Lecture 9 :

(Statistical) Hypothesis Testing:

Formal Definition: Hypothesis testing is a formal means of distinguishing between probability distributions on the basis of random variables generated from one of the distributions.

H_0 : → $dist_0$ ⟷ data $X_1, \ldots, X_n \sim dist_0$

H_a : → $dist_a$ ⟷ data $X_1, \ldots, X_n \sim dist_a$

The null hypothesis (H_0) is the claim that is initially assumed to be true.

→ our prior belief

The alternative hypothesis (H_a) is the claim that is contradictory to the null hypothesis H_0.

- The null hypothesis will be rejected in favor of the alternative hyp. only if the sample data/evidence strongly suggests that H_0 is false.

$H_0 \rightarrow$ prior belief

$H_a \rightarrow$ we reject H_0 in favor of H_a only if data suggests it strongly.

The two possible conclusions of a hypothesis testing analysis are "reject H_0" and "fail to reject H_0".

The simplest form of hypothesis testing:

- $X_1, \dots, X_n \sim f(x \mid \underline{\theta})$

  $\theta \in \mathbb{R}$

- θ is unknown but our prior belief is that the value of θ is θ_0.

- Hypothesis tesing problem:

prior believe (a given number)

example 1: $\begin{cases} H_0: \ \theta = \theta_0 \\ H_a: \ \theta > \theta_0 \end{cases}$

$f(x \mid \underline{\theta})$

example 4:

$\begin{cases} H_0: \ \theta = \theta_0 \\ H_a: \ \theta \in (\theta_0, \theta_0+5] \end{cases}$

example 2: $\begin{cases} H_0: \ \theta = \theta_0 \\ H_a: \ \theta < \theta_0 \end{cases}$

example 3: $\begin{cases} H_0: \ \theta = \theta_0 \\ H_a: \ \theta \neq \theta_0 \end{cases}$

The Neyman – Pearson Paradigm:

Recall that we need to analyse the hypotheses using sample data $X_1, \ldots, X_n$.
A decision to whether/not reject H_0 in favor of H_a is made based on a "statistic", $T(X_1, \ldots, X_n)$, which is a function of sample data and will be designed.

- The set of values of T for which H_0 is "accepted" is called the acceptance region, and the ~~the~~ set of values of T for which H_0 is rejected is called the "rejection region".

