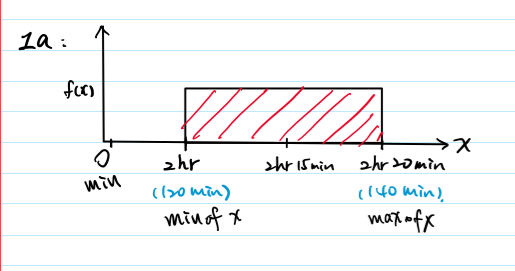


**Question 1 (3 points)**

RUBRIC: (-Write down all steps of the answer. Write down your final answers and explain as needed. Lack of each item results in 0.5-to-1-point deduction)

Delta Airlines quotes a flight time of 2 hours, 5 minutes for its flights from Cincinnati to Tampa. Suppose we believe that actual flight times are uniformly distributed between 2 hours and 2 hours, 20 minutes.

a) Show the graph of the probability density function for flight time.



b) What is the probability that the flight will be no more than 5 minutes late?

P(X <= x) = x-a/ b-a

P(X < 2hr 5mins + 5 mins) = P(X < 130) = 130 - 120 / 140 - 120 = 10 / 20 = .5

Ans: The probability that the flight will be no more than 5 minutes late is .5

c) What is the probability that the flight will be more than 10 minutes late?

P(X > x) = 1 - (x-a/ b-a)

P(X > 2hr 5mins + 10 mins) = P(X > 135) = 1-(135 - 120) / (140 -120)

= 1 - 15/20 = .25

Ans: the probability that the flight will be more than 10 minutes late is .25

d) What is the expected flight time?

expected fight time = the mean of flight time

mu = (120+140) / 2 = 130

Ans: the expected flight time is 130 mins.

**Question 2 (4 points)**

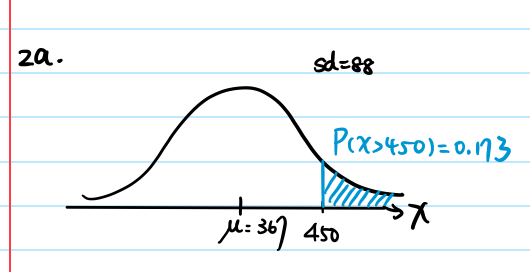
RUBRIC: (-Write down all steps and formula. -Write down (or copy) the normal distribution formula in Excel or Python. Write down your final answers. For each item Sketch or draw the normal distribution and highlight the desired area. Lack of providing the graph for each question results in 0.5-1-points grade deduction).

Automobile repair costs continue to rise with an average 2015 cost of $367 per repair (U.S. News & World Report website). Assume that the cost for an automobile repair is normally distributed with a standard deviation of $88. Answer the following questions about the cost of automobile repairs.

We know that the **mu = 367, sd = 88**, and the automobile repair costs follow a normal distribution.

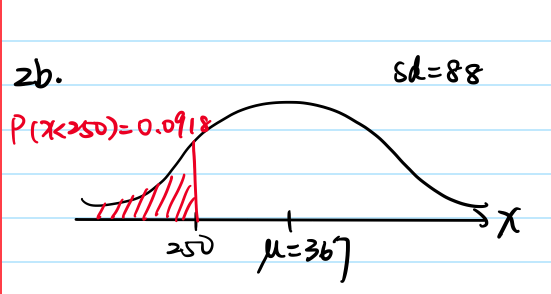
a) What is the probability that the cost will be more than $450?

P(x > 450) = 1- P(x<=450) = 1 - NORMDIST(450, 367, 88, 1) = 0.1727939558



b) What is the probability that the cost will be less than $250?

P(x < 250) = NORMDIST(250, 367, 88, 1) = 0.09183403998

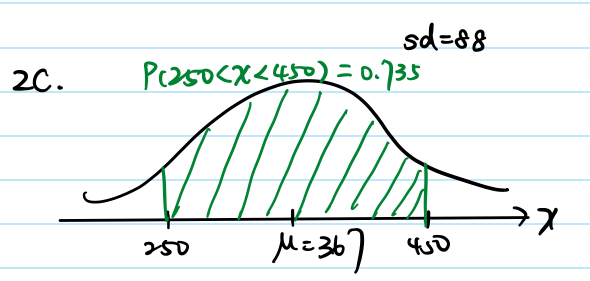


c) What is the probability that the cost will be between $250 and $450?

