

## Chapter 213

# Equivalence Tests for the Difference Between Two Proportions

---

### Introduction

This module provides power analysis and sample size calculation for equivalence tests of the difference 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 eight 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 of the new treatment is between 0.63 and 0.77, the new treatment would be adopted. The *margin of equivalence* is 0.07.

The developers must design an experiment to test the hypothesis that the response rate of the new treatment does not differ from that of the standard treatment by more than 0.07. The statistical hypothesis to be tested is

$$H_0 : |p_1 - p_2| > 0.07 \text{ versus } H_1 : |p_1 - p_2| \leq 0.07$$

---

### 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.

Approximate sample size formulas for equivalence tests for the difference between two proportions are presented in Julius and Campbell (2012), section 3.4. Only large sample (normal approximation) results are given there. It is

## Equivalence Tests for the Difference Between Two Proportions

also possible to calculate power based on the enumeration of all possible values in the binomial distribution. Both options are available in this procedure.

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  $\delta$  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

$$H_0: p_{1,0} - p_2 \leq \delta_L \quad \text{or} \quad p_{1,0} - p_2 \geq \delta_U \quad \text{versus} \quad H_1: \delta_L \leq p_{1,0} - p_2 \leq \delta_U$$

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

| <u>Parameter</u> | <u>Computation</u>           | <u>Alternative Hypotheses</u>                    |
|------------------|------------------------------|--|
| 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       | $\psi = Odds_{1,0} / Odds_2$ | $H_1: \psi_L \leq o_{1,0} / o_2 \leq \psi_U$     |

## Equivalence Tests for the Difference Between Two Proportions

### Difference

The difference is perhaps the most direct method of comparison between two proportions. It is easy to interpret and communicate. It gives the absolute impact of the treatment. However, there are subtle difficulties that can arise with its interpretation.

One difficulty arises when the event of interest is rare. If a difference of 0.001 occurs when the baseline probability is 0.40, it would be dismissed as being trivial. However, if the baseline probability of a disease is 0.002, a 0.001 decrease would represent a reduction of 50%. Thus interpretation of the difference depends on the baseline probability of the event.

Note that  $\delta_L < 0$  and  $\delta_U > 0$ . Usually,  $\delta_L = -\delta_U$ .

### Equivalence using a Difference

The following example might help you understand the concept of an *equivalence* test. Suppose 60% of patients respond to the current treatment method ( $p_2 = 0.60$ ). If the response rate of the new treatment is no less than five percentage points better or worse than the existing treatment, it will be considered to be equivalent. Substituting these figures into the statistical hypotheses gives

$$H_0: p_{1.0} - p_2 \leq -0.05 \text{ or } p_{1.0} - p_2 \geq 0.05 \text{ versus } H_1: -0.05 \leq p_{1.0} - p_2 \leq 0.05$$

Using the relationship

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

gives

$$H_0: p_{1.0} \leq 0.55 \text{ or } p_{1.0} \geq 0.65 \text{ versus } H_1: 0.55 \leq p_{1.0} \leq 0.65$$

In this example, when the null hypothesis is rejected, the concluded alternative is that the response rate is between 0.55 and 0.65.

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

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

## Equivalence Tests for the Difference Between Two Proportions

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  $\alpha_U$ .

### 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: 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.

### Z Test (Pooled)

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

$$z_t = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\hat{\sigma}_1}$$

## Equivalence Tests for the Difference Between Two Proportions

where

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

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

### Z Test (Unpooled)

This test statistic does not pool the two proportions in computing the standard error.

$$z_t = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\hat{\sigma}_2}$$

where

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

### Z Test with Continuity Correction (Pooled)

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

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

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

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

where  $F$  is -1 for lower-tailed hypotheses and 1 for upper-tailed hypotheses.

### Z Test with Continuity Correction (Unpooled)

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

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

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

where  $F$  is -1 for lower-tailed hypotheses and 1 for upper-tailed hypotheses.

## Equivalence Tests for the Difference Between Two Proportions

## T-Test

Because of a detailed, comparative study of the behavior of several tests, D'Agostino (1988) and Upton (1982) proposed using the usual two-sample t-test for testing whether the two proportions are equal. One substitutes a '1' for a success and a '0' for a failure in the usual, two-sample  $t$ -test formula.

## Miettinen and Nurminen's Likelihood Score Test

Miettinen and Nurminen (1985) proposed a test statistic for testing whether the difference is equal to a specified, non-zero, value,  $\delta_0$ . The regular MLE's,  $\hat{p}_1$  and  $\hat{p}_2$ , are used in the numerator of the score statistic while MLE's  $\tilde{p}_1$  and  $\tilde{p}_2$ , constrained so that  $\tilde{p}_1 - \tilde{p}_2 = \delta_0$ , are used in the denominator. 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 this test statistic is

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

where

$$\hat{\sigma}_{MND} = \sqrt{\left( \frac{\tilde{p}_1 \tilde{q}_1}{n_1} + \frac{\tilde{p}_2 \tilde{q}_2}{n_2} \right) \left( \frac{N}{N-1} \right)}$$

$$\tilde{p}_1 = \tilde{p}_2 + \delta_0$$

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

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

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

$$C = \frac{L_2^3}{27L_3^3} - \frac{L_1 L_2}{6L_3^2} + \frac{L_0}{2L_3}$$

$$L_0 = x_{21} \delta_0 (1 - \delta_0)$$

$$L_1 = [n_2 \delta_0 - N - 2x_{21}] \delta_0 + m_1$$

$$L_2 = (N + n_2) \delta_0 - N - m_1$$

$$L_3 = N$$

## Equivalence Tests for the Difference Between Two Proportions

### Farrington and Manning's Likelihood Score Test

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

The formula for computing the test statistic is

$$z_{FMD} = \frac{\hat{p}_1 - \hat{p}_2 - \delta_0}{\sqrt{\left(\frac{\tilde{p}_1 \tilde{q}_1}{n_1} + \frac{\tilde{p}_2 \tilde{q}_2}{n_2}\right)}}$$

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

$$(-\tilde{\gamma})z_{GND}^2 + (-1)z_{GND} + (z_{FMD}(\delta) + \tilde{\gamma}) = 0$$

where

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

## 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 differences, 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.

