Hypothesis Testing is a statistical method used to make inferences or decisions about a population based on sample of data. It helps you determine whether there is enough evidence to support a specific claim or belief about a population.

## Core Idea:

1.) **Null Hypothesis ($H_0$):** The default or status quo. (e.g. nothing has changed or there is no effect).

2) **Alternative Hypothesis:** ($H_1$ or $H_a$) - what you're trying to improve (e.g. something has changed, or there is an effect).

## 4 steps of Hypothesis Testing:

Formulate Hypothesis → Finalize the right test → Conduct a test → Make a decision

## Z-Test

2 approaches:

Rejection Region

P-Value

### Rejection Region:

$H_0$ = Housing inflation is 10%

$H_a$: Housing inflation is $> 10\%$

Population mean $(\mu) = 10$

Population S.D $= (\sigma) = 4$

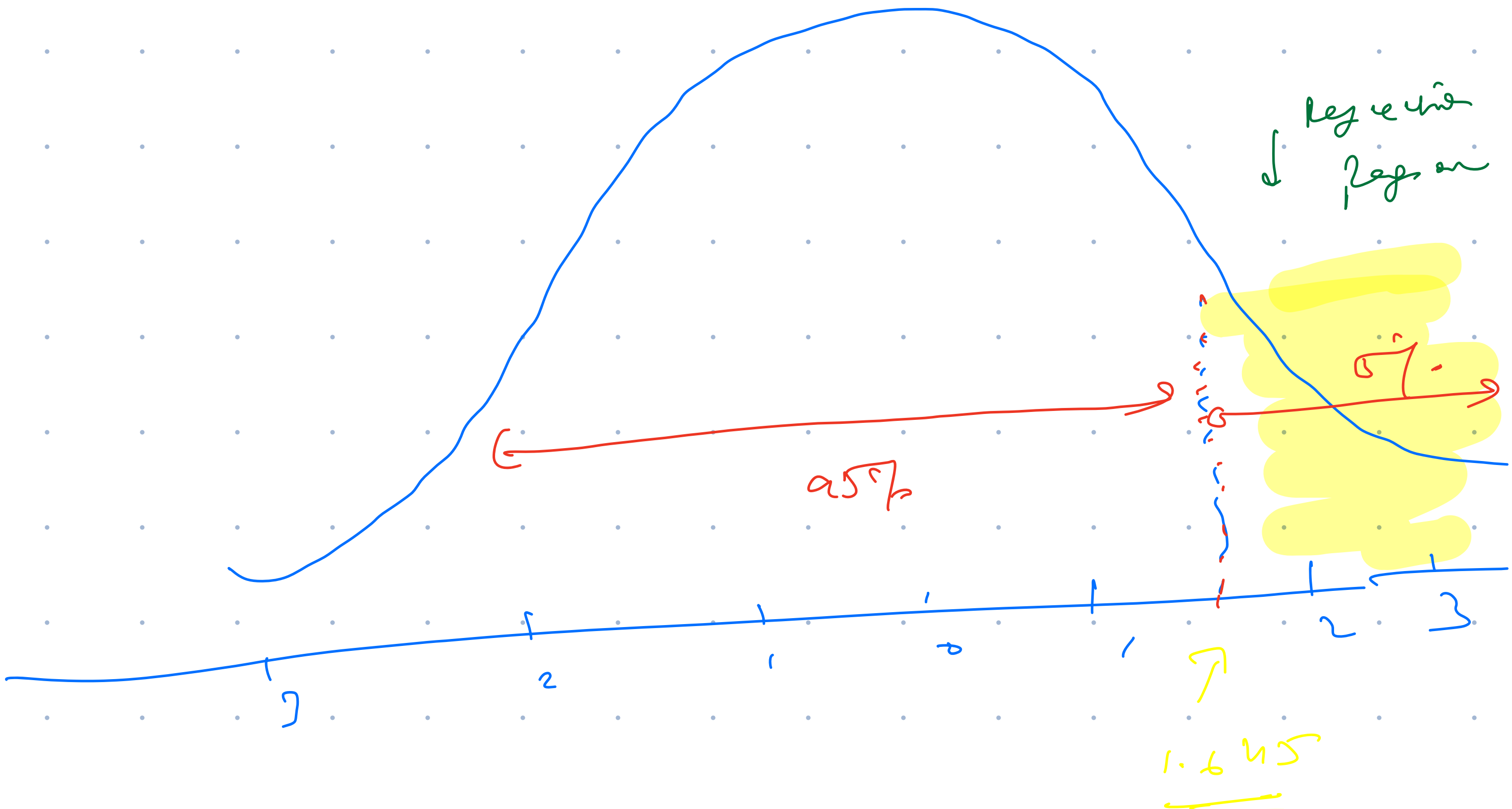
Sample mean $(\bar{x}) = 11$

Sample size $(n) = 100$

Let's say:

