

CHAPTER OVERVIEW

22: Modeling Categorical Relationships

Learning Objectives

- Describe the concept of a contingency table for categorical data.
- Describe the concept of the chi-squared test for association and compute it for a given contingency table.
- Describe Simpson's paradox and why it is important for categorical data analysis.

So far we have discussed the general concept of statistical modeling and hypothesis testing, and applied them to some simple analyses. In this chapter we will focus on the modeling of *categorical* relationships, by which we mean relationships between variables that are measured qualitatively. These data are usually expressed in terms of counts; that is, for each value of the variable (or combination of values of multiple variables), how many observations take that value? For example, when we count how many people from each major are in our class, we are fitting a categorical model to the data.

[22.1: Example- Candy Colors](#)

[22.2: Pearson's chi-squared Test](#)

[22.3: Contingency Tables and the Two-way Test](#)

[22.4: Standardized Residuals](#)

[22.5: Odds Ratios](#)

[22.6: Bayes Factor](#)

[22.7: Categorical Analysis Beyond the 2 X 2 Table](#)

[22.8: Beware of Simpson's Paradox](#)

[22.9: Additional Readings](#)

This page titled [22: Modeling Categorical Relationships](#) is shared under a [CC BY-NC 4.0](#) license and was authored, remixed, and/or curated by [Russell A. Poldrack](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.