## Equivalence Tests for the Difference Between Two Proportions

---

### 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  $N_1$  and  $N_2 \leq$  Maximum  $N_1$  or  $N_2$  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 $N_1$ or $N_2$ for Binomial Enumeration

When both  $N_1$  and  $N_2$  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  $N_1$  or  $N_2$  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.

---

### Test

#### Test Type

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

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



## Equivalence Tests for the Difference Between Two Proportions

---

### 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.

## Equivalence Tests for the Difference Between Two Proportions

### N1 (Sample Size, Group 1)

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

$N1$  is the number of items or individuals sampled from the Group 1 population.

$N1$  must be  $\geq 2$ . 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  $\geq 2$ . 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 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  $N2$ .

## Equivalence Tests for the Difference Between Two Proportions

- **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  $\geq 2$ . 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  $\geq 2$ . 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  $\geq 2$ . 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.

## Equivalence Tests for the Difference Between Two Proportions

### Percent in Group 1

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

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

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

### Effect Size

#### Input Type

Indicate what type of values to enter to specify the equivalence and actual differences. Regardless of the entry type chosen, the calculations are the same. This option is simply given for convenience in specifying the difference.

### Effect Size – Equivalence Differences

*These options are displayed only if Input Type = "Differences"*

#### D0.U & D0.L (Upper & Lower Equivalence Difference)

Specify the *margin of equivalence* by specifying the largest distance above (D0.U) and below (D0.L) P2 which will still result in the conclusion of equivalence. As long as the actual difference is between these two values, the difference is not considered to be large enough to be of practical importance.

The values of D0.U must be positive and the values of D0.L must be negative. D0.L can be set to '-D0.U,' which is usually what is 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:  $P1.0U = D0.U + P2$ .

You may enter a range of values for D0.U such as .03 .05 .10 or .05 to .20 by .05.

Note that if you enter values for D0.L (other than '-D0.U'), they are used in pairs with the values of D0.U. Thus, the first values of D0.U and D0.L are used together, then the second values of each are used, and so on.

RANGE:

D0.L must be between -1 and 0. D0.U must be between 0 and 1. Neither can take on the values -1, 0, or 1.

### Effect Size – Actual Difference

*This option is displayed only if Input Type = "Differences"*

#### D1 (Actual Difference)

This option specifies the actual difference between P1.1 (the actual value of P1) and P2. This is the value of the difference at which the power is calculated. In equivalence trials, this difference is often set to 0.

The power calculations assume that P1.1 is the actual value of the proportion in group 1 (experimental or treatment group). This difference is used with P2 to calculate the true value of P1 using the formula:  $P1.1 = D1 + P2$ .

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

## Equivalence Tests for the Difference Between Two Proportions

---

### Effect Size – Equivalence Proportions

*These options are displayed only if Input Type = “Proportions”*

#### **P1.0U & P1.0L (Upper & Lower Equivalence Proportion)**

Specify the *margin of equivalence* directly by giving the upper and lower bounds of P1.0. The two groups are assumed to be equivalent when P1.0 is between these values. Thus, P1.0U should be greater than P2 and P1.0L should be less than P2.

Note that the values of P1.0U and P1.0L are used in pairs. Thus, the first values of P1.0U and P1.0L are used together, then the second values of each are used, and so on.

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

Proportions must be between 0 and 1. They cannot take on the values 0 or 1. These values should surround P2.

---

### Effect Size – Actual Proportion

*This option is displayed only if Input Type = “Proportions”*

#### **P1.1 (Actual Proportion)**

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

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

---

### 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*.

## Equivalence Tests for the Difference Between Two Proportions

### 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 50% 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 its cure rate is within 15% of the standard treatment.

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 difference inside 15% when the actual cure rate of the new treatment ranges from 50% to 60%. 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 Difference Between Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Difference Between 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**.

| <u>Option</u>                             | <u>Value</u>                                |
|---|---|
| <b>Design Tab</b>                         |   |
| Solve For .....                           | <b>Power</b>                                |
| Power Calculation Method.....             | <b>Normal Approximation</b>                 |
| Test Type .....                           | <b>Likelihood Score (Farr. &amp; Mann.)</b> |
| Alpha.....                                | <b>0.05</b>                                 |
| Group Allocation .....                    | <b>Equal (N1 = N2)</b>                      |
| Sample Size Per Group .....               | <b>50 to 500 by 50</b>                      |
| Input Type .....                          | <b>Differences</b>                          |
| D0.U (Upper Equivalence Difference) ....  | <b>0.15</b>                                 |
| D0.L (Lower Equivalence Difference) ..... | <b>-D0.U</b>                                |
| D1 (Actual Difference) .....              | <b>0.00 0.05 0.10</b>                       |
| P2 (Group 2 Proportion) .....             | <b>0.5</b>                                  |

## Equivalence Tests for the Difference Between Two Proportions

## Output

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

## Numeric Results

## Numeric Results for Equivalence Tests for the Difference Between Two Proportions

Test Statistic: Farrington & Manning Likelihood Score Test

H0:  $P1 - P2 \leq D0.L$  or  $P1 - P2 \geq D0.U$ . H1:  $D0.L < P1 - P2 = D1 < D0.U$ .

| Power* | N1  | N2  | N   | Ref.<br>P2 | P1.0L | P1.0U | D0.L   | D0.U  | D1    | Alpha  |
|--------|-----|-----|-----|------------|-------|-------|--------|-------|-------|--------|
| 0.0000 | 50  | 50  | 100 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.3795 | 100 | 100 | 200 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.6689 | 150 | 150 | 300 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.8305 | 200 | 200 | 400 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.9160 | 250 | 250 | 500 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.9594 | 300 | 300 | 600 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.9808 | 350 | 350 | 700 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.9911 | 400 | 400 | 800 | 0.500      | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 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. This is 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.

