Chapter 491

Tests for the Matched-Pair Difference of Two Means 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 continuous (mean outcome). 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.

<table>
<thead>
<tr>
<th>Group</th>
<th>1 (Control)</th>
<th>2 (Treatment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size:</td>
<td>$K$</td>
<td>$K$</td>
</tr>
<tr>
<td>Means:</td>
<td>$\mu_1$</td>
<td>$\mu_2$</td>
</tr>
<tr>
<td>Standard Deviations:</td>
<td>$\sigma_1$</td>
<td>$\sigma_2$</td>
</tr>
</tbody>
</table>

(these are the within-cluster standard deviations)

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)

Cluster size: $M$ (Average number of individuals per cluster)
Hypotheses
The null and alternative hypotheses are

\[ H_0: \mu_{\text{diff}} = 0 \quad \text{vs.} \quad H_1: \mu_{\text{diff}} \neq 0 \]

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

\[ K = 2 + \frac{(z_{\alpha/2} + z_{\beta})^2}{2} \cdot \frac{\left(\sigma_1^2 + \sigma_2^2\right)/M + CV_M^2(\mu_1^2 + \mu_2^2)}{\mu_1^2 - \mu_2^2} \]

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

Let \( \mu_{ij} \) represent the true mean in the \( j \)th cluster \((j = 1,2)\) of the \( i \)th pair \((i = 1,\ldots,M)\) and \( \bar{x}_{ij} \) represent the corresponding observed mean. If the empirical variance of the \( i \)th pair is called \( s_i^2 \) then define

\[ s_m^2 = \frac{\sum s_i^2}{M} \]

as the average of the within-pair variances.

\( CV_M \) may be estimated from

\[ CV_M^2 = \frac{s_m^2 - Av(\sigma_i^2/n_{ij})}{Av(\bar{x}_{ij}^2)} \]

where \( n_{ij} \) is the number of individuals in the \( i \)th pair of the \( j \)th 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

\[ \text{Power} = 1 - \Phi\left( \frac{\sqrt{(K - 2)(\mu_1 - \mu_2)^2}}{\left(\sigma_1^2 + \sigma_2^2\right)/M + CV_M^2(\mu_1^2 + \mu_2^2)} - \frac{z_{\alpha/2}}{2} \right) \]

where \( z_{\alpha} = \Phi(x) \) is the standard normal distribution function.
Procedure Options
This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab
The Design tab contains most of the parameters and options that you will be concerned with.

Solve For
This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are Power, K, M, and μ2.

Under most situations, you will select either Power to calculate power or K to calculate the number of cluster pairs. Occasionally, you may want to fix the number of clusters and find the necessary cluster size.

The value selected here always appears as the vertical axis on the charts (when applicable).

When searching for M, some scenarios with small K’s are not feasible.

Test

Alternative Hypothesis
Specify whether the test is one-sided or two-sided. The one-sided option specifies a one-tailed test.

Power and Alpha

Power
This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected.

Values must be between zero and one. Historically, the value of 0.80 (beta = 0.20) was used for power. Now, 0.90 (beta = 0.10) is commonly used.

A single value may be entered or a range of values, such as 0.8 to 0.95 by 0.05, may be entered.

Alpha
This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected.

Values must be between zero and one. Usually, the value of 0.05 is used for two-sided tests and 0.025 is used for one-sided tests.

You may enter a range of values such as 0.01 0.05 0.10 or 0.01 to 0.10 by 0.01.
Sample Size – Number of Clusters & Cluster Size

K (Number of Cluster Pairs)
This is the number of cluster pairs. Each cluster pair consists of two clusters that are matched or paired based on one or more criteria. The total sample size is equal to the number of clusters times the cluster size. This value must be a positive number. You can enter a single value such as 20 or a series of values such as 10 20 30 or 10 to 30 by 5.

