

## Chapter 233

# Non-Zero Null Tests for the Difference of Two Proportions in a Cluster-Randomized Design

## Introduction

This module provides power analysis and sample size calculation for inequality tests of the difference with a non-zero null hypothesis 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,  $k = 1, 2, \dots, K_g$  is a cluster within group  $g$ , and  $m = 1, 2, \dots, M_g$  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

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

## Non-Zero Null Tests for the Difference of Two Proportions in a Cluster-Randomized Design

The factor  $[1 + (m_g - 1)\rho]$  is called the *inflation factor*. The Greek letter  $\rho$  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  $\rho$  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 205, "Tests for the Difference Between Two Proportions with a Non-Zero Null Hypothesis") 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 study the effectiveness of a new treatment. Historically, the standard treatment has enjoyed a 60% cure rate. The new treatment is more expensive and has more side effects than the standard treatment. Thus, the new treatment will be adopted if it is at least 8% better than the standard treatment.

The researchers will recruit patients from various hospitals. All patients at a particular hospital will receive the same treatment. They anticipate an average of 100 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 one-sided Farrington and Manning test at group cluster sizes ranging from 2 to 10 for detecting a difference of 0.08 when the actual cure rate of the new treatment ranges from 70% to 76%. 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 .....	<b>Power</b>
Alternative Hypothesis .....	<b>One-Sided (H1: P1 - P2 &gt; D0)</b>
Test Type .....	<b>Likelihood Score (Farr. &amp; Mann.)</b>
Alpha .....	<b>0.05</b>
K1 (Clusters in Group 1) .....	<b>2 4 6 8 10</b>
M1 (Average Cluster Size) .....	<b>100</b>
K2 (Clusters in Group 2) .....	<b>K1</b>
M2 (Average Cluster Size) .....	<b>M1</b>
Input Type .....	<b>Differences</b>
D0 (Difference H0 = P1.0 - P2) .....	<b>0.08</b>
D1 (Difference H1 = P1.1 - P2) .....	<b>0.10 to 0.16 by 0.01</b>
P2 (Group 2 Proportion) .....	<b>0.6</b>
ICC (Intracluster Correlation) .....	<b>0.002</b>

## Non-Zero Null Tests for the Difference of Two Proportions in a Cluster-Randomized Design

## 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:  $H_0: P_1 - P_2 \leq D_0$  vs.  $H_1: P_1 - P_2 > D_0$

Power	Number of Clusters			Cluster Size		Total Sample Size N	Proportions			Difference		Intraclass Correlation ICC	Alpha
	K1	K2	K	M1	M2		P1 H0 P1.0	P1 H1 P1.1	P2	Diff H0 D0	Diff H1 D1		
0.10341	2	2	4	100	100	400	0.68	0.70	0.6	0.08	0.10	0.002	0.05
0.13506	4	4	8	100	100	800	0.68	0.70	0.6	0.08	0.10	0.002	0.05
0.16346	6	6	12	100	100	1200	0.68	0.70	0.6	0.08	0.10	0.002	0.05
0.19021	8	8	16	100	100	1600	0.68	0.70	0.6	0.08	0.10	0.002	0.05
0.21588	10	10	20	100	100	2000	0.68	0.70	0.6	0.08	0.10	0.002	0.05
0.14261	2	2	4	100	100	400	0.68	0.71	0.6	0.08	0.11	0.002	0.05
0.20377	4	4	8	100	100	800	0.68	0.71	0.6	0.08	0.11	0.002	0.05
0.25988	6	6	12	100	100	1200	0.68	0.71	0.6	0.08	0.11	0.002	0.05
0.31271	8	8	16	100	100	1600	0.68	0.71	0.6	0.08	0.11	0.002	0.05
0.36274	10	10	20	100	100	2000	0.68	0.71	0.6	0.08	0.11	0.002	0.05
0.19129	2	2	4	100	100	400	0.68	0.72	0.6	0.08	0.12	0.002	0.05
0.29081	4	4	8	100	100	800	0.68	0.72	0.6	0.08	0.12	0.002	0.05
0.38063	6	6	12	100	100	1200	0.68	0.72	0.6	0.08	0.12	0.002	0.05
0.46199	8	8	16	100	100	1600	0.68	0.72	0.6	0.08	0.12	0.002	0.05
0.53513	10	10	20	100	100	2000	0.68	0.72	0.6	0.08	0.12	0.002	0.05
0.24963	2	2	4	100	100	400	0.68	0.73	0.6	0.08	0.13	0.002	0.05
0.39323	4	4	8	100	100	800	0.68	0.73	0.6	0.08	0.13	0.002	0.05
0.51582	6	6	12	100	100	1200	0.68	0.73	0.6	0.08	0.13	0.002	0.05
0.61848	8	8	16	100	100	1600	0.68	0.73	0.6	0.08	0.13	0.002	0.05
0.70270	10	10	20	100	100	2000	0.68	0.73	0.6	0.08	0.13	0.002	0.05
0.31705	2	2	4	100	100	400	0.68	0.74	0.6	0.08	0.14	0.002	0.05
0.50493	4	4	8	100	100	800	0.68	0.74	0.6	0.08	0.14	0.002	0.05
0.65065	6	6	12	100	100	1200	0.68	0.74	0.6	0.08	0.14	0.002	0.05
0.75896	8	8	16	100	100	1600	0.68	0.74	0.6	0.08	0.14	0.002	0.05
0.83677	10	10	20	100	100	2000	0.68	0.74	0.6	0.08	0.14	0.002	0.05
0.39208	2	2	4	100	100	400	0.68	0.75	0.6	0.08	0.15	0.002	0.05
0.61756	4	4	8	100	100	800	0.68	0.75	0.6	0.08	0.15	0.002	0.05
0.76997	6	6	12	100	100	1200	0.68	0.75	0.6	0.08	0.15	0.002	0.05
0.86637	8	8	16	100	100	1600	0.68	0.75	0.6	0.08	0.15	0.002	0.05
0.92449	10	10	20	100	100	2000	0.68	0.75	0.6	0.08	0.15	0.002	0.05
0.47241	2	2	4	100	100	400	0.68	0.76	0.6	0.08	0.16	0.002	0.05
0.72226	4	4	8	100	100	800	0.68	0.76	0.6	0.08	0.16	0.002	0.05
0.86326	6	6	12	100	100	1200	0.68	0.76	0.6	0.08	0.16	0.002	0.05
0.93590	8	8	16	100	100	1600	0.68	0.76	0.6	0.08	0.16	0.002	0.05
0.97106	10	10	20	100	100	2000	0.68	0.76	0.6	0.08	0.16	0.002	0.05

