Chapter 487

Tests for the Matched-Pair Difference of Two Event Rates in a Cluster-Randomized Design

Introduction

Cluster-randomized designs are those in which whole clusters of subjects (classes, hospitals, communities, etc.) are sampled, rather than individual subjects. This sample size and power procedure is used for the case where the subject responses are counts (event rates). To reduce the variation (and thus increase power), clusters are matched, with one cluster of each pair assigned to the control group, and the other assigned the treatment group. This procedure gives the number of pairs needed for the desired power requirement.

The formulas used here are based on Hayes and Bennett (1999) and Hayes and Moulton (2009). The methods are discussed in Donner and Klar (2000) and briefly in Campbell and Walters (2014).

Technical Details

Definition of Terms

The following table presents the various terms that are used.

Group	1 (Control)	2 (Treatment)
Sample size	K	K
Individual event rates	λ_1	λ_2

Coefficient of Variation:	CV_M (within-pair coefficient of variation between clusters in the absence of intervention)
Number of cluster pairs:	K ($2K$ is the total number of clusters)
Person-Years per cluster:	M

Hypotheses

The null and alternative hypotheses are

$$H_0: \lambda_{Diff} = 0 \quad \text{vs.} \quad H_1: \lambda_{Diff} \neq 0$$

Corresponding one-sided hypotheses may also be used. These hypotheses may be tested based on an appropriate paired difference test.

Sample Size and Power Calculations

Sample Size Calculation

The sample size calculation for the number of cluster pairs, as given in Hayes and Bennett (1999) and Hayes and Moulton (2009), is

$$K = 2 + \left(z_{\alpha/2} + z_\beta\right)^2 \frac{(\lambda_1 + \lambda_2)/M + CV_M^2(\lambda_1^2 + \lambda_2^2)}{(\lambda_1 - \lambda_2)^2}$$

Estimating CV$_M$

Hayes and Bennett (1999) suggest that CV_M may be estimated based on the following, if prior data is available:

Let λ_{ij} represent the true event rate in the jth cluster (j = 1,2) of the ith pair (i = 1, …, M) and r_{ij} represent the corresponding observed rate. The empirical variance of the ith pair is given by

$$s_i^2 = (r_{i2} - r_{i1})^2/2$$

Define

$$s_m^2 = \sum s_i^2/M$$

as the average of the within-pair variances.

CV_M may be estimated from

$$CV_M^2 = \frac{s_m^2 - Av(r_i/y_{ij})}{Av(r_i^2)}$$

where y_{ij} is the person-year of follow-up in the ith pair of the jth cluster, and $Av()$ indicates the mean over all M clusters.

According to Hayes and Bennett (1999), "If only unmatched data are available, a conservative approach is to use [the coefficient of variation (SD/Mean) between clusters within each group] as an upper limit for CV_M."

If no data is available, a series of plausible values, usually between 0 and 0.5, should be considered.

Power Calculation

The corresponding power calculation to the sample size calculation above is

$$Power = 1 - \Phi\left(\sqrt{\frac{(K-2)(\lambda_1 - \lambda_2)^2}{(\lambda_1 + \lambda_2)/M + CV_M^2(\lambda_1^2 + \lambda_2^2)}} - z_{\alpha/2}\right)$$

where $z_x = \Phi(x)$ is the standard normal distribution function.

Example 1 – Calculating Sample Size

One difficulty in calculating sample size for a matched-pair difference of two event rates in a cluster-randomized design is obtaining a value for CV_M, the within-pair coefficient of variation between clusters. This example shows how to enter a range of values to determine the effect of CV_M on sample size. Suppose that a cluster randomized study is to be conducted in which the control event rate is assumed to be $\lambda 1 = 0.80$, and the treatment event rate is $\lambda 2 = 0.58$. Each cluster is assumed to have about 80 individuals. Each will be followed for one year. A range of values between 0.05 and 0.5 for CV_M will be examined. The desired power and alpha are 0.9 and 0.05, respectively. The test will be 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 ....**K (Number of Cluster Pairs)**
Alternative Hypothesis ....**Two-Sided**
Power...**0.90**
Alpha...**0.05**
M (Person-Years per Cluster).........................**80**
λ1 (Event Rate of Group 1)............................**0.80**
Enter λ2, Diff, or Ratio for Group 2 ....**λ2 (Event Rate of Group 2)**
λ2 (Event Rate of Group 2)............................**0.58**
CVM (Within-Pair Coefficient of Variation).....**0.05 to 0.50 by 0.05**

Output

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

Numeric Results
───
Solve For: K (Number of Cluster Pairs)
Groups: 1 = Control, 2 = Treatment
Alternative Hypothesis: Two-Sided
───

Power	Number of Cluster Pairs K	Total Number of Clusters 2K	Person-Years Cluster M	Person-Years Total N	Event Rate Group 1 λ_1	Event Rate Group 2 λ_2	Event Rate Difference $\lambda_2 - \lambda_1$	Event Rate Ratio λ_2 / λ_1	Within-Pair Coefficient of Variation CVM	Alpha
0.9389	7	14	80	1120	0.8	0.58	-0.22	0.73	0.05	0.05
0.9064	8	16	80	1280	0.8	0.58	-0.22	0.73	0.10	0.05
0.9151	11	22	80	1760	0.8	0.58	-0.22	0.73	0.15	0.05
0.9167	15	30	80	2400	0.8	0.58	-0.22	0.73	0.20	0.05
0.9001	19	38	80	3040	0.8	0.58	-0.22	0.73	0.25	0.05
0.9022	25	50	80	4000	0.8	0.58	-0.22	0.73	0.30	0.05
0.9027	32	64	80	5120	0.8	0.58	-0.22	0.73	0.35	0.05
0.9025	40	80	80	6400	0.8	0.58	-0.22	0.73	0.40	0.05
0.9020	49	98	80	7840	0.8	0.58	-0.22	0.73	0.45	0.05
0.9013	59	118	80	9440	0.8	0.58	-0.22	0.73	0.50	0.05
───

Power   The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
K   Represents the number of cluster pairs needed in the design.
2K   The total number of clusters in the design, 2 clusters per pair.
M   The average number of person-years per cluster in all clusters.
N   The total number of person-years in the design. N = 2KM.
λ_1 The event (or incidence) rate of the control group. This is the baseline rate.
λ_2 The event (or incidence) rate of the treatment group.
$\lambda_2 - \lambda_1$ The difference between the treatment event rate and the control event rate.
λ_2 / λ_1 The ratio of the treatment event rate to the control event rate.
CVM   The within-pair coefficient of variation between clusters in the absence of intervention. See the documentation or
          references for suggestions for estimating CVM.
Alpha   The probability of rejecting a true null hypothesis, that is, rejecting H0 when the event rates are actually equal.

Summary Statements
───
A matched-cluster-pair cluster-randomized design will be used to test whether there is a difference in event rates
($\lambda_2 - \lambda_1$ = $\lambda_{treatment} - \lambda_{control}$). The comparison will be made using a two-sided, paired-difference Z-test, with a
Type I error rate (α) of 0.05. The within-pair coefficient of variation between clusters in the absence of intervention
is assumed to be 0.05. To detect an event rate difference of -0.22 (λ_2 = 0.58, λ_1 = 0.8) with 80 person-years per
cluster and 90% power, the number of needed cluster pairs is 7, or 14 total clusters (1120 total person-years).
───

References
───
Hayes, R.J. and Bennett, S. 1999. 'Simple sample size calculation for cluster-randomized trials'. International
    Journal of Epidemiology. Vol 28, pages 319-326.
Hayes, R.J. and Moulton, L.H. 2009. Cluster Randomised Trials. CRC Press. New York.
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.
───

