Chapter 230

Tests for Two Proportions in a Cluster-Randomized Design

Introduction

A *cluster (group) randomized design* is one in which whole units, or clusters, of subjects are randomized to the groups rather than the individual subjects in those clusters. However, the conclusions of the study concern individual subjects rather than the clusters. Examples of clusters are families, school classes, neighborhoods, and hospital wards.

Cluster-randomized designs are often adopted when there is a high risk of contamination if cluster members were randomized individually. For example, it may be difficult for an instructor to use two methods of teaching individuals in the same class. The price of randomizing by clusters is a loss of efficiency--the number of subjects needed to obtain a certain level of precision in a cluster-randomized trial is usually much larger than the number needed when the subjects are randomized individually. Hence, the standard methods of sample size estimation cannot be used.

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

$$\sigma^2_{S,g} = \frac{p_g\left(1 - p_g\right)}{n_g}$$

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

$$\sigma^2_{C,g} = \frac{p_g\left(1 - p_g\right)\left(1 + \left(m_g - 1\right)\rho\right)}{k_g m_g}$$

$$= \sigma^2_{S,g}\left[1 + \left(m_g - 1\right)\rho\right]$$

$$= F_{g,\rho}\sigma^2_{S,g}$$

The factor $\left[1 + \left(m_g - 1\right)\rho\right]$ 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 200, "Tests for Two Proportions") may be used with cluster-randomized designs as well, as long as an adjustment is made for the inflation factor. The basic form of the z-test becomes

$$z = \frac{\left|\hat{D} - \delta_0\right|}{\hat{\sigma}_{\hat{D}}(\delta_0)}$$

where

$$\hat{D} = \hat{p}_1 - \hat{p}_2$$

$$\delta_0 = p_1 - p_2 | H_0$$

$$\hat{\sigma}_{\hat{D}}(\delta_0) = \sqrt{\frac{\tilde{p}_1\left(1 - \tilde{p}_1\right)F_{1,\rho}}{n_1} + \frac{\tilde{p}_2\left(1 - \tilde{p}_2\right)F_{2,\rho}}{n_2}}$$

The quantities $\tilde{p}_1$ and $\tilde{p}_2$ are the maximum likelihood estimates constrained by $\tilde{p}_1 - \tilde{p}_2 = \delta_0$.

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.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design and Options tabs. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains the parameters associated with this test such as the proportions, sample sizes, alpha, and power.

Solve For

Solve For

This option specifies the parameter to be solved for using the other parameters. The parameters that may be selected are *Power*, *Sample Size (K1)*, *Sample Size (M1)*, *Effect Size*, and *ICC*. Under most situations, you will select either *Power* or *Sample Size (K1)*.

Select *Sample Size (K1)* when you want to calculate the number of clusters per group needed to achieve a given power and alpha level.

Select *Power* when you want to calculate the power of an experiment.

Test

Alternative Hypothesis

Specify the direction of the alternative hypothesis.

- **One-Sided (H1: P1 – P2 < 0)**

  Refers to a lower-tailed, one-sided test in which the alternative hypothesis is of the type H1: P1 – P2 < 0.

- **One-Sided (H1: P1 – P2 > 0)**

  Refers to an upper-tailed, one-sided test in which the alternative hypothesis is of the type H1: P1 – P2 > 0.

- **Two-Sided (H1: P1 - P2 ≠ 0)**

  Refers to a two-sided test in which the alternative hypothesis is of the type H1: P1 – P2 ≠ 0.

For one-sided tests, the direction you select must match the values of P1 (or D1) and P2 you enter. For example, if you select H1: P1 - P2 < 0, then the value of P1 must be less than the value of P2 (or D1 must be less than 0).

Test Type

Specify which test statistic is used in searching and reporting.

Power and Alpha

Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal proportions when in fact they are different.

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. For this procedure, a type-I error occurs when you reject the null hypothesis of equal proportions when in fact they are 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 – Group 1 (Treatment)

K1 (Clusters in Group 1)

Enter a value (or range of values) for the number of clusters in this group. You may enter a range of values such as *10 to 20 by 2*. The sample size for this group is equal to the number of clusters times the number of subjects per cluster.

M1 (Items per Cluster in Group 1)

This is the average number of items (subjects) per cluster in group one. This value must be a positive number that is at least one. You can use a list of values such as *100 150 200*.

Sample Size – Group 2 (Control)

K2 (Clusters in Group 2)

This is the number of clusters in group two. The sample size for this group is equal to the number of clusters times the number of subjects per cluster. This value must be a positive number.

