

Chapter 231

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Introduction

A *parallel cluster (group) randomized design* is one in which whole units, or clusters, of subjects are randomized to the treatment or control group 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, hospitals, etc.

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.

A *2×2 cluster-randomized cross-over design* is one in which each cluster receives both the treatment and control. The objective is to study the difference between the two. Each cluster crosses over from the treatment group to the control group (or vice-versa). It is assumed that there is a washout period between applications during which the response returns back to its baseline value. If this does not occur, there is said to be a carry-over effect.

A *stepped-wedge cluster-randomized design* is similar to a cross-over design in that each cluster receives both the treatment and control over time. In a stepped-wedge design, however, the clusters switch or cross-over in one direction only (usually from the control group to the treatment group). Once a cluster is randomized to the treatment group, it continues to receive the treatment for the duration of the study. In a typical stepped-wedge design, all of the clusters are assigned to the control group at the first time point and then individual clusters are progressively randomized to the treatment group over time. The stepped-wedge design is particularly useful for cases where it is logistically impractical to apply a particular treatment to half of the clusters at the same time.

This procedure computes power and sample size for tests for the difference between two proportions in cross-sectional stepped-wedge cluster-randomized designs. In cross-sectional designs, different subjects are measured within each cluster at each point in time. No one subject is measured more than once. (This is not to be confused with cohort studies (i.e., repeated measures) where individuals are measured at each point in time. The methods in this procedure should not be used for cohort or repeated measures designs.)

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Three Types of Cluster-Randomized Designs

In each design pattern matrix below, "0" represents the control and "1" represents the treatment.

Parallel		2x2 Cross-Over			Stepped-Wedge								
	Time		Time			Time							
	1		1	2		1	2	3	4	5	6	7	
Cluster	1	1	1	0	Cluster	1	0	1	1	1	1	1	
	2	1	2	1		0	2	0	0	1	1	1	1
	3	1	3	1		0	3	0	0	0	1	1	1
	4	0	4	0		1	4	0	0	0	0	1	1
	5	0	5	0		1	5	0	0	0	0	0	1
	6	0	6	0		1	6	0	0	0	0	0	0

Stepped-Wedge Cluster-Randomized Designs

There are two types of stepped-wedge designs that can be analyzed by this procedure: complete and incomplete (or custom).

Complete

In complete stepped-wedge designs, clusters sequentially switch from control to treatment in a balanced fashion over a fixed number of time periods (T). Once a cluster switches from control to treatment, the cluster continues to receive the treatment at each time point for the duration of the study. The number of clusters (K) is equal to number steps (S) multiplied by the number of clusters switching at each step (R), that is

$$K = S \times R.$$

The following stepped-wedge design pattern matrices illustrate complete designs (0 = Control, 1 = Treatment):

Complete Design		K = 6 Clusters, S = 6 Steps, T = 7 Time Periods, R = 1 Switch per Step						
	Cluster	Time						
		1	2	3	4	5	6	7
	1	0	1	1	1	1	1	1
	2	0	0	1	1	1	1	1
	3	0	0	0	1	1	1	1
	4	0	0	0	0	1	1	1
	5	0	0	0	0	0	1	1
	6	0	0	0	0	0	0	1

Complete Design		K = 6 Clusters, S = 3 Steps, T = 4 Time Periods, R = 2 Switches per Step			
	Cluster	Time			
		1	2	3	4
	1	0	1	1	1
	2	0	1	1	1
	3	0	0	1	1
	4	0	0	1	1
	5	0	0	0	1
	6	0	0	0	1

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Incomplete (Custom)

In incomplete (or custom) stepped-wedge designs, there is no balance required. Furthermore, incomplete designs also allow for time periods at which no observations are made. The only requirement is that once a cluster switches from control to treatment, the cluster continues to receive the treatment if and when an observation is made.

When specifying incomplete designs using the number of clusters (K) and the number steps (S) (or the number of time periods (T)), **PASS** searches among all possible design configurations that satisfy the design constraints (K, S, T) to find an optimal design that achieves the highest power. This search is controlled by the Incomplete (Custom) Design Pattern Matrix Search Options on the Options tab.

The following stepped-wedge design pattern matrices illustrate incomplete designs (0 = Control, 1 = Treatment, • = No Observation):

Incomplete Design K = 6 Clusters, S = 4 Steps, T = 5 Time Periods						
	Time					
	1	2	3	4	5	
Cluster	1	0	1	1	1	1
	2	0	1	1	1	1
	3	0	0	1	1	1
	4	0	0	0	1	1
	5	0	0	0	1	1
	6	0	0	0	0	1

Incomplete Design K = 5 Clusters, S = 6 Steps, T = 7 Time Periods							
	Time						
	1	2	3	4	5	6	7
Cluster	1	0	•	1	1	1	1
	2	0	0	•	1	1	1
	3	0	0	0	•	1	1
	4	0	0	0	0	•	1
	5	0	0	0	0	0	•

Incomplete Design K = 4 Clusters, S = 5 Steps, T = 6 Time Periods						
	Time					
	1	2	3	4	5	6
Cluster	1	0	1	1	•	•
	2	•	0	1	1	•
	3	•	•	0	1	1
	4	•	•	•	0	1

Technical Details

This procedure is based on the results outlined in Hussey and Hughes (2007). In the technical discussions that follow, we will adopt the following notation:

K	Number of clusters
S	Number of steps in the stepped-wedge design
T	Number of time periods in the stepped-wedge design
M	Number of subjects (or items) per cluster
m	Number of subjects (or items) per cluster per time period
N	Total number of subjects (or items) from all clusters and all time periods combined
P_1	The treatment group proportion, assuming the alternative hypothesis
P_2	The control, standard, reference, or baseline group proportion

The Linear Model

Linear mixed models are often used to model stepped-wedge cluster-randomized designs with time as a fixed factor at T levels and inter-cluster variation modeled as a random effect. As described in Hussey and Hughes (2007) and Hemming, Lilford, and Girling (2015) but using our notation and for a cross-sectional stepped-wedge design, let Y_{ikt} represent the response from individual i ($i = 1, \dots, m$) in cluster k ($k = 1, \dots, K$) at time t ($t = 1, \dots, T$). The linear mixed model can then be written as

$$Y_{ikt} = X_{kt}\theta + \alpha_k + \beta_t + e_{ikt}$$

$$\alpha_k \sim N(0, \tau^2)$$

$$e_{ikt} \sim N(0, \sigma_w^2)$$

with

$$\text{Var}(Y_{ikt}) = \sigma_y^2 = \tau^2 + \sigma_w^2,$$

where

X_{kt}	Indicator variable of the group assignment of cluster k at time t , with 0 = control and 1 = treatment
θ	Fixed treatment effect
α_k	Random effect for cluster k
β_t	Fixed effect for time t
e_{ikt}	Within-cluster error

The Treatment Effect

θ is the treatment effect and is equal to $D1$, the difference between the treatment mean, μ_1 , and the control group mean, μ_2 , such that

$$\theta = D1 = \mu_1 - \mu_2.$$

In the case of testing two proportions, with $\mu_1 = P_1$ and $\mu_2 = P_2$ as the treatment and control group proportions, respectively, then

$$\theta = D1 = P_1 - P_2.$$

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Between- and Within-Cluster Variation

From the model above, τ^2 is the between-cluster variance and σ_w^2 is the within-cluster variance. If τ^2 and σ_w^2 are known (as is generally assumed in power and sample size calculations), the model for the cell mean of cluster k at time t can then be written as

$$\bar{Y}_{kt} = X_{kt}\theta + \alpha_k + \beta_t + e_{kt}$$

$$\alpha_k \sim N(0, \tau^2)$$

$$e_{kt} \sim N\left(0, \frac{\sigma_w^2}{m}\right)$$

with

$$\text{Var}(\bar{Y}_{kt}) = \tau^2 + \frac{\sigma_w^2}{m}.$$

For a complete stepped-wedge design, the $K \times T$ block-diagonal variance-covariance matrix, \mathbf{V} , of cell means has the form

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{V}_k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{V}_K \end{bmatrix},$$

where each $T \times T$ block matrix, \mathbf{V}_k , represents a single cluster and has the form

$$\mathbf{V}_k = \begin{bmatrix} \tau^2 + \frac{\sigma_w^2}{m} & \tau^2 & \cdots & \tau^2 \\ \tau^2 & \tau^2 + \frac{\sigma_w^2}{m} & \cdots & \tau^2 \\ \vdots & \vdots & \ddots & \vdots \\ \tau^2 & \tau^2 & \cdots & \tau^2 + \frac{\sigma_w^2}{m} \end{bmatrix}.$$

An incomplete stepped-wedge design has a similar variance-covariance matrix structure with different dimensions depending on the incomplete design pattern matrix.

