****

9B18E014

chi-square GOODNESS-OF-FIT TEST AND ITS USAGE IN EXCEL

Kyle Maclean wrote this note solely to provide material for class discussion. The author does not intend to provide legal, tax, accounting or other professional advice. Such advice should be obtained from a qualified professional.

*This publication may not be transmitted, photocopied, digitized, or otherwise reproduced in any form or by any means without the permission of the copyright holder. Reproduction of this material is not covered under authorization by any reproduction rights organization. To order copies or request permission to reproduce materials, contact Ivey Publishing, Ivey Business School, Western University, London, Ontario, Canada, N6G 0N1; (t) 519.661.3208; (e)* [*cases@ivey.ca*](mailto:cases@ivey.ca)*;* [*www.iveycases.com*](http://www.iveycases.com)*.*

Copyright © 2018, Ivey Business School Foundation Version: 2018-09-14

The purpose of this note is to explain the application areas of the chi-square goodness-of-fit test and how to run the test in Microsoft Excel. Upon completion of this note and its associated activities, users will

* understand how a chi-square goodness-of-fit test is useful in business; and
* be able to conduct a chi-square goodness-of-fit test in Excel manually or by using an automatic function in Excel.

**CHI-Square GOODNESS-OF-FIT TEST[[1]](#footnote-2)**

The chi-square goodness-of-fit test is used to determine a *statistically significant* difference between two distributions. The term “goodness-of-fit” is used because we want to measure how well a hypothetical distribution fits with an observed distribution. In practice, this test is often used to compare *expected* frequencies and *observed* frequencies in one or more categories. For instance, suppose we flip a quarter 100 times. We count that it landed 55 times on “heads” and 45 times on “tails.” A keen gambler might speculate that this coin is unbalanced or weighted toward heads. If so, this might represent a profitable betting opportunity. However, the gambler also knows that pure luck may have caused an imbalance in outcomes. A chi-square goodness-of-fit test is a useful way of determining statistically if sufficient evidence exists to reject the hypothesis that the coin is balanced.

The chi-square goodness-of-fit test has many application areas in business and particularly in sports and entertainment, such as the following two uses:

* To determine if a process is fair (e.g., “Is a roulette wheel unbalanced?”)
* Using survey data, to determine if different customer segments have statistically different preferences (e.g., “Do frequent movie-goers prefer higher budget films?”)

**Running a chi-square Goodness-of-fit test[[2]](#footnote-3)**

To run a chi-square goodness-of-fit test, we must have both observed counts and expected counts for each cell or outcome that we are concerned about. If our data are in percentages, they must first be transformed into counts. Observed frequencies are simply the actual data that have been observed. Expected frequencies must be calculated based on some null hypothesis. A null hypothesis should be explicit; without one, we cannot find the expected frequencies. Usually, a null hypothesis is that the data are independent or the process is random, or something similar.

In the case of the potentially unbalanced quarter, we have the observed amounts below:

|  |  |
| --- | --- |
| **Observed** | |
| **Outcome** | **Count** |
| Heads | 55 |
| Tails | 45 |

We must calculate the expected frequencies under some null hypothesis. For the quarter being flipped, our null hypothesis is that the quarter is balanced and that there is a 50-per-cent chance of flipping heads and a 50-per-cent chance of flipping tails. In total, we know that there were 100 flips (45 + 55 = 100). Under the null hypothesis, we can then calculate the expected number for each outcome as noted below:

|  |  |
| --- | --- |
| **Expected** | |
| **Outcome** | **Count** |
| Heads | (100 × 50%) = 50 |
| Tails | (100 × 50%) = 50 |

A good rule of thumb is that after calculating expected counts, there should be at least five in each cell, otherwise we do not have enough data. Once we have the expected and observed quantities, we calculate the following expression for each cell:

The sum of the numbers will give us the chi-square statistic. A relatively small chi-square statistic indicates that our observed data are close to the expected distribution. A relatively large chi-square statistic indicates that it is different (and thus that our null hypothesis can probably safely be rejected).

