

## Chapter 389

# GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

---

## Introduction

This module calculates the power for testing for differences among the time-averaged responses (TAD) of two or more groups from correlated **binary** data that are analyzed using the GEE method. Such data can occur in two design types: clustered and longitudinal. This procedure emphasizes longitudinal designs.

GEE is different from mixed models in that it does not require the full specification of the joint distribution of the repeated measurements, as long as the marginal mean model is correctly specified. Estimation consistency is achieved even if the correlation matrix is incorrect. Also, the correlation matrix of the responses is specified directly, rather than using an intermediate, random effects model as is the case in MM. For clustered designs, GEE often uses a *compound symmetric* (CS) correlation structure. For longitudinal data, an *autoregressive* (AR(1)) correlation structure is often used.

Time-averaged response analysis is often used when the outcome to be measured varies with time. For example, suppose that you want to compare three treatment groups based on the means of a certain outcome such as whether the pressure is over 140. It is known that a person's blood pressure depends on several instantaneous factors such as amount of sleep, excitement level, mood, exercise, etc. If only a single measurement is taken from each patient, then the comparison of group values may be insensitive because of the large degree of variation in blood pressure levels within a patient. The precision of the experiment is increased by taking multiple measurements from each individual and comparing the time-averaged responses among the groups. Care must be taken in the analysis because of the correlation that is introduced when several measurements are taken from the same individual. The correlation structure may take on several forms depending on the nature of the experiment and the subjects involved.

---

## Missing Values

This procedure allows you to specify various patterns of incomplete (or missing) data. Subjects may miss some appointments but attend others. This phenomenon of incomplete data can be accounted for in the sample size calculation which can greatly reduce the overall sample size from that calculated by just omitting subjects with incomplete observations.

## Technical Details

### Theory and Notation

Technical details are given in Wang, Zhang, and Ahn (2018).

Suppose we have  $n_k$  ( $k = 1, \dots, G$ ) subjects in each of  $G$  groups for a total of  $N$  subjects, each measured on  $M$  occasions at times  $t_j$  ( $j = 1, \dots, M$ ). For convenience, we normalize these time points to the proportion of total time so that  $t_1 = 0$  and  $t_M = 1$ .

Let  $y_{kij}$  be the binary response of subject  $i$  in group  $k$  at time  $t_j$ . The response is modeled by the logistic model

$$Y_{kij} \sim \text{Bernoulli}(P_k)$$

The mean of  $Y_{kij}$  is modeled by

$$\text{logit}(P_k) = \log\left(\frac{P_k}{1 - P_k}\right) = \beta_k$$

This implies that

$$E(Y_{kij}) = P_k = \frac{e^{\beta_k}}{1 + e^{\beta_k}}$$

The GEE estimator of  $\beta_k$  is  $b_k$ , given by

$$b_k = \log\left(\frac{\sum_{i=1}^{n_k} \sum_{j=1}^M Y_{kij} / (n_k M)}{1 - \sum_{i=1}^{n_k} \sum_{j=1}^M Y_{kij} / (n_k M)}\right)$$

In this procedure, the primary interest is to test that the contrast based on the coefficients  $C = c_1, \dots, c_G$  is zero, that is, that  $H_0: \sum_{k=1}^G \beta_k c_k = 0$  against the alternative that it is non-zero.

GEE is used to estimate the  $\beta_i$ 's and test this hypothesis. The test statistic is

$$Z = \frac{C'b}{\sqrt{\text{Var}(C'b)}}$$

$H_0$  is rejected with a type I error  $\alpha$  if  $|Z| > z_{1-\alpha/2}$  where  $z_{1-\alpha/2}$  is the  $100(1 - \alpha/2)$ th percentile of a standard normal distribution.

## Correlation Patterns

In a longitudinal design with  $N$  subjects, each measured  $m$  times, observations from a single subject are correlated, and a pattern of those correlations through time needs to be specified. Several choices are available.

### Compound Symmetry

A compound symmetry correlation model assumes that all correlations are equal, and all diagonal are equal to one. That is

$$[\rho_{jj'}] = \begin{bmatrix} 1 & \rho & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \rho & \cdots & \rho \\ \rho & \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \rho & \cdots & 1 \end{bmatrix}_{M \times M}$$

where  $\rho$  is the baseline correlation.

### Banded(1)

A Banded(1) (banded order 1) correlation model assumes that diagonal elements are one, correlations for observations one time period apart are equal to  $\rho$ , and correlations for measurements greater than one time period apart are equal to zero. That is

$$[\rho_{jj'}] = \begin{bmatrix} 1 & \rho & 0 & 0 & \cdots & 0 \\ \rho & 1 & \rho & 0 & \cdots & 0 \\ 0 & \rho & 1 & \rho & \cdots & 0 \\ 0 & 0 & \rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{M \times M}$$

where  $\rho$  is the baseline correlation.

### Banded(2)

A Banded(2) (banded order 2) correlation model assumes that diagonal elements are one, correlations for observations one time period or two periods apart are equal to  $\rho$ , and correlations for measurements greater than one time period apart are equal to zero. That is

$$[\rho_{jj'}] = \begin{bmatrix} 1 & \rho & \rho & 0 & \cdots & 0 \\ \rho & 1 & \rho & \rho & \cdots & 0 \\ \rho & \rho & 1 & \rho & \cdots & 0 \\ 0 & \rho & \rho & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}_{M \times M}$$

where  $\rho$  is the baseline correlation.

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

**AR1 (Traditional)**

This version of AR1 (autoregressive order 1) correlation model assumes that correlations  $t$  time periods apart are equal to  $\rho^t$ . That is

$$[\rho_{jj'}] = \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 & \dots & \rho^{M-1} \\ \rho & 1 & \rho & \rho^2 & \dots & \rho^{M-2} \\ \rho^2 & \rho & 1 & \rho & \dots & \rho^{M-3} \\ \rho^3 & \rho^2 & \rho & 1 & \dots & \rho^{M-4} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{M-1} & \rho^{M-2} & \rho^{M-3} & \rho^{M-4} & \dots & 1 \end{bmatrix}_{M \times M}$$

where  $\rho$  is the baseline correlation.

