

## Chapter 519

# Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

---

### Introduction

This chapter documents three closely related procedures: non-inferiority tests, superiority (by a margin) tests, and equivalence tests. These procedures compute asymptotic confidence intervals and hypothesis tests for the difference, ratio, and odds ratio of two proportions.

#### Non-Inferiority Tests

Non-inferiority tests are one-sided hypothesis tests in which the null and alternative hypotheses are arranged to test whether one group is almost as good (not much worse) than the other group. So, we might be interested in showing that a new, less-expensive method of treatment is no worse than the current method of treatment.

#### Superiority Tests

Superiority tests are one-sided hypothesis tests in which the null and alternative hypotheses are arranged to test whether one group is better than the other group by more than a stated margin. So, we might be interested in showing that a new method of treatment is better by more than a clinically insignificant amount.

#### Equivalence Tests

Equivalence tests are used to show conclusively that two methods (i.e. drugs) are equivalent.

The conventional method of testing equivalence hypotheses is to perform two, one-sided tests (TOST) of hypotheses. The null hypothesis of non-equivalence is rejected in favor of the alternative hypothesis of equivalence if both one-sided tests are rejected. Unlike the common two-sided tests, however, the type I error rate is set directly at the nominal level (usually 0.05)—it is not split in half. So, to perform the test, two, one-sided tests are conducted at the significance level  $\alpha$ . If both are rejected, the alternative hypothesis is concluded at the significance level  $\alpha$ . Note that the p-value of the test is the maximum of the p-values of the two tests.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

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. Non-inferiority, superiority (by a margin) tests, and equivalence 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	$A + C$	$B + D$	$N$

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

<b>Quantity</b>	<b>Notation</b>
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.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 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.

### 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

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

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

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.

---

### 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

$$L = \inf \left\{ A_0 : \frac{\hat{\Delta} - A_0}{\tilde{\sigma}_{A_0}} < z \right\}$$

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

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

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

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

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

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

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

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

and  $l_2$  and  $u_2$  are the roots of

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

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

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.

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

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

---

## 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.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### Nam and Blackwelder Score

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

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

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

### Nam and Blackwelder Wald Z

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

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

## Hypothesis Tests

Several statistical tests are available for testing hypotheses about two correlated proportions. Some tests are based on the *difference* in the proportions while others are based on the *ratio* of proportions. In this section, some of these distinctions are explained.

### Types of Hypothesis Tests

The tests used in these procedures are large sample (or asymptotic) tests are based on the central limit theorem (CLT) which states that for large samples, the distribution of many of these test statistics approach the normal distribution. Hence, significance levels can be computed using the normal distribution which has been extensively tabulated and can now be easily computed.

A difficult determination when deciding whether to use a large sample test is whether or not the sample is large enough for the CLT to properly take effect. We suggest a minimum of 25 pairs of data points.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

---

### Non-Inferiority Tests

Non-inferiority tests are one-sided hypothesis tests in which the null and alternative hypotheses are arranged to test whether one group is almost as good (not much worse) than the other group. So, we might be interested in showing that a new, less-expensive method of treatment is no worse than the current method of treatment.

The statistical hypotheses used are as follows:

#### Differences

Assume that  $P_1$  and  $P_2$  are the event proportions of an experimental treatment and a control treatment, respectively. NIM is the non-inferiority margin.

##### If Lower Proportions are Worse

$$H_0: P_1 - P_2 \leq \text{NIM} \quad H_a: P_1 - P_2 > \text{NIM} \quad \text{NIM} < 0$$

NIM is the smallest amount that the new treatment's proportion is less than the current treatment's proportion and we still are willing to consider it not inferior.

##### If Lower Proportions are Better

$$H_0: P_1 - P_2 \geq \text{NIM} \quad H_a: P_1 - P_2 < \text{NIM} \quad \text{NIM} > 0$$

NIM is the largest amount that the new treatment's proportion is greater than the current treatment's proportion and we still are willing to consider it not inferior.

#### Ratios

Assume that  $P_1$  and  $P_2$  are the event proportions of an experimental treatment and a control treatment, respectively. NIR is the non-inferiority ratio.

