Chapter 153

Tests for Two Proportions in a Cluster-Randomized Design with Clustering in Only One Arm

Introduction

Cluster-randomized designs are those in which whole clusters of subjects (classes, hospitals, communities, etc.) are put into a treatment group or a control group. In this case, the proportions of the two arms (groups), where the first arm is made up of K_i clusters of M_{ij} individuals each and the second arm is made up of individuals, are to be tested using a modified *z* test.

In the trials analyzed by this procedure, subjects are treated together (nested) within clusters in one arm, but not in the other. In the other arm, subjects are treated individually. For example, one arm might receive individual intervention (such as medicine) while the other receives this medicine plus some type of group therapy session. The fact that they attend this therapy session implies that they are "clustered". Thus the clustering occurs in only one arm.

Technical Details

Our formulation comes from Moerbeek and Wong (2008). They combine two mixed models: one for the arm made up of individuals (group 2) and a second, more complicated model for the arm (group 1) which accounts for the clustering. Let y_i be a binary variable that is one of a certain outcome is observed and zero otherwise. The treatment effect is measured by $D = P_1 - P_2$. This can be estimated by

$$\widehat{D} = \bar{y}_1 - \bar{y}_2$$

where $\bar{y}_1$ and $\bar{y}_2$ are the estimators of the probabilities of the outcome $y_i = 1$, P_1 and P_2.

The treatment effect may be tested for statistical significance using

$$\widehat{D}/\sqrt{\mathrm{var}(\widehat{D})}$$

The variance of this estimator is

$$\mathrm{var}(\widehat{D}) = P_2(1-P_2)\left(\tau\frac{(\overline{m}-1)\rho+1}{\overline{m}k}+\frac{1}{N_2}\right)$$

where $\overline{m}$ is the average cluster size, N_2 is the number of subjects in the non-clustered arm, ρ is intracluster correlation coefficient, and τ is the ratio of the variances of the outcome when $\rho = 0$,

$$\tau = \frac{P_1(1-P_1)}{P_2(1-P_2)}.$$

Assume that $D = P_1 - P_2$ is to be tested using a z-test. The statistical hypotheses are $H_0: D = 0$ vs. $H_a: D \neq 0$. The test statistic

$$t = \frac{\overline{Y}_1 - \overline{Y}_2}{\sqrt{\mathrm{var}(\widehat{D})}}$$

has an approximate normal distribution for a *subject-level* analysis.

Using the above, a large sample formula for computing the power of a two-sided test of the significance of the treatment effect at significance level α can be derived from

$$\mathrm{var}(\widehat{D}) = \left(\frac{D}{z_{1-\frac{\alpha}{2}} + z_{Power}}\right)^2$$

The power of a one-sided test can be calculated similarly.

Example 1 – Calculating Sample Size

Suppose that a cluster randomized study is to be conducted in which one arm (group 2) will receive an individual medical intervention while the other arm (group 1) receives this medicine plus a special group therapy session conducted by a trained therapist. These therapy sessions will be treated as clusters. Here, group 2 (the non-clustered subjects) is assigned to the 'control group' and group 1 (the clustered subjects) is assigned to the treatment group. The researchers want to explore what happens as *R* is varied from 1 to 2.

The parameter values are set as follows: *alpha* = 0.05, *power* = 0.9, R = 1.0 1.5 2.0, *P1* = 0.25, *P2* = 0.4, *ρ* = 0.01, and *M1* = 10. Sample size is to be calculated for a two-sided 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 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	**Sample Size**
Alternative Hypothesis	**Two-Sided (H1: P1 - P2 ≠ 0)**
Power	**0.90**
Alpha	**0.05**
Group Allocation	**Enter R = (K1 × M1) / N2, solve for K1 and N2**
R (Allocation Ratio)	**1 1.5 2**
M1 (Average Cluster Size)	**10**
Input Type	**Proportions**
P1 (Group 1 Proportion|H1)	**0.25**
P2 (Group 2 Proportion)	**0.4**
ρ (Intracluster Correlation, ICC)	**0.01**

