Hypothesis Testing II

ANOVA and Chi-Square Test

ANOVA

ANOVA, or Analysis of Variance, is a statistical test used to compare the means of three or more groups to determine if there are statistically significant differences between them.

It is often used when you have multiple groups and want to know if there is a significant difference in the means of a continuous dependent variable across these groups.

ANOVA tests the null hypothesis that there are no significant differences among the group means. If you reject the null hypothesis, it means at least one group's mean is different from the others, but it doesn't tell you which one(s) specifically.

ANOVA

Factor: This is the categorical variable that defines the groups you want to compare. For example, in a study on the impact of diet on weight loss, the factor could be "Type of Diet" with three levels: A, B, and C.

Dependent Variable: This is the continuous variable you are measuring or comparing among the groups. In the weight loss study, the dependent variable would be "Weight Loss."

Null Hypothesis (H0): The null hypothesis states that there are no significant differences among the group means. In ANOVA, it's often written as H0: $\mu 1 = \mu 2 = \mu 3$ (and so on, where μ represents the population mean of each group).

Alternative Hypothesis (Ha): The alternative hypothesis suggests that at least one group's mean is different from the others. For example, Ha: At least one μi is different.

ANOVA

Types:

One-Way ANOVA: This is used when you have one categorical factor with three or more levels (groups).

Two-Way ANOVA: This is used when you have two categorical factors (e.g., Type of Diet and Gender) and want to analyze their combined effects on the dependent variable.

One-Way ANOVA:

Scenario: One-Way ANOVA is used when you have one categorical independent variable (factor) with more than two levels or groups, and you want to compare the means of a continuous dependent variable across these groups.

Purpose: It assesses whether there are statistically significant differences in the means of the dependent variable among the different levels or groups of the independent variable.

Example: You are comparing the test scores of students who have been taught using three different teaching methods (independent variable: teaching method) to see if there is a significant difference in their average scores (dependent variable: test scores).

Main Effects: One-Way ANOVA examines the main effect of the independent variable on the dependent variable but does not assess interactions between independent variables.

Two-Way ANOVA:

Scenario: Two-Way ANOVA is used when you have two categorical independent variables (factors) and you want to analyze their combined effects on a continuous dependent variable. It can examine not only the main effects of each factor but also their interaction.

Purpose: It assesses whether there are statistically significant differences in the means of the dependent variable due to each factor (main effects) and whether there is an interaction effect between the two factors.

Example: You are investigating the impact of both gender (independent variable: gender) and type of diet (independent variable: diet) on weight loss (dependent variable). In this case, you want to know if there are differences in weight loss between males and females, between different diet types, and if the effect of diet on weight loss depends on gender (interaction effect).

Main Effects and Interaction: Two-Way ANOVA examines both the main effects of each independent variable (e.g., the effect of gender and the effect of diet) and the interaction between them (e.g., whether the effect of diet on weight loss is different for males and females)

Hypothesis Testing Steps using ANOVA

Step 1: State the Hypotheses

Null Hypothesis (H0): There is no significant difference in the mean test scores among the three teaching methods.

H0: $\mu 1 = \mu 2 = \mu 3$ (where $\mu 1$, $\mu 2$, and $\mu 3$ are the population means for Methods A, B, and C, respectively)

Alternative Hypothesis (Ha): At least one teaching method has a different mean test score compared to the others.

Ha: At least one μi is different (where μi represents the population mean for each teaching method)

Hypothesis Testing Steps using ANOVA

Step 2: Set the Significance Level

Choose a significance level (α), which represents the probability of making a Type I error. Common choices are 0.05 (5%) or 0.01 (1%). For this example, let's use $\alpha = 0.05$.

Step 3: Collect and Analyze the Data

Gather the test scores from students who were taught using the three teaching methods. In this example, we'll assume we have the following data:

Method A: [85, 90, 88, 92, 87]

Method B: [78, 82, 80, 79, 85]

Method C: [92, 88, 86, 90, 94]

Hypothesis Testing Steps using ANOVA

Step 4: Calculate the Test Statistic (F-statistic)

Calculate the F-statistic, which compares the variances between the groups and within the groups. You can use statistical software or calculators to do this. The formula for the F-statistic is:

$$F_s = \frac{MS_{(between)}}{MS_{(within)}}$$

MSB (Mean Square Between): This measures the variance between the group means.