##### If Lower Proportions are Worse

$$H_0: P_1 / P_2 \leq \text{NIR} \quad H_a: P_1 / P_2 > \text{NIR} \quad \text{NIR} < 1$$

NIR is the smallest that the proportion ratio can be and we still are willing to consider the experimental treatment not inferior to the control treatment.

##### If Lower Proportions are Better

$$H_0: P_1 / P_2 \geq \text{NIR} \quad H_a: P_1 / P_2 < \text{NIR} \quad \text{NIR} > 1$$

NIR is the largest that the proportion ratio can be and we still are willing to consider the experimental treatment not inferior to the control treatment.

---

### Superiority Tests

Superiority tests are one-sided hypothesis tests in which the null and alternative hypotheses are arranged to test whether one group is better than the other group by more than a stated margin. So, we might be interested in showing that a new method of treatment is better by more than a clinically insignificant amount.

The statistical hypotheses used are as follows:

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### Differences

Assume that  $P_1$  and  $P_2$  are the event proportions of an experimental treatment and a control treatment, respectively.  $SM$  is the superiority margin.

#### If Lower Proportions are Worse

$$H_0: P_1 - P_2 \leq SM \quad H_a: P_1 - P_2 > SM \quad SM > 0$$

$NIM$  is the smallest amount that the new treatment's proportion is greater than the current treatment's proportion and we still are willing to consider it superior.

#### If Lower Proportions are Better

$$H_0: P_1 - P_2 \geq SM \quad H_a: P_1 - P_2 < SM \quad SM < 0$$

$NIM$  is the largest amount that the new treatment's proportion is less than the current treatment's proportion and we still are willing to consider it superior.

### Ratios

Assume that  $P_1$  and  $P_2$  are the event proportions of an experimental treatment and a control treatment, respectively.  $SR$  is the superiority ratio.

#### If Lower Proportions are Worse

$$H_0: P_1 / P_2 \leq SR \quad H_a: P_1 / P_2 > SR \quad SR > 1$$

$SR$  is the smallest that the proportion ratio can be and we still are willing to consider the experimental treatment superior to the control treatment.

#### If Lower Proportions are Better

$$H_0: P_1 / P_2 \geq SR \quad H_a: P_1 / P_2 < SR \quad SR < 1$$

$SR$  is the largest that the proportion ratio can be and we still are willing to consider the experimental treatment superior to the control treatment.

---

## Equivalence Tests

Equivalence tests are used to show conclusively that two methods (i.e. drugs) are equivalent.

The conventional method of testing equivalence hypotheses is to perform two, one-sided tests (TOST) of hypotheses. The null hypothesis of non-equivalence is rejected in favor of the alternative hypothesis of equivalence if both one-sided tests are rejected. Unlike the common two-sided tests, however, the type I error rate is set directly at the nominal level (usually 0.05)—it is not split in half. So, to perform the test, two, one-sided tests are conducted at the significance level  $\alpha$ . If both are rejected, the alternative hypothesis is concluded at the significance level  $\alpha$ . Note that the p-value of the test is the maximum of the p-values of the two tests.

### Differences

Assume that  $P_1$  and  $P_2$  are the event proportions of an experimental treatment and a control treatment, respectively.

$$H_0: P_1 - P_2 \leq LM \text{ or } P_1 - P_2 \geq UM \quad H_a: LM < P_1 - P_2 < UM \text{ (Equivalence)} \quad LM < 0, UM > 0$$

$LM$  and  $UM$  are the lower and upper margins of equivalence, respectively. If the proportion difference is between these bounds, we are willing to say that the two treatments are equivalent.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### Ratios

Assume that  $P_1$  and  $P_2$  are the event proportions of an experimental treatment and a control treatment, respectively.

$H_0: P_1 / P_2 \leq LR$  or  $P_1 / P_2 \geq UR$        $H_a: LR < P_1 / P_2 < UR$  (Equivalence)       $LR < 1, UR > 1$

$LR$  and  $UR$  are the lower and upper margins of equivalence, respectively. If the proportion ratio is between these bounds, we are willing to say that the two treatments are equivalent.

## Test Statistics

This section documents the various test statistics that are available.

