

## Chapter 220

# Group-Sequential Tests for Two Proportions

---

### Introduction

Clinical trials are longitudinal. They accumulate data sequentially through time. The participants cannot be enrolled and randomized on the same day. Instead, they are enrolled as they enter the study. It may take several years to enroll enough patients to meet sample size requirements. Because clinical trials are long term studies, it is in the interest of both the participants and the researchers to monitor the accumulating information for early convincing evidence of either harm or benefit. This permits early termination of the trial.

Group sequential methods allow statistical tests to be performed on accumulating data while a phase III clinical trial is ongoing. Statistical theory and practical experience with these designs have shown that making four or five *interim analyses* is almost as effective in detecting large differences between treatment groups as performing a new analysis after each new data value. Besides saving time and resources, such a strategy can reduce the experimental subject's exposure to an inferior treatment and make superior treatments available sooner.

When repeated significance testing occurs on the same data, adjustments have to be made to the hypothesis testing procedure to maintain overall significance and power levels. The landmark paper of Lan & DeMets (1983) provided the theory behind the *alpha spending function* approach to group sequential testing. This paper built upon the earlier work of Armitage, McPherson, & Rowe (1969), Pocock (1977), and O'Brien & Fleming (1979). *PASS* implements the methods given in Reboussin, DeMets, Kim, & Lan (1992) to calculate the power and sample sizes of various group sequential designs.

This module calculates sample size and power for group sequential designs used to compare two group proportions. Other modules perform similar analyses for the comparison of means and survival functions. The program allows you to vary the number and times of interim tests, the type of alpha spending function, and the test boundaries. It also gives you complete flexibility in solving for power, significance level, sample size, or effect size. The results are displayed in both numeric reports and informative graphics.

---

### Technical Details

Suppose the means of two samples of  $N1$  and  $N2$  individuals will be compared at various stages of a trial using the  $z_k$  statistic:

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

## Group-Sequential Tests for Two Proportions

The subscript  $k$  indicates that the computations use all data that are available at the time of the  $k^{\text{th}}$  interim analysis or  $k^{\text{th}}$  look ( $k$  goes from 1 to  $K$ ). This formula computes the standard z-test that is assumed to be normally distributed.

### Spending Functions

Lan and DeMets (1983) introduced alpha spending functions,  $\alpha(\tau)$ , that determine a set of boundaries  $b_1, b_2, \dots, b_K$  for the sequence of test statistics  $z_1, z_2, \dots, z_K$ . These boundaries are the critical values of the sequential hypothesis tests. That is, after each interim test, the trial is continued as long as  $|z_k| < b_k$ . When  $|z_k| \geq b_k$ , the hypothesis of equal means is rejected and the trial is stopped early.

The time argument  $\tau$  either represents the proportion of elapsed time to the maximum duration of the trial or the proportion of the sample that has been collected. When elapsed time is being used it is referred to as *calendar time*. When time is measured in terms of the sample, it is referred to as *information time*. Since it is a proportion,  $\tau$  can only vary between zero and one.

Alpha spending functions have the characteristics:

$$\alpha(0) = 0$$

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

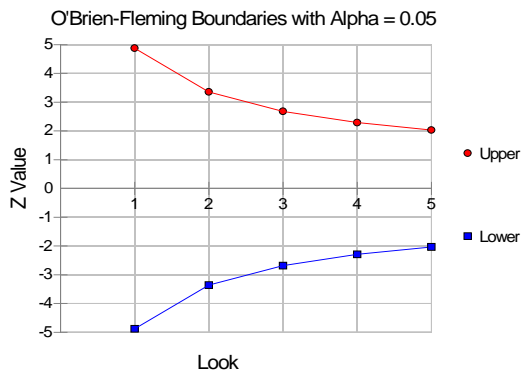
The last characteristic guarantees a fixed  $\alpha$  level when the trial is complete. That is,

$$\Pr(|z_1| \geq b_1 \text{ or } |z_2| \geq b_2 \text{ or } \dots \text{ or } |z_k| \geq b_k) = \alpha(\tau)$$

This methodology is very flexible since neither the times nor the number of analyses must be specified in advance. Only the functional form of  $\alpha(\tau)$  must be specified.

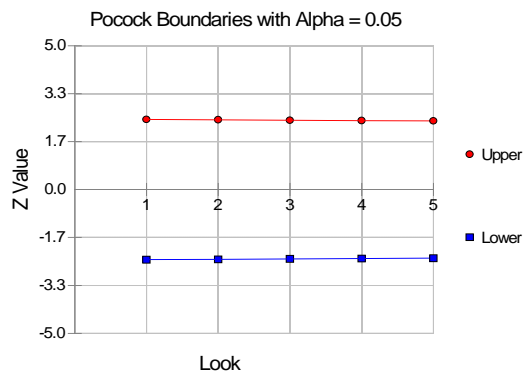
PASS provides five popular spending functions plus the ability to enter and analyze your own boundaries. These are calculated as follows:

