Chapter 161

Non-Inferiority Tests for the Ratio of Two Correlated Proportions

Introduction

This module computes power and sample size for non-inferiority tests of the ratio in which two dichotomous responses are measured on each subject. When one is interested in showing that the true proportions are different, the data are often analyzed with McNemar's test. However, we are interested in showing non-inferiority rather than difference. For example, suppose a diagnostic procedure is accurate, but is expensive to apply or has serious side effects. A replacement procedure is sought which is no less accurate, but is less expensive or has fewer side effects. In this case, we are not interested in showing that the two diagnostic procedures are different, but rather that the second is no worse than the first. *Non-inferiority tests* were designed for this situation.

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

Technical Details

The results of a study in which two dichotomous responses are measured on each subject can be displayed in a 2-by-2 table in which one response is shown across the columns and the other is shown down the rows. In the discussion to follow, the columns of the table represent the standard (reference or control) response and the rows represent the treatment (experimental) response. The outcome probabilities can be classified into the following table.

Experimental Diagnosis	Standard Diagnosis Yes	Standard Diagnosis No	Total
Yes	p_{11}	p_{10}	P_T
No	p_{01}	p_{00}	$1-P_T$
Total	P_S	$1-P_S$	1

In this table, $p_{ij} = p_{Treatment,\,Standard}$. That is, the first subscript represents the response of the new, experimental procedure while the second subscript represents the response of the standard procedure. Thus, p_{01} represents the proportion having a negative treatment response and a positive standard response.

Sensitivity, Specificity, and Prevalence

To aid in interpretation, analysts have developed a few proportions that summarize the table. Three of the most popular ratios are *sensitivity*, *specificity*, and *prevalence*.

Sensitivity

Sensitivity is the proportion of subjects with a positive standard response who also have a positive experimental response. In terms of proportions from a 2-by-2 table,

$$\text{Sensitivity} = p_{11} / \left(p_{01} + p_{11} \right) = p_{11} / P_{\mathrm{S}}$$

Specificity

Specificity is the proportion of subjects with a negative standard response who also have a negative experimental response. In terms of proportions from a 2-by-2 table,

$$\text{Specificity} = p_{00} / \left(p_{10} + p_{00} \right)$$

Prevalence

Prevalence is the overall proportion of individuals with the disease (or feature of interest). In terms of proportions from a 2-by-2 table,

$$\text{Prevalence} = P_{\mathrm{S}}$$

Table Probabilities

The outcome counts from a sample of *n* subjects can be classified into the following table.

Experimental	Standard	Diagnosis	
Diagnosis	Yes	No	Total
Yes	n_{11}	n_{10}	n_T
No	n_{01}	n_{00}	$n - n_T$
Total	n_S	$n - n_S$	n

Note that $n_{11} + n_{00}$ is the number of matches (*concordant pairs*) and $n_{10} + n_{01}$ is the number of *discordant pairs*.

The hypothesis of interest concerns the two marginal probabilities P_{T} and P_{S}. P_{S} represents the accuracy or success of the standard test and P_{T} represents the accuracy or success of the new, experimental test. Non-inferiority is defined in terms of either the difference, $D = P_{\mathrm{T}} - P_{\mathrm{S}}$, or the relative risk ratio, $R = P_{\mathrm{T}} / P_{\mathrm{S}}$, of these two proportions. The choice between *D* and *R* will usually lead to different sample sizes to achieve the same power.

Non-Inferiority Hypotheses using Ratios

The following is based on Nam and Blackwelder (2002). We refer you to this paper for the complete details of which we will only provide a brief summary here.