D0.L is the lowest difference that still results in the conclusion of equivalence. D0.U is the highest difference that still results in the conclusion of equivalence. D1 is the actual difference,  $P1 - P2$ , 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.

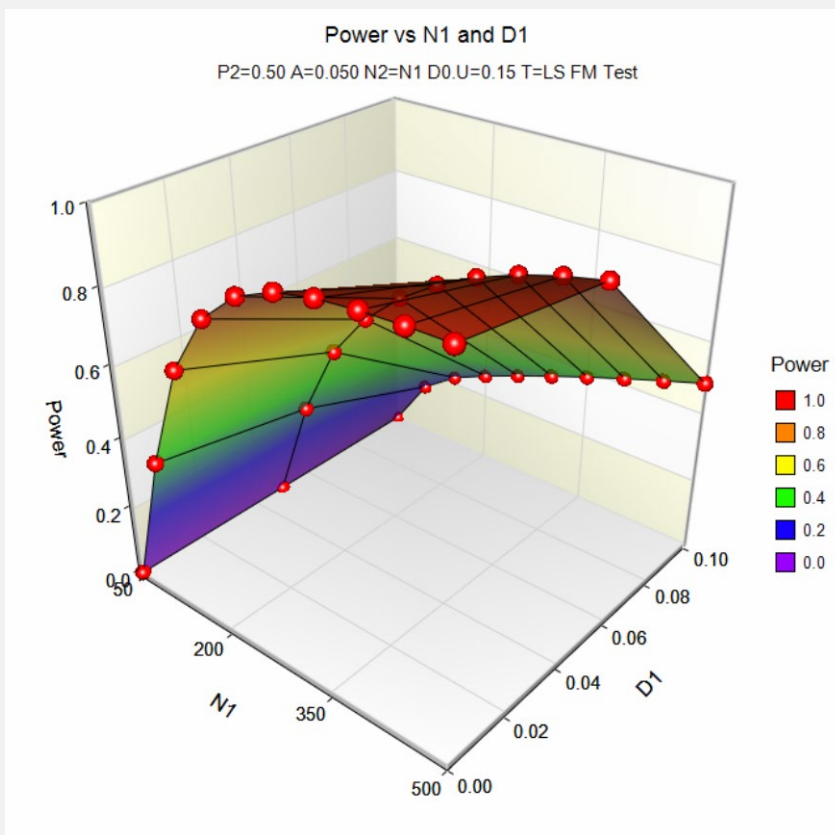
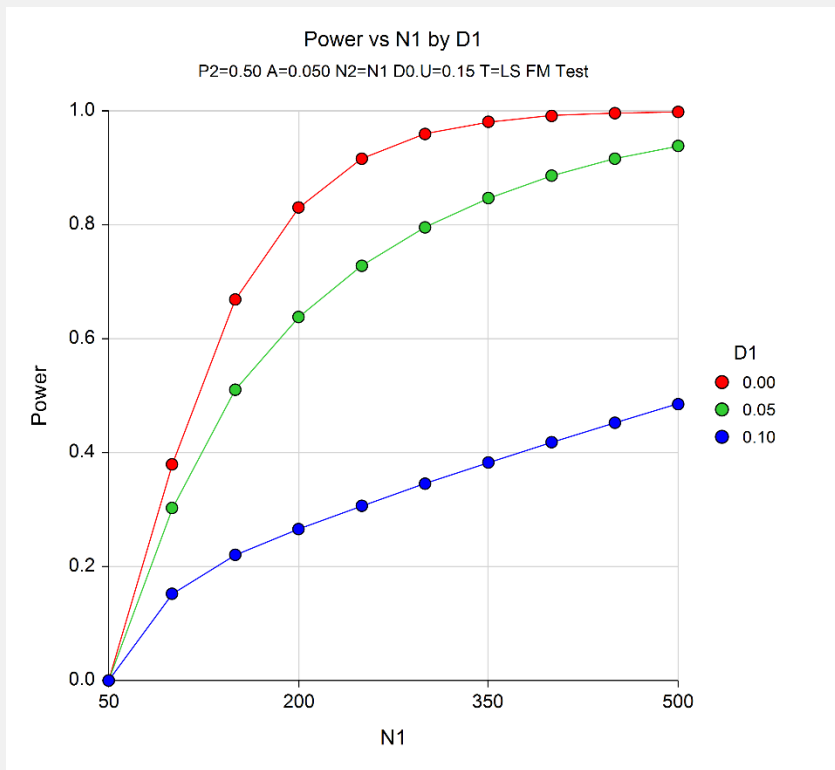
## Summary Statements

Sample sizes of 50 in the treatment group and 50 in the reference group achieve 0.00% power to detect equivalence. The margin of equivalence, given in terms of the difference, extends from -0.150 to 0.150. The actual difference is 0.000. The reference group proportion is 0.500. 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.

