# Introduction to Statistical Methods

Team ISM

BITS Pilani
Pilani|Dubai|Goa|Hyderabad

innovate achieve lead

BITS Pilani

# Session No 7

## Testing of Hypothesis

## (18th /19th June ,2022)

| Contact Session | List of Topic Title | Reference |
|---|---|---|
| CS - 7 | Testing of Hypothesis - Type I & II errors, Critical region, t – test, Chi – Square test and F – test(Introduce and discuss these tests) | T1:Chapter 7 ,8,9 & 10 |
| HW | Problems on Testing of Hypothesis | T1:Chapters 7 to 10 |

# Need for testing of hypothesis

Often the decisions are made based on samples estimates to generalize on population parameter (as described in sampling and estimation).

In this process, there may be a difference between the estimate and the parameter which needs to be examined.

The following possibilities might arise due to sampling

$$|\text{Estimate-Parameter}| = \begin{cases} 0 \\ Small \\ Large \end{cases}$$

# Need for testing of hypothesis

**Case(i):**

If the difference is zero, it is called unbiased

**Case(ii):**

If the difference is small, it may due to chance or sampling error (improper sampling technique used leads to sampling error)

**Case(iii):**

If the difference is large, it may a real one or due to sampling error

(improper sampling technique used leads to sampling error)

Hence, there is a need to test what type of difference is between estimate and parameter.

# Hypothesis

A statement which is yet to be proved/ established or a statement on the parameter(s) of the Probability distribution to be tested

**Null Hypothesis**

Hypothesis of no difference or neutral or may be due to Sampling variation

**Alternative Hypothesis**

Hypothesis of difference which is yet to be proved/ established

# Hypothesis

## Hypothesis testing (Non-statistical)


A suspected criminal is produced before jury. The Jury has to decide whether the defendant is innocent or guilty.


Jury must decide between two hypotheses

The null hypothesis → $H_0$: The defendant may be innocent

The alternative hypothesis → $H_1$: The defendant may be guilty

# Hypothesis

The jury do not know which hypothesis is true.

The jury should make a decision on the basis of evidence presented before them by the advocates .

# Hypothesis - Formulation

- A taxi company manager is trying to decide whether the use of radial tires or regular belted tires improves fuel economy.
- The variable measured is quantitative, therefore

# Hypothesis - Formulation


Regular Tyres


Radial Tyres

The variable measured fuel consumption

**Quantitative**

# Hypothesis - Formulation

$H_0$ The mean fuel consumption in cars fitted with radial tyres and regular belted tires will be same

$$H_0 : \mu_1 = \mu_2$$

**Note: $H_0$ can also be stated as one-tailed**

$H_1$ The mean fuel consumption in cars fitted with radial tyres may be inferior to regular belted tires

$$H_1 : \mu_1 < \mu_2$$

# Hypothesis - Formulation

$H_1$ → The mean fuel consumption in cars fitted with radial tires may be better than regular belted tires

$$H_1 : \mu_1 > \mu_2$$

$H_1$ → The mean fuel consumption in cars fitted with radial tires and regular belted tires may be different

$$H_1 : \mu_1 \neq \mu_2$$

# Hypothesis - Formulation

Two judges have to judge independently whether the defendant is innocent or guilty on the basis of evidence. Lack of sufficient evidence may lead to erroneous decisions like false positive or false negative. Suppose based on evidences, if we are interested in finding proportion of false positivity in the judgement, then the hypothesis to be tested, if variable measured is qualitative will be

# Hypothesis - Formulation



Judge 1

Judge 2

Suppose based on evidences, if we are interested in finding proportion of false positivity in the judgment of two Judges

Formulate the hypotheses

???

# Hypothesis - Formulation

$H_0$

The proportion of false positive judgement between Judges may be same

$$H_0 : P_1 = P_2$$

$H_1$

The proportion of false positive judgement by Judge 1 may be **lower than** proportion of false positive judgement by Judge 2

$$H_1 : P_1 < P_2$$

# Hypothesis - Formulation

$H_1$

The proportion of false positive judgement by Judge 1 may be **more than** proportion of false positive judgement by Judge 2

$$H_1 : P_1 > P_2$$

$H_1$

The proportion of false positive judgement between both Judges may be different

$$H_1 : P_1 \neq P_2$$

# Test

$$\text{Test} = \begin{cases} \mu_1 < \mu_2 \Rightarrow \text{One} - \text{tailed test} \\ \mu_1 > \mu_2 \Rightarrow \text{One} - \text{tailed test} \\ \mu_1 \neq \mu_2 \Rightarrow \text{Two} - \text{tailed test} \end{cases}$$

# Test

Test is a statistical rule which decides whether to accept the null hypothesis or not ?