Confidence Level = 95%

then Significance Level $(\alpha) = 1 - CI = \underline{5\%}$




z-score of sample

$$z = \frac{x - \mu}{\sigma} = \frac{11 - 10}{4/\sqrt{n}} = \frac{11 - 10}{4/\sqrt{100}} = \frac{11 - 10}{0.4}$$

$$= \frac{1}{0.4}$$

$$= \underline{2.5}$$

$z_{cr} = 1.645$

$z = 2.5$

Since $z > z_{cr}$, we can reject our null hypothesis

## When to use Z - Test.

→ Use when:

1.) σ known

2.) Sample size is large ( $n \geq 30$ )

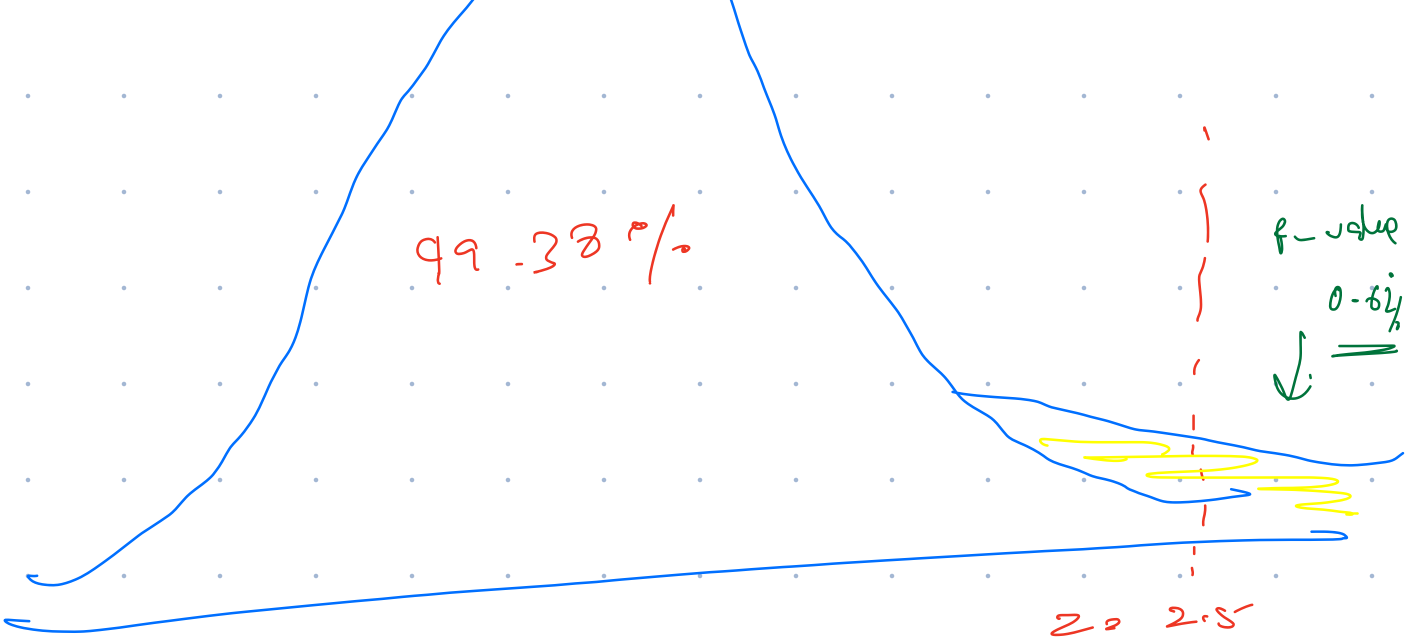
3.) mean Comparison ( Comparing sample mean to a population mean).

## ( p - value Method )

↳ what is the probability that the observed difference could have occurred by chance?

OR

Assuming the null hypothesis is correct, what is the probability of obtaining results as extreme as observed in the statistical test.




p-value = 0.62%

p-threshold = 5%

Since p-value < p-threshold
we reject the null hypothesis.

## Type 1 and Type 2 errors:

(1) Type 1: (False Positive):

You reject the null hypothesis even though it's actually true.

→ happens when your sample happens to show a rare outcome just by random chance.

Why it happens?

1.) Problem with the sampling process.

2.) Higher significance level ($\alpha$)

(2) Type 2 error :- ( false negative ) :

failing to reject the null hypothesis when its actually false.

→ Happens when your sample doesn't show strong evidence to detect a real effect.

Why this happens :

1.) small sample size

2.) lower significance level ($\alpha$)

Type 2 error is represented by $\beta$ :

$\beta$ = The probability of failing to reject $H_0$ when it is false.

## statistical Power

$1 - \beta$ = The probability of correctly rejecting a false null hypothesis.

