Chapter 475

Group-Sequential Tests for Two Means (Legacy)

This procedure is the original two-means group-sequential procedure in PASS. Power calculations and boundaries are generated from analytic calculations (simulation is not used). This procedure does not give any options for futility boundaries.

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 treatment means. Other modules perform similar analyses for the comparison of proportions 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:

$$z_k = \frac{\bar{X}_{1k} - \bar{X}_{2k}}{\sqrt{\frac{s_{1k}^2}{N_{1k}} + \frac{s_{2k}^2}{N_{2k}}}}$$

The subscript *k* indicates that the computations use all data that are available at the time of the k^{th} interim analysis or k^{th} *look* (*k* goes from 1 to *K*). This formula computes the standard *z* test that is appropriate when the variances of the two groups are different. The statistic, z_k, is assumed to be normally distributed.

Spending Functions

Lan and DeMets (1983) introduced alpha spending functions, $\alpha(\tau)$, that determine a set of boundaries $b_1, b_2, \ldots, b_K$ for the sequence of test statistics $z_1, z_2, \ldots, z_K$. 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 τ 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, τ can only vary between zero and one.

Alpha spending functions have the characteristics:

$$\alpha(0) = 0$$

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

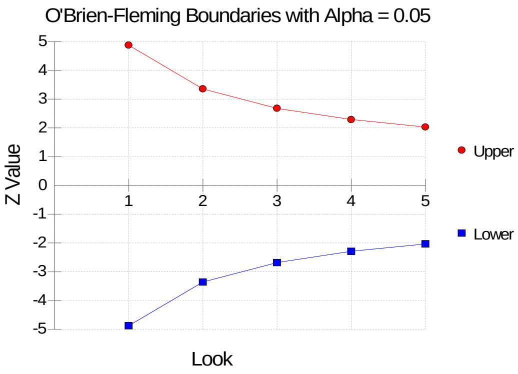
The last characteristic guarantees a fixed α level when the trial is complete. That is,

$$Pr(|z_1| \geq b_1 \; or \; |z_2| \geq b_2 \; or \ldots or \; |z_k| \geq b_k) = \alpha(\tau)$$

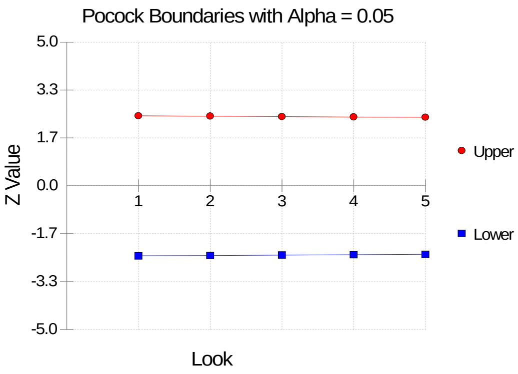
This methodology is very flexible since neither the times nor the number of analyses must be specified in advance. Only the functional form of $\alpha(\tau)$ must be specified.

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

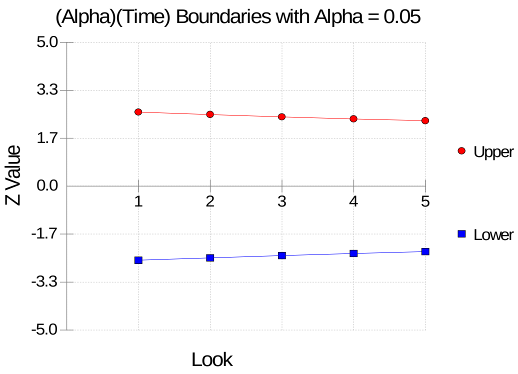
1. **O'Brien-Fleming** $2-2\Phi\left(\frac{Z_{\alpha/2}}{\sqrt{t}}\right)$