**Warning**

Decision is made based on the sample not on the population

Leads to possibility of **error** between the decision made and the reality

# Types of test

A statistical rule which decides whether to accept or reject the null hypothesis on the basis of data

**Parametric tests**

Based on the assumption of some probability distribution

**Non-parametric tests**

Not based on any assumption of probability distribution

# Parametric tests

It is assumed that the data do follow some probability. distribution which is characterized by any parameters.

| | |
|---|---|
| **Large Sample Test** | n ≥ 30 |
| **Standard Normal Test** | Z-Test |
| **Small Sample Test** | n < 30 |
| **Student's t-test** | Unpaired t-Test<br>Paired t-Test |
| **Analysis of Variance** | ANOVA<br>Rm ANOVA |

# Non - Parametric tests

It is assumed that the data do not follow any probability distribution which is not characterized by any parameters.

→ **Distribution - free tests**

- **Chi - Square Test**
- **Fisher's exact probabilities**
- **Mann – Whitney U test**
- **Wilcoxon Signed Rank Test**
- **Kruskal - WallisTest**
- **Friedman'sTest**

# Steps involved in Testing of Hypothesis

State null and alternative hypotheses

Specify the level of significance ‘α’

Define the probability distribution the data follows

Compute the test statistic based defined population

Define the rejection criteria/critical regional

Conclusion

## Errors in decision making

**Any example based on data**

**Statistical Example**

| Decision | Null Hypothesis ($H_0$) | |
|---|---|---|
| | True | False |
| Accept | | |
| Reject | | |

**Any example based on data**

**Statistical Example**

| Decision | Null Hypothesis ($H_0$) | |
|---|---|---|
| | True | False |
| Accept | | Type – II Error |
| Reject | Type – I Error | |

# Errors in decision making

| Any example based on data | | |
|---|---|---|
| Statistical Example | | |
| Decision | Null Hypothesis ($H_0$) | |
| | True | False |
| Accept | | β - error |
| Reject | α- Error | |

| Any example based on data | | |
|---|---|---|
| Statistical Example | | |
| Decision | Null Hypothesis ($H_0$) | |
| | True | False |
| Accept | Confidence Level (1-α) | |
| Reject | | Power (1-β) |

# Decision on α –error and β - error

Conventionally

- Fix α at 5%
- Minimize β → Maximize 1-β

# Parametric tests

**Z-test** → This is a test based on Standard Normal Distribution

Used for testing the

- Mean of a single population ($\mu$)
- Difference between means of two populations ($\mu_1 - \mu_2$)
- Proportion of a single population (P)
- Difference between proportions of two populations ($P_1 - P_2$)

# Assumptions on Z-test

- Samples are drawn from normal distribution
- The population variances should be known
- Z-test
- Two groups should be independent
- Subjects should be allocated randomly to both groups
- **The sample size should be more than 30 (i.e., n ≥ 30)**

# Testing mean of a single population

1. State null and alternative hypothesis
2. Specify the level of significance ‘α’
3. Standard Normal Distribution
4. Compute the test statistic
5. Define the critical region/ rejection criteria
6. Conclusion

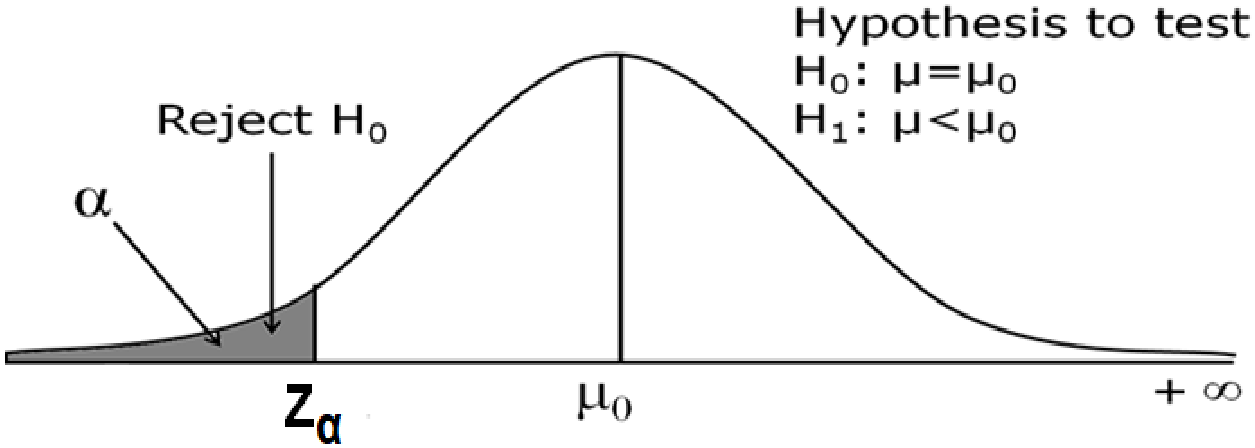
$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu < \mu_0$$
$$\text{or } H_1 : \mu > \mu_0$$
$$\text{or } H_1 : \mu \neq \mu_0$$

$$Z = \frac{\bar{x} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \cong N(0,1)$$

# Rejection criteria

**5** Define the critical region/ rejection criteria

**(i)** Reject $H_0$ if computed value of Z is less than the critical value, ie., $P(Z < -z_\alpha)$, otherwise do not reject $H_0$