If you simply want a multiple of the value for group one, you would enter the multiple followed by *K1*, with no blanks. If you want to use *K1* directly, you do not have to pre-multiply by *1*. For example, all of the following are valid entries: *10 K1 2K1 0.5K1*.

You can use a list of values such as *10 20 30* or *K1 2K1 3K1*.

M2 (Items per Cluster in Group 2)

This is the average number of items (subjects) per cluster in group two. This value must be at least one.

If you simply want a multiple of the value for group one, you would enter the multiple followed by *M1*, with no blanks. If you want to use *M1* directly, you do not have to pre-multiply by *1*. For example, all of the following are valid entries: *10 M1 2M1 0.5M1*.

You can use a list of values such as *10 20 30* or *M1 2M1 3M1*.

Effect Size

Input Type

Indicate what type of values to enter to specify the effect size. Regardless of the entry type chosen, the test statistics used in the power and sample size calculations are the same. This option is simply given for convenience in specifying the effect size.

P1 (Group 1 Proportion |H1)

This is the value of P1 under the alternative hypothesis. The power calculations assume that this is the actual value of the proportion in group 1. The difference D1 = P1 – P2.

You may enter a range of values such as *0.1 0.2 0.3* or *0.1 to 0.9 by 0.1*. Note that values must be between zero and one.

D1 (Difference|H1 = P1 – P2)

This option specifies the difference between P1 and P2. This difference is used with P2 to calculate the value of P1 using the formula: P1 = D1 + P2. Note that P1 here means the value of P1 under H1. Differences must be between -1 and 1. They cannot take on the values -1 or 1.

The power calculations assume that P1 is the actual value of the proportion in group 1 (experimental or treatment group).

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

R1 (Ratio|H1 = P1/P2)

This option specifies the ratio of P1 and P2, where P1 is the proportion in group 1 under the alternative hypothesis. This ratio is used with P2 to calculate the value of P1 using the formula: P1 = R1 x P2.The power calculations assume that P1 is the actual value of the proportion in group 1 (experimental or treatment group).

You may enter a range of values such as *0.5 0.6 0.7 0.8* or *1.25 to 2.0 by 0.25*. Ratios must be greater than zero. They cannot take on the value of one.

P2 (Group 2 Proportion)

Specify the value of p_2, the control, baseline, or standard group's proportion. The null hypothesis is that the two proportions differ by a specified amount.

Since p_2 is a proportion, these values must be between zero and one. You may enter a range of values such as *0.1,0.2,0.3* or *0.1 to 0.9 by 0.1*.

Effect Size – Intracluster Correlation

ICC (Intracluster Correlation)

Enter a value (or range of values) for the intracluster correlation. This correlation may be thought of as the simple correlation between any two observations in the same cluster. It may also be thought of as the proportion of total variance in the observations that can be attributed to difference between clusters.

Although the actual range for this value is from zero to one, typical values range from 0.002 to 0.05.

Example 1 – Finding Power

Two competing physical therapy treatments have been available for several years but have not yet been compared as to their effectiveness. The comparison of the two treatments is complicated by the sampling method that will be used. Instead of randomly assigning individuals to treatments, the researchers will randomly select two groups of physical therapists. The first group will be selected from those who use treatment 1. The second group will be selected from those who use treatment 2. The researchers will then follow up on the success or failure of the treatment for multiple patients of each physical therapist. The success rate of treatment 2 is known to be about 0.44. The researchers want to examine effect of the number of physical therapists used in each group and the number of patients for each physical therapist on the power of the test. They wish to determine the power if the treatments are at least 0.07 apart in proportion. They plan to use the Unpooled Z-Test to analyze the data. Based on similar studies, the intracluster correlation is estimated to be 0.02.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Proportions in a Cluster-Randomized Design** procedure window by expanding **Proportions**, then **Two Proportions – Cluster Randomized**, then clicking on **Test (Inequality)**, and then clicking on **Tests for Two Proportions in a Cluster-Randomized Design**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

Option   **Value**

Design Tab
Solve For .... **Power**
Alternative Hypothesis .... **Two-Sided (H1: P1 - P2 ≠ 0)**
Test Type... **Z-Test (Unpooled)**
Alpha... **0.05**
K1 (Clusters in Group 1) .... **10 to 25 by 5**
M1 (Items per Cluster in Group 1) .... **10 to 50 by 10**
K2 (Clusters in Group 2) **K1**
M2 (Items per Cluster in Group 2) **M1**
Input Type.. **Differences**
D1 (Difference|H1 = P1-P2).................... **0.07**
P2 (Group 2 Proportion) **0.44**
ICC (Intracluster Correlation).................. **0.02**

