

Chapter 211

Non-Inferiority Tests for the Ratio of Two Proportions

Introduction

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

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

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 ratio of the two treatments is no longer considered ignorable. After thoughtful discussion with several clinicians, it was decided that if the response rate ratio is no less than 0.9, the new treatment would be adopted. This ratio is called the *margin of non-inferiority*. The margin of non-inferiority in this example is 0.9.

The developers must design an experiment to test the hypothesis that the response rate ratio of the new treatment to the standard is at least 0.9. The statistical hypothesis to be tested is

$$H_0: p_1/p_2 \leq 0.9 \quad \text{versus} \quad H_1: p_1/p_2 > 0.9$$

Notice that when the null hypothesis is rejected, the conclusion is that the response rate ratio is at least 0.9. Note that even though the response rate of the current treatment is 0.70, the hypothesis test is about a response rate ratio of 0.9. 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 “Tests for Two Proportions,” and they will not be duplicated here. Instead, this chapter only discusses those changes necessary for non-inferiority tests.

This procedure has the capability for calculating power based on large sample (normal approximation) results and 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

Group	Success	Failure	Total
Treatment	x_{11}	x_{12}	n_1
Control	x_{21}	x_{22}	n_2
Totals	m_1	m_2	N

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 ϕ_0 represent the smallest ratio (margin of non-inferiority) 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, $\phi_0 < 1$. The set of statistical hypotheses that are tested is

$$H_0: p_1/p_2 \leq \phi_0 \quad \text{versus} \quad H_1: p_1/p_2 > \phi_0$$

which can be rearranged to give

$$H_0: p_1 \leq p_2 \phi_0 \quad \text{versus} \quad H_1: p_1 > p_2 \phi_0$$

There are three common methods of specifying the margin of non-inferiority. 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

<u>Parameter</u>	<u>Computation</u>	<u>Hypotheses</u>
Difference	$\delta_0 = p_{1,0} - p_2$	$H_0: p_1 - p_2 \leq \delta_0$ versus $H_1: p_1 - p_2 > \delta_0$
Ratio	$\phi_0 = p_{1,0}/p_2$	$H_0: p_1/p_2 \leq \phi_0$ versus $H_1: p_1/p_2 > \phi_0$
Odds Ratio	$\psi_0 = O_{1,0}/O_2$	$H_0: O_1/O_2 \leq \psi_0$ versus $H_1: O_1/O_2 > \psi_0$

Non-Inferiority Tests for the Ratio of Two Proportions

Ratio

The ratio, $\phi = p_1/p_2$, gives the relative change in the probability of the response. Testing non-inferiority uses the formulation

$$H_0: p_1/p_2 \leq \phi_0 \quad \text{versus} \quad H_1: p_1/p_2 > \phi_0$$

or equivalently

$$H_0: \phi \leq \phi_0 \quad \text{versus} \quad H_1: \phi > \phi_0.$$

For non-inferiority tests with higher proportions better, $\phi_0 < 1$. For non-inferiority tests with higher proportions worse, $\phi_0 > 1$.

Non-Inferiority

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

$$H_0: \phi \leq 0.90 \quad \text{versus} \quad H_1: \phi > 0.90.$$

In this example, when the null hypothesis is rejected 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.

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, $z_{critical}$, 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 $z_t > z_{critical}$, 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 $p_{1.1}$ and p_2 as

$$1 - \beta = \sum_A \binom{n_1}{x_{11}} p_{1.1}^{x_{11}} q_{1.1}^{n_1 - x_{11}} \binom{n_2}{x_{21}} p_2^{x_{21}} q_2^{n_2 - x_{21}}.$$

Non-Inferiority Tests for the Ratio of Two Proportions

5. Compute the actual value of alpha achieved by the design by substituting $p_{1.0}$ for $p_{1.1}$ to obtain

$$\alpha^* = \sum_A \binom{n_1}{x_{11}} p_{1.0}^{x_{11}} q_{1.0}^{n_1-x_{11}} \binom{n_2}{x_{21}} p_2^{x_{21}} q_2^{n_2-x_{21}}.$$

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 $p_{1.1}$ 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

Three test statistics have been proposed for testing whether the ratio is 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

$$z_t = \frac{\hat{p}_1 / \hat{p}_2 - \phi_0}{\hat{\sigma}}$$

In power calculations, the values of \hat{p}_1 and \hat{p}_2 are not known. The corresponding values of $p_{1.1}$ 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.

