Analysis of Variance (ANOVA)

1. Meaning

Analysis of Variance (ANOVA) is a **statistical technique** used to test whether there are **significant differences between the means** of three or more groups or populations.
It helps determine **if the observed variations among sample means are due to real differences** or just random chance.

- It is an **extension of the t-test**, which compares means of only two groups.
- ANOVA uses **variance (dispersion)** to infer whether group means differ significantly.

2. Purpose

ANOVA helps answer the question:

"Are the differences among sample means statistically significant or due to random sampling error?"

For example:
A company wants to test whether **three different training programs** produce different average performance scores among employees.

3. Types of ANOVA

Type	Description	Example
One-Way ANOVA	Compares means of three or more groups based on one factor.	Comparing average test scores of students taught using three different teaching methods.
Two-Way ANOVA	Compares means based on two factors simultaneously (also studies interaction effects).	Studying the effect of **teaching method (factor 1)** and **gender (factor 2)** on student performance.
MANOVA (Multivariate ANOVA)	Tests differences in multiple dependent variables across groups.	Comparing the effect of training programs on both productivity and job satisfaction.

4. Basic Idea

In ANOVA, the total variation in the data is divided into:

- **Between-group variation** – differences among the sample means.
- **Within-group variation** – differences within individual groups (random error).

If the **between-group variance** is much larger than **within-group variance**, the group means are likely to be different.

5. Assumptions of ANOVA

1. **Normality:** The data in each group are normally distributed.
2. **Independence:** Observations are independent of each other.
3. **Homogeneity of variance:** All groups have equal variances.
4. **Random sampling:** The samples are randomly selected.

6. Hypotheses in ANOVA

- **Null Hypothesis (H_0):** All population means are equal.
$$H_0: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k$$
- **Alternative Hypothesis (H_1):** At least one mean differs.
$$H_1\text{:At least one } \mu_i \text{ is different}$$

7. ANOVA Table and Formula

Source of Variation	Sum of Squares (SS)	Degrees of Freedom (df)	Mean Square (MS)	F-Ratio
Between Groups	$SS_B = \sum n_i(\bar{X}_i - \bar{X})^2$	k – 1	$MS_B = SS_B/(k–1)$	$F = MS_B/MS_W$
Within Groups	$SS_W = \sum(X_{ij} - \bar{X}_i)^2$	N – k	$MS_W = SS_W/(N–k)$	—
Total	$SS_T = SS_B + SS_W$	N – 1	—	—

Where:

- k= number of groups
- N= total number of observations
- $\bar{X}_i$= mean of group i
- $\bar{X}$= overall mean

8. Steps in Conducting ANOVA

1. **State Hypotheses**
   - H_0: All means are equal.
   - H_1: At least one mean is different.
2. **Set Level of Significance (α)**
   Commonly 0.05.
3. **Compute Sum of Squares (SS)**
   - Between groups (SSB)
   - Within groups (SSW)
4. **Calculate Mean Squares (MS)**
   - $MS_B = SS_B/(k-1)$
   - $MS_W = SS_W/(N-k)$
5. **Compute F-Ratio**
   - $F = MS_B/MS_W$
6. **Compare with Critical Value**
   - If calculated F > tabulated F → Reject H_0.
   - Otherwise, fail to reject H_0.
7. **Draw Conclusion**

9. Example (One-Way ANOVA)

Suppose three teaching methods (A, B, C) are tested on student performance.

Method	Scores
A	70, 65, 60
B	80, 75, 85
C	50, 55, 45

Step 1: Calculate group means and overall mean.

- Mean of A = 65
- Mean of B = 80
- Mean of C = 50
- Overall mean = 65

Step 2: Compute Between-Group Sum of Squares (SSB).

$$SSB = 3[(65-65)^2 + (80-65)^2 + (50-65)^2] = 3[0 + 225 + 225] = 1350$$

Step 3: Compute Within-Group Sum of Squares (SSW).

$$SSW = [(70-65)^2 + (65-65)^2 + (60-65)^2] + [(80-80)^2 + (75-80)^2 + (85-80)^2] + [(50-50)^2 + (55-50)^2 + (45-50)^2] = 50 + 50 + 50 = 150$$

Step 4: Calculate Mean Squares.

$$MS_B = 1350/(3-1) = 675$$
$$MS_W = 150/(9-3) = 25$$

Step 5: Compute F-ratio.

$$F = 675/25 = 27$$

Step 6: Decision
At $\alpha = 0.05$ and (2,6) d.f., critical $F \approx 5.14$.
Since $27 > 5.14$ → **Reject H_0.**

Conclusion: There is a significant difference between the three teaching methods.

Chi-Square (χ^2) Test

1. Meaning

The **Chi-Square test** is a **non-parametric statistical test** used to determine whether there is a **significant association or difference** between **categorical (qualitative)** variables.

It compares the **observed frequencies (O)** in a dataset with the **expected frequencies (E)** that would occur if there were **no relationship** between the variables.

