

Lab 3

PSTAT 115

2025-10-13

This lab will focus on the following topics:

- Derive posterior distribution from prior and likelihood function
- Common conjugate priors and the meaning of them.
- Posterior mean and MAP estimates

Estimating the probability of a female birth

The proportion of births that are female has long been a topic of interest both scientifically and to the lay public. Two hundred years ago it was established that the proportion of female births in European populations was less than 0.5. The currently accepted value of the proportion of female births in large European-race populations is 0.485.

- Assuming Bernoulli distribution for each child birth, and let p denote the probability of being a female. What is the probability of observing y female babies among n newborn babies (y is between 0 and n , inclusive)?
- Now assume we do not have any prior information about the probability of female birth, so a uniform prior on $[0, 1]$ is used. What is the posterior distribution under this prior?
- We know the uniform distribution is a special case of the Beta distribution. Now what if we apply a $Beta(2, 2)$ distribution as our prior? What is the posterior distribution now?

Common Conjugate priors

Conjugacy is formally defined as follows. A collection of pdfs (or pmfs) is called a conjugate prior family for a model $X \sim f(x|\theta), \theta \in \Theta$, if whenever a prior $\xi(\theta)$ is chosen from the collection, it leads to a posterior $\xi(\theta|x)$ that is also a member of the collection, for every observation $X = x$. For instance, in our above binomial model, the Beta distribution is a conjugate prior to the binomial distribution.

i.e. The prior and posterior distributions belong to the same family.

Conjugate priors permit fast posterior computations, which can be valuable in high dimension problems. Meanwhile, there are handy formulas for mean and mode of the posterior distributions. The following table summarizes the common distribution we might encounter throughout the course. The mean and variance of each posterior distribution are also of importance.

https://en.wikipedia.org/wiki/Conjugate_prior

Posterior mean and MAP estimates

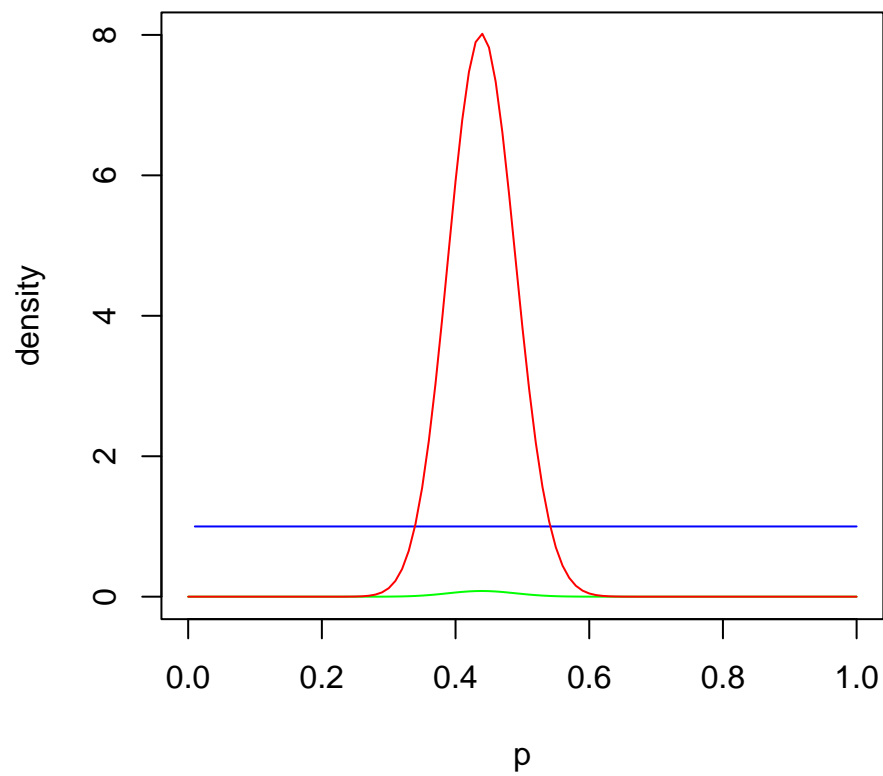
An early study concerning the sex of newborn Germany babies found that of a total of 98 births, 43 were female. What distribution will this follow? Assume we are using the uniform prior.

