

## Chapter 387

# Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)

---

### Introduction

This procedure calculates power and sample size for a two-level hierarchical mixed model in which clusters (groups, classes, hospitals, etc.) of subjects are measured one time (cross-sectional) on a continuous variable. The goal of the study is to compare the two group means.

In this design, the subjects are the level one units and the clusters are the level two units. Each subject in a particular cluster (level two unit) is randomized individually to one of two possible interventions. Note that a companion procedure power analyzes the other case in which the randomization occurs for the level two units (the clusters).

---

### Technical Details

Our formulation comes from Ahn, Heo, and Zhang (2015), chapter 5, section 5.3.2, pages 154-157. The hierarchical mixed model that is adopted is

$$Y_{ij} = \beta_0 + \delta X_{ij} + u_i + e_{ij}$$

where

- $Y_{ij}$  is the continuous response of the  $j^{\text{th}}$  subject in the  $i^{\text{th}}$  cluster.
- $\beta_0$  is the fixed intercept.
- $\delta$  is the treatment (intervention) effect of interest. It is the difference between the two treatment means.
- $X_{ij}$  is an indicator variable that is set to 1 if subject  $j$  in cluster  $i$  receives intervention 1 and 0 otherwise.

**Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)**

- $u_i$  is a random effect (subject-specific intercept) term for the  $i^{\text{th}}$  cluster that is distributed as  $N(0, \sigma_u^2)$ .
- $e_{ij}$  is a random effect for the  $j^{\text{th}}$  subject in the  $i^{\text{th}}$  cluster that is distributed as  $N(0, \sigma_e^2)$ .
- $\sigma_u^2$  is variance of the level two (cluster) random effects.
- $\sigma_e^2$  is variance of the level one (subject) random effects.
- $\sigma^2$  is the variance of  $Y$ , where  $\sigma^2 = \sigma_e^2 / (1 - \rho)$ .
- $\rho$  is the intraclass correlation (ICC), where  $\rho = \text{Corr}(Y_{ij}, Y_{ij'})$

The test of significance of the  $X_{ij}$  term in the mixed model analysis is the test statistic of interest. It tests the difference of the two treatment means. Since these are asymptotic results, the specific type of mixed model is not stated.

Assume that  $\delta = \mu_1 - \mu_2$  is to be tested using a z-test (large sample). The statistical hypotheses are  $H_0: \delta = 0$  vs.  $H_a: \delta \neq 0$ . When  $M_1 = M_2$ , the test statistic is

$$z = \frac{(\bar{Y}_1 - \bar{Y}_2)\sqrt{KM_1}}{\sigma\sqrt{2(1-\rho)}}$$

where

$$\bar{Y}_g = \frac{1}{KM_1} \sum_{i=1}^K \sum_{j=1}^{M_1} Y_{ij}, g = 1, 2$$

- $K$  is the number of clusters in the study.
- $M_1$  is the average number of subjects per cluster in group 1.
- $M_2$  is the average number of subjects per cluster in group 2.

This test statistic has an approximate normal distribution.

The power is calculated using

$$\text{Power} = \Phi\left\{\frac{\delta}{\sigma}\sqrt{KM_1M_2/[(1-\rho)(M_1+M_2)]}\right\} - \Phi(1-\alpha/2)$$

## 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

#### Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are  $\delta$ ,  $Power$ ,  $K$ , and  $M1$ .

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

## Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)

Note that the value selected here always appears as the vertical axis on the charts.

The program is set up to calculate power directly. To find appropriate values of the other parameters, a binary search is made using an iterative procedure until an appropriate value is found. This search considers integer values of M1 only.

---

### 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 also 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.

If your only interest is in determining the appropriate sample size for a confidence interval, set power to 0.5.

#### 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 alpha and this has become a standard. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

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 and Subjects – Number of Clusters

#### K (Number of Clusters)

This is the number of clusters (classes, hospitals, practices, etc.) in the study. Subjects in each of these clusters are randomly assigned to one of the two interventions (group 1 or 2). The study sample size is equal to the number of clusters times the number of subjects per cluster.

#### Range

This value must be a positive integer.

#### List

You can enter a list of values such as "10 20 30". A separate analysis will be run for each element in the list.

---

### Sample Size – Number of Clusters and Subjects – Number of Subjects Per Cluster

#### M1 (Group 1 Subjects Per Cluster)

This is the average number of items (subjects) per cluster receiving intervention (treatment) 1 and thus assigned to group 1. PASS arbitrarily calls group 1 the Treatment Group.

**Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)****Range**

This value must be a positive number that is at least 1.