## Equivalence Tests for the Difference Between Two Proportions

### 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 D1.



## Equivalence Tests for the Difference Between 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 D1 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 Difference Between Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Difference Between 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>                                |
|--|---|
| <b>Design Tab</b>                        |   |
| Solve For .....                          | <b>Sample Size</b>                          |
| Power Calculation Method.....            | <b>Normal Approximation</b>                 |
| Test Type.....                           | <b>Likelihood Score (Farr. &amp; Mann.)</b> |
| Power.....                               | <b>0.80</b>                                 |
| Alpha.....                               | <b>0.05</b>                                 |
| Group Allocation .....                   | <b>Equal (N1 = N2)</b>                      |
| Input Type.....                          | <b>Differences</b>                          |
| D0.U (Upper Equivalence Difference) .... | <b>0.15</b>                                 |
| D0.L (Lower Equivalence Difference)..... | <b>-D0.U</b>                                |
| D1 (Actual Difference) .....             | <b>0.00 0.05 0.10</b>                       |
| P2 (Group 2 Proportion) .....            | <b>0.5</b>                                  |

### Output

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

### Numeric Results

#### Numeric Results for Equivalence Tests for the Difference Between Two Proportions

Test Statistic: Farrington & Manning Likelihood Score Test

H0:  $P1 - P2 \leq D0.L$  or  $P1 - P2 \geq D0.U$ . H1:  $D0.L < P1 - P2 = D1 < D0.U$ .

| Target Power | Actual Power* | N1   | N2   | N    | Ref. P2 | P1.0L | P1.0U | D0.L   | D0.U  | D1    | Alpha  |
|--------------|---------------|------|------|------|---------|-------|-------|--------|-------|-------|--------|
| 0.80         | 0.8003        | 188  | 188  | 376  | 0.500   | 0.350 | 0.650 | -0.150 | 0.150 | 0.000 | 0.0500 |
| 0.80         | 0.8001        | 304  | 304  | 608  | 0.500   | 0.350 | 0.650 | -0.150 | 0.150 | 0.050 | 0.0500 |
| 0.80         | 0.8001        | 1202 | 1202 | 2404 | 0.500   | 0.350 | 0.650 | -0.150 | 0.150 | 0.100 | 0.0500 |

\* Power was computed using the normal approximation method.

The required sample size will depend a great deal on the value of D1. Any effort spent determining an accurate value for D1 will be worthwhile.

## Example 3 – Comparing the Power of Several Test Statistics

Continuing with Example 2, the researchers want to determine which of the eight possible test statistics to adopt by using the comparative reports and charts that **PASS** produces. They decide to compare the powers and actual alphas for various sample sizes between 50 and 200 when  $D1$  is 0.1.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Equivalence Tests for the Difference Between Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Difference Between 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>                                |
|--|---|
| <b>Design Tab</b>                        |   |
| Solve For .....                          | <b>Power</b>                                |
| Power Calculation Method.....            | <b>Binomial Enumeration</b>                 |
| Max N1 or N2 for Binomial Enumeration    | <b>5000</b>                                 |
| Zero Count Adjustment Method.....        | <b>Add to zero cells only</b>               |
| Zero Count Adjustment Value .....        | <b>0.0001</b>                               |
| Test Type.....                           | <b>Likelihood Score (Farr. &amp; Mann.)</b> |
| Alpha.....                               | <b>0.05</b>                                 |
| Group Allocation .....                   | <b>Equal (N1 = N2)</b>                      |
| Sample Size Per Group.....               | <b>50 to 200 by 50</b>                      |
| Input Type.....                          | <b>Differences</b>                          |
| D0.U (Upper Equivalence Difference) .... | <b>0.15</b>                                 |
| D0.L (Lower Equivalence Difference)..... | <b>-D0.U</b>                                |
| D1 (Actual Difference) .....             | <b>0.10</b>                                 |
| P2 (Group 2 Proportion) .....            | <b>0.5</b>                                  |
| <b>Reports Tab</b>                       |   |
| Show Comparative Reports.....            | <b>Checked</b>                              |
| <b>Comparative Plots Tab</b>             |   |
| Show Comparative Plots .....             | <b>Checked</b>                              |

Equivalence Tests for the Difference Between Two Proportions

Output

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

Numeric Results and Plots

Power Comparison of Eight Different Equivalence Tests for the Difference Between Two Proportions

H0:  $P_1 - P_2 \leq D_0.L$  or  $P_1 - P_2 \geq D_0.U$ . H1:  $D_0.L < P_1 - P_2 = D_1 < D_0.U$ .

| N1/N2   | P2     | D0.U   | Target Alpha | Z(P) Test Power | Z(UnP) Test Power | Z(P) CC Test Power | Z(UnP) CC Test Power | T Test Power | F.M. Score Power | M.N. Score Power | G.N. Score Power |
|---------|--------|--------|--------------|-----------------|-------------------|--------------------|----------------------|--------------|------------------|------------------|------------------|
| 50/50   | 0.5000 | 0.1500 | 0.0500       | 0.0000          | 0.0000            | 0.0000             | 0.0000               | 0.0000       | 0.0000           | 0.0000           | 0.0000           |
| 100/100 | 0.5000 | 0.1500 | 0.0500       | 0.1494          | 0.1494            | 0.1047             | 0.1047               | 0.1493       | 0.1495           | 0.1494           | 0.1494           |
| 150/150 | 0.5000 | 0.1500 | 0.0500       | 0.2208          | 0.2208            | 0.1863             | 0.1863               | 0.2208       | 0.2208           | 0.2208           | 0.2208           |
| 200/200 | 0.5000 | 0.1500 | 0.0500       | 0.2552          | 0.2553            | 0.2238             | 0.2239               | 0.2551       | 0.2566           | 0.2566           | 0.2560           |