- O'Brien-Fleming**  $2 - 2\Phi\left(\frac{Z_{\alpha/2}}{\sqrt{\tau}}\right)$

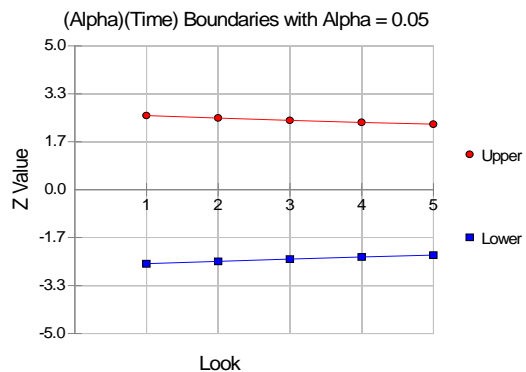


Group-Sequential Tests for Two Proportions

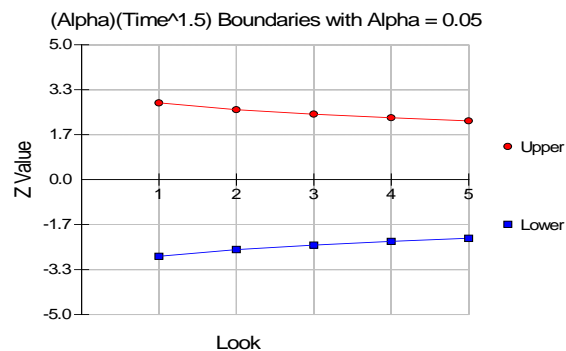
2. Pocock  $\alpha \ln(1 + (e - 1)\tau)$



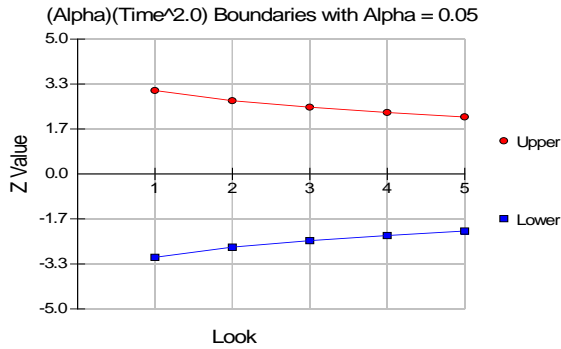
3. Alpha \* time  $\alpha\tau$



4. Alpha \* time^1.5  $\alpha\tau^{3/2}$



## Group-Sequential Tests for Two Proportions

5. Alpha \* time<sup>2</sup>  $\alpha\tau^2$ 

## 6. User Supplied

A custom set of boundaries may be entered.

The O'Brien-Fleming boundaries are commonly used because they do not significantly increase the overall sample size and because they are conservative early in the trial. Conservative in the sense that the proportions must be extremely different before statistical significance is indicated. The Pocock boundaries are nearly equal for all times. The Alpha\*t boundaries use equal amounts of alpha when the looks are equally spaced. You can enter your own set of boundaries using the User Supplied option.

---

## Theory

A detailed account of the methodology is contained in Lan & DeMets (1983), DeMets & Lan (1984), Lan & Zucker (1993), and DeMets & Lan (1994). The theoretical basis of the method will be presented here.

Group sequential procedures for interim analysis are based on their equivalence to discrete boundary crossing of a Brownian motion process with drift parameter  $\theta$ . The test statistics  $z_k$  follow the multivariate normal distribution with means  $\theta\sqrt{\tau_k}$  and, for  $j \leq k$ , covariances  $\sqrt{\tau_k / \tau_j}$ . The drift parameter is related to the parameters of the z-test through the equation

$$\theta = \frac{p_1 - p_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{N1} + \frac{\bar{p}(1-\bar{p})}{N2}}}$$

where

$$\bar{p} = \frac{N1p_1 + N2p_2}{N1 + N2}$$

Hence, the algorithm is as follows:

1. Compute boundary values based on a specified spending function and alpha value.
2. Calculate the drift parameter based on those boundary values and a specified power value.
3. Use the drift parameter and the above equation to calculate the appropriate sample size.

## Group-Sequential Tests for Two Proportions

---

## Procedure Tabs

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

---

### Design Tab

The Design tab contains the parameters associated with the z-test such as the proportions, sample sizes, alpha, and beta.

---

#### Solve For

##### Solve For

This option specifies the parameter to be solved for from the other parameters. The parameters that may be selected are *P1*, *P2*, *Alpha*, *Power*, *Sample Size (N1)* or *Sample Size (N2)*. Under most situations, you will select either *Power* or *Sample Size (N1)*.

Select *Sample Size (N1)* when you want to calculate the sample size needed to achieve a given power and alpha level.

Select *Power* when you want to calculate the power of an experiment.

---

#### Power and Alpha

##### Power

This option specifies one or more values for power. Power is the probability of rejecting a false null hypothesis, and is equal to one minus Beta. Beta is the probability of a type-II error, which occurs when a false null hypothesis is not rejected. In this procedure, a type-II error occurs when you fail to reject the null hypothesis of equal proportions when in fact they are different.

Values must be between zero and one. Historically, the value of 0.80 (Beta = 0.20) was used for power. Now, 0.90 (Beta = 0.10) is also commonly used.

A single value may be entered here or a range of values such as *0.8 to 0.95 by 0.05* may be entered.

##### Alpha

This option specifies one or more values for the probability of a type-I error. A type-I error occurs when a true null hypothesis is rejected. For this procedure, a type-I error occurs when you reject the null hypothesis of equal proportions when in fact they are equal.

Values must be between zero and one. Historically, the value of 0.05 has been used for alpha. This means that about one test in twenty will falsely reject the null hypothesis. You should pick a value for alpha that represents the risk of a type-I error you are willing to take in your experimental situation.

You may enter a range of values such as *0.01 0.05 0.10* or *0.01 to 0.10 by 0.01*.

## Group-Sequential Tests for Two Proportions

### 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.

#### **$N1$ (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.

$N1$  must be  $\geq 2$ . 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.

$N2$  must be  $\geq 2$ . 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  $\geq Y$ .

