Chapter 482

Tests for Two Means in a Cluster-Randomized Design

Introduction

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

Technical Details

Our formulation comes from Campbell and Walters (2014) and Ahn, Heo, and Zhang (2015). Denote an observation by Y_{ijk} where i = 1, 2 gives the group, j = 1, 2, ..., K_i gives the cluster within group i, and k = 1, 2, ..., m_{ij} denotes an individual in cluster j of group i.

We let σ^2 denote the variance of Y_{ijk}, which is $\sigma^2_{Between} + \sigma^2_{Within}$, where $\sigma^2_{Between}$ is the variation between clusters and σ^2_{Within} is the variation within clusters. Also, let ρ denote the intracluster correlation coefficient (ICC) which is $\sigma^2_{Between}/\left(\sigma^2_{Between} + \sigma^2_{Within}\right)$. This correlation is the simple correlation between any two observations in the same cluster.

For sample size calculation, we assume that the m_{ij} are distributed with a mean cluster size of M_i and a coefficient of variation cluster sizes of *COV*. The variance of the two group means, $\bar{Y}_i$, are approximated by

$$V_i = \frac{\sigma^2(DE_i)(RE_i)}{K_i M_i}$$

$$DE_i = 1 + (M_i - 1)\rho$$

$$RE_i = \frac{1}{1 - (COV)^2\lambda_i(1 - \lambda_i)}$$

$$\lambda_i = M_i\rho/(M_i\rho + 1 - \rho)$$

DE is called the *Design Effect* and RE is the *Relative Efficiency* of unequal to equal cluster sizes. Both are greater than or equal to one, so both inflate the variance.

Assume that $\delta = \mu_1 - \mu_2$ is to be tested using a t-test (small sample) or z-test (large sample). The statistical hypotheses are $H_0: \delta = 0$ vs. $H_a: \delta \neq 0$. The test statistic

$$t = \frac{\bar{Y}_1 - \bar{Y}_2}{\sqrt{\hat{V}_1 + \hat{V}_2}}$$

has an approximate t distribution with degrees of freedom $DF = K_1M_1 + K_2M_2 - 2$ for a *subject-level* analysis or $K_1 + K_2 - 2$ for a *cluster-level* analysis.

Let the noncentrality parameter $\Delta = \delta/\sigma_d$, where $\sigma_d = \sqrt{V_1 + V_2}$. We can define the two critical values based on a central t-distribution with DF degrees of freedom as follows.

$$X_1 = t_{\frac{\alpha}{2},DF}$$

$$X_2 = t_{1-\frac{\alpha}{2},DF}$$

The power can be found from the following to probabilities

$$P_1 = H_{X_1,DF,\Delta}$$

$$P_2 = H_{X_2,DF,\Delta}$$

$$\text{Power} = 1 - (P_2 - P_1)$$

where $H_{X,DF,\Delta}$ is the cumulative probability distribution of the noncentral-t distribution.

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

Example 1 – Calculating Power

Suppose that a cluster randomized study is to be conducted in which δ = 1; σ = 2; ρ = 0.01; *M1* and *M2* = 5 or 10; *COV* = 0.65; *alpha* = 0.05; and *K1* and *K2* = 5 to 20 by 5. Power 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	**Power**
Ha (Alternative Hypothesis)	**Two-Sided (Ha: δ ≠ 0)**
Test Statistic	**T-Test Based on Number of Subjects**
Alpha	**0.05**
K1 (Number of Clusters)	**5 10 15 20**
M1 (Average Cluster Size)	**5 10**
K2 (Number of Clusters)	**K1**
M2 (Average Cluster Size)	**M1**
COV of Cluster Sizes	**0.65**
δ (Mean Difference = μ1 - μ2)	**1**
σ (Standard Deviation)	**2**
ρ (Intracluster Correlation, ICC)	**0.01**

Output

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

Numeric Reports

Numeric Results for a Test of Mean Difference

Solve For: Power
Groups: 1 = Treatment, 2 = Control
Test Statistic: T-Test with DF based on number of subjects
Hypotheses: H0: δ = 0 vs. H1: δ ≠ 0

