**Spring 2016**

**BA 240**

**Statistical Analysis Winnie Li**

**Exam 2**

**May 2016**

**Name:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**I acknowledge and accept the Honor Code**

# (Signed):\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**Score: \_\_\_\_\_\_\_\_\_\_/100**

**Instructions:**

* **Do *NOT* Open the exam until instructed to do so.**
* You will have the whole class period (100 minutes) for this exam.
* This exam is Closed Book, Closed Notes. ***Only Two sheets (Four pages) of cheat sheets a clean z-table & t-table, and a calculator are allowed.***
* The exam has 6 pages, including this one (Title).
* **Round to the THIRD decimal place unless otherwise noted in questions.**
* **Read each question/part carefully, follow the instructions**, and write solutions in the space provided. If you need more space, use the back.
* **PLEASE SHOW ALL YOUR WORK COMPLETELY AND CLEARLY!!!**

 **Good Luck** 

**Problem 1:** The research department in Wall Street Journal states that 64% of U.S. workers take their lunch to work based on its reliable resource and extensive research.

1. Tupperware Corporation is considering serving free lunch as an added employee benefit. Before the final decision is made, the corporation surveyed 500 workers, where 340 indicated that they took their lunch to work with them. Find out 98% confidence interval for the true proportion of workers who take their lunch to work with them, and interpret the result. (8 points) **Note: use the z-score with the ACCURATE THIRD DECIMAL PLACE.**

n = 500 p (hat) = 340 out of 500 took their lunch to work with them

α = 0.02 α/2 = 0.01 z α/2 = 2.326

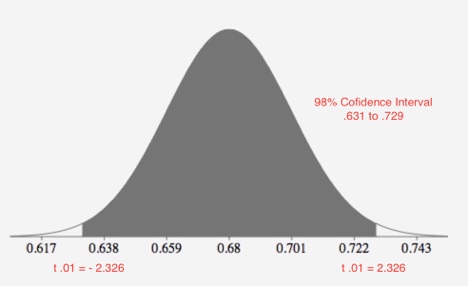
p(hat) ± z α/2 TBD

p(hat) = 340/500 = 0.68

.68 ± 2.326 √.68 x .32 / 500

.68 ± 2.326 x 0.0208 = [0.631, 0.729]

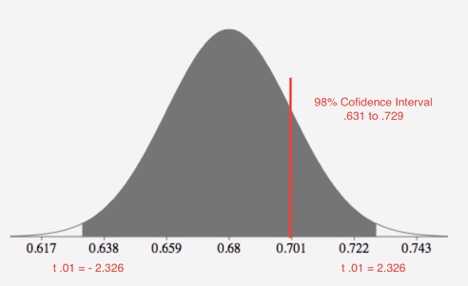




1. Based on the interval in part a), can Tupperware Corporation be 98 percent confident that more than 70 percent of workers take their lunch to work with them? Circle the correct answer and explain. (2 points)

**Yes X No Explain: 70% (.70) is in the 98% confidence interval**

We are 98% confident that the true proportion of workers who take their lunch to work with them is between 63.1% and 72.9%.



1. Determine the sample size needed in order to be 95 percent confident that the sample proportion of workers who take their lunch to work, is within .01 of p, the true proportion of workers who take their lunch to work. Interpret the result. **Note: State which value of *p* you should use and why.** (6 points)

The p value is .68 because we are asked to determined the sample size needed in order to be 95% confident that the sample proportion of workers who take their lunch to work is within .01 of p.

n = (**1.96)2** .68 x .32 / (.01)2 = 8359.32 = 8360

E2 = (.01)2 = 1

***z.025* = 1.96**



1. **Without any calculation**, how would the sample size in part c) change if we change from 95 percent confidence to 99 percent confidence while all other conditions keep the same? How about change to 90 percent? Circle the correct answer. (4 points)

**For 99%: Less Samples Needed More Samples Needed**

**For 90%: Less Samples Needed More Samples Needed**

**Problem 2:** The metropolitan airport commission is considering establishing limits on noise pollution around a local airport. At the present time the noise level per jet takeoff in one neighborhood near the airport is approximately normally distributed with mean of 100 decibels and a standard deviation of 6 decibels. **Note: For EACH of the following part, draw a normal curve, mark the x-axis accordingly and highlight the corresponding area(s). Round z-scores to the SECOND decimal place, and keep the ORIGINAL FOUR decimal places for the probability.**

