

SNS COLLEGE OF TECHNOLOGY

Kurumbapalayam (Po), Coimbatore – 641 035

An Autonomous Institution

Accredited by NAAC – UGC with ‘A++’ Grade

Approved by AICTE, New Delhi & Affiliated to Anna University, Chennai



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

**COURSE NAME: 23ADO201- DATA SCIENCE FUNDAMENTALS
(OPEN ELECTIVE)**

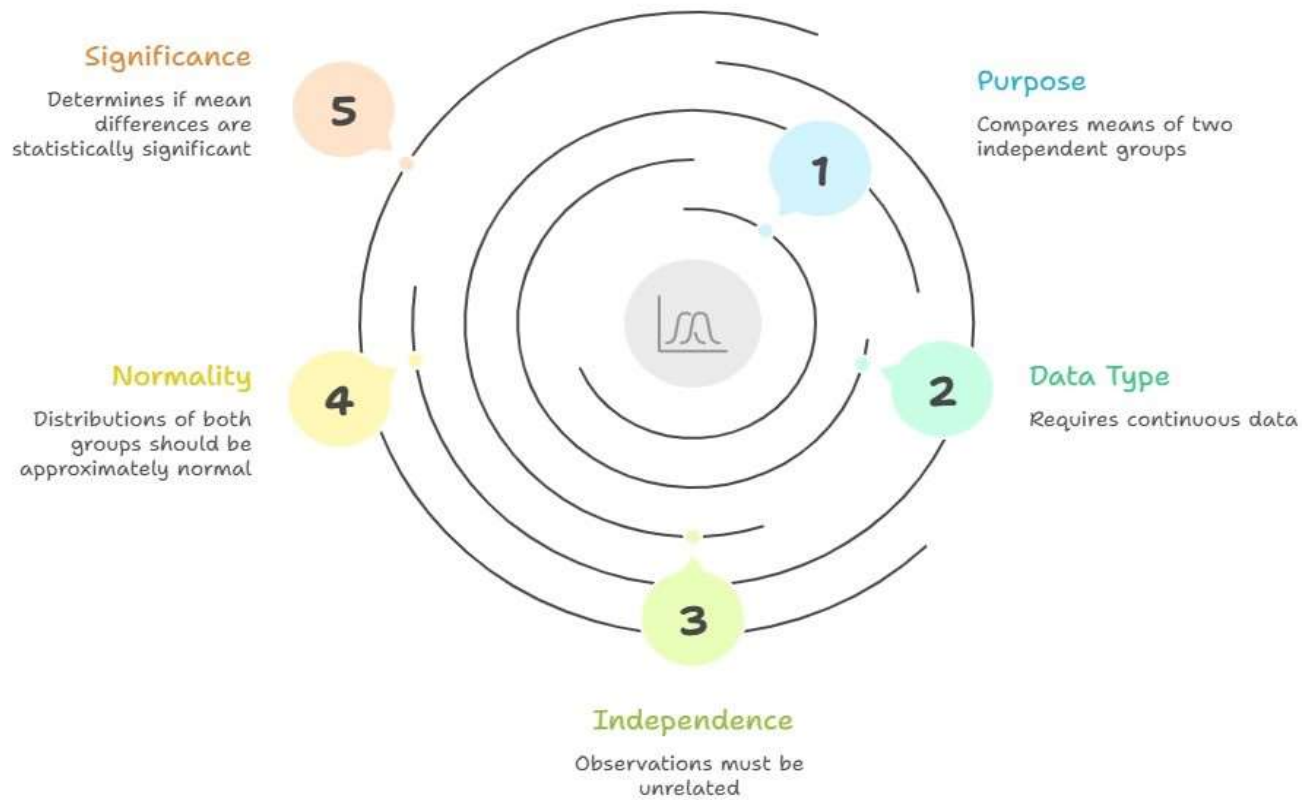
III YEAR / VI SEMESTER

1

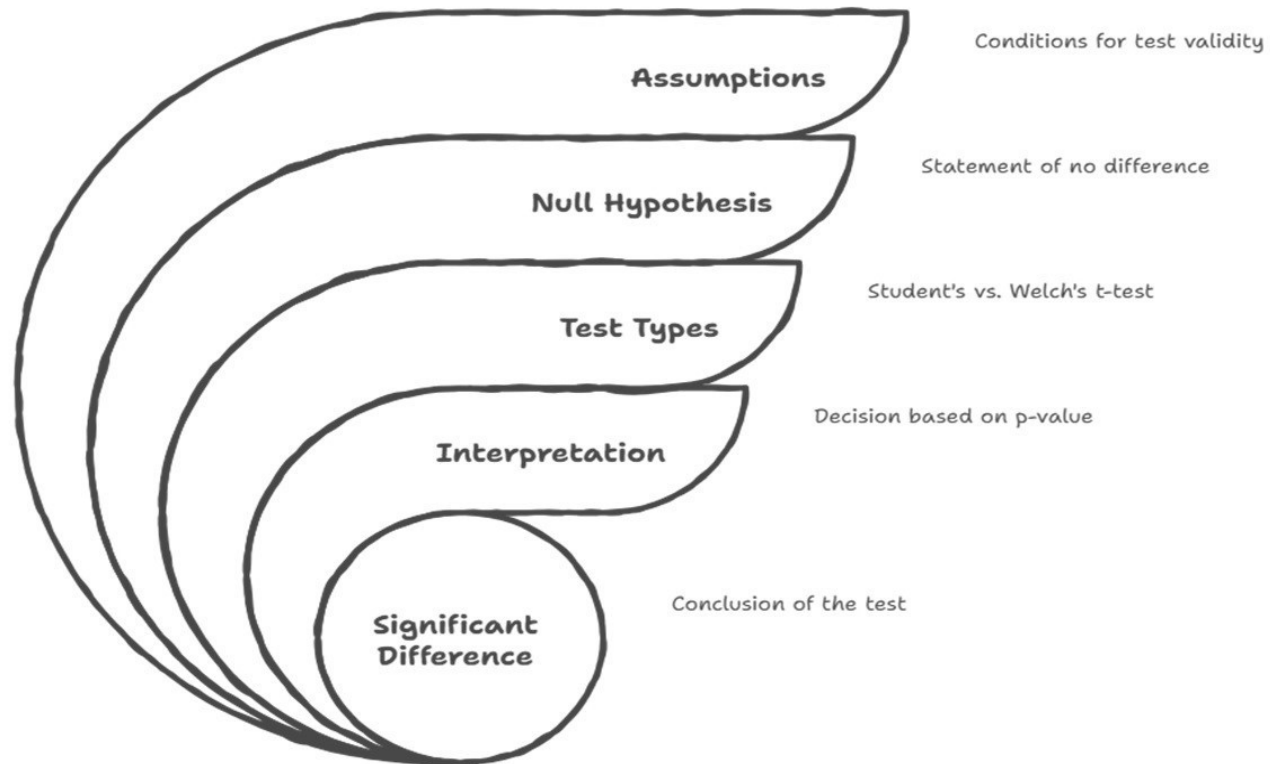
Unit 4 - ANALYSIS OF VARIANCE

Topic : T-test for two independent samples

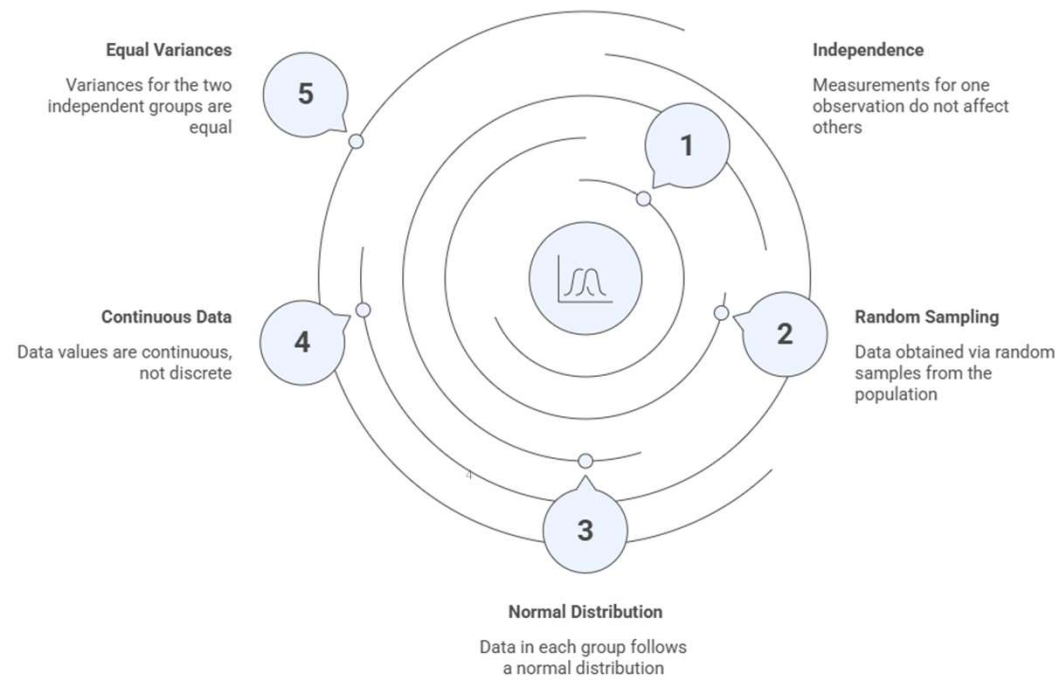
Understanding the Independent Two-Sample T-Test



