Ximan Liu
1935858
DSC 423
HW1

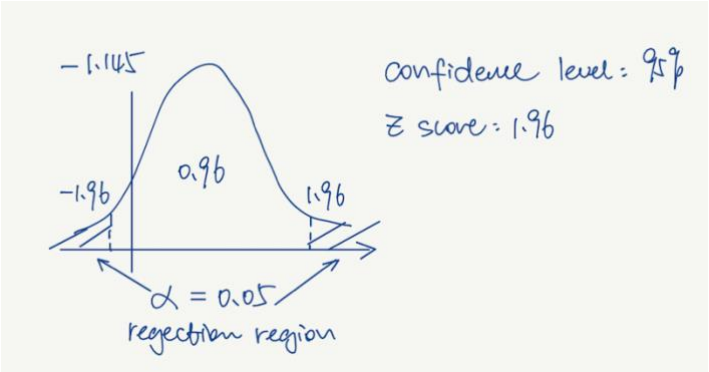
I have completed this work independently. The solutions given are entirely my own work.

1)
The option (b) "take a simple random sample of 100 graduate students studying Data Science" will have the smaller standard deviation for the amount spent on textbooks used for classes.
From the definition of standard deviation, we can tell that, it measures the average distance of the observations from the mean. Since we chose the sample of 100 graduate students studying in Data Science, it represents the samples are all selected from sample pools with similar backgrounds, which is destined to have relatively small differences between them. In that case, the average distance of the observations from the mean will be a smaller number.

2)
a)
The answer is 68%.
Since we have the 68-95-99.7 Rule, the mean is 28 and the standard deviation is 4, approximately 68% of the observations fall within σ of μ, which represents the area between the age between 24 and 32 years old.
b)
The answer is 2.5%.
[(100-95)/2] % = 2.5%

3)
Z = 2.326 is associated to 99th percentile.
2.326 * 35 + 150 = 231.41 thousand dollars.

4)



Ho: μ = 45
Ha: μ ≠ 45
Critical value = -1.96 or 1.96
Significant level α = 0.05
Test statistic: Z = (42 - 45) / (15.5 / √ 35) = -1.145
P-value = 0.125
Since |-1.145|< |-1.96| and 0.125 > 0.05, we fail to reject the null hypothesis Ho, the claim is correct.
If Ho is true, there is a 0.125 (12.5%) chance that we would see results at least as extreme as those in the sample; thus, because we saw results that are likely if Ho is true, we therefore do not have good evidence against H0 and in favor of Ha.