BITS Pilani
Pilani | Dubai | Goa | Hyderabad

Introduction to Statistical Methods

ISM Team

innovate achieve lead

BITS Pilani
Pilani Campus

Session 9

Testing of Hypothesis-I

(Session 9: 25th / 26th January 2025)

IMP Note to Self

Hypothesis Testing

Contact Session 9: Module 4: Hypothesis Testing

Contact Session	List of Topic Title	Reference
CS - 9	Basics of testing of hypothesis: Z-test for difference in means and proportions of two populations.	T1 & T2

Testing of Hypothesis

Need for Testing of Hypothesis

Often the decisions are made based on sample estimates to generalize to a 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.

Need for Testing of Hypothesis

The following possibilities might arise due to sampling:

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

Unbiased

Due to chance

Sampling error

Real difference

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 be due to chance or sampling error (improper sampling technique used leads to sampling error).

Case(iii): If the difference is large, it may be 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**.

Testing of Hypothesis

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

Null Hypothesis	Alternative Hypothesis
Hypothesis of no difference or neutral or may be due to **Sampling variation**	Hypothesis of difference which is yet to be proved/ established

Testing of Hypothesis

Choose the best brand out of two available brands

Brand A Tyres

Brand B Tyres

The variable measured fuel consumption

Quantitative

Testing of Hypothesis

H_0 → The mean milage of Brand A and Brand B tyres may be same

$$H_0 : \mu_1 = \mu_2$$

Note: H_0 can be stated as one-tailed or two tailed, depending on scenario.

H_1 → The mean milage of Brand A may be less than Brand B tyres.

$$H_1 : \mu_1 < \mu_2$$

Testing of Hypothesis

H_1 → The mean milage of Brand A may be more than Brand B tyres

$$H_1 : \mu_1 > \mu_2$$

H_1 → The mean milage of Brand A and Brand B tyres may be different

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

Testing of Hypothesis

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

Testing of Hypothesis

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

???

Testing of Hypothesis

H_0

The proportion of false positive judgement between the two 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** the proportion of false positive judgement by Judge 2

$$H_1 : P_1 < P_2$$

Testing of Hypothesis

H_1

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

$$H_1 : P_1 \neq P_2$$

Testing of Hypothesis

It 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 statistical test

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

Parametric tests	Non-parametric tests
Based on the assumption of some probability distribution	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 → $n < 30$ (Generally)

Student's t-test →
- Unpaired t-Test
- Paired t-Test

Analysis of Variance →
- 1-way ANOVA
- 2-way 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

Testing of Hypothesis

Decision	Null Hypothesis (H_0)	
	True	False
Do not reject		
Reject		

Decision	Alternative Hypothesis (H_1)	
	False	True
Reject		
Do not reject		

Errors in decision making

Decision	Null Hypothesis (H_0)	
	True	False
Do not reject	Correct decision	Error in decision
Reject	Error in decision	Correct decision

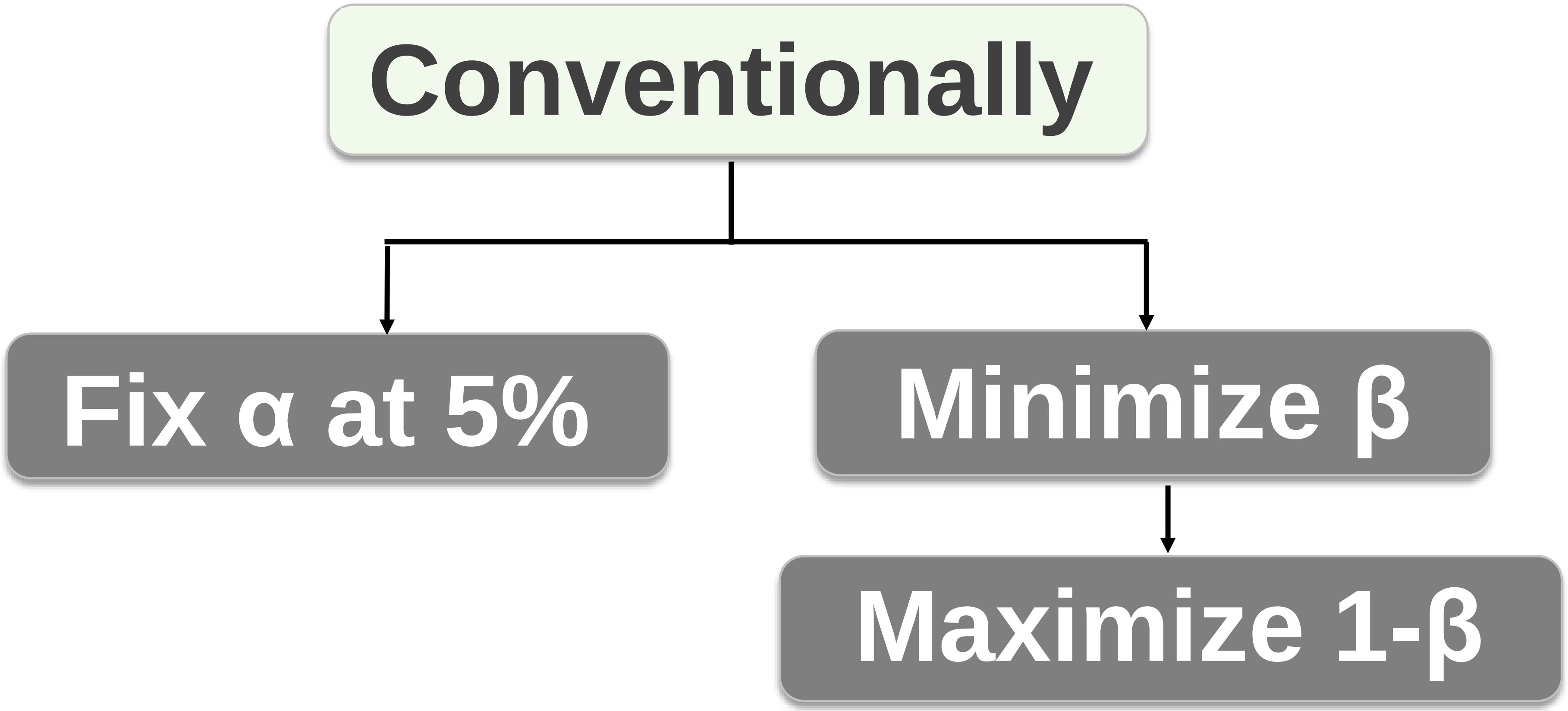
Decision	Alternative Hypothesis (H_1)	
	False	True
Reject	Correct decision	Type – II Error
Do not reject	Type - I Error	Correct decision

Errors in decision making

Decision	Null Hypothesis (H_0)	
	True	False
Do not reject	Confidence level (1-α)	β-error
Reject	α-error	Power 1-β

Decision	Alternative Hypothesis (H_1)	
	False	True
Reject		
Do not reject		Power 1-β

Decision on α – error and β – error



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

How do we take decision?

Based on the sample data the test-statistic should be computed using a specific Formula

Based on the computed test statistic, H0 will either not be rejected or rejected.

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 for means

- 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., $n \geq 30$)

Parametric tests

In practical situations most often we do not know the value of population variances. Hence, it is unlikely that the Z-test is applied for testing the mean of a single population or difference between two populations.