MSW (Mean Square Within): This measures the variance within each group.

Hypothesis Testing Steps using ANOVA

Step 5: Find the Critical Value or p-value

Determine the critical F-value or find the p-value associated with the calculated F-statistic. You can look up critical values from an F-distribution table or use software. For $\alpha = 0.05$ and degrees of freedom (df) of 2 (between groups) and 12 (total - 1), the critical F-value might be approximately 3.89.

Alternatively, you can calculate the p-value associated with the F-statistic. If the p-value is less than or equal to α, you reject the null hypothesis.

Hypothesis Testing Steps using ANOVA

Step 6: Make a Decision

If F-statistic $>$ critical F-value (or p-value $\leq \alpha$), reject the null hypothesis.

If F-statistic $\leq$ critical F-value (or p-value $> \alpha$), fail to reject the null hypothesis.

Step 7: Draw a Conclusion

Based on your decision in Step 6, draw a conclusion and interpret your results. If you reject the null hypothesis, it means there is a significant difference in the mean test scores among the teaching methods. If you fail to reject the null hypothesis, it means there is not enough evidence to conclude that there is a significant difference.

Basics of ANOVA

$$\text{ANOVA} = \frac{\text{Variance Between}}{\text{Variance Within}}$$

$$\text{Total Variance} = \text{Variance Between} + \text{Variance Within}$$

$$\frac{\text{Variance Between}}{\text{Variance Within}} > 1$$ Reject H_0

$$\frac{\text{Variance Between}}{\text{Variance Within}} < 1$$ Fail to Reject H_0

$$\frac{\text{Variance Between}}{\text{Variance Within}} = 1$$ Fail to Reject H_0

We want to see if three different studying methods can lead to different mean exam scores or not. To test this, we select 30 students and randomly assign 10 each to use a different studying method.

Basics of ANOVA

Sno	Method A	Method B	Method C
1.	10	8	9
2.	9	9	8
3.	8	10	7
4.	7.5	8	10
5.	8.5	8.5	9
6.	9	7	8
7.	10	9.5	7
8.	8	9	10
9.	8	7	9
10.	9	10	8
	8.7	8.6	8.5

Overall Mean 8.6

Between Group Variation =10*(8.7-8.6)^2+10*(8.6- 8.6)^2+10*(8.5-8.6)^2

Between Group Variation = **0.2**

Within Group Variation: $\Sigma(X_{ij} - X_j)^2$

Where:

Σ: a symbol that means "sum"

X_{ij}: the i^{th} observation in group j

X_j: the mean of group j

Method A: $(10-8.7)^2 + (9-8.7)^2 + (8-8.7)^2 + (7.5-8.7)^2 + (8.5-8.7)^2 + (9-8.7)^2 + (10-8.7)^2 + (8-8.7)^2 + (8-8.7)^2 + (9-8.7)^2 = 6.6$

Method B: $(8-8.6)^2 + (9-8.6)^2 + (10-8.6)^2 + (8-8.6)^2 + (8.5-8.6)^2 + (7-8.6)^2 + (9.5-8.6)^2 + (9-8.6)^2 + (7-8.6)^2 + (10-8.6)^2 = 10.9$

Method C: $(9-8.5)^2 + (8-8.5)^2 + (7-8.5)^2 + (10-8.5)^2 + (9.5-8.5)^2 + (8-8.5)^2 + (7-8.5)^2 + (10-8.5)^2 + (9-8.5)^2 + (8-8.5)^2 = 10.5$

Within Group Variation: 6.6+10.9+10.5 = 28

Method 1

$$\frac{\text{Variance Between}}{\text{Variance Within}} = \frac{0.2}{28} = 0.0071 < 1$$

