Bayesian Statistics VII

Matthieu Vignes

April 2024

Conjugate Bayesian Analysis

We introduce the idea of *conjugate prior* distributions for Bayesian inference for a continuous parameter.

Conjugate Priors for binomial proportion

Background

We still consider the elephant example. We just showed how to compute the posterior distribution for q, using a uniform prior distribution. We saw that, conveniently, the posterior distribution for q is a peculiar case of a Beta distribution.

Here we generalize this calculation to the case where the prior distribution on q is any Beta distribution. We will find that, in this case, the posterior distribution on q is also 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 that "The Beta distribution is the conjugate prior distribution for the binomial proportion" (or model).

Details

If we take the previous derivation of the posterior when the prior was uniform on $[0, 1]$ and we change the prior to be a $(Beta)(a, b)$, for any $a, b > 0$, i.e.

$$p(q) = q^{a-1}(1-q)^{b-1}, \text{ for } q \in [0, 1],$$

We can combine this with the likelihood $P(D|q) \propto q^{30}(1-q)^{70}$ to derive the posterior distribution for q:

$$P(q|D) \propto q^{30+a-1}(1-q)^{70+b-1}.$$

At this point we again apply the "trick" of recognising this density as the density of a Beta distribution. Specifically, this is the Beta distribution with parameters $(30 + a, 70 + b)$.

Generalisation

Of course, there is nothing special about the 30 `1` alleles and 70 `0` alleles we observed here. Suppose we observed n_1 of the `1` allele and n_0 of the `0` allele. Then the likelihood becomes $p(D|q) \propto q^{n_1}(1-q)^{n_0}$, and you should be able to show (Exercise) that the posterior is:

$$q|D \sim Beta(n1 + a, n0 + b)$$

.

This can be interpreted as a role of the prior to have an "equivalent sample size effect" including a number of `1` and `0` data points related to a and b, respectively.

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 that 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 parameters n and λ. Show that the posterior distribution on μ is also a Gamma distribution.

Hint: you should take the following steps. 1. write down the likelihood $p(X|\mu)$ for μ (look up the Poisson distribution if you cannot remember it). 2. Write down the prior density for μ (look up the density of a 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 the gamma distribution.