



JOHNS HOPKINS
BLOOMBERG SCHOOL
of PUBLIC HEALTH

Lecture 5

The classical linear regression model

The material in this video is subject to the copyright of the owners of the material and is being provided for educational purposes under rules of fair use for registered students in this course only. No additional copies of the copyrighted work may be made or distributed.

Review of key concepts from Lecture 3 and 4

- ▶ Simple linear regression model
 - ▶ $ARM = B_0 + B_1 (\text{age} - 6) + e$, $e \sim N(0, \sigma^2)$, independent



Review of key concepts from Lecture 3 and 4

- ▶ Sex adjusted relationship between ARM and age
 - ▶ $ARM = B_0 + B_1 (age - 6) + B_2 \text{ Female} + e$, $e \sim N(0, \sigma^2)$, independent



Review of key concepts from Lecture 3 and 4

- ▶ Height adjusted relationship between ARM and age
 - ▶ $ARM = B_0 + B_1 (age - 6) + B_2 (HT - 62) + e$, $e \sim N(0, \sigma^2)$, independent



Review of key concepts from Lecture 3 and 4

- ▶ Effect modification: Is the ARM vs. age relationship the same or different by sex
 - ▶ $ARM = B_0 + B_1 (age - 6) + B_2 \text{ Female} + B_3 (age - 6) \text{ Female} + e$, $e \sim N(0, \sigma^2)$, independent



Multiple Linear Regression Model

- ▶ Y is a random variable representing the outcome of interest in the population
- ▶ The explanatory variables, X_1, X_2, \dots, X_p are fixed/known (not random or measured with error)
- ▶ Sample of size n is observed, data are:

$$Y_i = \mu_i(\beta, X_i) + \varepsilon_i$$

- ▶ X is the design matrix
- ▶ X_i is the row of the design matrix corresponding to subject i



Multiple Linear Regression Model

$$Y_i = \mu_i(\beta, X_i) + \varepsilon_i$$

- ▶ Systematic component:
 - ▶ $\mu_i(\beta, X_i)$
- ▶ ε_i is the random components:
- ▶ The least squares solution finds the values of β that minimize:



Least squares solution: simple linear regression



Maximum likelihood inference in MLR

- ▶ Start with the MLR:

- ▶ Other notation:



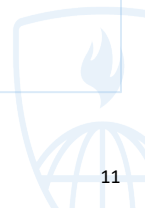
Likelihood function definition

- ▶ Model:
- ▶ Probability density function:
- ▶ Likelihood function:



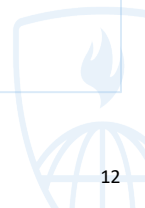
Maximum likelihood estimation under gaussian residuals

- ▶ Likelihood function



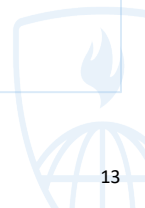
Maximum likelihood estimation under gaussian residuals

▶ Log Likelihood Function



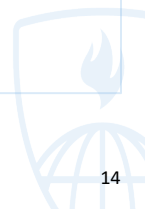
Maximum likelihood estimation under gaussian residuals

- ▶ Solution for β_j



Maximum likelihood estimation under gaussian residuals

- ▶ Solution for β_j

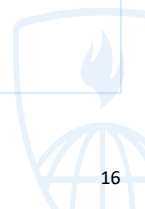


Maximum likelihood estimation under gaussian residuals

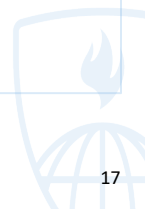
- ▶ Solution for σ^2



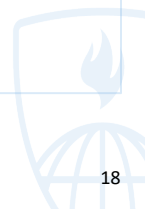
MLEs for simple linear regression



MLEs for simple linear regression



MLEs for simple linear regression



Take away messages



Take away messages



Next time....

- ▶ Vector / Matrix representation of MLR
- ▶ Geometry of least squares
- ▶ Distribution of MLEs for regression parameters

