

# Hyperparameter

From Wikipedia, the free encyclopedia

*This article is about hyperparameters in Bayesian statistics. For hyperparameters in machine learning, see Hyperparameter optimization.*

In Bayesian statistics, a **hyperparameter** is a parameter of a prior distribution; the term is used to distinguish them from parameters of the model for the underlying system under analysis.

For example, if one is using a beta distribution to model the distribution of the parameter  $p$  of a Bernoulli distribution, then:

- $p$  is a parameter of the underlying system (Bernoulli distribution), and
- $\alpha$  and  $\beta$  are parameters of the prior distribution (beta distribution), hence *hyperparameters*.

One may take a single value for a given hyperparameter, or one can iterate and take a probability distribution on the hyperparameter itself, called a hyperprior.

## Contents

- 1 Purpose
  - 1.1 Conjugate priors
  - 1.2 Sensitivity analysis
- 2 Hyperpriors
- 3 References

## Purpose

One often uses a prior which comes from a parametric family of probability distributions – this is done partly for explicitness (so one can write down a distribution, and choose the form by varying the hyperparameter, rather than trying to produce an arbitrary function), and partly so that one can *vary* the hyperparameter, particularly in the method of *conjugate priors*, or for *sensitivity analysis*.

## Conjugate priors

*Main article: Conjugate prior*

When using a conjugate prior, the posterior distribution will be from the same family, but will have different hyperparameters, which reflect the added information from the data: in subjective terms, one's beliefs have been updated. For a general prior distribution, this is computationally very involved, and the posterior may have an unusual or hard to describe form, but with a conjugate prior, there is generally a simple formula relating the values of the hyperparameters of the posterior to the values of the hyperparameters of the prior, and thus the computation of the posterior distribution is very easy.

## Sensitivity analysis

*Main article: sensitivity analysis*

A key concern of users of Bayesian statistics, and criticism by critics, is the dependence of the posterior distribution on one's prior. Hyperparameters address this by allowing one to easily vary them and see how the posterior distribution (and various statistics of it, such as credible intervals) vary: one can see how *sensitive* one's conclusions are to one's prior assumptions, and the process is called *sensitivity analysis*.

Similarly, one may use a prior distribution with a range for a hyperparameter, perhaps reflecting uncertainty in the correct prior to take, and reflect this in a range for final uncertainty.<sup>[1]</sup>

## Hyperpriors

*Main article: Hyperprior*

Instead of using a single value for a given hyperparameter, one can instead consider a probability distribution of the hyperparameter itself; this is called a "hyperprior." In principle, one may iterate this, calling parameters of a hyperprior "hyperhyperparameters," and so forth.

## References

1. Giulio D'Agostini, Purely subjective assessment of prior probabilities (<http://www.roma1.infn.it/~dagos/rpp/node30.html>), in Bayesian Inference in Processing Experimental Data: Principles and Basic Applications (<http://www.roma1.infn.it/~dagos/rpp/rpp.html>)
  - Bernardo, J.M., Smith, A.F.M. (2000) *Bayesian Theory*. Wiley. ISBN 0-471-49464-X

Retrieved from "<https://en.wikipedia.org/w/index.php?title=Hyperparameter&oldid=660674026>"

Categories: Bayesian statistics | Sensitivity analysis | Statistical terminology

- 
- This page was last modified on 3 May 2015, at 23:49.
  - 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.