When $R_E < 1$, the statistical hypotheses of non-inferiority are

$$H_0: P_T / P_S \le R_E \text{ versus } H_1: P_T / P_S > R_E$$

Test Statistics

The test statistic for an asymptotic test based on constrained maximum likelihood for large *n* is given by

$$Z(R_E) = \sqrt{\frac{n\left(\hat{P}_T - R_E\hat{P}_S\right)}{R_E\left(\tilde{p}_{10} + \tilde{p}_{01}\right)}}$$

where

$$\tilde{p}_{10} = \frac{-\hat{P}_T + R_E^2\left(\hat{P}_S + 2\hat{p}_{10}\right) + \sqrt{\left(\hat{P}_T - R_E^2\hat{P}_S\right)^2 + 4R_E^2\hat{p}_{10}\hat{p}_{01}}}{2R_E\left(R_E + 1\right)}$$

$$\tilde{p}_{01} = R_E\tilde{p}_{10} - \left(R_E - 1\right)\left(1 - \hat{p}_{00}\right)$$

$$\hat{p}_{01} = \frac{n_{01}}{n},\ \hat{p}_{10} = \frac{n_{10}}{n},\ \hat{P}_T = \frac{n_{10} + n_{11}}{n},\ \hat{P}_S = \frac{n_{01} + n_{11}}{n}$$

Power Formula

The power when the true value of the relative risk ratio is R_E can be evaluated exactly using the multinomial distribution. When *n* is large, we use a normal approximation to the multinomial distribution which leads to

$$\beta(R_A) = \Phi(c_U)$$

where

$$c_U = \frac{z_{1-\alpha}\sqrt{\overline{V}_0(T_0)} - E_1(T_0)}{\sqrt{V_1(T_0)}}$$

$$\overline{V}_0(T_0) = \frac{R_E\left(\overline{p}_{10} + \overline{p}_{01}\right)}{n}$$

$$E_1(T_0) = \left(R_A - R_E\right)P_S$$

$$V_1(T_0) = \frac{\left(R_A + R_E^2\right)P_S - 2R_E p_{11} - \left(R_A - R_E\right)^2 P_S^2}{n}$$

$$\overline{p}_{10} = \frac{-P_T + R_E^2\left(P_S + 2p_{10}\right) + \sqrt{\left(P_T - R_E^2 P_S\right)^2 + 4R_E^2 p_{10}p_{01}}}{2R_E\left(R_E + 1\right)}$$

$$\overline{p}_{01} = R_E\overline{p}_{10} - \left(R_E - 1\right)\left(1 - p_{00}\right)$$

Nuisance Parameter

Unfortunately, the 2-by-2 table includes four parameters p_{11}, p_{10}, p_{01}, and p_{00}, but the power specifications above only specify two parameters: P_S and D_A or R_A. A third parameter is defined implicitly since the sum of the four parameters is one. One parameter, known as a nuisance parameter, remains unaccounted for. This parameter must be addressed to fully specify the problem. This fourth parameter can be specified by specifying any one of the following: p_{11}, p_{10}, p_{01}, p_{00}, $p_{10} + p_{01}$, $p_{11} + p_{00}$, or the sensitivity of the experimental response, p_{11} / P_S.

It may be difficult to specify a reasonable value for the nuisance parameter since its value may not be even approximately known until after the study is conducted. Because of this, we suggest that you calculate power or sample size for a range of values of the nuisance parameter. This will allow you to determine how sensitive the results are to its value.

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, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *Power* or *Sample Size*.

Power Calculation

Power Calculation Method

Select the method to be used to calculate power.

The choices are

- **Multinomial Enumeration**

  Power is computed using multinomial enumeration of all possible outcomes when N ≤ Max N for Multinomial Enumeration (otherwise, the normal approximation is used). Multinomial enumeration of all outcomes is possible because of the discrete nature of the data.

- **Normal Approximation**

  Approximate power is computed using the normal approximation to the multinomial distribution.

The exact calculation using the multinomial distribution becomes very time consuming for N > 500. When N > 500, the difference between the multinomial and approximate calculations is small.

For small values of N (less than 100), the Multinomial Enumeration power may be overly optimistic because the discrete nature of the trinomial distribution results in the actual alpha value being higher than its target. To be on the safe side, we recommend that you use the approximate calculation.