$H_0 \quad \theta = \theta_0 \longrightarrow X_1, \ldots, X_n$

$H_a \quad \theta \neq \theta_0 \longrightarrow X_1, \ldots, X_n$

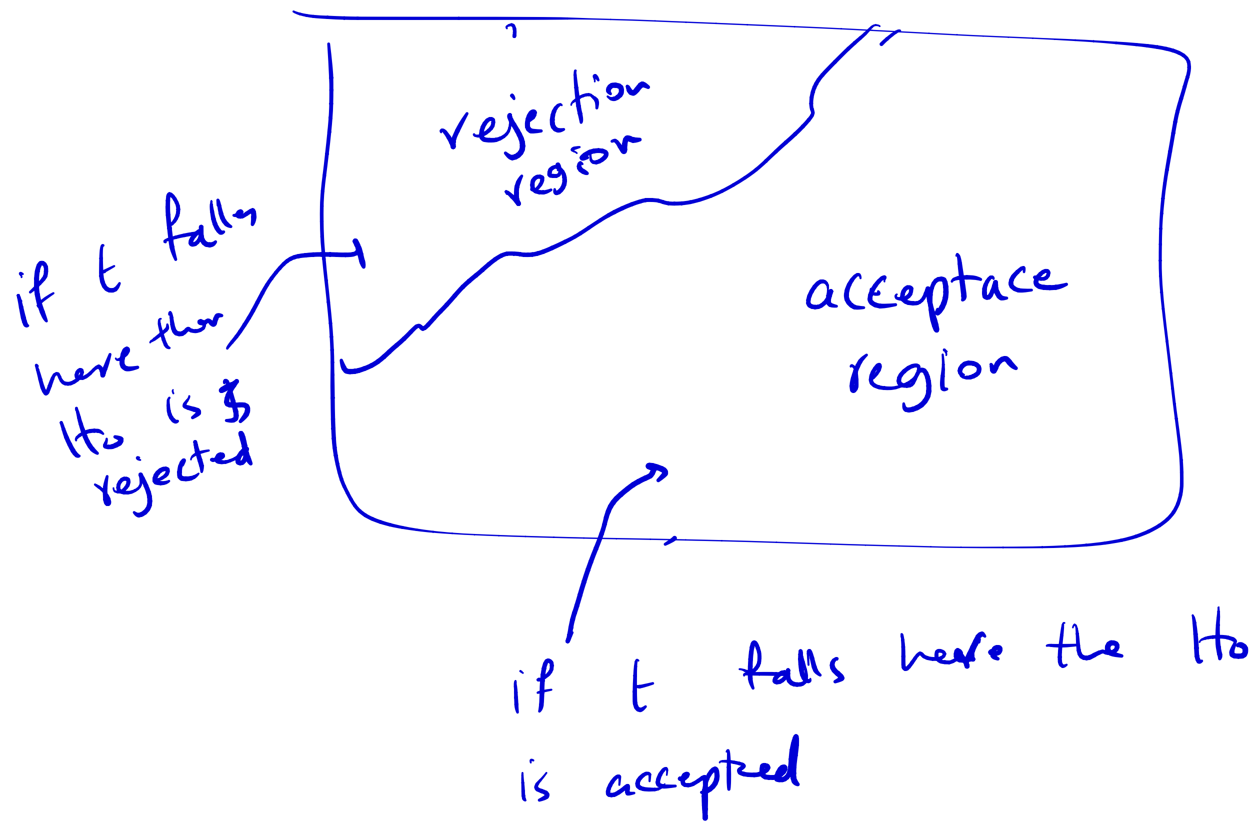
$\downarrow$

design a statistic

$\underbrace{T(X_1, \ldots, X_n)}$

let's denote the outcome of T by t:

$t \triangleq T(X_1, \ldots, X_n)$



- we should think about the statistic T as a function that takes as input the observations $X_1, \dots, X_n$ and extracts the necessary information to distinguish between H_0 and H_a.

Errors in Hypothesis Testing:

H_0

H_a

- A type I error consists of rejecting the null hypothesis H_0 when it is true.

- A type II error occurs when we accept the null hypothesis and it is false.

$$\alpha = \Pr\{\text{type I error}\}$$

$$= \Pr\{\text{reject } H_0 \mid H_0 \text{ is true}\}$$

$\beta = \Pr\{\text{type II error}\}$

$= \Pr\{\text{accept } H_0 \mid H_0 \text{ is false})$

- Type I error is typically more important that the type II error, because H_0 is formed according prior beliefs / domain knowledge / etc and rejection H_0 is equivalent to contradicting prior belief / domain knowledge / etc.

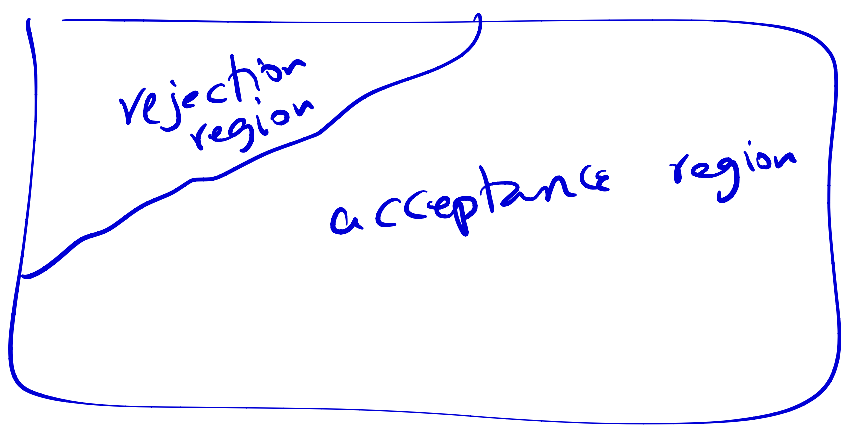
Ideally, we'd like to make both α and β as small as possible.

