Introduction to Statistical Methods

ISM Team

Session No: 9
Testing of Hypothesis
(5th /6th Aug ,2023)

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.

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

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

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	$n \geq 30$
Standard Normal Test	Z-Test
Small Sample Test	$n < 30$
Student's t-test	Unpaired t-Test Paired t-Test
Analysis of Variance	ANOVA 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 Hypothesis Testing

innovate achieve lead

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

Errors in decision making

Any example based on data		
Statistical Example		
Decision	Null Hypothesis (H_0)	
	True	False
Accept		
Reject		

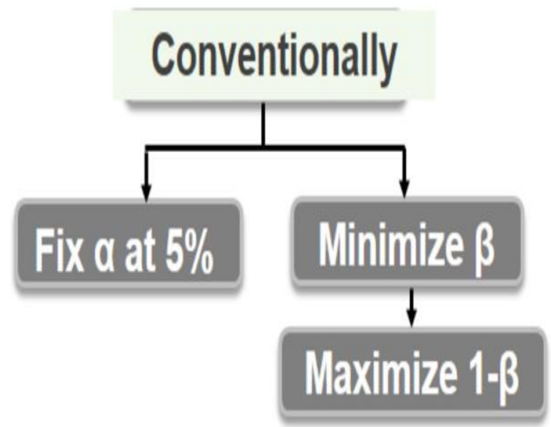
Any example based on data		
Statistical Example		
Decision	Null Hypothesis (H_0)	
	True	False
Accept		Type – II Error
Reject	Type – I Error	

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



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

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.

Univariate analysis

Dependent Variable

	Nominal	Ordinal	Interval/ Ratio
Statistical measure	Proportion	Median	Mean and Variance
Large sample tests ($n \geq 30$)	Z - test	Sign test	Z-test (If normality assumption holds and σ^2 known) t-test (If normality assumption holds and σ^2 unknown)
Small sample tests ($n < 30$)	Binomial test	Sign test	t-test (If normality assumption holds and σ^2 unknown)

Multivariate analysis

Dependent Variable

			Nominal	Ordinal	Interval/ Ratio
Statistical measure			Proportion	Median	Mean and Variance
Large sample tests ($n \geq 30$)	Independent variable	2 groups	Z - test	Mann-Whitney U-test	Z-test (If normality assumption holds and σ^2 known) t-test (If normality assumption holds and σ^2 unknown)
		> 2 groups	Chi-square test	Kruskal-Wallis test	ANOVA (If normality assumption holds)
Small sample tests ($n < 30$)	Independent variable	2 groups	Chi-square test	Mann-Whitney U-test	Z-test (If normality assumption holds and σ^2 known) t-test (If normality assumption holds and σ^2 unknown)
		> 2 groups	Chi-square test	Kruskal-Wallis test	ANOVA (If normality assumption holds)

Parametric tests

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

Used for testing the

- Mean of a single population (μ)
- 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

Z-test

- Samples are drawn from normal distribution
- The population variances should be known
- 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

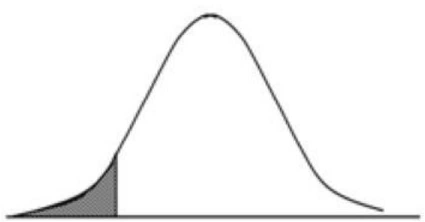
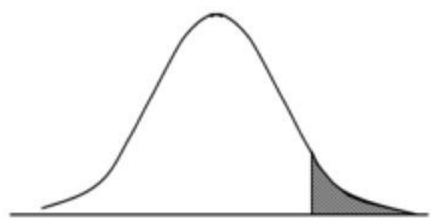
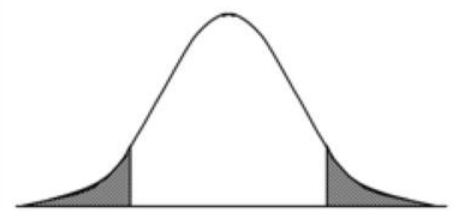
1. State null and alternative hypothesis
2. Specify the level of significance 'α'
3. Standard Normal Distribution
4. Compute the test statistic
5. Define the critical region/ rejection criteria
6. Conclusion

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

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

Rejection criteria

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 $H_1: \mu \neq \mu_0$
		

Example:1

- It is claimed that sports-car owners drive on the average 17000 kms per year. A consumer firm believes that the average milage is probably higher. To check, the consumer firm obtained information from randomly selected 40 sports-car owners that resulted in a sample mean of 17352 kms with a population standard deviation of 1348 kms. At what can be concluded about this claim at
- 5% level of significance (Critical value is 1.645)
- 1% level of significance (Critical value is 2.331)

