Chapter 210

Non-Inferiority Tests for Two Proportions

Introduction

This module provides power analysis and sample size calculation for non-inferiority and superiority tests in two-sample designs in which the outcome is binary. Users may choose from among eight popular test statistics commonly used for running the hypothesis test.

The power calculations assume that independent, random samples are drawn from two populations.

Four Procedures Documented Here

There are four procedures in the menus that use the program module described in this chapter. These procedures are identical except for the type of parameterization. The parameterization can be in terms of proportions, differences in proportions, ratios of proportions, and odds ratios. Each of these options is listed separately on the menus.

Example

A non-inferiority test example will set the stage for the discussion of the terminology that follows. Suppose that the current treatment for a disease works 70% of the time. Unfortunately, this treatment is expensive and occasionally exhibits serious side-effects. A promising new treatment has been developed to the point where it can be tested. One of the first questions that must be answered is whether the new treatment is as good as the current treatment. In other words, do at least 70% of treated subjects respond to the new treatment?

Because of the many benefits of the new treatment, clinicians are willing to adopt the new treatment even if it is slightly less effective than the current treatment. They must determine, however, how much less effective the new treatment can be and still be adopted. Should it be adopted if 69% respond? 68%? 65%? 60%? There is a percentage below 70% at which the difference between the two treatments is no longer considered ignorable. After thoughtful discussion with several clinicians, it was decided that if a response of at least 63% were achieved, the new treatment would be adopted. The difference between these two percentages is called the margin of equivalence. The margin of equivalence in this example is 7%.

The developers must design an experiment to test the hypothesis that the response rate of the new treatment is at least 0.63. The statistical hypothesis to be tested is

\[ H_0: p_1 - p_2 \leq -0.07 \quad \text{versus} \quad H_1: p_1 - p_2 > -0.07 \]

Notice that when the null hypothesis is rejected, the conclusion is that the response rate is at least 0.63. Note that even though the response rate of the current treatment is 0.70, the hypothesis test is about a response rate of 0.63. Also notice that a rejection of the null hypothesis results in the conclusion of interest.
Technical Details

The details of sample size calculation for the two-sample design for binary outcomes are presented in the chapter “Two Proportions Non-Null Case,” and they will not be duplicated here. Instead, this chapter only discusses those changes necessary for non-inferiority and superiority tests.

Approximate sample size formulas for non-inferiority tests of two proportions are presented in Chow et al. (2003), page 90. Only large sample (normal approximation) results are given there. The results available in this module use exact calculations based on the enumeration of all possible values in the binomial distribution.

Suppose you have two populations from which dichotomous (binary) responses will be recorded. Assume without loss of generality that the higher proportions are better. The probability (or risk) of cure in population 1 (the treatment group) is \(p_1\) and in population 2 (the reference group) is \(p_2\). Random samples of \(n_1\) and \(n_2\) individuals are obtained from these two populations. The data from these samples can be displayed in a 2-by-2 contingency table as follows:

<table>
<thead>
<tr>
<th>Group</th>
<th>Success</th>
<th>Failure</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>(x_{11})</td>
<td>(x_{12})</td>
<td>(n_1)</td>
</tr>
<tr>
<td>Control</td>
<td>(x_{21})</td>
<td>(x_{22})</td>
<td>(n_2)</td>
</tr>
<tr>
<td>Totals</td>
<td>(m_1)</td>
<td>(m_2)</td>
<td>(N)</td>
</tr>
</tbody>
</table>

The binomial proportions, \(p_1\) and \(p_2\), are estimated from these data using the formulae

\[
\hat{p}_1 = \frac{a}{m} = \frac{x_{11}}{n_1} \quad \text{and} \quad \hat{p}_2 = \frac{b}{n} = \frac{x_{21}}{n_2}
\]

Let \(p_{1,0}\) represent the group 1 proportion tested by the null hypothesis, \(H_0\). The power of a test is computed at a specific value of the proportion which we will call \(p_{1,1}\). Let \(\delta\) represent the smallest difference (margin of equivalence) between the two proportions that still results in the conclusion that the new treatment is not inferior to the current treatment. For a non-inferiority test, \(\delta < 0\). The set of statistical hypotheses that are tested is

\[
H_0: p_{1,0} - p_2 \leq \delta \quad \text{versus} \quad H_1: p_{1,0} - p_2 > \delta
\]

which can be rearranged to give

\[
H_0: p_{1,0} \leq p_2 + \delta \quad \text{versus} \quad H_1: p_{1,0} > p_2 + \delta
\]

There are three common methods of specifying the margin of equivalence. The most direct is to simply give values for \(p_2\) and \(p_{1,0}\). However, it is often more meaningful to give \(p_2\) and then specify \(p_{1,0}\) implicitly by specifying the difference, ratio, or odds ratio. Mathematically, the definitions of these parameterizations are

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Computation</th>
<th>Hypotheses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Difference</td>
<td>(\delta = p_{1,0} - p_2)</td>
<td>(H_0: p_{1,0} - p_2 \leq \delta_0 \ vs. \ H_1: p_{1,0} - p_2 &gt; \delta_0, \ \delta_0 &lt; 0)</td>
</tr>
<tr>
<td>Ratio</td>
<td>(\phi = p_{1,0} / p_2)</td>
<td>(H_0: p_1 / p_2 \leq \phi_0 \ vs. \ H_1: p_1 / p_2 &gt; \phi_0, \ \phi_0 &lt; 1)</td>
</tr>
<tr>
<td>Odds Ratio</td>
<td>(\psi = \text{Odds}_{1,0} / \text{Odds}_2)</td>
<td>(H_0: \text{Odds}_{1,0} / \text{Odds}<em>2 \leq \psi_0 \ vs. \ H_1: \text{Odds}</em>{1,0} / \text{Odds}_2 &gt; \psi_0, \ \psi_0 &lt; 1)</td>
</tr>
</tbody>
</table>
The difference is perhaps the most direct method of comparison between two proportions. It 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 difficulty arises when the event of interest is rare. If a difference of 0.001 occurs when the baseline probability is 0.40, it would be dismissed as being trivial. However, if the baseline probability of a disease is 0.002, a 0.001 decrease would represent a reduction of 50%. Thus, interpretation of the difference depends on the baseline probability of the event.

Note that if $\delta < 0$, the procedure is called a non-inferiority test while if $\delta > 0$ the procedure is called a superiority test.

### Non-Inferiority using a Difference

The following example might help you understand the concept of a non-inferiority test. Suppose 60% of patients respond to the current treatment method ($p_2 = 0.60$). If the response rate of the new treatment is no less than 5 percentage points worse ($\delta = -0.05$) than the existing treatment, it will be considered to be noninferior. Substituting these figures into the statistical hypotheses gives

$$H_0: \delta \leq -0.05 \text{ versus } H_1: \delta > -0.05$$

Using the relationship

$$p_{1.0} = p_2 + \delta$$

gives

$$H_0: p_{1.0} \leq 0.55 \text{ versus } H_1: p_{1.0} > 0.55$$

In this example, when the null hypothesis is rejected, the concluded alternative is that the response rate is at least 55%, which means that the new treatment is not inferior to the current treatment.

### Superiority using a Difference

The following example is intended to help you understand the concept of a superiority test. Suppose 60% of patients respond to the current treatment method ($p_2 = 0.60$). If the response rate of the new treatment is at least 10 percentage points better ($\delta = 0.10$), it will be considered to be superior to the existing treatment. Substituting these figures into the statistical hypotheses gives

$$H_0: \delta \leq 0.10 \text{ versus } H_1: \delta > 0.10$$

Using the relationship

$$p_{1.0} = p_2 + \delta$$

gives

$$H_0: p_{1.0} \leq 0.70 \text{ versus } H_1: p_{1.0} > 0.70$$

In this example, when the null hypothesis is rejected, the concluded alternative is that the response rate is at least 0.70. That is, the conclusion of superiority is that the new treatment’s response rate is at least 0.10 more than that of the existing treatment.
The ratio, \( \phi = \frac{p_1}{p_2} \), gives the relative change in the probability of the response. Testing non-inferiority and superiority use the formulation

\[
H_0: \frac{p_1}{p_2} \leq \phi_0 \quad \text{versus} \quad H_1: \frac{p_1}{p_2} > \phi_0
\]

The only subtlety is that for non-inferiority tests \( \phi_0 < 1 \), while for superiority tests \( \phi_0 > 1 \).

