Apr 15

Hypothesis Testing

① Two competing statements, called hypothesis

② They are tested based on given/available data to decide which of these statements is close to the data.

Eg:

① A physician wants

to decide whether a

to decide whether a patient is suffering from a disease based on biochemical counts.

2. Aeroplane take off decision based on certain observations.

etc.

<u>Hypothesis</u>:

These are statements to be tested.

① Null Hypothesis :
Denoted by H_0 :
This is referred as the statement of "no difference".

② Alternate Hypothesis :
Denoted by H_1 or H_A :
Usually H_1 is the negation of H_0.

Therefore, H_0 & H_1 divides the parameter space ⓗ into two parts

(i) $H_0 = \{\theta : \theta \in \Theta_0\}$

(ii) $H_1 = \{\theta : \theta \in \Theta_1\}$

where $\Theta_0 \cup \Theta_1 = \Theta$

and $\Theta_0 \cap \Theta_1 = \phi$.

Eg: Consider a population with distribution $N(\theta, 1)$.

$H_0 : \theta = 1 \; ; \Theta_0 = \{1\}$

$H_1 : \theta \neq 1 \; ; \Theta_1 = \mathbb{R} \setminus \{1\}$

Another way to define Class of Hypothesis

(I) Simple Hypothesis:
A hypothesis is said to simple if it is a singleton set.

(II) Composite hypothesis:
A hypothesis which is not simple; is said to to be a Composite hypothesis.

Eg. $N(\theta, 1)$

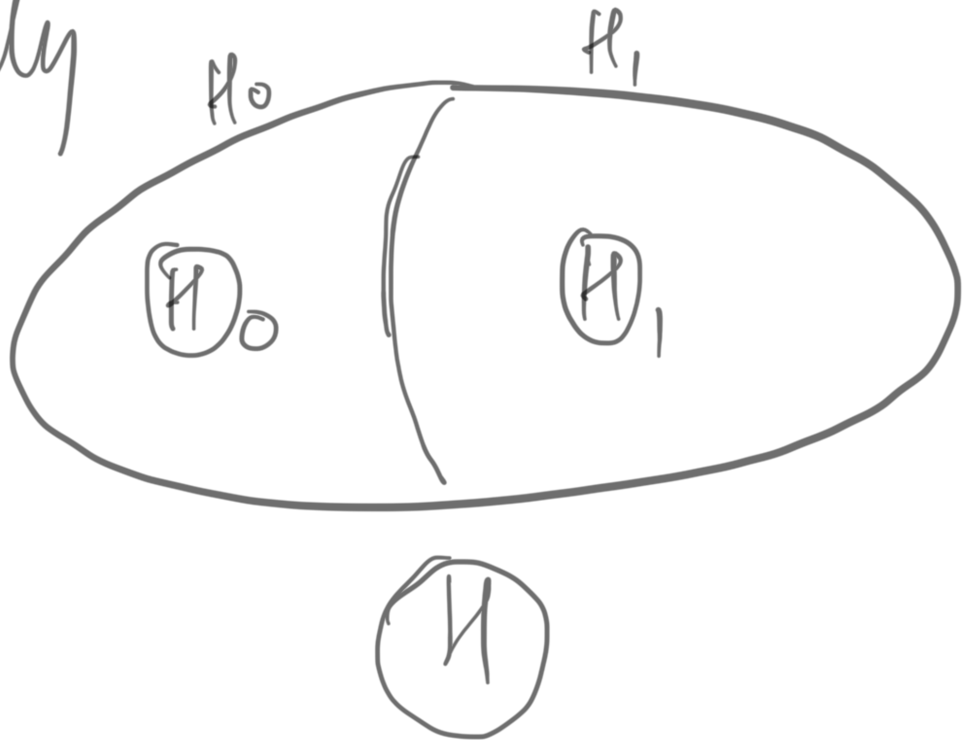
$H_1 : \theta = 1$

$H_2 : \theta = \{0, 3\}$

$H_3 : \theta \geq 1$

$H_4 : \theta \in (1,2) \cup \{5\}$

Graphically



For simplicity, let

$H_0: \theta = \theta_0$, θ_0 is fixed.

