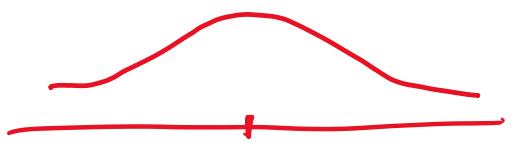


## Hypothesis Testing

Hypothesis = Belief.

e.g.: Hypothesis: (Avg. Sal. of Data Scientists have changed around the world Post COVID-19) ✓ earlier of \$1,50,000



$$\mu = \$1,50,000$$

How to test the belief/hypothesis  $\Rightarrow$  H.T.

Case I

$$\bar{x} \text{ (Avg. Sal. of a sample)} = \$90,000$$

Can we say that the avg. Sal. has changed?  $\rightarrow$   
(this deviation might be due to random variability).

$$90000 = 150000$$

~~✓~~

diff is not significant.

Case II

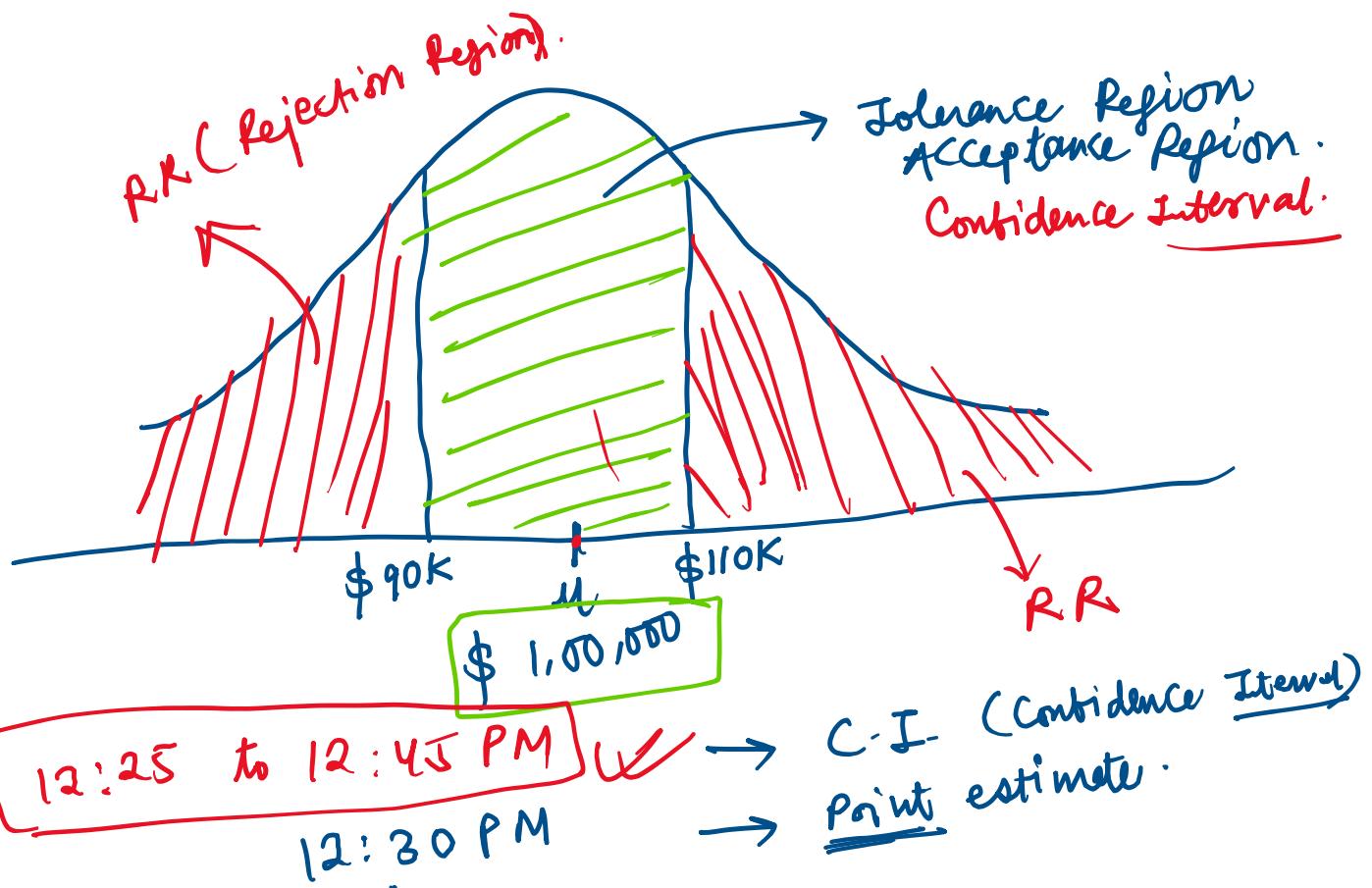
$$\bar{x} \text{ (Avg. Sal. of a sample)} = \$1,10,000$$