### Difference

These modules test statistical hypotheses about the difference in the two proportions:

1.  $H_{0L}: P_1 - P_2 \geq \Delta$  versus  $H_{aL}: P_1 - P_2 < \Delta$ ; this is a *one-tailed test*.
2.  $H_{0U}: P_1 - P_2 \leq \Delta$  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

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

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

and

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

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

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

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

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

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 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

These modules test three statistical hypotheses about the difference in the two proportions:

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

### Nam Test

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

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

where

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

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

---

### Equivalence Tests

Equivalence tests are hypothesis tests in which the alternative hypothesis, not the null hypothesis, is equality. For example, suppose an accurate diagnostic test has serious side effects, so a replacement test is sought. In this case, researchers are not interested in showing that the two tests are different, but rather that they are the same.

These tests are often divided into two categories: *equivalence* (two-sided) tests and *non-inferiority* (one-sided) tests. Here, the term *equivalence tests* means that we want to show that two tests are equivalent—that is, their accuracy is about the same. This requires a two-sided hypothesis test. On the other hand, *non-inferiority tests* are used when we want to show that a new (experimental) test is no worse than the existing (reference or gold-standard) test. This requires a one-sided hypothesis test.

In the discussion to follow, two ways of expressing difference are considered: the difference and the ratio. The simple difference between two proportions is perhaps the most straight forward way of expressing that these two proportions are different. A difference of zero means that the two proportions are equal. Unfortunately, this method does not work well near zero or one. For example, suppose a diagnostic test achieves 95% accuracy and

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

we wish to establish that a new test is within 7 percentage points of the original test. The acceptable range is from  $95\% - 7\% = 88\%$  to  $95\% + 7\% = 102\%$ . Of course, 102% is impossible.

A second method of setting up the hypotheses that does not suffer from this problem is to consider the ratio of the two proportions. A ratio of one indicates that the two proportions are equal. Using the ratio to define the hypotheses is less prone to impossible values.

### Equivalence Based on the Difference

The equivalence between two proportions may be tested using the following hypothesis which is based on the difference:

$$H_0: P_1 - P_2 \leq \Delta_L \text{ or } H_0: P_1 - P_2 \geq \Delta_U \text{ vs. } H_a: \Delta_L < P_1 - P_2 < \Delta_U$$

where the  $\Delta$  are the established equivalence limits.

This hypothesis is often tested at the alpha significance level by determining if  $100(1 - 2\alpha)\%$  confidence limits are both between  $\Delta_L$  and  $\Delta_U$ . A second method to test this hypothesis is to separate it into two sets of one-sided hypothesis tests (TOST).

#### First One-Sided Test

$$H_{0L}: P_1 - P_2 \leq \Delta_L \text{ versus } H_{aL}: P_1 - P_2 > \Delta_L.$$

#### Second One-Sided Test

$$H_{0U}: P_1 - P_2 \geq \Delta_U \text{ versus } H_{aU}: P_1 - P_2 < \Delta_U.$$

#### Nam Test

Both the two, one-sided hypotheses method and the confidence interval test are tested using the score test proposed by Nam (1997). These tests and the confidence interval formulas were presented above.

### Equivalence Based on the Ratio

The equivalence between two proportions may be tested using the following hypothesis which is based on the ratio:

$$H_0: P_1/P_2 \leq \phi_L \text{ or } H_0: P_1/P_2 \geq \phi_U \text{ vs. } H_a: \phi_L < P_1/P_2 < \phi_U$$

where the  $\phi$  are the established equivalence limits.

This hypothesis is often tested at the alpha significance level by determining if  $100(1 - 2\alpha)\%$  confidence limits are both between  $\phi_L$  and  $\phi_U$ . A second method to test this hypothesis is to separate it into two sets of one-sided hypotheses (TOST).

#### First One-Sided Test

$$H_{0L}: P_1/P_2 \leq \phi_L \text{ versus } H_{aL}: P_1/P_2 > \phi_L.$$

#### Second One-Sided Test

$$H_{0U}: P_1/P_2 \geq \phi_U \text{ versus } H_{aU}: P_1/P_2 < \phi_U.$$

#### Nam Test

Both the two, one-sided hypotheses method and the confidence interval test are tested using the score test proposed by Nam and Blackwelder (2002). These tests and the confidence interval were presented above.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 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.