Miettinen and Nurminen's Likelihood Score Test

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the ratio is equal to a specified value ϕ_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

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

Non-Inferiority Tests for the Ratio of Two Proportions

where

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

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

$$A = N\phi_0$$

$$B = -[n_1\phi_0 + x_{11} + n_2 + x_{21}\phi_0]$$

$$C = m_1$$

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

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

$$A = n_2(\psi_0 - 1)$$

$$B = n_1\psi_0 + n_2 - m_1(\psi_0 - 1)$$

$$C = -m_1$$

Farrington and Manning's Likelihood Score Test

Farrington and Manning (1990) proposed a test statistic for testing whether the ratio is equal to a specified value ϕ_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

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

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

Gart and Nam's Likelihood Score Test

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

$$(-\tilde{\varphi})z_{GNR}^2 + (-1)z_{GNR} + (z_{FMR}(\phi) + \tilde{\varphi}) = 0$$

Non-Inferiority Tests for the Ratio of Two Proportions

where

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

$$\tilde{u} = \frac{\tilde{q}_1}{n_1 \tilde{p}_1} + \frac{\tilde{q}_2}{n_2 \tilde{p}_2}$$

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

The Design tab contains the parameters associated with this test such as the proportions, ratios, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for using the other parameters. The parameters that may be selected are *Power*, *Sample Size*, and *Effect Size*.

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

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

Power Calculation

Power Calculation Method

Select the method to be used to calculate power. When the sample sizes are reasonably large (i.e. greater than 50) and the proportions are between 0.2 and 0.8 the two methods will give similar results. For smaller sample sizes and more extreme proportions (less than 0.2 or greater than 0.8), the normal approximation is not as accurate so the binomial calculations may be more appropriate.

The choices are

- **Binomial Enumeration**

Power for each test is computed using binomial enumeration of all possible outcomes when N1 and N2 ≤ Maximum N1 or N2 for Binomial Enumeration (otherwise, the normal approximation is used). Binomial enumeration of all outcomes is possible because of the discrete nature of the data.

- **Normal Approximation**

Approximate power for each test is computed using the normal approximation to the binomial distribution.

Actual alpha values are only computed when “Binomial Enumeration” is selected.

Non-Inferiority Tests for the Ratio of Two Proportions

Power Calculation – Binomial Enumeration Options

Only shown when Power Calculation Method = “Binomial Enumeration”

Maximum N1 or N2 for Binomial Enumeration

When both N1 and N2 are less than or equal to this amount, power calculations using the binomial distribution are made. The value of the “Actual Alpha” is only calculated when binomial power calculations are made.

When either N1 or N2 is larger than this amount, the normal approximation to the binomial is used for power calculations.

Zero Count Adjustment Method

Zero cell counts cause many calculation problems when enumerating binomial probabilities. To compensate for this, a small value (called the “Zero Count Adjustment Value”) may be added either to all cells or to all cells with zero counts. This option specifies which type of adjustment you want to use.

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 seems to work well.

Zero Count Adjustment Value

Zero cell counts cause many calculation problems when enumerating binomial probabilities. 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 will be affected by the amount specified here!

Test

Higher Proportions Are

Use this option to specify the direction of the test.

If Higher Proportions are “Better”, the alternative hypothesis is $H_1: P_1/P_2 > R_0$.

If Higher Proportions are “Worse”, the alternative hypothesis is $H_1: P_1/P_2 < R_0$.

Test Type

Specify which test statistic is used in searching and reporting.

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.

Non-Inferiority Tests for the Ratio of Two Proportions

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

$$N2 = R * N1.$$
- **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.

$N1$ (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.

Non-Inferiority Tests for the Ratio of Two Proportions

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,

$$R = N2 / N1.$$

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:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

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 $N - N1$.