P (250 < x < 450) = P(x < 450) - P(x < 250)

= NORMDIST(450, 367, 88, 1) - NORMDIST(250, 367, 88, 1)

= 0.7353720042

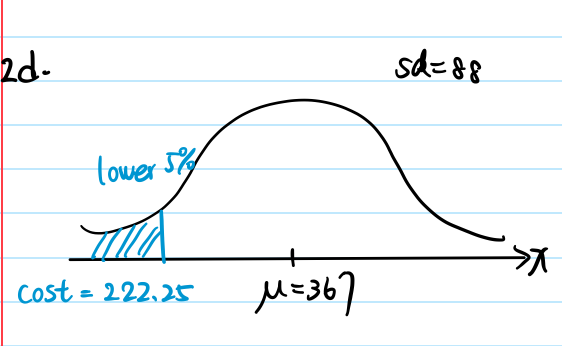


d) If the cost for your car repair is in the lower 5% of automobile repair charges, what is

your cost?

P(x < cost) = 0.05, cost = ?

cost =NORMINV(0.05, 367, 88) = 222.252881



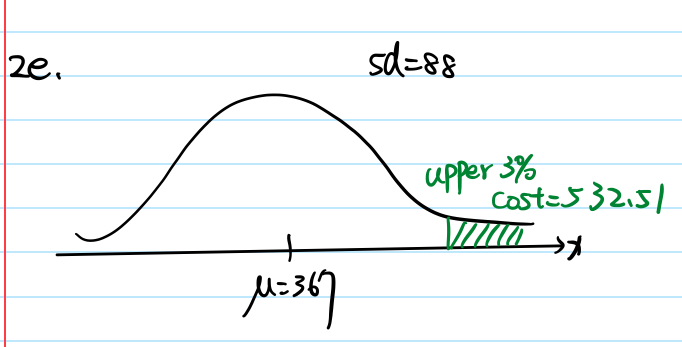
e) If the cost for your car repair is in the upper 3% of automobile repair charges, what is

your cost?

1 - upper 3% = lower 97 % (probability)

P(x < cost) = 0.97, cost = ?

cost =NORMINV(0.97, 367, 88) = 532.5098373

****

**Question 3 (3 points)**

RUBRIC: (-Write down all steps and formula. -Write down (or copy) the normal distribution

formula in Excel or Python. -Write down your final answers. -For each item Sketch or draw the

normal distribution and highlight the desired area. Lack of providing the graph for each question

results in 0.5-1-points grade deduction).

Sales personnel for Skillings Distributors submit weekly reports listing the customer contacts

made during the week. A sample of 65 weekly reports showed a sample mean of 19.5 customer

contacts per week. The standard deviation was 5.2.