However, as a matter of procedure let us learn how it is applied in case of testing the difference between two population means.

Parametric tests: Z-test for testing ($\mu_1 - \mu_2$)

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

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

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \cong N(0,1)$$

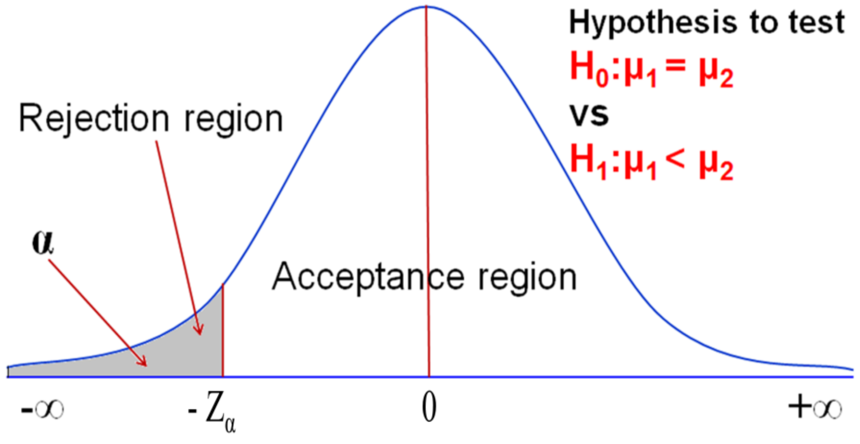
Note: Assuming population variances are known

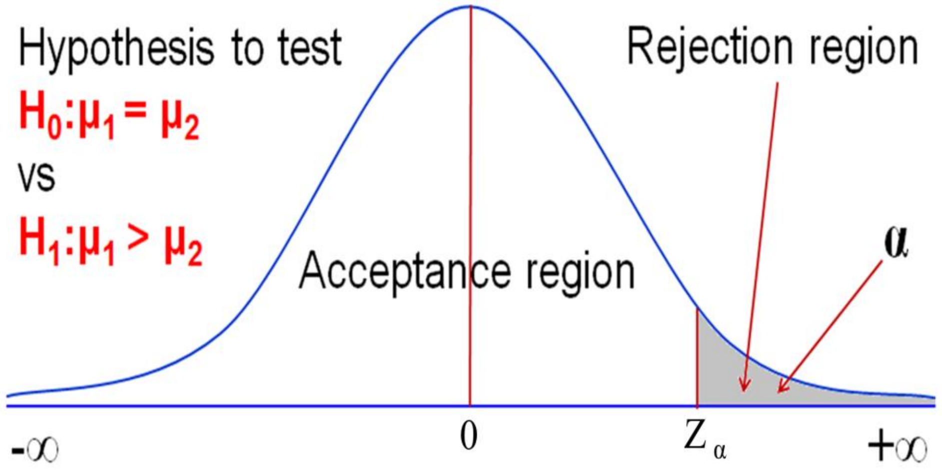
Critical region/ rejection criteria/ P - value

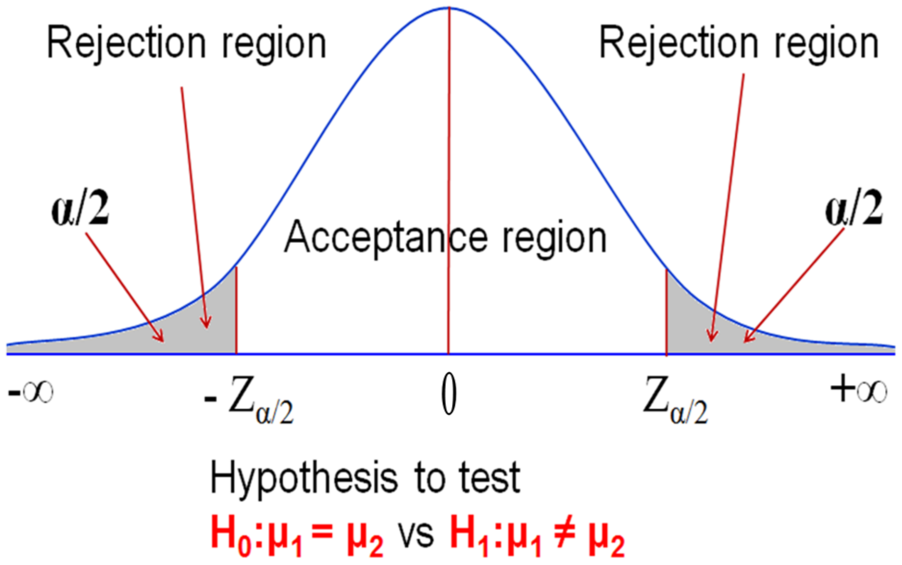
(i) Reject H_0, if computed value of Z is less than the critical value, i.e., **$P(Z<-z_\alpha)$**, otherwise, do not reject H_0.

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

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

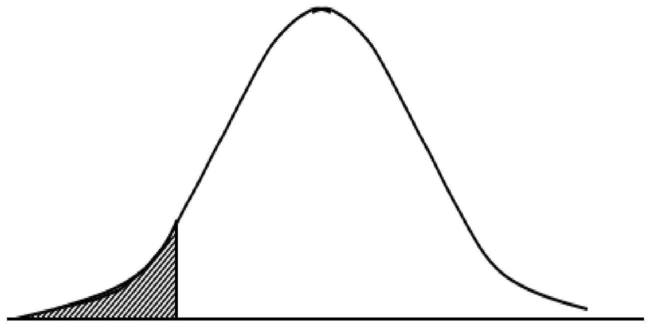
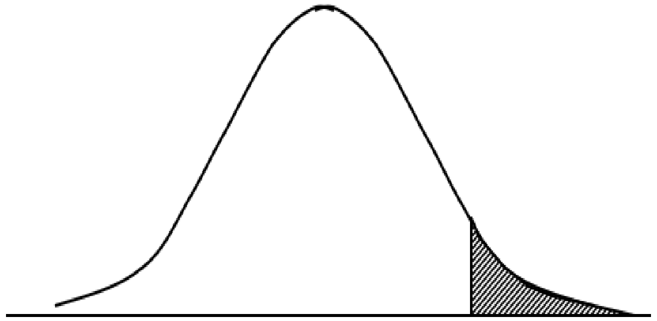
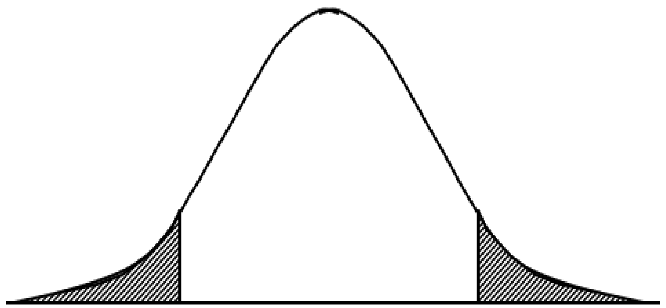






Parametric tests: Z-test for testing (μ_1 - μ_2)

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)
$H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 < \mu_2$	$H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 > \mu_2$	$H_0: \mu_1 = \mu_2$ vs $H_1: \mu_1 \neq \mu_2$
		

Parametric tests: Z-test for testing ($\mu_1 - \mu_2$)

innovate achieve lead

Example 1:

A taxi company is trying to decide whether to purchase brand A or brand B tyres for its fleet of taxis. To estimate difference in two brands, an experiment was conducted using 30 and 40 tyres of brand A and brand B, respectively. The tyres are run until they wear out.