In this case, we perform the above formula for each cell: and . After adding the two results, we obtain a chi-square statistic of 1. We are interested in how extreme this statistic is and what the probability is of obtaining a result as extreme, or more so, by chance. To determine this, we can use the CHISQ.DIST (test statistic, degrees of freedom, cumulative?) function in Excel.

CHISQ.DIST is an Excel function that returns the distribution of chi-square statistic.[[3]](#footnote-4) The first argument to the function is the chi-square statistic we just generated. The second argument is the “degrees of freedom.” The degrees of freedom can be conceptualized as the number of observations that can vary. In this case, we know that there were 100 flips of the coin; if we know that there were 70 “heads” results, we also know that there must have been 30 “tails” results. So, there is only one degree of freedom in our case. In examples like this, with a single column of observations, the degrees of freedom are simply the number of observations minus 1. In tables that are two-by-two or larger, the degrees of freedom can be calculated using the following equation:

The third argument asks whether we are interested in the cumulative density, with TRUE telling the function to return the cumulative density and FALSE otherwise. Recall that we are interested in finding the probability of obtaining a test statistic “as extreme or more” as the one we found (i.e., the right-hand side of the distribution). To find the area to the right, we can subtract the area to the left from one (i.e., =1-CHISQ.DIST; test statistic, degrees of freedom, TRUE). In this example, we key the following details into Excel, and we obtain the result shown:

=1-CHISQ.DIST(1, 1, TRUE) 🡪 0.3173

In plain language, we have an approximately 32-per-cent chance of obtaining results this different by chance. Note that the first two arguments are one, but this is merely by coincidence. The first value is the test statistic obtained, while the second value is the degrees of freedom. By statistical convention, we reject the null hypothesis if the value obtained is less than 5 per cent. In this case, the value is far greater, so we cannot reject the null hypothesis that the coin is evenly weighted, and it would most likely be unwise to assume that it is.

Let us try an example with multiple rows and columns. Suppose that we have obtained a customer survey of preferences toward two television shows: Show A and Show B. Respondents chose which of the two shows they preferred and selected whether they were a heavy television user (defined as an average viewing time of two or more hours per day) or a light television user (defined as an average viewing time of less than two hours).

|  |  |  |
| --- | --- | --- |
| **Observed** | | |
|  | **Heavy** | **Light** |
| Show A | 83 | 42 |
| Show B | 37 | 38 |

A manager might be interested in whether there is statistical evidence that heavy television users have different preferences than light television users. To do this, recall that we need both observed (which is given above) and expected frequencies. The null hypothesis would be that there is no difference between heavy television users and light television users. To find the expected amount, subtotalling the categories would be helpful. For instance, a total of 125 respondents selected Show A. The table below shows all subtotals:

|  |  |  |  |
| --- | --- | --- | --- |
| **Observed** | | | |
|  | **Heavy** | **Light** | **Subtotal** |
| **Show A** | 83 | 42 | 125 |
| **Show B** | 37 | 38 | 75 |
| **Subtotal** | 120 | 80 | 200 |

After the subtotals are calculated, we can determine the frequency under the null hypothesis. Of the complete sample of television users, 60 per cent (120 ÷ 200) are heavy users. Also, in the complete sample, 62.5 per cent (125 ÷ 200) preferred Show A. Under the null hypothesis, that there is no difference in preferences between heavy and light users, we would expect that of the complete sample, 75 heavy users prefer Show A, calculated as follows: 200 × 60% × 62.5% = 75. We then perform this calculation for all cells below. A mathematical trick that can be used is to multiply the two subtotals together and divide by the total. This is mathematically identical, but may be easier to remember. The mathematical trick is shown for the light users below.

|  |  |  |
| --- | --- | --- |
| **Calculating the Chi-Square Statistic** | | |
| **Outcome** | **Heavy** | **Light** |
| Show A | (83-75)^2 / 75 = 0.853 | (42-50)^2 / 50 = 1.28 |
| Show B | (37-45)^2 / 45 = 1.42 | (38-30)^2 / 30 = 2.13 |

To calculate the test statistic, we must calculate the formula above for each individual cell. We show this process below.

