THE POWER TO KNOW.

SAS/STAT® 9.2 User's Guide, Second Edition

The SEQDESIGN Procedure

(Book Excerpt)

SAS® Documentation

This document is an individual chapter from the *SAS/STAT® 9.2 User's Guide, Second Edition*.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc. 2009. *SAS/STAT® 9.2 User's Guide, Second Edition*. Cary, NC: SAS Institute Inc.

SAS/STAT® User's Guide, Second Edition

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st electronic book, September 2009

SAS® Publishing provides a complete selection of books and electronic products to help customers use SAS software to its fullest potential. For more information about our e-books, e-learning products, CDs, and hard-copy books, visit the SAS Publishing Web site at **support.sas.com/publishing** or call 1-800-727-3228.

Chapter 77

The SEQDESIGN Procedure

Contents

Overview: SEQDESIGN Procedure

The purpose of the SEQDESIGN procedure is to design interim analyses for clinical trials. Clinical trials are experiments on human subjects to demonstrate the efficacy and safety of new drugs or treatments. A simple example is a trial to test the effectiveness of a new drug in humans by comparing the outcomes in a group of patients who receive the new drug with the outcomes in a comparable group of patients who receive a placebo.

A clinical trial is conducted according to a plan called a *protocol*. A protocol details the objectives of the trial, the data collection process, and the analysis. The protocol specifies the null hypothesis and an alternative hypothesis, a test statistic, the probability α of a Type I error, the probability β of a Type II error, the sample size needed to attain a specified power of $1 - \beta$ at an alternative reference, and critical values that are associated with the test statistic.

In a fixed-sample trial, data about all individuals are first collected and then examined at the end of the study. Most major trials have committees that periodically monitor safety and efficacy data during the trial and recommend that a trial be stopped for safety concerns such as an unacceptable toxicity level. In certain situations, the committee might recommend that a trial be stopped for efficacy. In contrast to a fixed-sample trial, a group sequential trial provides for interim analyses before the completion of the trial while maintaining the specified overall Type I and Type II error probabilities.

A group sequential trial is most useful in situations where it is important to monitor the trial to prevent unnecessary exposure of patients to an unsafe new drug, or alternatively to a placebo treatment if the new drug shows significant improvement. In most cases, if a group sequential trial stops early

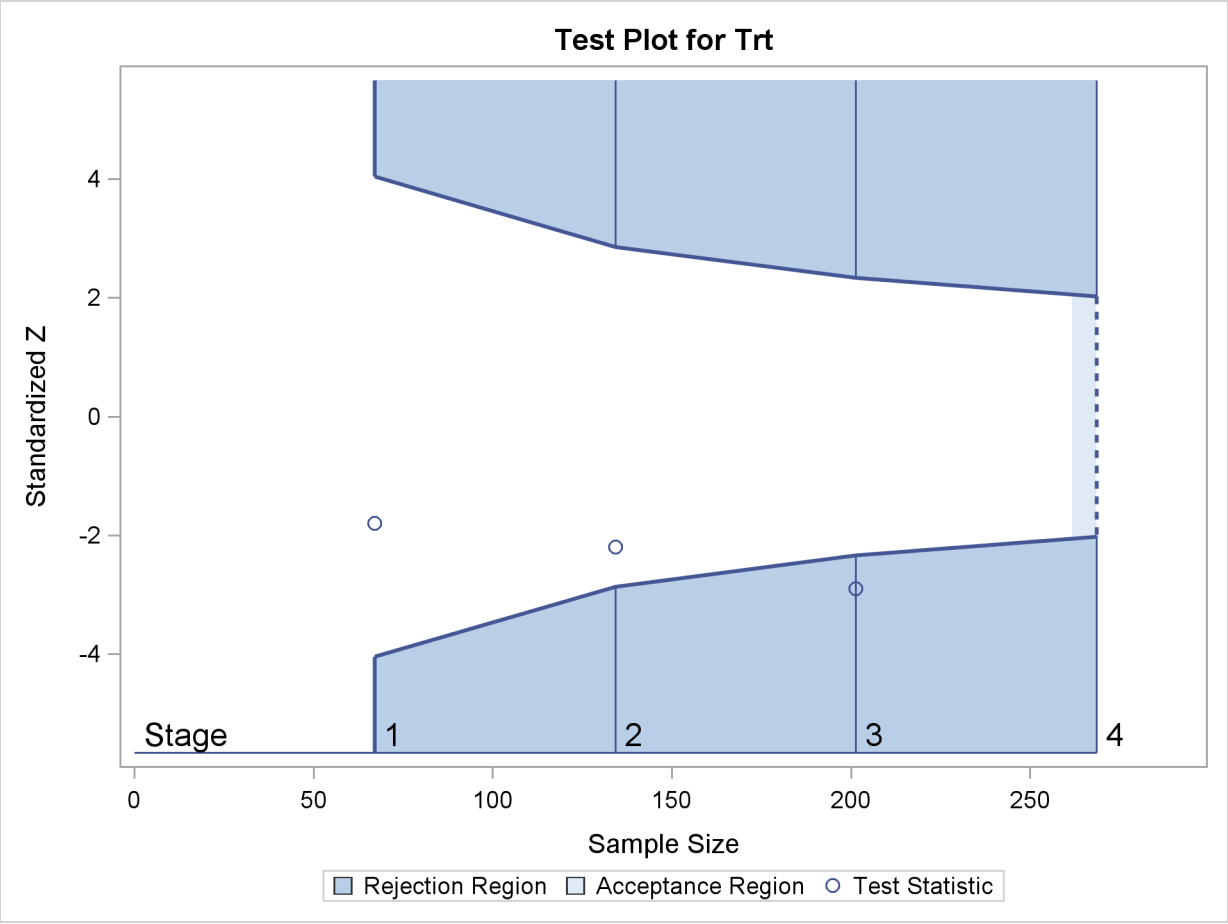
for safety concerns, fewer patients are exposed to the new treatment than in the fixed-sample trial. If a trial stops early for efficacy reasons, the new treatment is available sooner than it would be in a fixed-sample trial. Early stopping can also save time and resources.

A group sequential design provides detailed specifications for a group sequential trial. In addition to the usual specification for a fixed-sample design, it provides the total number of stages (the number of interim stages plus a final stage) and a stopping criterion to reject, to accept, or to either reject or accept the null hypothesis at each interim stage. It also provides critical values and the sample size at each stage for the trial.

At each interim stage, the data collected at the current stage in addition to the data collected at previous stages are analyzed, and statistics such as a maximum likelihood test statistic and its associated standard error are computed. The test statistic is then compared with critical values that are generated from the sequential design, and the trial is stopped or continued. If a trial continues to the final stage, the null hypothesis is either rejected or accepted. The critical values for each stage are chosen in such a way that the overall α and β are maintained at the specified levels.

Figure 77.1 shows a two-sided symmetric group sequential trial that stops early to reject the null hypothesis that the parameter Trt is zero.

Figure 77.1 Sequential Plot for Two-Sided Test



The trial has four stages, which are indicated by the vertical lines labeled 1, 2, 3, and 4. With early stopping to reject the null hypothesis, the lower rejection boundary is constructed by connecting the lower critical values for the stages. Similarly, the upper rejection boundary is constructed by connecting the upper critical values for the stages. The horizontal axis indicates the sample size for the group sequential trial, and the vertical axis indicates the values of the test statistic on the standardized Z scale.

At each interim stage, if the test statistic falls into a rejection region (the darker shaded areas in Figure 77.1), the trial stops and the null hypothesis is rejected. Otherwise, the trial continues to the next stage. At the final stage (stage 4), the null hypothesis is rejected if Z falls into a rejection region. Otherwise, the null hypothesis is not rejected. In Figure 77.1, the test statistic does not fall into the rejection regions for stages 1 and 2, and so the trial continues to stage 3. At stage 3, the test statistic falls into the rejection region, and the null hypothesis is rejected.

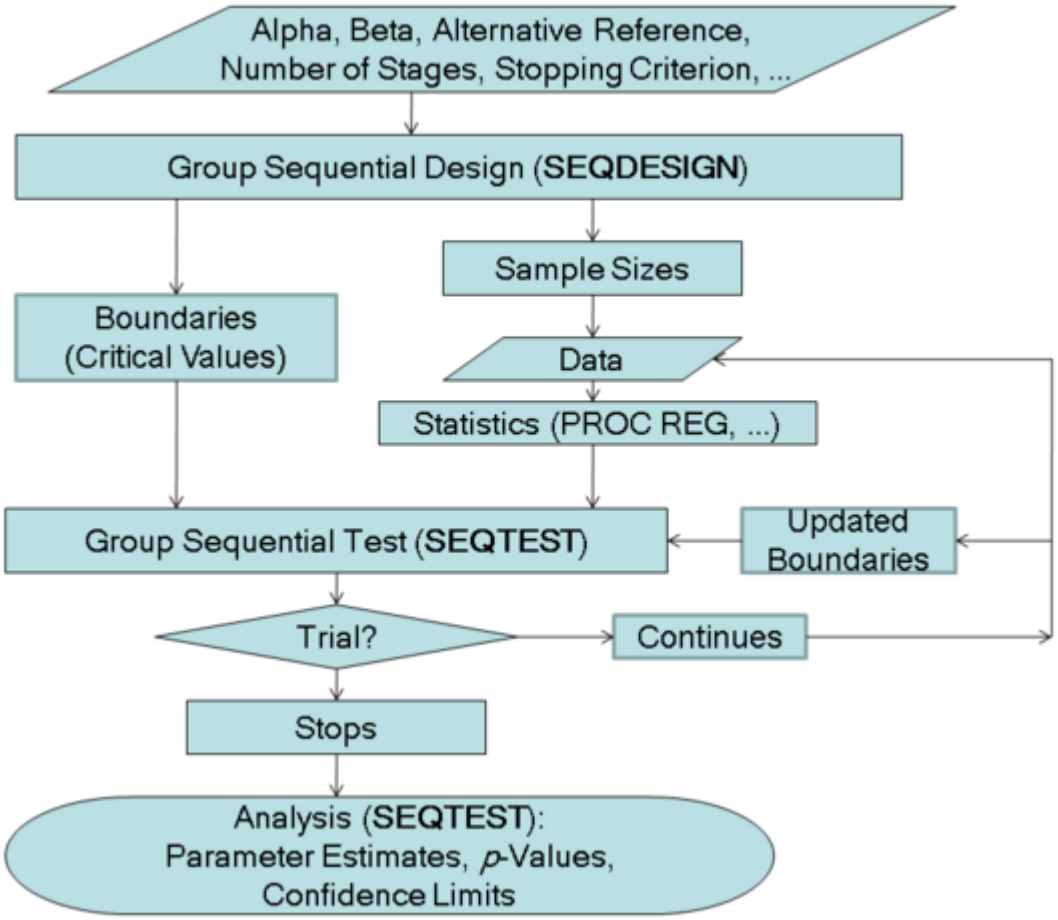
A group sequential trial usually involves six steps:

1. You specify the statistical details of the design, including the null and alternative hypotheses, a test statistic for the hypothesis test, the Type I and II error probabilities, a stopping criterion, the total number of stages, and the relative information level at each stage.
2. You compute the boundary values for the trial based on the specifications in Step 1. You also compute the sample size required at each stage for the specified hypothesis test.
3. At each stage, you collect additional data with the required sample sizes. The data available at each stage include the data collected at previous stages in addition to the data collected at the current stage.
4. At each stage, you analyze the available data with a procedure such as the REG procedure, and you compute the test statistic.
5. At each stage, you compare the test statistic with the corresponding boundary values. You stop the trial to reject or accept the hypothesis, or you continue the trial to the next stage. If you continue the trial to the final stage, you either accept or reject the hypothesis.
6. After the trial stops, you compute parameter estimates, confidence limits for the parameter, and a p-value for the hypothesis test.

You use the SEQDESIGN procedure at Step 2 to compute the initial boundary values and required sample sizes for the trial. You use the companion SEQTEST procedure at Step 5 to compare the test statistic with its boundary values. At stage 1, the boundary values are derived by using the boundary information tables created by the SEQDESIGN procedure. These boundary information tables are structured for input to the SEQTEST procedure. At each subsequent stage, the boundary values are derived by using the test information tables are created by the SEQTEST procedure at the previous stage. These test information tables are also structured for input to the SEQTEST procedure. You also use the SEQTEST procedure at Step 6 to compute parameter estimates, confidence limits, and p-values after the trial stops.

The flowchart in Figure 77.2 summarizes the steps in a typical group sequential trial and the relevant SAS procedures.

Figure 77.2 Group Sequential Trial



Features of the SEQDESIGN Procedure

The SEQDESIGN procedure assumes that the standardized Z test statistics for the null hypothesis $H_0 : \theta = 0$ at the stages have the joint canonical distribution with the information levels at the stages for the parameter θ. This implies that these test statistics are normally distributed. If the test statistic is not normally distributed, then it is assumed that the test statistic is computed from a large sample such that the statistic has an approximately normal distribution. See the section "Statistical Assumptions for Group Sequential Designs" on page 5829 for a detailed description of the joint canonical distribution.

You can use the SEQDESIGN procedure to compute required sample sizes for commonly used hypothesis tests. Note that for a fixed-sample design, you should use the POWER and GLMPOWER procedures to compute sample sizes.

The applicable tests include tests for binomial proportions and the log-rank test for two survival distributions. See the section "Applicable One-Sample Tests and Sample Size Computation" on page 5858, the section "Applicable Two-Sample Tests and Sample Size Computation" on page 5860, and the section "Applicable Regression Parameter Tests and Sample Size Computation" on page 5868 for examples of applicable tests in group sequential trials.

At each stage, the data are analyzed with a statistical procedure such as the REG procedure, and a test statistic and its associated information level are computed. The information level is the amount of information available about the unknown parameter. For a maximum likelihood statistic, the information level is the inverse of its variance.

At each stage, you use the SEQTEST procedure to derive the boundary values that correspond to the information level associated with the test statistic. You then use the SEQTEST procedure to compare the test statistic with these boundary values. When a trial is stopped at an interim stage or at the final stage, the SEQTEST procedure also derives parameter estimates, confidence limits for the parameter, and a p-value for hypothesis testing.

Output from the SEQDESIGN Procedure

In addition to computing the boundary values for a group sequential design, the SEQDESIGN procedure computes the following quantities:

- maximum sample size (as a percentage of the corresponding fixed-sample size) if the trial does not stop at an interim stage
- average sample numbers (as percentages of the corresponding fixed-sample sizes for nonsurvival data or fixed-sample numbers of events for survival data) under various hypothetical references, including the null and alternative references
- stopping probabilities at each stage under various hypothetical references to indicate how likely it is that the trial will stop at that stage
- sample sizes required at each stage for the specified hypothesis test
- numbers of events required at each stage for the specified hypothesis test with survival data

You can create more than one design with multiple DESIGN statements in the SEQDESIGN procedure and then choose the design with the most desirable features. The next two subsections introduce some basic aspects of group sequential designs that are useful for getting started with the SEQDESIGN procedure.

