How to Interpret Chi-Square Test Results in SPSS

Authored by

mohammed looti

January 7, 2026

RECOMMENDED CITATION

mohammed looti (2026). *How to Interpret Chi-Square Test Results in SPSS*. PSYCHOLOGICAL SCALES. Retrieved from https://scales.arabpsychology.com/?p=124917

The Chi-Square Test, or Pearson's Chi-Square Test of Independence, is an indispensable tool in statistical analysis, particularly within social sciences, epidemiology, and market research. Its primary function is to assess whether a significant relationship exists between two categorical variables. When conducting this analysis using software like SPSS (Statistical Package for the Social Sciences), understanding the output tables--especially the core components like the Chi-Square value and the associated p-value--is crucial for making valid research claims. This detailed guide provides expert instruction on accurately interpreting the results generated by SPSS.

Proper interpretation hinges on comparing the observed data distribution against the distribution expected if no relationship existed between the variables. The resulting Chi-Square statistic quantifies the magnitude of the disparity between these observed and expected frequencies. However, the true measure of significance comes from the p-value, which represents the probability of observing such a difference--or a greater one--purely by random chance, assuming the null hypothesis of independence is true. A comprehensive analysis also necessitates looking beyond the overall test statistic to examine the contributions of individual cells through standardized residuals, providing a nuanced understanding of where the association lies.

Understanding the Context of the Chi-Square Test

The core purpose of the Chi-Square Test of Independence is to test the null hypothesis that two categorical variables are statistically independent. This means that the classification of an observation into a category of one variable does not influence its classification in the categories of the second variable. Conversely, the alternative hypothesis posits that the variables are indeed associated. Before executing the test, researchers must ensure their data meets the necessary assumptions, primarily that observations are independent and that expected cell counts are not too small (typically, expected counts should be 5 or greater in at least 80% of cells).

We will demonstrate the complete process using a typical scenario analyzed within SPSS. Understanding this workflow is essential not only for executing the analysis correctly but also for framing the interpretation within the context of the original research question. The steps involve defining variables, generating the output, and meticulously reviewing three key tables provided by the software.

Setting Up the Chi-Square Test of Independence in SPSS: Example Scenario

To illustrate the interpretation process, let us consider a common research question: Is there a significant association between an individual's **Gender** and their **Political Party Preference**? This inquiry is perfect for the Chi-Square Test because both variables--Gender (e.g., Male/Female) and Party Preference (e.g., Party A/Party B/Party C)--are nominal categorical variables.

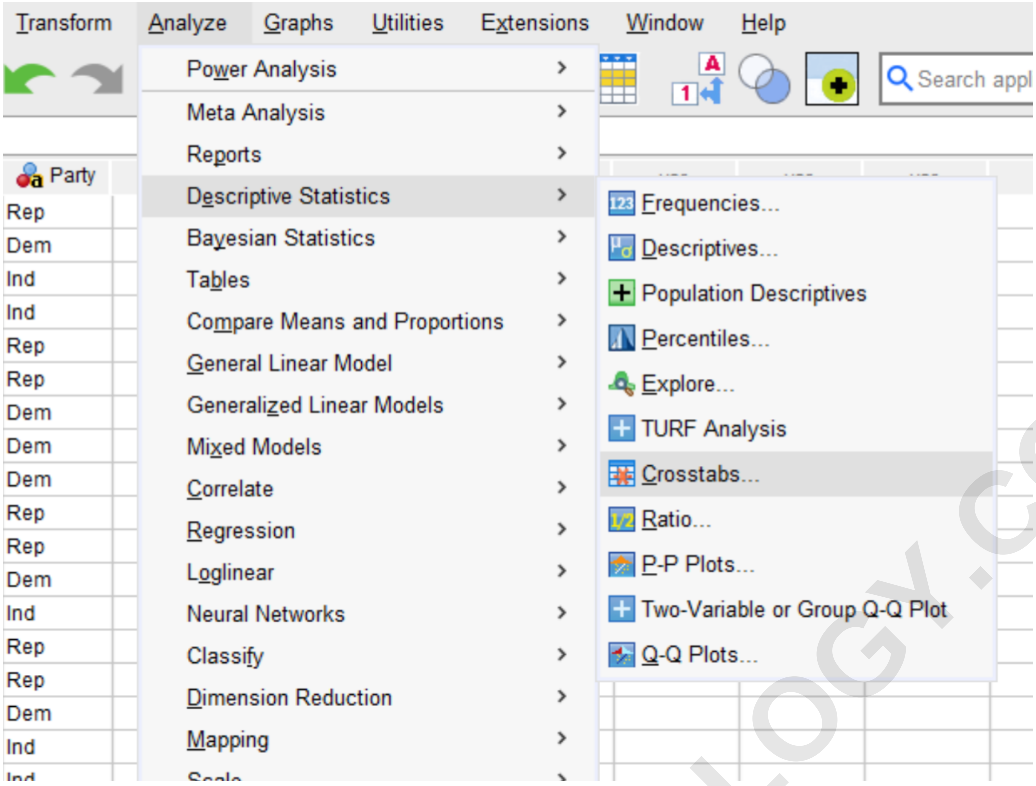
For this hypothetical scenario, we have collected data from a simple random sample of **50 voters**. Each voter was surveyed regarding their gender and their primary political affiliation. This type of cross-sectional data collection ensures the independence of observations, a fundamental requirement for the validity of the test. The structure of the raw data, ready for analysis in the SPSS Data View, is displayed below, illustrating how categorical responses are coded numerically for software processing.

