

Chapter 770

Group-Sequential Tests for Two Proportions (Simulation)

In this procedure, boundaries are calculated analytically, while simulation is used for the calculation of power (and sample size). A variety of futility boundary options are available.

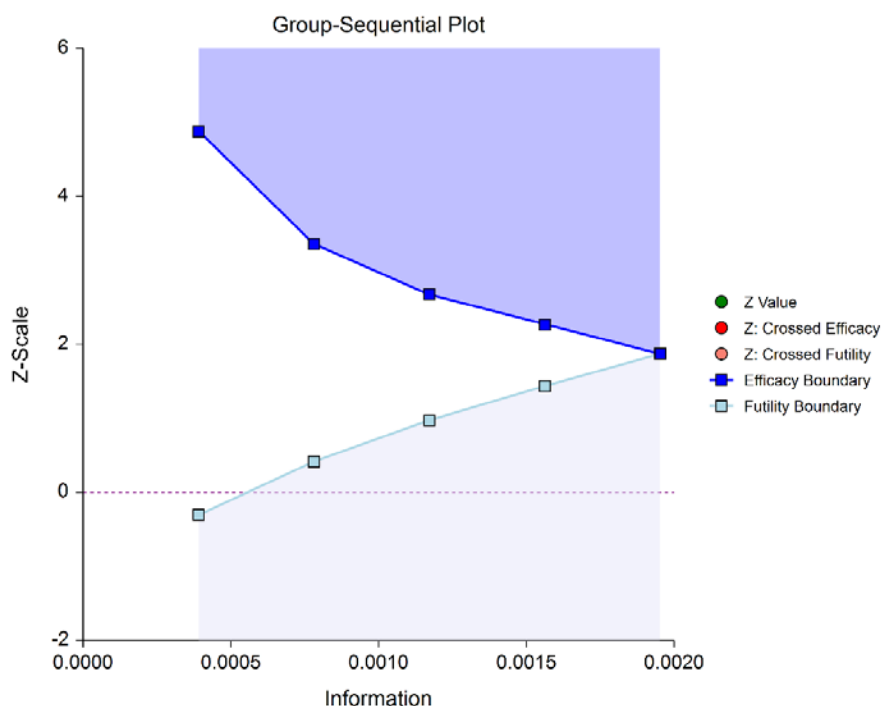
The corresponding analysis and sample size re-estimation procedure, found in NCSS Analysis and Graphics software, is Group-Sequential Analysis for Two Proportions.

Introduction

This procedure can be used to determine power, sample size and/or boundaries for group-sequential tests comparing the proportions of two groups. The available Z-tests are the common Wald Z-test using the unpooled variance estimate, with or without the continuity correction. For one- and two-sided tests, efficacy and/or futility boundaries can be generated. The spacing of the stages can be equal or custom specified. Individual stages may also be skipped. Boundaries can be computed based on popular alpha- and beta-spending functions (O'Brien-Fleming Analog, Pocock Analog, Hwang-Shih-DeCani Gamma family, linear) or custom spending functions, or boundaries may be input directly, if desired. Futility boundaries can be binding or non-binding. Corresponding P-Value boundaries are given for each boundary statistic. Alpha and/or beta spent at each stage is reported. Plots of boundaries are also produced.

This procedure is used as the planning tool for determining sample size and initial boundaries. Stage data, as it is obtained, can be evaluated using the companion procedure *Group-Sequential Analysis for Two Proportions*. The companion procedure also gives the option for sample-size re-estimation and updated boundaries for current-stage information. In that procedure, simulation can be used to evaluate boundary-crossing probabilities given the current stage results.

An example of a group-sequential boundary plot produced in this procedure is shown below.



Outline of a Group-Sequential Study

There are three basic phases of a group-sequential (interim analysis) study:

- Design
- Group-Sequential Analysis
- Reporting

Design Phase – Determine the Number of Subjects

To begin the group-sequential testing process, an initial calculation should be made to determine the sample size and target information if the final stage is reached (maximum information). The sample size calculation requires the specification of the following:

- Alpha
- Power
- Test Direction (two-sided or one-sided direction)
- Types of boundaries (efficacy, binding futility, non-binding futility)
- Maximum number of stages
- Proportion of maximum information at each stage
- Spending functions
- Assumed proportions

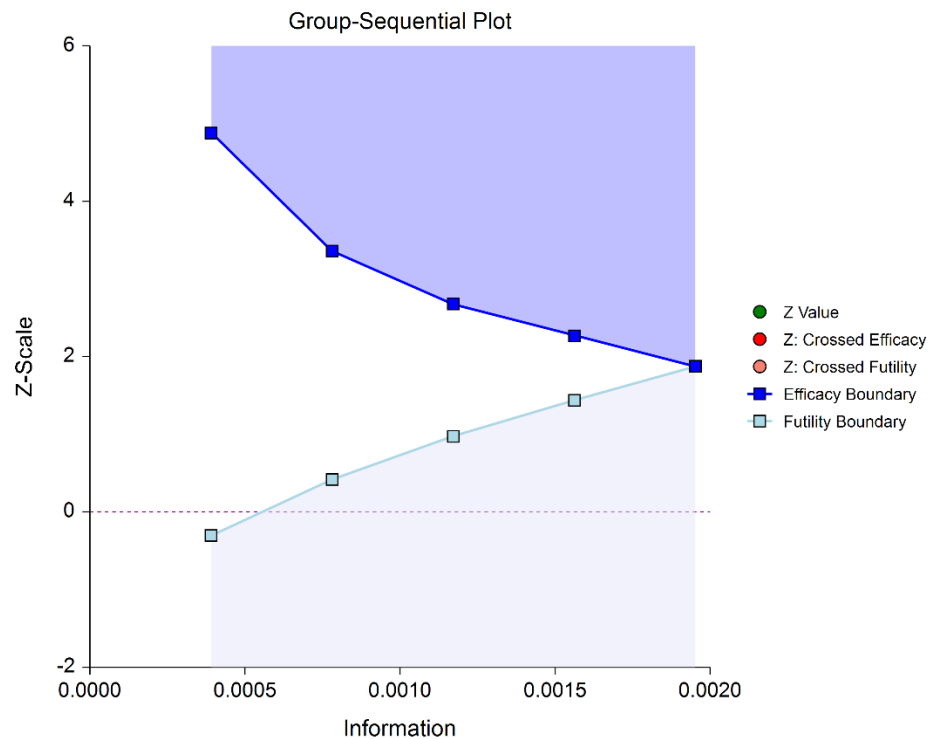
The design phase calculation is performed in this procedure. PASS software permits the user to easily try a range of proportion differences, as these values are typically not known in advance.

The resulting sample size of the sample size calculation also permits the calculation of the maximum information, which is the total information of the study if the final stage is reached (for calculation details, see the Information section later in this chapter).

Based on the maximum information, the target information and target sample size of each stage may be calculated. In particular, this permits the user to have a target sample size for the first stage.

Although it is likely to change over the course of the group-sequential analysis, a design group-sequential boundary plot can be a useful visual representation of the design:

Group-Sequential Tests for Two Proportions (Simulation)



Group-Sequential Analysis Phase

A group sequential analysis consists of a series of stages where a decision to stop or continue is made at each stage. This analysis can be performed using the companion (analysis) procedure to this sample size procedure in NCSS.

First Interim Stage

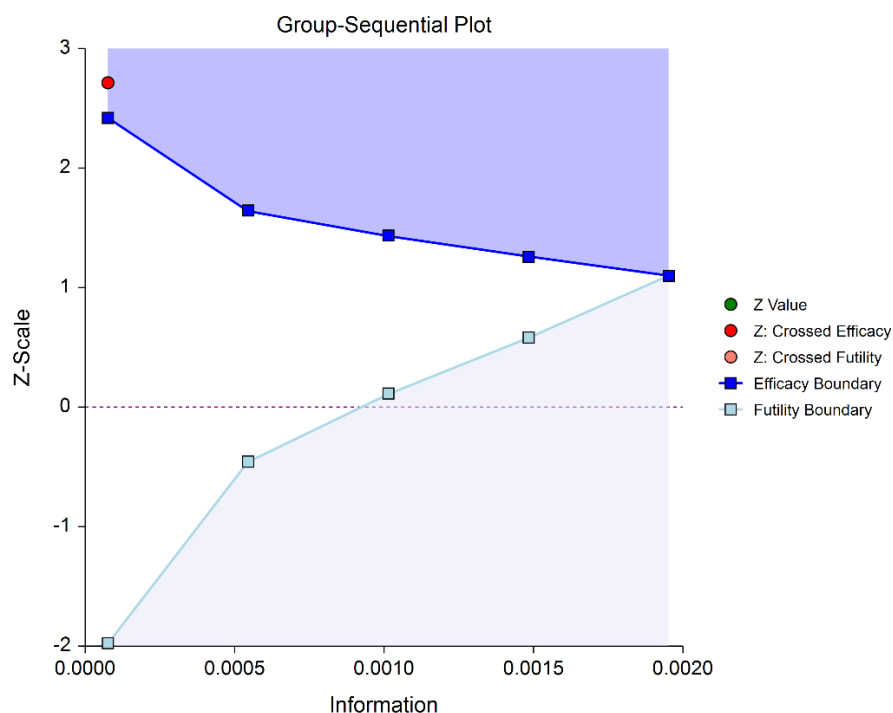
The design phase gives the target number of subjects for the first stage. The study begins, and response data is collected for subjects, moving toward the first-stage target number of subjects, until a decision to perform an analysis on the existing data is made. The analysis at this point is called the first stage.

Unless the number of subjects at the first stage matches the design target for the first stage, the calculated information at the first stage will not exactly match the design information for the first stage. Generally, the calculated information will not differ too greatly from the design information, but regardless, **spending function group-sequential analysis is well-suited to make appropriate adjustments for any differences.**

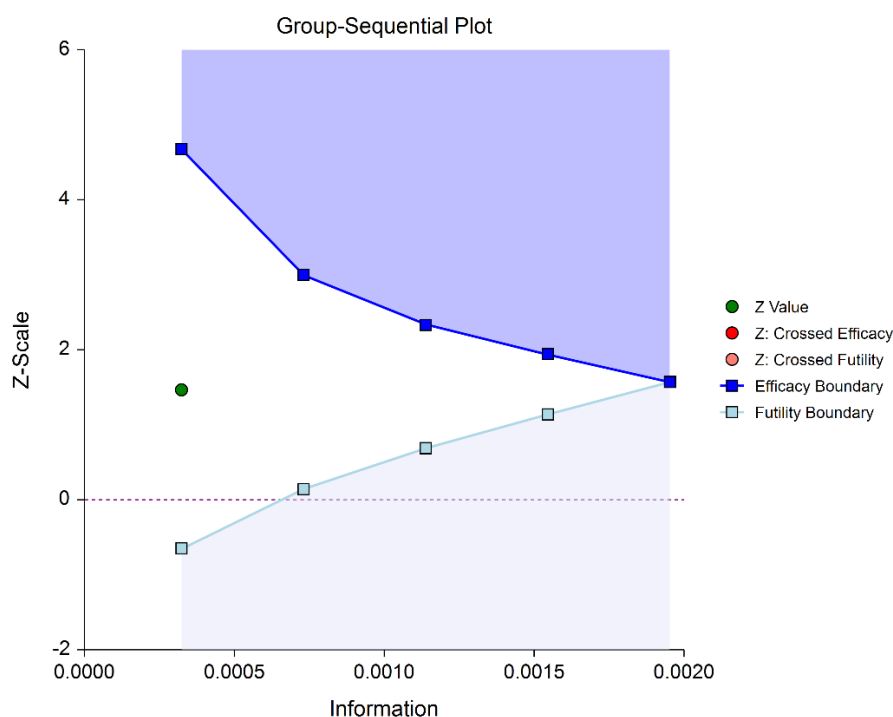
The first stage information is divided by the maximum information to obtain the stage one information proportion (or information fraction). This information proportion is used in conjunction with the spending function(s) to determine the alpha and/or beta spent at that stage. In turn, stage one boundaries, corresponding to the information proportion, are calculated.

A z-statistic is calculated from the raw proportion difference. The stage one z-statistic is compared to each of the stage one boundaries. Typically, if one of the boundaries is crossed, the study is stopped (non-binding futility boundaries may be an exception).

Group-Sequential Tests for Two Proportions (Simulation)



If none of the boundaries are crossed the study continues to the next stage.



If none of the boundaries are crossed it may also be useful to examine the conditional power or stopping probabilities of future stages, using the NCSS procedure. Conditional power and stopping probabilities are based on the user-specified supposed true difference.

Group-Sequential Tests for Two Proportions (Simulation)

Second and other interim stages (if reached)

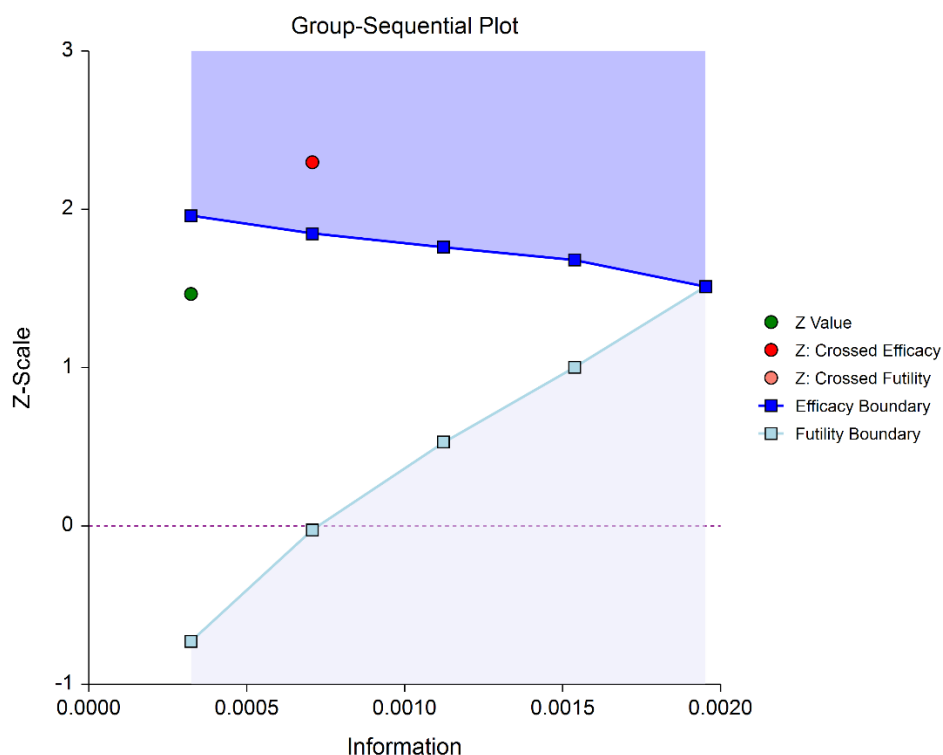
Since the first stage information proportion is not equal to the design information proportion, a designation must be made at this point as to the target information of the second stage. Two options are available in the NCSS procedure.

One option is to target the information proportion of the original design. For example, if the original design proportions of a four-stage design are 0.25, 0.50, 0.75, 1.0, and the stage one observed proportion is 0.22, the researcher might still opt to target 0.50 for the second stage, even though that now requires an additional information accumulation of 0.28 (proportion). The third and fourth stage targets would also remain 0.75 and 1.0.

A second option is to adjust the target information proportionally to the remaining proportions. For this option, if the design proportions are 0.25, 0.50, 0.75, 1.0, and 0.22 is observed, the remaining 0.78 is distributed proportionally to the remaining stages. In this example, the remaining target proportions become 0.48, 0.74, 1.0.

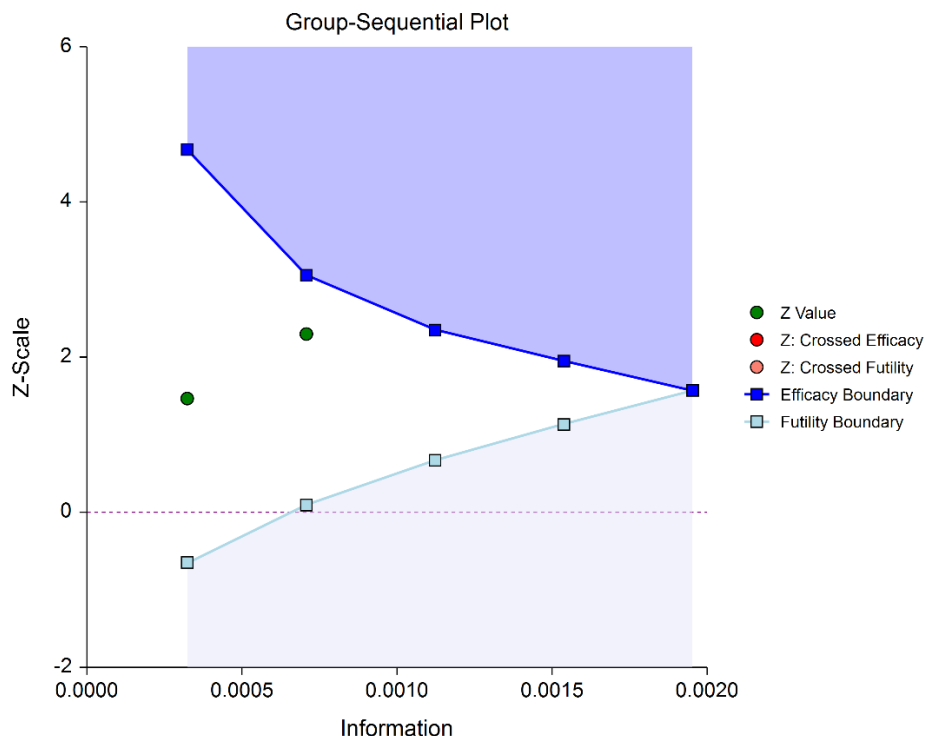
For either option, once the target information is determined for the next stage, revised target sample sizes are given (in the NCSS procedure), and the study continues until the decision is made to perform the next interim analysis on the cumulative response data. In the same manner as the first stage, the current stage information proportion is used with the spending function to determine alpha and/or beta spent at the current stage. The current stage boundaries are then computed. The z-statistic is calculated and compared to the boundaries, and a decision is made to stop or continue.

If a boundary is crossed, the study is typically stopped.



Group-Sequential Tests for Two Proportions (Simulation)

If none of the boundaries are crossed the study continues to the next stage.



Once again, if no boundary is crossed, conditional power and stopping probabilities may be considered based on a choice of a supposed true difference.

The study continues from stage to stage until the study is stopped for the crossing of a boundary, or until the final stage is reached.

Final Stage (if reached)

The final stage (if reached) is similar to all the interim stages, with a couple of exceptions. For all interim analyses the decision is made whether to stop for the crossing of a boundary, or to continue to the next stage. At the final stage, only the decision of efficacy or futility can be made.

Another intricacy of the final stage that does not apply to the interim stages is the calculation of the maximum information. At the final stage, the current information must become the maximum information, since the spending functions require that the proportion of information at the final look must be 1.0. If the current information at the final stage is less than the design maximum information, the scenario is sometimes described as *under-running*. Similarly, if the current information at the final stage is greater than the design maximum information, the result may be termed *over-running*.