Boundaries for Group Sequential Designs

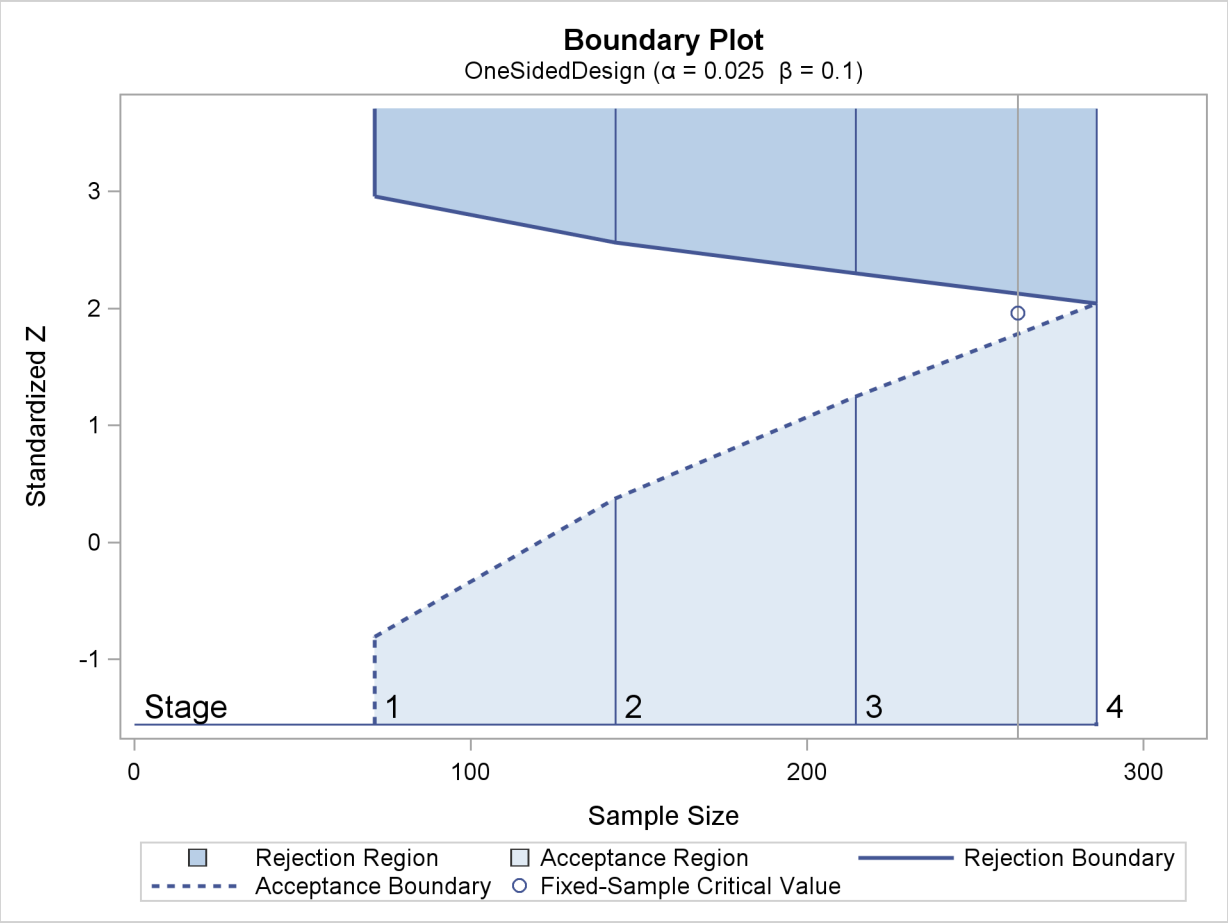
A one-sided test is a test of a hypothesis with either a lower alternative ($H_1 : \theta < 0$) or an upper alternative ($H_1 : \theta > 0$), and a two-sided test is a test with a two-sided alternative ($H_1 : \theta \neq 0$). The number of critical values for a test depends on whether the alternative is one-sided or two-sided, and it also depends on whether the trial is conducted with a fixed-sample design or a group sequential design.

For a fixed-sample trial, a one-sided test has one critical value and a two-sided test has two critical values. These critical values are computed with the specified Type I error probability α. In contrast, at each interim stage of a group sequential trial, a one-sided test has up to two critical values and a two-sided group sequential test has up to four critical values. Thus, there are two or four possible

boundaries for a group sequential design, and each boundary is a set of critical values, one from each stage.

Figure 77.3 illustrates the boundaries for a one-sided test with an upper alternative that allows for early stopping to either reject or accept the null hypothesis.

Figure 77.3 Boundary Plot for One-Sided Design



With an upper alternative, as in this example, the design has the following two boundaries: an upper α (rejection) boundary for the rejection region that consists of upper rejection critical values and an upper β (acceptance) boundary for the acceptance region that consists of upper acceptance critical values. The stages are indicated by vertical lines with accompanying stage numbers. In Figure 77.3, the horizontal axis indicates the cumulative sample size for the group sequential trial. The vertical axis indicates the critical values at each stage on the standardized Z scale. Other scales can be used for the vertical axis, including the MLE scale, score statistic scale, and p-value scale.

At each interim stage, if the test statistic is in the rejection region (darker shaded area in Figure 77.3), the trial stops and the null hypothesis is rejected. If the test statistic is in the acceptance region (lightly shaded area in Figure 77.3), the trial stops and the hypothesis is accepted. Otherwise, the trial continues to the next stage. If the trial proceeds to the final stage (stage 4), the upper α and upper β critical values are identical, and the trial stops to either reject or accept the null hypothesis.

Group Sequential Methods

For a group sequential design, there are two possible boundaries for a one-sided test and four possible boundaries for a two-sided test. Each boundary consists of one boundary value (critical value) for each stage. The SEQDESIGN procedure provides various methods for computing the boundary values.

The boundary value for a fixed-sample test cannot be applied to each stage of sequential design because, as shown by Armitage, McPherson, and Rowe (1969), repeated significance tests at a fixed level on accumulating data increase the probability of a Type I error. For example, with a fixed-sample two-sided test, the critical values ± 1.96 for a standardized Z statistic produce a Type I error probability level $\alpha = 0.05$. But for a two-sided group sequential test with two equally spaced stages, if the same critical values ± 1.96 are used to reject the null hypothesis at these two stages, the Type I error probability level is $\alpha = 0.083$, larger than the fixed-sample α.

Numerous methods are available for deriving the critical values for each boundary in a sequential design. Pocock (1977) applies repeated significance tests to group sequential trials with equal-size groups and derives a constant critical value on the standardized normal Z scale across all stages that maintains the specified Type I error probability level. O'Brien and Fleming (1979) propose a sequential procedure that has boundary values (in absolute value) decrease over the stages on the standardized normal Z scale.

The SEQDESIGN procedure provides the following three types of methods:

- fixed boundary shape methods, which derive boundaries with specified boundary shapes
- Whitehead methods, which adjust boundaries derived for continuous monitoring so that they apply to discrete monitoring
- error spending methods

Each type of methods uses a distinct approach to derive the boundary values for a group sequential trial. Whitehead methods require much less computation with resulting Type I error probability and power that are close but differ slightly from the specified values due to the approximations used in deriving the tests (Jennison and Turnbull 2000, p. 106). Fixed boundary shape methods derive boundary values by estimating a fixed number of parameters and require more computation. Error spending methods derive boundary values at each stage sequentially and require much more computation than other types of methods for group sequential trials with a large number of stages.

Within each type of methods, you can choose methods that creating boundary values range from conservative stopping boundary values at early stages to liberal stopping boundary values at very early stages.

You can use the SEQDESIGN procedure to specify methods from the same type for each design. A different method can be specified for each boundary separately, but all methods in a design must be of the same type.

Fixed Boundary Shape Methods

The fixed boundary shape methods include the unified family methods and the Haybittle-Peto method. The unified family methods (Kittelson and Emerson 1999) derive boundaries from specified boundary shapes. These methods include Pocock's method (Pocock 1977) and the O'Brien-Fleming method (O'Brien and Fleming 1979) as special cases.

The Haybittle-Peto method (Haybittle 1971; Peto et al. 1976) uses a value of 3 for the critical values in interim stages, so the critical value at the final stage is close to the critical value for the fixed-sample design. In the SEQDESIGN procedure, the Haybittle-Peto method has been generalized to allow for different boundary values at interim stages.

Whitehead Methods

Whitehead and Stratton (1983) and Whitehead (1997, 2001) develop triangular and straight-line boundaries by adapting tests constructed for continuous monitoring to discrete monitoring of group sequential tests. With continuous monitoring, the values for each boundary fall in a straight line when plotted on the score statistic scale. The discrete boundary is derived by subtracting the expected overshoot from the continuous boundary to obtain the desired Type I and Type II error probabilities. For a design with early stopping to either reject or accept the null hypothesis, the boundaries form a triangle when plotted on the score statistic scale. See the section "Score Statistic" on page 5819 for a detailed description of the score statistic.

Error Spending Methods

For every sequential design, the α and β errors at each stage can be computed from the boundary values. On the other hand, you can derive the boundary values from specified α and β errors for each stage. The error spending function approach (Lan and DeMets 1983) uses an error spending function to specify the errors at each stage for each boundary and then derives the boundary values.

Getting Started: SEQDESIGN Procedure

This section illustrates a clinical study design that uses a two-sided O'Brien-Fleming design (O'Brien and Fleming 1979) to stop the trial early for ethical concerns about possible harm or for unexpectedly strong efficacy of the new drug.

Suppose that a pharmaceutical company is conducting a clinical trial to test the efficacy of a new cholesterol-lowering drug. The primary focus is low-density lipoprotein (LDL), the so-called bad cholesterol, which is a risk factor for coronary heart disease. LDL is measured in mg/dL, milligrams per deciliter of blood.

The trial consists of two groups of equally allocated patients with elevated LDL levels: an experimental group given the new drug and a placebo control group. Suppose the changes in LDL level

after the treatment for individuals in the experimental and control groups are normally distributed with means μ_e and μ_c, respectively, and have a common variance σ^2. Then the null hypothesis of no effect for the new drug is $H_0 : \theta = 0$, where $\theta = \mu_e - \mu_c$.

For a fixed-sample design with a total sample size N, the MLE for θ is computed as $\hat{\theta} = \hat{\mu}_e - \hat{\mu}_c$, where $\hat{\mu}_e$ and $\hat{\mu}_c$ are the sample means of the decreases in LDL level in the experimental and control groups, respectively.

Following the derivation in the section "Test for the Difference between Two Normal Means" on page 5860, the statistic $\hat{\theta}$ has a normal distribution

$$\hat{\theta} \sim N\left(\theta, \frac{4\sigma^2}{N}\right)$$

Thus, under the null hypothesis $H_0 : \theta = 0$, the standardized statistic

$$Z = \frac{\hat{\theta}}{\sqrt{\frac{4\sigma^2}{N}}} \sim N(0, 1)$$

The Z statistic can be used to test the null hypothesis H_0. If the variance σ^2 is unknown, the sample variance can be used to compute the test statistic if it is assumed that the sample variance is computed from a large sample such that the Z statistic has an approximately standard normal distribution.

With a Type I error probability $\alpha = 0.05$, the critical values for the Z statistic are given by $\Phi^{-1}(\alpha/2) = -1.96$ and $\Phi^{-1}(1-\alpha/2) = 1.96$, where Φ is the cumulative standard normal distribution function. At the end of study, if $Z \geq 1.96$, the null hypothesis is rejected for harmful drug effect, and if $Z \leq -1.96$, the null hypothesis is rejected for efficacy of the new drug. Otherwise, the null hypothesis is not rejected and the drug effect is not significant.

Also suppose that for the trial, the alternative reference $\theta = -10$ is the clinically meaningful difference that the trial should detect with a high probability (power). Further suppose that a good estimate of the standard deviation for the changes in LDL level is $\hat{\sigma} = 20$. The following statements invoke the SEQDESIGN procedure and request a four-stage O'Brien-Fleming design for standardized normal test statistics:

```
ods graphics on;
 proc seqdesign altref=-10
                plots=boundary(hscale=samplesize)
                ;
   TwoSidedOBrienFleming: design nstages=4
                          method=obf
                          ;
   samplesize model=twosamplemean(stddev=20);
ods output Boundary=Bnd_LDL;
run;
ods graphics off;
```

The ALTREF= option specifies the alternative reference, and the actual maximum information is derived in the SEQDESIGN procedure. With the specified ODS GRAPHICS ON statement, the PLOTS=BOUNDARY option displays a boundary plot with the rejection and acceptance regions.

In the DESIGN statement, the label `TwoSidedOBrienFleming` identifies the design in the output tables. The default ALT=TWOSIDED option specifies a two-sided alternative hypothesis. The default STOP=REJECT option specifies early stopping in the interim stages only for the purpose of rejecting the null hypothesis. That is, at each interim stage, the trial will either be stopped to reject the null hypothesis or continue to the next stage.

The NSTAGES=4 option in the DESIGN statement specifies the total number of stages in the group sequential trial, including three interim stages and a final stage. In the SEQDESIGN procedure, the null hypothesis for the design is $H_0 : \theta = 0$. The default ALPHA=0.05 option specifies a Type I error probability $\alpha = 0.05$, and the default BETA=0.10 option specifies a Type II error probability $\beta = 0.10$, which corresponds to a power of $1-\beta = 0.90$ at the alternative reference $H_1 : \theta = -10$.

For a two-sided design with early stopping to reject the null hypothesis, there are two boundaries for the design: an upper α boundary that consists of upper rejection critical values and a lower α boundary that consists of lower rejection critical values. Each boundary is a set of critical values, one from each stage. With the METHOD=OBF option in the DESIGN statement, the O'Brien-Fleming method is used for the two boundaries for the design; see Figure 77.7.

A property of the boundaries constructed with the O'Brien-Fleming design is that the null hypothesis is more difficult to reject in the early stages than in the later stages. That is, the trial is rejected in the early stages only with overwhelming evidence, because in these stages there might not be a sufficient number of responses for a reliable estimate of the treatment effect.

The SAMPLESIZE statement with the MODEL=TWOSAMPLEMEAN option uses the derived maximum information to compute required sample sizes for a two-sample test for mean difference. The ODS OUTPUT statement with the BOUNDARY=BND_LDL option creates an output data set named BND_LDL which contains the resulting boundary information.

In a clinical trial, the amount of information about an unknown parameter available from the data can be measured by the Fisher information. For a maximum likelihood statistic, the information level is the inverse of its variance. See the section "Maximum Likelihood Estimator" on page 5818 for a detailed description of Fisher information. At each stage of the trial, data are collected and analyzed with a statistical procedure, and a test statistic and its corresponding information level are computed.

In this example, you can use the REG procedure to compute the maximum likelihood estimate $\hat{\theta}$ for the drug effect and the corresponding standard error for $\hat{\theta}$. At stage 1, you can use the SEQTEST procedure to compare the test statistic with adjusted boundaries derived from the boundary information stored in the BOUND_LDL data set. At each subsequent stage, you can use the SEQTEST procedure to compare the test statistic with adjusted boundaries derived from the boundary information stored in the test information table created by the SEQTEST procedure at the previous stage. The test information tables are structured for input to the SEQTEST procedure.

At each interim stage, the trial will either be stopped to reject the null hypothesis or continue to the next stage. At the final stage, the null hypothesis is either rejected or accepted.

By default, the SEQDESIGN procedure derives boundary values with equally spaced information levels for all stages—that is, the same information increment between successive stages. The "Design Information," "Method Information," and "Boundary Information" tables are displayed by default, as shown in Figure 77.4, Figure 77.5, and Figure 77.6, respectively.

The "Design Information" table in Figure 77.4 displays design specifications and four derived statistics: the actual maximum information, the maximum information, the average sample number under the null hypothesis (Null Ref ASN), and the average sample number under the alternative hypothesis (Alt Ref ASN). Except for the actual maximum information, each statistic is expressed as a percentage of the identical statistic for the corresponding fixed-sample information. The average sample number is the expected sample size (for nonsurvival data) or expected number of events (for survival data). Note that for a symmetric two-sided design, the ALTREF=−10 option implies a lower alternative reference of −10 and an upper alternative reference of 10.

Figure 77.4 O'Brien-Fleming Design Information

The SEQDESIGN Procedure
Design: TwoSidedOBrienFleming

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Method	O'Brien-Fleming
Boundary Key	Both
Alternative Reference	-10
Number of Stages	4
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	102.2163
Max Information	0.107403
Null Ref ASN (Percent of Fixed Sample)	101.5728
Alt Ref ASN (Percent of Fixed Sample)	76.7397

The maximum information is the information level at the final stage of the group sequential trial. The Max Information (Percent Fixed-Sample) is the maximum information for the sequential design expressed as a percentage of the information for the corresponding fixed-sample design. In Figure 77.4, the Max Information (Percent Fixed-Sample) is 102.22%, which means that the information needed for the group sequential trial is 2.22% more than that of the corresponding fixed-sample design if the trial does not stop at any interim stage.

The Null Ref ASN (Percent Fixed-Sample) is the average sample number (expected sample size) required under the null hypothesis for the group sequential design expressed as a percentage of the sample size for the corresponding fixed-sample design. In Figure 77.4, the Null Ref ASN is 101.57%, which means that the expected sample size for the group sequential trial is 1.57% greater than the corresponding fixed-sample size.

Similarly, the Alt Ref ASN (Percent Fixed-Sample) is the average sample number (expected sample size) required under the alternative hypothesis for the group sequential design expressed as a percentage of the sample size for the corresponding fixed-sample design. In Figure 77.4, the Alt Ref ASN is 76.74%, which means that the expected sample size for the group sequential trial is 76.74% of the corresponding fixed-sample size. That is, if the alternative hypothesis is true, then on average, only 76.74% of the fixed-sample size is needed for the group sequential trial.