## effect size:

→ a measure of how big the difference is b/w the population mean and your sample mean.

$$\text{effect size (cohen's d)} = \frac{|\bar{x} - \mu|}{\sigma}$$

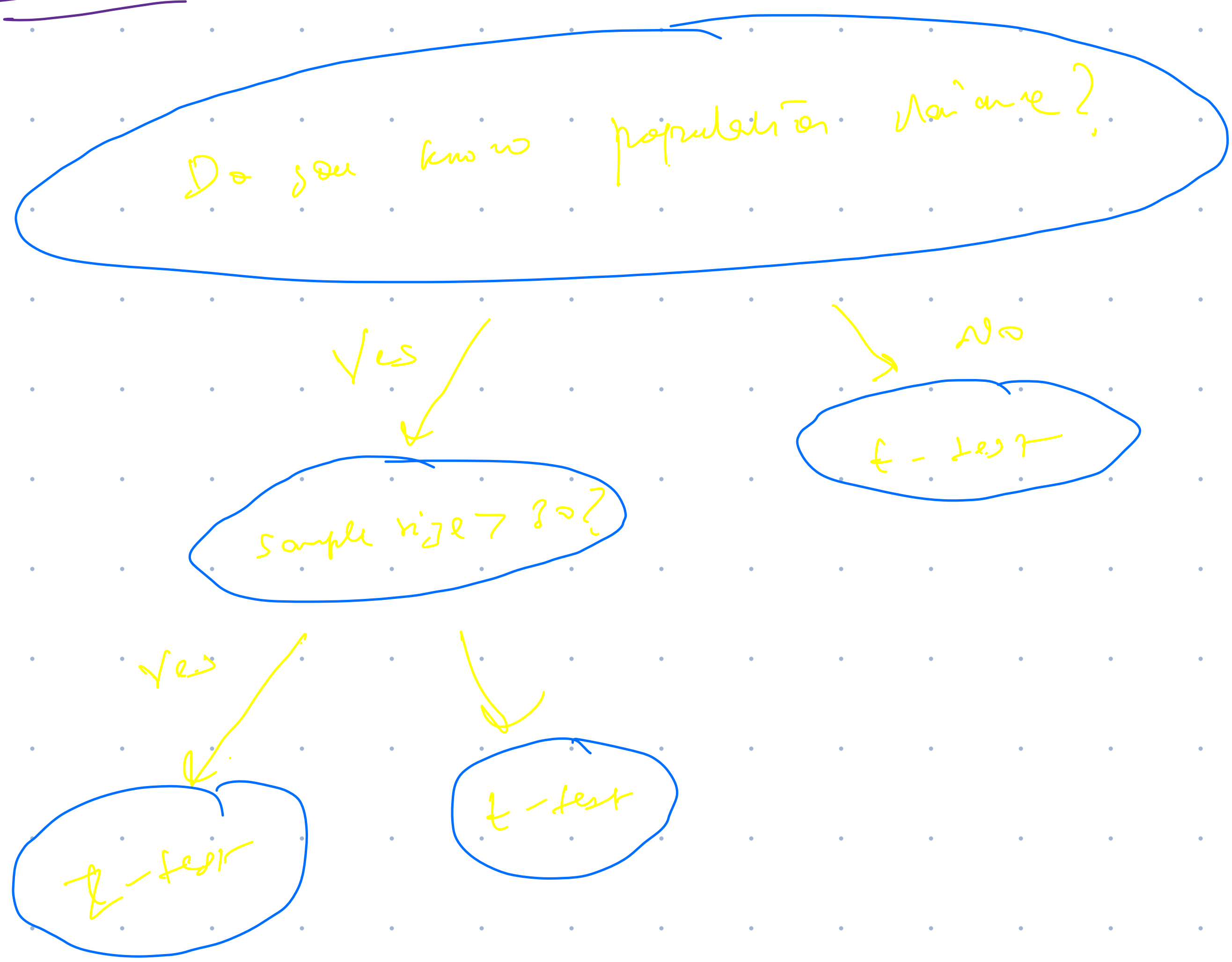
→ its unitless, help compares across studies

→ the larger the effect size, the easier it is to detect.

# A/B Testing:

- A/B testing is a method of comparing 2 versions (A and B) of something - like a webpage, feature or ad. - to see which one performs better.

## Z- Test:



Population mean ($\mu$) = 72

Population std ($\sigma$) = ?

6

sample size ($n$) = 20

Sample std dev ($s$) = 7.82

$$\sigma_{\bar{x}} = \frac{s}{\sqrt{n}}$$

$n = 20$ $\quad \bar{x} = 7.59$

| name | score |
| --- | --- |
| | |
| | |
| | |

$$t = \frac{\bar{x} - \mu}{\sigma_{\bar{x}}}$$

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

## Chi-squared Distribution :

→ is a right-skewed distribution used in hypothesis testing, particularly for categorical data.

→ Commonly used to test :-

1) goodness of fit : does the observed data match expected proportions.

2) independence :- Are 2 categorical variables related.