The following screenshot shows how to enter this data into SPSS:

	VoterID	Gender	Party	var	var
1	1	Male	Rep		
2	2	Male	Dem		
3	3	Female	Ind		
4	4	Male	Ind		
5	5	Female	Rep		
6	6	Male	Rep		
7	7	Male	Dem		
8	8	Male	Dem		
9	9	Female	Dem		
10	10	Male	Rep		
11	11	Male	Rep		
12	12	Female	Dem		
13	13	Female	Ind		
14	14	Male	Rep		
15	15	Male	Rep		
16	16	Female	Dem		
17	17	Male	Ind		
18	18	Female	Ind		
19	19	Female	Rep		
20	20	Male	Dem		
21	21	Male	Dem		
22	22	Female	Ind		

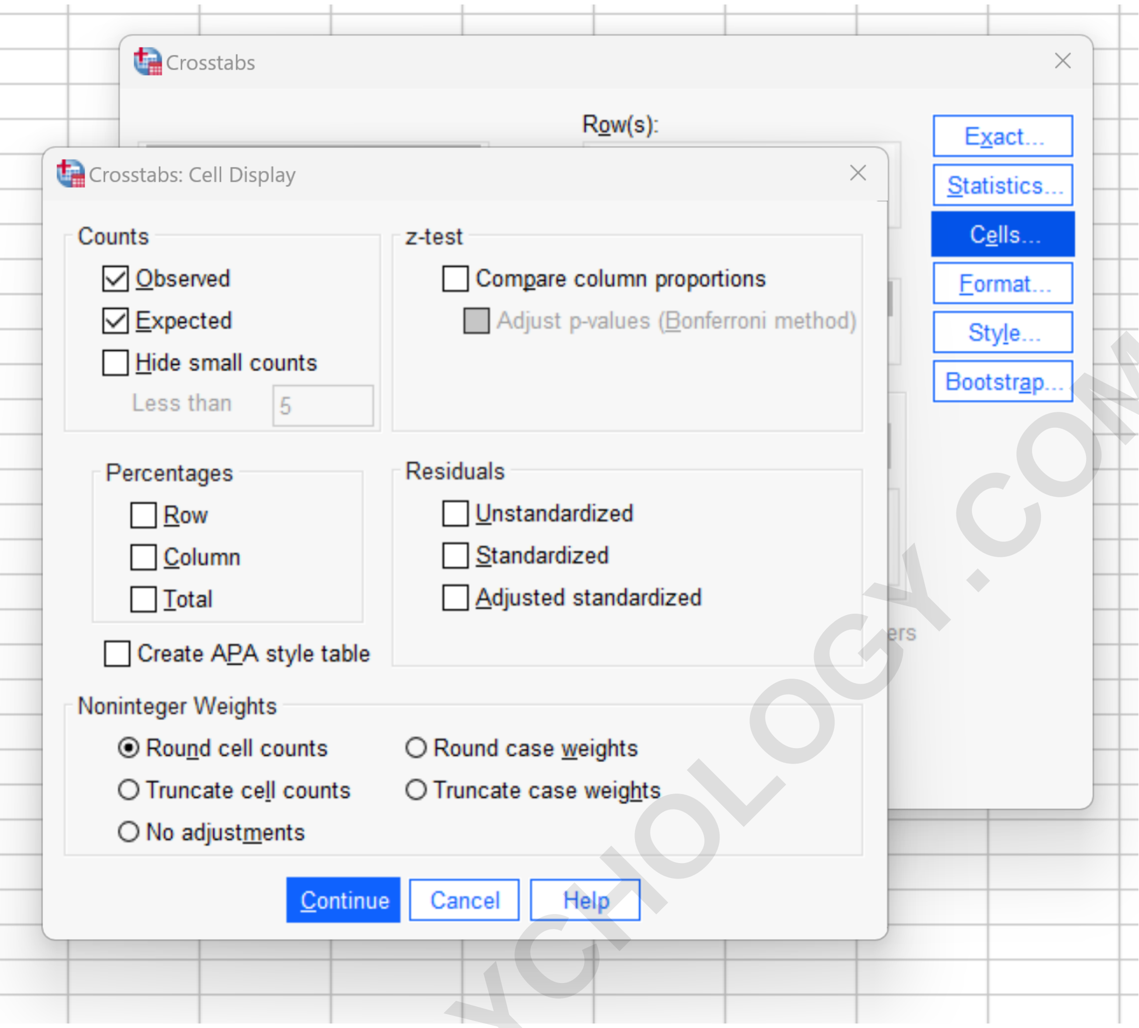
Executing the Analysis: Step-by-Step Procedure in SPSS

