Portfolio 2
PA 606: Seminar in Quantitative Techniques (Spring 2020)

25 points

Due: April 4, 2020

Topics:

t-Test (t), Chi Square (X^2), Correlation (r)

Data Sources:

`GSS2014_final.csv` (**data set**: 2014 General Social Survey – GSS)
`GSS_Codebook.pdf` (**codebook**: description of 2014 GSS and measurement of variables)
`GSS_Variables_and_Descriptions.pdf` (**descriptions**: names and titles for 2014 GSS)

Overview:

Answer each of the questions below, **in full sentences/paragraphs**, and show your tables/plots (if necessary) inline (interspersed within the text). For full credit, you must append (copy and paste) your R Script at the end of this portfolio as the final page(s).

Problems

Data Set Information

- Describe the data set, including the name and who administered the data/survey (e.g. which survey research firm), year administered, the unit of analysis, and the number of observations.

t-Test

In this test, you'll be examining mean differences in **occupational prestige** of a respondent's job (`sei10`) by their **sex** (`sex`).

- Describe the variables. Using the codebook (`GSS_Codebook.pdf`) and the list of variable descriptions (`GSS_Variables_and_Descriptions.pdf`) for the `GSS2014_final.csv` data set, for each variable, describe its text (e.g. the question asked for each variable), the level of measurement, and the values/categories within the variable.
- Define a research question for the variables (e.g. "*Is variation in X related to/associated with variation in Y; Is there a mean difference in Y by categories of X*").
- Define the null hypothesis (H_0) and the alternative hypothesis (H_1) for this test.

- Describe the various assumptions of the t-Test (t) and how you would assess them. Next, show all necessary tables/plots that demonstrate your assessment of whether or not you've met the assumptions of the test. (*Note: if showing tables, create them, do not simply copy and paste from your output.*)
- Run the test. Fully and correctly report the test. If you find significance, describe what you find in terms of mean differences by group (e.g. compare the group means).

Chi Square

In this test, you'll be examining the association between the **highest degree** a respondent completed (`degree`) by their belief about whether people should have police **permits** to purchase a gun (`gunlaw`).

- Describe the variables. Using the codebook (`GSS_Codebook.pdf`) and the list of variable descriptions (`GSS_Variables_and_Descriptions.pdf`) for the `GSS2014_final.csv` data set, for each variable, describe its text (e.g. the question asked for each variable), the level of measurement, and the values/categories within the variable.
- Define a research question for the variables (e.g. "*Is variation in X related to/associated with variation in Y; Is there a mean difference in Y by categories of X*").
- Define the null hypothesis (H_0) and the alternative hypothesis (H_1) for this test.
- Describe the various assumptions of the Chi Square (X^2) test and how you would assess them. Next, show all necessary tables/plots that demonstrate your assessment of whether or not you've met the assumptions of the test. (*Note: if showing tables, create them, do not simply copy and paste from your output.*)
- Run the test. Fully and correctly report the test. If you find significance, describe how the attributes of one variable align with attributes of the other.

Correlation

In this test, you'll be examining the relationship between the **level of education** (`educ`) by their belief about **whether American culture is undermined by immigrants** (`immcult`).

- Describe the variables. Using the codebook (`GSS_Codebook.pdf`) and the list of variable descriptions (`GSS_Variables_and_Descriptions.pdf`) for the `GSS2014_final.csv` data set, for each variable, describe its text (e.g. the question asked for each variable), the level of measurement, and the values/categories within the variable.
- Define a research question for the variables (e.g. "*Is variation in X related to/associated with variation in Y; Is there a mean difference in Y by categories of X*").
- Define the null hypothesis (H_0) and the alternative hypothesis (H_1) for this test.
- Describe the various assumptions of the Correlation (r) and how you would assess them. Next, show all necessary tables/plots that demonstrate your assessment of whether or not you've met the assumptions of the test. (*Note: if showing tables, create them, do not simply copy and paste from your output.*)
- Run the test. Fully and correctly report the test. If you find significance, also describe the strength and direction of relationship and the effect size (e.g. where you explain how much variation in Y is explained by variation in X).

Extra Credit: Bivariate Tests and Variables of Your Choice

- Select two variables from the data set for which you are able to test a bivariate relationship. These variables must be distinct from the variables used in the tests above.
- Describe the variables. Using the codebook (`GSS_Codebook.pdf`) and the list of variable descriptions (`GSS_Variables_and_Descriptions.pdf`) for the `GSS2014_final.csv` data set, for each variable, describe its text (e.g. the question asked for each variable), the level of measurement, and the values/categories within the variable.
- Define a research question for the variables (e.g. "*Is variation in X related to/associated with variation in Y; Is there a mean difference in Y by categories of X*").
- Given the level of measurement, describe the kind of bivariate test you ***WOULD*** run.