McNemar’s Test in Stata: A Step-by-Step Guide for Analyzing Paired Data

Authored by

Mohammed looti

November 8, 2025

McNemar's Test is a highly specialized, non-parametric statistical procedure essential for researchers working with dependent observations. Its primary purpose is to determine if there is a **statistically significant difference** between the proportions of two related **dichotomous (binary) variables**. Unlike tests designed for independent groups, McNemar's Test is specifically tailored to analyze **paired data**, making it invaluable in experimental designs such as 'before and after' studies, where the same subjects are measured repeatedly, or in tightly matched case-control studies.

For researchers leveraging powerful statistical packages, mastering the correct execution and interpretation of this test is fundamental. This comprehensive tutorial offers a detailed, step-by-step methodology for conducting and understanding McNemar's Test using the industry-leading statistical software, **Stata**. We will walk through data preparation, command execution, and the critical steps required for drawing robust conclusions from the output.

The Analytical Foundation of McNemar's Test

Selecting the appropriate statistical test hinges upon understanding the underlying structure of the data and the research design. McNemar's Test stands out because it addresses situations where observations are inherently dependent. This dependency means that the outcome of the first measurement is linked to the second, as both readings originate from the same individual or a carefully constructed matched pair. Standard Chi-Square tests, which assume independence, would be inappropriate and misleading in this context.

The core logic of McNemar's Test rests entirely on examining the discordance within the pairs. It ignores the subjects whose responses remained the same across both observations (the concordant pairs) and focuses exclusively on the subjects whose responses changed. By comparing the number of subjects who shifted in one direction (e.g., from 'Success' to 'Failure') against those who shifted in the opposite direction (e.g., from 'Failure' to 'Success'), the test determines if the observed change in proportion is random or systematic.

The underlying assumption, articulated in the **null hypothesis**, is that if no true effect exists, the proportion of individuals changing from Condition A to Condition B will be statistically identical to the proportion changing from Condition B to Condition A. Consequently, rejecting the null hypothesis provides strong evidence that the intervention, treatment, or condition introduced between the two measurements caused a measurable, non-random shift in the population's characteristics or opinions. This unique focus on change makes the test highly sensitive to treatment effects in longitudinal or paired designs.

Designing the Study: A Market Research Application

To illustrate the practical utility of McNemar's Test, consider a common scenario in market research or political science: evaluating the persuasive impact of a new piece of media content, specifically a digital marketing video promoting a specific piece of legislation. The primary goal is to assess whether exposure to the video effectively shifts public opinion.

The researchers adopt a classic 'before and after' experimental design, recruiting a sample of 100 participants. During the initial phase (the 'Before' measurement), every participant is surveyed to establish their current stance regarding the legislation: they either support the law or they do not support it. This establishes the baseline proportion of supporters.

Next, the intervention is administered: all 100 participants are immediately exposed to the marketing video. Following the completion of the viewing, the participants are surveyed a second time (the 'After' measurement) using the exact same dichotomous question about their support for the law. This methodology, measuring the same individuals at two distinct time points related by an intervention, successfully generates the dependent, paired data structure that McNemar's Test requires for valid analysis. The central research objective then becomes determining if the observed changes in opinion are substantial enough to be considered statistically significant, thereby confirming the video's influence.

Structuring the Paired Data for Statistical Analysis

The results derived from any paired survey, such as our market research example, must first be condensed and organized into a standard 2x2 **contingency table**. This table systematically categorizes the outcomes for all participants based on their responses in both the initial and final measurements. This structure is essential because it isolates the four possible combinations of results, making the discordant pairs immediately visible for the test calculation.

The four cells represent: those who held the same opinion both times (concordant pairs) and the two critical groups who shifted their opinions (discordant pairs). The following structure uses the aggregated counts from our hypothetical 100 participants, categorized by their opinions before and after viewing the marketing material:

	Before Marketing Video	
After Marketing Video	**Support**	**Do not support**
Support	30	40
Do not Support	12	18

Analyzing the cell counts reveals key insights: 30 individuals (Cell A) supported the law both times, and 18 individuals (Cell D) did not support the law both times--these are the concordant pairs, which do not influence the McNemar test statistic. The crucial data lies in the discordant cells: 40 people (Cell B) shifted from 'Do Not Support' (Before) to 'Support' (After), demonstrating a positive influence; while 12 people (Cell C) shifted from 'Support' (Before) to 'Do Not Support' (After), representing a negative or contrary influence. McNemar's Test precisely compares the counts of these two discordant cells (40 vs. 12) to assess the significance of the change.