It can be a decimal (fractional) number such as '2.7'. The resulting values of N1 and N2 will be rounded to integers.

**List**

You can use a list of values such as "10 15 20". A separate analysis will be run for each element in the list.

**M2 (Group 2 Subjects Per Cluster)**

This is the average number of items (subjects) per cluster receiving intervention (treatment) 2 and thus assigned to group 2. PASS arbitrarily calls group 2 the Control Group. The sample size for this group is equal to the number of clusters times the number of group 2 subjects per cluster.

**Range**

This value must be a positive number that results in a group sample size greater than 1. It can be a decimal (fractional) number such as '2.7'.

**Using 'MI'**

If you simply want to use a multiple of *MI*, the value for group 1, you would enter the multiple followed by "*MI*", with no blanks. If you want to use *MI* directly, you do not have to enter the leading "1". For example, all of the following are valid entries:

M1 2M1 1.5M1 0.5M1 10 15.

**List**

You can use a list of values such as "10 20 30" or "M1 2M1 3M1".

**Effect Size** **$\delta$  (Mean Difference =  $\mu_1 - \mu_2$ )**

Enter a value for the difference between the means of groups 1 and 2 at which the design is powered. That is, the power is the probability of detecting a difference of at least this amount. This value is not necessarily the true mean difference. Rather, it is the minimum difference you want to be able to detect.

**Range**

$\delta$  can be any non-zero value (positive or negative). Since this procedure uses a two-sided test statistic, you will get the same result with either positive or negative values.

The difference has the same scale as the standard deviation.

**Notes**

You can enter a single value such as 10 or a series of values such as 10 20 30 or 5 to 50 by 5.

When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

 **$\sigma$  (Standard Deviation)**

Enter the subject-to-subject standard deviation. This standard deviation applies for both groups.

Note that  $\sigma$  must be a positive number. You can enter a single value such as 5 or a series of values such as 1 3 5 7 9 or 1 to 9 by 2.

Press the small 'σ' button to the right to obtain calculation options for estimating the standard deviation.

## Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)

### $\rho$ (Intracluster Correlation, ICC)

This is the value of the intracluster (or intraclass) correlation coefficient. It 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 response that can be accounted for by the between-cluster variation.

Possible values are from 0 to just below 1. Typical values are between 0.0001 and 0.05.

You may enter a single value or a list of values.

## Example 1 – Calculating Power

Suppose that a two-level hierarchical design is planned in which there will be only one measurement per subject and treatments will be selected randomly for each subject (level-one unit). The analysis will be a mixed model of continuous data. The following parameter settings are to be used for the power analysis:  $\delta = 0.5$ ;  $\sigma = 1$ ;  $\rho = 0.01$ ;  $M1 = 5$  or  $10$ ;  $ICC = 0.01$ ;  $\alpha = 0.05$ ; and  $K = 5$  to  $20$  by  $5$ .

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)** 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**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alpha.....	<b>0.05</b>
K (Number of Clusters).....	<b>5 10 15 20</b>
M1 (Group 1 Subjects Per Cluster) .....	<b>5 10</b>
M2 (Group 2 Subjects Per Cluster) .....	<b>M1</b>
$\delta$ (Mean Difference = $\mu_1 - \mu_2$ ).....	<b>0.5</b>
$\sigma$ (Standard Deviation) .....	<b>1</b>
$\rho$ (Intracluster Correlation, ICC).....	<b>0.01</b>

### Annotated Output

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

### Numeric Results

Numeric Results for a Two-Level Hierarchical Mixed Model										
	Total Subject Count	Subject Count Grp 1	Subject Count Grp 2	Cluster Count	Subjects Per Clus in Grp 1	Subjects Per Clus in Grp 2	Diff $\mu_1 - \mu_2$	Std Dev	ICC	Alpha
Power	N	N1	N2	K	M1	M2	$\delta$	$\sigma$	$\rho$	
0.4273	50	25	25	5	5	5	0.50	1.00	0.010	0.050
0.7097	100	50	50	5	10	10	0.50	1.00	0.010	0.050
0.7097	100	50	50	10	5	5	0.50	1.00	0.010	0.050
0.9445	200	100	100	10	10	10	0.50	1.00	0.010	0.050
0.8681	150	75	75	15	5	5	0.50	1.00	0.010	0.050
0.9916	300	150	150	15	10	10	0.50	1.00	0.010	0.050
0.9445	200	100	100	20	5	5	0.50	1.00	0.010	0.050
0.9989	400	200	200	20	10	10	0.50	1.00	0.010	0.050

## Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)

### References

Ahn, C., Heo, M., and Zhang, S. 2015. Sample Size Calculations for Clustered and Longitudinal Outcomes in Clinical Research. CRC Press. New York.