Mean x (bar) = 100

Std dev = 6

1. What is the probability of a randomly chosen score to be no more than 85? (4 points)

P(x < 85) = P(z < -2.5) = 0.5 - .4938 = 0.0062

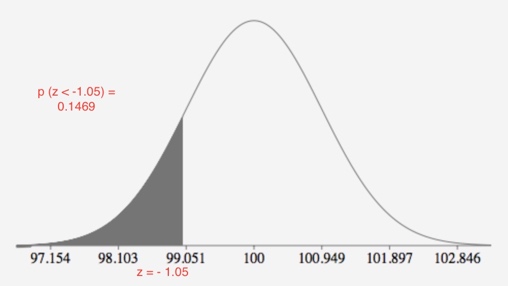
Z = 85 - 100 / 6 = -2.5

1. Suppose a randomly sample of 40 scores is selected, what is the probability that the mean score

is at most 99? (4 Points)

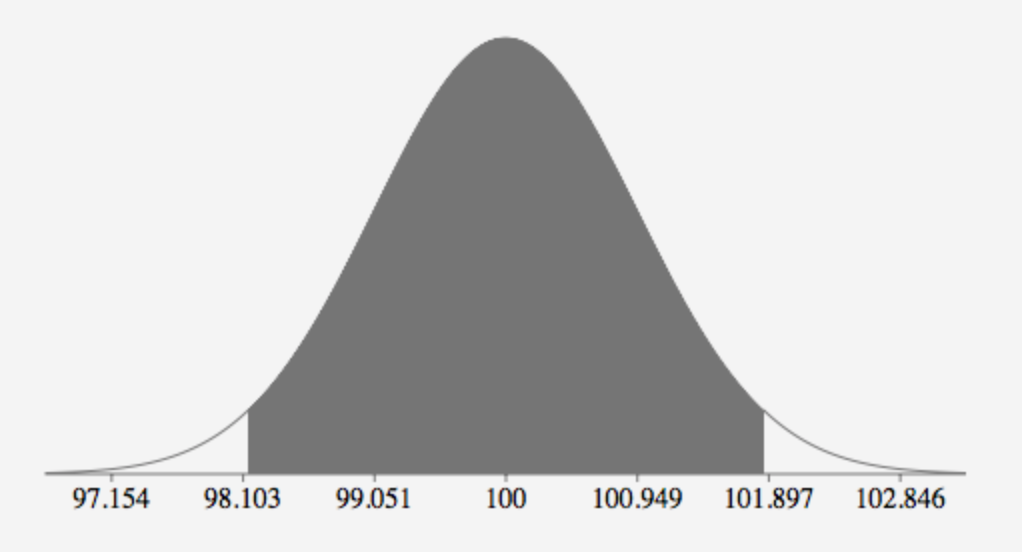
n = 40 std dev = 6/ √40

**z = 99 – 100 /** 6/ √40 = -1.05 p(x < 99) = p( z < - 1.05) = 0.5 - .3531 = 0.1469



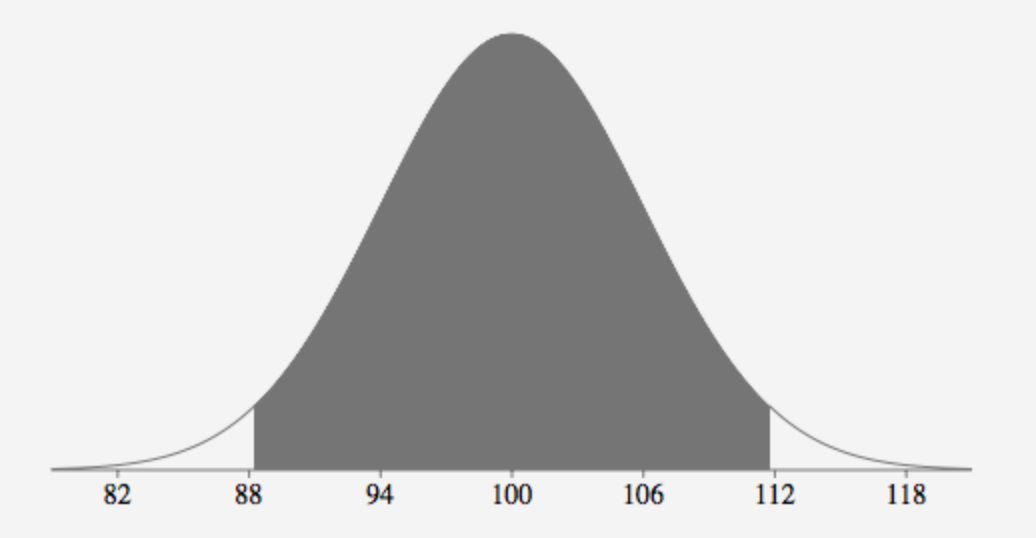
1. Suppose 95% of the scores are contained in an ***interval symmetrically around the mean***. Find this interval. (4 points) Assuming sample n = 40 from previous problem