But, is it possible to make both α and β arbitrarily small (even if we have very large n)?

$$\alpha \downarrow 0 \quad (n \to \infty)$$

$$\beta \downarrow 0 \quad (n \to \infty)$$

$$X_1, \dots, X_n \longrightarrow T(X_1, \dots, X_n)$$



$$\left\{\begin{array}{l} \underline{\alpha} = \Pr\{T(X_1, \dots, X_n) \in \underline{\text{rejection region}} \mid H_0 \text{ is true}\} \\ \\ \beta = \Pr\{T(X_1, \dots, X_n) \in \text{acceptance region} \mid H_0 \text{ is false}\} \end{array}\right.$$

$$\alpha \downarrow \Longrightarrow \text{rejection region} \downarrow \Longrightarrow \underbrace{\text{acceptance region} \uparrow}_{\beta \uparrow}$$

$\beta \downarrow \Longrightarrow$ acceptance region $\downarrow \Rightarrow \underbrace{\text{rejection region} \uparrow}_{\alpha \uparrow}$

So, there is always a "conflict" or "trade-off" between α and β.

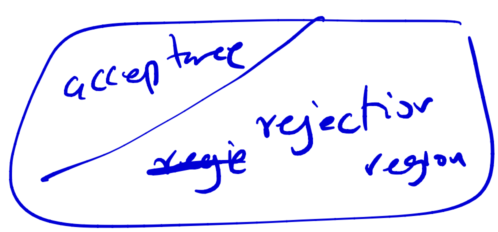
---

principles that we've learnt so far:

- decreasing α small is more important than decreasing β.

- $\alpha \downarrow \longleftrightarrow \beta \uparrow$ and $\beta \downarrow \longleftrightarrow \alpha \uparrow$

- our goal is to design T and



- In hypothesis testing, we typically make sure that α is smaller than a so-called "tolerance level" (e.g. $\alpha \leq 0.05$) and, within this constraint, we will try to minimize β (by using better choice of T).

Significance level: The largest value of α that can be tolerated.

H_0

H_a

$\overbrace{X_1, \ldots, X_n}^{\text{Data}} \longrightarrow$ Design T s.t.

- $\alpha \leq$ given significant level (e.g. $\alpha \leq 0.01$)
- minimize β given the constrain on α.

As statisticians we would like to design a statistic T with acceptance / rejection regions ~~ts~~ such that:

(1) $\alpha \leq$ significance level

(2) β is the smallest value given (1).

Tests about the Population Mean:

Data: $\underline{X_1, \dots, X_n} \sim N(\underline{\mu}, \underline{\sigma}^2)$

We assume that <u>we know</u> the value of σ, but we are not completely sure about the value of μ:

two hypotheses about μ:
$$\begin{cases} \underline{H}_0 : & \mu = \mu_0 \\ \underline{H}_a : & \mu > \mu_0 \end{cases}$$

we are also given a significance level α_0 (e.g. $\alpha_0 = 0.05$).

$$X_1, \ldots, X_n \sim N(\underline{\mu}, \overset{\text{known}}{\widetilde{\sigma^2}})$$

Step: Design T:

proposed statistic $T(X_1, \ldots, X_n) = \dfrac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$

Question if H_0 is the true hypothesis, what is the distribution of $T(X_1, \ldots, X_n)$.

$$T(X_1, \ldots, X_n) \sim N(0,1)$$

answer: $X_1, \ldots, X_n \sim N(\mu, \sigma^2)$

$$\bar{X} \sim N(\mu, \sigma^2/n) \Rightarrow \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

$\Downarrow$

if H_0 is true

then $\mu = \mu_0$

and hence

$$\frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}} \sim N(0,1)$$