Step 1:

Null Hypothesis: H_0 :- The data meets the expected distribution.

Alternate Hypothesis, H_1 :- The data does not meet the expected distribution.

Step 2:

$\alpha = 0.05$, C.I = 95%

Step 3:

DOF = No. of category - 1 = 3 - 1 = 2

Step 4:

Decision Boundary : → from chi square table for DOF = 2 & $\alpha = 0.05$

if $\chi^2 > 5.991$ → Reject Null Hypothesis

Decision Boundary = 5.991

Step 5:

Chi Square Test Statistics

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

f_o = observed value

f_e = Expected value

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

$$\chi^2 = 232.494$$

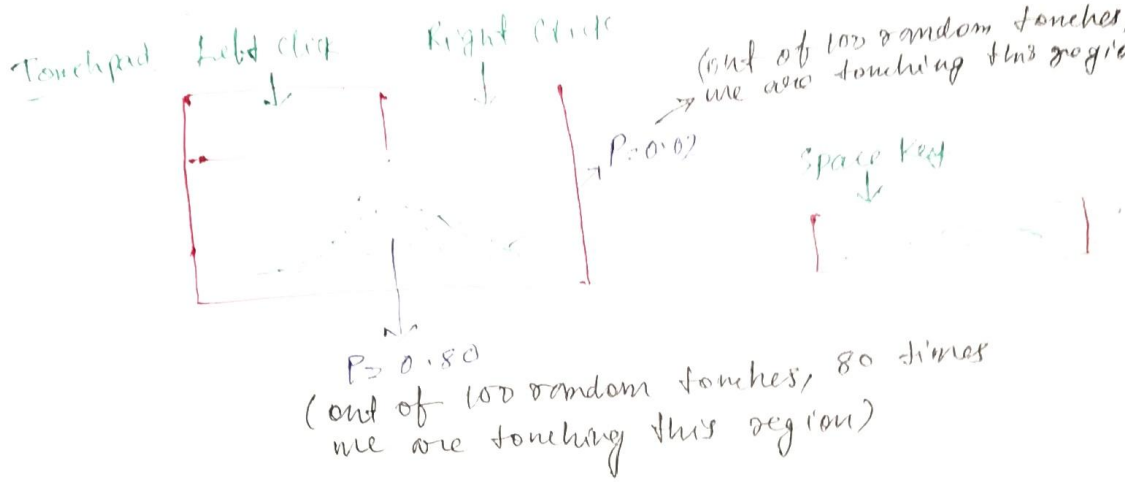
Conclusion

$\chi^2 > 5.991$ → Reject Null Hypothesis

It means, the data does not meet the expected distribution. The distribution has changed.

18-SEP-2022

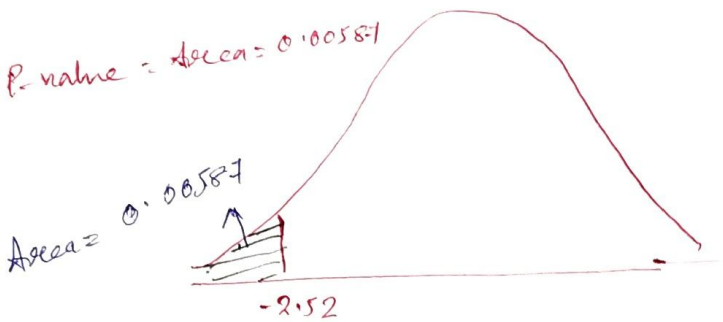
P-value



Note:

P-value and Significance values are completely different. Significance value is decided from confidence interval and P-value must be always less than or equal to significance value.

Eg: Previous to previous example



$\alpha = 0.02$

{ P-value $< \alpha$
$0.00587 < 0.02$
Reject Null Hypothesis. }

[if P-value $< \alpha \Rightarrow$ Reject Null Hypothesis
if P-value $\geq \alpha \Rightarrow$ Accept Null Hypothesis]

Que The average weight of all residents in a town xyz is 168 pounds. A researcher believes the true mean to be different. She measured the weight of 36 individuals and found the mean to be 169.5 pounds with a standard deviation of 3.9 Two tail → any. weight > 168 or any weight < 168

(a) Null Hypothesis and Alternate Hypothesis

(b) 95% CI. Is there enough evidence to discard the null hypothesis?