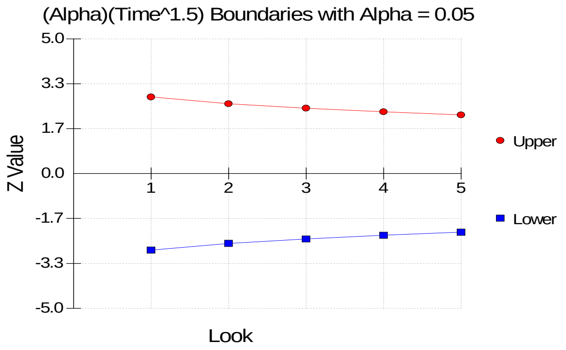
2. **Pocock** $\alpha \cdot \ln(1+(e-1)t)$



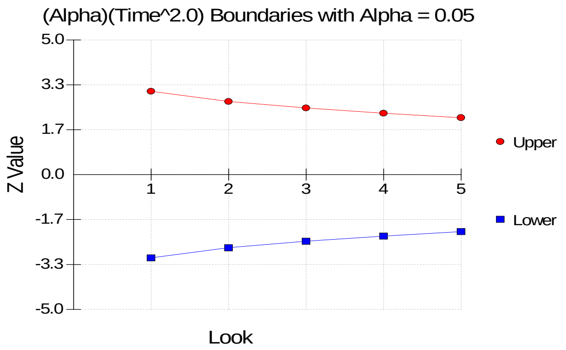
3. **Alpha * time** $\alpha \cdot t$



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



5. Alpha * time^2 $\alpha \cdot t^2$



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 means 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 θ. 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

$$\theta = \frac{\mu_1 - \mu_2}{\sqrt{\frac{\sigma_1^2}{N_1} + \frac{\sigma_2^2}{N_2}}}$$

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 estimates of the other parameters in the above equation to calculate the appropriate sample size.

Example 1 – Finding the Sample Size

A clinical trial is to be conducted over a two-year period to compare the mean response of a new treatment with the current treatment. The current mean is 127 with a standard deviation of 55.88. The health community will be interested in the new treatment if the mean response rate is increased by 20%. So that the sample size requirements for different effect sizes can be compared, it is also of interest to compute the sample size at 10%, 30%, 40%, 50%, 60%, and 70% increases in the response rates.

Testing will be done at the 0.05 significance level and the power should be set to 0.10. 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.

We could enter these amounts directly into the Group Sequential Means window. Since the base mean is 127, a 20% increase would translate to a new mean response of 127(120/100) = 152.4. The other mean response rates could be computed similarly. However, to make the results more meaningful, we will scale the input by dividing by the current mean. The scaled standard deviation will be 100(55.88)/127 = 44.00. We set Mean1 to zero since we are only interested in the changes in *Mean2*. The values of *Mean2* will then be 10, 20, 30, 40, 50, 60, and 70.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Sample Size**
Alternative Hypothesis .... **Two-Sided**
Power.. **0.90**
Alpha... **0.05**
Group Allocation .... **Equal (N1 = N2)**
Mean1 (Mean of Group 1)............................... **0**
Mean2 (Mean of Group 2)............................... **10 to 70 by 10**
S1 (Standard Deviation Group 1).................... **44**
S2 (Standard Deviation Group 2).................... **S1**
Number of Looks **4**
Spending Function... **O'Brien-Fleming**
Boundary Truncation...................................... **None**
Max Time **2**
Times... **Equally Spaced**
Informations .... ***Blank***

Output

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

Numeric Reports

Numeric Results for Two-Sided Hypothesis Test of Means

Solve For: Sample Size

Target Power	Actual Power	N1	N2	N	Mean1	Mean2	S1	S2	Alpha
0.9	0.90048	415	415	830	0	10	44	44	0.05
0.9	0.90117	104	104	208	0	20	44	44	0.05
0.9	0.90584	47	47	94	0	30	44	44	0.05
0.9	0.90117	26	26	52	0	40	44	44	0.05
0.9	0.90714	17	17	34	0	50	44	44	0.05
0.9	0.91156	12	12	24	0	60	44	44	0.05
0.9	0.91695	9	9	18	0	70	44	44	0.05

Target Power The desired power value (or values) entered in the procedure. Power is the probability of rejecting a false null hypothesis.
Actual Power 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 The number of items sampled from each population.
N The total sample size. N = N1 + N2.
Mean1 The mean of populations 1 and 2 under the null hypothesis of equality.
Mean2 The mean of population 2 under the alternative hypothesis.
S1 and S2 The population standard deviations of groups 1 and 2.
Alpha The probability of rejecting a true null hypothesis.