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

ICC (Intraclass Correlation Coefficient)

The correlation structure for the stepped-wedge cluster-randomized design is often characterized by the intraclass correlation coefficient, or ICC. The ICC may be interpreted as the correlation between any two observations in the same cluster. It may also be thought of as the proportion of the variation in the response that can be accounted for by the between-cluster variation. The ICC is calculated as

$$ICC = \frac{\tau^2}{\sigma_y^2} = \frac{\tau^2}{\tau^2 + \sigma_w^2}$$

and can be used along with the within-cluster variance, σ_w^2 , to calculate the between-cluster variance, τ^2 , as

$$\tau^2 = \frac{ICC \times \sigma_w^2}{1 - ICC}$$

COV (Coefficient of Variation of Outcomes)

The correlation structure for the stepped-wedge cluster-randomized design can also be characterized by the coefficient of variation, or COV, of outcomes. (Note that this is not the COV of cluster sizes as is often referenced in conjunction with cluster-randomized designs.) If $\mu_2 = P_2$ is the control group mean (i.e., a proportion for binary data), then the COV of outcomes for the control group is calculated as

$$COV = \frac{\tau}{\mu_2} = \frac{\tau}{P_2}$$

and can be used along with $\mu_2 = P_2$ to calculate τ^2 as

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times P_2^2.$$

Specifying the Total and Within-Cluster Variance for a Binary Response

In the case of testing two proportions where the responses, Y_{ikt} , are binary (0 or 1) and using the normal approximation, the variance can be calculated using three different methods, given as options in this procedure.

Method 1:	Null Variance	$\sigma^2 = P_2(1 - P_2)$
Method 2:	Pooled Variance	$\sigma^2 = \left(\frac{P_1 - P_2}{2}\right) \left(1 - \frac{P_1 - P_2}{2}\right)$
Method 3:	Group Variance Average	$\sigma^2 = \frac{P_1(1 - P_1) + P_2(1 - P_2)}{2}$

Method 1 is used in Hussey and Hughes (2007) and Hemming and Girling (2014). Method 2 is used in Baio et al. (2015). Method 3 is used in Hemming, Lilford, and Girling (2015).

The variance that is calculated from the two proportions, σ^2 , may be considered to be either the total variance, σ_y^2 , or the within-cluster variance, σ_w^2 . If σ^2 is considered to be the total variance, σ_y^2 , then using the ICC, the between-cluster variance, τ^2 , can be computed as

$$\tau^2 = ICC \times \sigma_y^2 = ICC \times \sigma^2,$$

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

and using the COV the between-cluster variance, τ^2 , can be computed as

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times P_2^2.$$

and, finally, the within-cluster variance can be calculated as

$$\sigma_w^2 = \sigma_y^2 - \tau^2 = \sigma^2 - \tau^2.$$

Otherwise, if σ^2 is considered to be the within-cluster variance, σ_w^2 , then

$$\sigma_w^2 = \sigma^2$$

and using the ICC, the between-cluster variance, τ^2 , can be computed as

$$\tau^2 = \frac{ICC \times \sigma_w^2}{1 - ICC} = \frac{ICC \times \sigma^2}{1 - ICC}$$

and using the COV, the between-cluster variance, τ^2 , can be computed as

$$\tau^2 = COV^2 \times \mu_2^2 = COV^2 \times P_2^2.$$

Test Statistic

Hussey and Hughes (2007) suggest that a statistical hypothesis test of $H_0: \theta = P_1 - P_2 = 0$ vs. $H_1: \theta = P_1 - P_2 \neq 0$ can be conducted using an asymptotic Wald test. The test statistic is

$$Z = \frac{\hat{\theta}}{\sqrt{\text{Var}(\hat{\theta})}} = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\text{Var}(\hat{\theta})}}.$$

$\hat{\theta}$ is the estimated treatment effect from a weighted least-squares analysis. $\text{Var}(\hat{\theta})$ is the appropriate element of $(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}$ from the weighted least-squares analysis. \mathbf{X} is the design matrix (recall that X_{kt} is an indicator variable of the group assignment of cluster k at time t , e.g., 0 = control and 1 = treatment).

If X_{kt} contains only 0's and 1's, no missing cells, and m is equal for all clusters at all time points then $\text{Var}(\hat{\theta})$ can be written in closed form as

$$\text{Var}(\hat{\theta}) = \frac{K \left(\frac{\sigma_w^2}{m} \right) \left(\frac{\sigma_w^2}{m} + T\tau^2 \right)}{\left(\frac{\sigma_w^2}{m} \right) (KU - W) + \tau^2 (U^2 + KTV - TW - KV)}$$

where

$$U = \sum_{kt} X_{kt}$$

$$V = \sum_k (\sum_t X_{kt})^2$$

$$W = \sum_t (\sum_k X_{kt})^2$$

Power Calculations

The power calculations available in this procedure for both complete and incomplete designs are based on the results outlined in Hussey and Hughes (2007). With τ^2 and σ_w^2 assumed to be known, the power for a two-sided Wald test is computed as

$$\begin{aligned} \text{Power} &= \Phi \left(\frac{\theta_A}{\sqrt{\text{Var}(\hat{\theta})}} - Z_{1-\alpha/2} \right) \\ &= \Phi \left(\frac{P_1 - P_2}{\sqrt{\{(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\}}} - Z_{1-\alpha/2} \right) \end{aligned}$$

where Φ is the cumulative standard Normal distribution, $\{(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\}$ is the appropriate element of the matrix, and $Z_{1-\alpha/2}$ is the $(1 - \alpha/2)$ quantile of the standard Normal distribution function. θ_A is the value for θ under the alternative hypothesis, i.e., $\theta_A = P_1 - P_2$.

When solving for sample size (number of clusters), sample size (cluster size), or effect size, a search is conducted based on this power formula to find a solution that fits all required conditions.

A Note on Power Calculations for Incomplete Designs

When specifying an incomplete stepped-wedge design pattern using the number of clusters (K) and the number of steps (S) or the number of time periods (T), there are a variety of ways that the clusters can be arranged in an actual design and still meet the design criteria. **PASS** utilizes the following logic when creating a design pattern matrix using K and S or T and calculating power:

1. Sequentially assign R complete cluster sets to a design pattern sub-matrix, where R is the largest integer that satisfies $(R \times S) < K$ and $((R + 1) \times S) > K$. (Note: If $(R \times S) = K$, then the design is complete and there are no "Extra" clusters no optimal design pattern search is required.)

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

2. Assign $J = K - (R \times S)$ extra clusters using the following options for the Design Pattern Assignment Type for Extra Clusters:

- **Balanced**

The J "Extra" clusters are assigned in a balanced fashion. When one extra cluster is given a particular design pattern, it cannot be repeated.

Balanced Assignment						
		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	1	1	1	1
	3	0	0	1	1	1
	4	0	0	1	1	1
	5	0	0	0	1	1
	6	0	0	0	0	1
	7	0	0	0	0	1

- **Unbalanced**

The J "Extra" clusters are assigned in a (potentially) unbalanced manner. When one extra cluster is given a particular design pattern, it may be repeated. This option results in maximum power but may result in a very unbalanced design.

Unbalanced Assignment						
		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	1	1	1	1
	3	0	1	1	1	1
	4	0	0	1	1	1
	5	0	0	0	1	1
	6	0	0	0	0	1
	7	0	0	0	0	1

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

- Sequential**

The J "Extra" clusters are assigned sequentially. The first extra cluster is given the same design pattern as the first already-assigned complete cluster, the second extra cluster is given the same design pattern as the second already-assigned complete cluster, and so on. When one extra cluster is given a particular design pattern, it cannot be repeated.

Sequential Assignment						
		Time				
		1	2	3	4	5
Cluster	1	0	1	1	1	1
	2	0	1	1	1	1
	3	0	0	1	1	1
	4	0	0	1	1	1
	5	0	0	0	1	1
	6	0	0	0	1	1
	7	0	0	0	0	1

3. Compute the power for all possible assignment combinations.
4. Return the design pattern matrix with the highest power.

Delayed Treatment Effect

Everything that has been discussed so far assumes that the full effect of the treatment occurs in the same time period in which the treatment is administered. This, however, might not always be the case. The full effect of the treatment may not be realized until several time periods after the treatment is applied. Hussey and Hughes (2007) propose that this situation can be modelled by simply altering the stepped-wedge design pattern matrix to include fractional numbers instead of 0's and 1's, where a fractional value indicates that the treatment is not fully efficacious at a particular time period.