Ans Given: $\mu = 168$
$n = 36$
$\bar{x} = 169.5$
$s = 3.9$

Step 1:

Null Hypothesis : $\mu = 168$

Alternate Hypothesis : $\mu \neq 168$

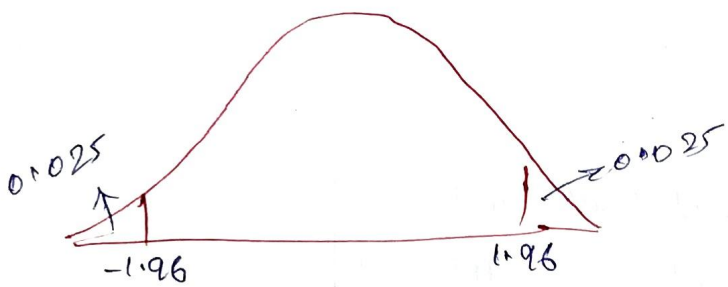
Step 2:

$CI = 95\%$

$\alpha = 1 - 0.95 = 0.05$

Step 3:

$n \geq 36$, we will use z-score.



z-score calculation

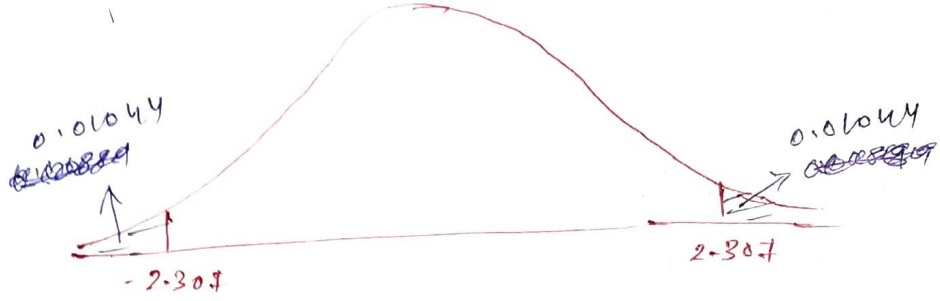
$$z\text{-score} = \frac{169.5 - 168}{\frac{3.9}{\sqrt{36}}} = 2.307$$

Conclusion

$2.307 > 1.96$

we will reject null hypothesis.

Using P-value



for $2.307 \approx 2.31$

Area = $1 - 0.99111, 0.98956 = 0.01044$ (from z-table)

~~$= 0.00889$ (from z-table)~~

for -2.307, ~~Area = 0.00889~~ ≈ 2.31, Area = P-value = ~~0.01044~~ 0.01044 (from z-table)

Conclusion

~~$0.00884 < 0.025$~~ P-value $= 0.01044 + 0.01044 = 0.02088$

~~P-value $< \alpha$ (S.V)~~ $0.02088 < 0.05$.

~~As~~ we will reject null hypothesis

T-Test

Que A company manufactures bikes batteries with an average life span of 2 year or more years. An Engineer believes this value to be less. Using 10 samples, he measures the average life span to be 1.8 years, with a standard deviation of 0.15. → one tail (left) T-test

(a) State the Null and Alternate Hypothesis?

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

Ans Given: $\mu = 2$, $n = 10$, $\bar{x} = 1.8$, $S = 0.15$.

Step 1:

H_0 : $\mu \geq 2$

H_1 : $\mu < 2$

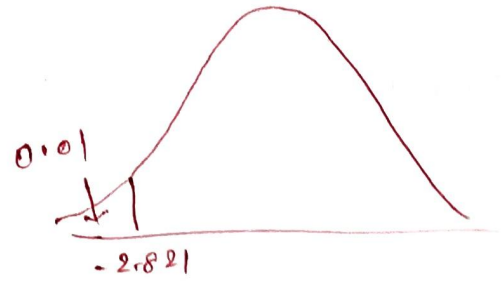
(left one tail test)

Step 2:

C.I = 99%

$\alpha = 1 - 0.99 = 0.01$

Step 3:



$n = 10 < 30$ and Sample standard deviation given, we will use t-test.

$DOF = n-1 = 10-1 = 9$

for $DOF = 9$ and $\alpha = 0.01$, (from t-table)

t-score $= -2.821$

t-test score calculation

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

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

$$= -4.216$$

Conclusion

$-4.216 < -2.821$

We will reject null hypothesis. The average life span of the batteries is less than 2 years.

Z-test with proportions

One A tech company believes that the percentage of residents in town xyz that owns a cell phone is 70%. A marketing manager believes that they value to be different. He conducts a survey of 200 individuals and found that 130 responded Yes owning a cell phone?

(a) State and null and alternate hypothesis

(b) At a 95% CI, is there enough evidence to reject the null hypothesis?