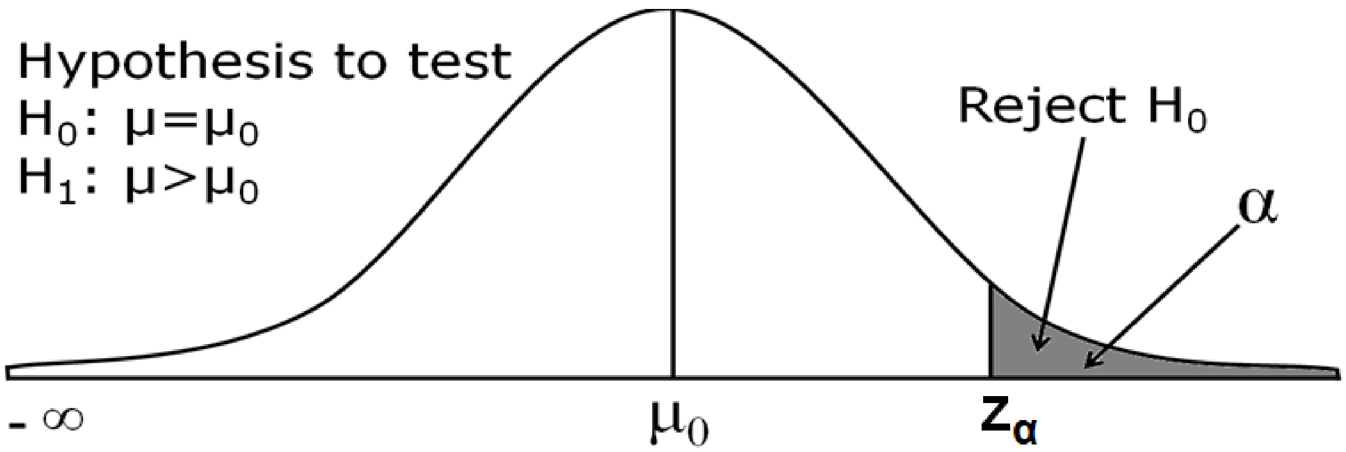



**6** Conclusion

# Rejection criteria

5 Define the critical region/ rejection criteria

(ii) Reject $H_0$ if computed value of Z is greater than the critical value, ie., $P(Z > z_\alpha)$, otherwise do not reject $H_0$

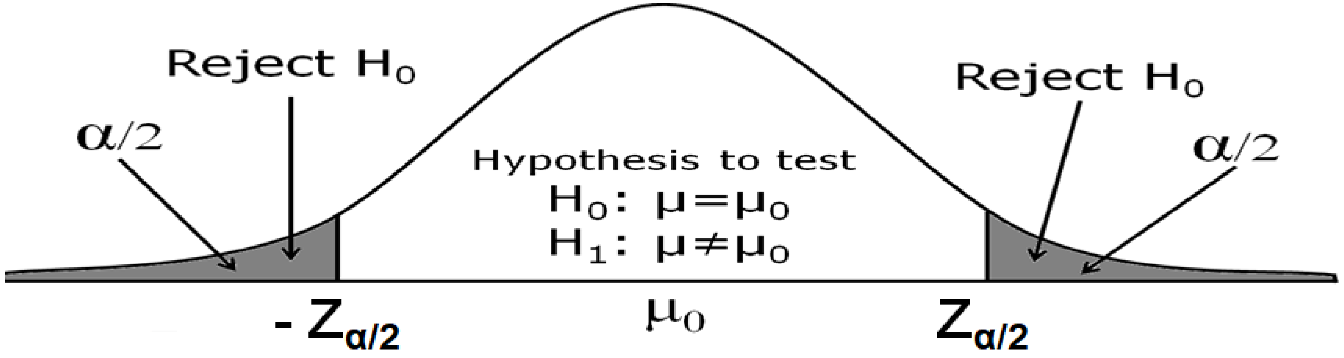



6 Conclusion

# Rejection criteria

**5** Define the critical region/ rejection criteria

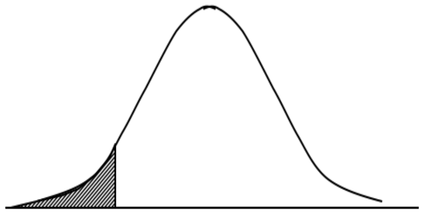
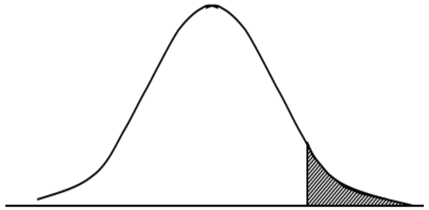
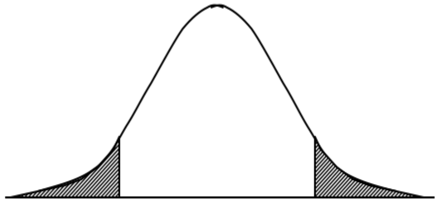
**(iii)** Reject $H_0$ if computed value of Z is less than or greater than the critical value, ie., $P(Z < -z_{\alpha/2})$ or $P(Z > z_{\alpha/2})$, otherwise do not reject $H_0$




**6** Conclusion

# Rejection criteria

**Summary of One- and Two-Tail Tests**

| One-Tail Test (left tail) | One-Tail Test (right tail) | Two-Tail Test (Either left or right tail) |
|---|---|---|
| $H_0: \mu = \mu_0$ vs $H_1: \mu < \mu_0$ | $H_0: \mu = \mu_0$ vs $H_1: \mu > \mu_0$ | $H_0: \mu = \mu_0$ vs $H_1: \mu \neq \mu_0$ |
|  |  |  |

## P - value

In hypothesis testing, the choice of the value of α is somewhat arbitrary. For the same data, if the test is based on two different values of α, the conclusion could be different. Many Statisticians prefer to compute the so called P-value, which is calculated based on the observed test statistic. For computing the P-value, it is not necessary to specify a value of α. We can use the given value data to obtain the P-value.