The following is an example of a design pattern matrix exhibiting a delayed treatment effect where the treatment is 50% effective in the first time period, 80% effective in the second time period, and 100% effective by the third time period.

Design Pattern Matrix with a Delayed Treatment Effect								
		Time						
		1	2	3	4	5	6	7
Cluster	1	0	0.5	0.8	1	1	1	1
	2	0	0	0.5	0.8	1	1	1
	3	0	0	0	0.5	0.8	1	1
	4	0	0	0	0	0.5	0.8	1

The test statistic and power calculations are the same as for the case where the design pattern matrix contains all 0's and 1's, except that the closed form representation of $\text{Var}(\hat{\theta})$ cannot be used. In this case, $\text{Var}(\hat{\theta})$ must be calculated using matrix operations. The delay has the overall effect of reducing the power.

Example 1 – Finding Power of a Complete Design (Validation using Hussey and Hughes (2007))

Hussey and Hughes (2007) presents an example related to a proposed program aimed at reducing the number of STDs in the state of Washington. In the study, the current prevalence of Chlamydia infection was $P_2 = 0.05$. The study involved $K = 24$ counties that were to be randomized to the intervention $R = 6$ at a time, such that the number of required time periods was $T = 5$, and the number of steps was $S = 4$. They planned to test $m = 100$ individuals per county per time interval. What kind of power would the study have for COV values of 0.3 and 0.5 and proportion ratios (P_1/P_2) between 0.5 and 0.8, with an alpha level of 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 1a** 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
Alpha.....	0.05
Design Type.....	Complete
Design Parameter Entry Type.....	Clusters (K) & Switches per Step (R)
K (Number of Clusters)	24
R (Number of Clusters Switching at each Step).....	6
Cluster Size Entry Type	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period).....	100
Input Type.....	Ratios
R1 (Ratio = P_1/P_2).....	0.5 to 0.8 by 0.05
P2 (Control Proportion)	0.05
Variance Calculation Formula.....	$\sigma^2 = P_2(1 - P_2)$
Use Calculated Variance as.....	Within-Cluster Variance ($\sigma^2 = \sigma_w^2$)
Between-Cluster Variability Input Type.....	COV
COV (Coefficient of Variation of Outcomes)	0.3 0.5

Reports Tab

ICC, COV, and Variances Decimals	4
--	----------

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Output

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

Numeric Reports

Numeric Results

Solve For: **Power**
 Design Type: Complete
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 / P_2 = 1$ vs. $H_1: P_1 / P_2 \neq 1$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Ratio R1	Coefficient of Variation COV	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.96458	4	5	6	24	500	100	12000	0.0250	0.05	0.50	0.3	0.05
0.94839	4	5	6	24	500	100	12000	0.0250	0.05	0.50	0.5	0.05
0.92361	4	5	6	24	500	100	12000	0.0275	0.05	0.55	0.3	0.05
0.89805	4	5	6	24	500	100	12000	0.0275	0.05	0.55	0.5	0.05
0.85387	4	5	6	24	500	100	12000	0.0300	0.05	0.60	0.3	0.05
0.81900	4	5	6	24	500	100	12000	0.0300	0.05	0.60	0.5	0.05
0.75065	4	5	6	24	500	100	12000	0.0325	0.05	0.65	0.3	0.05
0.70974	4	5	6	24	500	100	12000	0.0325	0.05	0.65	0.5	0.05
0.61788	4	5	6	24	500	100	12000	0.0350	0.05	0.70	0.3	0.05
0.57680	4	5	6	24	500	100	12000	0.0350	0.05	0.70	0.5	0.05
0.46947	4	5	6	24	500	100	12000	0.0375	0.05	0.75	0.3	0.05
0.43445	4	5	6	24	500	100	12000	0.0375	0.05	0.75	0.5	0.05
0.32539	4	5	6	24	500	100	12000	0.0400	0.05	0.80	0.3	0.05
0.30041	4	5	6	24	500	100	12000	0.0400	0.05	0.80	0.5	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the within-cluster variance ($\sigma^2 = \sigma_w^2$) in power computations.

Power The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
 S The number of steps in the study design. $S = T - 1$.
 T The number of time periods in the study, including the baseline. $T = S + 1$.
 R The number clusters switching from control to treatment at each step.
 K The total number of clusters to be randomized. $K = S \times R$.
 M The average number of subjects per cluster. $M = m \times T$.
 m The average number of subjects per cluster per time period. $M = m \times T$.
 N The total sample size from all clusters and time periods combined. $m = M / T$.
 P1 The treatment proportion, assuming the alternative hypothesis.
 P2 The control, standard, reference, or baseline proportion.
 R1 The ratio assuming the alternative hypothesis (H_1). $R1 = P1 / P2$.
 COV The coefficient of variation of outcomes.
 Alpha The probability of rejecting a true null hypothesis.

Summary Statements

A complete stepped-wedge cluster-randomized design with 5 time periods (including the baseline) and 4 steps (with 6 clusters switching from control to treatment at each step) will be used to test whether the Group 1 (treatment) proportion (P_1) is different from the Group 2 (control) proportion (P_2) ($H_0: P_1 / P_2 = 1$ versus $H_1: P_1 / P_2 \neq 1$). The comparison will be made using a two-sided Wald Z-test based on the proportion difference with variance calculation formula $\sigma^2 = P_2(1 - P_2)$ (considered to be the within-cluster variance), and with a Type I error rate (α) of 0.05. The control group proportion (P_2) is assumed to be 0.05. The coefficient of variation of outcomes for the control group is assumed to be 0.3. To detect a proportion ratio (P_1 / P_2) of 0.5 (or P_1 of 0.025) with 24 clusters with average cluster sizes of 500 subjects per cluster and an average of 100 subjects per cluster per time period (for a total sample size of 12000 subjects), the power is 0.96458.

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

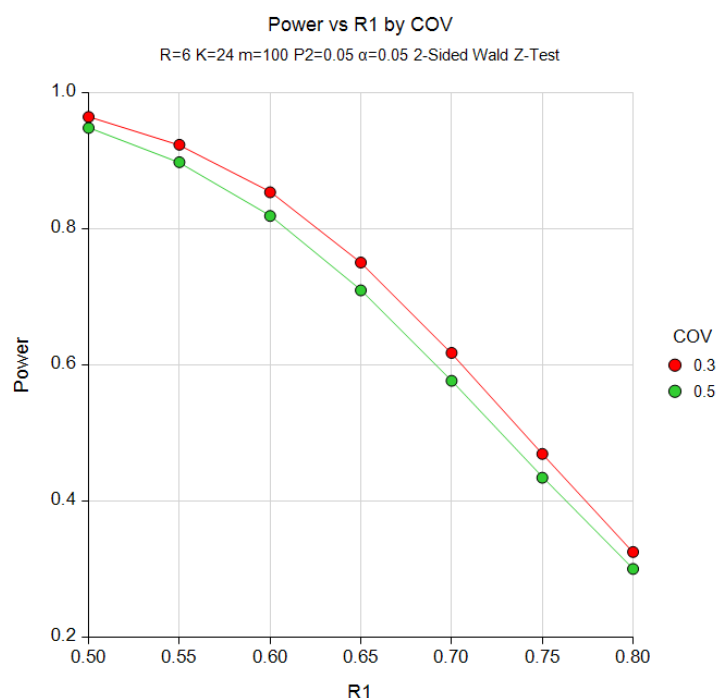
References

- Hussey, M.A., and Hughes, J.P. 2007. 'Design and analysis of stepped wedge cluster randomized trials'. Contemporary Clinical Trials, Volume 28, pages 182-191.
- Hemming, K., and Girling, A. 2014. 'A menu-driven facility for power and detectable-difference calculations in stepped-wedge cluster-randomized trials'. The Stata Journal, Volume 14, pages 363-380.
- Hemming, K., Lilford, R., and Girling A.J. 2015. 'Stepped-wedge cluster randomised controlled trials: a generic framework including parallel and multiple-level designs'. Statistics in Medicine, Volume 34, pages 181-196.
- Baio G., et al. 2015. 'Sample size calculation for a stepped wedge trial'. Trials, 16: 354.
- Hemming, K., and Taljaard, M. 2016. 'Sample size calculations for stepped wedge and cluster randomised trials: a unified approach'. Journal of Clinical Epidemiology, Volume 69, pages 137-146.

