during the century were based on the frequentist interpretation. However, with the advent of powerful computers and new [algorithms](#) like [Markov chain Monte Carlo](#), Bayesian methods have seen increasing use within statistics in the 21st century.<sup>[1][2]</sup>

## Contents [hide]

- [Bayes' theorem](#)
- [Outline of Bayesian methods](#)
  - [Bayesian inference](#)

## Part of a series on [Statistics](#) Bayesian statistics



### Theory

[Admissible decision rule](#) · [Bayesian efficiency](#) · [Bayesian probability](#) · [Probability interpretations](#) · [Bayes' theorem](#) · [Bayes factor](#) · [Bayesian inference](#) · [Bayesian network](#) · [Prior](#) · [Posterior](#) · [Likelihood](#) · [Conjugate prior](#) · [Posterior predictive](#) · [Hyperparameter](#) · [Hyperprior](#) · [Principle of indifference](#) · [Principle of maximum entropy](#) · [Empirical Bayes method](#) · [Cromwell's rule](#) · [Bernstein–von Mises theorem](#) · [Schwarz criterion](#) · [Credible interval](#) · [Maximum a posteriori estimation](#) · [Radical probabilism](#)

### Techniques

[Bayesian linear regression](#) · [Bayesian estimator](#) · [Approximate Bayesian computation](#) · [Markov chain Monte Carlo](#)

[Statistics portal](#)

V · T · E

- 2.2 [Statistical modeling](#)
- 2.3 [Design of experiments](#)
- 2.4 [Statistical graphics](#)
- 3 [References](#)
- 4 [Further reading](#)
- 5 [External links](#)

## Bayes' theorem [[edit](#)]

*Main article: [Bayes' theorem](#)*

Bayes' theorem is a fundamental theorem in Bayesian statistics, as it is used by Bayesian methods to update probabilities, which are degrees of belief, after obtaining new data. Given two events **A** and **B**, the conditional probability of **A** given that **B** is true is expressed as follows:<sup>[3]</sup>

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

where  $P(B) \neq 0$ . Although Bayes' theorem is a fundamental result of [probability theory](#), it has a specific interpretation in Bayesian statistics. In the above equation, **A** usually represents a [proposition](#) (such as the statement that a coin lands on heads fifty percent of the time) and **B** represents the evidence, or new data that is to be taken into account (such as the result of a series of coin flips).  $P(A)$  is the [prior probability](#) of **A** which expresses one's beliefs about **A** before evidence is taken into account. The prior probability may also quantify prior knowledge or information about **A**.  $P(B \mid A)$  is the [likelihood function](#), which can be interpreted as the probability of the evidence **B** given that **A** is true. The likelihood quantifies the extent to which the evidence **B** supports the proposition **A**.  $P(A \mid B)$  is the [posterior probability](#), the probability of the proposition **A** after taking the evidence **B** into account. Essentially, Bayes' theorem updates one's prior beliefs  $P(A)$  after considering the new evidence **B**.<sup>[1]</sup>

The probability of the evidence  $P(B)$  can be calculated using the [law of total probability](#). If  $\{A_1, A_2, \dots, A_n\}$  is a [partition](#) of the [sample space](#), which is the set of all [outcomes](#) of an experiment, then,<sup>[1][3]</sup>

$$P(B) = P(B \mid A_1)P(A_1) + P(B \mid A_2)P(A_2) + \dots + P(B \mid A_n)P(A_n) = \sum_i P(B \mid A_i)P(A_i)$$

When there are an infinite number of outcomes, it is necessary to [integrate](#) over all outcomes to calculate  $P(B)$  using the law of total probability. Often,  $P(B)$  is difficult to calculate as the calculation would involve sums or integrals that would be time-consuming to evaluate, so often only the product of the prior and likelihood is considered, since the evidence does not change in the same analysis. The posterior is proportional to this product:<sup>[1]</sup>

$$P(A \mid B) \propto P(B \mid A)P(A)$$

The [maximum a posteriori](#), which is the [mode](#) of the posterior and is often computed in Bayesian statistics using [mathematical optimization](#) methods, remains the same. The posterior can be approximated even without computing the exact value of  $P(B)$  with methods such as Markov chain Monte Carlo or [variational Bayesian methods](#).<sup>[1]</sup>

## Outline of Bayesian methods [[edit](#)]

The general set of statistical techniques can be divided into a number of activities, many of which have special Bayesian versions.

## Bayesian inference [[edit](#)]

*Main article: [Bayesian inference](#)*

Bayesian inference refers to [statistical inference](#) where uncertainty in inferences is quantified using probability. In classical [frequentist inference](#), model [parameters](#) and hypotheses are considered to be fixed. Probabilities are not assigned to parameters or hypotheses in frequentist inference. For example, it would not make sense in frequentist inference to directly assign a probability to an event that can only happen once, such as the result of the next flip of a fair coin. However, it would make sense to state that the proportion of heads [approaches one-half](#) as the number of coin flips increases.<sup>[4]</sup>