diff is not significant.

Case III

$$\bar{x} \text{ (Avg. Sal. of a sample)} = \$1,25,000$$

diff is significant.



### Technical Discussion -

Hypothesis

- ↳ Null Hypothesis
- ↳ Alternate hypothesis

eg: Null Hypothesis ( $H_0$ )  $\Rightarrow$   
Alt. " " ( $H_A$ )  $\Rightarrow$

Earth is flat  
Earth is not flat

$\mu = \$1,00,000$  ✓  
 $\mu \neq \$1,00,000$

Case I  
Reject null hypo

Case II  
Fail to Reject  
 $H_0$

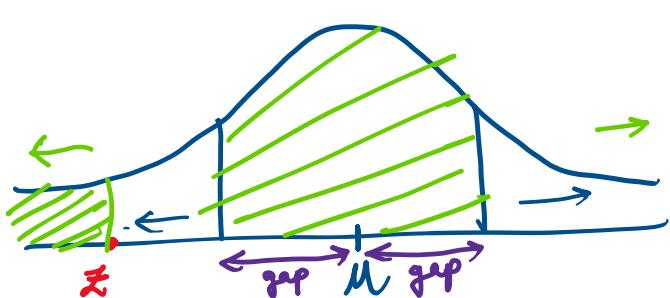
## Acceptance Region / Tolerance Region (AR/TR)

Q. How to Construct AR/TR?

$$M \pm \text{gap}$$

$$\mu \pm Z * \text{std. error}$$

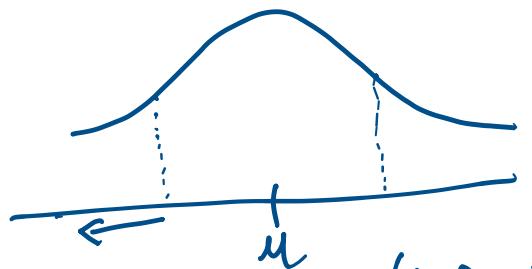
$$\frac{6}{\sqrt{n}}$$



At what Prob.  
I want to  
construct my  
significance.

whether this  
is a 1-tailed test  
or a two tailed  
test.

**Z-score** :- How far  
the  
↓  
**Prob. / Percentile** :- Z-table



(AR) Area ↑      (RR) Area ↓  
Z-Score ↑  
(Prob.) ART ↑      (Prob.) RR ↓

Q. How to decide Z-value in H.T.?

- ① significance level / level of significance.
- ② tail of the test : 1 or 2 tailed test.