In this example, the O'Brien-Fleming design requires only a slight increase in sample size if the trial proceeds to the final stage. On the other hand, if the alternative hypothesis is correct, this design provides a substantial saving in sample size on average.

The "Method Information" table in Figure 77.5 displays the computed Type I and Type II error probabilities α and β, and the derived drift parameter for the design. For a two-sided test with early stopping to reject the null hypothesis, both lower and upper α boundaries are created. With the specified ALTREF= option, the alternative references are also included.

With the zero null reference, the drift parameter is the standardized alternative reference at the final stage $\theta_1\sqrt{I_X}$, where θ_1 is the alternative reference and I_X is the maximum information. See the section "Specified and Derived Parameters" on page 5851 for a detailed description of the drift parameter. The drift parameters for the design are derived in the SEQDESIGN procedure even if the alternative reference is not specified or derived in the procedure.

Figure 77.5 Method Information

Method Information

Boundary	Method	Alpha	Beta	Unified Family Rho	Unified Family Tau	Unified Family C
Upper Alpha	O'Brien-Fleming	0.02500	0.10000	0.5	0	2.02429
Lower Alpha	O'Brien-Fleming	0.02500	0.10000	0.5	0	2.02429

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	10	3.277238
Lower Alpha	-10	-3.27724

The O'Brien-Fleming method belongs to the unified family of designs, which is parameterized by two parameters, ρ and τ, as implemented in the SEQDESIGN procedure. See Table 77.3 for parameter values of commonly used methods in the unified family. The "Method Information" table in Figure 77.5 displays the values of $\rho = 0.5$ and $\tau = 0$, which are the parameters for the O'Brien-Fleming method. The table also displays the derived parameter $C_\alpha = 2.0243$, which is used in the construction of symmetric lower and upper α boundaries; see the section "Unified Family Methods" on page 5839.

The "Boundary Information" table in Figure 77.6 displays the information level, including the proportion, actual level, and corresponding sample size (N) at each stage. The table also displays the lower and upper alternative references, and the lower and upper boundary values at each stage. By default, equally spaced information levels for all stages are used to derive boundary values.

Figure 77.6 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Information Level			Alternative Reference	
Stage	Proportion	Actual	N	Lower	Upper
1	0.2500	0.026851	42.96116	−1.63862	1.63862
2	0.5000	0.053701	85.92233	−2.31736	2.31736
3	0.7500	0.080552	128.8835	−2.83817	2.83817
4	1.0000	0.107403	171.8447	−3.27724	3.27724

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Boundary Values	
	Lower	Upper
Stage	Alpha	Alpha
1	−4.04859	4.04859
2	−2.86278	2.86278
3	−2.33745	2.33745
4	−2.02429	2.02429

The information proportion is the proportion of maximum information available at each stage and N is the corresponding sample size. The default BOUNDARYSCALE=STDZ option specifies that the standardized Z scale be used to display the boundary values in the boundary information table and the boundary plot. The alternative reference on the standardized Z scale at stage k is given by $\theta_1\sqrt{I_k}$, where θ_1 is the alternative reference and I_k is the information available at stage k, $k = 1, 2, 3, 4$. These standardized alternative references for the design are derived in the SEQDESIGN procedure even if the alternative reference is not specified or derived in the procedure.

In this example, a standardized Z statistic is computed by standardizing the parameter estimate of the effect in LDL level. A lower Z test statistic indicates a beneficial effect. Consequently, at each interim stage, if the standardized Z test statistic is less than or equal to the corresponding lower α boundary value, the hypothesis $H_0 : \theta = 0$ is rejected for efficacy. If the test statistic is greater than or equal to the corresponding upper α boundary value, the hypothesis H_0 is rejected for harmful effect. Otherwise, the process continues to the next stage. At the final stage (stage 4), the hypothesis H_0 is rejected for efficacy if the Z statistic is less than or equal to the corresponding lower α boundary value -2.0243, and the hypothesis H_0 is rejected for harmful effect if the Z statistic is greater than or equal to the corresponding upper α boundary value 2.0243. Otherwise, the hypothesis of no significant difference is accepted.

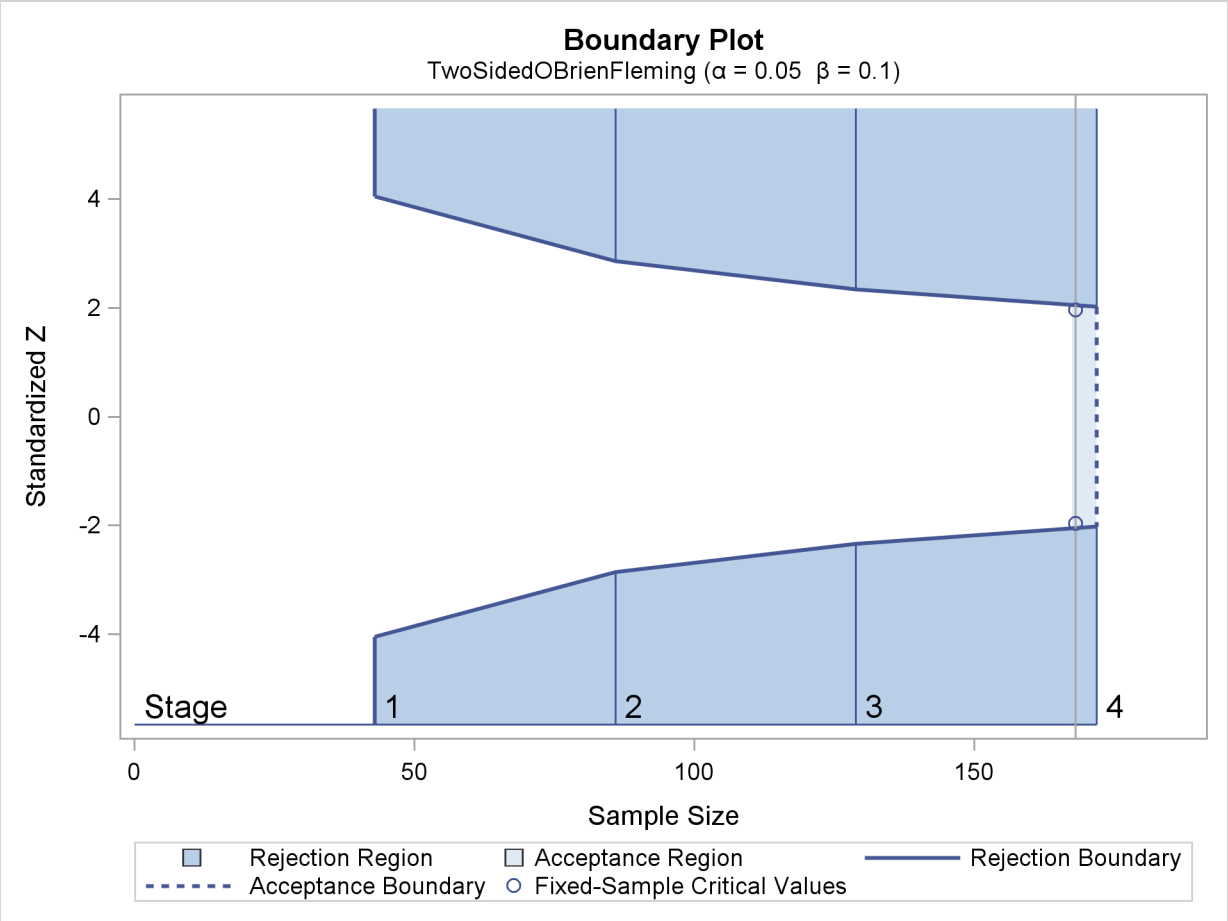
Note that in a typical trial, the actual information levels do not match the information levels specified in the design. The SEQTEST procedure modifies the boundary values stored in the BOUND_LDL data set to adjust for these new information levels.

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Figure 77.7. This plot displays the boundary values in the "Boundary Information" table in Figure 77.6. The stages are indicated by vertical lines

with accompanying stage numbers. The horizontal axis indicates the sample sizes for the stages. Note that comparing with a fixed-sample design, only a small increase in sample size is needed for the O'Brien-Fleming design, as shown in Figure 77.7.

If a test statistic at an interim stage is in the rejection region (shaded area), the trial stops and the null hypothesis is rejected. If the statistic is not in any rejection region, the trial continues to the next stage.

Figure 77.7 Boundary Plot



The boundary plot also displays critical values for the corresponding fixed-sample design. The symbol "∘" identifies the fixed-sample critical values of -1.96 and 1.96, and the accompanying vertical line indicates the required sample size for the fixed-sample design at the horizontal axis. Note that the boundary values ± 2.0243 at the final stage are close to the fixed-sample critical values ± 1.96.

When you specify the SAMPLESIZE statement, the maximum information (either explicitly specified or derived in the SEQDESIGN procedure) is used to compute the required sample sizes for the study. The MODEL=TWOSAMPLEMEAN(STDDEV=20) option specifies the test for the difference between two normal means. See the section "Test for the Difference between Two Normal Means" on page 5860 for a detailed derivation of these required sample sizes.

The "Sample Size Summary" table in Figure 77.8 displays the parameters for the sample size computation and the resulting maximum and expected sample sizes.

Figure 77.8 Sample Size Summary

Sample Size Summary

Test	Two-Sample Means
Mean Difference	-10
Standard Deviation	20
Max Sample Size	171.8447
Expected Sample Size (Null Ref)	170.7627
Expected Sample Size (Alt Ref)	129.0137

The "Sample Sizes (N)" table in Figure 77.9 displays the required sample sizes at each stage for the trial, in both fractional and integer numbers. The derived fractional sample sizes are displayed under the heading "Fractional N." These sample sizes are rounded up to integers under the heading "Ceiling N." With the default WEIGHT=1 option in the SAMPLESIZE statement, the sample sizes for the two groups are equal for the two-sample test.

Figure 77.9 Derived Sample Sizes

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Fractional N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	42.96	21.48	21.48	0.0269
2	85.92	42.96	42.96	0.0537
3	128.88	64.44	64.44	0.0806
4	171.84	85.92	85.92	0.1074

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Ceiling N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	44	22	22	0.0275
2	86	43	43	0.0538
3	130	65	65	0.0812
4	172	86	86	0.1075

In practice, integer sample sizes are used in the trial, and the resulting information levels increase slightly. Thus, 22, 43, 65, and 86 individuals are needed in each of the two groups for the four stages, respectively.

Syntax: SEQDESIGN Procedure

The following statements are available in PROC SEQDESIGN:

PROC SEQDESIGN < *options* > **;**
  < *label:* > **DESIGN** *options* **;**
  SAMPLESIZE < *MODEL= option* > **;**

The PROC SEQDESIGN statement and the DESIGN statement are required for the SEQDESIGN procedure. Each DESIGN statement requests a new group sequential design, and multiple DESIGN statements can be used to create more than one design for comparison of features. The label, which must be a valid SAS name, is used to identify the design in the output tables and graphics. The SAMPLESIZE statement computes the required sample sizes for the design specified in each DESIGN statement. With a selected design, the SAMPLESIZE statement computes the required sample sizes for the trial.

PROC SEQDESIGN Statement

PROC SEQDESIGN < *options* > **;**

Table 77.1 summarizes the options in the PROC SEQDESIGN statement.

Table 77.1 Summary of PROC SEQDESIGN Options

Option	Description
Design Parameters	
ALTREF=	specifies alternative reference
BOUNDARYSCALE=	specifies statistic scale for the boundary
MAXINFO=	specifies maximum information level
Table Output	
ERRSPEND	displays cumulative error spending at each stage
PSS	displays powers and expected sample sizes
STOPPROB	displays expected cumulative stopping probabilities
Graphics Output	
PLOTS=ASN	displays expected sample numbers plot
PLOTS=BOUNDARY	displays detailed boundary plot
PLOTS=COMBINEDBOUNDARY	displays combined boundary plot
PLOTS=ERRSPEND	displays error spending plot
PLOTS=POWER	displays powers plot

By default, the SEQDESIGN procedure displays tables of design information, method information, and boundary information for each specified design. If the ODS GRAPHICS ON statement is specified, it also displays a detailed boundary plot.

In addition, you can use output options to display output tables such as expected cumulative stopping probability at each stage under various hypothetical references. If the ODS GRAPHICS ON statement is specified, you can also use output options to display plots such as powers and expected sample sizes under various hypothetical references.

The following options can be used in the PROC SEQDESIGN statement to derive boundary values for all sequential designs in the procedure. They are listed in alphabetical order.

ALTREF=θ_1 < (< LOWER=θ_{1l} > < UPPER=θ_{1u} >) >

specifies the alternative reference—that is, the hypothetical reference under the alternative hypothesis at which the power is computed. The LOWER= and UPPER= options are applicable only for a two-sided design with different lower and upper alternative references.

For a one-sided design, $\theta_{1l} = -|\theta_1|$ is the lower alternative reference and $\theta_{1u} = |\theta_1|$ is the upper alternative reference. For a two-sided design, the specified θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively. If the LOWER= option is not specified, $\theta_{1l} = -|\theta_1|$, and if the UPPER= option is not specified, $\theta_{1u} = |\theta_1|$.

The specification of the ALTREF= option depends on the hypothesis used in the clinical trial. For example, suppose the null hypothesis $H_0 : \theta = 0$ with an alternative hypothesis $H_1 : \theta = \theta_1$ is used to compare two binomial populations, $p_a = p_b$. Then θ_1 is the proportion difference under H_1 if $\theta = p_a - p_b$, and θ_1 is the log odds ratio under H_1 if $\theta = \log\left(\frac{p_a(1-p_b)}{p_b(1-p_a)}\right)$.

If the ALTREF= option is not specified, the alternative reference θ_1 can also be specified or derived in the SAMPLESIZE statement. If θ_1 is specified or derived in the SAMPLESIZE statement, $\theta_{1l} = -|\theta_1|$ and $\theta_{1u} = |\theta_1|$ are the lower and upper alternative references, respectively.

Note that if the SAMPLESIZE statement is specified with a two-sided design, the sample sizes derived by using the lower and upper alternatives might be different. If θ_1 is specified or derived in the SAMPLESIZE statement, it is used to compute the sample sizes. Otherwise, the θ_1 specified in the ALTREF= option is used.

BOUNDARYSCALE=MLE | SCORE | STDZ | PVALUE

BSCALE=MLE | SCORE | STDZ | PVALUE

specifies the scale for the statistic that is displayed in the boundary table and boundary plots. The keywords MLE, SCORE, STDZ, and PVALUE correspond to the boundary with the maximum likelihood estimate scale, the score statistic scale, the standardized normal Z scale, and the p-value scale, respectively. The default is BOUNDARYSCALE=STDZ.

With the BOUNDARYSCALE=MLE or BOUNDARYSCALE=SCORE option, the maximum information must be either explicitly specified with the MAXINFO= option or derived in the SEQDESIGN procedure to provide the necessary information level at each stage to compute the boundary values. See the section "Boundary Scales" on page 5832 for a detailed description of the statistic scale for the boundary values.

Note that for a two-sided design, the p-value scale displays the one-sided fixed-sample p-value under the null hypothesis with a lower alternative hypothesis.

MAXINFO=*number*

specifies the maximum information level for the design. If the MAXINFO=option is specified and the alternative reference is either specified explicitly with the ALTREF= option or derived from the SAMPLESIZE statement, then the Type I and Type II error probability levels cannot be met simultaneously. In this case, the ALPHA= option in the DESIGN statement is applicable only with the default BOUNDARYKEY=ALPHA option in the DESIGN statement, and the Type II error probability β is derived. The BETA= option in the DESIGN statement is applicable only with the BOUNDARYKEY=BETA option in the DESIGN statement, and the Type I error probability α is derived.

Table Output Options

The following options can be used in the PROC SEQDESIGN statement to display addition table output. They are listed in alphabetical order.

ERRSPEND

displays the error spending at each stage for each boundary in the design.

PSS < (CREF= *numbers* **) >**

displays powers and expected sample sizes under various hypothetical references, where the numbers $c_i \geq 0$.

For a one-sided design, the power and expected sample sizes under hypotheses $\theta = c_i\,\theta_1$ are displayed, where θ_1 is the alternative reference and c_i are the values specified in the CREF= option.

For a two-sided design, the power and expected sample sizes under hypotheses $\theta = c_i\theta_{1l}$ and $\theta = c_i\theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively. The default is CREF= 0 0.5 1.0 1.5.

