CV 510^{0}_{1}
Modeling, Uncertainty, and Data for Engineers

(July – Nov 2025)

Dr. Prakash S Badal

1

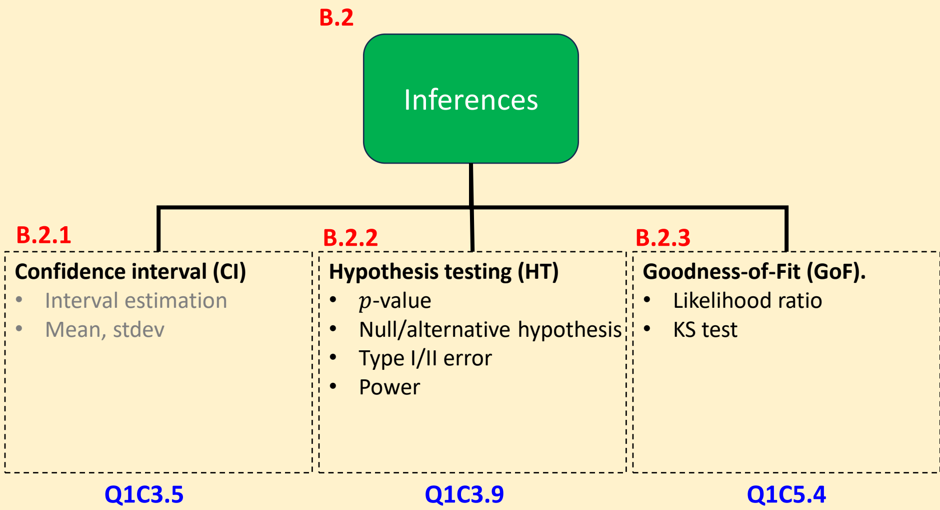
Parameter estimation

- Next few class:
  - Estimation: MoM, MLE
  - Inference: Confidence interval (CI)
  - Inference: Hypothesis testing (HT)
  - Goodness of fit (GoF): χ^2, KS

2

2

Module Overview



3

Flow

- CLT recap: demo
- Null/alternative hypothesis
- Type I/II error
- Power of test
- p-value

4

4

Central limit theorem: recap

- For large sample size n, sample mean approaches

a normal distribution,

with mean μ (same as population mean), and

variance σ^2/n (less than population Var.).

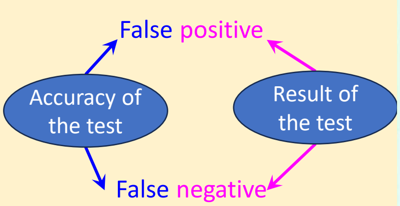
Spreadsheet demo!

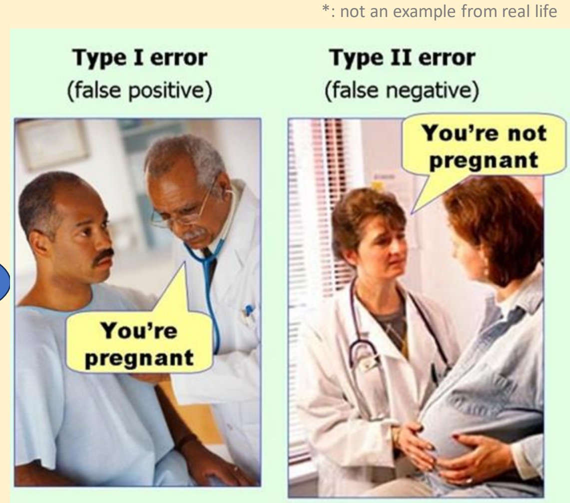
5

5

Spot the error*

*: not an example from real life





People are generally not pregnant!

Default	You are not pregnant	H_0
Alternative	You are pregnant	H_1

6

6

Justice system in a democracy

- How can someone be arrested if they really are presumed innocent?
- Why is a defendant pronounced "Not Guilty" instead of innocent?
- Why do citizens put up with a system that allows criminals to go free on technicalities?

7

7

Common challenges in justice, research (life?)

- A strong urge to believe in unusual event/discovery
  - A charged person (aka "defendant") *is* the convict
  - This new medicine *cures* Alzheimer's/cancer
  - New concrete mixing yields *higher strength/durability*
- Huge repercussions of wrong decision

You acted on your belief, BUT The belief turned out to be FALSE	You did NOT act on your belief, and the belief turned out to be TRUE
Innocent being punished	guilty going free
Patients getting the wrong medicine (side effect/death?)	Patients suffering/dying, despite a discovered medicine
Buildings constructed using inferior material	Wasted resources, unsustainable

- No sure-shot way to prove the unusual event
- Limited data

8

8

Hypothesis testing is the solution!