For both under-running and over-running, the mechanism for adjustment is the same, and is described in the Technical Details section, under Information and Total Information.

Aside from these two exceptions, the final stage analysis is made in the same way that interim analyses were made. The remaining alpha and beta to be spent are used to calculate the final stage boundaries. If the test is a one-sided test, then the final stage boundary is a single value. The final stage z -statistic is computed from the sample proportions of the complete data from each group. The z -statistic is compared to the boundary and a decision of efficacy or futility is made.

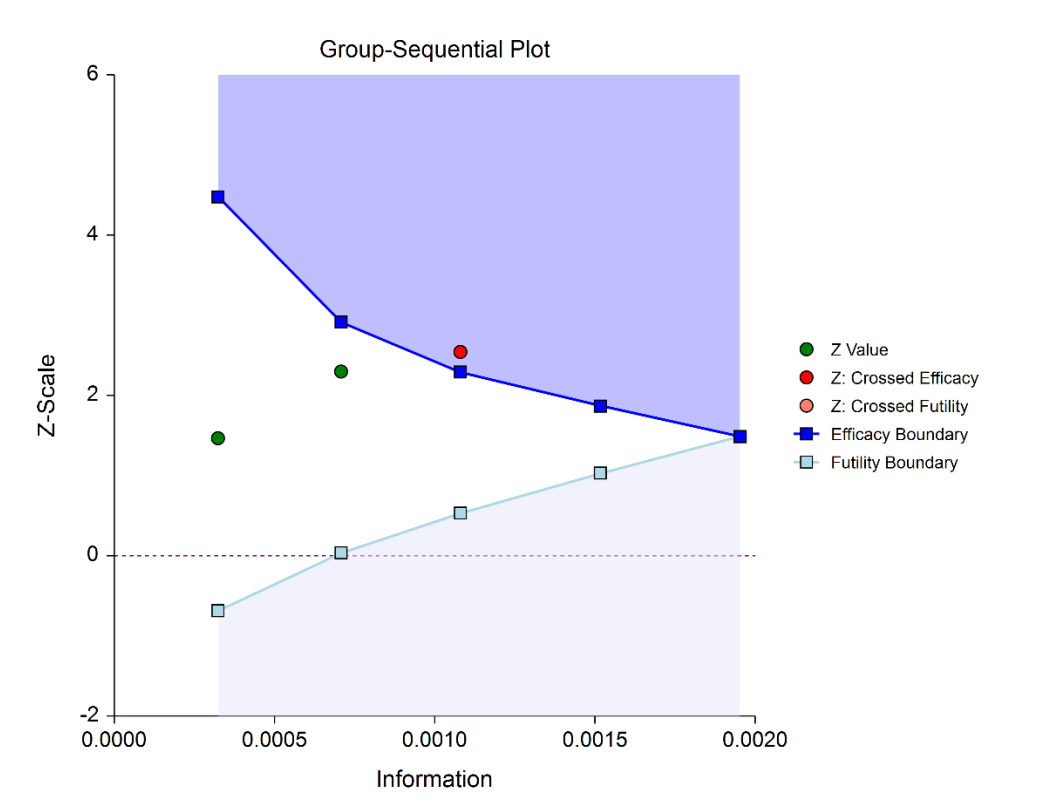
Reporting Phase

Once a group-sequential boundary is crossed and the decision is made to stop, there remains the need to properly summarize and communicate the study results. Some or all of the following may be reported:

- Boundary plot showing the crossed boundary
- Adjusted confidence interval and estimate of the proportion difference
- Sample size used

Boundary plot showing the crossed boundary

The boundary plot gives an appropriate visual summary of the process leading to the reported decision of the study.



Adjusted confidence interval and estimate of the proportion difference

Due to the bias that is introduced in the group-sequential analysis process, the raw data confidence interval of the difference in proportions should not be used. An adjusted confidence interval should be used instead.

Sample size used

The sample size at the point the study was stopped should be reported in addition to the sample size that would have been used had the final stage been reached.

Technical Details

Many articles and texts have been written about group sequential analysis. Details of many of the relevant topics are discussed below, but this is not intended to be a comprehensive review of group-sequential methods. One of the more influential works in the area of group-sequential analysis is Jennison and Turnbull (2000).

Null and Alternative Hypotheses

For comparing two proportions, the basic null hypothesis is that the proportions are equal,

$$H_0: P_1 = P_2$$

with three common alternative hypotheses,

$$H_a: P_1 \neq P_2 ,$$

$$H_a: P_1 < P_2 , \text{ or}$$

$$H_a: P_1 > P_2 ,$$

one of which is chosen according to the nature of the experiment or study.

These hypotheses may be specified equivalently as

$$H_0: P_1 - P_2 = 0$$

versus

$$H_a: P_1 - P_2 \neq 0$$

$$H_a: P_1 - P_2 < 0$$

$$H_a: P_1 - P_2 > 0$$

A slightly different set of null and alternative hypotheses are used if the goal of the test is to determine whether P_1 or P_2 is greater than or less than the other by a given amount.

The null hypothesis then takes on the form

$$H_0: P_1 - P_2 = \text{Hypothesized Difference}$$

and the alternative hypotheses,

$$H_a: P_1 - P_2 \neq \text{Hypothesized Difference}$$

$$H_a: P_1 - P_2 < \text{Hypothesized Difference}$$

$$H_a: P_1 - P_2 > \text{Hypothesized Difference}$$

For testing these hypotheses with a hypothesized difference, a superiority by a margin or non-inferiority test should be used instead.

Stages in Group-Sequential Testing

The potential to obtain the benefit from a group-sequential design and analysis occurs when the response data are collected over a period of weeks, months, or years rather than all at once. A typical example is the case where patients are enrolled in a study as they become available, as in many types of clinical trials.

A group-sequential testing stage is a point in the accumulation of the data where an interim analysis occurs, either by design or by necessity. At each stage, a test statistic is computed with all the accumulated data, and it is determined whether a boundary (efficacy or futility) is crossed. When an efficacy (or futility) boundary is crossed,

Group-Sequential Tests for Two Proportions (Simulation)

the study is usually concluded, and inference is made. If the final stage is reached, the group-sequential design forces a decision of efficacy or futility at this stage.

For the discussions below, a non-specific interim analysis stage is referenced as k , and the final stage is K .

Test Statistic

The z -statistic for any stage k is obtained from all the accumulated data up to and including that stage, using the unpooled variance estimate, and with or without the continuity correction.

Z Test (Unpooled)

This test statistic was first proposed by Karl Pearson in 1900. Although this test can be expressed as a Chi-Square statistic, it is expressed here as a z so that it can be used for one-sided hypothesis testing.

The formula for the test statistic is

$$z_k = \frac{\hat{p}_{1k} - \hat{p}_{2k}}{\hat{\sigma}_{D_k}}$$

with

$$\hat{\sigma}_{D_k} = \sqrt{\frac{\hat{p}_{1k}(1 - \hat{p}_{1k})}{n_{1k}} + \frac{\hat{p}_{2k}(1 - \hat{p}_{2k})}{n_{2k}}}$$

Continuity Correction

Frank Yates is credited with proposing a correction to the Pearson Chi-Square test for the lack of continuity in the binomial distribution. However, the correction was in common use when he proposed it in 1922.

The continuity corrected z -test is

$$z_k = \frac{(\hat{p}_{1k} - \hat{p}_{2k}) + \frac{F}{2} \left(\frac{1}{n_{1k}} + \frac{1}{n_{2k}} \right)}{\hat{\sigma}_{D_k}}$$

where F is -1 for upper-tailed, 1 for lower-tailed, and either -1 or 1 for two-sided hypotheses, depending on whether the numerator difference is positive or negative.

Group-Sequential Design Phase

In most group-sequential studies there is a design or planning phase prior to beginning response collection. In this phase, researchers specify the anticipated number and spacing of stages, the types of boundaries that will be used, the desired alpha and power levels, the spending functions, the anticipated proportions, and an estimate of the true difference in proportions.

Based on these input parameters, an initial set of boundaries is produced, an estimate of the total number of needed subjects is determined, and the anticipated total information at the final stage is calculated. This procedure can be used to make these planning phase sample size estimation calculations.

Information and Total Information

In the group-sequential design phase, the information at any stage k may be calculated from the specified proportions and the sample sizes, as

$$I_k = \frac{1}{\sigma_{D_k}^2}$$

where

$$\sigma_{D_k} = \sqrt{\frac{P_{1k}(1 - P_{1k})}{n_{1k}} + \frac{P_{2k}(1 - P_{2k})}{n_{2k}}}$$

The planning P_1 and P_2 are used for P_{1k} and P_{2k} , since realized values are not available before data is collected. When the analysis is carried out, the sample estimates \hat{p}_{1k} and \hat{p}_{2k} will be used in place of P_{1k} and P_{2k} . The final stage (K) or total (design) information is calculated from the specified proportions and the final sample sizes, as

$$I_K^* = \frac{1}{\sigma_{D_K}^2}$$

The proportion of the total information (or information fraction) at any stage is

$$prop_k = \frac{I_k}{I_K^*}$$

The information fractions are used in conjunction with the spending function(s) to define the alpha and/or beta to be spent at each stage.

To properly use the spending function at the final stage, it is required that $prop_K = 1$. However, if the final stage is reached, we see that

$$I_K = \frac{1}{\sigma_{D_{K_{achieved}}}^2} \neq I_K^* = \frac{1}{\sigma_{D_K}^2}$$

so that

$$prop_K = \frac{I_K}{I_K^*} \neq 1$$

$\sigma_{D_{K_{achieved}}}^2$ is based on $n_{1K_{achieved}}$ and $n_{2K_{achieved}}$.

When $I_K > I_K^*$, it is called over-running. When $I_K < I_K^*$, it is called under-running. In either case, the spending function is adjusted to accommodate the inequality, by redefining

$$I_K^* = I_K$$

See the discussion in Wassmer and Brannath (2016), pages 78-79, or Jennison and Turnbull (2000), pages 153-154, 162.

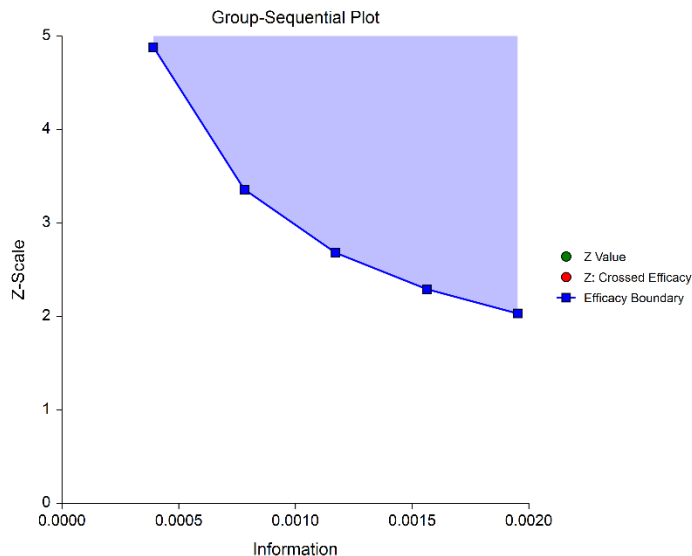
Group-Sequential Tests for Two Proportions (Simulation)

Types of Boundaries

A variety of boundary designs are available to reflect the needs of the study design.

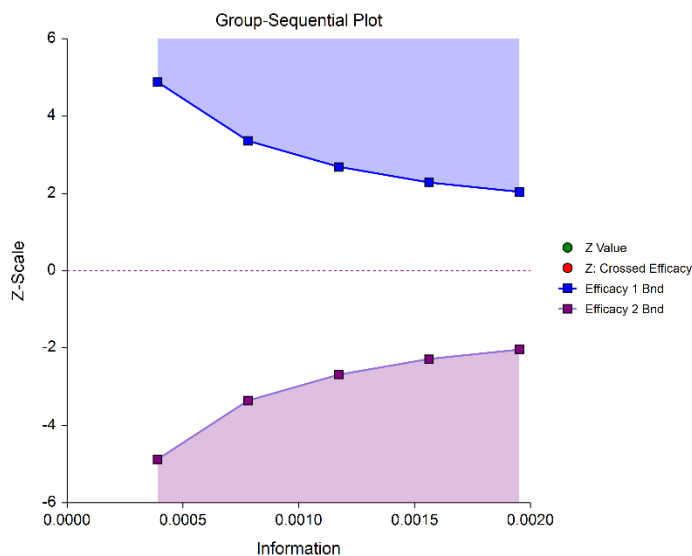
Efficacy Only (One-Sided)

The simplest group-sequential test involves a single set of stage boundaries with early stopping for efficacy.



Efficacy Only (Two-Sided, Symmetric)

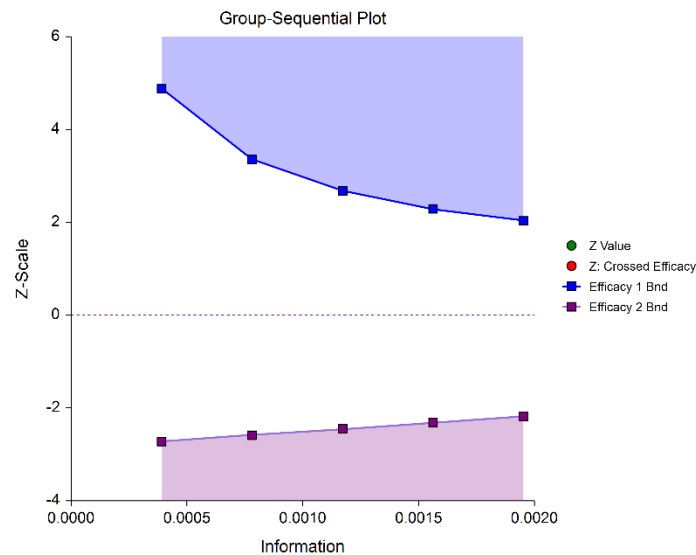
This boundary type would be used if the goal is to compare treatments, and it is not known in advance which treatment should be better.



Group-Sequential Tests for Two Proportions (Simulation)

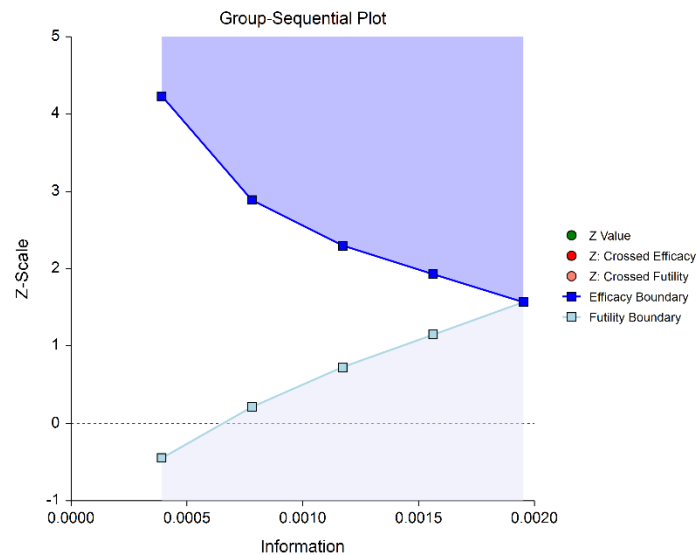
Efficacy 1 and Efficacy 2 / Harm (Two-Sided, Asymmetric)

These boundaries might be used to show efficacy on one side or harm on the other side. This design might be used in place of a one-sided efficacy and futility design if showing harm has additional benefit over stopping early for futility.



Efficacy and Binding Futility (One-Sided)

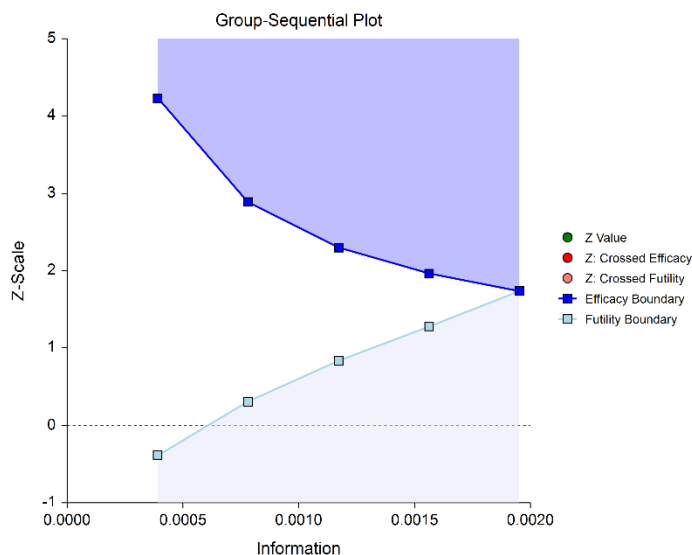
This design allows early stopping for either efficacy or futility. For binding futility designs, the Type I error protection (α) is only maintained if the study is strictly required to stop if either boundary is crossed.



Group-Sequential Tests for Two Proportions (Simulation)

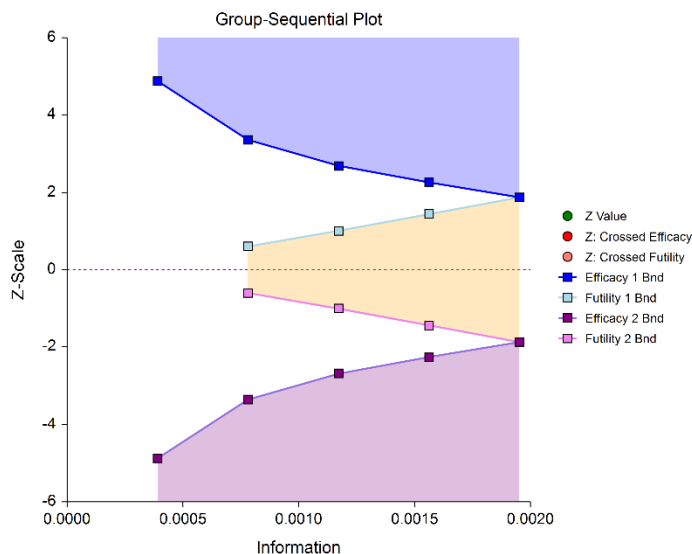
Efficacy and Non-Binding Futility (One-Sided)

This design also allows early stopping for either efficacy or futility. For non-binding futility designs, the Type I error protection (α) is maintained, regardless of whether the study continues after crossing a futility boundary. However, the effect is to make the test conservative (α is lower than the stated α and power is lower than the stated power).