Non-Inferiority using a Ratio

The following example might help you understand the concept of non-inferiority as defined by the ratio. Suppose that 60% of patients \( p_2 = 0.60 \) respond to the current treatment method. If a new treatment decreases the response rate by no more than 10% \( \phi_0 = 0.90 \), it will be considered to be noninferior to the standard treatment. Substituting these figures into the statistical hypotheses gives

\[
H_0: \phi \leq 0.90 \quad \text{versus} \quad H_1: \phi > 0.90
\]

Using the relationship

\[
p_1 = \phi_0 p_2
\]

gives

\[
H_0: p_1 \leq 0.54 \quad \text{versus} \quad H_1: p_1 > 0.54
\]

In this example, when the null hypothesis is rejected, the concluded alternative is that the response rate is at least 54%. That is, the conclusion of non-inferiority is that the new treatment’s response rate is no worse than 10% less than that of the standard treatment.

Odds Ratio

The odds ratio, \( \psi = \left( \frac{p_1}{1 - p_1} \right) / \left( \frac{p_2}{1 - p_2} \right) \), gives the relative change in the odds of the response. Testing non-inferiority and superiority use the same formulation

\[
H_0: \psi \leq \psi_0 \quad \text{versus} \quad H_1: \psi > \psi_0
\]

The only difference is that for non-inferiority tests \( \psi_0 < 1 \), while for superiority tests \( \psi_0 > 1 \).

A Note on Setting the Significance Level, Alpha

Setting the significance level has always been somewhat arbitrary. For planning purposes, the standard has become to set alpha to 0.05 for two-sided tests. Almost universally, when someone states that a result is statistically significant, they mean statistically significant at the 0.05 level.

Although 0.05 may be the standard for two-sided tests, it is not always the standard for one-sided tests, such as non-inferiority tests. Statisticians often recommend that the alpha level for one-sided tests be set at 0.025 since this is the amount put in each tail of a two-sided test.
Power Calculation

The power for a test statistic that is based on the normal approximation can be computed exactly using two binomial distributions. The following steps are taken to compute the power of these tests.

1. Find the critical value using the standard normal distribution. The critical value, \( z_{\text{critical}} \), is that value of \( z \) that leaves exactly the target value of alpha in the appropriate tail of the normal distribution.

2. Compute the value of the test statistic, \( z_t \), for every combination of \( x_{11} \) and \( x_{21} \). Note that \( x_{11} \) ranges from 0 to \( n_1 \), and \( x_{21} \) ranges from 0 to \( n_2 \). A small value (around 0.0001) can be added to the zero-cell counts to avoid numerical problems that occur when the cell value is zero.

3. If \( z_t > z_{\text{critical}} \), the combination is in the rejection region. Call all combinations of \( x_{11} \) and \( x_{21} \) that lead to a rejection the set \( A \).

4. Compute the power for given values of \( p_{11} \) and \( p_2 \) as

\[
1 - \beta = \sum_{A} \binom{n_1}{x_{11}} p_{11}^{x_{11}} q_{11}^{n_1-x_{11}} \binom{n_2}{x_{21}} p_2^{x_{21}} q_2^{n_2-x_{21}}
\]

5. Compute the actual value of alpha achieved by the design by substituting \( p_2 \) for \( p_{11} \) to obtain

\[
\alpha^* = \sum_{A} \binom{n_1}{x_{11}} \binom{n_2}{x_{21}} p_2^{x_{21}+x_{21}} q_2^{n_1+n_2-x_{11}-x_{21}}
\]

Asymptotic Approximations

When the values of \( n_1 \) and \( n_2 \) are large (say over 200), these formulas often take a long time to evaluate. In this case, a large sample approximation can be used. The large sample approximation is made by replacing the values of \( \hat{p}_1 \) and \( \hat{p}_2 \) in the \( z \) statistic with the corresponding values of \( p_{11} \) and \( p_2 \), and then computing the results based on the normal distribution. Note that in large samples, the Farrington and Manning statistic is substituted for the Gart and Nam statistic.

Test Statistics

Several test statistics have been proposed for testing whether the difference, ratio, or odds ratio are different from a specified value. The main difference among the several test statistics is in the formula used to compute the standard error used in the denominator. These tests are based on the following \( z \)-test

\[
z_t = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0 - c}{\hat{\sigma}}
\]

The constant, \( c \), represents a continuity correction that is applied in some cases. When the continuity correction is not used, \( c \) is zero. In power calculations, the values of \( \hat{p}_1 \) and \( \hat{p}_2 \) are not known. The corresponding values of \( p_{11} \) and \( p_2 \) may be reasonable substitutes.

Following is a list of the test statistics available in PASS. The availability of several test statistics begs the question of which test statistic one should use. The answer is simple: one should use the test statistic that will be used to analyze the data. You may choose a method because it is a standard in your industry, because it seems to have better statistical properties, or because your statistical package calculates it. Whatever your reasons for selecting a certain test statistic, you should use the same test statistic when doing the analysis after the data have been collected.
Non-Inferiority Tests for Two Proportions

Z Test (Pooled)
This test was first proposed by Karl Pearson in 1900. Although this test is usually expressed directly as a chi-square statistic, it is expressed here as a z statistic so that it can be more easily used for one-sided hypothesis testing. The proportions are pooled (averaged) in computing the standard error. The formula for the test statistic is

\[ z = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\hat{\sigma}_1} \]

where

\[ \hat{\sigma}_1 = \sqrt{\bar{p}(1 - \bar{p}) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} \]

\[ \bar{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} \]

Z Test (Unpooled)
This test statistic does not pool the two proportions in computing the standard error.

\[ z = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\hat{\sigma}_2} \]

where

\[ \hat{\sigma}_2 = \sqrt{\hat{p}_1(1 - \hat{p}_1) + \hat{p}_2(1 - \hat{p}_2) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} \]

Z Test with Continuity Correction (Pooled)
This test is the same as Z Test (Pooled), except that a continuity correction is used. Remember that in the null case, the continuity correction makes the results closer to those of Fisher’s Exact test.

\[ z = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0 + \frac{F}{2} \left( \frac{1}{n_1} + \frac{1}{n_2} \right)}{\hat{\sigma}_1} \]

\[ \hat{\sigma}_1 = \sqrt{\bar{p}(1 - \bar{p}) \left( \frac{1}{n_1} + \frac{1}{n_2} \right)} \]

\[ \bar{p} = \frac{n_1 \hat{p}_1 + n_2 \hat{p}_2}{n_1 + n_2} \]

where \( F \) is -1 for lower-tailed hypotheses and 1 for upper-tailed hypotheses.
Non-Inferiority Tests for Two Proportions

Z Test with Continuity Correction (Unpooled)
This test is the same as the Z Test (Unpooled), except that a continuity correction is used. Remember that in the null case, the continuity correction makes the results closer to those of Fisher’s Exact test.

\[
Z = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\hat{\sigma}}
\]

where \( \hat{\sigma} = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \)

\[
\hat{p}_1 \text{ and } \hat{p}_2 \text{ are used in the numerator of the score statistic while MLE’s } \tilde{p}_1 \text{ and } \tilde{p}_2 \text{, constrained so that } \tilde{p}_1 - \tilde{p}_2 = \delta_0 \text{, are used in the denominator. A correction factor of } \frac{N}{N-1} \text{ is applied to make the variance estimate less biased. The significance level of the test statistic is based on the asymptotic normality of the score statistic. The formula for computing this test statistic is }
\]

\[
z_{MND} = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\hat{\sigma}_{MND}}
\]

where

\[
\hat{\sigma}_{MND} = \left( \frac{\tilde{p}_1 q_1}{n_1} + \frac{\tilde{p}_2 q_2}{n_2} \right) \left( \frac{N}{N-1} \right)
\]

\[
\tilde{p}_1 = \tilde{p}_2 + \delta_0
\]

\[
\tilde{p}_1 = 2B \cos(A) - \frac{L_2}{3L_3}
\]

\[
A = \frac{1}{3} \left[ \pi + \cos^{-1} \left( \frac{C}{B^3} \right) \right]
\]

\[
B = \text{sign}(C) \sqrt{\frac{L_2^2}{9L_3^2} - \frac{L_1}{3L_3}}
\]

\[
C = \frac{L_2}{27L_3^2} - \frac{L_1L_2}{6L_3^2} + \frac{L_0}{2L_3}
\]

\[
L_0 = x_{21} \delta_0 (1 - \delta_0)
\]
Miettinen and Nurminen’s Likelihood Score Test of the Ratio

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the ratio is equal to a specified value $\phi_0$. The regular MLE’s, $\hat{p}_1$ and $\hat{p}_2$, are used in the numerator of the score statistic while MLE’s $\tilde{p}_1$ and $\tilde{p}_2$, constrained so that $\tilde{p}_1 / \tilde{p}_2 = \phi_0$, are used in the denominator. A correction factor of $N/(N-1)$ is applied to make the variance estimate less biased. The significance level of the test statistic is based on the asymptotic normality of the score statistic.

The formula for computing the test statistic is

$$z_{MNR} = \frac{\hat{p}_1 / \hat{p}_2 - \phi_0}{\sqrt{\left(\frac{\tilde{p}_1 \tilde{q}_1}{n_1} + \phi_0^2 \frac{\tilde{p}_2 \tilde{q}_2}{n_2}\right) \left(\frac{N}{N-1}\right)}}$$

where

$$\tilde{p}_1 = \tilde{p}_2 \phi_0$$

$$\tilde{p}_2 = -\frac{B - \sqrt{B^2 - 4AC}}{2A}$$

$$A = N\phi_0$$

$$B = -\left[N_1\phi_0 + x_{11} + N_2 + x_{21}\phi_0\right]$$

$$C = M_1$$

Miettinen and Nurminen’s Likelihood Score Test of the Odds Ratio

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the odds ratio is equal to a specified value, $\psi_0$. Because the approach they used with the difference and ratio does not easily extend to the odds ratio, they used a score statistic approach for the odds ratio. The regular MLE’s are $\hat{p}_1$ and $\hat{p}_2$. The constrained MLE’s are $\tilde{p}_1$ and $\tilde{p}_2$. These estimates are constrained so that $\tilde{\psi} = \psi_0$. A correction factor of $N/(N-1)$ is applied to make the variance estimate less biased. The significance level of the test statistic is based on the asymptotic normality of the score statistic.

The formula for computing the test statistic is

$$z_{MNO} = \frac{\left(\hat{p}_1 - \tilde{p}_1\right) - \left(\hat{p}_2 - \tilde{p}_2\right)}{\sqrt{\left(\frac{1}{N_2 \hat{p}_1 \hat{q}_1} + \frac{1}{N_2 \hat{p}_2 \hat{q}_2}\right) \left(\frac{N}{N-1}\right)}}$$
where

\[ \tilde{p}_1 = \frac{\tilde{p}_2 \psi_0}{1 + \tilde{p}_2 (\psi_0 - 1)} \]

\[ \tilde{p}_2 = \frac{-B + \sqrt{B^2 - 4AC}}{2A} \]

\[ A = N_2 (\psi_0 - 1) \]

\[ B = N_1 \psi_0 + N_2 - M_1 (\psi_0 - 1) \]

\[ C = -M_1 \]

**Farrington and Manning’s Likelihood Score Test of the Difference**

Farrington and Manning (1990) proposed a test statistic for testing whether the difference is equal to a specified value \( \delta_0 \). The regular MLE’s, \( \hat{p}_1 \) and \( \hat{p}_2 \), are used in the numerator of the score statistic while MLE’s \( \tilde{p}_1 \) and \( \tilde{p}_2 \), constrained so that \( \tilde{p}_1 - \tilde{p}_2 = \delta_0 \), are used in the denominator. The significance level of the test statistic is based on the asymptotic normality of the score statistic.

The formula for computing the test statistic is

\[ z_{FMD} = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\sqrt{\frac{\hat{p}_1 \tilde{q}_1 + \hat{p}_2 \tilde{q}_2}{n_1} + \frac{\hat{p}_1 \tilde{q}_1}{n_1} \tilde{p}_2 \tilde{q}_2}} \]

where the estimates \( \tilde{p}_1 \) and \( \tilde{p}_2 \) are computed as in the corresponding test of Miettinen and Nurminen (1985) given above.

**Farrington and Manning’s Likelihood Score Test of the Ratio**

Farrington and Manning (1990) proposed a test statistic for testing whether the ratio is equal to a specified value \( \phi_0 \). The regular MLE’s, \( \hat{p}_1 \) and \( \hat{p}_2 \), are used in the numerator of the score statistic while MLE’s \( \tilde{p}_1 \) and \( \tilde{p}_2 \), constrained so that \( \tilde{p}_1 / \tilde{p}_2 = \phi_0 \), are used in the denominator. A correction factor of \( N/(N-1) \) is applied to increase the variance estimate. The significance level of the test statistic is based on the asymptotic normality of the score statistic.

The formula for computing the test statistic is

\[ z_{FMR} = \frac{\hat{p}_1 / \hat{p}_2 - \phi_0}{\sqrt{\frac{\tilde{p}_1 \tilde{q}_1 + \phi_0 \tilde{p}_2 \tilde{q}_2}{n_1} + \frac{\phi_0^2 \tilde{p}_2 \tilde{q}_2}{n_2}}} \]

where the estimates \( \tilde{p}_1 \) and \( \tilde{p}_2 \) are computed as in the corresponding test of Miettinen and Nurminen (1985) given above.
Farrington and Manning’s Likelihood Score Test of the Odds Ratio

Farrington and Manning (1990) indicate that the Miettinen and Nurminen statistic may be modified by removing the factor \( N/(N-1) \).

The formula for computing this test statistic is

\[
Z_{FMO} = \frac{\left( \hat{p}_1 - \hat{p}_1 \right) - \left( \hat{p}_2 - \hat{p}_2 \right)}{\hat{p}_1 \hat{q}_1 - \hat{p}_2 \hat{q}_2}
\]

\[
\sqrt{\left( \frac{1}{N_1 \hat{p}_1 \hat{q}_1} + \frac{1}{N_2 \hat{p}_2 \hat{q}_2} \right)}
\]

where the estimates \( \hat{p}_1 \) and \( \hat{p}_2 \) are computed as in the corresponding test of Miettinen and Nurminen (1985) given above.

Gart and Nam’s Likelihood Score Test of the Difference

Gart and Nam (1990), page 638, proposed a modification to the Farrington and Manning (1988) difference test that corrects for skewness. Let \( z_{FMD}(\delta) \) stand for the Farrington and Manning difference test statistic described above. The skewness corrected test statistic, \( z_{GND} \), is the appropriate solution to the quadratic equation

\[
\left( -\tilde{\gamma} \right) z_{GND}^2 + \left( -1 \right) z_{GND} + \left( z_{FMD}(\delta) + \tilde{\gamma} \right) = 0
\]

where

\[
\tilde{\gamma} = \frac{\tilde{\gamma}^{3/2}(\delta)}{6} \left( \frac{-\hat{p}_1 \hat{q}_1 (\hat{q}_1 - \hat{p}_1)}{n_1^2} - \frac{-\hat{p}_2 \hat{q}_2 (\hat{q}_2 - \hat{p}_2)}{n_2^2} \right)
\]

Gart and Nam’s Likelihood Score Test of the Ratio

Gart and Nam (1988), page 329, proposed a modification to the Farrington and Manning (1988) ratio test that corrects for skewness. Let \( z_{FMR}(\phi) \) stand for the Farrington and Manning ratio test statistic described above. The skewness corrected test statistic, \( z_{GNR} \), is the appropriate solution to the quadratic equation

\[
\left( -\tilde{\phi} \right) z_{GNR}^2 + \left( -1 \right) z_{GNR} + \left( z_{FMR}(\phi) + \tilde{\phi} \right) = 0
\]

where

\[
\tilde{\phi} = \frac{1}{6\tilde{u}^{3/2}} \left( \frac{-\hat{q}_1 (\hat{q}_1 - \hat{p}_1)}{n_1^2 \hat{p}_1^2} - \frac{-\hat{q}_2 (\hat{q}_2 - \hat{p}_2)}{n_2^2 \hat{p}_2^2} \right)
\]

\[
\tilde{u} = \frac{\hat{q}_1}{n_1 \hat{p}_1} + \frac{\hat{q}_2}{n_2 \hat{p}_2}
\]
Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab (Common Options)

The Design tab contains the parameters associated with this test such as the proportions, sample sizes, alpha, and power. This chapter covers four procedures, each of which has different options. This section documents options that are common to all four procedures. Later, unique options for each procedure will be documented.