H_0 ⇨ The average milage of sports-car as claimed and the sample average milage may be same

$$H_0 : \mu = \mu_0 = 17000$$

H_1 ⇨ The average milage of sports-car as claimed may be higher than the sample average milage

$$H_1 : \mu > \mu_0 = 17000$$

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

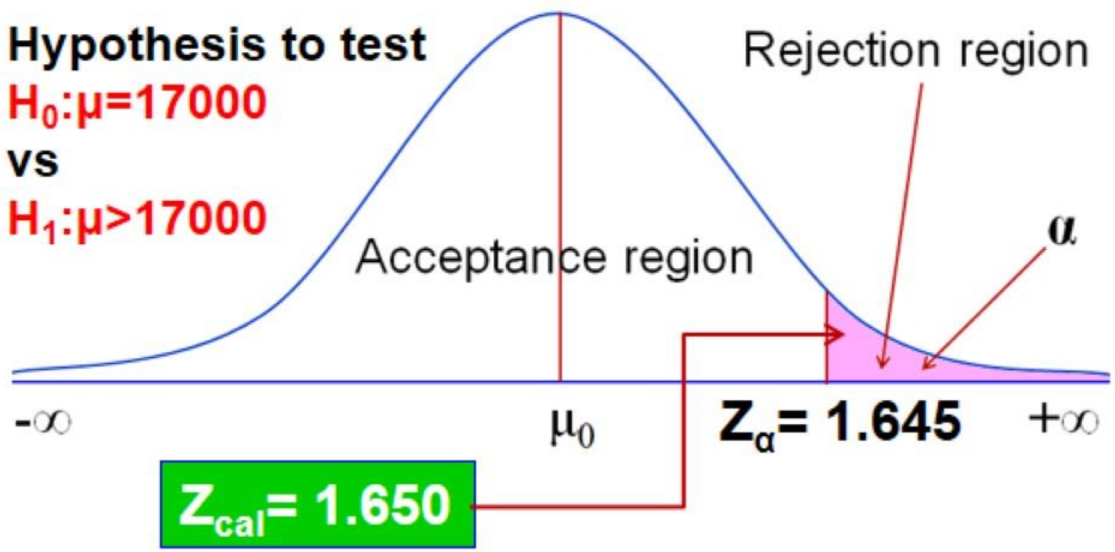
$$Z = \frac{17352 - 17000}{\frac{1348}{\sqrt{40}}} = 1.650$$

Hypothesis to test

$H_0: \mu = 17000$

vs

$H_1: \mu > 17000$



Critical value for

$\alpha = 0.05$ is 1.645

Since Z =1.650 > 1.645,

Reject H_0 and Accept H_1

(b) At 1% level of significance with critical value 2.331

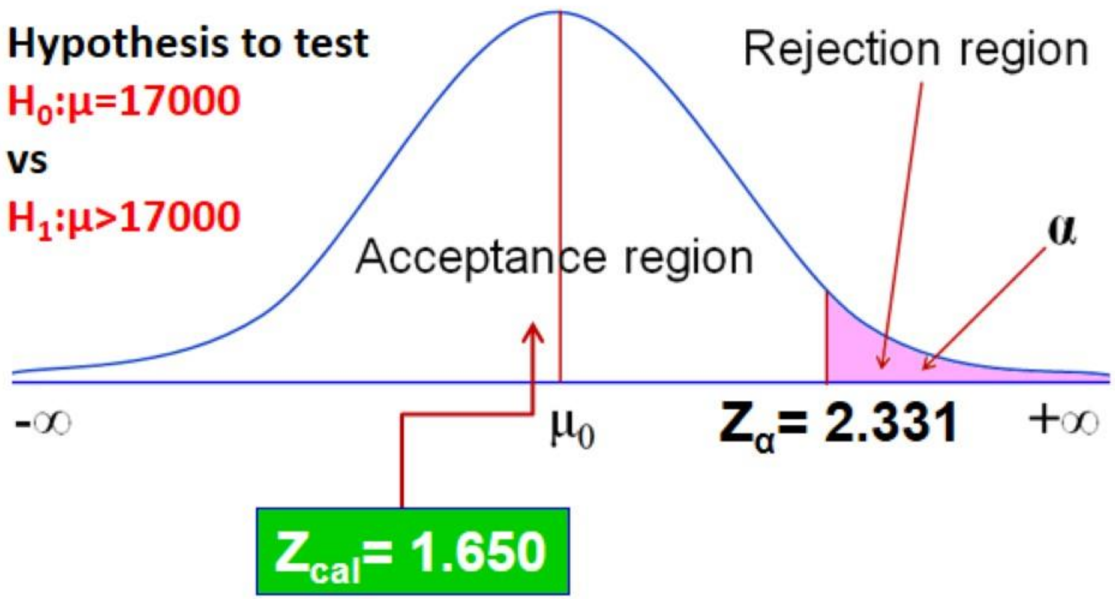
$$Z = \frac{17352 - 17000}{\frac{1348}{\sqrt{40}}} = 1.650$$

Hypothesis to test

$H_0: \mu = 17000$

vs

$H_1: \mu > 17000$



Critical value for

$\alpha = 0.01$ is 2.331

Since Z =1.624 < 2.330,

Accept H_0 Reject H_1

Example:2

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 20 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. It was observed from the past experience that the population 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 the means.

