

Chi-square Goodness of Fit Test

The chi-square test is designed to test differences whether one frequency is different from another frequency. The chi-square test is designed for use with data on a nominal scale, when all you know about people is the category they're in.

The chi-square test gives you a way of comparing the results that you should get if the number of passengers has no effect on the frequency of stopping to the results you actually got.

For example:

Having a coin come up heads or tails is a variable on a nominal scale. Heads is a different category from tails.

If we flipped a coin 20 times how many times **should** it come up tails? [10]

Say the actual number of tails was 18. Is this number far enough away from 10 for you to be suspicious that the coin was rigged. The chi-square test tests whether the odds are less than 5% of getting 18 tails when you know that 10 tails is what you would expect to get.

What we're going to do is to calculate a value for the chi-square statistic and then compare it to a value for chi-square that we can look up in the back of the stats book. That number will tell you how large the calculated value has to be in order to be confident that the result was not just due to chance.

To calculate the value for chi-square for the coin data you need to know what frequencies you would expect to get. 10 heads and 10 tails. You know that the actual values were 2 heads and 18 tails.

To get chi-square:

1. For every possible outcome, subtract the expected frequency from the observed, or actual, frequency.

$$2 - 10 = -8$$

$$18 - 10 = 8$$

2. Now square each of these differences:

$$64$$

$$64$$

3. Divide each of these squared differences by the original expected frequency for that cell:

$$64/10 = 6.4$$

$$64/10 = 6.4$$

4. Add these numbers up. This gives you the value for chi-square. [12.8]

The equation for calculating Chi-square is thus...

$$X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

The steps for calculating in table for are...

	f(o)	f(e)	f(o) - f(e)	[f(o) - f(e)] ²	[f(o) - f(e)] ² / f(e)
Heads	2	10	-8	64	6.4
Tails	18	10	+8	64	6.4
					X ² = 12.8

To get the critical value for Chi-square:

1. You need to know the number of degrees of freedom. This is equal to number of possible categories minus one. [2-1 = 1]

2. You need to know which column to look the value up in. If you are going to use 5% as the odds of the frequencies being different by chance, look the number up in the .05 column.

[3,48]

In this example, the number for chi-square that we calculated was larger than the comparison number we looked up in the book. This tells us that there is less a 5% chance that the percentages are different by chance. So you might decide that the coin is rigged.

APA format for writing the conclusion statement is:

The frequencies of the "tails" option in the coin flips is significantly greater from what one would expect to get by chance, $X^2(1, N = 20) = 12.8, p < .05$.

A Chi-square test when there is one variable is often referred to as a **Goodness of Fit Test**.

Chi-Square Test of Independence

Now let's say we observe drivers at an intersection with a stop sign. We record two pieces of information: whether the driver stops at the stop sign and whether the driver is talking on a cell phone. The question we want to answer is whether drivers are less likely to come to a complete stop at the stop sign when talking on a cell phone compared to when drivers are not talking on a cell phone.

Here are the data...

	Cell phone	No cell phone	
stop	5	15	20
don't stop	20	10	30
	25	25	N=50

The question is really whether the **percentage** of drivers who stopped while talking on a cell phone is significantly different from the **percentage** of drivers who stopped while not talking on a cell phone.

Obviously, 5/25 or 20% of drivers with no passengers came to a complete stop. 15/25, or 60% of drivers with one or more passengers came to a complete stop.

The chi-square test will still tell us if the observed frequencies are significantly different from the expected frequencies. We can use the numbers in the row and column margins to help us to compute the expected frequencies.

The expected frequency for each cell is computed using the following formula:

$$f(e) = [\text{row total})(\text{column total})] / N$$

For the top left cell, the expected frequency is:

$$(20)(25) / 50 = 500/50 = 10$$

top right:

$$(20)(25) / 50 = 500/50 = 10$$

bottom left:

$$(30)(25) / 50 = 750/50 = 15$$

bottom right:

$$(30)(25) / 50 = 750/50 = 15$$

- Once you have the expected frequencies, the calculation for chi-square is exactly the same except that you have four rows in the calculations instead of two.

	$f(o) - f(e)$	$[f(o) - f(e)]^2$	$[f(o) - f(e)]^2 / f(e)$
Stop – no pass	$5 - 10 = -5$	25	2.5
Stop -- pass	$5 - 10 = 5$	25	2.5
No stop – no pass	$20 - 15 = 5$	25	1.67
No Stop -- pass	$10 - 15 = -5$	25	1.67
			8.34

You compare this observed value for chi-square against a comparison value you look up in the chi-square table.

You still use the .05 column because we're using 5% as the odds we're using to make our decision. The number of degrees of freedom to use is found from the following equation:

$$df = (\# \text{ of rows} - 1)(\# \text{ of columns} - 1)$$

In this case we have $(2-1)(2-1) = (1)(1) = 1$, giving us a comparison value of 3.84.

