SOC 5050: Lab 10

Christopher Prener, Ph.D.

October 24th, 2016

Directions

Please complete all steps below. Your final work by do-file, log-file, plots, and markdown file with answers should be uploaded to your GitHub assignment repository by 4:20pm on Monday, October 31st, 2016. You can open these data by using the following command:

```
use http://www.ats.ucla.edu/stat/stata/notes/hsb2, clear
```

Part 1: One-sample T-test

1. Test to see whether the sample data in the variable `socst` comes from a population where the average score on the social science portion of a standardized test is $\mu = 54$. Be sure to provide a complete interpretation of the results.
2. Test to see whether the sample data in the variable `science` comes from a population where the average score on the science portion of a standardized test is $\mu = 54$. Be sure to provide a complete interpretation of the results.
3. Create a well-formatted plot showing the mean and confidence interval for the distribution of `science` and $\mu = 54$.

Part 2: Independent T-test

4. Using the variables `female` and `science`, test to see whether the assumption of homogeneity of variance holds.
5. Based on your answer to question 4, conduct the appropriate independent t-test to see whether there is a significant difference in science scores between men and women in this sample. Be sure to provide a complete interpretation of the results.
6. Based on your answer to question 4, calculate and interpret the appropriate effect size.

7. Create a well-formatted plot showing the means and confidence intervals of science scores between men and women in this sample.

Part 3: Dependent T-test

8. Since there is overlap between writing and social science skills, it is possible that these two scores are not independent. Test to see whether there is a significant difference in writing and social science scores in this sample. Be sure to provide a complete interpretation of the results.
9. Create a well-formatted plot showing the means and confidence intervals of social science and writing scores in this sample.

Part 4: Power Analysis

10. Using the means and standard deviation data in this lab as "pilot data"[1], conduct a power analysis using the following parameters: $\alpha = .05$, $\beta = .9$, d = small. What is the sample size needed to detect differences of that magnitude given the parameters above? This power analysis should focus on gender differences related to science scores.
11. Conduct a second power analysis using the following parameters: $\alpha = .05$, $\beta = .9$, d = moderate. What is the sample size needed to detect differences of that magnitude given the parameters?
12. Conduct a third power analysis using the following parameters: $\alpha = .05$, $\beta = .8$, d = large. What is the sample size needed to detect differences of that magnitude given the parameters?

[1] Imagine a situation where this relatively small sample was designed as a pilot study where questions were tested and concepts explored in terms of their viability for a larger study. This is a common research strategy in the health sciences, survey research, and larger social science projects. These pilot data are used to construct expectations and power analyses for a larger, grant-funded study.

Document Details

Document produced by Christopher Prener, Ph.D. for the Saint Louis University course SOC 5050 - QUANTITATIVE ANALYSIS: APPLIED INFERENTIAL STATISTICS. See the course wiki and the repository `README.md` file for additional details. Data are drawn from the ULCA Institute for Digital Research and Education.