Bayesian data analysis – reading instructions 8

Aki Vehtari

Chapter 8

In the earlier chapters it was assumed that the data collection is ignorable. Chapter 8 explains when data collection can be ignorable and when we need to model also the data collection. We don't have time to go through chapter 8 in BDA course at Aalto, but it is highly recommended that you would read it in the end or after the course. Most important parts are 8.1, 8.5, pp 220–222 of 8.6, and 8.8, and you can get back to the other sections later.

Outline of the chapter 8 (* denotes the most important parts)

- 8.1 Bayesian inference requires a model for data collection (*)
- 8.2 Data-collection models and ignorability
- 8.3 Sample surveys
- 8.4 Designed experiments
- 8.5 Sensitivity and the role of randomization (*)
- 8.6 Observational studies (* pp 220–222)
- 8.7 Censoring and truncation (*)

Most important terms in the chapter

- observed data
- complete data
- missing data
- stability assumption
- data model
- inclusion model
- complete data likelihood
- observed data likelihood
- finite-population and superpopulation inference
- ignorability
- ignorable designs
- propensity score
- sample surveys
- random sampling of a finite population
- stratified sampling
- cluster sampling
- designed experiments
- complete randomization
- randomized blocks and latin squares
- sequntial designs
- randomization given covariates

- observational studies
- censoring
- truncation
- missing completely at random

Gelman: “All contexts where the model is fit to data that are not necessarily representative of the population that is the target of study. The key idea is to include in the Bayesian model an inclusion variable with a probability distribution that represents the process by which data become observed.”