Plots

_____
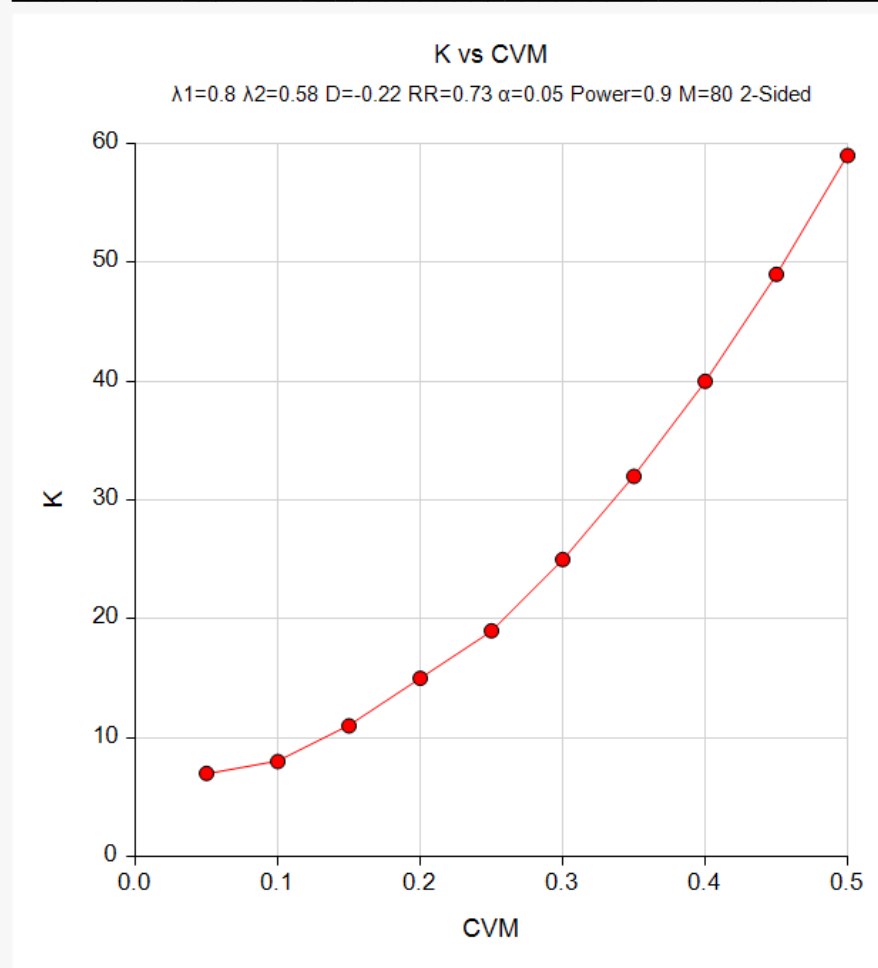
This report shows the needed number of cluster pairs for each of the coefficient of variation values.

Example 2 – Validation using Direct Calculation

Hayes and Bennett (1999) and Hayes and Moulton (2009) give formulas for sample size for tests of the matched-pair difference in a cluster-randomized design for comparing event rates, proportions, and means. However, the only example corresponds to the proportions case. Thus, we give an example where the number of cluster pairs needed is determined by direct calculation.

We will consider the (two-sided) case where *λ1 = 0.6; λ2 = 0.4; M = 200; alpha = 0.05; CVM = 0.25; Power = 0.80.*

$$K = 2 + \left(z_{\frac{\alpha}{2}} + z_{\beta}\right)^2 \frac{(\lambda_1 + \lambda_2)/M + CV_M^2(\lambda_1^2 + \lambda_2^2)}{(\lambda_1 - \lambda_2)^2}$$

$$= 2 + (1.9600 + 0.8416)^2 \frac{(0.6 + 0.4)/200 + 0.25^2(0.6^2 + 0.4^2)}{(0.4 - 0.6)^2}$$

$$= 2 + (7.84888)\left(\frac{0.005 + 0.0625(0.52)}{0.04}\right)$$

$$= 9.358$$

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 ....**K (Number of Cluster Pairs)**
Alternative Hypothesis ....**Two-Sided**
Power..**0.80**
Alpha...**0.05**
M (Person-Years per Cluster).........................**200**
λ1 (Event Rate of Group 1)............................**0.6**
Enter λ2, Diff, or Ratio for Group 2 ....**λ2 (Event Rate of Group 2)**
λ2 (Event Rate of Group 2)............................**0.4**
CVM (Within-Pair Coefficient of Variation).....**0.25**

Output

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

Numeric Results

Solve For: K (Number of Cluster Pairs)
Groups: 1 = Control, 2 = Treatment
Alternative Hypothesis: Two-Sided

| | Number of Cluster Pairs | Total Number of Clusters | Person-Years | | Event Rate | | | | Within-Pair Coefficient of Variation | |
| | | | Cluster | Total | Group 1 | Group 2 | Difference | Ratio | | |
Power	K	2K	M	N	$\lambda 1$	$\lambda 2$	$\lambda 2 - \lambda 1$	$\lambda 2 / \lambda 1$	CVM	Alpha
0.8318	10	20	200	4000	0.6	0.4	-0.2	0.67	0.25	0.05

PASS calculates the number of cluster pairs needed to be 10, which matches the rounded-up calculation value.