



A menu icon consisting of three horizontal grey bars of decreasing length from left to right.

OCTOBER 21, 2020 BY ZACH

How to Perform a Chi-Square Goodness of Fit Test in R

A **Chi-Square Goodness of Fit Test** is used to determine whether or not a categorical variable follows a hypothesized distribution.

This tutorial explains how to perform a Chi-Square Goodness of Fit Test in R.

Example: Chi-Square Goodness of Fit Test in R

A shop owner claims that an equal number of customers come into his shop each weekday. To test this hypothesis, a researcher records the number of customers that come into the shop in a given week and finds the following:

- **Monday:** 50 customers
- **Tuesday:** 60 customers
- **Wednesday:** 40 customers
- **Thursday:** 47 customers
- **Friday:** 53 customers

Use the following steps to perform a Chi-Square goodness of fit test in R to determine if the data is consistent with the shop owner's claim.

Step 1: Create the data.

First, we will create two arrays to hold our observed frequencies and our expected proportion of customers for each day:

```
observed <- c(50, 60, 40, 47, 53)
expected <- c(.2, .2, .2, .2, .2) #must add up to 1
```

Step 2: Perform the Chi-Square Goodness of Fit Test.

Next, we can perform the Chi-Square Goodness of Fit Test using the **chisq.test()** function, which uses the following syntax:

chisq.test(x, p)

where:

- **x**: A numerical vector of observed frequencies.
- **p**: A numerical vector of expected proportions.

The following code shows how to use this function in our example:

```
#perform Chi-Square Goodness of Fit Test
chisq.test(x=observed, p=expected)

Chi-squared test for given probabilities

data: observed
X-squared = 4.36, df = 4, p-value = 0.3595
```

The Chi-Square test statistic is found to be **4.36** and the corresponding p-value is **0.3595**.

Note that the p-value corresponds to a Chi-Square value with $n-1$ degrees of freedom (dof), where n is the number of different categories. In this case, dof = 5-1 = 4.

You can use the [Chi-Square to P Value Calculator](#) to confirm that the p-value that corresponds to $X^2 = 4.36$ with dof = 4 is **0.35947**.

Recall that a Chi-Square Goodness of Fit Test uses the following null and alternative hypotheses:

- **H_0 : (null hypothesis)** A variable follows a hypothesized distribution.
- **H_1 : (alternative hypothesis)** A variable does not follow a hypothesized distribution.

Since the p-value (.35947) is not less than 0.05, we fail to reject the null hypothesis. This means we do not have sufficient evidence to say that the true distribution of customers is different from the distribution that the shop owner claimed.

Additional Resources

[How to Perform a Chi-Square Test of Independence in R](#)

[How to Calculate the P-Value of a Chi-Square Statistic in R](#)

[How to Find the Chi-Square Critical Value in R](#)



Published by Zach