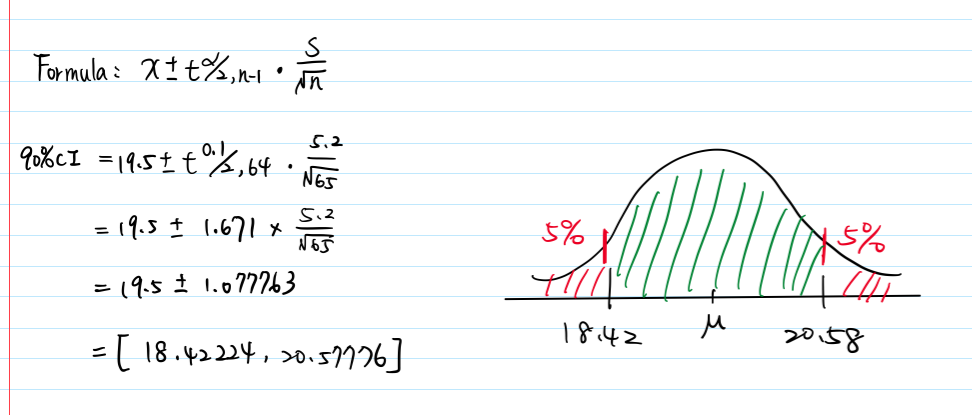
n = 65, x̄ = 19.5, sd = 5.2

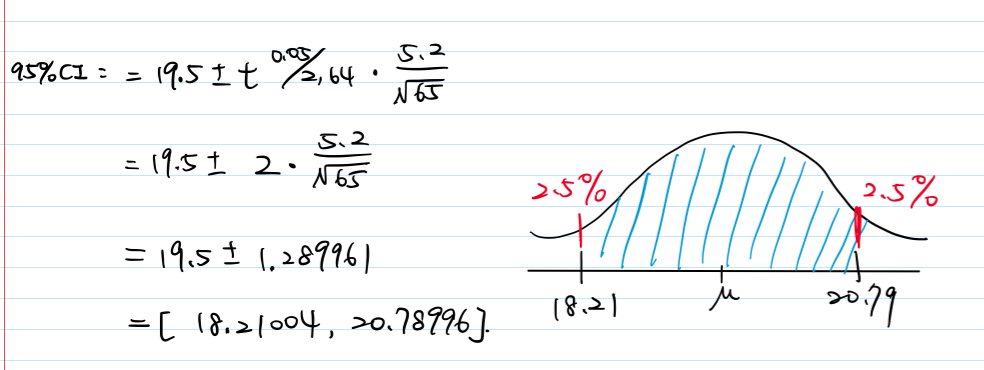
According to the t-table, we know that if the degrees of freedom is 64, the t-score is 1.671 when alpha = .1, and the t-score is 2 when alpha = .05. Also, confidence intervals are always two tailed. I use t-table to calculate part a.

T-table: <https://www.statisticshowto.com/tables/t-distribution-table/>

a) Provide 90% and 95% confidence intervals for the population mean number of weekly

customer contacts for the sales personnel.





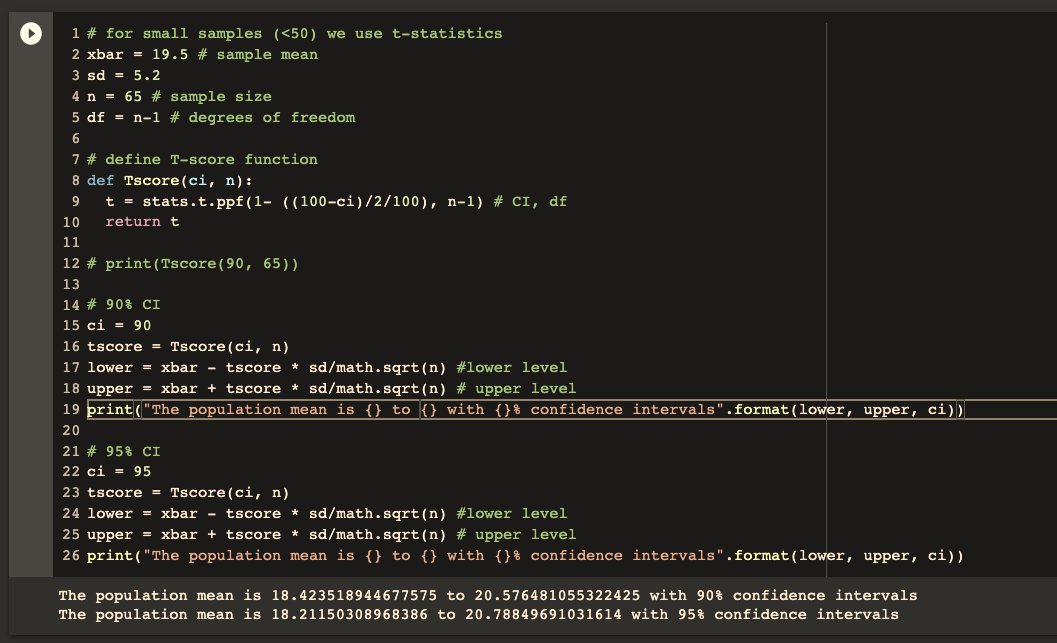
b) Provide the conclusion that you can make based on this interval estimation.

Based on the result, we have 90 % confidence that the population mean of weekly customer contacts is between 18.42 to 20.58. Also, we have 95 % confidence to say that the population mean of weekly customer contacts is between 18.21 to 20.79.

c) Repeat Part a) using a Python code. Provide the code file in a .py file and provide

comments within the code. For this problem in python use a t-test (as described in the

class).