Non-Inferiority Tests for the Ratio of Two Proportions

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,

$$R = N2/N1$$

Use this value to obtain $N2$ as a multiple (or proportion) of $N1$.

$N2$ is calculated from $N1$ using the formula:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

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.

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.

Non-Inferiority Tests for the Ratio of Two Proportions

Effect Size – Ratios

R0 (Non-Inferiority Ratio)

This option specifies the trivial ratio (also called the Relative Margin of Non-inferiority) between P1 and P2. The power calculations assume that P1.0 is the value of the P1 under the null hypothesis. This value is used with P2 to calculate the value of P1.0 using the formula: $P1.0 = R0 \times P2$.

When *Higher Proportions Are* is set to *Better*, the trivial ratio is the relative amount by which P1 can be less than P2 and still have the treatment group declared non-inferior to the reference group. R0 should be less than one for non-inferiority tests. The reverse is the case when *Higher Proportions Are* is set to *Worse*.

Ratios must be positive. R0 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: $P1.1 = R1 \times P2$. 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*.

Effect Size – Group 2 (Reference)

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*.

Non-Inferiority Tests for the Ratio of Two Proportions

Example 1 – Finding Power

A study is being designed to establish the non-inferiority of a new treatment compared to the current treatment. Historically, under the current treatment only 6% experience side effects. The new treatment is much less expensive and has been proven to be as effective. Thus, the new treatment will be adopted even if the rate of side effects is slightly worse than the current treatment. The researchers will recommend adoption of the new treatment if the rate ratio for side effects is less than 2.0.

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 200 to 1000 when the actual ratio ranges from 1 to 1.5. 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 the Ratio of Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Proportions**. 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**.

Option	Value
Design Tab	
Solve For	Power
Power Calculation Method.....	Normal Approximation
Higher Proportions Are	Worse (H1: P1/P2 < R0)
Test Type	Likelihood Score (Farr. & Mann.)
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	200 to 1000 by 200
R0 (Non-Inferiority Ratio).....	2.0
R1 (Actual Ratio)	1 1.25 1.5
P2 (Group 2 Proportion)	0.06

Non-Inferiority Tests for the Ratio of Two Proportions

Annotated Output

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

Numeric Results

Numeric Results

Test Statistic: Farrington & Manning Likelihood Score Test

Hypotheses: $H_0: P_1/P_2 \geq R_0$ vs. $H_1: P_1/P_2 < R_0$

Power*	N1	N2	N	Ref. P2	P1 H0 P1.0	P1 H1 P1.1	NI Ratio R0	Ratio R1	Alpha
0.43819	200	200	400	0.0600	0.1200	0.0600	2.000	1.000	0.0250
0.69368	400	400	800	0.0600	0.1200	0.0600	2.000	1.000	0.0250
0.84475	600	600	1200	0.0600	0.1200	0.0600	2.000	1.000	0.0250
0.92539	800	800	1600	0.0600	0.1200	0.0600	2.000	1.000	0.0250
0.96558	1000	1000	2000	0.0600	0.1200	0.0600	2.000	1.000	0.0250
0.26051	200	200	400	0.0600	0.1200	0.0750	2.000	1.250	0.0250
0.43785	400	400	800	0.0600	0.1200	0.0750	2.000	1.250	0.0250
0.58551	600	600	1200	0.0600	0.1200	0.0750	2.000	1.250	0.0250
0.70194	800	800	1600	0.0600	0.1200	0.0750	2.000	1.250	0.0250
0.79005	1000	1000	2000	0.0600	0.1200	0.0750	2.000	1.250	0.0250

(report continues)

* Power was computed using the normal approximation method.

