DAY - 5

Inferential Statistics

Hypothesis Testing: We make assumptions (or) conclusions regarding the population data Based on the Sample data. which we will be done by using "Hypothesis Testing".

steps of Hypothesis testing:

1) **Null Hypothesis**: By Default it will be true.

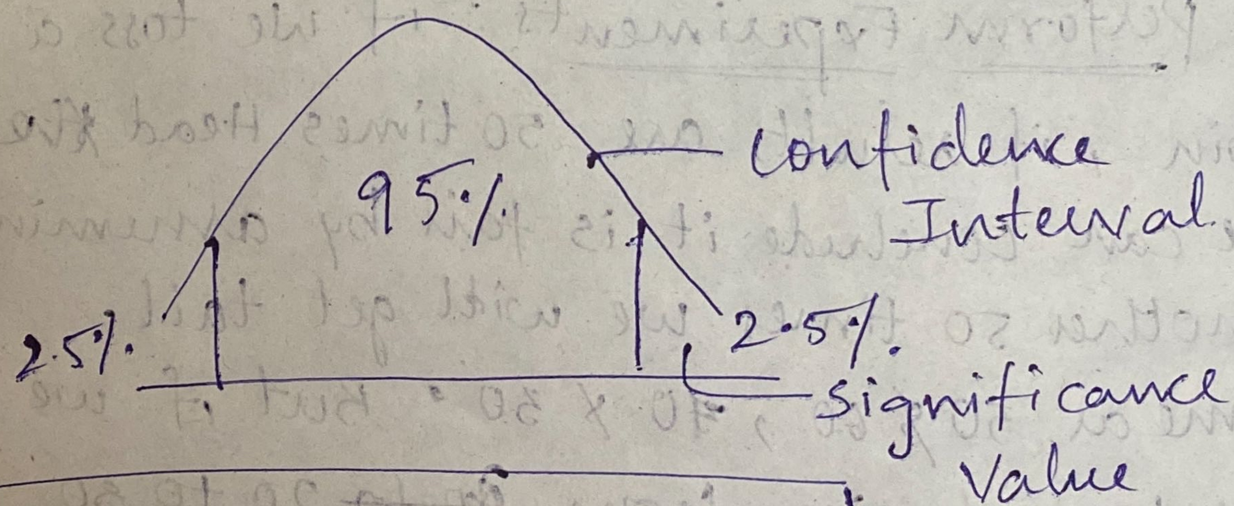
2) **Alternate Hypothesis**: It will be opposite to Null Hypothesis

3) **Perform Experiments**: if we toss a fair coin, if results are 50 times Head then we can conclude it is fair by assuming another 50 times we will get tail. Same as, 50 & 60, 70 & 30. But if we set our boundaries from ~~20 to~~ 20 to 80 then it is called as "Confidence Interval" which will be done by "Domain Expert".

If we get my coin values in defined range i.e, from 20-80. we can conclude coin is fair and we can also say Null Hypothesis is Accepted.

If we get my coin values in out of range i.e, 10 times Head we can conclude coin is not fair and we can also say Null Hypothesis is Rejected and Alternate Hypothesis is accepted.

Confidence Interval : If we take C.I as 95%. Significance value will be
$1-0.95 = \underline{0.05}$



$\boxed{\text{Significance value} = 1 - C.I}$

Point Estimate: The value of any statistics that Estimate the value of a parameter is called point Estimate

Ex: $\bar{x} \longrightarrow \mu$

$\bar{x}$ ↓ Sample mean μ ↓ population mean

with the help of sample mean we can assume the population.

Here, we will take Sample man as statistic and population mean as parameter.

Point Estimate ± Margin of Errom = parameter

Lower fence = point Estimate − Margin of error.

Higher fence = point Estimat + Margin of error

$$\text{Margin of error} = \boxed{Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}}$$

α = Significance value

$\sigma/\sqrt{n}$ = standard error

Problem: On the quant test of CAT Exam a sample of 25 test takers has a mean of 520 with a ~~sample~~ population standard deviation of 100. Construct a 95% C.I about the mean ?