The process for generating the necessary output tables in SPSS is straightforward, beginning with the menu navigation. To initiate the Chi-Square Test of Independence, navigate to the **Analyze** tab, select **Descriptive Statistics**, and finally click on **Crosstabs**. This sequence opens the primary dialog box where you define the structure of your two-way table.



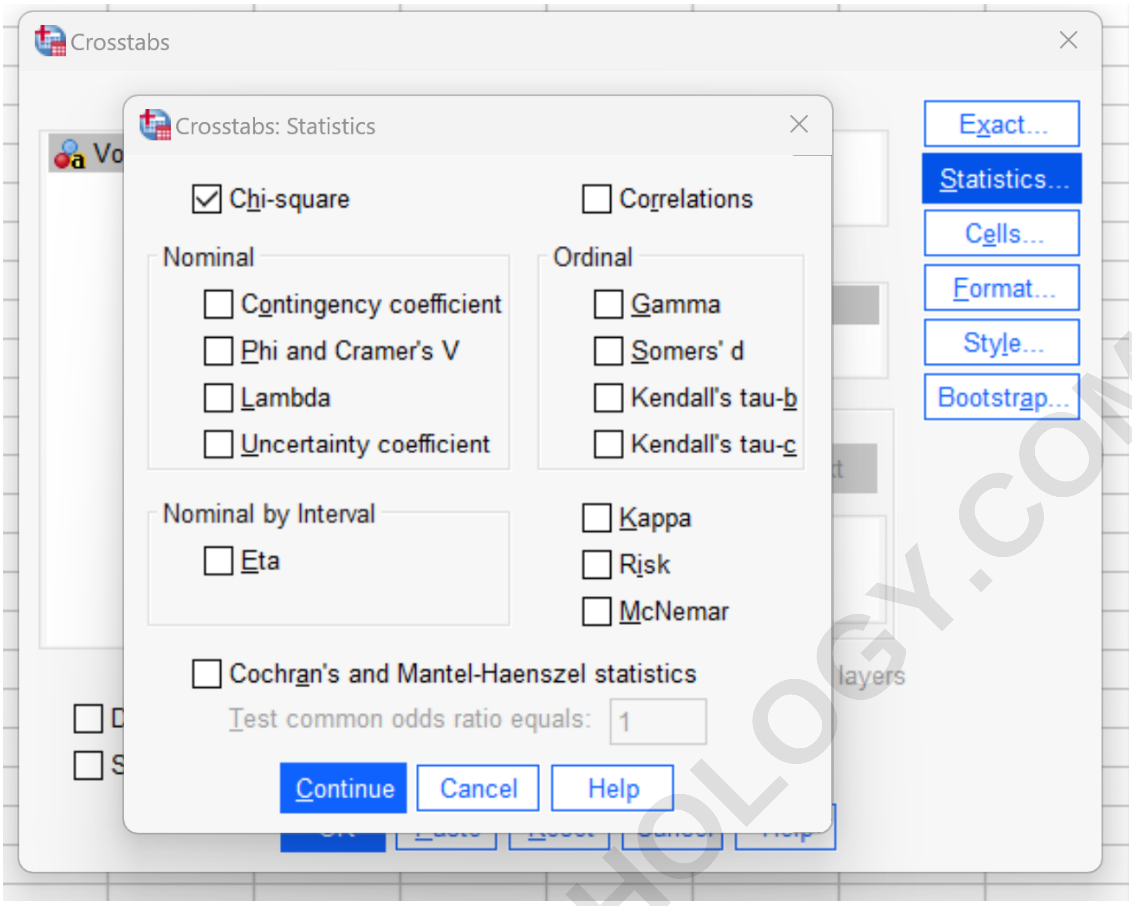
In the Crosstabs dialog box, you must assign your variables correctly. Conventionally, the variable that represents the independent or grouping factor (in this case, **Gender**) is dragged into the **Rows** panel, while the dependent variable (**Party** preference) is dragged into the **Columns** panel. While this arrangement does not affect the ultimate Chi-Square statistic, it standardizes the presentation of the resulting contingency table, making interpretation easier.

