# Hypothesis Testing

# Hypothesis

- Null: What is believed to exist or currently accepted or default
- Alternate: Challenges the Null, also called research hypothesis
- You either reject the Null hypothesis or you fail to reject the null hypothesis

# Let us note few basics

- Null Hypothesis will always have an = or ≤ or ≥ sign
- Alternate Hypothesis never contains equality
- When H0 has = sign, H1 will have ≠
- When H0 has ≤ sign, H1 will have >
- When H0 has ≥ sign, H1 will have <

# Let us note few basics

- Null Hypothesis will always have an = or ≤ or ≥ sign
- When H1 has ≠ sign, it is a two tailed test
- When H1 has < it is left tailed test
- When H1 has > it is right tailed test

# When we conduct a Hypothesis test

- We want to decide whether or not to reject the null hypothesis
- When we reject null, we say we have enough evidence to support the alternative
- When we fail to reject null, we say we do not have enough evidence to support the alternative
- To determine when to reject the null, we select a significance level $\alpha$
- 5% or 0.05 is the most commonly used $\alpha$ value
- 0.10 & 0.01 are also used

# Significance level

- The significance level $\alpha$ specifies the size of the rejection region where the null hypothesis should be rejected
- That is called the critical or rejection region
- For a two tailed test $\alpha$ is divided by two into both tails
- So we reject the null hypothesis in the tails and don't reject in the middle region
- For one tailed test $\alpha$ is not divided
- Rejection region for left tailed test lies in the left tail
- Rejection region for right tailed test lies in the right tail

# P Value

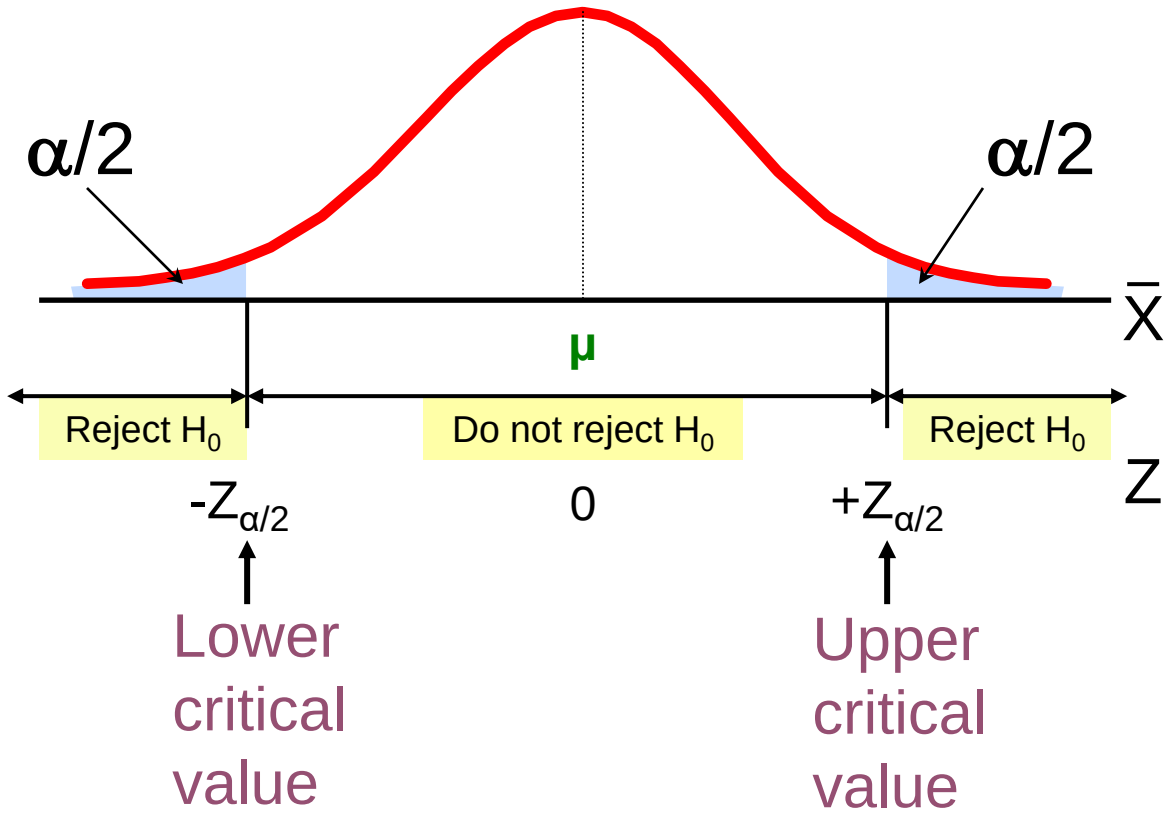
- The probability of getting an unusual result if the null were true
- Strength of the evidence provided by our sample against the null hypothesis
- Smaller the p-value, greater the evidence against the null hypothesis
- If we have a given significance level alpha then, we can reject the null hypothesis if the p-value is less than or equal to the alpha
- We can say there is significant evidence against null
- Roughly, if the p value is less than 0.01 the very strong evidence against null
- If it is between 0.01 to 0.05, then strong evidence against null
- If it is between 0.05 to 0.10, then weak evidence against null
- If it is greater than 0.1 then no evidence against null

