Chapter 214

Equivalence Tests for the Ratio of Two Proportions

Introduction

This module provides power analysis and sample size calculation for equivalence tests of the ratio in two-sample designs in which the outcome is binary. The equivalence test is usually carried out using the Two One-Sided Tests (TOST) method. This procedure computes power and sample size for the TOST equivalence test method. 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

An equivalence test example will set the stage for the discussion of the terminology that follows. Suppose that the response rate of the standard treatment of a disease is 0.70. 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 therapeutically equivalent to the standard treatment.

Because of the many benefits of the new treatment, clinicians are willing to adopt the new treatment even if its effectiveness is slightly different from the standard. After thoughtful discussion with several clinicians, it is decided that if the response rate ratio of the new treatment to the standard treatment is between 0.9 and 1.1, the new treatment would be adopted.

The developers must design an experiment to test the hypothesis that the response rate ratio of the new treatment to the standard is between 0.9 and 1.1. The statistical hypothesis to be tested is

$$H_0 : p_1 / p_2 < 0.9 \text{ or } p_1 / p_2 > 1.1 \text{ versus } H_1 : 0.9 \le p_1 / p_2 \le 1.1$$

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 equivalence 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.

Equivalence Tests for the Ratio of Two Proportions

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

Group	Success	Failure	Total
Treatment	a	c	m
Control	b	d	n
Totals	s	f	N

The following alternative notation is also used.

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

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

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 δ represent the smallest difference (margin of equivalence) between the two proportions that still results in the conclusion that the new treatment is equivalent to the current treatment. The set of statistical hypotheses that are tested is

$$H_0 : |p_{1.0} - p_2| \geq \delta \text{ versus } H_1 : |p_{1.0} - p_2| < \delta$$

These hypotheses can be rearranged to give

$$H_0 : p_{1.0} - p_2 \leq \delta_L \text{ or } p_{1.0} - p_2 \geq \delta_U \text{ versus } H_1 : \delta_L \leq p_{1.0} - p_2 \leq \delta_U$$

This composite hypothesis can be reduced to two one-sided hypotheses as follows

$$H_{0L} : p_{1.0} - p_2 \leq \delta_L \text{ versus } H_{1L} : \delta_L \leq p_{1.0} - p_2$$

$$H_{0U} : p_{1.0} - p_2 \geq \delta_U \text{ versus } H_{1U} : \delta_U \geq p_{1.0} - p_2$$

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 reporting the difference, ratio, or odds ratio. Mathematically, the definitions of these parameterizations are

Parameter	**Computation**	**Alternative Hypotheses**
Difference	$\delta = p_{1.0} - p_2$	$H_1 : \delta_L \leq p_{1.0} - p_2 \leq \delta_U$
Ratio	$\phi = p_{1.0} / p_2$	$H_1 : \phi_L \leq p_{1.0} / p_2 \leq \phi_U$
Odds Ratio	$\psi = Odds_{1.0} / Odds_2$	$H_1 : \psi_L \leq o_{1.0} / o_2 \leq \psi_U$

Ratio

The ratio, $\phi = p_{1.0} / p_2$, gives the relative change in the probability of the response. Testing equivalence uses the formulation

$$H_0: p_{1.0} / p_2 \le \phi_L \text{ or } p_{1.0} / p_2 \ge \phi_U \text{ versus } H_1: \phi_L \le p_{1.0} / p_2 \le \phi_U$$

The only subtlety is that for equivalence tests $\phi_L < 1$ and $\phi_U > 1$. Usually, $\phi_L = 1 / \phi_U$.

Equivalence using a Ratio

The following example might help you understand the concept of *equivalence* as defined by the ratio. Suppose that 60% of patients $(p_2 = 0.60)$ respond to the current treatment method. If the response rate of a new treatment is within 10% of 0.60, it will be considered to be equivalent to the standard treatment. Substituting these figures into the statistical hypotheses gives

$$H_0: p_{1.0} / p_2 \le 0.9 \text{ or } p_{1.0} / p_2 \ge 1.1 \text{ versus } H_1: 0.9 \le p_{1.0} / p_2 \le 1.1$$

Using the relationship

$$p_{1.0} = \phi_0 p_2$$

gives

$$H_0: p_{1.0} \le 0.54 \text{ or } p_{1.0} \ge 0.66 \text{ versus } H_1: 0.54 \le p_{1.0} \le 0.66$$

The equivalence test is usually carried out using the Two One-Sided Tests (TOST) method. This procedure computes power and sample size for the TOST equivalence test method.

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 values using the standard normal distribution. The critical values z_L and z_U are chosen as 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 $z_t > z_L$ and $z_t < z_U$, 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

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

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

$$\alpha_L = \sum_A \binom{n_1}{x_{11}} p_{1.0L}^{x_{11}} q_{1.0L}^{n_1 - x_{11}} \binom{n_2}{x_{21}} p_2^{x_{21}} q_2^{n_2 - x_{21}}$$

and

$$\alpha_U = \sum_A \binom{n_1}{x_{11}} p_{1.0U}^{x_{11}} q_{1.0U}^{n_1-x_{11}} \binom{n_2}{x_{21}} p_2^{x_{21}} q_2^{n_2-x_{21}}$$

The value of alpha is then computed as the maximum of α_L and α_U.

Asymptotic Approximations

When the values of n_1 and n_2 are large (say over 200), these formulas 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

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 $p_{1.1}$ and p_2 can 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: <u>one should use the test statistic that will be used to analyze the data</u>. 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

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

where

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

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

$$A = N\phi_0$$

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

$$C = m_1$$

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

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

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, equivalence and actual 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.

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 and the odds ratio will be affected by the amount specified here!

Test

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. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of unequal proportions when in fact they are equivalent.

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. In this procedure, a type-I error occurs when you reject the null hypothesis of unequal proportions when in fact they are not equal.

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.

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 <u>Not</u> 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,

$$R = N2/N1$$

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

N2 is calculated from *N1* using the formula:

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

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.

Effect Size – Equivalence Ratios

R0.U & R0.L (Upper & Lower Equivalence Ratio)

Specify the *margin of equivalence* by specifying the largest ratio (P1/P2) above (R0.U) and below (R0.L) 1 which will still result in the conclusion of equivalence. As long as the actual ratio is between these two values, the difference between the proportions is not considered to be large enough to be of practical importance.

The values of R0.U must be greater than 1 and the values of R0.L must be less than 1. R0.L can be set to '1/R0.U,' which is often desired.

The power calculations assume that P1.0 is the value of P1 under the null hypothesis. This value is used with P2 to calculate the value of P1.0 using the formula: $P1.0U = R0.U \times P2$.

You may enter a range of values for R0.U such as *1.1 1.5 1.8* or *1.1 to 2.1 by 0.2*.

Note that if you enter values for R0.L (other than '1/R0.U'), they are used in pairs with the values of R0.U. Thus, the first values of R0.U and R0.L are used together, then the second values of each are used, and so on.

R0.L must be between 0 and 1. R0.U must be greater than 1. Neither can take on the value 1.

Effect Size – Actual Ratio

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 equivalence 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 (Group 2 Proportion)

Specify the value of P2, 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*.

Example 1 – Finding Power

A study is being designed to establish the equivalence of a new treatment compared to the current treatment. Historically, the current treatment has enjoyed a 65% 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 the rate ratio of treatment to control is between 0.75 and 1.333.

The researchers plan to use the Farrington and Manning likelihood score test statistic to analyze the data. They want to study the power of the Farrington and Manning test at group sample sizes ranging from 50 to 500 for detecting a rate ratio between 0.75 and 1.333 when the actual rate ratio ranges from 1.0 to 1.2. The significance level will be 0.05.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Equivalence Tests for the Ratio of Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence 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**
Test Type... **Likelihood Score (Farr. & Mann.)**
Alpha.. **0.05**
Group Allocation .... **Equal (N1 = N2)**
Sample Size Per Group.......................... **50 to 500 by 50**
R0.U (Upper Equivalence Ratio) **1.333**
R0.L (Lower Equivalence Ratio)............. **1/R0.U**
R1 (Actual Ratio) **1.0 1.1 1.2**
P2 (Group 2 Proportion) **0.65**

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.L or P1/P2 ≥ R0.U   vs.   H1: R0.L < P1/P2 < R0.U

Power*	N1	N2	N	Ref. P2	P1.0L	P1.0U	R0.L	R0.U	R1	Alpha
0.2089	50	50	100	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.7120	100	100	200	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.9060	150	150	300	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.9715	200	200	400	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.9918	250	250	500	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.9978	300	300	600	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.9994	350	350	700	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.9998	400	400	800	0.650	0.488	0.866	0.750	1.333	1.000	0.0500

(Report continues)

* Power was computed using the normal approximation method.