100 ± 1.96 x 6/ √40 = [98.1406, 101.8594]



Population

100 ± 1.96 x 6 = [88.24, 111.76]

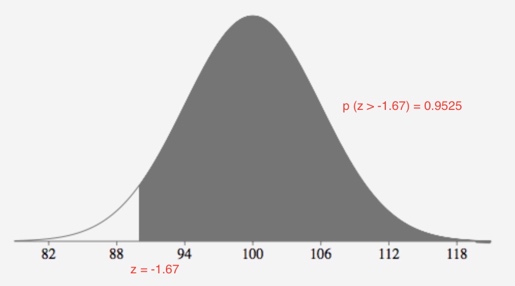


1. What is the probability of a randomly chosen score to be at least 90? (4 points)

**Population:**

**Z = 90 – 100 / 6 = -1.667 = -1.67**

**P(x > 90) = p(z > -1.67) = 0.5 + 0.4525 = 0.9525**

****

**Sample:**

**Z = 90 – 100 / 6**/ √40  **= -0.9486 = -0.95**

**P(x > 90) = p(z > -0.95) = 0.5 + 0.3829 = 0.8289**

**Problem 3**: The high U.S. unemployment rate in 2010 has been a major blow to the economic recovery. In April 2010, 52 percent of the unemployed had been out of work longer than six months. Policy makers felt that this rate declined during the past years as the job market improved. A random sample of 400 unemployed people was selected, and it was found that 190 had been out of work for more than six months.

1. State null hypothesis and alternative hypothesis. (4 points)

**Ho:** p0 = .52 The rate of unemployment remains the same than 52%

**Ha:** p0 < .52 The percent of unemployment decline to less than 52%

1. Calculate the value of the Test Statistic. Specify the Test Statistic “z” or “t” clearly. (5 points)

N = 400

p(hat) = 190/400 = 0.475

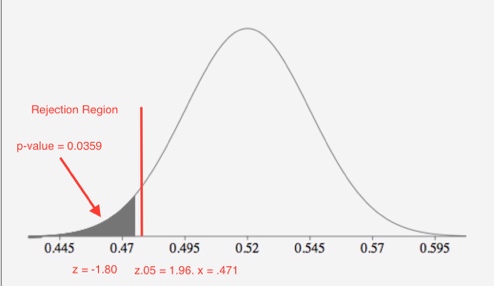
z = 0.475 – 0.52 / √.52 x 0.48 / 400 = - 1.8014

1. Find the p-value using ***appropriate table***. Draw the normal curve and highlight the corresponding area. (4 points)

z = -1.80 p( z < -1.80) = 0.5 - .4641 = 0.0359

1. At .05 significance level, do you reject the null hypothesis? How about at .01 significance level? Circle the correct answer and explain ***using the p-value found in part c).*** (4 points)

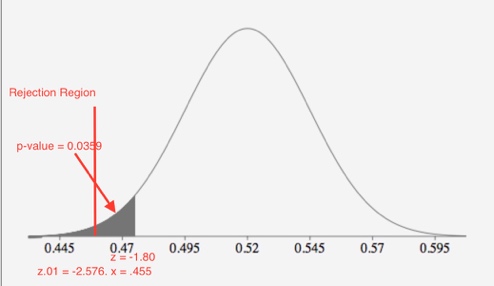
**At .05 significance level : Reject Ho / DO NOT Reject Ho Explain:**

****

At theα = 0.05 we reject the null hypothesis, **Ho**, because p-value of 0.0359 is less than 0.05.

**At .01 significance level : Reject Ho / DO NOT Reject Ho Explain:**

At theα = 0.01 we don not reject the null hypothesis, **Ho**, because p-value of 0.0359 is greater than 0.01.



