

Learn How to Perform a Chi-Square Goodness of Fit Test in Google Sheets: A Step-by-Step Guide

Authored by
Mohammed Iooti

November 5, 2025

RECOMMENDED CITATION

Mohammed Iooti (2025). *Learn How to Perform a Chi-Square Goodness of Fit Test in Google Sheets: A Step-by-Step Guide*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=10997>

The [Chi-Square Goodness of Fit Test](#) is an indispensable statistical method designed to assess whether observed frequency data for a [categorical variable](#) aligns significantly with a predefined theoretical or [hypothesized distribution](#). This powerful inferential tool allows researchers and analysts to formally determine if the discrepancies between the expected results and the actual empirical outcomes are merely due to random chance, or if the underlying population distribution genuinely differs from the initial assumption. Understanding this test is crucial for making data-driven decisions when dealing with nominal data across various fields, from marketing to biology.

In essence, the Chi-Square Goodness of Fit Test answers a fundamental question: does the observed sample accurately represent the population we assume it comes from? It is particularly valuable when analyzing data that falls into distinct, mutually exclusive categories, such as preferences, demographics, or, as we will explore, daily traffic counts. This guide provides a comprehensive, four-step process for executing this critical analysis efficiently using the capabilities of **Google Sheets**.

Understanding the Test's Application: A Practical Business Case

To illustrate the utility of the Chi-Square Goodness of Fit Test, consider a common retail scenario. A shop owner asserts that customer traffic is evenly distributed throughout the standard five-day business week (Monday through Friday). This assertion forms the basis of our statistical inquiry, specifically defining the [null hypothesis](#) (H₀): that the day of the week, as a categorical variable, follows a [uniform distribution](#). The purpose of our test is to gather evidence strong enough to potentially reject this established assumption.

To test this claim, an independent analyst monitors the store and records the actual number of customers entering the shop over a full week. These recorded values constitute the "observed" frequencies, which will then be rigorously compared against the shop owner's "expected" frequencies derived from the uniform distribution assumption. The core challenge is determining if the natural fluctuation in daily counts is statistically significant or just random noise.

The empirical data collected by the researcher shows the following **observed customer counts**:

Monday: 50 customers

Tuesday: 60 customers

Wednesday: 40 customers

Thursday: 47 customers

Friday: 53 customers

By applying the Chi-Square Goodness of Fit Test methodology, we can definitively assess whether the variation present in these observed data points provides sufficient statistical justification to reject the shop owner's claim of perfectly equal daily traffic. The following steps guide you through

the precise execution of this analysis within the familiar environment of Google Sheets.

Step 1: Organizing Data and Calculating Expected Frequencies

The analytical process begins with meticulous data organization in Google Sheets, followed by the calculation of the corresponding expected frequencies (E). These expected values are derived directly from the parameters specified by the [null hypothesis](#). Maintaining clarity in the spreadsheet structure is paramount, as it minimizes the potential for calculation errors in subsequent stages.

First, input the weekday categories and their associated observed customer counts (O) into dedicated columns. Next, calculate the total number of observations (N). In this specific business example, summing the daily counts (50 + 60 + 40 + 47 + 53) yields a grand total of **250 customers** observed throughout the week.

Since the shop owner claims a [uniform distribution](#)--meaning an equal likelihood of customers on any given day--the expected frequency (E) for each of the five categories is calculated by dividing the total number of customers by the number of categories: 250 total customers / 5 categories = **50 expected customers per day**. It is vital to clearly label the columns for Observed (O) and Expected (E) values to maintain transparency in the data structure, as shown below:

	A	B	C	D	E
1	Day	Observed	Expected		
2	Monday	50	50		
3	Tuesday	60	50		
4	Wednesday	40	50		
5	Thursday	47	50		
6	Friday	53	50		
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					

A critical note on Expected Counts: For a goodness of fit test where a uniform distribution is assumed, the expected count is simply the total sample size divided equally among the categories. However, if the null hypothesis proposed a different, non-uniform distribution (e.g., 20% traffic on

Monday, 30% on Tuesday, etc.), the expected counts (E) would be calculated by multiplying the total sample size (N) by those specific hypothesized proportions.

Step 2: Calculating the Chi-Square Test Statistic (X^2)

The central objective of the analysis is to calculate the **Chi-Square test statistic**, conventionally denoted as X^2 . This statistic provides a single numerical value that quantifies the overall magnitude of the difference between the observed data and the expected model. The mathematical foundation for this calculation is expressed by the following formula:

$$X^2 = \sum(O - E)^2 / E$$

The components of this formula are defined as follows:

Σ (Sigma): This symbol mandates summation. We must perform the calculation $(O-E)^2 / E$ for every single category and then sum these results to obtain the final X^2 value.

O: The **O**bserved frequency (the actual customer count) for a specific category (weekday).

E: The **E**xpected frequency for that same category, which is derived directly from the assumptions of the [null hypothesis](#).

To execute this efficiently within Google Sheets, a new column must be introduced specifically for the calculation of the component term: (Observed minus Expected, squared, divided by Expected). This step isolates the unique contribution of each category's deviation toward the final total test statistic. The procedure involves three sequential intermediate calculations for each row: (1) calculating the difference (O-E); (2) squaring that difference $(O-E)^2$ to ensure all values are positive and to exponentially penalize larger deviations; and (3) normalizing the squared difference by dividing by the expected value (E).

