MM 225 – AI and Data Science

Day 20: Hypothesis Testing 2

Instructors: Hina Gokhale, MP Gururajan, N. Vishwanathan

12 SEPTEMBER 2024

Probability of Type I error – P value

Recall the Classical approach

Classical approach is to

- fix α at the minimum possible level and,
- Probability of Type I error of the test should be $\leq \alpha$

The approach demonstrated equates the probability of Type I error with α and computes the rejection region

This approach is called *"fixed level of significance approach".*

It has its advantages as it directly leads to the concept of Type –II error.

P – Values in Hypothesis Testing

It has certain shortcomings:

- The decision taken does not give any information as to the test statistic was barely in the rejection region or was well within the critical region.
- This way the predefined level of significance is imposed on the users, who may not be comfortable with it.

To avoid these situation, P-value approach is adopted.

P-value is the probability that the test statistic will take on a value that is as extreme as the observed value the test statistic will take when the null hypothesis is true.

Case of N$\left(\mu, \sigma^2\right)$, when σ^2 is known

The critical region C in this case is

$$C = \{X_1, X_2, \dots, X_n \mid |\bar{X} - \mu_0| > c\}$$

When H_0 is true: $Z = \frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$

Therefore, P[type – I error] = P $\left\{\left|\frac{\bar{X}-\mu_0}{\sigma/\sqrt{n}}\right| > c'\right\} = P[|Z| > c']$

Let $z_0 = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}}$ is an observed value of test statistic Z

Hence, P-value = $P[|Z| > z_0] = 1 - P[-z_0 \le Z \le z_0]$

Example:

An aerospace industry is interested in buying certain super alloy rods from a foundry. The industry has been told that the super alloy would have yield strength of 1110 MPa with standard deviation 110 MPa. The industry takes a random sample of size 81 from the supplied lot and finds that the average yield strength to be 1129 MPa. Should the company accept the supply?

μ_0 = 1110 MPa, σ = 110 MPa, n= 81 and $\bar{x}$ = 1129 MPa

$H_0 : \mu = 1110$ vs $H_A : \mu \neq 1110$

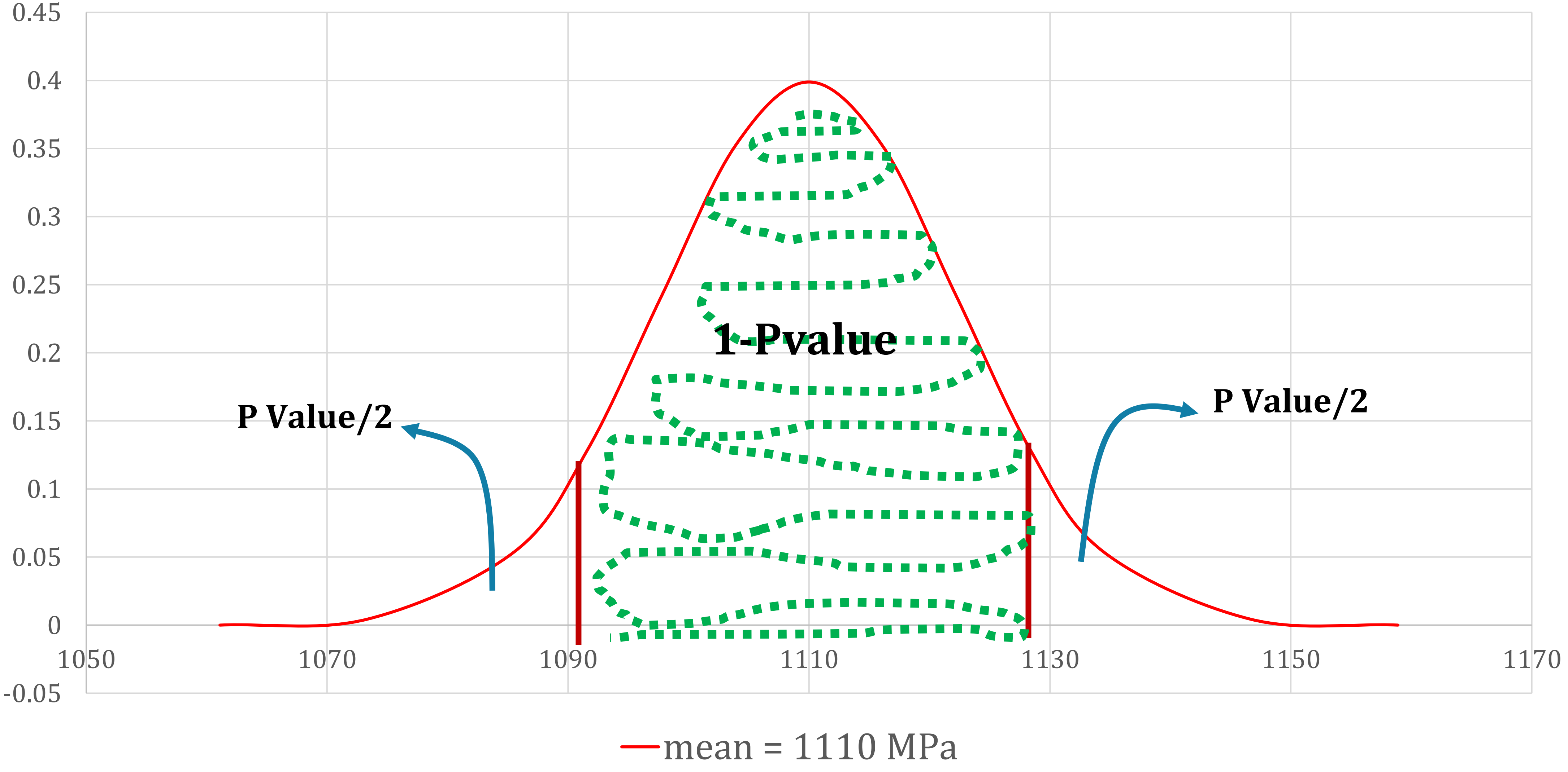
$$z_0 = \frac{1129 - 1110}{110/\sqrt{81}} = 1.56$$

P-value = 1- P[-1.56 ≤ Z ≤1.56] = 0.12

P-value > 0.05 the prefixed level of significance , the null hypothesis cannot be rejected.