Solution: $n = 25$ (Sample data)

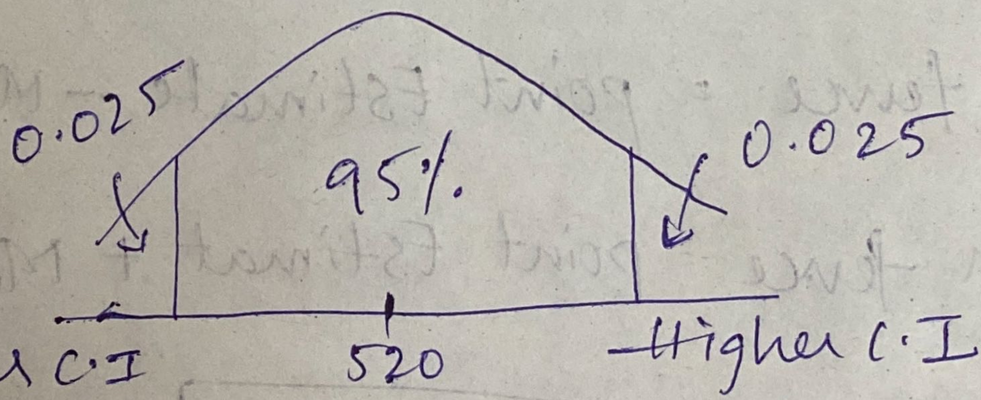
$\bar{x} = 520$ (Sample mean)

$\sigma = 100$ (Standard deviation)

C.I = 95% (Confidence Interval)

S.V = 1 − C.I (Significance value)

$= \underline{\underline{0.05}}$



Lower C.I = point Estimate − Margin of error

$$= 520 - Z_{0.05/2}\left[\frac{\sigma}{\sqrt{n}}\right]$$

$$= 520 - Z_{0.025}\left[\frac{\cancel{100}\ 20}{\cancel{5}\ 1}\right]$$

$= 520 - 1.96 \times 20 = \underline{\underline{480.8}}$

Higher C.I = Point Estimate + Margin of error

$= 520 + 1.96 \times 20 = \underline{\underline{559.2}}$

Now, the confidence Interval range is

$\boxed{480.8 - 559.2}$

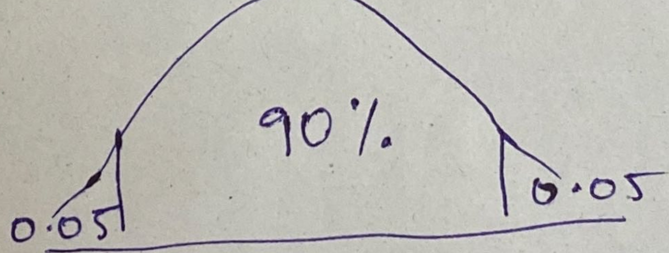
This will be only defined by Domain Expert

Problem: $\bar{x} = 480, \sigma = 85, n = 25$

$C.I = 90\% = 0.90, S.V = 1 - C.I$

$= 1 - 0.90$

$= \underline{\underline{0.10}}$



Lower C.I $= 480 - Z_{0.10/2}\left[\frac{\sigma}{\sqrt{n}}\right]$

$= 480 - Z_{0.05}\left[\frac{85}{25}\right]$

$= 480 - Z_{0.05}\left[\frac{85}{5}\right]$ $\sqrt{n}$ is nothing but $\sqrt{25} = \underline{\underline{5}}$

$= 480 - 1.64[17]$

$= \underline{452.12}$

Higher C.I = 480 + 1.64 [17]

= 507.88

Now, the confidence Interval range

is $\boxed{452.12 - 507.88}$

Domine Expert will give the C.I Range.

Problem: on the quant test of CAT Exam sample of 25 test takes has a mean of 520 with a sample standard deviation of 80 construct 95%. C.I ?

Solution: $\bar{x} = 520$, $s = 80$, C.I = 95%.

S.V = 1 - 0.95
= 0.05

Here we dont have population data s.d so we have to use t - test
if population s.d is give we have to use Z - table.
Sample s.d is given now, we have to change the value (or) formula

$$\bar{x} + t_{\alpha/2}\left[\frac{s}{\sqrt{n}}\right]$$

Whenever we want to test T-test we need to find degree of freedom

degree of freedom is given as $\boxed{n-1}$

$= 25 - 1$

$\boxed{n = 24}$

Lower C.I: $\bar{x} \mp t_{\alpha/2}\left[\frac{s}{\sqrt{n}}\right]$

$= 520 \mp t_{0.05/2}\left[\frac{80}{\sqrt{24}}\right]$

$= 520 \mp 0.025\left[\frac{80}{5}\right]$

$= 520 - 2.064 \times 16$

$= 486.976$

Higher C.I : $520 + 2.064 \times 16$

$= 553.024$

Now the confidence Interval range is $486.976 - 553.024$

Hypothesis Testing problems

A factory has a machine that fills 80ml of baby medicines in a bottle. An employee believes the average amount of medicine is not 80 ml using 40 samples he measures the average amount dispensed by the machine to be 78ml with a standard deviation of 2.5

a) State null & Alternate Hypothesis

b) At 95% C.I, is there enough evidence to support machine is working properly or not.