$$\text{CI} + \text{SL} = 100\% \text{ OR } 1$$

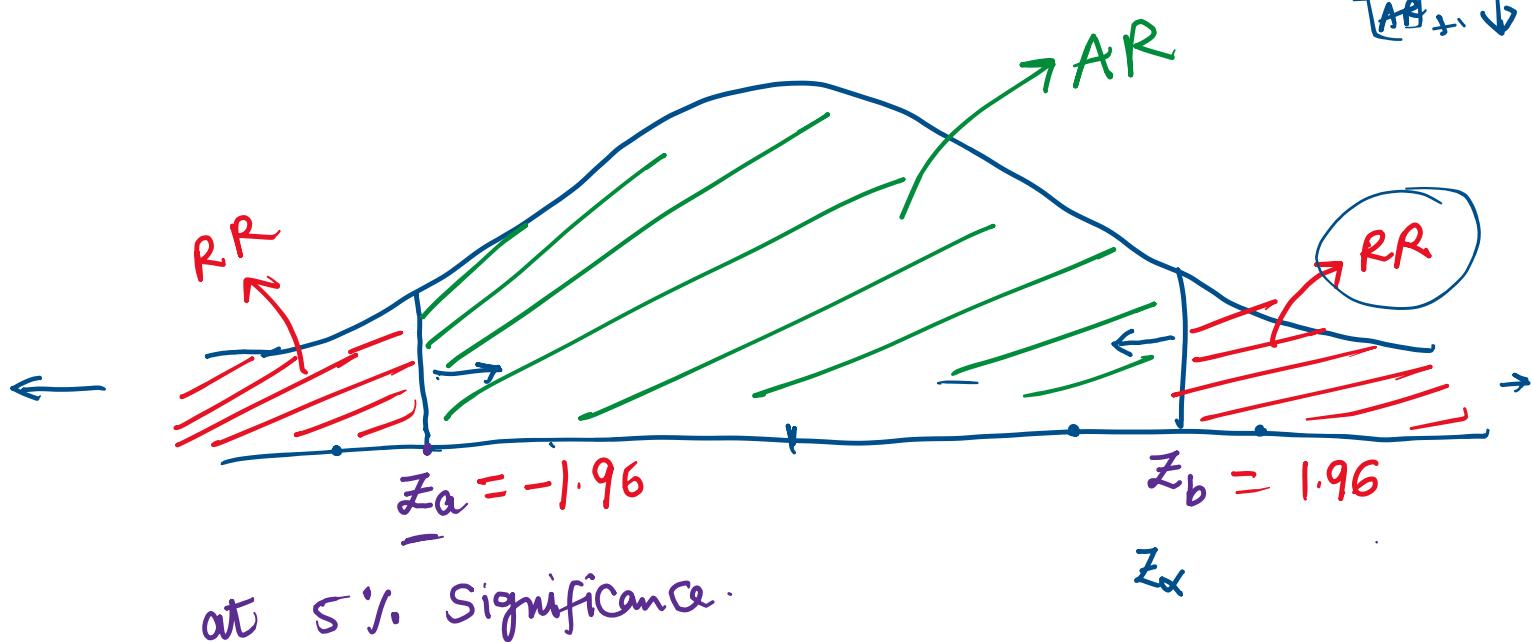
$$95\% + 5\% = 100\%$$

*RR*

*CI* + *SL* = 100% OR 1  
 95% + 5% = 100%

*CDF of body RR*

*CDF of body AR*

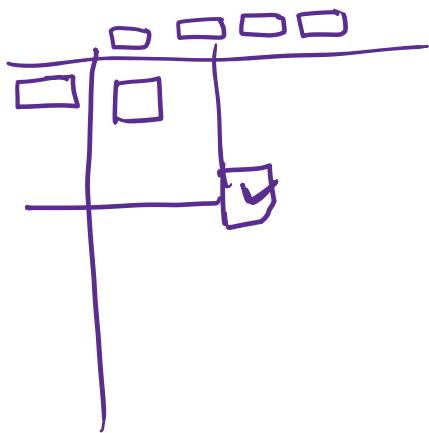


Soln:

