

Chapter 520

Two Correlated Proportions (McNemar Test)

Introduction

This procedure computes confidence intervals and hypothesis tests for the comparison of the marginal frequencies of two factors (each with two levels) based on a 2-by-2 table of n pairs. Confidence limits can be obtained for the marginal probability difference, ratio, or odds ratio. Inequality tests are available for the marginal probability difference and ratio.

Experimental Design

A typical design for this scenario involves N pairs of individuals where a dichotomous measurement of one factor is measured on one of the individuals of the pair (case), and a second dichotomous measurement based on a second factor is measured on the second individual of the pair (control). Or similarly N individuals are measured twice, once for each of two dichotomous factors.

Comparing Two Correlated Proportions

Suppose you have two dichotomous measurements Y_1 and Y_2 on each of N subjects (where in many cases the 'subject' may be a pair of matched individuals). The proportions P_1 and P_2 represent the success probabilities. That is,

$$P_1 = \Pr(Y_1 = 1)$$

$$P_2 = \Pr(Y_2 = 1)$$

The data from this design can be summarized in the following 2-by-2 table:

	$Y_2 = 1$ (Yes, Present)	$Y_2 = 0$ (No, Absent)	Total
$Y_1 = 1$ (Yes, Present)	A	B	$A + B$
$Y_1 = 0$ (No, Absent)	C	D	$C + D$
Total	$A + C$	$B + D$	N

Two Correlated Proportions (McNemar Test)

The marginal proportions P_1 and P_2 are estimated from these data using the formulae

$$p_1 = \frac{A+B}{N} \text{ and } p_2 = \frac{A+C}{N}$$

Three quantities which allow these proportions to be compared are

<u>Quantity</u>	<u>Notation</u>
Difference	$\Delta = P_1 - P_2$
Risk Ratio	$\phi = P_1 / P_2$
Odds Ratio	$\psi = \frac{O_1}{O_2}$

Although these three parameters are (non-linear) functions of each other, the choice of which is to be used should not be taken lightly. The associated tests and confidence intervals of each of these parameters can vary widely in power and coverage probability.

Difference

The proportion (risk) difference $\delta = P_1 - P_2$ is perhaps the most direct method of comparison between the two event probabilities. This parameter is easy to interpret and communicate. It gives the absolute impact of the treatment. However, there are subtle difficulties that can arise with its interpretation.

One interpretation difficulty occurs when the event of interest is rare. If a difference of 0.001 were reported for an event with a baseline probability of 0.40, we would probably dismiss this as being of little importance. That is, there usually is little interest in a treatment that decreases the probability from 0.400 to 0.399. However, if the baseline probability of a disease was 0.002 and 0.001 was the decrease in the disease probability, this would represent a reduction of 50%. Thus we see that interpretation depends on the baseline probability of the event.

A similar situation occurs when the amount of possible difference is considered. Consider two events, one with a baseline event rate of 0.40 and the other with a rate of 0.02. What is the maximum decrease that can occur? Obviously, the first event rate can be decreased by an absolute amount of 0.40 while the second can only be decreased by a maximum of 0.02.

So, although creating the simple difference is a useful method of comparison, care must be taken that it fits the situation.

Ratio

The proportion (risk) ratio $\phi = p_1 / p_2$ gives the relative change in risk in a treatment group (group 1) compared to a control group (group 2). This parameter is also direct and easy to interpret. To compare this with the difference, consider a treatment that reduces the risk of disease from 0.1437 to 0.0793. Which single number is most enlightening, the fact that the absolute risk of disease has been decreased by 0.0644, or the fact that risk of disease in the treatment group is only 55.18% of that in the control group? In many cases, the percentage (100 x risk ratio) communicates the impact of the treatment better than the absolute change.

Perhaps the biggest drawback of this parameter is that it cannot be calculated in one of the most common experimental designs: the case-control study. Another drawback, when compared to the odds ratio, is that the odds ratio occurs naturally in the likelihood equations and as a parameter in logistic regression, while the proportion ratio does not.