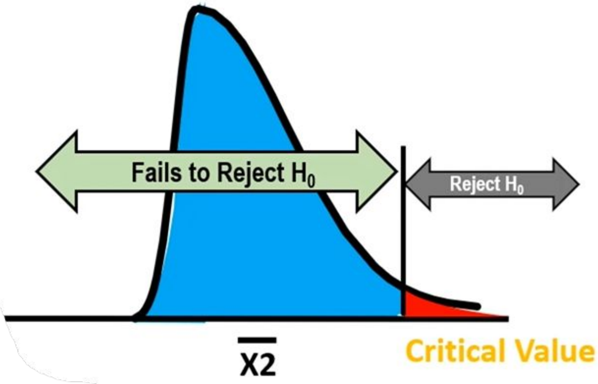
Fail to Reject H_0

"Means are very close to overall mean and distribution overlap is hard to distinguish".

Basics of ANOVA

$F_{Critical} > F_{Stat}$ Fail to Reject H_0

$F_{Critical} < F_{Stat}$ Reject H_0



Assuming $\alpha = 0.05$

$$F_{Stat} = \frac{\text{Variance Between}}{\text{Variance Within}}$$

$$\frac{0.2}{28} = 0.0071$$

$$F_{Critical} = \frac{\text{Numerator Degree of Freedom}}{\text{Denominator Degree of Freedom}}$$

Numerator Degree of Freedom = No. of Samples - 1 = 3 - 1 = 2

Denominator Degree of Freedom = $\sum(nj-1) = n_T - k = 30 - 3 = 27$

$F_{Critical} = F_{(2,27)} = 3.35$

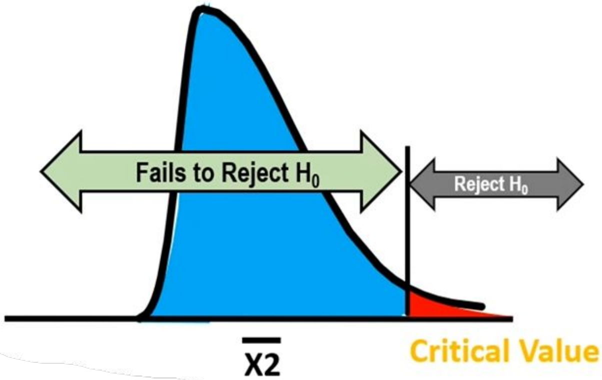
Method 2

F-table of Critical Values of α = 0.05 for F(df1, df2)																			
	DF1=1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**12**	**15**	**20**	**24**	**30**	**40**	**60**	**120**	**∞**
DF2=1	161.45	199.50	15.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88	243.91	245.95	248.01	249.05	250.10	251.14	252.20	253.25	254.31
2	18.51	19.00	9.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40	19.41	19.43	19.45	19.45	19.46	19.47	19.48	19.49	19.50
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00

Basics of ANOVA

$F_{Critical} > F_{Stat}$ Fail to Reject H_0

$F_{Critical} < F_{Stat}$ Reject H_0



Assuming $\alpha = 0.05$

$$F_{Stat} = \frac{\text{Variance Between}}{\text{Variance Within}}$$

$$\frac{0.2}{28} = 0.0071$$

$$F_{Critical} = \frac{\text{Numerator Degree of Freedom}}{\text{Denominator Degree of Freedom}}$$

Numerator Degree of Freedom = No. of Samples - 1 = 3-1 = 2

Denominator Degree of Freedom = $\sum(nj-1) = n_T - k$ = 30 -3 = 27

$$F_{Critical} = F_{(2,27)} = 3.35$$

$$F_{Critical} > F_{Stat}$$

Fail to Reject H_0

Motivating ANOVA

- A random sample of some quantitative trait was measured in individuals randomly sampled from population

- Genotyping of a single SNP
  - AA: 82, 83, 97
  - AG: 83, 78, 68
  - GG: 38, 59, 55

Rational of ANOVA

- Basic idea is to partition total variation of the data into two sources

    1. Variation within levels (groups)

    2. Variation between levels (groups)

- If H_0 is true the *standardized* variances are equal to one another

The Details

Our Data:

