Welcome Back everyone

1. One tail, 2 tail test
2. p-value
3. Chi-square test
4. Annova

Q ⇒ The avg. wt of residents in a society in lbs is 168 lbs. A nutritionist believes that this is not right - she measured wt of 36 ind. and found avg to be 169.5 lbs with a σ of 3.9.

a) State null & Alternate

b) Using a 95% C.I. is there enough evidence to discard null hypothesis?

Step 0 : $n=36$ $\bar{x} = 169.5$ $s = 3.9$ $\alpha = 0.05$

Step 1 : $H_0 \Rightarrow \mu = 168$

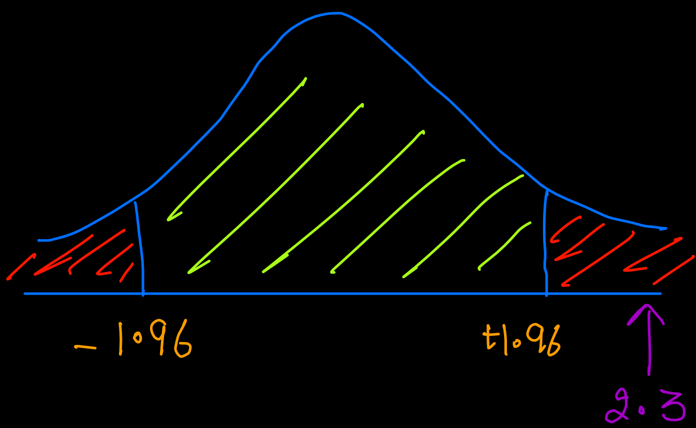
$H_1 \Rightarrow \mu \neq 168$

Step 3 : Decision Boundary

step 4 :- Calculate test statistics

$$Z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{169.5 - 168}{3.9/\sqrt{36}}$$

$$= 2.3$$



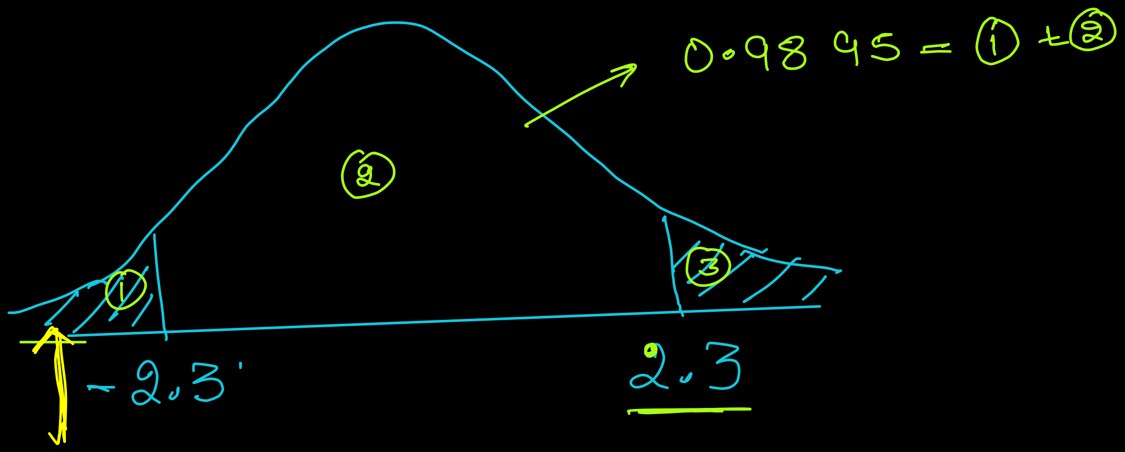
$2.3 > 1.96$

$\Rightarrow$ reject H_0

Step 5: Conclusion

With 95% CI, we can say that we reject null hypothesis

$\Rightarrow$ nutritionist is right



$$3 = 1 - \{1 + 2\} = 0.0105$$

$$1 + 3 = 0.02 = 0.0105 * 2$$

$$\boxed{0.02 < 0.05}$$

$\Downarrow$

reject H_0

Q :- A factory manufactures a car with warranty of 5 years on the engine. An engineer believes that engine will malfunction in less than 5 years. < 5

He tests sample of 40 cars and find avg to be 4.8 with a of 0.5

a) H_0 , H_1

b) Assume $\boxed{98\%\ CI}$, is there enough level to support the idea that warranty should be revised.