Two Correlated Proportions (McNemar Test)

Odds Ratio

Chances are usually communicated as long-term proportions or probabilities. In betting, chances are often given as odds. For example, the odds of a horse winning a race might be set at 10-to-1 or 3-to-2. How do you translate from odds to probability? An odds of 3-to-2 means that the event will occur three out of five times. That is, an odds of 3-to-2 (1.5) translates to a probability of winning of 0.60.

The odds of an event are calculated by dividing the event risk by the non-event risk. Thus, in our case of two populations, the odds are

$$O_1 = \frac{P_1}{1 - P_1} \text{ and } O_2 = \frac{P_2}{1 - P_2}$$

For example, if P_1 is 0.60, the odds are $0.60/0.40 = 1.5$. In some cases, rather than representing the odds as a decimal amount, it is re-scaled into whole numbers. Thus, instead of saying the odds are 1.5-to-1, we may equivalently say they are 3-to-2.

In this context, the comparison of proportions may be done by comparing the odds through the ratio of the odds. The odds ratio of two events is

$$\begin{aligned} \psi &= \frac{O_1}{O_2} \\ &= \frac{\frac{P_1}{1 - P_1}}{\frac{P_2}{1 - P_2}} \end{aligned}$$

In the case of two correlated proportions, the odds ratio is calculated as

$$\psi = \frac{B}{C}$$

Until one is accustomed to working with odds, the odds ratio is usually more difficult to interpret than the proportion (risk) ratio, but it is still the parameter of choice for many researchers. Reasons for this include the fact that the odds ratio can be accurately estimated from case-control studies, while the risk ratio cannot. Also, the odds ratio is the basis of logistic regression (used to study the influence of risk factors). Furthermore, the odds ratio is the natural parameter in the conditional likelihood of the two-group, binomial-response design. Finally, when the baseline event-rates are rare, the odds ratio provides a close approximation to the risk ratio since, in this case, $1 - P_1 \approx 1 - P_2$, so that

$$\psi = \frac{\frac{P_1}{1 - P_1}}{\frac{P_2}{1 - P_2}} \approx \frac{P_1}{P_2} = \phi$$

One benefit of the log of the odds ratio is its desirable statistical properties, such as its continuous range from negative infinity to positive infinity.

Confidence Intervals

Several methods for computing confidence intervals for proportion difference, proportion ratio, and odds ratio have been proposed. We now show the methods that are available in NCSS.

Two Correlated Proportions (McNemar Test)

Difference

Four methods are available for computing a confidence interval of the difference between the two proportions $\Delta = P_1 - P_2$. The lower (L) and upper (U) limits of these intervals are computed as follows. Note that $z = |z_{\alpha/2}|$ is the appropriate percentile from the standard normal distribution.

Newcombe (1998) conducted a comparative evaluation of ten confidence interval methods. He recommended that the modified Wilson score method be used instead of the Pearson Chi-square or the Yate's Corrected Chi-square.

Nam's Score

For details, see Nam (1997) or Tango (1998). The lower limit is the solution of

$$L = \inf \left\{ \Delta_0 : \frac{\hat{\Delta} - \Delta_0}{\tilde{\sigma}_{\Delta_0}} < z \right\}$$

and the upper limit is the solution of

$$U = \sup \left\{ \Delta_0 : \frac{\hat{\Delta} - \Delta_0}{\tilde{\sigma}_{\Delta_0}} > -z \right\}$$

where $\tilde{\sigma}_{\Delta_0}$ is given by

$$\tilde{\sigma}_{\Delta} = \frac{\tilde{p}_{21} + \tilde{p}_{12} - \Delta^2}{n}$$

$$\tilde{p}_{21} = \left\{ \frac{-e + \sqrt{e^2 - 8f}}{4} \right\}$$

$$\tilde{p}_{12} = \tilde{p}_{21} - \Delta$$

$$e = -\hat{\Delta}(1 - \Delta) - 2(\hat{p}_{21} + \Delta)$$

$$f = \Delta(1 + \Delta)\hat{p}_{21}$$

Wilson's Score as modified by Newcombe

For further details, see Newcombe (1998c), page 2639. This is Newcombe's method 10.