Crucially, you must request both the observed and expected counts to fully understand the results and verify assumptions. Click the **Cells** button and ensure that the boxes next to **Observed** and **Expected** are checked. The observed count is the actual frequency gathered from your sample, while the expected count is the hypothetical frequency you would observe if perfect independence existed between the two variables. This comparison is the foundation of the Chi-Square calculation. Click **Continue** to save these selections.



Then click **Continue**.

Next, click the **Statistics** button. Then check the box next to **Chi-square**. This command instructs SPSS to calculate Pearson's Chi-Square statistic, the degrees of freedom, and the associated asymptotic significance (p-value). After verifying this selection, click **Continue**.



Then click **Continue**.

Once all settings are confirmed, click **OK** in the main Crosstabs dialog box to run the analysis and generate the output viewer window.

Crosstabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Gender * Party	50	100.0%	0	0.0%	50	100.0%

Gender * Party Crosstabulation

			Party			
			Dem	Ind	Rep	Total
Gender	Female	Count	9	7	5	21
		Expected Count	8.0	6.3	6.7	21.0
	Male	Count	10	8	11	29
		Expected Count	11.0	8.7	9.3	29.0
Total		Count	19	15	16	50
		Expected Count	19.0	15.0	16.0	50.0

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.118[a]	2	.572
Likelihood Ratio	1.140	2	.566
N of Valid Cases	50		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6.30.

Initial Review of SPSS Output: Case Processing and Crosstabulation

The first step in interpreting the comprehensive SPSS output is to review the preliminary tables to ensure data integrity and verify expected frequencies. The output starts with the **Case Processing Summary** table.

The **Case Processing Summary** is crucial for understanding the size and completeness of the analyzed sample. This table reports the total number of cases used in the analysis (Valid) and the number of cases excluded due to missing data (Missing). In our example, we observe **50** valid cases and **0** missing cases. This confirms that all data collected was utilized, minimizing potential bias associated with listwise deletion of incomplete observations. If a substantial number of cases were missing, the reliability of the overall findings might be compromised.

Following the summary, the detailed **Crosstabulation** table presents the core structure of the analysis. This table is a visual matrix displaying the intersection of the two categorical variables: Gender (Rows) and Party Preference (Columns). Each cell within this matrix contains two critical values: the **Observed Count** (the actual frequency from the data) and the **Expected Count**.

The **Expected Count** is the frequency that would hypothetically occur in each cell if the null hypothesis of independence were perfectly true. It is calculated by multiplying the row total by the column total and dividing by the grand total sample size. Reviewing these expected counts is mandatory to validate the assumptions of the Chi-Square Test. If more than 20% of the cells have an expected count less than 5, or if any cell has an expected count less than 1, the standard Pearson Chi-Square result may be unreliable, and alternative tests, such as Fisher's Exact Test, should be considered.

Refer to external resources for a deeper explanation of how to calculate expected counts in a Chi-Square test.

Deciphering the Chi-Square Tests Table

The most important section for the primary inference is the **Chi-Square Tests** table. This table summarizes several variations of the Chi-Square statistic, but analysts typically focus on the row labeled "Pearson Chi-Square." This row contains the statistical measure that tests the relationship between the two variables.