# The test statistic

- Once the rejection region is set, we then calculate the test statistic
- If the test statistic falls within the rejection region, we reject the hypothesis otherwise we fail to reject it

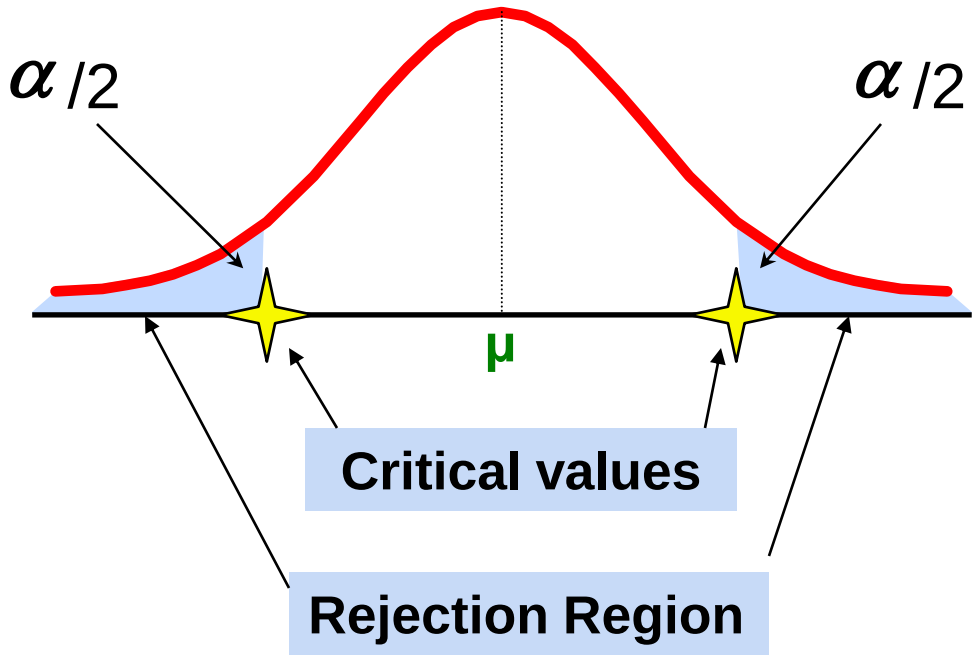
# Two-Tail Tests

- There are two cutoff values (critical values), defining the regions of rejection



# Level of Significance and the Rejection Region

Level of significance = $\alpha$




This is a two-tail test because there is a rejection region in both tails


# PG Students age as per previous data

- Historically, the mean age of a PG student has been 23 years, with a standard deviation of 2.4 and has a normal distribution.

- This year, a random sample of 42 students gave an average of 23.8
- Can we say at $\alpha = 0.05$ that the mean has changed?

# PG students age as per previous data

- $H_0$: The mean age of PG students = 23
- $H_1$: The mean age of PG students ≠ 23

Given

- This population has a standard deviation of 2.4 (The SD is known)
- The class has a normal distribution

- A random sample of 42 students gave an average of 23.8
- Can we say at $\alpha$ = 0.05 that the mean has changed?

# Deciding the test statistic to be used

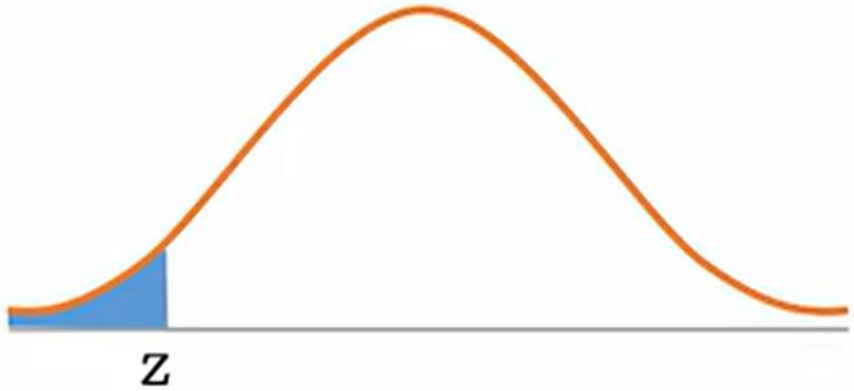
- N = 42
- $\mu = 23$
- $\overline{X} = 23.8$
- $\sigma = 2.4$
- $\alpha = 0.05$