**AR1 (Proportional)**

This version of AR1 (autoregressive order 1) correlation model is described in the book by Ahn et al. (2015). It assumes that correlations  $|t_j - t_k|$  time periods apart are equal to  $\rho^{|t_j - t_k|}$ . That is

$$[\rho_{jk}] = [\rho^{|t_j - t_k|}]_{M \times M}$$

where  $\rho$  is the baseline correlation. Note that in this pattern, the value of  $\rho$  is shown in the final column since in this case  $t_j = 0$  and  $t_k = 1$ , so  $|t_j - t_k| = 1$ .

**Damped Exponential**

A damped exponential is an extension of the AR(1) correlation model in which the exponents are raised to the power  $Dexp$  ( $\theta = Dexp$  in the diagram below). This causes the resulting correlations to be reduced (dampened). Here is an example

$$[\rho_{jj'}] = \begin{bmatrix} 1 & \rho & \rho^{2\theta} & \rho^{3\theta} & \dots & \rho^{(M-1)\theta} \\ \rho & 1 & \rho & \rho^{2\theta} & \dots & \rho^{(M-2)\theta} \\ \rho^{2\theta} & \rho & 1 & \rho & \dots & \rho^{(M-3)\theta} \\ \rho^{3\theta} & \rho^{2\theta} & \rho & 1 & \dots & \rho^{(M-4)\theta} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{(M-1)\theta} & \rho^{(M-2)\theta} & \rho^{(M-3)\theta} & \rho^{(M-4)\theta} & \dots & 1 \end{bmatrix}_{M \times M}$$

where  $\rho$  is the baseline correlation.

## Damped Exponential (Proportional)

This version of the damped exponential correlation model is described in the book by Ahn et al. (2015). It assumes that all variances on the diagonal are equal and that correlations  $|t_j - t_k|$  time periods apart are equal to  $\rho^{|t_j - t_k|^\theta}$ . That is

$$[\rho_{jk}] = \left[ \rho^{|t_j - t_k|^\theta} \right]_{M \times M}$$

where  $\rho$  is the baseline correlation. Note that in this pattern, the value of  $\rho^{|t_j - t_k|^\theta}$  turns up in the final column since in this case  $t_j = 0$  and  $t_k = 1$ , so  $|t_j - t_k| = 1$ .

## Linear Exponential Decay

A linear exponential decay correlation model is one in which the exponent of the correlation decays according to a linear equation from 1 at the *Base Time Proportion* to a final value,  $E_{max}$ . The resulting pattern looks similar to the damped exponential. Note that the exponents are applied to the absolute difference between the Measurement Time Proportions. This method allows you to easily construct comparable correlation matrices of different dimensions. Otherwise, differences in the resulting power would be more strongly due to differences in the correlation matrices.

Here is an example. Suppose  $M$  is 6,  $\rho = 0.5$ ,  $E_{max} = 3$ , the *Base Time Proportion* is 0.20, and the Measurement Time Proportions are (0, 0.2, 0.4, 0.6, 0.8, 1). The following correlation matrix would be obtained

$$[\rho_{jj'}] = \begin{bmatrix} 1 & 0.5 & 0.3536 & 0.25 & 0.1768 & 0.125 \\ 0.5 & 1 & 0.5 & 0.3536 & 0.25 & 0.1768 \\ 0.3536 & 0.5 & 1 & 0.5 & 0.3536 & 0.25 \\ 0.25 & 0.3536 & 0.5 & 1 & 0.5 & 0.3536 \\ 0.1768 & 0.25 & 0.3536 & 0.5 & 0.3536 & 0.5 \\ 0.125 & 0.1768 & 0.25 & 0.3536 & 0.5 & 1 \end{bmatrix}_{M \times M}$$

Note that in the top row, the correlation is 0.5 for the second (0.2 - 0) time point and 0.125 ( $0.5^3$ ) at the last (1 - 0) time points. The correlations are obtained by raising 0.5 to the appropriate exponent. The linear equation from 1 to 3 results in the exponents 1, 1.5, 2, 2.5, 3 correspondent to the time proportions 0, 0.2, 0.4, 0.6, 0.8, and 1.

As a further example, note that the correlation for the 0.4 time point is,  $0.5^{1.5} = 0.35355339 \approx 0.3536$ .

This method allows you to compare various values of  $M$  while keeping the correlation matrix similar. To see what we mean, consider what the correlation matrix looks like when  $M$  is reduced to 4 and the measurement time proportions are set to (0, 0.2, 0.6, 1). It becomes

$$[\rho_{jj'}] = \begin{bmatrix} 1 & 0.5 & 0.25 & 0.125 \\ 0.5 & 1 & 0.5 & 0.25 \\ 0.25 & 0.5 & 1 & 0.5 \\ 0.125 & 0.25 & 0.5 & 1 \end{bmatrix}_{M \times M}$$

Note that the correlation at a measurement time difference of 0.6 is equal to 0.25 in both matrices.

## Missing Data Patterns

The problem of missing data occurs for several reasons. In longitudinal studies in which a subject is measured multiple times, missing data becomes more complicated to model because it is possible that a subject is measured only some of the time. In fact, it is probably more common for data to be incomplete than complete. The approach of omitting subjects with incomplete data during the planning phase is very inaccurate. Certainly, subjects with partial measurements are included in the analysis. This procedure provides several missing data patterns to choose from so that your sample size calculations are more realistic.

In the presentation to following, we denote the percent of subjects with a missing response at time point  $t_j$  as  $\kappa_j$ . The proportion non-missing at a particular time point is  $\phi_j = 1 - \kappa_j$ . We will refer to  $\phi_j$  as the *marginal observant probability* at time  $t_j$  and  $\phi_{jj'}$  as a *joint observant probability* at times  $t_j$  and  $t_{j'}$ .

