Stats Lesson 70 - Inference for Linear Regression – Confidence Intervals

Confidence Interval for Slope

We are going to learn how to construct a confidence interval to estimate the unknown slope *b* of the population (true) regression line because that is usually the most important parameter in the regression problem.

Step 1: State the name of the confidence interval and the parameter of interest that we are estimating.

Step 2: **Conditions for Regression Inference (Confidence Interval and Hypothesis Test)**

Linear – The actual relationship between *x* and *y* is **L**inear. For any fixed value of *x*, the mean response μ_y falls on the population (true) regression line $\mu_y = \alpha + \beta x$ The slope *b* and intercept *a* are unknown parameters.

Independent – Individual observations are **I**ndependent of each other.

Normal – For any fixed value of x, the response *y* varies according to a **N**ormal distribution. In other words the residuals should have an approximately normal distribution. A stemplot of the residuals will show this.

Equal variance – The standard deviation of *y* is the same for all values of x. In other words, the residuals have **E**qual variance. A residual plot will show this – We want the residuals to appear random and we do not want the residuals to follow a pattern (growing, shrinking, etc.)

Random – The data comes from a well-designed **R**andom sample or a randomized experiment.

**********Remember **L.I.N.E.R**.**********

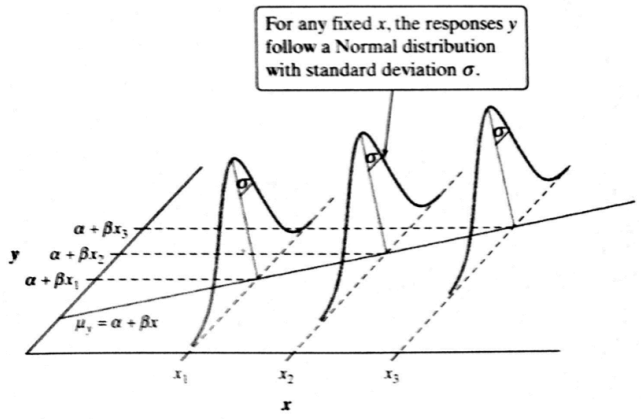


FIGURE 12.6 The regression model when the conditions for inference are met. The line is the population (true) regression line, which shows how the mean response μ_y changes as the explanatory variable *x* changes. For any fixed value of *x*, the observed response *y* varies according to a Normal distribution having mean μ_y and standard deviation σ.

Step 3: Calculations

From the AP stat packet: statistic ± (critical value) · (standard deviation of statistic)

IN AP STAT PROGRAM: Run Linear Regression, then Error Analysis

Error: 1.76 • .0007414 = .0013 Interval: (-.00475, -.00214)

Step 4: Interpretation

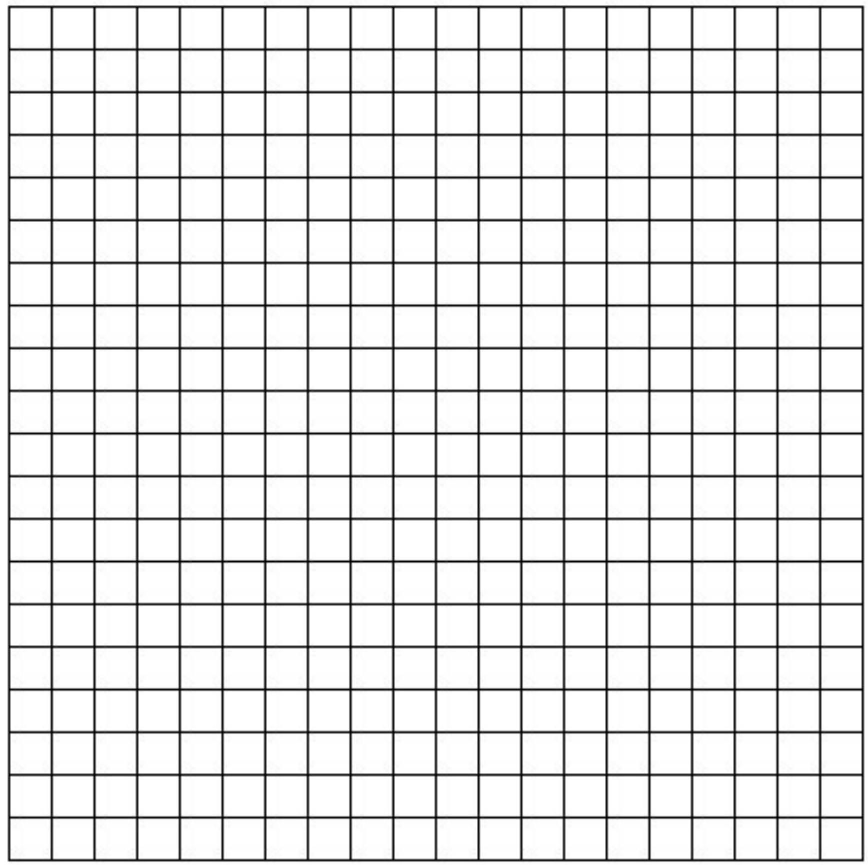
We are 90% confident that the slope of the true regression line is captured by the interval: (-.00475, -.00214). zero is not in the interval, so we have evidence that a negative association exists.

Significant results are reached when zero is NOT in the interval.

Class Data

Suggested Variables:
- How many days of school you have missed this year.
- Hours of sleep you had last night.

Days missed	Sleep last night



Class Data as Homework

<u>Suggested Variables</u> – class will decide the variables