Test at 5% level of significance, is there any significant difference between the two brands of tyres?

Parametric tests: Z-test for testing ($\mu_1 - \mu_2$)

Suppose some data is given below, now how will you decide to purchase the brand?

Brands	Sample size	Mean (kms)	SD (kms)
A	30	38600	5000
B	40	35450	6100

Parametric tests: Z-test for testing ($\mu_1 - \mu_2$)

innovate achieve lead

H_0 → The mean milage of Brand A and Brand B tyres may be same

$$H_0 : \mu_1 = \mu_2$$

H_1 → The mean milage of Brand A and Brand B tyres may be different

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

Parametric tests

95% CI for $\mu_1 - \mu_2$ is (472.013, 5827.987)

At 5% (0.05) level of significance with critical value ± 1.96

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(38600 - 35450)}{\sqrt{\frac{5000^2}{30} + \frac{6100^2}{40}}} = 2.30$$

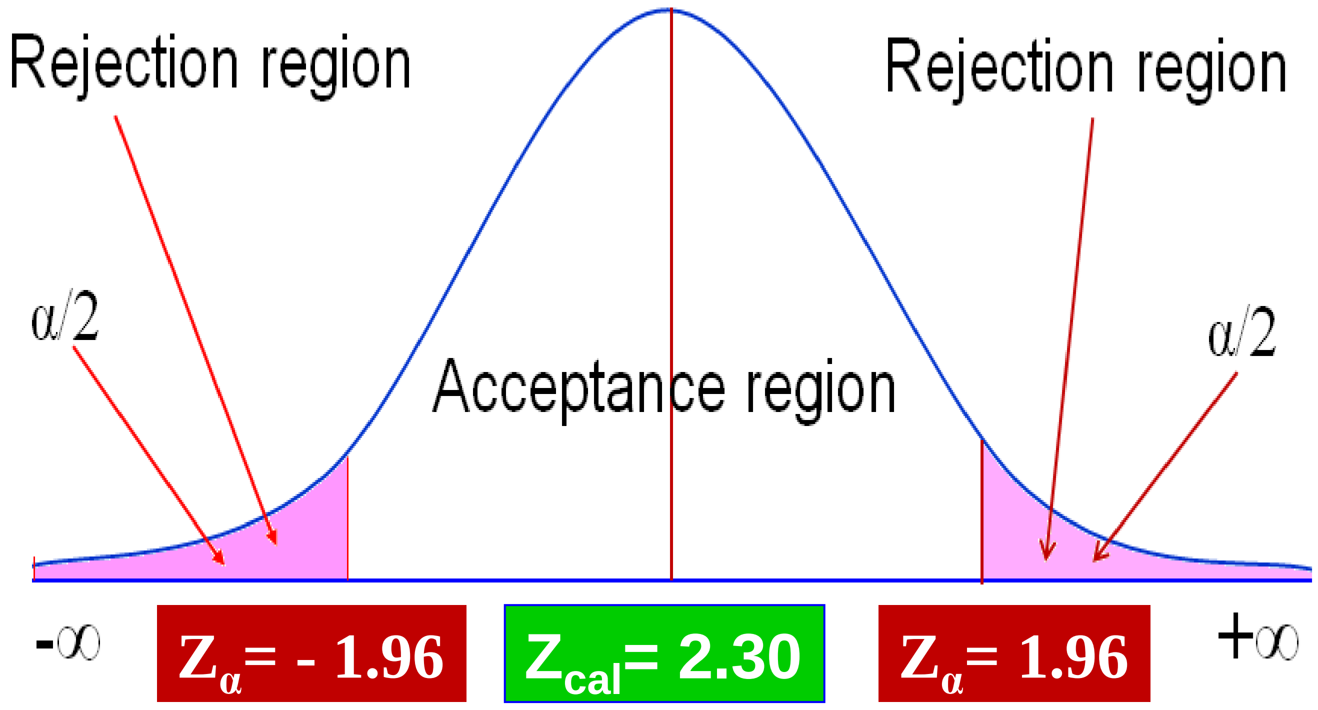
Hypothesis to test

$H_0: \mu_1 - \mu_2 = 0$

vs

$H_1: \mu_1 - \mu_2 \neq 0$

Critical value for α = 0.05 is 1.96. Since Z = 2.30 > 1.96, **Reject H_0**, Don't reject H_1.



Parametric tests

95% CI for $\mu_1 - \mu_2$ is (- 375.105, 6675.105)

At 1% (0.01) level of significance with critical value ± 2.58

Hypothesis to test

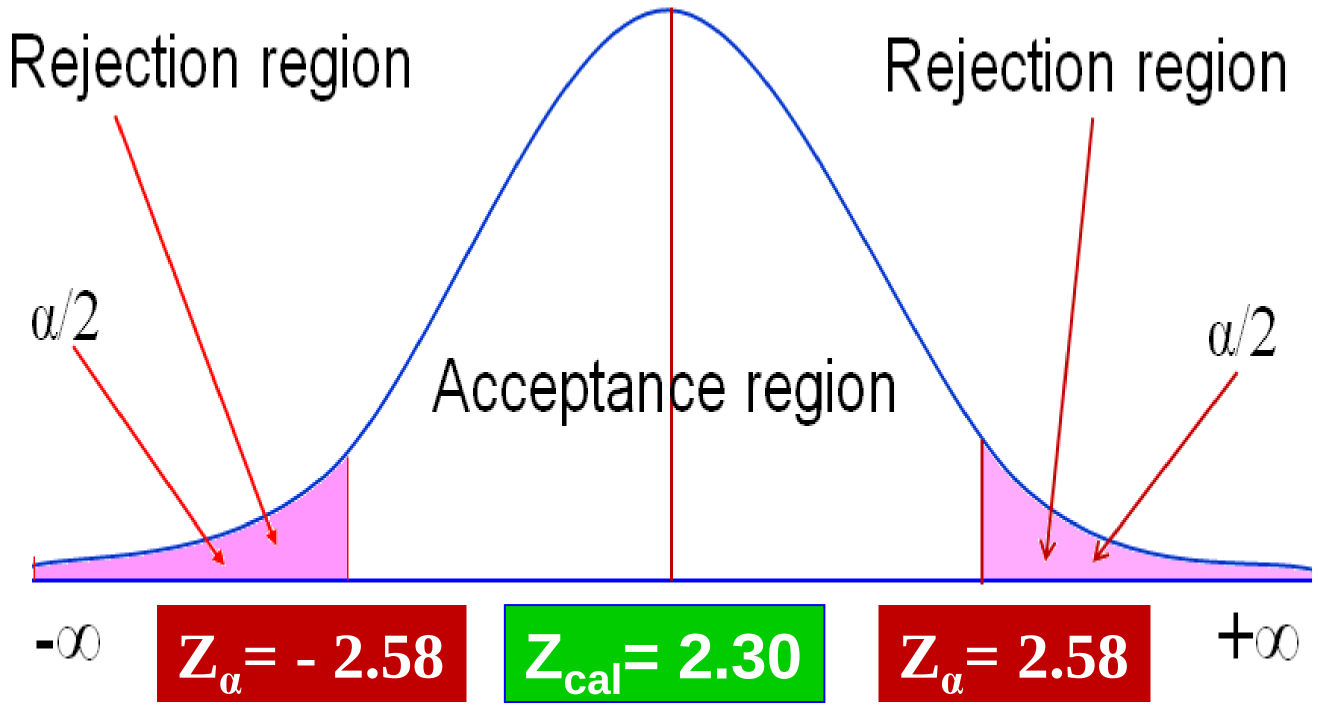
$H_0: \mu_1 - \mu_2 = 0$

vs

$H_1: \mu_1 - \mu_2 \neq 0$

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(38600 - 35450)}{\sqrt{\frac{5000^2}{30} + \frac{6100^2}{40}}} = 2.30$$