Summary Statements

Sample sizes of 415 and 415 achieve 90% power to detect a mean difference of 10 with standard deviations of 44 and 44 with an overall Type I error rate (α) of 0.05 using a two-sided z-test. These results assume 4 sequential tests are made if the final stage is reached. The O'Brien-Fleming spending function was used to determine the test boundaries.

This report shows the values of each of the parameters, one scenario per row. Note that 104 participants in each arm of the study are required to meet the 90% power requirement when the mean increase is 20%.

The values from this table are in the chart below. Note that this plot actually is found farther down in the report.

Details Section

Details when Spending = O'Brien-Fleming, N1 = 415, N2 = 415, S1 = 44, S2 = 44, Diff = -10

Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.5	-4.33263	4.33263	0.000	0.000	0.000	0.00351	0.004
2	1.0	-2.96311	2.96311	0.003	0.003	0.003	0.25500	0.259
3	1.5	-2.35902	2.35902	0.018	0.016	0.019	0.42760	0.686
4	2.0	-2.01406	2.01406	0.044	0.031	0.050	0.21437	0.900

Drift = 3.27383

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 0.0183. However, the entry is 0.0162. 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

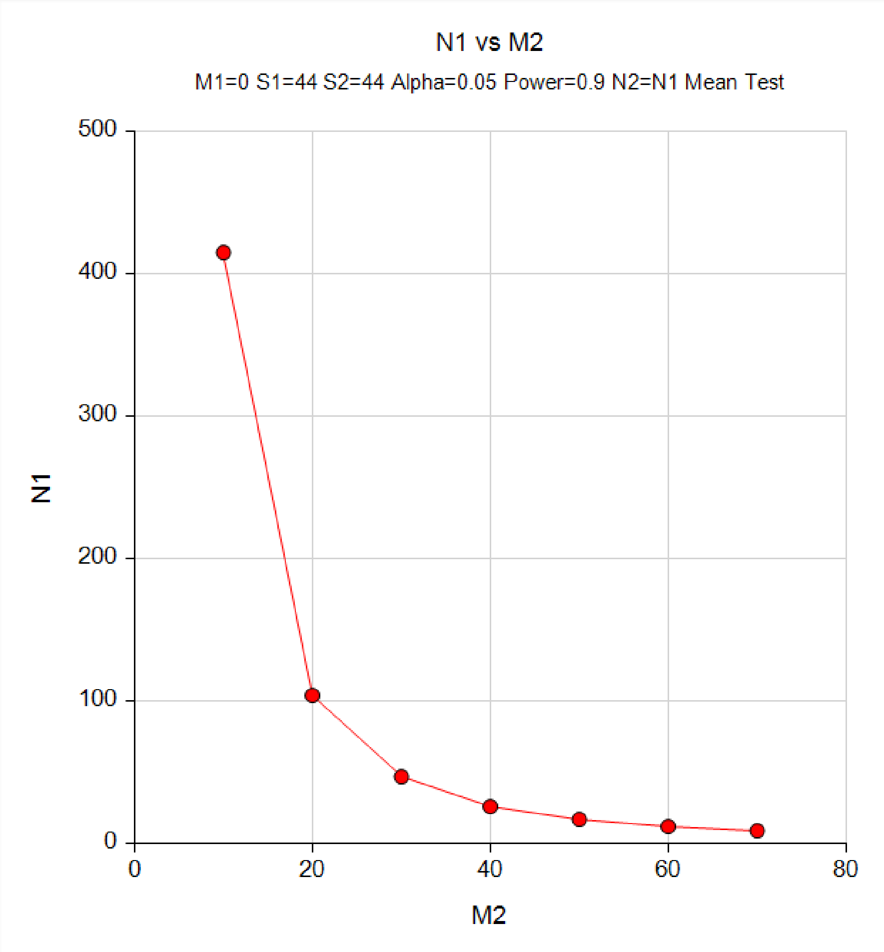
These are the cumulative power values. They are also the cumulative exit probabilities. That is, they are the probability that the trial is stopped at or before the corresponding time.

