

## Chapter 243

# GEE Tests for Two Means in a Stratified Cluster-Randomized Design

---

### Introduction

This procedure calculates power and sample size for tests of two means in a stratified cluster-randomized design in which the outcome variable is continuous. It uses the work of Wang, Zhang, and Ahn (2017) which give the power in a size-stratified cluster-randomized design in which the cluster size is allowed to vary within strata. The analysis is of a simple means model fit with the GEE method.

---

### Technical Details

The following discussion summarizes the results in Wang, Zhang, and Ahn (2017).

Suppose you are interested in comparing the means of two groups (treatment and control). Further suppose that the response is known to be related to other covariates (such as age, race, or gender) and so their impact needs to be adjusted for. This may be accomplished by stratifying on the covariates and forming hypotheses about a common mean difference across all clusters and strata. Often, the stratification is based on cluster size, but this is not required.

Let  $Y_{kji}$  be the continuous outcome of the  $i^{\text{th}}$  ( $i = 1, \dots, n_{kj}$ ) subject in the  $j^{\text{th}}$  ( $j = 1, \dots, J_k$ ) cluster of the  $k^{\text{th}}$  ( $k = 1, \dots, K$ ) stratum. Let  $X_{kj} = 0$  or 1 depending on whether the cluster is assigned to the control (0) or treatment (1) group. Let  $R = P(X_{kj} = 1) \times 100$  be the percentage of clusters assigned to the treatment group. Hence,  $R = 50$  indicates balanced randomization.

Suppose the data will be analyzed using the GEE (generalized estimating equation) method. The mean model is

$$E(Y_{kji}) = \beta_0 + \beta_1 X_{kj}$$

where  $\beta_0 = \mu_2$  is the mean of the control group,  $\mu_1$  is the mean of the treatment group, and  $\beta_1 = \mu_1 - \mu_2 = \delta$  is the treatment effect.

The null hypothesis is  $H_0: \beta_1 = 0$  which is paired with the alternative hypothesis  $H_1: \beta_1 \neq 0$ .

The variance matrix of  $Y_{kji}$  has  $\sigma^2$  on the diagonal and  $\rho\sigma^2$  on the off-diagonal for subjects within the same cluster. Hence  $\sigma^2$  is the variance of random error and  $\rho$  is the intracluster correlation coefficient (ICC).

If the GEE estimators of  $\beta_0$  and  $\beta_1$ , found using an independent working correlation structure, are  $\hat{\beta}_0$  and  $\hat{\beta}_1$ , then  $\sqrt{N}(\hat{\beta}_1 - \beta_1)$  follows a normal distribution.

## GEE Tests for Two Means in a Stratified Cluster-Randomized Design

Let  $\alpha$  is the type I error rate,  $\delta$  is the actual, non-zero, value of the treatment effect, and

$$v_1 = \frac{\sigma^2 \sum_{k=1}^K J_k \theta_k^2 \left[ \frac{1}{\theta_k} (1-\rho) + (1+\xi_k^2) \rho \right]}{R/100 (\sum_{k=1}^K J_k \theta_k)^2},$$

$$v_2 = \frac{\sigma^2 \sum_{k=1}^K J_k \theta_k^2 \left[ \frac{1}{\theta_k} (1-\rho) + (1+\xi_k^2) \rho \right]}{(1-R/100) (\sum_{k=1}^K J_k \theta_k)^2},$$

$$E(n_{kj}) = \theta_k \text{ (Average of cluster sizes in stratum } k),$$

$$\text{Var}(n_{kj}) = \tau_k^2 \text{ (Variance of cluster sizes in stratum } k),$$

$$\text{CV}(n_{kj}) = \frac{\tau_k}{\theta_k} = \xi_k \text{ (Variance of cluster sizes in stratum } k),$$

$$N = \sum_{k=1}^K J_k \theta_k \text{ (Total number of subjects in study),}$$

$$f_k = 100 \frac{J_k \theta_k}{N} \text{ (The percentage of the total sample in stratum } k).$$

Three cases of statistical hypotheses are available. If  $\Phi(x)$  is the standard normal cumulative distribution function, the power formula for each case is