**P – value:** The strength of the evidence against the null hypothesis that the true difference in the population is zero

In other words

Corresponding to an observed value of a test statistic, the P-value (or attained level of significance) is the lowest level of significance at which the null hypothesis would have been rejected.

# P-value

## P - value

01 Possibility that the observed differences were a chance event

02 Entire population need to be studied to know that a difference is really present with certainty

03 Research community and statisticians had to pick a level of uncertainty at which they could live

# P-value

If the P-value is less than 1% (< 0.01),

**Overwhelming evidence** that supports the alternative hypothesis

If the P-value is between 5% and 10%,

**Weak evidence** that supports the alternative hypothesis

If the P-value is between 1% and 5%,

**Strong evidence** that supports the alternative hypothesis

If the P-value exceeds 10%,

**No evidence** that supports the alternative hypothesis.

# Testing of Hypothesis → Student's t-test

Unpaired t-test **OR** Independent sample t-test **AND** Paired t-test

## Independent Sample t-test (Unpaired t-test)

- Testing mean of a single population
- Testing difference between means of two populations

innovate achieve lead

t-test → Testing mean of single population (μ)

- Samples are drawn from normal population
- The population variance should be unknown
- The sample size should be less than 30 (i.e., $n < 30$)

t-test

- Sample should be allocated randomly
- However even if sample size more than 30 (i.e., $n > 30$) and population variance unknown, t-test should be continue to apply, because of central limit theorem it approaches normal.

innovate achieve lead

t-test

## Degrees of freedom (df): No. of independent observations

### Suppose

a+b = 20. If we assign a=9 then b=11 or vice-versa. ∴ df=(2-1)=1

a+b+c = 20. If we assign a=9 and b=6 then c=5. ∴ df=(3-1)=2

**In general, if there are n observations df = n-1**

# Mean of a single population using t-test

t-test → Testing mean of a single population ($\mu$)

**Assumptions**

- Assume that the samples are drawn from normal distribution
- The population variance may be unknown
- The sample size should be less than 30 ($n < 30$)
- Subjects should be selected randomly

1 State null and alternative hypothesis

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu < \mu_0$$
$$\text{or } H_1 : \mu > \mu_0$$
$$\text{or } H_1 : \mu \neq \mu_0$$

2 Specify the level of significance 'α'

3 Student's t-distribution

4 Compute the test statistic

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} \cong t_{(\alpha, n-1)}$$

5 Define the critical region/ rejection criteria

6 Conclusion

Note: Rejection criteria may be based on critical value or P-value

## 5 Define the critical region/ rejection criteria

**(i)** Reject $H_0$, if computed value of t is less than the critical value, ie., $P(t < -t_\alpha)$, otherwise do not reject $H_0$

**(ii)** Reject $H_0$, if computed value of t is greater than the critical value, ie., $P(t > t_\alpha)$, otherwise do not reject $H_0$

By combining both (i) and (ii), Reject $H_0$, if computed value of $|t|$ is greater than the critical value, ie., $P(|t| > t_\alpha)$, otherwise do not reject $H_0$. Besides α, the df is also important.

## Conclusion

## 5 Define the critical region/ rejection criteria

(iii) Reject $H_0$, if computed value of t is less than or greater than the critical value, ie., $P(t < -t_{\alpha/2})$ or $P(t > t_{\alpha/2})$, otherwise do not reject $H_0$

Alternatively, reject $H_0$, if computed value of $|t|$ is greater than the critical value, ie., $P(|t| > t_{\alpha/2})$, otherwise do not reject $H_0$. Besides α, the degrees of freedom is also important.

## Conclusion

# Example 1

It is claimed that sports-car owners drive on the average 18580 kms per year. A consumer firm believes that the average milage is probably higher. To check, the consumer firm obtained information from randomly selected 10 sports-car owners that resulted in a sample mean of 17352 kms with a sample standard deviation of 2012 kms. What can be concluded about this claim at

- 5% level of significance
- 1% level of significance

$H_0$

The average milage of sports-car as claimed and the sample average milage may be same

$$H_0 : \mu = \mu_0 = 18580$$

$H_1$

The average milage of sports-car as claimed may be higher than the sample average milage

$$H_1 : \mu > \mu_0 = 18580$$

(a) At 5% level of significance with critical value 1.645

$$|t| = \frac{|17352 - 18580|}{\frac{2012}{\sqrt{10}}} = 1.929$$

95% CI for $\mu$ is

$$\bar{x} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} = [16184.91, 18519.09]$$

$$P - \text{value} = 0.0428$$

Hypothesis to test

$H_0: \mu = 18580$ vs $H_1: \mu > 18580$

Critical value for $\alpha$ = 0.05 is 1.833 for 9 degree of freedom

Since | t | =1.929 > 1.833, Reject $H_0$ and Accept $H_1$

# EXAMPLE 2

The management of a local health club claims that its members lose on the average 7 kgs or more within 3 months after joining the club. To check this claim, a consumer agency took a random sample of 15 members of this health club and found that they lost an average of 6.26 kgs within the first three months of membership. The sample standard deviation 1.91 kgs.