AA:	82, 83, 97	$\bar{x}_{1.} = (82 + 83 + 97)/3 = 87.3$
AG:	83, 78, 68	$\bar{x}_{2.} = (83 + 78 + 68)/3 = 76.3$
GG:	38, 59, 55	$\bar{x}_{3.} = (38 + 59 + 55)/3 = 50.6$

- Let X_{ij} denote the data from the i^{th} level and j^{th} observation

- Overall, or ***grand mean***, is:

$$\bar{x}_{..} = \sum_{i=1}^{K} \sum_{j=1}^{J} \frac{x_{ij}}{N}$$

$$\bar{x}_{..} = \frac{82 + 83 + 97 + 83 + 78 + 68 + 38 + 59 + 55}{9} = 71.4$$

Partitioning Total Variation

- Recall, variation is simply average squared deviations from the mean

$$SST = SST_G + SST_E$$

$$\sum_{i=1}^{K}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{..})^2 \qquad \sum_{i=1}^{K} n_i \bullet (\bar{x}_{i.} - \bar{x}_{..})^2 \qquad \sum_{i=1}^{K}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{i.})^2$$

Sum of squared deviations about the grand mean across all N observations

Sum of squared deviations for each group mean about the grand mean

Sum of squared deviations for all observations within each group from that group mean, summed across all groups

In Our Example

$$SST = SST_G + SST_E$$

$$\sum_{i=1}^{K}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{..})^2 \qquad \sum_{i=1}^{K} n_i \bullet (\bar{x}_{i.} - \bar{x}_{..})^2 \qquad \sum_{i=1}^{K}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{i.})^2$$

$$\begin{aligned}&(82-71.4)^2 + (83-71.4)^2 + (97-71.4)^2 +\\ &(83-71.4)^2 + (78-71.4)^2 + (68-71.4)^2 +\\ &(38-71.4)^2 + (59-71.4)^2 + (55-71.4)^2 =\end{aligned}$$

$$\begin{aligned}&3 \bullet (87.3-71.4)^2 +\\ &3 \bullet (76.3-71.4)^2 +\\ &3 \bullet (50.6-71.4)^2 =\end{aligned}$$

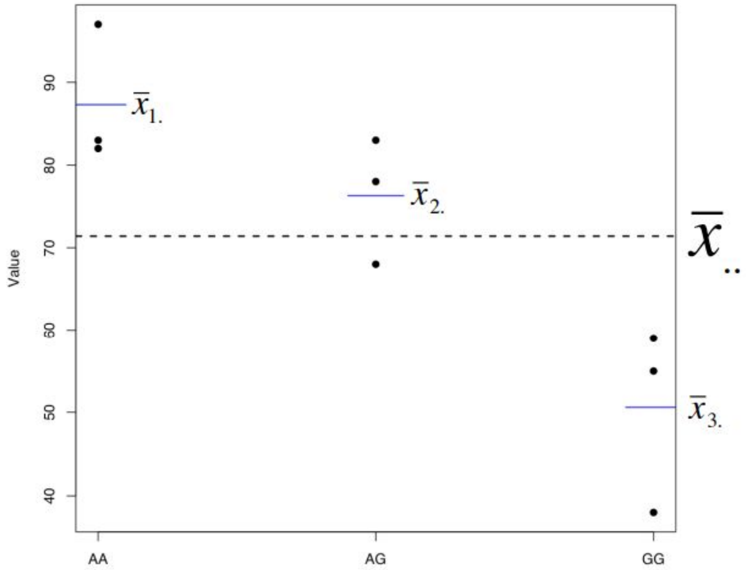
$$\begin{aligned}&(82-87.3)^2 + (83-87.3)^2 + (97-87.3)^2 +\\ &(83-76.3)^2 + (78-76.3)^2 + (68-76.3)^2 +\\ &(38-50.6)^2 + (59-50.6)^2 + (55-50.6)^2 =\end{aligned}$$