Plots Tab
2D Plot: X (Horizontal) Axis Parameter .. **K1**
3D Plot: X (Horizontal) Axis Parameter .. **K1**

Output

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

Numeric Results

Numeric Results for Testing Two Proportions (Cluster-Randomized) ————————————————
Test Statistic: Z-Test (Unpooled)
Hypotheses: H0: P1 - P2 = 0 vs. H1: P1 - P2 ≠ 0

Power	Group 1 Clusters/ Items K1/M1	Group 2 Clusters/ Items K2/M2	Group 1 Prop P1	Group 2 Prop P2	Diff D1	Intra- Cluster Corr. ICC	Alpha
0.14998	10/10	10/10	0.5100	0.4400	0.0700	0.0200	0.050
0.22330	10/20	10/20	0.5100	0.4400	0.0700	0.0200	0.050
0.27777	10/30	10/30	0.5100	0.4400	0.0700	0.0200	0.050
0.31930	10/40	10/40	0.5100	0.4400	0.0700	0.0200	0.050
0.35181	10/50	10/50	0.5100	0.4400	0.0700	0.0200	0.050
0.20158	15/10	15/10	0.5100	0.4400	0.0700	0.0200	0.050
0.31063	15/20	15/20	0.5100	0.4400	0.0700	0.0200	0.050
0.38870	15/30	15/30	0.5100	0.4400	0.0700	0.0200	0.050
0.44611	15/40	15/40	0.5100	0.4400	0.0700	0.0200	0.050
0.48963	15/50	15/50	0.5100	0.4400	0.0700	0.0200	0.050
0.25316	20/10	20/10	0.5100	0.4400	0.0700	0.0200	0.050
0.39437	20/20	20/20	0.5100	0.4400	0.0700	0.0200	0.050
0.49061	20/30	20/30	0.5100	0.4400	0.0700	0.0200	0.050
0.55827	20/40	20/40	0.5100	0.4400	0.0700	0.0200	0.050
0.60761	20/50	20/50	0.5100	0.4400	0.0700	0.0200	0.050
0.30407	25/10	25/10	0.5100	0.4400	0.0700	0.0200	0.050
0.47271	25/20	25/20	0.5100	0.4400	0.0700	0.0200	0.050
0.58119	25/30	25/30	0.5100	0.4400	0.0700	0.0200	0.050
0.65365	25/40	25/40	0.5100	0.4400	0.0700	0.0200	0.050
0.70424	25/50	25/50	0.5100	0.4400	0.0700	0.0200	0.050

References
Donner, A. and Klar, N. 2000. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold.
  London.

Report Definitions
Power is the probability of rejecting a false null hypothesis. It should be close to one.
K1 and K2 are the number of clusters in groups 1 and 2, respectively.
M1 and M2 are the average number of items (subjects) per cluster in groups 1 and 2, respectively.
P1 is the proportion for group 1 (the treatment group) assuming the alternative hypothesis (H1).
P2 is the proportion for group 2, the standard, reference, baseline, or control group.
D1 = P1 - P2 is the difference assuming the alternative hypothesis (H1) at which power is calculated.
ICC is the intracluster correlation.
Alpha is the probability of rejecting a true null hypothesis.
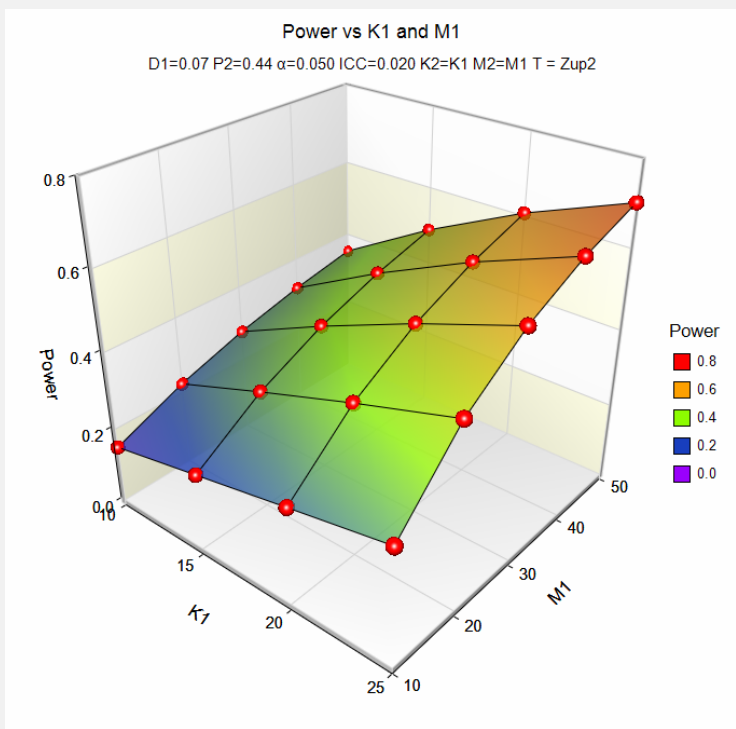
Summary Statements ————————————————————————————————————
Sample sizes of 100 in group 1 and 100 in group 2, which were obtained by sampling 10 clusters
with 10 subjects each in group 1 and 10 clusters with 10 subjects each in group 2, achieve
14.998% power to detect a difference between the group proportions of 0.0700. The proportion in
group 1 (the treatment group) is assumed to be 0.4400 under the null hypothesis and 0.5100
under the alternative hypothesis. The proportion in group 2 (the control group) is 0.4400. The
test statistic used is the two-sided Z-Test (Unpooled). The intracluster correlation is 0.0200,
and the significance level of the test is 0.050.

