Chapter 8.2 Modeling Measurements

Jim Albert and Monika Hu

Chapter 8 Modeling Measurement and Count Data

Introduction

- Assume we have hypothetical population of individuals of interest.
- There is a continuous-valued measurement Y associated with each individual.
- Represent the collection of measurements from all individuals by means of a continuous probability density $f(y)$.
- The mean value μ is a typical value of the continuous measurement Y.

Sampling

- A random sample of individuals $Y_1, ..., Y_n$ will be taken.
- Also have prior beliefs about location of μ.
- The inferential problem is to use these measurements together with any prior beliefs to learn about the population mean μ.

College applications

- How many college applications does a high school senior in the United States complete?
- Imagine a population of all American high school seniors and the measurement is the number of completed college applications
- The unknown quantity is the mean number of applications μ completed by these high school seniors
- Conduct a survey to a sample of high school seniors and use information from the survey to perform inference about the mean number of college applications.

Household spending

- How much does a household in San Francisco spend on housing every month?
- Consider the population of households in San Francisco and the continuous measurement is the amount of money spent on housing .
- To learn about the mean value of housing μ of all San Francisco residents, a sample survey is conducted.
- The mean value of the housing costs $\bar{y}$ from this sample of surveyed households is informative about the mean housing cost μ for all residents.

The general approach

Recall the three general steps of Bayesian inference in the context of an unknown proportion p.

- Step 1: **Prior** We express an opinion about the location of the proportion p before sampling.
- Step 2: **Data/Likelihood** We take the sample and record the observed proportion.
- Step 3: **Posterior** We use Bayes' rule to sharpen and update the previous opinion about p given the information from the sample.

Normal population

- Have a continuous population of measurements represented by Y with density function $f(y)$.
- Assume that this population has a Normal shape with mean μ and standard deviation σ.
- Single measurement Y is assume to come from the density function

$$f(y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(y-\mu)^2}{2\sigma^2}\right\}, -\infty < y < \infty.$$

- Assume that the standard deviation σ of the measurement distribution is known and the objective is to learn about the single mean measurement μ.

Prior

- Step 1 in Bayesian inference is to express an opinion about the parameter.
- One constructs a prior for the mean parameter μ that expresses one's opinion about the location of this mean.
- One attractive discrete approach for expressing this prior opinion constructs a list of possible values of μ, and then one assigns probabilities to the possible values to reflect one's belief.
- Alternatively, one can use of a continuous prior to represent one's belief for μ.

Data

- Step 2 of our process is to collect measurements from a random sample.
- In our first situation, one collects the number of applications from a sample of 100 high school seniors.
- If measurements are viewed as independent observations from a Normal sampling density with mean μ, then one constructs a likelihood function which is the joint density of the sampled measurements viewed as a function of the unknown parameter.

Update opinion

- Step 3 applies Bayes' rule to update one's prior opinion to obtain a posterior distribution for the mean μ.
- The algebraic implementation of Bayes' rule is a bit more tedious when dealing with continuous data with a Normal sampling density.
- But we will see there is a simple procedure for computing the posterior mean and standard deviation.