- Test at 1% level of significance whether the claim made by management of a local health club is acceptable or not?
- Also find the P-value of this test.

$H_0$

The average weight loss as claimed by the health club management is 7

$$H_0 : \mu = \mu_0 \geq 7$$

$H_1$

The average weight loss as claimed by the health club management of 7 may be higher than the sample average weight loss

$$H_1 : \mu = \mu_0 < 7$$

At 1% (0.01) level of significance with critical value 2.624

$$|t| = \frac{|\bar{x} - \mu_0|}{\frac{S}{\sqrt{n}}} = \frac{|6.26 - 7|}{\frac{1.91}{\sqrt{15}}} = 1.501$$

Hypothesis to test

$H_0: \mu = \mu_0 \geq 7$

vs

$H_1: \mu = \mu_0 < 7$

95% CI for μ

[5.203, 7.317]

includes $\mu_0 = 7$

P-value is

$$0.05 < P < 0.1$$

**99% CI for μ is (4.792, 7.728)**

Critical value for α = 0.01 is 2.624 for df=14. $|t| = 1.501 < 2.624$, accept $H_0$ & Reject $H_1$

# Testing the difference between means

1. State null and alternative hypothesis
2. Specify the level of significance 'α'
3. Standard Normal Distribution
4. Compute the test statistic
5. Define the critical region/ rejection criteria
6. Conclusion

$$H_0 : \mu_1 = \mu_2 \text{ vs } H_1 : \mu_1 < \mu_2$$
$$\text{or } H_1 : \mu_1 > \mu_2$$
$$\text{or } H_1 : \mu_1 \neq \mu_2$$

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{S_1^{\;2}}{n_1} + \frac{S_2^{\;2}}{n_2}}} \cong t_{(\alpha, n_1+n_2-2)}$$

Note: If sample sizes are unequal compute pooled SE

Note: Rejection criteria may be based on critical value or P-value

## t-test → Difference between means of two populations ($\mu_1 - \mu_2$)

- Samples are drawn from normal populations
- The population variances should be unknown
- The sample size should be less than 30 (i.e., $n < 30$)
- The population variances should be equal
- Two groups should be independent
- Subjects should be allocated randomly to both groups

**t-test**

- However even if sample size more than 30 (i.e., $n > 30$) and population variances unknown, t-test should be continue to apply, because of central limit theorem it approaches normal.

# Example 3

Random samples of 15 and 10 were selected from two thermocouples. The sample means were 315, 303 and sample standard deviations were 3.8, 4.9 respectively.

- Construct 95% CI for difference in means
- Test whether there is any significant difference in the means of two thermocouples at 5% level of significance
- Find the P-value

$H_0$

The mean of two thermocouples may be same

$$H_0 : \mu_1 = \mu_2$$

$H_1$

The mean of two thermocouples may be different

$$H_1 : \mu_1 \neq \mu_2$$

At 5% (0.05) level of significance with critical value

$$|t| = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^{\,2}}{n_1} + \frac{S_2^{\,2}}{n_2}}} = \frac{315 - 303}{\sqrt{\frac{(3.8)^2}{15} + \frac{(4.9)^2}{10}}} = 3.571$$

Hypothesis to test

**$H_0 : \mu_1 - \mu_2 = 0$**

vs

**$H_1 : \mu_1 - \mu_2 > 0$**

**???**

**95% CI for μ is [6.24, 17.76] not includes 0**

**95% CI for μ is [6.24, 17.76]**

**Critical value for α = 0.05 is 1.714. Since | t | = 3.571 > 1.714, Reject $H_0$ & Accept $H_1$**

# PROBLEM:

The manager of a courier service believes that packets delivered at the beginning of the month are heavier than those delivered at the end of month. As an experiment, he weighed a random sample of 15 packets at the beginning of the month and found that the mean weight was 5.25 kg. A randomly selected 10 packets at the end of the month had a mean weight of 4.56 kg. It was observed from the past experience that the sample variances are 1.20 kg and 1.15 kg.

- At 5% level of significance, can it be concluded that the packets delivered at the beginning of the month weigh more?
- Also find P-value and 95% confidence interval for the difference between themeans.