$$\text{CI} + \underline{\text{SL}} = 100\% \text{ OR } 1$$

$$5\%$$

$$2.5\% \quad 2.5\%$$



$$2.5\% = \underline{0.025}$$

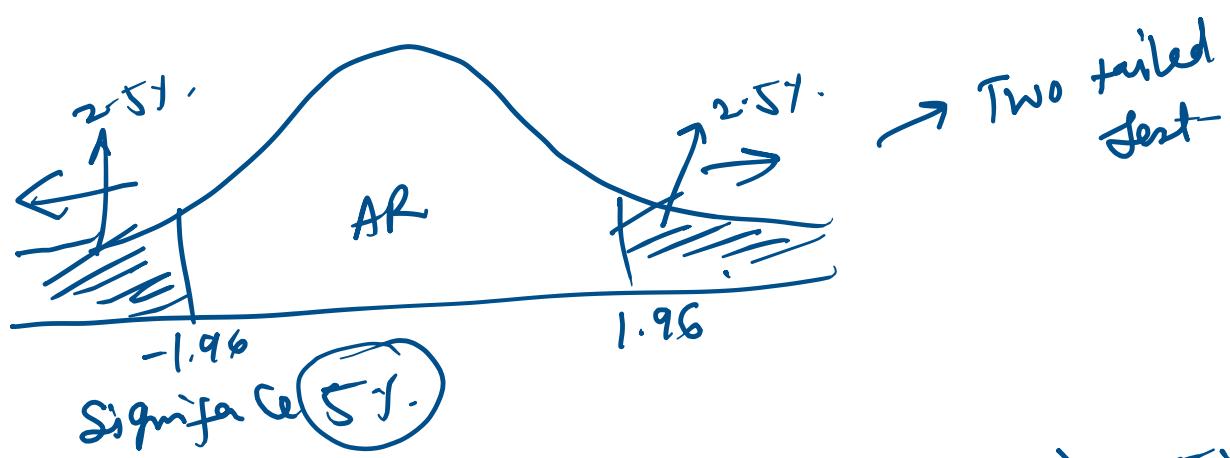
one tailed or two tailed test :

Case I

$$\begin{aligned} H_0 &\Rightarrow \mu = \$100000 \\ * H_A &\Rightarrow \mu \neq \$100000 \end{aligned}$$