Note that for a symmetric two-sided design, only the power and expected sample sizes under hypotheses $\theta = c_i\,\theta_{1u}$ are derived. See the section "Type I and Type II Errors" on page 5836 for a detailed description of the power computation. See the section "Powers and Expected Sample Sizes" on page 5878 for a detailed description of the expected sample size computation.

STOPPROB < (CREF= *numbers* **) >**

displays expected cumulative stopping probabilities under various hypothetical references, where the numbers $c_i \geq 0$.

For a one-sided design, expected cumulative stopping probabilities at each stage under hypotheses $\theta = c_i\,\theta_1$ are displayed, where θ_1 is the alternative reference and c_i are the values specified in the CREF= option.

For a two-sided design, expected cumulative stopping probabilities at each stage under hypotheses $\theta = c_i\theta_{1l}$ and $\theta = c_i\theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively. Note that for a symmetric two-sided design, only expected cumulative stopping probabilities under hypotheses $\theta = c_i\,\theta_{1u}$ are derived. The default is CREF= 0 0.5 1.0 1.5.

Graphics Output Options

This section describes the options for using ODS Graphics with the SEQDESIGN procedure to create plots. To request these graphs, you must specify the ODS GRAPHICS ON statement in addition to the following options in the PROC SEQDESIGN statement. For more information about the ODS GRAPHICS statement, see Chapter 21, "Statistical Graphics Using ODS."

The following options can be used in the PROC SEQDESIGN statement to display graphs with ODS Graphics. They are listed in alphabetical order.

PLOTS < (ONLY) > < = *plot-request* >

PLOTS < (ONLY) > < = (*plot-request* < . . . *plot-request* >) >

specifies options that control the details of the plots. The default is PLOTS=BOUNDARY. The global plot option ONLY suppresses the default plots and displays only plots specifically requested.

The plot request options are as follows.

ALL

produces all appropriate plots.

ASN < (CREF= *numbers*) >

displays a plot of the average sample numbers (expected sample sizes for nonsurvival data or expected numbers of events for survival data) under various hypothetical references, where the numbers $c_i \geq 0$. These average sample numbers are displayed as percentages of the average sample numbers for the corresponding fixed-sample design.

For a one-sided design, expected sample numbers under hypotheses $\theta = c_i\ \theta_1$ are displayed, where θ_1 is the alternative reference and c_i are the values specified in the CREF= option.

For a two-sided design, expected sample numbers under hypotheses $\theta = c_i \theta_{1l}$ and $\theta = c_i \theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively. Note that for a symmetric two-sided design, only the average sample numbers under hypotheses $\theta = c_i\ \theta_{1u}$ are derived. The default is CREF= 0 to 1.5 by 0.01.

BOUNDARY < (HSCALE=INFO | SAMPLESIZE) >

displays a plot of the resulting sequential boundaries with the acceptance and rejection regions for each design. Either the information level (HSCALE=INFO) or the sample size (HSCALE=SAMPLESIZE) is displayed on the horizontal axis. If the maximum information is not available for the design, the information in percentage of its corresponding fixed-sample design are used in the plot. The stage number for each stage is displayed inside the plot. The default is HSCALE=INFO.

If the HSCALE=SAMPLESIZE option is specified. the SAMPLESIZE statement must also be specified. The options MODEL=INPUTNEVENTS, MODEL=TWOSAMPLESURVIVAL, and MODEL=PHREG in the SAMPLESIZE statement indicate survival data. For a sample that does not contain survival data, the sample size at each stage is displayed on the horizontal axis. For survival data, the number of events is displayed on the horizontal axis at each stage. The critical values for the corresponding fixed-sample design are also displayed in the plot.

COMBINEDBOUNDARY < (HSCALE=INFO | SAMPLESIZE | STAGE) >

displays a plot of the resulting sequential boundaries for all designs simultaneously. You can

display the information level (HSCALE=INFO), the sample size (HSCALE=SAMPLESIZE), or the stage number (HSCALE=STAGE) on the horizontal axis. The default is HSCALE=INFO. With HSCALE=INFO, if the maximum information is not available for the design, then the information in percentage of its corresponding fixed-sample design is used in the plot.

If the HSCALE=SAMPLESIZE option is specified, the SAMPLESIZE statement must also be specified. The options MODEL=INPUTNEVENTS, MODEL=TWOSAMPLESURVIVAL, and MODEL=PHREG in the SAMPLESIZE statement indicate survival data. For a sample that does not contain survival data, the sample size at each stage is displayed on the horizontal axis. For survival data, the number of events is displayed on the horizontal axis at each stage.

ERRSPEND < (HSCALE=INFO | STAGE) >

displays a plot of the error spending for all sequential boundaries in the designs simultaneously. You can display the information level (HSCALE=INFO) or the stage number (HSCALE=STAGE) on the horizontal axis. With HSCALE=INFO, the information fractions are used in the plot. The default is HSCALE=STAGE.

NONE

suppresses all plots.

POWER < (CREF= *numbers* **) >**

displays a plot of the power curves under various hypothetical references, where the numbers $c_i \geq 0$.

For a one-sided design, powers under hypotheses $\theta = c_i\,\theta_1$ are displayed, where θ_1 is the alternative reference and c_i are the values specified in the CREF= option.

For a two-sided design, powers under hypotheses $\theta = c_i\theta_{1l}$ and $\theta = c_i\theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively. Note that for a symmetric two-sided design, only powers under hypotheses $\theta = c_i\,\theta_{1u}$ are derived. The default is CREF= 0 to 1.5 by 0.01.

DESIGN Statement

< label: > **DESIGN** *< options >* **;**

The DESIGN statement requests a new group sequential design. You can use multiple DESIGN statements, and each DESIGN statement corresponds to a separate group sequential design.

Table 77.2 lists the options available in the DESIGN statement.

Table 77.2 Design Statement Options

Option	Description
Design Parameters	
ALPHA=	specifies Type I error probability level α
ALT=	specifies type of alternative hypothesis
BETA=	specifies Type II error probability level β

Table 77.2 *continued*

Option	Description
BETAOVERLAP=	checks for overlapping of the lower and upper β boundaries in a two-sided design with error spending methods
BOUNDARYKEY=	specifies type of error probability to maintain
INFO=	specifies information levels
NSTAGES=	specifies number of stages
STOP=	specifies condition for early stopping
Boundary Methods	
METHOD=	specifies methods for boundary values

The required NSTAGES= option specifies the number of stages. The METHOD= option is required if the number of stages specified in the NSTAGES= option is greater than one. The following options can be used in the DESIGN statement. They are listed in alphabetical order.

ALPHA=α < (< LOWER=α_l > < UPPER=α_u >) >

specifies the Type I error probability α. The default is $\alpha = 0.05$. The LOWER= and UPPER= options are applicable only for the two-sided design. The LOWER= option specifies the lower Type I error probability α_l, and the upper Type I error probability is computed as $\alpha_u = \alpha - \alpha_l$. The UPPER= option specifies the upper Type I error probability α_u, and the lower Type I error probability is computed as $\alpha_l = \alpha - \alpha_u$. If both LOWER= and UPPER= options are not specified, $\alpha_l = \alpha_u = \alpha/2$.

If both the MAXINFO= and ALTREF= options are specified, then the Type I and Type II error probability levels cannot be met simultaneously. In this case, the ALPHA= option is applicable only with the default BOUNDARYKEY=ALPHA option, and the Type II error probability β is derived.

ALT=LOWER | UPPER | TWOSIDED

specifies the type of alternative hypothesis in the design. For a test of $H_0 : \theta = 0$, the keywords LOWER, UPPER, and TWOSIDED correspond to the alternatives of $\theta < 0$, $\theta > 0$, and $\theta \neq 0$, respectively. The default is ALT=TWOSIDED.

BETA=β < (< LOWER=β_l > < UPPER=β_u >) >

specifies the Type II error probability level β. The default is $\beta = 0.10$. The LOWER= and UPPER= options are applicable only for the two-sided design. The LOWER= option specifies the lower Type II error probability level β_l, and the UPPER= option specifies the upper Type II error probability level β_u. If the LOWER= or UPPER= option is not specified, β is used.

If both the MAXINFO= and ALTREF= options are specified, then the Type I and Type II error probability levels cannot be met simultaneously. In this case, the BETA= option is applicable only with the BOUNDARYKEY=BETA option, and the Type I error probability α is derived.

BETAOVERLAP=ADJUST | NOADJUST

OVERLAP=ADJUST | NOADJUST

specifies whether to check for overlapping of the lower and upper β boundaries for the two corresponding one-sided tests. This option applies to two-sided designs with STOP=ACCEPT or STOP=BOTH that are constructed with error spending methods, and this

type of overlapping might result from a small β spending at an interim stage. When you specify BETAOVERLAP=ADJUST, the procedure checks for this type of overlapping. If such overlapping is found, the β boundaries for the two-sided design at that stage are set to missing, and the β spending values at subsequent stages are adjusted, as described in the section "Boundary Adjustments for Overlapping Lower and Upper β Boundaries" on page 5850".

You can specify BETAOVERLAP=NOADJUST to request that no adjustment be made. The default is BETAOVERLAP=ADJUST.

BOUNDARYKEY=ALPHA | BETA | BOTH | NONE

specifies types of errors to be maintained in the resulting boundary. The default is BOUNDARYKEY=ALPHA if both ALTREF= and MAXINFO= options are specified. Otherwise, the default is BOUNDARYKEY=NONE for Whitehead methods with the STOP=BOTH option, and it is BOUNDARYKEY=BOTH for others.

See the section "Applicable Boundary Keys" on page 5853 for a detailed description of applicable boundary keys.

INFO=EQUAL

INFO=CUM(*numbers* **)**

specifies relative information levels for all stages in the design. The INFO=EQUAL option specifies equally spaced information levels, and the INFO=CUM option specifies cumulative relative information levels. The default is INFO=EQUAL.

If the number of information levels specified in the INFO=CUM option is less than the number of stages specified in the NSTAGES= option, the last available information increment is used as the information increment for each subsequent stage.

METHOD=WHITEHEAD < (TAU=τ < (<LOWER=τ_l > <UPPER=τ_u >) >) >

METHOD=*method*

METHOD(*boundary***) =** *method*

specifies the methods for the boundaries in the design, where $0 \leq \tau < 0.5$.

For a one-sided design, an α boundary is created with the STOP=REJECT or STOP=BOTH option, and a β boundary is created with the STOP=ACCEPT or STOP=BOTH option. For a two-sided design, lower and upper α boundaries are created with the STOP=REJECT or STOP=BOTH option, and lower and upper β boundaries are created with the STOP=ACCEPT or STOP=BOTH option.

There are three types of methods available in the SEQDESIGN procedure. The unified family methods and Haybittle-Peto methods derive boundary values with fixed boundary shape; the Whitehead methods derive boundary values by adjusting the boundary values generated from continuous monitoring; and the error spending methods derive the boundary values from the specified errors used at each stage. You can specify different methods for the same design, but all methods must be from the same group.

For a design with early stopping to reject or accept the null hypothesis, the METHOD=WHITEHEAD option uses Whitehead's triangular design and double-triangular design for a one-sided design and two-sided design, respectively (Whitehead and Stratton 1983; Whitehead 1997, 2001). For a design with early stopping only to reject the null hypothesis or only to accept the null hypothesis, you can specify the slope of the boundary line in the score statistic scale with the TAU= τ option. The default is TAU=0.25. See the section "Whitehead Methods" on page 5845 for a detailed description of the Whitehead methods.

The following options specify available error spending methods for the boundary. Each of these methods can be specified with the METHOD= option for all boundaries, or with the METHOD(*boundary*) = option for an individual boundary. See the section "Error Spending Methods" on page 5848 for a detailed description of these error spending methods.

ERRFUNCGAMMA < (GAMMA= γ **) >**
specifies a gamma cumulative error spending function for the boundary (Hwang, Shih, and DeCani 1990). The GAMMA= option specifies the gamma parameter γ in the function, where $\gamma \leq 3$. The boundaries created with $\gamma = 1$ are similar to the boundaries from the Pocock method, and the boundaries created with $\gamma = -4$ or $\gamma = -5$ are similar to the boundaries from the O'Brien-Fleming method. The default is GAMMA=−2, which is the average of $\gamma = 1$ and $\gamma = -5$.

ERRFUNCOBF
specifies the O'Brien-Fleming-type cumulative error spending function for the boundary (Lan and DeMets 1983).

ERRFUNCPOC
specifies the Pocock-type cumulative error spending function for the boundary (Lan and DeMets 1983).

ERRFUNCPOW < (RHO=ρ **) >**
specifies a power cumulative error spending function for the boundary (Jennison and Turnbull 2000, p. 148). The RHO= option specifies the power parameter ρ in the function, where $\rho \geq 0.25$. The boundaries created with $\rho = 1$ are similar to the boundaries from the Pocock method, and the boundaries created with $\rho = 3$ are similar to the boundaries from the O'Brien-Fleming method. The default is RHO=2, which is the average of $\rho = 1$ and $\rho = 3$.

ERRSPEND (*numbers* **)**
specifies the relative cumulative error spending at each stage.

With a fixed boundary shape, you can use the following available Haybittle-Peto methods and unified family methods to derive the boundary. You can specify each of these methods in the METHOD= option for all boundaries, or in the METHOD(*boundary*) = option for an individual boundary. See the section "Haybittle-Peto Method" on page 5844 for a detailed description of the Haybittle-Peto methods, and see the section "Unified Family Methods" on page 5839 for a detailed description of unified family methods.

HP | HAYBITTLE | PETO < (Z= *numbers* **| PVALUE=** *numbers* **) >**
specifies the Haybittle-Peto method (Haybittle 1971; Peto et al. 1976). The values specified are used to create the boundary values. The boundary value at the final stage

can be derived in the procedure to maintain the Type I and Type II error probability levels. The default is Z=3.

OBF | OBRIENFLEMING
specifies the O'Brien-Fleming method (O'Brien and Fleming 1979). The O'Brien-Fleming method is equivalent to a power family method with RHO=0.5.

POC | POCOCK
specifies the Pocock method (Pocock 1977). The Pocock method is equivalent to a power family method with RHO=0.

POW | POWER < (RHO=ρ) >
specifies a power family method (Wang and Tsiatis 1987; Emerson and Fleming 1989; Pampallona and Tsiatis 1994). The RHO= option specifies the power parameter ρ in the power family method, where $\rho \geq -0.25$. The power family method with $\rho = 0$ corresponds to the Pocock method, and the power family method with $\rho = 0.5$ corresponds to the O'Brien-Fleming method. The default is RHO=0.25, a value halfway between the Pocock and O'Brien-Fleming methods. A power family method is equivalent to a unified family method with RHO=ρ and TAU=0.

TRI | TRIANGULAR < (TAU=τ) >
specifies a unified family triangular method (Kittelson and Emerson 1999), where $0 \leq \tau \leq 1$. The default is TAU=1.0. The triangular method is identical to the unified family method with RHO=0.5 and TAU=τ. Note that this unified family triangular method is different from Whitehead's triangular method.

UNI | UNIFIED < (< TAU=τ > < RHO=ρ >) >
specifies a unified family method (Kittelson and Emerson 1999). The TAU= and RHO= options specify the τ and ρ parameters in a unified family method, respectively, where $\rho \geq 0$ and $0 \leq \tau \leq 2\rho$. The defaults are TAU=0 and RHO=0.25. See the section "Unified Family Methods" on page 5839 for a detailed description of the unified family methods.

The O'Brien-Fleming, Pocock, power family, and triangular methods are all special cases of the unified family methods. Table 77.3 summarizes the corresponding parameters in the unified family for these methods.

Table 77.3 Parameters in the Unified Family for Various Methods

Method	Option	Unified Family	
		Rho	Tau
Pocock	POC	0	0
O'Brien-Fleming	OBF	0.5	0
Power family	POW (RHO=ρ)	ρ	0
Triangular	TRI (TAU=τ)	0.5	τ

Note that the power parameter $\rho = 1/2-\Delta = \rho^*-1/2$, where Δ is the power parameter used in Jennison and Turnbull (2000) and Wang and Tsiatis (1987) and ρ^* is the power parameter used in Kittelson and Emerson (1999).

If a method with specified parameters is used for all boundaries in the design, you can use the METHOD= option to specify the method. Otherwise, you can use the following METHOD(*boundary*)= options to specify different methods from the same group for the boundaries.