Max N for Multinomial Enumeration

Only shown when Power Calculation Method = "Multinomial Enumeration"

Specify the maximum value of N (sample size) that uses the exact power calculation based on the multinomial distribution. N's greater than this value will use the asymptotic approximation.

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. Here, a type-II error occurs when you fail to conclude non-inferiority when in fact it is true.

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. Here, a type-I error occurs when you falsely conclude non-inferiority.

Sample Size

N (Sample Size)

Enter a value for the sample size. This value must be greater than two. You may enter a range of values such as *10 to100 by 10*.

Effect Size – Ratios

Rni (Non-Inferiority Ratio)

Rni is the minimum size of the relative risk ratio, P_T / P_S, that will still result in the conclusion of non-inferiority. Non-inferiority trials use a value that is less than one. Typical values for this ratio are 0.8 or 0.9.

Ra (Actual Ratio)

Enter a value for Ra, the actual relative risk ratio P_T / P_S. This value is used to generate the value of P_T using the formula $P_T = P_S R_a$. Often this value is set equal to one, but this is not necessary.

Effect Size – Standard Proportion

Ps (Standard Proportion)

This is the proportion of yes's (or successes), P_S , when subjects received the standard treatment. This value or a good estimate is often available from previous studies.

You may enter a set of values separated by blank spaces. For example, you could enter '0.50 0.60 0.70'. Values between, but not including, 0 and 1 are permitted.

Effect Size – Nuisance Parameter

Nuisance Parameter Type

Enter the type of nuisance parameter here. Unfortunately, the 2-by-2 table cannot be completely specified by using only the parameters Ps and Da or Ps and Ra. One other parameter must be specified. This additional parameter is called a "nuisance" parameter. It will be assumed to be a known quantity. Several possible choices are available. This option lets you specify which parameter you want to use. In all cases, the value you specify is a proportion.

- **P11 (% Positive Matches)**
  The proportion of subjects that are positive on both tests.

- **P00 (% Negative Matches)**
  The proportion of subjects that are negative on both tests.

- **P01 (% -Trt +Std)**
  The proportion of subjects that are negative on the treatment, but positive on the standard.

- **P10 (% +Trt -Std)**
  The proportion of subjects that are positive on the treatment, but negative on the standard.

- **P11+P00 (% Matches)**
  The proportion of matches (concordant pairs).

- **P01+P10 (% Disagree)**
  The proportion of non-matches (discordant pairs).

- **P11/Ps (Sensitivity)**
  The sensitivity.

Nuisance Parameter Value

Enter the value of the nuisance parameter that you specified in the "Nuisance Parameter Type" box. This value is a proportion, so it must be between 0 and 1.

Example 1 – Finding Power

Researchers have developed a new treatment for migraine headaches which is less expensive than a current standard. The researchers need to show that the proportion of individuals who respond to the new treatment is not inferior to the standard treatment. The new treatment will be considered non-inferior if its success rate is no less than 95% of the success rate of the standard, which is about 0.65. They want to study the power for various sample sizes between 500 and 4000 at the 5% significance level. They'll study various values of the nuisance parameter: P11/Ps = sensitivity (0.5 to 0.9).

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 Correlated Proportions** procedure window by expanding **Proportions**, then **Two Correlated Proportions**, then clicking on **Non-Inferiority**, and then clicking on **Non-Inferiority Tests for the Ratio of Two Correlated 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**
Alpha	**0.05**
N (Sample Size)	**500 to 4000 by 500**
Rni (Non-Inferiority Ratio)	**0.95**
Ra (Actual Ratio)	**1.0**
Ps (Standard Proportion)	**0.65**
Nuisance Parameter Type	**P11/Ps (Sensitivity)**
Nuisance Parameter Value	**0.5 to 0.9 by 0.1**