Output

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

Numeric Reports

Numeric Results

Solve For: Sample Size
Clustering: Group 1 has Clustering, Group 2 has No Clustering
Group Allocation: Enter R = (K1 × M1) / N2, solve for K1 and N2
Hypotheses: H0: D = 0 vs. H1: D ≠ 0

	Group 1 Clusters		Sample Size (Subjects)				Proportions				
Power	Number of Clusters K1	Average Cluster Size M1	Group 1 N1	Group 2 N2	Total N	Allocation Ratio R	P1	P2	Difference D1	ICC ρ	Alpha
0.90326	21	10	210	210	420	1.0	0.25	0.4	-0.15	0.01	0.05
0.90665	27	10	270	180	450	1.5	0.25	0.4	-0.15	0.01	0.05
0.90027	32	10	320	160	480	2.0	0.25	0.4	-0.15	0.01	0.05

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
K1 The number of clusters in group 1.
M1 The average cluster size (number of subjects) of the clusters in group 1.
N1, N2, and N The number of subjects in groups 1 and 2, and their total. Note that N1 = K1 × M1.
R The target value of the allocation ratio of the number of subjects in group 1 and number of subjects in group 2. R = (K1 × M1) / N2. The actual ratio value may not be exactly R because N1 and N2 are integers.
P1 The value of the response proportion of group 1 assumed by the alternative hypothesis, H1.
P2 The value of the response proportion of group 2 assumed by both H0 and H1.
D1 The difference in the response proportions at which the power is calculated. D1 = P1 - P2.
ρ The intracluster correlation (ICC). The correlation between a pair of subjects within a cluster.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A parallel, two-group design with cluster-randomized subjects in Group 1 only (and no clustering in Group 2) will be used to test whether the Group 1 proportion (P1) is different from the Group 2 proportion (P2) (H0: D = 0 versus H1: D ≠ 0, D = P1 - P2). The comparison will be made using a two-sided mixed model Z-test with a Type I error rate (α) of 0.05. The intracluster correlation coefficient for Group 1 is assumed to be 0.01. To detect a difference (P1 - P2) of -0.15 (P1 = 0.25 and P2 = 0.4), with 10 subjects per cluster in Group 1 and 210 subjects in Group 2, with 90% power, the number of needed clusters in Group 1 is 21 (totaling 210 subjects in Group 1).

