

Hypothesis Testing



What is Hypothesis?

- - Hypothesis is a predictive statement, capable of being tested by scientific methods, that relates an independent variables to some dependent variable.
 - A hypothesis states what we are looking for and it is a proportion which can be put to a test to determine its validity
- e.g.
- Students who receive counseling will show a greater increase in creativity than students not receiving counseling

Characteristics of Hypothesis

- Clear and precise.
- Capable of being tested.
- Stated relationship between variables.
- limited in scope and must be specific.
- Stated as far as possible in most simple terms so that the same is easily understood by all concerned. But one must remember that simplicity of hypothesis has nothing to do with its significance.
- Consistent with most known facts.
- Responsive to testing within a reasonable time. One can't spend a lifetime collecting data to test it.
- Explain what it claims to explain; it should have empirical reference.

Types of Hypothesis

- **Null Hypothesis (H_0)**
- **Alternative Hypothesis (H_a or H_1)**

$H_0 : \mu = \mu_0$	$H_a : \mu \neq \mu_0$
$H_0 : \mu \leq \mu_0$	$H_a : \mu > \mu_0$
$H_0 : \mu \geq \mu_0$	$H_a : \mu < \mu_0$

Null Hypothesis

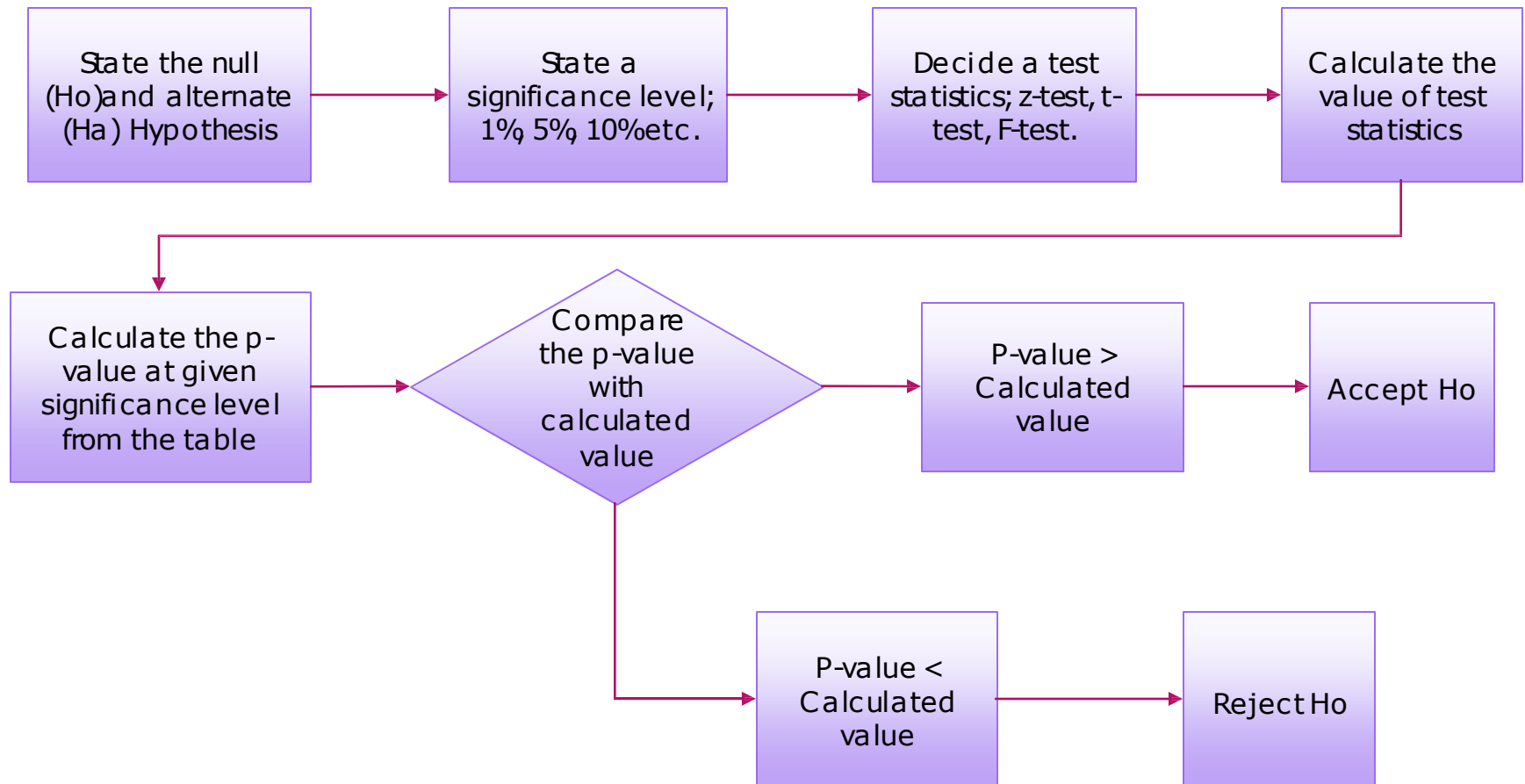
- It is an assertion that we hold as true unless we have sufficient statistical evidence to conclude otherwise.
- Null Hypothesis is denoted by H_0
- If a population mean is equal to hypothesised mean then Null Hypothesis can be written as

