Conjugate Bayesian analysis

Matthew Stephens
University of Chicago
January 12, 2026

See here for a PDF version of this vignette.

Overview

This vignette introduces the idea of "conjugate prior" distributions for Bayesian inference for a continuous parameter. You should be familiar with Bayesian inference for a binomial proportion.

Conjugate Priors for binomial proportion

Background

In this example we considered the following problem.

Suppose we sample 100 elephants from a population, and measure their DNA at a location ("locus") in their genome, where there are two types ("alleles"). We label these alleles as "0" and "1".

In my sample, I observe that 30 of the elephants have the 1 allele and 70 have the 0 allele. What can I say about the frequency, q, of the 1 allele in the population?

The example showed how to compute the posterior distribution for q using a *uniform* prior distribution. We saw that, conveniently, the posterior distribution for q was a Beta distribution.

Here we generalize this calculation to the case where the prior distribution on q is a Beta distribution. We will find that, in this case, the posterior distribution on q is again a Beta distribution. The property where the posterior distribution comes from the same family as the prior distribution is very convenient, and so has a special name: it is called "conjugacy". We say, "the Beta distribution is the conjugate prior distribution for the binomial proportion."

Details

As before, we use Bayes Theorem, which we can write in words as

$$\text{posterior} \propto \text{likelihood} \times \text{prior},$$

or in mathematical notation as

$$p(q \mid D) \propto p(D \mid q)\, p(q),$$

where D denotes the observed data. In this case, the likelihood $p(D \mid q)$ is given by

$$p(D \mid q) \propto q^{30}(1-q)^{70}.$$

If our prior distribution on q is a Beta distribution, say, $\text{Beta}(a, b)$, then the prior density $p(q)$ is

$$p(q) \propto q^{a-1}(1-q)^{b-1} \qquad (q \in [0,1]).$$

Combining these two, we get

$$\begin{aligned} p(q \mid D) &\propto q^{30}(1-q)^{70} q^{a-1}(1-q)^{b-1} \\ &\propto q^{30+a-1}(1-q)^{70+b-1} \end{aligned}$$

At this point, we again apply the trick of recognizing this density as the density of a Beta distribution — specifically, the Beta distribution with parameters $30 + a$, $70 + b$, written as $\text{Beta}(30 + a, 70 + b)$.

Generalization

There is nothing special about the 30 "1" alleles and 70 "0" alleles we observed. Suppose we observed n_1 of the 1 allele and n_0 of the 0 allele. Then the likelihood becomes

$$p(D \mid q) \propto q^{n_1}(1-q)^{n_0},$$

and you should be able to show — *this is an exercise* — that the posterior is

$$q \mid D \sim \text{Beta}(n_1 + a, n_0 + b).$$

Summary

When doing Bayesian inference for a binomial proportion, q, if the prior distribution is a Beta distribution, then the posterior distribution is also Beta. We say, "the Beta distribution is the conjugate prior for a binomial proportion."

Exercise

Show that the Gamma distribution is the conjugate prior for a Poisson mean.

That is, suppose we have observations X that are Poisson distributed, $X \sim \text{Poisson}(\mu)$. Assume that your prior distribution on μ is a Gamma distribution with scale a and shape b. Show that the posterior distribution on μ is also a Gamma distribution.

Hint: You should take the following steps. 1. Write down the likelihood $p(X \mid \mu)$. (Look up the density function of the Poisson distribution if you cannot remember it.) 2. Write down the prior density for μ. (Look up the density of the Gamma distribution if you cannot remember it.) 3. Multiply them together to obtain the posterior density (up to a constant of proportionality), and notice that it has the same form as a Gamma distribution.