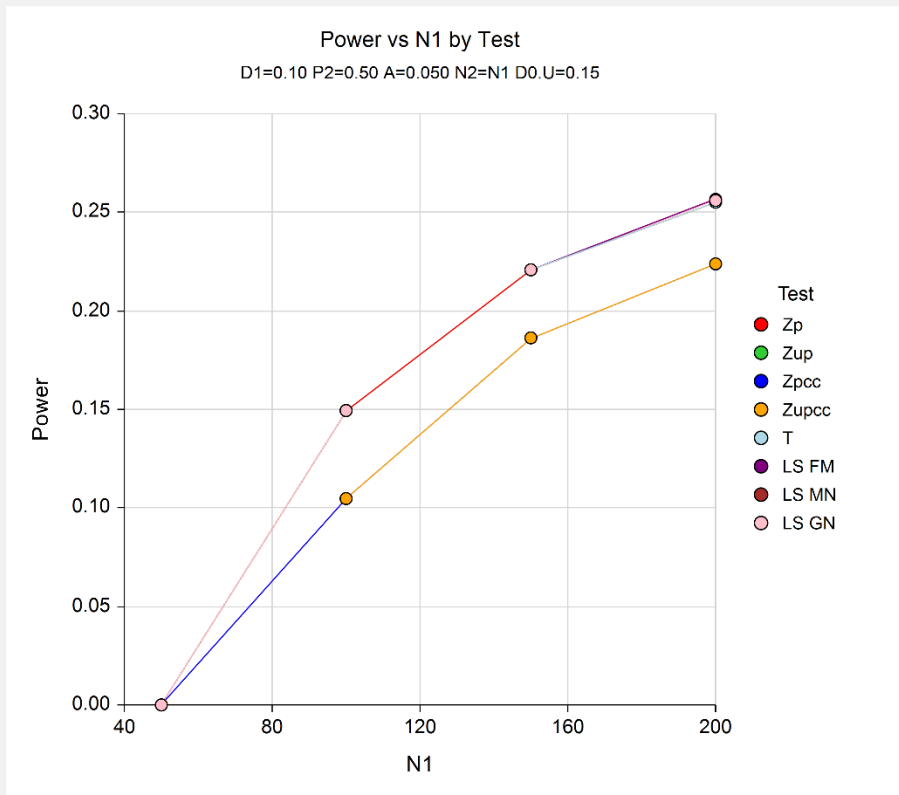
Note: Power was computed using binomial enumeration of all possible outcomes.

Actual Alpha Comparison of Eight Different Equivalence Tests for the Difference Between Two Proportions

H0:  $P_1 - P_2 \leq D_0.L$  or  $P_1 - P_2 \geq D_0.U$ . H1:  $D_0.L < P_1 - P_2 = D_1 < D_0.U$ .

| N1/N2   | P2     | D0.U   | Target Alpha | Z(P) Test Alpha | Z(UnP) Test Alpha | Z(P) CC Test Alpha | Z(UnP) CC Test Alpha | T Test Alpha | F.M. Score Alpha | M.N. Score Alpha | G.N. Score Alpha |
|---------|--------|--------|--------------|-----------------|-------------------|--------------------|----------------------|--------------|------------------|------------------|------------------|
| 50/50   | 0.5000 | 0.1500 | 0.0500       | 0.0515          | 0.0515            | 0.0334             | 0.0334               | 0.0514       | 0.0515           | 0.0515           | 0.0515           |
| 100/100 | 0.5000 | 0.1500 | 0.0500       | 0.0486          | 0.0486            | 0.0358             | 0.0358               | 0.0485       | 0.0489           | 0.0487           | 0.0487           |
| 150/150 | 0.5000 | 0.1500 | 0.0500       | 0.0495          | 0.0495            | 0.0386             | 0.0386               | 0.0495       | 0.0495           | 0.0495           | 0.0495           |
| 200/200 | 0.5000 | 0.1500 | 0.0500       | 0.0465          | 0.0468            | 0.0376             | 0.0378               | 0.0464       | 0.0488           | 0.0488           | 0.0481           |

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



It is interesting to note that the powers of the continuity-corrected test statistics are consistently lower than the other tests. This occurs because the actual alpha achieved by these tests is lower than for the other tests. An interesting finding of this example is that the regular *t*-test performed about as well as the *z*-test.

## Equivalence Tests for the Difference Between 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 **Equivalence Tests for the Difference Between Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Difference Between 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>                                |
|--|---|
| <b>Design Tab</b>                        |   |
| Solve For .....                          | <b>Power</b>                                |
| Power Calculation Method.....            | <b>Normal Approximation</b>                 |
| Test Type.....                           | <b>Likelihood Score (Farr. &amp; Mann.)</b> |
| Alpha.....                               | <b>0.05</b>                                 |
| Group Allocation .....                   | <b>Equal (N1 = N2)</b>                      |
| Sample Size Per Group.....               | <b>50 to 200 by 50</b>                      |
| Input Type.....                          | <b>Differences</b>                          |
| D0.U (Upper Equivalence Difference) .... | <b>0.15</b>                                 |
| D0.L (Lower Equivalence Difference)..... | <b>-D0.U</b>                                |
| D1 (Actual Difference) .....             | <b>0.10</b>                                 |
| P2 (Group 2 Proportion) .....            | <b>0.5</b>                                  |
| <b>Reports Tab</b>                       |   |
| Show Power Detail Report .....           | <b>Checked</b>                              |

**Output**

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

**Numeric Results and Plots****Power Detail Report for Equivalence Tests for the Difference Between Two Proportions**

Test Statistic: Farrington & Manning Likelihood Score Test

H0:  $P1 - P2 \leq D0.L$  or  $P1 - P2 \geq D0.U$ . H1:  $D0.L < P1 - P2 = D1 < D0.U$ .