9

Hypothesis testing (& justice system)

No numerical values in courts, but they share four common features:

1. **The alternative hypothesis:** This is why a *criminal is arrested*.
   - The police, of course, do not think that the criminal is innocent.
   - The researchers think that their treatment is effective. H_1 or H_A.
2. **The null hypothesis:** The *presumption of innocence*.
   - The suspect or treatment didn't do anything. H_0 is the logical opposite of H_1.
3. **A standard of justice:** *A reasonable doubt*. A test score!
   - No possibility of absolute proof. So, a standard has to be set.
   - Reject the null hypothesis beyond a reasonable doubt.
4. **A data sample:** Evaluation of *partial information.*
   - Eye-witnesses/fingerprints/DNA analysis/experimental/numerical data of treatment.
   - Getting the "whole truth and nothing but the truth" is often impossible.

Hypothesis testing (& justice system)

Both statistical testing and the justice system:

1. Concentrate on rejecting the null hypothesis
   - It's much easier
   - Rejection of presumption of innocence ≡ defendant is pronounced guilty
2. Consider a failure to reject the null hypothesis
   - As "Not guilty" verdict.
   - "medicine does not treat cancer/concrete is not stronger".
   - Proving H_0 (the null hypothesis of innocence) will take endless evidence.

11

Hypothesis testing (& justice system)

Neither statistical testing and the justice system are perfect:

- Sometimes, the jury makes an error.
  - An innocent person goes to jail
  - Statisticians call it a Type I error
- Sometimes, a guilty person is set free
  - Statisticians call it a Type II error

Which one is worse?

- Citizens find Type II error disturbing but not as horrifying as Type I errors.

$$\text{Type I error} \equiv \left\{ \begin{array}{c} \text{an innocent person goes to jail} \\ + \\ \text{a guilty person walks free} \end{array} \right\}$$

- In a sense, a Type I error is twice as bad as a Type II error

12

Errors in Hypothesis testing (& justice system)

- An innocent person goes to jail (Type I error)
- A guilty person is set free (Type II error)

$$\text{Type I error} \equiv \begin{Bmatrix} \text{an innocent person goes to jail} \\ + \\ \text{a guilty person walks free} \end{Bmatrix}$$

HT (& justice system) puts a lot of emphasis on avoiding Type I error.

Errors in Hypothesis testing (& justice system)

HT & justice lady put a lot of emphasis on avoiding Type I error.

Product example:

- Null hypothesis, H_0: a product satisfies customer requirements.
- If H_0 is rejected, do not sell the product to customers.

Type I error: Rejecting a good batch by mistake. A very expensive error.
Type II error: Failing to reject a bad batch and shipping to customers (losing customer and tarnishing company's image)

Which one is worse? Type II error

Example 1

- We are interested in the yield strength of steel bars
  - Yield strength is a RV that can be described by a PDF
  - Suppose we want to focus on the mean yield strength (a parameter of this PDF)
- Specifically, we are interested in deciding whether or not the mean yield strength is 500 MPa
- We may express this formally as

$$\boxed{\begin{aligned} &H_0: \mu = 500 \text{ MPa} \\ &H_1: \mu \neq 500 \text{ MPa} \end{aligned}}$$

15

Type I and Type II Errors

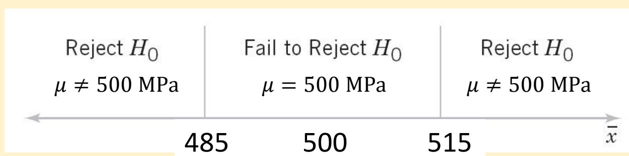
- Type I error
  - Rejecting the null hypothesis H_0 when it is true is defined as a type I error
    - It may happen that, even though the true mean strength is 500 MPa, we could select a RS that gives us a sample mean $\bar{x}$ that falls into the critical region
- Type II error
  - Failing to reject the null hypothesis H_0 when it is false is defined as a type II error
    - It may happen that, even though the true mean strength is not 500 MPa, we could select a RS that gives us a sample mean $\bar{x}$ that falls into the acceptance region

16

Test of a Hypothesis

- Consider the steel strength problem, where we wish to test
  - $H_0: \mu = 500$ MPa and $H_1: \mu \neq 500$ MPa
- Suppose that a sample of $n = 10$ specimens is tested and that the sample mean strength $\bar{x}$ is observed
- We know that $\bar{x}$ is an estimate of μ

- If $\bar{x}$ is close to the hypothesized value μ, this does not conflict with H_0
- If $\bar{x}$ is considerably different from μ, then it is evidence in support of H_1



17

Decision in HT

- four different situations determine whether the final decision is correct or in error