## Group-Sequential Tests for Two Proportions

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 Not Solving for Sample Size)

### 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  $N2$ .
- **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. 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  $\geq 2$ . 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.

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

## Group-Sequential Tests for Two Proportions

### 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.

$N2$  must be  $\geq 2$ . 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.

### 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 (Proportion in Group 1)

Enter value(s) for the response proportion in the first group under both hypotheses and the response proportion of the second group under the null hypothesis of equal proportions. The values must be between zero and one.

You may enter a range of values such as  $0.1$ ,  $0.2$ ,  $0.3$  or  $0.1$  to  $0.9$  by  $0.2$ .

### P2 (Proportion in Group 2)

Enter value(s) for the response proportion of the second group under the alternative hypothesis. You may enter a range of values such as  $0.1$ ,  $0.2$ ,  $0.3$  or  $0.1$  to  $0.9$  by  $0.2$ .



## Group-Sequential Tests for Two Proportions

### Look Details

This box contains the parameters associated with Group Sequential Design such as the type of spending function, the times, and so on.

#### Number of Looks

This is the number of interim analyses (including the final analysis). For example, a five here means that four interim analyses will be run in addition to the final analysis.

#### Boundary Truncation

You can truncate the boundary values at a specified value. For example, you might decide that no boundaries should be larger than 4.0. If you want to implement a boundary limit, enter the value here.

If you do not want a boundary limit, enter *None* here.

#### Spending Function

Specify which alpha spending function to use. The most popular is the O'Brien-Fleming boundary that makes early tests very conservative. Select *User Specified* if you want to enter your own set of boundaries.

#### Max Time

This is the total running time of the trial. It is used to convert the values in the Times box to fractions. The units (months or years) do not matter, as long as they are consistent with those entered in the Times box.

For example, suppose Max Time = 3 and Times = 1, 2, 3. Interim analyses would be assumed to have occurred at 0.33, 0.67, and 1.00.

#### Times

Enter a list of time values here at which the interim analyses will occur. These values are scaled according to the value of the Max Time option.

For example, suppose a 48-month trial calls for interim analyses at 12, 24, 36, and 48 months. You could set Max Time to 48 and enter *12,24,36,48* here or you could set Max Time to *1.0* and enter *0.25,0.50,0.75,1.00* here.

The number of times entered here must match the value of the Number of Looks.

- **Equally Spaced**

If you are planning to conduct the interim analyses at equally spaced points in time, you can enter *Equally Spaced* and the program will generate the appropriate time values for you.

#### Informations

You can weight the interim analyses on the amount of information obtained at each time point rather than on actual calendar time. If you would like to do this, enter the information amounts here. Usually, these values are the sample sizes obtained up to the time of the analysis.

For example, you might enter *50, 76, 103, 150* to indicate that 50 individuals were included in the first interim analysis, 76 in the second, and so on.

#### Upper and Lower Boundaries (Spending = User)

If the Spending Function is set to *User Supplied* you can enter a set of lower test boundaries, one for each interim analysis. The lower boundaries should be negative and the upper boundaries should be positive. Typical entries are *4,3,3,3,2* and *4,3,2,2,2*.

- **Symmetric**

If you only want to enter the upper boundaries and have them copied with a change in sign to the lower boundaries, enter *Symmetric* for the lower boundaries.

## Group-Sequential Tests for Two Proportions

### Test

#### Alternative Hypothesis

Specify whether the test is one-sided or two-sided. When a two-sided hypothesis is selected, the value of alpha is halved. Everything else remains the same.

Note that the accepted procedure is to use Two-Sided unless you can justify using a one-sided test.

#### Continuity Correction

Specify whether to use the Continuity Correction. This option applies an adjustment to the sample sizes that is recommend by Fleiss(1981) page 45 to make the alpha and beta values more accurate. The formula for the adjustment is

$$N1_{new} = \frac{N1_{old}}{4} \left( 1 + \sqrt{1 + \frac{2(R+1)}{R(N1_{old})|P1-P2|}} \right)^2$$

### Options Tab

The Options tab controls the convergence of the various iterative algorithms used in the calculations.

#### Maximum Iterations

##### Maximum Iterations (Lan-Demets Algorithm)

This is the maximum number of iterations used in the Lan-DeMets algorithm during its search routine. We recommend a value of at least 200.

#### Tolerance

##### Probability Tolerance

During the calculation of the probabilities associated with a set of boundary values, probabilities less than this are assumed to be zero.

We suggest a value of 0.0000000001.

##### Power Tolerance

This is the convergence level for the search for the spending function values that achieve a certain power. Once the iteration changes are less than this amount, convergence is assumed. We suggest a value of 0.0000001.

If the search is too time consuming, you might try increasing this value.

##### Alpha Tolerance

This is the convergence level for the search for a given alpha value. Once the changes in the computed alpha value are less than this amount, convergence is assumed and iterations stop. We suggest a value of 0.0001.

This option is only used when you are searching for alpha.

If the search is too time consuming, you can try increasing this value.

## Group-Sequential Tests for Two Proportions

**Example 1 – Finding the Sample Size**

A clinical trial is to be conducted over a two-year period to compare the proportion response of a new treatment to that of the current treatment. The current response proportion is 0.53. Although the researchers do not know the true proportion of patients that will survive with the new treatment, they would like to examine the power that is achieved if the proportion under the new treatment is 0.63. In order to compare the sample size requirements for different effect sizes, it is also of interest to compute the sample size at response rates of 0.60, 0.65, 0.70, and 0.75.