Critical value for α = 0.05 is 1.96. Since Z = 2.30 < 2.58, **Don't Reject H_0,** Reject H_1



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

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

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%,

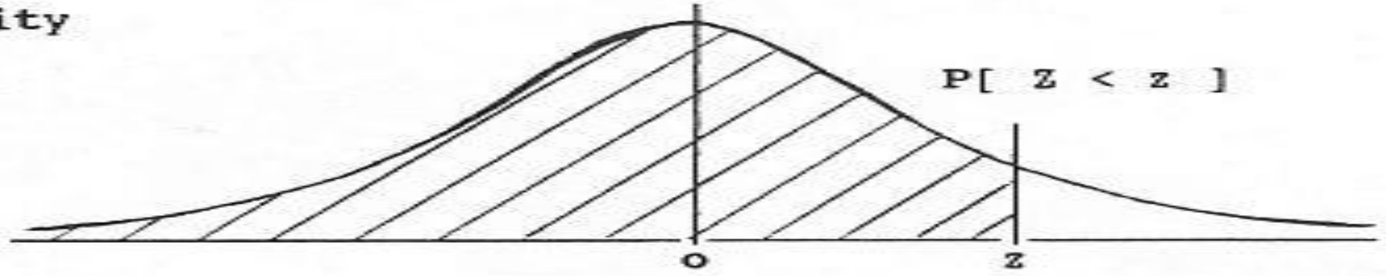
No evidence that supports the alternative hypothesis.

STANDARD STATISTICAL TABLES

1. Areas under the Normal Distribution

The table gives the cumulative probability up to the standardised normal value z
i.e.

$$P[\ Z < z\ ] = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp(-\tfrac{1}{2}Z^2)\ dZ$$



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5159	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7854
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8804	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9924	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

z	3.00	3.10	3.20	3.30	3.40	3.50	3.60	3.70	3.80	3.90
P	0.9986	0.9990	0.9993	0.9995	0.9997	0.9998	0.9998	0.9999	0.9999	1.0000

Parametric tests

95% CI for $\mu_1 - \mu_2$ is (472.013, 5827.987)

At 5% (0.05) level of significance with critical value ± 1.96

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} = \frac{(38600 - 35450)}{\sqrt{\frac{5000^2}{30} + \frac{6100^2}{40}}} = 2.30$$

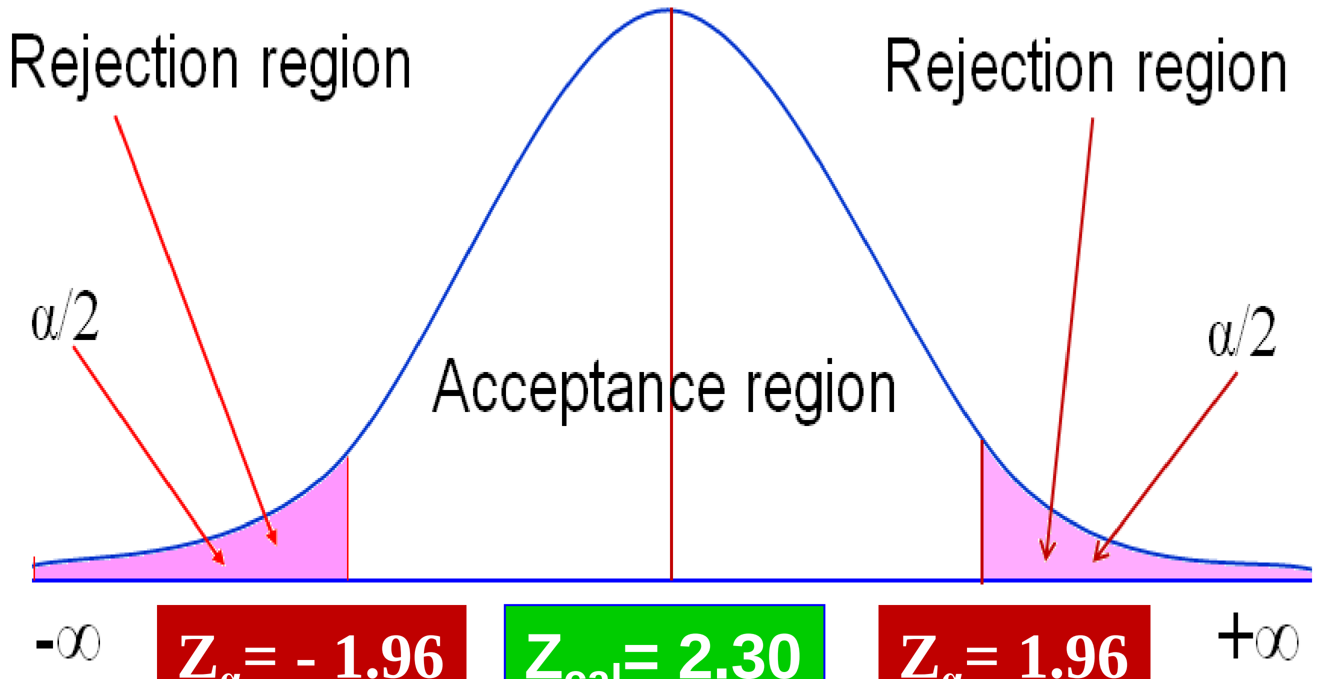
Hypothesis to test

$H_0: \mu_1 - \mu_2 = 0$

vs

$H_1: \mu_1 - \mu_2 \neq 0$

Critical value for α = 0.05 is 1.96. Since Z = 2.30 > 1.96, **Reject H_0**, Don't reject H_1



$Z_\alpha = -1.96$ | $Z_{cal} = 2.30$ | $Z_\alpha = 1.96$

P(Z ≥ 2.30) =0.0107
=> p-value=0.0214

Parametric tests: Z-test for testing (P_1 - P_2)

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

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

$$Z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{\frac{p_1 q_1}{n_1} + \frac{p_2 q_2}{n_2}}} \cong N(0, 1)$$

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

Note: Standard error of (p_1-p_2) depends on n_1 & n_2

Study the following examples

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

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

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

(ii) If the sample sizes are unequal, then SE (P_1-P_2) is calculated by

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

Study the following examples

Example 2:

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 whether Judge A has committed more false positivity in the judgement compared to the other Judge. Use $\alpha = 0.05$.