Efficacy and Binding Futility (Two-Sided, Symmetric)

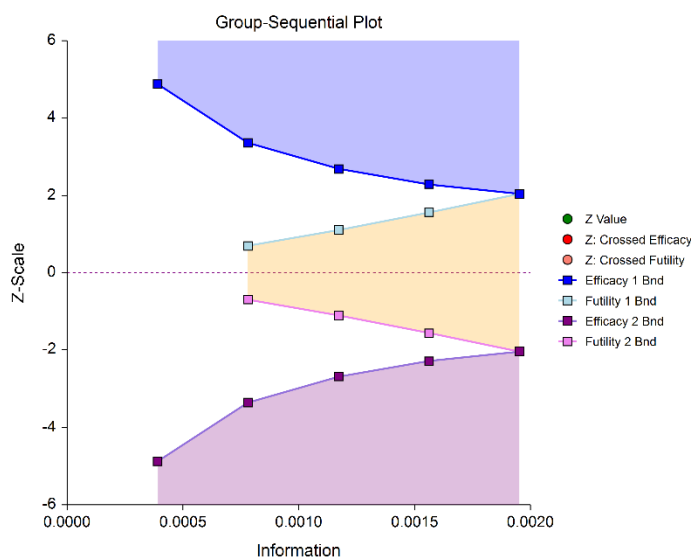
This design allows early stopping for either efficacy or futility on either side. Alpha is preserved only if crossing of futility boundaries strictly leads to early stopping for futility. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



Group-Sequential Tests for Two Proportions (Simulation)

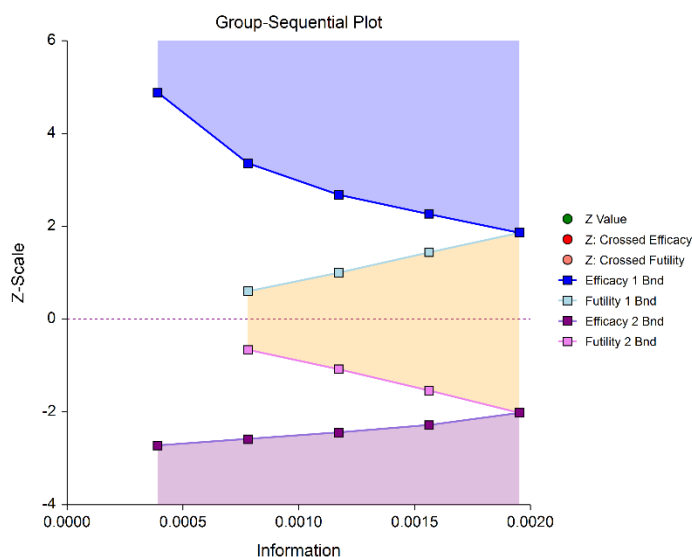
Efficacy and Non-Binding Futility (Two-Sided, Symmetric)

This design allows early stopping for either efficacy or futility on either side. Alpha is preserved even when the study is allowed to continue after crossing a futility boundary. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



Efficacy 1, Efficacy 2 / Harm, and Binding Futility (Two-Sided, Asymmetric)

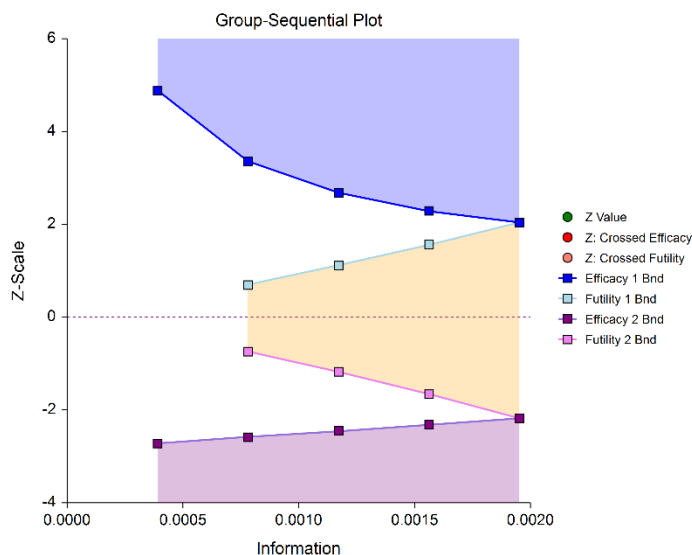
This design allows early stopping for efficacy and efficacy futility, and for harm and harm futility (or efficacy 2 and efficacy 2 futility). Binding futility boundaries require that the study is stopped when a binding futility boundary is crossed. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



Group-Sequential Tests for Two Proportions (Simulation)

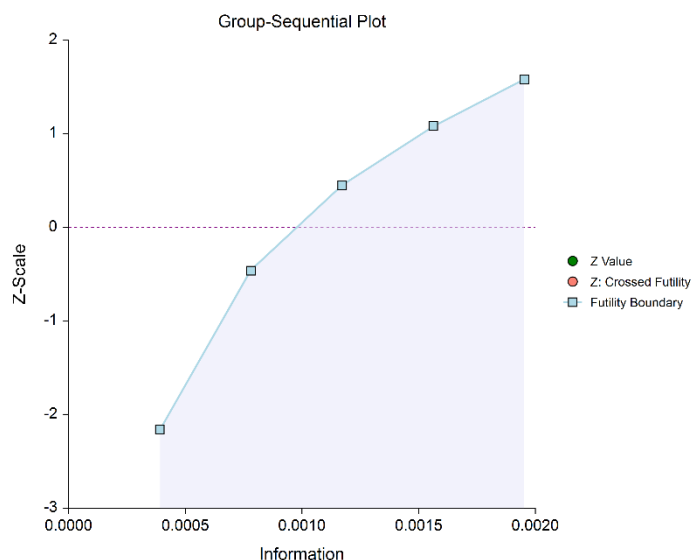
Efficacy 1, Efficacy 2 / Harm, and Non-Binding Futility (Two-Sided, Asymmetric)

This design allows early stopping for efficacy and efficacy futility, and for harm and harm futility (or efficacy 2 and efficacy 2 futility). Non-binding futility boundaries do not require that the study is stopped when a binding futility boundary is crossed, but the study design is conservative. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.



Futility Only (One-Sided)

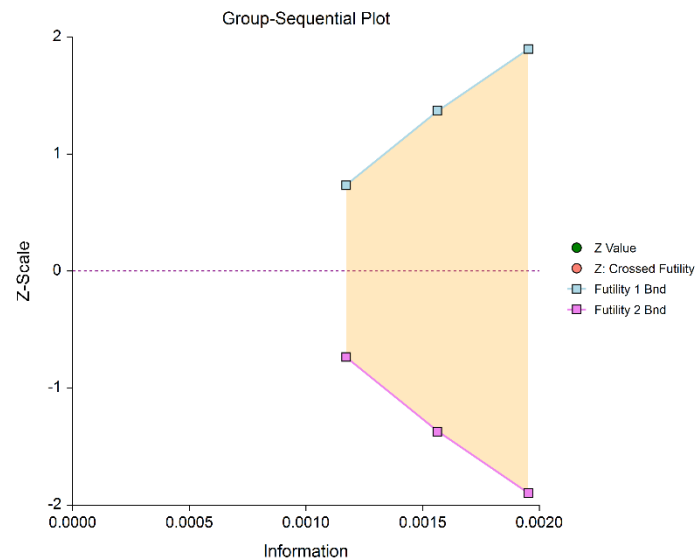
In this design, the interim analyses are used only for futility. Please be aware that, due to computational complexity, these boundaries may take several minutes to compute, particularly when some stages are skipped.



Group-Sequential Tests for Two Proportions (Simulation)

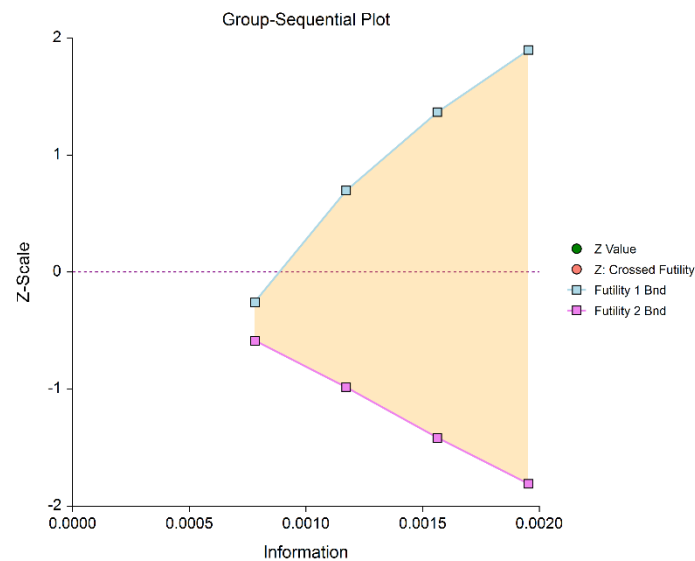
Futility Only (Two-Sided, Symmetric)

In this design, the study is stopped early only for futility. Overlapping futility boundaries may be skipped or left as they are. Please be aware that, due to computational complexity, these boundaries may take several minutes to compute, particularly when overlapping boundaries are removed or some stages are skipped.



Futility Only (Two-Sided, Asymmetric)

In this design, all stages previous to the final stage are used only for futility. Overlapping futility boundaries may be skipped or left as they are. Please be aware that, due to computational complexity, these boundaries may take several minutes to compute, particularly when overlapping boundaries are removed or some stages are skipped.



Boundary Calculations

The foundation of the spending function approach used in this procedure is given in Lan & DeMets (1983). This procedure implements the methods given in Reboussin, DeMets, Kim, & Lan (1992) to calculate the boundaries and stopping probabilities of the various group sequential designs. Some adjustments are made to these methods to facilitate the calculation of futility boundaries.

Binding vs. Non-Binding Futility Boundaries

Futility boundaries are used to facilitate the early stopping of studies when early evidence leans to lack of efficacy. When binding futility boundaries are to be used, the calculation of the futility and efficacy boundaries assumes that the study will be strictly stopped at any stage where a futility or efficacy boundary is crossed. If strict adherence is not maintained, then the Type I and Type II error probabilities associated with the boundaries are no longer valid. One (perhaps undesirable) effect of using binding futility boundaries is that the resulting final stage boundary may be lower than the boundary given in the corresponding fixed-sample design.

When non-binding futility boundaries are calculated, the efficacy boundaries are first calculated ignoring futility boundaries completely. This is done so that alpha may be maintained whether or not a study continues after crossing a futility boundary. One (perhaps undesirable) effect of using non-binding futility boundaries is that the overall group-sequential test becomes conservative (alpha is lower than the stated alpha and power is lower than the stated power).

Spending Functions

Spending functions are used to distribute portions of alpha (or beta) to the stages according to the proportion of accumulated information at each look.

Spending Function Characteristics

- Spending functions give a value of zero when the proportion of accumulated information is zero.

$$\alpha(0) = 0 \text{ (for alpha-spending)}$$

$$\beta(0) = 0 \text{ (for beta-spending)}$$

- Spending functions are increasing functions.
- Spending functions give a value of alpha (or beta) when the proportion of accumulated information is one.

$$\alpha(1) = \alpha \text{ (for alpha-spending)}$$

$$\beta(1) = \beta \text{ (for beta-spending)}$$

Using spending functions in group-sequential analyses is very flexible in that neither the information proportions nor the number of stages need be specified in advance to maintain Type I and Type II error protection.

Group-Sequential Tests for Two Proportions (Simulation)

Spending Functions Available in this Procedure

The following spending functions are shown as alpha-spending functions. The corresponding beta-spending function is given by replacing α with β .

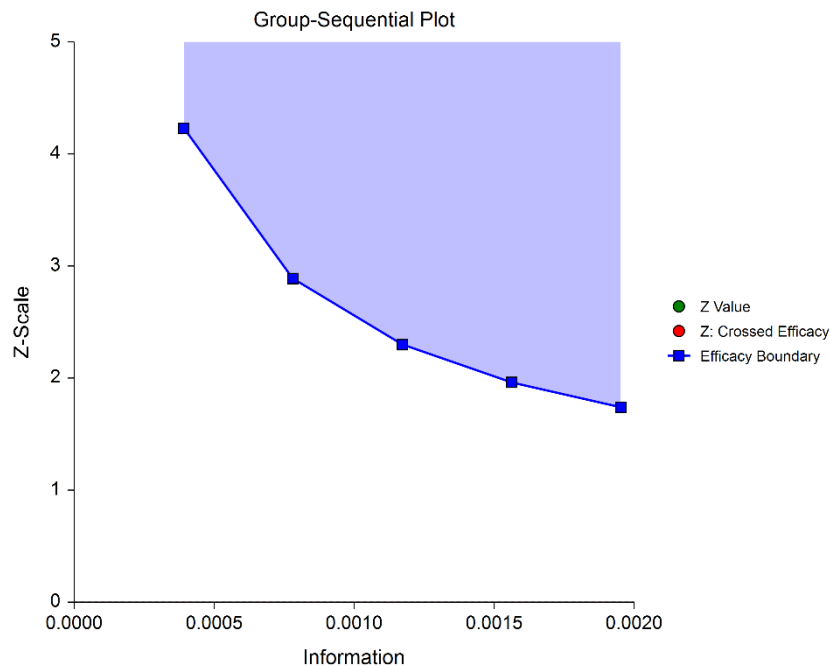
O'Brien-Fleming Analog

The O'Brien Fleming Analog (Lan & DeMets, 1983) roughly mimics the O'Brien-Fleming (non-spending function) design, with the key attribute that only a small proportion of alpha is spent early. Its popularity comes from it proportioning enough alpha to the final stage that the final stage boundary is not too different from the fixed-sample (non-group-sequential) boundary.

$$\alpha(0) = 0$$

$$\alpha(p_k) = 2 - 2\Phi\left(\frac{Z_{1-\alpha/2}}{\sqrt{p_k}}\right)$$

$$\alpha(1) = \alpha$$



Group-Sequential Tests for Two Proportions (Simulation)

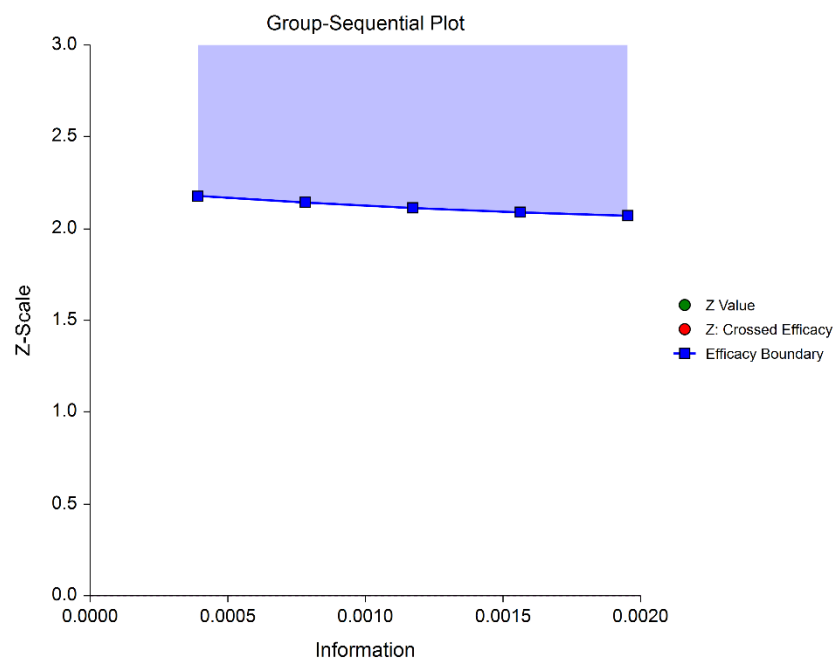
Pocock Analog

The Pocock Analog (Lan & DeMets, 1983) roughly mimics the Pocock (non-spending function) design, with the key attribute that alpha is spent roughly equally across all stages.

$$\alpha(0) = 0$$

$$\alpha(p_k) = \alpha \ln(1 + (e - 1)p_k)$$

$$\alpha(1) = \alpha$$



Group-Sequential Tests for Two Proportions (Simulation)

Power Family

The power family of spending functions has a ρ parameter that gives flexibility in the spending function shape.

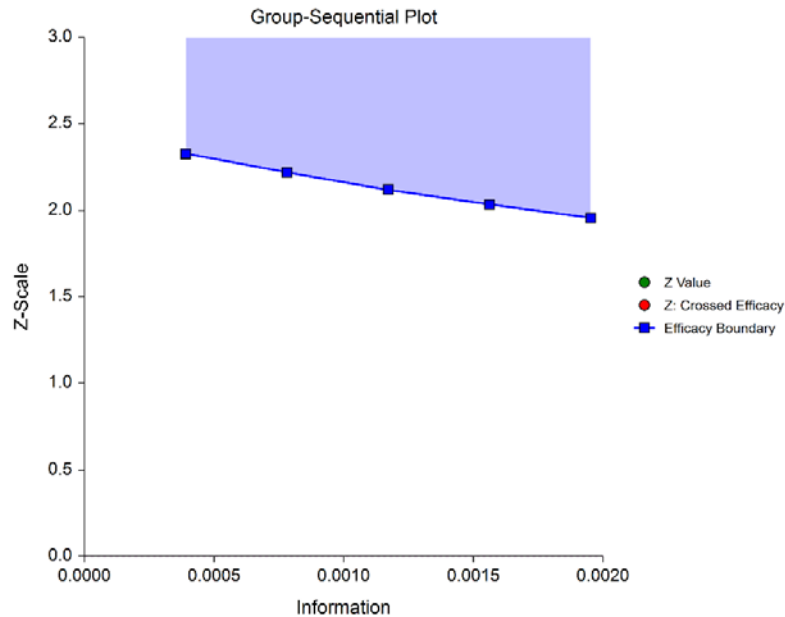
$$\alpha(0) = 0$$

$$\alpha(p_k) = p_k^\rho, \quad \rho > 0$$

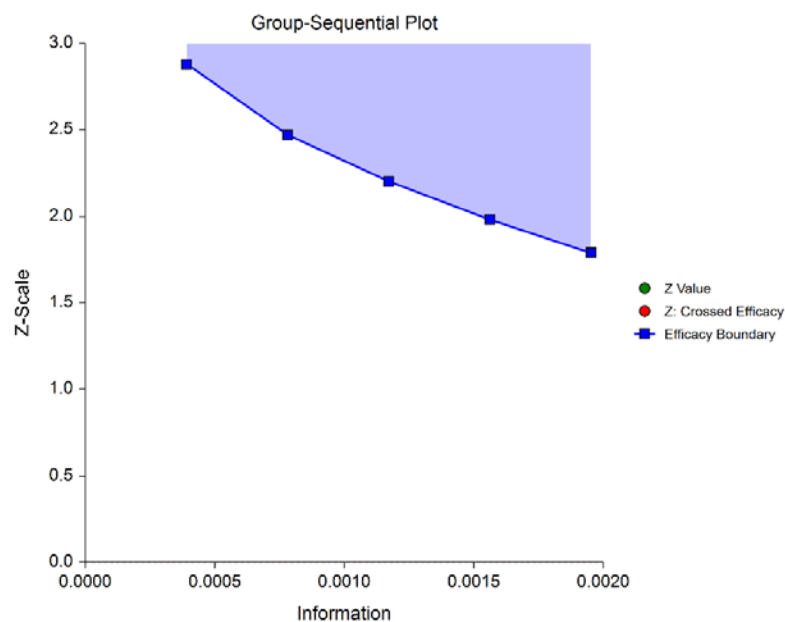
$$\alpha(1) = \alpha$$

A power family spending function with a ρ of 1 is similar to a Pocock design, while a power family spending function with a ρ of 3 is more similar to an O'Brien-Fleming design.

$$\rho = 1$$

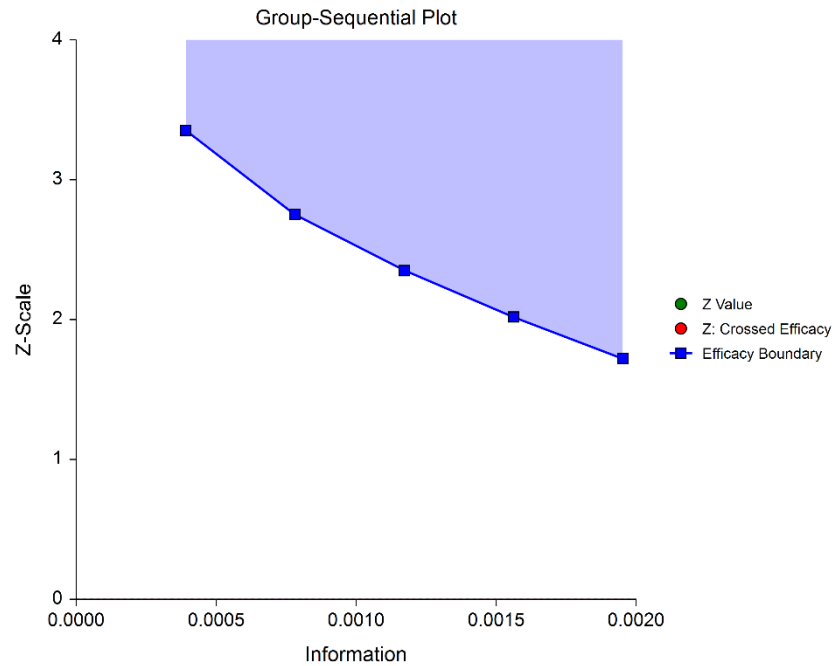


$$\rho = 2$$



Group-Sequential Tests for Two Proportions (Simulation)

$$\rho = 3$$



Group-Sequential Tests for Two Proportions (Simulation)

Hwang-Shih-DeCani (Gamma Family)

The Hwang-Shih-DeCani gamma family of spending function has a γ parameter that allows for a variety of spending functions.