Testing will be done at the 0.05 significance level and the power should be set to 0.90. A total of four tests are going to be performed on the data as they are obtained. The O'Brien-Fleming boundaries will be used.

Find the necessary sample sizes and test boundaries assuming equal sample sizes per arm and two-sided hypothesis tests.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Group-Sequential Tests for Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Tests for Two Proportions**. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Continuity Correction .....	<b>Checked</b>
Power .....	<b>0.90</b>
Alpha .....	<b>0.05</b>
Group Allocation .....	<b>Equal (N1 = N2)</b>
P1 (Proportion in Group 1) .....	<b>0.53</b>
P2 (Proportion in Group 2) .....	<b>0.60 0.63 0.65 0.70 0.75</b>
Number of Looks .....	<b>4</b>
Spending Function .....	<b>O'Brien-Fleming</b>
Boundary Truncation .....	<b>None</b>
Max Time .....	<b>2</b>
Times .....	<b>Equally Spaced</b>
Informations .....	<i>Blank</i>

## Group-Sequential Tests for Two Proportions

## Annotated Output

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

## Numeric Results

## Numeric Results for Two-Sided Test of Proportions. Continuity Correction Applied.

Target Power	Actual Power	N1	N2	N	P1	P2	Alpha
0.90	0.900168	1102	1102	2204	0.53	0.60	0.050
0.90	0.900930	542	542	1084	0.53	0.63	0.050
0.90	0.900408	376	376	752	0.53	0.65	0.050
0.90	0.901093	187	187	374	0.53	0.70	0.050
0.90	0.903126	111	111	222	0.53	0.75	0.050

## Report Definitions

Target Power is the desired power value (or values) entered in the procedure. Power is the probability of rejecting a false null hypothesis.

Actual Power is the power obtained in this scenario. Because N1 and N2 are discrete, this value is often (slightly) larger than the target power.

N1 and N2 are the number of items sampled from each population.

N is the total sample size, N1 + N2.

P1 is the proportion of populations 1 and 2 under the null hypothesis of equality.

P2 is the proportion of population 2 at which power and sample size calculations are made.

Alpha is the probability of rejecting a true null hypothesis.

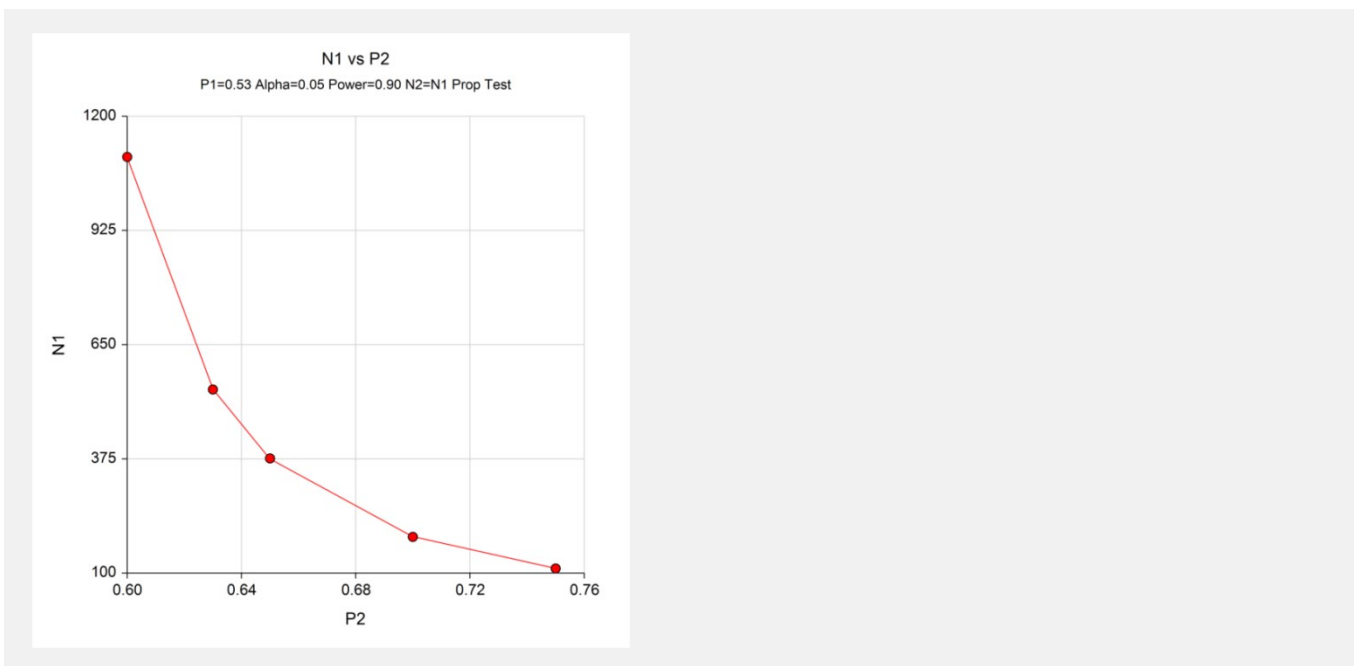
## Summary Statements

Sample sizes of 1102 and 1102 achieve 90% power to detect a difference of 0.07 between the group proportions of 0.53 and 0.60 at a significance level (alpha) of 0.0500 using a two-sided z-test with continuity correction. These results assume that 4 sequential tests are made using the O'Brien-Fleming spending function to determine the test boundaries.

This report shows the values of each of the parameters, one scenario per row. Note that 542 participants in each arm of the study are required to meet the 90% power requirement when the proportion is 0.63.

The values from this table are in the chart below. Note that this plot actually occurs further down in the report.