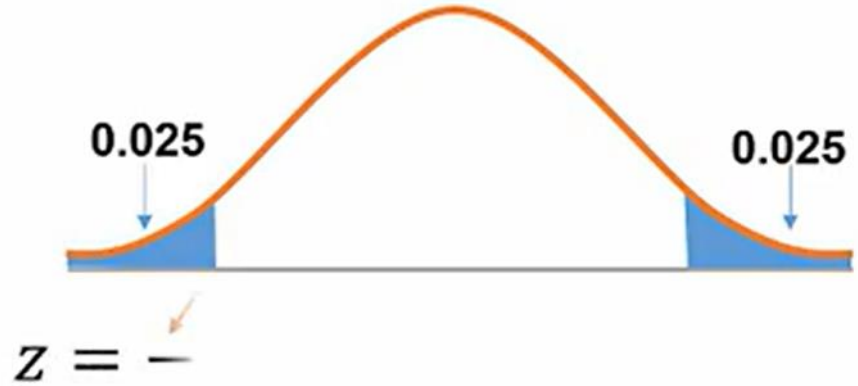
We use Z test here, since the Standard deviation is known.

When it is not known, we use T – test.

# Conducting Z Test

$$Z_{STAT} = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$




| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| -2.4 | 0.0082 | 0.0080 | 0.0078 | 0.0075 | 0.0073 | 0.0071 | 0.0069 | 0.0068 | 0.0066 | 0.0064 |
| -2.3 | 0.0107 | 0.0104 | 0.0102 | 0.0099 | 0.0096 | 0.0094 | 0.0091 | 0.0089 | 0.0087 | 0.0084 |
| -2.2 | 0.0139 | 0.0136 | 0.0132 | 0.0129 | 0.0125 | 0.0122 | 0.0119 | 0.0116 | 0.0113 | 0.0110 |
| -2.1 | 0.0179 | 0.0174 | 0.0170 | 0.0166 | 0.0162 | 0.0158 | 0.0154 | 0.0150 | 0.0146 | 0.0143 |
| -2.0 | 0.0228 | 0.0222 | 0.0217 | 0.0212 | 0.0207 | 0.0202 | 0.0197 | 0.0192 | 0.0188 | 0.0183 |
| -1.9 | 0.0287 | 0.0281 | 0.0274 | 0.0268 | 0.0262 | 0.0256 | 0.0250 | 0.0244 | 0.0239 | 0.0233 |
| -1.8 | 0.0359 | 0.0351 | 0.0344 | 0.0336 | 0.0329 | 0.0322 | 0.0314 | 0.0307 | 0.0301 | 0.0294 |
| -1.7 | 0.0446 | 0.0436 | 0.0427 | 0.0418 | 0.0409 | 0.0401 | 0.0392 | 0.0384 | 0.0375 | 0.0367 |
| -1.6 | 0.0548 | 0.0537 | 0.0526 | 0.0516 | 0.0505 | 0.0495 | 0.0485 | 0.0475 | 0.0465 | 0.0455 |
| -1.5 | 0.0668 | 0.0655 | 0.0643 | 0.0630 | 0.0618 | 0.0606 | 0.0594 | 0.0582 | 0.0571 | 0.0559 |

# Conducting Z Test

$$Z_{STAT} = \frac{23.8-23}{\frac{2.4}{\sqrt{42}}} = 2.16$$

Since this value 2.16 > 1.96 at 0.05 level confidence, we reject the null hypothesis and support the alternate. Which means the average age is changed.

# Meaning of Hypothesis Testing

Hypothesis testing is a statistical method used to make inferences about a population based on sample data.

It involves formulating a hypothesis, collecting data, and analyzing the data to determine the validity of the hypothesis.

# Basic Concepts of Hypothesis Testing

- Null Hypothesis (H0): The hypothesis to be tested, often assuming no significant difference or relationship.
- Alternative Hypothesis (Ha or H1): The alternative to the null hypothesis, suggesting a significant difference or relationship.
- Significance Level (α): The predetermined threshold for rejecting the null hypothesis.
- Test Statistic: A numerical value calculated from the sample data to determine the likelihood of obtaining the observed result under the null hypothesis.

# Flow Diagram of Hypothesis Testing

- State the null and alternative hypotheses.
- Set the significance level (α).
- Collect and analyze the sample data.
- Calculate the test statistic.
- Compare the test statistic to the critical value or p-value.
- Draw conclusions and make inferences based on the results.

# Power of a Hypothesis Test

Power is the probability of correctly rejecting a false null hypothesis.

It measures the ability of the test to detect a true difference or relationship.

Power is influenced by the sample size, effect size, significance level, and variability of the data.

# Limitations of Tests of the Hypothesis

- Results may not be applicable to the entire population.
- Tests assume specific distributions and assumptions that may not hold in practice.
- Sample size and variability can impact the reliability of the results.
- Tests can only provide evidence against the null hypothesis but cannot prove its truth or falsehood.