### Pairwise Missing Pattern

The program provides three options for how the pairwise (joint) observant probabilities  $\phi_{jj'}$  are calculated. These are

**Independent (Ind):**  $\phi_{jj'} = \phi_j \phi_{j'}, \phi_{jj} = \phi_j$

**Monotonic (Mon):**  $\phi_{jj'} = \phi_k$  where  $k = \max(j, j')$

**Mixture:**  $\phi_{jj'} = W(\text{Ind}) + (1 - W)(\text{Mon})$  for weighting factor  $W$ .

### Missing Input Type

There are several ways in which the missing value pattern can be specified. Each missing value pattern is a list of missing proportions at each of the  $M$  time points. Each value in the list must be non-negative and less than 1. Possible input choices are

- **Constant = 0**  
All missing proportions are set to 0. That is, there are no missing values.
- **Constant**  
All missing proportions are set to constant value.
- **Piecewise Constant on Spreadsheet**  
A set of missing proportions are defined for several time intervals using the spreadsheet. One column contains the missing proportions for the interval, going down the rows. Another column defines the corresponding upper limit of time proportion of the interval. The lower limit is implied by the limit given immediately above. The program assumes that the first-time interval starts at 0 percent.
- **Linear (Steady Change)**  
The missing proportions fall along a straight-line between 0 and 1 elapsed time. Only the first and last proportions are entered.
- **Piecewise Linear on Spreadsheet**  
The missing proportions fall along a set of connected straight-lines that are defined by two columns on the spreadsheet.

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

- **List**

Enter a list of M missing proportions, one for each time point.

- **Multiple Lists on Spreadsheet**

Select multiple columns containing vertical lists of missing proportions. Each column contains a set of missing proportions in rows, one for each time point.

- **Pairwise Observed Proportions on Spreadsheet**

Enter an  $M \times M$  matrix of observant probabilities by selecting M columns. These observant probabilities are the proportion of the responses for both the row and column time points that are observed.

---

## Sample Size Calculations

The details of the calculation of sample size and power is given in Wang, Zhang, and Ahn (2018). The formula for the sample size is

$$N = \frac{C'VC \left( z_{1-\frac{\alpha}{2}} + z_{1-\gamma} \right)^2}{\left( C'\underline{\beta} \right)^2}$$

where

$\gamma$             1 – power

$\alpha$             significance level

$z_{1-\alpha/2}$     is the  $100(1 - \alpha/2)^{\text{th}}$  percentile of a standard normal distribution.

$V$             is a diagonal matrix of elements  $\left\{ h \frac{(1+e^{\beta_k})^2}{r_k e^{\beta_k}} \right\}$

$h$             is  $\frac{\sum_{j=1}^M \sum_{j'=1}^M \phi_{jj'} \rho_{jj'}}{(\sum_{k=1}^G \phi_k)^2}$

$r_k$             is the proportion of subjects in group  $k$ .

The above formula is easily rearranged to obtain a formula for power.

## Example 1 – Determining Sample Size

Researchers are planning a study comparing three heart-rate medications: a standard drug and two experimental drugs. The experimental drugs are expected to have about the same impact on heart rate. Each subject will receive four applications of only one drug, two days apart. The response that will be analyzed is whether the heart rate is above 60 bpm. With the standard drug, the percentage is 40%. The researchers want a sample size large enough to detect a difference of 20 percentage points in the time-averaged response. They will use the response percentages of 40, 20, and 20 to represent this difference. The contrast coefficients that they will use are -2, 1, 1.

Similar studies have shown an autocorrelation between adjacent measurements on the same individual of 0.7, so the researchers want to try values of 0.6, 0.7, and 0.8. The researchers assume that first-order autocorrelation adequately represents the autocorrelation pattern. The test will be conducted at the 0.05 significance level and powered at 90%. The subjects will be randomly split equally among the three groups.

The researchers anticipate that the missing pattern across time will begin at 0% missing and increase steadily to 20% at the fourth measurement. They assume that the pairwise missing is *independent*.

What are the sample size requirements for this study?

### 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 1** 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 .....	<b>Sample Size</b>
Power.....	<b>0.90</b>
Alpha.....	<b>0.05</b>
G (Number of Groups) .....	<b>3</b>
Group Allocation Input Type .....	<b>Equal (n1 = n2 = ... = nG = n)</b>
Measurement Time Input Type .....	<b>Equally Spaced Measurement Times</b>
M (Number of Measurement Times) .....	<b>4</b>
Pi's Input Type .....	<b>P1, P2, ..., PG</b>
P1, P2, ..., PG .....	<b>0.4 0.2 0.2</b>
Contrast Input Type .....	<b>List of Contrast Coefficients</b>
Contrast Coefficients.....	<b>-2 1 1</b>
Pattern of p's Across Time .....	<b>AR1 (Traditional)</b>
p (Base Correlation).....	<b>0.6 0.7 0.8</b>
Missing Input Type.....	<b>Linear (Steady Change)</b>
Pairwise Missing Pattern.....	<b>Independent (Ind)</b>
First Missing Proportion (Ind).....	<b>0</b>
Last Missing Proportion (Ind).....	<b>0.2</b>



## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Output

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

## Numeric Reports

## Numeric Results for a Time-Averaged Response Test of Binary Data using GEE

Solve For: [Sample Size](#)  
 Contrast: List of Coefficients  
 Measurement Times: Equally spaced  
 Correlation: AR1:  $\rho(j,k) = \rho^{|j-k|}$   
 Missing Pattern: Range of missing proportions  
 Observant Proportions: Assume independence  
 Number of Groups: 3

Power	Total Sample Size N	Group Allocation Proportions ri	Number of Measurement Times M	Time- Averaged Responses Pi	Contrast of Pi		Base Correlation p	First Row of Correlation Matrix	Missing Data Proportions	Measurement Times	Alpha
					Coefficients Ci	Value  Ci'Pi					
0.9037	150	ri(1)	4	Pi(1)	Con(1)	0.4	0.6	p1(1)	Ms1(1)	T(1)	0.05
0.9042	171	ri(1)	4	Pi(1)	Con(1)	0.4	0.7	p2(1)	Ms1(1)	T(1)	0.05
0.9009	192	ri(1)	4	Pi(1)	Con(1)	0.4	0.8	p3(1)	Ms1(1)	T(1)	0.05