METHOD(ALPHA)=*method*

METHOD(REJECT)=*method*

specifies the method for the α boundary of a one-sided design or the lower and upper α boundaries for a two-sided design.

METHOD(LOWERALPHA)=*method*

METHOD(LOWERREJECT)=*method*

specifies the method for the lower α boundary of a two-sided design.

METHOD(UPPERALPHA)=*method*

METHOD(UPPERREJECT)=*method*

specifies the method for the upper α boundary of a two-sided design.

METHOD(BETA)=*method*

METHOD(ACCEPT)=*method*

specifies the method for the β boundary of a one-sided design or the lower and upper β boundaries for a two-sided design.

METHOD(LOWERBETA)=*method*

METHOD(LOWERACCEPT)=*method*

specifies the method for the lower β boundary of a two-sided design.

METHOD(UPPERBETA)=*method*

METHOD(UPPERACCEPT)=*method*

specifies the method for the upper β boundary of a two-sided design.

NSTAGES=*number*

specifies the number of stages for the design. This option is required in the DESIGN statement, and the maximum allowed number of stages is 25.

STOP=ACCEPT | REJECT | BOTH

specifies the condition of early stopping for the design. The keywords ACCEPT, REJECT, and BOTH correspond to early stopping only to accept, only to reject, and either to accept or reject the null hypothesis H_0, respectively. The default is STOP=REJECT.

SAMPLESIZE Statement

SAMPLESIZE < *MODEL= option* > **;**

If each observation in the data set provides one unit of information in a hypothesis testing such as a one-sample test for the mean, the SAMPLESIZE statement computes the required sample sizes for the sequential design specified in each DESIGN statement. However, for a survival analysis, an individual in the survival time data might provide only partial information because of censoring. For this hypothesis, the SAMPLESIZE statement computes the required numbers of events. With additional accrual information in a survival analysis, the sample sizes can also be computed.

Only one SAMPLESIZE statement can be specified. For each specified group sequential design, the SAMPLESIZE statement computes the required sample sizes or numbers of events. The SAMPLESIZE statement is not required if the SEQDESIGN procedure is used only to compare features among different designs. Table 77.4 lists the options available in the SAMPLESIZE statement.

Table 77.4 SAMPLESIZE Statement Options

Option	Description
Fixed-Sample Models	
INPUTNOBS	specifies sample size for fixed-sample design
INPUTNEVENTS	specifies number of events for fixed-sample design
One-Sample Models	
ONESAMPLEMEAN	specifies one-sample Z test for mean
ONESAMPLEFREQ	specifies one-sample test for binomial proportion
Two-Sample Models	
TWOSAMPLEMEAN	specifies two-sample Z test for mean difference
TWOSAMPLEFREQ	specifies two-sample test for binomial proportions
TWOSAMPLESURVIVAL	specifies log-rank test for two survival distributions
Regression Models	
REG	specifies test for a regression parameter
LOGISTIC	specifies test for a logistic regression parameter
PHREG	specifies test for a proportional hazards regression parameter

The MODEL= option specifies the input sample size or number of events from a fixed-sample study, or it specifies a statistical model to compute the required sample size. The MODEL=INPUTNOBS option specifies the input sample size from a fixed-sample study of nonsurvival data, and the MODEL=INPUTNEVENTS option specifies the number of events from a fixed-sample study of survival data. The remaining MODEL= options specify the statistical models used to compute the required sample size. The default is MODEL=TWOSAMPLEMEAN, the two-sample Z test for the mean difference.

With the MODEL=INPUTNOBS or MODEL=INPUTNEVENTS option, the required sample size or number of events for the group sequential trial is computed by multiplying the input sample size or number of events by the ratio between the design information level and its corresponding fixed-sample information level. This ratio can be obtained by dividing the Max Information (Percent

Fixed-Sample) in the "Design Information" table by 100. See the section "Design Information" on page 5877 for a description of the "Design Information" table.

Fixed-Sample Models

The following two options compute the required sample size or number of events for a group sequential trial by using the sample size or number of events for the fixed-sample design.

MODEL=INPUTNOBS < (*options*) >
specifies the sample size information for a fixed-sample design. The available options are as follows:

- N= n
- SAMPLE= ONE | TWO
- WEIGHT= w_a < w_b >
- MATCHNOBS= YES | NO

The required N=n option specifies the sample size n for the fixed-sample design. The SAMPLE=ONE option specifies a one-sample test, and the SAMPLE=TWO option specifies a two-sample test. The default is SAMPLE=ONE.

With a two-sample test, the WEIGHT= option specifies the sample size allocation weights for the two groups. If w_b is not specified, $w_b = 1$ is used. The default is WEIGHT=1, equal allocation for the two groups. The derived fractional sample sizes are rounded up to integers, and the MATCHNOBS=YES option requests these integer sample sizes to match the sample size allocation.

See the section "Input Sample Size for Fixed-Sample Design" on page 5856 for a detailed description of the input sample size for the fixed-sample design in sample size computation.

MODEL=INPUTNEVENTS < (*options*) >
specifies the number of events D for a fixed-sample survival test. The available options are as follows:

- D= d
- SAMPLE= ONE | TWO

The required D=d option specifies the fixed-sample number of events d. The SAMPLE=ONE option specifies a one-sample test, and the SAMPLE=TWO option specifies a two-sample test. The default is SAMPLE=ONE.

In order to derive the sample size, addition options are needed. The available options for the sample size computation are as follows:

- HAZARD= h_a < h_b >
- MEDSURVTIME= t_a < t_b >
- WEIGHT= w_a < w_b >
- ACCRATE= r_a
- ACCTIME= T_a

- FOLTIME= T_f
- TOTALTIME= T

The hazard rates are needed for the sample size computation. For a one-sample test, the HAZARD= h_a option specifies the hazard rate h_a explicitly, and the MEDSURVTIME=t_a option specifies the hazard rate implicitly through the median survival time t_a. Similarly, for a two-sample test, the HAZARD= h_a h_b option specifies the hazard rates h_a and h_b for groups A and B explicitly, and the MEDSURVTIME=t_a t_b option specifies hazard rates for groups A and B implicitly through the median survival times t_a and t_b. Also, for a two-sample test, $h_b = h_a$ if h_b is not specified and $t_b = t_a$ if t_b is not specified.

With a two-sample test, the WEIGHT= option specifies the sample size allocation weights for the two groups. If w_b is not specified, $w_b = 1$ is used. The default is WEIGHT=1.

Assuming that the hazard rates are constant and the individual accrual is uniform in the accrual time T_a with a constant accrual rate r_a, the sample size and study time can be derived.

The ACCRATE= option specifies the constant accrual rate r_a, and the ACCTIME= and FOLTIME= options specify the accrual time T_a and follow-up time T_f, respectively. The TOTALTIME= option specifies the total study time, $T = T_a + T_f$.

If the ACCRATE= option is specified, then one of the ACCTIME=, FOLTIME, and TOTALTIME options is required for the sample size computation. Otherwise, two of the ACCTIME=, FOLTIME, and TOTALTIME options are required to compute the accrual rate and sample size.

See the section "Input Number of Events for Fixed-Sample Design" on page 5856 for a detailed description of the input number of events for the fixed-sample design in sample size computation.

One-Sample Models

The following two options compute the required sample size for a one-sample group sequential test.

MODEL=ONESAMPLEMEAN < (*options*) >

specifies the one-sample Z test for mean. The available options are as follows:

- MEAN= μ_1
- STDDEV= σ

The MEAN= option specifies the alternative reference μ_1 and is required if the alternative reference is not specified or derived in the procedure. If the MEAN=option is not specified, the specified or derived alternative reference is used.

The STDDEV= option specifies the standard deviation σ. The default is STDDEV=1. See the section "Test for a Normal Mean" on page 5858 for a detailed description of the one-sample Z test for mean.

Note that the one-sample Z test for mean also includes the paired difference in two-treatment comparison (Jennison and Turnbull 2000, pp. 51–52), where μ_1 is the mean of differences within pairs under the alternative hypothesis and σ is the standard deviation for the mean of differences within pairs.

MODEL=ONESAMPLEFREQ < (*options*) >

specifies the one-sample test for binomial proportion with the null hypothesis $H_0 : \theta = 0$ and the alternative hypothesis $H_1 : \theta = \theta_1$, where $\theta = p - p_0$ and $\theta_1 = p_1 - p_0$. The available options are as follows:

- NULLPROP= p_0
- PROP= p_1
- REF= NULLPROP | PROP

The NULLPROP= and PROP= options specify the proportions under the null and alternative hypotheses, respectively. The default for the null reference is NULLPROP=0.5. The PROP= option is required if the alternative reference is not specified or derived in the procedure. If the PROP= option is not specified, the specified or derived alternative reference θ_1 is used to compute the alternative reference $p_1 = p_0 + \theta_1$.

The REF= option specifies the hypothesis under which the proportion is used in the sample size computation. The REF=NULLPROP option uses the null hypothesis, and the REF=PROP option uses the alternative hypothesis to compute the sample size. The default is REF=PROP. See the section "Test for a Binomial Proportion" on page 5859 for a detailed description of the one-sample tests for proportion.

Two-Sample Models

The following three options compute the required sample size or number of events for a two-sample group sequential trial.

MODEL=TWOSAMPLEMEAN < (*options*) >

specifies the two-sample Z test for mean difference. The available options are as follows:

- MEANDIFF= θ_1
- STDDEV= σ_a $< \sigma_b >$
- WEIGHT= w_a $< w_b >$
- MATCHNOBS= YES | NO

The MEANDIFF= option specifies the alternative reference θ_1 and is required if the alternative reference is not specified or derived in the procedure. If the MEANDIFF= option is not specified, the specified or derived alternative reference is used.

The STDDEV= option specifies the standard deviations σ_a and σ_b. If σ_b is not specified, $\sigma_b = \sigma_a$. The default is STDDEV=1.

The WEIGHT= option specifies the sample size allocation weights for the two groups. If w_b is not specified, $w_b = 1$ is used. The default is WEIGHT=1, equal sample size for the two groups. The derived fractional sample sizes are rounded up to integers, and the MATCHNOBS=YES option requests these integer sample sizes to match the sample size allocation. The default is MATCHNOBS=NO.

See the section "Test for the Difference between Two Normal Means" on page 5860 for a detailed description of the two-sample Z test for mean difference.

MODEL=TWOSAMPLEFREQ < (*options*) >

specifies the two-sample test for binomial proportions. The available options are as follows:

- NULLPROP= $p_{0a} < p_{0b} >$
- PROP= p_{1a}
- TEST= PROP | LOGOR | LOGRR
- REF= NULLPROP | PROP | AVGNULLPROP | AVGPROP
- WEIGHT= $w_a < w_b >$
- MATCHNOBS= YES | NO

The NULLPROP= option specifies proportions $p_a = p_{0a}$ and $p_b = p_{0b}$ in groups A and B, respectively, under the null hypothesis. If p_{0b} is not specified, $p_{0b} = p_{0a}$. The default is NULLPROP=0.5.

The PROP= option specifies proportion $p_a = p_{1a}$ in group A under the alternative hypothesis. The proportion p_{1b} in group B under the alternative hypothesis is given by $p_{1b} = p_{0b}$. The PROP= option is required if the alternative reference is not specified or derived in the procedure. If the PROP= option is not specified, the specified or derived alternative reference is used to compute p_{1a}, the proportion in group A under the alternative hypothesis.

The TEST= option specified the null hypothesis $H_0 : \theta = 0$ in the test. The TEST=PROP option uses the difference in proportions $\theta = (p_a - p_b) - (p_{0a} - p_{0b})$, the TEST=LOGOR option uses the log odds-ratio test $\theta = \delta - \delta_0$, where

$$\delta = \log\left(\frac{p_a(1-p_b)}{p_b(1-p_a)}\right) \qquad \delta_0 = \log\left(\frac{p_{0a}(1-p_{0b})}{p_{0b}(1-p_{0a})}\right)$$

and the TEST=LOGRR option uses the log relative risk test with $\theta = \delta - \delta_0$, where

$$\delta = \log\left(\frac{p_a}{p_b}\right) \qquad \delta_0 = \log\left(\frac{p_{0a}}{p_{0b}}\right)$$

The default is TEST=LOGOR.

The REF= option specifies the hypothesis under which the proportions are used in the sample size computation. The REF=NULLPROP option uses the null proportions p_{0a} and p_{0b}, the REF=PROP option uses the alternative proportions p_{1a} and p_{1b}, the REF=AVGNULLPROP option uses the average null proportion, and the REF=AVGPROP option uses the average alternative proportion. The default is REF=PROP.

The WEIGHT= option specifies the sample size allocation weights for the two groups. If w_b is not specified, $w_b = 1$ is used. The default is WEIGHT=1, equal sample size for the two groups. The derived fractional sample sizes are also rounded up to integers, and the MATCHNOBS=YES option requests that these integer sample sizes match the sample size allocation.

See the section "Test for the Difference between Two Binomial Proportions" on page 5862, the section "Test for Two Binomial Proportions with a Log Odds Ratio Statistic" on page 5864, and the section "Test for Two Binomial Proportions with a Log Relative Risk Statistic" on page 5865 for a detailed description of the two-sample tests for proportions.

MODEL=TWOSAMPLESURVIVAL < (*options*) >

MODEL=TWOSAMPLESURV < (*options*) >

specifies the log-rank test for two survival distributions with the null hypothesis $H_0 : \theta = \delta - \delta_0 = 0$, where the parameter $\delta = -\log(h_a/h_b)$, δ_0 is the value of δ under the null hypothesis and the values h_a and h_b are the hazard rates for groups A and B, respectively.

The available options for the number of events are as follows:

- NULLHAZARD= $h_{0a} < h_{0b} >$
- NULLMEDSURVTIME= $t_{0a} < t_{0b} >$
- HAZARD= h_{1a}
- MEDSURVTIME= t_{1a}
- HAZARDRATIO= λ_1

The NULLHAZARD= option specifies hazard rates $h_a = h_{0a}$ and $h_b = h_{0b}$ for groups A and B, respectively, under the null hypothesis. If h_{0b} is not specified, $h_{0b} = h_{0a}$. The NULLMEDSURVTIME= option specifies the median survival times $t_a = t_{0a}$ and $t_b = t_{0b}$ under the null hypothesis. If t_{0b} is not specified, $t_{0b} = t_{0a}$. If both NULLHAZARD= and NULLMEDSURVTIME= option are not specified, NULLHAZARD=0.06931, which corresponds to NULLMEDSURVTIME=10, is used.

The hazard rate for group B under the alternative hypothesis $h_{1b} = h_{0b}$, as the hazard rate under the null hypothesis. The HAZARD=, MEDSURVTIME=, and HAZARDRATIO= options specify the group A hazard rate h_{1a}, the group A median survival time t_{1a}, and the hazard ratio $\lambda_1 = h_{1a}/h_{1b}$, respectively, under the alternative hypothesis. The HAZARD=, MEDSURVTIME=, or HAZARDRATIO= option is required if the alternative reference is not specified or derived in the procedure. If these three options are not specified, the specified or derived alternative reference θ_1 is used to compute h_{1a} from the equation:

$$\theta_1 = -\log(\frac{h_{1a}}{h_{1b}}) - (-\log(\frac{h_{0a}}{h_{0b}})) = -\log(\frac{h_{1a}}{h_{0a}})$$

In order to derive the sample size, additional options are needed. The available options for the sample size computation are as follows:

- REF= NULLHAZARD | HAZARD
- WEIGHT= $w_a < w_b >$
- ACCRATE= r_a
- ACCTIME= T_a
- FOLTIME= T_f
- TOTALTIME= T

The REF= option specifies the hypothesis under which the hazard is used in the sample size computation. The REF=NULLHAZARD option uses the null hypothesis, and the REF=HAZARD option uses the alternative hypothesis. The default is REF=HAZARD.

The WEIGHT= option specifies the sample size allocation weights for the two groups. If w_b is not specified, $w_b = 1$ is used. The default is WEIGHT=1, equal sample size for the two groups.

With the available maximum information, the number of events can be derived for the specified hypothesis. Assuming that the hazard rates are constant and the individual accrual is uniform in the accrual time T_a with a constant accrual rate r_a, the sample size and study time can be derived.

The ACCRATE= option specifies the constant accrual rate r_a, and the ACCTIME= and FOLTIME= options specify the accrual time T_a and follow-up time T_f, respectively. The TOTALTIME= option specifies the total study time, $T = T_a + T_f$.