Solve For

This option specifies the parameter to be solved for using the other parameters. The parameters that may be selected are $P_1$, $D_1$, $R_1$, or $OR_1$, $Alpha$, $Power$, $Sample Size (N_1)$, and $Sample Size (N_2)$. Under most situations, you will select either $Power$ or $Sample Size (N_1)$.

Select $Sample Size (N_1)$ when you want to calculate the sample size needed to achieve a given power and alpha level.

Select $Power$ when you want to calculate the power of an experiment.

Test

Higher Proportions Are

This option specifies whether proportions represent successes (better) or failures (worse).

- Better (Successes)
  When proportions represent successes, higher proportions are better. A non-inferior treatment is one whose proportion is at least almost as high as that of the reference group.

  For testing non-inferiority, $D_0$ is negative, $R_0$ is less than 1, and $OR_0$ is less than 1. For testing superiority, $D_0$ is positive, $R_0$ is greater than 1, and $OR_0$ is greater than 1.

- Worse (Failures)
  When proportions represent failures, lower proportions are better. A non-inferior treatment is one whose proportion is at most almost as low as that of the reference group.

  For testing non-inferiority, $D_0$ is positive, $R_0$ is greater than 1, and $OR_0$ is greater than 1. For testing superiority, $D_0$ is negative, $R_0$ is less than 1, and $OR_0$ is less than 1.

Test Type

Specify which test statistic is used in searching and reporting. Although the pooled $z$-test is commonly shown in elementary statistics books, the likelihood score test is arguably the best choice.

Note that C.C. is an abbreviation for Continuity Correction. This refers to the adding or subtracting $1/(2n)$ to (or from) the numerator of the $z$-value to bring the normal approximation closer to the binomial distribution.
Power and Alpha

Power
This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as 0.8 to 0.95 by 0.05 may be entered.

Alpha
This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as 0.01 0.05 0.10 or 0.01 to 0.10 by 0.01.

Sample Size (When Solving for Sample Size)

Group Allocation
Select the option that describes the constraints on \(N1\) or \(N2\) or both.

The options are

- **Equal (\(N1 = N2\))**
  
  This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.

- **Enter \(N1\), solve for \(N2\)**
  
  Select this option when you wish to fix \(N1\) at some value (or values), and then solve only for \(N2\). Please note that for some values of \(N1\), there may not be a value of \(N2\) that is large enough to obtain the desired power.

- **Enter \(N2\), solve for \(N1\)**
  
  Select this option when you wish to fix \(N2\) at some value (or values), and then solve only for \(N1\). Please note that for some values of \(N2\), there may not be a value of \(N1\) that is large enough to obtain the desired power.

- **Enter \(R = N2/N1\), solve for \(N1\) and \(N2\)**
  
  For this choice, you set a value for the ratio of \(N2\) to \(N1\), and then PASS determines the needed \(N1\) and \(N2\), with this ratio, to obtain the desired power. An equivalent representation of the ratio, \(R\), is
  \[
  N2 = R \times N1.
  \]

- **Enter percentage in Group 1, solve for \(N1\) and \(N2\)**
  
  For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed \(N1\) and \(N2\) with this percentage to obtain the desired power.
Non-Inferiority Tests for Two Proportions

**N1 (Sample Size, Group 1)**

*This option is displayed if Group Allocation = “Enter N1, solve for N2”*

N1 is the number of items or individuals sampled from the Group 1 population. N1 must be ≥ 2. You can enter a single value or a series of values.

**N2 (Sample Size, Group 2)**

*This option is displayed if Group Allocation = “Enter N2, solve for N1”*

N2 is the number of items or individuals sampled from the Group 2 population. N2 must be ≥ 2. You can enter a single value or a series of values.

**R (Group Sample Size Ratio)**

*This option is displayed only if Group Allocation = “Enter R = N2/N1, solve for N1 and N2.”*

R is the ratio of N2 to N1. That is, 

\[ R = \frac{N2}{N1}. \]

Use this value to fix the ratio of N2 to N1 while solving for N1 and N2. Only sample size combinations with this ratio are considered.

N2 is related to N1 by the formula:

\[ N2 = \lceil R \times N1 \rceil, \]

where the value \( \lceil Y \rceil \) is the next integer ≥ Y.

For example, setting \( R = 2.0 \) results in a Group 2 sample size that is double the sample size in Group 1 (e.g., \( N1 = 10 \) and \( N2 = 20 \), or \( N1 = 50 \) and \( N2 = 100 \)).

R must be greater than 0. If \( R < 1 \), then N2 will be less than N1; if \( R > 1 \), then N2 will be greater than N1. You can enter a single or a series of values.

**Percent in Group 1**

*This option is displayed only if Group Allocation = “Enter percentage in Group 1, solve for N1 and N2.”*

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for N1 and N2. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

---

**Sample Size (When Not Solving for Sample Size)**

**Group Allocation**

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal (N1 = N2)**
  
  This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter N1 and N2 individually**
  
  This choice permits you to enter different values for N1 and N2.
Enter N1 and R, where N2 = R * N1
Choose this option to specify a value (or values) for N1, and obtain N2 as a ratio (multiple) of N1.

Enter total sample size and percentage in Group 1
Choose this option to specify a value (or values) for the total sample size (N), obtain N1 as a percentage of N, and then N2 as N - N1.

Sample Size Per Group
This option is displayed only if Group Allocation = “Equal (N1 = N2).”
The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations. Since the sample sizes are the same in each group, this value is the value for N1, and also the value for N2.
The Sample Size Per Group must be ≥ 2. You can enter a single value or a series of values.

N1 (Sample Size, Group 1)
This option is displayed if Group Allocation = “Enter N1 and N2 individually” or “Enter N1 and R, where N2 = R * N1.”
N1 is the number of items or individuals sampled from the Group 1 population.
N1 must be ≥ 2. You can enter a single value or a series of values.

N2 (Sample Size, Group 2)
This option is displayed only if Group Allocation = “Enter N1 and N2 individually.”
N2 is the number of items or individuals sampled from the Group 2 population.
N2 must be ≥ 2. You can enter a single value or a series of values.

R (Group Sample Size Ratio)
This option is displayed only if Group Allocation = “Enter N1 and R, where N2 = R * N1.”
R is the ratio of N2 to N1. That is,

\[ R = \frac{N2}{N1} \]

Use this value to obtain N2 as a multiple (or proportion) of N1.
N2 is calculated from N1 using the formula:

\[ N2 = \lceil R \times N1 \rceil, \]

where the value \( \lceil Y \rceil \) is the next integer ≥ Y.
For example, setting R = 2.0 results in a Group 2 sample size that is double the sample size in Group 1.
R must be greater than 0. If R < 1, then N2 will be less than N1; if R > 1, then N2 will be greater than N1. You can enter a single value or a series of values.

Total Sample Size (N)
This option is displayed only if Group Allocation = “Enter total sample size and percentage in Group 1.”
This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines N1 and N2.
The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.
You can enter a single value or a series of values.
Non-Inferiority Tests for Two Proportions

Percent in Group 1

This option is displayed only if Group Allocation = “Enter total sample size and percentage in Group 1.”

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size – Reference (Group 2)

P2 (Reference Group Proportion)

Specify the value of \( p_2 \), the reference, baseline, or control group’s proportion. The null hypothesis is that the two proportions differ by no more than a specified amount. Since P2 is a proportion, these values must be between 0 and 1.

You may enter a range of values such as 0.1 0.2 0.3 or 0.1 to 0.9 by 0.1.

Design Tab (Proportions)

This section documents options that are used when the parameterization is in terms of the values of the two proportions, P1 and P2. P1.0 is the value of the P1 assumed by the null hypothesis and P1.1 is the value of P1 at which the power is calculated.

Effect Size – Treatment (Group 1)

P1.0 (Equivalence Proportion)

This option allows you to specify the value P1.0 directly. This is that value of treatment group’s proportion above which the treatment group is considered noninferior to the reference group.

When Higher Proportions Are is set to Better, the trivial proportion is the smallest value of P1 for which the treatment group is declared noninferior to the reference group. In this case, P1.0 should be less than P2 for non-inferiority tests and greater than P2 for superiority tests. The reverse is the case when Higher Proportions Are is set to Worse.

Proportions must be between 0 and 1. They cannot take on the values 0 or 1. This value should not be set to exactly the value of P2.