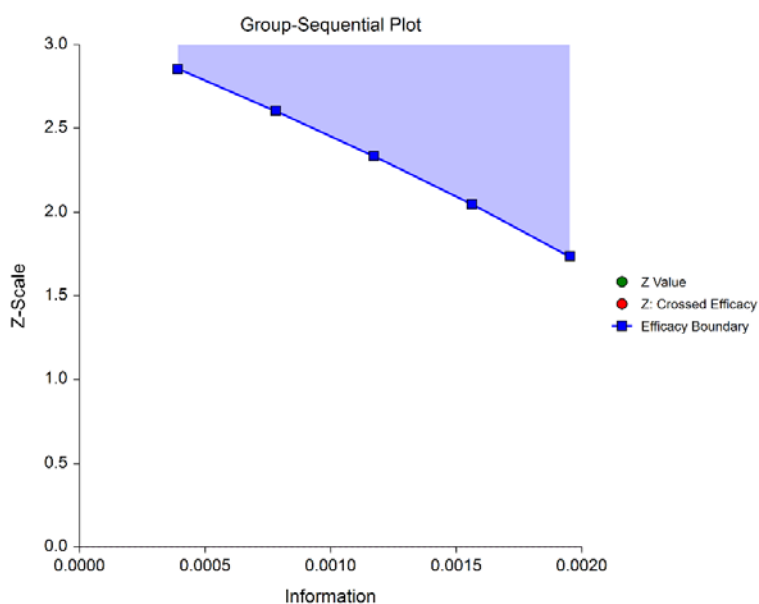
$$\alpha(0) = 0$$

$$\alpha(p_k) = \alpha \left(\frac{1 - e^{-\gamma p_k}}{1 - e^{-\gamma}} \right), \quad \gamma \neq 0$$

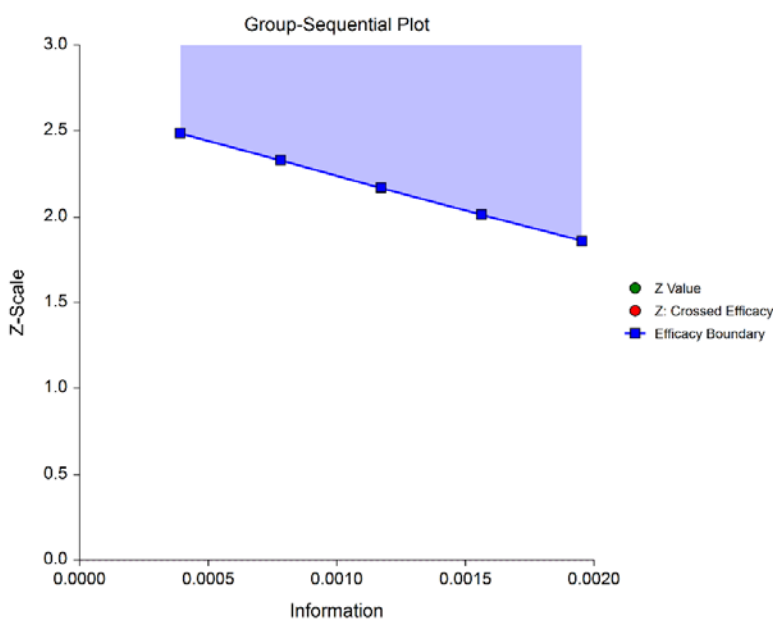
$$\alpha(p_k) = \alpha p_k, \quad \gamma = 0$$

$$\alpha(1) = \alpha$$

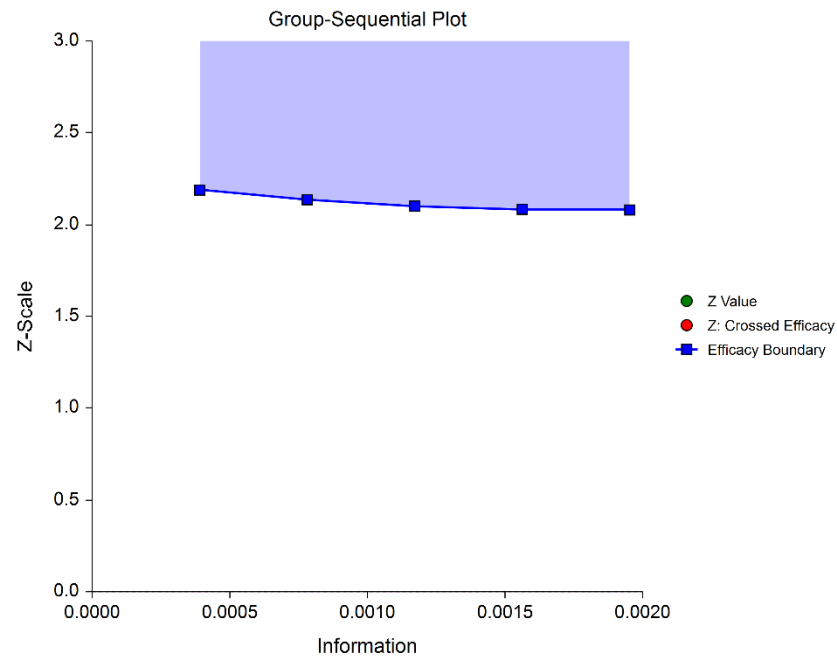
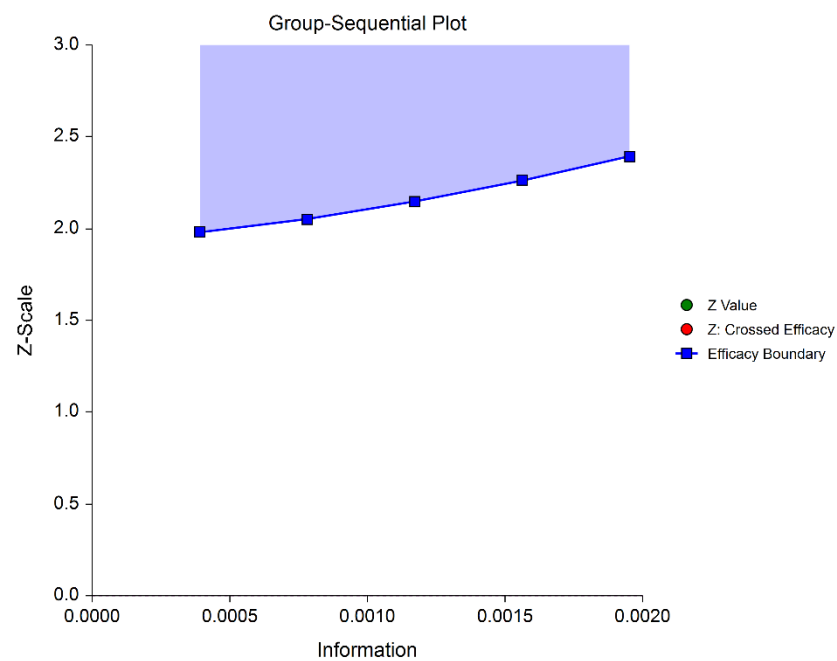
$$\gamma = -3$$



$$\gamma = -1$$



Group-Sequential Tests for Two Proportions (Simulation)

 $\gamma = 1$  $\gamma = 3$ 

Using Simulation to obtain Boundary Crossing Probabilities

In addition to providing an overall estimate of power, it can be useful to researchers to know the probability of crossing each of the group-sequential boundaries, given a specified assumed value for the proportions. The following steps are used to estimate these probabilities using simulation:

1. Determine the target (cumulative) sample sizes for each stage, including the final stage. Fractional sample sizes are rounded up to the next integer.
2. For each simulation, obtain a simulated data set with the final stage sample sizes. Simulated values are generated from Bernoulli distributions with user-specified proportions.
3. Determine whether simulation Z-values are ‘held out’ after crossing a boundary, or whether simulation Z-values are ‘left in’ (compared to boundaries at all future stages, regardless of whether a boundary was crossed at a previous stage).
 - a. If simulation Z-values are ‘held out’ after crossing a boundary, it is determined for each simulation which boundary was crossed first (except in the case of non-binding futility boundaries).
 - b. If simulation Z-values are ‘left in’ after crossing a boundary, it is determined for each simulation all the boundaries where the Z-value is across the boundary.
4. The proportion of simulations crossing each boundary provides an estimate of the probability of crossing each boundary, given the specified assumed proportions.
5. Overall power and alpha calculations are also based on the specification of ‘held out’ or ‘left in’.
 - a. When Hold Out is selected, power and alpha are calculated as the sum of all efficacy boundary proportions.
 - b. When Leave In is selected, power and alpha are calculated as the efficacy boundary proportion of the final stage.

Non-binding Futility Boundaries

When non-binding futility boundaries are used, the study may continue when a futility boundary is crossed. The simulation proportions will have a slightly different interpretation when this is the case.

Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure in general, go to the Procedures chapter.

Design Tab

Solve For

Solve For

This option specifies the parameter to be solved for from the other parameters. This parameter will be displayed on the vertical axis of any summary plots.

Group-Sequential Tests for Two Proportions (Simulation)

Power and Alpha

Power

In the case of group-sequential testing, power corresponds to the cumulative probability, across all stages, of rejecting the null hypothesis when it is false. Please note that a separate value is entered for beta (Design Beta) for the calculation of futility boundaries. Perhaps in many cases, Design Beta will be set to equal 1 minus Power. The valid range is 0 to 1. Usually, a power between 0.8 and 0.95 is desired. Different disciplines have different standards for setting power. One common choice is 0.90, but 0.80 is also popular. You can enter a single value such as *0.90* or a series of values such as *0.70 0.80 0.90* or *0.70 to 0.90 by 0.1*. When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

Alpha

In the case of group-sequential testing, alpha corresponds to the cumulative probability, across all stages, of rejecting the null hypothesis when it is true. As such, alpha is divided among the stages of the study according to the spending function. Alpha, combined with the alpha-spending function, is used in the calculation of the efficacy boundaries. Since Alpha is a probability, it is bounded by 0 and 1. Commonly, it is between 0.001 and 0.10. Alpha is often set to 0.05 for two-sided tests and to 0.025 for one-sided tests. You can enter a single value such as *0.05* or a series of values such as *0.05 0.10 0.15* or *0.05 to 0.15 by 0.01*. When a series of values is entered, PASS will generate a separate calculation result for each value of the series.

Sample Size (When Solving for Sample Size)

Group Allocation

Select the option that describes the constraints on $N1$ or $N2$ or both.

The options are

- **Equal ($N1 = N2$)**
This selection is used when you wish to have equal sample sizes in each group. Since you are solving for both sample sizes at once, no additional sample size parameters need to be entered.
- **Enter $N1$, solve for $N2$**
Select this option when you wish to fix $N1$ at some value (or values), and then solve only for $N2$. Please note that for some values of $N1$, there may not be a value of $N2$ that is large enough to obtain the desired power.
- **Enter $N2$, solve for $N1$**
Select this option when you wish to fix $N2$ at some value (or values), and then solve only for $N1$. Please note that for some values of $N2$, there may not be a value of $N1$ that is large enough to obtain the desired power.
- **Enter $R = N2/N1$, solve for $N1$ and $N2$**
For this choice, you set a value for the ratio of $N2$ to $N1$, and then PASS determines the needed $N1$ and $N2$, with this ratio, to obtain the desired power. An equivalent representation of the ratio, R , is
$$N2 = R * N1.$$
- **Enter percentage in Group 1, solve for $N1$ and $N2$**
For this choice, you set a value for the percentage of the total sample size that is in Group 1, and then PASS determines the needed $N1$ and $N2$ with this percentage to obtain the desired power.

Group-Sequential Tests for Two Proportions (Simulation)

N1 (Sample Size, Group 1)

This option is displayed if Group Allocation = "Enter N1, solve for N2"

$N1$ is the number of items or individuals sampled from the Group 1 population, if the final stage is reached.

$N1$ must be ≥ 2 . You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed if Group Allocation = "Enter N2, solve for N1"

$N2$ is the number of items or individuals sampled from the Group 2 population, if the final stage is reached.

$N2$ must be ≥ 2 . You can enter a single value or a series of values.

R (Group Sample Size Ratio)

This option is displayed only if Group Allocation = "Enter $R = N2/N1$, solve for $N1$ and $N2$."

R is the ratio of $N2$ to $N1$. That is,

$$R = N2 / N1.$$

Use this value to fix the ratio of $N2$ to $N1$ while solving for $N1$ and $N2$. Only sample size combinations with this ratio are considered.

$N2$ is related to $N1$ by the formula:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1 (e.g., $N1 = 10$ and $N2 = 20$, or $N1 = 50$ and $N2 = 100$).

R must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single or a series of values.

Percent in Group 1

This option is displayed only if Group Allocation = "Enter percentage in Group 1, solve for $N1$ and $N2$."

Use this value to fix the percentage of the total sample size allocated to Group 1 while solving for $N1$ and $N2$. Only sample size combinations with this Group 1 percentage are considered. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single or a series of values.

Sample Size (When Solving for Power)

Group Allocation

Select the option that describes how individuals in the study will be allocated to Group 1 and to Group 2.

The options are

- **Equal ($N1 = N2$)**

This selection is used when you wish to have equal sample sizes in each group. A single per group sample size will be entered.

- **Enter $N1$ and $N2$ individually**

This choice permits you to enter different values for $N1$ and $N2$.

Group-Sequential Tests for Two Proportions (Simulation)

- **Enter N1 and R, where $N2 = R * N1$**

Choose this option to specify a value (or values) for $N1$, and obtain $N2$ as a ratio (multiple) of $N1$.

- **Enter total sample size and percentage in Group 1**

Choose this option to specify a value (or values) for the total sample size (N), obtain $N1$ as a percentage of N , and then $N2$ as $N - N1$.

Sample Size Per Group

This option is displayed only if Group Allocation = "Equal ($N1 = N2$)."

The Sample Size Per Group is the number of items or individuals sampled from each of the Group 1 and Group 2 populations, if the final stage is reached. Since the sample sizes are the same in each group, this value is the value for $N1$, and also the value for $N2$.

The Sample Size Per Group must be ≥ 2 . You can enter a single value or a series of values.

N1 (Sample Size, Group 1)

*This option is displayed if Group Allocation = "Enter N1 and N2 individually" or "Enter N1 and R, where $N2 = R * N1$."*

$N1$ is the number of items or individuals sampled from the Group 1 population, if the final stage is reached.

$N1$ must be ≥ 2 . You can enter a single value or a series of values.

N2 (Sample Size, Group 2)

This option is displayed only if Group Allocation = "Enter N1 and N2 individually."

$N2$ is the number of items or individuals sampled from the Group 2 population, if the final stage is reached.

$N2$ must be ≥ 2 . You can enter a single value or a series of values.

R (Group Sample Size Ratio)

*This option is displayed only if Group Allocation = "Enter N1 and R, where $N2 = R * N1$."*

R is the ratio of $N2$ to $N1$. That is,

$$R = N2/N1$$

Use this value to obtain $N2$ as a multiple (or proportion) of $N1$.

$N2$ is calculated from $N1$ using the formula:

$$N2 = [R \times N1],$$

where the value $[Y]$ is the next integer $\geq Y$.

For example, setting $R = 2.0$ results in a Group 2 sample size that is double the sample size in Group 1.

R must be greater than 0. If $R < 1$, then $N2$ will be less than $N1$; if $R > 1$, then $N2$ will be greater than $N1$. You can enter a single value or a series of values.

Total Sample Size (N)

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This is the total sample size, or the sum of the two group sample sizes. This value, along with the percentage of the total sample size in Group 1, implicitly defines $N1$ and $N2$.

The total sample size must be greater than one, but practically, must be greater than 3, since each group sample size needs to be at least 2.

You can enter a single value or a series of values.

Group-Sequential Tests for Two Proportions (Simulation)

Percent in Group 1

This option is displayed only if Group Allocation = "Enter total sample size and percentage in Group 1."

This value fixes the percentage of the total sample size allocated to Group 1. Small variations from the specified percentage may occur due to the discrete nature of sample sizes.

The Percent in Group 1 must be greater than 0 and less than 100. You can enter a single value or a series of values.

Effect Size

P1

Enter the value for the proportion of group 1. Power calculation simulations for group 1 will be generated from a Bernoulli distribution with proportion P1. You can enter a single value such as *0.4* or a series of values such as *0.4 0.5 0.6* or *0.3 to 0.8 by 0.1*.

P2

Enter the value for the proportion of group 2. Power calculation simulations for group 2 will be generated from a Bernoulli distribution with proportion P2. You can enter a single value such as *0.4* or a series of values such as *0.4 0.5 0.6* or *0.3 to 0.8 by 0.1*.

Design

Maximum Number of Stages (K)

Enter an integer value indicating the number of stages in the planning design if the final stage is reached.

