Section for Applied Statistics and Data Analysis

TA: Cong Mu

Office Hour: Wednesday 10:00AM - 12:00PM

September 27, 2019

Overview

1. Some Statistics
   - Properties of Least Squares Estimates
   - Inference in Regression
2. Some Programming
   - Examples in Faraway
   - Exercises in Faraway

Review

Recall

For constant matrices A, B, C and matrices X, Y with random variables

$$\mathbb{E}[AXB + C] = A\mathbb{E}[X]B + C.$$

$$\mathbb{E}[AX + BY] = A\mathbb{E}[X] + B\mathbb{E}[Y].$$

$$Cov[AX, BY] = ACov[X, Y]B^{\top}.$$

$$Var[AX] = Cov[AX, AX] = ACov[X, X]A^{\top} = AVar[X]A^{\top}.$$

Properties of Least Squares Estimates

- General Formulation

$$y = X\beta + \epsilon \qquad \text{where} \qquad \epsilon \sim \mathcal{N}(0, \sigma^2 I)$$

- Least Squares Estimates

$$\widehat{\beta} = \left(X^\top X\right)^{-1} X^\top y.$$

- Properties

$$\mathbb{E}\left[y\right] = X\beta, \qquad \mathbb{V}ar\left[y\right] = \sigma^2 I.$$

$$\mathbb{E}\left[\widehat{\beta}\right] = \beta, \qquad \mathbb{V}ar\left[\widehat{\beta}\right] = \sigma^2 \left(X^\top X\right)^{-1}.$$

Inference in Regression

- Recall Sum of Squares Error

$$SSE = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

- Hypothesis Tests to Compare Models

$$H_0 : \text{Smaller Model } \omega \text{ with dimension } q$$

$$H_A : \text{Larger Model } \Omega \text{ with dimension } p$$

$$F = \frac{(SSE_\omega - SSE_\Omega)/(p-q)}{SSE_\Omega/(n-p)} \sim F(p-q, n-p).$$

$$\text{Reject } H_0 \text{ if } F > F_\alpha(p-q, n-p).$$

Examples and Exercises in Faraway Chapter 3

- **Example:** `gala` dataset
- **Exercise 1:** `prostate` dataset

Thanks for listening!