$$L = \hat{\Delta} - \delta$$

$$U = \hat{\Delta} + \varepsilon$$

where

$$\delta = \sqrt{f_2^2 - 2\hat{\phi}f_2g_3 + g_3^2}$$

$$\varepsilon = \sqrt{g_2^2 - 2\hat{\phi}g_2f_3 + f_3^2}$$

$$f_2 = \frac{(A + B)}{N} - l_2$$

Two Correlated Proportions (McNemar Test)

$$g_2 = u_2 - \frac{(A+B)}{N}$$

$$f_3 = \frac{(A+C)}{N} - l_3$$

$$g_3 = u_3 - \frac{(A+C)}{N}$$

and l_2 and u_2 are the roots of

$$\left| x - \frac{A+B}{N} \right| = z \sqrt{\frac{x(1-x)}{N}}$$

and l_3 and u_3 are the roots of

$$\left| x - \frac{A+C}{N} \right| = z \sqrt{\frac{x(1-x)}{N}}$$

$\hat{\phi}$

$$\hat{\phi} = \begin{cases} \frac{\max(AD - BC - N/2, 0)}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} & \text{if } AD > BC \\ \frac{AD - BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} & \text{otherwise} \end{cases}$$

Note that if the denominator of $\hat{\phi}$ is zero, $\hat{\phi}$ is set to zero.

Wald Z Method

For further details, see Newcombe (1998c), page 2638

$$L = \hat{\Delta} - z s_W$$

$$U = \hat{\Delta} + z s_W$$

where

$$\hat{\Delta} = p_1 - p_2 = (B - C) / N$$

$$s_W^2 = \frac{(A+D)(B+C) + 4BC}{N^3}$$

Wald Z Method with Continuity Correction

For details, see Newcombe (1998c), page 2638.

$$L = \hat{\Delta} - z s_W - \frac{1}{N}$$

$$U = \hat{\Delta} + z s_W + \frac{1}{N}$$

Two Correlated Proportions (McNemar Test)

Ratio

Two methods are available for computing a confidence interval of the risk ratio $\phi = P_1 / P_2$. Note that $z = |z_{\alpha/2}|$ is the appropriate percentile from the standard normal distribution.

Nam and Blackwelder (2002) present two methods for computing confidence intervals for the risk ratio. These are presented here. Note that the score method is recommended.

Score (Nam and Blackwelder)

For details, see Nam and Blackwelder (2002), page 691. The lower limit is the solution of

$$z(\phi) = |z_{\alpha/2}|$$

and the upper limit is the solution of

$$z(\phi) = -|z_{\alpha/2}|$$

where

$$z(\phi) = \frac{\sqrt{N}(p_1 - \phi p_2)}{\sqrt{\phi(\tilde{p}_{12} + \tilde{p}_{21})}}$$

and

$$\tilde{p}_{12} = \frac{-p_1 + \phi^2(p_2 + 2p_{12}) + \sqrt{(p_1 - \phi p_2)^2 + 4\phi^2 p_{12} p_{12}}}{2\phi(\phi + 1)}$$

$$\tilde{p}_{21} = \phi \tilde{p}_{12} - (\phi - 1)(1 - p_{22})$$

Wald Z (Nam and Blackwelder)

For details, see Nam and Blackwelder (2002), page 692. The lower limit is the solution of

$$z_w(\phi) = |z_{\alpha/2}|$$

and the upper limit is the solution of

$$z_w(\phi) = -|z_{\alpha/2}|$$

where

$$z_w(\phi) = \frac{\sqrt{N}(\hat{p}_1 - \phi \hat{p}_2)}{\sqrt{\phi(\hat{p}_{12} + \hat{p}_{21})}}$$