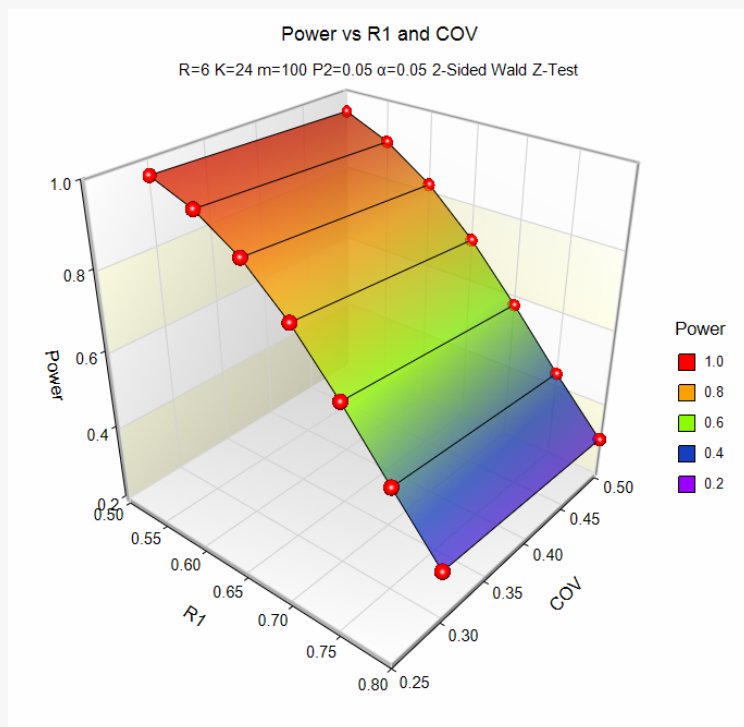
This report shows the values of each of the parameters, one scenario per row. The values from this table are shown in the plot below.

Plots Section

Plots



Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design



The values from the table are displayed on the plots. These plots give a quick look at the power that is achieved for various combinations of the effect size (R_1) and COV. If you look on page 187 of Hussey and Hughes (2007), you'll see that these plots look very similar to the plot in Figure 2, with the X-axis reversed. This is because the effect size in Hussey and Hughes (2007) is equal to $1 - R_1$, as specified here.

Design Details

Design 1 Details

Design Type: Complete

Power	Design Parameters			Number of Clusters K	Cluster Size		Proportion		Variance			Intraclass Correlation Coefficient ICC	Coefficient of Variation COV
	S	T	R		M	m	Treatment P1	Control P2	Total σ^2	Between-Cluster τ^2	Within-Cluster σ^2		
0.96458	4	5	6	24	500	100	0.025	0.05	0.0477	0.0002	0.0475	0.0047	0.3

Design Pattern (Complete)

Cluster	T1	T2	T3	T4	T5
1	0	1	1	1	1
2	0	1	1	1	1
3	0	1	1	1	1
4	0	1	1	1	1
5	0	1	1	1	1
6	0	1	1	1	1
7	0	0	1	1	1
8	0	0	1	1	1
9	0	0	1	1	1
10	0	0	1	1	1
11	0	0	1	1	1
12	0	0	1	1	1

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

13	0	0	0	1	1
14	0	0	0	1	1
15	0	0	0	1	1
16	0	0	0	1	1
17	0	0	0	1	1
18	0	0	0	1	1
19	0	0	0	0	1
20	0	0	0	0	1
21	0	0	0	0	1
22	0	0	0	0	1
23	0	0	0	0	1
24	0	0	0	0	1

(A Design Details report is given for each additional scenario)

This report gives the details about each design for which power was calculated. The design pattern matrix is also printed, showing exactly what design is being analyzed.

Procedure Input Settings

Procedure Input Settings

Autosave Inactive

Design Tab

Solve For	Power
Alternative Hypothesis	Two-Sided
Alpha	0.05
Design Type	Complete
Design Parameter Entry Type	Clusters (K) & Switches per Step (R)
K (Number of Clusters)	24
R (Number of Clusters Switching at each Step)	6
Cluster Size Entry Type	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period)	100
Input Type	Ratios
R1 (Ratio = P1/P2)	0.5 to 0.8 by 0.05
P2 (Control Proportion)	0.05
Variance Calculation Formula	$\sigma^2 = P2(1 - P2)$
Use Calculated Variance as	Within-Cluster Variance ($\sigma^2 = \sigma w^2$)
Between-Cluster Variability Input Type	COV
COV (Coefficient of Variation of Outcomes)	0.3 0.5

Options Tab

Design Pattern Assignment Type for Extra Clusters	Balanced
Maximum Combinations for Incomplete Design Pattern Search	10000

This report displays the critical procedure input settings that were used to generate the report. This report is given primarily for the purpose of record-keeping.

Additional Validation

Hussey and Hughes (2007) reports that for $COV = 0.3$ and $P1 = 0.032$, the power is “about 80%”. To replicate this, you may make the appropriate entries as listed below, or open **Example 1b** by going to the **File** menu and choosing **Open Example Template**.

Design Tab

Solve For **Power**
 Alternative Hypothesis **Two-Sided**
 Alpha..... **0.05**
 Design Type..... **Complete**
 Design Parameter Entry Type..... **Clusters (K) & Switches per Step (R)**
 K (Number of Clusters) **24**
 R (Number of Clusters Switching at each Step)..... **6**
 Cluster Size Entry Type **Subjects per Cluster per Time Period (m)**
 m (Ave. Subjects per Cluster per Time Period)..... **100**
 Input Type..... **Proportions**
 P1 (Treatment Proportion) **0.032**
 P2 (Control Proportion) **0.05**
 Variance Calculation Formula **$\sigma^2 = P2(1 - P2)$**
 Use Calculated Variance as..... **Within-Cluster Variance ($\sigma^2 = \sigma w^2$)**
 Between-Cluster Variability Input Type **COV**
 COV (Coefficient of Variation of Outcomes) **0.3**

Reports Tab

ICC, COV, and Variances Decimals **4**

Output

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

Numeric Results

Solve For: **Power**
 Design Type: **Complete**
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: **Wald Z-Test**
 Hypotheses: **H0: P1 - P2 = 0 vs. H1: P1 - P2 \neq 0**

	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Coefficient of Variation COV	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
Power												
0.77393	4	5	6	24	500	100	12000	0.032	0.05	-0.018	0.3	0.05

Note: The variance is calculated as $\sigma^2 = P2(1 - P2)$ and is considered to be the within-cluster variance ($\sigma^2 = \sigma w^2$) in power computations.

PASS calculated a value of 0.77393 for the power, which matches the statement in Hussey and Hughes (2007) that the power is “about 80%”.

Example 2 – Finding Power of a Complete Design (Validation using Hemming and Girling (2014))

Hemming and Girling (2014) presents an example where power is calculated for a complete design where $P_1 = 0.5$, $P_2 = 0.4$ (Note that in the paper, Proportion 1 is the standard proportion and Proportion 2 is the treatment proportion, so they are reversed), $\alpha = 0.05$, $ICC = 0.01$, $m = 12$, $R = 1$, and $T = 11$ (Note that they state that the number of time periods not including the baseline is 10). They use $\sigma^2 = P_2(1 - P_2)$ as the variance calculation formula and assume it is the total variance.

They compute power for this scenario to be 0.6998, a total sample size of 1320, a τ^2 of 0.0024, a COV of 0.12, and 10 steps in the final design.

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
Alpha.....	0.05
Design Type.....	Complete
Design Parameter Entry Type.....	Time Periods (T) & Switches per Step (R)
T (Number of Time Periods including Baseline).....	11
R (Number of Clusters Switching at each Step).....	1
Cluster Size Entry Type	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period).....	12
Input Type.....	Proportions
P1 (Treatment Proportion)	0.5
P2 (Control Proportion)	0.4
Variance Calculation Formula.....	$\sigma^2 = P_2(1 - P_2)$
Use Calculated Variance as.....	Total Variance ($\sigma^2 = \sigma_y^2 = \tau^2 + \sigma_w^2$)
Between-Cluster Variability Input Type	ICC
ICC (Intraclass Correlation Coefficient)	0.01

Reports Tab

ICC, COV, and Variances Decimals	4
--	----------

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Output

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

Numeric Results

Solve For: [Power](#)
 Design Type: Complete
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 - P_2 = 0$ vs. $H_1: P_1 - P_2 \neq 0$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Intraclass Correlation Coefficient ICC	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.69978	10	11	1	10	132	12	1320	0.5	0.4	0.1	0.01	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

Design Details

Design Type: Complete

Power	Design Parameters			Number of Clusters K	Cluster Size		Proportion		Variance			Intraclass Correlation Coefficient ICC	Coefficient of Variation COV
	S	T	R		M	m	Treatment P1	Control P2	Total σ_y^2	Between-Cluster τ^2	Within-Cluster σ_w^2		
0.69978	10	11	1	10	132	12	0.5	0.4	0.24	0.0024	0.2376	0.01	0.1225

Design Pattern (Complete)

Cluster	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
1	0	1	1	1	1	1	1	1	1	1	1
2	0	0	1	1	1	1	1	1	1	1	1
3	0	0	0	1	1	1	1	1	1	1	1
4	0	0	0	0	1	1	1	1	1	1	1
5	0	0	0	0	0	1	1	1	1	1	1
6	0	0	0	0	0	0	1	1	1	1	1
7	0	0	0	0	0	0	0	1	1	1	1
8	0	0	0	0	0	0	0	0	1	1	1
9	0	0	0	0	0	0	0	0	0	1	1
10	0	0	0	0	0	0	0	0	0	0	1

The power value of 0.69978 calculated by **PASS** matches the result of Hemming and Girling (2014) exactly, as do the values for N, τ^2 , COV, and S. The design pattern shows exactly what the design looks like.

