

# *SOC 4015/5050: PS-03 - Foundations for Inference*

*Christopher Prener, Ph.D.*

*Fall 2018*

## *Directions*

Please complete all steps below. Your your work “by hand” should be scanned and included in your PS-03 assignment submission. All work should be uploaded to your GitHub assignment repository by 4:15pm on Monday, October 8<sup>th</sup>, 2018.

**Use the following scenario:** CDC data suggests that the population average cholesterol level for adults is 196. Assume that this distribution has a standard deviation of 20.

## *Analysis Development: Create a Project Folder System*

1. Using RStudio, add an R Project to the *existing* directory in your assignments repository named Lab-05.
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 sentence or two describing the purpose of this project. Then create an outline using bullets of the contents of the project itself.<sup>1</sup>
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 of the assignment!

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

<sup>1</sup> See my write-up of the Markdown syntax in *Sociospatial Data Science* for details on creating lists.

## *Part 1: Sampling Distributions*

7. Assume you draw repeated random samples of  $n=1250$  students. What is the standard error of these repeated samples?
8. If you were to draw repeated random samples of  $n=1250$  students, what proportion of these samples will have sample means greater-than or equal to 225?

9. What sample size would we need to have a sample mean that is within 10 points of the population's?

*Part 2: Predictive Intervals*

10. Calculate and interpret a 95% predictive interval for  $x$ .
11. Calculate and interpret a 99% predictive interval for  $x$ .
12. Calculate and interpret a 95% predictive interval for  $\bar{x}$ .
13. Calculate and interpret a 99.9% predictive interval for  $\bar{x}$ .

*Part 3: Confidence Intervals*

14. Calculate and interpret a 95% confidence interval assuming we draw a random sample of  $n=1000$  adults.
15. Calculate and interpret a 99.9% confidence interval assuming we draw a random sample of  $n=1000$  adults.
16. Calculate and interpret a 95% confidence interval assuming we draw a random sample of  $n=1500$  adults.
17. Calculate and interpret a 99% confidence interval assuming we draw a random sample of  $n=1500$  adults.

*Rubric*

Individual Questions					
Part 1		Part 2		Part 3	
Question	Points	Question	Points	Question	Points
7 through 9	2	10 through 13	2	14 through 17	2
<i>Points Possible</i>	6		8		8

Note: Partial credit possible

Project Organization		
Category	Details	Points
Excellent	PS-03 organized following workflow without error	3
Good	Minor concerns with organization	2.5
Improvement Needed	Significant concerns with organization	1.5
Unsatisfactory	No organizational strategy used	0
<i>Points Possible</i>		3

Notebook Formatting & RMarkdown		
Category	Details	Points
Excellent	Syntax used appropriately & without error	5
Good	Minor concerns with syntax use	4.5
Improvement Needed	Significant concerns with syntax	3
Unsatisfactory	No RMarkdown used	0
<i>Points Possible</i>		5

Literate Programming		
Category	Details	Points
Excellent	Narrative throughout with great detail	5
Good	Some narrative with inconsistent detail	4.5
Improvement Needed	Limited narrative with little detail	3
Unsatisfactory	No narrative included	0
<i>Points Possible</i>		5