| N1/N2   | P2     | D0.U  | D1    | Normal Approximation |        | Binomial Enumeration |        |
|---------|--------|-------|-------|----------------------|--------|----------------------|--------|
|         |        |       |       | Power                | Alpha  | Power                | Alpha  |
| 50/50   | 0.5000 | 0.150 | 0.100 | 0.0000               | 0.0500 | 0.0000               | 0.0515 |
| 100/100 | 0.5000 | 0.150 | 0.100 | 0.1523               | 0.0500 | 0.1495               | 0.0489 |
| 150/150 | 0.5000 | 0.150 | 0.100 | 0.2206               | 0.0500 | 0.2208               | 0.0495 |
| 200/200 | 0.5000 | 0.150 | 0.100 | 0.2659               | 0.0500 | 0.2566               | 0.0488 |

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

## Example 5 – Computing the Power after Completing an Experiment

Researchers are testing a generic drug to determine if it is equivalent to the name-brand alternative. Equivalence is declared if the success rate of the generic brand is no more than 5% from that of the name-brand drug. In a study with 1000 individuals in each group, they find that 774, or 77.4%, are successfully treated using the name-brand drug, and 700, or 70%, respond to the generic drug. An equivalence test (exact test) with  $\alpha = 0.05$  failed to declare that the two drugs are equivalent. The researchers would now like to compute the power for actual differences ranging from 0 to 4%. Suppose that the true value for the response rate for the name-brand drug is 77%.

Note that the power is not calculated at the difference observed in the study, 77.4%. In fact, the difference observed in the study is larger than the proposed equivalence difference, 5%. It would make no sense to perform a power calculation for a difference larger than the equivalence difference. It is more informative to study a range of values smaller than or equal to the equivalence difference.

### 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 Difference Between Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Difference Between Two Proportions**. You may then make the appropriate entries as listed below, or open **Example 5** by going to the **File** menu and choosing **Open Example Template**.

| <u>Option</u>                            | <u>Value</u>                                |
|--|---|
| <b>Design Tab</b>                        |   |
| Solve For .....                          | <b>Power</b>                                |
| Power Calculation Method.....            | <b>Normal Approximation</b>                 |
| Test Type.....                           | <b>Likelihood Score (Farr. &amp; Mann.)</b> |
| Alpha.....                               | <b>0.05</b>                                 |
| Group Allocation .....                   | <b>Equal (N1 = N2)</b>                      |
| Sample Size Per Group.....               | <b>1000</b>                                 |
| Input Type.....                          | <b>Differences</b>                          |
| D0.U (Upper Equivalence Difference) .... | <b>0.05</b>                                 |
| D0.L (Lower Equivalence Difference)..... | <b>-D0.U</b>                                |
| D1 (Actual Difference) .....             | <b>0.00 to 0.04 by 0.01</b>                 |
| P2 (Group 2 Proportion) .....            | <b>0.77</b>                                 |

## Equivalence Tests for the Difference Between Two Proportions

## Output

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

## Numeric Results

## Numeric Results for Equivalence Tests for the Difference Between Two Proportions

Test Statistic: Farrington & Manning Likelihood Score Test

H0:  $P_1 - P_2 \leq D_0.L$  or  $P_1 - P_2 \geq D_0.U$ . H1:  $D_0.L < P_1 - P_2 = D_1 < D_0.U$ .

| Power* | N1   | N2   | N    | Ref.<br>P2 | P1.0L | P1.0U | D0.L   | D0.U  | D1    | Alpha  |
|--------|------|------|------|------------|-------|-------|--------|-------|-------|--------|
| 0.6875 | 1000 | 1000 | 2000 | 0.770      | 0.720 | 0.820 | -0.050 | 0.050 | 0.000 | 0.0500 |
| 0.6313 | 1000 | 1000 | 2000 | 0.770      | 0.720 | 0.820 | -0.050 | 0.050 | 0.010 | 0.0500 |
| 0.4731 | 1000 | 1000 | 2000 | 0.770      | 0.720 | 0.820 | -0.050 | 0.050 | 0.020 | 0.0500 |
| 0.2857 | 1000 | 1000 | 2000 | 0.770      | 0.720 | 0.820 | -0.050 | 0.050 | 0.030 | 0.0500 |
| 0.1362 | 1000 | 1000 | 2000 | 0.770      | 0.720 | 0.820 | -0.050 | 0.050 | 0.040 | 0.0500 |

\* Power was computed using the normal approximation method.

The power of the test ranges from 68.75% if the true difference is actually 0.0% to 13.62% if the true difference is 4%.

## Example 6 – Finding the Sample Size using Proportions

A study is being designed to prove the equivalence of a new drug to the current standard. The current drug is effective in 85% of cases. The new drug, however, is cheaper to produce. The new drug will be deemed equivalent to the standard if its success rate is between 78% and 92%. What sample sizes are necessary to obtain 80% or 90% power for actual success rates ranging from 80% to 90%? The researchers will test at a significance level of 0.05 using the Farrington and Manning likelihood score test.

## Setup

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

| Option                                  | Value                                       |
|---|---|
| <b>Design Tab</b>                       |   |
| Solve For .....                         | <b>Sample Size</b>                          |
| Power Calculation Method.....           | <b>Normal Approximation</b>                 |
| Test Type.....                          | <b>Likelihood Score (Farr. &amp; Mann.)</b> |
| Power.....                              | <b>0.80 0.90</b>                            |
| Alpha.....                              | <b>0.05</b>                                 |
| Group Allocation .....                  | <b>Equal (N1 = N2)</b>                      |
| Input Type.....                         | <b>Proportions</b>                          |
| P1.0U (Upper Equivalence Proportion) .. | <b>0.92</b>                                 |
| P1.0L (Lower Equivalence Proportion)... | <b>0.78</b>                                 |
| P1.1 (Actual Proportion) .....          | <b>0.80 to 0.90 by 0.02</b>                 |
| P2 (Group 2 Proportion) .....           | <b>0.85</b>                                 |

## Equivalence Tests for the Difference Between Two Proportions

### Output

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

### Numeric Results

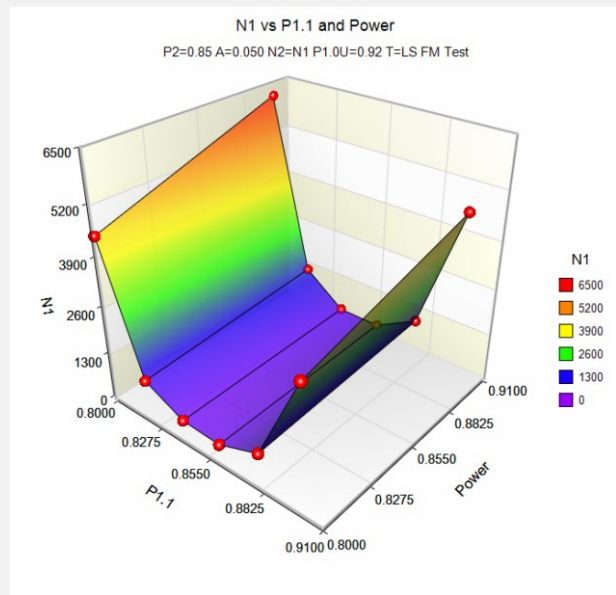
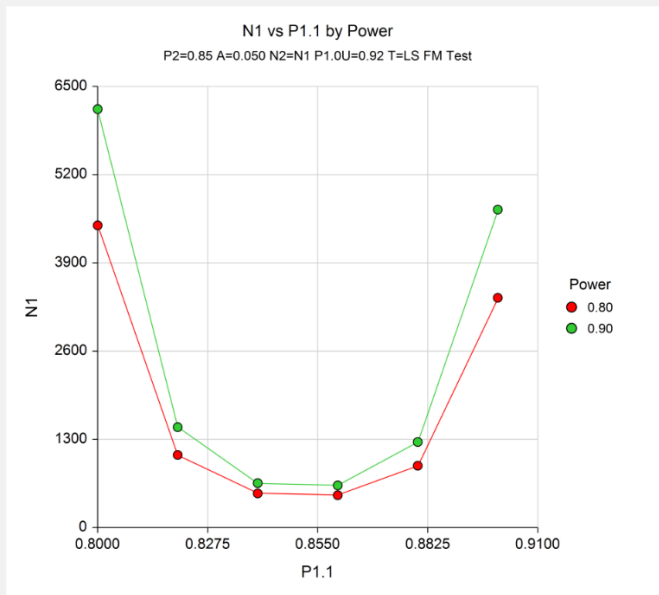
#### Numeric Results for Equivalence Tests for the Difference Between Two Proportions

Test Statistic: Farrington & Manning Likelihood Score Test

H0:  $P_1 - P_2 \leq D_{0.L}$  or  $P_1 - P_2 \geq D_{0.U}$ . H1:  $D_{0.L} < P_1 - P_2 = D_1 < D_{0.U}$ .

| Target Power | Actual Power* | N1   | N2   | N     | Ref. P2 | P1.0L | P1.0U | D0.L   | D0.U  | D1     | Alpha  |
|--------------|---------------|------|------|-------|---------|-------|-------|--------|-------|--------|--------|
| 0.80         | 0.8001        | 4453 | 4453 | 8906  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | -0.050 | 0.0500 |
| 0.90         | 0.9000        | 6166 | 6166 | 12332 | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | -0.050 | 0.0500 |
| 0.80         | 0.8002        | 1070 | 1070 | 2140  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | -0.030 | 0.0500 |
| 0.90         | 0.9000        | 1480 | 1480 | 2960  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | -0.030 | 0.0500 |
| 0.80         | 0.8008        | 503  | 503  | 1006  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | -0.010 | 0.0500 |
| 0.90         | 0.9001        | 655  | 655  | 1310  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | -0.010 | 0.0500 |
| 0.80         | 0.8004        | 477  | 477  | 954   | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | 0.010  | 0.0500 |
| 0.90         | 0.9004        | 622  | 622  | 1244  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | 0.010  | 0.0500 |
| 0.80         | 0.8002        | 912  | 912  | 1824  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | 0.030  | 0.0500 |
| 0.90         | 0.9002        | 1261 | 1261 | 2522  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | 0.030  | 0.0500 |
| 0.80         | 0.8000        | 3386 | 3386 | 6772  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | 0.050  | 0.0500 |
| 0.90         | 0.9000        | 4685 | 4685 | 9370  | 0.850   | 0.780 | 0.920 | -0.070 | 0.070 | 0.050  | 0.0500 |

\* Power was computed using the normal approximation method.



It is evident from these results that the sample sizes required to achieve 80% and 90% power depend a great deal on the actual value of the success rate, P1.1.

## Example 7 – Validation of Sample Size Calculation for the Unpooled Z-Test using Julius and Campbell (2012)

Julius and Campbell (2012) presents Table XVI that gives the results of sample size calculations for an unpooled Z-test for equivalence with  $P_2$  between 0.7 and 0.9,  $|D_0|$  between 0.05 and 0.20 and  $D_1$  between -0.05 and 0.05. Sample sizes are calculated for 90% power and  $\alpha = 0.025$ . This example will replicate all values of  $D_1$  for  $P_2 = 0.70$  and  $|D_0| = 0.20$  in the table.

The sample sizes reported in the table for  $D_1$  between -0.05 and 0.05 are 205, 180, 161, 148, 140, 137, 138, 143, 152, 167, and 186.

