SOC 4015/5050: Lab-04 - Working with Random Variables

Christopher Prener, Ph.D.

Fall 2018

Directions

Please complete all steps below. The final parts of this lab use the `auto17` data from the `testDriveR` package. Your answers "by hand" should be scanned and submitted as a `pdf` image. This assignment should be uploaded to your Assignments Repository by 4:15pm on Monday, October 1st, 2018.

Analysis Development: Create a Project Folder System

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

1. Using RStudio, add an R Project to the *existing* directory in your assignments repository named `Lab-04`.
2. Add a new folder named `docs` to you project.
3. Create a new text file for your `README.md`. In the body of your `README.md` file, use Markdown formatting to write a sentance or two describing the purpose of this project. Then create an outline using bullets of the contents of the project itself.[1]
4. Create a new notebook with an expanded YAML heading.
5. Make sure your notebook has *completed* introductory, package loading, and data loading sections before proceeding with the parts below.
6. Be sure to "knit" your notebook at the end!

[1] See my write-up of the Markdown syntax in *Sociospatial Data Science* for details on creating lists.

Part 1: Binomial Distribution

Complete this section in R/RStudio.

7. What is the probability of more than 24 successes occurring in a sequence of 250 independent trials with a binary outcome where the probability of success is .4 for each trial?

8. What is the probability of 25 or fewer successes occurring in a sequence of 250 independent trials with a binary outcome where the probability of success is .4 for each trial?
9. What is the probability of exactly 25 successes occurring in a sequence of 250 independent trials with a binary outcome where the the probability of success is .4 for each trial?

Part 2: Poisson Distribution

Complete this section in R/RStudio.

10. The probability of a catastrophic failure of a rocket carrying satellites into space is .025. Over 1,000 launches, what is the probability of observing more than 4 failures?
11. What is the probability of observing exactly 18 failures?
12. What is the probability of observing 15 for fewer failures?

Part 3: Normal Distribution

Complete this section in R/RStudio.

13. A literature review shows the distribution of literacy test scores on a given instrument to be normally distributed. The population average test score is 21 with a standard deviation of 3. What is the probability of drawing a individual whose score is a 25?

Part 4: Skew and Kurtosis

Complete this section by hand.

14. The following are a distribution of scores on a simple functional capacity task for individuals recovering from a stroke: 1, 4, 3, 2, 4, 2, 1, 4, 3, 3. What is the skewness and kurtosis of this distribution of scores?

Part 5: Normality Testing in R

Complete this section in R/RStudio.

15. Use the variable `fuelCost` from the `auto17` data set in the `testDriveR` package to conduct a full set of normality tests:

    (a) What is the variable's skew?

    (b) What is the variable's kurtosis?

    (c) Create and interpret a q-q plot.

    (d) What are the results of a Shapiro-Francia test?