Drift

This is the value of the Brownian motion drift parameter.

Plots Section

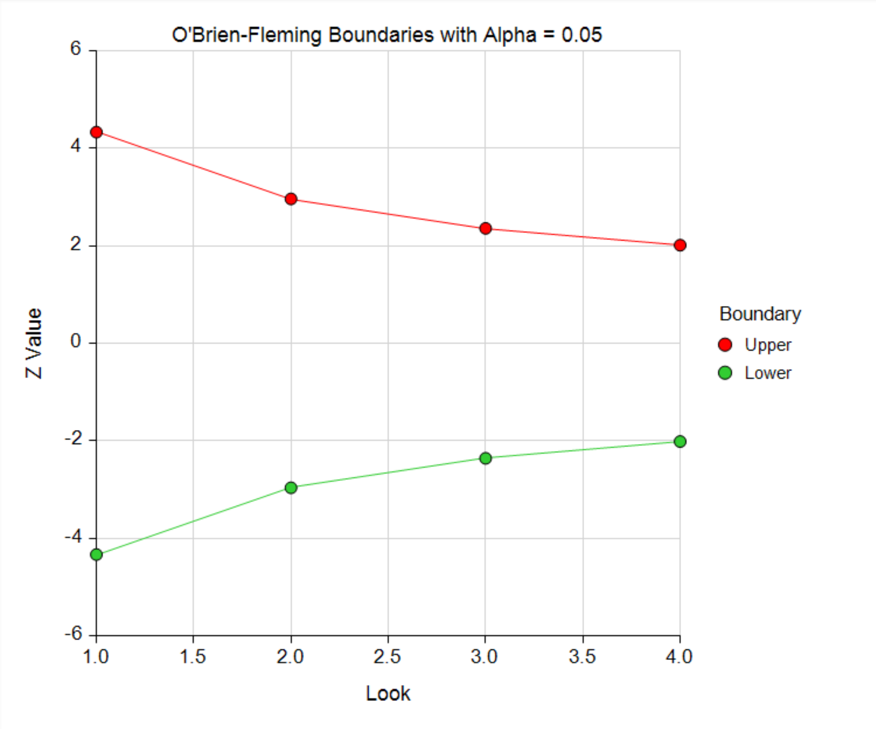
Plots



This plot shows that a large increase in sample size is necessary to test mean differences below 20%.

Boundary Plots

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.

Example 2 – Finding the Power

A clinical trial is to be conducted over a two-year period to compare the mean response of a new treatment with the current treatment. The current mean is 127 with a standard deviation of 55.88. The health community will be interested in the new treatment if the mean response rate is increased by 20%. The researcher wishes to calculate the power of the design at sample sizes 20, 60, 100, 140, 180, and 220. Testing will be done at the 0.01, 0.05, 0.10 significance levels and the overall power will be set to 0.10. 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 power of these sample sizes and test boundaries assuming equal sample sizes per arm and two-sided hypothesis tests.

Proceeding as in Example1, we decide to translate the mean and standard deviation into a percent of mean scale.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Power**
Alternative Hypothesis .... **Two-Sided**
Alpha.. **0.01 0.05 0.10**
Group Allocation .... **Equal (N1 = N2)**
Sample Size Per Group .... **20 to 220 by 40**
Mean1 (Mean of Group 1).............................. **0**
Mean2 (Mean of Group 2).............................. **20**
S1 (Standard Deviation Group 1).................. **44**
S2 (Standard Deviation Group 2).................. **S1**
Number of Looks **4**
Spending Function.. **O'Brien-Fleming**
Boundary Truncation..................................... **None**
Max Time **2**
Times .... **Equally Spaced**
Informations .... ***Blank***