Item	Values
ri(1)	0.333, 0.333, 0.333
Pi(1)	0.4, 0.2, 0.2
Con(1)	-2, 1, 1
p1(1)	1, 0.6, 0.36, 0.216
p2(1)	1, 0.7, 0.49, 0.343
p3(1)	1, 0.8, 0.64, 0.512
Ms1(1)	0, 0.07, 0.13, 0.2
T(1)	0, 0.33, 0.67, 1

Power	The probability of rejecting a false null hypothesis when the alternative hypothesis is true.
N	The total number of subjects in the study.
ri	The Group Allocation Proportions gives the name of the set containing group allocation proportions.
M	The number of time points at which each subject is scheduled to be measured.
Pi	The Time-Averaged Response gives the name of the set containing the time-averaged response probabilities for each group.
Ci	The Contrast Coefficient gives the name of the set containing the contrast coefficients that are combined with the group response probabilities.
Ci'Pi	The Contrast Value of Ci and Pi gives the linear combination of the contrast coefficients and the group response probabilities.
p	The base correlation between two responses on the same subject. It may be transformed based on the correlation pattern.
First Row of Correlation Matrix	Presents the top row of the correlation matrix.
Missing Data Proportions	Gives the name of the set containing the individual missing value proportions for each time period.
Measurement Times	The Measurement Times gives the name of the set containing the time proportions. The values represent the proportion of the total study time that has elapsed just before the measurement.
Alpha	The probability of rejecting a true null hypothesis.

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Summary Statements

A 3-group repeated measures design (with a binary response and with 4 measurements for each subject) will be used to test whether there is a group difference in time-averaged response probabilities based on a contrast with group contrast coefficients -2, 1, 1. The test will be performed by comparing the contrast of the GEE estimates of the logit-transformed proportions to the null value zero, with a Type I error rate ( $\alpha$ ) of 0.05. The (repeated) measurements of each subject will be made at the following 4 times, expressed as proportions of the total study time: 0, 0.33, 0.67, 1. Missing values are assumed to occur completely at random (MCAR). The missing value proportions will be combined to form the pairwise observant probabilities using an independent pairwise missing pattern. The anticipated proportions missing at each measurement time are 0, 0.07, 0.13, 0.2. The first row of the autocorrelation matrix of the responses within a subject is assumed to be 1, 0.6, 0.36, 0.216, with subsequent rows following the same pattern (AR1:  $\rho(j,k) = \rho^{|j-k|}$ ). To detect time-averaged group response probabilities of 0.4, 0.2, 0.2 with 90% power, the total number of needed subjects is 150 (divided into the 3 groups according to the proportions: 0.333, 0.333, 0.333).

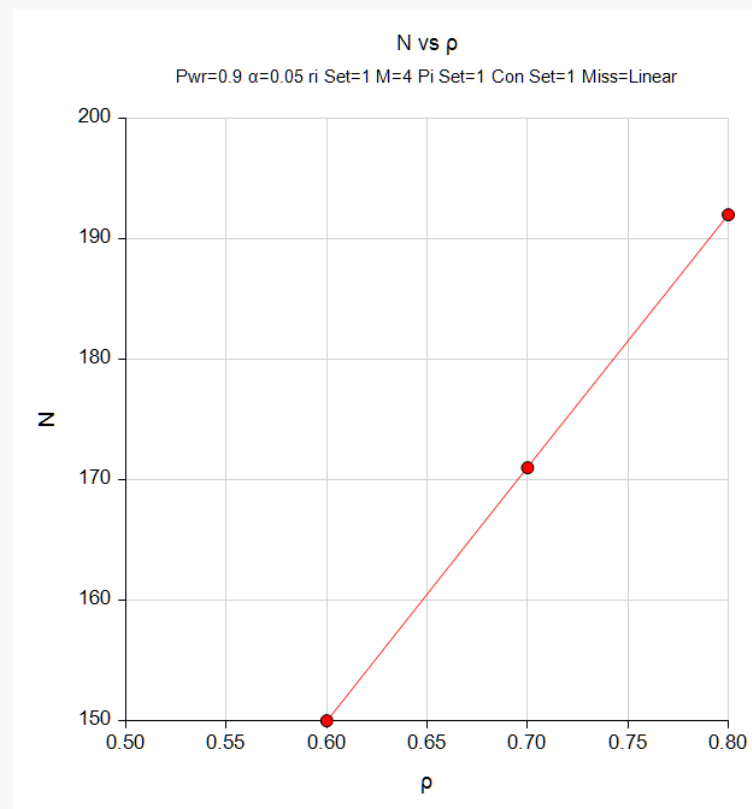
## References

Wang, J., Zhang, S., and Ahn, C. 2018. Sample Size Calculations for Comparing Time-Averaged Responses in K-group Repeated Binary Outcomes. (To appear in) Communications for Statistical Applications and Methods.

This report gives the sample size for each value of the other parameters.

## Plots Section

## Plots



This shows the relationship among the design parameters.

## Autocorrelation Matrices

**Autocorrelation Matrix for Report Row 1**

Time	T(0)	T(0.33)	T(0.67)	T(1)
T(0)	1.000	0.60	0.36	0.216
T(0.33)	0.600	1.00	0.60	0.360
T(0.67)	0.360	0.60	1.00	0.600
T(1)	0.216	0.36	0.60	1.000

**Autocorrelation Matrix for Report Row 2**

Time	T(0)	T(0.33)	T(0.67)	T(1)
T(0)	1.000	0.70	0.49	0.343
T(0.33)	0.700	1.00	0.70	0.490
T(0.67)	0.490	0.70	1.00	0.700
T(1)	0.343	0.49	0.70	1.000

**Autocorrelation Matrix for Report Row 3**

Time	T(0)	T(0.33)	T(0.67)	T(1)
T(0)	1.000	0.80	0.64	0.512
T(0.33)	0.800	1.00	0.80	0.640
T(0.67)	0.640	0.80	1.00	0.800
T(1)	0.512	0.64	0.80	1.000

These reports show the autocorrelation matrix for the indicated row of the report.

