

# G-test of Goodness of Fit: Definition + Example?

Authored by  
**stats writer**

December 6, 2025

## RECOMMENDED CITATION

stats writer (2025). *G-test of Goodness of Fit: Definition + Example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=106504>

The G-test of Goodness of Fit is a powerful statistical test used to determine whether observed frequencies differ significantly from expected frequencies in a given data set. It is an alternative to the chi-squared test and is primarily employed to compare observed data distributions against hypothesized or theoretical distributions. For instance, in a retail environment, if management hypothesizes an equal preference for five different colors of a new product, the G-test can be utilized to compare the actual observed sales figures (frequencies) of each color against the expected sales figures, thereby assessing the validity of the initial hypothesis.

## Understanding the G-test of Goodness of Fit

In statistics, the **G-test of Goodness of Fit** is fundamentally used to ascertain whether a specific categorical variable follows a predefined or hypothesized distribution. This test is essential for researchers across various disciplines--from biology and economics to marketing--whenever there is a need to validate a theoretical expectation against real-world observations. The core utility of the G-test lies in its ability to evaluate discrepancies between what we expect to see and what we actually record in our sample data.

The test derives its statistical power from the maximum likelihood theory, making it particularly adept at handling complex data sets. Unlike some simpler frequency comparison tests, the G-test uses logarithms to quantify the discrepancy, providing a statistic that is asymptotically distributed according to the chi-squared distribution, which simplifies the determination of the associated p-value and the final decision regarding the hypothesis.

## G-test vs. The Chi-Squared Test: Key Differences

The G-test of Goodness of Fit often serves as a robust alternative to the traditional chi-squared test (Pearson's Chi-Squared Test). While both tests evaluate the same underlying null hypothesis--that the observed data fit the expected distribution--the G-test is preferred in several important scenarios. Specifically, the G-test is based on the likelihood ratio statistic, which often exhibits better performance when dealing with extremely large data sets or when analyzing data where outliers might exert undue influence on the sum of squared differences inherent in the chi-squared calculation.

Historically, the chi-squared test was favored due to the ease of manual calculation. However, with the widespread availability of computational tools, the G-test has gained prominence because its test statistic, derived from natural logarithms, provides a more theoretically sound measure of difference based on information theory. Furthermore, the G-test is particularly advantageous in situations where log-linear models are employed, as the statistic generated by the G-test is directly related to these complex models, making it a versatile tool for advanced statistical analysis.

## Formulating the Hypotheses for the G-test

Like all frequentist hypothesis tests, the G-test requires the formulation of two competing statements: the null hypothesis and the alternative hypothesis. These hypotheses formally define the relationship between the observed data and the theoretical distribution we are testing. Establishing these statements clearly is the critical first step in performing any G-test.

The G-Test of Goodness of Fit uses the following null and alternative hypotheses:

**H0 (Null Hypothesis):** The variable follows the specified hypothesized distribution. This is the statement of no difference or no effect that the researcher aims to disprove.

**HA (Alternative Hypothesis):** The variable *does not* follow the specified hypothesized distribution. This hypothesis suggests that there is a statistically significant difference between the observed frequencies and the expected frequencies.

The goal of the test is to gather sufficient evidence from the sample data to determine whether the deviation from the expected frequencies is large enough to reject the null hypothesis (H0) in favor of the alternative hypothesis (HA).

## The Mathematical Foundation: Calculating the G Statistic

The core of the G-test lies in calculating the G statistic, which measures the log-likelihood ratio between the observed data and the expected model. This statistic quantifies the degree of divergence from the hypothesized distribution. The larger the G value, the greater the evidence against the null hypothesis.

The test statistic is calculated using the following general formula, which incorporates the natural logarithm (ln):

$$G = 2 * \sum$$

where the summation ( $\Sigma$ ) is carried out across all categories or cells in the data:

**O:** Represents the **observed count** (frequency) recorded in a specific cell or category.

**E:** Represents the **expected count** (frequency) in that same cell, calculated based on the assumed distribution specified in the null hypothesis.

The multiplication by two ensures that the resulting statistic approximates the chi-squared distribution, allowing for standard statistical tables or software functions to be used for determining the p-value. It is crucial that both O and E are defined accurately across all categories to ensure the validity of the G statistic calculation.