You may enter a range of values such as 0.03 0.05 0.10 or 0.01 to 0.05 by 0.01.

P1.1 (Actual Proportion)

This option specifies the value of P1.1 which is the value of the treatment proportion at which the power is to be calculated. Proportions must be between 0 and 1. They cannot take on the values 0 or 1.

You may enter a range of values such as 0.03 0.05 0.10 or 0.01 to 0.05 by 0.01.
Design Tab (Differences)

This section documents options that are used when the parameterization is in terms of the difference, $P_1 - P_2$. $P_{1.0}$ is the value of $P_1$ assumed by the null hypothesis and $P_{1.1}$ is the value of $P_1$ at which the power is calculated. Once $P_2$, $D_0$, and $D_1$ are given, the values of $P_{1.1}$ and $P_{1.0}$ can be calculated.

Effect Size – Differences

D0 (Equivalence Difference)
This option specifies the trivial difference (often called the margin of error) between $P_{1.0}$ (the value of $P_1$ under H0) and $P_2$. This difference is used with $P_2$ to calculate the value of $P_{1.0}$ using the formula: $P_{1.0} = P_2 + D_0$.

When Higher Proportions Are is set to Better, the trivial difference is that amount by which $P_1$ can be less than $P_2$ and still have the treatment group declared noninferior to the reference group. In this case, $D_0$ should be negative for non-inferiority tests and positive for superiority tests.

The reverse is the case when Higher Proportions Are is set to worse.

You may enter a range of values such as -.03 -.05 -.10 or -.05 to -.01 by .01. Differences must be between -1 and 1. $D_0$ cannot take on the values -1, 0, or 1.

D1 (Actual Difference)
This option specifies the actual difference between $P_{1.1}$ (the actual value of $P_1$) and $P_2$. This is the value of the difference at which the power is calculated. In non-inferiority trials, this difference is often set to 0.

The power calculations assume that $P_{1.1}$ is the actual value of the proportion in group 1 (experimental or treatment group). This difference is used with $P_2$ to calculate the value of $P_1$ using the formula: $P_{1.1} = D_1 + P_2$.

You may enter a range of values such as -.05 0 .5 or -.05 to .05 by .02. Actual differences must be between -1 and 1. They cannot take on the values -1 or 1.

Design Tab (Ratios)

This section documents options that are used when the parameterization is in terms of the ratio, $P_1 / P_2$. $P_{1.0}$ is the value of $P_1$ assumed by the null hypothesis and $P_{1.1}$ is the value of $P_1$ at which the power is calculated. Once $P_2$, $R_0$, and $R_1$ are given, the values of $P_{1.0}$ and $P_{1.1}$ can be calculated.

Effect Size – Ratios

R0 (Equivalence Ratio)
This option specifies the trivial ratio (also called the Relative Margin of Equivalence) between $P_{1.0}$ and $P_2$. The power calculations assume that $P_{1.0}$ is the value of $P_1$ under the null hypothesis. This value is used with $P_2$ to calculate the value of $P_{1.0}$ using the formula: $P_{1.0} = R_0 \times P_2$.

When Higher Proportions Are is set to Better, the trivial ratio is the relative amount by which $P_1$ can be less than $P_2$ and still have the treatment group declared noninferior to the reference group. In this case, $R_0$ should be less than one for non-inferiority tests and greater than 1 for superiority tests. The reverse is the case when Higher Proportions Are is set to Worse.

Ratios must be positive. $R_0$ cannot take on the value of 1. You may enter a range of values such as 0.95 .97 .99 or .91 to .99 by .02.
**R1 (Actual Ratio)**

This option specifies the ratio of P1.1 and P2, where P1.1 is the actual proportion in the treatment group. The power calculations assume that P1.1 is the actual value of the proportion in group 1. This difference is used with P2 to calculate the value of P1 using the formula: P1.1 = R1 x P2. In non-inferiority trials, this ratio is often set to 1.

Ratios must be positive. You may enter a range of values such as 0.95 1 1.05 or 0.9 to 1.9 by 0.02.

---

**Design Tab (Odds Ratios)**

This section documents options that are used when the parameterization is in terms of the odds ratios, O1.1 / O2 and O1.0 / O2. Note that the odds are defined as O2 = P2 / (1 – P2), O1.0 = P1.0 / (1 – P1.0), etc. P1.0 is the value of P1 assumed by the null hypothesis and P1.1 is the value of P1 at which the power is calculated. Once P2, OR0, and OR1 are given, the values of P1.1 and P1.0 can be calculated.

---

**Effect Size – Odds Ratios**

**OR0 (Equivalence Odds Ratio)**

This option specifies the trivial odds ratio between P1.0 and P2. The power calculations assume that P1.0 is the value of the P1 under the null hypothesis. OR0 is used with P2 to calculate the value of P1.0.

When *Higher Proportions Are* is set to *Better*, the trivial odds ratio implicitly gives the amount by which P1 can be less than P2 and still have the treatment group declared noninferior to the reference group. In this case, OR0 should be less than 1 for non-inferiority tests and greater than 1 for superiority tests. The reverse is the case when *Higher Proportions Are* is set to *Worse*.

Odds ratios must be positive. OR0 cannot take on the value of 1.

You may enter a range of values such as 0.95 0.97 0.99 or 0.91 to 0.99 by 0.02.

**OR1 (Actual Odds Ratio)**

This option specifies the odds ratio of P1.1 and P2, where P1.1 is the actual proportion in the treatment group. The power calculations assume that P1.1 is the actual value of the proportion in group 1. This value is used with P2 to calculate the value of P1. In non-inferiority trials, this odds ratio is often set to 1.

Odds ratios must be positive. You may enter a range of values such as 0.95 1 1.05 or 0.9 to 1 by 0.02.

---

**Options Tab**

The Options tab contains various limits and options.

---

**Zero Counts**

**Zero Count Adjustment Method**

Zero cell counts cause many calculation problems. To compensate for this, a small value (called the Zero Count Adjustment Value) can 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. We recommend that you use the option ‘Add to zero cells only.’

Zero cell values often do not occur in practice. However, since power calculations are based on total enumeration, they will occur in power and sample size estimation.

Adding a small value is controversial, but can be necessary for computational considerations. Statisticians have recommended adding various fractions to zero counts. We have found that adding 0.0001 seems to work well.
Zero Count Adjustment Value

Zero cell counts cause many calculation problems when computing power or sample size. To compensate for this, a small value may be added either to all cells or to all zero cells. This is the amount that is added. We have found that 0.0001 works well.

Be warned that the value of the ratio and the odds ratio will be affected by the amount specified here!

---

Exact Test Options

Maximum N1 or N2 for Exact Calculations

When either N1 or N2 is above this amount, power calculations are based on the normal approximation to the binomial. In this case, the actual value of alpha is not calculated. Currently, for three-gigahertz computers, a value near 200 is reasonable. As computers get faster, this number may be increased.
Example 1 – Finding Power

A study is being designed to establish the non-inferiority of a new treatment compared to the current treatment. Historically, the current treatment has enjoyed a 60% cure rate. The new treatment reduces the seriousness of certain side effects that occur with the current treatment. Thus, the new treatment will be adopted even if it is slightly less effective than the current treatment. The researchers will recommend adoption of the new treatment if it has a cure rate of at least 55%.