Information Proportion at each Stage

Specify whether stages are intended to reflect evenly spaced increments in information, or whether custom spacing is to be used.

- **Equally incremented**

For this selection, the increment in information at each stage is the same for all stages. For example, if the maximum number of stages is 4, the information proportion at each stage is planned to be 0.25, 0.50, 0.75, 1.00.

- **Custom (Enter cumulative information proportions)**

Select this option to enter specific cumulative information proportions for each stage. The values entered should be between 0 and 1, increasing in value, and the final value should be 1. The number of proportions entered should equal the Maximum Number of Stages.

Cumulative Information Proportions

When Information Proportion at each stage is set to Custom, this option appears to allow the user to enter specific cumulative information proportions for each stage.

The values entered should be between 0 and 1, increasing in value, and the final value should be 1. The number of proportions should equal the Maximum Number of Stages. The values may be separated by spaces or commas.

Group-Sequential Tests for Two Proportions (Simulation)

Continuity Correction and Zero Count Adjustment

Continuity Correction

Specify whether to use the continuity correction in the calculation of Z-statistics for the simulations. See the documentation section 'Test Statistic' for details.

- **Yes**

For this selection, a small correction value is added to (or subtracted from) the Z-statistic numerator.

- **No**

For this choice, no correction value is used. The Z-statistic numerator is calculated as the simple difference in sample proportions.

Zero Count Adjustment Method

In the simulations, zero cell counts can cause division by zero problems with the Z-statistic calculations. To compensate for this, a small value (called the Zero Count Adjustment Value) may be added either to all cells or to all cells with zero counts. This option specifies whether you want to use the adjustment and which type of adjustment you want to use. Adding a small value is controversial, but may be necessary. Some statisticians recommend adding 0.5 while others recommend 0.25. Smaller values may work as well.

Zero Count Adjustment Value

In the simulations, zero cell counts can cause division by zero problems with the Z-statistic calculations. To compensate for this, a small value may be added either to all cells or to all zero cells. The Zero Count Adjustment Value is the amount that is added. Some statisticians recommend adding 0.5 while others recommend 0.25. Smaller values may work as well.

Boundaries

This section is used to specify the details of the types of boundaries, the direction of the alternative hypothesis, and alpha and beta spending.

Boundaries – Boundary Structure

Boundaries Used

Use this option to determine whether boundaries are for efficacy only, futility only, or both. Also use this option to indicate whether the boundaries are one-sided or two-sided. For two-sided scenarios, you may specify whether the two sides are symmetric or specified separately (asymmetric). To see example images of each of the boundary types, see the Type of Boundaries section of the documentation.

- **One-Sided Efficacy Only**

The simplest group-sequential test involves a single set of stage boundaries with early stopping for efficacy (Rejecting H_0).

- **Two-Sided Efficacy Only (Symmetric)**

This selection gives two, symmetric sets of efficacy boundaries. This boundary type might be used if the goal is to compare treatments, and it is not known in advance which treatment should be better.

- **Two-Sided Efficacy Only (Asymmetric)**

These boundaries might be used to show efficacy on one side or harm on the other side. This design might be used in place of a one-sided efficacy and futility design if showing harm has additional benefit over stopping early for futility.

Group-Sequential Tests for Two Proportions (Simulation)

- **One-Sided Efficacy with Futility**

This design allows early stopping for either efficacy (Reject H_0) or futility (Accept H_0). The final stage boundary is equal for the two sets of boundaries.

- **Two-Sided Efficacy with Futility (Symmetric)**

This design allows early stopping for either efficacy or futility on either side. In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.

- **Two-Sided Efficacy with Futility (Asymmetric)**

This design allows early stopping for efficacy and efficacy futility, and for harm and harm futility (or efficacy 2 and efficacy 2 futility). In early looks of this design, the futility boundaries may overlap. Overlapping futility boundaries may be skipped or left as they are.

- **One-Sided Futility Only**

In this one-sided design, the interim analyses are used only for futility. Please be aware that these boundaries may take several minutes to compute, particularly when one or more stages are skipped.

- **Two-Sided Futility Only (Symmetric)**

In this two-sided design, the study is stopped early only for futility. Overlapping futility boundaries may be skipped or left as they are. Please be aware that these boundaries may take several minutes to compute, particularly when overlapping boundaries are removed or some stages are skipped.

- **Two-Sided Futility Only (Asymmetric)**

In this design, all stage previous to the final stage are used only for futility. The two sides are specified separately. Overlapping futility boundaries may be skipped or left as they are. Please be aware that these boundaries may take several minutes to compute, particularly when overlapping boundaries are removed or some stages are skipped.

Hypothesis Direction (One-Sided)

Specify the direction of the alternative hypothesis.

- **Ha: $P_1 - P_2 < 0$**

For this hypothesis direction the efficacy boundaries (if applicable) are lower boundaries and futility boundaries (if applicable) are upper boundaries.

- **Ha: $P_1 - P_2 > 0$**

For this hypothesis direction the efficacy boundaries (if applicable) are upper boundaries and futility boundaries (if applicable) are lower boundaries.

Hypothesis Direction (Two-Sided, Asymmetric)

Specify the efficacy 1 and efficacy 2 / harm sides of the hypotheses.

- **Efficacy 1 is Ha: $P_1 - P_2 < 0$, Efficacy 2 is Ha: $P_1 - P_2 > 0$**

For this hypothesis direction, the primary efficacy boundaries (if applicable) are lower boundaries, with primary futility boundaries (if applicable) as upper boundaries. The secondary (efficacy or harm) boundaries (if applicable) are upper boundaries, with secondary futility boundaries (if applicable) as lower boundaries.

Group-Sequential Tests for Two Proportions (Simulation)

- **Efficacy 1 is $H_a: P_1 - P_2 > 0$, Efficacy 2 is $H_a: P_1 - P_2 < 0$**

For this hypothesis direction, the primary efficacy boundaries (if applicable) are upper boundaries, with primary futility boundaries (if applicable) as lower boundaries. The secondary (efficacy or harm) boundaries (if applicable) are lower boundaries, with secondary futility boundaries (if applicable) as upper boundaries.

Boundary Specification

Specify whether the boundaries are calculated based on spending functions or input directly.

- **Spending Function Calculation**

Spending function methods are now mainstream methods since they are flexible to the proportion of information at each stage, and there is a wide variety of spending functions available.

- **Enter Boundaries Directly**

In some cases, it may be useful to enter specific Z -value boundaries. This selection provides that option.

Beta

Beta is the probability of a Type II error during the course of the study. Power is one minus beta. Beta is divided among the stages of the study according to the selected spending function. Beta values of 0.2 and 0.1 are common, but any value between 0 and 1 is eligible.

Spending Function (Alpha or Beta)

The alpha- and beta-spending functions determine how alpha and/or beta are distributed across the stages. See the Spending Function section of the documentation for more details.

- **Hwang-Shih-DeCani (Gamma Family)**

The Hwang-Shih-DeCani gamma family of spending function has a γ parameter that allows for a variety of spending functions. The γ parameter can be positive, negative, or 0.

- **O'Brien-Fleming Analog**

The O'Brien Fleming Analog (Lan & DeMets, 1983) roughly mimics the O'Brien-Fleming (non-spending function) design, with the key attribute that only a small proportion of alpha or beta is spent early. Its popularity comes from it proportioning enough alpha (or beta) to the final stage that the final stage boundary is not too different from the fixed-sample (non-group-sequential) boundary.

- **Pocock Analog**

The Pocock Analog (Lan & DeMets, 1983) roughly mimics the Pocock (non-spending function) design, with the key attribute that alpha and/or beta are spent roughly equally across all stages.

- **Power Family**

The power family of spending functions has a ρ parameter that gives flexibility in the spending function shape. The ρ parameter must be greater than 0. A power family spending function with a ρ of 1 is similar to a Pocock design, while a power family spending function with a ρ of 3 is more similar to an O'Brien-Fleming design.

- **Custom**

Select this option to enter specific cumulative alpha (or beta) values for each stage. The values entered should be between 0 and alpha (or half alpha for two-sided tests) or beta, increasing in value, and the final value should be alpha (or half alpha) or beta. The number of custom cumulative alpha (or beta) spent values should equal the Maximum Number of Stages.

Group-Sequential Tests for Two Proportions (Simulation)

γ (Gamma)

γ is used to define the Hwang-Shih-DeCani (γ) spending function. Negative values of γ spend more of alpha or beta at later looks, values near 0 spend alpha or beta evenly, and positive values of γ spend more of alpha or beta at earlier looks.

ρ

ρ is used to define the power family spending function. Only positive values for ρ are permitted. Values of ρ near zero spend more of alpha or beta at earlier looks, values near 1 spend alpha or beta evenly, and larger values of ρ spend more of alpha or beta at later looks. A power family spending function with a ρ of 1 is similar to a Pocock design, while a power family spending function with a ρ of 3 is more similar to an O'Brien-Fleming design.

Custom Cumulative Alpha or Beta Spent

Enter a series of cumulative alpha or beta values, separated by spaces. The number of values should equal the Maximum Number of Stages. As cumulative values, the values must be increasing. The final value could be set to equal alpha or beta. Otherwise, the values will be scaled so that the final value is equal to alpha or beta.

Boundary Z Values (Entered Directly)

Enter a Z value for each stage of this boundary, separated by spaces or commas. The number of Z values entered should match the Maximum Number of Stages (K).

Skipped Stages

Specify the stages that will be skipped, if any, for this boundary. The final stage cannot be skipped. The stage numbers should be separated by spaces or commas. For example, if the maximum number of stages is 5, but a boundary analysis at stages 2 and 4 is to be skipped, 2 4 may be entered here. The spending function accounts for the amount of alpha or beta that should be spent at a stage following a skipped stage.

Binding or Non-Binding Futility

This option specifies whether the futility boundaries are binding or non-binding.

- **Binding**

Binding futility boundaries are computed in concert with significance boundaries. They are called binding because they require the stopping of a trial if they are crossed. If the trial is not stopped, the probability of a false positive will exceed alpha.

- **Non-Binding**

When Non-binding futility boundaries are computed, the significance boundaries are first computed, ignoring the futility boundaries. The futility boundaries are then computed. These futility boundaries are non-binding because continuing the trial after they are crossed will not affect the overall probability of a false positive declaration. One effect of using non-binding futility boundaries is that the overall group-sequential test becomes conservative (alpha is lower than the stated alpha and power is lower than the stated power).

Overlapped Futility Boundaries

For two-sided tests with futility boundaries, there exists the possibility that one or more pairs of the early futility boundaries are crossed (the futility boundary for the upper side is lower than the futility boundary of the lower side). In this procedure, there are two ways to handle this scenario.

- **Remove (skip) overlapped futility boundaries**

For this selection, futility boundaries that overlap are removed completely (skipped). The spending function pushes the spending of beta to the first stage where the futility boundaries are not overlapped.

Group-Sequential Tests for Two Proportions (Simulation)

- **Keep overlapped futility boundaries**

For this choice, the boundaries are left as is. However, this creates a scenario where a z -statistic crossing a futility boundary from one side may also be in a location where it is inside an efficacy boundary from the other side. This can be difficult to reconcile, particularly if the futility boundary is binding.

Options Tab

The options on this panel specify some simulation and search details.

Simulations

Number of Simulations

Increasing the number of simulations improves the accuracy of the estimated probabilities of boundary crossing, but it also increases the computation time. One suggestion is to first run the procedure with a small number of simulations, such as 1,000 or 10,000. The simulation time can then be estimated and a reasonable increase in simulations can be made before a second run. A decent lower number of simulations is 100,000, but more is better when feasible.

After Boundary Crossing

This option specifies whether simulation Z -values are 'held out' after crossing a boundary, or whether simulation Z -values are compared to boundaries at all future stages, regardless of whether a boundary was crossed at a previous stage. For two-sided boundary scenarios, such as 'Two-sided Efficacy with Futility (Asymmetric)', the comparisons of the simulation Z -statistics to the two (upper and lower) sides are made independently.

- **Hold out**

For this selection, once a simulation Z -statistic crosses an efficacy boundary or a binding futility boundary, it is removed from consideration for future stages. For non-binding futility boundaries, simulation Z -statistics remain in the simulation for future stages, regardless of whether 'After Boundary Crossing' is set to 'Hold out' or 'Leave in'. When Hold Out is selected, power and alpha are calculated as the sum of all efficacy boundary proportions.

- **Leave in**

For this choice, simulation Z -statistics at all future stages are compared to the boundary values, regardless of whether a boundary was crossed at a previous stage. When Leave In is selected, power and alpha are calculated as the efficacy boundary proportion of the final stage.

Maximum N1 Before Search Termination

Specify the maximum $N1$ before the search for $N1$ is aborted. Since simulations for large sample sizes are very computationally intensive and hence time-consuming, this value can be used to stop searches when $N1$ is larger than reasonable sample sizes for the study. This applies only when "Solve For" is set to Sample Size. The procedure uses a binary search when searching for $N1$. If a value for $N1$ is tried that exceeds this value, and the power is not reached, a warning message will be shown on the output indicating the desired power was not reached.

Boundary Reports Tab

The options on this panel specify which boundary reports will be included in the output.

Boundary Reports for Each Scenario

Boundaries using Z Scale

This report gives the Z-Statistic boundary values and the information proportion at each stage.

Boundaries using P-Value Scale

This report gives the boundary p-values and the information proportion at each stage.

Information Report

The Information Report gives the target information, target information proportion, and target sample sizes for each stage.

Alpha Spending

This report shows the projected amount and percentage of alpha spent at each stage, both individually and cumulatively. The report is blank if boundaries are input directly.

Beta Spending

This report shows the projected amount and percentage of beta spent at each stage, both individually and cumulatively. The report is blank if there are no futility boundaries or if boundaries are input directly.

Decimal Places for Boundary Reports

Decimals

Specify the number of decimal places displayed for this item on the boundary reports.

Boundary Plots Tab

The options on this panel control the inclusion and appearance of the plots. The set of available plots depends upon the choice of 'Boundaries Used'.

Group-Sequential Boundary Plots

Z-Statistic vs Information / Stage / N

Check the box to display the plot. Click the plot format button (large square button) to change the plot settings. To edit the plot with live data, check the Live Edit box near the top-right corner before running the procedure. Use care when adding Stage Notes (notes for each stage just above the X-axis). These notes should only be added with live data, as their position is based on the X-axis values.

Live Edit

When checked, the procedure will stop while it is running to allow additional plot formatting. This is useful when the user wishes to format the plot with actual values rather than randomly generated data. The plot format changes will persist in the procedure settings, but are not automatically saved.

Summary Reports Tab

The options on this panel specify which summary reports will be included in the output.

Select Summary Report Output

Show Summary Reports

This report gives the summary of all the scenarios, one scenario per line.

Summary Plots Tab

The plots of the Summary Plots tab are used to visualize the (final stage) sample size or overall power in terms of the other parameters.

Example 1 – Sample Size and Initial Boundaries for a Group-Sequential Test

A childbirth study is to be conducted to determine whether a new approach during labor results in a lower proportion of C-sections than the standard techniques. The response for each patient is C-section or no C-section. A one-sided test with alpha equal to 0.025 is used. The Z-Test will use unpooled variance estimation with no continuity correction.

The new approach is assigned to Group 1, and the standard is assigned to Group 2, so that the null and alternative hypotheses are

$$H_0: P_1 - P_2 = 0 \quad (H_0: P_{New} = P_{Std})$$

versus

$$H_a: P_1 - P_2 < 0 \quad (H_a: P_{New} < P_{Std})$$

The design calls for five equally spaced stages if the final stage is reached. A power of 0.90 is needed. The assumed caesarian proportion for the standard approach is 0.31. Researchers wish to examine the sample sizes needed for new approach proportions of 0.21, 0.24, and 0.27. Both efficacy and non-binding futility boundaries are intended. The efficacy (alpha-spending) spending function used is the O'Brien-Fleming analog. The Hwang-Shih-DeCani (Gamma) beta-spending function with gamma parameter 1.5 is used for futility.

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 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	Sample Size
Power	0.90
Alpha	0.025
Group Allocation	Equal (N1 = N2)
P1	0.21 0.24 0.27
P2	0.31
Maximum Number of Stages (K)	5
Continuity Correction	No
Zero Count Adjustment Method	None
Info. Proportion at each Stage	Equally incremented
Boundaries Used	One-sided Efficacy with Futility
Hypothesis Direction	Ha: P1 – P2 < 0
Boundary Specification	Spending Function Calculation
Alpha Spending Function	O'Brien-Fleming Analog
Skipped Efficacy Stages	Blank
Design Beta	0.10
Beta Spending Function	Hwang-Shih-DeCani (γ)
γ	1.5
Skipped Futility Stages	Blank
Binding or Non-Binding Futility	Non-Binding

Group-Sequential Tests for Two Proportions (Simulation)**Options Tab**

Number of Simulations **100,000**
 After Boundary Crossing **Hold out**

Boundary Reports Tab

All Reports **Checked**

Boundary Plots Tab

Z-Statistic vs Information **Checked**
 Z-Statistic vs Stage **Checked**
 Z-Statistic vs N **Checked**

Summary Reports Tab

All Reports **Checked**

Summary Plots Tab

All plots **Checked [only 2D will be used]**

Annotated Output

Click the Calculate button to perform the calculations and generate the following output. Due to simulation time, this run will take a few minutes. The simulation results will differ slightly for each separate run.

Run Summary Report – Scenario 1

This report can be used to confirm that the input was processed as intended.

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	0
Alternative Hypothesis:	$P_1 - P_2 < 0$ (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani ($\gamma = 1.5$)
Futility Boundaries:	Non-Binding
Target Alpha:	0.0250
Alpha (from simulations):	0.02552
P1:	0.21
P2:	0.31
N1 (if final stage reached):	409
N2 (if final stage reached):	409
Target Power:	0.9
Power (from simulations):	0.90066
Maximum Information:	1076.8826

Z-Value Boundaries

This section gives the planning stage Z-statistic boundaries, numerically. These values are reflected in the group-sequential boundary plot.