In fact it can only be rejected at any level of significance greater than or equal to 0.12.

Acceptance Region when H_0 is true



Case of $\mathbf{N}(\mu, \sigma^2)$, when σ^2 is unknown

Critical region in this case is C = $\{X_1, X_2, \ldots, X_n \mid \left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| > c\}$

when H_0 is true T = $\frac{\bar{X}-\mu_0}{S/\sqrt{n}} \sim t_{n-1}$

Therefore, P[type – I error] = P $\left\{\left|\frac{\bar{X}-\mu_0}{S/\sqrt{n}}\right| > c'\right\} = P[|T| > c']$

$w_0 = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$ is the observed value of the test statistic T

As per definition of P-value = $P[|T| > w_0] = 1 - P[-w_0 \leq T \leq w_0]$

Example

An aerospace industry is interested in buying certain super alloy rods from a foundry. The industry has been told that the super alloy would have yield strength of 1110 MPa. The industry takes a random sample of size 81 from the supplied lot and finds that the average yield strength to be 1129 MPa with standard deviation of 112 MPa. Should the company accept the supply?

μ_0 = 1110 MPa, n= 81, $\bar{x}$ = 1129 MPa and s = 112 MPa

$H_0 : \mu = 1110$ vs $H_A : \mu \neq 1110$

$w_0 = 1.53 = \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$

P-value = 1 – P[-1.53 ≤ T ≤ 1.53] = 0.14 and T ~ t(80) when H_0 is true.

P-value > 0.05 the prefixed level of significance , hence the null hypothesis cannot be rejected.

In fact it can only be rejected at any level of significance greater than or equal to 0.14.

Type II Error

Type II Error = Accepting H_0, when it is not true

P[Type II error] = β = P[Accepting $H_0|H_A$ is true]

Consider the case of $H_0 : \mu = \mu_0$ vs. $H_A : \mu \neq \mu_0$ for N(μ, σ^2) population with σ^2 known

$\beta = P[Acceptance\ of\ H_0, when\ population\ mean = \mu]$

This shows that β is a function of μ that is β(μ)

Example

Now, in the example with σ^2 known :

- $H_0 : \mu = 1110 \, MPa$ vs. $H_A : \mu \neq 1110 \, MPa$ and σ^2 = 110 MPa, n=81
- Let us take specific case of . $H_A : \mu = 1120 \, MPa$

$$\beta = P\left[-1.56 \leq \frac{\bar{X} - 1110}{110/\sqrt{81}} \leq 1.56 \mid \mu = 1120\right]$$

$$= P[1091 \leq \bar{X} \leq 1129 | \, \mu = 1120]$$

$$= P\left[\frac{1090.96 - 1120}{12.22} \leq \frac{\bar{X} - 1120}{12.22} \leq \frac{1129.04 - 1120}{12.22}\right]$$

$$= P(-2.37 \leq Z \leq 0.74)$$

$$= P(Z < 0.74) - P(Z < -2.37) = 0.76$$

Problem 5 : Testing of Binomial parameter p

A poll has been conducted covering 150 students to determine if more than two third of the student population is willing to postpone the examination. 105 students agreed for postponement. Can this be taken as evidence to postpone the examination?

Solution

X = number of students support the postponement

$X \sim Bin(n, p)$, n = 150 and p unknown

$H_0: p \leq \frac{2}{3}$ $vs. H_A: p > \frac{2}{3}$

this can be translated in terms of RV X as follows

Since E(X) = np when H_0 is true $X \leq np = 100$ we get

$H_0: X \leq 100$ $vs. H_A: X > 100$

Test statistic is X as X/150 is an unbiased estimator of p.

Hence, when H_0 is true we have

$P\{X \sim Bin(150, 0.667) \geq x_q\} = q - value,$

$$where\ x_q is\ a\ sample\ value\ of\ X = 105$$

$$q\ value = \sum_{x=105}^{150} \binom{150}{x} \left(\frac{2}{3}\right)^{x} \left(\frac{1}{3}\right)^{150-x} = 0.22$$

q-value is large, hence it can be said that there is not enough evidence to reject the hypothesis.

Note: Excel calculates

$$1 - q = \sum_{x=0}^{104} \binom{150}{x} \left(\frac{2}{3}\right)^{x} \left(\frac{1}{3}\right)^{150-x} = 0.78$$

Another way

Note that when n is large binomial distribution can be approximated by normal distribution as follows:

$X \sim Bin(n, p)$, n = 150 and p unknown

Then for large n applying CLT we have Z $= \frac{X-np}{\sqrt{np(1-p)}} \sim N(0,1)$

q-value can be obtained by

$q = P\{Z > \ z_q\}, where\ z_q\ is\ a\ sample\ value\ of\ Z$

Thus $z_q = \frac{105-100}{\sqrt{150(0.667)(0.333)}} = 0.866$

q = 0.193

Again null hypothesis cannot be rejected.

Summary

Concept of P-value introduced

Probability of Type II error (β) as function of mean of Normal Population

Testing of Hypothesis on percentage

Thank you