Independent T-Test Structure



Assumptions for a Two-Sample *t*-Test



Made with Napkin

Two-sample *t*-test example

One way to measure a person's fitness is to measure their body fat percentage. Average body fat percentages vary by age, but according to some guidelines, the normal range for men is 15-20% body fat, and the normal range for women is 20-25% body fat.

Our sample data is from a group of men and women who did workouts at a gym three times a week for a year. Then, their trainer measured the body fat. The table below shows the data.

Table 1: Body fat percentage data grouped by gender

up	Body Fat Percentages				
Men	13.3	6.0	20.0	8.0	14.0
	19.0	18.0	25.0	16.0	24.0
	15.0	1.0	15.0		5
Women	22.0	16.0	21.7	21.0	30.0
	26.0	12.0	23.2	28.0	23.0

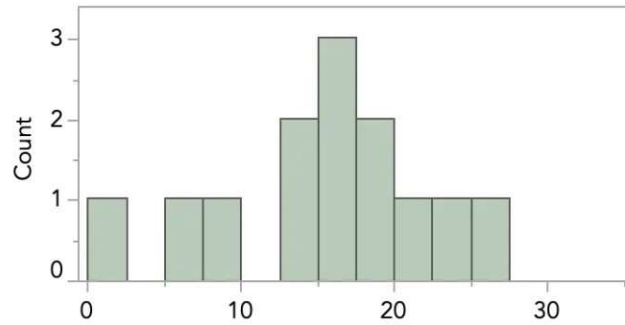
Checking the data

Let's start by answering: Is the two-sample *t*-test an appropriate method to evaluate the difference in body fat between men and women?

- The data values are independent. The body fat for any one person does not depend on the body fat for another person.
- We assume the people measured represent a simple random sample from the population of members of the gym.
- We assume the data are normally distributed, and we can check this assumption.
- The data values are body fat measurements. The measurements are continuous.
- We assume the variances for men and women are equal, and we can check this assumption.

Group=Men

Body Fat Percentage

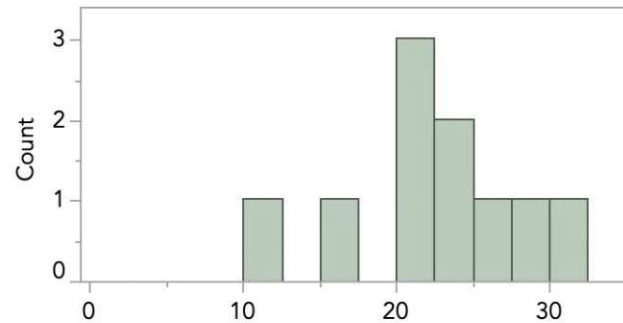


Summary Statistics

Mean	14.95
Std Dev	6.84
Std Err Mean	1.90
Upper 95% Mean	19.08
Lower 95% Mean	10.81
N	13.00

Group=Women

Body Fat Percentage



Summary Statistics

Mean	22.29
Std Dev	5.32
Std Err Mean	1.68
Upper 95% Mean	26.10
Lower 95% Mean	18.48
N	10.00

4/8/2026

4/8/2026

How to perform the two-sample t -test

For each group, we need the average, standard deviation and sample size. These are shown in the table below.

Table 2: Average, standard deviation and sample size statistics grouped by gender

up	Sample Size (n)	Average (\bar{X})	Standard deviation (s)
Women	10	22.29	5.32
Men	13	14.95	6.84

We'll further explain the principles underlying the two sample t -test in the statistical details section below, but let's first proceed through the steps from beginning to end. We start by calculating our test statistic. This calculation begins with finding the difference between the two averages:

$$22.29 - 14.95 = 7.34$$

This difference in our samples estimates the difference between the population means for the two groups.

