STAT 535: Introduction

David B. Hitchcock
E-Mail: hitchcock@stat.sc.edu

Spring 2022

Chapter 1: Introduction to Bayesian Data Analysis

- **Bayesian** statistical inference uses **Bayes' Rule** (a.k.a. Bayes' Law or Bayes' Theorem) to combine prior information and sample data to make conclusions about a parameter of interest.
- Bayesian inference differs from classical **frequentist** inference in that it specifies a **probability distribution** for the parameter(s) of interest.
- Why use Bayesian methods? Some reasons:
  1. We wish to specifically incorporate previous knowledge we have about a parameter of interest.
  2. To logically update our knowledge about the parameter after observing sample data
  3. To make formal probability statements about the parameter of interest.
  4. To specify model assumptions and check model quality and sensitivity to these assumptions in a straightforward way.

- Why do people use classical methods?
  1. If the parameter(s) of interest is/are truly fixed (without the possibility of changing), as is possible in a highly controlled experiment
  2. If there is no prior information available about the parameter(s)
  3. If they prefer “cookbook”-type formulas with little input from the scientist/researcher
- Many reasons classical methods are more common than Bayesian methods are historical:
  1. Many methods were developed in the context of controlled experiments.
  2. Bayesian methods require a bit more mathematical formalism.
  3. Historically (but not now) realistic Bayesian analyses had been infeasible due to a lack of computing power.

Motivation for Bayesian Modeling

- Bayesians treat unobserved data and unknown parameters in similar ways.
- They describe each with a probability distribution.
- As their model, Bayesians specify:
  1. A **joint density function**, which describes the form of the distribution of the **full sample** of data (given the parameter values)
  2. A **prior distribution**, which describes the behavior of the parameter(s) **unconditional** on the data
- The prior could reflect:
  1. Uncertainty about a parameter that is actually fixed
     OR
  2. the variety of values that a truly stochastic parameter could take.

How Probability is Treated

- The Bayesian approach treats probability as the **relative plausibility** of an event.
- The frequentist approach treats probability as the *long-run relative frequency* of an event that can be repeated.
- Consider the probability of getting a “head” when flipping a coin. The Bayesian would ask, “How likely is a “head” relative to the likelihood of a “tail”?
- The frequentist would ask, “If we flipped the coin many times, what proportion of the flips would result in heads?”
- The probabilities might be the same in each case, but the ways they are conceived of are different.

Non-repeatable Events

- For some events, it's hard to imagine them as repeatable when thinking of probability.
  1. The probability of the Democrats winning re-election in the 2024 presidential election.
  2. The probability of the Gamecocks women's basketball team winning the national championship in 2024.
  3. The probability of rain tomorrow.
  4. The probability that humans will one day live on Mars.

A Thought Experiment

- For each of these situations, think of whether the probability of a "success" is greater than 0.5:
    1. A person claims he can predict the outcome of an ordinary coin flip. In seven flips of a fair coin, the person correctly predicts the outcome all 7 times.
    2. A lady claims she can tell, by tasting tea, whether milk has been added to tea, or whether the milk was poured first and the tea added to it. With seven cups of tea, the lady correctly identifies the order of the pouring all 7 times.
    3. A consumer claims she can tell whether a coffee contains sugar or artificial sweetener. With seven cups of coffee, the consumer correctly names the type of sweetener all 7 times.
    4. A music student claims he can tell whether a piece of music was written by Mozart or Haydn. In seven pieces of music, he correctly names the composer all 7 times.

A Thought Experiment: Conclusions

- If your answers were different for these four situations, then you are using **prior information** in your assessment.
- Bayesian analyses balance **prior information** and **information from the data**.

Two Approaches to Testing Hypotheses

- Example from *Bayes Rules!* textbook:
- Imagine that you tested positive for a rare disease and only got to ask the doctor one question: (a) what's the chance that I actually have the disease?, or (b) if in fact I do not have the disease, what's the chance that I would've gotten this positive test result?
- A Bayesian hypothesis test seeks to answer: In light of the observed data, what's the chance that the hypothesis is correct? [Question (a)]
- A frequentist hypothesis test seeks to answer: If in fact the hypothesis is incorrect, what's the chance I'd have observed this, or even more extreme, data? [Question (b)]
- This "chance" is simply a frequentist p-value (a quantity that is often misinterpreted).
- Which answer is more helpful?

Rise of Bayesian Thinking

- We earlier mentioned why classical frequentist statistics is used in practice.
- In recent years, Bayesian approaches have enjoyed a resurgence. Why?
  - Advances in computing that allow more complicated Bayesian model to be practically implemented.
  - A gradual departure from tradition — most people use classical statistics simply because this is what they are taught, but this is **slowly** changing.
  - A greater acceptance of subjectivity in science. Since the Enlightenment, objective analyses have been the gold standard, but recently people have recognized that incorporating subjectivity can improve analyses.

Outline of Course

- Bayesian Foundations and Simple Bayesian Models
- Posterior Simulation and Analysis, Including MCMC Methods
- Bayesian Regression and Classification Methods
- Hierarchical Bayesian Models