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, ..., N gives the subject, and k = 1, 2, ..., 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.

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.

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**.

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 = µ1 - µ2).......................**0.8**
σ (Standard Deviation)...................................**5 6**
Correlation Matrix Structure ....**One Within-Subject Correlation (ρ)**
ρ (Within-Subject Correlation).......................**0.42**

Output

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

Numeric Reports

Numeric Results
───
Solve For: Power
Groups: 1 = Treatment, 2 = Control
Difference: δ = µ1 - µ2
───

Power	Total Sample Size N	Sites per Group per Subject M	Mean Difference δ	Standard Deviation σ	Effect Size δ / σ	Within-Subject Correlation ρ	Alpha
0.7301	50	6	0.8	5	0.160	0.42	0.05
0.5731	50	6	0.8	6	0.133	0.42	0.05
0.8832	75	6	0.8	5	0.160	0.42	0.05
0.7473	75	6	0.8	6	0.133	0.42	0.05
0.9534	100	6	0.8	5	0.160	0.42	0.05
0.8582	100	6	0.8	6	0.133	0.42	0.05
0.9825	125	6	0.8	5	0.160	0.42	0.05
0.9237	125	6	0.8	6	0.133	0.42	0.05
0.9937	150	6	0.8	5	0.160	0.42	0.05
0.9603	150	6	0.8	6	0.133	0.42	0.05

Power   The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N   The total number of subjects in the study.
M   The number of sites per subject in the treatment group and the control group.
δ   The mean difference between the treatment group and the control group. δ = μ1 - μ2.
σ   The standard deviation of random error (residuals) within a group.
δ / σ   Effect Size. It is the ratio of δ and σ.
ρ   The correlation between sites within a subject.
Alpha   The probability of rejecting a true null hypothesis.

Summary Statements

A split-mouth (split-cluster) design will be used to test whether the Group 1 (treatment) mean (μ1) is different from
the Group 2 (control) mean (μ2). The comparison will be made using a two-sided generalized estimating equation
(GEE) regression coefficient test, with a Type I error rate (α) of 0.05. The standard deviation of the random error
term (residuals) is assumed to be 5 and is assumed to be equal for both groups. The correlation matrix is
assumed to be compound symmetric with a within-subject correlation of 0.42. To detect a difference (μ1 - μ2) of
0.8 with a total of 50 subjects, with 6 sites per group per subject, the power is 0.7301.

Dropout-Inflated Sample Size

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	50	63	13
20%	75	94	19
20%	100	125	25
20%	125	157	32
20%	150	188	38

Dropout Rate   The percentage of subjects (or items) that are expected to be lost at random during the course of the study
               and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N   The evaluable sample size at which power is computed (as entered by the user). If N subjects are evaluated
               out of the N' subjects that are enrolled in the study, the design will achieve the stated power.
N'   The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects,
               based on the assumed dropout rate. N' is calculated by inflating N using the formula N' = N / (1 - DR), with
               N' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and
               Lokhnygina, Y. (2018) pages 32-33.)
D   The expected number of dropouts. D = N' - N.