$$• H_0: \mu = \mu_0$$

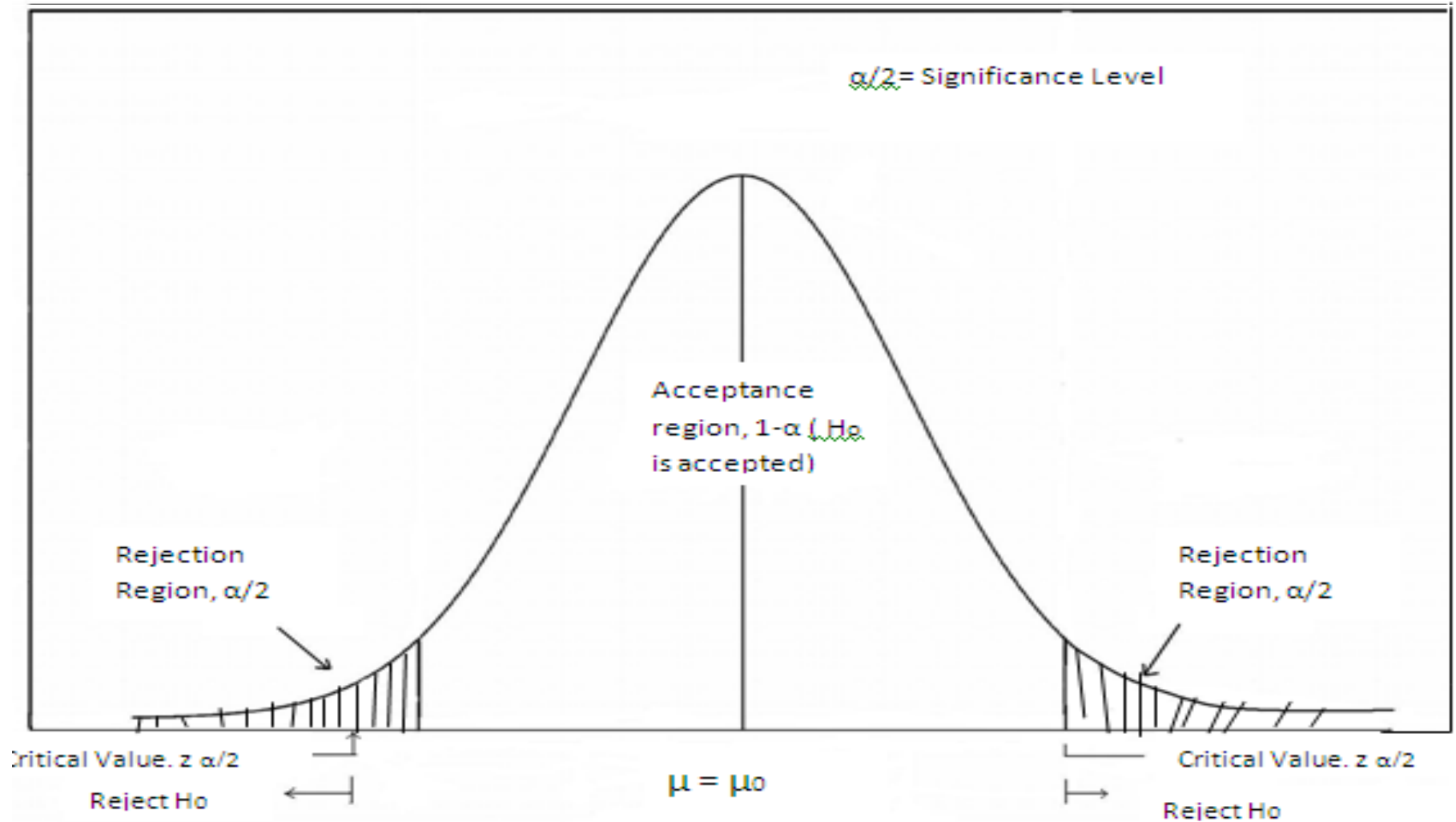
Alternative Hypothesis

- The Alternative hypothesis is negation of null hypothesis and is denoted by H_a
- If Null is given as $H_0: \mu = \mu_0$
- Then alternative Hypothesis can be written as
 - $H_a: \mu \neq \mu_0$
 - $H_a: \mu > \mu_0$
 - $H_a: \mu < \mu_0$

Procedure for Hypothesis Testing



Establish Critical or Rejection region



Areas of Accepted and Rejection of H_0 (Two -Tailed test)

Select the Suitable Test of significance or Test Statistic

- Whether the test involves one sample, two samples, or samples?
- Whether two or more samples used are independent or related?
- Is the measurement scale nominal, ordinal, interval, or ratio?

The choice of a probability distribution of a sample statistics is guided by the sample size n and the value of population standard deviation σ as shown in the table.



Sample Size n	Population Standard Deviation σ	
	Known	Unknown
$n > 30$	Normal Distribution	Normal Distribution
$n \leq 30$	Normal Distribution	t-distribution



Is the Population normal?

No

Yes

Is the Sample Size, $n \geq 30$?

No

Yes

Use a non-parametric test

Use Z-test statistic

Is the Population Standard deviation, σ known?

No

Yes

Use t-test statistic

Use z-test statistic

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}} \quad \text{or} \quad \frac{\bar{X} - \mu}{s / \sqrt{n}}$$

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

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$



Types of error		
Type of decision	H_0 true	H_0 false
Reject H_0	Type I error (α)	Correct decision ($1-\beta$)
Accept H_0	Correct decision ($1-\alpha$)	Type II error (β)

Summary of Certain Values for Sample Statistics Z

Rejection region	Level of Significance, α percent			
	$\alpha = 0.10$	$\alpha = 0.05$	$\alpha = 0.01$	$\alpha = 0.005$
One - tailed region	± 1.28	± 1.645	± 2.33	± 2.58
Two - tailed region	± 1.645	± 1.96	± 2.58	± 2.81