$\boxed{\text{Step - 1}}$

Solution: Null Hypothesis:

$\mu = 80$, {machine is working properly}

Alternate Hypothesis

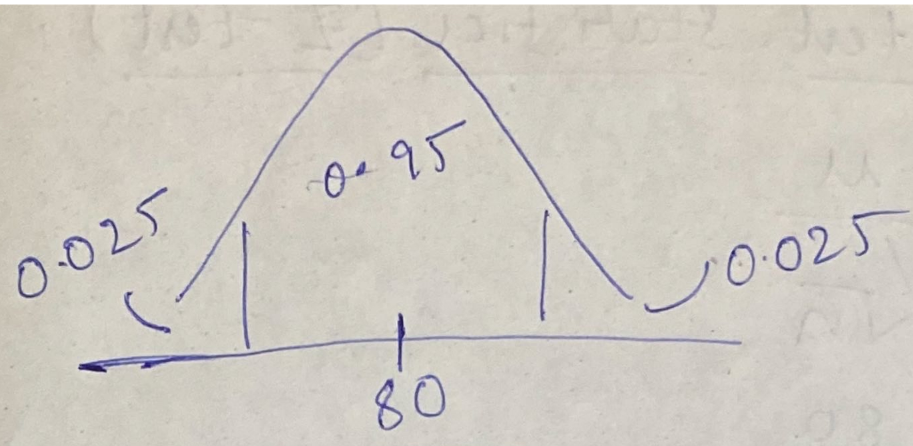
$\mu \neq 80$ {machine is not working properly}

$\boxed{\text{Step - 2}}$

$\mu = 80$, $n = 40$, $\bar{x} = 78$, $S = 2.5$, $C.I = .95$

S.L - 0.05

It is Two tail test



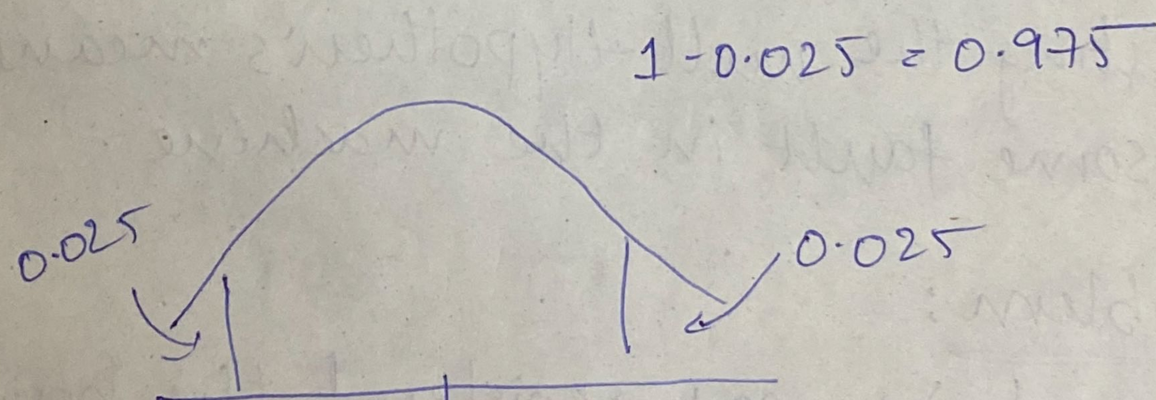
$\boxed{\text{Step - 3}}$

If $n \geq 30$ (or) population standard Deviation we have to use Z - test.

If $n < 30$ and sample standard deviation we have to us T- test.

$\boxed{\text{Step 4}}$

Experiment;



"Decision Boundary"

* Calculate test Statistics (Z-test):

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$= \frac{78 - 80}{\frac{2.5}{\sqrt{40}}}$$

$$= -5.05$$

* Conclusion:

Decision rule: If $Z = -5.05$ is less than -1.96 or greater that 1.96, reject the null Hypothesis with 95%. C.I

Rejecting the null-Hypothesis means there is some fault in the machine.