If the ACCRATE= option is specified, then one of the ACCTIME=, FOLTIME=, and TOTALTIME= options is required for the sample size computation. Otherwise, two of the ACCTIME=, FOLTIME=, and TOTALTIME= options are required to compute the accrual rate and sample size.

See the section "Test for Two Survival Distributions with a Log-Rank Test" on page 5866 for a detailed description of the two-sample log-rank test for survival data.

Regression Models

The following three options compute the required sample size or number of events for group sequential tests on a regression parameter.

MODEL=REG < (*options*) >

specifies the Z test for a normal regression parameter. The available options are as follows:

- BETA= β_1
- VARIANCE= σ_y^2
- XVARIANCE= σ_x^2
- XRSQUARE= r_x^2

The BETA= option specifies the alternative reference β_1 and is required if the alternative reference is not specified or derived in the procedure. If the BETA= option is not specified, $\beta_1 = \theta_1$, the specified or derived alternative reference.

The VARIANCE= and XVARIANCE= options specify the variances for the response variable Y and covariate X, respectively. The defaults are VARIANCE=1 and XVARIANCE=1. For a model with more than one covariate, the XRSQUARE= option can be used to derive the variance of X after adjusting for other covariates. The default is XRSQUARE=0.

See the section "Test for a Parameter in the Regression Model" on page 5869 for a detailed description of the Z test for the regression parameter.

MODEL=LOGISTIC < (*options*) >

specifies the Z test for a logistic regression parameter. The available options are as follows:

- BETA= β_1
- PROP= p
- XVARIANCE= σ_x^2
- XRSQUARE= r_x^2

The BETA= option specifies the alternative reference β_1 and is required if the alternative reference is not specified or derived in the procedure. If the BETA= option is not specified, $\beta_1 = \theta_1$, the specified or derived alternative reference.

The PROP= option specifies the proportion of the binary response variable Y. The default is PROP=0.5. The XVARIANCE= option specifies the variance of the covariate X. The default is XVARIANCE=1. For a model with more than one covariate, the XRSQUARE= option can be used to derive the variance of X after adjusting for other covariates. The default is XRSQUARE=0.

See the section "Test for a Parameter in the Logistic Regression Model" on page 5870 for a detailed description of the Z test for the logistic regression parameter.

MODEL=PHREG < (*options*) >

specifies the Z test for a proportional hazards regression parameter. The available options for the number of events are as follows:

- BETA= β_1
- XVARIANCE= σ_x^2
- XRSQUARE= r_x^2

The BETA= option specifies the alternative reference β_1 and is required if the alternative reference is not specified or derived in the procedure. If the BETA= option is not specified, $\beta_1 = \theta_1$, the specified or derived alternative reference.

The XVARIANCE= option specifies the variance of the covariate X. The default is XVARIANCE=1. For a model with more than one covariate, the XRSQUARE= option can be used to derive the variance of X after adjusting for other covariates. The default is XRSQUARE=0.

In order to derive the sample size, additional options are needed. The available options for the sample size computation are as follows:

- HAZARD= h_a
- MEDSURVTIME= t_a
- ACCRATE= r_a
- ACCTIME= T_a
- FOLTIME= T_f
- TOTALTIME= T

The hazard rate is required for the sample size computation. The HAZARD= h_a option specifies the hazard rate h_a explicitly, and the MEDSURVTIME=t_a option specifies the hazard rate implicitly through the median survival time t_a.

Assuming that the hazard rates are constant and the individual accrual is uniform in the accrual time T_a with a constant accrual rate r_a, the sample size and study time can be derived.

The ACCRATE= option specifies the constant accrual rate r_a, the ACCTIME= option specifies the accrual time T_a, and the FOLTIME= option specifies the follow-up time T_f. The TOTALTIME= option specifies the total study time, $T = T_a + T_f$.

If the ACCRATE= option is specified, then one of the ACCTIME=, FOLTIME=, and TOTALTIME= options is required for the sample size computation. Otherwise, two of the ACCTIME=, FOLTIME=, and TOTALTIME= options are required to compute the accrual rate and sample size.

See the section "Test for a Parameter in the Proportional Hazards Regression Model" on page 5871 for a detailed description of the Z test for the proportional hazards regression parameter.

Details: SEQDESIGN Procedure

Fixed-Sample Clinical Trials

A clinical trial is a research study in consenting human beings to answer specific health questions. One type of trial is a treatment trial, which tests the effectiveness of an experimental treatment. An example is a planned experiment designed to assess the efficacy of a treatment in humans by comparing the outcomes in a group of patients who receive the test treatment with the outcomes in a comparable group of patients who receive a placebo control treatment, where patients in both groups are enrolled, treated, and followed over the same time period.

A clinical trial is conducted according to a plan called a protocol. The protocol provides detailed description of the study. For a fixed-sample trial, the study protocol contains detailed information such as the null hypothesis, the one-sided or two-sided test, and the Type I and II error probability levels. It also includes the test statistic and its associated critical values in the hypothesis testing.

Generally, the efficacy of a new treatment is demonstrated by testing a hypothesis $H_0 : \theta = 0$ in a clinical trial, where θ is the parameter of interest. For example, to test whether a population mean μ is greater than a specified value μ_0, $\theta = \mu - \mu_0$ can be used with an alternative $\theta > 0$.

A one-sided test is a test of the hypothesis with either an upper (greater) or a lower (lesser) alternative, and a two-sided test is a test of the hypothesis with a two-sided alternative. The drug industry often prefers to use a one-sided test to demonstrate clinical superiority based on the argument that a study should not be run if the test drug would be worse (Chow, Shao, and Wang 2003, p. 28). But in practice, two-sided tests are commonly performed in drug development (Senn 1997, p. 161). For a fixed Type I error probability α, the sample sizes required by one-sided and two-sided tests are

different. Refer to Senn (1997, pp. 161–167) for a detailed description of issues involving one-sided and two-sided tests.

For independent and identically distributed observations $y_1, y_2, \ldots, y_n$ of a random variable, the likelihood function for θ is

$$L(\theta) = \prod_{j=1}^{n} L_i(\theta)$$

where θ is the population parameter and $L_i(\theta)$ is the probability or probability density of y_i. Using the likelihood function, two statistics can be derived that are useful for inference: the maximum likelihood estimator and the score statistic.

Maximum Likelihood Estimator

The maximum likelihood estimate (MLE) of θ is the value $\hat{\theta}$ that maximizes the likelihood function for θ. Under mild regularity conditions, $\hat{\theta}$ is an asymptotically unbiased estimate of θ with variance $1/E_\theta(I(\theta))$, where $I(\theta)$ is the Fisher information

$$I(\theta) = -\frac{\partial^2 \log(L(\theta))}{\partial \theta^2}$$

and $E_\theta(I(\theta))$ is the expected Fisher information (Diggle et al. 2002, p. 340)

$$E_\theta(I(\theta)) = -E_\theta\left(\frac{\partial^2 \log(L(\theta))}{\partial \theta^2}\right)$$

The score function for θ is defined as

$$S(\theta) = \frac{\partial \log(L(\theta))}{\partial \theta}$$

and usually, the MLE can be derived by solving the likelihood equation $S(\theta) = 0$. Asymptotically, the MLE is normally distributed (Lindgren 1976, p. 272):

$$\hat{\theta} \sim N\left(\theta, \frac{1}{E_\theta(I(\theta))}\right)$$

If the Fisher information $I(\theta)$ does not depend on θ, then $I(\theta)$ is known. Otherwise, either the expected information evaluated at the MLE $\hat{\theta}$ ($E_{\theta=\hat{\theta}}(I(\theta))$) or the observed information $I(\hat{\theta})$ can be used for the Fisher information (Cox and Hinkley, 1974, p. 302; Efron and Hinkley, 1978, p. 458), where the observed Fisher information

$$I(\hat{\theta}) = -\left(\frac{\partial^2 \log(L(\theta))}{\partial \theta^2} \mid \theta = \hat{\theta}\right)$$

If the Fisher information $I(\theta)$ does depend on θ, the observed Fisher information is recommended for the variance of the maximum likelihood estimator (Efron and Hinkley, 1978, p. 457).

Thus, asymptotically, for large n,

$$\hat{\theta} \sim N\left(\theta, \frac{1}{I}\right)$$

where I is the information, either the expected Fisher information $E_{\theta=0}(I(\theta))$ or the observed Fisher information $I\hat{\theta})$.

So to test $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$, you can use the standardized Z test statistic

$$Z = \frac{\hat{\theta}}{\sqrt{\text{Var}(\hat{\theta})}} = \hat{\theta}\sqrt{I} \sim N\,(0,\ 1)$$

and the two-sided p-value is given by

$$\text{Prob}(|Z| > |z_0|) = 1 - 2\Phi(|z_0|)$$

where Φ is the cumulative standard normal distribution function and z_0 is the observed Z statistic.

If the BOUNDARYSCALE=SCORE is specified in the SEQDESIGN procedure, the boundary values for the test statistic are displayed in the score statistic scale. With the standardized Z statistic, the score statistic $S = Z\sqrt{I} = \hat{\theta}I$ and

$$S \sim N\,(0,\ I)$$

Score Statistic

The score statistic is based on the score function for θ,

$$S(\theta) = \frac{\partial \log(L(\theta))}{\partial\theta}$$

Under the null hypothesis $H_0 : \theta = 0$, the score statistic $S(0)$ is the first derivative of the log likelihood evaluated at the null reference 0:

$$S(0) = \frac{\partial \log(L(\theta))}{\partial\theta} \,|\, \theta = 0$$

Under regularity conditions, $S(0)$ is asymptotically normally distributed with mean zero and variance $E_{\theta=0}(I(\theta))$, the expected Fisher information evaluated at the null hypothesis $\theta = 0$ (Kalbfleisch and Prentice, 1980, p. 45), where $I(\theta)$ is the Fisher information

$$I(\theta) = -E\left(\frac{\partial^2 \log(L(\theta))}{\partial\theta^2}\right)$$

That is, for large n,

$$S(0) \sim N\,(0,\ E_{\theta=0}(I(\theta)))$$

Asymptotically, the variance of the score statistic $S(0)$, $E_{\theta=0}(I(\theta))$, can also be replaced by the expected Fisher information evaluated at the MLE $\theta = \hat{\theta}$ ($E_{\theta=\hat{\theta}}(I(\theta))$), the observed Fisher information evaluated at the null hypothesis $\theta = 0$ ($I(0)$), or the observed Fisher information evaluated at the MLE $\theta = \hat{\theta}$ ($I(\hat{\theta})$) (Kalbfleisch and Prentice, 1980, p. 46), where

$$I(0) = -\left(\frac{\partial^2 \log(L(\theta))}{\partial \theta^2} \mid \theta = 0\right)$$

$$I(\hat{\theta}) = -\left(\frac{\partial^2 \log(L(\theta))}{\partial \theta^2} \mid \theta = \hat{\theta}\right)$$

Thus, asymptotically, for large n,

$$S(0) \sim N\,(0,\ I)$$

where I is the information, either an expected Fisher information ($E_{\theta=0}(I(\theta))$ or $E_{\theta=\hat{\theta}}(I(\theta))$) or a observed Fisher information ($I(0)$ or $I(\hat{\theta})$).

So to test $H_0 : \theta = 0$ versus $H_1 : \theta \neq 0$, you can use the standardized Z test statistic

$$Z = \frac{S(0)}{\sqrt{I}}$$

If the BOUNDARYSCALE=MLE is specified in the SEQDESIGN procedure, the boundary values for the test statistic are displayed in the MLE scale. With the standardized Z statistic, the MLE statistic $\hat{\theta} = Z/\sqrt{I} = U(0)/I$ and

$$\hat{\theta} \sim N\left(0,\ \frac{1}{I}\right)$$

One-Sample Test for Mean

The following one-sample test for mean is used to demonstrate fixed-sample clinical trials in the section "One-Sided Fixed-Sample Tests in Clinical Trials" on page 5821 and the section "Two-Sided Fixed-Sample Tests in Clinical Trials" on page 5824.

Suppose $y_1, y_2, \ldots, y_n$ are n observations of a response variable Y from a normal distribution

$$y_i \sim N\left(\theta,\ \sigma^2\right)$$

where θ is the unknown mean and σ^2 is the known variance.

Then the log likelihood function for θ is

$$\log(L(\theta)) = \sum_{j=1}^{n} -\frac{1}{2}\frac{(y_j - \theta)^2}{\sigma^2} + c$$

where c is a constant. The first derivative is

$$\frac{\partial \log(L(\theta))}{\partial \theta} = \frac{1}{\sigma^2}\sum_{j=1}^{n}(y_j - \theta) = \frac{n}{\sigma^2}(\overline{y} - \theta)$$

where $\overline{y}$ is the sample mean.

Setting the first derivative to zero, the MLE of θ is $\hat{\theta} = \overline{y}$, the sample mean. The variance for $\hat{\theta}$ can be derived from the Fisher information

$$I(\theta) = -\frac{\partial^2 \log(L(\theta))}{\partial \theta^2} = \frac{n}{\sigma^2}$$

Since the Fisher information $I_0 = I(\theta)$ does not depend on θ in this case, $1/I_0$ is used as the variance for $\hat{\theta}$. Thus the sample mean $\overline{y}$ has a normal distribution with mean θ and variance σ^2/n:

$$\hat{\theta} = \overline{y} \sim N\left(\theta, \frac{1}{I_0}\right) = N\left(\theta, \frac{\sigma^2}{n}\right)$$

Under the null hypothesis $H_0 : \theta = 0$, the score statistic

$$S(0) = \frac{\partial \log(L(\theta))}{\partial \theta} | \theta = 0 = \frac{n}{\sigma^2}\overline{y}$$

has a mean zero and variance

$$I(\theta) = -\frac{\partial^2 \log(L(\theta))}{\partial \theta^2} = \frac{n}{\sigma^2}$$

With the MLE $\hat{\theta}$, the corresponding standardized statistic is computed as $Z = \hat{\theta}\sqrt{I_0} = \overline{y}/(\sigma/\sqrt{n})$, which has a normal distribution with variance 1:

$$Z \sim N\left(\theta\sqrt{I_0}, 1\right) = N\left(\frac{\theta}{\sigma/\sqrt{n}}, 1\right)$$

Also, the corresponding score statistic is computed as $S = \hat{\theta} I_0 = n\overline{y}/\sigma^2$ and

$$S \sim N\left(\theta I_0, I_0\right) = N\left(\frac{n\theta}{\sigma^2}, \frac{n}{\sigma^2}\right)$$

which is identical to $S(0)$ computed under the null hypothesis $H_0 : \theta = 0$.

Note that if the variable Y does not have a normal distribution, then it is assumed that the sample size n is large such that the sample mean has an approximately normal distribution.

One-Sided Fixed-Sample Tests in Clinical Trials

A one-sided test has either an upper (greater) or a lower (lesser) alternative. This section describes one-sided tests with upper alternatives only. Corresponding results for one-sided tests with lower alternatives can be derived similarly.

For a one-sided test of $H_0 : \delta \leq \delta_0$ with an upper alternative $H_1 : \delta > \delta_0$, an equivalent null hypothesis is $H_0 : \theta \leq 0$ with an upper alternative $H_1 : \theta > 0$, where $\theta = \delta - \delta_0$. A fixed-sample test rejects H_0 if the standardized test statistic $Z_0 = \hat{\theta}\sqrt{I_0} \geq C_\alpha$, where $\hat{\theta}$ is the sample estimate of θ and $C_\alpha = \Phi^{-1}(1-\alpha)$ is the critical value.

The p-value of the test is given by $1 - \Phi(Z_0)$, and the hypothesis H_0 is rejected if the p-value is less than α. An upper $(1-\alpha)$ confidence interval has the lower limit

$$\theta_l = \hat{\theta} - \frac{\Phi^{-1}(1-\alpha)}{\sqrt{I_0}} = \frac{Z_0 - \Phi^{-1}(1-\alpha)}{\sqrt{I_0}}$$

The hypothesis H_0 is rejected if the confidence interval for the parameter θ does not contain zero—that is, if the lower limit θ_l is greater than 0.