## Plots Section



This plot shows that a large increase in sample size is necessary when the detectable proportion in group two is less than 0.63.

## Group-Sequential Tests for Two Proportions

## Details Section

Details when Spending = O'Brien-Fleming, N1 = 542, N2 = 542, P1 = 0.53, P2 = 0.63, Continuity Correction.

Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.50	-4.33263	4.33263	0.000015	0.000015	0.000015	0.003525	0.003525
2	1.00	-2.96311	2.96311	0.003045	0.003036	0.003051	0.255573	0.259098
3	1.50	-2.35902	2.35902	0.018323	0.016248	0.019299	0.427801	0.686899
4	2.00	-2.01406	2.01406	0.044003	0.030701	0.050000	0.214031	0.900930
Drift	3.27640							

This report shows information about the individual interim tests. One report is generated for each scenario.

**Look**

These are the sequence numbers of the interim tests.

**Time**

These are the time points at which the interim tests are conducted. Since the Max Time was set to 2 (for two years), these time values are in years. Hence, the first interim test is at half a year, the second at one year, and so on.

We could have set Max Time to 24 so that the time scale was in months.

**Lower and Upper Boundary**

These are the test boundaries. If the computed value of the test statistic  $z$  is between these values, the trial should continue. Otherwise, the trial can be stopped.

**Nominal Alpha**

This is the value of alpha for these boundaries if they were used for a single, standalone, test. Hence, this is the significance level that must be found for this look in a standard statistical package that does not adjust for multiple looks.

**Inc Alpha**

This is the amount of alpha that is *spent* by this interim test. It is close to, but not equal to, the value of alpha that would be achieved if only a single test was conducted. For example, if we lookup the third value, 2.35902, in normal probability tables, we find that this corresponds to a (two-sided) alpha of 0.018323. However, the entry is 0.016248. The difference is due to the correction that must be made for multiple tests.

**Total Alpha**

This is the total amount of alpha that is used up to and including the current test.

**Inc Power**

These are the amounts that are added to the total power at each interim test. They are often called the exit probabilities because they give the probability that significance is found and the trial is stopped, given the alternative hypothesis.

**Total Power**

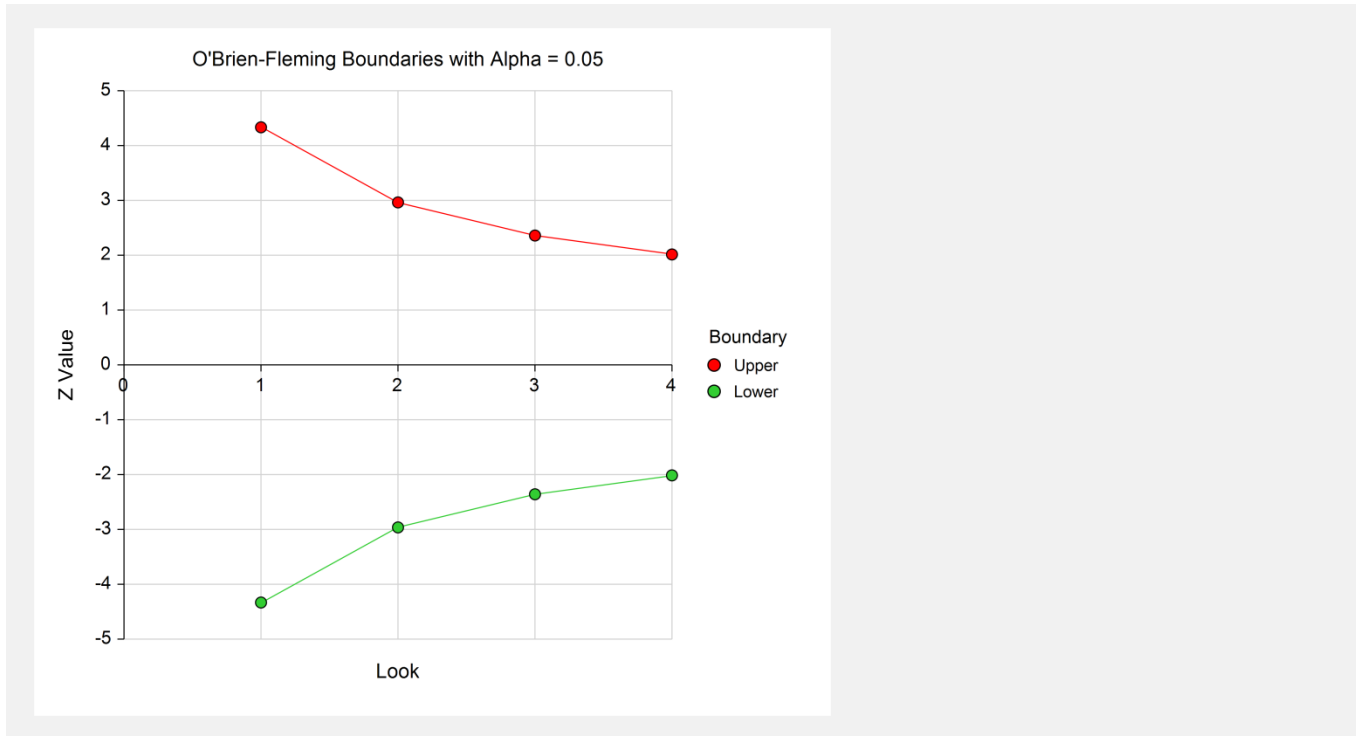
These are the cumulative power values. They are also the cumulative exit probabilities. That is, they are the probability that the trial is stopped at or before the corresponding time.