The sum of all four cells is approximately 5.69. The number of degrees of freedom is (2 – 1) × (2 – 1) = 1. Recall that a *p-* value tells us the probability of obtaining a result as extreme or more under the null hypothesis. With this information, we can now use the CHISQ.DIST function to find the *p*- value as follows: =1-CHISQ.DIST(5.69, 1, TRUE) 🡪 0.017.

Using the statistical convention of a 5-per-cent rejection threshold, we note that 1.7 per cent is less than 5 per cent, and we can reject the null hypothesis and conclude that heavy and light television users do have

different preferences.

|  |  |  |  |
| --- | --- | --- | --- |
| **Expected** | | | |
|  | **Heavy** | **Light** | **Subtotal** |
| **Show A** | 200(120 / 200) (125 / 200) = 75 | 80 125 / 200 = 50 | 125 |
| **Show B** | 200(120 / 200) (75 / 200) = 45 | 80 75 / 200 = 30 | 75 |
| **Subtotal** | 120 | 80 | 200 |

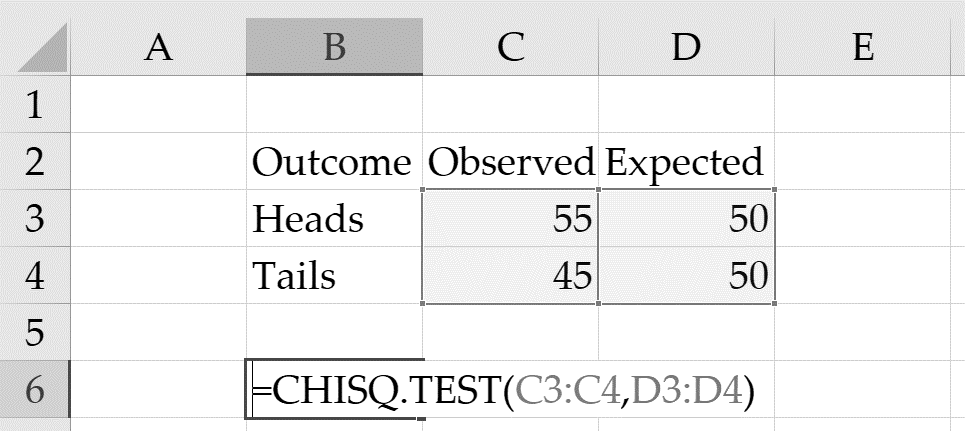
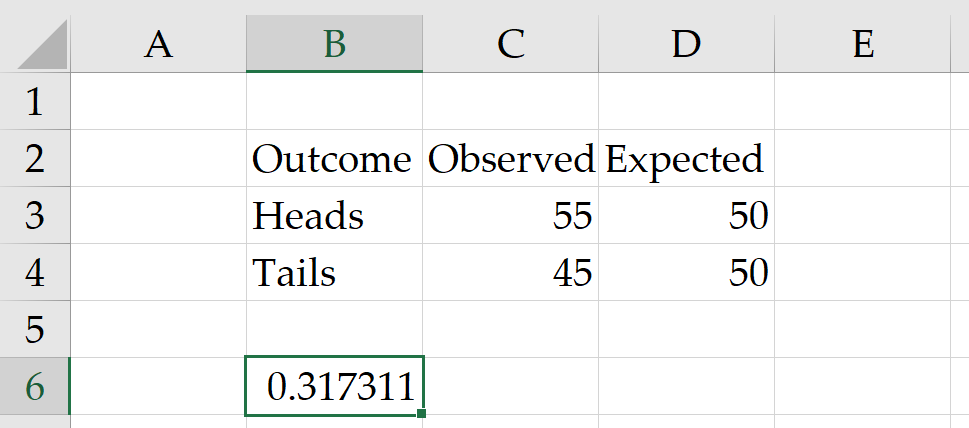
**CHISQ.TEST (Actual Range, Expected Range)** **EXCEL FUNCTION**

The preceding examples and discussion showed how the chi-square goodness-of-fit test can be calculated semi-manually. For Excel 2013, Microsoft added a new function that automates much of this process.