Judges	No. of defendants (n)	No. of false positives	False positive rate
1	3000	90	3.00%
2	2500	48	1.92%

Parametric tests

At 5% (0.05) level of significance with critical value ± 1.645 (for 1-tailed)

$$Z = \frac{(p_1 - p_2) - (P_1 - P_2)}{\sqrt{p * q\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{(0.030 - 0.019)}{\sqrt{0.025 * 0.975 * 00007}} = 2.55$$

Hypothesis to test

H_0: $\mu_1 = \mu_2$

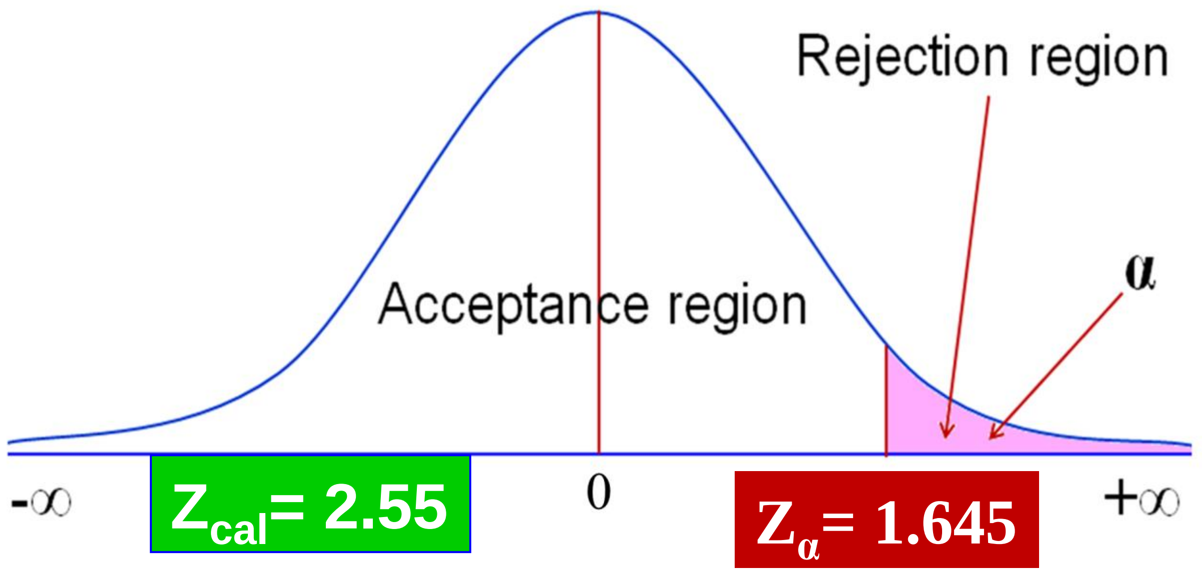
vs

H_1: $\mu_1 > \mu_2$

How to find CI for the one-tailed test?

Critical value for α = 0.05 is 1.645. Since Z = 2.55 > 1.645, **Reject H_0**, Don't reject H_1.

P(Z ≥ 2.55) =0.0054 (<0.05)



Parametric tests: Student's t – test

Background:

In the application of Z–test for testing the difference in means of two populations, an important assumption stated was that the population variances should be known.

But most often, population variances are unknown as the population itself may be unknown or infinite. Hence, as an alternative to Z – test for means, the Student's t – test will be applied.

Parametric tests: Student’s t – test

The Student’s t – test is based on the sampling distribution of means.

Since the sample variance contains n – 1 number of observations, the degrees of freedom is required to be known for t – test.

Parametric tests: Student's t – test

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

Parametric tests: 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

Testing the mean of a single population

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.**

Testing the mean of a single population

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

Testing the mean of a single population

innovate achieve lead

5 Define the critical region/ rejection criteria

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

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

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.

6 Conclusion

Testing the mean of a single population

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, i.e., $P(t < -t_{\alpha/2})$ or $P(t > t_{\alpha/2})$, otherwise do not reject H_0 (two tailed)

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.

6 Conclusion

Testing the mean of a single population

It is claimed that the average time spent on social media by teenagers is 3 hours per day. A research organization believes that the average time is probably higher. To check this claim, the organization surveyed a randomly selected sample of 10 teenagers, resulting in a sample mean of 3.3 hours per day with a sample standard deviation of 0.5 hours.

(a) 5% level of significance

(b) 2.5% level of significance

Testing the mean of a single population

H_0

The average time spent on social media as claimed and the actual average time may be same

$$H_0 : \mu = 3$$

H_1

The average time spent on social media may be higher than the actual average time

$$H_1 : \mu > 3$$

Testing the mean of a single population

(a) At 5% level of significance with critical value 1.833

$$|t| = \frac{|3.3-3|}{\frac{0.5}{\sqrt{10}}} = 1.897$$

Hypothesis to test

$H_0: \mu = 3$ vs $H_1: \mu > 3$

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

Since $|t| = 1.897 > 1.833$, Reject H_0 and Accept H_1

Testing the mean of a single population

innovate achieve lead

(b) At 2.5% level of significance with critical value 2.262

$$|t| = \frac{|3.3 - 3|}{\frac{0.5}{\sqrt{10}}} = 1.897$$

Hypothesis to test

H_0:μ=3 vs H_1:μ>3

Critical value for α = 0.01 is **2.262** for 9 degree of freedom

Since | t | =**1.897 < 2.262**, Don't accept H_0 and reject H_1

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Manual method of finding P – value

α	0.1	0.05	0.025
1	3.078	6.314	12.076
2	1.886	2.920	4.303
3	1.638	2.353	3.182
4	1.533	2.132	2.776
5	1.476	2.015	2.571
6	1.440	1.943	2.447
7	1.415	1.895	2.365
8	1.397	1.860	2.306
9	1.383	1.833	2.262
10	1.372	1.812	2.228

0.05 > P > 0.025

t_{obs}=1.897

$$\frac{2.262 - 1.833}{1.897 - 1.833} = \frac{0.05 - 0.025}{P - 0.025} \Rightarrow P = 0.0287$$

Testing the mean of a single population

Finding P - Value and Critical value		Excel code
Level of significance (α)	0.05	
Degrees of freedom	9	
t -statistic ($t_{(observed)}$)	1.897	
P - value (1-tailed)	0.029	T.DIST.RT(t, df)
Critical value (1-tailed)	1.833	T.INV(α, df)
P - value (2-tailed)	0.058	T.DIST.2T(t, df)
Critical value (2-tailed)	2.262	T.INV.2T(α, df)

Testing the mean of a single population

(b) At 5% level of significance with critical value 1.833

$$|t| = \frac{|3.3-3|}{\frac{0.5}{\sqrt{10}}} = 1.897$$

Hypothesis to test
$H_0:\mu=3$ vs $H_1:\mu>3$

Critical value for α = 0.05 is **1.833** for 9 degree of freedom

Since | t | =**1.897 > 1.833**, Don't Reject H_0 and Accept H_1.

P-value is ???
0.0287

Testing the deference between means of two population ($\mu_1 - \mu_2$)

If there are two independent populations whose population variances are unknown, then the choice of the test will be Independent sample t – test (Unpaired t – test).

Testing the deference between means of two population ($\mu_1 - \mu_2$)

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 should be 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.

Testing the deference between means of two population ($\mu_1 - \mu_2$)

innovate achieve lead

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

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

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \simeq t_{(n_1+n_2-2)}$$

Note: If sample sizes are unequal compute pooled SE

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

Testing the deference between means of two population ($\mu_1 - \mu_2$)

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 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 the means.