The design pattern matrix shown on page 191 is as follows:

[illegible]

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 3a** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Solve For	Power
Alternative Hypothesis	Two-Sided
Alpha.....	0.05
Design Type.....	Incomplete (Custom)
Design Parameter Entry Type.....	Custom Design Pattern Matrix
Design Pattern Matrix Columns	ALL

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

R (Number of Design Pattern Replicates)..... **1**
 Cluster Size Entry Type **Subjects per Cluster per Time Period (m)**
 m (Ave. Subjects per Cluster per Time Period)..... **12**
 Input Type..... **Proportions**
 P1 (Treatment Proportion) **0.5**
 P2 (Control Proportion) **0.4**
 Variance Calculation Formula..... **$\sigma^2 = P2(1 - P2)$**
 Use Calculated Variance as..... **Total Variance ($\sigma^2 = \tau^2 + \sigma w^2$)**
 Between-Cluster Variability Input Type..... **ICC**
 ICC (Intraclass Correlation Coefficient) **0.01**

Input Spreadsheet Data

Section 1

Row	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
1	0	0	0	0	0	0	0	0	0	0	0	0	.	1
2			0	0	0	0	0	0	0	0	0	0	0	0
3				0	0	0	0	0	0	0	0	0	0	0
4					0	0	0	0	0	0	0	0	0	0
5						0	0	0	0	0	0	0	0	0
6							0	0	0	0	0	0	0	0
7								0	0	0	0	0	0	0
8									0	0	0	0	0	0
9										0	0	0	0	0
10											0	0	0	0

Section 2

Row	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	C25	C26	C27
1	1	1	1	1	1	1	1	1	1	1	1		
2	.	1	1	1	1	1	1	1	1	1	1	1	1
3	0	.	1	1	1	1	1	1	1	1	1	1	1
4	0	0	.	1	1	1	1	1	1	1	1	1	1
5	0	0	0	.	1	1	1	1	1	1	1	1	1
6	0	0	0	0	.	1	1	1	1	1	1	1	1
7	0	0	0	0	0	0	.	1	1	1	1	1	1
8	0	0	0	0	0	0	0	0	.	1	1	1	1
9	0	0	0	0	0	0	0	0	0	0	0	.	1
10	0	0	0	0	0	0	0	0	0	0	0	0	.

Section 3

Row	C28	C29	C30	C31	C32	C33	C34	C35	C36	C37	C38	C39
1												
2												
3	1											
4	1	1										
5	1	1	1									
6	1	1	1	1								
7	1	1	1	1	1	1						
8	1	1	1	1	1	1	1	1				
9	1	1	1	1	1	1	1	1	1	1	1	
10	1	1	1	1	1	1	1	1	1	1	1	1

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Output

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

Numeric Results

Solve For: [Power](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 - P_2 = 0$ vs. $H_1: P_1 - P_2 \neq 0$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Intraclass Correlation Coefficient ICC	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.78826	38	39	1	10	288	12	2880	0.5	0.4	0.1	0.01	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

The power value of 0.78826 and total sample size of $N = 2880$ calculated by **PASS** matches the results in Hemming, Lilford, and Girling (2014).

Now, change the variance calculation formula to $\sigma^2 = [P_1(1 - P_1) + P_2(1 - P_2)]/2$ or open **Example 3b** by going to the **File** menu and choosing **Open Example Template**.

Design Tab

Variance Calculation Formula..... $\sigma^2 = [P_1(1 - P_1) + P_2(1 - P_2)]/2$

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

Numeric Results

Solve For: [Power](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 - P_2 = 0$ vs. $H_1: P_1 - P_2 \neq 0$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Intraclass Correlation Coefficient ICC	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.77997	38	39	1	10	288	12	2880	0.5	0.4	0.1	0.01	0.05

Note: The variance is calculated as $\sigma^2 = [P_1(1 - P_1) + P_2(1 - P_2)]/2$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

The power value of 0.77997 and total sample size of $N = 2880$ calculated by **PASS** also matches the results in Hemming, Lilford, and Girling (2014).

Example 4 – Finding Sample Size (Number of Clusters) for an Incomplete Design (Validation using Baio et al. (2015))

Baio et al. (2015) presents an example in Table 1 of various sample size calculations for the case where the baseline probability, P_2 , is 0.26 and the odds ratio, OR_1 , is 0.56. They assume $m = 20$ individuals per cluster per time interval, $T = 6$ time intervals, and compute the required sample sizes for 80% power with ICC values from 0 to 0.5. The alpha level is 0.05. Baio et al. (2015) uses the variance calculation formula of $\sigma^2 = [(P_1 + P_2)/2] \times [1 - (P_1 + P_2)/2]$ and uses it as the total variance. In this example we'll need to search among incomplete designs to match their results.

Using a design effect adjustment method for computing number of clusters, they report K values of 10, 13, 12, 11, 10, and 8 for ICC values of 0, 0.1, 0.2, 0.3, 0.4, and 0.5, respectively.

Note that **PASS** does not solve for number of clusters using the same method that Baio et al. (2015) uses. Instead **PASS** performs an exhaustive search of all design pattern matrix models to compute the required number of clusters.

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 4a** 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 [K (Number of Clusters)]
Alternative Hypothesis	Two-Sided
Power.....	0.80
Alpha.....	0.05
Design Type.....	Incomplete (Custom)
Design Constraint Entry Type	Fixed Number of Time Periods (T)
T (Number of Time Periods including Baseline).....	6
Cluster Size Entry Type	Subjects per Cluster per Time Period (m)
m (Ave. Subjects per Cluster per Time Period).....	20
Input Type.....	Odds Ratios
OR_1 (Odds Ratio = O_1/O_2)	0.56
P_2 (Control Proportion)	0.26
Variance Calculation Formula.....	$\sigma^2 = [(P_1 + P_2)/2] \times [1 - (P_1 + P_2)/2]$
Use Calculated Variance as.....	Total Variance ($\sigma^2 = \sigma_y^2 = \tau^2 + \sigma_w^2$)
Between-Cluster Variability Input Type	ICC
ICC (Intraclass Correlation Coefficient)	0 to 0.5 by 0.1

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Output

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

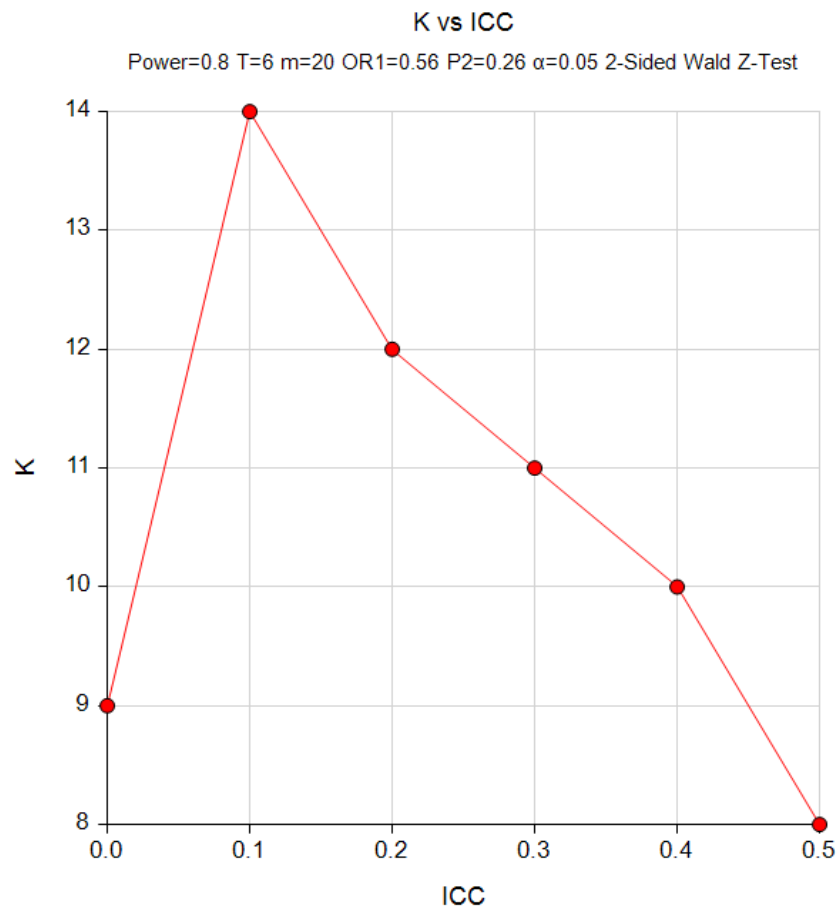
Numeric Results

Solve For: Sample Size [K (Number of Clusters)]
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: OR = 1$ vs. $H_1: OR \neq 1$

