Chapter 240

Equivalence Tests for the Difference of Two Proportions in a Cluster-Randomized Design

Introduction

This module provides power analysis and sample size calculation for equivalence tests of the difference in two-sample, cluster-randomized designs in which the outcome is binary.

Technical Details

Our formulation comes from Donner and Klar (2000). Denote a binary observation by Y_{gkm} where *g* = 1 or 2 is the group, *k* = 1, 2, ..., K_g is a cluster within group *g*, and *m* = 1, 2, ..., M_g is an individual in cluster *k* of group *g*. The results that follow assume an equal number of individuals per cluster. When the number of subjects from cluster to cluster are about the same, the power and sample size values should be fairly accurate. In these cases, the average number of subjects per cluster can be used.

The statistical hypothesis that is tested concerns the difference between the two group proportions, p_1 and p_2. When necessary, we assume that group 1 is the treatment group and group 2 is the control group. With a simple modification, all of the large-sample sample size formulas that are listed in the module for testing two proportions can be used here.

When the individual subjects are randomly assigned to one of the two groups, the variance of the sample proportion is

$$\sigma_{S,g}^2 = \frac{p_g(1-p_g)}{n_g}$$

When the randomization is by clusters of subjects, the variance of the sample proportion is

$$\begin{aligned}\sigma_{C,g}^2 &= \frac{p_g(1-p_g)\left(1+(m_g-1)\rho\right)}{k_g m_g} \\ &= \sigma_{S,g}^2\left[1+(m_g-1)\rho\right] \\ &= F_{g,\rho}\sigma_{S,g}^2\end{aligned}$$

The factor $[1 + (m_g - 1)\rho]$ is called the *inflation factor*. The Greek letter ρ is used to represent the *intracluster correlation coefficient (ICC)*. This correlation may be thought of as the simple correlation between any two subjects within the same cluster. If we stipulate that ρ is positive, it may also be interpreted as the proportion of total variability that is attributable to differences between clusters. This value is critical to the sample size calculation.

The asymptotic formulas that were used in comparing two proportions (see Chapter 213, "Equivalence Tests for the Difference Between Two Proportions") may be used with cluster-randomized designs as well, as long as an adjustment is made for the inflation factor.

Power Calculations

A large sample approximation may be used that is most accurate when the values of n_1 and n_2 are large. The large 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 and p_2 under the alternative hypothesis, and then computing the results based on the normal distribution.

Note that in this case, exact calculations are not possible.

Example 1 – Finding Power

A study is being designed to establish the equivalence of a new treatment compared to the current treatment. Historically, the standard treatment has enjoyed a 60% cure rate. The new treatment reduces the seriousness of certain side effects that occur with the standard treatment. Thus, the new treatment will be adopted even if it is slightly less effective than the standard treatment. The researchers will recommend adoption of the new treatment if its cure rate is within 0.15 of the standard treatment.

The researchers will recruit patients from various hospitals. All patients at a particular hospital will receive the same treatment. They anticipate enlisting an average of 50 patients per hospital. Based on similar studies, they estimate the intracluster correlation to be 0.002.

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 two, one-sided tests proposed by Farrington and Manning when the number of clusters per groups ranges from 2 to 10. They want to investigate the behavior of this test when the actual cure rate of the new treatment ranges from 60% to 66%. 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**
Test Type	**Likelihood Score (Farr. & Mann.)**
Alpha	**0.05**
K1 (Clusters in Group 1)	**2 4 6 8 10**
M1 (Average Cluster Size)	**50**
K2 (Clusters in Group 2)	**K1**
M2 (Average Cluster Size)	**M1**
Input Type	**Differences**
D0.U (Upper Equivalence Difference)	**0.15**
D0.L (Lower Equivalence Difference)	**-D0.U**
D1 (Actual Difference)	**0 0.03 0.06**
P2 (Group 2 Proportion)	**0.6**
ICC (Intracluster Correlation)	**0.002**

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: Likelihood Score Test (Farrington & Manning)
Hypotheses: H0: P1 - P2 ≤ D0.L or P1 - P2 ≥ D0.U vs. H1: D0.L < P1 - P2 < D0.U