### 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.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

---

### 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).

---

### 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.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 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.

---

### Non-Inferiority Tests of the Difference (P1 – P2)

Use the check boxes to specify which tests are desired.

#### Lower Proportions are

This option specifies whether lower proportions are **better** or **worse**. This implicitly defines the direction of the test. For example, if the event of interest is whether an experimental treatment cured a certain disease, then lower proportions are ‘worse’. On the other hand, if the event of interest is a failure, such as a smartphone quits working, then lower proportions are ‘better’.

#### Non-Inferiority Margin (NIM)

Specify the difference margin (P1-P2). If lower proportions are *better*, this is the amount that P1 can be greater than P2 and you still conclude that variable 1 is not inferior to variable 2. This value should be between 0.0 and 0.3.

If lower proportions are *worse*, this is the amount that P1 can be less than P2 and you still conclude that variable 1 is not inferior to variable 2. This value should be between -0.3 and 0.0.

Often, NIM will be determined as a percentage of P2 (the variable 2 proportion). Common percentages are 20% and 25%. For example, if lower proportions are better, P2 is 0.7, and 20% is selected, the value of NIM would be  $0.7 \times 0.2 = 0.14$ .

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

---

### Superiority Tests of the Difference (P1 – P2)

Use the check boxes to specify which tests are desired.

#### Lower Proportions are

This option specifies whether lower proportions are **better** or **worse**. This implicitly defines the direction of the test. For example, if the event of interest is whether an experimental treatment cured a certain disease, then lower proportions are ‘worse’. On the other hand, if the event of interest is a failure, such as a smartphone quits working, then lower proportions are ‘better’.

#### Superiority Margin (SM)

Specify the difference margin (P1-P2). If lower proportions are *better*, this is the minimum amount that P1 must be less than P2 and you conclude that variable 1 is superior to variable 2. This value should be between -0.3 and 0.0.

If lower proportions are *worse*, this is the minimum amount that P1 can be less than P2 and you still conclude that variable 1 is superior to variable 2. This value should be between 0.0 and 0.3.

---

### Equivalence Tests Based on the Difference

Use the check boxes to specify which tests are desired.

#### Upper Equivalence Bound

Specify the upper bound for the equivalence test of the difference in proportions. If the difference in the two proportions is greater than this value, the two variables are not equivalent. This value is sometimes called the *Margin of Equivalence*. The possible range of values is between 0 and 1. Typical values are between 0.05 and 0.25.

#### Lower Equivalence Bound

Specify the lower bound for the equivalence test of the difference in population proportions. If the difference in the two population proportions is less than this value, the two variables are not equivalent. This value is sometimes called the *Margin of Equivalence*. The possible range of values is between 0 and 1. Typical values are between 0.05 and 0.25.

If this value is left blank, then the negative of the Upper Equivalence Bound is used.

---

## Ratio Reports Tab

---

### Confidence Intervals of the Ratio (P1/P2)

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

---

### Non-Inferiority Tests of the Ratio (P1/P2)

Use the check boxes to specify which tests are desired.

#### Lower Proportions are

This option specifies whether lower proportions are **better** or **worse**. This implicitly defines the direction of the test. For example, if the event of interest is whether an experimental treatment cured a certain disease, then lower proportions are ‘worse’. On the other hand, if the event of interest is a failure, such as a smartphone quits working, then lower proportions are ‘better’.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### Non-Inferiority Ratio (NIR)

Specify the non-inferiority ratio ( $P1/P2$ ). If lower proportions are *better*, this is the largest the true ratio can be and you still conclude that variable 1 is not inferior to variable 2. The range is between 1 and about 1.3.

If lower proportions are *worse*, this is the smallest the true ratio can be and you still conclude that variable 1 is not inferior to variable 2. The range is between 0.5 and 1.0.

---

### Superiority Tests of the Ratio (P1/P2)

Use the check boxes to specify which tests are desired.

#### Lower Proportions are

This option specifies whether lower proportions are *better* or *worse*. This implicitly defines the direction of the test. For example, if the event of interest is whether an experimental treatment cured a certain disease, then lower proportions are worse. On the other hand, if the event of interest is a failure, such as a smartphone quits working, then lower proportions are better.