2. Purpose

The Chi-square test is mainly used to:

1. Test the **goodness of fit** (how well observed data fit an expected distribution).
2. Test the **independence of attributes** (whether two categorical variables are related).
3. Test **homogeneity** (whether different populations have the same distribution).

3. Basic Idea

It measures how much the **observed frequencies** deviate from the **expected frequencies** due to chance.

If the deviation is large, it indicates that the difference is **not due to chance**, and the variables are likely **dependent** or the fit is **not good**.

4. Formula

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Where:

- **O** = Observed frequency
- **E** = Expected frequency

5. Assumptions of Chi-Square Test

1. The data are in **frequency form** (counts, not percentages).
2. Observations are **independent**.
3. The sample size is **sufficiently large** ($E \geq 5$ for each cell).
4. Categories are **mutually exclusive** and **exhaustive**.

6. Types of Chi-Square Tests

A. Chi-Square Test of Goodness of Fit

It tests whether the sample data fit a particular **theoretical or expected distribution**.

Hypotheses

- **H_0:** The observed distribution fits the expected distribution.
- **H_1:** The observed distribution does not fit the expected distribution.

Formula

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Degrees of freedom (df) = (n – 1)

Example

A die is rolled 60 times, and the results are as follows:

Face	1	2	3	4	5	6
Observed (O)	8	9	10	12	11	10

If the die is fair, expected frequency for each face $E = 60/6 = 10$

$$\chi^2 = \sum \frac{(O - E)^2}{E} = \frac{(8-10)^2}{10} + \frac{(9-10)^2}{10} + \frac{(10-10)^2}{10} + \frac{(12-10)^2}{10} + \frac{(11-10)^2}{10} + \frac{(10-10)^2}{10}$$

$$\chi^2 = \frac{4+1+0+4+1+0}{10} = 1.0$$

df = 5;
Critical value at 5% level = 11.07

✅ Since 1.0 < 11.07 → **Fail to reject H_0.**
The die is **fair**.

B. Chi-Square Test of Independence

It tests whether **two categorical variables** are **independent or associated**.

Hypotheses

- **H_0:** The two attributes are independent.
- **H_1:** The two attributes are not independent (they are related).

Formula for Expected Frequency

$$E_{ij} = \frac{(Row\ Total_i) \times (Column\ Total_j)}{Grand\ Total}$$

Example

A survey was conducted among 100 people to study the relationship between **gender** and **preference for a product**.

Gender / Preference	Like	Dislike	Total
Male	20	30	50
Female	10	40	50
Total	30	70	100

Step 1: Calculate Expected Frequencies

For Male–Like:

$$E = \frac{(50 \times 30)}{100} = 15$$

Similarly,

- Male–Dislike = 35
- Female–Like = 15
- Female–Dislike = 35

Step 2: Apply Formula

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$
$$\chi^2 = \frac{(20 - 15)^2}{15} + \frac{(30 - 35)^2}{35} + \frac{(10 - 15)^2}{15} + \frac{(40 - 35)^2}{35}$$
$$\chi^2 = \frac{25}{15} + \frac{25}{35} + \frac{25}{15} + \frac{25}{35} = 1.67 + 0.71 + 1.67 + 0.71 = 4.76$$

Step 3: Degrees of Freedom

$$df = (r - 1)(c - 1) = (2 - 1)(2 - 1) = 1$$

Critical value ($\chi^2_{0.05,1}$) = 3.84

Since 4.76 > 3.84 → **Reject H_0.**
There is a **significant association** between gender and product preference.

C. Chi-Square Test of Homogeneity

Tests whether **different samples** come from the **same population**.
Used when comparing distributions across multiple groups (similar to test of independence but for multiple populations).

7. Interpretation

- If **calculated χ^2 > critical χ^2**, reject H_0 → There is a **significant difference or association**.
- If **calculated $\chi^2 \leq$ critical χ^2**, fail to reject H_0 → No significant difference or relationship.

8. Advantages

- Simple to apply and interpret.
- No need for population parameters (non-parametric).
- Can be used for nominal data.

1. Sign Test

Meaning

The **Sign Test** is the **simplest non-parametric test** used to test the **median difference** between **paired (dependent)** observations.
It ignores the magnitude of differences and only considers their **direction (+ or –)**.

Purpose

Used when:

- Data are **paired** (before–after, husband–wife, etc.)
- Measurement is **ordinal** or **nominal**
- Distribution is **not normal**

Hypotheses

- **H_0:** Median difference = 0 (no change or effect)
- **H_1:** Median difference ≠ 0 (there is change)

Steps

1. Compute the difference between pairs.
2. Assign + if the second value > first value, – if lower, and ignore zeros.
3. Count number of + and – signs.
4. Apply **binomial test** (n, 0.5) or use **z-test** for large n.

Example

A company measures sales before and after training for 10 salesmen.