Testing the deference between means of two population ($\mu_1 - \mu_2$)

At 5% (0.05) level of significance with critical value 2.069 (dF=23)

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

Hypothesis to test

$H_0: \mu_1 - \mu_2 = 0$

vs

$H_1: \mu_1 - \mu_2 \neq 0$

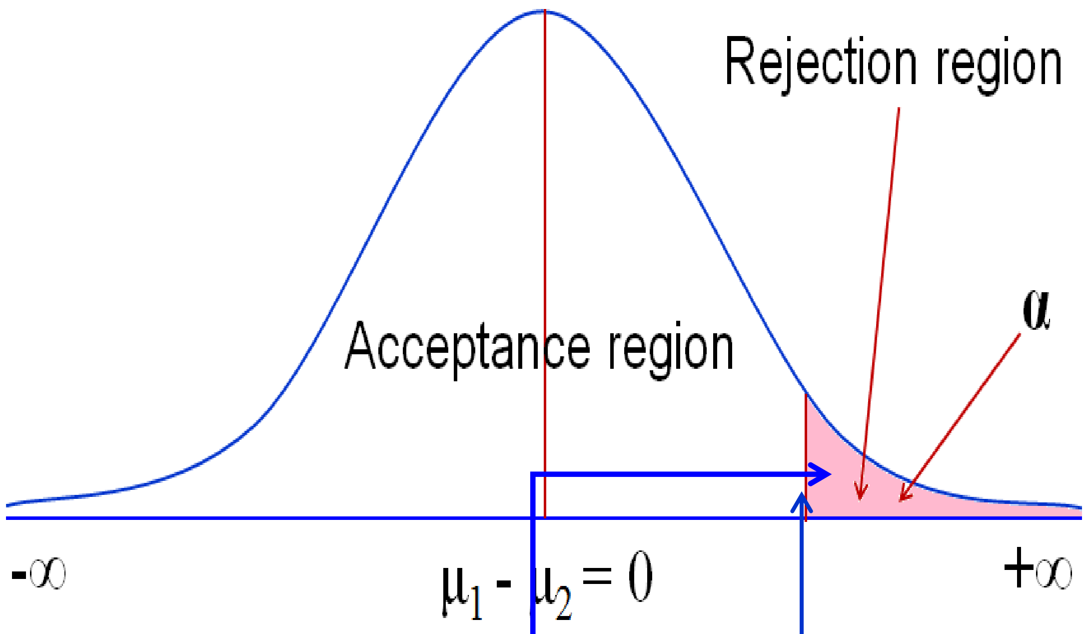
(Two tailed)

$$0.02 \le P \le 0.05$$

???

(μ_1 - μ_2) =0 not included in

95% CI for μ_1- μ_2 is (0.073, 1.907)



t_{cal}= 2.233

$t_{\alpha/2}$= 2.069

Since t = 2.233 > 2.069, Reject $H_{0,}$ Don’t reject H_1

Testing the deference between means of two population ($\mu_1 - \mu_2$)

Finding P - Value and Critical value		Excel code
Level of significance (α)	0.05	
Degrees of freedom	23	
t -statistic ($t_{(observed)}$)	2.233	
P - value (1-tailed)	0.018	T.DIST.RT(t, df)
Critical value (1-tailed)	1.714	T.INV(α, df)
P - value (2-tailed)	0.036	T.DIST.2T(t, df)
Critical value (2-tailed)	2.069	T.INV.2T(α, df)

Testing the deference between means of paired observations

If the two observations are made on the sample, it forms paired observations which are related/ dependent, then in such cases, unpaired t – test is an ideal test instead Paired t – test should be applied.

Testing the difference between means of paired observations

t-test

Testing mean before and after observations of a single population (μ_d)

Assumptions

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

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

Subjects should be selected randomly

Testing the difference between means of paired observations

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

Testing the difference between means of paired observations

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

Test, at 5% level of significance, if there is any difference in the test score after training

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

Testing the difference between means of paired observations

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

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

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

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

Testing the difference between means of paired observations

At 5% (0.05) level of significance the critical value is 1.833

$$|t| = \frac{\bar{d} - \mu_d}{\frac{S_d}{\sqrt{n}}} = \frac{5-0}{\frac{9.21}{\sqrt{9}}} = 1.629$$

Hypothesis to test

$H_0: \mu_d = 0$

vs

$H_1: \mu_d > 0$

95% CI for μ_d is (- 0.627, 10.627) includes 0

???

95% CI for μ_d is (- 0.627, 10.627)

P-value is

$$P = 0.073$$

Critical value for α = 0.05 is 1.833. Since | t | = 1.629 < 1.833, Don't reject H_0 & Reject H_1

Testing the difference between means of paired observations

Finding P - Value and Critical value		Excel code
Level of significance (α)	0.05	
Degrees of freedom	8	
t -statistic ($t_{(observed)}$)	1.629	
P - value (1-tailed)	0.071	T.DIST.RT(t, df)
Critical value (1-tailed)	1.860	T.INV(α, df)
P - value (2-tailed)	0.142	T.DIST.2T(t, df)
Critical value (2-tailed)	2.306	T.INV.2T(α, df)

Chi-square test

Chi-square distribution is a sampling distribution of variance of a single population. A test based on this will be called Chi-square test.

The objective of Chi-square test is to test

(a) Independence between two categorical variables

(b) Goodness-of-fit

In both cases difference between observed and expected frequency will be tested to find out whether it is due to chance or not.

Chi-square test

Chi-square test should be applied only for frequencies. We cannot use them for data in terms of percentages, proportions, means or similar statistical contents.

Chi-squared is generally applied for nominal rather than ordinal data.

The data will be presented in an r x c contingency table, where r is the number of rows and c is the number of columns. If r = 2 and c = 2, then it is called 2 x 2 contingency table. In Machine learning it is called as Confusion matrix.

Chi-square test

Chi-square test measures the relationship between two categorical variables, but as this relation may be based on chance, it is appropriate to call it as **ASSOCIATION** rather than RELATION

Chi-square test

It is very important to verify the 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

The Fisher's exact Probability

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

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 and columns meaningfully

Chi-square test

An example of a 2 x 2 contingency table

<table>
<tr><td rowspan="2">Categorical Variable 1</td><td colspan="2">Categorical variable 2</td><td rowspan="2">Total</td></tr>
<tr><td>Response 1</td><td>Response 2</td></tr>
<tr><td>Response 1</td><td>O_1</td><td>O_2</td><td>r_1</td></tr>
<tr><td>Response 2</td><td>O_3</td><td>O_4</td><td>r_2</td></tr>
<tr><td>Total</td><td>c_1</td><td>c_2</td><td>n</td></tr>
</table>

O_i's are called observed frequency.

Chi-square test

2 x 2 Contingency Table for testing independence

Categorical Variable 1	Categorical variable 2		Total
	Present	Absent	
Present	O_1 E_1	O_2 E_2	r_1
Absent	O_3 E_3	O_4 E_4	r_2
Total	c_1	c_2	n

Calculation of expected frequencies

$$E_1 = \frac{r_1 c_1}{n}$$

$$E_2 = \frac{r_1 c_2}{n}$$

$$E_3 = \frac{r_2 c_1}{n}$$

$$E_4 = \frac{r_2 c_2}{n}$$

Chi-square test

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)

Observed = Expected

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

Observed ≠ Expected

Chi-square test

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

$$\chi^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

$$E_{ij} = \frac{r_i c_j}{n},$$

$$\text{for } i = 1, 2, \ldots, m; \; j = 1, 2, \ldots, n$$

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

Chi-square test

Example:

An RTO department has observed that, on a highway, a greater number of accidents have taken place in the early hours of the day than other timings and wish to associate the outcome of the accidents with timings. A survey findings show that, of the 400 accident cases studied, 280 had met with accident in the early hours and 99 of them were fatal. Further, those who met with accident in the early hours and died was 80. Does this data indicate any association between the time of accident and fatality of the accident. Use $\alpha = 0.05$.