Dropout Summary Statements

Anticipating a 20% dropout rate, 63 subjects should be enrolled to obtain a final sample size of 50 subjects.
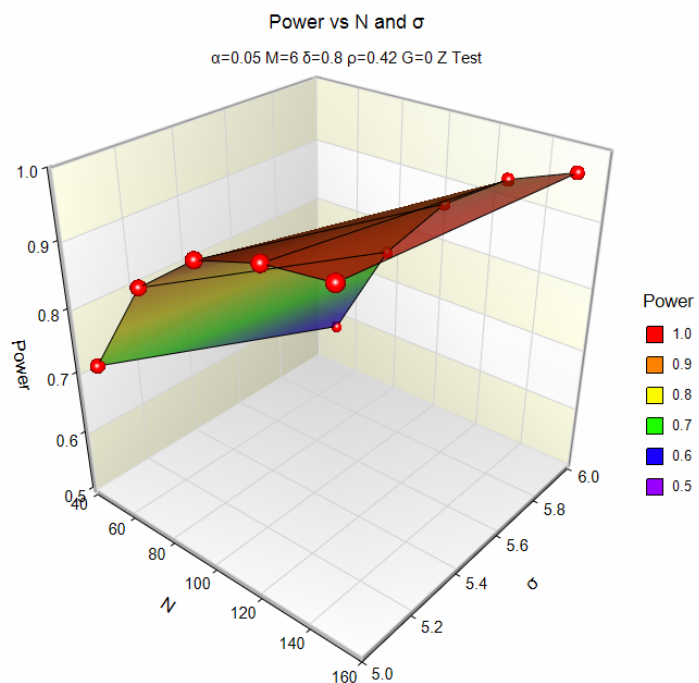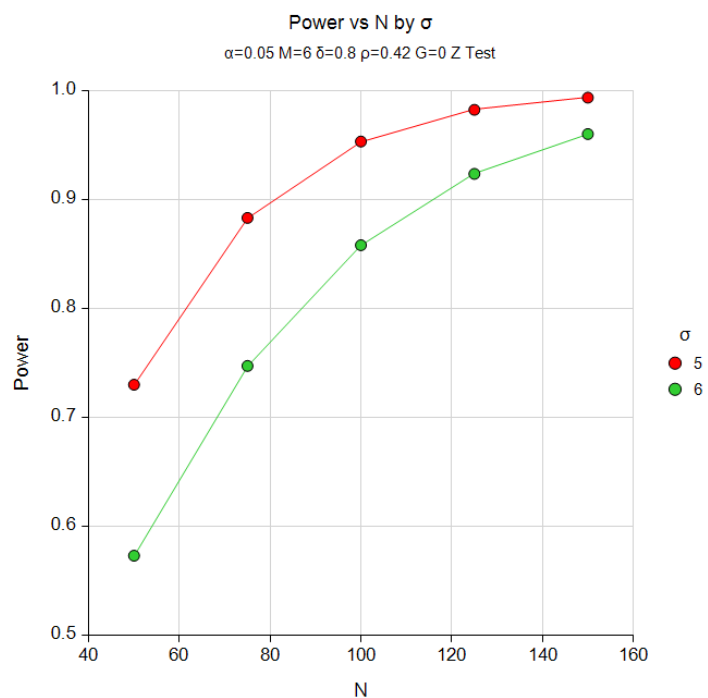
References

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.

This report gives the power for each of the scenarios.

Plots Section

Plots
_____

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, δ = 0.2; σ = 0.7071; ρʙ = 0.05, 0.1, 0.15; ρw = 0.1; M = 3; alpha = 0.05; and power = 0.8. The resulting N's are 69, 59, and 49.

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 ....**Sample Size**

Power...**0.8**

Alpha (Two-Sided Test) **0.05**

M (Sites Per Group Per Subject) **3**

δ (Mean Difference = μ1 - μ2).......................**0.2**

σ (Standard Deviation)..................................**0.7071**

Correlation Matrix Structure ....**Two Within-Subject Correlations (ρʙ, ρw)**

ρʙ (Between-Segment Correlation)...............**0.05 0.1 0.15**

ρw (Within-Segment Correlation)..................**0.1**

Output

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

Numeric Results

Solve For: Sample Size
Groups: 1 = Treatment, 2 = Control
Difference: δ = μ1 - μ2

| | | | | | | Correlation | | |
| | Total Sample Size | Sites per Group per Subject | Mean Difference | Standard Deviation | Effect Size | Between-Segment | Within-Segment | |
Power	N	M	δ	σ	δ / σ	ρв	ρw	Alpha
0.8018	69	3	0.2	0.71	0.283	0.05	0.1	0.05
0.8009	59	3	0.2	0.71	0.283	0.10	0.1	0.05
0.8074	50	3	0.2	0.71	0.283	0.15	0.1	0.05

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.