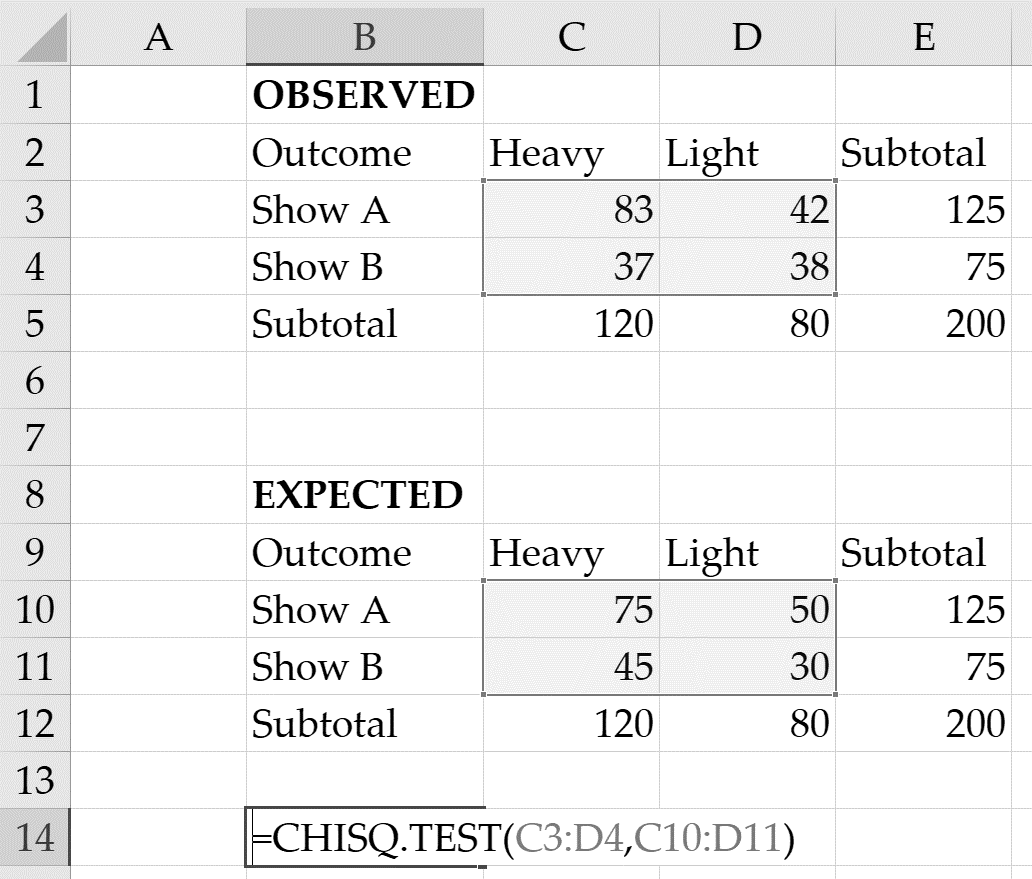
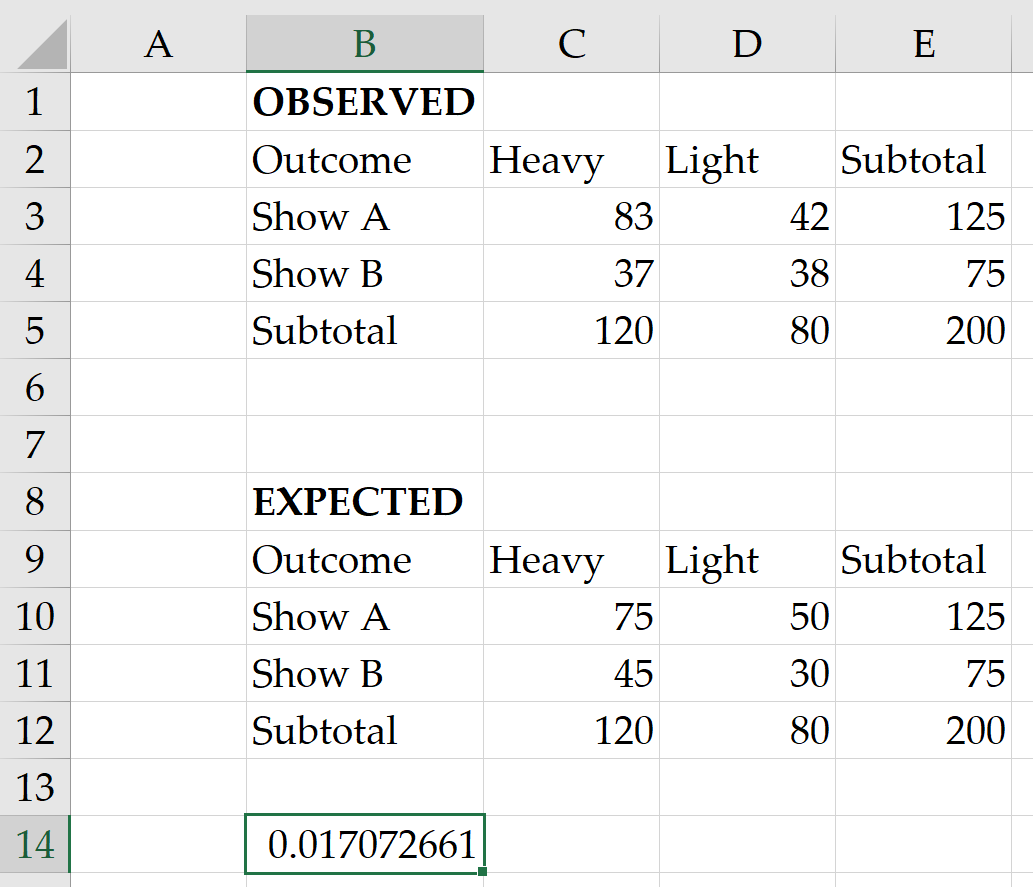
CHISQ.TEST calculates the *p*-value automatically by providing both the actual data and the expected data.[[4]](#footnote-5)