Odds Ratio

Sahai and Khurshid (1995) present two methods for computing confidence intervals of the odds ratio $\psi = O_1 / O_2$. Note that the maximum likelihood estimate of this is given by

$$\hat{\psi} = B / C$$

Two Correlated Proportions (McNemar Test)

Exact Binomial

The lower limit is

$$\psi_L = \frac{B}{(C+1)F_{\alpha/2, 2C+2, 2B}}$$

and the upper limit

$$\psi_U = \frac{B+1}{CF_{\alpha/2, 2B+2, 2C}}$$

where F is the ordinate of the F distribution.

Maximum Likelihood

The lower limit is

$$\psi_L = \exp\{\ln(\hat{\psi}) - z_{\alpha/2}s_{\hat{\psi}}\}$$

and the upper limit

$$\psi_U = \exp\{\ln(\hat{\psi}) + z_{\alpha/2}s_{\hat{\psi}}\}$$

where

$$s_{\hat{\psi}} = \sqrt{\frac{1}{B} + \frac{1}{C}}$$

Hypothesis Tests

Difference

This module tests three statistical hypotheses about the difference in the two proportions:

1. $H_0 : P_1 - P_2 = \Delta$ versus $H_a : P_1 - P_2 \neq \Delta$; this is a *two-tailed test*.
2. $H_{0L} : P_1 - P_2 \geq \Delta$ versus $H_{aL} : P_1 - P_2 < \Delta$; this is a *one-tailed test*.
3. $H_{0U} : P_1 - P_2 \leq \Delta$ versus $H_{aU} : P_1 - P_2 > \Delta$; this is a *one-tailed test*.

Nam Test

Liu *et al.* (2002) recommend a likelihood score test which was originally proposed by Nam (1997). The tests are calculated as

$$z_L = \frac{\hat{\Delta} + \Delta}{\tilde{\sigma}_L} \quad \text{and} \quad z_U = \frac{\hat{\Delta} - \Delta}{\tilde{\sigma}_U}$$

where

$$\tilde{\sigma}_L = \sigma_{-\Delta}$$

$$\tilde{\sigma}_U = \sigma_{\Delta}$$

Two Correlated Proportions (McNemar Test)

and

$$\sigma_D = \frac{\tilde{p}_{21} + \tilde{p}_{12} - D^2}{N}$$

$$\tilde{p}_{21} = \left\{ \frac{-e + \sqrt{e^2 - 8f}}{4} \right\}$$

$$\tilde{p}_{12} = \tilde{p}_{21} - D$$

$$e = -\hat{\Delta}(1 - D) - 2(p_{21} + D)$$

$$f = D(1 + D)p_{21}$$

McNemar Test

Fleiss (1981) presents a test that is attributed to McNemar for testing the two-tailed null hypothesis. This is calculated as

$$\chi_1^2 = \frac{(B - C)^2}{B + C}$$

McNemar Test with Continuity Correction

Fleiss (1981) also presents a continuity-corrected version of McNemar test. This is calculated as

$$\chi_1^2 = \frac{(|B - C| - 1)^2}{B + C}$$

Wald Test

Liu *et al.* (2002) present a pair of large-sample, Wald-type z tests for testing the two one-tailed hypothesis about the difference $p_1 - p_2 = \Delta$. These are calculated as

$$z_L = \frac{\hat{\Delta} + \Delta - \frac{1}{2N}}{\hat{\sigma}} \quad \text{and} \quad z_U = \frac{\hat{\Delta} - \Delta + \frac{1}{2N}}{\hat{\sigma}}$$

where

$$\hat{\sigma}^2 = \frac{p_{21} + p_{12} - \hat{\Delta}^2}{N}$$

$$\hat{\Delta} = p_1 - p_2$$

Ratio

This module tests three statistical hypotheses about the difference in the two proportions:

1. $H_0 : P_1 / P_2 = \phi$ versus $H_a : P_1 / P_2 \neq \phi$; this is a *two-tailed test*.
2. $H_{0L} : P_1 / P_2 \geq \phi$ versus $H_{aL} : P_1 / P_2 < \phi$; this is a *one-tailed test*.
3. $H_{0U} : P_1 / P_2 \leq \phi$ versus $H_{aU} : P_1 / P_2 > \phi$; this is a *one-tailed test*.