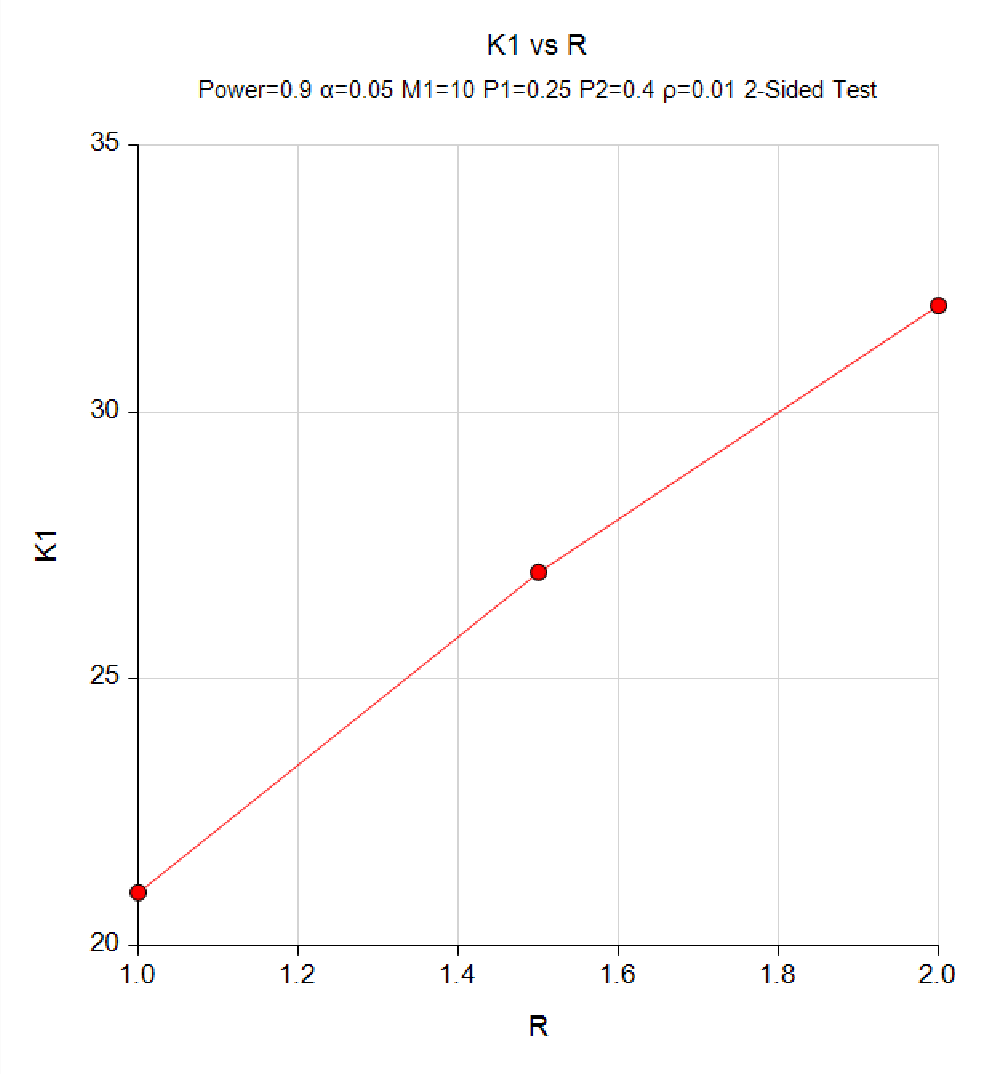
References

Moerbeek, M. and Wong, W.K. 2008. 'Sample size formulae for trials comparing group and individual treatments in a multilevel model.' Statistics in Medicine, Vol. 27, pages 2850-2864.
Donner, A. and Klar, N. 2000. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold. London.

This report shows the results for each of the scenarios.

Plots Section

Plots



This plot shows the number of clusters versus the sample size allocation ratios.

Example 2 – Validation using Moerbeek and Wong (2008)

Moerbeek and Wong (2008) pages 2858 and 2859 provide an example which we will used to validate this procedure. When alpha is 0.05, *P2* is 0.243, P1 is 0.397, ρ is 0.05, N2 is 146, M1 is 8, and K1 is 23, they calculate a power of approximately 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 .... **Power**
Alternative Hypothesis .... **Two-Sided (H1: P1 - P2 ≠ 0)**
Alpha.. **0.05**
Group Allocation .... **Enter K1 and N2 individually**
K1 (Number of Clusters) **23**
M1 (Average Cluster Size)............................ **8**
N2 (Number of Subjects) **146**
Input Type... **Proportions**
P1 (Group 1 Proportion|H1) **0.397**
P2 (Group 2 Proportion)................................ **0.243**
ρ (Intracluster Correlation, ICC)................... **0.05**

Output

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

Numeric Results

Solve For: Power
Clustering: Group 1 has Clustering, Group 2 has No Clustering
Group Allocation: Enter K1 and N2 individually
Hypotheses: H0: D = 0 vs. H1: D ≠ 0

	Group 1 Clusters		Sample Size (Subjects)			Proportions				
Power	Number of Clusters K1	Average Cluster Size M1	Group 1 N1	Group 2 N2	Total N	P1	P2	Difference D1	ICC ρ	Alpha
0.80062	23	8	184	146	330	0.397	0.243	0.154	0.05	0.05

PASS calculates a power of approximately 0.80. Thus the procedure is validated.