$$\mu < 100000 \checkmark$$

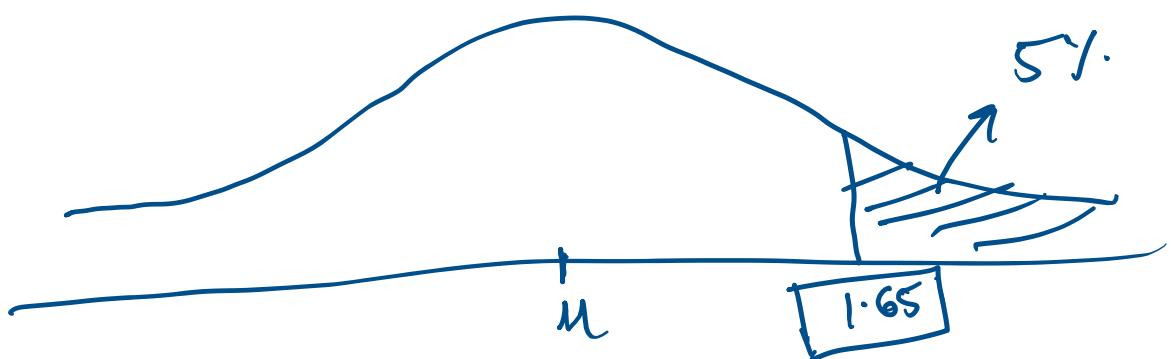
$$\mu > 100000$$



Case II

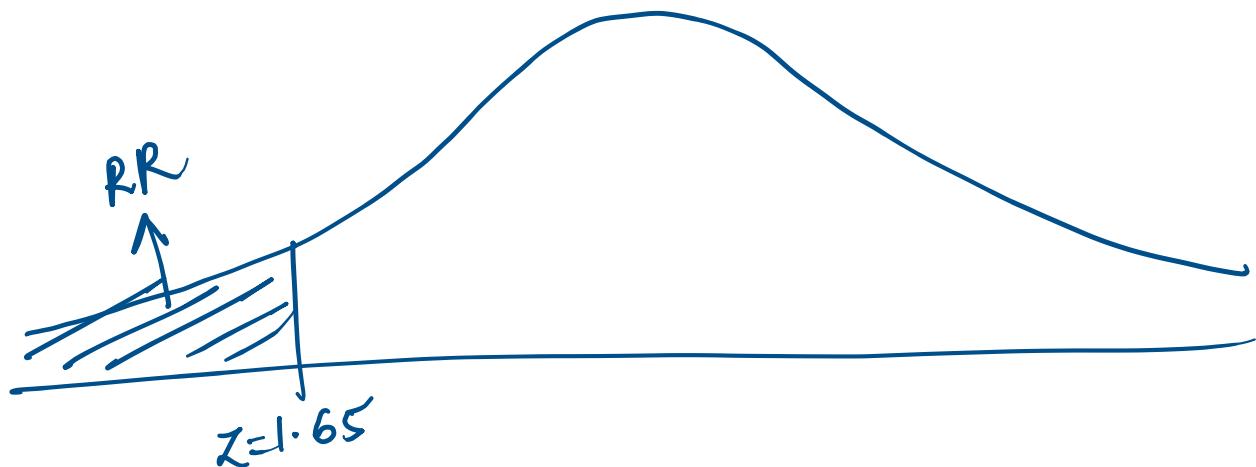
$$\begin{aligned} H_0 &\Rightarrow \mu \leq \$150K \\ * H_A &\Rightarrow \mu > \$100K \end{aligned}$$

$$\text{Sign}(\alpha) = 5\%$$



Case III

$H_0 \Rightarrow \mu \geq \$100K$   
 $H_A \Rightarrow \mu < \$100K$



$\alpha$ $= 5\%$	1T	1.65	5% -
$= 10\%$	2T	$= 1.96$	2.5% 2.5%
	1T	1.28	10%
	2T	1.65	5% & 5%



Tests:  $Z$  & T-test

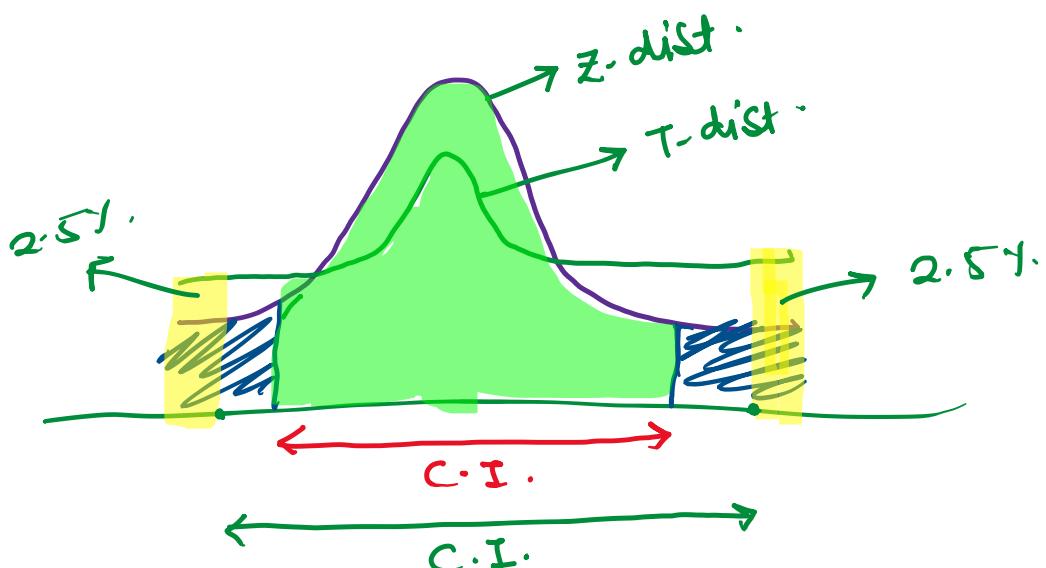
$\sigma \rightarrow S$

$$\frac{\bar{X}}{\sqrt{n}} \sim \frac{S}{\sqrt{n}}$$

When to use T-test

- ① when  $\text{popn } \sigma$  (STD) is unknown.
- ② when  $n < 30$ .

$n > 30 \Rightarrow Z = T$



Ques: A researcher has agreed upon a data on a daily return of portfolio of call option over a recent 250 days period. The mean daily return has been 0.1% and the standard deviation of daily return portfolio is 0.25%.

The researchers believe that mean daily portfolio return is not 0. Construct a hypothesis test of researchers belief at 95% Confidence Interval.

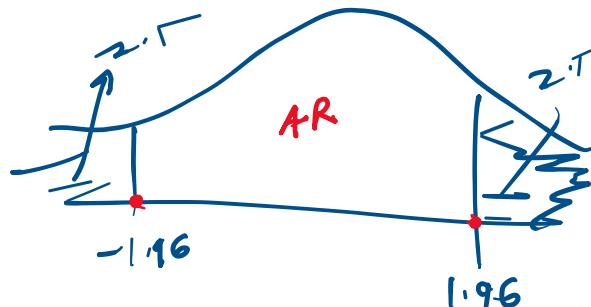
....

- Step 1: Define Hypotheses  $\rightarrow H_0 \& H_A$
- Step 2: Choosing type of test  $\rightarrow 1 \text{ or } 2 \text{ tailed}$
- Step 3: Construct AR/RR
- Step 4: decision.

Step 1:

$$H_0 \Rightarrow \mu = 0$$

$$H_A \Rightarrow \mu \neq 0$$



Step 2: Two Tailed.

Step 3: (a)  $\alpha = 1 - C.I. = 1 - 95\% = 5\%$ .

(b) Choice of Test  $\Rightarrow \bar{Z}$

Constructing AR

$$AR = \mu \pm (Z \times S.E.)$$

$$= \mu \pm Z \times \frac{S}{\sqrt{n}}$$

$$= 0 \pm 1.96 \times \frac{0.0025}{\sqrt{250}}$$

LL  
•

UL  
•

Step 4: will check if my  $\bar{x}$  lies bet<sup>n</sup> this Rng<sup>n</sup>.

If yes  $H_0 \Rightarrow$  Accept (Fail to Reject)  
If No  $H_0 \Rightarrow$  Reject

$\bar{x} = 112.5$ , since it falls outside of AR  
Hence  $H_0$  is Rejected.

Question: A principal at a school claims that the students in his school are above average in terms of intelligence. A random sample of 30 students' IQ scores have a mean of 112.5. The mean population IQ is 100 with a STD of 15. Test the hypothesis of Principal's claim.

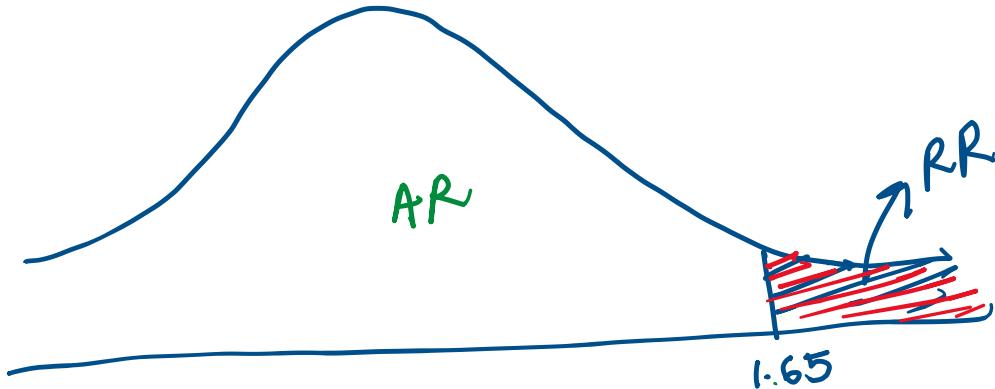
$$① H_0 = \text{Mean} \leq 100$$

$$H_A = \text{Mean} > 100 \rightarrow \text{one tailed} \rightarrow \text{Right tailed.}$$

$$② \text{Choice of test} \rightarrow Z$$

$$③ \underline{\alpha = 5\% \text{ or } 0.05} \quad | \quad Z_{\alpha} = \underline{1.65}$$

Method-I (critical value method)



$$Z_{\text{cal}} = \frac{\bar{x} - \mu}{(6/\sqrt{n})} = \frac{112.5 - 100}{15/\sqrt{30}} = \underline{4.56}$$

$$Z_{\text{cal}} > Z_{\alpha}$$

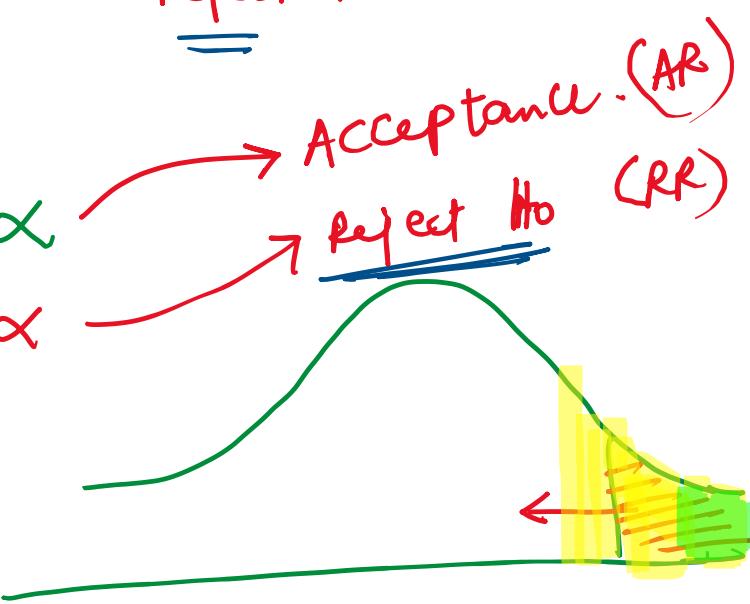
hence Reject the  $H_0$

Method 2

P-value method

$$(P\text{-value})_{\text{cal}} > \alpha$$

$$(P\text{-value})_{\text{cal}} < \alpha$$



~~99.7%~~

$$P\text{-value for } Z = 1 - A_L(4.56)$$

$$= 1 - 99.9\%$$

$$= 0.000001$$

$$\alpha = 0.05$$

Method 3

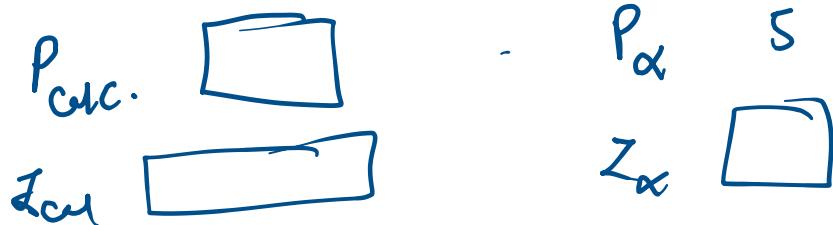
AR - Method.

(Already discussed in previous Q.)

Q. P-value

If the null hypothesis is true, what is the Prob. that we'd observe a more extreme value test statistics in the dirn of Alt. Hypothesis

Same as  $\Rightarrow$  If the defendant is innocent what is the chance that we'd observe such extreme criminal behaviour.



### Points to Ponder -

- ① Hypothesis & H.T.
- ②  $H_0$  &  $H_A \Rightarrow$  How to frame?
- ③ TR/AR & RR
- ④ Methods
  - ↳ critical value method / Z-score method
  - ↳ p-value method
  - ↳ TR/AR - RR Method
- ⑤ Reverse lookup of Z-table
- ⑥ Tails  $\rightarrow$  one tailed / two tailed
- ⑦ C.I. & Point estimate.

## ⑧ Level of Significance ( $\alpha$ )