Hypothesis Testing :-

* A Systematic procedure is used to predict whether the results obtained from a study supports a particular theory that is related to the population is known as hypothesis testing.

* Hypothesis test is also called as significance tests.

* It uses sample data in order to evaluate the hypothesis of the population.

* This test 2 hypothesis they are null hypothesis, alternate hypothesis

* Hypothesis testing was invented as a way to protect from being fooled by random chance, it literally means misinterpreting randomness.

Null Hypothesis :- (H_0)

* Null hypothesis is a comprehensive statement or default states that there is zero happening.

* In other words difference between 2 situations is zero. (no correlation)

* Null hypothesis is a hypothesis in which the sample observations results from the chance.

* It is generally assumed that the null hypothesis is true until any good other proof has been brought into the light to deny the hypothesis.

Alternate hypothesis :- (H_1)

* Its just a contradictory to null hypothesis
* the hypothesis which we are seeking to find evidence should be a Alternate hypothesis.
* It's a hypothesis that a random cause may influence the observed data or sample.
* It has 3 types :-

i) Left tailed :-

* It is expected that sample proportion (π) is less than a specified value (π_0)

$$H_1 : \pi < \pi_0$$

ii) Right tailed

* sample proportion (π) greater than some value

$$H_1 : \pi > \pi_0$$

iii) Two - tailed

* sample proportion (π) not equal to some value

$$H_1 : \pi \neq \pi_0$$

* For all these hypothesis the null hypothesis will be $H_0 : \pi = \pi_0$.

Statistical Significance:-

* In stats, significance means "not by chance" or probably true.

* Level of significance (α) is defined as the fixed probability of wrong elimination of null hypothesis when in fact, it is true.

* α is probability of Type I error.

* It defines whether Ho is assumed to be rejected or accepted.

Type I error:-

* Type I error appears when Ho of an experiment is true but still it is rejected.

* It is False positive (also called as), where a real hit was rejected by the test and is observed as a miss.

Type II error:-

* It occurs when a true alternative hypothesis is not acknowledged

* It means H_1 is rejected when it is True.

* It is also called as False negative.

ex:-

H_0 : Defendent is innocent

H_1 : Defendent is guilty

COURT DECISION

REALITY	Not guilty	Guilty	
not guilty (H_0 is true)	$1-\alpha$ TN	Type I Error α FP	$\Rightarrow 1-\alpha+\alpha=1$
Guilty (H_1 is true)	Type II Error β FN	Power $1-\beta$ TP	$\Rightarrow \beta+1-\beta=1$

Power of test

* the statistical ~~hy~~power of hypothesis is the probability that the test correctly rejects the Ho.

