

Chapter 541

GEE Tests for Two Means in a Split-Mouth Design

Introduction

This procedure assumes that continuous data will be obtained from a study that uses a *split-mouth* design. The GEE method is used to analyze the repeated measures model that is assumed. The sample size formula is derived in Zhu, Zhang, and Ahn (2017).

A *split-mouth* design is used in dental trials in which treatments are randomized over segments of the mouth within each subject. In this design, the mouth is divided into two or more *segments* or regions. For example, the segments might be top and bottom, left and right, or a combination of both. Within each segment, specific *sites* (e.g. teeth) are identified. The same treatment is applied to all sites within a segment.

The split-mouth design, also called the *split-cluster* design, is occasionally used in other areas such as dermatology and animal studies. Although the design may be used in other experiments, the terminology of the split-mouth design will be used in this procedure.

Technical Details

Our formulation comes from Zhu, Zhang, and Ahn (2017). The goal of this analysis will be to provide a test comparing the means of two groups. We assume that the two groups are labelled “1” for treatment and “2” for control.

Denote an observation by Y_{ijk} where $i = 1, 2$ gives the group, $j = 1, 2, \dots, N$ gives the subject, and $k = 1, 2, \dots, M$ gives the site. Note that the segments themselves are not specified directly in the model. Also note that the total number of sites per subject is $2M$.

The following linear regression model will be analyzed using GEE.

$$Y_{ijk} = \beta_0 + \beta_1 x_{ij} + \epsilon_{ijk}$$

where β_0 is the intercept, β_1 is the difference between the group means ($\delta = \mu_1 - \mu_2$), $x_{ij} = 1$ if the site is in group 1 and $x_{ij} = 0$ if the site is in group 2, and ϵ_{ijk} is random error. The variance of ϵ_{ijk} is σ^2 .

Two alternative models are available to model the correlation between Y_{ijk} and $Y_{i'jk'}$ (responses within the same subject). The simpler model assumes that a common, within-subject, correlation ρ is at work. A more general model assumes that there are two, within-subject, correlations. These are ρ_W , which is the correlation between responses of sites within the same segment, and ρ_B , which is the correlation between responses of sites that are not in the same segment.

GEE Tests for Two Means in a Split-Mouth Design

Zhu, Zhang, and Ahn (2017) provide formulas for each correlation structure. For the one, within-subject correlation model, the formula for sample size is

$$N = \frac{2\sigma^2(1 + (M - 1)\rho_W - M\rho_B) \left(z_{1-\frac{\alpha}{2}} + z_{1-\beta} \right)^2}{M\delta^2}$$

For the two, within-subject correlations model, the formula for sample size is

$$N = \frac{2\sigma^2(1 - \rho) \left(z_{1-\frac{\alpha}{2}} + z_{1-\beta} \right)^2}{M\delta^2}$$

where α is the probability of a type-I error and β is the probability of a type-II error. The test is assumed to be two-sided.

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 *Power*, *Sample Size*, or *Effect Size*.

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.

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.

GEE Tests for Two Means in a Split-Mouth Design

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.

Note that this procedure assumes a two-sided test.

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

N (Subjects)

Enter one or more values for N, the number of subjects in the study. If you enter a list of values, a separate analysis is done for each value.

Range

$N > 1$

Examples: *10 20 30 40* or *20 to 200 by 20*.

M (Sites Per Group Per Subject)

Enter one or more values of the number of sites (e.g. teeth) per group per subject. The design is balanced so that M is the number of sites in all segments assigned to the treatment group and also M is the number of sites in all segments assigned to the control group.

The split-mouth design assumes that the mouth is subdivided into two or more segments (e.g. left and right, top and bottom, etc.), each made up of a number of sites. The two treatments are randomly assigned to segments in such a way that both groups include M sites.

Range

This value must be a positive integer that is at least 2.

List

You can use a list of values such as "4 6 8". A separate analysis will be run for each element in the list.

Effect Size

δ (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 does not represent the true difference between the means. Rather, it is the value at which you want the sample size computed.

Range

δ can be any non-zero value (positive or negative), but it should be consistent with the choice of the alternative hypothesis. Also, it should have 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 result for each value of the series.

GEE Tests for Two Means in a Split-Mouth Design

σ (Standard Deviation)

Enter a value (or range of values) for the standard deviation of the random error value (the residuals). The analysis assumes that this standard deviation is the same for both groups.

Range

These values must be positive.

You can enter a single value such as *10* or a series of values such as *10 20 30 40 50* or *1 to 5 by 0.5*.

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

Correlation Matrix Structure

Specify the structure of the correlation matrix of sites within a subject. The possible choices are

- **One Within-Subject Correlation (ρ)**

A common correlation is assumed for all outcome pairs within the same subject.

