

Perform a Chi-Square Goodness of Fit Test in SAS

Authored by
Mohammed loot

November 1, 2025

RECOMMENDED CITATION

Mohammed loot (2025). *Perform a Chi-Square Goodness of Fit Test in SAS*.
PSYCHOLOGICAL STATISTICS. Retrieved from
<https://statistics.arabpsychology.com/?p=7681>

The [Chi-Square Goodness of Fit Test](#) represents a core **statistical procedure** used widely across data analysis fields. Its primary function is to rigorously evaluate whether the observed frequency distribution of a single categorical variable aligns significantly with a predefined, [hypothesized distribution](#). This test is indispensable when researchers need to validate foundational assumptions regarding population parameters based solely on sample data collected from various sources.

For data scientists and statistical analysts, mastering the execution of this crucial test is paramount. The following guide provides a comprehensive, step-by-step example detailing how to perform the **Chi-Square Goodness of Fit Test** using the powerful statistical software, [SAS](#). We will walk through data preparation, code execution, and formal interpretation of the results generated by the SAS output.

Example Scenario: Testing Customer Traffic Uniformity in SAS

To illustrate the application of this test, we consider a practical scenario involving a retail business. A shop owner hypothesizes that customer traffic is distributed uniformly and equally across the five standard business weekdays: Monday through Friday. This belief forms the basis of our test's hypothesized uniform distribution--meaning 20% of customers should arrive on each day. To empirically test this claim against reality, a researcher carefully documents the total observed number of customers visiting the shop during one representative week.

The collection of observed frequencies, which will serve as our input data for analysis in [SAS](#), is summarized below:

Monday: 50 customers

Tuesday: 60 customers

Wednesday: 40 customers

Thursday: 47 customers

Friday: 53 customers

Our objective is to employ the [Chi-Square Goodness of Fit Test](#) within [SAS](#) to statistically determine if this observed customer data is consistent with the shop owner's assertion of an equal, 20% probability distribution across all weekdays. If the data significantly deviates from this expected pattern, the null hypothesis must be rejected.

Step 1: Creating and Structuring the Dataset in SAS

The foundational step in any statistical analysis within [SAS](#) requires defining the data structure accurately. Since the data is already aggregated--we possess observed counts (frequencies) for each categorical level (weekday)--we must input both the categorical variable and its

corresponding frequency count directly into the dataset. This contrasts with raw data where each row represents a single observation.

We initiate this process by utilizing the `DATA` step to create a dataset named `my_data`. The subsequent `INPUT` statement explicitly defines our two variables: `Day`, which is a character variable (indicated by the `$` sign) representing the weekday, and `Customers`, a numeric variable representing the observed frequency count. The actual data points are supplied immediately afterward using the `DATALINES` statement, ensuring clarity and transparency in the data entry process.

```
/*create dataset*/  
data my_data;  
input Day $ Customers;  
datalines;  
Mon 50  
Tue 60  
Wed 40  
Thur 47  
Fri 53  
;  
run;  
  
/*print dataset*/  
proc print data=my_data;
```

Executing the `PROC PRINT` procedure is a crucial quality assurance measure. It allows the researcher to visually confirm that the `my_data` dataset has been correctly populated, ensuring the categorical variable (`Day`) and the frequency variable (`Customers`) are properly aligned before proceeding to the computationally intensive statistical analysis phase.

Obs	Day	Customers
1	Mon	50
2	Tue	60
3	Wed	40
4	Thur	47
5	Fri	53

Step 2: Executing the Chi-Square Goodness of Fit Test in SAS

The [Chi-Square Goodness of Fit Test](#) is performed in [SAS](#) using the highly versatile `PROC FREQ` procedure. While `PROC FREQ` is primarily known for generating simple frequency tables, its power is extended through specific options to handle distributional tests on categorical data effectively.

Two elements are absolutely essential for executing the goodness of fit test correctly within this procedure: the `WEIGHT` statement and the `CHISQ` option. The `WEIGHT Customers` statement is critical because it instructs **SAS** that the values listed in the `Customers` column are not individual observations, but rather the pre-calculated, observed frequencies (counts) corresponding to each level of the `Day` variable. Failure to include this statement would lead to a completely erroneous calculation.