### Report Definitions

Power is the probability of rejecting a false null hypothesis. It should be close to one.

N, N1, and N2 are the total number of subjects and the number in groups 1 and 2, respectively.

K is the number of clusters.

M1 is the average number of items (subjects) per cluster in group one, which is designated as the treatment group.

M2 is the average number of items (subjects) per cluster in group two, which is designated as the control group.

$\delta$  is the mean difference ( $\mu_1 - \mu_2$ ) in the response at which the power is calculated.

$\sigma$  is the standard deviation of the subject responses.

$\rho$  (ICC) is the intraclass correlation.

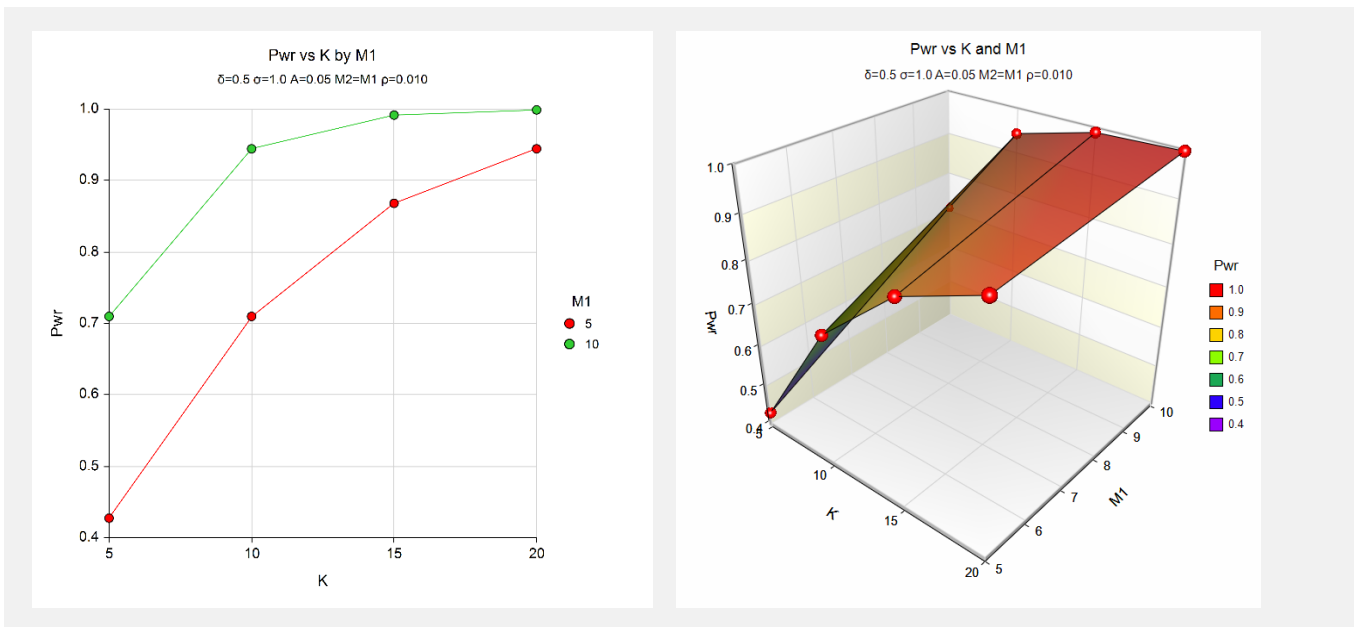
Alpha is the probability of rejecting a true null hypothesis, that is, rejecting when the means are actually equal.

### Summary Statements

Sample sizes of 25 subjects in group one and 25 subjects in group two, which were obtained by sampling 5 clusters with an average of 5 subjects per cluster in group one and 5 subjects per cluster in group two, achieve 43% power to detect a difference between the group means of at least 1. The standard deviation of subjects is 1.00. The intraclass correlation coefficient is 0.010. A test based on a mixed-model analysis is anticipated at a significance level of 0.050.

This report shows the power for each of the scenarios.

### Plots Section



These plots show the power versus the cluster count for the two cluster size values.