This report shows the values of each of the parameters, one scenario per row. The values from this table are plotted in the chart below.

Plots Section

Power vs K1 by M1

D1=0.07 P2=0.44 α=0.050 ICC=0.020 K2=K1 M2=M1 T = Zup2

Power vs K1 and M1

D1=0.07 P2=0.44 α=0.050 ICC=0.020 K2=K1 M2=M1 T = Zup2
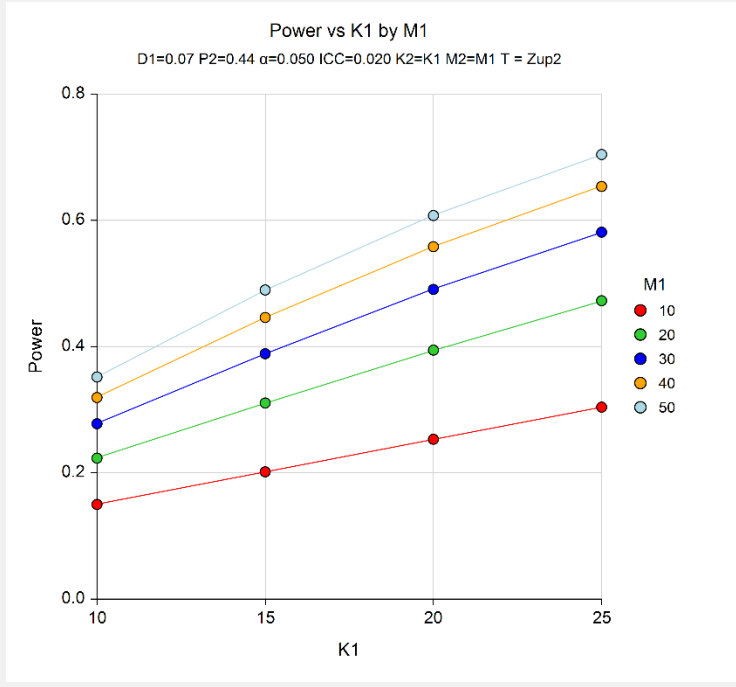
The values from the table are displayed on the above charts. These charts give a quick look at the power that is achieved for various combinations of cluster size and numbers of clusters.

Example 2 – Finding Sample Size

Continuing with Example 1, the maximum number of therapists the researchers hope to use is 25 for each treatment. They decide to determine how many patients each therapist would have to treat to achieve 90% power if the maximum number of therapists is used.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Proportions in a Cluster-Randomized Design** procedure window by expanding **Proportions**, then **Two Proportions – Cluster Randomized**, then clicking on **Test (Inequality)**, and then clicking on **Tests for Two Proportions in a Cluster-Randomized Design**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

Option   **Value**

Design Tab
Solve For .... **Sample Size (M1)**
Alternative Hypothesis .... **Two-Sided (H1: P1 - P2 ≠ 0)**
Test Type .... **Z-Test (Unpooled)**
Power **0.90**
Alpha **0.05**
K1 (Clusters in Group 1) **25**
K2 (Clusters in Group 2) **K1**
M2 (Items per Cluster in Group 2) **M1**
Input Type .... **Differences**
D1 (Difference|H1 = P1-P2) **0.07**
P2 (Group 2 Proportion) **0.44**
ICC (Intracluster Correlation) **0.02**

Output

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

Numeric Results

Numeric Results for Testing Two Proportions (Cluster-Randomized) ————————————————————————
Test Statistic: Z-Test (Unpooled)
Hypotheses: H0: P1 - P2 = 0 vs. H1: P1 - P2 ≠ 0

Power	Group 1 Clusters/ Items K1/M1	Group 2 Clusters/ Items K2/M2	Group 1 Prop P1	Group 2 Prop P2	Diff D1	Intra-Cluster Corr. ICC	Alpha
0.90004	25/281	25/281	0.5100	0.4400	0.0700	0.0200	0.050

To achieve 90% power, each therapist would need to be evaluated on 281 patients.