M (Cluster Size)
This is the average number of individuals or items per cluster. This value must be a positive number that is at least 1. It can be a decimal (fractional) number such as 20.7. You can enter a single value such as 50 or a series of values such as 40 50 60 or 40 to 60 by 5.

Effect Size – Means

µ1 (Mean for Group 1)
Enter a value (or range of values) for the group 1 (control group) mean. You can enter a single value such as 14 or a series of values such as 5 6 7 or 5 to 7 by 0.5.

Enter µ2, Diff, or Ratio for Group 2
This option lets you indicate how µ2 is entered. The options are

- µ2 (Mean for Group 2)
  Enter the value of µ2 directly.

- Diff (Difference Between Means)
  Enter values for the difference between the means (Diff = µ2 - µ1). The value of µ2 is equal to µ1 + Diff.

- Ratio (Ratio of Means)
  Enter values for the ratio of the means (Ratio = µ2 / µ1). The value of µ2 is equal to µ1 * Ratio. Note that the hypothesis still concerns the difference. This is just a convenient way of specifying a value.

µ2 (Mean for Group 2)
This option is displayed only if Enter µ2, Diff, or Ratio for Group 2 = “µ2 (Mean of Group 2).”
Enter a value (or range of values) for the mean in group 2 (treatment group). The value must be different from µ1. You can enter a single value such as 14 or a series of values such as 5 6 7 or 5 to 7 by 0.5.

Diff (Difference Between Means)
This option is displayed only if Enter µ2, Diff, or Ratio for Group 2 = “Diff (Difference Between Means).”
Enter values for the difference between the means (Diff = µ2 - µ1). The value of µ2 is equal to µ1 + Diff. The range is any positive or negative value other than 0 (Diff ≠ 0). You can enter a single value such as 5 or a series of values such as 5 6 7 or 5 to 7 by 5.

Ratio (Ratio of Means)
This option is displayed only if Enter µ2, Diff, or Ratio for Group 2 = “Ratio (Ratio of Means).”
Enter values for the ratio of the means (Ratio = µ2 / µ1) at which the power is to be calculated. The value of µ2 is calculated using µ1 * Ratio. The range is Ratio > 0, and Ratio ≠ 1. Typically, Ratio is between 0.5 and 2.0. You can enter a single value such as 1.15 or a series of values such as 1.05 1.10 1.15 1.20 or 1.05 to 1.20 by 0.05.
Effect Size – Within-Cluster Standard Deviations

\( \sigma_1 \) (Standard Deviation for Group 1)

Enter a value (or range of values) for the group 1 (control group) within cluster standard deviation. The value must be greater than zero. You can enter a single value such as 5 or a series of values such as 5 6 7 or 5 to 7 by 0.5.

\( \sigma_2 \) (Standard Deviation for Group 2)

Enter a value (or range of values) for the group 2 (treatment group) within cluster standard deviation. The value must be greater than zero. You can enter a single value such as 5 or a series of values such as 5 6 7 or 5 to 7 by 0.5.

Effect Size – Within-Pair Coefficient of Variation between Clusters

\( CVM \) (Within-Pair Coefficient of Variation)

Enter a value or values for the within-pair coefficient of variation (SD/Mean) between clusters in the absence of intervention. See the documentation or references for suggestions for estimating CVM. If prior information is not available, Hayes and Bennett (1999) suggest that CVM is usually less than 0.25 and seldom greater than 0.50. You can enter a single value such as 0.15 or a series of values such as 0.05 0.10 0.15 0.20 or 0.05 to 0.30 by 0.05.
### Example 1 – Calculating Sample Size

One difficulty in calculating sample size for a matched-pair difference of two means 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 mean is assumed to be $\mu_1 = 8.4$, and the treatment mean is $\mu_2 = 7.1$. The within-cluster standard deviation for both groups is assumed to be 2.8. Each cluster is assumed to have about 120 individuals. 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