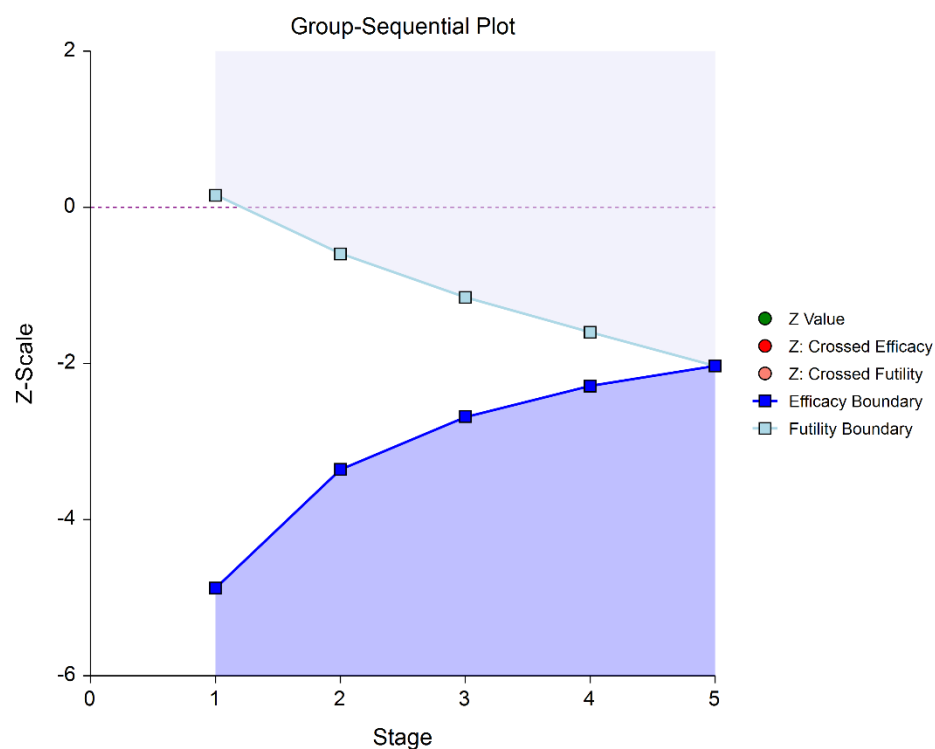
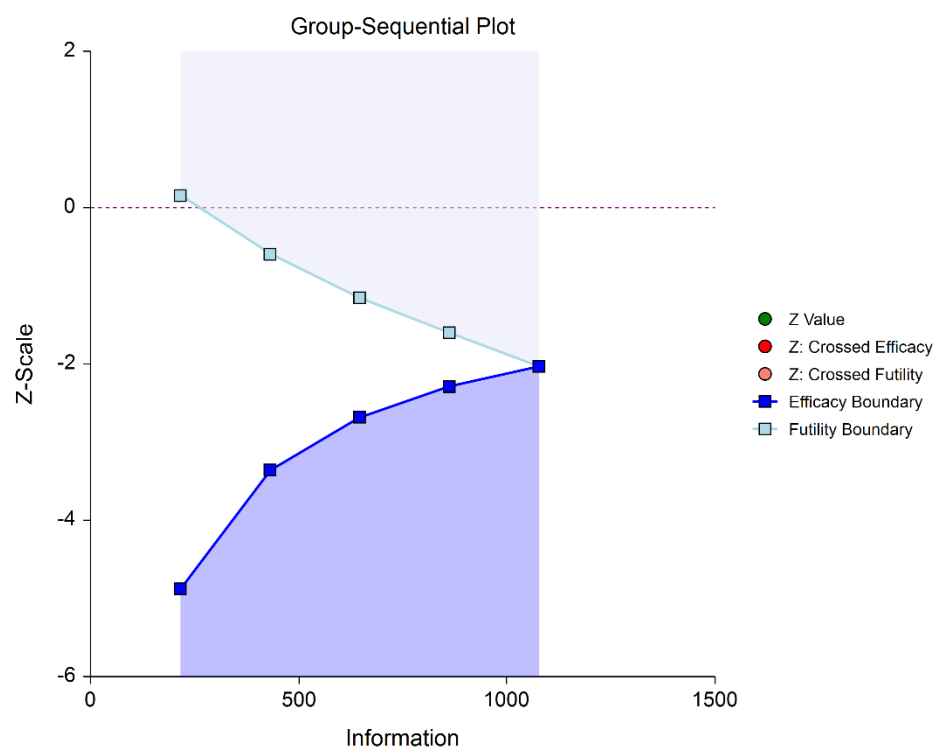
Maximum Information: 1076.8826
 Alternative Hypothesis: $P_1 - P_2 < 0$ (one-sided)
 Futility Boundaries: Non-Binding

Stage	Boundaries		Information Proportion
	Efficacy	Futility	
1	-4.8769	0.1534	0.2000
2	-3.3569	-0.5982	0.4000
3	-2.6803	-1.1542	0.6000
4	-2.2898	-1.6011	0.8000
5	-2.0310	-2.0310	1.0000

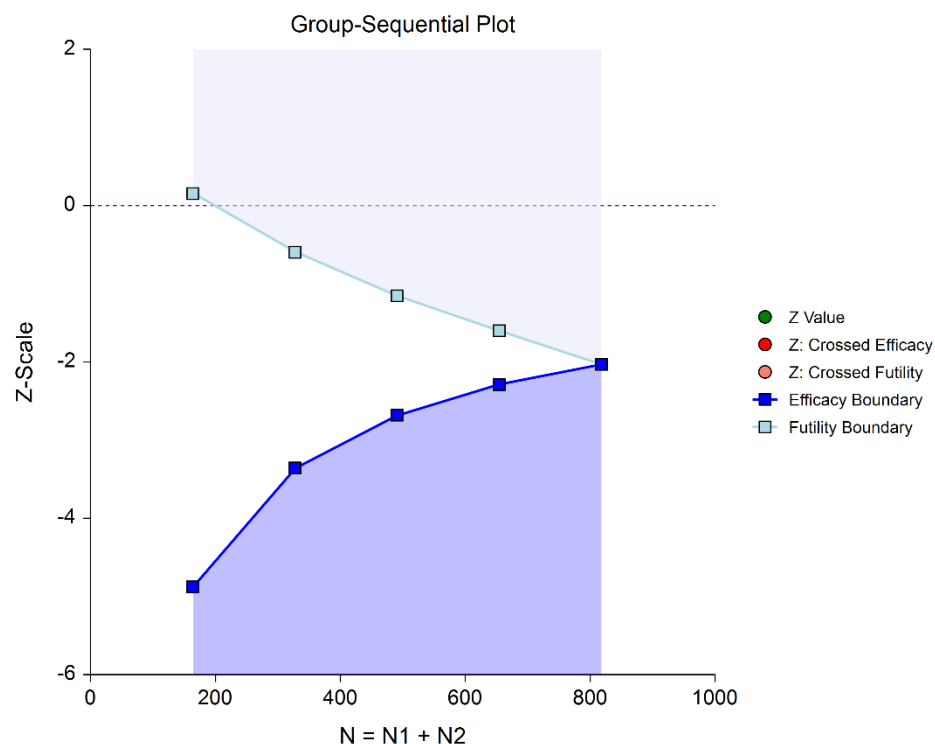
Group-Sequential Tests for Two Proportions (Simulation)

Boundary Plot(s)

This plot shows the efficacy and futility Z-statistic planning boundaries. It is anticipated that these boundaries will adjust to the actual information proportions as the data for each stage is realized.



Group-Sequential Tests for Two Proportions (Simulation)



P-Value Boundaries

This section reflects the conversion of the Z-value boundaries to the corresponding P-value boundaries.

Maximum Information: 1076.8826
 Alternative Hypothesis: $P_1 - P_2 < 0$ (one-sided)
 Futility Boundaries: Non-Binding
 P-value boundaries are one-sided values.

Stage	Boundaries		Information Proportion
	Efficacy	Futility	
1	0.00000	0.56095	0.2000
2	0.00039	0.27484	0.4000
3	0.00368	0.12421	0.6000
4	0.01102	0.05468	0.8000
5	0.02113	0.02113	1.0000

Group-Sequential Tests for Two Proportions (Simulation)

Information Report

This section gives the target information for each stage, as well as the sample sizes and proportions used to calculate those informations.

Maximum Information: 1076.8826
 Alternative Hypothesis: $P_1 - P_2 < 0$ (one-sided)
 Alpha: 0.0250

Stage	Target Information Proportion	Target Information	— Target Sample Size —		P1	P2
			N1	N2		
1	0.2000	215.3765	81.80	81.80	0.2100	0.3100
2	0.4000	430.7530	163.60	163.60	0.2100	0.3100
3	0.6000	646.1295	245.40	245.40	0.2100	0.3100
4	0.8000	861.5061	327.20	327.20	0.2100	0.3100
5	1.0000	1076.8826	409.00	409.00	0.2100	0.3100

Alpha Spending

This section shows how alpha is anticipated to be spent across the stages.

Target Final Stage Alpha: 0.0250
 Spending Function: O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	Nominal (Boundary) Alpha	Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1 *	0.2000	0.0000	0.0000	0.000001	0.0	0.0
2 *	0.4000	0.0004	0.0004	0.000394	1.6	1.6
3 *	0.6000	0.0034	0.0038	0.003678	13.7	15.2
4 *	0.8000	0.0084	0.0122	0.011017	33.6	48.8
5 *	1.0000	0.0128	0.0250	0.021128	51.2	100.0

* projected

Beta Spending for Futility

This section shows how beta is anticipated to be spent across the stages.

Target Cumulative Beta at Final Stage: 0.1000
 Spending Function for Futility: Hwang-Shih-DeCani ($\gamma = 1.5$)

Stage	Information Proportion	Beta Spent this Stage	Cumulative Beta Spent	Nominal (Boundary) Beta	Percentage Beta Spent this Stage	Cumulative Percentage Beta Spent
1 *	0.2000	0.0334	0.0334	0.560952	33.4	33.4
2 *	0.4000	0.0247	0.0581	0.274840	24.7	58.1
3 *	0.6000	0.0183	0.0764	0.124207	18.3	76.4
4 *	0.8000	0.0136	0.0900	0.054676	13.6	90.0
5 *	1.0000	0.0100	0.1000	0.021128	10.0	100.0

* projected

Group-Sequential Tests for Two Proportions (Simulation)

Boundary Probabilities for $\delta = -0.1$

Using simulation based on the specified proportions, this section gives the estimated probabilities of crossing each of the boundaries. Values given here will vary for each simulation.

Number of Simulations: 100000
 Futility Boundaries: Non-Binding
 After Efficacy Boundary Crossing: Hold Out
 After Non-Binding Futility Boundary Crossing: Leave In
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 P1: 0.21
 P2: 0.31
 δ : -0.1

Stage	N1	N2	Efficacy		Futility	
			Boundary	Probability	Boundary	Probability
1	*81.80	*81.80	-4.8769	0.0007	0.1534	0.0586
2	*163.60	*163.60	-3.3569	0.1073	-0.5982	0.0650
3	*245.40	*245.40	-2.6803	0.3414	-1.1542	0.0819
4	*327.20	*327.20	-2.2898	0.2997	-1.6011	0.0902
5	*409.00	*409.00	-2.0310	0.1515	-2.0310	0.1057

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 302.057134
 Average N2: 302.057134

Boundary Probabilities for $\delta = 0$ (Alpha)

This section estimates the probabilities of crossing each boundary if the difference for the remaining stages is assumed to be zero (the proportions are assumed to be the same).

Number of Simulations: 100000
 Futility Boundaries: Non-Binding
 After Efficacy Boundary Crossing: Hold Out
 After Non-Binding Futility Boundary Crossing: Leave In
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 P1: 0.31
 P2: 0.31
 δ : 0

Stage	N1	N2	Efficacy		Futility	
			Boundary	Probability	Boundary	Probability
1	*81.80	*81.80	-4.8769	0.0000	0.1534	0.4639
2	*163.60	*163.60	-3.3569	0.0004	-0.5982	0.7233
3	*245.40	*245.40	-2.6803	0.0039	-1.1542	0.8717
4	*327.20	*327.20	-2.2898	0.0086	-1.6011	0.9448
5	*409.00	*409.00	-2.0310	0.0126	-2.0310	0.9786

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 407.555412
 Average N2: 407.555412

Scenario 2

All of the same boundary reports are given for Scenario 2, corresponding to a P1 value of 0.24.

Scenario 3

All of the same boundary reports are given for Scenario 3, corresponding to a P1 value of 0.27.

Group-Sequential Tests for Two Proportions (Simulation)

Power and Sample Size Summary

Maximum Number of Stages: 5
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 Alpha Spending Function: O'Brien-Fleming Analog
 Beta Spending Function: Hwang-Shih-DeCani ($\gamma = 1.5$)

Target Power	Sim Power	N1	N2	N	P1	P2	Target Alpha	Sim Alpha
0.90	0.90066	409	409	818	0.2100	0.3100	0.0250	0.0262
0.90	0.90139	873	873	1746	0.2400	0.3100	0.0250	0.0246
0.90	0.90002	2776	2776	5552	0.2700	0.3100	0.0250	0.0251

References

Jennison, C. and Turnbull, B.W. 2000. Group Sequential Methods with Applications to Clinical Trials. Chapman and Hall/CRC. Boca Raton.
 Lan, K.K.G. and DeMets, D.L. 1983. 'Discrete sequential boundaries for clinical trials.' Biometrika, 70, pages 659-663.
 Reboussin, D.M.; DeMets, D.L.; Kim, K; and Lan, K.K.G. 1992. 'Programs for computing group sequential boundaries using the Lan-DeMets Method.' Technical Report 60, Department of Biostatistics, University of Wisconsin-Madison.

Report Definitions

Target Power is the desired power value (or values) entered in the procedure.
 Sim Power is the proportion of simulation z-values that cross an efficacy boundary. Because 'After Boundary Crossing' is set to 'Hold out', it is the sum of the individual boundary crossing proportions.
 N1 and N2 are the anticipated number of individuals in each group if the final stage is reached.
 N is the total sample size, $N1 + N2$, if the final stage is reached.
 P1 is the assumed proportion of population 1 for power calculation simulations.
 P2 is the assumed proportion of population 2 for power calculation simulations. P2 is also the assumed proportion of populations 1 and 2 for alpha calculation simulations.
 Target Alpha is the alpha used in the computation of the boundaries, and is the desired overall probability of a Type 1 error.
 Sim Alpha is the proportion of null simulation z-values that cross an efficacy boundary. Because 'After Boundary Crossing' is set to 'Hold out', it is the sum of the individual boundary crossing proportions.

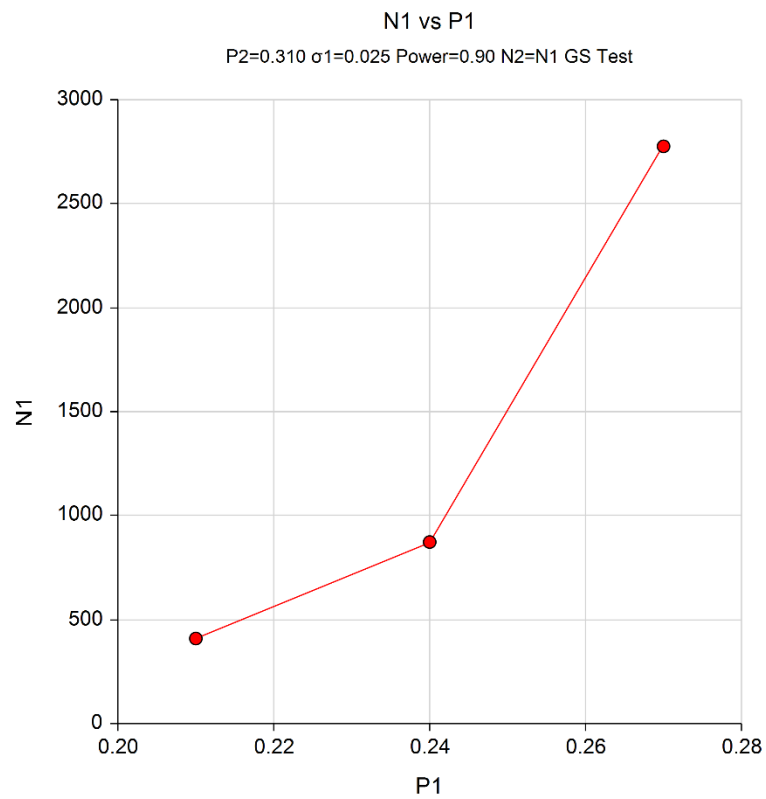
Summary Statements

Final stage sample sizes of 409 and 409 achieve 90.066% power to detect a proportion difference of 0.1000 at a target alpha-level of 0.025. The assumed population proportions are 0.2100 and 0.3100. These results are based on one-sided Z-test group sequential testing with 5 stages and boundary values calculated from spending functions.

This report shows the values of each of the parameters, one scenario per row. The values may vary slightly due to the variation in simulations. The details for each of the rows of this report are given in the earlier boundary reports. The values from this table are exhibited in the chart below.

Group-Sequential Tests for Two Proportions (Simulation)

Chart Section for Power and Sample Size Summary



Example 2 – Skipping Stage Boundaries

Suppose that the scenario is exactly as in Example 1, except that the first two futility boundaries are to be skipped.

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 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>
Design Tab	
Solve For	Sample Size
Power.....	0.90
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
P1.....	0.21 0.24 0.27
P2.....	0.31
Maximum Number of Stages (K)	5
Continuity Correction	No
Zero Count Adjustment Method.....	None
Info. Proportion at each Stage.....	Equally incremented
Boundaries Used	One-sided Efficacy with Futility
Hypothesis Direction.....	Ha: P1 – P2 < 0
Boundary Specification	Spending Function Calculation
Alpha Spending Function	O'Brien-Fleming Analog
Skipped Efficacy Stages	<i>Blank</i>
Design Beta	0.10
Beta Spending Function	Hwang-Shih-DeCani (γ)
γ	1.5
Skipped Futility Stages	1 2
Binding or Non-Binding Futility	Non-Binding
Options Tab	
Number of Simulations	100,000
After Boundary Crossing	Hold out
Boundary Reports Tab	
All Reports	Checked
Boundary Plots Tab	
Z-Statistic vs Information	Checked
Z-Statistic vs Stage.....	Checked
Z-Statistic vs N.....	Checked
Summary Reports Tab	
All Reports	Checked
Summary Plots Tab	
All plots	Checked [only 2D will be used]

Group-Sequential Tests for Two Proportions (Simulation)

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Run Summary Report

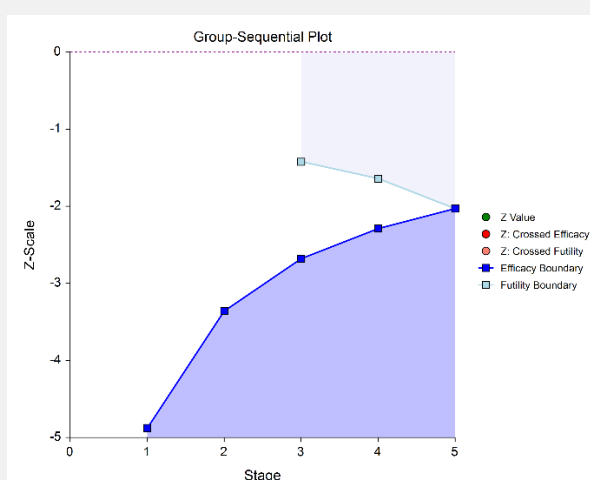
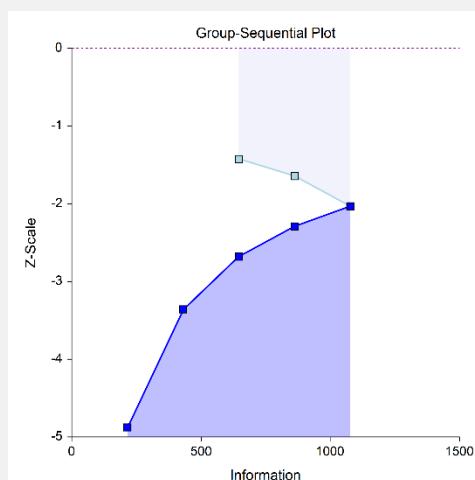
Item	Value
Maximum Number of Stages (Design):	5
Skipped Futility Stage(s):	1 2
Current Stage:	0
Alternative Hypothesis:	$P_1 - P_2 < 0$ (one-sided)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani ($\gamma = 1.5$)
Futility Boundaries:	Non-Binding
Target Alpha:	0.0250
Alpha (from simulations):	0.02552
P1:	0.21
P2:	0.31
N1 (if final stage reached):	409
N2 (if final stage reached):	409
Target Power:	0.9
Power (from simulations):	0.90066
Maximum Information:	1076.8826

Z-Value Boundaries

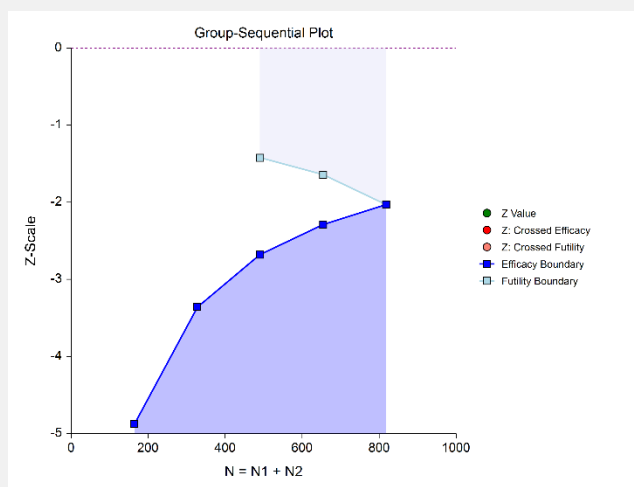
Maximum Information: 1076.8826
 Alternative Hypothesis: $P_1 - P_2 < 0$ (one-sided)
 Futility Boundaries: Non-Binding

Stage	Boundaries		Information Proportion
	Efficacy	Futility	
1	-4.8769		0.2000
2	-3.3569		0.4000
3	-2.6803	-1.4232	0.6000
4	-2.2898	-1.6443	0.8000
5	-2.0310	-2.0310	1.0000

Boundary Plot(s)



Group-Sequential Tests for Two Proportions (Simulation)



P-Value Boundaries

Maximum Information: 1076.8826
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 Futility Boundaries: Non-Binding
 P-value boundaries are one-sided values.