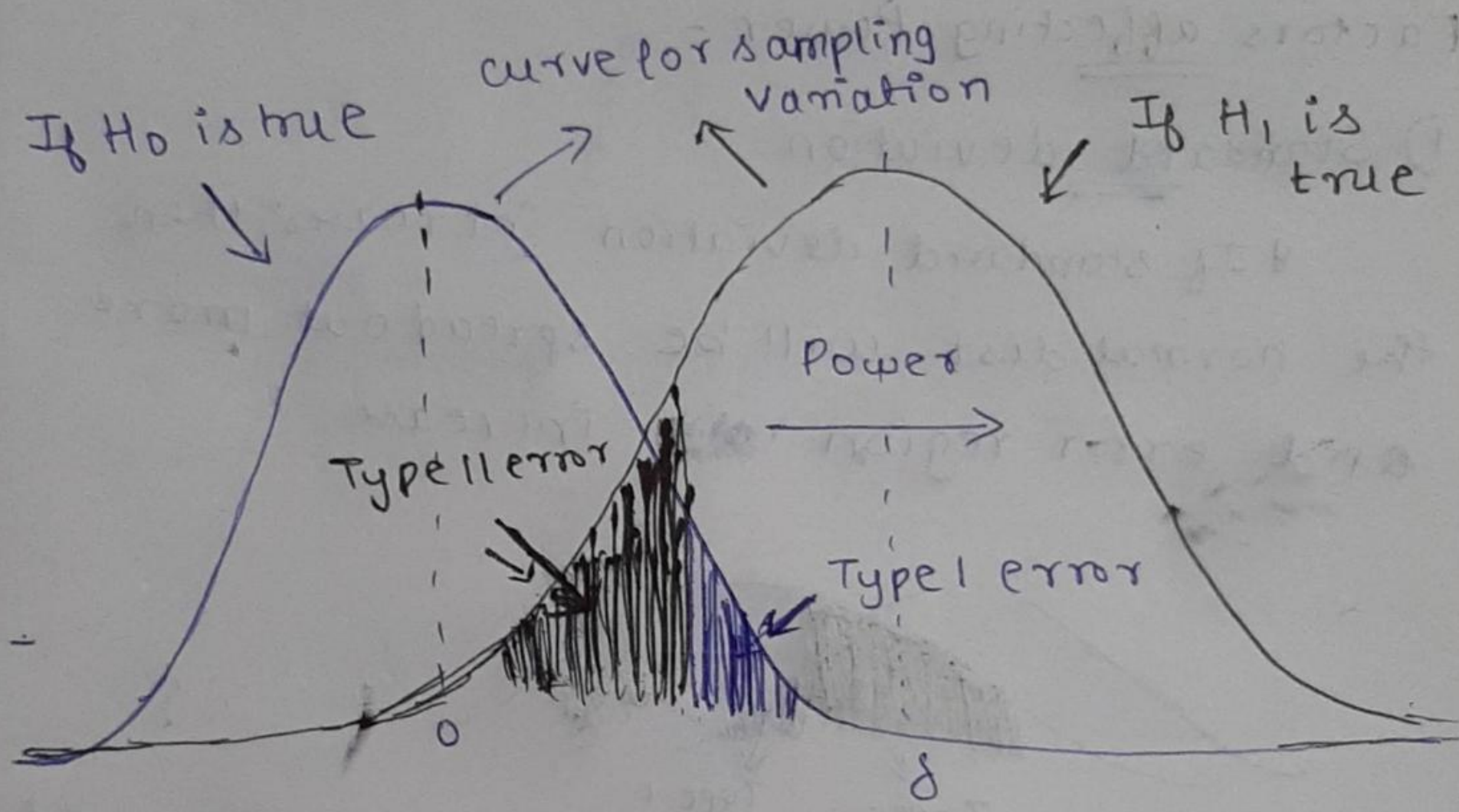
* If is $(1-\beta)$

Example:- (Depiction of possibilities through graph)

H_0 : No effect of smoking cessation on lung function ($\delta = 0$)

H_1 : Smoking cessation improves lung function ($\delta > 0$)

(smoking cessation means stopping the habit of smoking)



Explanation:-

* The Blue distribution represents the distribution if the H_0 is true means $\delta = 0$, means there is no change or difference.

* The level of significance (α) is set, It will usually 5-1. ($\alpha = 0.05$)

* The blue shaded region is type I error.

* The Black distribution represents the distribution if the H_1 is true means δ has some different value, means the chance has some effect.

✓ Since H_1 is ($\delta > 0$), the normal dist has the mean of δ.