Next, we calculate the pooled standard deviation. This builds a combined estimate of the overall standard deviation. The estimate adjusts for different group sizes. First, we calculate the pooled variance:

$$s_p^2 = \frac{((n_1-1)s_1^2) + ((n_2-1)s_2^2)}{n_1+n_2-2}$$

$$s_p^2 = \frac{((10-1)5.32^2) + ((13-1)6.84^2)}{(10+13-2)}$$

$$= \frac{(9 \times 28.30) + (12 \times 46.82)}{21}$$

$$= \frac{(254.7 + 561.85)}{21}$$

$$= \frac{816.55}{21} = 38.88$$

Next, we take the square root of the pooled variance to get the pooled standard deviation. This is:

$$\sqrt{38.88} = 6.24$$

We now have all the pieces for our test statistic. We have the difference of the averages, the pooled standard deviation and the sample sizes. We calculate our test statistic as follows:

$$t = \frac{\text{difference of group averages}}{\text{standard error of difference}} = \frac{7.34}{(6.24 \times \sqrt{(1/10 + 1/13)})} = \frac{7.34}{2.62} = 2.80$$

Activity



Activity 2: Step-by-Step Hypothesis Testing

Task:

Given data:

Group 1: 45, 50, 55, 60, 65

Group 2: 40, 42, 48, 52, 46

Steps:

- 1.State H_0 and H_1
- 2.Calculate mean and standard deviation
- 3.Compute t-value
- 4.Compare with table value
- 5.Draw conclusion

4/8/2026

8

MCQ



MCQs on T-Test for Two Independent Samples (with answers):

1. What is the main purpose of a two independent sample t-test?

- A) To compare means of the same group
- B) To compare means of two different groups
- C) To compare variances
- D) To test correlation

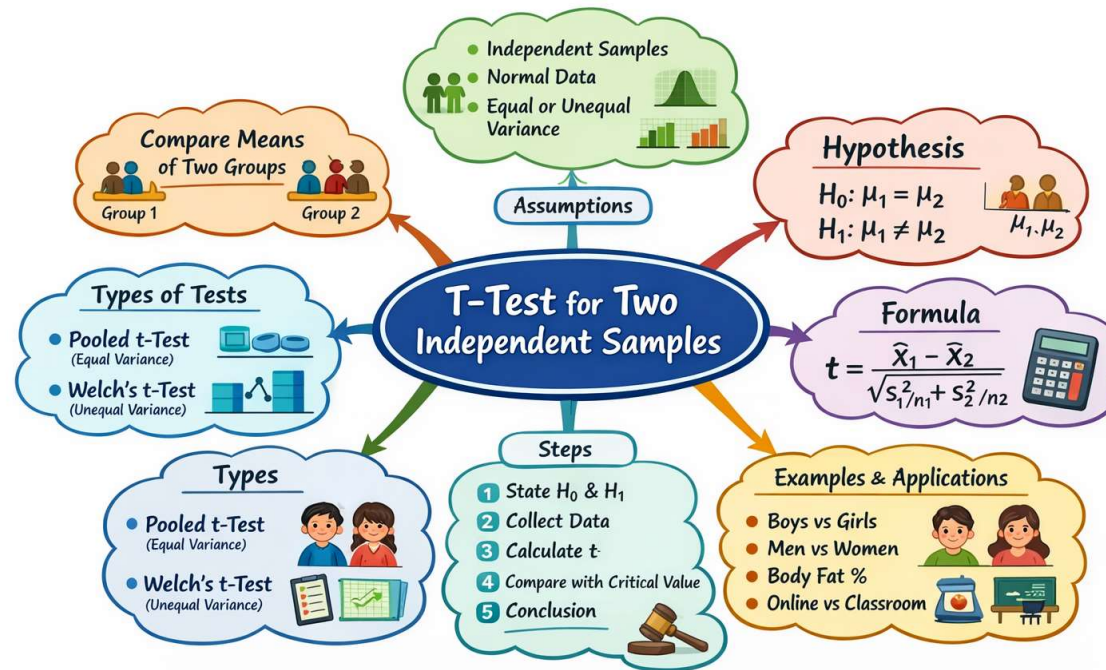
Answer: B

2. Which condition is required for a two-sample t-test?

- A) Dependent samples
- B) Independent samples
- C) Same population
- D) Equal medians only

Answer: B

MINDMAP



TEXT BOOKS	
1.	Srinivasan Desikan and Gopaldaswamy Ramesh, —Software Testing – Principles and Practices, Pearson Education, 2006.
REFERENCES	
1.	Ilene Burnstein, —Practical Software Testing, Springer International Edition, 2003
2.	Edward Kit, Software Testing in the Real World – Improving the Process, Pearson Education, 1995
3.	Boris Beizer, Software Testing Techniques – 2nd Edition, Van Nostrand Reinhold, New York, 1990.
4.	Aditya P. Mathur, —Foundations of Software Testing _ Fundamental Algorithms and Techniques, Dorling Kindersley (India) Pvt. Ltd., Pearson Education, 2008.

THANK YOU!

12