Annotated Output

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

Numeric Results

Numeric Results for a Non-Inferiority (One-Sided) Test of a Ratio

Power	Sample Size N	Non-Inf. Ratio Rni	Actual Ratio Ra	Treatment Proportion Pt	Standard Proportion Ps	Nuisance Parameter P11/Ps	Alpha
0.23552	500	0.950	1.000	0.650	0.650	0.500	0.050
0.27009	500	0.950	1.000	0.650	0.650	0.600	0.050
0.32464	500	0.950	1.000	0.650	0.650	0.700	0.050
0.42349	500	0.950	1.000	0.650	0.650	0.800	0.050
0.64700	500	0.950	1.000	0.650	0.650	0.900	0.050
0.36772	1000	0.950	1.000	0.650	0.650	0.500	0.050
0.42680	1000	0.950	1.000	0.650	0.650	0.600	0.050
0.51572	1000	0.950	1.000	0.650	0.650	0.700	0.050
0.65984	1000	0.950	1.000	0.650	0.650	0.800	0.050
0.89100	1000	0.950	1.000	0.650	0.650	0.900	0.050
0.48241	1500	0.950	1.000	0.650	0.650	0.500	0.050
0.55722	1500	0.950	1.000	0.650	0.650	0.600	0.050
0.66224	1500	0.950	1.000	0.650	0.650	0.700	0.050
0.80955	1500	0.950	1.000	0.650	0.650	0.800	0.050

(report continues)

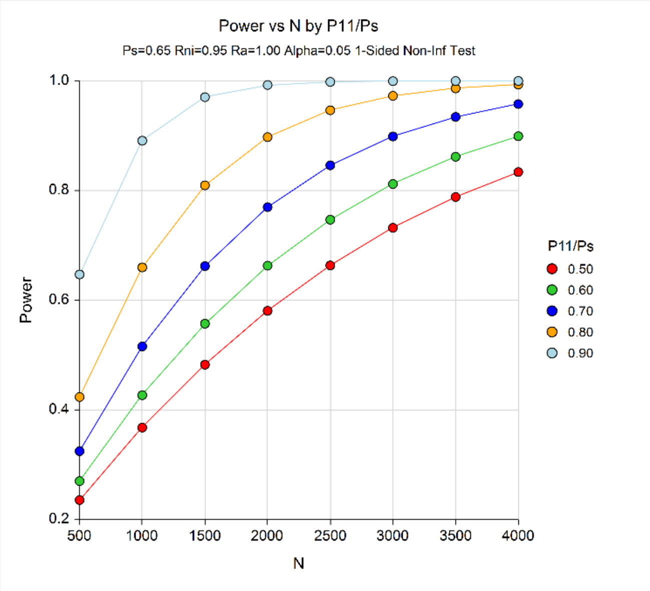
* Power was computed using the normal approximation method.

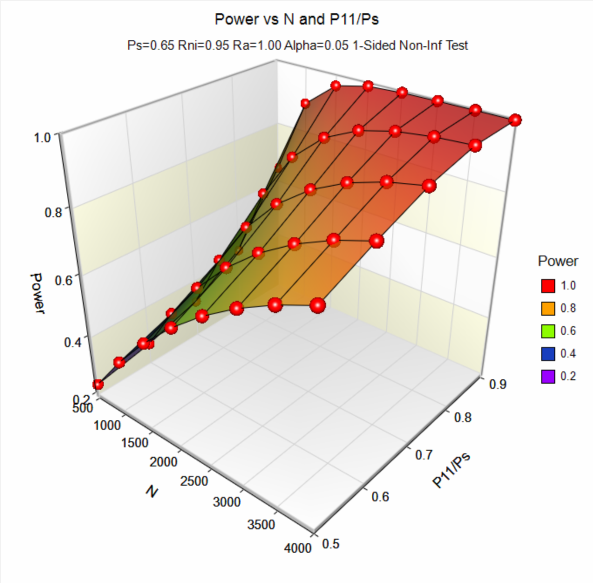
Report Definitions
Power is the probability of rejecting a false null hypothesis.
N is the number of subjects, the sample size.
Rni is the maximum ratio between Pt and Ps that is still called 'non-inferior'.
Ra is the actual ratio between Pt and Ps. That is, Ra = Pt/Ps.
Pt is the response proportion to the treatment (experimental or new) test.
Ps is the response proportion to the standard (reference or old) test.
The Nuisance Parameter is a value that is needed, but is not a direct part of the hypothesis.
Alpha is the probability of rejecting a true null hypothesis.