Output

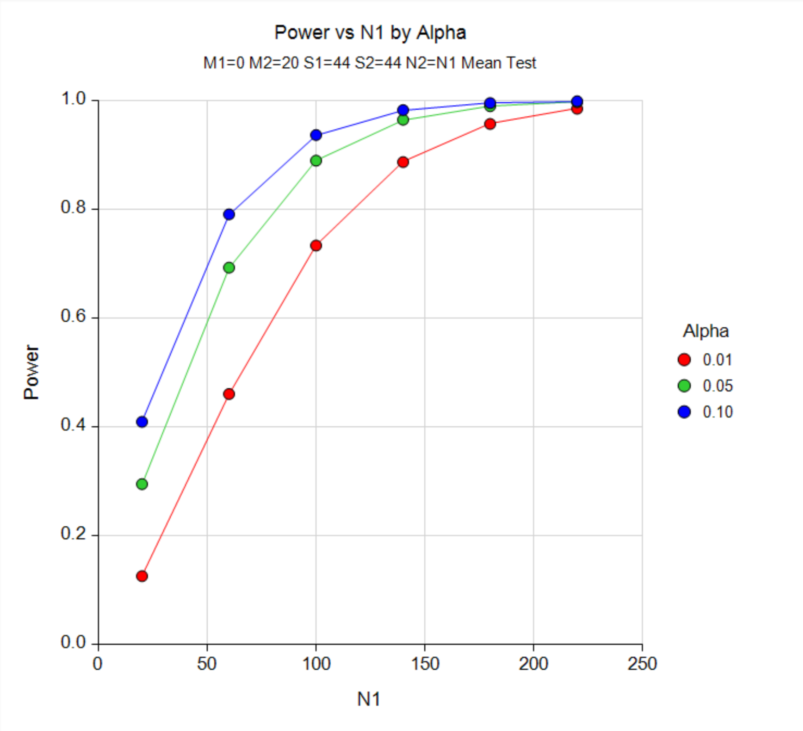
Click the Calculate button to perform the calculations and generate the following output.

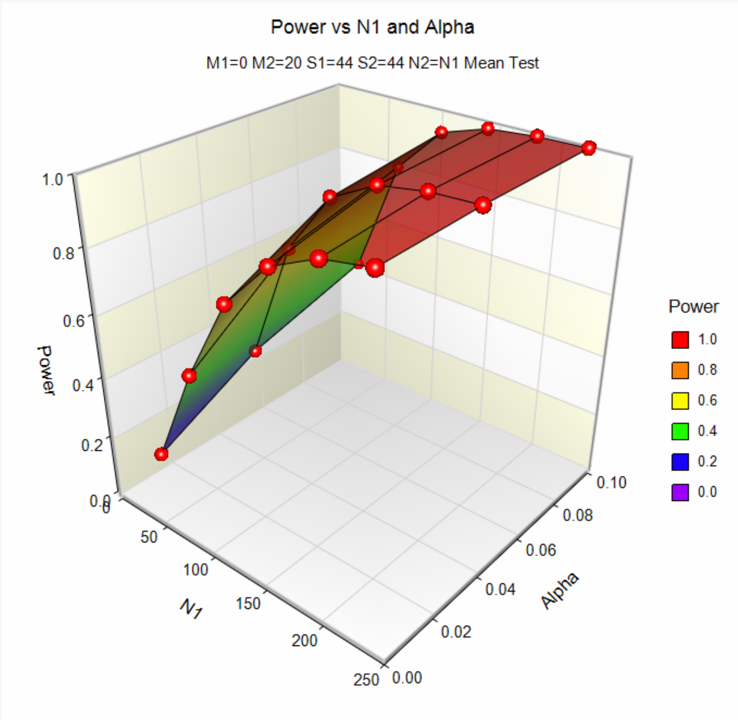
Numeric Results for Two-Sided Hypothesis Test of Means

Solve For: Power

Power	N1	N2	N	Mean1	Mean2	S1	S2	Alpha
0.12561	20	20	40	0	20	44	44	0.01
0.46051	60	60	120	0	20	44	44	0.01
0.73350	100	100	200	0	20	44	44	0.01
0.88707	140	140	280	0	20	44	44	0.01
0.95723	180	180	360	0	20	44	44	0.01
0.98514	220	220	440	0	20	44	44	0.01
0.29481	20	20	40	0	20	44	44	0.05
0.69292	60	60	120	0	20	44	44	0.05
0.88969	100	100	200	0	20	44	44	0.05
0.96495	140	140	280	0	20	44	44	0.05
0.98980	180	180	360	0	20	44	44	0.05
0.99722	220	220	440	0	20	44	44	0.05
0.40936	20	20	40	0	20	44	44	0.10
0.79088	60	60	120	0	20	44	44	0.10
0.93682	100	100	200	0	20	44	44	0.10
0.98274	140	140	280	0	20	44	44	0.10
0.99561	180	180	360	0	20	44	44	0.10
0.99893	220	220	440	0	20	44	44	0.10

