**Week 8 Assignment**

Michael K. Zoucha

College of Science & Information Technology, Bellevue University

DSC 510: Introduction to Programming

Professor Eller

22 October 2021

**Week 8 Assignment**

Exercise 10-1:

Import necessary packages, and grab classes for dffmeans, correlation, and preglengthtest given in the exercise above.

**Text

Description automatically generatedText

Description automatically generated**

Function RunTests takes the sample and number of iterations, then compares the pregnancy lengths, pregnancy weights, correlation, and chi-squared of lengths, and returns all values on one line:

Text

Description automatically generated

Iterate through the RunTests function 7 times with different sizes of samples to see its effect on the power of a hypothesis test:  
Graphical user interface, text, application

Description automatically generated

Results:  
 n Test 1 Test 2 Test 3 Test 4  
9148 0.16 0.00 0.00 0.00  
4574 0.10 0.01 0.00 0.00  
2287 0.25 0.06 0.00 0.00  
1143 0.24 0.03 0.39 0.03  
571 0.81 0.00 0.04 0.04  
285 0.57 0.41 0.48 0.83  
142 0.45 0.08 0.60 0.04

Exercise 10-1:

Import necessary packages, and grab functions for sampling, resampling, and summarizing given in the exercise above.

Text

Description automatically generated

Import the BRFSS data and extract the heights and weights, also convert weights to log weights:  
Text

Description automatically generated

Estimate intercept and slope using LeastSquares from tinkstats2 package:

Graphical user interface, text, application, website

Description automatically generated  
Scatterplot #1: (Log scale)  
Chart

Description automatically generated

Scatterplot #2: (Linear scale)

Chart

Description automatically generated

Calculate residuals using thinkstats2 package, then add the residuals to the dataframe. Bin the data into groups and calculate CDF and mean for the groups, then plot the mean and CDF:  
A picture containing text

Description automatically generated

Correlation using thinkstats2 package:  
Graphical user interface, application

Description automatically generated

Coefficient of determination using thinkstats2 package:  
Graphical user interface, application

Description automatically generated

Standard deviation of predictions with and without height:  
Graphical user interface

Description automatically generated Graphical user interface, application

Description automatically generated

Difference in RMSE with and without height information:  
Graphical user interface, application

Description automatically generated

Resample 100 times and collect the sampling distribution for slope to create a CDF plot:  
Graphical user interface, text, application, website

Description automatically generated Chart

Description automatically generated

P – Value of slope: CI of Slope:  
 Graphical user interface

Description automatically generated Graphical user interface, text, application

Description automatically generated

Mean: Standard Error:

Graphical user interface, application, website

Description automatically generated Graphical user interface, application

Description automatically generated

Resample without weights:  
Graphical user interface, text, website

Description automatically generated

Resample with weights:  
Graphical user interface, text, application, website

Description automatically generated

Results:  
The mean height with weights included is almost 2 cm taller than the mean height without the weights factored in, which is significantly larger than the standard deviation for each.