Power	Design Parameters		Number of Clusters K	Cluster Size		Sample Size N	Proportion		Odds Ratio OR1	Intraclass Correlation Coefficient ICC	Alpha
	S	T		M	m		Treatment P1	Control P2			
0.81965	5	6	9	120	20	1080	0.1644	0.26	0.56	0.0	0.05
0.82622	5	6	14	120	20	1680	0.1644	0.26	0.56	0.1	0.05
0.80496	5	6	12	120	20	1440	0.1644	0.26	0.56	0.2	0.05
0.81516	5	6	11	120	20	1320	0.1644	0.26	0.56	0.3	0.05
0.83368	5	6	10	120	20	1200	0.1644	0.26	0.56	0.4	0.05
0.81935	5	6	8	120	20	960	0.1644	0.26	0.56	0.5	0.05

Note: The variance is calculated as $\sigma^2 = [(P_1 + P_2)/2] \times [1 - (P_1 + P_2)/2]$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

Plots



Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Design 1 Details

Design Type: Incomplete (Custom)

Power	Design Parameters		Number of Clusters K	Cluster Size		Proportion		Variance			Intraclass Correlation Coefficient ICC	Coefficient of Variation COV
	S	T		M	m	Treatment P1	Control P2	Total σ^2	Between-Cluster τ^2	Within-Cluster σ_w^2		
0.81965	5	6	9	120	20	0.1644	0.26	0.167	0	0.167	0	0

Design Pattern (Balanced, Optimal)

Cluster	T1	T2	T3	T4	T5	T6
1	0	1	1	1	1	1
2	0	1	1	1	1	1
3	0	0	1	1	1	1
4	0	0	1	1	1	1
5	0	0	0	1	1	1
6	0	0	0	0	1	1
7	0	0	0	0	1	1
8	0	0	0	0	0	1
9	0	0	0	0	0	1

(A Design Details report is given for each additional scenario)

PASS computed K values of 9, 14, 12, 11, 10, and 8, which match those given in Table 1 of Baio et al. (2015). There are two differences, however, when ICC = 0 and ICC = 0.1. When ICC = 0, Baio et al. (2015) reports a value of 10, while **PASS** computes a value of 8. When ICC = 0.1, Baio et al. (2015) reports a value of 13, while **PASS** computes a value of 14. These difference may be due to rounding, but it is important to note that **PASS** uses a different method to arrive at the required number of clusters. **PASS** uses the power along with a search for the optimal (balanced) design pattern matrix to find the sample size. Baio et al. (2015) does not specify what final model is being used and the actual power that is achieved. **PASS** always achieves the desired level of power in the result. The power value of 0.81965, for example, indicates that K = 9 with ICC = 0 does achieve the desired power and is, therefore, the correct balanced-design solution.

It's interesting to note that if you search among unbalanced design pattern matrices in **PASS**, the reported number of clusters for ICC = 0 and ICC = 0.1 are even lower... 8 and 13, respectively. The value of 13 for ICC = 0.1 then matches Table 1 of Baio et al. (2015), but the value for ICC = 0 is lower and even further from the Table 1 result of 10. To replicate this, you may make the appropriate changes as listed below, or open **Example 4b** by going to the **File** menu and choosing **Open Example Template**.

Options Tab

Design Pattern Assignment Type for Extra Clusters..... **Unbalanced**

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

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Numeric Results

Solve For: Sample Size [K (Number of Clusters)]
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: H0: OR = 1 vs. H1: OR ≠ 1

Power	Design Parameters		Number of Clusters K	Cluster Size		Sample Size N	Proportion		Odds Ratio OR1	Intraclass Correlation Coefficient ICC	Alpha
	S	T		M	m		Treatment P1	Control P2			
0.80381	5	6	8	120	20	960	0.1644	0.26	0.56	0.0	0.05
0.80057	5	6	13	120	20	1560	0.1644	0.26	0.56	0.1	0.05
0.80496	5	6	12	120	20	1440	0.1644	0.26	0.56	0.2	0.05
0.81516	5	6	11	120	20	1320	0.1644	0.26	0.56	0.3	0.05
0.83368	5	6	10	120	20	1200	0.1644	0.26	0.56	0.4	0.05
0.81935	5	6	8	120	20	960	0.1644	0.26	0.56	0.5	0.05

Note: The variance is calculated as $\sigma^2 = [(P1 + P2)/2] \times [1 - (P1 + P2)/2]$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

PASS computed K values of 8, 13, 12, 11, 10, and 8, which match those given in Table 1 of Baio et al. (2015), except for ICC = 0. The power value of 0.80381, however, indicates that K = 8 with ICC = 0 does achieve the desired power and is, therefore, the correct unbalanced-design solution.

Example 5 – Finding Effect Size for an Incomplete Design (Validation using Hemming and Girling (2014))

Hemming and Girling (2014) presents an example (Example 2) where they calculate the effect size required to achieve 80% power in an incomplete design with an alpha level of 0.05 when $P_2 = 0.4$ (Note that in the paper, Proportion 1 is the standard proportion and Proportion 2 is the treatment proportion, so they are reversed), $ICC = 0.01$, and $m = 12$. They use $\sigma^2 = P_2(1 - P_2)$ as the variance calculation formula and assume it is the total variance.

The design pattern matrix shown on page 375 is as follows:

		Week																							
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21		
Team (Cluster)	1	0	1	1	1	1	1	1	1	1	1	1	1	1											
	2	0	0	1	1	1	1	1	1	1	1	1	1	1	1										
	3	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1									
	4	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1								
	5	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1						
	6	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1					
	7	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1				
	8	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1			
	9	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1		
	10	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	

They compute the detectable difference to be 0.1096 ($P_1 = 0.5096$ if searching above P_2 and $P_1 = 0.2904$ if searching below P_2) and a total sample size of 2100.

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 5a** 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 **Effect Size (P1, D1, R1, OR1)**
 Alternative Hypothesis **Two-Sided**
 Power..... **0.80**
 Alpha..... **0.05**
 Design Type..... **Incomplete (Custom)**
 Design Parameter Entry Type..... **Custom Design Pattern Matrix**
 Design Pattern Matrix Columns **ALL**
 R (Number of Design Pattern Replicates)..... **1**
 Cluster Size Entry Type **Subjects per Cluster per Time Period (m)**
 m (Ave. Subjects per Cluster per Time Period)..... **12**

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Input Type..... **Proportions**
 P1 (Treatment Proportion) **Search > P2**
 P2 (Control Proportion) **0.4**
 Variance Calculation Formula..... **$\sigma^2 = P2(1 - P2)$**
 Use Calculated Variance as..... **Total Variance ($\sigma^2 = \sigma_y^2 = \tau^2 + \sigma_w^2$)**
 Between-Cluster Variability Input Type..... **ICC**
 ICC (Intraclass Correlation Coefficient) **0.01**

Input Spreadsheet Data

Section 1

Row	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14
1	0	1	1	1	1	1	1	1	1	1	1	1	1	
2	0	0	1	1	1	1	1	1	1	1	1	1	1	1
3	0	0	0	1	1	1	1	1	1	1	1	1	1	1
4	0	0	0	0	1	1	1	1	1	1	1	1	1	1
5	0	0	0	0	0	1	1	1	1	1	1	1	1	1
6	0	0	0	0	0	0	1	1	1	1	1	1	1	1
7	0	0	0	0	0	0	0	1	1	1	1	1	1	1
8	0	0	0	0	0	0	0	0	1	1	1	1	1	1
9	0	0	0	0	0	0	0	0	0	1	1	1	1	1
10	0	0	0	0	0	0	0	0	0	0	1	1	1	1

Section 2

Row	C15	C16	C17	C18	C19	C20	C21	C22
1								
2								
3	1							
4	1	1						
5	1	1	1					
6	1	1	1	1				
7	1	1	1	1	1			
8	1	1	1	1	1	1		
9	1	1	1	1	1	1	1	
10	1	1	1	1	1	1	1	1

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Output

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

Numeric Results

Solve For: [Effect Size \(P1, D1, R1, OR1\)](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 - P_2 = 0$ vs. $H_1: P_1 - P_2 \neq 0$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Intraclass Correlation Coefficient ICC	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.8	21	22	1	10	210	12	2100	0.5096	0.4	0.1096	0.01	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

The detectable difference value of $D1 = 0.1096$, treatment proportion of 0.5096, and total sample size of $N = 2100$ calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.