**Drift**

This is the value of the Brownian motion drift parameter.

## Group-Sequential Tests for Two Proportions

## Boundary Plots



This plot shows the interim boundaries for each look. This plot shows very dramatically that the results must be extremely significant at early looks, but that they are near the single test boundary (1.96 and -1.96) at the last look.

## Group-Sequential Tests for Two Proportions

**Example 2 – Finding the Power**

Continuing the scenario began in Example 1, the researcher wishes to calculate the power of the design at sample sizes 200, 400, 600, 800, 1000. Testing will be done at the 0.01, 0.05, 0.10 significance levels and the overall power will be set to 0.10. Find the power of these sample sizes and test boundaries assuming equal sample sizes per arm and two-sided hypothesis tests.

Proceeding as in Example 1, we decide to translate the mean and standard deviation into a percent of mean scale.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Group-Sequential Tests for Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Tests for Two Proportions**. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Continuity Correction .....	<b>Checked</b>
Alpha .....	<b>0.01 0.05 0.10</b>
Group Allocation .....	<b>Equal (N1 = N2)</b>
Sample Size Per Group .....	<b>200 to 1000 by 200</b>
P1 (Proportion in Group 1) .....	<b>0.53</b>
P2 (Proportion in Group 2) .....	<b>0.63</b>
Number of Looks .....	<b>4</b>
Spending Function .....	<b>O'Brien-Fleming</b>
Boundary Truncation .....	<b>None</b>
Max Time .....	<b>2</b>
Times .....	<b>Equally Spaced</b>
Informations .....	<i>Blank</i>