#### Superiority Ratio (SR)

Specify the superiority ratio ( $P1/P2$ ). If lower proportions are *better*, this is the smallest the true ratio can be and you still conclude that variable 1 is superior to variable 2. The range is between 0.5 and 1.

If lower proportions are *worse*, this is the smallest the true ratio can be and you still conclude that variable 1 is superior variable 2. The range is between 1.0 and 1.3.

---

### Equivalence Tests Based on the Ratio

Use the check boxes to specify which tests are desired.

#### Upper Equivalence Bound

Specify the upper bound for the equivalence test of the ratio of proportions. If the ratio of the two population proportions is greater than this value, the two variables are not equivalent. The possible range of values is greater than 1. Typical values are between 1 and 2.

#### Lower Equivalence Bound

Specify the lower bound for the equivalence test of the ratio in proportions. If the ratio of the two population proportions is less than this value, the two variables are not equivalent. The possible range of values is between 0 and 1. Typical values are between 0.5 and 0.99.

If this value is left blank, then the inverse of the Upper Equivalence Bound is used.

---

## 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.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 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.

#### 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.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

- **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 it's 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.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### Example 1 – Non-Inferiority Test of Two Correlated Proportions

A study was made of 100 matched pairs of individuals. One individual from each pair was given a new treatment and the other individual in the pair was given a standard treatment. The choice was made by the flip of a coin. The response of each individual is either recovery or no recovery. The researchers wish to show that the new treatment is **no worse** than the standard treatment. The results were tabulated into the following table. Note that the table is arranged so that the new treatment is shown as the rows of the table and the standard treatment is shown as the columns of the table.

New Treatment	Standard Treatment	
	Recovery	No Recovery
Recovery	53	23
No Recovery	15	9

The researchers wish to show that the proportion recovering with the new treatment is no less than 0.05 below the proportion recovering with the standard treatment. The value for alpha to be used is 0.025.

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 – Non-Inferiority Tests window.

#### 1 Open the Two Correlated Proportions – Non-Inferiority Tests window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Two Correlated Proportions – Non-Inferiority Tests** 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 **New**.
- In the **Variable 1, Label of 1<sup>st</sup> Value** box, enter **Recovery**.
- In the **Variable 1, Label of 2<sup>nd</sup> Value** box, enter **No Recovery**.
- In the **Variable 2, Heading** box, enter **Standard**.
- In the **Variable 2, Label of 1<sup>st</sup> Value** box, enter **Recovery**.
- In the **Variable 2, Label of 2<sup>nd</sup> Value** box, enter **No Recovery**.
- In the **Var1 = Yes, Var2 = Yes** box, enter **53**.
- In the **Var1 = Yes, Var2 = No** box, enter **23**.
- In the **Var1 = No, Var2 = Yes** box, enter **15**.
- In the **Var1 = No, Var2 = No** box, enter **9**.

#### 3 Specify the Summary Reports.

- Select the **Summary Reports** tab.
- Set **Alpha for Tests** to **0.025**.
- Check **Counts and Proportions**.
- Check **Proportions Analysis**.

#### 4 Specify the Difference Reports.

- Select the **Difference Reports** tab.
- Set **Lower Proportions are to Worse**.
- Set **Non-Inferiority Margin (NIM)** to **-0.05**.
- Check **Nam Score Test** under **Non-Inferiority Tests of the Difference**

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 5 Run the procedure.

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

## Counts and Proportions Sections

Counts and Proportions			
<u>Standard</u>			
<u>New</u>	Recovery	No Recovery	Total
	Count	Count	Count
Recovery	53	23	76
No Recovery	15	9	24
Total	68	32	100

$p1 = (76/100) = 0.7600$   
 $p2 = (68/100) = 0.6800$

Proportions Analysis	
Statistic	Value
Variable 1 Event Rate (p1)	0.7600
Variable 2 Event Rate (p2)	0.6800
Proportion Matching	0.6200
Proportion Not Matching	0.3800
Absolute Risk Difference $ p1 - p2 $	0.0800
Number Needed to Treat $1/ p1 - p2 $	12.5000
Relative Risk Reduction $ p1 - p2 /p2$	0.1176
Relative Risk $p1/p2$	1.1176
Odds Ratio $o1/o2$	1.5333

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

## Non-Inferiority Test Report

Upper Non-Inferiority Tests of the Difference (P1 - P2)						
H0: P1 - P2 ≤ -0.05 vs. Ha: P1 - P2 > -0.05						
Test			Difference	Test	Prob	Reject
Statistic	p1	p2	p1 - p2	Statistic	Level	H0 at
Name				Value		$\alpha = 0.025?$
Nam Score*	0.7600	0.6800	0.0800	2.088	0.0184	Yes

\*It is recommended that N be at least 25 when using the Nam Score test.