Problem:

A complain was registered, the boys in a Govt School are underfed. Average weight of the boys of age 10 is 32 kgs with S.D 9 kgs. a sample of 25 boys were selected from the govt school and the avg weight

is found to be 29.5 kgs, with CI 95%.
check it is True or False ?

solution: Conditions for Z-test:

1) we know the population s.d (σ)
2) we do not know the population s.d but our sample is larger than 30 i.e, $n \geq 30$

Conditions for t-test:

1) we don't know the population s.d
2) our sample size is smaller than 30 ie, $n < 30$
3) sample s.d is given.

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}} \longrightarrow Z \text{ test}.$$

$\boxed{\text{step - 1}}$

Null Hypothesis (H_0):

$\mu = 32$ {fed well}

Alternative Hypothesis (H_1):

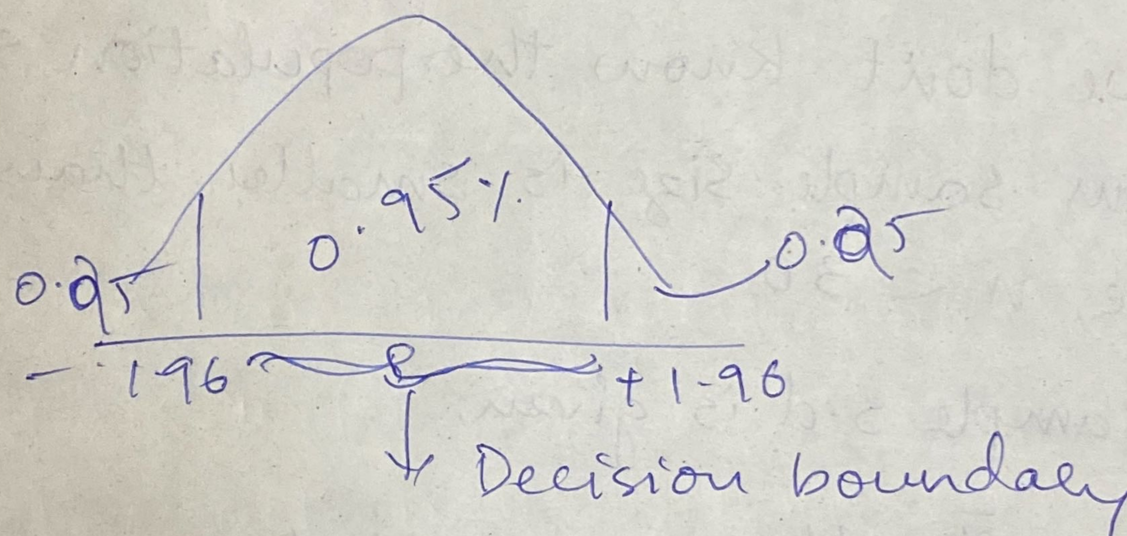
$\mu \neq 32$ {not fed well}

Step 2

C.I = .95 , S.V - 1 - C.I
(α) = 1 - .95
= 0.05

Step 3

Z test (this is 1 tail test)



Z-score = $\frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$ = $\frac{29.5 - 32}{9 / \sqrt{25}}$

= -1.39

Conclusion: -1.39 > -1.96

Accept the null Hypothesis 9.5% C.I
we fail to reject the null Hypothesis
The boys are fedwell.

Problem: A factory manufactures the cars with a warranty of 5 or more years on the engine & transmission. An engineer believes that the engine or transmission will malfunction in less than 5 years. He tests a sample of 40 cars & finds the average time to be 4.8 years with standard Deviation of 0.50.

a) State the null & alternate hypothesis

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

Solution: $\mu = 5$, $n = 40$, $\bar{x} = 4.8$

$\boxed{\text{Step 1}}$

$H_0 - \mu \geq 5$

$H_1 - \mu < 5$

$\boxed{\text{Step-2}}$

$C.I = 0.98$, $\alpha = 0.02$

$\boxed{\text{Step 3}}$

Z-test (this is 1 tail test)

$$Z\text{-score} = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{4.8 - 5}{0.50/\sqrt{40}}$$

$$= \frac{-0.2}{0.0790}$$

$$= \underline{\underline{-2.53}}$$

$$Z_\alpha = -2.33$$

$$\therefore -2.5 < -2.33$$

<u>Conclusion</u>: we accept the null Hypothesis with 98% of C.I and reject the alternate Hypothesis i.e, yes the warning is True.