Chapter 12.2a Bayesian Multiple Regression

Jim Albert and Monika Hu

Chapter 12 Bayesian Multiple Regression and Logistic Models

Example: expenditures of U.S. households

- The U.S. Bureau of Labor Statistics (BLS) conducts the Consumer Expenditure (CE) Surveys through which the BLS collects data on expenditures, income, and tax statistics about households across the United States.
- The Quarterly Interview Survey aims to capture large purchases. The Diary Survey focuses on capturing small purchases (such as food, beverages, tobacco).
- Consider a sample of the Quarterly Interview Survey in 2017. This sample contains 1000 consumer units (CU), and provides information of the CU's total expenditures in last quarter, the amount of CU income before taxes in past 12 months, and the CU's urban/rural status.

Variable Description for the CE Sample

Variable	Description
Expenditure	Continuous; CU's total expenditures in last quarter
Income	Continuous; the amount of CU income before taxes in past 12 months
UrbanRural	Binary; the urban/rural status of CU: 1 = Urban, 2 = Rural

A Prediction Problem

- Suppose someone is interested in predicting a CU's expenditure from its urban/rural status and its income before taxes.
- Think of expenditure as the response variable and the other two variables as predictors.
- One needs to develop a model to express the relationship between expenditure and the other two predictors.
- Want to extend the simple linear regression model introduced in Chapter 11 to the case with multiple predictors.
- This extension is known as multiple linear regression.

Questions

- How to set up a multiple linear regression model?
- How to specify prior distributions for regression coefficients of multiple predictors?
- How to make Bayesian inferences and predictions?
- Consider a situation where one of the explanatory variables is continuous (like income) and one explanatory variable is categorical Ilike urban/rural status).