Stage	Boundaries		Information Proportion
	Efficacy	Futility	
1	0.00000		0.2000
2	0.00039		0.4000
3	0.00368	0.07733	0.6000
4	0.01102	0.05006	0.8000
5	0.02113	0.02113	1.0000

Information Report

Maximum Information: 1076.8826
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 Alpha: 0.0250

Stage	Target Information Proportion	Target Information	— Target Sample Size —		P1	P2
			N1	N2		
1	0.2000	215.3765	81.80	81.80	0.2100	0.3100
2	0.4000	430.7530	163.60	163.60	0.2100	0.3100
3	0.6000	646.1295	245.40	245.40	0.2100	0.3100
4	0.8000	861.5061	327.20	327.20	0.2100	0.3100
5	1.0000	1076.8826	409.00	409.00	0.2100	0.3100

Group-Sequential Tests for Two Proportions (Simulation)

Alpha Spending

Target Final Stage Alpha: 0.0250
Spending Function: O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	Nominal (Boundary) Alpha	Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1 *	0.2000	0.0000	0.0000	0.000001	0.0	0.0
2 *	0.4000	0.0004	0.0004	0.000394	1.6	1.6
3 *	0.6000	0.0034	0.0038	0.003678	13.7	15.2
4 *	0.8000	0.0084	0.0122	0.011017	33.6	48.8
5 *	1.0000	0.0128	0.0250	0.021128	51.2	100.0

* projected

Beta Spending for Futility

Target Cumulative Beta at Final Stage: 0.1000
Spending Function for Futility: Hwang-Shih-DeCani ($\gamma = 1.5$)

Stage	Information Proportion	Beta Spent this Stage	Cumulative Beta Spent	Nominal (Boundary) Beta	Percentage Beta Spent this Stage	Cumulative Percentage Beta Spent
1 *	0.2000	0.0000	0.0000	1.000000	0.0	0.0
2 *	0.4000	0.0000	0.0000	1.000000	0.0	0.0
3 *	0.6000	0.0764	0.0764	0.077334	76.4	76.4
4 *	0.8000	0.0136	0.0900	0.050056	13.6	90.0
5 *	1.0000	0.0100	0.1000	0.021128	10.0	100.0

* projected

Boundary Probabilities for $\delta = -0.1$

Number of Simulations: 100000
Futility Boundaries: Non-Binding
After Efficacy Boundary Crossing: Hold Out
After Non-Binding Futility Boundary Crossing: Leave In
Alternative Hypothesis: $P_1 - P_2 < 0$ (one-sided)
P1: 0.21
P2: 0.31
 δ : -0.1

Stage	N1	N2	Efficacy		Futility	
			Boundary	Probability	Boundary	Probability
1	*81.80	*81.80	-4.8769	0.0007		0.0000
2	*163.60	*163.60	-3.3569	0.1073		0.0000
3	*245.40	*245.40	-2.6803	0.3414	-1.4232	0.1285
4	*327.20	*327.20	-2.2898	0.2997	-1.6443	0.0969
5	*409.00	*409.00	-2.0310	0.1515	-2.0310	0.1057

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 302.057134
Average N2: 302.057134

Group-Sequential Tests for Two Proportions (Simulation)

Boundary Probabilities for $\delta = 0$ (Alpha)

Number of Simulations: 100000
 Futility Boundaries: Non-Binding
 After Efficacy Boundary Crossing: Hold Out
 After Non-Binding Futility Boundary Crossing: Leave In
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 $P1$: 0.31
 $P2$: 0.31
 δ : 0

Stage	N1	N2	Efficacy		Futility	
			Boundary	Probability	Boundary	Probability
1	*81.80	*81.80	-4.8769	0.0000		0.0000
2	*163.60	*163.60	-3.3569	0.0004		0.0000
3	*245.40	*245.40	-2.6803	0.0039	-1.4232	0.9205
4	*327.20	*327.20	-2.2898	0.0086	-1.6443	0.9493
5	*409.00	*409.00	-2.0310	0.0126	-2.0310	0.9786

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 407.555412
 Average N2: 407.555412

Scenario 2

All of the same boundary reports are given for Scenario 2, corresponding to a $P1$ value of 0.24.

Scenario 3

All of the same boundary reports are given for Scenario 3, corresponding to a $P1$ value of 0.27.

Power and Sample Size Summary

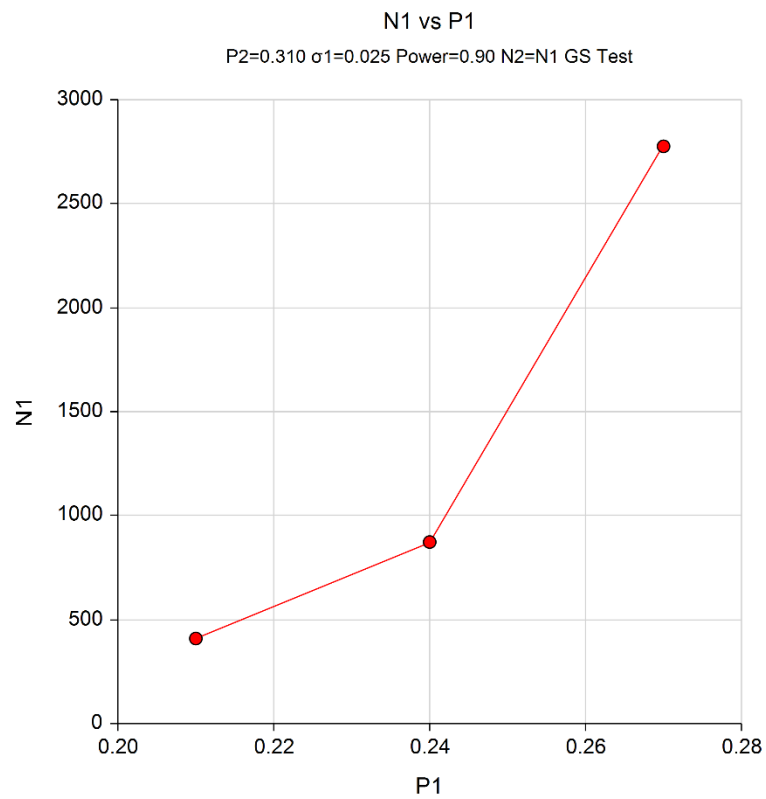
Maximum Number of Stages: 5
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 Alpha Spending Function: O'Brien-Fleming Analog
 Beta Spending Function: Hwang-Shih-DeCani ($\gamma = 1.5$)

Target Power	Sim Power	N1	N2	N	P1	P2	Target Alpha	Sim Alpha
0.90	0.90066	409	409	818	0.2100	0.3100	0.0250	0.0262
0.90	0.90139	873	873	1746	0.2400	0.3100	0.0250	0.0246
0.90	0.90002	2776	2776	5552	0.2700	0.3100	0.0250	0.0251

This report shows no change in overall sample size.

Group-Sequential Tests for Two Proportions (Simulation)

Chart Section for Power and Sample Size Summary



Example 3 – Finding Power

Suppose that the scenario is similar to the setup of Example 1, except that now we will solve for power for various sample sizes of 200, 400, 600, 800, and 1000 per group.

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 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>
Design Tab	
Solve For	Power
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	200 400 600 800 1000
P1.....	0.21 0.24 0.27
P2.....	0.31
Maximum Number of Stages (K)	5
Continuity Correction	No
Zero Count Adjustment Method.....	None
Info. Proportion at each Stage.....	Equally incremented
Boundaries Used	One-sided Efficacy with Futility
Hypothesis Direction.....	Ha: P1 – P2 < 0
Boundary Specification	Spending Function Calculation
Alpha Spending Function	O'Brien-Fleming Analog
Skipped Efficacy Stages	Blank
Design Beta	0.10
Beta Spending Function	Hwang-Shih-DeCani (γ)
γ	1.5
Skipped Futility Stages	Blank
Binding or Non-Binding Futility	Non-Binding
Options Tab	
Number of Simulations	100,000
After Boundary Crossing	Hold out
Boundary Reports Tab	
All Reports	Checked
Boundary Plots Tab	
Z-Statistic vs Information	Checked
Z-Statistic vs Stage.....	Checked
Z-Statistic vs N.....	Checked
Summary Reports Tab	
All Reports	Checked
Summary Plots Tab	
All plots	Checked

Group-Sequential Tests for Two Proportions (Simulation)

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Scenario Reports

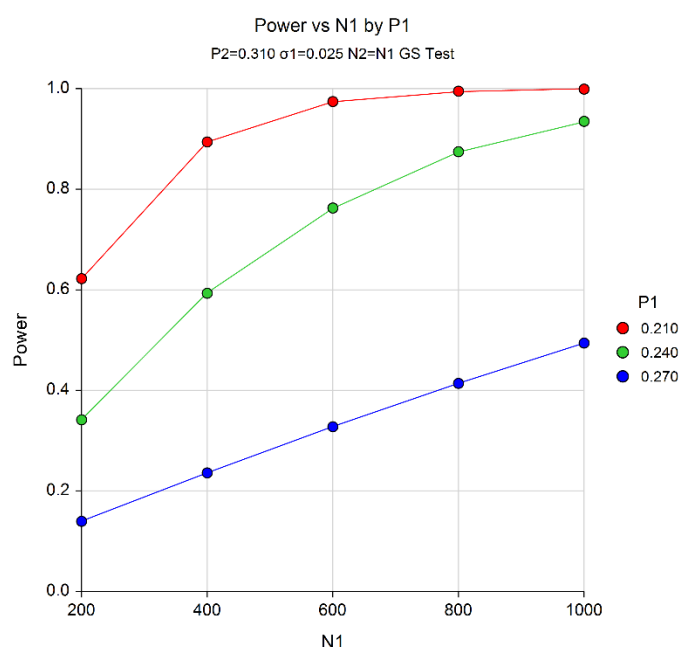
All of the scenario reports for each of the 15 scenarios are generated in the output, but they are not shown here.

Power and Sample Size Summary

Maximum Number of Stages: 5
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 Alpha Spending Function: O'Brien-Fleming Analog
 Beta Spending Function: Hwang-Shih-DeCani ($\gamma = 1.5$)

Sim Power	N1	N2	N	P1	P2	Target Alpha	Sim Alpha
0.62215	200	200	400	0.2100	0.3100	0.0250	0.0254
0.89406	400	400	800	0.2100	0.3100	0.0250	0.0248
0.97439	600	600	1200	0.2100	0.3100	0.0250	0.0239
0.99500	800	800	1600	0.2100	0.3100	0.0250	0.0244
0.99907	1000	1000	2000	0.2100	0.3100	0.0250	0.0256
0.34133	200	200	400	0.2400	0.3100	0.0250	0.0256
0.59307	400	400	800	0.2400	0.3100	0.0250	0.0254
0.76264	600	600	1200	0.2400	0.3100	0.0250	0.0243
0.87456	800	800	1600	0.2400	0.3100	0.0250	0.0244
0.93496	1000	1000	2000	0.2400	0.3100	0.0250	0.0252
0.13940	200	200	400	0.2700	0.3100	0.0250	0.0264
0.23579	400	400	800	0.2700	0.3100	0.0250	0.0255
0.32761	600	600	1200	0.2700	0.3100	0.0250	0.0242
0.41381	800	800	1600	0.2700	0.3100	0.0250	0.0248
0.49399	1000	1000	2000	0.2700	0.3100	0.0250	0.0247

Chart Section for Power and Sample Size Summary



The power curve plot shows the effect of sample size and proportion difference on the power for each scenario.

Example 4 – Finding Power with Binding Futility Boundaries

Following the setup of Example 3, we wish to see the effect on power of changing from non-binding futility boundaries to binding futility boundaries.

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 make the appropriate entries as listed below or open **Example 4** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	200 400 600 800 1000
P1.....	0.21 0.24 0.27
P2.....	0.31
Maximum Number of Stages (K)	5
Continuity Correction	No
Zero Count Adjustment Method.....	None
Info. Proportion at each Stage.....	Equally incremented
Boundaries Used	One-sided Efficacy with Futility
Hypothesis Direction.....	Ha: P1 – P2 < 0
Boundary Specification	Spending Function Calculation
Alpha Spending Function	O'Brien-Fleming Analog
Skipped Efficacy Stages	<i>Blank</i>
Design Beta	0.10
Beta Spending Function	Hwang-Shih-DeCani (γ)
γ	1.5
Skipped Futility Stages	<i>Blank</i>
Binding or Non-Binding Futility	Binding
Options Tab	
Number of Simulations	100,000
After Boundary Crossing	Hold out
Boundary Reports Tab	
All Reports	Checked
Boundary Plots Tab	
Z-Statistic vs Information	Checked
Z-Statistic vs Stage.....	Checked
Z-Statistic vs N.....	Checked
Summary Reports Tab	
All Reports	Checked
Summary Plots Tab	
All plots	Checked

Group-Sequential Tests for Two Proportions (Simulation)

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Scenario Reports

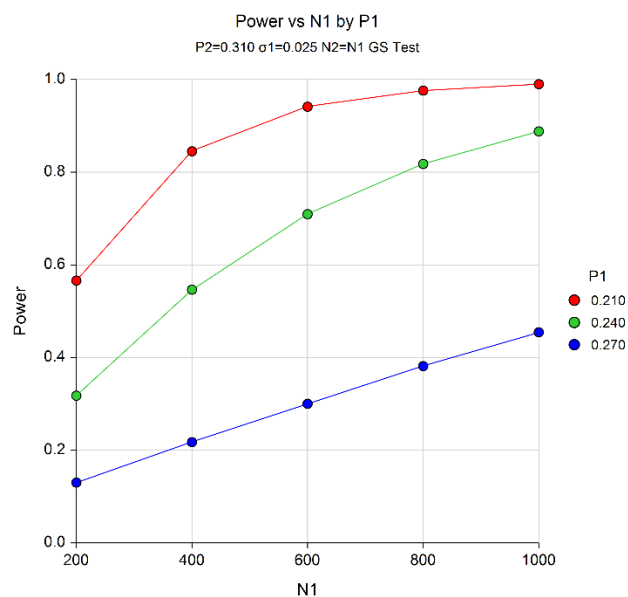
All the scenario reports for each of the 15 scenarios are generated in the output, but they are not shown here.

Power and Sample Size Summary

Maximum Number of Stages: 5
 Alternative Hypothesis: $P1 - P2 < 0$ (one-sided)
 Alpha Spending Function: O'Brien-Fleming Analog
 Beta Spending Function: Hwang-Shih-DeCani ($\gamma = 1.5$)

Sim Power	N1	N2	N	P1	P2	Target Alpha	Sim Alpha
0.56600	200	200	400	0.2100	0.3100	0.0250	0.0275
0.84501	400	400	800	0.2100	0.3100	0.0250	0.0257
0.94122	600	600	1200	0.2100	0.3100	0.0250	0.0243
0.97602	800	800	1600	0.2100	0.3100	0.0250	0.0247
0.98994	1000	1000	2000	0.2100	0.3100	0.0250	0.0252
0.31767	200	200	400	0.2400	0.3100	0.0250	0.0263
0.54648	400	400	800	0.2400	0.3100	0.0250	0.0255
0.70937	600	600	1200	0.2400	0.3100	0.0250	0.0246
0.81774	800	800	1600	0.2400	0.3100	0.0250	0.0248
0.88801	1000	1000	2000	0.2400	0.3100	0.0250	0.0246
0.13016	200	200	400	0.2700	0.3100	0.0250	0.0262
0.21741	400	400	800	0.2700	0.3100	0.0250	0.0254
0.30031	600	600	1200	0.2700	0.3100	0.0250	0.0248
0.38148	800	800	1600	0.2700	0.3100	0.0250	0.0247
0.45450	1000	1000	2000	0.2700	0.3100	0.0250	0.0251

Chart Section for Power and Sample Size Summary



If the power results are compared to those with non-binding futility boundaries in Example 3, it is seen that the power for binding futility boundaries is several percent lower. Higher numbers of simulations might be used to fine-tune these differences.

Example 5 – Comparing Numbers of Stages

Following the setup of Example 3, we wish to see the effect on power of changing the number of stages. This requires multiple runs with different numbers of stages. The numbers of stages examined here are 2, 3, 4, 5, 10 and 20. A proportion for Group 1 of 0.24 and a sample size of 600 per group will be considered.