- Plotting the prior distribution (in blue lines), binomial likelihood (in green lines) and the posterior distribution (in red lines).

```
# prior
curve(p/p, from = 0, to = 1, xname = "p", xlab = "p", ylab = "density",
      ylim = c(0, 8), col = 'blue')

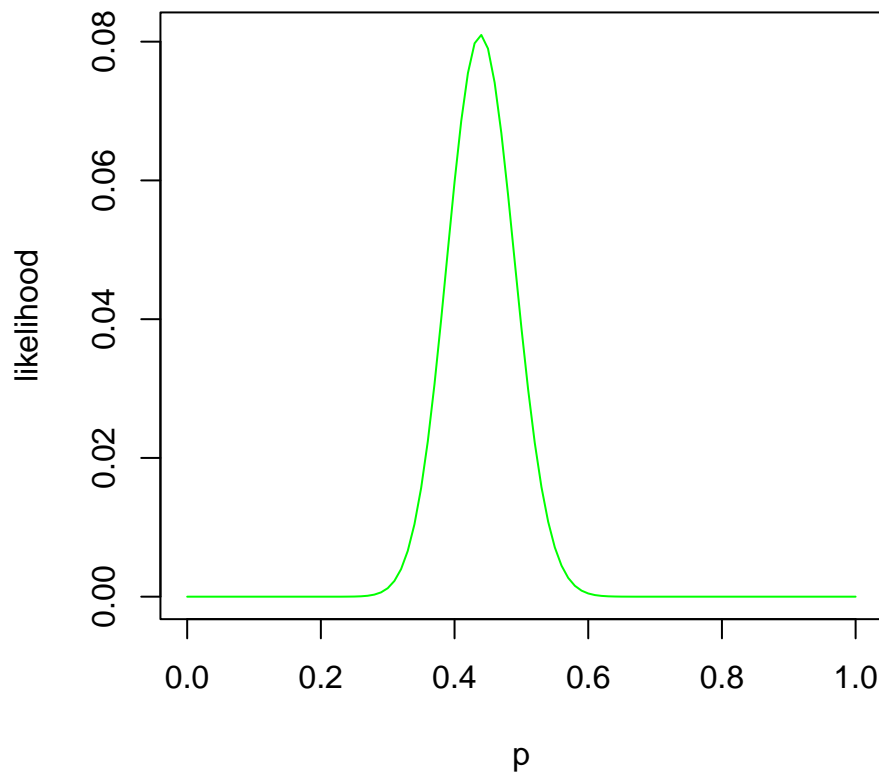
# likelihood
curve(choose(98, 43) * p^43 * (1-p)^(98-43),
      from = 0, to = 1, xname = "p",
      xlab = "p", ylab = "likelihood", add = TRUE, ylim = c(0, 8), col = 'green')

# posterior
a_post = 1 + 43
b_post = 1 + 98 - 43
curve(gamma(a_post + b_post)/gamma(a_post)/gamma(b_post) *
      p^(a_post - 1) * (1-p)^(b_post - 1), from = 0, to = 1, xname = "p",
      xlab = "p", ylab = "density", add = TRUE, ylim = c(0, 8), col = 'red')
```



It helps to plot the likelihood alone.

```
# likelihood
curve(choose(98, 43) * p^43 * (1-p)^(98-43),
      from = 0, to = 1, xname = "p",
      xlab = "p", ylab = "likelihood", col = 'green')
```



The posterior probability distribution contains all the current information about the parameter θ . A graphical report on the entire posterior distribution is definitely meaningful and useful. For many practical cases, however, various numerical summaries of the distributions are desirable.

- As for all distributions, the mean of the distribution is an important location summary. What is the mean of the posterior distribution of our above example?
- Meanwhile, the MAP (maximum a posteriori), which is the mode of the posterior distribution, can be interpreted as the single “most likely” value of the parameter, given the data and the model. What is the MAP estimate in our posterior distribution?
- Is there always a unique MAP value in the posterior distribution?