

# A Step-by-Step Guide to Chi-Square Goodness of Fit Tests in Excel

Authored by  
**Mohammed looti**

November 8, 2025

## RECOMMENDED CITATION

Mohammed looti (2025). *A Step-by-Step Guide to Chi-Square Goodness of Fit Tests in Excel*. PSYCHOLOGICAL STATISTICS. Retrieved from <https://statistics.arabpsychology.com/?p=13184>

The [Chi-Square Goodness of Fit Test](#) is a powerful and indispensable [non-parametric statistical](#) procedure used to determine if the observed frequency distribution of a [categorical variable](#) significantly deviates from a known or theoretically [hypothesized distribution](#). Fundamentally, this test allows researchers and analysts to assess whether the discrepancies between the data collected from a sample (observed data) and the data expected from a population (expected data) are merely the result of random chance or represent a genuine statistical misalignment.

This comprehensive guide is designed for professionals and students seeking to master the execution and interpretation of the **Chi-Square Goodness of Fit Test** utilizing the robust computational environment offered by [Microsoft Excel](#). By following these structured steps, you will ensure both the accuracy of your calculation and the clarity of your statistical conclusion.

## The Core Principles and Hypotheses of the Test

Before launching into data calculation, it is crucial to understand the conceptual framework underpinning the Goodness of Fit test. Statistical analysis frequently requires verifying whether a collected sample is statistically representative of a larger, predefined population model. This test specifically addresses this requirement by comparing actual observed counts across various categories against the counts that would be mathematically expected if the null hypothesis were perfectly true.

A key feature of the Chi-Square test is its status as a non-parametric test. This means it avoids making restrictive assumptions about the parameters (like the mean or standard deviation) of the population distribution, focusing instead solely on the counts or frequencies within discrete, mutually exclusive categories. This flexibility makes it widely applicable across various fields, from social sciences to quality control.

The formal establishment of hypotheses is the essential starting point for any inferential test. The [null hypothesis](#) ( $H_0$ ) always asserts that there is no meaningful difference between the observed distribution and the expected distribution--in essence, the categorical variable precisely follows the hypothesized model. Conversely, the **alternative hypothesis** ( $H_a$ ) contends that a significant difference exists, suggesting the variable does not conform to the hypothesized distribution. The ultimate goal of the Chi-Square procedure is to produce evidence that either supports the rejection of  $H_0$  or suggests insufficient evidence to reject it.

## Practical Application: Setting Up the Data for Analysis in Excel

To illustrate the practical steps involved in conducting this test, let us consider a common business scenario. Imagine a retail shop owner who claims that customer traffic is spread **uniformly** across the five business weekdays (Monday through Friday). To objectively test this assertion of equal distribution, an independent researcher tracks and records the total number of customers entering

the shop over the course of one full week. This raw data represents our observed frequencies.

The observed frequency data (O) collected by the researcher is summarized below:

**Monday:** 50 customers

**Tuesday:** 60 customers

**Wednesday:** 40 customers

**Thursday:** 47 customers

**Friday:** 53 customers

Our subsequent statistical objective is to utilize the **Chi-Square Goodness of Fit Test** to determine if this observed sample data provides compelling statistical evidence necessary to reject the shop owner's assertion of perfect uniformity. If the test yields a sufficiently large test statistic, implying a high degree of variation, we will be able to statistically contradict the claim of an equal customer volume distribution.

## Step 1: Inputting Observed and Calculating Expected Frequencies

The initial and most fundamental phase of any Chi-Square analysis involves structuring the raw data meticulously within the Excel environment. We must establish two distinct, category-aligned columns: one for the **Observed Frequencies** (O) and one for the **Expected Frequencies** (E). In our example, these categories correspond precisely to the five weekdays.

While the observed values (O) are straightforwardly transcribed from the collected data, the calculation of the **Expected values** (E) is critical, as they are derived directly from the assumption stated in the null hypothesis (H<sub>0</sub>). Since the shop owner hypothesized an equal, or uniform, distribution, we must calculate the total number of customers observed and distribute that total equally across all five days.

For this dataset, the total number of customers observed is calculated as:  $50 + 60 + 40 + 47 + 53 = 250$ . Given the claim of equality (uniform distribution), the expected value (E) for each individual day is thus calculated by dividing the total number of observations (250) by the total number of categories (5), yielding an expected count of 50 customers per day. We then structure our Excel worksheet to display the weekday categories, the observed counts, and the derived expected counts side-by-side:

	A	B	C	D	E	F
1	<b>Day</b>	<b>Observed</b>	<b>Expected</b>			
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						

It is paramount to note that the calculation of the **expected values** must precisely reflect the stated  $H_0$ . In this specific scenario, since the total sample size was 250 customers and the shop owner hypothesized an equal distribution across 5 days, the expected frequency (E) for each day is derived by dividing the total (250) by the number of categories (5), resulting in an expected count of 50 customers per day.

