

McNemar's Change Test -- Analysis of 2-Within-Group Data with a Qualitative Response Variable

Application: This statistic has two applications that can appear very different, but are really just two variations of the same statistical question. In one application the same qualitative (binary) variable is measured at two different times from the same sample (or from two samples that have been matched on one or more important variables). In the other application, two comparable qualitative variables are measured from the same sample (usually at the same time). In both applications, McNemar's Change Test is used to compare the distributions of the two qualitative variables.

There are two specific versions of the H_0 , depending upon whether one characterizes the different conditions as representing a single population under two different circumstances (as in the example below) or as representing comparable variables measured from a single population (e.g., comparing the diagnoses of a set of patients made by two different specialists). Here are versions of the H_0 : statement for each of these characterizations.

H_0 : The two populations represented by the conditions have the same distributions of scores across the two categories of the response variable. Another way to state this null that emphasizes that the interest is in *changes* across conditions is -- those cases that do change category from one condition to another show no pattern of change (i.e., the probability of change from one category to the other is the same for both categories).

To reject H_0 : is to say that the populations have different distributions of response variable categories across the conditions (or that there is a pattern to the changes in response variable categories).

H_0 : Within this population, cases that are assigned to the comparable categories of the two variables. To express this to emphasize an interest in change across the variables we could say, the variables show no pattern of differential categorization (i.e., the probability of shifting from one category to the other across the two variables is the same for each category).

To reject H_0 : is to say that within this population cases have different distributions of categories across the two variables.

The data: The researcher who had collected the Pet Shop data learned that a new and inexpensive filtration system for saltwater aquariums had been marketed and was expected to dramatically increase the number of saltwater aquariums sold in this country. Given this, the researcher wanted to know whether more stores would begin to carry both freshwater and saltwater fish. The researcher went back to each store (approximately 1 year after the initial data were collected) and recorded whether they carried only freshwater or both saltwater and freshwater fish at that time. In this analysis the one variable is before vs. after the introduction of the new filtration system, and the response variable is the type of fish displayed. From our database (with the new data added, called fishdep2), we use two variables **fishdept** (1 = display freshwater fish only, 2 = display both saltwater and freshwater fish) and **fishdep2** (same codes). These scores are shown for the 12 stores below (fishdept, fishdep2).

1,1 2,2 2,1 2,2 1,2 2,2 2,1 1,2 1,2 1,2 2,2 1,1

Research Hypothesis: The researcher hypothesized that pet shops that had displayed both saltwater and freshwater fish during the period of the first data collection would still do so, but that pet shops that displayed only freshwater fish at that time would be more likely to currently display both saltwater and freshwater fish.

H_0 : for this analysis: Pet stores have the same distribution of type of fish displayed before and after the introduction of the inexpensive saltwater filtration system.

Step 1 Organize the data into a contingency table. For McNemar's test, this will always be a 2x2 table.

fishdept	fishdep2	
	freshwater	salt- & freshwater
freshwater		
salt- & freshwater		

Each store's data will be collated into one of the four cells. For example, a store that displayed only freshwater fish during the first data collection but that displayed both saltwater and freshwater fish during the second data collection would be tallied in the upper right cell. Below is the contingency table filled with the responses from the 12 stores.

fishdept	fishdep2	
	freshwater	salt- & freshwater
freshwater	2	4
salt- & freshwater	2	4

Step 3 Compute Chi-square

Remember that we are primarily interested in those stores that changed the type of fish they displayed (i.e., that are in different categories of fishdept and fishdep2). The two of the cells in the contingency table that represent changes from the first data collection to the second are the upper right cell (displaying only freshwater fish during the first data collection but displaying both saltwater and freshwater fish the during the second) and the lower left cell (displaying both saltwater and freshwater fish the during the first data collection but displaying only freshwater fish during the second). **Only the values in these cells that represent changes from the first data collection to the second will be used to compute the Chi-square.**

We will label these two cells C_1 and C_2 (it doesn't matter which of the cells is given which label), and apply the following formula.