- **Two Within-Subject Correlations (ρ_B, ρ_W)**

Two within-subject correlations are needed:

ρ_B is the correlation between sites in different segments within a subject.

ρ_W is the correlation between sites within a particular segment within a subject.

ρ (Within-Subject Correlation)

Enter one or more values for ρ , the common correlation of site outcomes within the same subject. At least one value must be entered. If multiple values are entered, a separate analysis is performed for each value.

Range

$-1 < \rho < 1$. A value near 0 indicates low correlation. A value near 1 indicates high correlation.

Note that negative values may cause numerical trouble and should be avoided.

You can enter a single value such as *0.5* or a series of values such as *0.5 0.6 0.7* or *0 to 0.8 by 0.1*.

ρ_B (Between-Segment Correlation)

Enter one or more values for ρ_B , the correlation of site outcomes that are in different segments within a subject. At least one value must be entered. If multiple values are entered, a separate analysis is performed for each value.

Range

$-1 < \rho < 1$. A value near 0 indicates low correlation. A value near 1 indicates high correlation.

Note that negative values may cause numerical trouble and should be avoided.

You can enter a single value such as *0.5* or a series of values such as *0.5 0.6 0.7* or *0 to 0.8 by 0.1*.

ρ_W (Within-Segment Correlation)

Enter one or more values for ρ_W , the correlation of site outcomes that are in the same segments within a subject. At least one value must be entered. If multiple values are entered, a separate analysis is performed for each value.

Range

$-1 < \rho < 1$. A value near 0 indicates low correlation. A value near 1 indicates high correlation.

Note that negative values may cause numerical trouble and should be avoided.

You can enter a single value such as *0.5* or a series of values such as *0.5 0.6 0.7* or *0 to 0.8 by 0.1*.

GEE Tests for Two Means in a Split-Mouth Design

Example 1 – Calculating Power

Suppose a dental clinical trial is being planned in which the mouth is divided into four segments with three sites in each. The trial is being conducted to compare a new treatment for gingivitis with a common existing treatment. In the study, suppose $\delta = 0.8$; $\sigma = 5$ or 6 ; $\rho = 0.42$; $M = 6$; $alpha = 0.05$; and $N = 50$ 75 100 125 150 . Power is to be calculated for 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 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>
Design Tab	
Solve For	Power
Alpha (Two-Sided Test)	0.05
N (Subjects)	50 75 100 125 150
M (Sites Per Group Per Subject)	6
δ (Mean Difference = $\mu_1 - \mu_2$)	0.8
σ (Standard Deviation)	5 6
Correlation Matrix Structure	One Within-Subject Correlation (ρ)
ρ (Within-Subject Correlation)	0.42

Annotated Output

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

Numeric Results

Numeric Results

	Total Sample Size	Sites Per Group Per Subject	Mean Diff $\mu_1 - \mu_2$	Std Dev σ	Effect Size δ/σ	Within Subj Corr ρ	Alpha
Power	N	M	δ	σ	δ/σ	ρ	
0.7301	50	6	0.80	5.00	0.160	0.420	0.050
0.5731	50	6	0.80	6.00	0.133	0.420	0.050
0.8832	75	6	0.80	5.00	0.160	0.420	0.050
0.7473	75	6	0.80	6.00	0.133	0.420	0.050
0.9534	100	6	0.80	5.00	0.160	0.420	0.050
0.8582	100	6	0.80	6.00	0.133	0.420	0.050
0.9825	125	6	0.80	5.00	0.160	0.420	0.050
0.9237	125	6	0.80	6.00	0.133	0.420	0.050
0.9937	150	6	0.80	5.00	0.160	0.420	0.050
0.9603	150	6	0.80	6.00	0.133	0.420	0.050

References

Zhu, H., Zhang, S., Ahn, C. 2017. 'Sample size considerations for split-mouth design'. Statistical Methods in Medical Research. Volume 26(6). Pages 2543-2551.

GEE Tests for Two Means in a Split-Mouth Design

Report Definitions

Power is the probability of rejecting a false null hypothesis.

Total Sample Size N is the total number of subjects in the study.

Sites Per Group Per Subject M is the number of sites per subject in the treatment group and the control group.

Mean Diff $\mu_1 - \mu_2$ δ is the mean difference between the treatment group and the control group.

Std Dev σ is the standard deviation of random error (residuals) within a group.

Effect Size δ/σ is a measure of the effect size. It is the ratio of δ and σ .

Within Subj Corr ρ is the correlation between sites within a subject.