**Question 4 (4 points)**

RUBRIC: (-Write down all steps and formula. -Write down (or copy) the normal distribution

formula in Excel or Python. -Write down your final answers).

Many medical professionals believe that eating too much red meat increases the risk of heart

disease and cancer. Suppose you would like to conduct a survey to determine the yearly

consumption of beef by a typical American and want to use **3 pounds** as the desired margin of

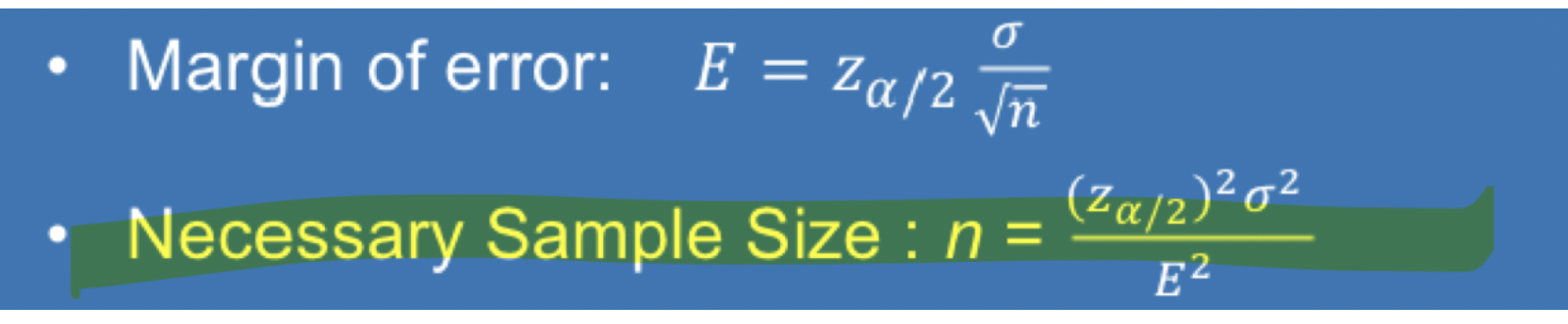
error for a confidence interval estimate of the **population mean** amount of beef consumed

annually. Use **25 pounds** as a planning value for the population standard deviation and

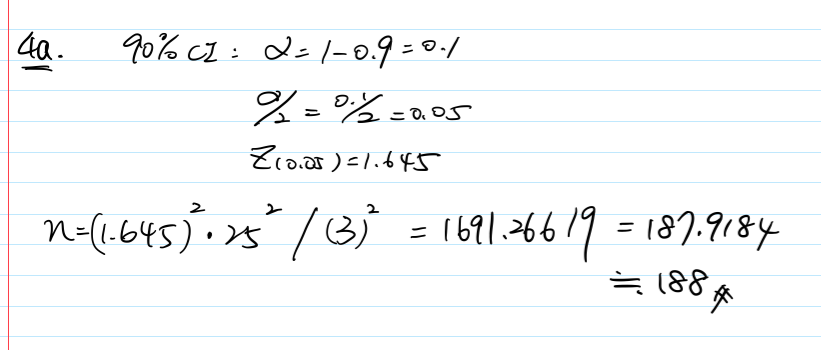
**recommend a sample size** for each of the following situations.

Note: margin of error = 3, sd = 25

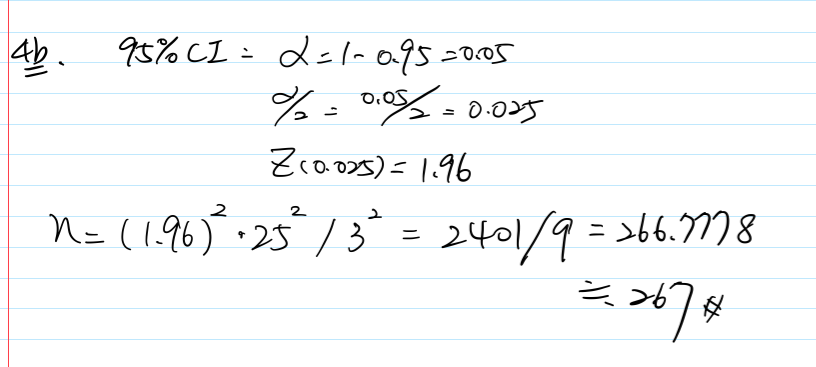
Z-score: <https://www.medcalc.org/manual/values_of_the_normal_distribution.php>



