

## Math 140 Project 3 Instructions Spring 2020 / Teachout

### Project 3: Two-population Mean Average Hypothesis Test

#### Claim

In part 1 of the project, you collected some categorical and quantitative data and separated your quantitative data by groups. Come up with a claim about the population mean average for group 1 and group 2. Maybe you think that the population mean average age of smokers is less than for non-smokers. Maybe you think that the population mean average age of smokers is greater than for non-smokers. Maybe you think the population ages are the same.

#### Null and Alternative Hypothesis

- Label your population means. Use your claim to write your null and alternative hypothesis. Which is the claim? Is this a right-tailed, left-tailed or two-tailed test?
- *Sample null and alternative hypothesis*

$\mu_1$ : The population mean average age of smokers.

$\mu_2$ : The population mean average age of non-smokers.

$H_0$ :  $\mu_1 = \mu_2$  (The population mean average age is the same for smokers and non-smokers. This also indicates that there is no relationship between the age and smoking status.)

$H_A$ :  $\mu_1 < \mu_2$  (The population mean average age for smokers is less than for non-smokers. This also indicates that there is a relationship between the age and smoking status.) **CLAIM**

**This is a left-tailed test!**

#### Statistics Software Printout

Use Statcato or Statcrunch to perform the two-population mean hypothesis test. Your hypothesis test printout should have the critical value, the T-test statistic, the significance level 5% and the P-value.

#### Histograms

Create two histograms, one for each of your two separated quantitative data sets. (For example a histogram for the ages of smokers and a histogram for the ages of non-smokers.)

#### Assumptions for 2 population Mean

Check the following assumptions.

- *The two samples were collected randomly or represents the population*
- *Individuals within and between the samples are independent of each other*
- *Both sample sizes are at least 30 or nearly normal.*

#### T-Test Statistic

Give the two-population mean T-test statistic. Write a two-population mean T-test statistic sentence to explain what the T-test statistic tells you about the sample means. Did the test statistic fall in a tail determined by a critical value?

#### Significance

Does the sample data significantly disagree with the null hypothesis? Why or why not? Are the sample means significantly different? Why or why not?

#### P-value

Give the P-value proportion and convert it into a percentage. Write a sentence to explain the P-value. Explain why the P-value is lower or higher than the significance level.

### Sampling Variability (Random Chance)

Is the null hypothesis was true, could this sample data occur because of sampling variability or is it unlikely? Explain why.

### Reject $H_0$ or Fail to reject $H_0$ ?

Does the sample data indicate that we should reject the null hypothesis or fail to reject the null hypothesis? Explain.

### Conclusion

Write a conclusion in context addressing your claim and evidence.

### Related or not related

Does the hypothesis test indicate that the categorical data you collected is related to the quantitative data you collected or not related? Explain.

### **Grading Rubric**

- Statistics software printout with test statistic, critical value, significance level 5% and P-value (7% of grade)
- Two histograms (5% each)
- Claim (5% of grade)
- Null hypothesis (7% of grade)
- Alternative Hypothesis (7% of grade)
- What tailed test is it? (5% of grade)
- Checking all three assumptions (5% each)
- Test statistic sentence (7% of grade)
- Significance explanation (7% of grade)
- P-value sentence (7% of grade)
- Sampling Variability explanation (7% of grade)
- Reject or Fail to reject the null hypothesis with explanation (7% of grade)
- Conclusion (7% of grade)
- Related or not related explanation (7% of grade)