## Example 2 – Finding the Power

Continuing with Example 1, the researchers want to determine the power corresponding to group sample sizes ranging from 40 to 70 for the middle values of the other parameters.

### 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 .....	<b>Power</b>
Alpha.....	<b>0.05</b>
G (Number of Groups) .....	<b>3</b>
Group Allocation Input Type .....	<b>Equal (n1 = n2 = ... = nG = n)</b>
n (Sample Size Per Group).....	<b>40 50 60 70</b>
Measurement Time Input Type .....	<b>Equally Spaced Measurement Times</b>
M (Number of Measurement Times) .....	<b>4</b>
Pi's Input Type .....	<b>P1, P2, ..., PG</b>
P1, P2, ..., PG .....	<b>0.4 0.2 0.2</b>
Contrast Input Type .....	<b>List of Contrast Coefficients</b>
Contrast Coefficients.....	<b>-2 1 1</b>
Pattern of p's Across Time .....	<b>AR1 (Traditional)</b>
$\rho$ (Base Correlation).....	<b>0.7</b>
Missing Input Type.....	<b>Linear (Steady Change)</b>
Pairwise Missing Pattern.....	<b>Independent (Ind)</b>
First Missing Proportion (Ind).....	<b>0</b>
Last Missing Proportion (Ind).....	<b>0.2</b>

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Output

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

**Numeric Results for a Time-Averaged Response Test of Binary Data using GEE**

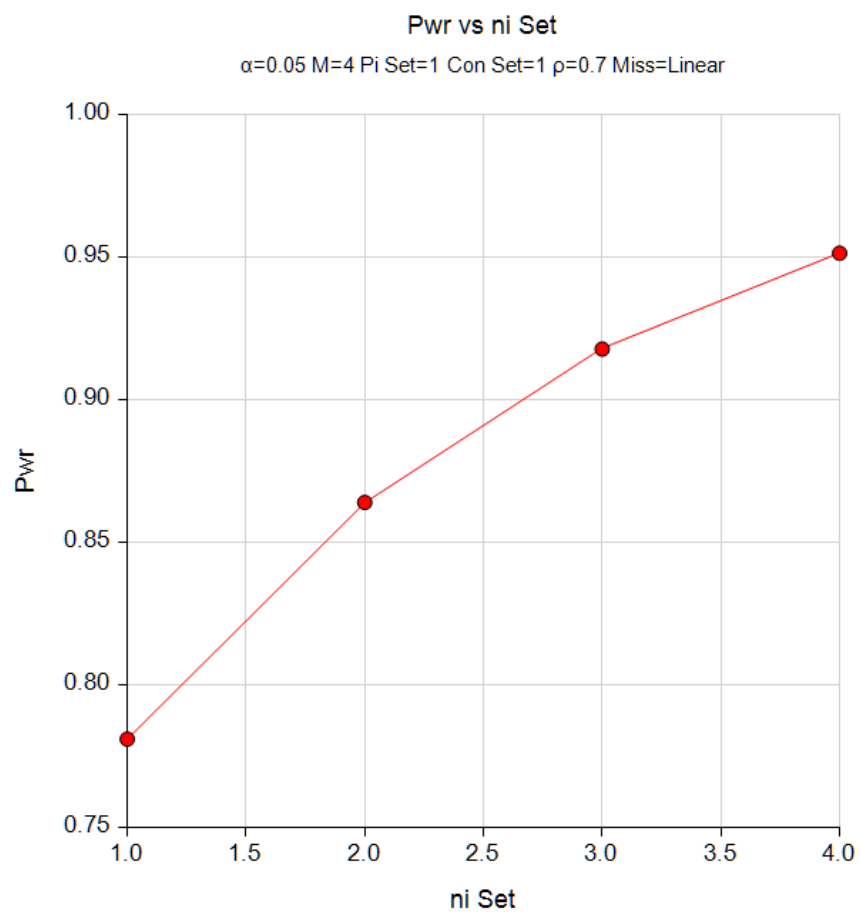
Solve For: [Power](#)  
 Contrast: List of Coefficients  
 Measurement Times: Equally spaced  
 Correlation: AR1:  $\rho(j,k) = \rho^{|j-k|}$   
 Missing Pattern: Range of missing proportions  
 Observant Proportions: Assume independence  
 Number of Groups: 3

Power	Total Sample Size	Group Sample Sizes	Number of Measurement Times	Time-Averaged Responses	Contrast of Pi		Base Correlation	First Row of Correlation Matrix	Missing Data Proportions	Measurement Times	Alpha
	N	ni	M	Pi	Coefficients	Value	$\rho$				
0.7810	120	ni(1)	4	Pi(1)	Con(1)	0.4	0.7	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.8640	150	ni(2)	4	Pi(1)	Con(1)	0.4	0.7	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9178	180	ni(3)	4	Pi(1)	Con(1)	0.4	0.7	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9514	210	ni(4)	4	Pi(1)	Con(1)	0.4	0.7	$\rho_1(1)$	Ms1(1)	T(1)	0.05

Item	Values
ni(1)	40, 40, 40
ni(2)	50, 50, 50
ni(3)	60, 60, 60
ni(4)	70, 70, 70
Pi(1)	0.4, 0.2, 0.2
Con(1)	-2, 1, 1
$\rho_1(1)$	1, 0.7, 0.49, 0.343
Ms1(1)	0, 0.07, 0.13, 0.2
T(1)	0, 0.33, 0.67, 1

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Plots



Note that the horizontal axis is scaled according to the *ni* footnote index. Hence, 1 is for *ni*(1), 2 is for *ni*(2), and so on. The same is true for the *Pi* Set and the *Con* Set in the subtitle.

## Example 3 – Comparing Various Effect Sizes

Continuing with Examples 1 and 2, the researchers want to compare the sample size of various sets of group response probabilities. To do this, they will compare the four sets of probabilities shown in the following table. The first row is the set of time-averaged response probabilities for the standard medication. The second and third rows give the anticipated response probabilities for the experimental medications.

The values in this table must be loaded into the spreadsheet.