$H_0$

The mean weight of packets delivered at the early in the month and at the end of month may be same

$$H_0 : \mu_1 = \mu_2$$

$H_1$

The mean weight of packets delivered at the early in the month may be higher than at the end of month

$$H_1 : \mu_1 > \mu_2$$

# Estimation → Confidence interval for $(\mu_1-\mu_2)$ based t-test

**Finding Confidence Interval for difference between two population means $(\mu_1 - \mu_2)$**

**The100 (1-α)% confidence interval for difference between two means**

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\, SE\, (\bar{x}_1 - \bar{x}_2)$$

$$(\bar{x}_1 - \bar{x}_2) - t_{\alpha/2}\, SE(\bar{x}_1 - \bar{x}_2) \le (\mu_1 - \mu_2) \le (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2}\, SE(\bar{x}_1 - \bar{x}_2)$$

## 95% Confidence Interval for difference between two population means ($\mu_1$ - $\mu_2$)

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\, SE\,(\bar{x}_1 - \bar{x}_2) = (5.25 - 4.26) \pm 2.069 * 0.443$$

$$= (0.073, 1.907)$$

# At 5% (0.05) level of significance with critical value 1.714

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{5.25-4.26}{0.443} = 2.233$$

Hypothesis to test

$H_0: \mu_1 - \mu_2 = 0$

vs

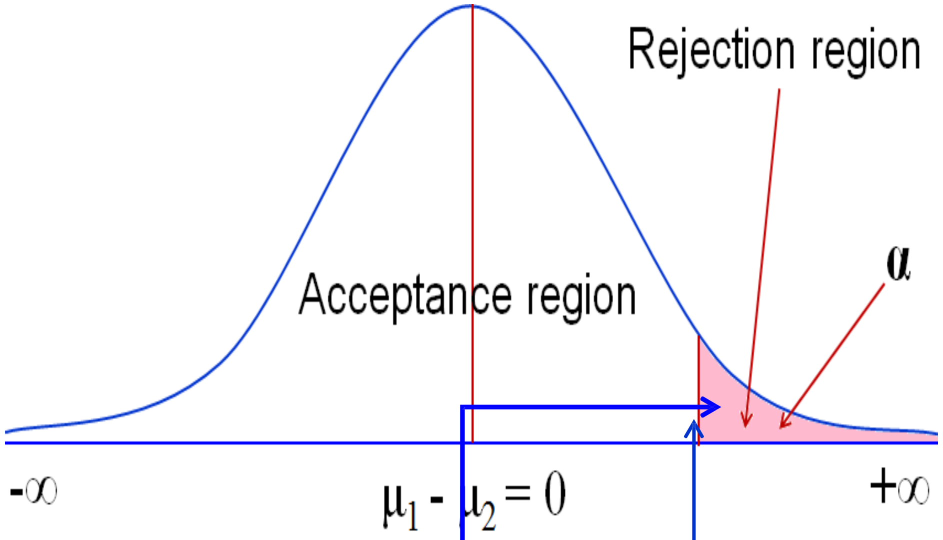
$H_1: \mu_1 - \mu_2 > 0$

???

$$0.025 \le P \le 0.01$$

**$(\mu_1 - \mu_2) = 0$ not included in**

**95% CI for $\mu_1 - \mu_2$ is (0.073, 1.907)**




**$t_{cal} = 2.233$**

$Z_\alpha = 1.714$

Critical value for $\alpha = 0.05$ is 1.714. Since $t = 2.233 > 1.714$, Reject $H_0$, Don't reject $H_1$

# Student's paired t-test

**t-test**

**Testing mean before and after observations of a single population ($\mu_d$)**

## Assumptions

Assume that the difference between before and after observations follow normal distribution

**The sample size should be less than 30 (n < 30)**

1 **State null and alternative hypothesis**

$$H_0 : \mu_d = 0 \text{ vs } H_1 : \mu_d < 0$$
$$\text{or } H_1 : \mu_d > 0$$
$$\text{or } H_1 : \mu_d \neq 0$$

2 **Specify the level of significance 'α'**

3 **Student's t-distribution**

4 **Compute the test statistic**

$$t = \frac{\bar{d} - \mu_d}{\frac{s_d}{\sqrt{n}}} \cong t_{(\alpha, n-1)}$$

**But $\mu_d$ under $H_0$ will be 0**

5 **Define the critical region/ rejection criteria**

6 **Conclusion**

**Note: Rejection criteria may be based on critical value or P-value**

# Student’s paired t-test

- The HRD manager wishes to see if there has been any change in the ability of trainees after a specific training programme.
- The trainees take a aptitude test

Before and after training programme.

| Subjects | Before (x) | After (y) |
|---|---|---|
| 1 | 75 | 70 |
| 2 | 70 | 77 |
| 3 | 46 | 57 |
| 4 | 68 | 60 |
| 5 | 68 | 79 |
| 6 | 43 | 64 |
| 7 | 55 | 55 |
| 8 | 68 | 77 |
| 9 | 77 | 76 |

| Subjects | Before (x) | After (y) | d = y-x | (d-mean)$^2$ |
|---|---|---|---|---|
| 1 | 75 | 70 | -5 | 100 |
| 2 | 70 | 77 | 7 | 4 |
| 3 | 46 | 57 | 11 | 36 |
| 4 | 68 | 60 | -8 | 169 |
| 5 | 68 | 79 | 11 | 36 |
| 6 | 43 | 64 | 21 | 256 |
| 7 | 55 | 55 | 0 | 25 |
| 8 | 68 | 77 | 9 | 16 |
| 9 | 77 | 76 | -1 | 36 |
| Total | | | 45 | 678 |