$H_1: \theta \neq \theta_0$

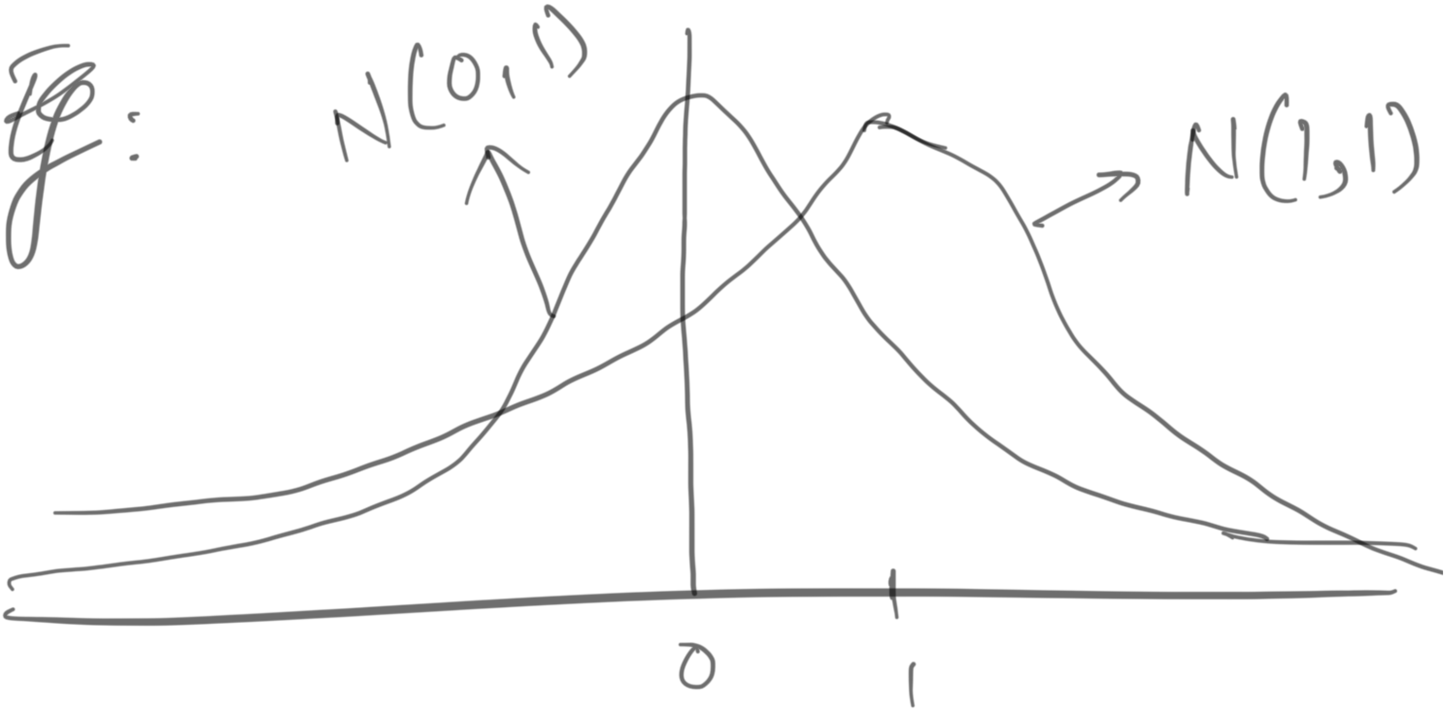
<u>Steps</u>.

(i) We have IID observations $X_1, \ldots, X_n$ from the pop.

(ii) We compute a point estimator $T = T(X_1, \ldots, X_n)$ for θ.

(iii) We check whether T is close to θ or not.

Eg:



$$\mathbb{P}((X_1, \dots x_n) \in N(0,1))$$

$$\mathbb{P}((X_1, \dots, x_n) \in N(1,1))$$

Error's

Type-I error :

"Type-II error":

H_0	True	False
Accept	✓	Type-II
Reject	Type-I	✓

Probability of type-I error

It is denoted by 'α'.

Also referred as "Level of

significance" or
"size of test".

Probability of type-II error

It is denoted by β.

Neyman-Pearson way:

Fix type-I error probability and minimize type-II error probability.

Eg: let $X_1, \ldots, X_n$ are from

IID observations from
$N(\theta, 1)$. We want to
test
$H_0 : \theta = 0$
against $H_1 : \theta \neq 0$.