Salesman	Before	After	Difference	Sign
1	45	50	+5	+
2	42	40	-2	-
3	50	55	+5	+
4	38	36	-2	-
5	44	48	+4	+
6	41	45	+4	+
7	48	49	+1	+
8	43	44	+1	+
9	40	42	+2	+
10	46	47	+1	+

+ = 8, - = 2

At 5% significance, p-value > 0.05
Accept H_0: No significant median difference.

2. Wilcoxon Signed-Rank Test

Meaning

The **Wilcoxon Signed-Rank Test** is a **non-parametric alternative to the paired t-test**, which considers both **direction** and **magnitude** of differences using **ranks**.

Purpose

Used for:

- Paired samples
- Ordinal data
- To test **median differences**

Hypotheses

- **H_0:** Median difference = 0
- **H_1:** Median difference ≠ 0

Formula

$$W = \text{Smaller of (Sum of + ranks, Sum of – ranks)}$$

Steps

1. Calculate the difference between paired data.
2. Ignore zero differences.

3. Rank absolute differences.
4. Assign sign (+ or –).
5. Compute sum of + and – ranks.
6. Compare **W** with **critical value** from Wilcoxon table.

3. Mann–Whitney U Test

Meaning

The **Mann–Whitney U Test** (or **Wilcoxon Rank-Sum Test**) is a **non-parametric alternative to the independent samples t-test**.
It tests whether **two independent samples** come from the same population.

Purpose

Used when:

- Data are **independent** and **ordinal**
- Population distribution is unknown

Hypotheses

- **H_0:** The two samples come from the same population.
- **H_1:** The two samples come from different populations.

Formula

$$U_A = n_A n_B + \frac{n_A(n_A + 1)}{2} - R_A$$
$$U_B = n_A n_B - U_A$$

Where R_A is the sum of ranks of group A.

Example

Scores of two teaching methods:

Method A	**12**	**15**	**14**
Method B	8	9	10

Combined ranks:
8(1), 9(2), 10(3), 12(4), 14(5), 15(6)

Sum of ranks:
$R_a = 4 + 5 + 6 = 15$
$R_\beta = 1 + 2 + 3 = 6$

$$U_A = 3 \times 3 + \frac{3 \times 4}{2} - 15 = 9 + 6 - 15 = 0$$
$$U_B = 9 - 0 = 9$$

Critical value ($n_1 = 3, n_2 = 3$) = 2 → **Reject H_0**
The two methods differ significantly.

4. Kruskal–Wallis Test

Meaning

The **Kruskal–Wallis Test** is the **non-parametric equivalent of the One-Way ANOVA**.
It compares **three or more independent samples** to check if they come from the same distribution.

Hypotheses

- **H_0:** All samples come from the same population.
- **H_1:** At least one sample differs.

Formula

$$H = \frac{12}{N(N+1)} \sum \frac{R_i^2}{n_i} - 3(N+1)$$

Where:

- R_i= sum of ranks in group i
- n_i= number of observations in group i
- N= total number of observations

Example

Three teaching methods:

Group A	14	15	16
Group B	18	19	20
Group C	10	12	11

Ranks (smallest = 1):
C(1–3), A(4–6), B(7–9)

Sum of ranks:
R_1=15, R_2=24, R_3=6, N=9

$$H = \frac{12}{9(10)}[(15^2/3) + (24^2/3) + (6^2/3)] - 30$$
$$H = \frac{12}{90}(75 + 192 + 12) - 30 = 0.133(279) - 30 = 7.2$$

Critical $\chi^2(2) = 5.99$ → **Reject H_0**
At least one teaching method differs significantly.

5. Kolmogorov–Smirnov (K–S) Test

Meaning

The **K–S Test** compares a sample's distribution with a **theoretical (expected)** distribution or compares **two samples' distributions**.
It is based on the **maximum difference (D)** between cumulative frequencies.

Purpose

Used to test:

1. **Goodness of fit** of a sample distribution to a theoretical one (e.g., normal).
2. **Equality of distributions** between two samples.

Formula

$$D = \text{Maximum} \ | F_o(X) - F_e(X) |$$

Where:

- $F_o(X)$= observed cumulative frequency
- $F_e(X)$= expected cumulative frequency

Example

For a sample of 20, the observed and expected cumulative probabilities differ by **D = 0.14**.
Critical value $D_{0.05} = 0.19$.

Since $D < D_{0.05}$ → **Accept H_0**.
The sample **follows the theoretical distribution**.

Summary Table

Test	Parametric Equivalent	Samples	Data Type	Purpose
Sign Test	Paired t-test	Paired	Nominal/Ordinal	Median difference
Wilcoxon Signed-Rank Test	Paired t-test	Paired	Ordinal	Median difference using ranks
Mann–Whitney U Test	Independent t-test	2 Independent	Ordinal	Compare two groups
Kruskal–Wallis Test	One-way ANOVA	3+ Independent	Ordinal	Compare 3 or more groups
Kolmogorov–Smirnov Test	—	1 or 2 Samples	Continuous	Test distribution fit