SOC 5050: Lab 06

Christopher Prener, Ph.D.

September 5th, 2016

Directions

Please complete all steps below. The final parts of this lab use the 2011 CPS data. Your final do-file, log-file, plots, and markdown file with answers should be uploaded to your GitHub assignment repository by 4:20pm on Monday, October 3rd, 2016.

Part 1: Random Variables

1. 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?
2. What is the probability of drawing a individual whose score is a 19?
3. What is the probability of 25 or more successes occurring in a sequence of 250 independent trials with a binary outcome where the probability of success is .08 for each trial?
4. 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 .08 for each trial?
5. 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 .08 for each trial?
6. The probability of a catastrophic failure of a rocket carrying satellites into space is .025. Over 800 launches, what is the probability of observing 5 or more failures?
7. What is the probability of observing exactly 18 failures?
8. What is the probability of 40 or more successes occurring in a sequence of 50 trials with a binary outcome where the probability of success is .3?

9. What is the probability of 40 successes occurring in a sequence of 50 trials with a binary outcome where the probability of success is .8?
10. The probability of becoming infected with a particular virus is .01 in a village of 3,000 residents. What is the probability of observing exactly 24 infections in that village?
11. What is the probability of observing 12 or fewer infections in that village?
12. What is the probability of observing 40 or more infections in that village?

Part 2: Skew and Kurtosis by Hand

13. 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 3: Normality Testing in Stata

14. Use the variable `HRNUMHOU` 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 well laid-out p-p plot.
    (d) Create and interpret a well laid-out q-q plot.
    (e) Is the variable `HRNUMHOU` appropriate for using either the Shapiro-Wilk or Shapiro-Francia normality tests?
    (f) Regardless of your answer to the above question, run and interpret both hypothesis tests.

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.