1.  $H_0: \delta \geq 0$  versus  $H_1: \delta < 0$ .

$$\text{Power} = \Phi \left( z_\alpha - \frac{\delta}{\sqrt{v_1 + v_2}} \right)$$

2.  $H_0: \delta \leq 0$  versus  $H_1: \delta > 0$ .

$$\text{Power} = \Phi \left( \frac{\delta}{\sqrt{v_1 + v_2}} - z_{1-\alpha} \right)$$

3.  $H_0: \delta = 0$  versus  $H_1: \delta \neq 0$ .

$$\text{Power} = \Phi \left( z_{\alpha/2} - \frac{\delta}{\sqrt{v_1 + v_2}} \right) + \Phi \left( \frac{\delta}{\sqrt{v_1 + v_2}} - z_{1-\alpha/2} \right)$$

These power formulas are used to conduct a binary search for sample size or odds ratio.

---

## 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, refer to the Procedure Window chapter.

---

## Design Tab

The Design tab contain most of the parameters and options of interest for this procedure.

---

### Solve For

#### Solve For

This option specifies the parameter to be solved for using the other parameters. The parameters that may be selected are  $\delta$ , *Power*, or *Sample Size*. Select *Sample Size* when you want to calculate the sample size needed to achieve a given power and alpha level. Select *Power* when you want to calculate the power.

---

## Test

### Alternative Hypothesis

Specify the alternative hypothesis, H1.

- **One-Sided (H1:  $\delta < 0$ )**  
Refers to a lower-tailed, one-sided test with alternative hypothesis H1:  $\delta < 0$ .
- **One-Sided (H1:  $\delta > 0$ )**  
Refers to an upper-tailed, one-sided test with alternative hypothesis H1:  $\delta > 0$ .
- **Two-Sided (H1:  $\delta \neq 0$ )**  
Refers to a two-sided test with alternative hypothesis H1:  $\delta \neq 0$ .

Note that the alternative hypothesis enters into power calculations by specifying the rejection region of the hypothesis test. Its accuracy is critical.

---

## 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. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal proportions when in fact they are different.

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 here 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. For this procedure, a type-I error occurs when you reject the null hypothesis of equal group proportions when in fact they are equal.

Values must be between zero and one. Historically, an alpha value of 0.05 has been recommended for two-sided tests and 0.025 for one-sided tests. 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

### N (Sample Size)

Enter one or more values of N, the total number of subjects in the study. This parameter is the sum of the subjects in all clusters or strata. The number of treatment and control subjects depends on the Group Allocation option.

The N subjects are divided among the strata using the 'Percent of N in Each Stratum' setting. These are the  $f_k$  values defined above

### Example

Suppose that N is 1000 and three strata are used with 'Percent of N in Each Stratum' set to 25, 25, 50. The three stratum sample sizes would be 250, 250, and 500.

The number of clusters is based on this number and the 'Average of Cluster Sizes' settings.

## GEE Tests for Two Means in a Stratified Cluster-Randomized Design

### Group Allocation

Select how clusters are allocated to the treatment and control groups.

The choices are

- **Equal (Treatment = Control = 50%)**  
The clusters are allocated equally to the treatment and control groups. This is accomplished by setting the percentage allocated to the treatment group (R) to 50.
- **Enter R, the Percentage in Group 1**  
The clusters are allocated so that R% of them are allocated to the treatment group. The remaining (100-R)% of the clusters are assigned to the control group.

### R (Percentage in Group 1)

Enter one or more values of the percentage of clusters within each stratum that are allocated to group 1 (treatment). All subjects so allocated will receive the treatment. All remaining clusters are allocated to group 2 (control).

#### Notes

A balanced allocation occurs when this parameter is set to 50.

Since this is a percentage and since at least some clusters must be allocated to each group, this value can range from 1 to 99.

Realistically, this value will be between 20 and 80.

The smallest total sample size will result from setting this value to 50.

---

### 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 difference you want to be able to detect.

$\delta$  can be any non-zero value (positive or negative). The difference has the same scale as the standard deviation.

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. It may also be thought of as the standard deviation of the random error term in the GEE model.

$\sigma$  must be greater than zero.

