

Lesson 5

Background

Let us re-visit the investigation of the gender wage gap started in Lesson 2. The dataset we are using is based on a portion of the Current Population Survey Uniform Extracts from the Center for Economic and Policy Research (Center, 2019).¹ For this analysis we will focus on a random sample of 1,400 *adult earners with Bachelors degrees who made \$150,000 or less per year*.

Step 1: Ask a research question. What is the research question?

Step 2: Design a study and collect data. Use no more than two sentences to describe the data. Be sure to indicate how many observational units are in the analysis and what values the variables take on.

Step 3: Explore the data. Use boxplots to plot the conditional distributions of income for males and females. Consider whether there appears to be an association between the two variables.

Step 4: Draw inferences beyond the data. The difference in income is practically meaningful, but is it beyond what we would expect to observe by chance? That is, is the difference in income statistically significant at the $\alpha = 0.05$ significance level? Write the null and alternative hypothesis in words and symbols.

We can use two approaches to determine statistical significance: the 3S strategy or the theory-based approach.

A. 3S Strategy

Use the Comparing Groups applet. Make sure the explanatory and response variables are ordered to match the button above the data and press ‘Use Data’. Check the box to ‘Show Groups’. Check the ‘Show Shuffle Options’ box and select the ‘Plot’ radio button. Use at least 1000 shuffles.

- a) What is the sample statistic (difference in means)?

¹All variable modifications are documented here.

b) What is the standardized statistic?

c) Is there evidence of a statistically significant difference in the mean income for males and females?

d) Interpret the 95% confidence interval.

B. Theory-based approach

a. What theory-based approach could we use to determine if there was a statistically significant difference in incomes between the sexes?

b. Are the validity conditions for this theory-based approach satisfied?

c. Are the results for the theory-based approach consistent with those from the 3S Strategy?

Step 5: Formulate conclusions. Summarize the conclusions we could draw from this study. How far can we generalize our conclusions?

Step 6: Look back and ahead. Identify limitations to our approach. What could we do to address them in a future study?

References

Center for Economic and Policy Research. March CPS Uniform Extracts, Version 1.1. Washington, DC. <https://ceprdata.org/cps-uniform-data-extracts/cps-outgoing-rotation-group/cps-org-data/>. (2019). Accessed: 2020-08-12.