References

- Chow, S.C., Shao, J., and Wang, H. 2008. Sample Size Calculations in Clinical Research, Second Edition. Chapman & Hall/CRC. Boca Raton, Florida.
- Farrington, C. P. and Manning, G. 1990. 'Test Statistics and Sample Size Formulae for Comparative Binomial Trials with Null Hypothesis of Non-Zero Risk Difference or Non-Unity Relative Risk.' *Statistics in Medicine*, Vol. 9, pages 1447-1454.
- Fleiss, J. L., Levin, B., Paik, M.C. 2003. *Statistical Methods for Rates and Proportions*. Third Edition. John Wiley & Sons. New York.
- Gart, John J. and Nam, Jun-mo. 1988. 'Approximate Interval Estimation of the Ratio in Binomial Parameters: A Review and Corrections for Skewness.' *Biometrics*, Volume 44, Issue 2, 323-338.
- Gart, John J. and Nam, Jun-mo. 1990. 'Approximate Interval Estimation of the Difference in Binomial Parameters: Correction for Skewness and Extension to Multiple Tables.' *Biometrics*, Volume 46, Issue 3, 637-643.
- Julious, S. A. and Campbell, M. J. 2012. 'Tutorial in biostatistics: sample sizes for parallel group clinical trials with binary data.' *Statistics in Medicine*, 31:2904-2936.
- Lachin, John M. 2000. *Biostatistical Methods*. John Wiley & Sons. New York.
- Machin, D., Campbell, M., Fayers, P., and Pinol, A. 1997. *Sample Size Tables for Clinical Studies*, 2nd Edition. Blackwell Science. Malden, Mass.
- Miettinen, O.S. and Nurminen, M. 1985. 'Comparative analysis of two rates.' *Statistics in Medicine* 4: 213-226.

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, $N_1 + N_2$.

P2 is the proportion for Group 2, which is the standard, reference, or control group.

P1 is the proportion for Group 1, which is the treatment or experimental group. P1.0 is the largest Group 1 proportion that still yields a non-inferiority conclusion. P1.1 is the proportion for Group 1 under the alternative hypothesis at which power and sample size calculations are made.

R0 is the non-inferiority ratio, P_1 / P_2 , assuming H_0 . R1 is the ratio under the alternative hypothesis used for power and sample size calculations.

Alpha is the probability of rejecting a true null hypothesis.

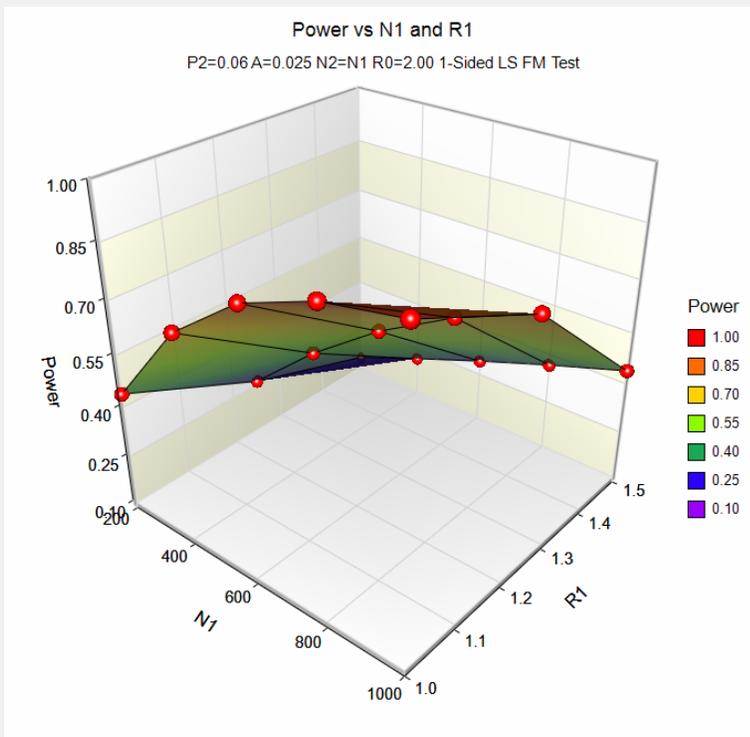
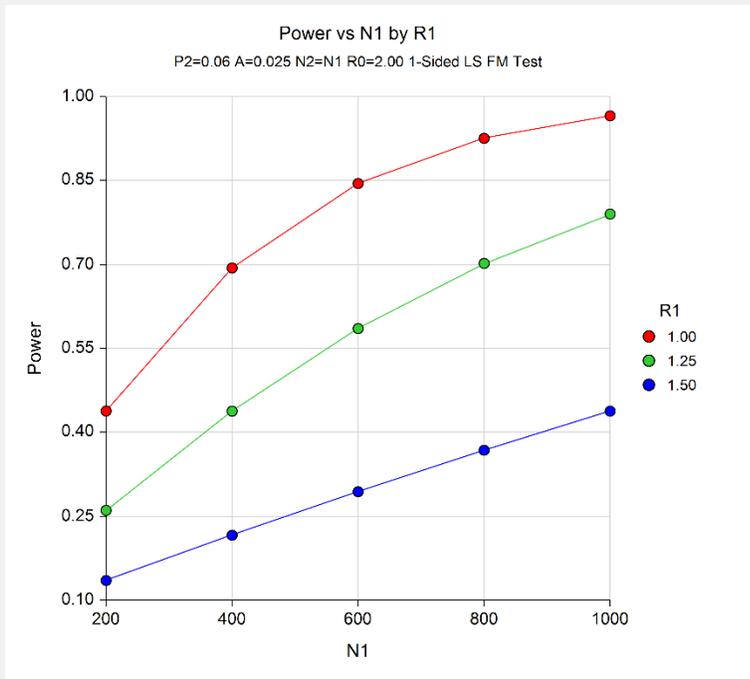
Summary Statements