With an alternative reference $\theta = \theta_1$, $\theta_1 > 0$, a Type II error probability is defined as

$$\beta = P_{\theta=\theta_1}(Z_0 < C_\alpha)$$

which is equivalent to

$$\beta = P_{\theta=\theta_1}\left(Z_0 - \theta_1\sqrt{I_0} \; < \; C_\alpha - \theta_1\sqrt{I_0}\right) = \Phi\left(C_\alpha - \theta_1\sqrt{I_0}\right)$$

Thus, $\Phi^{-1}(\beta) = C_\alpha - \theta_1\sqrt{I_0}$. Then, with $C_\alpha = \Phi^{-1}(1-\alpha)$,

$$\theta_1\sqrt{I_0} = \Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta)$$

The drift parameter $\theta_1\sqrt{I_0}$ can be computed for specified α and β and the maximum information is given by

$$I_0 = \left(\frac{\Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta)}{\theta_1}\right)^2$$

If the maximum information is available, then the required sample size can be derived. For example, in a one-sample test for the mean with a specific standard deviation σ, the sample size n required for the test is

$$n = \sigma^2 I_0 = \sigma^2\left(\frac{\Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta)}{\theta_1}\right)^2$$

On the other hand, if the alternative reference θ_1, standard deviation σ, and sample size n are all specified, then α can be derived for a given β and, similarly, β can be derived for a given α.

With an alternative reference $\theta = \theta_1$, $\theta_1 > 0$, the power $1 - \beta$ is the probability of correctly rejecting the null hypothesis H_0 at θ_1:

$$1 - \beta = 1 - P_{\theta=\theta_1}(Z_0 < C_\alpha) = \Phi\left(\theta_1\sqrt{I_0} - C_\alpha\right)$$

Superiority Trials

A superiority trial that tests the response to a new drug is clinically superior to a comparative placebo control or active control therapy. If a positive value indicates a beneficial effect, a test for superiority has

$$H_0 : \theta \leq 0 \qquad H_1 : \theta > 0$$

where H_0 is the hypothesis of nonsuperiority and H_1 is the alternative hypothesis of superiority.

The superiority test rejects the hypothesis H_0 and declares superiority if the standardized statistic $Z_0 = \hat{\theta}\sqrt{I_0} \geq C_\alpha$, where the critical value $C_\alpha = \Phi^{-1}(1-\alpha)$.

For example, if θ is the response difference between the treatment and placebo control groups, then a superiority trial can be

$$H_0 : \theta \leq 0 \qquad H_1 : \theta = 6$$

with a Type I error probability level $\alpha = 0.025$ and a power $1 - \beta = 0.90$ at $\theta_1 = 6$.

Noninferiority Trials

A noninferiority trial does not compare the response to a new treatment with the response to a placebo. Instead, it demonstrates the effectiveness of a new treatment compared with that of a nonexisting placebo by showing that the response of a new treatment is not clinically inferior to the response of a standard therapy with an established effect. That is, this type of trial attempts to demonstrate that the new treatment effect is not worse than the standard therapy effect by an acceptable margin. These trials are often performed when there is an existing effective therapy for a serious disease, and therefore a placebo control group cannot be ethically included.

It can be difficult to specify an appropriate noninferiority margin. One practice is to choose with reference to the effect of the active control in historical placebo-controlled trials (Snapinn 2000, p. 20). With this practice, there is some basis to imply that the new treatment is better than the placebo for a positive noninferiority trial.

If a positive value indicates a beneficial effect, a test for noninferiority has a null hypothesis $\delta \leq -\delta_0$ and an alternative hypothesis $\delta = \delta_1 > -\delta_0$, where $\delta_0 > 0$ is the specified noninferiority margin.

An equivalent test has

$$H_0 : \theta \leq 0 \qquad H_1 : \theta = \theta_1 > 0$$

where the parameter $\theta = \delta + \delta_0$, H_0 is the null hypothesis of inferiority, and H_1 is the alternative hypothesis of noninferiority,

The noninferiority test rejects the hypothesis H_0 and declares noninferiority if the standardized statistic $Z_0 = \hat{\theta}\sqrt{I_0} = (\hat{\delta} + \delta_0)\sqrt{I_0} \geq C_\alpha$, where the critical value $C_\alpha = \Phi^{-1}(1-\alpha)$.

For example, if δ is the response difference between the treatment and active control groups and $\delta_0 = 2$ is the noninferiority margin, then a noninferiority trial with a power $1-\beta = 0.90$ at $\delta_1 = 1$ might be

$$H_0 : \theta \leq 0 \qquad H_1 : \theta = 3$$

where $\theta = \delta + \delta_0 = \delta + 2$.

Two-Sided Fixed-Sample Tests in Clinical Trials

A two-sided test is a test of a hypothesis with a two-sided alternative. Two-sided tests include simple symmetric tests and more complicated asymmetric tests that might have distinct lower and upper alternative references.

Symmetric Two-Sided Tests for Equality

For a symmetric two-sided test with the null hypothesis $\delta = \delta_0$ against the alternative $\delta \neq \delta_0$, an equivalent null hypothesis is $H_0 : \theta = 0$ with a two-sided alternative $H_1 : \theta \neq 0$, where $\theta = \delta - \delta_0$. A fixed-sample test rejects H_0 if $|\hat{\theta}\sqrt{I_0}| \geq C_{\alpha/2}$, where $\hat{\theta}$ is a sample estimate of θ and $C_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$ is the critical value.

A common two-sided test is the test for the response difference between a treatment group and a control group. The null and alternative hypotheses are $H_0 : \theta = 0$ and $H_1 : \theta \neq 0$, respectively, where θ is the response difference between the two groups. If a greater value indicates a beneficial effect, then there are three possible results:

- The test rejects the hypothesis H_0 of equality and indicates that the treatment is significantly better if the standardized statistic $Z_0 = \hat{\theta}\sqrt{I_0} \geq C_{\alpha/2}$.
- The test rejects the hypothesis H_0 and indicates the treatment is significantly worse if the standardized statistic $Z_0 = \hat{\theta}\sqrt{I_0} \leq -C_{\alpha/2}$.
- The test indicates no significant difference between the two responses if $-C_{\alpha/2} < \hat{\theta}\sqrt{I_0} < C_{\alpha/2}$.

The p-value of the test is $2(1-\Phi(Z_0))$ if $Z_0 > 0$ and $2\Phi(Z_0)$ if $Z_0 \leq 0$. The hypothesis H_0 is rejected if the p-value of the test is less than α—that is, if $1-\Phi(Z_0) < \alpha/2$ or $\Phi(Z_0) < \alpha/2$. A symmetric $(1-\alpha)$ confidence interval for θ has lower and upper limits

$$\left(\hat{\theta} - \frac{C_{\alpha/2}}{\sqrt{I_0}}\,,\ \hat{\theta} + \frac{C_{\alpha/2}}{\sqrt{I_0}}\right)$$

which is

$$\left(\frac{1}{\sqrt{I_0}}\left(Z_0 - C_{\alpha/2}\right)\,,\ \frac{1}{\sqrt{I_0}}\left(Z_0 + C_{\alpha/2}\right)\right)$$

The hypothesis H_0 is rejected if the confidence interval for the parameter θ does not contain zero. That is, the lower limit is greater than zero or the upper limit is less than zero.

With an alternative reference $\theta = \theta_1 > 0$, a Type II error probability is defined as

$$\beta = P_{\theta=\theta_1}(-C_{\alpha/2} \ < \ Z_0 \ < \ C_{\alpha/2})$$

which is

$$\beta = P_{\theta=\theta_1}\left((-C_{\alpha/2} - \theta_1\sqrt{I_0}) \ < \ (Z_0 - \theta_1\sqrt{I_0}) \ < \ (C_{\alpha/2} - \theta_1\sqrt{I_0})\right)$$

Thus

$$\beta = \Phi\left(C_{\alpha/2} - \theta_1\sqrt{I_0}\right) - \Phi\left(-C_{\alpha/2} - \theta_1\sqrt{I_0}\right)$$

The resulting power $1 - \beta$ is the probability of correctly rejecting the null hypothesis, which includes the probability for the lower alternative and the probability for the upper alternative. The SEQDESIGN procedure uses only the probability of correctly rejecting the null hypothesis for the correct alternative in the power computation.

Thus, under the upper alternative hypothesis, the power in the SEQDESIGN procedure is computed as the probability of rejecting the null hypothesis for the upper alternative, $1. - \Phi\left(C_{\alpha/2} - \theta_1\sqrt{I_0}\right) = \Phi\left(\theta_1\sqrt{I_0} - C_{\alpha/2}\right)$, and a very small probability of rejecting the null hypothesis for the lower alternative, $\Phi\left(-C_{\alpha/2} - \theta_1\sqrt{I_0}\right)$, is ignored. This power computation is more rational than the power based on the probability of correctly rejecting the null hypothesis (Whitehead 1997, p. 75).

That is,

$$\beta = P_{\theta=\theta_1}\left((Z_0 - \theta_1\sqrt{I_0}) \ < \ (C_{\alpha/2} - \theta_1\sqrt{I_0})\right) = \Phi\left(C_{\alpha/2} - \theta_1\sqrt{I_0}\right)$$

Then with $\Phi^{-1}(\beta) = C_{\alpha/2} - \theta_1\sqrt{I_0}$,

$$\theta_1\sqrt{I_0} = C_{\alpha/2} - \Phi^{-1}(\beta) = \Phi^{-1}(1-\alpha/2) + \Phi^{-1}(1-\beta)$$

The drift parameter $\theta_1\sqrt{I_0}$ can be derived for specified α and β, and the maximum information is given by

$$I_0 = \left(\frac{\Phi^{-1}(1-\alpha/2) + \Phi^{-1}(1-\beta)}{\theta_1}\right)^2$$

If the maximum information is available, then the required sample size can be derived. For example, in a one-sample test for mean, if the standard deviation σ is known, the sample size n required for the test is

$$n = \sigma^2 I_0 = \sigma^2\left(\frac{\Phi^{-1}(1-\alpha/2) + \Phi^{-1}(1-\beta)}{\theta_1}\right)^2$$

On the other hand, if the alternative reference θ_1, standard deviation σ, and sample size n are all known, then α can be derived with a given β and, similarly, β can be derived with a given α.

Generalized Two-Sided Tests for Equality

For a generalized two-sided test with the null hypothesis $\delta = \delta_0$ against the alternative $\delta \neq \delta_0$, an equivalent null hypothesis is $H_0 : \theta \leq 0$ with a two-sided alternative $H_1 : \theta \neq 0$, where $\theta = \delta - \delta_0$. A fixed-sample test rejects H_0 if the standardized statistic $Z_0 = \hat{\theta}\sqrt{I_0} < -C_{\alpha_l}$ or $Z_0 = \hat{\theta}\sqrt{I_0} > C_{\alpha_u}$, where the critical values $C_{\alpha_l} = \Phi^{-1}(1-\alpha_l)$ and $C_{\alpha_u} = \Phi^{-1}(1-\alpha_u)$.

With the lower alternative reference $\theta_{1l} < 0$, a lower Type II error probability is defined as

$$\beta_l = P_{\theta=\theta_{1l}}\left(-C_{\alpha_l} \leq Z_{0l}\sqrt{I_0}\right) = P_{\theta=\theta_{1l}}\left(-C_{\alpha_l} - \theta_{1l}\sqrt{I_0} \leq Z_{0l}\sqrt{I_0} - \theta_{1l}\sqrt{I_0}\right)$$

This implies

$$\beta_l = 1 - \Phi(-C_{\alpha_l} - \theta_{1l}\sqrt{I_0})$$

and the power is the probability of correctly rejecting the null hypothesis for the lower alternative,

$$1 - \beta_l = \Phi(-C_{\alpha_l} - \theta_{1l}\sqrt{I_0})$$

The lower drift parameter is derived as

$$\theta_{1l}\sqrt{I_0} = -\left(\Phi^{-1}(1-\alpha_l) + \Phi^{-1}(1-\beta_l)\right)$$

Then, with specified α_l and β_l, if the maximum information is known, the lower alternative reference θ_{1l} can be derived. If the maximum information is unknown, then with the specified lower alternative reference θ_{1l}, the maximum information required is

$$I_{0l} = \left(\frac{\Phi^{-1}(1-\alpha_l) + \Phi^{-1}(1-\beta_l)}{-\theta_{1l}}\right)^2$$

Similarly, the upper drift parameter is derived as

$$\theta_{1u}\sqrt{I_0} = \Phi^{-1}(1-\alpha_u) + \Phi^{-1}(1-\beta_u)$$

For a given α_u, β_u, and the upper alternative reference θ_{1u}, the maximum information required is

$$I_{0u} = \left(\frac{\Phi^{-1}(1-\alpha_u) + \Phi^{-1}(1-\beta_u)}{\theta_{1u}}\right)^2$$

Thus, the maximum information required for the design is given by

$$I_0 = \max\left(I_{0l},\ I_{0u}\right)$$

Note that with the maximum information level I_0, if $I_{0l} < I_0$, then the derived power from the lower alternative is larger than the specified $1-\beta_l$. Similarly, if $I_{0u} < I_0$, then the derived power from the upper alternative is larger than the specified $1-\beta_u$.

If maximum information is available, the required sample size can be derived. For example, in a one-sample test for mean, if the standard deviation σ is known, the sample size n required for the test is $n = \sigma^2 I_0$.

On the other hand, if the alternative references, Type I error probabilities α_l and α_u, standard deviation σ, and sample size n are all specified, then the Type II error probabilities β_l and β_u and the corresponding powers can be derived.

Group Sequential Methods

A group sequential design provides interim analyses before the formal completion of a trial. The monitoring process provides possible early stopping for either positive or negative results and thus reduces the time to complete the trial. With a specified number of stages, the design creates critical values such that at each interim analysis, a hypothesis can be rejected, accepted, or continued to the next time point. At the final stage, a hypothesis is either rejected or accepted. Usually, the critical values are derived such that the specified overall Type I and Type II error probability levels are maintained in the design.

For example, to test a null hypothesis H_0 with an upper alternative in a fixed-sample design, a critical value c_α is created. The null hypothesis H_0 is rejected if the test statistic is greater than or equal to the critical value c_α. Otherwise, H_0 is accepted. But, for a group sequential design with early stopping to reject or accept the null hypothesis H_0, there are two critical values created at each interim analysis: an α critical value $c_{\alpha k}$ to reject the null hypothesis and a β critical value $c_{\beta k}$ to accept the null hypothesis. The null hypothesis H_0 is rejected if the test statistic is greater than or equal to the α critical value $c_{\alpha k}$, and H_0 is accepted if the test statistic is less than the β critical value $c_{\beta k}$. If the test statistic is between these two critical values, the process continues to the next stage. At the final stage, the two critical values are equal, and the hypothesis is either rejected or accepted.

Armitage, McPherson, and Rowe (1969) showed that repeated significance tests at a fixed level on accumulating data increase the probability of obtaining a significant result under the null hypothesis. For example, with a significance level 0.05 in a two-sided fixed-sample test, the critical value is 1.96. If this value is used in a five-stage group sequential trial with early stopping to reject the null hypothesis, then the probability of rejecting the null hypothesis at or before the fifth stage is 0.14169, much larger than the nominal value 0.05 (Armitage, McPherson, and Rowe 1969, p. 239).

Pocock (1977) applied these repeated significance tests to group sequential trials with equally spaced information levels and derives a constant critical value on the standardized normal Z scale across all stages that maintains the Type I error probability level. For example, with a significance level 0.05 in a two-sided test, the derived critical value at each stage is 2.413 on the standardized normal Z scale, larger than the fixed-sample critical value 1.96. The corresponding nominal p-value is 0.0158, which is smaller than the fixed-sample p-value 0.025 (Pocock 1977, p. 193).

O'Brien and Fleming (1979) proposed a sequential procedure that has boundary values decrease over the stages on the standardized normal Z scale to make the early stop less likely. The procedure has conservative stopping boundary values at very early stages, and boundary values at the final stage are close to the fixed-sample design. For example, with a significance level 0.05 in a two-sided test, the derived critical values at these five stages on the standardized normal Z scale are 4.562, 3.226, 2.634, 2.281, and 2.040.

Wang and Tsiatis (1987), Emerson and Fleming (1989) and Pampallona and Tsiatis (1994) generalized the Pocock and O'Brien-Fleming methods to the power family, where a power parameter is used to allow a continuous set of designs between the Pocock and O'Brien-Fleming methods.

Kittelson and Emerson (1999) extended the methods in the power family even further to the unified family, which also includes the exact triangular method. The shape and location of each of the four boundaries can be independently specified in the unified family methods.