The researchers plan to use the Farrington and Manning likelihood score test statistic to analyze the data that will be (or has been) obtained. They want to study the power of the Farrington and Manning test at group sample sizes ranging from 50 to 500 for detecting a difference of -0.05 when the actual cure rate of the new treatment ranges from 57% to 70%. The significance level will be 0.025.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Non-Inferiority Tests for Two Proportions using Differences procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Non-Inferiority, and then clicking on Non-Inferiority Tests for Two Proportions using Differences. You may then make the appropriate entries as listed below, or open Example 1 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solve For</td>
<td>Power</td>
</tr>
<tr>
<td>Higher Proportions Are</td>
<td>Better</td>
</tr>
<tr>
<td>Test Type</td>
<td>Likelihood Score (Farr. &amp; Mann.)</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.025</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Sample Size Per Group</td>
<td>50 to 500 by 50</td>
</tr>
<tr>
<td>D0 (Non-Inferiority Difference)</td>
<td>-0.05</td>
</tr>
<tr>
<td>D1 (Actual Difference)</td>
<td>-0.03 0.00 0.05 0.10</td>
</tr>
<tr>
<td>P2 (Reference Group Proportion)</td>
<td>0.6</td>
</tr>
<tr>
<td>Maximum N1 or N2 for Exact Calc.</td>
<td>300</td>
</tr>
</tbody>
</table>
Non-Inferiority Tests for Two Proportions

Annotated Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

<table>
<thead>
<tr>
<th>Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P2</th>
<th>P1</th>
<th>P0</th>
<th>P1</th>
<th>P1</th>
<th>Ref. D0</th>
<th>N1 Diff</th>
<th>Diff</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.03798</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0236</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.04942</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0267</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05252</td>
<td>150</td>
<td>150</td>
<td>300</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0241</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.05980</td>
<td>200</td>
<td>200</td>
<td>400</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0244</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.06503</td>
<td>250</td>
<td>250</td>
<td>500</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0241</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.07349</td>
<td>300</td>
<td>300</td>
<td>600</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0261</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.07762</td>
<td>350</td>
<td>350</td>
<td>700</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08315</td>
<td>400</td>
<td>400</td>
<td>800</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.08862</td>
<td>450</td>
<td>450</td>
<td>900</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>-0.0300</td>
<td>0.0250</td>
<td>0.0261</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: exact results based on the binomial were only calculated when both N1 and N2 were less than 300.

Report Definitions

Power is the probability of rejecting a false null hypothesis.
N1 and N2 are the number of items sampled from each population.
N is the total sample size, N1 + N2.
P2 is the proportion for Group 2. This is the standard, reference, or control group.
P1 is the treatment or experimental group proportion. P1.0 is the smallest treatment-group response rate that still yields a non-inferiority conclusion. P1.1 is the proportion for Group 1 at which power and sample size calculations are made.
D0 is the non-inferiority margin. It is the difference P1 - P2, assuming H0. D1 is the difference P1 - P2 assumed for power and sample size calculations.
Target Alpha is the input probability of rejecting a true null hypothesis. Actual Alpha is the value of alpha that is actually achieved.

Summary Statements

Sample sizes of 50 in group one and 50 in group two achieve 4% power to detect a non-inferiority margin difference between the group proportions of -0.0500. The reference group proportion is 0.6000. The treatment group proportion is assumed to be 0.5500 under the null hypothesis of inferiority. The power was computed at for the case when the actual treatment group proportion is 0.5700. The test statistic used is the one-sided Score test (Farrington & Manning). The significance level of the test was targeted at 0.0250. The significance level actually achieved by this design is 0.0236.

This report shows the values of each of the parameters, one scenario per row. Note that the actual alpha value is blank for sample sizes greater than 300, which was the limit set for exact computation.

Most of the report columns have obvious interpretations. Those that may not be obvious are presented here. Note that the discussion below assumes that higher response rates are better and that non-inferiority testing (rather than superiority testing) is planned.

Prop Grp 2 P2
This is the value of P2, the response rate in the control group.

Equiv. Grp 1 Prop P1.0
This is the value of P1.0, the response rate of the treatment group, as specified by the null hypothesis of inferiority. Values of P1 less than this amount are considered different from P2. Values of P1 greater than this are considered noninferior to the reference group. The difference between this value and P2 is the value of the null hypothesis.
Actual Grp 1 Prop P1.1
This is the value of P1.1, the response rate of the treatment group, at which the power is computed. This is the value of P1 under the alternative hypothesis. The difference between this value and P2 is the value of the alternative hypothesis.

Equiv. Margin Diff D0
This is the value of D0, the difference between the two group proportions under the null hypothesis. This value is often called the margin of non-inferiority.

Actual Margin Diff D1
This is the value of D1, the difference between the two group proportions at which the power is computed. This is the value of the difference under the alternative hypothesis.

Target Alpha
This is the value of alpha that was targeted by the design. Note that the target alpha is not usually achieved exactly. For one-sided tests, this value should usually be 0.025.

Actual Alpha
This is the value of alpha that was actually achieved by this design. Note that since the limit on exact calculations was set to 300, and since this value is calculated exactly, it is not shown for values of N1 greater than 300.

The difference between the Target Alpha and the Actual Alpha is caused by the discrete nature of the binomial distribution and the use of the normal approximation to the binomial in determining the critical value of the test statistic.

Plots Section
The values from the table are displayed in the above chart. These charts give us a quick look at the sample size that will be required for various values of $D_1$. 
Example 2 – Finding the Sample Size

Continuing with the scenario given in Example 1, the researchers want to determine the sample size necessary for each value of D1 to achieve a power of 0.80. To cut down on the runtime, they decide to look at approximate values whenever N1 is greater than 100.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Non-Inferiority Tests for Two Proportions using Differences procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Non-Inferiority, and then clicking on Non-Inferiority Tests for Two Proportions using Differences. You may then make the appropriate entries as listed below, or open Example 2 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Sample Size</td>
</tr>
<tr>
<td>Higher Proportions Are</td>
<td>Better</td>
</tr>
<tr>
<td>Test Type</td>
<td>Likelihood Score (Farr. &amp; Mann.)</td>
</tr>
<tr>
<td>Power</td>
<td>0.8</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.025</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>D0 (Non-Inferiority Difference)</td>
<td>-0.05</td>
</tr>
<tr>
<td>D1 (Actual Difference)</td>
<td>-0.03 0.00 0.05 0.10</td>
</tr>
<tr>
<td>P2 (Reference Group Proportion)</td>
<td>0.6</td>
</tr>
</tbody>
</table>

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

<table>
<thead>
<tr>
<th>Numeric Results for Non-Inferiority Tests Based on the Difference: P1 - P2</th>
</tr>
</thead>
<tbody>
<tr>
<td>H0: P1 - P2 ≤ D0. H1: P1 - P2 = D1 &gt; D0. Test Statistic: Score test (Farrington &amp; Manning)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Target Power</th>
<th>Actual Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P2</th>
<th>P1</th>
<th>H0</th>
<th>P1</th>
<th>H1</th>
<th>NI Diff</th>
<th>Diff D1</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80</td>
<td>0.80002</td>
<td>9509</td>
<td>9509</td>
<td>19018</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.5700</td>
<td>-0.0500</td>
<td>0.0300</td>
<td>0.0250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.80</td>
<td>0.80008</td>
<td>1505</td>
<td>1505</td>
<td>3010</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.6000</td>
<td>-0.0500</td>
<td>0.0000</td>
<td>0.0250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.80</td>
<td>0.80075</td>
<td>368</td>
<td>368</td>
<td>736</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.6500</td>
<td>-0.0500</td>
<td>0.0500</td>
<td>0.0250</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.80</td>
<td>0.80187</td>
<td>159</td>
<td>159</td>
<td>318</td>
<td>0.6000</td>
<td>0.5500</td>
<td>0.7000</td>
<td>-0.0500</td>
<td>0.1000</td>
<td>0.0250</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The required sample size will depend a great deal on the value of D1. Any effort spent determining an accurate value for D1 will be worthwhile.
Example 3 – Comparing the Power of Several Test Statistics

Continuing with Example 1, the researchers want to determine which of the eight possible test statistics to adopt by using the comparative reports and charts that PASS produces. They decide to compare the powers and actual alphas for various sample sizes between 50 and 200 when D1 is 0.1.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Non-Inferiority Tests for Two Proportions using Differences procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Non-Inferiority, and then clicking on Non-Inferiority Tests for Two Proportions using Differences. You may then make the appropriate entries as listed below, or open Example 3 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solve For:</td>
<td>Power</td>
</tr>
<tr>
<td>Higher Proportions Are:</td>
<td>Better</td>
</tr>
<tr>
<td>Test Type:</td>
<td>Likelihood Score (Farr. &amp; Mann.)</td>
</tr>
<tr>
<td>Alpha:</td>
<td>0.025</td>
</tr>
<tr>
<td>Group Allocation:</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Sample Size Per Group:</td>
<td>50 100 150 200</td>
</tr>
<tr>
<td>D0 (Non-Inferiority Difference):</td>
<td>-0.05</td>
</tr>
<tr>
<td>D1 (Actual Difference):</td>
<td>0.10</td>
</tr>
<tr>
<td>P2 (Reference Group Proportion):</td>
<td>0.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Options Tab</th>
<th>Not checked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum N1 or N2 for Exact Calc.</td>
<td>300</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reports Tab</th>
<th>Checked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show Numeric Report:</td>
<td>Not checked</td>
</tr>
<tr>
<td>Show Comparative Reports:</td>
<td>Checked</td>
</tr>
<tr>
<td>Show Summary Statements:</td>
<td>Not checked</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Plots Tab</th>
<th>Not checked</th>
</tr>
</thead>
<tbody>
<tr>
<td>Show Plots:</td>
<td>Not checked</td>
</tr>
<tr>
<td>Show Comparative Plots:</td>
<td>Checked</td>
</tr>
</tbody>
</table>
Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results and Plots

Power Comparison of Non-Inferiority Tests Based on the Difference: P1 - P2
H0: P1 - P2 ≤ D0.  H1: P1 - P2 = D1 > D0.

