Type II error!

Accept Null Hypotheses when it is false.

Power of a Test: POT is defined as the probability of rejecting the NH when it is false.

→ To avoid Type II error, one need to maximize the power.

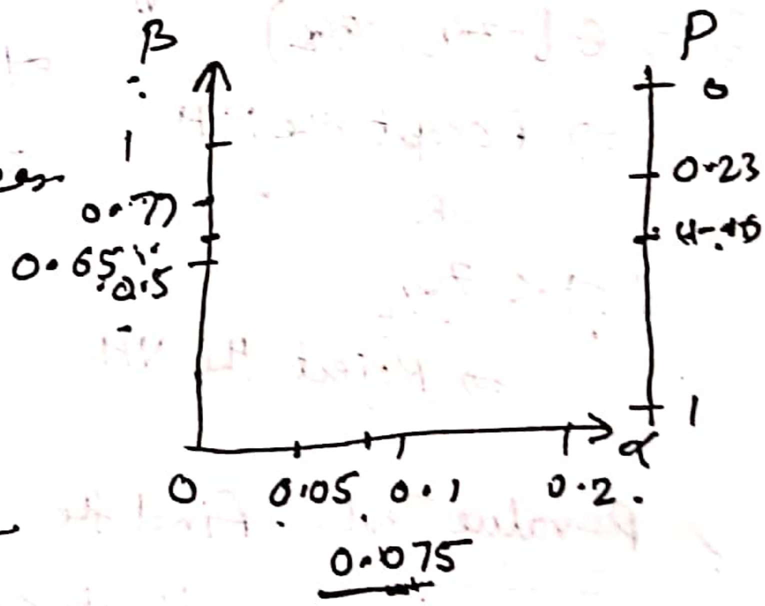
→ Probability of Type II error is always represented by β.

→ $power = 1-\beta \equiv 1 - P(\text{Type II error})$

* The power or β depends on multiple factors, and some of the most common factors are as follows.

(i) LOS:

→ if α increases, β decreases



(ii) Alternative hypotheses:

True value in H_0 and the hypothesized value

*(iii) Sample size:- A large sample size will help in reducing Type II error.

(IV) Standard deviation: A sample with low std then the Type II error is minimal.

Hypotheses Testing Procedure:

1) Construct Null Hypotheses. (H_0)
2) Define Alternative Hypotheses (H_1) after a careful study of the problem given. Also define the nature of the HT problem (one tail or two tail probem).
3) Consider LOS or take from the given problem, if specified.
4) Find the test statistic: (T)
5) Find the critical point and decide wheter to accept or reject the ~~N~~NT by comparing T and critical point.

If $T \in [-Z_{\alpha/2}, Z_{\alpha/2})$

$\Rightarrow$ Accept the NH

OR

$|T| < Z_{\alpha/2}$

$\Rightarrow$ Reject the NH

(Method of Confidence Interval)

OR

* P-value Test: Find the p-value and reject the NH if $p < \alpha$.

(Two tail Analysis)

Tests of Hypotheses: Summary (w.r to Mean)

S. No	Test For	Null Hypothesis (H_0)	Test Statistic (Z)	Distribution	Use When	p-value	Reject H_0
1	Population mean (μ)	$\mu = \mu_0$	$Z = \dfrac{\bar{X} - \mu}{\frac{\sigma_0}{\sqrt{n}}}$	Z	Normal distribution or $n > 30$; σ known	$p = 2P(Z \geq \lvert T \rvert)$	$p < \alpha$ $\lvert T \rvert > z_{\alpha/2}$
2	Population mean (μ)	$\mu = \mu_0$	$t = \dfrac{\bar{X} - \mu}{\frac{S_0}{\sqrt{n}}}$	t $n-1$ dof	$n < 30$, and/or σ unknown	$p = 2P(T_{n-1} \geq \lvert T \rvert)$	$p < \alpha$ $\lvert T \rvert > t_{\alpha/2}$
3	Difference of two means ($\mu_1 - \mu_2$)	$\mu_1 - \mu_2 = 0$ or $\mu_1 = \mu_2$	$\dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	Z	Both normal distributions, or $n_1, n_2 \geq 30$; σ_1, σ_2 known	$p = 2P(Z \geq \lvert T \rvert)$	$p < \alpha$ $\lvert T \rvert > z_{\alpha/2}$
4	Difference of two means ($\mu_1 - \mu_2$)	$\mu_1 - \mu_2 = 0$	$t = \dfrac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$	t distribution with df = the Sum ~~smaller of~~ $n_1 - 1$ and $n_2 - 1$	$n_1, n_2 < 30$; and/or σ_1, σ_2 unknown $\sigma_1 = \sigma_2$ $S_1 = S_2 = S$	$p = 2P(T_{dof} > \lvert T \rvert)$ dof: $n_1 + n_2 - 2$	$p < \alpha$ $\lvert T \rvert > t_{\alpha/2}$