Three critical values are reported in this row: the Chi-Square statistic itself, the degrees of freedom (df), and the Asymptotic Significance (2-sided), which is the p-value. In our example, the calculated Chi-Square test statistic is **1.118**. This value represents the aggregate difference between the observed frequencies and the frequencies expected under the condition of independence. A larger Chi-Square value indicates a greater deviation from independence, suggesting a stronger association.

The degrees of freedom (df) are calculated as (Number of Rows - 1) * (Number of Columns - 1). This value is essential because it determines the specific Chi-Square distribution used to calculate the p-value. The degrees of freedom adjust the distribution shape based on the complexity of the contingency table. Finally, the Asymptotic Significance, or the p-value, is **.572**. This is the probability that the observed disparity (or greater) could occur purely by chance if Gender and Party Preference were truly independent in the population.

Interpreting the Results: Hypothesis Testing and Decision Rule

Statistical inference relies on formally stating the null and alternative hypotheses before applying the decision rule. For the Chi-Square Test of Independence, these hypotheses are defined as follows:

H0 (Null Hypothesis): The two variables (Gender and Political Party Preference) are statistically independent in the population.
HA (Alternative Hypothesis): The two variables are *not* independent; they are significantly associated.

To make a decision, we compare the calculated p-value against a predetermined **significance level** ($alpha$), which is typically set at 0.05. This significance level dictates the threshold for statistical significance--the maximum risk we are willing to accept for rejecting a true null hypothesis (Type I error).

The formal decision rule states: If the p-value is less than $alpha$ (e.g., 0.05), we **reject** the null hypothesis, concluding that a significant association exists. If the p-value is greater than or equal to $alpha$, we **fail to reject** the null hypothesis, concluding that there is insufficient evidence to claim an association.

In our analysis, the Asymptotic Significance (p-value) is **.572**. Comparing this to the standard significance level of 0.05: $0.572 > 0.05$. Therefore, we **fail to reject the null hypothesis**. This failure to reject means we do not have sufficient statistical evidence, based on the sample of 50 voters, to conclude that Gender and Political Party Preference are associated. The observed differences between the genders' political preferences are likely due merely to random sampling variation.

Advanced Interpretation: Utilizing Standardized Residuals

While the Pearson Chi-Square test determines if an overall association exists, it does not specify which particular categories or cells contribute most significantly to that association. For this deeper insight, analysts must examine the **standardized residuals**, which can be requested within the SPSS Crosstabs 'Cells' option (under 'Residuals').

A standardized residual represents the difference between the observed count and the expected count, standardized by the cell's standard error. This standardization allows for meaningful comparison across different cell sizes. If the absolute value of the standardized residual exceeds 1.96 (corresponding roughly to a two-tailed $alpha = 0.05$ threshold), that specific cell is considered a significant contributor to the overall Chi-Square statistic.

If, for instance, a cell showed a high positive standardized residual (e.g., +3.0), it would indicate that the observed count in that specific category pairing is significantly higher than expected under the assumption of independence. Conversely, a highly negative residual (e.g., -2.5) suggests the observed count is significantly lower than expected. Although our primary test failed to find an overall association (p=0.572), had the result been significant (p<0.05), the standardized residuals would be the next logical step to precisely locate where the dependencies lie within the crosstabulation matrix.

Summary of Interpretation and Reporting Guidelines

Interpreting the results of the Chi-Square Test in SPSS involves a systematic review of the output tables, moving from data validation to hypothesis testing. The interpretation is complete only when both the overall statistical significance (from the Chi-Square Tests table) and the practical relevance (often inferred from the Crosstabulation and standardized residuals) have been assessed.

For academic or professional reporting, the results must be communicated clearly, including the test statistic (chi^2), the degrees of freedom (df), and the p-value. Based on our example, the official reporting statement would be: "A Chi-Square Test of Independence indicated that there was no significant association between Gender and Political Party Preference, $chi^2(2) = 1.118$, p = 0.572. We failed to reject the null hypothesis of independence."

Mastering this interpretation allows researchers to confidently draw conclusions about the relationship between two categorical variables, ensuring that statistical claims are grounded in robust evidence and adherence to methodological requirements, such as expected cell count assumptions.

The following tutorials explain how to perform other common tasks in SPSS: