McNemar's test

- The *McNemar test* is a test on a 2x2 classification table when we want to test the difference between paired proportions, e.g. in studies in which patients serve as their own control, or in studies with "before and after" design.
- The test is used to determine whether the row and column marginal frequencies are equal (that is, whether there is "marginal homogeneity")

	Test 2 positive	Test 2 negative	Row total
Test 1 positive	*a*	*b*	*a* + *b*
Test 1 negative	*c*	*d*	*c* + *d*
Column total	*a* + *c*	*b* + *d*	*n*

- For example, in the table above, the null hypothesis is $\pi_a + \pi_b = \pi_a + \pi_c$
- We can cancel π_a on both sides, and the hypotheses to be tested are:

$$H_0: \pi_b = \pi_c \quad H_1: \pi_b \neq \pi_c$$

Test statistic: X^2

- The McNemar test statistic is:

$$X^2 = \frac{(b-c)^2}{b+c}$$

The test statistics follows a chi-square distribution with 1 Degree Of Freedom.

- **Note**: if b + c < 25, the test statistic is not well approximated by the χ^2 distribution, we need to use an exact binomial test to test the hypotheses.
- It is also important to know that $X^2 = z^2$, that is, the chi-square distribution with one degree of freedom is the square of the z–statistic which follows a standard normal distribution, therefore we can take the square root of X^2 and test whether the difference is significant or not.

$$z = \frac{b-c}{\sqrt{(b+c)}}$$

- If the result is significant, we reject the null hypothesis and conclude that the marginal proportions are significantly different for the two groups.

Example

A study involving a cohort of women in Birmingham, AL examined revision surgery involving silicone gel breast implants (Brown and Pennello, 2002). Of 165 women with surgical records who had reported having surgery, the following information was obtained.

- In 69 cases, both self report and surgical records said there was a rupture or leak.
- In 63 cases, both self report and surgical records said there was no rupture or leak.
- In 28 cases, the self report said there was a rupture or leak, but the surgical records did not report one.
- In 5 cases, the self report said there was no rupture or leak, but the surgical records reported one.

The data are summarized in the Table. Present refers to a rupture or leak, Absent refers to no rupture or leak.

		Surgical Record		
		Present	Absent	
Self Report	Present	69	28	97
	Absent	5	63	68
		74	91	165

Example(cont.)

- To test if there is a difference between the probability of a self-reported rupture or leak and the probability of a rupture or leak documented in the surgical record, we should use McNemar's test.

$$H_0: the\ probabilities\ are\ equal \qquad \mathrm{H}_1: the\ probabilities\ are\ different$$

$$z = \frac{28 - 5}{\sqrt{28 + 5}} = \frac{23}{5.74} = 4 > z_{0.025} = 1.96$$

- We can reject the null hypothesis and conclude that the probabilities are statistically significantly different.
- **Stata code: mcci 69 28 5 63**

Stata code: mcci 69 28 5 63

```
. mcci 69 28 5 63

                    | Controls
Cases               |  Exposed   Unexposed |      Total
--------------------+----------------------+-----------
          Exposed   |       69          28 |         97
        Unexposed   |        5          63 |         68
--------------------+----------------------+-----------
            Total   |       74          91 |        165

McNemar's chi2(1) =     16.03    Prob > chi2 = 0.0001
Exact McNemar significance probability      = 0.0001

Proportion with factor
        Cases         .5878788
        Controls      .4484848     [95% Conf. Interval]
                    ------------   --------------------
        difference    .1393939      .0684956   .2102923
        ratio         1.310811      1.147691   1.497114
        rel. diff.    .2527473      .1457932   .3597014

        odds ratio         5.6      2.134896    18.5738   (exact)
```

As can be seen, the test statistics is the same to the result you would obtain from the hand calculation. (Please be careful that here the test statistics is Chi-square , which is the square of the z-score.)