	Number of Clusters			Cluster Size			Sample Size			Mean Difference	Standard Deviation	ICC	
Power	K1	K2	K	M1	M2	COV	N1	N2	N	δ	σ	ρ	Alpha
0.3908	5	5	10	5	5	0.65	25	25	50	1	2	0.01	0.05
0.6439	5	5	10	10	10	0.65	50	50	100	1	2	0.01	0.05
0.6714	10	10	20	5	5	0.65	50	50	100	1	2	0.01	0.05
0.9115	10	10	20	10	10	0.65	100	100	200	1	2	0.01	0.05
0.8399	15	15	30	5	5	0.65	75	75	150	1	2	0.01	0.05
0.9822	15	15	30	10	10	0.65	150	150	300	1	2	0.01	0.05
0.9274	20	20	40	5	5	0.65	100	100	200	1	2	0.01	0.05
0.9969	20	20	40	10	10	0.65	200	200	400	1	2	0.01	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 and 2, and their total.
M1 and M2	The average number of items (subjects) per cluster in groups 1 and 2, respectively.
COV	The coefficient of variation of the cluster sizes.
N1, N2, and N	The number of subjects in groups 1 and 2, and their total.
δ	The mean difference in the response at which the power is calculated. δ = μ1 - μ2.
σ	The standard deviation of the subject responses.
ρ	The intra-cluster 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 cluster-randomized design will be used to test whether the Group 1 (treatment) mean (μ1) is different from the Group 2 (control) mean (μ2) (H0: δ = 0 versus H1: δ ≠ 0, δ = μ1 - μ2). The comparison will be made using a two-sided t-test with the degrees of freedom based on the total number of subjects (see Campbell and Walters, 2014, and Ahn, Heo, and Zhang, 2015), with a Type I error rate (α) of 0.05. The common subject-to-subject standard deviation for both groups is assumed to be 2, the intracluster correlation coefficient is assumed to be 0.01, and the coefficient of variation of cluster sizes is assumed to be 0.65. To detect a mean difference (μ1 - μ2) of 1, with 5 clusters of 5 subjects per cluster in Group 1 (totaling 25 subjects) and 5 clusters of 5 subjects per cluster in Group 2 (totaling 25 subjects), the power is 0.3908.

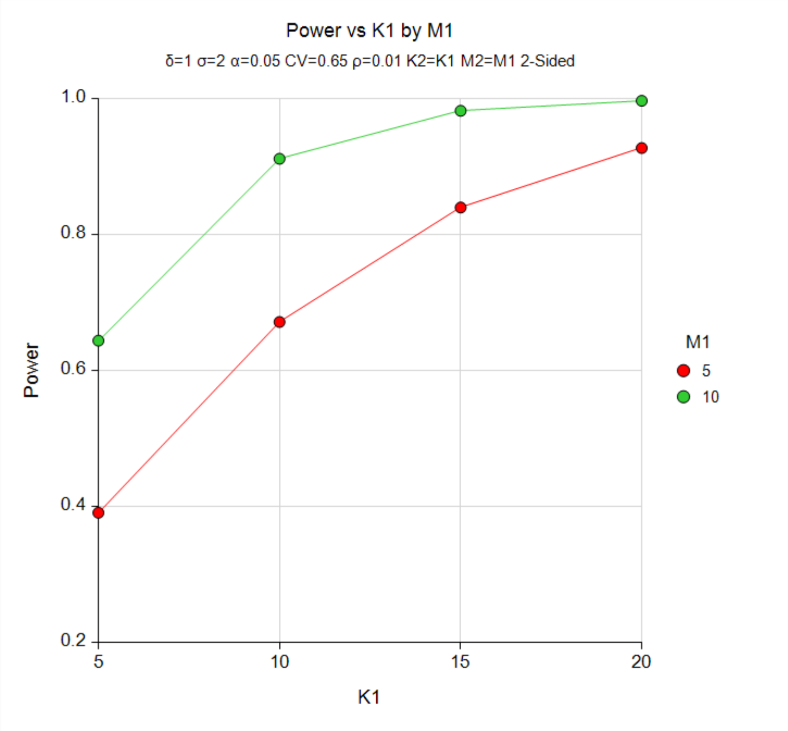
References

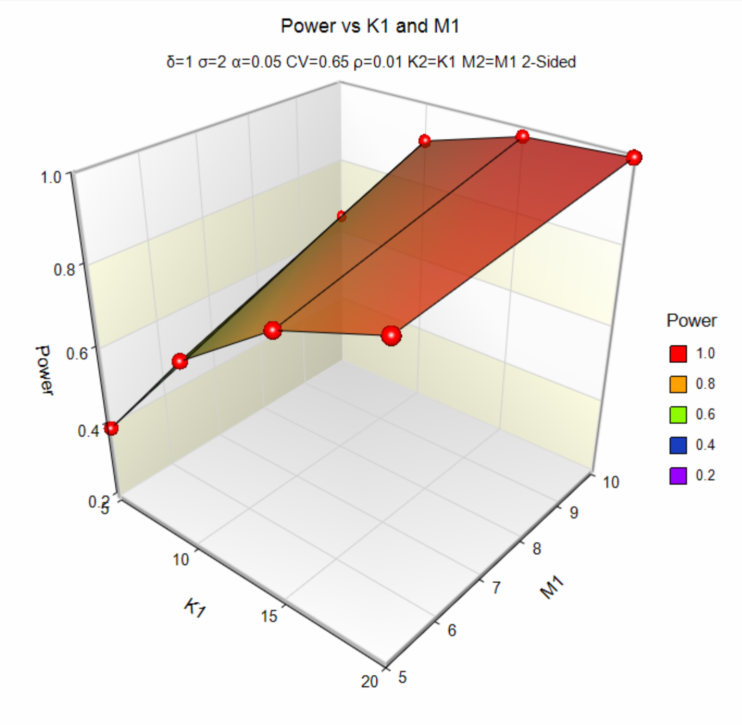
Ahn, C., Heo, M., and Zhang, S. 2015. Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research. CRC Press. New York.
Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. 2018. Sample Size Calculations in Clinical Research, 3rd Edition. Chapman & Hall/CRC. Boca Raton, FL. Pages 86-88.
Campbell, M.J. and Walters, S.J. 2014. How to Design, Analyse and Report Cluster Randomised Trials in Medicine and Health Related Research. Wiley. New York.
Donner, A. and Klar, N. 1996. 'Statistical Considerations in the Design and Analysis of Community Intervention Trials'. J. Clin. Epidemiol. Vol 49, No. 4, pages 435-439.
Donner, A. and Klar, N. 2000. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold. London.

