Intervals for Prediction

READING:	2.4
EXERCISES:	CH 2. 11abcd
ASSIGNED:	HW 6
PRODUCER:	DR. MARIO

IMG CREDIT: ALEX RIEGERT-WATERS

Example: Fatalities

- Question: *Is there a linear relationship between the **percent of young drivers** in a state and the **number of vehicle fatalities per 1,000 people** in a state?*
- Question: *What would we predict the number of vehicle fatalities of a state to be where **20%** of the state's drives are young drivers?*
- Recall: A Young Driver is Between 15 and 24 (inclusive)

Supplement for Lecture 12

- Estimate Linear Regression Model
- Correlation test
- Extract R^2 and Interpret
- Predict for $x^* = 20$
- Interpretation of Prediction when $x^* = 20$

Truth About The Fitted Line

- Line Represents the **Average Value** of *Y* (μ_Y) for a Given Value of *X*
- CI's for β_0 and β_1 Indicate **Uncertainty** in the Fit of the Line
- Therefore, We Have Uncertainty about μ_Y for a Given Value of *X*
- Standard Error of Regression $\hat{\sigma}_\epsilon$ Measures Are Uncertainty a **Little**
- We Would Be **More Confident** in Making Predictions of μ_Y for **Typical** Values of *X*

Confidence Interval for μ_Y

- Formula:

$$\hat{y} \pm t_{0.025,n-2} SE_{\hat{\mu}}$$

- Standard Error

$$SE_{\hat{\mu}} = \hat{\sigma}_{\epsilon} \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum (x - \bar{x})^2}}$$

Making Predictions for Single Observation

- Suppose We Fit a Simple Linear Regression

$$\widehat{Grade} = 75 + 4(Hours\ Spent\ Studying)$$

- You Studied 1 Hour But You May Not Be a **Typical/Average** Student
- I May Predict You to Get a 79%, but I Am **More Uncertain**
- Why? Different Students Who Studied 1 Hour Will All Not Get the Same Grade Because of ϵ in the Population Model

$$Grade = \beta_0 + \beta_1(Hours\ Spent\ Studying) + \epsilon$$

Prediction Interval for Single Observation

- Formula:

$$\hat{y} \pm t_{0.025,n-2} SE_{\hat{y}}$$

- Standard Error

$$SE_{\hat{y}} = \hat{\sigma}_{\epsilon} \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{\sum (x - \bar{x})^2}}$$

Supplement for Lecture 12

- Predict for $x^* = 20$
  - Confidence Interval for Mean of *Y*
  - Prediction Interval for Y for a Specific State
- Interpretation of 95% Confidence Interval
- Interpretation of 95% Prediction Interval
- Visual Comparing the Intervals Around the Fitted Regression Line

Conclusions

- Confidence Intervals for μ_Y are Always Smaller Than Prediction Intervals for *Y* Assuming We are Predicting for the Same Value of *X*
- Confidence Intervals for μ_Y Represent Where We Believe the True Line to be if Fitted to All the Data in the Population
- Prediction Intervals for *Y* Represent Where We Would Predict an Individual *Y* Value to be for Different Values of X

Thank You

Make Reasonable Decisions