STAT 431 — Applied Bayesian Analysis — Course Notes

Population Proportion: Obtaining a Posterior

Fall 2022

In a population, suppose an unknown proportion θ of individuals have a certain characteristic.

Example: What proportion of people like us have pets?

How do we usually estimate θ?

What kind of data would that require?

Suppose we have a sample of size n from the population.

Let

$$y \;=\; \text{number in the sample having the characteristic}$$

Then the usual estimate of θ is

$$\hat{\theta} \;=\; \frac{y}{n} \;=\; \text{the sample proportion}$$

But what would a Bayesian do?

Bayesians want a **posterior distribution** for θ.

Step 1: Define the Data Model(s)

Example: proportion of people like us who own pets

Assuming a random sample of given size n, let

$$Y = \text{number in sample having pets}$$

so that

$$Y \mid \theta \sim \mathrm{Binomial}(n, \theta)$$

$$\theta = \text{(unknown) population proportion with pets}$$

Density (PMF) for data:

$$f(y \mid \theta) = \binom{n}{y} \theta^{y} (1-\theta)^{n-y}, \quad y = 0, 1, \ldots n$$

R Example 1.1(a):

Population Proportion Models

Step 2: Obtain the Likelihood Function

Data (from a survey):

$$y \;=\; 12 \qquad \text{out of} \qquad n \;=\; 70$$

The likelihood function:

$$\begin{aligned} f(y = 12 \mid \theta) \;&=\; \binom{70}{12}\,\theta^{12}\,(1-\theta)^{58} \\ &\propto\; \theta^{12}\,(1-\theta)^{58} \end{aligned}$$

R Example 1.1(b):

Binomial Model Likelihood

Step 3: Specify the Prior Density

We will try several prior densities for θ.

Eg: Flat ("Noninformative") Prior

$$\theta \sim \mathrm{Uniform}(0,1)$$

$$\pi(\theta) = \begin{cases} 1 & 0 < \theta < 1 \\ 0 & \text{otherwise} \end{cases}$$

But there is a more gereral and flexible choice ...

Eg: Beta Prior (see BSM Appendix A.1)

$$\theta \quad \sim \quad \mathrm{Beta}(\alpha, \beta)$$

$$\pi(\theta) \quad = \quad \begin{cases} \dfrac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\, \theta^{\alpha-1}\,(1-\theta)^{\beta-1} & 0 < \theta < 1 \\ 0 & \text{otherwise} \end{cases}$$

As we will later see, this distribution is **conjugate**:
It combines with the likelihood to produce a beta posterior.

(Q: What α and β make this prior uniform?)

Some guesses about θ (from various people) gave

average: 0.236 sample std. deviation: 0.147

We can get the beta prior to have

$$\mathrm{E}(\theta) \approx 0.236 \qquad \mathrm{Var}(\theta) \approx (0.147)^2$$

by choosing

$$\alpha \approx 1.74 \qquad \beta \approx 5.63$$

(verify — BSM Appendix A.1)

Alternative way: Find interval I containing, say, 95% of people's guesses, then choose α and β such that

$$\mathrm{Prob}(\theta \in I) \;\; = \;\; 0.95$$

I is a "prior 95% interval."

Eg: Discrete Prior

Assign probabilities to a few discrete points, e.g.

θ:	0.05	0.15	0.25	0.35	0.5
$\pi(\theta)$:	0.10	0.30	0.30	0.15	0.15

May seem strange, since θ is a "continuous" parameter, but can be useful in high-dimensional problems.

Eg: Histogram Prior

Assign probabilities to ranges of values, e.g.

R	$\text{Prob}(\theta \in R)$
$[0, 0.15)$	0.25
$[0.15, 0.25)$	0.30
$[0.25, 0.35)$	0.25
$[0.35, 0.45)$	0.05
$[0.45, 1)$	0.15

Then let the density be uniform on each range.

R Example 1.1(c):

Prior Densities

Step 4: Compute the Posterior Density

Eg: Beta Prior (including uniform)

Conjugacy permits an analytical solution:

$$
\begin{aligned}
p(\theta \mid y = 12) &\propto f(y = 12 \mid \theta)\, \pi(\theta) \\
&\propto \theta^{12}\, (1-\theta)^{58} \;\cdot\; \theta^{\alpha-1}\, (1-\theta)^{\beta-1} \\
&= \theta^{12+\alpha-1}\, (1-\theta)^{58+\beta-1} \qquad \text{for } 0 < \theta < 1
\end{aligned}
$$

We recognize this as a $\mathrm{Beta}(12+\alpha, 58+\beta)$ **kernel**: the density except for multiplicative constants.

It follows that

$$\theta \mid Y = 12 \quad \sim \quad \text{Beta}(12 + \alpha, 58 + \beta)$$

e.g. uniform prior ($\alpha = \beta = 1$) gives

$$\text{Beta}(13, 59)$$

e.g. our beta prior with $\alpha = 1.74$, $\beta = 5.63$ gives

$$\text{Beta}(13.74,\ 63.63)$$

For our discrete prior and our histogram prior, we will let the computer produce the posterior ...

R Example 1.1(d):

Posterior Densities