Summary Statements
A sample size of 500 subjects achieves 23.552% power at a significance level of 0.050 using a one-sided non-inferiority test of correlated proportions when the standard proportion is 0.650, the maximum ratio of these proportions that still results in non-inferiority (the range of non-inferiority) is 0.950, and the actual ratio of the proportions is 1.000.

This report shows the power for the indicated scenarios. All of the columns are defined in the 'Report Definitions' section.

Plots Section





These plots show the power versus the sample size for the various values of sensitivity. In this example, we see that the value of the nuisance parameter has a large effect on the calculated power.

Example 2 – Finding Sample Size

Continuing with Example 1, the analysts want to determine the exact sample size necessary to achieve 90% power for all values of the nuisance parameter.

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

Option	Value
Design Tab	
Find	**Sample Size**
Power Calculation Method	**Normal Approximation**
Power	**0.90**
Alpha	**0.05**
Rni (Non-Inferiority Ratio)	**0.95**
Ra (Actual Ratio)	**1.0**
Ps (Standard Proportion)	**0.65**
Nuisance Parameter Type	**P11/Ps (Sensitivity)**
Nuisance Parameter Value	**0.5 to 0.9 by 0.1**

Output

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

Numeric Results for a Non-Inferiority (One-Sided) Test of a Ratio

Power	Sample Size N	Non-Inf. Ratio Rni	Actual Ratio Ra	Treatment Proportion Pt	Standard Proportion Ps	Nuisance Parameter P11/Ps	Alpha
0.90004	5013	0.950	1.000	0.650	0.650	0.500	0.050
0.90001	4013	0.950	1.000	0.650	0.650	0.600	0.050
0.90004	3015	0.950	1.000	0.650	0.650	0.700	0.050
0.90011	2020	0.950	1.000	0.650	0.650	0.800	0.050
0.90016	1035	0.950	1.000	0.650	0.650	0.900	0.050

* Power was computed using the normal approximation method.

These scenarios require a large sample size. In fact, the first two rows are blank because the sample size is so large it can't be determined.

Example 3 – Validation using Nam and Blackwelder (2002)

Nam and Blackwelder (2002) give an example in which Ps is 0.80, P10 is 0.05, Ra is 1.00, Rni is 0.80, the significance level is 0.05, and the power is 80%. From their Table III, the sample size is 34. Note that their calculations use the approximate formula.

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

Option	Value
Design Tab	
Find	**Sample Size**
Power Calculation Method	**Normal Approximation**
Power	**0.80**
Alpha	**0.05**
Rni (Non-Inferiority Ratio)	**0.80**
Ra (Actual Ratio)	**1.0**
Ps (Standard Proportion)	**0.80**
Nuisance Parameter Type	**P10 (% +Trt -Std)**
Nuisance Parameter Value	**0.05**

Output

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

Numeric Results

Numeric Results for a Non-Inferiority (One-Sided) Test of a Ratio

Power	Sample Size N	Non-Inf. Ratio Rni	Actual Ratio Ra	Treatment Proportion Pt	Standard Proportion Ps	Nuisance Parameter P10	Alpha
0.80050	34	0.800	1.000	0.800	0.800	0.050	0.050

* Power was computed using the normal approximation method.

The calculated sample size of 34 matches the results of Nam and Blackwelder (2002).