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n} = \frac{45}{9} = 5$$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n} (d_i - \bar{d})^2}{n-1}}$$

$$S_d = \sqrt{\frac{678}{8}} = 9.21$$

At 5% (0.05) level of significance with critical value is 3.31

$$|t| = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}} = \frac{5-0}{\frac{9.21}{\sqrt{9}}} = 3.07$$

Hypothesis to test

$H_0: \mu_d = 0$

vs

$H_1: \mu_d > 0$

???

**95% CI for μ is [[-2.52, 12.52] not includes 0**

**95% CI for μ is [-2.52, 12.52]**

Critical value for α = 0.05 is 1.895. Since | t | = 1.63 < 2.31, Accept $H_0$ & Reject $H_1$

# Exercise

Diet-modification Program


Ten individuals have participated

| Subject | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| Weight Before | 195 | 213 | 247 | 201 | 187 | 210 | 215 | 246 | 294 | 310 |
| Weight After | 187 | 195 | 221 | 190 | 175 | 197 | 199 | 221 | 278 | 285 |

Is there sufficient evidence to support claim that this program is effective in reducing weight?

Use $\alpha = 0.05$.

Construct 95% confidence interval for mean difference.

Test 1

Test 2

Is there sufficient evidence to conclude that both tests give the same mean impurity level

| Specimen | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| Test 1 | 1.2 | 1.3 | 1.5 | 1.4 | 1.7 | 1.8 | 1.4 | 1.3 |
| Test 2 | 1.4 | 1.7 | 1.5 | 1.3 | 2.0 | 2.1 | 1.7 | 1.6 |

Using $\alpha = 0.01$

Construct 99% confidence interval for mean difference

# Errors in Hypothesis Testing

## Type I error calculation

$\boldsymbol{\alpha}$: denotes the probability of making a Type I error

$$\boldsymbol{\alpha} = \mathbf{P}(\text{Rejecting } H_0 | H_0 \text{ is true})$$

## Type II error calculation

$\boldsymbol{\beta}$: denotes the probability of making a Type II error

$$\boldsymbol{\beta} = \mathbf{P}(\text{Accepting } H_0 | H_0 \text{ is false})$$

**Note:**

$\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are not independent of each other as one increases, the other decreases

When the sample size increases, both to decrease since sampling error is reduced.

In general, we focus on Type I error, but Type II error is also important, particularly when sample size is small.

# Errors in Hypothesis Testing

In hypothesis testing, there are two types of errors.

**Type I error:** A type I error occurs when we incorrectly reject $H_0$ (i.e., we reject the null hypothesis, when $H_0$ is true).

**Type II error:** A type II error occurs when we incorrectly fail to reject $H_0$ (i.e., we accept $H_0$ when it is not true).

| | Observation | |
|---|---|---|
| **Decision** | $H_0$ is true | $H_0$ is false |
| $H_0$ is accepted | Decision is correct | Type II error |
| $H_0$ is rejected | Type I error | Decision is correct |

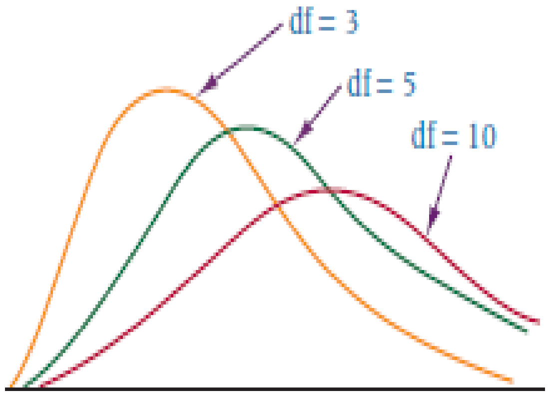
# Chi-square Statistic

- Although the technique is still rather widely presented as a mechanism for constructing confidence intervals to estimate a population variance, you should proceed with extreme caution and **apply** the technique **only** in cases **where** the **population** is known to be **normally distributed**. We can say that this technique lacks robustness.