A "nice" Statistic for θ is

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

And we know, Using CLT

$$\frac{\bar{X}_n - \theta}{1/\sqrt{n}} \sim N(0,1)$$

Let's denote

$\bar{X}_n - \theta$

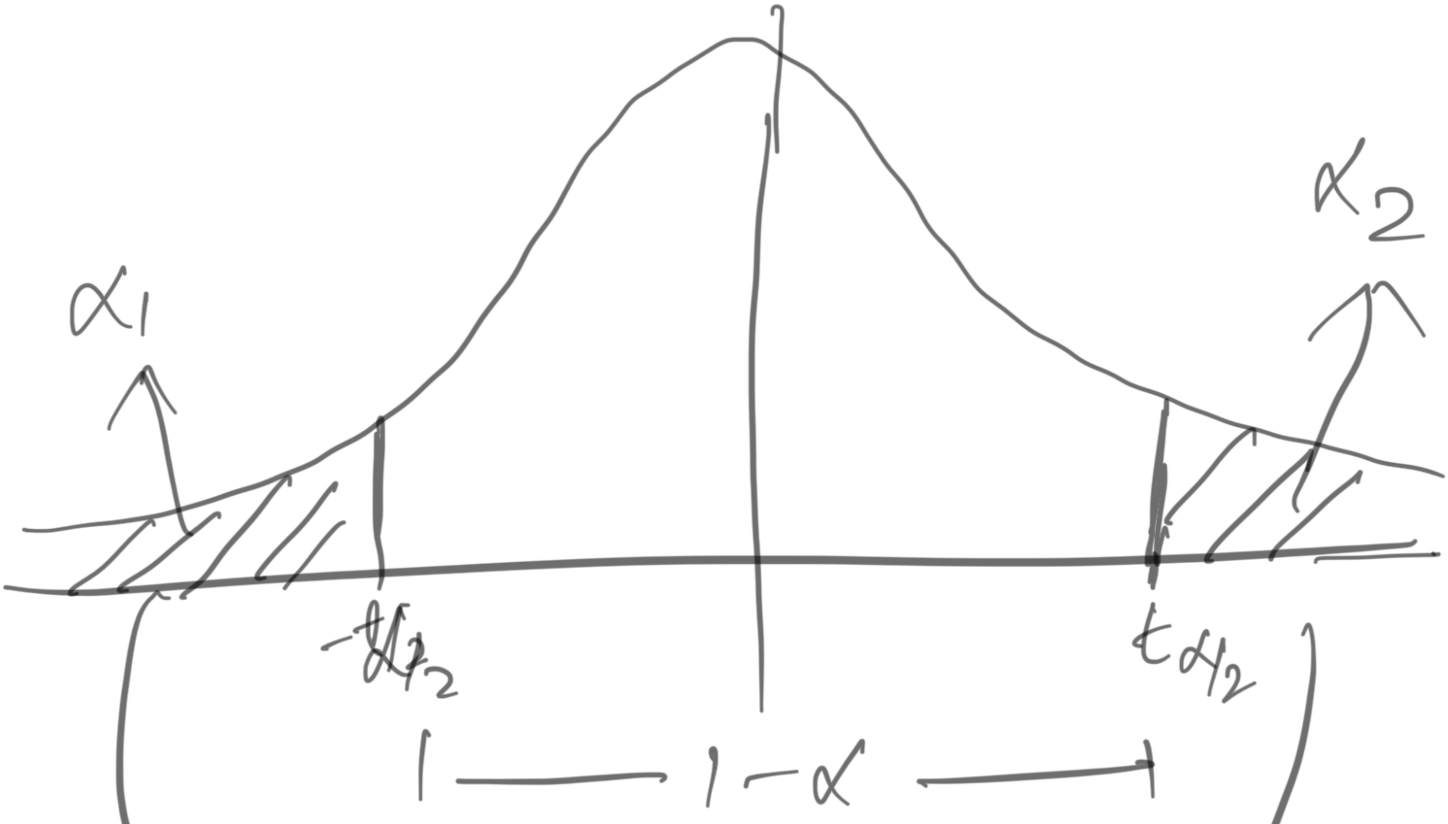
$$T_n(\theta) = \frac{\bar{X}_n - \theta}{1/\sqrt{n}}$$

Under the H_0,

$$T_n(0) = \bar{X}_n \cdot \sqrt{n} \sim N(0,1)$$

we want

$$\mathbb{P}(\text{type-I error}) = \alpha$$



$$\alpha_1 + \alpha_2 = \alpha \in (0,1)$$

$\alpha_1 + \alpha_2 = \alpha$

if $\alpha_1 = \alpha_2 = \alpha/2$

$$\Rightarrow \mathbb{P}\left(|T_n^{(0)}| > t_{\alpha/2} \right) = \alpha$$

That is, if we take the decision rule

① we reject H_0 if

$$|T_n(0)| > t_{\alpha/2}$$

② we donot reject H_0

if $$|T_n(0)| < t_{\alpha/2}$$

$$\Rightarrow -t_{\alpha/2} < T_n(0) < t_{\alpha/2}$$