**H0** An abbreviation for the null hypothesis. This is the hypothesis being evaluated by the statistical test.  
**H1** An abbreviation for the alternative hypothesis. This hypothesis gives the "true" parameter values.  
**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 number of subjects in the study.  $N = (K1 \times M1) + (K2 \times M2)$ .  
**P1.0** The proportion for group 1 (the treatment group) assuming the null hypothesis (H0).  
**P1.1** The proportion for group 1 (the treatment group) assuming the alternative hypothesis (H1).  
**P2** The proportion for group 2, the standard, reference, baseline, or control group.  
**D0** The difference in proportions assuming the null hypothesis (H0).  $D0 = P1.0 - P2$ .  
**D1** The difference in proportions at which the power is calculated.  $D1 = P1.1 - P2$ .  
**ICC** The intraclass correlation.  
**Alpha** The probability of rejecting a true null hypothesis.

## Non-Zero Null Tests for the Difference of Two Proportions in a Cluster-Randomized Design

**Summary Statements**

A parallel two-group cluster-randomized design will be used to test whether the Group 1 (treatment) proportion ( $P_1$ ) is greater than the Group 2 (reference) proportion ( $P_2$ ) by a margin, with a non-zero null margin of 0.08 ( $H_0: P_1 - P_2 \leq 0.08$  versus  $H_1: P_1 - P_2 > 0.08$ ). The comparison will be made using a one-sided Likelihood Score Test (Farrington & Manning) based on the proportion difference ( $P_1 - P_2$ ), with a Type I error rate ( $\alpha$ ) of 0.05. The reference group proportion ( $P_2$ ) is assumed to be 0.6. The intracluster correlation is assumed to be 0.002. To detect a proportion difference ( $P_1 - P_2$ ) of 0.1 (or  $P_1$  of 0.7), with 2 clusters of 100 subjects per cluster in Group 1 and 2 clusters of 100 subjects per cluster in Group 2, the power is 0.10341.

