

11.10: Putting It Together- Chi-Square Tests

Let's Summarize

In this module, *Chi-Square Tests*, we discussed three different hypothesis tests using the chi-square test statistic:

- Goodness-of-fit for a one-way table
- Test of independence for a two-way table
- Test of homogeneity for a two-way table

Goodness-of-Fit test for a One-Way Table

- In a goodness-of-fit test, we consider one population and one categorical variable.
- The goodness-of-fit test expands the z-test for a population proportion that we learned in *Inference for One Proportion* by looking at the distribution of proportions for all categories defined by the categorical variable.
- The goodness-of-fit test determines whether a set of categorical data comes from a claimed distribution. The null hypothesis is that the proportion in each category in the population has a specific distribution. The alternative hypothesis says that the proportions in the population are not distributed as stated in the null hypothesis.
- To test our hypotheses, we select a random sample from the population and gather data for one categorical variable.

Test of Independence for a Two-Way Table

- In the test of independence, we consider one population and two categorical variables.
- In *Probability and Probability Distribution*, we learned that two events are independent if $P(A|B) = P(A)$, but we did not pay attention to variability in the sample. With the chi-square test of independence, we have a method for deciding whether our observed $P(A|B)$ is “too far” from our observed $P(A)$ to infer independence in the population.
- The null hypothesis says the two variables are independent (or not associated). The alternative hypothesis says the two variables are dependent (or associated).
- To test our hypotheses, we select a single random sample and gather data for two different categorical variables.

Test of Homogeneity for a Two-Way Table

- In the test of homogeneity we consider two or more populations (or two or more subgroups of a population) and a single categorical variable.
- The test of homogeneity expands on the test for a difference in two population proportions that we learned in *Inference for Two Proportions* by comparing the distribution of the categorical variable across multiple groups or populations.
- The null hypothesis says that the distribution of proportions for all categories is the same in each group or population. The alternative hypothesis says that the distributions differ.
- To test our hypotheses, we select a random sample from each population or subgroup independently. We gather data for one categorical variable.

The Chi-Square Test Statistic and Distribution

For all chi-square tests, the chi-square test statistic χ^2 is the same. It measures how far the observed data are from the null hypothesis by comparing observed counts and expected counts. *Expected counts* are the counts we expect to see if the null hypothesis is true.

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

The chi-square model is a family of curves that depend on degrees of freedom. For a one-way table the degrees of freedom equals $(r - 1)$. For a two-way table, the degrees of freedom equals $(r - 1)(c - 1)$. All chi-square curves are skewed to the right with a mean equal to the degrees of freedom.

A chi-square model is a good fit for the distribution of the chi-square test statistic only if the following conditions are met:

- The sample is randomly selected.
- All expected counts are 5 or greater.

If these conditions are met, we use the chi-square distribution to find the P-value. We use the same logic that we have used in all hypothesis tests to draw a conclusion based on the P-value. If the P-value is at least as small as the significance level, we reject the

null hypothesis and accept the alternative hypothesis. The P-value is the likelihood that results from random samples have a χ^2 value equal to or greater than that calculated from the data if the null hypothesis is true.

CC licensed content, Shared previously

- Concepts in Statistics. **Provided by:** Open Learning Initiative. **Located at:** <http://oli.cmu.edu>. **License:** [CC BY: Attribution](#)

This page titled [11.10: Putting It Together- Chi-Square Tests](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [Bill Pelz](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.