## Example 2 – Calculating Sample Size (Number of Clusters)

Continuing with the last example, suppose the researchers want to determine the number of clusters needed to achieve 90% power for both values of M1.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)** 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**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>K (Number of Clusters)</b>
Power.....	<b>0.90</b>
Alpha.....	<b>0.05</b>
M1 (Group 1 Subjects Per Cluster) .....	<b>5 10</b>
M2 (Group 2 Subjects Per Cluster) .....	<b>M1</b>
$\delta$ (Mean Difference = $\mu_1 - \mu_2$ ).....	<b>0.5</b>
$\sigma$ (Standard Deviation) .....	<b>1</b>
$\rho$ (Intracluster Correlation, ICC).....	<b>0.01</b>

### Output

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

#### Numeric Results

Numeric Results for a Two-Level Hierarchical Mixed Model											
	Total Subject Count	Subject Count Grp 1	Subject Count Grp 2	Cluster Count	Subjects Per Clus in Grp 1	Subjects Per Clus in Grp 2	Diff $\mu_1 - \mu_2$	Std Dev	ICC		
Power	N	N1	N2	K	M1	M2	$\delta$	$\sigma$	$\rho$	Alpha	
0.9059	170	85	85	17	5	5	0.50	1.00	0.010	0.050	
0.9209	180	90	90	9	10	10	0.50	1.00	0.010	0.050	

This report shows the number of clusters needed for each scenario.

## Example 3 – Calculating Sample Size (Number of Subjects)

Continuing with the last example, suppose the researchers want to determine the number of subjects needed to achieve 90% power for all values of K.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)** procedure window. You may then make the appropriate entries as listed below, or open **Example 3** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>M1 (Group 1 Subjects Per Cluster)</b>
Power.....	<b>0.90</b>
Alpha.....	<b>0.05</b>
K (Number of Clusters).....	<b>5 10 15 20</b>
M2 (Group 2 Subjects Per Cluster) .....	<b>M1</b>
$\delta$ (Mean Difference = $\mu_1 - \mu_2$ ).....	<b>0.5</b>
$\sigma$ (Standard Deviation) .....	<b>1</b>
$\rho$ (Intracluster Correlation, ICC).....	<b>0.01</b>

### Output

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

### Numeric Results

Numeric Results for a Two-Level Hierarchical Mixed Model											
	Total Subject Count	Subject Count Grp 1	Subject Count Grp 2	Cluster Count	Subjects Per Clus in Grp 1	Subjects Per Clus in Grp 2	Diff $\mu_1 - \mu_2$	Std Dev	ICC		
Power	N	N1	N2	K	M1	M2	$\delta$	$\sigma$	$\rho$	Alpha	
0.9059	170	85	85	5	17	17	0.50	1.00	0.010	0.050	
0.9209	180	90	90	10	9	9	0.50	1.00	0.010	0.050	
0.9209	180	90	90	15	6	6	0.50	1.00	0.010	0.050	
0.9445	200	100	100	20	5	5	0.50	1.00	0.010	0.050	

This report shows the values of M1 and M2 needed for each value of K.



## Example 4 – Validation using Ahn, Heo, and Zhang (2015)

Ahn, Heo, and Zhang (2015) page 157 provide a table in which several scenarios are reported. We will validate this procedure by duplicating two of the table entries. The following parameter settings are used for the power analysis:  $\delta = 0.4$ ;  $\sigma = 1$ ;  $\rho = 0.1$ ;  $K = 10$ ;  $M1=M2 = 10$ ;  $ICC = 0.1$ ; and  $\alpha = 0.05$ . The reported value of the power is 0.846.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Mixed Models Tests for Two Means in a 2-Level Hierarchical Design (Level-1 Randomization)** procedure window. You may then make the appropriate entries as listed below, or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

#### Design Tab

Solve For ..... **Power**  
 Alpha..... **0.05**  
 K (Number of Clusters)..... **10**  
 M1 (Group 1 Subjects Per Cluster) ..... **10**  
 $\delta$  (Mean Difference =  $\mu_1 - \mu_2$ )..... **0.4**  
 $\sigma$  (Standard Deviation) ..... **1**  
 $\rho$  (Intracluster Correlation, ICC)..... **0.1**

### Output

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

#### Numeric Results

Numeric Results for a Two-Level Hierarchical Mixed Model											
	Total Subject Count	Subject Count Grp 1	Subject Count Grp 2	Cluster Count	Subjects Per Clus in Grp 1	Subjects Per Clus in Grp 2	Diff $\mu_1 - \mu_2$	Std Dev	ICC	Alpha	Power
	N	N1	N2	K	M1	M2	$\delta$	$\sigma$	$\rho$		
	200	100	100	10	10	10	0.40	1.00	0.100	0.050	0.8465

PASS also calculates the power at 0.846.