### 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 Difference Between Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Equivalence**, and then clicking on **Equivalence Tests for the Difference Between Two Proportions**. You may then make the appropriate entries as listed below, or open **Example 7** by going to the **File** menu and choosing **Open Example Template**.

| <u>Option</u>                            | <u>Value</u>                 |
|--|------------------------------|
| <b>Design Tab</b>                        |                              |
| Solve For .....                          | <b>Sample Size</b>           |
| Power Calculation Method.....            | <b>Normal Approximation</b>  |
| Test Type.....                           | <b>Z-Test (Unpooled)</b>     |
| Power.....                               | <b>0.90</b>                  |
| Alpha.....                               | <b>0.025</b>                 |
| Group Allocation .....                   | <b>Equal (N1 = N2)</b>       |
| Input Type.....                          | <b>Differences</b>           |
| D0.U (Upper Equivalence Difference) .... | <b>0.20</b>                  |
| D0.L (Lower Equivalence Difference)..... | <b>-D0.U</b>                 |
| D1 (Actual Difference) .....             | <b>-0.05 to 0.05 by 0.01</b> |
| P2 (Group 2 Proportion) .....            | <b>0.7</b>                   |



## Equivalence Tests for the Difference Between Two Proportions

## Output

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

## Numeric Results

## Numeric Results for Equivalence Tests for the Difference Between Two Proportions

Test Statistic: Z-Test with Unpooled Variance

H0:  $P1 - P2 \leq D0.L$  or  $P1 - P2 \geq D0.U$ . H1:  $D0.L < P1 - P2 = D1 < D0.U$ .

| Target Power | Actual Power* | N1  | N2  | N   | Ref. P2 | P1.0L | P1.0U | D0.L   | D0.U  | D1     | Alpha  |
|--------------|---------------|-----|-----|-----|---------|-------|-------|--------|-------|--------|--------|
| 0.90         | 0.9007        | 205 | 205 | 410 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | -0.050 | 0.0250 |
| 0.90         | 0.9010        | 180 | 180 | 360 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | -0.040 | 0.0250 |
| 0.90         | 0.9010        | 161 | 161 | 322 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | -0.030 | 0.0250 |
| 0.90         | 0.9011        | 148 | 148 | 296 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | -0.020 | 0.0250 |
| 0.90         | 0.9006        | 140 | 140 | 280 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | -0.010 | 0.0250 |
| 0.90         | 0.9015        | 137 | 137 | 274 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | 0.000  | 0.0250 |
| 0.90         | 0.9023        | 138 | 138 | 276 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | 0.010  | 0.0250 |
| 0.90         | 0.9024        | 143 | 143 | 286 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | 0.020  | 0.0250 |
| 0.90         | 0.9009        | 152 | 152 | 304 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | 0.030  | 0.0250 |
| 0.90         | 0.9014        | 167 | 167 | 334 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | 0.040  | 0.0250 |
| 0.90         | 0.9003        | 186 | 186 | 372 | 0.700   | 0.500 | 0.900 | -0.200 | 0.200 | 0.050  | 0.0250 |

\* Power was computed using the normal approximation method.

The sample sizes from **PASS** match Table XVI of Julius and Campbell (2012) exactly. You can replicate the other values in the table by changing the values for P2 and D0.U.

## Equivalence Tests for the Difference Between Two Proportions

## Example 8 – Validation of Sample Size Calculation for the Pooled Z-Test using Tuber-Bitter et al. (2000)

Tuber-Bitter et al. (2000), page 1271, present a sample size study in which  $P_2 = 0.1$ ;  $D_{0.U} = 0.01, 0.02, 0.03$ ;  $D_{0.L} = -D_{0.U}$ ;  $D_1 = 0.0$ ;  $\alpha = 0.05$ ; and  $\beta = 0.1$ . Using the pooled Z test statistic, they found the sample sizes to be 19484, 4871, and 2165 in each group.

### Setup

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

| <u>Option</u>                            | <u>Value</u>                |
|--|-----------------------------|
| <b>Design Tab</b>                        |                             |
| Solve For .....                          | <b>Sample Size</b>          |
| Power Calculation Method.....            | <b>Normal Approximation</b> |
| Test Type.....                           | <b>Z-Test (Pooled)</b>      |
| Power.....                               | <b>0.90</b>                 |
| Alpha.....                               | <b>0.05</b>                 |
| Group Allocation .....                   | <b>Equal (N1 = N2)</b>      |
| Input Type.....                          | <b>Differences</b>          |
| D0.U (Upper Equivalence Difference) .... | <b>0.01 0.02 0.03</b>       |
| D0.L (Lower Equivalence Difference)..... | <b>-D0.U</b>                |
| D1 (Actual Difference) .....             | <b>0.0</b>                  |
| P2 (Group 2 Proportion) .....            | <b>0.1</b>                  |

### Output

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

### Numeric Results

#### Numeric Results for Equivalence Tests for the Difference Between Two Proportions

Test Statistic: Z-Test with Pooled Variance

$H_0: P_1 - P_2 \leq D_{0.L}$  or  $P_1 - P_2 \geq D_{0.U}$ .  $H_1: D_{0.L} < P_1 - P_2 < D_{0.U}$ .

| Target Power | Actual Power* | N1    | N2    | N     | Ref. P2 | P1.0L | P1.0U | D0.L   | D0.U  | D1    | Alpha  |
|--------------|---------------|-------|-------|-------|---------|-------|-------|--------|-------|-------|--------|
| 0.90         | 0.9000        | 19480 | 19480 | 38960 | 0.100   | 0.090 | 0.110 | -0.010 | 0.010 | 0.000 | 0.0500 |
| 0.90         | 0.9000        | 4870  | 4870  | 9740  | 0.100   | 0.080 | 0.120 | -0.020 | 0.020 | 0.000 | 0.0500 |
| 0.90         | 0.9001        | 2165  | 2165  | 4330  | 0.100   | 0.070 | 0.130 | -0.030 | 0.030 | 0.000 | 0.0500 |

PASS found the required sample sizes to within rounding error of Tuber-Bitter et al. (2000).