## Practical Application: A Biological Example

To illustrate the practical use of the G-test, consider a scenario involving population ecology. A biologist claims that three distinct species of turtles (Species A, B, and C) exist in equal proportions within a specific wetland area. To verify this assertion, an independent researcher conducts a comprehensive survey, meticulously counting the number of each species encountered.

The observed counts are recorded as follows:

Species A: **80** individuals

Species B: **125** individuals

Species C: **95** individuals

The researcher must now employ the G-test of Goodness of Fit to determine if this collected data is statistically consistent with the biologist's claim of an equal distribution. Since there are three categories (species), and the null hypothesis dictates an equal proportion, the expected distribution (E) for each species must be calculated based on the total observed count ( $N = 80 + 125 + 95 = 300$ ).

### Step-by-Step Analysis of the Turtle Data

#### Step 1: State the Null and Alternative Hypotheses

The researcher begins by formally defining the hypotheses based on the biologist's claim:

**H<sub>0</sub>:** An equal proportion (1/3 for each) of the three species of turtles exists in this area.

**H<sub>A</sub>:** An equal proportion of the three species of turtles *does not* exist in this area (i.e., the population distribution is significantly different from 1:1:1).

#### Step 2: Calculate the G Test Statistic

The formula used to calculate the test statistic remains:

$$G = 2 * \sum$$

In this specific example, there are 300 total observed turtles. If the proportions were truly equal (as per H<sub>0</sub>), we would expect 100 turtles from each species ( $E = 300/3 = 100$ ). We can now substitute the observed (O) and expected (E) counts into the formula:

$$G = 2 *$$

Performing the logarithmic calculations and summation yields the final test statistic: **G = 10.337**. This numerical value represents the degree of deviation of the observed data from the expected

equal proportions.

## Interpreting the P-Value and Drawing Conclusions

### Step 3: Calculate the P-value of the Test Statistic

Once the G statistic is calculated, we must determine the probability of observing such a value (or more extreme) if the null hypothesis were true. This probability is the p-value. The G statistic follows a chi-squared distribution with specific degrees of freedom (df), calculated as the number of categories minus one. In this case,  $df = 3 \text{ categories} - 1 = 2$ .

According to the chi-squared distribution table (or statistical software), the p-value associated with a G test statistic of 10.337 and 2 degrees of freedom is calculated to be **0.005693**.

To make a decision, this p-value must be compared to a predefined significance level (alpha, typically set at 0.05). Since 0.005693 is significantly less than 0.05, the researcher has sufficient statistical evidence to reject the null hypothesis. This conclusive rejection implies that the observed proportions of the three turtle species are statistically different from the hypothesized equal proportion (1:1:1) in this particular area.

### Implementing the G-test in R (Bonus Section)

For efficiency and accuracy, researchers commonly utilize statistical programming languages like R to perform the G-test. The **GTest()** function, available within the widely used DescTools package, allows for the rapid execution of the G-test of Goodness of Fit, bypassing the need for manual logarithmic calculations.

The following code demonstrates how to replicate the analysis for the turtle example within the R environment:

```
#load the DescTools library
```

```
library(DescTools)
```

```
#perform the G-test
```

```
GTest(x = c(80, 125, 95), #observed values (O)
```

```
p = c(1/3, 1/3, 1/3), #expected proportions (1/3 for equal distribution)
```

```
correct = "none")
```

```
Log likelihood ratio (G-test) goodness of fit test
```

```
data: c(80, 125, 95)
```

```
G = 10.337, X-squared df = 2, p-value = 0.005693
```

The output confirms that the calculated G test statistic is precisely **10.337** and the corresponding p-value is **0.005693**. Since this p-value is smaller than the conventional alpha level of 0.05, the software output leads to the identical conclusion: rejection of the null hypothesis. This automated process validates the manual calculations and streamlines the workflow for analyzing larger and more complex categorical variable datasets.

## Summary and Additional Resources

The G-test of Goodness of Fit is an invaluable tool for testing distributional hypotheses when dealing with frequency data. Its reliance on the log-likelihood ratio makes it a statistically powerful and efficient alternative to the standard chi-squared test, particularly in modern computational statistics.

Feel free to utilize specialized calculators or statistical software to automatically perform a G-test for any dataset, ensuring that the necessary observed counts and expected proportions are accurately defined. Consulting reliable statistical resources is recommended for further understanding the theoretical underpinnings of likelihood ratio tests.