Step 0: $n = 40$ $\mu = 5$ $\bar{x} = 4.8$ $s = 0.5$

$\alpha = 1 - C.I = 1 - 98\% = 0.02$

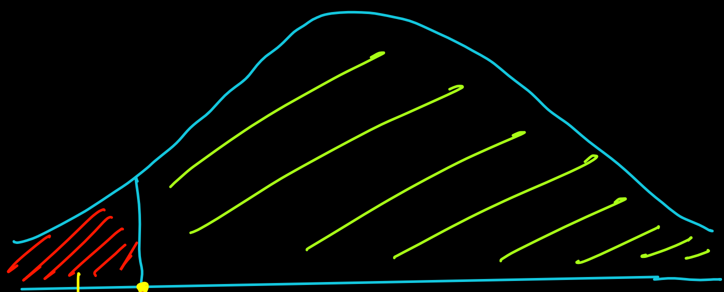
Step 1: $H_0 \Rightarrow \mu \geq 5$

$H_1 \Rightarrow \mu < 5$

Step 2: Calculate decision boundary

$\Rightarrow$ z-test as $n \geq 30$

$\Rightarrow$ one tail



Step 3 : Calculate test statistics

$$z = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{4.8 - 5}{0.5/\sqrt{40}}$$

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

$$= -2.53$$

$\Rightarrow$ Reject null hypotheses

p-value = 0.0057

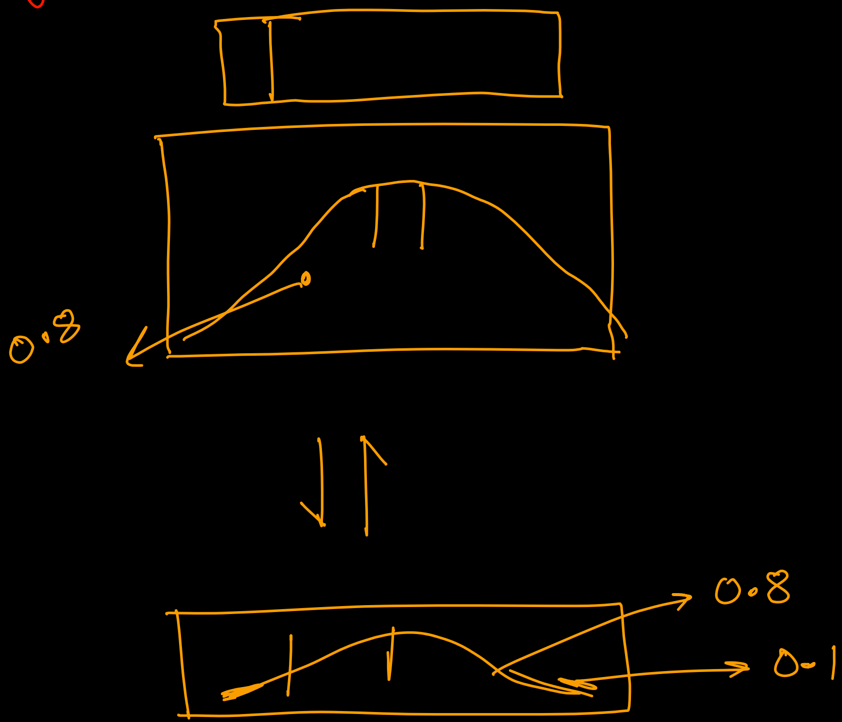
α = 0.02

p-value < α

$\Rightarrow$ reject Ho

P-value

Probability value



$H_0 \Rightarrow$ It is an hypotheses which assumes everything ~~fair~~ and equal.

p-value indicates the probability of obtaining the observed result if null hypothesis is true.

Significant value $\Rightarrow$ domain expert (given)

p-value is calculated (calculated)

if our p value $< \alpha$

reject null hypothesis

if our p-value $> \alpha$

then we accept the null hypothesis

Q-3) A company man. bikes with an average life of 2 or more yrs. an engineer believes this value to be less and he thus takes a sample of $\boxed{10 \text{ bikes}}$ & found avg. life to be 1.8 with stand. dev. of $\boxed{0.15}$

(a) H_0, H_1

(b) At a 99% CI, is their enough evidence to discard H_0?

$n = 10$

$H_0 \Rightarrow \mu \geq 2$

$H_1 \Rightarrow \mu < 2$

$n = 10$

$\bar{x} = 1.8$

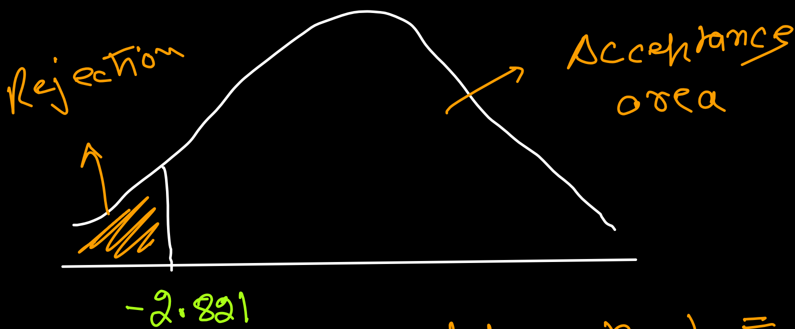
$s = 0.15$

$\alpha = 0.01$

$\Rightarrow$ as $n < 30$ & sample s.d is given

$\Rightarrow$ t - test

3 Decision boundary



$$dof = n - 1 = 10 - 1 = 9$$

(4) Calculate test statistics

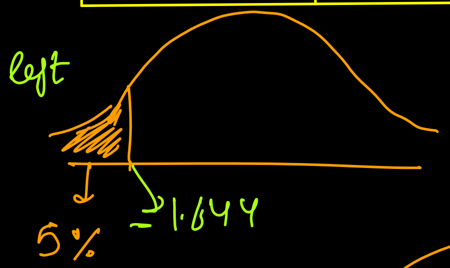
$$t = \frac{x - \bar{\mu}}{s/\sqrt{n}} =$$

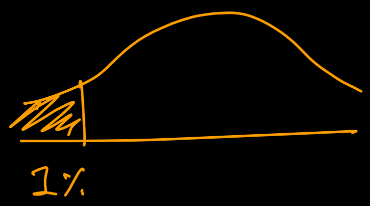
$$\frac{1.8 - 2}{0.15/\sqrt{10}} = -4.216$$

$$-4.216 < -2.821$$

$\Rightarrow$ reject null hypothesis

Z value	$\alpha = 5\%$	$\alpha = 1\%$
two tail	± 1.96	± 2.58
One tail	1.644 (right) -1.644 (left)	2.33 (right) 2.3 (left)





Chi-Square test (χ^2)

Chi-square is a statistical test which is used to determine if there is a significant association between categorical variables.

Used to compare observed data with expected data

It is a non parametric test

Q: In a 2000 US census, ages of individual in a small town were found to be below.

age	<18	18-35	>35
percentage	20%	30%	50%

In 2010, ages of 500 individuals were sampled & below were results

<18	18-35	>35
121	288	91

Using $\alpha = 0.05$, would you conclude that the population distribution has changed in last 10 yrs?

Ages	<18	18-35	35+
Expected %	20 %	30 %	50%
(2010) Observed	121	288	91
(Expected no.	100	150	250

Step 1.

$H_0 \Rightarrow$ data observed meets the data expected.

$H_1 \Rightarrow$ data do not meet expected.

② $\alpha = 0.05$

③ dof $= 3 - 1 = 2$
(no. of categories)

decision boundary from chi-square table 5.99

(4) Calculate test statistics

$$\chi^2 = \sum_{i=0}^{n} \frac{(f_o - f_e)^2}{f_e}$$

Ages	<18	18-35	35+
Expected %	20 %	30 %	50%
(2010) Observed	121	288	91
(Expected no.	100	150	250

f_o = observed value

f_e = expected value

$$= \frac{(121-100)^2}{100} + \frac{(288-150)^2}{150} + \frac{(91-250)^2}{250}$$

$$= \frac{21^2}{100} + \frac{138^2}{150} + \frac{(159)^2}{250}$$

$$\chi_c^2 = 232.494$$

If $\chi_c^2 >$ (decision boundary)

then we reject null hypothesis

$232.494 > 5.99$

$\Rightarrow$ reject null hypothesis

1. A poker-dealing machine is supposed to deal cards at random, as if from an infinite deck.

In a test, you counted 1600 cards, and observed the following:

Spades	404
Hearts	420
Diamonds	400
Clubs	376

Could it be that the suits are equally likely? Or are these discrepancies too much to be random?

Z - test with proportion

Q ⇒ A tech company believes that % of residence in Society Prestige that own mobile is 70%.

A marketing manager believes it is wrong & conducts a survey of 200 people out of which 130 response yes to owing a mobile

1. H_0 H_1
2. At 95% CI, enough evidence to reject null hypothesis

① $H_0 \Rightarrow p_0 = 70\%$

$H_1 \Rightarrow p_0 \neq 70\%$

$n = 200$

$x = 130$

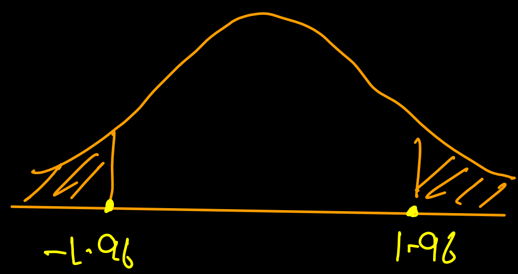
$p_0 = 70\%$

$\alpha = 0.05$

② $\hat{p} = \frac{x}{n} = \frac{130}{200} = 0.65$

$p_0 = 0.7 \qquad q_0 = 1 - p_0 = 0.3$

(3)



(4) Calculate statistic of test

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0 q_0 / n}}$$

$$= \frac{0.65 - 0.7}{\sqrt{\frac{0.7 \cdot 0.3}{200}}} = \frac{-0.05}{\sqrt{\frac{0.21}{200}}}$$

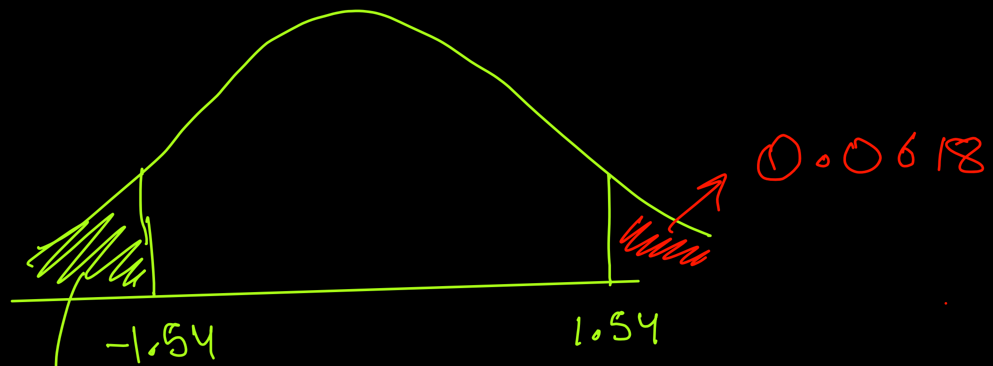
$$= -\frac{0.05 \times \sqrt{200}}{\sqrt{0.21}}$$

$$= -1.54$$

-1.54 $\qquad$ -1.9

as $\boxed{-1.54 > 1.96}$

we accept H_0



this area is 0.0618

$$p\text{-value} \Rightarrow 0.0618 * 2$$

$$\Rightarrow 0.1236$$

$$\alpha = 0.05$$

$$p > \alpha$$

$\Rightarrow$ accept null hypothesis

$\boxed{\text{Thanks a lot for staying.}}$:)