



Section 6.4: Testing for Independence in Two-Way Tables

Diez, D. M., Çetinkaya-Rundel, M., Barr, C. D. (2019). OpenIntro Statistics (4th ed.). OpenIntro.
<https://www.openintro.org/book/os/> CC BY-SA 3.0

STAT 1201
Introduction to Probability and Statistics

ONLINE AND DISTANCE EDUCATION

6.4 Testing for independence in two-way tables

We all buy used products – cars, computers, textbooks, and so on – and we sometimes assume the sellers of those products will be forthright about any underlying problems with what they’re selling. This is not something we should take for granted. Researchers recruited 219 participants in a study where they would sell a used iPod⁴⁰ that was known to have frozen twice in the past. The participants were incentivized to get as much money as they could for the iPod since they would receive a 5% cut of the sale on top of \$10 for participating. The researchers wanted to understand what types of questions would elicit the seller to disclose the freezing issue.

Unbeknownst to the participants who were the sellers in the study, the buyers were collaborating with the researchers to evaluate the influence of different questions on the likelihood of getting the sellers to disclose the past issues with the iPod. The scripted buyers started with “Okay, I guess I’m supposed to go first. So you’ve had the iPod for 2 years ...” and ended with one of three questions:

- General: What can you tell me about it?
- Positive Assumption: It doesn’t have any problems, does it?
- Negative Assumption: What problems does it have?

The question is the treatment given to the sellers, and the response is whether the question prompted them to disclose the freezing issue with the iPod. The results are shown in Figure 6.14, and the data suggest that asking the, *What problems does it have?*, was the most effective at getting the seller to disclose the past freezing issues. However, you should also be asking yourself: could we see these results due to chance alone, or is this in fact evidence that some questions are more effective for getting at the truth?

	General	Positive Assumption	Negative Assumption	Total
Disclose Problem	2	23	36	61
Hide Problem	71	50	37	158
Total	73	73	73	219

Figure 6.14: Summary of the iPod study, where a question was posed to the study participant who acted

DIFFERENCES OF ONE-WAY TABLES VS TWO-WAY TABLES

A one-way table describes counts for each outcome in a single variable. A two-way table describes counts for *combinations* of outcomes for two variables. When we consider a two-way table, we often would like to know, are these variables related in any way? That is, are they dependent (versus independent)?

The hypothesis test for the iPod experiment is really about assessing whether there is statistically significant evidence that the success each question had on getting the participant to disclose the problem with the iPod. In other words, the goal is to check whether the buyer’s question was independent of whether the seller disclosed a problem.

⁴⁰For readers not as old as the authors, an iPod is basically an iPhone without any cellular service, assuming it was one of the later generations. Earlier generations were more basic.

6.4.1 Expected counts in two-way tables

Like with one-way tables, we will need to compute estimated counts for each cell in a two-way table.

EXAMPLE 6.38

From the experiment, we can compute the proportion of all sellers who disclosed the freezing problem as $61/219 = 0.2785$. If there really is no difference among the questions and 27.85% of sellers were going to disclose the freezing problem no matter the question that was put to them, how many of the 73 people in the **General** group would we have expected to disclose the freezing problem?

We would predict that $0.2785 \times 73 = 20.33$ sellers would disclose the problem. Obviously we observed fewer than this, though it is not yet clear if that is due to chance variation or whether that is because the questions vary in how effective they are at getting to the truth.

GUIDED PRACTICE 6.39

If the questions were actually equally effective, meaning about 27.85% of respondents would disclose the freezing issue regardless of what question they were asked, about how many sellers would we expect to *hide* the freezing problem from the Positive Assumption group?⁴¹

We can compute the expected number of sellers who we would expect to disclose or hide the freezing issue for all groups, if the questions had no impact on what they disclosed, using the same strategy employed in Example 6.38 and Guided Practice 6.39. These expected counts were used to construct Figure 6.15, which is the same as Figure 6.14, except now the expected counts have been added in parentheses.

	General	Positive Assumption	Negative Assumption	Total
Disclose Problem	2 (20.33)	23 (20.33)	36 (20.33)	61
Hide Problem	71 (52.67)	50 (52.67)	37 (52.67)	158
Total	73	73	73	219

Figure 6.15: The observed counts and the (expected counts).

The examples and exercises above provided some help in computing expected counts. In general, expected counts for a two-way table may be computed using the row totals, column totals, and the table total. For instance, if there was no difference between the groups, then about 27.85% of each column should be in the first row:

$$0.2785 \times (\text{column 1 total}) = 20.33$$

$$0.2785 \times (\text{column 2 total}) = 20.33$$

$$0.2785 \times (\text{column 3 total}) = 20.33$$

Looking back to how 0.2785 was computed – as the fraction of sellers who disclosed the freezing issue ($158/219$) – these three expected counts could have been computed as

$$\left(\frac{\text{row 1 total}}{\text{table total}} \right) (\text{column 1 total}) = 20.33$$

$$\left(\frac{\text{row 1 total}}{\text{table total}} \right) (\text{column 2 total}) = 20.33$$

$$\left(\frac{\text{row 1 total}}{\text{table total}} \right) (\text{column 3 total}) = 20.33$$

This leads us to a general formula for computing expected counts in a two-way table when we would like to test whether there is strong evidence of an association between the column variable and row variable.

⁴¹We would expect $(1 - 0.2785) \times 73 = 52.67$. It is okay that this result, like the result from Example 6.38, is a fraction.

