Today Topics

1. Type 1 and Type 2 Error ✓
2. One Tailed and 2 Tailed Test ✓
3. Confidence Interval ✓
4. Z-test, t-test, Chi-Square Test

1. Type 1 and Type 2 Error

Null Hypothesis (H_0) = Coin is fair

Alternate Hypothesis (H_1) = Coin is not fair

Reality check

Null Hypothesis is True or Null Hypothesis is False

Decision

Null Hypothesis is True or Null Hypothesis is False

Outcome 1 :
We reject the Null Hypothesis, when in reality it is false → Yes

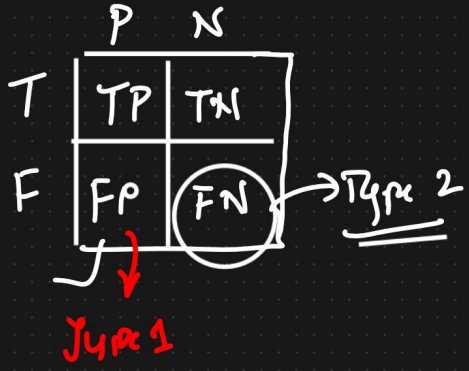
Movie

Person - Death Sentence

Outcome 2 : We reject the Null Hypothesis, When in reality it is true → Type 1 Error

Outcome 3 : We Accept/retain the Null Hypothesis when in reality it is false → Type 2 Error

Outcome 4 : We Accept the Null Hypothesis when in reality it is true → Good



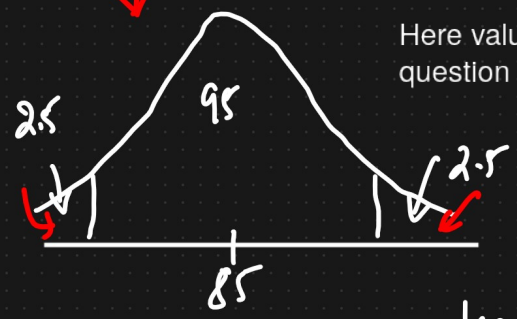
② 1 Tail and 2 tail Test

Eg: Colleges in Karnataka have an 85% placement rate. A new college was recently opened and it was found that a sample of 150 students had a placement rate of 88%. With a standard deviation 4%. Does this college has a different placement rate ? $\alpha = 0.05$

↳ 85%

(Significance value given is 0.05)

2 tailed Test



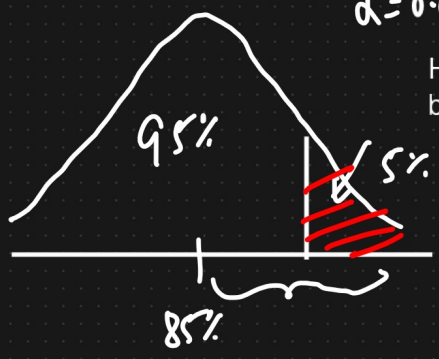
Here value can be greater than 85 or less than 85 so question will use 2 tailed test

If we modify the question:

Does this college have a placement rate greater than 85% ?

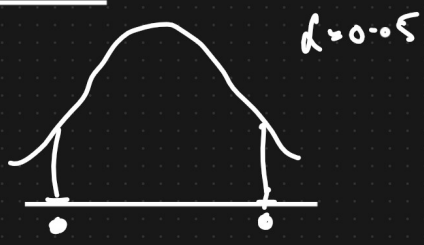
lenes

$\alpha = 0.05$



Here we are just checking greater(or either just lower) then it will be 1 tailed test(as only one region will be checked)

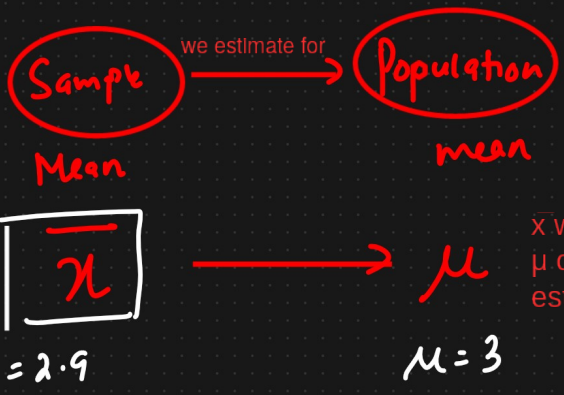
③ Confidence Intervals



Point Estimate

The value of any statistic that estimates the value of a parameter

In Inferential Stats



$\bar{x}$ will estimate μ(population mean), $\bar{x}$ may be approx equal to μ or it may be less or greater too. So we can say $\bar{x}$ is point estimate that will be estimating μ value.