References
Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York.
    Volume 2, 1367-1372.
Chow, S.C. and Liu, J.P. 1999. Design and Analysis of Bioavailability and Bioequivalence Studies. Marcel
    Dekker. New York.
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.
Tubert-Bitter, P., Manfredi,R., Lellouch, J., Begaud, B. 2000. 'Sample size calculations for risk equivalence
    testing in pharmacoepidemiology.' Journal of Clinical Epidemiology 53, 1268-1274.

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, the standard, reference, or control group.
P1.0L is the smallest treatment-group response rate that still yields an equivalence conclusion. P1.0U is the
    largest treatment-group response rate that still yields an equivalence conclusion.
R0.L is the lowest ratio that still results in the conclusion of equivalence. R0.U is the highest ratio that
    still results in the conclusion of equivalence. R1 is the actual ratio, P1/P2, at which the power is
    calculated.
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 the treatment group and 50 in the reference group achieve 20.89% power to
detect equivalence. The margin of equivalence, given in terms of the ratio, extends from 0.750
to 1.333. The actual ratio is 1.000. The reference group proportion is 0.650. The calculations
assume that two, one-sided likelihood score (Farrington & Manning) tests are used. The
significance level of the test is 0.0500.

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

Plots Section
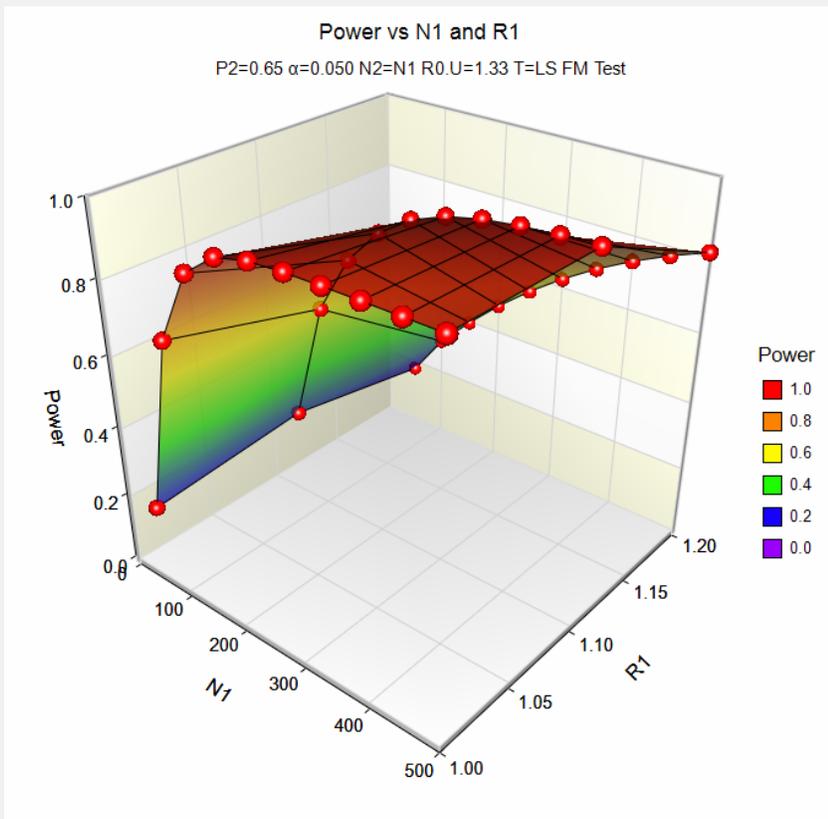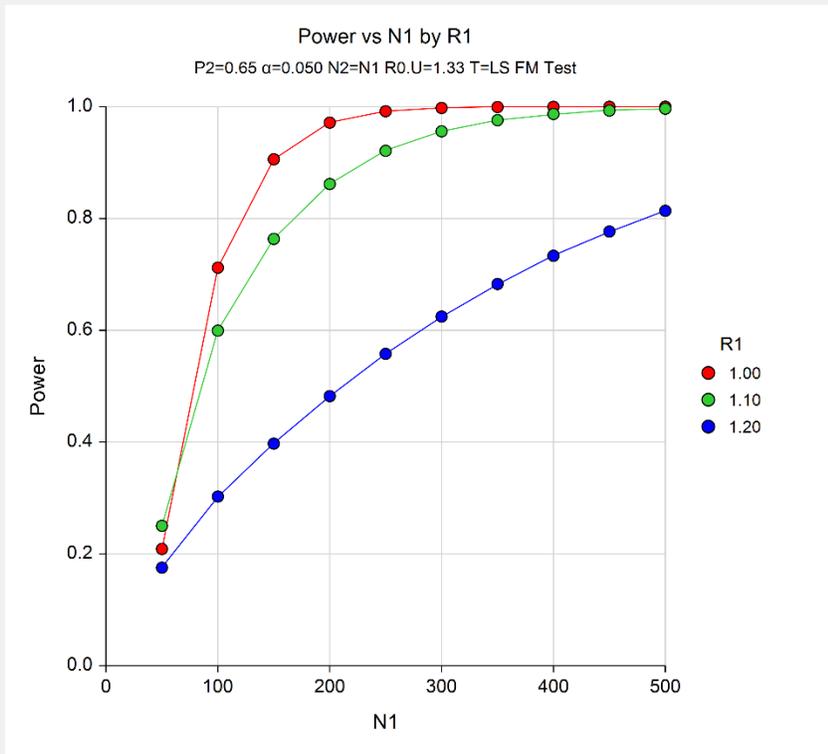
The values from the table are displayed in the above charts. These charts give a quick look at the sample size that will be required for various values of R1.