1. Interpret both of the results from part d). (3 points)

**We have strong evidence to suggest that null hypothesis is false and we can reject the null hypothesis in favor of the alternative hypothesis at** α = 0.05**.**

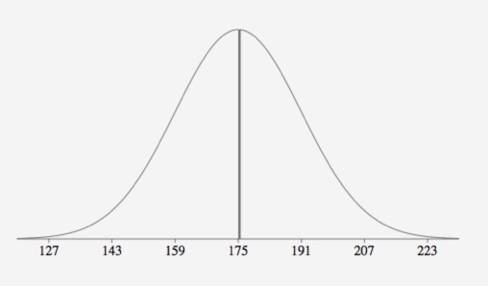
**We do not have sufficient evidence to suggest that null hypothesis is false and we can not reject the null hypothesis in favor of the alternative hypothesis at** α = 0.01**.**

**Problem 4:** Assumes the room rate for a three-star hotel in Paris follows a normal distribution with a mean of 175 Euros and a standard deviation of 16 Euros. **Note: For EACH of the following part, draw a normal curve, mark the x-axis accordingly and highlight the corresponding area(s).**

1. What is the median room rate? (3 points).

Median and mean = 175

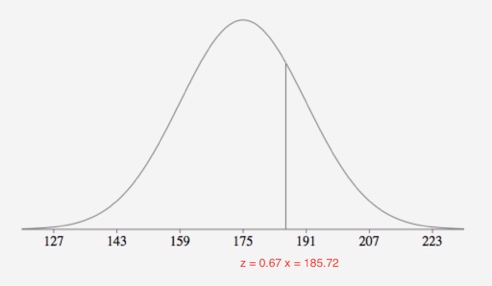
The mean and median is the same if normally distributed.



1. What is the room rate in the third quartile (Q3)? (3 points). **Note: Round z-score to the SECOND decimal place.**

p(z > X) = 0.75 = 0.5 + 0.25 = z = 0.67

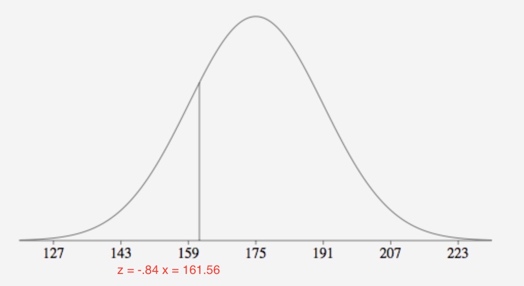
175 + (16 x 0.67) = 175 + 10.72 = 185.72



1. What is the room rate in the 20th percentile? (3 points). **Note: Use the z-score with the ACCURATE SECOND DECIMAL PLACE.**

P (z > X ) = 0.20 = 0.5 - 0.3 z = - .84

175 + (16 x -.84) = 161.56



**Problem 5**: The Glen Valley Steel Company manufactures steel bars. If the production process is working properly, it turns out steel bars that are normally distributed with mean length of

2.6 feet. Either longer steel bars or shorter bars must be scrapped. The research department is wondering if the company should adjust the production equipment. A random sample of 24 bars was selected, and the mean length is 2.37 feed with the variance of 0.09 foot.

1. State null hypothesis and alternative hypothesis. (4 points)

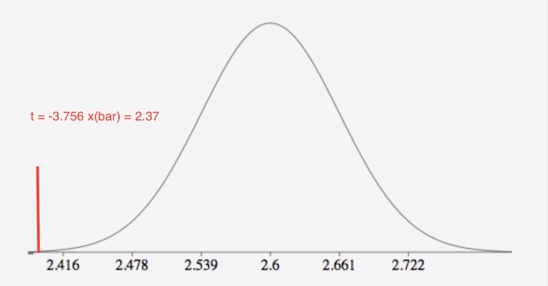
Two tail test mean = 2.37 variance = 0.09 n = 24 std dev = 0.3

**Ho:**  μ =2.6 the production process is working properly and production equipment does not need to be adjusted

**Ha:** μ ≠ 2.6 the production process is not working properly and production equipment does need to be adjusted

1. Calculate the value of the Test Statistic. Specify the Test Statistic “z” or “t” clearly. (5 points)

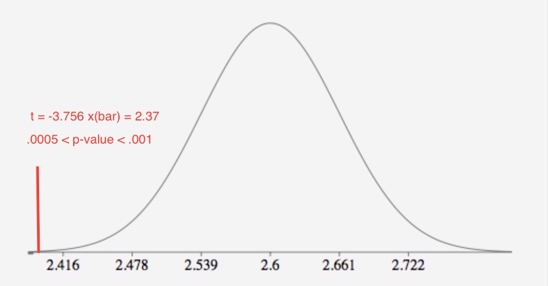
t = x(bar) - μ / s/√n = 2.37 – 2.6 / 0.3/√24 = -0.23 / 0.06123 = - 3.7558 = -3.756



