

17.5: 2 x 2 Simulation

Learning Objectives

- Explore the accuracy of the Chi Square test
- Determine the extent to which the Yates's correction affects the accuracy of the test

Instructions

The Chi Square Test is an approximate not an exact test. This simulation allows you to examine the accuracy and power of the Chi Square Test in a variety of situations. The significance of the difference between the proportion who succeed in Condition 1 and the proportion who succeed in Condition 2 is tested.

With the default parameters, the probability of success is the same (0.60) in both conditions so the null hypothesis is true. The default sample size is 10 per condition and the default significance level is 0.05.

Some authors have suggested that a correction called the "Yates Correction" be done whenever the expected frequency of any cell is below five. A test without this correction and one with the correction (if a cell has an expected frequency below five) is conducted for each simulated experiment.

If you click the "Simulate 1" button, one simulated experiment will be conducted. The observed and expected frequencies are presented as well as the Chi Square Test with and without the Yates correction. A tally of the number of times the tests were significant is shown below.

If you click the "Simulate 1000" (or 5000) button then 1000 (or 5000) simulated experiments are conducted and the numbers of significant and non-significant tests are shown.

1. Using the default values, click on the "Simulate 1" button. Note the number of successes and failures in each of the two conditions. Note the value of the Chi Square with and without the Yates's correction. Was either significant? Check the lower section that shows the count of the number significant. Most likely, it will show one non-significant.
2. Test whether the **Type I** error rate is close to the nominal significance level of 0.05. Do this by clicking the "Simulate 5000" button several times. Compare the proportion significant to 0.05. Look at the results both when the Yates correction is never used and when it is used when an expected cell frequency is less than 5. Is the test conservative (proportion significant < 0.05) or is it liberal (proportion significant > 0.05)
3. Redo the previous simulations when the probability of success is 0.50 for each condition. Are the results similar? Try making one of the sample sizes 10 and the other one 6.
4. Try to find a set of parameters such that the proportion significant is greater than 0.06. (Make sure the null hypothesis is always true -- that the probability of success is the same for both conditions.) Could you find such a set of parameters? Are there many circumstances in which the test is that liberal? Do you think the Yates correction is a good procedure?
5. Now consider cases in which the probability of success is different for the two conditions. Here the null hypothesis is false so the higher the proportion significant the better. What is the effect of using the Yates correction on rejecting a false null hypothesis?

Illustrated Instructions

The demo begins by running a single simulation, which turns out to be not statistically significant. The video concludes by running 1,000 simulations and then 5,000 simulations.

Video Demo

Try changing the proportions of the conditions?

This page titled [17.5: 2 x 2 Simulation](#) is shared under a [Public Domain](#) license and was authored, remixed, and/or curated by [David Lane](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.