Chapter 10.1 Introduction

Jim Albert and Monika Hu

Chapter 10 Bayesian Hierarchical Modeling

Observations in groups

- Chapters 7, 8, and 9 make an underlying assumption about the source of data: observations are assumed to be identically and independently distributed (*i.i.d.*) following a single distribution with one or more unknown parameters
- Sometimes *i.i.d.* is not sensible

Observations in groups cont'd

- In the dining our example in Chapter 7:
  - dining preferences for students may be different from dining performances of senior citizens
  - it would not make sense to use a single success probability for a combined group of students and senior citizens
- In many applications
  - some observations share characteristics which distinguish them from other observations
  - multiple distinct groups are observed

Standardized test scores

- Consider a study in which students' scores of a standardized test such as the SAT are collected from five different senior high schools in a given year
- Instead of Y_i of student i ($i = 1, \cdots, n$, where n is the total number of students from all five schools)
- Let Y_{ij} denote the SAT score of student i in school j
  - $j = 1, \cdots, 5$, and $i = 1, \cdots, n_j$
  - n_j is the number of students in school j
  - $n = \sum_{j=1}^{5} n_j$

Standardized test scores cont'd

- SAT scores are continuous: normal sampling model
- Within school j, assume that SAT scores are *i.i.d.** from a normal data model with a mean and standard deviation depending on the school
- Specifically, assume a school-specific mean μ_j and a school-specific standard deviation σ_j for the normal data model for school j

$$Y_{ij} \overset{i.i.d.}{\sim} \mathrm{Normal}(\mu_j, \sigma_j),$$

where $j = 1, \cdots, 5$ and $j = 1, \cdots, n_j$

Separate estimates?

- Focuse on the observations in school j
  - $\{Y_{1j}, Y_{2j}, \cdots, Y_{n_j j}\}$
  - choose a prior distribution $\pi(\mu_j, \sigma_j)$ for the mean and the standard deviation parameters
  - follow the Bayesian inference procedure in Chapter 9 and obtain posterior inference on μ_j and σ_j
- Cons: in many cases, one school's information provides insight about another school

Combined estimates?

- Ignore the school variable and simply assume that the SAT scores

$$Y_i \overset{i.i.d.}{\sim} \mathrm{Normal}(\mu, \sigma)$$

- $i = 1, \cdots, n$ where n is the total number of students from all five schools
- follow the Bayesian inference procedure in Chapter 9 and obtain posterior inference on μ and σ
- Cons: ignore any differences between the five schools

A two-stage prior leading to compromise estimates

- Data model:
$$Y_{ij} \overset{i.i.d.}{\sim} \mathrm{Normal}(\mu_j, \sigma_j)$$
- Prior stage 1:
$$\mu_j \mid \mu, \tau \sim \mathrm{Normal}(\mu, \tau),\ j = 1, ..., 5$$
- Prior stage 2:
$$\mu, \tau \sim \pi(\mu, \tau)$$

A two-stage prior leading to compromise estimates cont'd

- μ_j's are not the same value
- μ_j's *a priori* are related and come from the same distribution
    - τ large: μ_j's can be very different from each other *a priori*
    - τ small: μ_j's can be very similar to each other *a priori*
    - $\tau = 0$: "combined estimates"; $\tau = \infty$: "seperate estimates"
- μ and τ are hyperparameters; need hyperpriors