	Number of Clusters			Average Cluster Size		Total Sample Size	Proportion			Difference		Intracluster Correlation	
							Equivalence	Actual	Reference	Equivalence	Actual		
Power	K1	K2	K	M1	M2	N	P1.L | P1.U	P1.1	P2	D0.L | D0.U	D1	ICC	Alpha
0.33835	2	2	4	50	50	200	0.45 | 0.75	0.60	0.6	-0.15 | 0.15	0.00	0.002	0.05
0.80415	4	4	8	50	50	400	0.45 | 0.75	0.60	0.6	-0.15 | 0.15	0.00	0.002	0.05
0.94885	6	6	12	50	50	600	0.45 | 0.75	0.60	0.6	-0.15 | 0.15	0.00	0.002	0.05
0.98771	8	8	16	50	50	800	0.45 | 0.75	0.60	0.6	-0.15 | 0.15	0.00	0.002	0.05
0.99722	10	10	20	50	50	1000	0.45 | 0.75	0.60	0.6	-0.15 | 0.15	0.00	0.002	0.05
0.32109	2	2	4	50	50	200	0.45 | 0.75	0.63	0.6	-0.15 | 0.15	0.03	0.002	0.05
0.73719	4	4	8	50	50	400	0.45 | 0.75	0.63	0.6	-0.15 | 0.15	0.03	0.002	0.05
0.89163	6	6	12	50	50	600	0.45 | 0.75	0.63	0.6	-0.15 | 0.15	0.03	0.002	0.05
0.95510	8	8	16	50	50	800	0.45 | 0.75	0.63	0.6	-0.15 | 0.15	0.03	0.002	0.05
0.98181	10	10	20	50	50	1000	0.45 | 0.75	0.63	0.6	-0.15 | 0.15	0.03	0.002	0.05
0.25939	2	2	4	50	50	200	0.45 | 0.75	0.66	0.6	-0.15 | 0.15	0.06	0.002	0.05
0.55419	4	4	8	50	50	400	0.45 | 0.75	0.66	0.6	-0.15 | 0.15	0.06	0.002	0.05
0.70910	6	6	12	50	50	600	0.45 | 0.75	0.66	0.6	-0.15 | 0.15	0.06	0.002	0.05
0.81306	8	8	16	50	50	800	0.45 | 0.75	0.66	0.6	-0.15 | 0.15	0.06	0.002	0.05
0.88235	10	10	20	50	50	1000	0.45 | 0.75	0.66	0.6	-0.15 | 0.15	0.06	0.002	0.05

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
K1, K2, and K	The number of clusters in groups 1, 2, and both, respectively.
M1 and M2	The average number of items (subjects) per cluster in groups 1 and 2, respectively.
N	The total sample size for the study. N = (K1 x M1) + (K2 x M2).
P1.L	The smallest proportion for group 1 (treatment group) that still results in the conclusion of equivalence.
P1.U	The largest proportion for group 1 (treatment group) that still results in the conclusion of equivalence.
P1.1	The actual value of the proportion for group 1. This is the value at which the power is calculated.
P2	The proportion for group 2, the standard, reference, baseline, or control group.
D0.L	The lower bound on the difference that results in the conclusion of equivalence. D0.L = P1.L - P2.
D0.U	The upper bound on the difference that results in the conclusion of equivalence. D0.U = P1.U - P2.
D1	The actual difference at which the power is calculated. D1 = P1.1 - P2
ICC	The intracluster correlation. This is the correlation between any two subjects within a cluster.
Alpha	The probability of rejecting a true null hypothesis.