You can enter a single value such as '5' or a series of values such as '1 3 5 7 9' or '1 to 10 by 1'.

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

Press the small ' $\sigma$ ' button to the right to obtain help in estimating the standard deviation.

#### ICC (Intraclass Correlation)

Enter a value for ICC, the intraclass correlation coefficient. This may be thought of as the correlation between any two observations in the same cluster.

Note that occasionally,  $ICC < 0$ . Since this is a correlation coefficient, it is possible for it to range from -1 to 1. You may enter a single value or a list of values, such as *0.01 0.02 0.03*.

---

## Strata Information

### Cluster Size Variation Input

There are two ways in which you can specify the variation of the cluster sizes within a stratum. These are

- **Standard Deviation**

You will enter the population standard deviation of the cluster sizes within each stratum.

- **Coefficient of Variation**

You will enter the population coefficient of variation of the cluster sizes within each stratum.

### Set

This is an identification number used on the reports.

### Number of Strata

Specify the number of strata specified on this line. Usually, you will enter a '1' to specify a single stratum, or you will enter a '0' to ignore this line. However, this option lets you specify several strata that have the same parameter values.

The total number of strata is equal to the sum of these values.

### Examples

0 which means 'ignore this line'.

1 which means 'one stratum defined by this line'.

2 which means 'two strata defined by this line'.

### Percent of N in Each Stratum

Enter the percentage that the number of subjects in a stratum is of the total number of subjects. If this line defines more than one stratum, this is the percentage FOR EACH stratum.

These percentages must sum to 100. If they don't, the program will adjust them, so they do.

For example, suppose you enter two sets and specify a Number of Strata for both sets of 5. Further suppose that you enter 25 for the Percent of N in Each Stratum in both sets. The program would realize that there is a total of 10 strata in the experiment with an equal percentage of the total in each stratum. Thus, the individual percentages would be adjusted from 25 to 10.

Note that you should only enter one number, even if there are more than one stratum being defined by the line.

### Range

These values can be any positive values. The values will be adjusted so that the percentages sum to 100.

### Average of Cluster Sizes

Enter the average cluster size of the clusters in a stratum in this set. This number applies to clusters in both the treatment and the control groups.

The value must be greater than or equal to 1.

## GEE Tests for Two Means in a Stratified Cluster-Randomized Design

### Standard Deviation of Cluster Sizes

Enter the population standard deviation of the cluster sizes in this set. This number applies to both the treatment and the control groups.

The values must be than or equal to zero.

### Example

Suppose the five cluster sizes in a particular stratum are 225, 314, 197, 266, and 248. The average cluster size is 250 and the population standard deviation (using  $N = 5$  in the denominator) is 39.4715.

### Coefficient of Variation of Cluster Sizes

Enter the coefficient of variation of the cluster sizes in this set. This number will apply to both the treatment and the control groups.

### Coefficient of Variation (CV)

The coefficient of variation used here is the population standard deviation divided by the mean. It is used to calculate the standard deviation of the cluster sizes. It is often used because the standard deviation of cluster sizes is often roughly proportional to the cluster size mean.

The values must be greater than zero.

### Show More Strata Sets

Check this box to show ten more Strata Information sets.

If this option is not checked, any active strata sets (Number of Strata  $> 0$ ) with set identification numbers  $> 5$  will be ignored.

## Example 1 – Finding Sample Size

A study is being planned to investigate whether a new intervention will decrease a certain response variable. For a number of reasons, the researchers decide to administer the intervention to whole clusters (clinics) rather than randomize the treatment to individuals within the cluster. Past experience has shown that clinics can be separated into small, medium, and large according to their size. They want to obtain an equal number of subjects in each stratum. The number of clinics receiving each treatment will be balanced.

The average number of subjects per clinic of the small, medium, and large strata are 6, 21, and 73, respectively.

The coefficient of variation of all strata is 0.42.

Prior studies have obtained a mean score of 80 with a standard deviation of 23 and an ICC of 0.03. The researchers want to compare the necessary sample sizes when the change is -6, -8, and -10. They also want to consider an ICC value of 0.06.

