Hypothesis testing and Statistical analysis:

1. Z Test (Data that are dealing with average data) => Z Table => Z Score and P Value

2. T Test (Data that are dealing with average data) => T Table => T Score and P Value

3. Chi Square Test (Categorical Data)

4. Annova Test (Variance)

Z Test:

i) Population standard deviation should be known

ii) n >= 30, i.e, Sample size should be greater than or equal to thirty

Q. The average heights of all residents in a city is 168 cm with a . A doctor believes the mean to be different. He measured the heights of 36 individuals and found the average height to be 169.5 cm.

a) State Null and Alternate Hypothesis

b) At a 95% confidence level, is there enough evidence to reject the null hypothesis.

Given,

n = 36, i.e, n >= 30

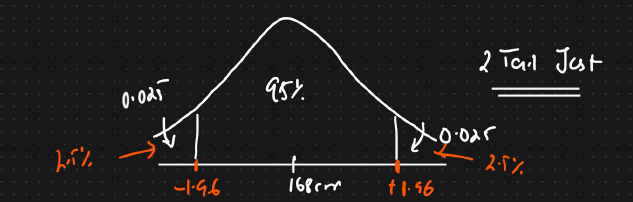
CI = 0.95

Significance value = = 1 – CI = 1 – 0.95 = 0.05

1. Null Hypothesis () =

2. Alternate Hypothesis () =

3. Based on confidence interval, we will draw Decision Boundary



We know how much area of the distribution is covered by each standard deviation in a normal distribution.

We have to cover 95% of the distribution. So on both ends, the probability would be 2.5% each.

If we get value within the confidence interval, we will state that the null hypothesis is accepted or that we fail to reject the null hypothesis.

If we fall in either of the extremes, we will reject the null hypothesis.

Calculating Z Score:

1 – 0.025 = 0.9750 = Area

Now we will refer to the Z table

We get the Z score = 1.96 after cross referencing with the table

If any value from the statistical analysis comes within this range of -1.96 to +1.96, then the null hypothesis is considered as true.

If it comes in the extreme ends, we will reject the null hypothesis.

If Z is less than -1.96 or greater than +1.96, then we have to reject the null hypothesis. Otherwise we accept it.

Z Test:

= = = 2.3

According to the Central Limit theorem, the formula for Z score is . But this is for population data.

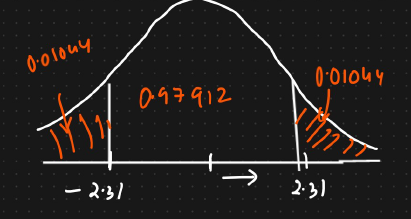
If we are using the formula for sample data, then it gets modified to

Since 2.3 is greater than +1.96, we reject the null hypothesis.

Hence we can confirm P < 0.05 (Significance Value)

Now constructing a new graph based on our calculations

Cross referencing the Z Table for area under the curve



In order to calculate P Value, there are two ways, either from Z Score as we have done before

Or P value can be calculated by adding the two regions

P value = 0.01044 + 0.01044 = 0.02088 < 5

Hence we reject the null hypothesis

Final conclusion: The average height is not equal to 168 cm. The average height seems to be increasing based on sample data.

Q. A factory manufactures bulbs with an average warranty of 5 years with standard deviation of 0.50. A worker believes that the bulb will malfunction in less than 5 years. He tests a sample of 40 bulbs and find the average time to be 4.8 years

a) State the null and alternate hypothesis.

b) At a 2% significance level, is there enough evidence to support the idea that the warranty should be revised.

Solution:

Given,

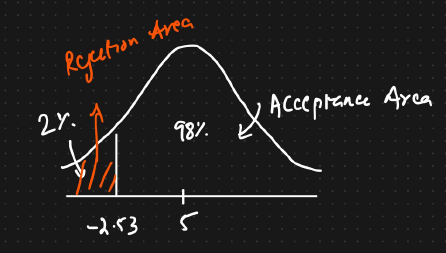
n = 40, i.e, n >= 30

Significance value = = 0.25

1. Null Hypothesis () =

2. Alternate Hypothesis () =

3. Based on confidence interval, we will draw Decision Boundary



Z Test:

= = = -2.53

Area under the curve with Z score -2.53 = 0.00570 = P Value

Comparing P value with significance value = 0.2,

Since P value is greater than significance value, we cannot reject the null hypothesis as P value falls under acceptance value.

So it will be a one tailed test only.

My Answer: Incorrect

1) Incorrect significance value taken

2) Incorrect calculation by chatgpt for Z test

Given,

n = 40, i.e, n >= 30

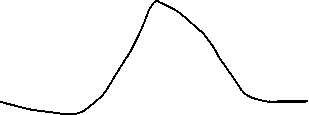
Significance value = = 0.25

CI = 1 – Significance Value = 0.75

1. Null Hypothesis () =

2. Alternate Hypothesis () =

3. Based on confidence interval, we will draw Decision Boundary



We know how much area of the distribution is covered by each standard deviation in a normal distribution.

We have to cover 75% of the distribution. So on both ends, the probability would be 12.5% each.

If we get value within the confidence interval, we will state that the null hypothesis is accepted or that we fail to reject the null hypothesis.

If we fall in either of the extremes, we will reject the null hypothesis.

Calculating Z Score:

1 – 0.125 = 0.875 = Area

Now we will refer to the Z table

We get the Z score = 1.15 after cross referencing with the table

If any value from the statistical analysis comes within this range of -1.15 to +1.15, then the null hypothesis is considered as true.

If it comes in the extreme ends, we will reject the null hypothesis.

If Z is less than -1.15 or greater than +1.15, then we have to reject the null hypothesis. Otherwise we accept it.

Z Test:

= = = -0.0632

Since -0.0632 is greater than -1.15, we fail to reject the null hypothesis.

Hence we can confirm P is not less than 0.25 (Significance Value)