Now, change the search direction for P_1 to Search < P_2 or open **Example 5b** by going to the **File** menu and choosing **Open Example Template**.

Design Tab

P_1 (Treatment Proportion) **Search < P_2**

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

Numeric Results

Solve For: [Effect Size \(P1, D1, R1, OR1\)](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 - P_2 = 0$ vs. $H_1: P_1 - P_2 \neq 0$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Intraclass Correlation Coefficient ICC	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.8	21	22	1	10	210	12	2100	0.2904	0.4	-0.1096	0.01	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

The detectable difference value of $D1 = -0.1096$, treatment proportion of 0.2904, and total sample size of $N = 2100$ calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.

Example 6 – Finding Effect Size for an Incomplete Design with a Transition Period (Validation using Hemming and Girling (2014))

Hemming and Girling (2014) presents an example (Example 3) where they calculate the effect size required to achieve 80% power with an alpha level of 0.05 in an incomplete design that has a transition period when $P_2 = 0.12$ (Note that in the paper, Proportion 1 is the standard proportion and Proportion 2 is the treatment proportion, so they are reversed), $COV = 0.3$, and $m = 1250$. They use $\sigma^2 = P_2(1 - P_2)$ as the variance calculation formula and assume it is the total variance.

The design pattern matrix shown on page 376 is as follows:

		Month							
		0-3	3-6	6-9	9-12	12-15	15-18	18-21	21-24
Hospital (Cluster)	1	0	0	•	1	1			
	2	0	0	•	1	1			
	3	0	0	•	1	1			
	4		0	0	•	1	1		
	5		0	0	•	1	1		
	6		0	0	•	1	1		
	7			0	0	•	1	1	
	8			0	0	•	1	1	
	9			0	0	•	1	1	
	10				0	0	•	1	1
	11				0	0	•	1	1
	12				0	0	•	1	1

They compute the detectable difference to be 0.0241 ($P_1 = 0.1441$ if searching above P_2 and $P_1 = 0.0959$ if searching below P_2) and a total sample size of 60000.

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 6a** 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 **Effect Size (P1, D1, R1, OR1)**
 Alternative Hypothesis **Two-Sided**
 Power..... **0.80**
 Alpha..... **0.05**
 Design Type..... **Incomplete (Custom)**
 Design Parameter Entry Type..... **Custom Design Pattern Matrix**

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Design Pattern Matrix Columns **ALL**
 R (Number of Design Pattern Replicates)..... **3**
 Cluster Size Entry Type **Subjects per Cluster per Time Period (m)**
 m (Ave. Subjects per Cluster per Time Period)..... **1250**
 Input Type..... **Proportions**
 P1 (Treatment Proportion) **Search > P2**
 P2 (Control Proportion) **0.12**
 Variance Calculation Formula..... **$\sigma^2 = P2(1 - P2)$**
 Use Calculated Variance as..... **Total Variance ($\sigma^2 = \sigma_y^2 = \tau^2 + \sigma_w^2$)**
 Between-Cluster Variability Input Type..... **COV**
 COV (Coefficient of Variation of Outcomes) **0.3**

Input Spreadsheet Data

Row	C1	C2	C3	C4	C5	C6	C7	C8
1	0	0	.	1	1			
2		0	0	.	1	1		
3			0	0	.	1	1	
4				0	0	.	1	1

Output

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

Numeric Results

Solve For: [Effect Size \(P1, D1, R1, OR1\)](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: H0: P1 - P2 = 0 vs. H1: P1 - P2 \neq 0

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Coefficient of Variation COV	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.8	7	8	3	12	5000	1250	60000	0.1441	0.12	0.0241	0.3	0.05

Note: The variance is calculated as $\sigma^2 = P2(1 - P2)$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

The detectable difference value of D1 = 0.0241, treatment proportion of 0.1441, and total sample size of N = 60000 calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.

Now, change the search direction for P1 to Search < P2 or open **Example 6b** by going to the **File** menu and choosing **Open Example Template**.

Design Tab

P1 (Treatment Proportion) **Search < P2**

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

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Numeric Results

Solve For: [Effect Size \(P1, D1, R1, OR1\)](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 - P_2 = 0$ vs. $H_1: P_1 - P_2 \neq 0$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Difference D1	Coefficient of Variation COV	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.8	7	8	3	12	5000	1250	60000	0.0959	0.12	-0.0241	0.3	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the total variance ($\sigma^2 = \sigma_y^2$) in power computations.

The detectable difference value of $D_1 = -0.0241$, treatment proportion of 0.0959, and total sample size of $N = 60000$ calculated by **PASS** all match the results in Hemming and Girling (2014) exactly.

Example 7 – Finding Power for an Incomplete Design with a Delayed Treatment Effect

Continuing with the situation described in Example 1 above, Hussey and Hughes (2007) presents an example in section 3.6 that demonstrates power calculations with a delayed treatment effect. A delayed treatment effect occurs when the full effect of the treatment does not occur in the same time period where it is introduced. Delayed treatment effects are modeled with fractional numbers in the design pattern matrix.

Motivated by the example in Hussey and Hughes (2007) but not using the same design, in this example we'll investigate the power achieved by the following design pattern matrix, with each row replicated 6 times.

		Time						
		1	2	3	4	5	6	7
Cluster	1	0	0.5	0.8	1	1	1	1
	2	0	0	0.5	0.8	1	1	1
	3	0	0	0	0.5	0.8	1	1
	4	0	0	0	0	0.5	0.8	1

What is the power for the parameters $P_2 = 0.05$, $R_1 = P_1/P_2 = 0.7$, $COV = 0.02$ to 0.5 , and $\text{Alpha} = 0.05$?

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 7a** 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**
 Alpha..... **0.05**
 Design Type..... **Incomplete (Custom)**
 Design Parameter Entry Type..... **Custom Design Pattern Matrix**
 Design Pattern Matrix Columns **ALL**
 R (Number of Design Pattern Replicates)..... **6**
 Cluster Size Entry Type **Subjects per Cluster per Time Period (m)**
 m (Ave. Subjects per Cluster per Time Period)..... **100**
 Input Type..... **Ratios**
 R1 (Ratio = P_1/P_2)..... **0.7**
 P2 (Control Proportion) **0.05**
 Variance Calculation Formula..... **$\sigma^2 = P_2(1 - P_2)$**
 Use Calculated Variance as..... **Within-Cluster Variance ($\sigma^2 = \sigma_w^2$)**
 Between-Cluster Variability Input Type **COV**
 COV (Coefficient of Variation of Outcomes) **0.02 to 0.5 by 0.02**

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Input Spreadsheet Data

Row	C1	C2	C3	C4	C5	C6	C7
1	0	0.5	0.8	1.0	1.0	1.0	1
2	0	0.0	0.5	0.8	1.0	1.0	1
3	0	0.0	0.0	0.5	0.8	1.0	1
4	0	0.0	0.0	0.0	0.5	0.8	1

Output

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

Numeric Results

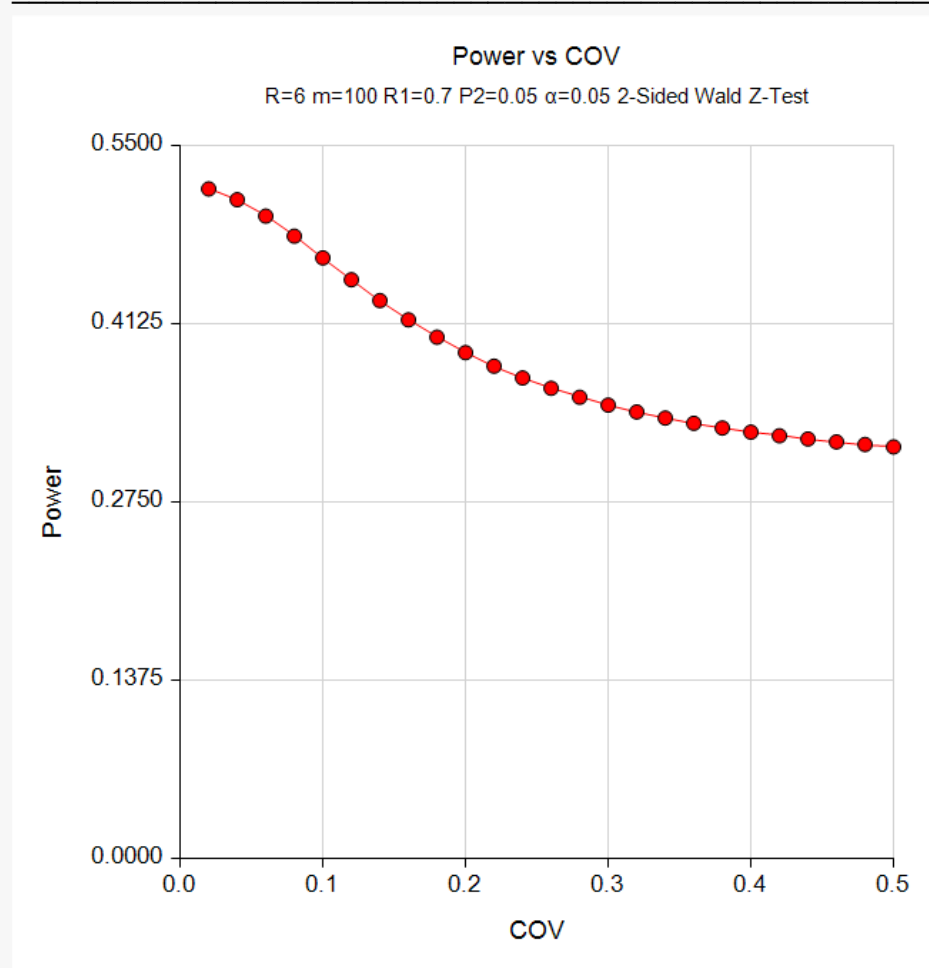
Solve For: [Power](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 / P_2 = 1$ vs. $H_1: P_1 / P_2 \neq 1$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Ratio R1	Coefficient of Variation COV	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.51663	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.02	0.05
0.50827	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.04	0.05
0.49562	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.06	0.05
0.48017	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.08	0.05
0.46341	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.10	0.05
0.44656	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.12	0.05
0.43046	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.14	0.05
0.41561	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.16	0.05
0.40222	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.18	0.05
0.39034	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.20	0.05
0.37990	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.22	0.05
0.37077	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.24	0.05
0.36280	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.26	0.05
0.35586	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.28	0.05
0.34980	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.30	0.05
0.34451	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.32	0.05
0.33987	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.34	0.05
0.33579	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.36	0.05
0.33220	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.38	0.05
0.32901	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.40	0.05
0.32619	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.42	0.05
0.32368	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.44	0.05
0.32143	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.46	0.05
0.31942	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.48	0.05
0.31761	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.50	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the within-cluster variance ($\sigma^2 = \sigma w^2$) in power computations.

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Plots



The results display the power achieved by the design for values of COV between 0.02 and 0.5. The power ranges from about 0.32 to 0.52.

To compare this against the power without a delayed treatment effect, go back and change the custom design pattern matrix by replacing all fractional numbers by 1's or open **Example 7b** by going to the **File** menu and choosing **Open Example Template**.

		Time						
		1	2	3	4	5	6	7
Cluster	1	0	1	1	1	1	1	1
	2	0	0	1	1	1	1	1
	3	0	0	0	1	1	1	1
	4	0	0	0	0	1	1	1

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Input Spreadsheet Data

Row	C1	C2	C3	C4	C5	C6	C7
1	0	1	1	1	1	1	1
2	0	0	1	1	1	1	1
3	0	0	0	1	1	1	1
4	0	0	0	0	1	1	1

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

Numeric Results

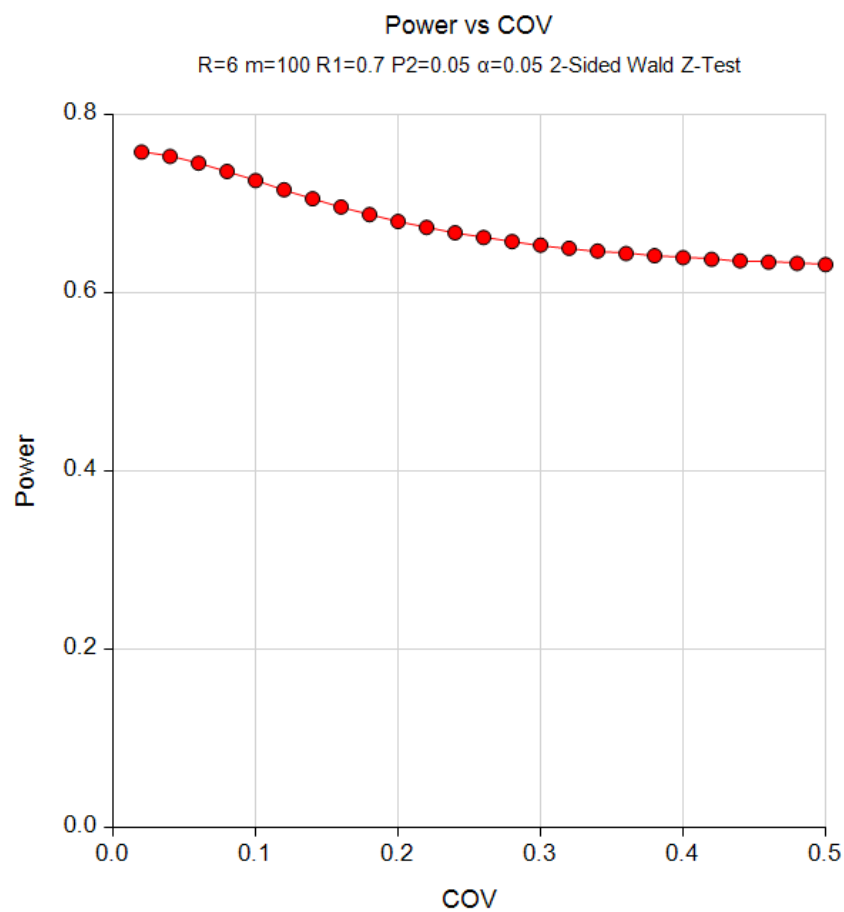
Solve For: [Power](#)
 Design Type: Incomplete (Custom)
 Groups: 1 = Treatment, 2 = Control
 Test Statistic: Wald Z-Test
 Hypotheses: $H_0: P_1 / P_2 = 1$ vs. $H_1: P_1 / P_2 \neq 1$

Power	Design Parameters			Number of Clusters K	Cluster Size		Sample Size N	Proportion		Ratio R1	Coefficient of Variation COV	Alpha
	S	T	R		M	m		Treatment P1	Control P2			
0.75806	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.02	0.05
0.75312	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.04	0.05
0.74558	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.06	0.05
0.73629	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.08	0.05
0.72609	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.10	0.05
0.71572	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.12	0.05
0.70570	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.14	0.05
0.69634	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.16	0.05
0.68781	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.18	0.05
0.68018	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.20	0.05
0.67340	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.22	0.05
0.66743	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.24	0.05
0.66219	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.26	0.05
0.65760	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.28	0.05
0.65356	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.30	0.05
0.65002	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.32	0.05
0.64691	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.34	0.05
0.64416	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.36	0.05
0.64173	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.38	0.05
0.63957	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.40	0.05
0.63765	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.42	0.05
0.63594	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.44	0.05
0.63440	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.46	0.05
0.63303	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.48	0.05
0.63179	6	7	6	24	700	100	16800	0.035	0.05	0.7	0.50	0.05

Note: The variance is calculated as $\sigma^2 = P_2(1 - P_2)$ and is considered to be the within-cluster variance ($\sigma^2 = \sigma_w^2$) in power computations.

Tests for Two Proportions in a Stepped-Wedge Cluster-Randomized Design

Plots



The power ranges from about 0.63 to 0.76. As you can see, the power is much lower in the design with the delayed treatment effect than it is with this design.