## Step 2: Determining the Chi-Square Test Statistic ( $X^2$ )

The central element of the test is the **Chi-Square test statistic** ( $X^2$ ), which quantifies the overall disparity between the observed data and the expected model. This statistic is computed by summing the standardized squared differences between the observed and expected values for every single category. The mathematical formula governing this calculation is:

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

Understanding the components of this formula is essential for accurate computation:

**$\Sigma$ :** This is the summation operator, directing us to calculate the value for each category and then aggregate the results into a single final test statistic.

**O:** The **Observed frequency** count for a specific category (e.g., the 60 customers recorded on Tuesday).

**E:** The **Expected frequency** count for that same category, derived under the assumption that the

null hypothesis is true (e.g., 50 customers for Tuesday).

To calculate the total  $X^2$  value efficiently, we introduce a new column in Excel dedicated to calculating the contribution of each category:  $(O - E)^2 / E$ . For instance, the calculation for Wednesday is  $(40 - 50)^2 / 50 = (-10)^2 / 50 = 100 / 50 = 2.0$ . Similarly, the calculation for Friday is  $(53 - 50)^2 / 50 = 9 / 50 = 0.18$ . By performing this calculation for all five weekdays, we isolate the variance attributed to each observation.

The following illustration demonstrates the necessary Excel formulas used to compute these individual contributions prior to summation:

	A	B	C	D	E	F	G
1	Day	Observed	Expected	$(O-E)^2/E$			
2	Monday	50	50	0	$=(B2-C2)^2/C2$		
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							

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

After calculating the contribution of each category, the final **Chi-Square test statistic** ( $X^2$ ) is obtained by summing the values in the contribution column. While this  $X^2$  value gives us the magnitude of the difference, it cannot be interpreted in isolation. To make a definitive statistical decision, this raw statistic must be converted into a corresponding **p-value**.

The p-value represents the probability of observing our calculated test statistic (or something even more extreme) if the null hypothesis were genuinely true. To calculate this probability, we must reference the theoretical Chi-Square distribution curve, which requires two inputs: the calculated  $X^2$  value and the appropriate **degrees of freedom** (df).

For the Chi-Square Goodness of Fit Test, the **degrees of freedom** (df) are calculated simply as the number of categories (n) minus one ( $df = n - 1$ ). Given that our example involves 5 distinct

categories (weekdays), the degrees of freedom are  $5 - 1 = 4$ . Excel significantly simplifies this final step through the use of the specialized function **CHISQ.DIST.RT**, which calculates the right-tailed probability (our required p-value) based on the test statistic and the df. The structure of the function is **CHISQ.DIST.RT(x, deg\_freedom)**.

The following visual demonstrates the culmination of the analysis, showing the summed  $X^2$  statistic and the direct calculation of the associated p-value:

	A	B	C	D	E	F
1	<b>Day</b>	<b>Observed</b>	<b>Expected</b>	<b>(O-E)<sup>2</sup>/E</b>		
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			<b><math>\chi^2</math></b>	4.36	=SUM(D2:D6)	
8			<b>p-value</b>	0.3595	=CHISQ.DIST.RT(D7, 4)	
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						

As illustrated, the final calculation uses the formula **=CHISQ.DIST.RT(G10, 4)**. Cell G10 holds the total calculated Chi-Square test statistic (4.36), and 4 is the calculated degrees of freedom for this five-category test.

#### Step 4: Interpreting the Final Results and Drawing Conclusions

The final phase of the hypothesis test involves synthesizing the calculated values--the  $X^2$  test statistic of **4.36** and the resulting p-value of approximately **0.3595**--and comparing the p-value against the established **significance level**, conventionally denoted as alpha ( $\alpha$ ). In most statistical contexts,  $\alpha$  is set at 0.05, representing a 5% risk of incorrectly rejecting a true null hypothesis.

The established decision rule for inferential statistical testing is as follows:

If the calculated p-value is less than or equal to  $\alpha$  (0.05), we conclude that there is sufficient evidence to **reject the null hypothesis** ( $H_0$ ). This result implies that the observed pattern is highly unlikely to have occurred by random chance alone, confirming a significant difference from the expected distribution.

If the calculated p-value is greater than  $\alpha$  (0.05), we consequently **fail to reject the null hypothesis** ( $H_0$ ). This result indicates that the observed deviations are statistically insignificant and are likely attributed to expected random sampling fluctuation, meaning we cannot conclude that the true distribution differs from the hypothesized one.

Applying this rule to our shop owner example, the calculated p-value (0.3595) is substantially larger than the standard significance level of 0.05. Therefore, based on the statistical evidence, we must **fail to reject the null hypothesis**. The formal conclusion is that the collected data does not provide sufficient statistical power to contradict the shop owner's original assertion of a uniform distribution of customer traffic across the weekdays. The observed variations in customer counts are not large enough to be considered anything more than typical random fluctuations.