**Table of Four Sets of Time-Averaged Response Probabilities**

C1	C2	C3	C4
0.40	0.40	0.40	0.40
0.10	0.20	0.30	0.35
0.10	0.20	0.30	0.35

### 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 3** 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..... **0.90**  
 Alpha..... **0.05**  
 G (Number of Groups) ..... **3**  
 Group Allocation Input Type ..... **Equal (n1 = n2 = ... = nG = n)**  
 Measurement Time Input Type ..... **Equally Spaced Measurement Times**  
 M (Number of Measurement Times) ..... **4**  
 Pi's Input Type ..... **Columns containing sets of Pi's**  
 Columns Containing Sets of Pi's..... **C1-C4**  
 Contrast Input Type ..... **List of Contrast Coefficients**  
 Contrast Coefficients..... **-2 1 1**  
 Pattern of p's Across Time ..... **AR1 (Traditional)**  
 $\rho$  (Base Correlation)..... **0.7**  
 Missing Input Type..... **Linear (Steady Change)**  
 Pairwise Missing Pattern..... **Independent (Ind)**  
 First Missing Proportion (Ind)..... **0**  
 Last Missing Proportion (Ind)..... **0.2**

#### Input Spreadsheet Data

Row	C1	C2	C3	C4
1	0.4	0.4	0.4	0.40
2	0.1	0.2	0.3	0.35
3	0.1	0.2	0.3	0.35

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Output

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

## Numeric Results for a Time-Averaged Response Test of Binary Data using GEE

Solve For: [Sample Size](#)  
 Contrast: List of Coefficients  
 Measurement Times: Equally spaced  
 Correlation: AR1:  $\rho(j,k) = \rho^{|j-k|}$   
 Missing Pattern: Range of missing proportions  
 Observant Proportions: Assume independence  
 Number of Groups: 3

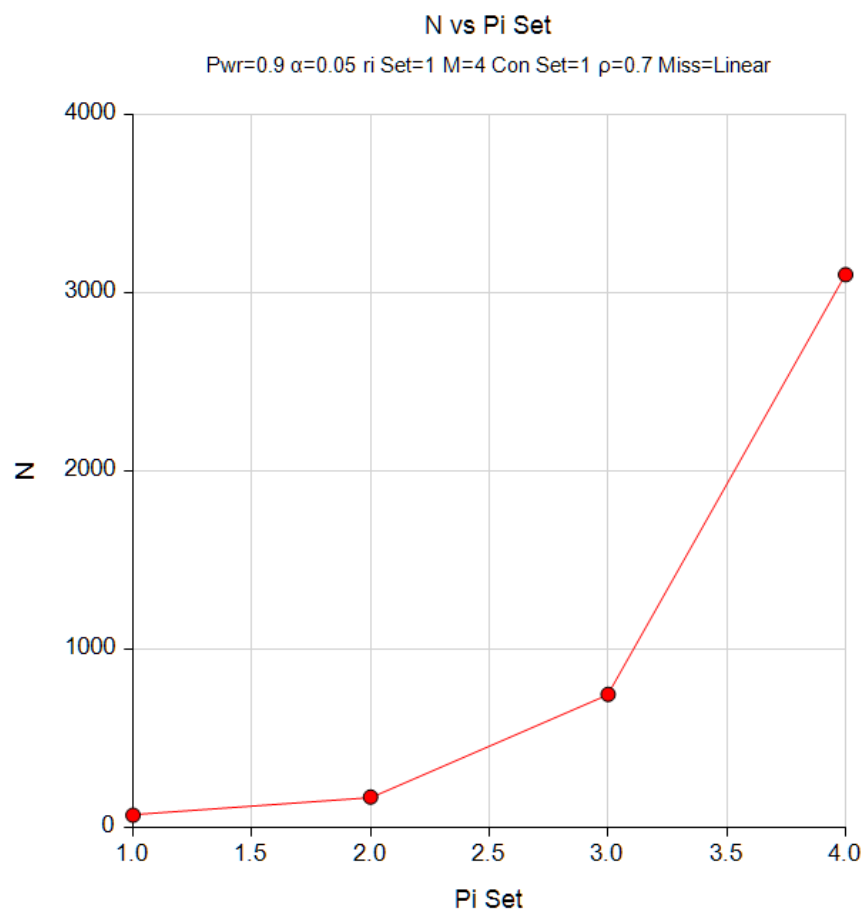
Power	Total Sample Size N	Group Allocation Proportions ri	Number of Measurement Times M	Time- Averaged Responses Pi	Contrast of Pi		Base Correlation ρ	First Row of Correlation Matrix	Missing Data Proportions	Measurement Times Alpha	
					Coefficients Ci	Value  Ci Pi					
0.9069	69	ri(1)	4	C1(1)	Con(1)	0.6	0.7	p1(1)	Ms1(1)	T(1)	0.05
0.9042	171	ri(1)	4	C2(2)	Con(1)	0.4	0.7	p1(1)	Ms1(1)	T(1)	0.05
0.9005	747	ri(1)	4	C3(3)	Con(1)	0.2	0.7	p1(1)	Ms1(1)	T(1)	0.05
0.9000	3102	ri(1)	4	C4(4)	Con(1)	0.1	0.7	p1(1)	Ms1(1)	T(1)	0.05

Item	Values
ri(1)	0.333, 0.333, 0.333
C1(1)	0.4, 0.1, 0.1
C2(2)	0.4, 0.2, 0.2
C3(3)	0.4, 0.3, 0.3
C4(4)	0.4, 0.35, 0.35
Con(1)	-2, 1, 1
$\rho_1(1)$	1, 0.7, 0.49, 0.343
Ms1(1)	0, 0.07, 0.13, 0.2
T(1)	0, 0.33, 0.67, 1



## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Plots



Note that the horizontal axis gives the index of the probability set footnote. Thus, '1.0' is column C1(1), '2.0' is column C2(2), and so on.