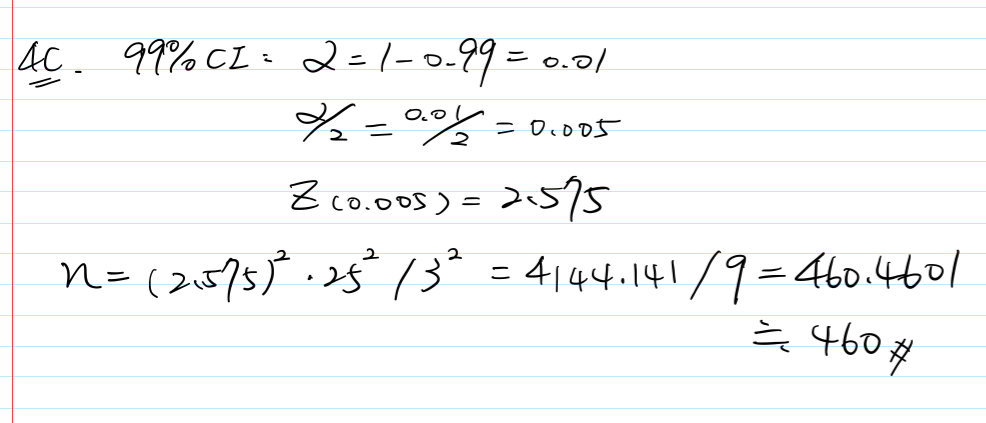
a) A 90% confidence interval is desired for the mean amount of beef consumed.



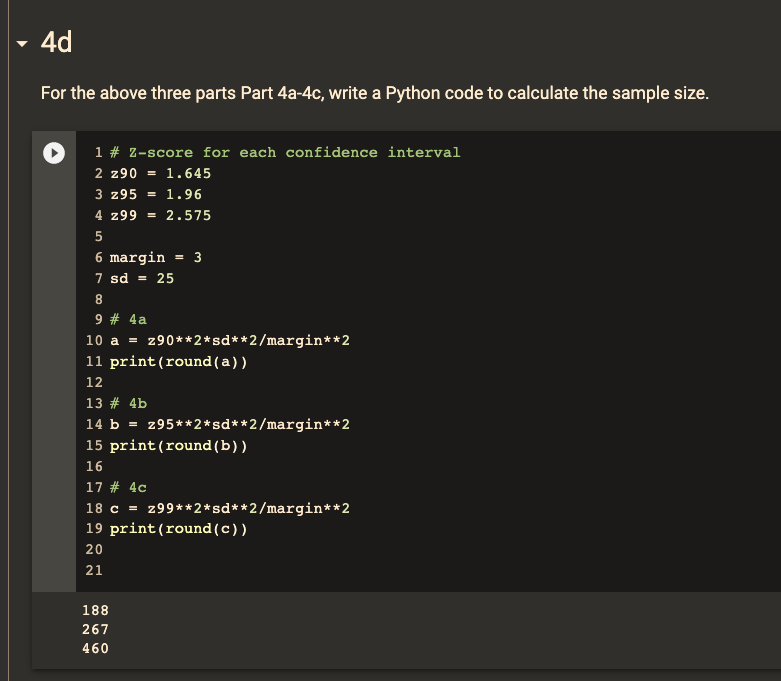
b) A 95% confidence interval is desired for the mean amount of beef consumed.



c) A 99% confidence interval is desired for the mean amount of beef consumed.



d) For the above three parts Part a-c, write a Python code to calculate the sample size.



e) When the desired margin of error is set, what happens to the sample size as the

confidence level is increased? Would you recommend using a 99% confidence interval in

this case? Discuss.

As we can see from the result above, as the confidence level increases, the appropriate sample size increases in order to keep the margin of error constant. The 99 percent confidence interval is the most precise when compared to the 90 percent, 95 percent, and 99 percent confidence intervals. The 99% CI also recommends 460 sample sizes, which is not a large number, implying that we can consider using a 99 percent confidence interval to determine a typical American's yearly consumption of beef.