Summary Statements

A parallel two-group cluster-randomized design will be used to test whether the Group 1 (treatment) proportion (P1) is equivalent to the Group 2 (reference) proportion (P2), with difference equivalence bounds of -0.15 and 0.15 (H0: P1 - P2 ≤ -0.15 or P1 - P2 ≥ 0.15 versus H1: -0.15 < P1 - P2 < 0.15). The comparison will be made using two one-sided Likelihood Score Tests (Farrington & Manning), with an overall Type I error rate (α) of 0.05. The reference group proportion (P2) is assumed to be 0.6. The intracluster correlation is assumed to be 0.002. To detect a proportion difference (P1 - P2) of 0 (or P1 of 0.6), with 2 clusters of 50 subjects per cluster in Group 1 and 2 clusters of 50 subjects per cluster in Group 2, the power is 0.33835.

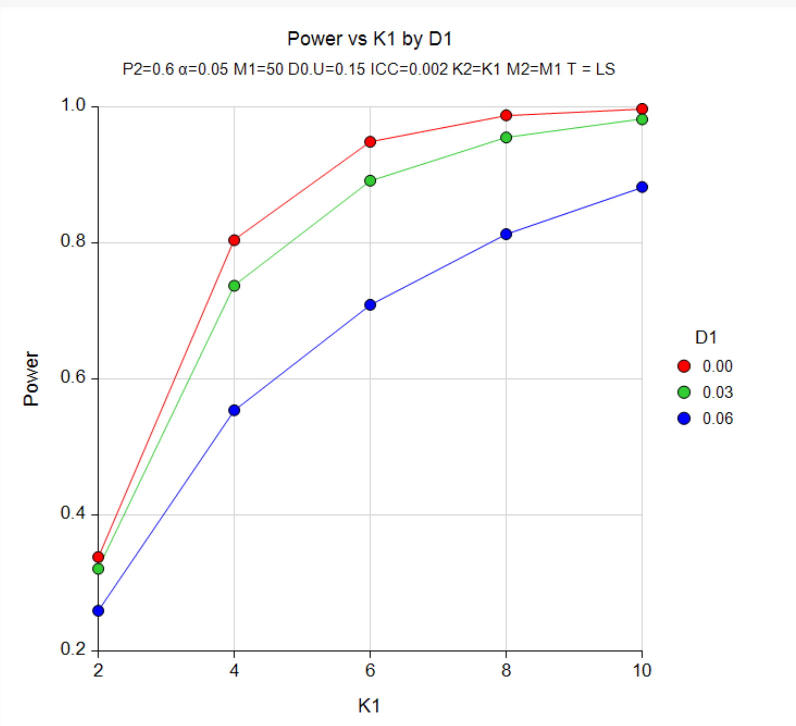
References

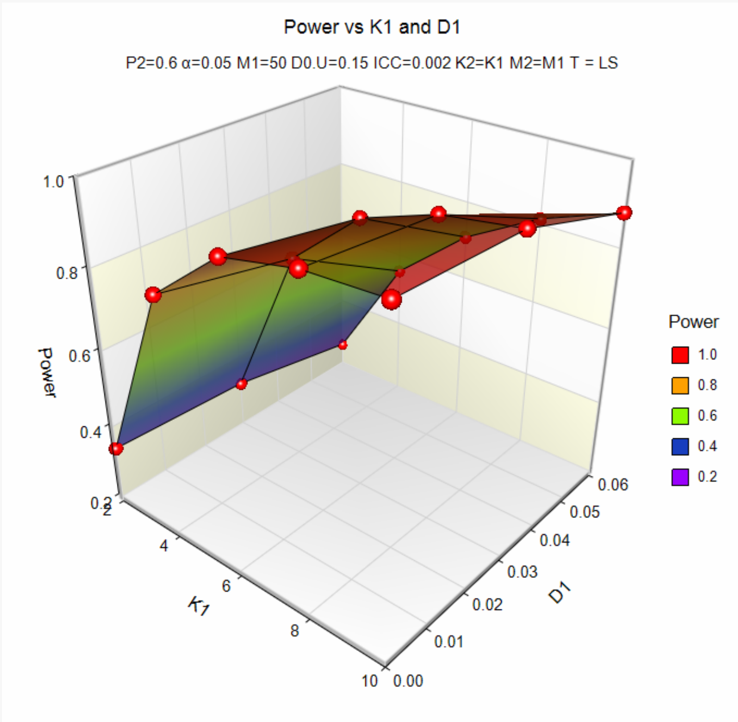
Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.
Chow, S.C. and Liu, J.P. 1999. Design and Analysis of Bioavailability and Bioequivalence Studies. Marcel Dekker. New York.
Donner, A. and Klar, N. 2000. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold. London.
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.
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.
Miettinen, O.S. and Nurminen, M. 1985. 'Comparative analysis of two rates.' Statistics in Medicine 4: 213-226.