## Group-Sequential Tests for Two Proportions

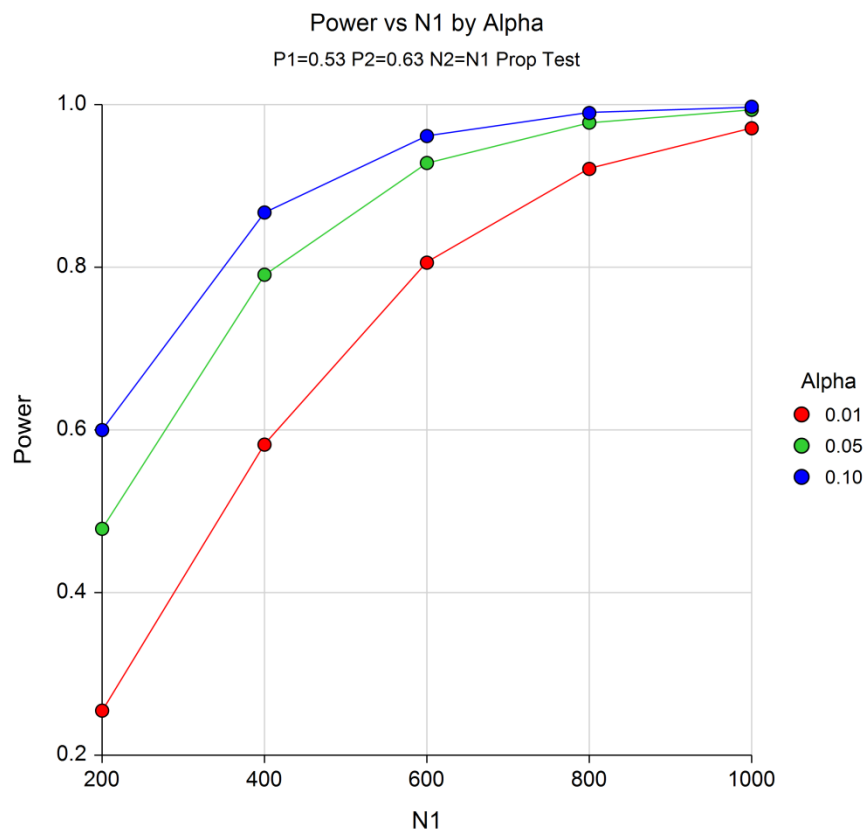
## Output

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

## Numeric Results and Plots

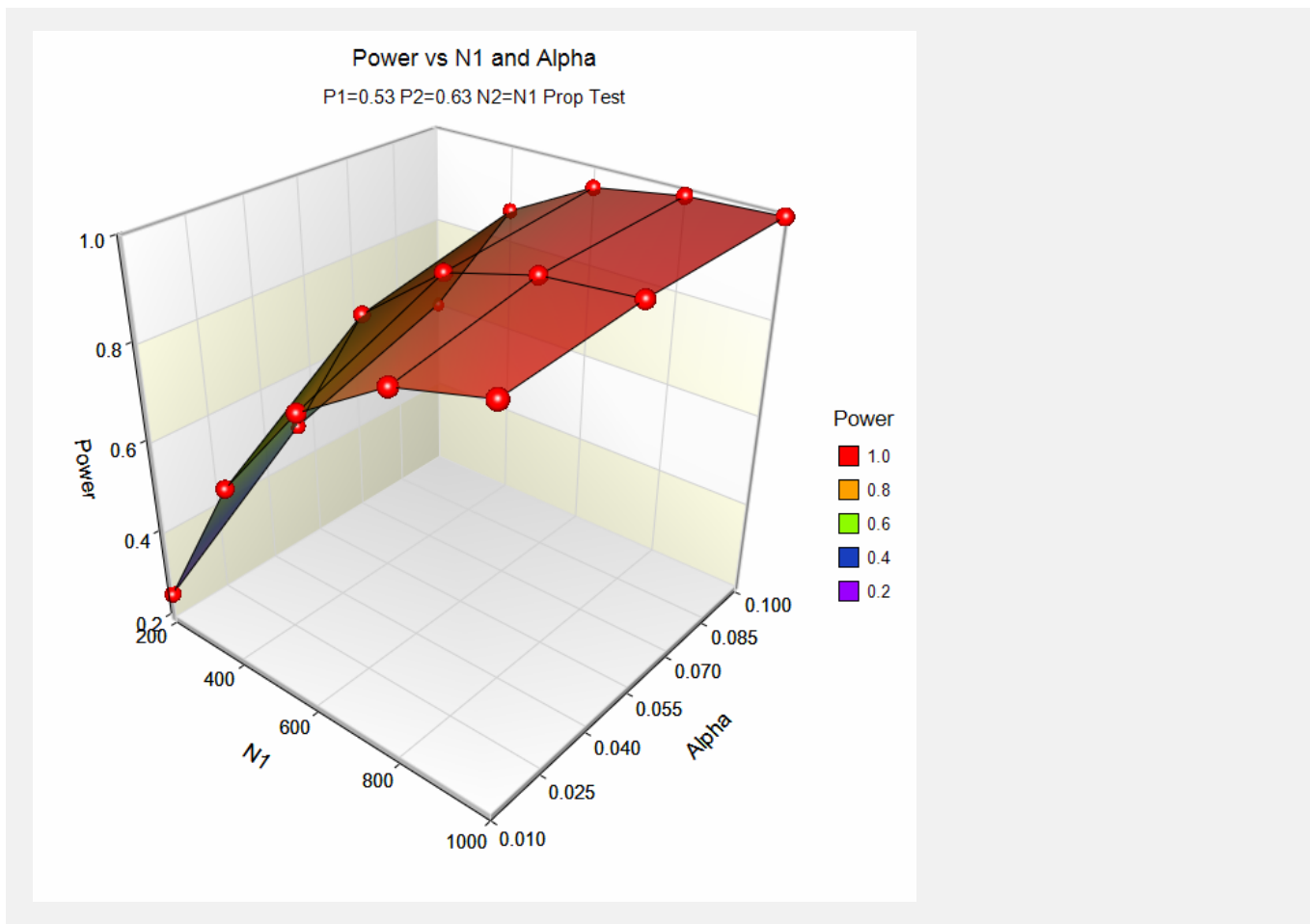
Numeric Results for Two-Sided Test of Proportions. Continuity Correction Applied.

Power	N1	N2	N	P1	P2	Alpha
0.254797	200	200	400	0.53	0.63	0.010
0.581920	400	400	800	0.53	0.63	0.010
0.805679	600	600	1200	0.53	0.63	0.010
0.920951	800	800	1600	0.53	0.63	0.010
0.970898	1000	1000	2000	0.53	0.63	0.010
0.478378	200	200	400	0.53	0.63	0.050
0.790815	400	400	800	0.53	0.63	0.050
0.928267	600	600	1200	0.53	0.63	0.050
0.977859	800	800	1600	0.53	0.63	0.050
0.993673	1000	1000	2000	0.53	0.63	0.050
0.599827	200	200	400	0.53	0.63	0.100
0.867330	400	400	800	0.53	0.63	0.100
0.961331	600	600	1200	0.53	0.63	0.100
0.989659	800	800	1600	0.53	0.63	0.100
0.997396	1000	1000	2000	0.53	0.63	0.100





Group-Sequential Tests for Two Proportions



These data show the power for various sample sizes and alphas. It is interesting to note that once the sample size is greater than 700, the value of alpha makes little difference on the value of power.

## Group-Sequential Tests for Two Proportions

**Example 3 – Effect of Number of Looks**

Continuing with examples one and two, it is interesting to determine the impact of the number of looks on power. *PASS* allows only one value for the Number of Looks parameter per run, so it will be necessary to run several analyses. To conduct this study, set alpha to 0.05, N1 to 500, and leave the other parameters as before. Run the analysis with Number of Looks equal to 1, 2, 3, 4, 6, 8, 10, and 20. Record the power for each run.

**Setup**

This section presents the values of each of the parameters needed to run this example. First, from the *PASS* Home window, load the **Group-Sequential Tests for Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Tests for Two Proportions**. You may then 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>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Continuity Correction .....	<b>Not checked</b>
Alpha .....	<b>0.05</b>
Group Allocation .....	<b>Equal (N1 = N2)</b>
Sample Size Per Group .....	<b>500</b>
P1 (Proportion in Group 1) .....	<b>0.53</b>
P2 (Proportion in Group 2) .....	<b>0.63</b>
Number of Looks .....	<b>1 (Also run with 2, 3, 4, 6, 8, 10, and 20)</b>
Spending Function .....	<b>O'Brien-Fleming</b>
Boundary Truncation .....	<b>None</b>
Max Time .....	<b>2</b>
Times .....	<b>Equally Spaced</b>
Informations .....	<i>Blank</i>

**Output**

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

**Numeric Results**

Numeric Results for Two-Sided Test of Proportions							
Power	N1	N2	N	P1	P2	Alpha	Looks
0.893174	500	500	1000	0.53	0.63	0.050	1
0.892118	500	500	1000	0.53	0.63	0.050	2
0.889620	500	500	1000	0.53	0.63	0.050	3
0.887691	500	500	1000	0.53	0.63	0.050	4
0.885125	500	500	1000	0.53	0.63	0.050	6
0.883535	500	500	1000	0.53	0.63	0.050	8
0.882456	500	500	1000	0.53	0.63	0.050	10
0.879929	500	500	1000	0.53	0.63	0.050	20

This analysis shows how little the number of looks impacts the power of the design. The power of a study with no interim looks is 0.893174. When twenty interim looks are made, the power falls to 0.879929—a very small change.