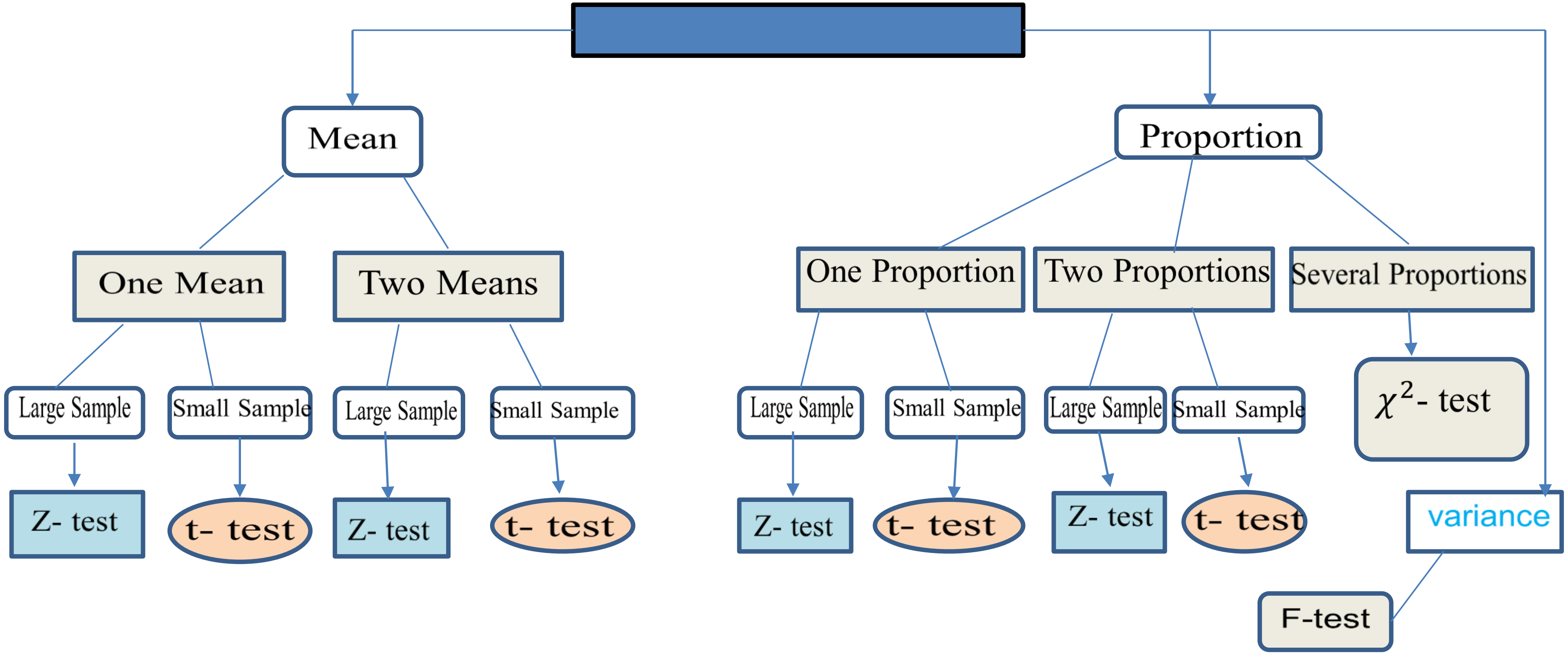
$\chi^2$ FORMULA FOR SINGLE VARIANCE (8.5)

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$\text{df} = n - 1$$




**Three Chi-Square Distributions**


Mean
Proportion
One Mean
Two Means
One Proportion
Two Proportions
Several Proportions
Large Sample
Small Sample
Large Sample
Small Sample
Large Sample
Small Sample
Large Sample
Small Sample
$\chi^2$- test
Z- test
t- test
Z- test
t- test
Z- test
t- test
Z- test
t- test
variance
F-test

| Critical value | Level of significance ∝ | | |
|---|---|---|---|
| | 1% | 5% | 10% |
| Two tailed test | $z_{\alpha/2} = 2.58$ | $z_{\alpha/2} = 1.96$ | $z_{\alpha/2} = 1.645$ |
| One tailed test | $z_{\alpha} = 2.33$ | $z_{\alpha} = 1.645$ | $z_{\alpha} = 1.28$ |

# PROBLEM

The Edison Electric Institute has published figures on the annual number of kilowatt-hours expended by various home appliances. It is claimed that a vacuum cleaner expends an average of 46 kilowatt-hoursper year. If a random sample of 20 homes included in a planned study indicates that vacuum cleaners expends an average of 42 kilowatt-hours per year with a standard deviation of 11.9 kilowatt-hours.

- Does this suggest at the 0.05 level of significance that vacuum cleaners expends, on an average less than 46 kilowatt-hors annually?
- Assume population of kilowatt- hours to be normal.

# PROBLEM

The manager of a courier service believes that packets delivered at the end of the month are heavier than those delivered early in the month.

As an experiment, he weighed a random sample of 10 packets at the beginning of the month and found that the mean weight was 5.25 kg. A randomly selected 10 packets at the end of the month had a mean weight of 4.96 kg. The respective sample standard deviations are 1.20 kg and 1.15 kg.

- At 5% level of significance, can it be concluded that the packets delivered at the end of the month weigh more?
- Also find P- value and 95% confidence interval for the difference between the means.

# Problem:

. Suppose $\mu_1$ and $\mu_2$ are true mean stopping distances at 50 mph for cars of a certain type equipped with two different types of braking systems.

- Use the two-sample t test at significance level .01 to test $H_0$:$\mu_1$-$\mu_2 = -10$ versus $H_0$:$\mu_1$-$\mu_2 < -10$ for the following data: m=6,$\bar{x} = 115.7, s_1 = 5.03, n = 6,$
  $\bar{y} = 129.3 \; and \; s_2 = 5.38$

# Thanks