## Example 4 – Validation of Sample Size Calculation using Wang, Zhang, and Ahn (2018)

Wang, Zhang, and Ahn (2018) present an example in which  $G = 4$ ,  $m = 6$ ,  $P_1 = \exp(0)/(1 + \exp(0)) = 0.5$ ,  $P_2 = P_3 = P_4 = \exp(0.5)/(1 + \exp(0.5)) = 0.62245933$ , compound symmetry,  $\rho = 0.3$  and  $0.5$ , missing = {0.0, 0.05, 0.10, 0.15, 0.20, 0.25} and assumes independence,  $\alpha = 0.05$ , and power = 0.8. The contrast uses the coefficients -3, 1, 1, 1. They obtained sample sizes of 300 and 413.

### 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 4** 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 .....	<b>Sample Size</b>
Power.....	<b>0.8</b>
Alpha.....	<b>0.05</b>
G (Number of Groups) .....	<b>4</b>
Group Allocation Input Type .....	<b>Equal (n1 = n2 = ... = nG = n)</b>
Measurement Time Input Type .....	<b>Equally Spaced Measurement Times</b>
M (Number of Measurement Times) .....	<b>6</b>
Pi's Input Type .....	<b>P1, P2, ..., PG</b>
P1, P2, ..., PG .....	<b>0.5 0.62245933</b>
Contrast Input Type .....	<b>List of Contrast Coefficients</b>
Contrast Coefficients.....	<b>-3 1 1 1</b>
Pattern of $\rho$ 's Across Time .....	<b>Compound Symmetry (All <math>\rho</math>'s Equal)</b>
$\rho$ (Base Correlation).....	<b>0.3 0.5</b>
Missing Input Type.....	<b>Linear (Steady Change)</b>
Pairwise Missing Pattern.....	<b>Independent (Ind)</b>
First Missing Proportion (Ind).....	<b>0</b>
Last Missing Proportion (Ind).....	<b>0.25</b>

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Output

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

### Numeric Results for a Time-Averaged Response Test of Binary Data using GEE

Solve For: [Sample Size](#)  
 Contrast: List of Coefficients  
 Measurement Times: Equally spaced  
 Correlation: Compound symmetry (all p's equal)  
 Missing Pattern: Range of missing proportions  
 Observant Proportions: Assume independence  
 Number of Groups: 4

Power	Total Sample Size N	Group Allocation Proportions r <sub>i</sub>	Number of Measurement Times M	Time- Averaged Responses P <sub>i</sub>	Contrast of P <sub>i</sub>		Base Correlation ρ	First Row of Correlation Matrix	Missing Data Proportions	Measurement Times	Alpha
					Coefficients C <sub>i</sub>	Value  C <sub>i</sub> P <sub>i</sub>					
0.8008	300	r <sub>i</sub> (1)	6	P <sub>i</sub> (1)	Con(1)	0.37	0.3	p <sub>1</sub> (1)	Ms <sub>1</sub> (1)	T(1)	0.05
0.8032	416	r <sub>i</sub> (1)	6	P <sub>i</sub> (1)	Con(1)	0.37	0.5	p <sub>2</sub> (1)	Ms <sub>1</sub> (1)	T(1)	0.05

Item	Values
ri(1)	0.25, 0.25, 0.25, 0.25
Pi(1)	0.5, 0.62, 0.62, 0.62
Con(1)	-3, 1, 1, 1
p1(1)	1, 0.3, 0.3, 0.3, 0.3, 0.3
p2(1)	1, 0.5, 0.5, 0.5, 0.5, 0.5
Ms1(1)	0, 0.05, 0.1, 0.15, 0.2, 0.25
T(1)	0, 0.2, 0.4, 0.6, 0.8, 1

The first sample size of 300 matches the validation answer exactly, but the second value of 416 is slightly higher than 413 since **PASS's** sample size search enforces the equal group allocation requirement.

## Example 5 – Comparing Contrast Coefficients

This example will show how important the contrast coefficients are in achieving a certain power.

This example uses  $G = 4$ ,  $M = 6$ , and  $\alpha = 0.05$ . The missing input type will be set to Linear from 0 to 30% and the pairwise missing assumption is independent. Group  $P_i$ 's are 0.5, 0.5, 0.6, 0.7. The per group sample sizes range from 25 to 225. The correlation pattern will be Linear Exponential Decay with a base correlation of 0.8, Base Time Proportion of 0.1, and Emax set to 4. The Table of Contrasts shows the contrasts that will be compared. Note that C4 was found by finding the Maximum Power contrast in an initial run.

**Table of Contrasts**

C1	C2	C3	C4
-3	-3	1	-0.6
1	-1	1	-0.6
1	1	1	0.2
1	3	-3	1.0

### 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 5** 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 ..... **Power**  
 Alpha..... **0.05**  
 G (Number of Groups) ..... **4**  
 Group Allocation Input Type ..... **Equal ( $n_1 = n_2 = \dots = n_G = n$ )**  
 n (Sample Size Per Group)..... **25 75 125 175 225**  
 Measurement Time Input Type ..... **Equally Spaced Measurement Times**  
 M (Number of Measurement Times) ..... **6**  
 $P_i$ 's Input Type ..... **P1, P2, ..., PG**  
 P1, P2, ..., PG..... **0.5, 0.5, 0.6, 0.7**  
 Contrast Input Type ..... **Multiple Lists of Contrast Coefficients**  
 Multiple Lists of Coefficients ..... **C1-C4**  
 Pattern of  $\rho$ 's Across Time ..... **Linear Exponential Decay**  
 $\rho$  (Base Correlation)..... **0.8**  
 Base Time Proportion ..... **0.10**  
 Emax (Max Decay Exponent) ..... **4**  
 Missing Input Type..... **Linear (Steady Change)**  
 Pairwise Missing Pattern..... **Independent (Ind)**  
 First Missing Proportion (Ind)..... **0**  
 Last Missing Proportion (Ind)..... **0.30**

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Input Spreadsheet Data

Row	C1	C2	C3	C4
1	-3	-3	1	-0.6
2	1	-1	1	-0.6
3	1	1	1	0.2
4	1	3	-3	1.0

## Output

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

## Numeric Results for a Time-Averaged Response Test of Binary Data using GEE

Solve For: [Power](#)  
 Contrast: Multiple Lists of Coefficients: {C1-C4}  
 Measurement Times: Equally spaced  
 Correlation: Linear exponential decay, with Emax = 4 and Base Time Prop = 0.10  
 Missing Pattern: Range of missing proportions  
 Observant Proportions: Assume independence  
 Number of Groups: 4