This report provides the Nam score, non-inferiority test. The p-value of the test is the Prob Level.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### Example 2 – Superiority Test of Two Correlated Proportions

A study was made of 100 matched pairs of individuals. One individual from each pair was given a new treatment and the other individual in the pair was given a standard treatment. The choice was made by the flip of a coin. The response of each individual is either recovery or no recovery. The researchers wish to show that the new treatment is **definitely better** than the standard treatment. The results were tabulated into the following table. Note that the table is arranged so that the new treatment is shown as the rows of the table and the standard treatment is shown as the columns of the table.

New Treatment	Standard Treatment	
	Recovery	No Recovery
Recovery	50	26
No Recovery	9	15

The researchers wish to show that the proportion recovering with the new treatment is at least 0.05 above the proportion recovering with the standard treatment. The value for alpha to be used is 0.025.

You may follow along here by making the appropriate entries or load the completed template **Example 2** by clicking on Open Example Template from the File menu of the Two Correlated Proportions – Superiority Tests window.

#### 1 Open the Two Correlated Proportions – Superiority Tests window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Two Correlated Proportions – Superiority Tests** 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 **New**.
- In the **Variable 1, Label of 1<sup>st</sup> Value** box, enter **Recovery**.
- In the **Variable 1, Label of 2<sup>nd</sup> Value** box, enter **No Recovery**.
- In the **Variable 2, Heading** box, enter **Standard**.
- In the **Variable 2, Label of 1<sup>st</sup> Value** box, enter **Recovery**.
- In the **Variable 2, Label of 2<sup>nd</sup> Value** box, enter **No Recovery**.
- In the **Var1 = Yes, Var2 = Yes** box, enter **50**.
- In the **Var1 = Yes, Var2 = No** box, enter **26**.
- In the **Var1 = No, Var2 = Yes** box, enter **9**.
- In the **Var1 = No, Var2 = No** box, enter **15**.

#### 3 Specify the Summary Reports.

- Select the **Summary Reports** tab.
- Set **Alpha for Tests** to **0.025**.
- Check **Counts and Proportions**.
- Check **Proportions Analysis**.

#### 4 Specify the Difference Reports.

- Select the **Difference Reports** tab.
- Set **Lower Proportions are to Worse**.
- Set **Superiority Margin (SM)** to **0.05**.
- Check **Nam Score Test** under **Superiority Tests of the Difference**

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 5 Run the procedure.

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

## Counts and Proportions Sections

Counts and Proportions			
<u>Standard</u>			
<u>New</u>	Recovery	No Recovery	Total
	Count	Count	Count
Recovery	50	26	76
No Recovery	9	15	24
Total	59	41	100

$p1 = (76/100) = 0.7600$   
 $p2 = (68/100) = 0.5900$

Proportions Analysis	
Statistic	Value
Variable 1 Event Rate (p1)	0.7600
Variable 2 Event Rate (p2)	0.5900
Proportion Matching	0.6500
Proportion Not Matching	0.3500
Absolute Risk Difference $ p1 - p2 $	0.1700
Number Needed to Treat $1/ p1 - p2 $	5.8824
Relative Risk Reduction $ p1 - p2 /p2$	0.2881
Relative Risk $p1/p2$	1.2881
Odds Ratio $o1/o2$	2.8889

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

## Superiority Test Report

Upper Superiority Tests of the Difference (P1 - P2)						
H0: P1 - P2 ≤ 0.05 vs. Ha: P1 - P2 > 0.05						
Test			Difference	Test	Prob	Reject
Statistic	p1	p2	p1 - p2	Statistic	Level	H0 at
Name				Value		$\alpha = 0.025?$
Nam Score*	0.7600	0.5900	0.1700	2.071	0.0192	Yes

\*It is recommended that N be at least 25 when using the Nam Score test.

