

Final Exam – Part 2

Jasmeet Singh Saini - 0758054

2023-04-17

Murderous Poverty

The table below shows the output of a linear model annual murders per million (annual_murders_per_mil) from percentage living in poverty (perc_pov) in a random sample of 20 metropolitan areas.

- a. Write the equation of the regression line. Interpret the slope and intercept in context.
- b. What are the hypotheses for evaluating whether the slope of the model predicting annual murder rate from poverty percentage is different than 0?
- c. What are the degrees of freedom of the test statistic?
- d. State the conclusion of the hypothesis test from part (b) in context of the data. What does this say about whether poverty percentage is a useful predictor of annual murder rate?
- e. Calculate a 95% confidence interval for the slope of poverty percentage, and interpret it in context of the data.

Highschool Senior Citizens

The National Center of Education Statistics conducted a survey of high school seniors, collecting test data on reading, writing, and several other subjects. Here we examine a simple random sample of 200 students from this survey. Side-by-side box plots of reading and writing scores as well as a histogram of the differences in scores are shown below. Also provided below is a histogram of randomized averages of paired differences of scores (read - write), with the observed difference ($\bar{x}_{\text{read-write}} = -0.545$) marked with a red vertical line. The randomization distribution was produced by doing the following 1000 times: for each student, the two scores were randomly assigned to either read or write, and the average was taken across all students in the sample.

- a. Is there a clear difference in the average reading and writing scores?
- b. Are the reading and writing scores of each student independent of each other?
- c. Create hypotheses appropriate for the following research question: is there an evident difference in the average scores of students in the reading and writing exam?
- d. Is the average of the observed difference in scores ($\bar{x}_{\text{read-write}} = -0.545$) consistent with the distribution of randomized average differences? Explain.
- e. Do these data provide convincing evidence of a difference between the average scores on the two exams? Estimate the p-value from the randomization test, and conclude the hypothesis test using words like “score on reading test” and “score on writing test.”