**Question 5 (3 points)**

RUBRIC: (-Write down all steps and formula. -Write down (or copy) the normal distribution

formula in Excel or Python. -Write down your final answers. -For the hypothesis test provide a

graph of critical values. Lack of providing the graph for each question results in 0.5-1-points

grade deduction).

A shareholders’ group, in lodging a protest, claimed that the mean tenure for a chief executive

office (CEO) was **at least nine years**. A survey of companies reported in The Wall Street Journal

found a s**ample mean tenure of x = 7.27** years for CEOs with a standard deviation of **s = 6.38**

years.

Based on the information above, we know that the sample mean xbar is 7.27, sd = 6.38, mu = 9

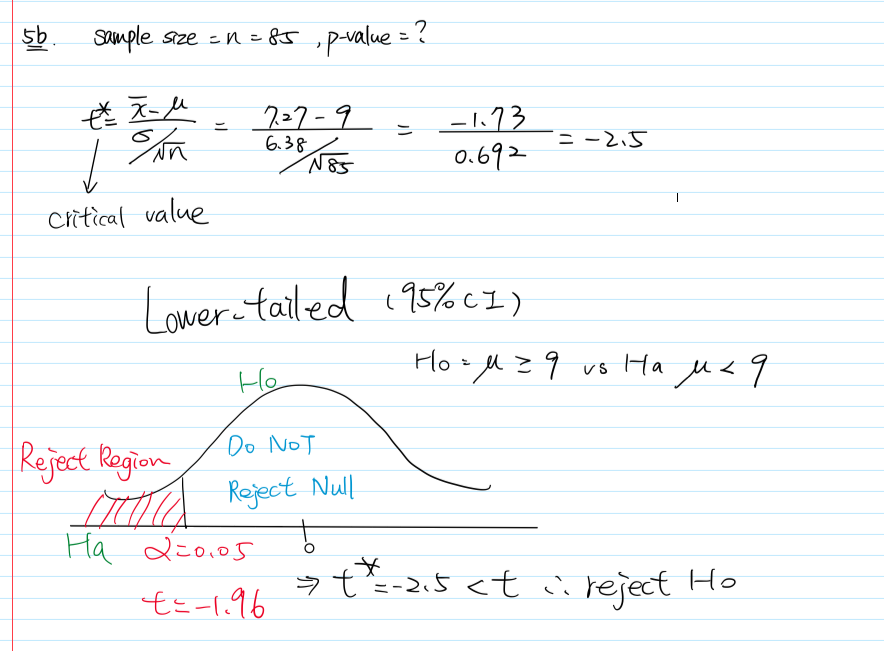
a) Formulate hypotheses that can be used to challenge the validity of the claim made by the

shareholders’ group.

Hypothesis testing: Ho: mu >= 9 years vs Ha: mu < 9 years. Also, this is a one-sided t-test.

b) Assume 85 companies were included in the sample. What is the p-value for your

hypothesis test?



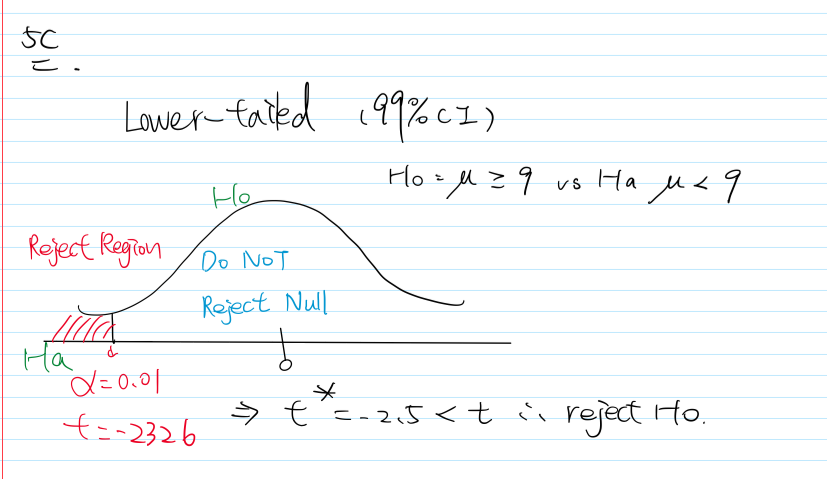
Now we know, the t is -2.5. And the degrees of freedom is n -1 = 84. The T-score calculator is then used to calculate a p-value with 84 degrees of freedom and a -2.5 t-value on a one-sided t-test. Finally, the p-value is 0.007182 with 95 % confidence interval. With the small p-value (< 0.05), we have evidence to reject the null in favor of the alternative hypothesis, meaning that the mean tenure for a CEO was NOT at least nine years or we can say the mean tenure for a CEO was less than nine years.