This report provides the Nam score, Superiority test. The p-value of the test is the Prob Level.

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### Example 3 – Equivalence Test of Two Correlated Proportions

A study was made of 200 matched pairs of individuals. One individual from each pair was given a new treatment and the other individual in the pair was given a standard treatment. The choice was made by the flip of a coin. The response of each individual is either recovery or no recovery. The researchers wish to show that the new treatment is **equivalent** to the standard treatment. The results were tabulated into the following table. Note that the table is arranged so that the new treatment is shown as the rows of the table and the standard treatment is shown as the columns of the table.

New Treatment	Standard Treatment	
	Recovery	No Recovery
Recovery	106	36
No Recovery	40	18

The researchers wish to show that the proportion recovering with the new treatment is no less than 0.1 above or below the proportion recovering with the standard treatment. The value for alpha to be used is 0.05.

You may follow along here by making the appropriate entries or load the completed template **Example 3** by clicking on Open Example Template from the File menu of the Two Correlated Proportions – Equivalence Tests window.

#### 1 Open the Two Correlated Proportions – Equivalence Tests window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Two Correlated Proportions – Equivalence Tests** 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 **New**.
- In the **Variable 1, Label of 1<sup>st</sup> Value** box, enter **Recovery**.
- In the **Variable 1, Label of 2<sup>nd</sup> Value** box, enter **No Recovery**.
- In the **Variable 2, Heading** box, enter **Standard**.
- In the **Variable 2, Label of 1<sup>st</sup> Value** box, enter **Recovery**.
- In the **Variable 2, Label of 2<sup>nd</sup> Value** box, enter **No Recovery**.
- In the **Var1 = Yes, Var2 = Yes** box, enter **106**.
- In the **Var1 = Yes, Var2 = No** box, enter **36**.
- In the **Var1 = No, Var2 = Yes** box, enter **40**.
- In the **Var1 = No, Var2 = No** box, enter **18**.

#### 3 Specify the Summary Reports.

- Select the **Summary Reports** tab.
- Set **Alpha for Tests** to **0.05**.
- Check **Counts and Proportions**.
- Check **Proportions Analysis**.

#### 4 Specify the Difference Reports.

- Select the **Difference Reports** tab.
- Set **Upper Equivalence Bound** to **0.1**.
- Set **Lower Equivalence Bound** to **-0.1**.
- Check **Equivalence Test** under **Equivalence Tests of the Difference**

## Two Correlated Proportions – Non-Inferiority, Superiority, and Equivalence Tests

### 5 Run the procedure.

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

## Counts and Proportions Sections

Counts and Proportions			
<u>Standard</u>			
<u>New</u>	Recovery	No Recovery	Total
	Count	Count	Count
Recovery	106	36	142
No Recovery	40	18	58
Total	146	54	200

$p1 = (142/200) = 0.7100$   
 $p2 = (146/200) = 0.7300$

Proportions Analysis	
Statistic	Value
Variable 1 Event Rate (p1)	0.7100
Variable 2 Event Rate (p2)	0.7300
Proportion Matching	0.6200
Proportion Not Matching	0.3800
Absolute Risk Difference $ p1 - p2 $	0.0200
Number Needed to Treat $1/ p1 - p2 $	50.0000
Relative Risk Reduction $ p1 - p2 /p2$	0.0274
Relative Risk $p1/p2$	0.9726
Odds Ratio $o1/o2$	0.9000

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

## Equivalence Test Report

Equivalence Test of the Difference (P1 - P2)							
H0: $P1 - P2 \leq -0.1$ or $P1 - P2 \geq 0.1$ vs. Ha: Equivalence							
Test Name	Difference p1 - p2	Lower Test Statistic	Lower Prob Level	Upper Test Statistic	Upper Prob Level	TOST Prob Level	Reject H0 at $\alpha = 0.05?$
Nam Score*	-0.0200	1.829	0.0337	-2.722	0.0032	0.0337	Yes

\*It is recommended that N be at least 25 when using the Nam Score test.

This report provides the Nam score equivalence test. The p-value of the test is the TOST Prob Level.