Two Correlated Proportions (McNemar Test)

Nam Test

For details, see Nam and Blackwelder (2002), page 691. The test statistic for testing a specific value of ϕ is

$$z(\phi) = \frac{\sqrt{N}(p_1 - \phi p_2)}{\sqrt{\phi(\tilde{p}_{12} + \tilde{p}_{21})}}$$

where

$$\tilde{p}_{12} = \frac{-p_1 + \phi^2(p_2 + 2p_{12}) + \sqrt{(p_1 - \phi p_2)^2 + 4\phi^2 p_{12} p_{12}}}{2\phi(\phi + 1)}$$

$$\tilde{p}_{21} = \phi \tilde{p}_{12} - (\phi - 1)(1 - p_{22})$$

Data Structure

This procedure can summarize data from a database or summarized count values can be entered directly into the procedure panel.

Procedure Options

This section describes the options available in this procedure.

Data Tab

The data values can be entered directly on the panel as count totals or tabulated from columns of a database.

Type of Data Input

Choose from among two possible ways of entering the data.

- **Enter Table of Counts**

For this selection, each of the four response counts is entered directly.

Var1	Var2	
	Yes	No
Yes	43	157
No	65	135

The labels *Yes* and *No* are used here. Alternatives might instead be *Success* and *Failure*, *Presence* and *Absence*, *Positive* and *Negative*, *Disease* and *No Disease*, *1* or *0*, or something else, depending on the scenario.

- **Tabulate Counts from the Database**

Use this option when you have raw data that must be tabulated. You will be asked to select two columns on the database, one containing the values of variable 1 (such as *1* and *0* or *Yes* and *No*) and a second variable containing the values of variable 2 (such as *1* and *0* or *Yes* and *No*).

The data in these columns will be read and summarized.

Two Correlated Proportions (McNemar Test)

Headings and Labels (Used for Summary Tables)

Heading

Enter headings for variable 1 and variable 2. These headings will be used on the reports. They should be kept short so the report can be formatted correctly.

Labels

Enter labels for the first and second variables. Since these are often paired variables, the labels will often be the same for both variables.

Counts (Enter the Individual Cells)

Counts

Enter the counts in each of the four cells of the 2-by-2 table. Since these are counts, they must be a non-negative numbers. Usually, they will be integers, but this is not required.

Database Input

Variable 1

Specify one or more categorical variables used to define variable 1. If more than one variable is specified, a separate analysis is performed for each.

This procedure analyzes two values at a time. If a variable contains more than two unique values, a separate analysis is created for each pair of values.

Sorting

The values in each variable are sorted alpha-numerically. The first value after sorting becomes value one and the next value becomes value two. If you want the values to be analyzed in a different order, specify a custom *Value Order* for the column using the *Column Info Table* on the *Data Window*.

Variable 2

Specify one or more categorical variables used to define variable 2. If more than one variable is specified, a separate analysis is performed for each.

This procedure analyzes two values at a time. If a variable contains more than two unique values, a separate analysis is created for each pair of values.

Sorting

The values in each variable are sorted alpha-numerically. The first value after sorting becomes value one and the next value becomes value two. If you want the values to be analyzed in a different order, specify a custom *Value Order* for the column using the *Column Info Table* on the *Data Window*.

Frequency Variable

Specify an optional column containing the number of observations (cases) represented by each row. If this option is left blank, each row of the dataset is assumed to represent one observation.

Break Variables

Enter up to five categorical break variables. The values in these variables are used to break the output up into separate reports and plots. A separate set of reports is generated for each unique value (or unique combination of values if multiple break variables are specified).