<table>
<thead>
<tr>
<th>N1/N2</th>
<th>P2</th>
<th>P1</th>
<th>Target Alpha</th>
<th>Z(P) Test Power</th>
<th>Z(UnP) Test Power</th>
<th>Z(P) CC Test Power</th>
<th>Z(UnP) CC Test Power</th>
<th>T Test Power</th>
<th>F.M. Score Power</th>
<th>M.N. Score Power</th>
<th>G.N. Score Power</th>
</tr>
</thead>
<tbody>
<tr>
<td>50/50</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.3581</td>
<td>0.3670</td>
<td>0.2782</td>
<td>0.2945</td>
<td>0.3464</td>
<td>0.3581</td>
<td>0.3464</td>
<td>0.3581</td>
</tr>
<tr>
<td>100/100</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.6030</td>
<td>0.6088</td>
<td>0.5474</td>
<td>0.5475</td>
<td>0.5982</td>
<td>0.6030</td>
<td>0.6030</td>
<td>0.6030</td>
</tr>
<tr>
<td>150/150</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.7821</td>
<td>0.7837</td>
<td>0.7453</td>
<td>0.7474</td>
<td>0.7821</td>
<td>0.7837</td>
<td>0.7821</td>
<td>0.7821</td>
</tr>
<tr>
<td>200/200</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.8949</td>
<td>0.8857</td>
<td>0.8635</td>
<td>0.8649</td>
<td>0.8849</td>
<td>0.8849</td>
<td>0.8849</td>
<td>0.8849</td>
</tr>
</tbody>
</table>

Actual Alpha Comparison of Non-Inferiority Tests Based on the Difference: P1 - P2
H0: P1 - P2 ≤ D0.  H1: P1 - P2 = D1 > D0.

<table>
<thead>
<tr>
<th>N1/N2</th>
<th>P1</th>
<th>P2</th>
<th>Target Alpha</th>
<th>Z(P) Test Alpha</th>
<th>Z(UnP) Test Alpha</th>
<th>Z(P) CC Test Alpha</th>
<th>Z(UnP) CC Test Alpha</th>
<th>T Test Alpha</th>
<th>F.M. Score Alpha</th>
<th>M.N. Score Alpha</th>
<th>G.N. Score Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>50/50</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.0236</td>
<td>0.0253</td>
<td>0.0140</td>
<td>0.0161</td>
<td>0.0225</td>
<td>0.0236</td>
<td>0.0225</td>
<td>0.0236</td>
</tr>
<tr>
<td>100/100</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.0267</td>
<td>0.0267</td>
<td>0.0190</td>
<td>0.0190</td>
<td>0.0266</td>
<td>0.0267</td>
<td>0.0267</td>
<td>0.0267</td>
</tr>
<tr>
<td>150/150</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.0239</td>
<td>0.0241</td>
<td>0.0181</td>
<td>0.0183</td>
<td>0.0239</td>
<td>0.0241</td>
<td>0.0239</td>
<td>0.0239</td>
</tr>
<tr>
<td>200/200</td>
<td>0.6000</td>
<td>0.7000</td>
<td>0.0250</td>
<td>0.0243</td>
<td>0.0244</td>
<td>0.0191</td>
<td>0.0191</td>
<td>0.0243</td>
<td>0.0244</td>
<td>0.0243</td>
<td>0.0243</td>
</tr>
</tbody>
</table>

It is interesting to note that the powers of the continuity-corrected test statistics are consistently lower than the other tests. This occurs because the actual alpha achieved by these tests is lower than for the other tests. An interesting finding of this example is that the regular t-test performed about as well as the z-test.
Example 4 – Validation using Machin with Equal Sample Sizes

Machin et al. (1997), page 106, present a sample size study in which $P_2 = 0.5$, $D_0 = -0.2$, $D_1=0$, one-sided alpha = 0.1, and beta = 0.2. Using the Farrington and Manning test statistic, they found the sample size to be 55 in each group.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Non-Inferiority Tests for Two Proportions using Differences procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Non-Inferiority, and then clicking on Non-Inferiority Tests for Two Proportions using Differences. You may then make the appropriate entries as listed below, or open Example 4 by going to the File menu and choosing Open Example Template.

Option

<table>
<thead>
<tr>
<th>Design Tab</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Solve For</td>
<td>Sample Size (N1)</td>
</tr>
<tr>
<td>Higher Proportions Are</td>
<td>Better</td>
</tr>
<tr>
<td>Test Type</td>
<td>Likelihood Score (Farr. &amp; Mann.)</td>
</tr>
<tr>
<td>Power</td>
<td>0.8</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.1</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>$D_0$ (Non-Inferiority Difference)</td>
<td>-0.2</td>
</tr>
<tr>
<td>$D_1$ (Actual Difference)</td>
<td>0.0</td>
</tr>
<tr>
<td>$P_2$ (Reference Group Proportion)</td>
<td>0.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Options Tab</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum N1 or N2 for Exact Calc.</td>
<td>2 (Set low for a rapid search.)</td>
</tr>
</tbody>
</table>

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

| Numeric Results for Non-Inferiority Tests Based on the Difference: P1 - P2 |
|--------------------------|--------------------------|
| H0: P1 - P2 ≤ D0. H1: P1 - P2 = D1 > D0. Test Statistic: Score test (Farrington & Manning) |

<table>
<thead>
<tr>
<th>Target Power</th>
<th>Actual Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P</th>
<th>Ref. P1/H0</th>
<th>Ref. P1/H1</th>
<th>NI Diff</th>
<th>Diff</th>
<th>D0</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80</td>
<td>0.80009</td>
<td>55</td>
<td>55</td>
<td>110</td>
<td>0.5000</td>
<td>0.3000</td>
<td>0.5000</td>
<td>-0.2000</td>
<td>0.0000</td>
<td>0.1000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PASS found the required sample size to be 55 which corresponds to Machin.
Example 5 – Validation of a Superiority Test using Farrington and Manning

Farrington and Manning (1990), page 1451, present a sample size study for a superiority test in which \( P_2 = 0.05 \), \( D_0 = 0.2 \), \( D_1 = 0.35 \), one-sided alpha = 0.05, and beta = 0.20. Using the Farrington and Manning test statistic, they found the sample size to be 80 in each group. They mention that the true power is 0.813.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Non-Inferiority Tests for Two Proportions using Differences procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Non-Inferiority, and then clicking on Non-Inferiority Tests for Two Proportions using Differences. You may then make the appropriate entries as listed below, or open Example 5 by going to the File menu and choosing Open Example Template.

Option | Value
---|---
Design Tab | Solve For ................. Sample Size
| Higher Proportions Are ............. Better
| Test Type .................. Likelihood Score (Farr. & Mann.)
| Power ................... 0.80
| Alpha .................. 0.05
| Group Allocation .................. Equal (N1 = N2)
| \( D_0 \) (Non-Inferiority Difference) ............... 0.2
| \( D_1 \) (Actual Difference) .................. 0.35
| \( P_2 \) (Reference Group Proportion) ........... 0.05

Options Tab | Maximum N1 or N2 for Exact Calc. ....... 2 (Set low for a rapid search.)