Confidence Intervals

As we know $\bar{x}$ can be equal or not equal to population mean, so we use confidence interval so we come towards the value of population mean

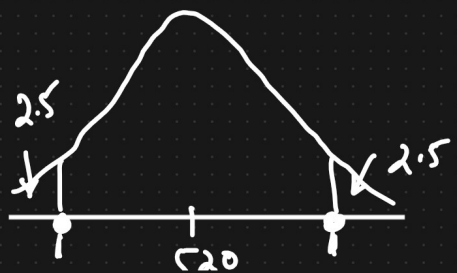
Formula for CI: Point Estimate ± Margin of Error

Test problem:

Q) On the quant test of CAT Exam, the standard deviation is known to be 100. A sample of 25 test takers has a mean of 520 score. Construct a 95% CI about the mean?

Ans) $\sigma = 100$ $n = 25$ $\alpha = 0.05$ $\bar{x} = 520$ (as given in question)

$\alpha = 1 - 0.95 = 0.05$



Our graph looks like

whenever:

{ ① Population std is given } → Z test

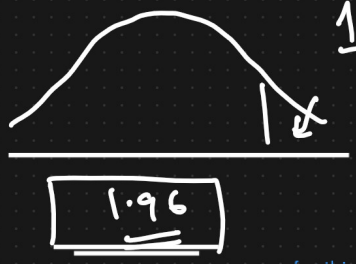
② $n \geq 30$ =

Point Estimate ± Margin of Error

$$\bar{x} \pm Z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$ → Standard Error

$$Z_{\frac{0.05}{2}} = Z_{0.025}$$

$$\frac{1 - 0.025}{2} = 0.975$$



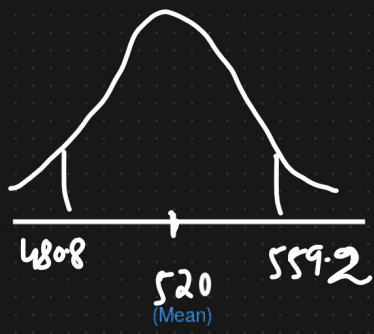
for this from z table we get 1.96

Upper bound = $\bar{x} + Z_{\frac{0.05}{2}} \frac{100}{\sqrt{25}}$ (we have taken 25 to make calculation simple)

Lower bound = $\bar{x} - Z_{\frac{0.05}{2}} \frac{100}{\sqrt{25}}$

Upper = $520 + 1.96\,(20) = 559.2$

Lower = $520 - 1.96\,(20) = 480.8$



Stats

One stats interview question:

Find the average size of the sharks throughout the world?

σ, $\bar{x}$, n, $\alpha = 0.05$

:We can assume these values with or without looking on internet and solve it just like above

Q) On the quant test of CAT exam, a sample of 25 test takers has a mean of 520 with a standard deviation of 80. Construct 95% confidence interval about the mean?

Ans) Condition $n = 25$ $\bar{x} = 520$ $S = 80$ $\alpha = 0.05$

Here population sd is not given → we use t-test

Point Estimate ± Margin of Error

$$\bar{x} \pm t_{\alpha/2}\left(\frac{s}{\sqrt{n}}\right) \rightarrow \text{Standard Error}$$

$$t_{\frac{0.05}{2}} = 2.064$$

$$\text{Upper bound} = \bar{x} + t_{0.05/2}\left(\frac{s}{\sqrt{n}}\right)$$

To calculate value of t, we first calculate: $\text{Degree of freedom} = n - 1 = 25 - 1 = 24$

$$= 520 + 2.064\left(\frac{80}{5}\right)$$

Now look in t table wrt 24 degree of freedom and 0.05 value, we get 2.064

$$= 553.024$$

$$\text{Lower bound} = \bar{x} - t_{0.05/2}\left(\frac{s}{\sqrt{n}}\right)$$

$$= 520 - 2.064\left(\frac{80}{5}\right)$$

$$= 486.97$$

$$[486.97 \longleftrightarrow 553.024]$$

① One Sample Z-Test

① Population sd is given
② Sample size $n \geq 30$

: we should use Z-test when

example:

④ In the population, the average IQ is 100 with a sd of 15. Researchers wants to test a new medication to see if there is positive or negative effect on intelligence, or no effect at all. A sample of 30 participants who have taken the medication has a mean of 140. Did the medication affect the intelligence? $\alpha = 0.05$ (CI = 95%)

→ 110 ✓

How to perform hypothesis test:

Ans) 1) Define Null Hypothesis

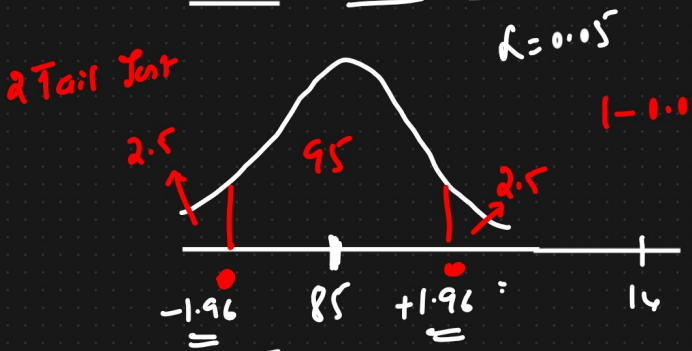
$H_0 = \mu = 100$ Mean is 100

② Alternate Hypothesis H_1 $\mu \neq 100$ Mean is not 100

③ State Alpha value

$\alpha = 0.05$

④ State Decision Rule



Z table

$1 - 0.025 = 0.975$

As question is asking whether the medication affect the intellogence which means it can either increase or decrease, so two tailed test will be used

$P \leq 0.05$

⑤ Calculate Test Statistics

Sample data

$$Z = \frac{\bar{x} - \mu}{\left\{\frac{\sigma}{\sqrt{n}}\right\}} \rightarrow \text{Standard Error}$$

$$= \frac{140 - 100}{\frac{15}{\sqrt{30}}} = \frac{40}{15} \times \sqrt{30} = 14.60$$

Next step 6)

State our Decision

$\{Z = -0.2\}$

$Z = 14.60$

$14.60 > 1.96$

If Z is less than -1.96 or greater than 1.96, reject the null hypothesis

Medication improve the intelligence or decrease ?!

Improve the intelligence

As 14 is greater than than 1.96 and hence it is increasing mean

② One Sample t-test

Z-test ⇒ population std

t-test ⇒ unknown population std

EQ) Population the average IQ = 100

$n = 30 \quad \bar{x} = 140 \quad s = 20$

Did the medication affect intelligence?

$\alpha = 0.05$

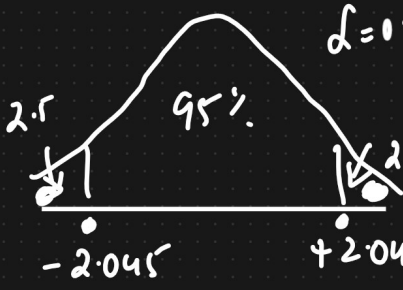
Ans) ① $H_0 = \mu = 100$

② $H_1 \quad \mu \neq 100$

③ Calculate the degree of freedom

$n - 1 = 30 - 1 = 29$

④ State Decision Rule



$\alpha = 0.05 \quad t = 10.96 > 2.045$

Reject Null Hypothesis

$P \leq$ significance value

⇓

{ Increased the intelligence. }

⑤ T Test

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

$\bar{x} = 140$

$\mu = 100$

$s = 20$

$n = 30$

$= 10.96$

Reject the Null Hypothesis

Real World Problem