T-score calculator: <https://www.socscistatistics.com/pvalues/tdistribution.aspx>

c) At alpha = .01, what is your conclusion? Draw the graph and show critical values.

If the p-value is less than the significance level, we reject the null hypothesis. In the question 5b, the significance level is .05. Because .007182 < .05 so we reject the null in question 5b.

Now, in question 5c, the significance level is .01. Because .007182 is also less than .01, so we still can reject the null hypothesis, meaning that there is no sufficient evidence to conclude that the mean tenure for a CEO was at least nine years with 99 % confidence interval.



**Question 6 (3 points)**

RUBRIC: (-Write down all steps and formula. -Write down (or copy) the normal distribution

formula in Excel or Python. -Write down your final answers. -For the hypothesis test provide a

graph of critical values. Lack of providing the graph for each question results in 0.5-1-points

grade deduction).

The U.S. Bureau of Labor Statistics reports that 11.3% of U.S. workers belonged to unions in 2013. Suppose a sample of 400 U.S. workers is collected in 2018 to determine whether union efforts to organize have **increased** union membership.

a) Formulate the hypothesis that can be used to determine whether union membership

increased in 2018. Clearly write down the Hypothesis test including H0 and Ha.

Because the question indicates the 11.3% of U.S. workers belonged to unions in 2013. In other words, the probability is 0.113. The question assumes workers belonged to unions is greater than 0.113, so we set workers > 0.113 as our alternative hypothesis and workers = 0.113 as our null hypothesis.

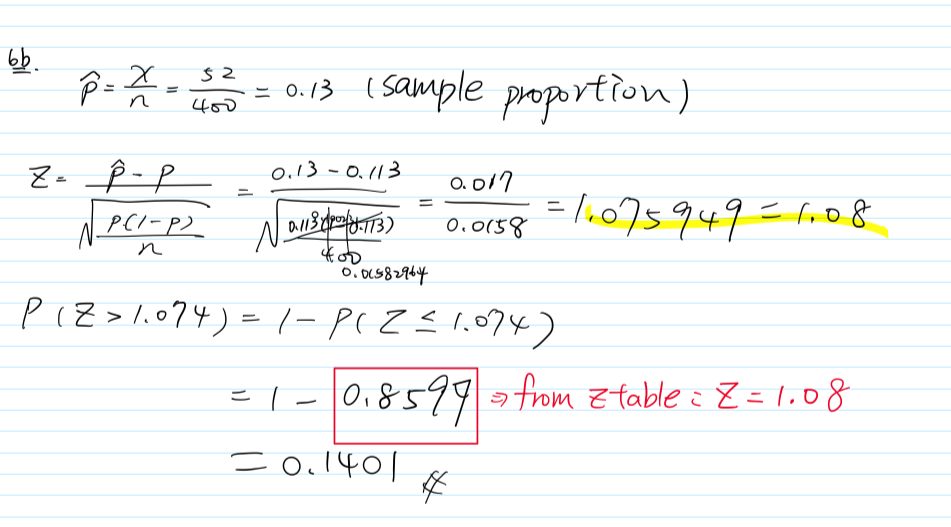
Hypothesis testing: Ho: workers = 0.113 vs Ha: workers > 0.113.

