

16.7: RM Chi-Square- The McNemar Test

Suppose you've been hired to work for the *Generic Political Party* (GPP), and part of your job is to find out how effective the GPP political advertisements are. So, what you do, is you put together a sample of $N=100$ people, and ask them to watch the GPP ads. Before they see anything, you ask them if they intend to vote for the GPP; and then after showing the ads, you ask them again, to see if anyone has changed their minds. Obviously, if you're any good at your job, you'd also do a whole lot of other things too, but let's consider just this one simple experiment. One way to describe your data is via the following contingency table:

Table 16.7.1- Voting and Advertisement Counts

Voting and Ads	Before	After	Total
Yes Vote	30	10	40
No Vote	70	90	160
Total	100	100	200

At first pass, you might think that this situation lends itself to the Pearson χ^2 Test of Independence. However, a little bit of thought reveals that we've got a problem: we have 100 participants, but 200 observations. This is because each person has provided us with an answer in *both* the before column and the after column. What this means is that the 200 observations aren't independent of each other: if voter A says "yes" the first time and voter B says "no" the first time, then you'd expect that voter A is more likely to say "yes" the second time than voter B. The consequence of this is that the usual χ^2 test won't give trustworthy answers due to the violation of the independence assumption (found in the section on [Assumptions of Chi-Square tests](#)). Now, if this were a really uncommon situation, I wouldn't be bothering to waste your time talking about it. But it's not uncommon at all: this is a *standard* repeated measures design, and none of the tests we've considered so far can handle it. (You might immediately think about the Phi correlation, Dr. MO certainly did! But according to [MathCracker.com](#), Phi is a χ^2 but with an extra step, so it would have the same assumptions as all Chi-Square analyses- no dependent data).

Eek.

The solution to the problem was published by McNemar (1947). The trick is to start by tabulating your data in a slightly different way:

Table 16.7.2- Rearranged Voting and Advertisement Counts

	Before: Yes	Before: No	Total
After: Yes	5	5	10
After: No	25	65	90
Total	30	70	100

This is exactly the same data, but it's been rewritten so that each of our 100 participants appears in only one cell. Because we've written our data this way, the independence assumption is now satisfied, and this is a contingency table that we *can* use to construct an χ^2 Goodness of Fit statistic. However, as we'll see, we need to do it in a slightly nonstandard way. To see what's going on, it helps to label the entries in our table a little differently:

Table 16.7.3- Cells Labeled

	Before: Yes	Before: No	Total
After: Yes	a	b	a+b
After: No	c	d	c+d
Total	a+c	b+d	n

Next, let's think about what our null hypothesis is: it's that the "before" test and the "after" test have the same proportion of people saying "Yes, I will vote for GPP". Because of the way that we have rewritten the data, it means that we're now testing the hypothesis that the *row totals* and *column totals* come from the same distribution. Thus, the null hypothesis in McNemar's test is

that we have “marginal homogeneity,” meaning that the row totals and column totals have the same distribution: $P_a + P_b = P_a + P_c$, and similarly that $P_c + P_d = P_b + P_d$. Notice that this means that the null hypothesis actually simplifies to $P_b = P_c$. In other words, as far as the McNemar test is concerned, it’s only the off-diagonal entries in this table (i.e., b and c) that matter! After noticing this, the McNemar test of marginal homogeneity is not that different to a usual χ^2 test.

Since the calculation is so similar to χ^2 we won't be going over it. If we ran a McNemar's test to determine if people were just as likely to vote GPP after the ads as they were before hand, we would find statistically significant difference ($\chi^2(1)=12.04, p<.001$), suggesting that the groups were not just as likely to vote GPP after the as as before. But look closely before you recommend dumping money into the advertising budget! It looks like the ads had a *negative* effect: people were *less* likely to vote GPP after seeing the ads. (Which makes a lot of sense when you consider the quality of a typical political advertisement.)

As always, if you are doing statistics for graduate school or your job, you'll have software that will do all of this for you. For now, you are learning the formulas for two reasons:

1. The formulas show you what is happening (mathematically) so that you understand the results better.
2. Being able to work through a formula helps with your logic, reasoning, and critical thinking skills.

Speaking of critical thinking, let's get to the final section of this chapter: Choosing the Correct Statistical Analysis!

Reference

McNemar, Q. (1947) Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12, 153-157.

This page titled [16.7: RM Chi-Square- The McNemar Test](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Michelle Oja](#).

- [12.8: The McNemar Test](#) by [Danielle Navarro](#) is licensed [CC BY-SA 4.0](#). Original source: <https://bookdown.org/ekothe/navarro26/>.
- [Current page](#) by [Michelle Oja](#) is licensed [CC BY-SA 4.0](#).