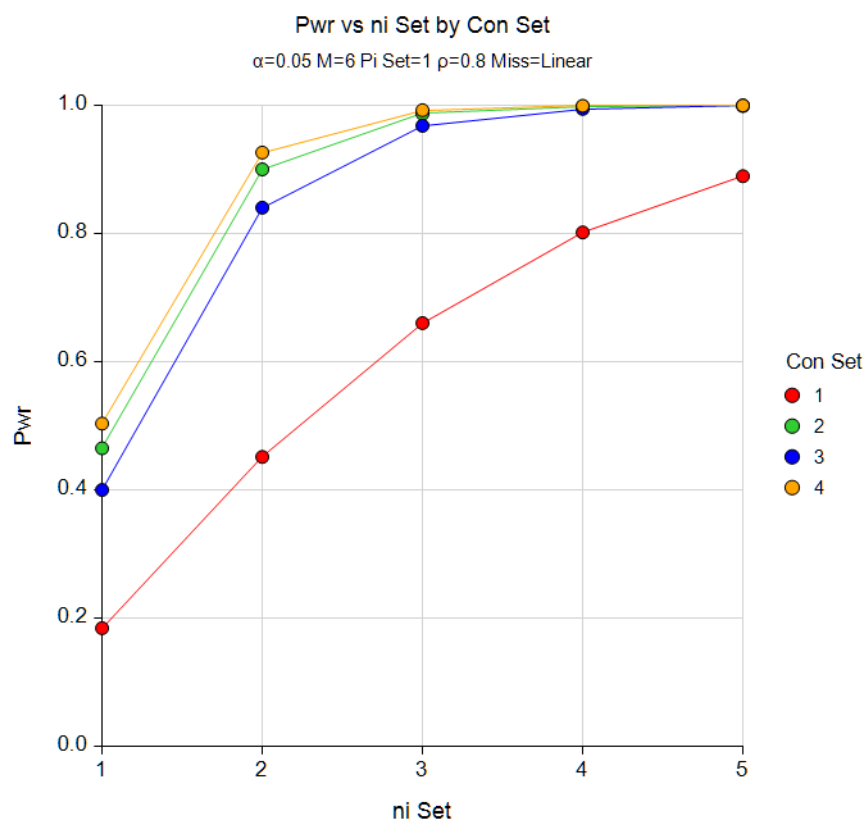
Power	Total Sample Size N	Group Sample Sizes ni	Number of Measurement Times M	Time-Averaged Responses Pi	Contrast of Pi		Base Correlation $\rho$	First Row of Correlation Matrix	Missing Data Proportions	Measurement Times	Alpha
					Coefficients Ci	Value  Ci*Pi					
0.1845	100	ni(1)	6	Pi(1)	C1(1)	0.30	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.4652	100	ni(1)	6	Pi(1)	C2(2)	0.70	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.4002	100	ni(1)	6	Pi(1)	C3(3)	0.50	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.5037	100	ni(1)	6	Pi(1)	C4(4)	0.22	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.4517	300	ni(2)	6	Pi(1)	C1(1)	0.30	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9003	300	ni(2)	6	Pi(1)	C2(2)	0.70	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.8406	300	ni(2)	6	Pi(1)	C3(3)	0.50	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9266	300	ni(2)	6	Pi(1)	C4(4)	0.22	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.6604	500	ni(3)	6	Pi(1)	C1(1)	0.30	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9870	500	ni(3)	6	Pi(1)	C2(2)	0.70	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9684	500	ni(3)	6	Pi(1)	C3(3)	0.50	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9927	500	ni(3)	6	Pi(1)	C4(4)	0.22	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.8019	700	ni(4)	6	Pi(1)	C1(1)	0.30	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9986	700	ni(4)	6	Pi(1)	C2(2)	0.70	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9947	700	ni(4)	6	Pi(1)	C3(3)	0.50	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9994	700	ni(4)	6	Pi(1)	C4(4)	0.22	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.8896	900	ni(5)	6	Pi(1)	C1(1)	0.30	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9999	900	ni(5)	6	Pi(1)	C2(2)	0.70	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
0.9992	900	ni(5)	6	Pi(1)	C3(3)	0.50	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05
1.0000	900	ni(5)	6	Pi(1)	C4(4)	0.22	0.8	$\rho_1(1)$	Ms1(1)	T(1)	0.05

## Item Values

ni(1)	25, 25, 25, 25
ni(2)	75, 75, 75, 75
ni(3)	125, 125, 125, 125
ni(4)	175, 175, 175, 175
ni(5)	225, 225, 225, 225
Pi(1)	0.5, 0.5, 0.6, 0.7
C1(1)	-3, 1, 1, 1
C2(2)	-3, -1, 1, 3
C3(3)	1, 1, 1, -3
C4(4)	-0.6, -0.6, 0.2, 1
$\rho_1(1)$	1, 0.7427, 0.64, 0.5515, 0.4753, 0.4096
Ms1(1)	0, 0.06, 0.12, 0.18, 0.24, 0.3
T(1)	0, 0.2, 0.4, 0.6, 0.8, 1

## GEE Tests for the TAD of Multiple Groups in a Repeated Measures Design (Binary Outcome)

## Plots



Note that the horizontal axis is scaled according to the *ni set* footnote index. Hence, 1 is for *ni(1)*, 2 is for *ni(2)*, and so on. The same is true for the *Pi Set* and the *Linear Miss* in the subtitle.

This example shows the huge impact that the choice of contrast coefficients can have on the power.

## Further Examples of GEE Options

The **PASS** GEE procedures offer many options that allow you to investigate various designs in detail. These are available at the end of Chapter 399, “GEE Tests for the Slope of Two Groups in a Repeated Measures Design (Continuous Outcome).” We suggest that you take time to look through those examples.