Ans **Step 1:**

Null Hypothesis : $P_0 = 0.70$

Alternate Hypothesis: $P_0 \neq 0.70$

Given: $n = 200$ $x = 130$

$$\hat{p} = \frac{130}{200} = 0.65 \quad \text{(proportion w.r.t Yes)}$$

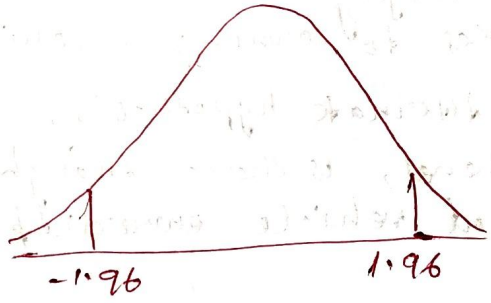
$$q_0 = 1 - P_0 = 1 - 0.70 = 0.30 \quad \text{(Remaining Proportion)}$$

Step 2:

$CI = 0.95$

$\alpha = 1 - 0.95 = 0.05$

Step 3:



→ -1.96 for $\alpha = 0.05$

Calculation

$$Z_{test} = \frac{\hat{p} - P_0}{\sqrt{\frac{P_0 q_0}{n}}}$$

$$Z\ test = \frac{0.65 - 0.70}{\sqrt{\frac{0.7 \times 0.3}{200}}} \simeq -1.54$$

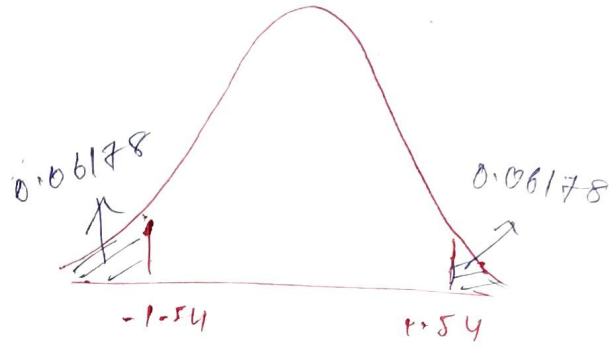
Conclusion

$-1.54 > -1.96$

Hence, we accept Null hypothesis. It means 70% of the people have cell phone.

Using P. value

P-value = 0.06178 + 0.06178
= 0.12356



P-value > Significance value
0.12356 > 0.05

Hence, we accept Null hypothesis. It means 70% of the people have cell phone.

Q. A car company believes that the percentage of residents in city ABC that owns a vehicle is 60% or less. A sales manager disagree with this. He conducts a hypothesis testing surveying 250 residents and found that 170 responded Yes to owning a vehicle?

(a) State the Null and Alternate hypothesis.

(b) At 10% significance level, is there enough evidence to support the idea that vehicle ownership in city ABC is 60% or less?

Ans Step 1:

Null Hypothesis : $P_0 \geq 0.60$

Alternate Hypothesis: $P_0 \neq 0.60$

Given: $n = 250$ $\quad x = 170$

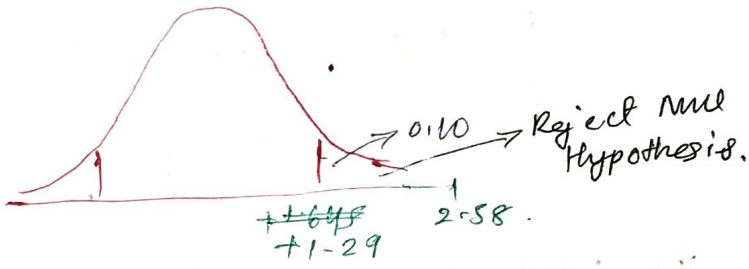
$$\hat{p} = \frac{170}{250} = 0.68$$

$$q_0 = 1 - P_0 = 1 - 0.60 = 0.40$$

Step 2:

$$\alpha = 0.10$$

Step 3:



Calculation

$$Z_{test} = \frac{\hat{P} - P_0}{\sqrt{\frac{P_0 q_0}{n}}}$$

$$= \frac{0.68 - 0.60}{\sqrt{\frac{0.60 \times 0.40}{250}}}$$

$$= 2.58$$

Conclusion

$2.58 > 1.29$,

→ Reject Null Hypothesis. It means that vehicle ownership in city ABC is 60% or less.

Chi Square Test

- Chi Square Test checks about population proportion.
- It is a non-parametric test that is performed on categorical data.
  - Ordinal data
  - Nominal data
- Here we don't consider 1 tail or 2 tail test.

Ex: In the 2000 U.S census the age of individuals in a small town found to be the following

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

→ These categories in Chi Square test should be used.

In 2010, ages of n = 500 individuals were sampled. Below are the result.

< 18	18-35	> 35
121	288	91

Using $\alpha = 0.05$, would you conclude, the population distribution of ages has changed in the last 10 years.

Ans

	< 18	18-35	> 35
Expected	20%	30%	50%

n = 500

	< 18	18-35	> 35
Observed	121	288	91
Expected	100	150	250
	↓ 20% of 500	↓ 30% of 500	↓ 50% of 500