Plots





These data show the power for various sample sizes and alphas. It is interesting to note that once the sample size is greater than 150, the value of alpha makes little difference on the value of power.

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 100, 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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For	**Power**
Alternative Hypothesis	**Two-Sided**
Alpha	**0.05**
Group Allocation	**Equal (N1 = N2)**
Sample Size Per Group	**100**
Mean1 (Mean of Group 1)	**0**
Mean2 (Mean of Group 2)	**20**
S1 (Standard Deviation Group 1)	**44**
S2 (Standard Deviation Group 2)	**S1**
Number of Looks	**1 (Also run with 2, 3, 4, 6, 8, 10, and 20)**
Spending Function	**O'Brien-Fleming**
Boundary Truncation	**None**
Max Time	**2**
Times	**Equally Spaced**
Informations	***Blank***

Output

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

Numeric Results for Two-Sided Hypothesis Test of Means

Solve For: Power

Power	N1	N2	N	Mean1	Mean2	S1	S2	Alpha	Looks
0.89511	100	100	200	0.00	20.00	44.0	44.0	0.050	1
0.89406	100	100	200	0.00	20.00	44.0	44.0	0.050	2
0.89159	100	100	200	0.00	20.00	44.0	44.0	0.050	3
0.88969	100	100	200	0.00	20.00	44.0	44.0	0.050	4
0.88715	100	100	200	0.00	20.00	44.0	44.0	0.050	6
0.88557	100	100	200	0.00	20.00	44.0	44.0	0.050	8
0.88451	100	100	200	0.00	20.00	44.0	44.0	0.050	10
0.88200	100	100	200	0.00	20.00	44.0	44.0	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 0.8951. When twenty interim looks are made, the power falls just 0.0131, to 0.8820—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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 4** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For .... **Power**
Alternative Hypothesis .... **Two-Sided**
Alpha.. **0.05 (will be calculated from boundaries)**
Group Allocation .... **Equal (N1 = N2)**
Sample Size Per Group **100**
Mean1 (Mean of Group 1)............................. **0**
Mean2 (Mean of Group 2)............................. **20**
S1 (Standard Deviation Group 1).................. **44**
S2 (Standard Deviation Group 2).................. **S1**
Number of Looks **5**
Spending Function.. **User Supplied**
Boundary Truncation..................................... **None**
Max Time **2**
Times .... **Equally Spaced**
Informations .... ***Blank***
Upper Boundaries... **3.5 3.5 3.0 2.5 2.0**
Lower Boundaries... **Symmetric**

Output

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

Numeric Results for Two-Sided Hypothesis Test of Means

Solve For: Power

Power	N1	N2	N	Mean1	Mean2	S1	S2	Alpha
0.88979	100	100	200	0	20	44	44	0.0482

Details when Spending = User Supplied, N1 = 100, N2 = 100, S1 = 44, S2 = 44, Diff = -20

Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.4	-3.5	3.5	0.0005	0.0005	0.0005	0.01958	0.0196
2	0.8	-3.5	3.5	0.0005	0.0004	0.0009	0.05884	0.0784
3	1.2	-3.0	3.0	0.0027	0.0024	0.0033	0.23249	0.3109
4	1.6	-2.5	2.5	0.0124	0.0103	0.0136	0.33997	0.6509
5	2.0	-2.0	2.0	0.0455	0.0345	0.0482	0.23893	0.8898

Drift = 3.21412

The power for this design is about 0.89. This value depends on both the boundaries and the sample size. The alpha level is 0.048157. This value depends only upon the boundaries.

