

Lecture 34: Chapter 13, Section 1

Two Quantitative Variables

Inference for Regression

- Regression for Sample vs. Population
- Population Model; Parameters and Estimates
- Regression Hypotheses
- Test about Slope; Interpreting Output
- Confidence Interval for Slope

Looking Back: *Review*

□ 4 Stages of Statistics

- Data Production (discussed in Lectures 1-4)
- Displaying and Summarizing (Lectures 5-12)
- Probability (discussed in Lectures 13-20)
- Statistical Inference
 - 1 categorical (discussed in Lectures 21-23)
 - 1 quantitative (discussed in Lectures 24-27)
 - cat and quan: paired, 2-sample, several-sample (Lectures 28-31)
 - 2 categorical (discussed in Lectures 32-33)
 - 2 quantitative

Regression Line and Residuals (*Review*)

Summarize linear relationship between explanatory (x) and response (y) values with line $\hat{y} = b_0 + b_1x$ minimizing sum of squared prediction errors $y_i - \hat{y}_i$ (called *residuals*). Typical residual size is

$$s = \sqrt{\frac{(y_1 - \hat{y}_1)^2 + \dots + (y_n - \hat{y}_n)^2}{n-2}}$$

- **Slope:** predicted change in response y for every unit increase in explanatory value x
- **Intercept:** predicted response for $x=0$

Note: this is the line that best fits the *sampled* points.