The following graphic illustrates the application of the formula $= (B2 - C2)^2 / C2$ for the first category (Monday) and shows how this formula is then copied down to calculate the individual contributions for all five weekdays:

	A	B	C	D	E
1	Day	Observed	Expected	(O-E)²/E	
2	Monday	50	50		<code>=(B2-C2)^2/C2</code>
3	Tuesday	60	50	2	
4	Wednesday	40	50	2	
5	Thursday	47	50	0.18	
6	Friday	53	50	0.18	
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					

Once these individual component values are calculated for all categories, the final X^2 [test statistic](#) is determined by using the `SUM()` function on the entire component column (Column D in the illustration above).

Step 3: Calculating the P-Value and Degrees of Freedom

Following the successful computation of the Chi-Square test statistic (X^2), the subsequent crucial step is to determine the associated [P-value](#). The P-value holds immense significance in hypothesis testing, as it represents the probability of observing data that is as extreme as, or even more extreme than, the calculated X^2 statistic, assuming that the [null hypothesis](#) is, in fact, true.

Google Sheets simplifies this calculation through the use of the powerful built-in function, `CHISQ.DIST.RT`. This function requires two essential inputs to return the P-value: the precise calculated X^2 value and the appropriate measure of [degrees of freedom](#) (df).

For the Chi-Square Goodness of Fit Test, the [degrees of freedom](#) are calculated straightforwardly as the total number of categories (n) minus one ($df = n - 1$). In the context of our retail example, where we analyze 5 distinct weekdays, the degrees of freedom are $5 - 1 = 4$.

Assuming, for instructional purposes, that the sum of the components calculated in Step 2 resulted in an X^2 value of 4.36, we would input the necessary formula into Google Sheets as illustrated

below:

	A	B	C	D	E	F	G
1	Day	Observed	Expected	(O-E)²/E			
2	Monday	50	50	0			
3	Tuesday	60	50	2			
4	Wednesday	40	50	2			
5	Thursday	47	50	0.18			
6	Friday	53	50	0.18			
7			X²	4.36	=SUM(D2:D6)		
8			p-value	0.3595	=CHISQ.DIST.RT(D7, COUNT(D2:D6)-1)		
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							

The `CHISQ.DIST.RT(x, deg_freedom)` function specifically calculates the right-tailed probability associated with the Chi-Square distribution, which is the standard approach for this type of hypothesis test. Using the resulting X^2 of 4.36 and 4 degrees of freedom, the formula executed is `=CHISQ.DIST.RT(4.36, 4)`. This calculation yields an X^2 [test statistic](#) of **4.36**, corresponding to an approximate [P-value](#) of **0.3595**.

Step 4: Interpreting the Results and Drawing a Conclusion

The final stage involves interpreting the calculated P-value by comparing it against a predefined significance level, typically denoted as alpha (α). In most scientific and business applications, alpha is set at 0.05, representing a 5% risk of incorrectly rejecting a true null hypothesis. This comparison dictates the ultimate statistical decision concerning the hypothesized [uniform distribution](#) of customer traffic.

The standard rules for statistical inference using the P-value method are clear and definitive:

If the P-value is less than the significance level ($P\text{-value} < 0.05$), we possess sufficient evidence to **reject the null hypothesis**. This signifies that the observed data significantly deviates from the expected distribution.

If the P-value is greater than or equal to the significance level ($P\text{-value} \geq 0.05$), we **fail to reject**

the null hypothesis. This indicates that the observed differences are likely attributable to normal, random chance.

In the case of our retail analysis, the calculated P-value of 0.3595 is substantially greater than the conventional alpha level of 0.05. Consequently, we must conclude that we **fail to reject the null hypothesis.**

This statistical conclusion suggests that there is insufficient evidence to assert that the true distribution of customer traffic differs from the shop owner's claim of a uniform daily distribution. The variations seen in the observed customer counts--such as the difference between 60 customers on Tuesday and 40 on Wednesday--are not statistically significant; they are most likely the product of expected, random daily fluctuations rather than an indication of a true, underlying difference in traffic patterns across the business week.

Conclusion: Mastering the Chi-Square Goodness of Fit

The [Chi-Square Goodness of Fit Test](#) is an invaluable statistical technique for comparing observed frequency data against theoretical expectations for a [categorical variable](#). Successful execution of this test in Google Sheets requires structured data organization, precise calculation of the X^2 test statistic using the summation formula, and the correct application of the `CHISQ.DIST.RT` function, always ensuring meticulous attention to the number of [degrees of freedom](#). Mastery of this process enables analysts to make robust, evidence-based statements about whether a sample distribution matches a hypothesized population model.

For professionals seeking to expand their analytical toolkit, exploring related statistical methodologies, such as the Chi-Square Test of Independence (which compares two categorical variables instead of one observed distribution against a theoretical one), offers further insights. Ultimately, the ability to correctly calculate and interpret the [P-value](#) relative to the significance level remains the cornerstone of sound statistical reporting.

Additional Resources for Advanced Statistical Analysis

To deepen your understanding of hypothesis testing and statistical inference, consider exploring official statistical documentation regarding the properties of the Chi-Square distribution and the theoretical limitations of using observed versus expected frequencies. Further learning on these topics ensures that the conclusions drawn from tools like Google Sheets are both mathematically sound and contextually appropriate for your data.