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Tests for the Matched-Pair Difference of Two Means in a Cluster-Randomized Design** procedure window. You may then make the appropriate entries as listed below, or open Example 1 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design Tab</strong></td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>$K$ (Number of Cluster Pairs)</td>
</tr>
<tr>
<td>Alternative Hypothesis</td>
<td>Two-Sided</td>
</tr>
<tr>
<td>Power</td>
<td>0.90</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>$M$ (Cluster Size)</td>
<td>120</td>
</tr>
<tr>
<td>$\mu_1$ (Mean for Group 1)</td>
<td>8.4</td>
</tr>
<tr>
<td>Enter $\mu_2$, Diff, or Ratio for Group 2</td>
<td>$\mu_2$ (Mean for Group 2)</td>
</tr>
<tr>
<td>$\mu_2$ (Mean for Group 2)</td>
<td>7.1</td>
</tr>
<tr>
<td>$\sigma_1$ (Standard Deviation for Group 1)</td>
<td>2.8</td>
</tr>
<tr>
<td>$\sigma_2$ (Standard Deviation for Group 2)</td>
<td>2.8</td>
</tr>
<tr>
<td>CVM (Within-Pair Coefficient of Variation)</td>
<td>0.05 to 0.50 by 0.05</td>
</tr>
</tbody>
</table>

### Annotated Output

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

### Numeric Results

<table>
<thead>
<tr>
<th>Number of Cluster Pairs</th>
<th>Total Number of Clusters</th>
<th>Cluster Size M</th>
<th>Total Size N = 2KM</th>
<th>Mean Gr 1 $\mu_1$</th>
<th>Mean Gr 2 $\mu_2$</th>
<th>Mean Diff. $\mu_2 - \mu_1$</th>
<th>Mean Ratio $\mu_2/\mu_1$</th>
<th>$\sigma_1$</th>
<th>$\sigma_2$</th>
<th>Coef. of Var. CVM</th>
<th>Two-Sided Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9281</td>
<td>5</td>
<td>10</td>
<td>120</td>
<td>1200</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.05</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9205</td>
<td>11</td>
<td>22</td>
<td>120</td>
<td>2640</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.100</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9042</td>
<td>20</td>
<td>40</td>
<td>120</td>
<td>4800</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.150</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9009</td>
<td>33</td>
<td>66</td>
<td>120</td>
<td>7920</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.200</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9011</td>
<td>50</td>
<td>100</td>
<td>120</td>
<td>12000</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.250</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9020</td>
<td>71</td>
<td>142</td>
<td>120</td>
<td>17040</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.300</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9002</td>
<td>95</td>
<td>190</td>
<td>120</td>
<td>22800</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.350</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9020</td>
<td>124</td>
<td>248</td>
<td>120</td>
<td>29760</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.400</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9016</td>
<td>156</td>
<td>312</td>
<td>120</td>
<td>37440</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.450</td>
<td>0.05</td>
</tr>
<tr>
<td>0.9002</td>
<td>191</td>
<td>382</td>
<td>120</td>
<td>45840</td>
<td>8.4</td>
<td>-1.3</td>
<td>0.85</td>
<td>2.8</td>
<td>2.8</td>
<td>0.500</td>
<td>0.05</td>
</tr>
</tbody>
</table>
Tests for the Matched-Pair Difference of Two Means in a Cluster-Randomized Design

References

Report Definitions
Power is the probability of rejecting a false null hypothesis.
K represents the number of cluster pairs needed in the design.
2K is the total number of clusters in the design, 2 clusters per pair.
M is the average number of individuals or items in each cluster.
N is the total number of individuals or items in the design. N = 2KM.
μ1 is the mean for the control group.
μ2 is the mean for the treatment group.
μ2 - μ1 is the difference between the treatment mean and the control mean.
μ2 / μ1 is the ratio of the treatment mean to the control mean.
CVM is the within-pair coefficient of variation between clusters in the absence of intervention. See the documentation or references for suggestions for estimating CVM.
Alpha is the probability of rejecting a true null hypothesis, that is, rejecting H0 when the means are actually equal.