This report shows the values of each of the parameters, one scenario per row. The total number of items sampled in group 1 is N1 = K1 x M1. The total number of items sampled in group 2 is N2 = K2 x M2.

Plots Section

Plots





The values from the table are displayed on the above plots. These plots give a quick look at the sample size that will be required for various values of D1.

Example 2 – Finding the Sample Size (Number of Clusters)

Continuing with the scenario given in Example 1, the researchers want to determine the number of clusters necessary for each value of D1 when the target power is set to 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 (K1)**
Test Type...**Likelihood Score (Farr. & Mann.)**
Power..**0.80**
Alpha...**0.05**
M1 (Average Cluster Size)............................**50**
K2 (Clusters in Group 2) **K1**
M2 (Average Cluster Size)............................**M1**
Input Type..**Differences**
D0.U (Upper Equivalence Difference)...........**0.15**
D0.L (Lower Equivalence Difference) ....**-D0.U**
D1 (Actual Difference)..................................**0 0.03 0.06**
P2 (Group 2 Proportion)................................**0.6**
ICC (Intracluster Correlation).........................**0.002**

Output

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

Numeric Results

Solve For: Sample Size (K1)
Groups: 1 = Treatment, 2 = Reference
Test Statistic: Likelihood Score Test (Farrington & Manning)
Hypotheses: H0: P1 - P2 ≤ D0.L or P1 - P2 ≥ D0.U vs. H1: D0.L < P1 - P2 < D0.U

	Number of Clusters			Average Cluster Size		Total Sample Size	Proportion Equivalence	Proportion Actual	Proportion Reference	Difference Equivalence	Difference Actual	Intracluster Correlation	
Power	K1	K2	K	M1	M2	N	P1.L | P1.U	P1.1	P2	D0.L | D0.U	D1	ICC	Alpha
0.80415	4	4	8	50	50	400	0.45 | 0.75	0.60	0.6	-0.15 | 0.15	0.00	0.002	0.05
0.83188	5	5	10	50	50	500	0.45 | 0.75	0.63	0.6	-0.15 | 0.15	0.03	0.002	0.05
0.81306	8	8	16	50	50	800	0.45 | 0.75	0.66	0.6	-0.15 | 0.15	0.06	0.002	0.05

The required sample size depends a great deal on the value of D1. The researchers should spend time determining the most appropriate value for D1.

Example 3 – Finding Power after an Experiment

Individuals promoting a new, more expensive treatment claim that it achieves better results than the current treatment without citing statistical evidence. A group of researchers attempted to show the claim was false through a study involving 12 hospitals. Two hundred patients at each of 6 randomly chosen hospitals were given the current treatment. Two hundred patients at each of the remaining 6 hospitals were given the new treatment. It was agreed before the experiment that if a difference of less than 0.05 in proportion of success could be shown, the two treatments would be deemed equivalent. The proportion of patients responding properly to the current treatment was 540/1200 = 0.450. The proportion of patients responding properly to the new treatment was 570/1200 = 0.475. This result did not show significant equivalence at the 0.05 level. The researchers want to know the power of their equivalence test. They decide to use the intracluster correlation coefficient estimated from the data, which was 0.0043. Although the observed difference in proportions is 0.475 – 0.450 = 0.025, the equivalence difference is still 0.05. This value is used in the power calculation.