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

Option	**Value**
Design Tab	
Solve For	**Sample Size**
Power Calculation Method......................	**Normal Approximation**
Test Type..	**Likelihood Score (Farr. & Mann.)**
Power..	**0.80**
Alpha...	**0.05**
Group Allocation	**Equal (N1 = N2)**
R0.U (Upper Equivalence Ratio)	**1.333**
R0.L (Lower Equivalence Ratio).............	**1/R0.U**
R1 (Actual Ratio)	**1.0 1.1 1.2**
P2 (Group 2 Proportion)	**0.65**

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.L or P1/P2 ≥ R0.U   vs.   H1: R0.L < P1/P2 < R0.U

Target Power	Actual Power*	N1	N2	N	Ref. P2	P1.0L	P1.0U	R0.L	R0.U	R1	Alpha
0.80	0.8012	117	117	234	0.650	0.488	0.866	0.750	1.333	1.000	0.0500
0.80	0.8003	166	166	332	0.650	0.488	0.866	0.750	1.333	1.100	0.0500
0.80	0.8004	481	481	962	0.650	0.488	0.866	0.750	1.333	1.200	0.0500

* 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 and actual alphas for various sample sizes between 50 and 200 when R1 is 1.0.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Equivalence Tests for the Ratio of Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence 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**.

Option	**Value**
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**
Test Type...	**Likelihood Score (Farr. & Mann.)**
Alpha..	**0.05**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group......................................	**50 to 200 by 50**
R0.U (Upper Equivalence Ratio)	**1.333**
R0.L (Lower Equivalence Ratio)..........................	**1/R0.U**
R1 (Actual Ratio)	**1.0**
P2 (Group 2 Proportion)	**0.65**
Reports Tab	
Show Comparative Reports................................	**Checked**
Comparative Plots Tab	
Show Comparative Plots	**Checked**

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: H0: P1/P2 ≤ R0.L or P1/P2 ≥ R0.U   vs.   H1: R0.L < P1/P2 < R0.U

| | | | | F.M.
Score | M.N.
Score | G.N.
Score |
| | | | Target | Score | Score | Score |
N1/N2	P2	R0.U	Alpha	Power	Power	Power
50/50	0.6500	1.3330	0.0500	0.2135	0.2135	0.2135
100/100	0.6500	1.3330	0.0500	0.7108	0.7108	0.7108
150/150	0.6500	1.3330	0.0500	0.9064	0.9064	0.9064
200/200	0.6500	1.3330	0.0500	0.9715	0.9714	0.9714

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

Actual Alpha Comparison of Three Different Tests ———————————————————
Hypotheses: H0: P1/P2 ≤ R0.L or P1/P2 ≥ R0.U   vs.   H1: R0.L < P1/P2 < R0.U

| | | | | F.M.
Score | M.N.
Score | G.N.
Score |
| | | | Target | Score | Score | Score |
N1/N2	P2	R0.U	Alpha	Alpha	Alpha	Alpha
50/50	0.6500	1.3330	0.0500	0.0516	0.0516	0.0516
100/100	0.6500	1.3330	0.0500	0.0509	0.0509	0.0509
150/150	0.6500	1.3330	0.0500	0.0510	0.0508	0.0508
200/200	0.6500	1.3330	0.0500	0.0505	0.0500	0.0502