2630.2   2124.2   506

In Our Example

$$SST \quad = \quad SST_G \quad + \quad SST_E$$

$$\sum_{i=1}^{K}\sum_{j=1}^{J}(x_{ij}-\bar{x}_{..})^2 \qquad \sum_{i=1}^{K} n_i \bullet (\bar{x}_{i.}-\bar{x}_{..})^2 \qquad \sum_{i=1}^{K}\sum_{j=1}^{J}(x_{ij}-\bar{x}_{i.})^2$$



Calculating Mean Squares

- To make the sum of squares comparable, we divide each one by their associated degrees of freedom
    - $SST_G = k - 1$ $(3 - 1 = 2)$
    - $SST_E = N - k$ $(9 - 3 = 6)$
    - $SST_T = N - 1$ $(9 - 1 = 8)$

- $MST_G = 2124.2 / 2 = 1062.1$

- $MST_E = 506 / 6 = 84.3$

Almost There… Calculating F Statistic

- The test statistic is the ratio of group and error mean squares

$$F = \frac{MST_G}{MST_E} = \frac{1062.2}{84.3} = 12.59$$

- If H_0 is true MST_G and MST_E are equal
- Critical value for rejection region is $F_{\alpha, k-1, N-k}$
- If we define $\alpha = 0.05$, then $F_{0.05, 2, 6} = 5.14$

Accept or Reject H_o?????

ANOVA Table

Source of Variation	df	Sum of Squares	MS	F
Group	k-1	SST_G	$\frac{SST_G}{k-1}$	$\frac{SST_G}{k-1} \Big/ \frac{SST_E}{N-k}$
Error	N-k	SST_E	$\frac{SST_E}{N-k}$	
Total	N-1	SST		

Chi-Square Test

The Chi-Square (χ^2) test is a statistical test used to determine if there is a significant association between two categorical variables.

It assesses whether the observed frequencies of categories in a contingency table differ from the expected frequencies under a null hypothesis of independence.

In simpler terms, it helps answer questions like, "Is there a relationship between two categorical variables, or are they independent?"

Why Use Chi-Square Test?

1. Test for Independence

One of the primary uses of the Chi-Square test is to determine whether there is a significant association or relationship between two categorical variables.

It helps answer questions like: Are two variables independent, or is there a dependency or association between them?

For example, in medical research, you might use a Chi-Square test to assess if there is an association between smoking status (smoker or non-smoker) and the incidence of lung cancer.

Why Use Chi-Square Test?

2. Test for Goodness of Fit

The Chi-Square test can be employed to test whether observed data fits a theoretical or expected distribution.

It helps researchers evaluate whether the observed frequencies of events match what would be expected under a specific hypothesis or model.

For instance, in genetics, you can use it to determine if observed genetic ratios (e.g., in a Punnett square) match expected Mendelian ratios.

Why Use Chi-Square Test?

3. Assess Homogeneity

Chi-Square tests can be used to assess whether the distribution of a categorical variable is consistent across different groups or populations.

Researchers can compare observed frequencies of categories to determine if they are significantly different across groups.

This can be useful in social sciences to examine whether preferences or behaviors vary among different demographic groups.

Why Use Chi-Square Test?

4. Variable/Feature Selection

In feature selection for machine learning and predictive modeling, Chi-Square tests can help identify which categorical variables are most relevant for prediction.

Variables that show significant associations with the outcome variable are often selected for inclusion in predictive models.

Why Use Chi-Square Test?

5. Quality Control and Testing Hypotheses

In quality control and manufacturing processes, Chi-Square tests can be used to assess whether observed defects or deviations from specifications are statistically significant.

Researchers can test hypotheses related to product quality and process improvements.

6. Survey and Market Research

In survey research and market analysis, Chi-Square tests can help analyze responses to categorical survey questions and investigate whether there are significant differences in preferences or opinions among different segments of the population.

7. Biology and Ecology

In biology and ecology, Chi-Square tests are used to analyze data related to species distribution, genetics, and ecological studies.

Perform a Chi-Square Test

1. **Formulate Hypotheses:** Start with null (H0) and alternative (Ha) hypotheses. For testing independence, H0 assumes no association, while Ha suggests an association.