Sample sizes of 200 in Group 1 and 200 in Group 2 achieve 43.819% power to detect a ratio of 1.000 when the non-inferiority ratio is 2.000. The reference group proportion is 0.0600. The treatment group proportion is assumed to be 0.1200 under the null hypothesis. The power was computed for the case when the actual treatment group proportion is 0.0600. The test statistic used is the one-sided Score test (Farrington & Manning). The significance level of the test is 0.0250.

This report shows the values of each of the parameters, one scenario per row.

Non-Inferiority Tests for the Ratio of Two Proportions

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 R1.

Non-Inferiority Tests for the Ratio of Two Proportions

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 R1 to achieve a power of 0.80.

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 the Ratio of Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Proportions**. 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Power Calculation Method.....	Normal Approximation
Higher Proportions Are	Worse (H1: P1/P2 < R0)
Test Type.....	Likelihood Score (Farr. & Mann.)
Power.....	0.80
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
R0 (Non-Inferiority Ratio).....	2.0
R1 (Actual Ratio)	1 1.25 1.5
P2 (Group 2 Proportion)	0.06

Output

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

Numeric Results

Numeric Results

Test Statistic: Farrington & Manning Likelihood Score Test
Hypotheses: H0: P1/P2 ≥ R0 vs. H1: P1/P2 < R0

Target Power	Actual Power*	N1	N2	N	Ref. P2	P1 H0 P1.0	P1 H1 P1.1	NI Ratio R0	Ratio R1	Alpha
0.80	0.80033	528	528	1056	0.0600	0.1200	0.0600	2.000	1.000	0.0250
0.80	0.80003	1027	1027	2054	0.0600	0.1200	0.0750	2.000	1.250	0.0250
0.80	0.80015	2508	2508	5016	0.0600	0.1200	0.0900	2.000	1.500	0.0250

* Power was computed using the normal approximation method.

The required sample size will depend a great deal on the value of R1. Any effort spent determining an accurate value for R1 will be worthwhile.

Example 3 – Comparing the Power of the Three Test Statistics

Continuing with Example 2, the researchers want to determine which of the three possible test statistics to adopt by using the comparative reports and charts that **PASS** produces. They decide to compare the powers from binomial enumeration and actual alphas for various sample sizes between 1000 and 1200 when R_1 is 1.25.