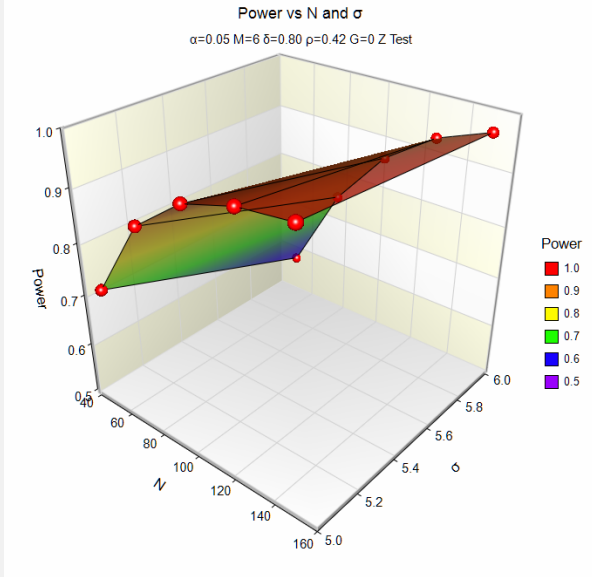
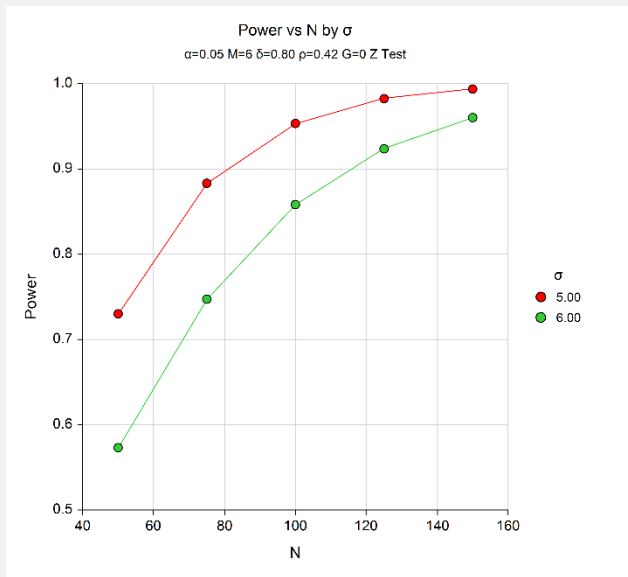
Alpha is the significance level of the test: the probability of rejecting the null hypothesis of equal means when it is true.

Summary Statements

A split-mouth (split-cluster) design is used to test the difference between two means. The data will be analyzed using a GEE regression model. A sample of 50 subjects with 6 sites per group per subject achieves 73% power at a significance level of 0.050 to detect a difference of 0.80. The correlation matrix is assumed to be compound symmetric with a within-subject correlation of 0.420. The standard deviation of the random error is 5.00.

This report gives the power for each of the scenarios.

Plots Section



These plots show the power versus the sample size for the two standard deviations.

Example 2 – Finding Sample Size and Validation using Zhu, Zhang, and Ahn (2017)

Zhu, Zhang, and Ahn (2017) page 2549 provide a table in which several sample size values are presented. We will use a few entries from this table to validate this procedure.

In the first row of Table 1, $\delta = 0.2$; $\sigma = 0.7071$; $\rho_B = 0.05, 0.1, 0.15$; $\rho_W = 0.1$; $M = 3$; $\alpha = 0.05$; and power = 0.8. The resulting N's are 69, 59, and 49.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the 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**.

Option

Value

Design Tab

Solve For	Sample Size
Power.....	0.8
Alpha (Two-Sided Test).....	0.05
M (Sites Per Group Per Subject).....	3
δ (Mean Difference = $\mu_1 - \mu_2$)	0.2
σ (Standard Deviation)	0.7071
Correlation Matrix Structure.....	Two Within-Subject Correlations (ρ_B, ρ_W)
ρ_B (Between-Segment Correlation)	0.05 0.1 0.15
ρ_W (Within-Segment Correlation)	0.1

Output

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

Numeric Results									
	Total Sample Size N	Sites Per Group Per Subject M	Mean Diff $\mu_1 - \mu_2$ δ	Std Dev σ	Effect Size δ/σ	Between Segment Corr ρ_B	Within Segment Corr ρ_W	Alpha	
Power	69	3	0.20	0.71	0.283	0.05	0.10	0.050	
	59	3	0.20	0.71	0.283	0.10	0.10	0.050	
	50	3	0.20	0.71	0.283	0.15	0.10	0.050	

PASS matches the first to scenarios exactly. Table 1 gives $N = 49$ for the third row, but PASS obtains 50. We checked the power for $N = 49$ and obtained a power of 0.7996 which is just slightly less than the required value of 0.8. The difference is due to rounding. Thus, PASS reports $N = 50$ which has a power greater than 0.8.