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

<table>
<thead>
<tr>
<th>Target Power</th>
<th>Actual Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P2</th>
<th>P1</th>
<th>H0</th>
<th>P1</th>
<th>H1</th>
<th>P1 Diff</th>
<th>D0</th>
<th>Diff D1</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.80</td>
<td>0.80068</td>
<td>80</td>
<td>80</td>
<td>160</td>
<td>0.0500</td>
<td>0.2500</td>
<td>0.4000</td>
<td>0.2000</td>
<td>0.3500</td>
<td>0.0500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PASS also calculated the required sample size to be 80.
Next, to calculate the exact power for this sample size, we make the following changes to the template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design Tab</strong></td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Power</td>
</tr>
<tr>
<td>Sample Size Per Group</td>
<td>80</td>
</tr>
<tr>
<td><strong>Options Tab</strong></td>
<td></td>
</tr>
<tr>
<td>Maximum N1 or N2 for Exact Calc.</td>
<td>200 (Set &gt;80 to force exact calculation.)</td>
</tr>
</tbody>
</table>

**Numeric Results**

<table>
<thead>
<tr>
<th>Numeric Results for Non-Inferiority Tests Based on the Difference: P1 - P2</th>
</tr>
</thead>
<tbody>
<tr>
<td>H0: P1 - P2 ≤ D0.  H1: P1 - P2 = D1 &gt; D0.  Test Statistic: Score test (Farrington &amp; Manning)</td>
</tr>
<tr>
<td>Power</td>
</tr>
<tr>
<td>-------</td>
</tr>
<tr>
<td>0.81320</td>
</tr>
</tbody>
</table>

PASS also calculated the exact power to be 0.813.

**Example 6 – Validation of Risk Ratio Calculations using Blackwelder**

Blackwelder (1993), page 695, presents a table of power values for several scenarios using the risk ratio. The second line of the table presents the results for the following scenario: P2 = 0.04, R0 = 0.3, R1=0.1, N1=N2=1044, one-sided alpha = 0.05, and beta = 0.20. Using the Farrington and Manning likelihood-score test statistic, he found the exact power to be 0.812, the exact alpha to be 0.044, and, using the asymptotic formula, the approximate power to be 0.794.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Non-Inferiority Tests for Two Proportions using Ratios** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for Two Proportions using Ratios**. You may then make the appropriate entries as listed below, or open Example 6 by going to the File menu and choosing **Open Example Template**.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design Tab</strong></td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Power</td>
</tr>
<tr>
<td>Higher Proportions Are</td>
<td>Worse</td>
</tr>
<tr>
<td>Test Type</td>
<td>Likelihood Score (Farr. &amp; Mann.)</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Sample Size Per Group</td>
<td>1044</td>
</tr>
<tr>
<td>R0 (Non-Inferiority Ratio)</td>
<td>0.3</td>
</tr>
<tr>
<td>R1 (Actual Ratio)</td>
<td>0.1</td>
</tr>
<tr>
<td>P2 (Reference Group Proportion)</td>
<td>0.04</td>
</tr>
<tr>
<td><strong>Options Tab</strong></td>
<td></td>
</tr>
<tr>
<td>Maximum N1 or N2 for Exact Calc.</td>
<td>2000 (Set high for exact results.)</td>
</tr>
</tbody>
</table>
Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

<table>
<thead>
<tr>
<th>Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P2</th>
<th>P1</th>
<th>H0</th>
<th>P1</th>
<th>H1</th>
<th>NI Ratio</th>
<th>Ratio</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.81178</td>
<td>1044</td>
<td>1044</td>
<td>2088</td>
<td>0.0400</td>
<td>0.0120</td>
<td>0.0040</td>
<td>0.300</td>
<td>0.100</td>
<td>0.0500</td>
<td>0.0444</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PASS also calculated the power to be 0.812 and the actual alpha to be 0.044, within rounding.

Next, to calculate the asymptotic power, we make the following changes to the template.

Option

Options Tab
Maximum N1 or N2 for Exact Calc. ........ 2 (Set < 1044 to force asymptotic calculation.)

Numeric Results

<table>
<thead>
<tr>
<th>Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P2</th>
<th>P1</th>
<th>H0</th>
<th>P1</th>
<th>H1</th>
<th>NI Ratio</th>
<th>Ratio</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.79373</td>
<td>1044</td>
<td>1044</td>
<td>2088</td>
<td>0.0400</td>
<td>0.0120</td>
<td>0.0040</td>
<td>0.300</td>
<td>0.100</td>
<td>0.0500</td>
<td>0.0444</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PASS also calculated the power to be 0.794.
Example 7 – Finding Power following an Experiment

In an effort to show a new treatment non-inferior to the current standard, researchers randomly assigned 80 subjects to each treatment. The new treatment was to be considered non-inferior if the odds ratio (treatment to standard) was at least 0.80. Using the Farrington and Manning Likelihood Score test, non-inferiority could not be concluded. The researchers now want to see the power of the test. The control proportion was 0.625.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Non-Inferiority Tests for Two Proportions using Odds Ratios procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Non-Inferiority, and then clicking on Non-Inferiority Tests for Two Proportions using Odds Ratios. You may then make the appropriate entries as listed below, or open Example 7 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design Tab</strong></td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>Power</td>
</tr>
<tr>
<td>Higher Proportions Are</td>
<td>Better</td>
</tr>
<tr>
<td>Test Type</td>
<td>Likelihood Score (Farr. &amp; Mann.)</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Sample Size Per Group</td>
<td>80</td>
</tr>
<tr>
<td>OR0 (Non-Inferiority Odds Ratio)</td>
<td>0.80</td>
</tr>
<tr>
<td>OR1 (Actual Odds Ratio)</td>
<td>1.0</td>
</tr>
<tr>
<td>P2 (Reference Group Proportion)</td>
<td>0.625</td>
</tr>
</tbody>
</table>

Output

Click the Calculate button to perform the calculations and generate the following output.

**Numeric Results**

<table>
<thead>
<tr>
<th>Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P2</th>
<th>P1</th>
<th>H0</th>
<th>P1</th>
<th>0.0</th>
<th>P1</th>
<th>H1</th>
<th>NI</th>
<th>O.R.</th>
<th>OR0</th>
<th>O.R.</th>
<th>0.0</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.18453</td>
<td>80</td>
<td>80</td>
<td>160</td>
<td>0.6250</td>
<td>0.5714</td>
<td>0.6250</td>
<td>0.800</td>
<td>1.000</td>
<td>0.0500</td>
<td>0.0571</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The power of a test with 80 receiving each treatment is only 0.18453.
Example 8 – Finding True Proportion Difference

Researchers have developed a new treatment with minimal side effects compared to the standard treatment. The researchers are limited by the number of subjects (140 per group) they can use to show the new treatment is non-inferior. The new treatment will be deemed non-inferior if it is at least 0.10 below the success rate of the standard treatment. The standard treatment has a success rate of about 0.75. The researchers want to know how much more successful the new treatment must be (in truth) to yield a test which has 90% power. The test statistic used will be the pooled Z test.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Non-Inferiority Tests for Two Proportions using Differences procedure window by expanding Proportions, then Two Independent Proportions, then clicking on Non-Inferiority, and then clicking on Non-Inferiority Tests for Two Proportions using Differences. You may then make the appropriate entries as listed below, or open Example 8 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Design Tab</td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>D1</td>
</tr>
<tr>
<td>Higher Proportions Are</td>
<td>Better</td>
</tr>
<tr>
<td>Test Type</td>
<td>Z Test (Pooled)</td>
</tr>
<tr>
<td>Power</td>
<td>0.90</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>Group Allocation</td>
<td>Equal (N1 = N2)</td>
</tr>
<tr>
<td>Sample Size Per Group</td>
<td>140</td>
</tr>
<tr>
<td>D0 (Non-Inferiority Difference)</td>
<td>-0.10</td>
</tr>
<tr>
<td>P2 (Reference Group Proportion)</td>
<td>0.75</td>
</tr>
<tr>
<td>Options Tab</td>
<td></td>
</tr>
<tr>
<td>Maximum N1 or N2 for Exact Calc.</td>
<td>500 (Set high for exact results.)</td>
</tr>
</tbody>
</table>

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

<table>
<thead>
<tr>
<th>Power</th>
<th>N1</th>
<th>N2</th>
<th>N</th>
<th>Ref. P2</th>
<th>P1</th>
<th>H0</th>
<th>P1.0</th>
<th>P1</th>
<th>H1</th>
<th>P1.1</th>
<th>NI Diff</th>
<th>Diff D0</th>
<th>Diff D1</th>
<th>Target Alpha</th>
<th>Actual Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.90000</td>
<td>140</td>
<td>140</td>
<td>280</td>
<td>0.7500</td>
<td>0.6500</td>
<td>0.7961</td>
<td>-0.1000</td>
<td>0.0461</td>
<td>0.0500</td>
<td>0.0505</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

With 140 subjects in each group, the new treatment must have a success rate 0.0461 higher than the current treatment (or about 0.7961) to have 90% power in the test of non-inferiority.