Anthony Tummillo

HW2 Report

Problem 1.

(Part 1)

mean_study_data: mean = 15.0415, standard deviation = 5.0279

(Part 2)

function was written and submitted

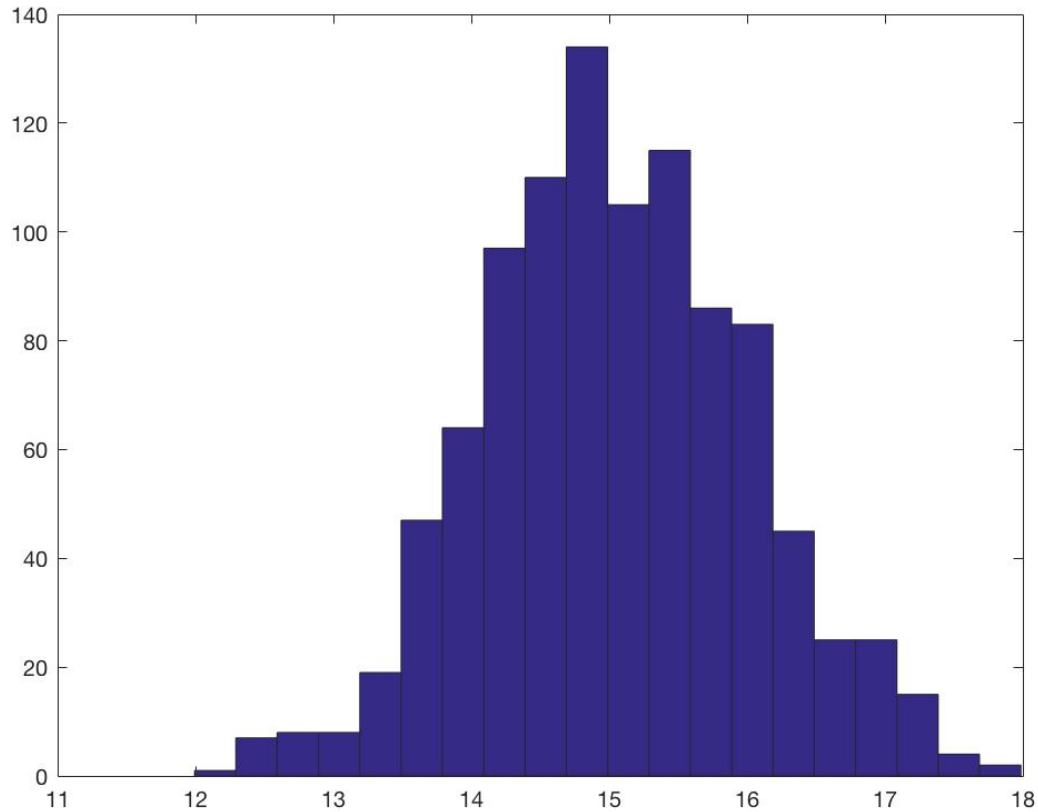
(Part 3)

completed

(Part 4)

When analyzing the mean I found in (part 1) and the histogram of means from (part 3) I observed that the histogram has the highest density of observations around the mean from part 1 (15.0415). The farther you go from 15.0415 in either direction on the graph the lower the number of observations there are. This shows with a random subset of a dataset you are most likely to get a mean that is close to the mean of the whole set, but it is still possible to calculate a mean that differs significantly.