Example 3 – Investigating the Impact of the Intracluster Correlation

Continuing with the scenario given in Example 2, the researchers now want to investigate the required number of therapists if each treats 300 patients and also look at effect of the intracluster correlation on the sample size. They decide to vary ICC between 0.01 and 0.03 to look at the effect.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for Two Proportions in a Cluster-Randomized Design** procedure window by expanding **Proportions**, then **Two Proportions – Cluster Randomized**, then clicking on **Test (Inequality),** and then clicking on **Tests for Two Proportions in a Cluster-Randomized Design**. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

Option   **Value**

Design Tab
Solve For .... **Sample Size (K1)**
Alternative Hypothesis .... **Two-Sided (H1: P1 - P2 ≠ 0)**
Test Type .... **Z-Test (Unpooled)**
Power **0.90**
Alpha **0.05**
M1 (Items per Cluster in Group 1) **300**
K2 (Clusters in Group 2) **K1**
M2 (Items per Cluster in Group 2) **M1**
Input Type .... **Differences**
D1 (Difference|H1 = P1-P2) **0.07**
P2 (Group 2 Proportion) **0.44**
ICC (Intracluster Correlation) .... **0.01 to 0.03 by 0.005**

Output

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

Numeric Results

Numeric Results for Testing Two Proportions (Cluster-Randomized) ————————————————————————
Test Statistic: Z-Test (Unpooled)
Hypotheses: H0: P1 - P2 = 0 vs. H1: P1 - P2 ≠ 0

Power	Group 1 Clusters/ Items K1/M1	Group 2 Clusters/ Items K2/M2	Group 1 Prop P1	Group 2 Prop P2	Diff D1	Intra-Cluster Corr. ICC	Alpha
0.91574	15/300	15/300	0.5100	0.4400	0.0700	0.0100	0.050
0.90764	20/300	20/300	0.5100	0.4400	0.0700	0.0150	0.050
0.90270	25/300	25/300	0.5100	0.4400	0.0700	0.0200	0.050
0.90850	31/300	31/300	0.5100	0.4400	0.0700	0.0250	0.050
0.90496	36/300	36/300	0.5100	0.4400	0.0700	0.0300	0.050

This report shows that the required number of therapists doubles when the ICC is changed from 0.01 to 0.03.

Example 4 – Validation using Donner and Klar (2000)

Donner and Klar (2000), page 63, presents a sample size study in which P2 = 0.06, D1 = -0.02, D0=0, ICC = 0.01, M1=M2=100, two-sided alpha = 0.05, and beta = 0.20. Using the pooled z test statistic, they found the number of subjects to be 3698 in each group, which they round off to 38 clusters per 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 **Tests for Two Proportions in a Cluster-Randomized Design** procedure window by expanding **Proportions**, then **Two Proportions – Cluster Randomized**, then clicking on **Test (Inequality)**, and then clicking on **Tests for Two Proportions in a Cluster-Randomized Design**. You may then make the appropriate entries as listed below, or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

Option	Value
Design Tab	
Solve For	**Sample Size (K1)**
Alternative Hypothesis	**Two-Sided (H1: P1 - P2 ≠ 0)**
Test Type...	**Z-Test (Pooled)**
Power...	**0.80**
Alpha..	**0.05**
M1 (Items per Cluster in Group 1)	**100**
K2 (Clusters in Group 2)	**K1**
M2 (Items per Cluster in Group 2)	**M1**
Input Type..	**Differences**
D1 (Difference|H1 = P1-P2)....................	**-0.02**
P2 (Group 2 Proportion)	**0.06**
ICC (Intracluster Correlation).................	**0.01**

Output

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

Numeric Results

Numeric Results for Testing Two Proportions (Cluster-Randomized) ————————————————————————
Test Statistic: Z-Test (Pooled)
Hypotheses: H0: P1 - P2 = 0 vs. H1: P1 - P2 ≠ 0

Power	Group 1 Clusters/ Items K1/M1	Group 2 Clusters/ Items K2/M2	Group 1 Prop P1	Group 2 Prop P2	Diff D1	Intra-Cluster Corr. ICC	Alpha
0.80962	38/100	38/100	0.0400	0.0600	-0.0200	0.0100	0.050

PASS has also found the required sample size to be 38 clusters.