**References**

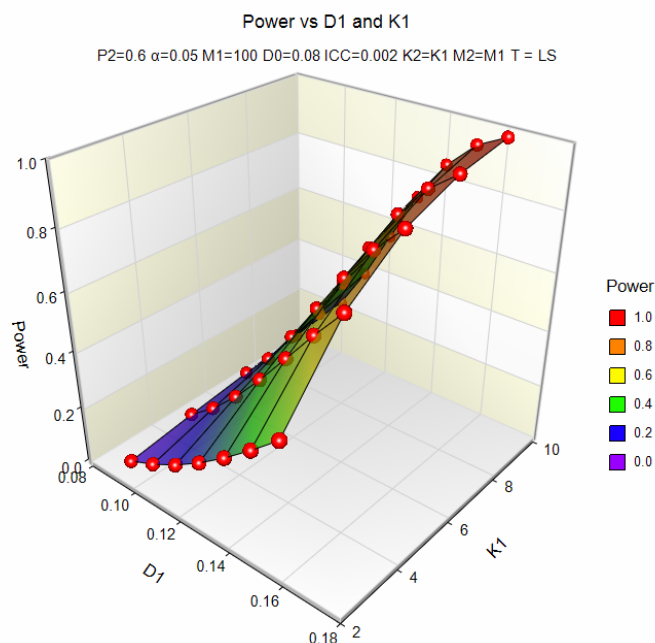
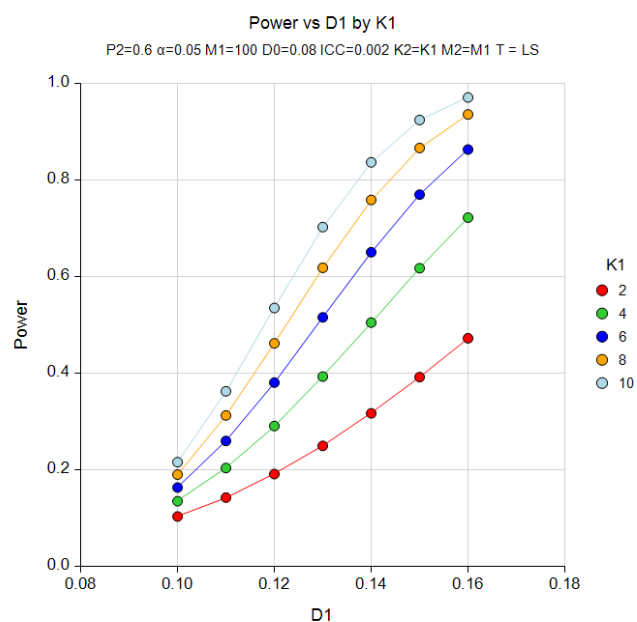
- Blackwelder, W.C. 1998. 'Equivalence Trials.' In Encyclopedia of Biostatistics, John Wiley and Sons. New York. Volume 2, 1367-1372.
- Chow, S.C. and Liu, J.P. 1999. Design and Analysis of Bioavailability and Bioequivalence Studies. Marcel Dekker. New York.
- Donner, A. and Klar, N. 2000. Design and 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  $N_1 = K_1 \times M_1$ . The total number of items sampled in group 2 is  $N_2 = K_2 \times M_2$ .

## Non-Zero Null Tests for the Difference of Two Proportions in a Cluster-Randomized Design

## Plots Section

## Plots



The values from the table are displayed on the above plots. These plots give us a quick look at the sample sizes 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)**  
 Alternative Hypothesis ..... **One-Sided (H1: P1 - P2 > D0)**  
 Test Type ..... **Likelihood Score (Farr. & Mann.)**  
 Power ..... **0.80**  
 Alpha ..... **0.05**  
 M1 (Average Cluster Size) ..... **100**  
 K2 (Clusters in Group 2) ..... **K1**  
 M2 (Average Cluster Size) ..... **M1**  
 Input Type ..... **Differences**  
 D0 (Difference|H0 = P1.0 - P2) ..... **0.08**  
 D1 (Difference|H1 = P1.1 - P2) ..... **0.10 to 0.16 by 0.01**  
 P2 (Group 2 Proportion) ..... **0.6**  
 ICC (Intraclass 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 vs. H1: P1 - P2 > D0

Power	Number of Clusters			Cluster Size		Total Sample Size N	Proportions			Difference		Intraclass Correlation ICC	Alpha
	K1	K2	K	M1	M2		P1 H0 P1.0	P1 H1 P1.1	P2	Diff H0 D0	Diff H1 D1		
0.80205	84	84	168	100	100	16800	0.68	0.70	0.6	0.08	0.10	0.002	0.05
0.80167	37	37	74	100	100	7400	0.68	0.71	0.6	0.08	0.11	0.002	0.05
0.80764	21	21	42	100	100	4200	0.68	0.72	0.6	0.08	0.12	0.002	0.05
0.82446	14	14	28	100	100	2800	0.68	0.73	0.6	0.08	0.13	0.002	0.05
0.80122	9	9	18	100	100	1800	0.68	0.74	0.6	0.08	0.14	0.002	0.05
0.82400	7	7	14	100	100	1400	0.68	0.75	0.6	0.08	0.15	0.002	0.05
0.80374	5	5	10	100	100	1000	0.68	0.76	0.6	0.08	0.16	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 – Validation

This procedure uses the same mechanics as the Tests for Two Proportions in a Cluster-Randomized Design procedure. We refer the user to Example 4 of Chapter 230 for the validation.