Two Correlated Proportions (McNemar Test)

Zero Count Adjustment

Add a small adjustment value for zero counts

When zero counts are present, calculation problems for some formulas may result. Check this box to specify how you wish to add a small value either to all cells, or to all cells with zero counts. Adding a small value to cells is controversial, but may be necessary for obtaining results.

Zero Count Adjustment Method

Zero cell counts cause many calculation problems with ratios and odds ratios. To compensate for this, a small value (called the Zero Adjustment Value) may be added either to all cells or to all cells with zero counts. This option specifies whether you want to use the adjustment and which type of adjustment you want to use.

Zero Count Adjustment Value

Zero cell counts cause many calculation problems. To compensate for this, a small value may be added either to all cells or to all zero cells. The Zero Count Adjustment Value is the amount that is added. Adding a small value is controversial, but may be necessary. Some statisticians recommend adding 0.5 while others recommend 0.25. We have found that adding values as small as 0.0001 may also work well.

Summary Reports Tab

Test Alpha and Confidence Level

Alpha for Tests

Enter the value of alpha to be used for all hypothesis tests in this procedure. The probability level (p -value) is compared to alpha to determine whether to reject the null hypothesis.

Confidence Level

This is the confidence level for all confidence interval reports selected. The confidence level reflects the percent of the times that the confidence intervals would contain the true proportion difference if many samples were taken. Typical confidence levels are 90%, 95%, and 99%, with 95% being the most common.

Data Summary Reports

Use these check boxes to specify which summary reports are desired.

Difference Reports Tab

Confidence Intervals of the Difference ($P1 - P2$)

Use these check boxes to specify which confidence intervals are desired.

Inequality Tests of the Difference ($P1 - P2$)

Use these check boxes to specify which tests are desired.

Test Direction

Use these drop-downs to specify the direction of the test.

Two Correlated Proportions (McNemar Test)

Ratio Reports Tab

Confidence Intervals of the Ratio (P1/P2)

Use these check boxes to specify which confidence intervals are desired.

Inequality Tests of the Ratio (P1/P2)

Use these check boxes to specify which tests are desired.

Test Direction

Use these drop-downs to specify the direction of the test.

Odds Ratio Reports Tab

Confidence Intervals of the Odds Ratio (O1/O2)

Use these check boxes to specify which confidence intervals are desired.

Report Options Tab

This tab is used to specify the hypothesis test alpha, the data reports, and the report decimal places.

Report Options

These options only apply when the *Type of Data Input* option on the *Data* tab is set to *Tabulate Counts from the Database*.

Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

Value Labels

This option lets you select whether to display data values, value labels, or both. Use this option if you want the output to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

Report Decimal Places

Alpha – Percentages

These options specify the number of decimal places to be displayed when the data of that type is displayed on the output. This is the number of digits to the right of the decimal place to display for each type of value.

If one of the Auto options is used, the ending zero digits are not shown. For example, if *Auto (Up to 7)* is chosen, *0.0500* is displayed as *0.05* and *1.314583689* is displayed as *1.314584*.

The output formatting system is not designed to accommodate *Auto (Up to 13)*, and if chosen, this will likely lead to lines that run on to a second line. This option is included, however, for the rare case when a very large number of decimals is desired.

Table Formatting

These options only apply when *Individual Tables* or *Combined Tables* are selected on the *Summary Reports* tab.

Two Correlated Proportions (McNemar Test)

Column Justification

Specify whether data columns in the contingency tables will be left or right justified.

Column Widths

Specify how the widths of columns in the contingency tables will be determined.

The options are

- **Autosize to Minimum Widths**

Each data column is individually resized to the smallest width required to display the data in the column. This usually results in columns with different widths. This option produces the most compact table possible, displaying the most data per page.

- **Autosize to Equal Minimum Width**

The smallest width of each data column is calculated and then all columns are resized to the width of the widest column. This results in the most compact table possible where all data columns have the same width. This is the default setting.