Chi-square test

Hypothesis

H_0: Timings of accident and its outcome are independent (Not associated)

H_1: Timings of accident and its outcome are dependent (Associated)

Chi-square test

Early hours of accident	Outcome of accident		Total
	Fatal	Non-fatal	
Yes	80	200	280
No	19	101	120
Total	99	301	400

$$E_1 = \frac{r_1 c_1}{n} = \frac{280 * 99}{400} = 69.3$$

$$E_3 = \frac{r_2 c_1}{n} = \frac{120 * 99}{400} = 29.7$$

$$E_2 = \frac{r_1 c_2}{n} = \frac{280 * 301}{400} = 210.7$$

$$E_4 = \frac{r_2 c_2}{n} = \frac{120 * 301}{400} = 90.3$$

Chi-square test

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	80	69.3	10.7	114.49	1.652
2	200	210.7	-10.7	114.49	0.543
3	19	29.7	-10.7	114.49	3.855
4	101	90.3	10.7	114.49	1.268
Total	400	400	Chi-square statistic		$\chi^2 = 7.318$

Chi-square test

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

Chi-square test

Hypothesis

H_0: Timings of accident and its outcome are independent (Not associated)

H_1: Timings of accident and its outcome are dependent (Associated)

$\chi^2 = 7.31$

df = 1, Critical value at $\alpha = 0.05$ is 3.841, P = 0.007

Inference: There may be an association between timings of accident and its outcome

Chi-square test

In an experiment to study the dependence of hypertension on smoking habits, the following data were taken on 180 individuals

Hypertension	Non smokers	Moderate smokers	Heavy smokers	Total
Yes	21	36	30	87
No	48	26	19	93
Total	69	62	49	180

Chi-square test

Hypothesis

H_0: Smoking habit and hypertension may be independent (may not be associated)

H_1: Smoking habit and hypertension may not be independent (may be associated)

Chi-square test

Calculation of expected frequencies

- $E_1 = (r_1 \text{ x } c_1)/n = (87 \text{ x } 69)/180 = 33.35$
- $E_2 = (r_1 \text{ x } c_2)/n = (87 \text{ x } 62)/180 = 29.97$
- $E_3 = (r_1 \text{ x } c_3)/n = (87 \text{ x } 49)/180 = 23.68$
- $E4 = (r_2 \text{ x } c_1)/n = (93 \text{ x } 69)/180 = 35.65$
- $E_5 = (r_2 \text{ x } c_2)/n = (93 \text{ x } 62)/180 = 32.03$
- $E_6 = (r_2 \text{ x } c_3)/n = (93 \text{ x } 49)/180 = 25.32$

Chi-square test

Calculation of Chi-square statistic

Sl No	(O_i)	(E_i)	$(O_i - E_i)$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	21	33.35	- 12.35	152.52	4.57
2	36	29.97	6.03	36.36	1.21
3	30	23.68	6.32	39.94	1.69
4	48	35.65	12.35	152.52	4.28
5	26	32.03	- 6.03	36.36	1.14
6	19	25.32	- 6.32	39.94	1.58
Total	180	180	Chi-square statistic		$\chi^2 = 14.46$

Chi-square test

Interpretation

H_0: Smoking habit and hypertension may be independent (may not be associated)

H_1: Smoking habit and hypertension may not be independent (may be associated)

$\chi^2 = 14.46$

df = 2, Critical value at $\alpha = 0.05$ is 5.99, $P < 0.001$

Inference: There may be an association between smoking and Hypertension

Chi-square test for Goodness-of-fit

A powerful test for testing the significance of the discrepancy between theory and experiment was given by Karl Pearson known as “Chi-square test for Goodness – of – fit”. It enables to find the deviation of the experiment from theory is just by chance or is it really due to the inadequacy of the theory to fit the observed data.

Chi-square test for Goodness-of-fit

If O_i (i = 1, 2, …, n) is a set of observed (experimental) Frequencies and E_i (i = 1, 2, …, n) is the corresponding set of expected (theoretical or hypothetical) frequencies, then the Chi-square test statistic is given by

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} \approx \chi^2_{(n-1)}$$

follows Chi-square distribution with n – 1 degree of freedom.

Chi-square test for Goodness-of-fit

Example

The following data shows the distribution of digits in numbers chosen at random from a telephone directory. The digits are:

Digits	0	1	2	3	4	5	6	7	8	9	Total
f	1026	1107	997	966	1075	933	1107	972	954	853	10000

H_0: The digits occur uniformly frequently in the directory

H_1: The digits do not occur uniformly frequently in the directory

Chi-square test for Goodness-of-fit

Under the null hypothesis, the expected frequency for each of the digits 0, 1, …, 9 is 10000 ÷ 10 = 1000. The Chi-square value is

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i} = \frac{(1026 - 1000)^2}{1000} + \ldots + \frac{(853 - 1000)^2}{1000}$$

$$\chi^2 = 58.542$$

Since $\chi^2 = 58.542 < 16.919$ (critical value, df = 9), it can be infer that the digits are not uniformly distributed.

Chi-square test for Goodness-of-fit

A consultant was employed by a city council to study the pattern of bus arrival and departure at a very busy interstate bus terminus. She collected data from the arrival of 200 buses. Based on the data, the average arrival time was found to be $\lambda = 2.96$. She divided the arrivals into 6 categories. Assuming that the arrivals follow Poisson distribution test whether the arrival distribution follows Poisson law. Use $\alpha = 0.01$.

Chi-square test for Goodness-of-fit

innovate achieve lead

The probabilities are to be calculated using Poisson distribution with λ = 2.96 and x = 0, 1, 2, 3, 4, and ≥ 5. The results are as follows:

No. of arrivals	O_i	p_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$
0	10	0.0524	10.48	0.0220
1	13	0.1545	30.90	10.3693
2	45	0.2277	45.54	0.0064
3	49	0.2238	44.76	0.4016
4	32	0.1651	33.02	0.0315
≥ 5	41	0.1765	35.30	0.9204
			χ^2=	11.7512

Since χ^2 = 3.402 < 13.27 (critical value at df = k-2 = 4), it can be inferred that the arrivals and departures follow Poisson law.

Exercises 1

Aircrew escape system are powered by a solid propellant. The burning rate of this propellant is an important product characteristic. Specification require that the mean burning rate must be 50 cm/s. From past experience it is known that the population SD is 2 cm/s. A sample of 25 solid propellant were selected randomly to reconfirm the specification stated. The sample mean found was 51.3 cm/s. At 5% level of significance what conclusion should be drawn regarding burning rate?