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 make the appropriate entries as listed below or open **Example 5** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Power.....	0.90
Alpha.....	0.025
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	600
P1.....	0.24
P2.....	0.31
Maximum Number of Stages (K)	2 (Also run with 3, 4, 5, 10, and 20)
Continuity Correction	No
Zero Count Adjustment Method.....	None
Info. Proportion at each Stage.....	Equally incremented
Boundaries Used	One-sided Efficacy with Futility
Hypothesis Direction.....	Ha: P1 – P2 < 0
Boundary Specification	Spending Function Calculation
Alpha Spending Function	O'Brien-Fleming Analog
Skipped Efficacy Stages	<i>Blank</i>
Design Beta	0.10
Beta Spending Function	Hwang-Shih-DeCani (γ)
γ	1.5
Skipped Futility Stages	<i>Blank</i>
Binding or Non-Binding Futility	Non-Binding
Options Tab	
Number of Simulations	100,000
After Boundary Crossing	Hold out
Boundary Reports Tab	
All Reports	Checked
Boundary Plots Tab	
All 3 Plots.....	Checked
Summary Reports Tab	
All Reports	Checked
Summary Plots Tab	
All plots	Checked [only 2D will be used]

Group-Sequential Tests for Two Proportions (Simulation)

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Scenario Reports

All the scenario reports for each of the 6 scenarios are generated in the output, but they are not shown here.

Power and Sample Size Summary

Maximum Number of Stages: 2 3 4 5 10 20 [separate runs]
 Alternative Hypothesis: $P_1 - P_2 < 0$ (one-sided)
 Alpha Spending Function: O'Brien-Fleming Analog
 Beta Spending Function: Hwang-Shih-DeCani ($\gamma = 1.5$)

Sim Power	N1	N2	N	P1	P2	Target Alpha	Sim Alpha	Number of Stages
0.78150	600	600	1200	0.2400	0.3100	0.0250	0.0251	2
0.77779	600	600	1200	0.2400	0.3100	0.0250	0.0254	3
0.76647	600	600	1200	0.2400	0.3100	0.0250	0.0243	4
0.76255	600	600	1200	0.2400	0.3100	0.0250	0.0232	5
0.76086	600	600	1200	0.2400	0.3100	0.0250	0.0258	10
0.75808	600	600	1200	0.2400	0.3100	0.0250	0.0242	20

It is seen that the impact of the number of stages on the overall power is very minor.

Example 6 – Two-Sided Boundaries

Suppose that the scenario is similar to the setup of Example 3, but the boundary structure is changed to two-sided boundaries with an alpha of 0.05.

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 make the appropriate entries as listed below or open **Example 6** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Power
Power.....	0.90
Alpha.....	0.05
Group Allocation	Equal (N1 = N2)
Sample Size Per Group	200 400 600 800 1000
P1.....	0.21 0.24 0.27
P2.....	0.31
Maximum Number of Stages (K)	5
Continuity Correction	No
Zero Count Adjustment Method.....	None
Info. Proportion at each Stage.....	Equally incremented
Boundaries Used	Two-sided Efficacy with Futility (Symmetric)
Boundary Specification	Spending Function Calculation
Alpha Spending Function	O'Brien-Fleming Analog
Skipped Efficacy Stages	<i>Blank</i>
Design Beta	0.10
Beta Spending Function	Hwang-Shih-DeCani (γ)
γ	1.5
Skipped Futility Stages	<i>Blank</i>
Binding or Non-Binding Futility	Non-Binding
Overlapped Futility Boundaries	Remove (skip) overlapped futility boundaries
Options Tab	
Number of Simulations	100,000
After Boundary Crossing	Hold out
Boundary Reports Tab	
All Reports	Checked
Boundary Plots Tab	
All 3 Plots.....	Checked
Summary Reports Tab	
All Reports	Checked
Summary Plots Tab	
All plots	Checked

Group-Sequential Tests for Two Proportions (Simulation)

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Run Summary Report – Scenario 1

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	0
Alternative Hypothesis:	$P_1 - P_2 \neq 0$ (two-sided, symmetric)
Alpha Spending Function:	O'Brien-Fleming Analog
Beta Spending Function:	Hwang-Shih-DeCani ($\gamma = 1.5$)
Futility Boundaries:	Non-Binding
Target Alpha:	0.0500
Alpha (from simulations):	0.03717
P1:	0.21
P2:	0.31
N1 (if final stage reached):	200
N2 (if final stage reached):	200
Power (from simulations):	0.51148
Maximum Information:	526.5929

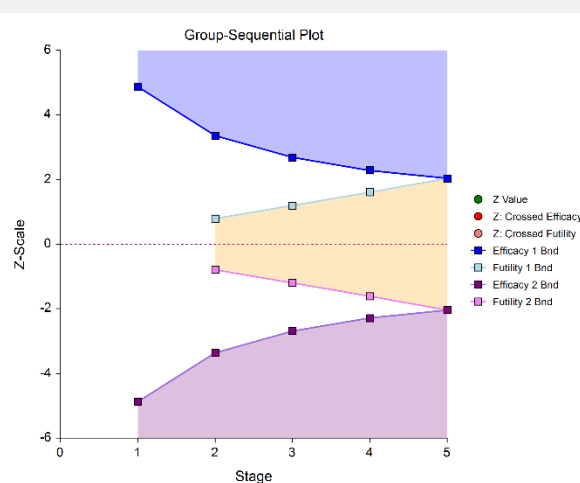
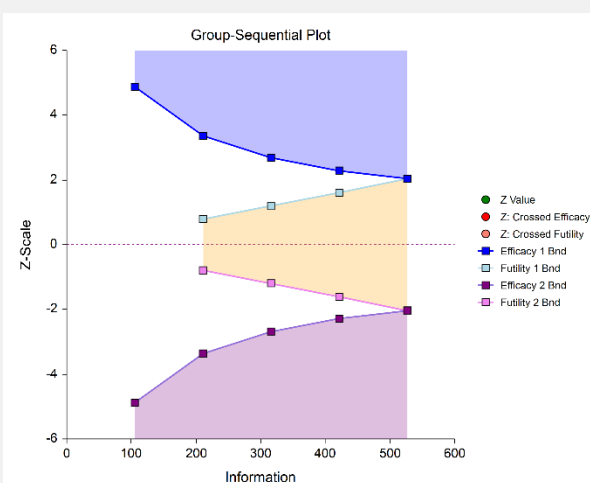
Z-Value Boundaries – Scenario 1

Maximum Information: 0.0480

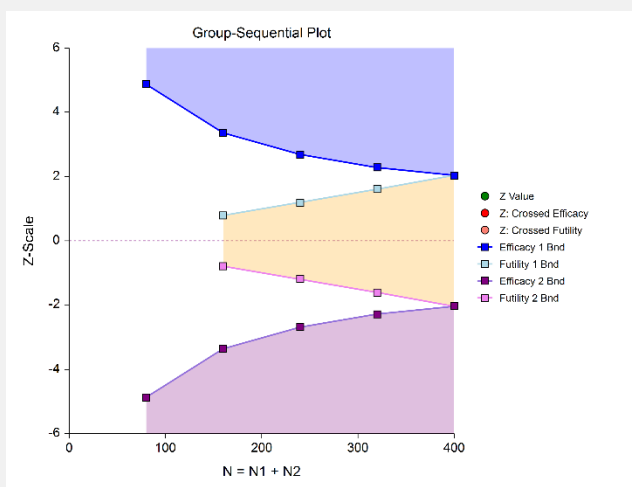
Alternative Hypothesis: $\mu_1 - \mu_2 \neq 0$ (two-sided, symmetric)

	Boundaries				Information Proportion
Stage	Upper Side		Lower Side		
	Efficacy 1	Futility 1	Efficacy 2	Futility 2	
1	4.8769		-4.8769		0.2000
2	3.3569	0.7907	-3.3569	-0.7907	0.4000
3	2.6803	1.1896	-2.6803	-1.1896	0.6000
4	2.2898	1.6104	-2.2898	-1.6104	0.8000
5	2.0310	2.0310	-2.0310	-2.0310	1.0000

Boundary Plot(s) – Scenario 1



Group-Sequential Tests for Two Proportions (Simulation)



Other Scenario Reports

All of the scenario reports for each of the 15 scenarios are generated in the output, but they are not shown here.

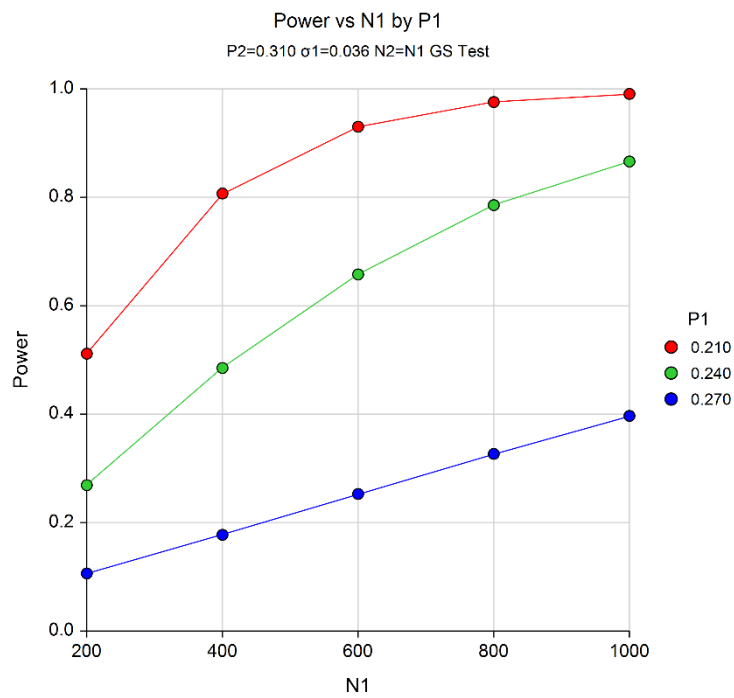
Power and Sample Size Summary

Maximum Number of Stages: 5
 Alternative Hypothesis: $P1 - P2 \neq 0$ (two-sided, symmetric)
 Alpha Spending Function: O'Brien-Fleming Analog
 Beta Spending Function: Hwang-Shih-DeCani ($\gamma = 1.5$)

Sim Power	N1	N2	N	P1	P2	Target Alpha	Sim Alpha
0.51148	200	200	400	0.2100	0.3100	0.0500	0.0372
0.80697	400	400	800	0.2100	0.3100	0.0500	0.0371
0.93008	600	600	1200	0.2100	0.3100	0.0500	0.0384
0.97564	800	800	1600	0.2100	0.3100	0.0500	0.0367
0.99067	1000	1000	2000	0.2100	0.3100	0.0500	0.0365
0.26900	200	200	400	0.2400	0.3100	0.0500	0.0389
0.48506	400	400	800	0.2400	0.3100	0.0500	0.0371
0.65772	600	600	1200	0.2400	0.3100	0.0500	0.0378
0.78552	800	800	1600	0.2400	0.3100	0.0500	0.0367
0.86599	1000	1000	2000	0.2400	0.3100	0.0500	0.0363
0.10613	200	200	400	0.2700	0.3100	0.0500	0.0387
0.17735	400	400	800	0.2700	0.3100	0.0500	0.0375
0.25274	600	600	1200	0.2700	0.3100	0.0500	0.0355
0.32664	800	800	1600	0.2700	0.3100	0.0500	0.0360
0.39670	1000	1000	2000	0.2700	0.3100	0.0500	0.0365

Group-Sequential Tests for Two Proportions (Simulation)

Chart Section for Power and Sample Size Summary



The power curve plot shows the effect of sample size and proportion difference on the power for each scenario.

Example 7 – Validation Using O’Brien-Fleming Analog Boundaries

Reboussin (1992) presents an example for Binomial distributed data for a design with two-sided O’Brien-Fleming boundaries, 5 stages, $\alpha = 0.05$, $\beta = 0.10$, $P_1 = 0.1100$, $P_2 = 0.0825$. They compute a sample size of 2381.78 per group. The upper boundaries are: 4.8769, 3.3569, 2.6803, 2.2898, 2.0310. We noticed a rounding issue with the calculations in Reboussin (1992). The sample size calculation of 2381.78 came from 3-digit rounding:

$$n_K = \frac{2(0.096)(0.904)(3.28)^2}{(0.028)^2} = 2381.78$$

With 4-digit rounding, we get

$$n_K = \frac{2(0.09625)(0.90375)(3.27939)^2}{(0.0275)^2} = 2474.00$$

To compare the result of Reboussin (1992) to that of PASS, enter the following.

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 make the appropriate entries as listed below or open **Example 7** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	Sample Size
Power.....	0.90
Alpha.....	0.05
Group Allocation	Equal (N1 = N2)
P1.....	0.1100
P2.....	0.0825
Maximum Number of Stages (K)	5
Info. Proportion at each Stage.....	Equally incremented
Continuity Correction	No
Zero Count Adjustment Method.....	None
Boundaries Used	Two-sided Efficacy Only (Symmetric)
Boundary Specification	Spending Function Calculation
Alpha Spending Function	O’Brien-Fleming Analog
Skipped Efficacy Stages.....	<i>Blank</i>
Options Tab	
Number of Simulations	100,000
After Boundary Crossing	Hold out
Boundary Reports Tab	
Boundaries using Z Scale.....	Checked
Information Report	Checked
Alpha Spending	Checked
Boundary Plots Tab	
Z-Statistic vs N.....	Checked
Summary Reports Tab	
All Reports	Checked

Group-Sequential Tests for Two Proportions (Simulation)

Output

Click the Calculate button to perform the calculations and generate the following output. The simulation results will differ slightly for each separate run.

Run Summary Report

Item	Value
Maximum Number of Stages (Design):	5
Current Stage:	0
Alternative Hypothesis:	$P1 - P2 \neq 0$ (two-sided, symmetric)
Alpha Spending Function:	O'Brien-Fleming Analog
Target Alpha:	0.0500
Alpha (from simulations):	0.05167
P1:	0.11
P2:	0.0825
N1 (if final stage reached):	2476
N2 (if final stage reached):	2476
Target Power:	0.9
Power (from simulations):	0.9008
Maximum Information:	14263.1863

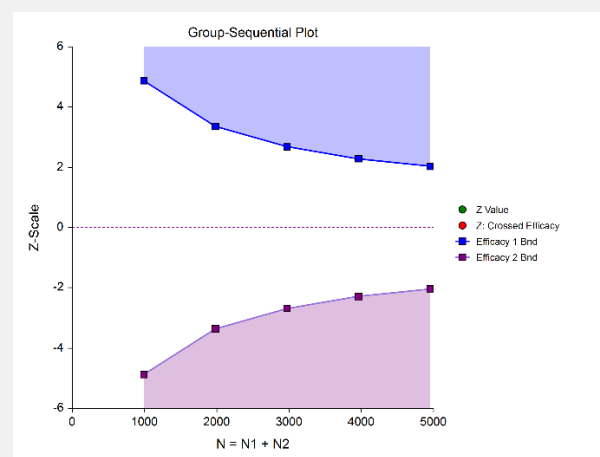
Z-Value Boundaries

Maximum Information: 14263.1863
Alternative Hypothesis: $P1 - P2 \neq 0$ (two-sided, symmetric)

Stage	Boundaries		Information Proportion
	Efficacy 1	Efficacy 2	
1	4.8769	-4.8769	0.2000
2	3.3569	-3.3569	0.4000
3	2.6803	-2.6803	0.6000
4	2.2898	-2.2898	0.8000
5	2.0310	-2.0310	1.0000

The boundary values match those of Reboussin (1992) exactly.

Boundary Plot(s)



Group-Sequential Tests for Two Proportions (Simulation)

Information Report

Maximum Information: 14263.1863

Alternative Hypothesis: $P_1 - P_2 \neq 0$ (two-sided, symmetric)

Alpha: 0.0500

Stage	Target Information Proportion	Target Information	— Target Sample Size —		P1	P2
			N1	N2		
1	0.2000	2852.6373	495.20	495.20	0.1100	0.0825
2	0.4000	5705.2745	990.40	990.40	0.1100	0.0825
3	0.6000	8557.9118	1485.60	1485.60	0.1100	0.0825
4	0.8000	11410.5491	1980.80	1980.80	0.1100	0.0825
5	1.0000	14263.1863	2476.00	2476.00	0.1100	0.0825

Alpha Spending

Target Final Stage Alpha: 0.0500

Spending Function: O'Brien-Fleming Analog

Stage	Information Proportion	Alpha Spent this Stage	Cumulative Alpha Spent	Nominal (Boundary) Alpha	(1-sided) Percentage Alpha Spent this Stage	Cumulative Percentage Alpha Spent
1 *	0.2000	0.0000	0.0000	0.000001	0.0	0.0
2 *	0.4000	0.0008	0.0008	0.000394	1.6	1.6
3 *	0.6000	0.0068	0.0076	0.003678	13.7	15.2
4 *	0.8000	0.0168	0.0244	0.011017	33.6	48.8
5 *	1.0000	0.0256	0.0500	0.021128	51.2	100.0

* projected

Boundary Probabilities for $\delta = 0.0275$

Number of Simulations: 100000

After Efficacy Boundary Crossing: Hold Out

Alternative Hypothesis: $P_1 - P_2 \neq 0$ (two-sided, symmetric)

P1: 0.11

P2: 0.0825

 δ : 0.0275

Stage	N1	N2	Efficacy 1		Efficacy 2	
			Boundary	Probability	Boundary	Probability
1	*495.20	*495.20	4.8769	0.0003	-4.8769	0.0000
2	*990.40	*990.40	3.3569	0.1000	-3.3569	0.0000
3	*1485.60	*1485.60	2.6803	0.3514	-2.6803	0.0000
4	*1980.80	*1980.80	2.2898	0.2966	-2.2898	0.0000
5	*2476.00	*2476.00	2.0310	0.1526	-2.0310	0.0000

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 1832.027064

Average N2: 1832.027064

Group-Sequential Tests for Two Proportions (Simulation)

Boundary Probabilities for $\delta = 0$ (Alpha)

Number of Simulations: 100000
 After Boundary Crossing: Hold Out
 Alternative Hypothesis: $P1 - P2 \neq 0$ (two-sided, symmetric)
 $P1$: 0.0825
 $P2$: 0.0825
 δ : 0

Stage	N1	N2	Efficacy 1		Efficacy 2	
			Boundary	Probability	Boundary	Probability
1	*495.20	*495.20	4.8769	0.0000	-4.8769	0.0000
2	*990.40	*990.40	3.3569	0.0005	-3.3569	0.0004
3	*1485.60	*1485.60	2.6803	0.0036	-2.6803	0.0036
4	*1980.80	*1980.80	2.2898	0.0084	-2.2898	0.0093
5	*2476.00	*2476.00	2.0310	0.0123	-2.0310	0.0136

* Simulation sample size (Non-integer sample sizes were rounded to the next highest integer.)

Average N1: 2458.7918
 Average N2: 2458.7918

Power and Sample Size Summary

Maximum Number of Stages: 5
 Alternative Hypothesis: $P1 - P2 \neq 0$ (two-sided, symmetric)
 Alpha Spending Function: O'Brien-Fleming Analog

Target Power	Sim Power	N1	N2	N	P1	P2	Target Alpha	Sim Alpha
0.90	0.90080	2476	2476	4952	0.1100	0.0825	0.0500	0.0517

The sample sizes of 2476 per group are very close to the values of 2474 per group, calculated by 4-digit rounding.