1. Find the p-value using appropriate table. Draw the normal curve and highlight the corresponding area. (3 points)

df = 24 – 1 = 23

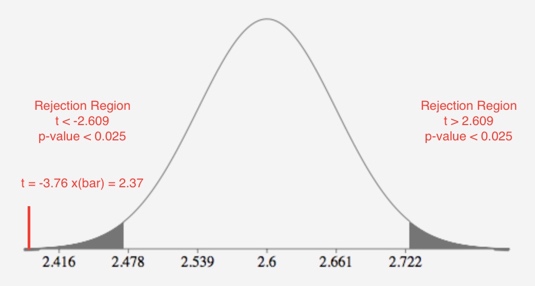
.0005 < p-value < .001



1. At .05 significance level, do you reject or retain the null hypothesis? How about at .01 significance level? Circle the correct answer and explain ***using the p-value found in part c).*** (4 points)

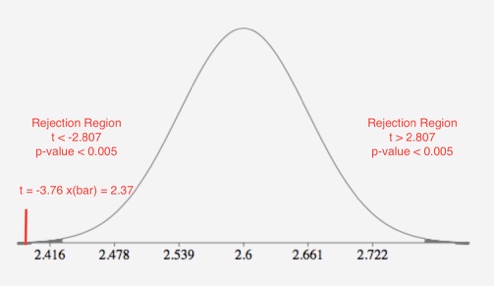
**At .05 significance level: Reject Ho / DO NOT Reject Ho Explain:**

At theα = 0.05 we reject the null hypothesis, **Ho**, because .0005 < p-value < .001. The p-value is less than 0.025



**At .01 significance level : Reject Ho / DO NOT Reject Ho Explain:**

At theα = 0.01 we reject the null hypothesis, **Ho**, because .0005 < p-value < .001. The p-value is less than 0.005



1. Interpret both of the results from part d). (4 points)

**We have very strong evidence to suggest that null hypothesis is false and we can reject the null hypothesis in favor of the alternative hypothesis at** α = 0.05**.**

**We have very strong evidence to suggest that null hypothesis is false and we can reject the null hypothesis in favor of the alternative hypothesis at** α = 0.01**.**

**Problem 5:** A study of the pay of corporate chief executive officers (CEOs) examined the increase in

cash compensation of the CEOs of 24 companies in year 2012. The mean increase was 6.9% and the standard deviation of the increases was 5.5%. Find a 99% confidence interval for the true increase in cash compensation of the CEOs of all companies in year 2012. Interpret the result. (5 points)

n = 24

mean x(bar) = 0.069

st dev = 0.055

α = .01 α/2 = .005

tα/2 = 2.807

x(bar) ± tα/2 x s/ √n = .069 ± 2.807 x 0.055 x√ 24 = 0.069 ± 0.031 =

[0.03748, 0.1005]

We are 99% confident that the true mean increase is between 3.75% and 10.05%



**Problem 7:** Circle the correct answer. (1 point each)

**1:** Suppose a 95% confidence interval for the population mean turns out to be (1000, 2100). If this interval was based on a sample of size of n = 22, explain what assumption is necessary for this interval to be valid.

1. The sampling distribution must be biased with 21 degrees of freedom.
2. The sampling distribution of the sample mean must have a normal distribution.
3. The population must have an approximate t distribution.
4. The population must have an approximately normal distribution.

**2:** Suppose you are studying different attributes of several breeds of dogs. You are interested in estimating the mean shoulder height of Labrador retrievers. What is the target parameter?

A) *~~x~~* B) µ C)  D) 𝑝 E) 𝑝̂

**3:** Forty-five CEOs from the electronics industry were randomly sampled and a 99% confidence interval for the average salary of all electronics CEOs was constructed. The interval was ($101,866, $115,016). To make more useful inferences from the data, it is desired to increase the width of the confidence interval. What two ways will result in an increased interval width? \_\_\_\_\_decrease Sample Size and \_\_\_\_\_ Confidence Level

A) Increase, Increase B) Increase, Decrease C) Decrease, Decrease D) Decrease, Increase

**4:** A hypothesis test is used to prevent a machine from underfilling or overfilling quart bottles of beer. On the basis of sample, the null hypothesis is rejected and the machine is shut down for inspection. A thorough examination reveals there is nothing wrong with the filling machine. From a statistics point of view:

A) Both Type I and Type II errors were made. B) A correct decision was made.

C) A Type I error was made. D) A Type II error was made. E) None of above

**5:** The Central Limit Theorem says the sampling distribution of the sample mean is approximately normal under certain conditions. What is a necessary condition for the C.L T. to be used?

1. The population from which we are sampling must be normally distributed.
2. The population from which we are sampling must not be normally distributed.
3. The population size must be large (e.g. at least 30).
4. The sample size must be large (e.g. at least 30).

**Problem 8:** Attach your 2 sheets of cheat sheet at the BACK of the exam, and Staple together.

***DO NOT attach the z table or t table!*** (5 Points)