The two-sided significance level is set to 0.05 and the power is set to 0.8.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load this procedure. 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>Sample Size</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\delta \neq 0</math>)</b>
Power .....	<b>0.8</b>
Alpha .....	<b>0.05</b>
Group Allocation .....	<b>Equal (Treatment = Control = 50%)</b>
$\delta$ (Mean Difference = $\mu_1 - \mu_2$ ) .....	<b>-6 -8 -10</b>
$\sigma$ (Standard Deviation) .....	<b>23</b>
ICC (Intraclass Correlation) .....	<b>0.03 0.06</b>
Cluster Size Variation Input .....	<b>Coefficient of Variation</b>
Set 1 Number of Strata .....	<b>1</b>
Set 1 Percent of N in Each Stratum .....	<b>33</b>
Set 1 Average of Cluster Sizes .....	<b>6</b>
Set 1 Coefficient of Variation of Cluster Sizes .....	<b>0.42</b>
Set 2 Number of Strata .....	<b>1</b>
Set 2 Percent of N in Each Stratum .....	<b>33</b>
Set 2 Average of Cluster Sizes .....	<b>21</b>
Set 2 Coefficient of Variation of Cluster Sizes .....	<b>0.42</b>
Set 3 Number of Strata .....	<b>1</b>
Set 3 Percent of N in Each Stratum .....	<b>33</b>
Set 3 Average of Cluster Sizes .....	<b>73</b>
Set 3 Coefficient of Variation of Cluster Sizes .....	<b>0.42</b>

## GEE Tests for Two Means in a Stratified Cluster-Randomized Design

Set 4 Number of Strata ..... 0  
 Set 5 Number of Strata ..... 0  
 Show More Strata Sets..... **Unchecked**

## Annotated Output

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

### Numeric Results

#### Numeric Results

Hypotheses: H0:  $\delta = 0$  vs. H1:  $\delta \neq 0$   
 Number of Strata: 3

Power	Total Sample Size N	Expected Number of Clusters	Percent Allocated to Treatment Group R	Actual Mean Difference $\delta = \mu_1 - \mu_2$	Standard Deviation $\sigma$	Intracluster Correlation Coefficient ICC	Alpha
0.8000	356	28	50.00	-10.00	23.00	0.03	0.050
0.8000	547	41	50.00	-10.00	23.00	0.06	0.050
0.8000	557	43	50.00	-8.00	23.00	0.03	0.050
0.8000	854	65	50.00	-8.00	23.00	0.06	0.050
0.8000	990	76	50.00	-6.00	23.00	0.03	0.050
0.8000	1519	115	50.00	-6.00	23.00	0.06	0.050

#### References

Wang, J., Zhang, S., and Ahn, C. 2017. 'Power analysis for stratified cluster randomisation trials with cluster size being the stratifying factor.' *Statistical Theory and Related Fields*, Volume 1, Number 1, pages 121-127.

#### Report Definitions

Power is the probability of rejecting a false null hypothesis. It should be close to one.  
 N is the total sample size, i.e., the total number of subjects summed across all strata.  
 Expected Number of Clusters is an estimate of the total number of clusters of all sizes.  
 R is the percent of clusters in each stratum that are allocated to the treatment group.  
 $\delta = \mu_1 - \mu_2$  is the mean difference assumed by the alternative hypothesis. This is the difference at which the power is computed.  
 $\sigma$  is the standard deviation of the response. It is the standard deviation of the random error component of the GEE means model.  
 ICC is the intracluster correlation coefficient. This is the correlation between any two responses within a particular cluster.  
 Alpha is the probability of rejecting a true null hypothesis.

#### Summary Statements

A stratified cluster-randomized design, which divides the sample among 3 strata, is analyzed using a two-sided test of the mean difference which is obtained from fitting a means model using GEE. The total sample size of 356 subjects, which is obtained from 28 clusters, achieves 80% power to reject the mean difference of zero assumed by the null hypothesis when the difference is actually -10.00. The standard deviation of the random error is 23.00. The intracluster correlation coefficient is 0.03. The significance level of the test is 0.050. Note that each cluster is completely allocated to either the treatment or control and that 50.00% of the clusters in each stratum are in the treatment group.