The CHISQ.TEST function is sensitive to the “shape” of data you provide; this is how it determines the degrees of freedom. Consequently, double-checking your results by conducting the calculation manually is recommended. Below, we show how the function can be used for the two examples above, noting that the results obtained are identical to those found manually:

Coin Example

🡪 

Television Example

🡪 

**SUMMARY**

The purpose of this note was to illustrate the use of chi-square goodness-of-fit tests for business purposes and how to run the tests in Excel. By using two examples, we showed how to manually calculate the test statistic, how to interpret the resulting *p*-values, and how to calculate the *p*-value automatically using a recently-added Excel function.

1. R. A. Fisher, “On the Interpretation of X2 from Contingency Tables, and the Calculation of P,” *Journal of the Royal Statistical Society* 85, no. 1 (January 1922): 87–94. [↑](#footnote-ref-2)
2. Phillip E. Pfeifer, *Chi-square Goodness-of-Fit Test* (Charlottesville, VA: Darden Business Publishing, 2006). Available from Ivey Publishing, product no. UVAQA0692. [↑](#footnote-ref-3)
3. “CHISQ.DIST Function,” Microsoft: Support, accessed July 23, 2018, https://support.office.com/en-us/article/chisq-dist-function-8486b05e-5c05-4942-a9ea-f6b341518732. [↑](#footnote-ref-4)
4. “CHISQ.TEST Function,” Microsoft: Support, accessed July 23, 2018, https://support.office.com/en-us/article/chisq-test-function-2e8a7861-b14a-4985-aa93-fb88de3f260f. [↑](#footnote-ref-5)