Executing McNemar's Test Using Stata's mcci Command

To formally test whether the difference observed between the proportion of supporters before and after the marketing video is statistically significant, we turn to the specialized commands within Stata. Stata provides the highly efficient `mcci` command (McNemar's Chi-squared test for count data) specifically designed to handle data already summarized in the 2x2 matrix format. This command bypasses the need to load raw data, requiring only the four cell counts as input.

When utilizing the `mcci` command, it is imperative to enter the cell counts (A, B, C, D) in the correct sequence, starting from the top-left cell and proceeding across each row from left to right. Based on the data derived from our case study (A=30, B=40, C=12, D=18), the required Stata syntax is notably straightforward and precise, ensuring the program correctly identifies the discordant and concordant pairs:

mcci 30 40 12 18

Upon execution, this single command instructs Stata to perform the McNemar calculation, generating a complete statistical output that includes the reconstructed table, the calculated test statistic, and the associated probability value, which is necessary for hypothesis testing.

```
. mcci 30 40 12 18

                 | Controls               |
Cases            |   Exposed   Unexposed  |      Total
-----------------+------------------------+-----------
         Exposed |        30          40  |         70
       Unexposed |        12          18  |         30
-----------------+------------------------+-----------
           Total |        42          58  |        100

McNemar's chi2(1) =     15.08    Prob > chi2 = 0.0001
Exact McNemar significance probability       = 0.0001

Proportion with factor
        Cases               .7
        Controls           .42     [95% Conf. Interval]
                     ---------     --------------------
        difference         .28      .1397545   .4202455
        ratio         1.666667      1.284237   2.162979
        rel. diff.    .4827586      .3075044   .6580128

        odds ratio    3.333333       1.71442    6.97982   (exact)
```

Interpreting the Statistical Output and Drawing Conclusions

The output generated by the `mcci` command contains three core components that are necessary for deriving meaningful conclusions regarding the effectiveness of the intervention. These components are the verification table, the **Chi-Square test statistic**, and the critical **p-value**. Researchers must carefully analyze each section to confirm the integrity of the data and the significance of the results.

The initial section of the Stata output displays the reconstructed **2x2 table**. While the labels may vary (often using 'exposed/unexposed' or 'case/control' terminology common in epidemiology), this table serves as immediate confirmation that the input counts were entered correctly (30, 40, 12, 18). Crucially, the table confirms that the calculation utilized a total sample size of 100 participants, ensuring alignment with the study design.

The next key piece of data is the **McNemar's chi2(1)** value, which represents the calculated **Chi-Square test statistic**. This statistic quantifies the magnitude of the difference between the two discordant cell counts (B and C). For our market research example, the calculated value is 15.08. Stata derives this figure by using the standard formula specific to McNemar's Test, which focuses solely on the discordant pairs: (B-C)2 / (B+C). In our specific case, the calculation yields (40-12)2 / (40+12) = 784 / 52, resulting in approximately 15.0769.

The most pivotal element for hypothesis testing is the **Prob > chi2**, which is the associated **p-value**. This value indicates the probability of observing a difference as large as the one calculated (15.08) if the null hypothesis were true--that is, if the marketing video had absolutely no effect on opinion. In this analysis, the p-value is reported as 0.0001. Since this calculated p-value (0.0001) is dramatically lower than the conventionally accepted statistical significance level of 0.05 (or even 0.01), we possess compelling statistical evidence to reject the null hypothesis. The conclusion, therefore, is that the marketing video successfully induced a statistically significant shift in opinion, specifically increasing the proportion of participants who supported the legislation.

Understanding the Chi-Square Approximation and Correction

While the calculation (B-C)2 / (B+C) provides the Chi-Square test statistic used by Stata, it is important to remember the theoretical structure of the data it operates on. McNemar's Test is fundamentally comparing the counts in the discordant cells, B and C, which represent the two opposing directions of change. The general structure is always defined as follows:

	Before Marketing Video	
After Marketing Video	**Support**	**Do not support**
Support	A	B
Do not Support	C	D

The validity of the Chi-Square approximation used by McNemar's Test relies on the assumption that the expected counts are sufficiently large. When the counts within the discordant cells (B and C) are particularly small, the continuous Chi-Square distribution may poorly approximate the discrete nature of the count data, leading to potentially inflated or unreliable p-values.

Consequently, when dealing with limited data, specifically when the sum of the discordant cells (B + C) is less than 20 or when either B or C is less than 5, some statistical guidelines recommend applying a **continuity correction**. This correction adjusts the numerator of the formula to better account for the discrete nature of the counts: (|B-C| - 1)2 / (B+C). This adjustment generally results in a slightly smaller Chi-Square value and a slightly higher, more conservative p-value. Although the continuity correction is a valuable theoretical consideration for small sample sizes, it is typically deemed unnecessary for large samples like the one utilized in our detailed case study.