SOC 4015/5050: Lab-15 - Chi-Squared

Christopher Prener, Ph.D.

Fall 2018

Directions

Please complete all steps below. All work should be uploaded to your GitHub assignment repository by 4:00pm on Monday, December 17th, 2018. All data can be obtained from the `testDriveR` package's `auto17` data set.

Analysis Development

Using RStudio and your operating system's file manager, create an R Project in the *existing* directory in your assignments repository named `Lab-14`. Add a `README.md` file, notebook, and all necessary folders before beginning.[1]

[1] This initial section follows the project workflow that is available in the `lecture-03` repo!

Data Preparation

1. Create a new logical variable that is `TRUE` if the vehicle is a "German" vehicle (i.e. one made by BMW, Mercedes, Porsche, and Volkswagen) and `FALSE` otherwise.
2. Subset your data so that it contains only the `id`, your new logical variable, and the `driveStr` variables.

Create Tables

Using the data created in Part 1, answer the following questions.

3. Create a two-way table of the logical variable you created above and `driveStr` using `janitor` that includes
   (a) a total row at the bottom and a total column,
   (b) properly formatted row percents that are display three decimal places,
   (c) and frequency values in the "front" position.

Fit the Chi-Square and Check Assumptions

Using the data created in Part 1, answer the following questions.

4. Fit and interpret the results of a chi-squared test comparing the relationship of German vehicles to drivetrains. Is there a meaningful relationship between these two variables?

5. Does this model violate the Cochran conditions assumption?

6. Regardless of your answer above, fit a Fisher's Exact Test on these same data. Does this change your interpretation of question 6?