- **Custom (User-Specified)**

Specify the widths (in inches) of the columns directly instead of having the software calculate them for you.

Custom Widths (Single Value or List)

Enter one or more values for the widths (in inches) of columns in the contingency tables. This option is only displayed if Column Widths is set to “Custom (User-Specified)”.

- **Single Value**

If you enter a single value, that value will be used as the width for all data columns in the table.

- **List of Values**

Enter a list of values separated by spaces corresponding to the widths of each column. The first value is used for the width of the first data column, the second for the width of the second data column, and so forth. Extra values will be ignored. If you enter fewer values than the number of columns, the last value in your list will be used for the remaining columns.

Type the word “Autosize” for any column to cause the program to calculate its width for you. For example, enter “1 Autosize 0.7” to make column 1 be 1 inch wide, column 2 be sized by the program, and column 3 be 0.7 inches wide.

Plots Tab

The options on this panel allow you to select and control the appearance of the plots output by this procedure.

Select and Format Plots

To display a plot for a table statistic, check the corresponding checkbox. The plots to choose from are:

- **Total Counts**
- **Total Proportions**

Click the appropriate plot format button to change the corresponding plot display settings.

Show Break as Title

Specify whether to display the values of the break variables as the second title line on the plots.

Example 1 – Analysis of Two Correlated Proportions

This section presents an example of how to run an analysis on hypothetical data. In this example, two dichotomous variables were measured on each of fifty subjects; 30 subjects scored ‘yes’ on both variables, 9 subjects scored ‘no’ on both variables, 6 scored a ‘yes’ and then a ‘no’, and 5 scored a ‘no and then a ‘yes’.

You may follow along here by making the appropriate entries or load the completed template **Example 1** by clicking on Open Example Template from the File menu of the Two Correlated Proportions (McNemar Test) window.

1 Open the Two Correlated Proportions (McNemar Test) window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Two Correlated Proportions (McNemar Test)** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

2 Specify the Data.

- Select the **Data** tab.
- Set **Type of Data Input** to **Enter Table of Counts**.

- In the **Variable 1, Heading** box, enter **Var1**.
- In the **Variable 1, Label of 1st Value** box, enter **Yes**.
- In the **Variable 1, Label of 2nd Value** box, enter **No**.
- In the **Variable 2, Heading** box, enter **Var2**.
- In the **Variable 2, Label of 1st Value** box, enter **Yes**.
- In the **Variable 2, Label of 2nd Value** box, enter **No**.

- In the **Var1 = Yes, Var2 = Yes** box, enter **30**.
- In the **Var1 = Yes, Var2 = No** box, enter **6**.
- In the **Var1 = No, Var2 = Yes** box, enter **5**.
- In the **Var1 = No, Var2 = No** box, enter **9**.

3 Specify the Summary Reports.

- Select the **Summary Reports** tab.
- Check **Counts and Proportions**.
- Check **Proportions Analysis**.

4 Specify the Difference Reports.

- Select the **Difference Reports** tab.
- Check **Nam RMLE Score** under **Confidence Intervals of the Difference**.
- Check **Wilson Score** under **Confidence Intervals of the Difference**.
- Set **Test Direction** to **Two-Sided**.
- Set **H0 Difference (D0)** to 0.0.
- Check **McNemar Test** under **Inequality Tests of the Difference**

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

Two Correlated Proportions (McNemar Test)

Counts and Proportions Sections

Counts and Proportions			
	<u>Var2</u>		
<u>Var1</u>	Yes	No	Total
	Count	Count	Count
Yes	30	6	36
No	5	9	14
Total	35	15	50

$p1 = (36/50) = 0.7200$
 $p2 = (14/50) = 0.2800$

Proportions Analysis	
Statistic	Value
Variable 1 Event Rate (p1)	0.7200
Variable 2 Event Rate (p2)	0.2800
Proportion Matching	0.7800
Proportion Not Matching	0.2200
Absolute Risk Difference $ p1 - p2 $	0.0200
Number Needed to Treat $1/ p1 - p2 $	50.0000
Relative Risk Reduction $ p1 - p2 /p2$	0.0286
Relative Risk $p1/p2$	1.0286
Odds Ratio $o1/o2$	1.2000

These reports document the values that were input, and give various statistics of these values.