GEE Tests for Two Means in a Stratified Cluster-Randomized Design

Strata-Detail Report

Strata-Detail Report

Set	Number of Strata	Percent of N in Each Stratum	Average of Cluster Sizes	Standard Deviation of Cluster Sizes	Coefficient of Variation of Cluster Sizes
1	1	33.33	6.0	2.52	0.420
2	1	33.33	21.0	8.82	0.420
3	1	33.33	73.0	30.66	0.420

Strata-Detail Report Definitions

Set is an identification number on the input screen.

Number of Strata is the number of strata defined by this set (input line).

Percent of N in Each Stratum is the percentage of the total sample size that is allocated to each stratum in this set.

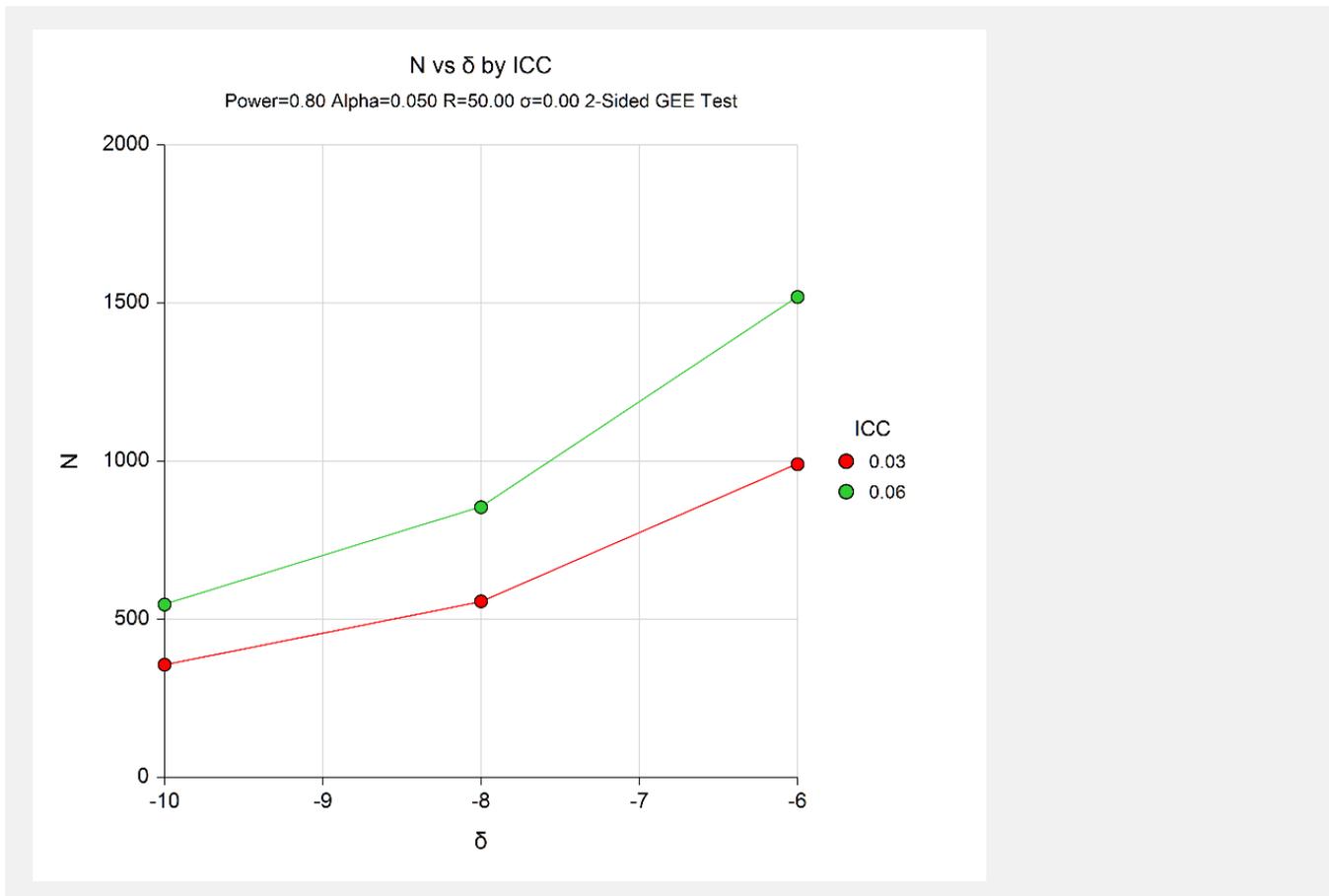
Average of Cluster Sizes is the average cluster size (number of subjects) in each stratum of this set.