Summary Statements
A sample of 5 cluster pairs (10 clusters) with 120 individuals per cluster achieves 93% power to detect a difference of -1.3 between the treatment mean 7.1 and the control mean 8.4. The within group standard deviation for group 1 is 2.8 and the within group standard deviation for group 2 is 2.8. The within-pair coefficient of variation between clusters is 0.050. A two-sided paired test of the mean difference is assumed, with a significance level of 0.050.

Chart Section
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 \( \mu_1 = 4.5; \mu_2 = 5.7; \sigma_1 = 3.3; \sigma_2 = 3.9; M = 200; \) alpha = 0.05; CVM = 0.25; Power = 0.80.

\[
K = 2 + \left( z_{\alpha/2} + z_{\beta} \right)^2 \frac{\left( \sigma_1^2 + \sigma_2^2 \right) / M + CV_M^2 (\mu_1^2 + \mu_2^2)}{(\mu_1 - \mu_2)^2}
\]

\[
= 2 + (1.9600 + 0.8416)^2 \frac{\left( 3.3^2 + 3.9^2 \right) / 200 + 0.25^2 (4.5^2 + 5.7^2)}{(4.5 - 5.7)^2}
\]

\[
= 2 + 7.84888 \frac{0.1305 + 0.0625 (52.74)}{1.44} = 20.678
\]

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the Tests for the Matched-Pair Difference of Two Means in a Cluster-Randomized Design procedure window. You may then make the appropriate entries as listed below, or open Example 2 by going to the File menu and choosing Open Example Template.

<table>
<thead>
<tr>
<th>Option</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Design Tab</strong></td>
<td></td>
</tr>
<tr>
<td>Solve For</td>
<td>K (Number of Cluster Pairs)</td>
</tr>
<tr>
<td>Alternative Hypothesis</td>
<td>Two-Sided</td>
</tr>
<tr>
<td>Power</td>
<td>0.80</td>
</tr>
<tr>
<td>Alpha</td>
<td>0.05</td>
</tr>
<tr>
<td>M (Cluster Size)</td>
<td>120</td>
</tr>
<tr>
<td>( \mu_1 ) (Mean for Group 1)</td>
<td>4.5</td>
</tr>
<tr>
<td>Enter ( \mu_2 ), Diff, or Ratio for Group 2</td>
<td>( \mu_2 ) (Mean for Group 2)</td>
</tr>
<tr>
<td>( \mu_2 ) (Mean for Group 2)</td>
<td>5.7</td>
</tr>
<tr>
<td>( \sigma_1 ) (Standard Deviation for Group 1)</td>
<td>3.3</td>
</tr>
<tr>
<td>( \sigma_2 ) (Standard Deviation for Group 2)</td>
<td>3.9</td>
</tr>
<tr>
<td>CVM (Within-Pair Coefficient of Variation)</td>
<td>0.25</td>
</tr>
</tbody>
</table>
Output

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

**Numeric Results**

<table>
<thead>
<tr>
<th>Number of Cluster Pairs</th>
<th>Total Number of Clusters</th>
<th>Cluster Size</th>
<th>Total Size</th>
<th>Mean Gr 1</th>
<th>Mean Gr 2</th>
<th>Mean Diff.</th>
<th>Mean Ratio</th>
<th>Power</th>
<th>K</th>
<th>2K</th>
<th>M</th>
<th>N = 2KM</th>
<th>μ1</th>
<th>μ2</th>
<th>μ2-μ1</th>
<th>μ2/μ1</th>
<th>σ1</th>
<th>σ2</th>
<th>CVM</th>
<th>Two-Sided Alpha</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8067</td>
<td>21</td>
<td>42</td>
<td>200</td>
<td>8400</td>
<td>4.5</td>
<td>5.7</td>
<td>1.2</td>
<td>1.27</td>
<td>3.3</td>
<td>3.9</td>
<td>0.250</td>
<td>0.050</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PASS calculates the number of cluster pairs needed to be 21, which matches the rounded-up calculation value.