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**
Test Type	**Likelihood Score (Farr. & Mann.)**
Alpha	**0.05**
K1 (Clusters in Group 1)	**6**
M1 (Average Cluster Size)	**200**
K2 (Clusters in Group 2)	**K1**
M2 (Average Cluster Size)	**M1**
Input Type	**Differences**
D0.U (Upper Equivalence Difference)	**0.05**
D0.L (Lower Equivalence Difference)	**-D0.U**
D1 (Actual Difference)	**0.0**
P2 (Group 2 Proportion)	**0.45**
ICC (Intracluster Correlation)	**0.0043**

Output

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

Numeric Results

Solve For: Power
Groups: 1 = Treatment, 2 = Reference
Test Statistic: Likelihood Score Test (Farrington & Manning)
Hypotheses: H0: P1 - P2 ≤ D0.L or P1 - P2 ≥ D0.U vs. H1: D0.L < P1 - P2 < D0.U

	Number of Clusters			Average Cluster Size		Total Sample Size	Proportion			Difference		Intracluster Correlation	
							Equivalence	Actual	Reference	Equivalence	Actual		
Power	K1	K2	K	M1	M2	N	P1.L | P1.U	P1.1	P2	D0.L | D0.U	D1	ICC	Alpha
0.13053	6	6	12	200	200	2400	0.4 | 0.5	0.45	0.45	-0.05 | 0.05	0	0.0043	0.05

The power of the test of equivalence was only 0.1305.

Example 4 – Finding Sample Size (Individuals within Clusters)

An agency would like to show the proportion of success is the same for two treatments. Eight doctors are available for the study. Four will be randomly chosen to be trained to administer treatment 1. The remaining four will administer treatment 2. The treatments will be considered equivalent if the proportion of success of treatment 1 is within 0.10 of treatment 2 success. The agency would like to know the number of patients that need to be treated by each doctor to achieve 80% power for the equivalence test. Various values for the intracluster correlation coefficient will be used since its true value is unknown. It is expected that the two treatments will have a success rate near 0.70. Alpha is set at 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 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 .... **Sample Size (M1)**
Test Type.. **Likelihood Score (Farr. & Mann.)**
Power... **0.80**
Alpha.. **0.05**
K1 (Clusters in Group 1) **4**
K2 (Clusters in Group 2) **K1**
M2 (Average Cluster Size)............................ **M1**
Input Type... **Differences**
D0.U (Upper Equivalence Difference)........... **0.1**
D0.L (Lower Equivalence Difference) .... **-D0.U**
D1 (Actual Difference).................................. **0.0**
P2 (Group 2 Proportion)................................ **0.7**
ICC (Intracluster Correlation)........................ **0.001 to 0.01 by 0.001**

Output

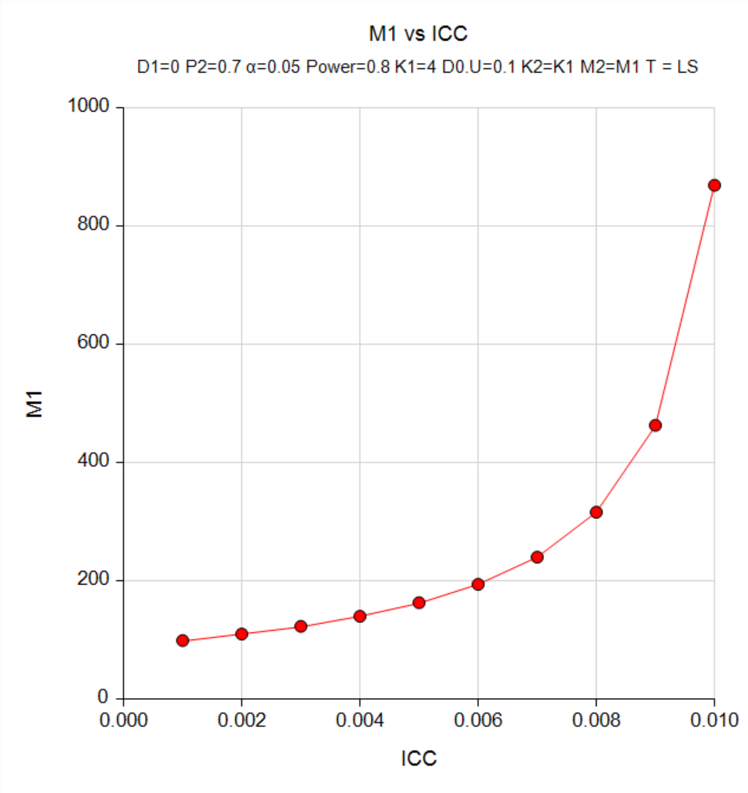
Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Solve For: Sample Size (M1)
Groups: 1 = Treatment, 2 = Reference
Test Statistic: Likelihood Score Test (Farrington & Manning)
Hypotheses: H0: P1 - P2 ≤ D0.L or P1 - P2 ≥ D0.U vs. H1: D0.L < P1 - P2 < D0.U