2. **Collect Data:** Gather data and organize it into a contingency table, which shows the counts or frequencies of observations for each combination of categories.

3. **Calculate Expected Frequencies:** Compute the expected frequencies for each cell in the table assuming independence. Typically, it's (row total * column total) / grand total.

Perform a Chi-Square Test

4. **Compute the Chi-Square Statistic:** Calculate the Chi-Square statistic using the formula:

The Formula for Chi-Square Is

$$\chi_c^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where:

$c =$ Degrees of freedom

$O =$ Observed value(s)

$E =$ Expected value(s)

Perform a Chi-Square Test

5. **Determine Degrees of Freedom:** Calculate the degrees of freedom (df), which depend on the dimensions of the contingency table. For a 2x2 table, df = 1.
6. **Find Critical Value or P-value:** Look up the critical Chi-Square value from a Chi-Square distribution table or use software. Alternatively, calculate the p-value associated with the Chi-Square statistic.
7. **Make a Decision:** Compare the calculated Chi-Square statistic to the critical value or p-value:
   - If $\chi^2 >$ Critical χ^2 or p-value $< \alpha$ (chosen significance level), reject the null hypothesis.
   - If $\chi^2 \leq$ Critical χ^2 or p-value $\geq \alpha$, fail to reject the null hypothesis.
8. **Draw a Conclusion:** Based on the decision, conclude whether there is a significant association between the variables or if they are independent.

Limitations of Chi-Square Test:

Assumes Independence: Chi-Square tests for independence assume that the variables are independent. If there are confounding factors or a causal relationship, the test might yield incorrect results.

Sample Size: For small sample sizes or sparse data, the Chi-Square test can be less reliable. It's important to have sufficient data in each cell of the contingency table.

Categorical Variables: It's designed for categorical variables. It's not suitable for continuous variables without discretization.

Limited to Association: Chi-Square can tell you if there's an association, but it doesn't provide information about the strength or direction of the relationship.

Not Suitable for All Data Types: Chi-Square is mainly used for nominal and ordinal data. For interval or ratio data, other tests like t-tests or ANOVA are more appropriate.

Aspect	Parametric Tests	Nonparametric Tests
Assumptions	- Assume a specific data distribution, often normal (e.g., Gaussian).	- Do not assume any specific data distribution.
	- Assume homoscedasticity (equal variances among groups).	- Do not assume equal variances (heteroscedasticity is acceptable).
Data Type	- Typically used for continuous data (interval or ratio).	- Applicable to a wide range of data types, including nominal, ordinal, interval, and ratio.
Examples	- t-Test, ANOVA, Pearson Correlation, Linear Regression, etc.	- Mann-Whitney U Test, Kruskal-Wallis Test, Chi-Square Test, etc.
Strengths	- More powerful and efficient when assumptions are met.	- Robust and applicable when assumptions are violated.
	- Provide more precise parameter estimates and smaller standard errors.	- Suitable for non-continuous data types and non-normally distributed data.
Limitations	- Strict assumptions may not hold in all cases.	- Generally, less powerful than parametric tests under ideal conditions.
	- May not be suitable for non-normally distributed data.	- May not provide as much information about underlying population parameters.
Use Cases	- Often used when data follows a known distribution and assumptions are met.	- Used when assumptions are violated or for non-continuous data.
	- Useful for inferential statistics and parameter estimation.	- Commonly applied in exploratory data analysis and hypothesis testing.

Nonparametric Test	What it does	Parametric Counterpart
Wilcoxon Signed Rank	Compares 1 Median to a specified value	z-test, 1-Sample t-test
	Compares 2 Dependent (Paired) Medians	Paired (Dependent) Samples t-test
Mann-Whitney	Compares 2 Independent Medians	2 (Independent) Samples t-test
Kruskal-Wallis	Compares 3 or more Medians, 1 Variable	1-way ANOVA
Friedman	Compares 3 or more Medians, 2 Variables	2-way ANOVA
Chi-Square Test of Independence	Tests 2 Categorical Variables for Independence (lack of Association)	none