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 the Ratio of Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Proportions**. 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Power Calculation Method.....	Binomial Enumeration
Maximum N1 or N2 for Binomial Enumeration	5000
Zero Count Adjustment Method.....	Add to zero cells only
Zero Count Adjustment Value	0.0001
Higher Proportions Are	Worse ($H_1: P_1/P_2 < R_0$)
Test Type.....	Likelihood Score (Farr. & Mann.)
Alpha.....	0.025
Group Allocation	Equal ($N_1 = N_2$)
Sample Size Per Group.....	1000 1100 1200
R_0 (Non-Inferiority Ratio).....	2.0
R_1 (Actual Ratio)	1.25
P_2 (Group 2 Proportion)	0.06
Reports Tab	
Show Comparative Reports.....	Checked
Comparative Plots Tab	
Show Comparative Plots	Checked

Non-Inferiority Tests for the Ratio of Two Proportions

Output

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

Numeric Results and Plots

Power Comparison of Three Different Tests

Hypotheses: $H_0: P_1/P_2 \geq R_0$ vs. $H_1: P_1/P_2 < R_0$

N1/N2	P2	P1.1	Target Alpha	F.M. Score Power	M.N. Score Power	G.N. Score Power
1000/1000	0.0600	0.0750	0.0250	0.7923	0.7923	0.7847
1100/1100	0.0600	0.0750	0.0250	0.8276	0.8276	0.8220
1200/1200	0.0600	0.0750	0.0250	0.8578	0.8578	0.8527

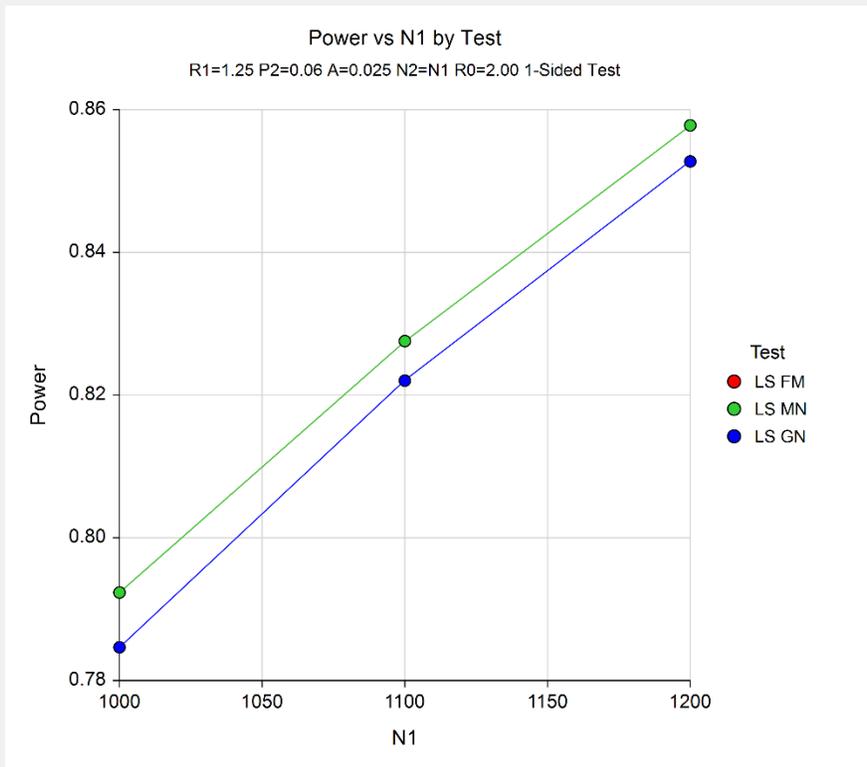
Note: Power was computed using binomial enumeration of all possible outcomes.

Actual Alpha Comparison of Three Different Tests

Hypotheses: $H_0: P_1/P_2 \geq R_0$ vs. $H_1: P_1/P_2 < R_0$

N1/N2	P2	P1.1	Target Alpha	F.M. Score Alpha	M.N. Score Alpha	G.N. Score Alpha
1000/1000	0.0600	0.0750	0.0250	0.0264	0.0264	0.0248
1100/1100	0.0600	0.0750	0.0250	0.0265	0.0262	0.0250
1200/1200	0.0600	0.0750	0.0250	0.0262	0.0262	0.0250

Note: Actual alpha was computed using binomial enumeration of all possible outcomes.



