STAT 431 — Applied Bayesian Analysis — Course Notes

Population Proportion: Getting a Posterior

Spring 2019

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

Example: What proportion of people like us have pets?

How do we usually estimate π?

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 π is

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

But what would a Bayesian do?

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

Step 1: Define the Data Model(s)

Example: proportion of people like us with pets

Assuming a random sample of given size n, let

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

so that

$$Y \mid \pi \quad \sim \quad \text{binomial}(n, \pi)$$

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

Density (pmf) for data:

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

R Example 3.1(a):

Population Proportion Models

Step 2: Obtain the Likelihood Function

Data (from our survey):

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

The likelihood function:

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

R Example 3.1(b):

Binomial Model Likelihood

Step 3: Specify the Prior

If we want a posterior distribution for π, we must specify a prior distribution for π ...

Eg: Flat ("Noninformative") Prior

$$\begin{aligned} \pi &\sim \text{uniform}(0,1) \\ p(\pi) &= \begin{cases} 1 & 0 < \pi < 1 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

Eg: Beta Prior (see Cowles, Table A.2)

$$\pi \sim \mathrm{beta}(\alpha, \beta)$$

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

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

Generally, a type of distribution is **conjugate** for a likelihood if a prior of that type produces a posterior of that same type.

Your guesses at π from the survey had

average: 0.236 sample std. deviation: 0.147

We can get the beta prior to have

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

by choosing

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

(verify — Cowles, Table A.2)

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

$$\mathrm{P}(\pi \in I) \ = \ 0.95$$

I is a "prior 95% interval." (See Cowles, Sec. 3.3.2, 3.5.2)

Eg: Discrete Prior

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

π:	0.05	0.15	0.25	0.35	0.5
$p(\pi)$:	0.10	0.30	0.30	0.15	0.15

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

Eg: Histogram Prior

Assign probabilities to ranges of values, e.g.

R	$\mathrm{P}(\pi \in R)$
$[0, 0.15)$	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 3.1(c):

Prior Densities

Step 4: Compute the Posterior

Eg: Beta Prior (including uniform)

Conjugacy permits an analytical solution:

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

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

It follows that

$$\pi \mid Y = 12 \quad \sim \quad \mathrm{beta}(12 + \alpha, 58 + \beta)$$

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

$$\mathrm{beta}(13, 59)$$

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

$$\mathrm{beta}(13.74,\ 63.63)$$

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

R Example 3.1(d):

Posterior Densities