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 end of the month may be higher than at the early of month

$$H_1 : \mu_1 > \mu_2$$

At 5% (0.05) level of significance with critical value 1.645

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{5.25 - 4.96}{\sqrt{\frac{(1.20)^2}{20} + \frac{(1.15)^2}{10}}} = 0.642$$

$$P(Z < 0.642) = 0.7389$$

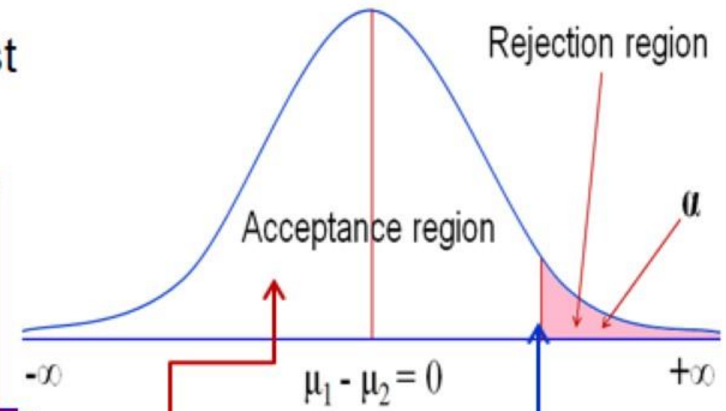
Hypothesis to test

$H_0: \mu_1 - \mu_2 = 0$

vs

$H_1: \mu_1 - \mu_2 > 0$

???



95% CI for μ is [- 454, 1.033]

$Z_{cal} = 0.642$

$Z_{\alpha} = 1.645$

Critical value for α = 0.05 is 1.645. Since Z = 0.642 < 1.645, Accept H_0 Reject H_1

Z-test for proportion of a single population

Z-test → **Proportion of a single population (P)**

Assumptions

- Assume that the samples are drawn from normal distribution
- The sample size should be more than or equal to 30
- Subjects should be selected randomly

1 **State null and alternative hypothesis**

$$H_0 : P = P_0 \text{ vs } H_1 : P < P_0$$
$$\text{or } H_1 : P > P_0$$
$$\text{or } H_1 : P \neq P_0$$

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

3 **Standard Normal Distribution**

4 **Compute the test statistic**

$$Z = \frac{p - P_0}{\sqrt{\frac{pq}{n}}} \cong N(0,1)$$

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

6 **Conclusion**

Note: Rejection criteria same as in one sample test or P-value

Example:

A researcher claims that Republican Party will win in the next Senate elections, especially in Florida State. Statistical data reported that 23% voted for Republican Party in the last election. To test the claim a researcher surveyed 80 people and found 22 said they voted for Republican Party in the last election. Is there enough evidence at $\alpha=0.05$ to support this claim?

- $P_0=0.23$
- $n=80$
- $\hat{p}=22/80=0.275$

Then, define the Null and Alternative hypothesis

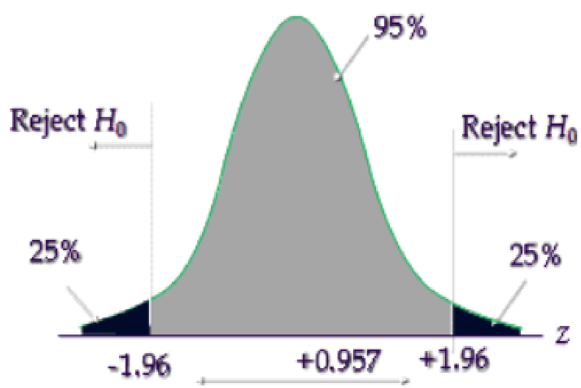
- Null Hypothesis: $p= 0.23$
- Alternative Hypothesis: $p\neq 0.23$

Then, calculate the test statistic

$$z = \frac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/_n}}$$

$$z = \frac{0.275 - 0.23}{\sqrt{0.23 * 0.77/_{80}}}$$

= 0.045 / 0.047 = 0.957



Since the calculated value is between -1.96 and 1.96
and it is not in the critical region, hence failed to reject the null hypothesis

Z-test for difference between proportions

1 State null and alternative hypothesis

$$H_0 : P_1 = P_2 \text{ vs } H_1 : P_1 < P_2$$
$$\text{or } H_1 : P_1 > P_2$$
$$\text{or } H_1 : P_1 \neq P_2$$

2 Specify the level of significance 'α'

3 Standard Normal Distribution

4 Compute the test statistic

$$Z = \frac{p_1 - p_2}{\sqrt{\dfrac{p_1 q_1}{n_1} + \dfrac{p_2 q_2}{n_2}}} \cong N(0,1)$$

5 Define the critical region/ rejection criteria

6 Conclusion

Note: Rejection criteria same as in one sample test or P-value

Testing of Hypothesis ➡ Z-test for proportions of two popln.

Based on sample size standard error for proportion may be calculated:

If the sample sizes are equal then SE (P_1-P_2) is calculated by

$$SE(p_1 - p_2) = \sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}$$

Based on sample size standard error for proportion may be calculated:

If the sample sizes are equal then SE (P_1-P_2) is calculated by

$$SE(p_1 - p_2) = \sqrt{pq\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}, p = \frac{x_1 + x_2}{n_1 + n_2} \text{ and } q = 1 - p$$

Testing of Hypothesis ⇒ Z-test for proportions of two popln.

A cigarette manufacturing company claims that its brand A cigarettes outsells its brand B cigarettes by 8%. If it is found that 42 out of a random sample of 200 smokers prefer brand A and 18 out of 100 smokers prefer brand B, test at 5% level of significance, whether 8% difference a valid claim. Also construct 95% CI for (P_1-P_2) and find P-value.

Testing of Hypothesis ➡ Z-test for proportions of two popln.

H_0 ⇨ There may be 8% difference in the sale of two brands of cigarettes may be a valid claim

$$H_0 : P_1 - P_2 = 0.08$$

H_1 ⇨ There may be 8% difference in the sale of two brands of cigarettes may not be a valid claim

$$H_1 : P_1 - P_2 \neq 0.08$$

Finding Confidence Interval for difference in population proportion ($P_1 - P_2$)

The100 (1-α)% confidence interval for difference between two population proportions ($P_1 - P_2$)

$$(p_1 - p_2) \pm Z_{\alpha/2}\, SE\,(p_1 - p_2)$$

Estimation ➡ Interval estimate

95% confidence interval for difference between two population proportions $(P_1 – P_2)$

$$(p_1 - p_2) \pm Z_{\alpha/2}\ SE\ (p_1 - p_2) = (0.21 - 0.18)\ \pm 1.96 * 0.049$$
$$= (-0.066,\ 0.126)$$

Testing of Hypothesis ⇒ Z-test for proportions of two popln.

At 5% level of significance with critical value ±1.96

$$Z = \frac{(0.21 - 0.18) - 0.08}{\sqrt{0.2 * 0.8\left(\frac{1}{200} + \frac{1}{100}\right)}} = -1.02$$

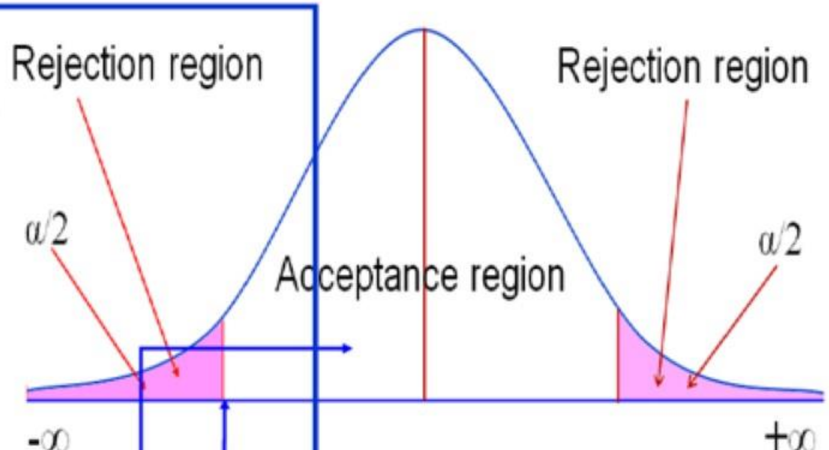
Hypothesis to test

$H_0: P_1 - P_2 = 0.08$

vs

$H_1: P_1 - P_2 \neq 0.08$

$P(Z < -2.074) = 0.1539$



Critical value for $\alpha = 0.05$ is -1.96

Since $Z = -1.02 > -1.96$, Don't reject H_0 and reject H_1

$Z_{cal} = -1.02$

$Z_{0.025} = -1.96$

95% CI for P is (-0.066, 0.126)

Testing of Hypothesis ➡ Z-test for proportions of two popln.

Problem on Z-test on difference in proportions

Use $\alpha = 1\%$.

Bangalore 2019

387/1500 BSNL telephones connections were disconnected

Delhi 2019

310/1200 BSNL telephones connections were disconnected

- Is there a significant difference between the two proportion of telephones disconnection by BSNL?

Construct 99% confidence interval for difference in proportions

Testing of Hypothesis ➡ Student's t-test

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

t-test

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

1. State null and alternative hypothesis
2. Specify the level of significance 'α'
3. Student's t-distribution
4. Compute the test statistic
5. Define the critical region/ rejection criteria
6. Conclusion

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

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

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

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 μ 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

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 ($\mu_1 - \mu_2$)**

t-test

- 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 shouldbe independent
- Subjects should be allocated randomly to both groups

- 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

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]

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

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

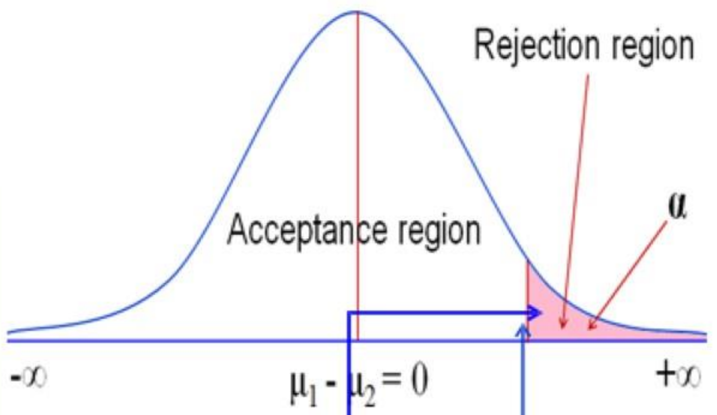
$H_1: \mu_1 - \mu_2 > 0$

$$0.025 \leq P \leq 0.01$$

???

$(\mu_1 - \mu_2) = 0$ 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

1. State null and alternative hypothesis
2. Specify the level of significance 'α'
3. Student's t-distribution
4. Compute the test statistic
5. Define the critical region/ rejection criteria
6. Conclusion

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

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

But μ_d under H_0 will be 0

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 $\alpha = 0.05$ 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

Problem-1

The target thickness for silicon wafers used in a certain type of integrated circuit is 245mm. A sample of 50 wafers is obtained and the thickness of each one is determined, resulting in a sample mean thickness of 246.18mm and a sample standard deviation of 3.60mm.

Does this data suggest that true average wafer thickness is something other than the target value?

Problem 2

In a survey of buying habits. 400 women shopper's are chosen at random in super Market A located in a certain city. Their average weekly food expenditure is Rs.250 with a SD of Rs.40. For 400 women shopper's are chosen at random in super Market B in another city. The average weekly food expenditure is Rs.220 with a SD of Rs.55.

Test at 1% LOS whether the average weekly food expenditure of the two populations of shoppers are equal.

Problem 3

- Two types of new cars produced in INDIA are tested for petrol mileage. One sample consisting of 42 cars averaged 15 kmpl while the other sample consisting of 80 cars averaged 11.5 kmpl with population variance as ${\sigma_1}^2 = 2$ and ${\sigma_2}^2 = 1.5$ respectively.

- Test whether there is any significance difference in the petrol consumption of these two types of cars.

Problem 4

Studying the flow of traffic at two busy intersections between 4pm and 6pm to determine the possible need for turn signals. It was found that on 40 week days there were on the average 247.3 cars approaching the first intersection from the south which made left turn, while on 30 week days there were on the average 254.1 cars approaching the second intersection from the south made left turns. The corresponding sample SD are 15.2 and 12.Test the significance between the difference of two means at 5% level.

Problem 5

- Is there any systematic tendency for part-time college faculty to hold their students to different standards than do full-time faculty? The article “Are There Instructional Differences Between Full-Time and Part-Time Faculty?” (College Teaching, 2009: 23–26) reported that for a sample of 125 courses taught by full-time faculty, the mean course GPA was 2.7186 and the standard deviation was .63342, whereas for a sample of 88 courses taught by part-timers, the mean and standard deviation were 2.8639 and .49241, respectively.

- Does it appear that true average course GPA for part-time faculty differs from that for faculty teaching full-time? Test the appropriate hypotheses at significance level 1%.

Problem 6

It is thought that the front cover and the nature of the first question on mail surveys influence the response rate. The article “The Impact of Cover Design and First Questions on Response Rates for a Mail Survey of Skydivers” (*Leisure Sciences,* 1991: 67–76) tested this theory by experimenting with different cover designs. One cover was plain; the other used a picture of a skydiver. The researchers speculated that the return rate would be lower for the plain cover.

Cover	Number Sent	Number Returned
Plain	207	104
Skydiver	213	109

Does this data support the researchers’ hypothesis?

Problem 7

A spectrophotometer used for measuring CO concentration [ppm (parts per million) by volume] is checked for accuracy by taking readings on a manufactured gas (called span gas) in which the CO concentration is very precisely controlled at 70 ppm. If the readings suggest that the spectrophotometer is not working properly, it will have to be recalibrated.

Assume that if it is properly calibrated, measured concentration for span gas samples is normally distributed.

On the

basis of the six readings—85, 77, 82, 68, 72, and 69—is recalibration necessary?

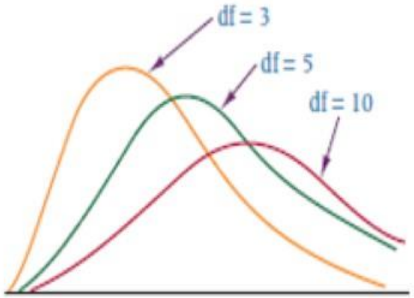
Chi-square Statistic

- 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.

χ^2 FORMULA FOR SINGLE VARIANCE (8.5)

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

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



Three Chi-Square Distributions

Testing of Hypothesis ➡ Chi-square test: Independence

Chi-square test

Chi-square Test

- Independence
- Goodness-of-fit

- Should be applied ONLY for Frequencies
- Not for percentages, ratios, mean etc.

Testing of Hypothesis ➡ Chi-square test: Independence

Based on attributes used to test

(a) INDEPENCE of two different categorical variables

or

(b) GOODNESS OF FIT

Caution:

Should be applied ONLY for FREQUENCIES not for

percentages, ratios, mean etc

Testing of Hypothesis ⇨ Chi-square test: Independence

Hypothesis for testing independence

The hypothesis to be tested for independence will be

H_0: The two categorical variables may be independent (may not be associated)

H_1: The two categorical variables may not be independent (may be associated)

Testing of Hypothesis ➡ Chi-square test: Independence

Procedure for testing independence

To check the independence (no association) between the two categorical variables, the statistical test used is Chi-square test given by

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}, k = r \text{ x } c \text{ \# of cells}$$

The test-statistic follows Chi-square distribution with (r-1)(c-1) degrees of freedom. r = # of rows, c = # of columns

Testing of Hypothesis ➡ Chi-square test: Independence

Expected frequencies

$$E_{ij} = \frac{r_i c_j}{n},$$
$$\text{for } i = 1, 2, \ldots, m;\ j = 1, 2, \ldots, n$$

Chi-square is calculated by

$$\chi^2 = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i} \approx \chi^2_{[(r-1)(c-1)]}$$

where k = r x c is the total number of cells in the r x c contingency table, r = total no. of rows and c is total no. of columns.

Testing of Hypothesis ➡ Chi-square test: Independence

Assumptions of Chi-square test

If the expected cell frequencies is < 5

Yate's correction should be applied for continuity

In a 2 x 2 contingency table, if one or more of the cell has the expected cell frequencies is < 5,

Fisher's exact probabilities should be computed

For the use of Chi-square test

The sample size should not be less than 20.

For an r x c table, if the expected frequencies in any cells are < 5, merge the rows are columns meaningfully

The Fisher's exact Probability

$$P = \frac{1}{n!}\frac{r_1!}{a!}\frac{r_2!}{b!}\frac{c_1!}{c!}\frac{c_2!}{d!}$$

Testing of Hypothesis ➡ Chi-square test: Independence

A study to find the association between smoking and ca. lung has revealed the following data? Find is there any association exists between smoking and ca. lung?

Smoking	Carcinoma of lung		Total
	Present	Absent	
Smokers	69	2431	2500
Non-smokers	24	1476	1500
Total	93	3907	4000

Testing of Hypothesis ➡ Chi-square test: Independence

Calculation of expected frequencies

$$E_1 = \frac{c_1 r_1}{n} = \frac{93 * 2500}{4000} = 58.125$$

$$E_2 = \frac{c_2 r_1}{n} = \frac{3907 * 2500}{4000} = 2441.875$$

$$E_3 = \frac{c_1 r_2}{n} = \frac{93 * 1500}{4000} = 34.875$$

$$E_4 = \frac{c_2 r_1}{n} = \frac{3907 * 1500}{4000} = 1465.125$$

Testing of Hypothesis ➡ Chi-square test: Independence

Calculation of Chi-square statistic - $\chi 2$

Sl No	Observed frequencies (O_i)	Expected frequencies (E_i)	$(O_i - E_i)$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	69	58.125	10.875	118.266	2.035
2	2431	2441.875	-10.875	118.266	0.048
3	24	34.875	-10.875	118.266	3.391
4	1476	1465.125	10.875	118.266	0.081
Total	4000	4000			$\chi^2 = 5.555$

Testing of Hypothesis ➡ Chi-square test: Independence

Percentage Points of the Chi-Square Distribution

Degrees of Freedom	Probability of a larger value of x^2								
	0.99	0.95	0.90	0.75	0.50	0.25	0.10	0.05	0.01
1	0.000	0.004	0.016	0.102	0.455	1.32	2.71	3.84	6.63
2	0.020	0.103	0.211	0.575	1.386	2.77	4.61	5.99	9.21
3	0.115	0.352	0.584	1.212	2.366	4.11	6.25	7.81	11.34
4	0.297	0.711	1.064	1.923	3.357	5.39	7.78	9.49	13.28
5	0.554	1.145	1.610	2.675	4.351	6.63	9.24	11.07	15.09
6	0.872	1.635	2.204	3.455	5.348	7.84	10.64	12.59	16.81
7	1.239	2.167	2.833	4.255	6.346	9.04	12.02	14.07	18.48
8	1.647	2.733	3.490	5.071	7.344	10.22	13.36	15.51	20.09
9	2.088	3.325	4.168	5.899	8.343	11.39	14.68	16.92	21.67
10	2.558	3.940	4.865	6.737	9.342	12.55	15.99	18.31	23.21

Testing of Hypothesis ➡ Chi-square test: Independence

Calculation of P – value

$$\frac{6.63 - 3.84}{5.56 \ - 3.84} = \frac{0.05 - 0.01}{\mathrm{P} \ - 0.01}$$

$$\mathrm{P} = 0.01 + \frac{(0.05 \ - 0.01) * (5.56 - 3.84)}{(6.63 \ - 3.84)} = \mathbf{0.035}$$

Testing of Hypothesis ➡ Chi-square test: Independence

Interpretation

H_0: Smoking habit and Cancer of lung may be independent (may not be associated)

H_1: Smoking habit and Cancer of lung may not be independent (may be associated)

$\chi^2 = 5.555$

df = 1, Critical value at $\alpha = 0.05$ is 3.841, P = 0.035

Inference: There may be an association between smoking and Cancer of lung

Testing of Hypothesis ➡ Chi-square test: Independence

By replacing O_1, O_2, O_3, and O_4, by a, b, c, and d the 2 x 2 contingency table can also be written as

Categorical variable 1	Categorical variable 2		Total
	Response 1	Response 2	
Response 1	**a**	**b**	r_1
Response 2	**c**	**d**	r_2
Total	c_1	c_2	n

Testing of Hypothesis ➡ Chi-square test: Independence

Alternate formula for calculation of chi-square statistic

When the expected frequencies in all the cells are more than 5, alternatively Chi-square statistic can be calculated using the formula for 2 x 2 table only by

$$\chi^2 = n \frac{(ad - bc)^2}{r_1 r_2 c_1 c_2}$$

Where a, b, c, d are cell frequencies; r_1 and r_2 are row totals; c_1 and c_2 are column totals

Testing of Hypothesis → Chi-square test: Independence

To assess the length of hospital stay and the type of insurance, data were taken on 70 individuals

Type of Insurance	Length of Hospital Stay (days)		Total
	≤10	>10	
Type 1	42	3	45
Type 2	18	7	25
Total	60	10	70

Examine whether Chi-square test can be applied to this data to test the independence between type of insurance and length of hospital stay?

Testing of Hypothesis ➡ Chi-square test: Independence

Calculation of expected frequencies

$$E_1 = \frac{c_1 r_1}{n} = \frac{60*45}{70} = 38.75$$

$$E_2 = \frac{c_2 r_1}{n} = \frac{10*45}{70} = 6.43$$

$$E_3 = \frac{c_1 r_2}{n} = \frac{60*25}{70} = 21.43$$

$$E_4 = \frac{c_2 r_1}{n} = \frac{10*25}{70} = 3.57$$

Testing of Hypothesis ➡ Chi-square test: Independence

Calculation of Chi-square statistic - $\chi 2$

Sl No	Observed frequencies (O_i)	Expected frequencies (E_i)	$(O_i - E_i)$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	42	38.57	-	-	-
2	3	6.43	-	-	-
3	18	21.43	-	-	-
4	7	3.57	-	-	-
Total	70	70			???

Testing of Hypothesis ➡ Chi-square test: Independence

Since the expected frequency in the 4 is less than 5 Chi-square cannot be applied and hence the Fisher's exact probabilities has to be calculated.

Testing of Hypothesis ➡ Chi-square test: Independence

H_0: Type of insurance plan and length of hospital stay may be independent

H_1: Type of insurance plan and length of hospital stay may be associated

Computed Probability

$$P = \frac{1}{70\,!}\frac{45\,!}{42\,!}\frac{25\,!}{3!}\frac{60\,!}{18\,!}\frac{10\,!}{7\,!}$$

One tailed: P=0.0201 Two tailed: P=0.0282

Conclusion: H_0 may be rejected and hence the type of insurance plan and length at may not be independent (may be associated)

Testing of Hypothesis ➡ Chi-square test: Independence

To assess the length of hospital stay and the type of insurance, data were taken on 70 individuals

Type of Insurance	Length of Hospital Stay (days)		Total
	≤10	>10	
Type 1	42	3	45
Type 2	13	12	25
Total	55	15	70

Examine whether Chi-square test can be applied to this data to test the independence between type of insurance and length of hospital stay?

Testing of Hypothesis ➡ Chi-square test: Independence

Calculation of expected frequencies

$$E_1 = \frac{r_1 c_1}{n} = \frac{45 * 55}{70} = 35.36$$

$$E_2 = \frac{r_1 c_2}{n} = \frac{45 * 15}{70} = 9.64$$

$$E_3 = \frac{r_2 c_1}{n} = \frac{25 * 55}{70} = 19.64$$

$$E_4 = \frac{r_2 c_2}{n} = \frac{25 * 15}{70} = 5.36$$

$$\chi^2 = \frac{(42 - 35.36)^2}{35.36} + \frac{(3 - 9.64)^2}{9.64} + \frac{(13 - 19.64)^2}{19.64} + \frac{(12 - 5.36)^2}{5.36} = 16.307$$

Testing of Hypothesis ➡ Chi-square test: Independence

- H_0: Duration of hospital stay and type of insurance plan may be independent (not associated)
- H_1: Duration of hospital stay and type of insurance plan may not be independent (Associated)
- $\chi^2 = 16.307$
- df=1
- $P < 0.001$
- Inference: Reject H_0, which shows duration of hospital stay and type of insurance may be associated

Testing of Hypothesis ➡ Chi-square test: Independence

Hypertension	Non smokers	Moderate smokers	Heavy smokers	Total
A	O_1	O_2	O_3	r_1
B	O_4	O_5	O_6	r_2
Total	c_1	c_2	c_3	n

The expected frequencies and Chi-square statistic are computed by

$$E_1 = \frac{r_1 c_1}{n}, \quad E_2 = \frac{r_1 c_2}{n}, \quad E_3 = \frac{r_1 c_3}{n}, \quad E_4 = \frac{r_2 c_1}{n},$$

$$E_5 = \frac{r_2 c_2}{n}, \quad E_6 = \frac{r_2 c_3}{n},$$ **and** $$\chi^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$$

Testing of Hypothesis ➡ Chi-square test: Independence

Under the null hypothesis, the observed frequencies and the calculated expected frequencies will be as follows:

Hypertension	Non smokers	Moderate smokers	Heavy smokers	Total
A	O_1 E_1	O_2 E_2	O_3 E_3	r_1
B	O_4 E_4	O_5 E_5	O_6 E_6	r_2
Total	c_1	c_2	c_3	n

Testing of Hypothesis ➡ Chi-square test: Independence

Three pension plans

Independent of job classification

Use $\alpha = 0.05$

The opinion of a random sample of 500 employees are shown below

Job Classification	Pension Plan			Total
	1	2	3	
Salaried workers	166	86	68	320
Hourly workers	84	64	32	180
Total	250	150	100	500

$$E_1 = \frac{r_1c_1}{n} = \frac{320 \text{ x } 250}{500} = 106.24$$

$$E_2 = \frac{r_1c_2}{n} = \frac{320 \text{ x } 150}{500} = 96.00$$

$$E_3 = \frac{r_1c_3}{n} = \frac{320 \text{ x } 100}{500} = 64.00$$

$$E_4 = \frac{r_2c_1}{n} = \frac{180 \text{ x } 250}{500} = 90.00$$

$$E_5 = \frac{r_2c_2}{n} = \frac{180 \text{ x } 150}{500} = 54.60$$

$$E_6 = \frac{r_2c_3}{n} = \frac{180 \text{ x } 100}{500} = 36.00$$

Testing of Hypothesis ➡ Chi-square test: Independence

Sl No	(O_i)	(E_i)	(O_i-E_i)	(O_i-E_i)^2	(O_i-E_i)^2/E_i
1	166	106.24	59.76	3571.26	33.62
2	86	96.00	-10.00	100.00	1.04
3	68	64.00	4.00	16.00	0.25
4	84	90.00	-6.00	36.00	0.40
5	64	54.60	9.40	88.36	1.62
6	32	36.00	-4.00	16.00	0.44
Total	180	180	Chi-square value		37.37

Testing of Hypothesis ➡ Chi-square test: Independence

- H_0: Job satisfaction and pension plan may be independently distributed (not associated)
- H_1: Job satisfaction and pension plan may not be independently distributed (Associated)
- $\chi^2 = 37.37$
- df=2
- $P<0.001$
- Inference: Reject H_0, which shows Job satisfaction and pension plan are associated

Problem

innovate achieve lead

A sample analysis of examination results of 500 students was made. It was found that 220 students had failed, 170 had secured a third class, 90 were placed in second class and 20 got a first class.

Do these figures commensurate with the general examination result which is in ratio of 4:3:2:1 for the various categories respectively.

H_0: The observed results commensurate with the general examination results.

H_1: The observed results not commensurate with the general examination results.

Expected frequencies are in the ratio of 4:3:2:1

Total frequency = 500

If we divide total frequency 500 in the ratio 4:3:2:1, We get the Expected frequencies

As 200$\left(500\ X\ \frac{4}{10}\right)$, 150$\left(500\ X\ \frac{3}{10}\right)$,100$\left(500\ X\ \frac{2}{10}\right)$,50$\left(500\ X\ \frac{1}{10}\right)$

S. NO	(Oi)	(Ei)	(Oi-Ei)	$(O-E)^2$	$(O-E)^2$/Ei
1	220	200	20	400	2
2	170	150	20	400	2.667
3	90	100	10	100	1
4	20	50	30	900	18

The Chi-square value is

- $\chi^2 = \sum_1^4 \frac{(O-E)^2}{E}$ = 2 + 2.667 + 1 +18

  = 23.667

Dof is 4-1 = 3 and Level of significance $\propto$ = 0.05

χ^2 tab value is 7.81

χ^2 calculated = 23.667 >χ^2 tab = 7.81 at $\propto$ = 0.05 for 3 dof

- Decision : we reject the Null hypothesis H_0 at α = 5 % LOS.

i.e, we accept H_1.

i.e, The observed results not commensurate with the general examination results

Hypothesis Testing

(Problems for Discussion)

EXAMPLE

The mean lifetime of a sample of 100 items of a product by a company is 1,560 hours with a S.D of 90 hours.

Test the hypothesis that the mean lifetime of the product is 1,580 hours.

EXAMPLE:

The mean life time of a sample of 400 fluoresent light bulbs produced by a company is found to be 1600 hours with a S.D of 150 hours.

Test the hypothesis that the mean life time of the bulbs produced is higher than 1570 hours at $\alpha = 0.01$.

Example:- Discussion

It is claimed that the mean of the population is 67 gms at 5% level of significance. Mean obtained from a random sample of size 100 is 64 with SD 3. Validate the claim.

$H_0 : \mu = 67$

$\boxed{H_1 : \mu \neq 67}$

$\alpha = 5\%$

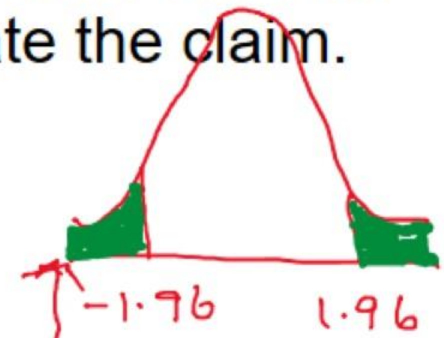
population

$\boxed{\mu = 64}$

$\boxed{n = 100 \quad \bar{x} = 64 \quad s = 3}$

sample

$$z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} = \frac{64 - 67}{3/\sqrt{100}} = \boxed{-10}$$



So H_0 is rejected

Example

There is an assumption that there is no significant difference between boys and girls wrt intelligence. Tests are conducted on two groups and the following are the observations

	Mean.	SD.	Size
Girls.	75.	8.	60
Boys.	73.	10.	100

Validate this at 5% LoS

Example

An automobile tyre manufacturer claims that the average life of a particular grade of tyre is more than 20,000 km. A random sample of 16 tyres is having mean 22,000 km with a standard deviation of 5000 km.

Validate the claim of the manufacturer at 5% LoS.

Example

As a manager of finance you are assigned the task of choosing the best product in terms of life of the product.

Product A: Mean 1456 hours with SD 423, size 10

Product B: Mean 1280 hours with SD 398, size 17

Use 5% level of significance.

Example :-

From the data available, it is observed that 400 out of 850 customers purchased the groceries online.

Can we say that most of the customers are moving towards online shopping even for groceries?

Example:-

It is found 290 errors in the randomly selected 400 lines of code from Team A; and 160 errors in 300 lines of code from Team B.

Can we assume that team B's performance is superior to that of team A : ?

Example:

Following is the record of number of accidents took place during the various days of the week.

Monday	Tuesday	wednesday	Thursday	Friday	Satday	Sunday
184	148	145	153	150	154	116

Can we conclude that accident is independent of the day in a week?

Example:

	with cancer	without cancer	total
smokers	400	300	700
Non-smokers	300	500	800
total	700	800	1500

Can we conclude that smoking causes cancer?

Example

- 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
- Test whether there is any significant difference in the means of two thermocouples at 5% level of significance

Thanks