The command `TABLES Day / CHISQ;` specifies that the distribution of the `Day` variable is the focus of the analysis and requests the calculation of the **Chi-Square** statistic. Since we did not manually input expected probabilities (e.g., using a `TESTP` option), **SAS** automatically defaults to testing for an **equal distribution** across all specified categories. This default behavior perfectly matches our research hypothesis that customer traffic is uniform across the five weekdays.

```
/*perform Chi-Square Goodness of Fit test*/  
proc freq data=my_data;  
tables Day / chisq;  
weight Customers;  
run;
```

Upon execution, this code block generates a detailed output table. This table includes all necessary components for hypothesis testing, specifically the calculated **Chi-Square test statistic**, the corresponding degrees of freedom, and, most importantly, the associated [p-value](#) necessary for our formal decision-making process.

The FREQ Procedure

Day	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Fri	53	21.20	53	21.20
Mon	50	20.00	103	41.20
Thur	47	18.80	150	60.00
Tue	60	24.00	210	84.00
Wed	40	16.00	250	100.00

Chi-Square Test for Equal Proportions	
Chi-Square	4.3600
DF	4
Pr > ChiSq	0.3595

Interpreting the SAS Output and Hypothesis Formulation

Statistical testing fundamentally involves comparing empirical observations against a theoretical expectation defined by a set of competing hypotheses. For the [Chi-Square Goodness of Fit Test](#), these hypotheses are formally structured as follows:

H0 (Null Hypothesis): The categorical variable (customer traffic distribution) follows the specified hypothesized distribution (uniformity across weekdays).

HA (Alternative Hypothesis): The categorical variable (customer traffic distribution) does not follow the specified hypothesized distribution.

We must now carefully scrutinize the output generated by [SAS](#) within the context of this framework. The decision to reject or fail to reject the [null hypothesis](#) hinges upon two critical values extracted from the results:

The calculated **Chi-Square test statistic** (χ^2): **4.36**

The corresponding **p-value** (labeled Pr > ChiSq in the output): **0.3595**

The **Chi-Square statistic** (4.36) quantifies the overall magnitude of the difference between the observed customer counts and the counts we would theoretically expect if the distribution were perfectly uniform. While a larger value indicates greater deviation, the definitive statistical conclusion is always governed by the **p-value**. The **p-value** represents the probability of observing our collected data (or data more extreme than ours) assuming that the [null hypothesis](#) (H_0) is, in fact, true.

Drawing the Statistical Conclusion

The final step in hypothesis testing involves comparing the calculated probability value against a pre-established significance threshold, conventionally known as α . For the vast majority of statistical investigations, this significance level is set at $\alpha = 0.05$. This 0.05 threshold defines the maximum acceptable risk of making a Type I error--the error of incorrectly rejecting a true **null hypothesis**.

The decision rule is unambiguous: If the calculated **p-value** is less than or equal to 0.05, we possess sufficient evidence to reject the **null hypothesis** (H_0). Conversely, if the **p-value** exceeds 0.05, we must consequently fail to reject H_0 .

In this analysis of customer traffic, the calculated **p-value** (0.3595) is substantially greater than the standard α level of 0.05. Therefore, the statistical imperative is to **fail to reject the null hypothesis**. This crucial finding signifies that, based on the observed weekly customer counts, there is insufficient statistical evidence to conclude that the true distribution of customer traffic across the weekdays deviates significantly from the uniform distribution claimed by the shop owner.

In practical terms, the minor fluctuations observed in the daily counts (such as Tuesday's high of 60 versus Wednesday's low of 40) are statistically attributable to normal random sampling variation rather than an underlying, systematic non-uniform pattern in customer behavior. The shop owner's assertion of uniform traffic distribution is therefore statistically supported by this sample data.

Resources for Advanced Chi-Square Analysis

For researchers and analysts seeking a deeper theoretical grasp or further practical examples related to the **Chi-Square Goodness of Fit Test**, continuous study is highly recommended. These additional resources can provide valuable context on essential concepts such as manually calculating expected values, accurately determining degrees of freedom, and executing tests for non-uniform hypothesized distributions where specific proportions (rather than equal proportions) are expected across categories.