Hint: $\mu=50$, $\sigma=2$, $n=25$, $\bar{X}=51.3$, $H_0: \mu=50$, $H_1: \mu\neq50$, (Two tailed test),

$Z = (\bar{X} - \mu)/(\sigma/5)$, $Z_{critical} = \pm 1.96$

Exercises 2

A builder claims that solar water heater are installed in 70% of all homes being constructed today in a city. Would you agree with this claim if a random sample of new homes in this city shows that 28 out of 55 had heat pumps installed? What P-value and confidence interval are related in this situation?

Hint: H_0 : p=0.7

H_1: p ≠0.7.

$\hat{p}$=28/55, SE=$\sqrt{pq/n}$, choose 95% confidence interval for a two tailed test.

CI= $\hat{p} \pm \mathrm{Z} \times \mathrm{SE}$

Exercises 3

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.

Hint: n1=200, n2=100, x1=42, x2=18,

Null hypothesis: p1-p2=0.08, alternative hypothesis: p1-p2 ≠0.08,

Find pooled SE where, p=0.2 and q=0.8,

CI=0.03 $\pm$ 0.096

Exercises 4

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.

Hint: $\alpha=0.01$, one tailed

H0: $\mu \geq 7$, H1: $\mu < 7$, n=15 (<30 => t-test), dF=15-1=14, $\bar{x} = 6.26$, s=1.91, p-value= 0.075 (approx value via t table between 0.05 to 0.10 at $|t|=1.5$, df=14)

Exercises 5

A taxi company manager is trying to decide whether the use of radial tires instead of regular belted tires improves fuel economy. 12 cars were equipped with radial tires and driven over a prescribed test course. Without changing drivers, the same cars were then equipped with regular belted tires and driven once again over the test course. The gasoline consumption in kilometers per liter was recorded. Assume that the populations are normally distributed

Radial tyres	4.2	5.7	2.6	7.0	6.7	4.5	5.7	5.1	7.4	6.9	6.1	3.2
Belted tyres	4.3	3.9	5.2	5.9	5.8	4.4	4.7	5.8	5.9	3.7	4.9	4.9

- (i) What are the underlying assumptions of Student's unpaired t-test?
- (ii) Based on the data can we conclude that cars equipped with radial tires give better fuel economy than those equipped with belted tires?
- (iii) Find P-value.
- (iv) Construct 95% CI for difference in means.
- (v) What is your inference?

Exercises 5: Hints

H_0: There is no difference in fuel economy between radial and regular belted tires.

H_1: Radial tires give better fuel economy than belted tires.

dF=11,

One tailed test with 95% CI

$\Rightarrow \alpha=0.05$ and $t(0.05, 11)=1.796$

Calculate t using student's unpaired t-test and compare with 1.796 to accept or reject the null hypothesis.

Exercises 6

Three different analytical tests can be used to determine the impurity level in steel alloys. Each specimen is tested using two procedures and the results are shown in the following tabulation. Is there sufficient evidence to conclude that the three tests give the same mean impurity level using $\alpha = 0.01$

Specimen	1	2	3	4	5	6	7	8
Test 1	1.2	1.3	1.5	1.4	1.7	1.3	1.6	
Test 2	1.4	1.7	1.5	1.3	2.0	2.1	1.7	1.9

Hint: H_0: Mean impurity level for all 3 tests are same.

H_1: Mean impurity level for atleast one test is different.

Find mean and variance for both data, and compute the test statistic (t-test), consider df=(15-1)+(6-1)=19.

Exercises 7

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.

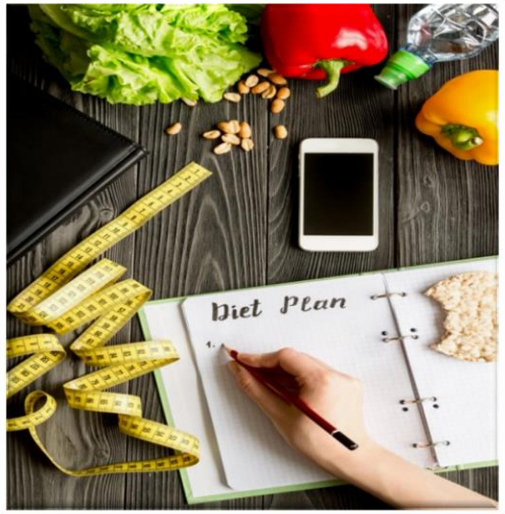
(a) Construct 95% CI for difference in means

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

(c) Find the P-value

Exercises 8

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?

Assume that the difference between before and after follow normal distribution. Use $\alpha = 0.05$.

Exercises 9

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?

Exercises 10

Three pension plans

Independent of job classification

Use **α = 0.05**

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

Examine whether there is any association between job classification and pension plan

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

Exercises 11

Jaswant is interested in breeding flowers of a certain species. The experimental breeding can result in four possible types of flowers

(a) Megenta flowers with green stigma (MG)

(b) Megenta flowers with red stigma (MR)

(c) Red flowers with green stigma (RG)

(d) Red flowers with red stigma (RR)

Exercises 11 (contd)

According to Mendel's law, these four kinds of flowers should come out in the ratio of 9 : 3 : 3 : 1. Jaswant found that under her experiment, out of 160 flowers that bloomed the number of flowers with types MG, MR, RG, and RR were 84, 35, 28, and 13. She wants to find out, whether these data are compatible with Mendel's law. Use $\alpha = 0.05$.

Exercises 12

A chemical company wishes to know if its sales of a liquid chemical are normally distributed. This information will help them in planning and controlling the inventory. The sales record for a random sample of 200 days are as follow: Using α 5% level test whether the company's sales normally distributed. The sample mean and sample standard deviation are 40 and 2.5 respectively.

Sales in '000 liters	< 34	34.0 - 35.5	35.5 - 37.0	37.0 - 38.5	38.5 - 40.0	40.0 - 41.5	41.5 - 43.0	43.0 - 44.5	44.5 - 46.0	> 46
No. of days	0	13	20	35	43	51	27	10	1	0

Glossary

Alpha The probability of a Type I error or the level of significance. Its symbol is the Greek letter α

Alternate hypothesis The conclusion we accept when we demonstrate that the null hypothesis is false. It is also called the research hypothesis.

Critical value A value that is the dividing point between the region where the null hypothesis is not rejected and the region where it is rejected.

Degrees of freedom The number of items in a sample that are free to vary. Suppose there are two items in a sample, and we know the mean. We are free to specify only one of the two values, because the other value is automatically determined (since the two values total twice the mean

Hypothesis A statement or claim about the value of a population parameter.

Hypothesis testing A statistical procedure, based on sample evidence and probability theory, used to determine whether the statement about the population parameter is reasonable.

Null hypothesis A statement about the value of a population parameter that is developed for testing in the face of numerical evidence. It is written as $H0$.

One-tailed test Used when the alternate hypothesis states a direction, such as read “the population mean is greater than 40.” Here the rejection region is only in one tail (the right tail).

Proportion A fraction or percentage of a sample or a population having a particular trait. If 5 out of 50 in a sample liked a new cereal, the proportion is 550, or .10.

***p*-value** The probability of computing a value of the test statistic at least as extreme as the one found in the sample data when the null hypothesis is true.

Significance level The probability of rejecting the null hypothesis when it is true.

Two-tailed test Used when the alternate hypothesis does not state a direction, such as read “the population mean is not equal to 75.” There is a region of rejection in each tail.

Type I error Occurs when a true is rejected.

Type II error Occurs when a false is accepted.

Thank You