Decision	H_0 Is True	H_0 Is False
Fail to reject H_0		
Reject H_0		

- Associate probabilities with Type I and Type II errors
  - The probability of making a type I error is denoted by α
  - The probability of making a type II error is denoted by β

18

Probability of Type I Error, α

- $\alpha = \Pr(\text{Type I error})$

$$= P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$$

Sometimes the type I error probability is called the significance level, or the α-error

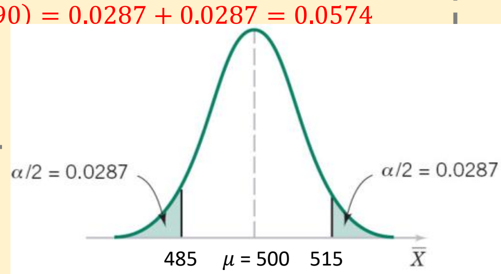
19

Probability of Type I Error, α

- In the steel strength example, a type I error will occur when either $\bar{x} > 515$ or $\bar{x} < 485$ for $\mu = 500$ MPa.
- Suppose the std of the steel strength is $\sigma = 25$ Mpa
- Find the probability of Type I error?
  - CLT ⇒ $\bar{x}$ is approx. normal with

    $\mu = 500$ MPa and std = $\sigma/\sqrt{n} = 7.9$ MPa
  - $\alpha = P(\bar{X} < 485 \text{ when } \mu = 500) + P(\bar{X} > 515 \text{ when } \mu = 500)$
  - z-values for the critical values of 485 and 515 are -1.90 and 1.90, respectively.

    $\alpha = P(Z < -1.90) + P(Z > 1.90) = 0.0287 + 0.0287 = 0.0574$



20

Probability of Type I Error, α

- $\alpha = \Pr(\text{Type I error})$
    - We can reduce α by widening the acceptance region
    - We could reduce α by increasing the sample size n

21

Probability of Type II Error, β

- $\beta = \Pr(\text{Type II error})$

$$= \Pr(\text{fail to reject } H_0 \text{ when } H_0 \text{ is false})$$

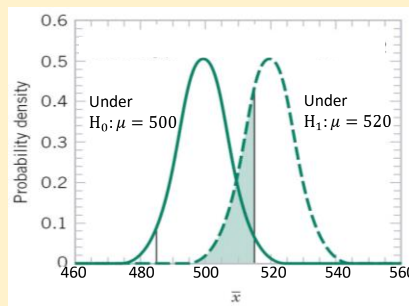
- β-error
- To calculate β we must have a specific alternative hypothesis, that is, we must have a particular value for μ
    - For example:
        - We can choose some critical value, which we might select to be, e.g., 520 MPa, so that $H_1{:}\, \mu = 520 \text{ MPa}$

22

Probability of Type II Error, β

- $\beta = \Pr(\text{Type II error})$
  - A type II error will be committed if the sample mean falls between 485 and 515 (the critical region boundaries) when $\mu = 520$

$$\beta = P(485 \le \bar{x} \le 515 \text{ when } \mu = 520)$$



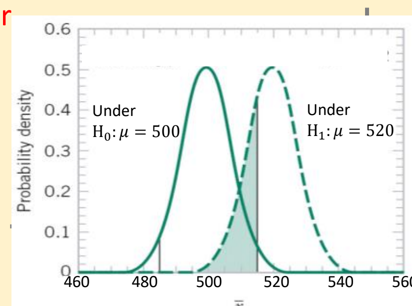
23

Probability of Type II Error, β

- The z-values that correspond to the critical values 485 and 515 are -4.43 and -0.63, respectively. Therefore

$$\beta = P(-4.43 \le Z \le -0.63) = P(Z < -0.63) - P(Z \le -4.43)$$
$$= 0.2643 - 0.0000 = 0.2643$$

  - If we are testing $H_0: \mu = 500$ MPa against $H_1: \mu \neq 500$ MPa with $n = 10$, and the true value of the mean is $\mu = 520$ MPa ,
  - the probability that we will fail to reject the false null hypothesis is 0.2643
- We can always reduce the type II error by increasing the sample size n



24

Probability of Type II Error, β

1. The size of the critical region, and consequently, the probability of a type I error, i.e., α can be reduced by appropriate selection of the critical values
2. Type I and type II errors are related:
   - A decrease in the probability of one type of error always results in an increase in the probability of the other, provided that the sample size n does not change
3. An increase in the sample size reduces β, provided that α is held constant
4. When the null hypothesis is false:
   - β increases as the true value of the parameter approaches the value hypothesized in the null hypothesis
   - The value of β decreases as the difference between the true mean and the hypothesized value increases

25

Questions, comments, or concerns?

26

26