	Number of Clusters			Average Cluster Size		Total Sample Size	Proportion			Difference		Intracluster Correlation	
							Equivalence	Actual	Reference	Equivalence	Actual		
Power	K1	K2	K	M1	M2	N	P1.L | P1.U	P1.1	P2	D0.L | D0.U	D1	ICC	Alpha
0.80228	4	4	8	99	99	792	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.001	0.05
0.80312	4	4	8	110	110	880	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.002	0.05
0.80160	4	4	8	123	123	984	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.003	0.05
0.80121	4	4	8	140	140	1120	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.004	0.05
0.80167	4	4	8	163	163	1304	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.005	0.05
0.80077	4	4	8	194	194	1552	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.006	0.05
0.80014	4	4	8	240	240	1920	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.007	0.05
0.80006	4	4	8	316	316	2528	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.008	0.05
0.80001	4	4	8	463	463	3704	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.009	0.05
0.80006	4	4	8	869	869	6952	0.6 | 0.8	0.7	0.7	-0.1 | 0.1	0	0.010	0.05

Plots



The number of patients needed to be treated by each doctor ranges from 99 to 869 depending on the value of the intracluster correlation coefficient.

Example 5 – Validation

We could not find an example of this type of analysis in the literature. Therefore, we will validate the procedure by comparing the results to those computed by the "Equivalence Tests for the Difference Between Two Proportions" procedure since both should give identical results for the same sample sizes when the ICC is set to zero. We ran the case when N1 = N2 = 200, P2 = 0.6, D0.U = 0.15, D0.L = -D0.U, D1 = 0, and Alpha = 0.05. In this procedure, set M1 = 1 and set K1 = 200. The Equivalence Tests for the Difference Between Two Proportions procedure calculates the power to be 0.84823 for the Likelihood Score (Farrington & Manning) test.

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 5** 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**
Test Type	**Likelihood Score (Farr. & Mann.)**
Alpha	**0.05**
K1 (Clusters in Group 1)	**200**
M1 (Average Cluster Size)	**1**
K2 (Clusters in Group 2)	**K1**
M2 (Average Cluster Size)	**M1**
Input Type	**Differences**
D0.U (Upper Equivalence Difference)	**0.15**
D0.L (Lower Equivalence Difference)	**-D0.U**
D1 (Actual Difference)	**0**
P2 (Group 2 Proportion)	**0.6**
ICC (Intracluster Correlation)	**0.0**

Output

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

Numeric Results

Solve For: Power
Groups: 1 = Treatment, 2 = Reference
Test Statistic: Likelihood Score Test (Farrington & Manning)
Hypotheses: H0: P1 - P2 ≤ D0.L or P1 - P2 ≥ D0.U vs. H1: D0.L < P1 - P2 < D0.U

	Number of Clusters			Average Cluster Size		Total Sample Size	Proportion			Difference		Intracluster Correlation	
							Equivalence	Actual	Reference	Equivalence	Actual		
Power	K1	K2	K	M1	M2	N	P1.L | P1.U	P1.1	P2	D0.L | D0.U	D1	ICC	Alpha
0.84823	200	200	400	1	1	400	0.45 | 0.75	0.6	0.6	-0.15 | 0.15	0	0	0.05

The power computed by this procedure is also 0.84823 when ICC = 0.