We can say that the percentages found in the first column are significantly different from the percentages found in the second column.

Drivers on a cell phone are significantly less likely to stop at a stop sign than drivers not talking on a cell phone, $\chi^2 (1, N = 50) = 8.34, p < .05$.

Chi-Square Practice Problems

For each of the problems below, ...

- Please state both the null and alternative hypotheses for this question.
- Provide the decision rule for making this decision. Use an alpha level of .05.
- Show all of the work necessary to calculate the appropriate statistic.
- What conclusion are you allowed to draw? Write a conclusions sentence in APA format (i.e., it includes the appropriate statistical information).

- From Thorne (1989): At a state university, the student population is approximately one-third male and two-thirds female. Over a two-day period the gender of each student entering the student union is recorded, with the following results: males, 452; females, 1548. Determine whether males and females are as likely to enter the union as would be predicted on the basis of their percentage in the student population.
- A chimpanzee was trained to make same-different judgments about pairs of stimuli. For pictures of objects, the animal was correct on 23 out of 24 trials. Did the chimp perform significantly better than chance?
- From Sprinthall (1990). A researcher is interested in whether or not a significant trend exists regarding the popularity of certain work shifts among police officers. A random sample of 60 police officers is selected from a large metropolitan police force. The officers are asked to indicate which of three work shifts they preferred. The results show that 40 officers prefer the first shift, 10 prefer the second shirt, and 10 prefer the third shift. Do the results deviate significantly from what would be expected due to chance?
- A professor wants to determine whether her department should keep the requirement of college algebra as a prerequisite for an Introductory Statistics course. Accordingly, she allows some students to register for the course on a pass-fail basis regardless of whether or not they have had the prerequisite. Of the 70 students in the class, 40 have had algebra and 30 have not. At the end of the semester, the professor compares the number of students passing or failing the class with whether or not they had algebra. The results are presented below. Are students more likely to pass the course if they have taken college algebra?

	Pass	Fail
Algebra	34	6
No algebra	12	18

- In a study of intraspecific aggression (aggression directed toward other members of the

same species), an experimenter finds that 16 out of 23 animals tested in species A exhibit aggression, while only 6 of 25 are aggressive in species B. Do the two species differ significantly in terms of intraspecific aggression?

6. In a study of the effectiveness of an antipsychotic drug, patients treated with the drug were compared to patients receiving a placebo. In terms of the number relapsing, 698 of 1,068 patients relapsed after taking the placebo while 639 out of 2,127 patients relapsed after taking the antipsychotic drug. Test the prediction that the antipsychotic is significantly more effective in preventing relapse than the placebo.

Answers:

1. HO: Males and females are equally likely to enter the student union.
H1: Males and females are not equally likely to enter the student union.

Decision rule: If $X^2 \geq +3.84$ or if $X^2 \leq -3.84$, reject HO.

Conclusion sentence: Males and females are not equally likely to enter the student union, $X^2(1, N = 2000) = 103.68, p < .05$.

2. HO: The chimp did not perform significantly better than chance.
H1: The chimp performed significantly better than chance.

Decision rule: If $X^2 \geq +3.84$ or if $X^2 \leq -3.84$, reject HO.

Conclusion sentence: The chimp performed significantly better than chance, $X^2(1, N = 24) = 20.16, p < .05$.

3. HO: Police officers do not have a preference in terms of the shift they work
H1: Police officers have a preference in terms of the shift they work

Decision rule: If $X^2 \geq +5.99$ or if $X^2 \leq -5.99$, reject HO.

Conclusion sentence: Police officers have a preference in terms of the shift they work, $X^2(1, N = 60) = 30.0, p < .05$.

4. HO: Students are not more likely to pass the course if they have taken college algebra
H1: Students are more likely to pass the course if they have taken college algebra

Decision rule: If $X^2 \geq +3.84$ or if $X^2 \leq -3.84$, reject H_0 .

Conclusion sentence: Students are more likely to pass the course if they have taken college algebra, $X^2(1, N = 70) = 15.39, p < .05$.

5. H_0 : The two species do not differ significantly in terms of intraspecific aggression.
 H_1 : The two species differ significantly in terms of intraspecific aggression.

Decision rule: If $X^2 \geq +3.84$ or if $X^2 \leq -3.84$, reject H_0 .

Conclusion sentence: The two species differ significantly in terms of intraspecific aggression, $X^2(1, N = 48) = 10.02, p < .05$.

6. H_0 : The antipsychotic is not significantly more effective in preventing relapse than the placebo.
 H_1 : The antipsychotic is significantly more effective in preventing relapse than the placebo.

Decision rule: If $X^2 \geq +3.84$ or if $X^2 \leq -3.84$, reject H_0 .

Conclusion sentence: The antipsychotic is significantly more effective in preventing relapse than the placebo, $X^2(1, N = 3195) = 364.32, p < .05$.