It is interesting to note that the powers of the Gart & Nam test statistics are consistently lower than the other tests. Notice, however, that the actual alpha levels for the Gart & Nam tests are consistently lower than the other two tests and achieve the target alpha level.

Non-Inferiority Tests for the Ratio of Two Proportions

Example 4 – Comparing Power Calculation Methods

Continuing with Example 3, let's see how the results compare if we were to use approximate power calculations instead of power calculations based on binomial enumeration.

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 the Ratio of Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Proportions**. 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**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Power Calculation Method.....	Normal Approximation
Higher Proportions Are	Worse (H1: P1/P2 < R0)
Test Type.....	Likelihood Score (Farr. & Mann.)
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	1000 1100 1200
R0 (Non-Inferiority Ratio).....	2.0
R1 (Actual Ratio)	1.25
P2 (Group 2 Proportion)	0.06
Reports Tab	
Show Power Detail Report	Checked

Output

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

Numeric Results and Plots**Power Detail Report**

Test Statistic: Farrington & Manning Likelihood Score Test
Hypotheses: H0: P1/P2 ≥ R0 vs. H1: P1/P2 < R0

N1/N2	P2	R0	R1	Normal Approximation		Binomial Enumeration	
				Power	Alpha	Power	Alpha
1000/1000	0.0600	2.000	1.250	0.79005	0.0250	0.79234	0.0264
1100/1100	0.0600	2.000	1.250	0.82497	0.0250	0.82757	0.0265
1200/1200	0.0600	2.000	1.250	0.85467	0.0250	0.85780	0.0262

Notice that the approximate power values are pretty close to the binomial enumeration values for all sample sizes.

Non-Inferiority Tests for the Ratio of Two Proportions

Example 5 – Validation of Power Calculations using Blackwelder (1993)

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: $P_2 = 0.04$, $R_0 = 0.3$, $R_1 = 0.1$, $N_1 = N_2 = 1044$, one-sided $\alpha = 0.05$, and $\beta = 0.20$. Using the Farrington and Manning likelihood-score test statistic, he found the binomial enumeration power to be 0.812, the actual 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 the Ratio of Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Proportions**. You may then make the appropriate entries as listed below, or open **Example 5(a or b)** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Power Calculation Method.....	Binomial Enumeration
Maximum N1 or N2 for Binomial Enumeration	5000
Zero Count Adjustment Method.....	Add to zero cells only
Zero Count Adjustment Value	0.0001
Higher Proportions Are	Worse (H1: $P_1/P_2 < R_0$)
Test Type.....	Likelihood Score (Farr. & Mann.)
Alpha.....	0.05
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	1044
R0 (Non-Inferiority Ratio).....	0.3
R1 (Actual Ratio)	0.1
P2 (Group 2 Proportion)	0.04

Output

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

Numeric Results

Numeric Results

Test Statistic: Farrington & Manning Likelihood Score Test
Hypotheses: $H_0: P_1/P_2 \geq R_0$ vs. $H_1: P_1/P_2 < R_0$

Power*	N1	N2	N	Ref. P2	P1 H0 P1.0	P1 H1 P1.1	NI Ratio R0	Ratio R1	Target Alpha	Actual Alpha*
0.81178	1044	1044	2088	0.0400	0.0120	0.0040	0.300	0.100	0.0500	0.0444

* Power and actual alpha were computed using binomial enumeration of all possible outcomes.

PASS calculated the power to be 0.81178 and the actual alpha to be 0.0444, which round to Blackwelder's values.

Non-Inferiority Tests for the Ratio of Two Proportions

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

<u>Option</u>	<u>Value</u>
Power Calculation Method.....	Normal Approximation

Numeric Results

Numeric Results _____

Test Statistic: Farrington & Manning Likelihood Score Test
 Hypotheses: H0: P1/P2 ≥ R0 vs. H1: P1/P2 < R0

Power*	N1	N2	N	Ref. P2	P1 H0 P1.0	P1 H1 P1.1	NI Ratio R0	Ratio R1	Alpha
0.79373	1044	1044	2088	0.0400	0.0120	0.0040	0.300	0.100	0.0500

* Power was computed using the normal approximation method.

PASS also calculated the power to be 0.794.