

## 12.10: Summary

The key ideas discussed in this chapter are:

- The chi-square goodness of fit test (Section 12.1) is used when you have a table of observed frequencies of different categories; and the null hypothesis gives you a set of “known” probabilities to compare them to. You can either use the `goodnessOfFitTest()` function in the `lsr` package to run this test, or the `chisq.test()` function.
- The chi-square test of independence (Section 12.2) is used when you have a contingency table (cross-tabulation) of two categorical variables. The null hypothesis is that there is no relationship/association between the variables. You can either use the `associationTest()` function in the `lsr` package, or you can use `chisq.test()`.
- Effect size for a contingency table can be measured in several ways (Section 12.4). In particular we noted the Cramer’s V statistic, which can be calculated using `cramersV()`. This is also part of the output produced by `associationTest()`.
- Both versions of the Pearson test rely on two assumptions: that the expected frequencies are sufficiently large, and that the observations are independent (Section 12.5). The Fisher exact test (Section 12.7) can be used when the expected frequencies are small, `fisher.test(x = contingency.table)`. The McNemar test (Section 12.8) can be used for some kinds of violations of independence, `mcnemar.test(x = contingency.table)`.

If you’re interested in learning more about categorical data analysis, a good first choice would be Agresti (1996) which, as the title suggests, provides an *Introduction to Categorical Data Analysis*. If the introductory book isn’t enough for you (or can’t solve the problem you’re working on) you could consider Agresti (2002), *Categorical Data Analysis*. The latter is a more advanced text, so it’s probably not wise to jump straight from this book to that one.

### References

- Pearson, K. 1900. “On the Criterion That a Given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling.” *Philosophical Magazine* 50: 157–75.
- Fisher, R. A. 1922a. “On the Interpretation of  $\chi^2$  from Contingency Tables, and the Calculation of p.” *Journal of the Royal Statistical Society* 84: 87–94.
- Yates, F. 1934. “Contingency Tables Involving Small Numbers and the  $\chi^2$  Test.” *Supplement to the Journal of the Royal Statistical Society* 1: 217–35.
- Cramér, H. 1946. *Mathematical Methods of Statistics*. Princeton: Princeton University Press.
- McNemar, Q. 1947. “Note on the Sampling Error of the Difference Between Correlated Proportions or Percentages.” *Psychometrika* 12: 153–57.
- Agresti, A. 1996. *An Introduction to Categorical Data Analysis*. Hoboken, NJ: Wiley.
- Agresti, A. 2002. *Categorical Data Analysis*. 2nd ed. Hoboken, NJ: Wiley.

170. I should point out that this issue does complicate the story somewhat: I’m not going to cover it in this book, but there’s a sneaky trick that you can do to rewrite the equation for the goodness of fit statistic as a sum over  $k-1$  independent things. When we do so we get the “proper” sampling distribution, which is chi-square with  $k-1$  degrees of freedom. In fact, in order to get the maths to work out properly, you actually have to rewrite things that way. But it’s beyond the scope of an introductory book to show the maths in that much detail: all I wanted to do is give you a sense of why the goodness of fit statistic is associated with the chi-squared distribution.
171. I feel obliged to point out that this is an over-simplification. It works nicely for quite a few situations; but every now and then we’ll come across degrees of freedom values that aren’t whole numbers. Don’t let this worry you too much – when you come across this, just remind yourself that “degrees of freedom” is actually a bit of a messy concept, and that the nice simple story that I’m telling you here isn’t the whole story. For an introductory class, it’s usually best to stick to the simple story: but I figure it’s best to warn you to expect this simple story to fall apart. If I didn’t give you this warning, you might start getting confused when you see  $df=3.4$  or something; and (incorrectly) thinking that you had misunderstood something that I’ve taught you, rather than (correctly) realising that there’s something that I haven’t told you.

172. In practice, the sample size isn't always fixed... e.g., we might run the experiment over a fixed period of time, and the number of people participating depends on how many people show up. That doesn't matter for the current purposes.
173. Well, sort of. The conventions for how statistics should be reported tend to differ somewhat from discipline to discipline; I've tended to stick with how things are done in psychology, since that's what I do. But the general principle of providing enough information to the reader to allow them to check your results is pretty universal, I think.
174. To some people, this advice might sound odd, or at least in conflict with the "usual" advice on how to write a technical report. Very typically, students are told that the "results" section of a report is for describing the data and reporting statistical analysis; and the "discussion" section is for providing interpretation. That's true as far as it goes, but I think people often interpret it way too literally. The way I usually approach it is to provide a quick and simple interpretation of the data in the results section, so that my reader understands what the data are telling us. Then, in the discussion, I try to tell a bigger story; about how my results fit with the rest of the scientific literature. In short; don't let the "interpretation goes in the discussion" advice turn your results section into incomprehensible garbage. Being understood by your reader is *much* more important.
175. Complicating matters, the G-test is a special case of a whole class of tests that are known as *likelihood ratio tests*. I don't cover LRTs in this book, but they are quite handy things to know about.
176. A technical note. The way I've described the test pretends that the column totals are fixed (i.e., the researcher intended to survey 87 robots and 93 humans) and the row totals are random (i.e., it just turned out that 28 people chose the puppy). To use the terminology from my mathematical statistics textbook (Hogg, McKean, and Craig 2005) I should technically refer to this situation as a chi-square test of homogeneity; and reserve the term chi-square test of independence for the situation where both the row and column totals are random outcomes of the experiment. In the initial drafts of this book that's exactly what I did. However, it turns out that these two tests are identical; and so I've collapsed them together.
177. Technically,  $E_{ij}$  here is an estimate, so I should probably write it  $\hat{E}_{ij}$ . But since no-one else does, I won't either.
178. A problem many of us worry about in real life.
179. Though I do feel that it's worth mentioning the `assocstats()` function in the `vcd` package. If you install and load the `vcd` package, then a command like `assocstats( chapekFrequencies )` will run the  $\chi^2$  test as well as the likelihood ratio test (not discussed here); and then report three different measures of effect size:  $\phi^2$ , Cram'er's V, and the contingency coefficient (not discussed here)
180. Not really.
181. This example is based on a joke article published in the *Journal of Irreproducible Results*.
182. The R functions for this distribution are `dhyper()`, `phyper()`, `qhyper()` and `rhyper()`, though you don't need them for this book, and I haven't given you enough information to use these to perform the Fisher exact test the long way.
183. Not surprisingly, the Fisher exact test is motivated by Fisher's interpretation of a p-value, not Neyman's!

---

This page titled [12.10: Summary](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Danielle Navarro](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.