COMPUTING EXPECTED COUNTS IN A TWO-WAY TABLE

To identify the expected count for the i^{th} row and j^{th} column, compute

$$\text{Expected Count}_{\text{row } i, \text{ col } j} = \frac{(\text{row } i \text{ total}) \times (\text{column } j \text{ total})}{\text{table total}}$$

6.4.2 The chi-square test for two-way tables

The chi-square test statistic for a two-way table is found the same way it is found for a one-way table. For each table count, compute

General formula	$\frac{(\text{observed count} - \text{expected count})^2}{\text{expected count}}$
Row 1, Col 1	$\frac{(2 - 20.33)^2}{20.33} = 16.53$
Row 1, Col 2	$\frac{(23 - 20.33)^2}{20.33} = 0.35$
⋮	⋮
Row 2, Col 3	$\frac{(37 - 52.67)^2}{52.67} = 4.66$

Adding the computed value for each cell gives the chi-square test statistic X^2 :

$$X^2 = 16.53 + 0.35 + \cdots + 4.66 = 40.13$$

Just like before, this test statistic follows a chi-square distribution. However, the degrees of freedom are computed a little differently for a two-way table.⁴² For two way tables, the degrees of freedom is equal to

$$df = (\text{number of rows minus } 1) \times (\text{number of columns minus } 1)$$

In our example, the degrees of freedom parameter is

$$df = (2 - 1) \times (3 - 1) = 2$$

If the null hypothesis is true (i.e. the questions had no impact on the sellers in the experiment), then the test statistic $X^2 = 40.13$ closely follows a chi-square distribution with 2 degrees of freedom. Using this information, we can compute the p-value for the test, which is depicted in Figure 6.16.

COMPUTING DEGREES OF FREEDOM FOR A TWO-WAY TABLE

When applying the chi-square test to a two-way table, we use

$$df = (R - 1) \times (C - 1)$$

where R is the number of rows in the table and C is the number of columns.

When analyzing 2-by-2 contingency tables, one guideline is to use the two-proportion methods introduced in Section 6.2.

⁴²Recall: in the one-way table, the degrees of freedom was the number of cells minus 1.

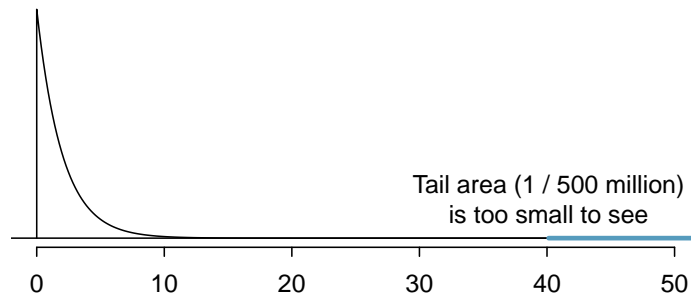


Figure 6.16: Visualization of the p-value for $X^2 = 40.13$ when $df = 2$.

EXAMPLE 6.40

Find the p-value and draw a conclusion about whether the question affects the sellers likelihood of reporting the freezing problem.

E

Using a computer, we can compute a very precise value for the tail area above $X^2 = 40.13$ for a chi-square distribution with 2 degrees of freedom: 0.000000002. (If using the table in Appendix C.3, we would identify the p-value is smaller than 0.001.) Using a significance level of $\alpha = 0.05$, the null hypothesis is rejected since the p-value is smaller. That is, the data provide convincing evidence that the question asked did affect a seller's likelihood to tell the truth about problems with the iPod.

EXAMPLE 6.41

Figure 6.17 summarizes the results of an experiment evaluating three treatments for Type 2 Diabetes in patients aged 10-17 who were being treated with metformin. The three treatments considered were continued treatment with metformin (**met**), treatment with metformin combined with rosiglitazone (**rosi**), or a lifestyle intervention program. Each patient had a primary outcome, which was either lacked glycemic control (failure) or did not lack that control (success). What are appropriate hypotheses for this test?

E

H_0 : There is no difference in the effectiveness of the three treatments.

H_A : There is some difference in effectiveness between the three treatments, e.g. perhaps the **rosi** treatment performed better than **lifestyle**.

	Failure	Success	Total
lifestyle	109	125	234
met	120	112	232
rosi	90	143	233
Total	319	380	699

Figure 6.17: Results for the Type 2 Diabetes study.

GUIDED PRACTICE 6.42

- ⓐ A chi-square test for a two-way table may be used to test the hypotheses in Example 6.41. As a first step, compute the expected values for each of the six table cells.⁴³

GUIDED PRACTICE 6.43

- ⓐ Compute the chi-square test statistic for the data in Figure 6.17.⁴⁴

GUIDED PRACTICE 6.44

- ⓐ Because there are 3 rows and 2 columns, the degrees of freedom for the test is $df = (3-1) \times (2-1) = 2$. Use $X^2 = 8.16$, $df = 2$, evaluate whether to reject the null hypothesis using a significance level of 0.05.⁴⁵

⁴³The expected count for row one / column one is found by multiplying the row one total (234) and column one total (319), then dividing by the table total (699): $\frac{234 \times 319}{699} = 106.8$. Similarly for the second column and the first row: $\frac{234 \times 380}{699} = 127.2$. Row 2: 105.9 and 126.1. Row 3: 106.3 and 126.7.

⁴⁴For each cell, compute $\frac{(\text{obs} - \text{exp})^2}{\text{exp}}$. For instance, the first row and first column: $\frac{(109 - 106.8)^2}{106.8} = 0.05$. Adding the results of each cell gives the chi-square test statistic: $X^2 = 0.05 + \dots + 2.11 = 8.16$.

⁴⁵If using a computer, we can identify the p-value as 0.017. That is, we reject the null hypothesis because the p-value is less than 0.05, and we conclude that at least one of the treatments is more or less effective than the others at treating Type 2 Diabetes for glycemic control.