ST 740: Bayesian Basics

Alyson Wilson

Department of Statistics
North Carolina State University

August 21, 2013

What is Bayesian statistics?

What is classical/frequentist statistics?

Building blocks

- data $\mathbf{Y}$ (observable vector)
- parameter θ (unobservable vector)
- covariate $\mathbf{X}$ (observable vector)

We also have

- $f(\mathbf{y} \mid \theta, \mathbf{X})$, a pdf or pmf specifying the distribution of $\mathbf{Y}$.
  Note: Bayesians usually call this the *sampling distribution* of the data.
- "$L(\theta; \mathbf{y})$", the likelihood function, $f(\mathbf{y} \mid \theta, \mathbf{X})$ evaluated at the observed $\mathbf{Y} = \mathbf{y}$ and considered as a function of θ

What is Bayesian statistics?

Classical or *frequentist* statistics uses the likelihood alone to guide inference and data analysis.

Think maximum *likelihood* estimation.

What is Bayesian statistics?

Bayesian statistics uses a broader set of information to guide inference and data analysis.

In particular

- $f(\mathbf{y} \mid \theta, \mathbf{X})$, the sampling distribution
- $\pi(\theta)$, the prior distribution

The prior distribution captures "pre-data" information about the parameter.

Prior Distributions

The addition of a prior distribution is a fundamental difference between Bayesian and classical methods. Why?

- For a Bayesian, a parameter is a *random variable*.
- This is not because Bayesians think that parameters are random.
- Rather, parameters are quantities about which we are uncertain. We choose to quantify our uncertainty about the parameters using probability.

Example

Classical Inference for the Binomial Proportion

There are a number of different confidence intervals for the binomial proportion (Agresti and Coull 1998). Before observing the data, these are random intervals with a 95% chance of containing the proportion. After observing the data, there's nothing random left. The interval does or does not contain the proportion, and we don't know which. Because of this, we can't actually say what we'd like to say, which is "there's a 95% chance that the parameter is in the inerval."

- Wald confidence interval: $\hat{p} \pm z_{\frac{\alpha}{2}} \sqrt{\hat{p}(1-\hat{p})/n}$
- Clopper-Pearson confidence interval:
$$\left(\left[1 + \frac{n-x+1}{xF_{2x,2(n-x+1),1-\alpha/2}} \right]^{-1}, \left[1 + \frac{n-x}{(x+1)F_{2(x+1),2(n-x),\alpha/2}} \right]^{-1} \right)$$

Example

Classical Inference for the Binomial Proportion

For this problem, we can use the observed data to calculate the following 95% confidence intervals:

- Wald confidence interval: (0.059, 0.174)
- Clopper-Pearson confidence interval: (0.065, 0.188)

Example

Bayesian Inference for the Binomial Proportion

When I talk about a *statistical model*, I mean **BOTH** the sampling distribution and the prior distribution. Bayesian inference always starts by specifying the statistical model.

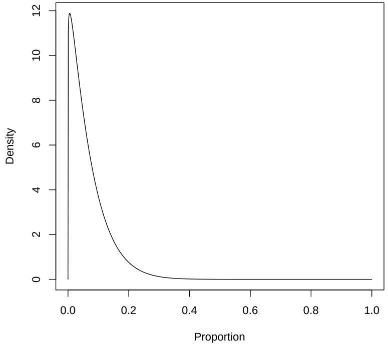
Example

Bayesian Inference for the Binomial Proportion

When I talk about a *statistical model*, I mean **BOTH** the sampling distribution and the prior distribution. Bayesian inference always starts by specifying the statistical model.

$$
\begin{aligned}
Y &\sim \text{Binomial}(n = 120, p) \\
p &\sim \text{Beta}(1.08, 14.96)
\end{aligned}
$$

The median of the prior distribution for p is 0.05, and the 0.9 quantile is 0.15.



Bayes Theorem

A good way to remember Bayes Theorem:

The posterior is proportional to the prior times the likelihood.