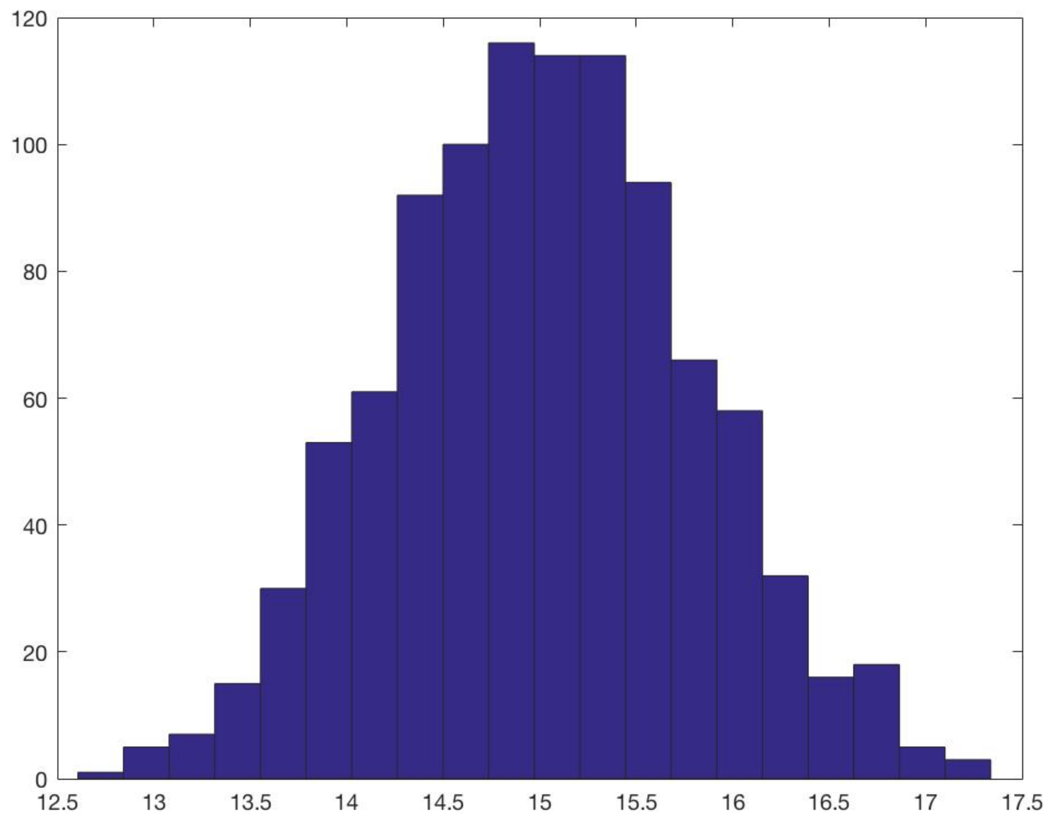
The graph is included below.



(Part 5)

The histogram for means of subsamples of size 40 is included below.

When using subsamples of size 40 the resulting histogram had an even higher density about the mean found in (part 1) (15.0415) than did the histogram from subsamples of size 25. This new histogram also had a much smaller range of means that did the one from (part 3). We can conclude that as the number of entries in a subset of a dataset increases, you are much more likely to get a mean value that is very close to that of the entire dataset. As the subset size increases the means of the subsets also vary much less from the mean of the entire dataset.



Problem 2.

(Part 1)

function was written and submitted.

(Part 2)

fold 1: mean: 3.9938, std: 4.4391
fold 2: mean: 1.8277, std: 3.6272
fold 3: mean: 2.1446, std: 2.3504
fold 4: mean: 1.7954, std: 3.1598
fold 5: mean: 2.0849, std: 3.3794
fold 6: mean: 1.7627, std: 3.2640
fold 7: mean: 2.1046, std: 3.4622
fold 8: mean: 1.0343, std: 2.5800
fold 9: mean: 1.5837, std: 3.4184
fold 10: mean: 2.4246, std: 2.2831

Problem 3.

(a)

solution: $2x$

(b)

solution: $1 + 6x^2$

(c)

solution: e^x

(d)

solution: $2x*\cos(x^2)$

(e)

solution: $-1/x^2$

(f)

solution: $-(2x+1)/(x^2+x)^2$

(g)

solution: $5/x$

(h)

solution: 0

(j)

solution: $n!/x$