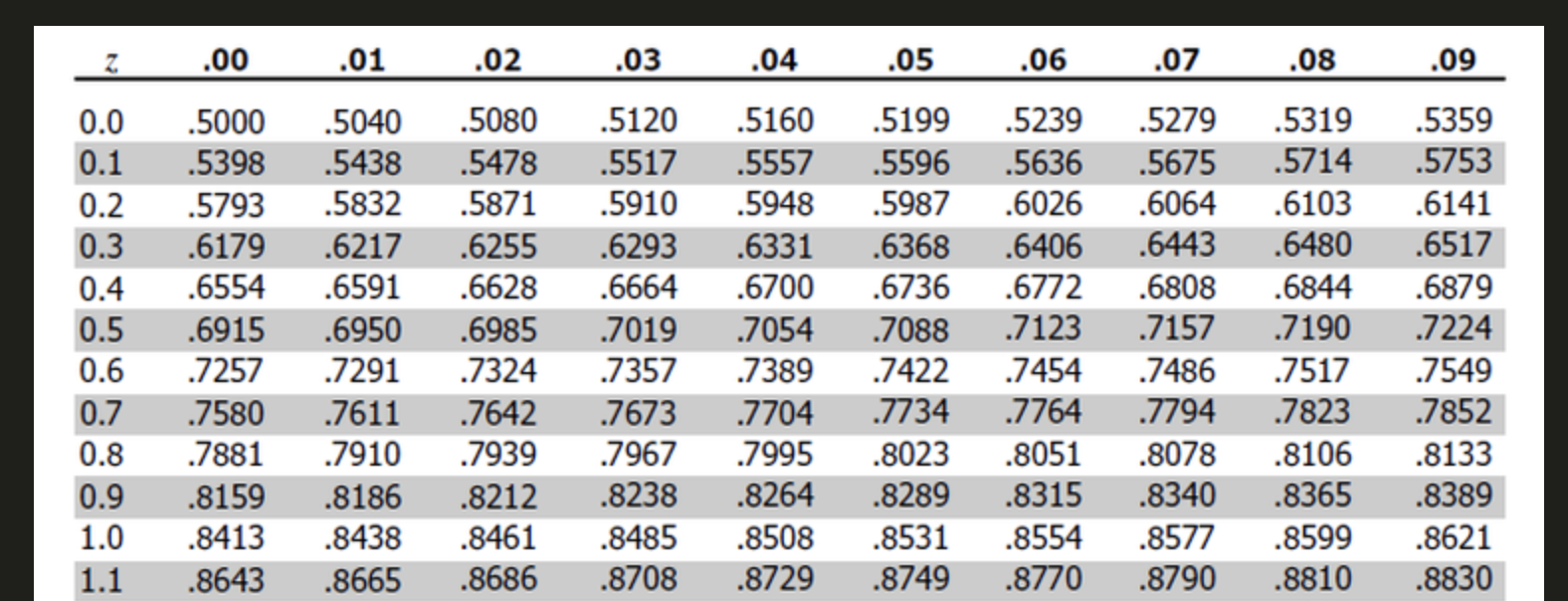
b) If the sample results show that 52 of the workers belonged to unions, what is the p-value

for your hypothesis test?

n = 400, sample results = x = 52, workers = p = 0.113, significant level = alpha = 0.05

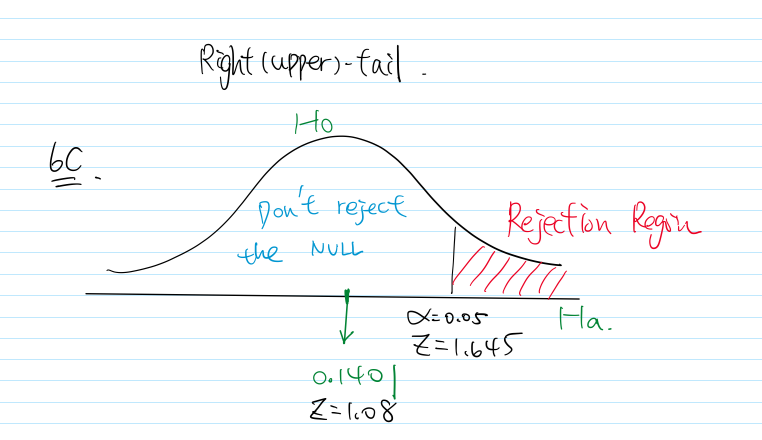
The right-tail p-value is 0.1401.





c) At alpha = .05, what is your conclusion?

As previously stated, the p-value is 0.1401, which is greater than alpha = 0.05. Thus, we fail to reject the null hypothesis, meaning that there is no evidence to conclude that the union efforts to organize have increased union membership.



d) Write a Python code to determine and address the hypothesis tests that you developed in

Part a) and manually answered in Part b).