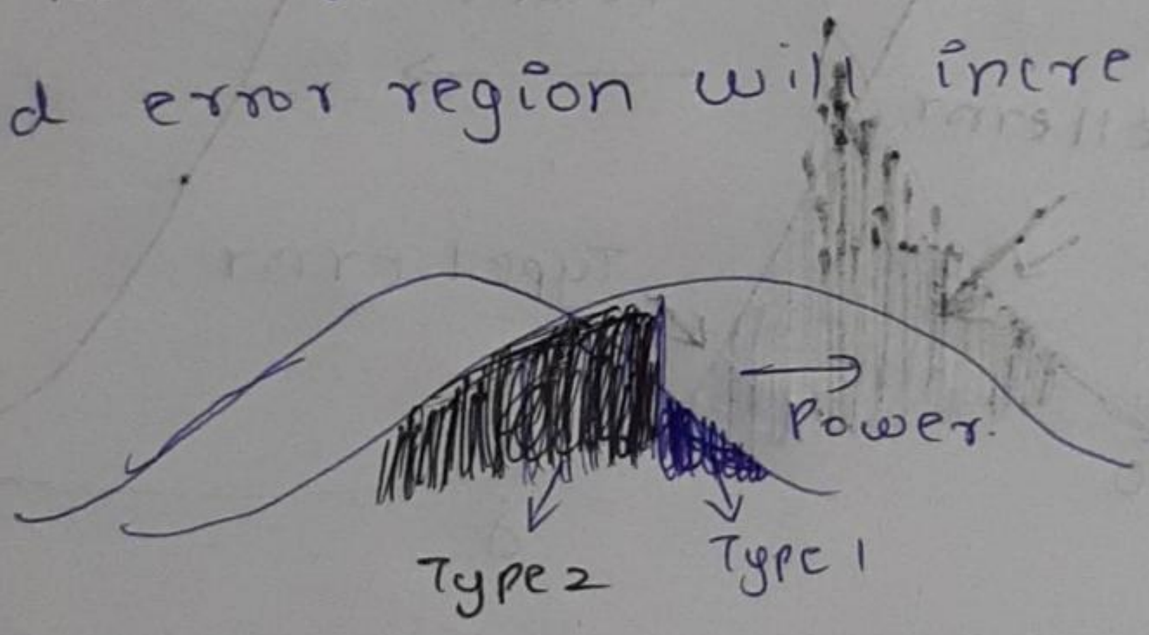
✓ The shaded region in black is Type II error, it can be set from the starting part of Type I error shaded region (that intersection part.)

✓ The Remaining region in Black curve is called the power of the test

Factors affecting Power:-

i) Standard deviation.

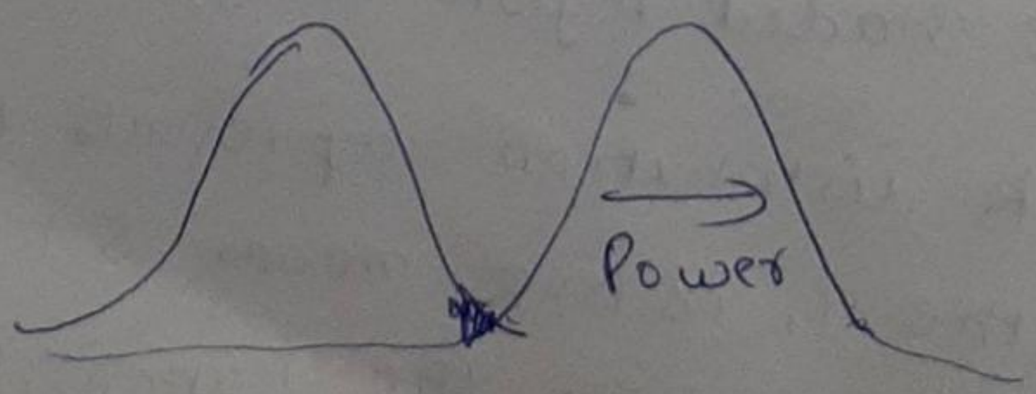
✓ If standard deviation increases then the normal dist will be spread out more and error region will increase.



✓ If S.D increases, Power decreases

ii) No. of. observations in sample (n)

✓ If no. of. observations in sample increases, the normal dist will become narrow and error region will decrease



* If n increases, Power increases

iii) δ (Difference between Hypothesis)

* If δ is large then both the distributions will be fall more apart from each other.

* If both dist completely apart from each other then it has more power.

(no overlapping of distributions)

* If δ increases, Power increases.