## Example 4 – Studying a Boundary Set

Continuing with the previous examples, suppose that you are presented with a set of boundaries and want to find the quality of the design (as measured by alpha and power). This is easy to do with **PASS**. Suppose that the analysis is to be run with five interim looks at equally spaced time points. The upper boundaries to be studied are 3.5, 3.5, 3.0, 2.5, 2.0. The lower boundaries are symmetric. The analysis would be run as follows.

### Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Group-Sequential Tests for Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Tests for Two Proportions**. You may then 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>
<b>Design Tab</b>	
Solve For .....	<b>Power</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Continuity Correction .....	<b>Not checked</b>
Alpha.....	<b>0.05 (will be calculated from boundaries)</b>
Group Allocation .....	<b>Equal (N1 = N2)</b>
Sample Size Per Group.....	<b>500</b>
P1 (Proportion in Group 1) .....	<b>0.53</b>
P2 (Proportion in Group 2) .....	<b>0.63</b>
Number of Looks .....	<b>5</b>
Spending Function.....	<b>User Supplied</b>
Boundary Truncation .....	<b>None</b>
Max Time .....	<b>2</b>
Times .....	<b>Equally Spaced</b>
Informations .....	<i>Blank</i>
Upper Boundaries.....	<b>3.5 3.5 3.0 2.5 2.0</b>
Lower Boundaries.....	<b>Symmetric</b>

### Output

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

#### Numeric Results

Numeric Results for Two-Sided Test of Proportions								
Power	N1	N2	N	P1	P2	Alpha		
0.887792	500	500	1000	0.53	0.63	0.048157		
Details when Spending = User Supplied, N1 = 500, N2 =500, P1 = 0.53, P2 = 0.63								
Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.40	-3.50000	3.50000	0.000465	0.000465	0.000465	0.019352	0.019352
2	0.80	-3.50000	3.50000	0.000465	0.000408	0.000874	0.058108	0.077460
3	1.20	-3.00000	3.00000	0.002700	0.002410	0.003284	0.230567	0.308026
4	1.60	-2.50000	2.50000	0.012419	0.010331	0.013615	0.339341	0.647367
5	2.00	-2.00000	2.00000	0.045500	0.034542	0.048157	0.240425	0.887792
Drift	3.20355							

## Group-Sequential Tests for Two Proportions

The power for this design is about 0.89. This value depends on both the boundaries and the sample size. The alpha level is about 0.048. This value only depends on the boundaries.

---

### Example 5 – Validation using O’Brien-Fleming Boundaries

Reboussin (1992) presents an example for binomial distributed data for a design with two-sided O’Brien-Fleming boundaries, looks = 5, alpha = 0.05, beta = 0.10, P1 = 0.1100, P2 = 0.0825. They compute a drift of 3.28 and a sample size of 2381.78 per group. The upper boundaries are: 4.8769, 3.3569, 2.6803, 2.2898, 2.0310.

To test that PASS provides the same result, 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 **Group-Sequential Tests for Two Proportions** procedure window by expanding **Proportions**, then **Two Independent Proportions**, then clicking on **Group-Sequential**, and then clicking on **Group-Sequential Tests for Two Proportions**. You may then 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>
<b>Design Tab</b>	
Solve For .....	<b>Sample Size</b>
Alternative Hypothesis .....	<b>Two-Sided</b>
Continuity Correction .....	<b>Not checked</b>
Power .....	<b>0.90</b>
Alpha .....	<b>0.05</b>
Group Allocation .....	<b>Equal (N1 = N2)</b>
P1 (Proportion in Group 1) .....	<b>0.1100</b>
P2 (Proportion in Group 2) .....	<b>0.0825</b>
Number of Looks .....	<b>5</b>
Spending Function .....	<b>O’Brien-Fleming</b>
Boundary Truncation .....	<b>None</b>
Max Time .....	<b>1</b>
Times .....	<b>Equally Spaced</b>
Informations .....	<i>Blank</i>

## Group-Sequential Tests for Two Proportions

## Output

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

## Numeric Results

Numeric Results for Two-Sided Test of Proportions								
Target Power	Actual Power	N1	N2	N	P1	P2	Alpha	
0.90	0.90011	2474	2474	4948	0.1100	0.0825	0.050000	
Details when Spending = O'Brien-Fleming, N1 = 2468, N2 = 2468, P1 = 0.1100, P2 = 0.0825								
Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.20	-4.87688	4.87688	0.000001	0.000001	0.000001	0.000324	0.000324
2	0.40	-3.35695	3.35695	0.000788	0.000787	0.000788	0.099454	0.099778
3	0.60	-2.68026	2.68026	0.007357	0.006828	0.007616	0.346699	0.446477
4	0.80	-2.28979	2.28979	0.022034	0.016807	0.024424	0.299644	0.746120
5	1.00	-2.03100	2.03100	0.042255	0.025576	0.050000	0.153985	0.900105
Drift	3.27939							

The difference in the sample sizes (2474 versus 2382) is due to rounding errors in the Reboussin article. Reboussin rounds from four-digits to three-digits, which caused a large difference. **PASS** uses more accurate routines.

To see that the results are equal to within rounding error, we will compute the sample size using Reboussin's results, but with more decimal places in the intermediate steps. They had

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

When we compute this without rounding, we get

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

A sample size of 2474 is the result obtained in **PASS**.