[Statistical models](#) specify a set of statistical assumptions and processes that represent how the sample data is generated. Statistical models have a number of parameters that can be modified. For example, a coin can be represented as samples from a [Bernoulli distribution](#), which models two possible outcomes. The Bernoulli distribution has a single parameter equal to the probability of one outcome, which in most cases is the probability of landing on heads. Devising a good model for the data is central in Bayesian inference. In most cases, models only approximate the true process, and may not take into account certain factors influencing the data.<sup>[1]</sup> In Bayesian inference, probabilities can be assigned to model parameters. Parameters can be represented as [random variables](#). Bayesian inference uses Bayes' theorem to update probabilities after more evidence is obtained or known.<sup>[1][5]</sup>

## Statistical modeling [\[ edit \]](#)

The formulation of [statistical models](#) using Bayesian statistics has the identifying feature of requiring the specification of [prior distributions](#) for any unknown parameters. Indeed, parameters of prior distributions may themselves have prior distributions, leading to [Bayesian hierarchical modeling](#)<sup>[6]</sup>, or may be interrelated, leading to [Bayesian networks](#).


## Design of experiments [\[ edit \]](#)

The [Bayesian design of experiments](#) includes a concept called 'influence of prior beliefs'. This approach uses [sequential analysis](#) techniques to include the outcome of earlier experiments in the design of the next experiment. This is achieved by updating 'beliefs' through the use of prior and [posterior distribution](#). This allows the design of experiments to make good use of resources of all types. An example of this is the [multi-armed bandit problem](#).

## Statistical graphics [\[ edit \]](#)

[Statistical graphics](#) includes methods for data exploration, for model validation, etc. The use of certain modern computational techniques for Bayesian inference, specifically the various types of [Markov chain Monte Carlo](#) techniques, have led to the need for checks, often made in graphical form, on the validity of such computations in expressing the required posterior distributions.

## References [\[ edit \]](#)

- <sup>^</sup> <sup>[a](#)</sup> <sup>[b](#)</sup> <sup>[c](#)</sup> <sup>[d](#)</sup> <sup>[e](#)</sup> <sup>[f](#)</sup> <sup>[g](#)</sup> <sup>[h](#)</sup> <sup>[i](#)</sup> Gelman, Andrew; Carlin, John B.; Stern, Hal S.; Dunson, David B.; Vehtari, Aki; Rubin, Donald B. (2013). *Bayesian Data Analysis, Third Edition*. Chapman and Hall/CRC. ISBN 978-1-4398-4095-5.
- <sup>^</sup> Fienberg, Stephen E. (2006). "When Did Bayesian Inference Become "Bayesian"?" [↗](#). *Bayesian Analysis*. **1** (1): 1–40.
- <sup>^</sup> <sup>[a](#)</sup> <sup>[b](#)</sup> Grinstead, Charles M.; Snell, J. Laurie (2006). *Introduction to probability* (2nd ed.). Providence, RI: American Mathematical Society. ISBN 978-0-8218-9414-9.
- <sup>^</sup> Wakefield, Jon (2013). *Bayesian and frequentist regression methods*. New York, NY: Springer. ISBN 978-1-4419-0924-4.
- <sup>^</sup> Congdon, Peter (2014). *Applied Bayesian modelling* (2nd ed.). Wiley. ISBN 978-1119951513.
- <sup>^</sup> Hajiramezanali, E. & Dadaneh, S. Z. & Karbalayghareh, A. & Zhou, Z. & Qian, X. Bayesian multi-domain learning for cancer subtype discovery from next-generation sequencing count data. 32nd Conference on Neural Information Processing Systems (NIPS 2018), Montréal, Canada. <https://arxiv.org/pdf/1810.09433.pdf> 

## Further reading [\[ edit \]](#)

- [Think Bayes](#), Allen B. Downey
- [Bayesian Statistics: Why and How](#)
- Puga JL, Krzywinski M, Altman N (May 2015). "[Bayesian Statistics](#)". Points of Significance. *Nature Methods*. **12** (5): 377–8. doi:10.1038/nmeth.3368. Retrieved 31 May 2016.

## External links [\[ edit \]](#)

- Eliezer S. Yudkowsky. "[An Intuitive Explanation of Bayes' Theorem](#)" (webpage). Retrieved 2015-06-15.
- Theo Kypraios. "[A Gentle Tutorial in Bayesian Statistics](#)" (PDF). Retrieved 2013-11-03.
- Jordi Vallverdu. "[Bayesians Versus Frequentists A Philosophical Debate on Statistical Reasoning](#)".
- [Bayesian statistics](#) [David Spiegelhalter](#), Kenneth Rice *Scholarpedia* 4(8):5230. doi:10.4249/scholarpedia.5230
- [Bayesian modeling book](#) and examples available for downloading.
- Rens Van De Schoot. "[A Gentle Introduction to Bayesian Analysis](#)" (PDF).
- [Bayesian A/B Testing Calculator](#) [Dynamic Yield](#)



Wikiversity has learning resources about **Bayesian statistics**

Categories: [Bayesian statistics](#)

This page was last edited on 18 May 2019, at 15:48 (UTC).

Text is available under the [Creative Commons Attribution-ShareAlike License](#); additional terms may apply. By using this site, you agree to the [Terms of Use](#) and [Privacy Policy](#). Wikipedia® is a registered trademark of the [Wikimedia Foundation, Inc.](#), a non-profit organization.

[Privacy policy](#) [About Wikipedia](#) [Disclaimers](#) [Contact Wikipedia](#) [Developers](#) [Cookie statement](#) [Mobile view](#)