This report shows the power for each of the scenarios.

Plots Section

Plots





These plots show the power versus the cluster counts for the two M1 values.

Example 2 – Validation using Donner and Klar (1996)

Donner and Klar (1996) page 436 provide a table in which several power values are calculated for a two-sided test. When alpha is 0.05, δ is 0.2, ρ is 0.001, σ is 1.0, and $K1 = K2 = 3$, they calculate a power of 0.43 for an $M1=M2$ of 100, 0.79 for an $M1=M2$ of 300, and 0.91 for an $M1=M2$ of 500. COV = 0.

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**
Ha (Alternative Hypothesis) .... **Two-Sided (Ha: δ ≠ 0)**
Test Statistic .... **T-Test Based on Number of Clusters**
Alpha.. **0.05**
K1 (Number of Clusters) **3**
M1 (Average Cluster Size)............................ **100 300 500**
K2 (Number of Clusters) **K1**
M2 (Average Cluster Size)............................ **M1**
COV of Cluster Sizes.................................... **0**
δ (Mean Difference = μ1 - μ2)....................... **0.2**
σ (Standard Deviation).................................. **1**
ρ (Intracluster Correlation, ICC)..................... **0.001**

Output

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

Numeric Results for a Test of Mean Difference

Solve For: Power
Groups: 1 = Treatment, 2 = Control
Test Statistic: T-Test with DF based on number of clusters
Hypotheses: H0: δ = 0 vs. H1: δ ≠ 0

	Number of Clusters			Cluster Size			Sample Size			Mean Difference	Standard Deviation	ICC	
Power	K1	K2	K	M1	M2	COV	N1	N2	N	δ	σ	ρ	Alpha
0.4301	3	3	6	100	100	0	300	300	600	0.2	1	0.001	0.05
0.7924	3	3	6	300	300	0	900	900	1800	0.2	1	0.001	0.05
0.9091	3	3	6	500	500	0	1500	1500	3000	0.2	1	0.001	0.05

PASS calculates the same power values as Donner and Klar (1996).

Example 3 – Validation using Campbell and Walters (2014)

Campbell and Walters (2014) page 71 give an example of the effect of variable cluster size when a subject-level analysis is run. In this example, they set alpha to 0.05, and ρ to 0.05. They don't specifically give values for the mean and standard deviation. By trial and error, we found that for a power of 0.9, δ of 0.3247 and σ of 1 gave results that matched theirs for a two-sided test. They indicate that when COV = 0, K1 = 29. When COV = 0.725, K1 = 33.

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	**K1 (Number of Clusters)**
Ha (Alternative Hypothesis)	**Two-Sided (Ha: δ ≠ 0)**
Test Statistic	**T-Test Based on Number of Subjects**
Power	**0.90**
Alpha	**0.05**
M1 (Average Cluster Size)	**10**
K2 (Number of Clusters)	**K1**
M2 (Average Cluster Size)	**M1**
COV of Cluster Sizes	**0 0.725**
δ (Mean Difference = μ1 - μ2)	**0.3247**
σ (Standard Deviation)	**1**
ρ (Intracluster Correlation, ICC)	**0.05**

Output

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

Numeric Results for a Test of Mean Difference

Solve For: K1 (Number of Clusters)
Groups: 1 = Treatment, 2 = Control
Test Statistic: T-Test with DF based on number of subjects
Hypotheses: H0: δ = 0 vs. H1: δ ≠ 0

	Number of Clusters			Cluster Size			Sample Size			Mean Difference	Standard Deviation	ICC	
Power	K1	K2	K	M1	M2	COV	N1	N2	N	δ	σ	ρ	Alpha
0.9000	29	29	58	10	10	0.000	290	290	580	0.32	1	0.05	0.05
0.9016	33	33	66	10	10	0.725	330	330	660	0.32	1	0.05	0.05

PASS calculates the same values of K1 as Campbell and Walters (2014).