



The McNemar test is used to analyze paired observations of a dichotomous variable, i.e. where there can only be one of two possible values such as: success/failure, +/-, 0/1, etc. and it is of interest to examine if the paired values are associated or are independent.

To be precise, consider the possible outcome from testing fifty specimens of sputum cultured on two different culture media, A and B, with the intention of detecting a particular bacterium. The four possible outcomes were as follows.

Type	Medium A	Medium B	Number
Both	+	+	20
A only	+	-	12
B only	-	+	2
Neither	-	-	16

These data can be arranged as a 2 by 2 contingency table, such as this table contained in SIMFIT test file `mcnemar.tfl`.

	B +	B -	Total
A +	20	12	32
A -	2	16	18
Total	22	28	50

Here, the cell frequencies are ($f_{11} = 20, f_{12} = 12, f_{21} = 2, f_{22} = 16$), the sum of row frequencies known as row marginals are (32, 18), the sum of column frequencies known as column marginals are (22, 28), and obviously the row and column marginals must separately both add up to the total number of frequencies ($n = 50$).

From the main SIMFIT menu choose [Statistics] then [Standard tests] and analyze the above data using the McNemar option to get the following table.

McNemar test 1		
H_0 : Expected value of $[(f(1,2) - (f(2,1)))/n] = 0$		
Title: Data for 2 by 2 McNemar test		
Number of rows/columns	2	
Chi-square test statistic C	5.786	
Number of degrees of freedom	1	
$P(\chi^2 \geq C)$	0.0162	Reject H_0 at 5% level
Upper tail 5% point	3.841	
Upper tail 1% point	6.635	
Continuity correction used in chi-square		

The frequencies f_{ij} in this table are analyzed by calculating the χ^2 test statistic given by

$$\chi^2 = \frac{(|f_{12} - f_{21}| - 1)^2}{f_{12} + f_{21}}$$

which has an approximate chi-square distribution with 1 degree of freedom. The outcome emphasizes the obvious fact that culture medium A is more effective than culture medium B.

Note that this test does not perform so well with small frequencies and, in particular, if $r = 2$ and

$$f_{12} + f_{21} \leq 20$$

a warning will be displayed. In such cases it may be preferable to do a binomial test using $N = f_{12} + f_{21}$ then $X = f_{12}$ or $X = f_{21}$ to check if $\hat{p} = X/N$ is consistent with a binomial distribution with parameters N and $p = 0.5$. Since in the 2 by 2 case the McNemar test is equivalent to testing if two sample estimates for a binomial probability parameter differ significantly, we can use SIMFIT to calculate exact 95% confidence limits as follows

$$\text{For } 2/14 : 0.0178 \leq \hat{p} = 0.1429 \leq 0.4281$$

$$\text{For } 12/14 : 0.5719 \leq \hat{p} = 0.8571 \leq 0.9822$$

which convincingly demonstrates the superiority of culture medium A over culture medium B.

To explain the logic behind this analysis, note that the overall proportion of successes with medium A is $(f_{11} + f_{12})/n$, while the overall proportion of successes with medium B is $(f_{11} + f_{21})/n$, so that the difference between these estimates for the probability of success depends only on $f_{12} - f_{21}$, and the null hypothesis for such a 2 by 2 table can be expressed as expectations in several equivalent ways without using the diagonal frequencies f_{ii} except in the sample size n , such as

$$H_0 : E\left(\frac{f_{12} - f_{21}}{n}\right) = 0, \text{ or}$$

$$H_0 : E\left(\frac{f_{12}}{f_{21}}\right) = 1.$$

Note that it is important that tables for the McNemar test are set up correctly to reflect the pairwise nature of the data, so an additional example is given using data in test file `mcnemar.tf2` for the case where medication A was applied to one arm and medication B to the other arm with subjects suffering from a rash on both arms.

	B worked	B failed	Total
A worked	11	6	17
A failed	10	24	34
Total	21	30	50

Chi-square test statistic C	0.5625	
Number of degrees of freedom	1	
$P(\chi^2 \geq C)$	0.4533	Consider accepting H_0
Upper tail 5% point	3.841	
Upper tail 1% point	6.635	

The outcome is that there is no evidence to support a significant difference between medications A and B in this experiment.

More generally, for larger r by r tables with identical paired row and column categorical variables, the continuity correction is not used, and the appropriate test statistic is

$$\chi^2 = \sum_{i=1}^r \sum_{j>i} \frac{(f_{ij} - f_{ji})^2}{f_{ij} + f_{ji}}$$

with $r(r - 1)/2$ degrees of freedom. Unlike the normal contingency table analysis where the null hypothesis is independence of rows and columns, with this test there is intentional association between rows and columns. The test statistic does not use the diagonal frequencies f_{ii} and is testing whether the upper right corner of the table is symmetrical with the lower left corner. The SIMFIT test file `mcnemar.tf3` contains data for a such a 3 by 3 McNemar table.