Note: Actual alpha was computed using binomial enumeration of all possible outcomes.
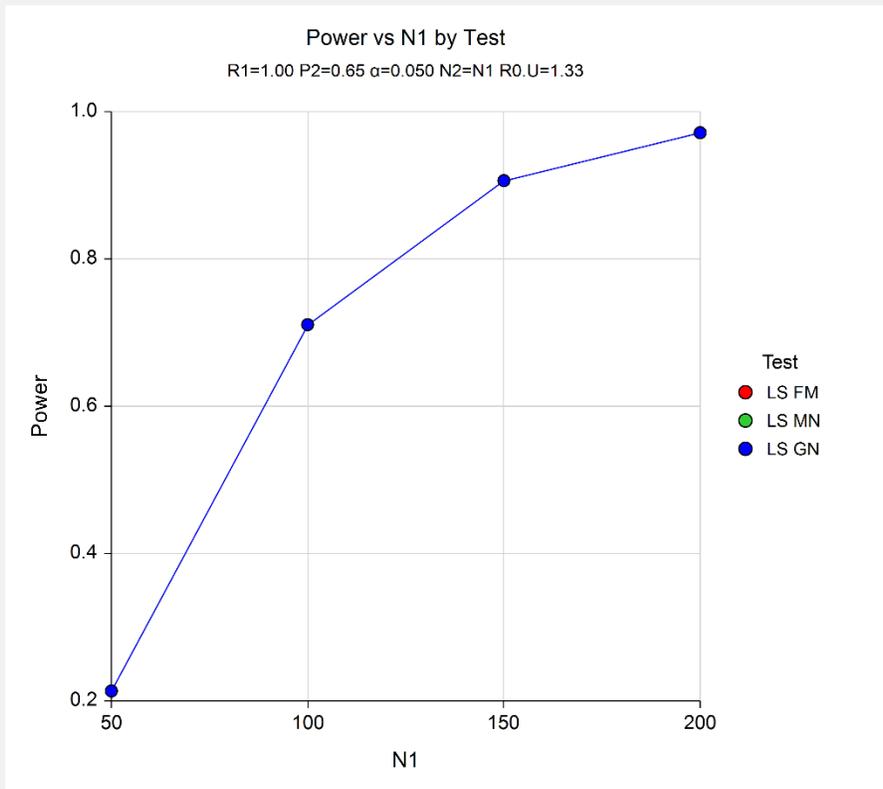
All three test statistics have about the same power for all sample sizes studied.

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

Option   **Value**

Design Tab
Solve For .... **Power**
Power Calculation Method...................... **Normal Approximation**
Test Type... **Likelihood Score (Farr. & Mann.)**
Alpha.. **0.05**
Group Allocation .... **Equal (N1 = N2)**
Sample Size Per Group.......................... **50 to 200 by 50**
R0.U (Upper Equivalence Ratio) **1.333**
R0.L (Lower Equivalence Ratio)............. **1/R0.U**
R1 (Actual Ratio) **1.0**
P2 (Group 2 Proportion) **0.65**

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.L or P1/P2 ≥ R0.U   vs.   H1: R0.L < P1/P2 < R0.U

				Normal Approximation		Binomial Enumeration	
N1/N2	**P2**	**R0.U**	**R1**	**Power**	**Alpha**	**Power**	**Alpha**
50/50	0.6500	1.333	1.000	0.2089	0.0500	0.2135	0.0516
100/100	0.6500	1.333	1.000	0.7120	0.0500	0.7108	0.0509
150/150	0.6500	1.333	1.000	0.9060	0.0500	0.9064	0.0510
200/200	0.6500	1.333	1.000	0.9715	0.0500	0.9715	0.0505

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

Example 5 – Validation

We could not find a validation example for an equivalence test for the ratio of two proportions. The calculations are basically the same as those for a non-inferiority test of the ratio of two proportions, which has been validated using Blackwelder (1993). We refer you to Example 5 of Chapter 211, "Non-Inferiority Tests for the Ratio of Two Proportions," for a validation example.