Confidence Interval of the Difference (P1 - P2)

Confidence Intervals of the Difference (P1 - P2)						
Confidence Interval			Difference	Lower 95% C.L. of	Upper 95% C.L. of	Confidence Interval
Name	p1	p2	p1 - p2	P1 - P2	P1 - P2	Width
Nam RMLE Score*	0.7200	0.7000	0.0200	-0.1197	0.1606	0.2802
Wilson Score	0.7200	0.7000	0.0200	-0.1145	0.1537	0.2683

This report provides large sample confidence intervals of the difference.

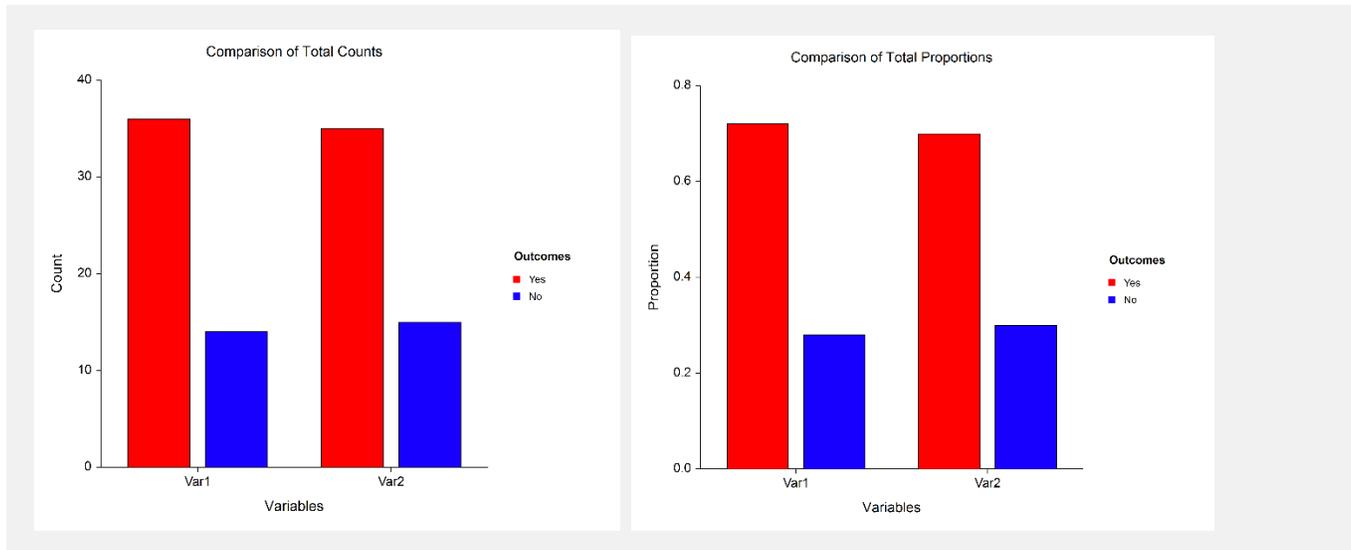
McNemar Inequality Test

Two-Sided Tests of the Difference (P1 - P2)						
H0: P1 - P2 = 0 vs. Ha: P1 - P2 ≠ 0						
Test Statistic			Difference	Test Statistic	Prob Level	Reject H0 at $\alpha = 0.05?$
Name	p1	p2	p1 - p2	Value	Level	
McNemar	0.7200	0.7000	0.0200	0.091	0.7630	No

This report provides the McNemar test. The p-value of the test is the Prob Level.

Two Correlated Proportions (McNemar Test)

Plots Section



These reports show the marginal totals and proportions of the two variables.