$$\begin{aligned}
 X^2 &= \frac{(|C_1 - C_2| - 1)^2}{C_1 + C_2} = \frac{(|2 - 4| - 1)^2}{2 + 4} = \frac{(|-2| - 1)^2}{6} = \\
 &= \frac{(2 - 1)^2}{6} = \frac{1}{6} = .167
 \end{aligned}$$

Step 4 Use the Chi-square table to determine the critical X^2 value for $df = 1$ (df is always 1 for the McNemar test) and $p = .05$

$$X^2(df=1, p=.05) = 3.84$$

Step 5 Compare the obtained X^2 and critical X^2 , and determine whether or not there is a statistically significant relationship between the two categorical variables.

- if the obtained X^2 is less than the critical X^2 , then retain the null hypothesis -- conclude that there is no relationship between subject's values on one categorical variable and their values on the other categorical variable, in the population represented by the sample
- if the obtained X^2 is greater than the critical X^2 , then reject the null hypothesis -- conclude that there is a relationship between the subject's values on one categorical variable and their values on the other categorical variable, in the population represented in the sample.

For the example data, we would decide to retain the null hypothesis, because the obtained Chi-square value of .167 is smaller than the critical Chi-square value of 3.84.

By the way: This test should only be applied when at least 10 of the cases change category across the two variables. When there are fewer changes, as in this example, then the test is likely to yield inaccurate results.

Step 6 IF you reject the null hypothesis, determine whether the pattern of the data in the contingency table supports, partially supports, or does not support the research hypothesis.

- IF you reject the null hypothesis, AND if the pattern of data in the contingency table agrees exactly with the research hypothesis, then the research hypothesis is completely supported.
- IF you reject the null hypothesis, AND if part of the pattern of data in the contingency table agrees with the research hypothesis, BUT part of the pattern of data does not, then the research hypothesis is partially supported.
- IF you retain the null hypothesis, OR you reject the null BUT NO PART of the pattern of data in the contingency table agrees with the research hypothesis, then the research hypothesis is not at all supported.

By the way: Usually the researcher hypothesizes that there is a pattern of relationship between the variables. Sometimes, however, the research hypothesis is that there is NO pattern of relationship. If so, the research hypothesis and H_0 are the same! When this is the case, retaining H_0 provides support for the research hypothesis, whereas rejecting H_0 provides evidence that research hypothesis is incorrect.

Since we retained the null hypothesis, we know that we did not find the pattern of results we hypothesized. As expected, there were some stores that displayed only freshwater fish during the first data collection that then displayed both saltwater and freshwater fish during the second data collection. However, there were also stores that showed the opposite pattern of change.

Step 7 Reporting the results

It is important to describe the univariate data before telling whether or not there is a pattern of relationship between the variables. Report the number (or percentage) that fall into the categories of each variable (showing the contingency table will help the reader). As for the other statistical tests, the report includes the "wordy" part and the statistical values upon which you made your statistical decision. If H_0 is rejected, be sure to describe the pattern of the relationship.

Four of the stores changed from displaying only freshwater fish during the first data collection to displaying both saltwater and freshwater fish during the second, whereas two of the stores changed from displaying both saltwater and freshwater fish to displaying only freshwater fish. The pattern of change was not significant ($X^2(1) = .167, p > .05$).

C² Critical values of Chi-Square

df	$\alpha = .05$	$\alpha = .01$
1	3.84	6.63
2	5.99	9.21
3	7.81	11.34
4	9.49	13.28
5	11.07	15.09
6	12.59	16.81
7	14.07	18.48
8	15.51	20.09
9	16.92	21.67
10	18.31	23.21
11	19.68	24.72
12	21.03	26.22
13	22.36	27.69
14	23.68	29.14
15	25.00	30.58
16	26.30	32.00
17	27.59	33.41
18	28.87	34.81
19	30.14	36.19
20	31.41	37.57
21	32.67	38.93
22	33.92	40.29
23	35.17	41.64
24	36.42	42.98
25	37.65	44.31
26	38.89	45.64
27	40.11	46.96
28	41.34	48.28
29	42.56	49.59
¥	43.77	50.89