Standard Deviation of Cluster Sizes is the standard deviation of the cluster sizes in each stratum of this set.

Coefficient of Variation of Cluster Sizes is the coefficient of variation (SD/Mean) of the cluster sizes in each stratum of this set.

This report shows the values of the individual, strata-level parameters that were used.

Plots Section



The values from the Numeric Results report are displayed in this plot.

## Example 2 – Validation using Wang, Zhang, and Ahn (2017)

Wang, Zhang, and Ahn (2017) provide an example of the power analysis of a size-stratified, cluster-randomized study. We will use their results to validate this procedure.

The number of strata were 200, 510, and 1300. We will enter these values and let the program rescale them into percentages. The total number of subjects is 2010. The average number of subjects per cluster were 5, 17, and 65. The corresponding standard deviations were 2.44949, 5.00000, and 22.36068.

The mean difference to be detected is 3. The standard deviation is 12. The value of ICC is 0.05. The significance level is set to 0.05. They calculate a power of 0.8432.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load this procedure. 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>Power</b>
Alternative Hypothesis .....	<b>Two-Sided (H1: <math>\delta \neq 0</math>)</b>
Alpha .....	<b>0.05</b>
N (Sample Size).....	<b>2010</b>
Group Allocation .....	<b>Equal (Treatment = Control = 50%)</b>
$\delta$ (Mean Difference = $\mu_1 - \mu_2$ ) .....	<b>3</b>
$\sigma$ (Standard Deviation) .....	<b>12</b>
ICC (Intracluster Correlation).....	<b>0.05</b>
Cluster Size Variation Input .....	<b>Standard Deviation</b>
Set 1 Number of Strata .....	<b>1</b>
Set 1 Percent of N in Each Stratum.....	<b>200</b>
Set 1 Average of Cluster Sizes.....	<b>5</b>
Set 1 Standard Deviation of Cluster Sizes .....	<b>2.44949</b>
Set 2 Number of Strata .....	<b>1</b>
Set 2 Percent of N in Each Stratum.....	<b>510</b>
Set 2 Average of Cluster Sizes.....	<b>17</b>
Set 2 Standard Deviation of Cluster Sizes .....	<b>5.00000</b>
Set 3 Number of Strata .....	<b>1</b>
Set 3 Percent of N in Each Stratum.....	<b>1300</b>
Set 3 Average of Cluster Sizes.....	<b>65</b>
Set 3 Standard Deviation of Cluster Sizes .....	<b>22.36068</b>
Set 4 Number of Strata .....	<b>0</b>
Set 5 Number of Strata .....	<b>0</b>
Show More Strata Sets.....	<b>Unchecked</b>

## GEE Tests for Two Means in a Stratified Cluster-Randomized Design

**Output**

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

**Numeric Results with Continuity Correction**

Numeric Results								
Hypotheses:		H0: $\delta = 0$ vs. H1: $\delta \neq 0$						
Number of Strata:		3						
Power	Total Sample Size N	Expected Number of Clusters	Percent Allocated to Treatment Group R	Actual Mean Difference $\delta = \mu_1 - \mu_2$	Standard Deviation $\sigma$	Intracluster Correlation Coefficient ICC	Alpha	
0.8432	2010	90	50.00	3.00	12.00	0.05	0.050	
Strata-Detail Report								
Set	Number of Strata	Percent of N in Each Stratum	Average of Cluster Sizes	Standard Deviation of Cluster Sizes	Coefficient of Variation of Cluster Sizes			
1	1	9.95	5.0	2.45	0.490			
2	1	25.37	17.0	5.00	0.294			
3	1	64.68	65.0	22.36	0.344			

PASS has also obtained a power of 0.8432 which validates this procedure.