Example 5 – Validation using O'Brien-Fleming Boundaries

Reboussin (1992) presents an example for normally distributed data for a design with two-sided O'Brien-Fleming boundaries, looks = 5, alpha = 0.05, beta = 0.10, *Mean1* = 220, *Mean2* = 200, standard deviation = 30. They compute a drift of 3.28 and a sample size of 48.41 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

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 5** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Sample Size**
Alternative Hypothesis ....**Two-Sided**
Power..**0.90**
Alpha..**0.05**
Group Allocation ....**Equal (N1 = N2)**
Mean1 (Mean of Group 1).............................**220**
Mean2 (Mean of Group 2).............................**200**
S1 (Standard Deviation Group 1)..................**30**
S2 (Standard Deviation Group 2)..................**S1**
Number of Looks **5**
Spending Function..**O'Brien-Fleming**
Boundary Truncation.....................................**None**
Max Time **1**
Times...**Equally Spaced**
Informations ....***Blank***

Output

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

Numeric Results for Two-Sided Hypothesis Test of Means

Solve For: Sample Size

Target Power	Actual Power	N1	N2	N	Mean1	Mean2	S1	S2	Alpha
0.9	0.90362	49	49	98	220	200	30	30	0.05

Details when Spending = O'Brien-Fleming, N1 = 49, N2 = 49, S1 = 30, S2 = 30, Diff = 20

Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.2	-4.87688	4.87688	0.000	0.000	0.000	0.00034	0.000
2	0.4	-3.35695	3.35695	0.001	0.001	0.001	0.10173	0.102
3	0.6	-2.68026	2.68026	0.007	0.007	0.008	0.35067	0.453
4	0.8	-2.28979	2.28979	0.022	0.017	0.024	0.29919	0.752
5	1.0	-2.03100	2.03100	0.042	0.026	0.050	0.15170	0.904

Drift = 3.29983

The slight difference in the power and the drift parameter is attributable to the rounding of the sample size from 48.41 to 49.

Example 6 – Validation with Pocock Boundaries

Reboussin (1992) presents an example for normally distributed data for a design with two-sided Pocock boundaries, looks = 5, alpha = 0.05, beta = 0.10, *Mean1* = 220, *Mean2* = 200, standard deviation = 30. They compute a drift of 3.55 and a sample size of 56.71 per group. The upper boundaries are: 2.4380, 2.4268, 2.4101, 2.3966, and 2.3859.

To test that **PASS** provides the same result, enter the following.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 6** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Sample Size**
Alternative Hypothesis ....**Two-Sided**
Power..**0.90**
Alpha...**0.05**
Group Allocation ....**Equal (N1 = N2)**
Mean1 (Mean of Group 1).............................**220**
Mean2 (Mean of Group 2).............................**200**
S1 (Standard Deviation Group 1)..................**30**
S2 (Standard Deviation Group 2)..................**S1**
Number of Looks **5**
Spending Function..**Pocock**
Boundary Truncation.....................................**None**
Max Time **1**
Times...**Equally Spaced**
Informations ....**Blank**

Output

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

Numeric Results for Two-Sided Hypothesis Test of Means

Solve For: Sample Size

Target Power	Actual Power	N1	N2	N	Mean1	Mean2	S1	S2	Alpha
0.9	0.90326	57	57	114	220	200	30	30	0.05

Details when Spending = Pocock, N1 = 57, N2 = 57, S1 = 30, S2 = 30, Diff = 20

Look	Time	Lower Bndry	Upper Bndry	Nominal Alpha	Inc Alpha	Total Alpha	Inc Power	Total Power
1	0.2	-2.43798	2.43798	0.015	0.015	0.015	0.19871	0.199
2	0.4	-2.42677	2.42677	0.015	0.011	0.026	0.26060	0.459
3	0.6	-2.41014	2.41014	0.016	0.009	0.035	0.21412	0.673
4	0.8	-2.39658	2.39658	0.017	0.008	0.043	0.14379	0.817
5	1.0	-2.38591	2.38591	0.017	0.007	0.050	0.08605	0.903

Drift = 3.55903

The slight difference in the power and the drift parameter is attributable to the rounding of the sample size from 56.71 to 57.