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

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

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.

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

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

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 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

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

These hypotheses can be rearranged to give

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

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

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

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

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

$$H_0: p_{1.0} / p_2 \leq \phi_L \quad \text{or} \quad p_{1.0} / p_2 \geq \phi_U \quad \text{versus} \quad H_1: \phi_L \leq p_{1.0} / p_2 \leq \phi_U$$

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

$$H_0: p_{1.0} / p_2 \leq 0.9 \quad \text{or} \quad p_{1.0} / p_2 \geq 1.1 \quad \text{versus} \quad H_1: 0.9 \leq p_{1.0} / p_2 \leq 1.1$$

Using the relationship

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

gives

$$H_0: p_{1.0} \leq 0.54 \quad \text{or} \quad p_{1.0} \geq 0.66 \quad \text{versus} \quad H_1: 0.54 \leq p_{1.0} \leq 0.66$$

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 $\alpha_L$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

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 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: 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

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

where

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

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

$$A = N\phi_0$$

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

$$C = m_1$$

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. 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

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

where

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

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

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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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 Reports

Numeric Results

Solve For: Power
Groups: 1 = Treatment, 2 = Reference
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

	Sample Size			Proportions Equivalence				Ratio Equivalence			
Power*	N1	N2	N	Lower P1.0L	Upper P1.0U	Actual P1.1	Reference P2	Lower R0.L	Upper R0.U	Actual R1	Alpha
0.2089	50	50	100	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.7120	100	100	200	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.9060	150	150	300	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.9715	200	200	400	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.9918	250	250	500	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.9978	300	300	600	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.9994	350	350	700	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.9998	400	400	800	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
1.0000	450	450	900	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
1.0000	500	500	1000	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.2502	50	50	100	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.5994	100	100	200	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.7635	150	150	300	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.8617	200	200	400	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.9211	250	250	500	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.9560	300	300	600	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.9759	350	350	700	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.9870	400	400	800	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.9931	450	450	900	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.9964	500	500	1000	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.1754	50	50	100	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.3026	100	100	200	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.3973	150	150	300	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.4822	200	200	400	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.5579	250	250	500	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.6246	300	300	600	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.6828	350	350	700	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.7333	400	400	800	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.7767	450	450	900	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05
0.8137	500	500	1000	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05

* Power was computed using the normal approximation method.

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N1 and N2	The number of items sampled from each population.
N	The total sample size. N = N1 + N2.
P1.0L	The smallest treatment-group response rate that still yields an equivalence conclusion.
P1.0U	The largest treatment-group response rate that still yields an equivalence conclusion.
P1.1	The proportion for group 1 assumed by the alternative hypothesis, H1. Group 1 is the treatment group. P1.1 = P1|H1.
P2	The proportion for group 2. Group 2 is the standard, reference, or control group.
R0.L	The lowest ratio that still results in the conclusion of equivalence.
R0.U	The highest ratio that still results in the conclusion of equivalence.
R1	The actual ratio, P1 / P2, at which the power is calculated.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group design will be used to test whether the Group 1 (treatment) proportion (P1) is equivalent to the Group 2 (reference) proportion (P2), with ratio equivalence bounds of 0.75 and 1.333 (H0: P1 / P2 ≤ 0.75 or P1 / P2 ≥ 1.333 versus H1: 0.75 < P1 / P2 < 1.333). The comparison will be made using two one-sided, two-sample likelihood score (Farrington & Manning) tests with an overall Type I error rate (α) of 0.05. The reference group proportion is assumed to be 0.65. To detect a proportion ratio (P1 / P2) of 1 (or P1 of 0.65) with sample sizes of 50 for Group 1 (treatment) and 50 for Group 2 (reference), the power is 0.2089.

Dropout-Inflated Sample Size

	Sample Size			Dropout-Inflated Enrollment Sample Size			Expected Number of Dropouts		
Dropout Rate	N1	N2	N	N1'	N2'	N'	D1	D2	D
20%	50	50	100	63	63	126	13	13	26
20%	100	100	200	125	125	250	25	25	50
20%	150	150	300	188	188	376	38	38	76
20%	200	200	400	250	250	500	50	50	100
20%	250	250	500	313	313	626	63	63	126
20%	300	300	600	375	375	750	75	75	150
20%	350	350	700	438	438	876	88	88	176
20%	400	400	800	500	500	1000	100	100	200
20%	450	450	900	563	563	1126	113	113	226
20%	500	500	1000	625	625	1250	125	125	250

Dropout Rate — The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N1, N2, and N — The evaluable sample sizes at which power is computed (as entered by the user). If N1 and N2 subjects are evaluated out of the N1' and N2' subjects that are enrolled in the study, the design will achieve the stated power.
N1', N2', and N' — The number of subjects that should be enrolled in the study in order to obtain N1, N2, and N evaluable subjects, based on the assumed dropout rate. N1' and N2' are calculated by inflating N1 and N2 using the formulas N1' = N1 / (1 - DR) and N2' = N2 / (1 - DR), with N1' and N2' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D1, D2, and D — The expected number of dropouts. D1 = N1' - N1, D2 = N2' - N2, and D = D1 + D2.

Dropout Summary Statements

Anticipating a 20% dropout rate, 63 subjects should be enrolled in Group 1, and 63 in Group 2, to obtain final group sample sizes of 50 and 50, respectively.

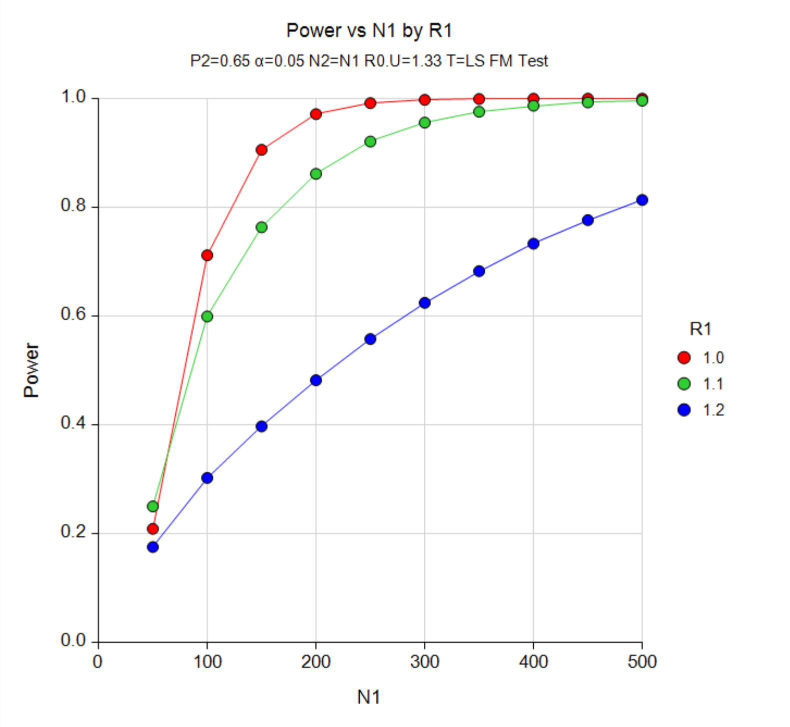
References

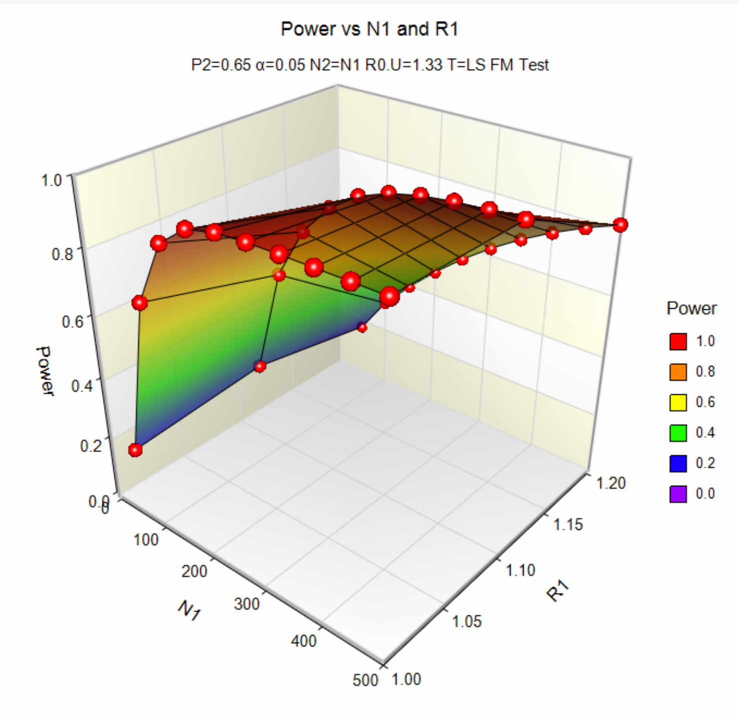
Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.
Chow, S.C. and Liu, J.P. 1999. Design and Analysis of Bioavailability and Bioequivalence Studies. Marcel Dekker. New York.
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, 3rd Edition. Chapman & Hall/CRC. Boca Raton, FL. Pages 86-88.
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.

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

Plots Section

Plots





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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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

Solve For: Sample Size
Groups: 1 = Treatment, 2 = Reference
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		Sample Size			Proportions				Ratio			
					Equivalence				Equivalence			
Target	Actual*	N1	N2	N	Lower P1.0L	Upper P1.0U	Actual P1.1	Reference P2	Lower R0.L	Upper R0.U	Actual R1	Alpha
0.8	0.8012	117	117	234	0.488	0.866	0.650	0.65	0.75	1.333	1.0	0.05
0.8	0.8003	166	166	332	0.488	0.866	0.715	0.65	0.75	1.333	1.1	0.05
0.8	0.8004	481	481	962	0.488	0.866	0.780	0.65	0.75	1.333	1.2	0.05

* 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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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.

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

Sample Size								Power		
N1	N2	N	P2	R0.L	R0.U	R1	Target Alpha	F.M. Score	M.N. Score	G.N. Score
50	50	100	0.65	0.75	1.333	1	0.05	0.2135	0.2135	0.2135
100	100	200	0.65	0.75	1.333	1	0.05	0.7108	0.7108	0.7108
150	150	300	0.65	0.75	1.333	1	0.05	0.9064	0.9064	0.9064
200	200	400	0.65	0.75	1.333	1	0.05	0.9715	0.9714	0.9714

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

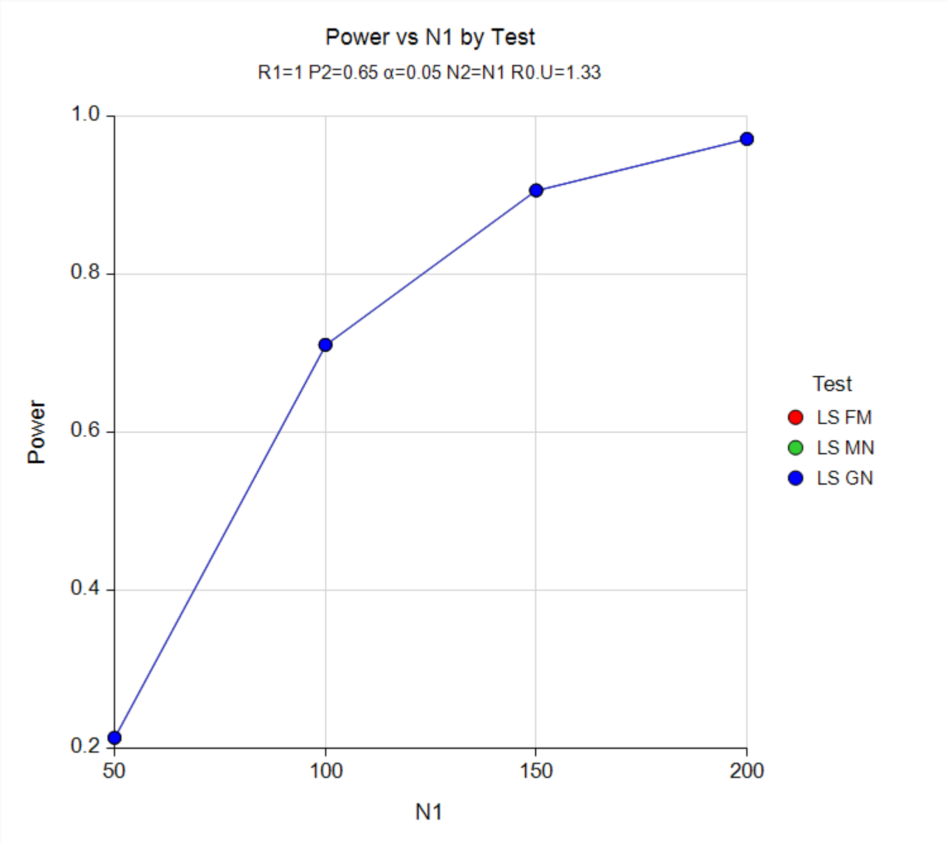
Actual Alpha Comparison of Three Different Tests

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

Sample Size							Alpha			
N1	N2	N	P2	R0.L	R0.U	R1	Target	F.M. Score	M.N. Score	G.N. Score
50	50	100	0.65	0.75	1.333	1	0.05	0.0516	0.0516	0.0516
100	100	200	0.65	0.75	1.333	1	0.05	0.0509	0.0509	0.0509
150	150	300	0.65	0.75	1.333	1	0.05	0.0510	0.0508	0.0508
200	200	400	0.65	0.75	1.333	1	0.05	0.0505	0.0500	0.0502

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

Plots



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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 4** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

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.

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

Sample Size							Normal Approximation		Binomial Enumeration	
N1	**N2**	**N**	**P2**	**R0.L**	**R0.U**	**R1**	**Power**	**Alpha**	**Power**	**Alpha**
50	50	100	0.65	0.75	1.333	1	0.2089	0.05	0.2135	0.0516
100	100	200	0.65	0.75	1.333	1	0.7120	0.05	0.7108	0.0509
150	150	300	0.65	0.75	1.333	1	0.9060	0.05	0.9064	0.0510
200	200	400	0.65	0.75	1.333	1	0.9715	0.05	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.