1) Sample size (n)
- $n \geq 30 \rightarrow$ CLT $\rightarrow$ ND
- $n < 30$ ($\not\rightarrow$ CLT $\not\rightarrow$ ND) (t distribution) Dof

2) Standard deviation:
- σ is known $\rightarrow$ SND $\quad \dfrac{\bar{X} - \mu}{\sigma} \sim N(0,1)$
- σ is unknown
  - $\hookrightarrow$ t-distribution.

$p = 2P(Z \geq \lvert T \rvert)$

Q) A sample of 100 students taken from IIT Jodhpur. The mean height of the students in this sample is 160 cm. What is the validity of statement that the mean height of the population is 165 cm. Given the SD is 10 cm.

Soln

S: $\bar{x} = 160$ $n = 100$

P: $\mu = 165$ cm $\sigma = 10$ cm

1) $H_0 : \mu = 165$

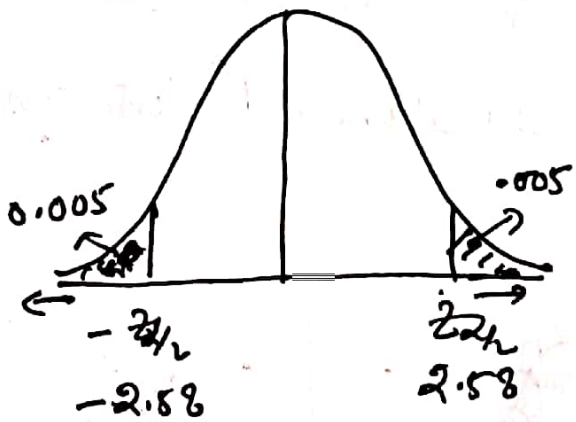
2) $H_1 : \mu \neq 165$ } (Two tail Analysis)

3) $\alpha = 0.01$ (1%) LOS

4) $z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{160 - 165}{10/\sqrt{100}} = -5$

5) (a) By z table

$\boxed{z_{\alpha/2} = 2.58}$



$\Rightarrow |-5| = |5| > z_{\alpha/2} = 2.58$

$\Rightarrow$ Ho is Rejected.

$\Rightarrow$ Mean heigh of the population is not 165 cm.

OR

(b) $p = 2P(z \geq |5|)$

$= 2P(z \geq 5) = ($ $)$ HW

$\Rightarrow p < 0.01$

$\Rightarrow$ Ho is Rejected.

Q2 Let X denote the growth of a tree in 15 days. It is known that average growth is 4mm. A sample of 20 trees is selected and has the mean of 3.8 mm with s.d of 0.3 mm. What can you say about the population mean.

Sol^n S: $\bar{x} = 3.8$ mm , $s = 0.3$ mm $n = 20$

P: $\mu = 4$ mm

1) $H_0: \mu = 4$ mm

2) $H_1: \mu \neq 4$ mm $\Rightarrow$ Two tail analysis

3) LOS $\alpha = 0.01$

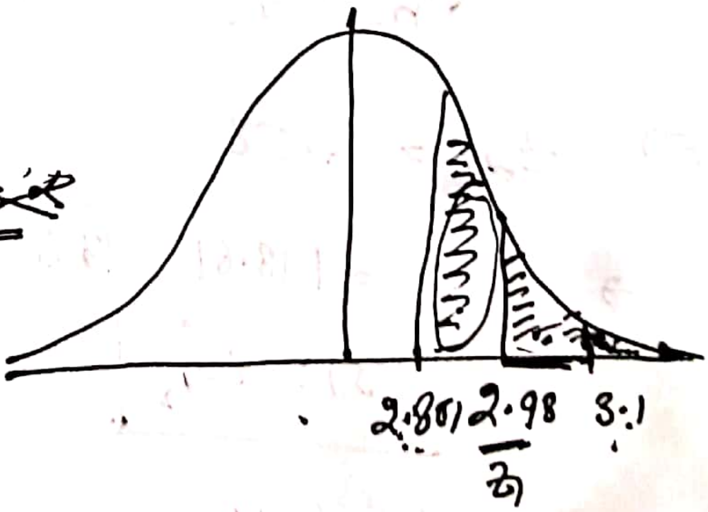
4) $t = \dfrac{\bar{x} - \mu}{s/\sqrt{n}} = \dfrac{3.8 - 4}{0.3/\sqrt{20}} = -2.98$

5) $p = 2P(t_{19} \geq \underset{t_1}{2.98}) < 2P(t_{19} \geq \underset{t_2}{2.861})$

$= 2 \times 0.005 = 0.01$

$\Rightarrow p = \alpha$ or $p \leq \alpha$

$\Rightarrow$ Reject NH H_0



Q3 In a random sample of size 500, the mean is found to be 20. In another independent sample of size 400, the mean is 15. What can you say about the samples whether they are drawn from same population with $\sigma = 4$.

Sol: Given S_1: $\overset{\mu_1}{\bar{x}_1} = 20$ $n_1 = 500$

S_2: $\overset{\mu_2}{\bar{x}_2} = 15$ $n_2 = 400$

P: $\sigma = 4$

1) $H_0: \bar{x}_1 = \bar{x}_2$ ||| $\mu_1 = \mu_2$

2) $H_1: \bar{x}_1 \neq \bar{x}_2$ ||| $\mu_1 \neq \mu_2$ ⇒ Two tail Analysis

3) LOS $\alpha = 0.01$

4) $$J = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{20-15}{4\sqrt{\frac{1}{500} + \frac{1}{400}}} = 18.6$$

5) $z_{\alpha/2} = 2.58$

⇒ $|J| = |18.6| = 18.6 > 2.58 \equiv z_{\alpha/2}$

⇒ $\boxed{|J| > z_{\alpha/2}}$

⇒ Reject H_0

⇒ Both the samples are not drawn from same population.

n_2

Power of Test Revisited:

$\text{Power} = 1 - \beta = 1 - P(\text{Type II error})$

$= 1 - P(\text{Accepting } H_0 \text{ when it is false})$

$= P(\text{Rejecting } H_0 \text{ when it is false})$

Q Suppose we are about to randomly sample 16 values for a ND population with $\sigma = 8$. Consider the following hypotheses.

$H_0: \mu = 75$

$H_1: \mu \neq 75$

At $\alpha = 0.05$, what is the power of the test if true mean is 76.

P: $\sigma = 8$

S: $n = 16$, $\sigma_S = 8/\sqrt{16} = 2$

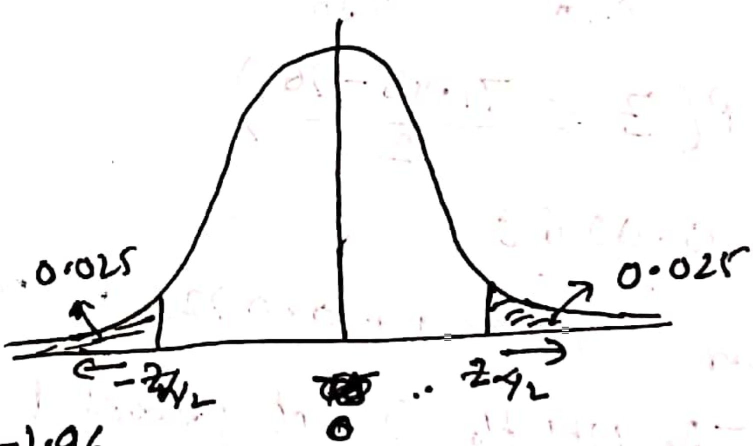
Sol

1) $H_0: \mu = 75$

2) $H_1: \mu \neq 75 \Rightarrow$ HT → Two tail analysis

3) Test statistic:

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



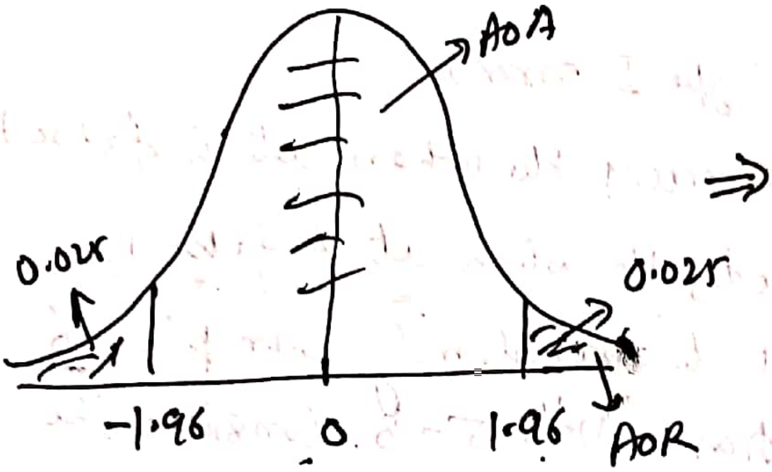
Reject H_0 when

$T \geq 1.96$ or $T \leq -1.96$

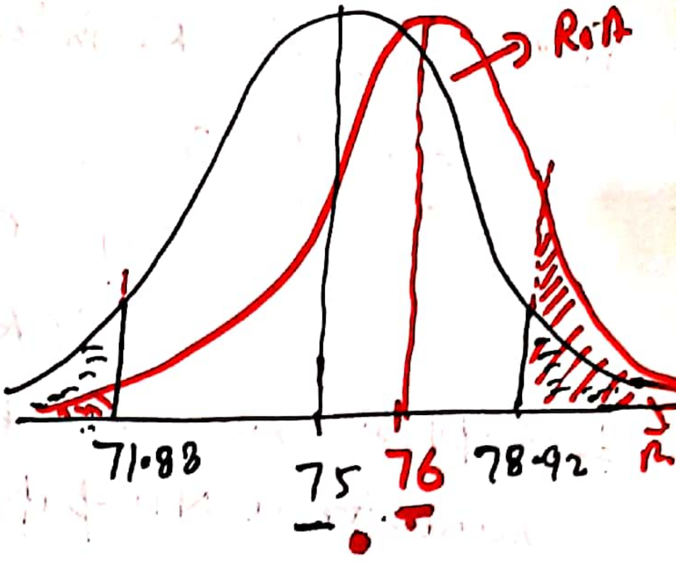
$\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \geq 1.96$ $\qquad$ $\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq -1.96$

$\Rightarrow \frac{\bar{X} - 75}{2} \geq 1.96$ $\qquad$ $\Rightarrow \frac{\bar{X} - 75}{2} \leq -1.96$

$\Rightarrow \bar{X} \geq 78.92$ $\qquad$ $\Rightarrow \bar{X} \leq 71.08$



⇒



4) Find the power:

Power = P(Reject H_0 | $\mu = 76$)

$= P(\bar{X} \leq 71.88 \mid \mu = 76) + P(\bar{X} \geq 78.92 \mid \mu = 76)$

(1) (2) —— (1)

(I)

$P(\bar{X} \leq 71.88 \mid \mu = 76)$

$= P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{71.88 - \mu}{\sigma/\sqrt{n}} \mid \mu = 76\right)$

$= P\left(Z \leq \frac{71.88 - 76}{2}\right)$

$= 0.0069$

(II) $P(\bar{X} \geq 78.92 \mid \mu = 76)$

$= P\left(Z \geq \frac{78.92 - 76}{2}\right)$

$= 0.0721$

Power $= 0.0069 + 0.0721 = .079$.

⇒ The NH H_0 is not powerful and has only 7.9% chances to detect the actual difference between the hypothesized and true mean.

⇒ P(Type II error) = 1 − Power

$= 1 - .079$

$= 0.921$ ⇒ ~92%.