Whitehead and Stratton (1983) and Whitehead (1997, 2001) developed triangular methods by adapting tests for continuous monitoring to discrete monitoring. With early stopping to reject or accept the null hypothesis in a one-sided test, the derived continuation region has a triangular shape for the score-scaled boundaries. Only elementary calculations are needed to derive the boundary values for Whitehead's triangular methods.

For a sequential design, you can derive the α and β error probabilities at each stage from the boundaries. On the other hand, you can derive the boundaries from specified α and β error probabilities at each stage. The error spending function approach (Lan and DeMets 1983) uses the error spending function to specify the error probabilities at each stage and then uses these probabilities to derive the boundaries. You can specify α and β explicitly or implicitly with an error spending function for the cumulative probabilities.

Refer to Jennison and Turnbull (2000, pp. 5–11) for a more detailed history of group sequential methods.

The following three types of methods are available in the SEQDESIGN procedure to derive boundaries in a sequential design:

- fixed boundary shape methods, which derive boundaries with specified boundary shapes. These include the unified family method and Haybittle-Peto method.
- Whitehead methods, which adjust the boundaries from continuous monitoring for discrete monitoring
- error spending methods

You can use the SEQDESIGN procedure to specify methods from the same group for each design. A different method can be specified for each boundary separately, but all methods in a design must be from the same group.

Fixed Boundary Shape Methods

The fixed boundary shape methods include the unified family method (Kittelson and Emerson 1999) and the Haybittle-Peto method (Haybittle 1971; Peto et al. 1976). The unified family methods derive the boundary values with the specified boundary shape. The unified family methods include the Pocock method (Pocock 1977), the O'Brien-Fleming method (O'Brien and Fleming 1979), the power family method (Wang and Tsiatis 1987; Emerson and Fleming 1989; Pampallona and Tsiatis 1994), and the triangular method (Kittelson and Emerson 1999). See the section "Unified Family Methods" on page 5839 for a detailed description of the methods that use the unified family approach.

The Haybittle-Peto method uses a value of 3 for the critical values in interim stages, so that the critical value at the final stage is close to the original design without interim monitoring. In the SEQDESIGN procedure, the Haybittle-Peto method has been generalized to allow for different boundary values at different stages. See the section "Haybittle-Peto Method" on page 5844 for a detailed description of the Haybittle-Peto method.

Whitehead Methods

The Whitehead methods (Whitehead and Stratton 1983; Whitehead 1997, 2001) derive the boundary values by adapting the continuous monitoring tests to the discrete monitoring of group sequential tests. The Type I error probability and power corresponding to the resulting boundaries are extremely close but differ slightly from the specified values because of the approximations used in deriving the tests (Jennison and Turnbull 2000, p. 106). The SEQDESIGN procedure provides the BOUNDARYKEY= option to adjust the boundary value at the final stage for the exact Type I or Type II error probability level. See the section "Whitehead Methods" on page 5845 for a detailed description of Whitehead's methods.

Error Spending Methods

An error spending method (Lan and DeMets 1983) uses the error spending function to specify the error spending at each stage and then uses these error probabilities to derive the boundary values. You can specify these errors explicitly or with an error spending function for these cumulative errors. See the section "Error Spending Methods" on page 5848 for a detailed description of the error spending methods.

Error spending methods derive boundary values at each stage sequentially and require much more computation than other types of methods for group sequential trials with a large number of stages, especially for a two-sided asymmetric design with early stopping to accept H_0, or to reject or accept H_0.

The sample size requirement for some applicable tests can also be computed in the procedure. After the actual data from a clinical trial are collected, you can then use the boundary information created in the SEQDESIGN procedure to perform a group sequential test in the SEQTEST procedure.

Statistical Assumptions for Group Sequential Designs

The SEQDESIGN procedure assumes that with a total number of stages K, the sequence of the standardized test statistics $\{Z_1, Z_2, \ldots, Z_K\}$ has the canonical joint distribution with information levels $\{I_1, I_2, \ldots, I_K\}$ for the parameter θ (Jennison and Turnbull 2000, p. 49):

- $(Z_1, Z_2, \ldots, Z_K)$ is multivariate normal
- $Z_k \sim N\left(\theta\sqrt{I_k}, 1\right), \; k = 1, 2, \ldots, K$
- $\mathrm{Cov}(Z_{k_1}, Z_{k_2}) = \sqrt{(I_{k_1}/I_{k_2})}, \quad 1 \le k_1 \le k_2 \le K$

In terms of the maximum likelihood estimator, $\hat{\theta}_k = Z_k/\sqrt{I_k}, k = 1, 2, \ldots, K$, the canonical joint distribution can be expressed as follows:

- $(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_K)$ is multivariate normal
- $\hat{\theta}_k \sim N\left(\theta,\ 1/I_k\right),\ k = 1, 2, \ldots, K$
- $\mathrm{Cov}(\hat{\theta}_{k_1}, \hat{\theta}_{k_2}) = 1/I_{k_2},\quad 1 \le k_1 \le k_2 \le K$

Furthermore, in terms of the score statistics $S_k = Z_k\sqrt{I_k}, k = 1, 2, \ldots, K$, the canonical joint distribution can be expressed as follows:

- $(S_1, S_2, \ldots, S_K)$ is multivariate normal
- $S_k \sim N\left(\theta\, I_k,\ I_k\right),\ k = 1, 2, \ldots, K$
- $\mathrm{Cov}(S_{k_1}, S_{k_2}) = \mathrm{Var}(S_{k_1}) = I_{k_1},\quad 1 \le k_1 \le k_2 \le K$

That is, the increments S_1, $S_2 - S_1$, …, and $S_K - S_{(K-1)}$ are independently distributed.

If the test statistic is computed from the data that are not from a normal distribution, such as a binomial distribution, then it is assumed that the test statistic is computed from a large sample such that the statistic has an approximately normal distribution.

If the increments S_1, $S_2 - S_1$, …, and $S_K - S_{(K-1)}$ are not independently distributed, then it is inappropriate to use group sequential methods in the SEQDESIGN procedure. One such example is the Gehan statistic, which is a weighted log-rank statistic for censored data. Refer to Jennison and Turnbull (2000 pp. 232–233, 276–277) and Proschan, Lan, and Wittes (2006 pp. 150–151) for a description of statistics with nonindependent increments.

If a trial stops at an early interim stage with only a small number of responses observed, it can lead to a distrust of the statistical findings, which rely on the assumption that the sample is large (Whitehead 1997, p. 167). A group sequential design can be specified such that at the first interim analysis, there are a sufficient number of responses to ensure that the analysis to be conducted is both reliable and persuasive (Whitehead 1997, p. 167).

Alternatively, a method such as the O'Brien-Fleming method can be used to derive conservative stopping boundary values at very early stages to make the early stop less likely. That is, the trial is stopped in early stages only with overwhelming evidence.

A simple example of the group sequential tests is the test for a normal mean, $\mu = \mu_0$. Suppose $y_1, y_2, \ldots, y_n$ are n observations of a response variable Y in a data set from a normal distribution with an unknown mean μ and a known variance σ^2. Then the maximum likelihood estimate of μ is the sample mean

$$\overline{y} = \frac{1}{n}\sum_{j=1}^{n} y_j$$

The sample mean has a normal distribution with mean μ and variance σ^2/n:

$$\overline{y} \sim N\left(\mu,\ \frac{\sigma^2}{n}\right)$$

An equivalent hypothesis for $\mu = \mu_0$ is $H_0 : \theta = 0$, where $\theta = \mu - \mu_0$. The MLE statistic for θ,

$$\hat{\theta} = \overline{y} - \mu_0 \sim N\left(\theta,\ {I_0}^{-1}\right)$$

where the information $I_0 = n/\sigma^2$.

For a group sequential test with K stages, there are $N_1, N_2, \ldots, N_K$ observations available at these stages. At stage k, the sample mean is computed as

$$\overline{y}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} y_{kj}$$

where y_{kj} is the value of the jth observation available at the kth stage and N_k is the cumulative sample size at stage k, which includes the N_{k-1} observations collected at previous stages and the $N_k - N_{k-1}$ observations collected at the current stage.

The maximum likelihood estimate

$$\hat{\theta}_k = \overline{y}_k - \mu_0 \sim N\left(\theta,\ {I_k}^{-1}\right)$$

where the information

$$I_k = \frac{1}{\mathrm{Var}(\overline{y}_k)} = \frac{N_k}{\sigma^2}$$

is the inverse of the variance.

Thus, the standardized statistic

$$Z_k = \hat{\theta}_k \sqrt{I_k} = (\overline{y}_k - \mu_0)\sqrt{I_k} \sim N\left(\theta \sqrt{I_k},\ 1\right)$$

The covariance of Z_{k_1} and Z_{k_2}, $1 \le k_1 \le k_2 \le K$ can be expressed as

$$\mathrm{Cov}(Z_{k_1}, Z_{k_2}) = \frac{1}{\sqrt{(I_{k_1} I_{k_2})}} \mathrm{Cov}(S_{k_1}, S_{k_2})$$

where $S_{k_1} = Z_{k_1}\sqrt{I_{k_1}}$ and $S_{k_2} = Z_{k_2}\sqrt{I_{k_2}}$.

Since $S_{k_2} - S_{k_1}$ is independent of S_{k_1}, $\mathrm{Cov}(S_{k_1}, S_{k_2}) = \mathrm{Var}(S_{k_1}) = I_{k_1}$ and

$$\mathrm{Cov}(Z_{k_1}, Z_{k_2}) = \frac{1}{\sqrt{(I_{k_1} I_{k_2})}} I_{k_1} = \sqrt{I_{k_1}/I_{k_2}}$$

Thus the statistics $\{Z_1, Z_2, \ldots, Z_K\}$ has the canonical joint distribution with information levels $\{I_1, I_2, \ldots, I_K\}$ for the parameter μ. See the section "Applicable One-Sample Tests and Sample Size Computation" on page 5858, the section "Applicable Two-Sample Tests and Sample Size Computation" on page 5860, and the section "Applicable Regression Parameter Tests and Sample Size Computation" on page 5868 for more examples of applicable tests in group sequential trials.

Boundary Scales

The boundaries computed by the SEQDESIGN procedure are applied to test statistics computed during the analysis, and so generally, the scale you select for the boundaries is determined by the scale of the statistics that you will be using.

The following scales are available in the SEQDESIGN procedure:

- MLE, maximum likelihood estimate
- standardized Z
- score statistic S
- p-value

These scales are all equivalent for a given set of boundary values—that is, there exists a unique transformation between any two of these scales. If you know the boundary values in terms of statistics from one scale, you can uniquely derive the boundary values of statistics for other scales. You can specify the scale with the BOUNDARYSCALE= option, and the default is the standardized Z scale.

You can also select the boundary scale to better examine the features of an individual group sequential design or to compare features among multiple designs. For example, with the standardized Z scale, the boundary values for the Pocock design are identical across all stages, and the O'Brien-Fleming design has boundary values (in absolute value) that decrease over the stages.

The remaining section demonstrates the transformations from one scale to the other scales. If the maximum likelihood estimate $\hat{\theta}$ is computed by the analysis, then

$$\hat{\theta} \sim N\left(\theta,\ \frac{1}{I}\right)$$

where I is the Fisher information if it does not depend on θ. Otherwise, I is either the expected Fisher information evaluated at $\hat{\theta}$ or the observed Fisher information. See the section "Maximum Likelihood Estimator" on page 5818 for a detailed description of these statistics.

With the MLE statistic $\hat{\theta}$, the corresponding standardized Z statistic is computed as

$$Z = \hat{\theta}\sqrt{I} \sim N\left(\theta\sqrt{I},\ 1\right)$$

and the corresponding score statistic is computed as

$$S = \hat{\theta}\, I \sim N\left(\theta\, I,\ I\right)$$

Similarly, if a score statistic S is computed by the analysis, then with

$$S \sim N\left(\theta\, I,\ I\right)$$

where I is the information, either an expected Fisher information ($E_{\theta=0}(I(\theta))$ or $E_{\theta=\hat{\theta}}(I(\theta))$) or an observed Fisher information ($I(0)$ or $I(\hat{\theta})$).

The corresponding standardized Z statistic is computed as

$$Z = \frac{S}{\sqrt{I}} \sim N\left(\theta\sqrt{I},\ 1\right)$$

and the corresponding MLE-scaled statistic is computed as

$$\hat{\theta} = \frac{S}{I} \sim N\left(\theta,\ \frac{1}{I}\right)$$

With a standardized normal Z statistic, the corresponding fixed-sample nominal p-value depends on the type of alternative hypothesis. With an upper alternative, the nominal p-value is defined as the one-sided p-value under the null hypothesis $H_0 : \theta = 0$ with an upper alternative:

$$p_k = 1 - \Phi(Z)$$

With a lower alternative or a two-sided alternative, the nominal p-value is defined as the one-sided p-value under the null hypothesis $H_0 : \theta = 0$ with a lower alternative:

$$p_k = \Phi(Z)$$

which is an increasing function of the standardized Z statistic (Emerson, Kittelson, and Gillen 2005, p. 12).

The BOUNDARYSCALE= MLE, STDZ, SCORE, and PVALUE options display the boundary values in the MLE, standardize Z, score, and p-value scales, respectively. For example, suppose $y_{k1}, y_{k2}, \ldots, y_{kn_k}$ are n_k observations of a response variable Y in a data set from a normal distribution with an unknown mean μ and a known variance σ^2. Then

$$y_{kj} \sim N\left(\mu,\ \sigma^2\right)$$

for $k = 1, 2, \ldots, K$, where K is the number of groups and n_k is the number of observations at group k.

If N_k is the cumulative number of observations for the first k groups, then the sample mean from these N_k observations

$$\overline{y}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} y_{kj}$$

has a normal distribution with mean θ and variance σ^2/N_k:

$$\overline{y}_k \sim N\left(\theta,\ \frac{\sigma^2}{N_k}\right)$$

To test the null hypothesis $\mu = \mu_0$, $H_0 : \theta = 0$, where $\theta = \mu - \mu_0$ can be used. The MLE of θ is $\hat{\theta}_k = \overline{y}_k - \mu_0$ and

$$\hat{\theta}_k \sim N\left(\theta,\ \frac{1}{I_k}\right)$$

where the information is the inverse of the variance of $\overline{y}_k$,

$$I_k = \frac{N_k}{\sigma^2}$$

The corresponding standardized Z statistic is

$$Z_k = \hat{\theta}_k \, I_k^{\frac{1}{2}} \sim N\left(\theta \, I_k^{\frac{1}{2}}, \, 1\right)$$

The score statistic in the SEQDESIGN procedure is then given by

$$S_k = \hat{\theta}_k \, I_k = Z_k \, I_k^{\frac{1}{2}} \sim N\left(\theta \, I_k, \, I_k\right)$$

For a null hypothesis $H_0 : \theta = 0$ with an upper alternative, the nominal p-value of the standardized Z statistic is $p_k = 1 - \Phi(Z_k)$. For a null hypothesis $H_0 : \theta = 0$ with a lower alternative or a two-sided alternative, the nominal p-value of the standardized Z statistic is $p_k = \Phi(Z_k)$.

Boundary Variables

The boundaries created in group sequential trials depend on the type of the alternative hypothesis and the early stopping criterion. Table 77.5 shows the boundaries created with various design specifications.

Table 77.5 Boundary Variables

Specifications		Boundary Variables			
Alternative		**Lower**		**Upper**	
Hypothesis	**Early Stopping**	**Alpha**	**Beta**	**Beta**	**Alpha**
Lower	Accept H_0		X		
	Reject H_0	X			
	Accept/Reject H_0	X	X		
Upper	Accept H_0			X	
	Reject H_0				X
	Accept/Reject H_0			X	X
Two-sided	Accept H_0		X	X	
	Reject H_0	X			X
	Accept/Reject H_0	X	X	X	X

Up to four boundaries can be generated in a group sequential design:

- the upper α boundary, to reject the null hypothesis for the upper alternative
- the upper β boundary, to accept the null hypothesis with an upper alternative

- the lower β boundary, to accept the null hypothesis with a lower alternative
- the lower α boundary, to reject the null hypothesis for the lower alternative

For a two-sided design, the null hypothesis is accepted only if both the hypothesis is accepted with an upper alternative and the hypothesis is accepted with a lower alternative.

For a one-sided design with a lower alternative, only the lower boundaries are created. Similarly, for a one-sided design with an upper alternative, only the upper boundaries are created. For example, Figure 77.10 shows the boundary plot for a one-sided test with an upper alternative.

Figure 77.10 Boundary Plot for One-Sided Test

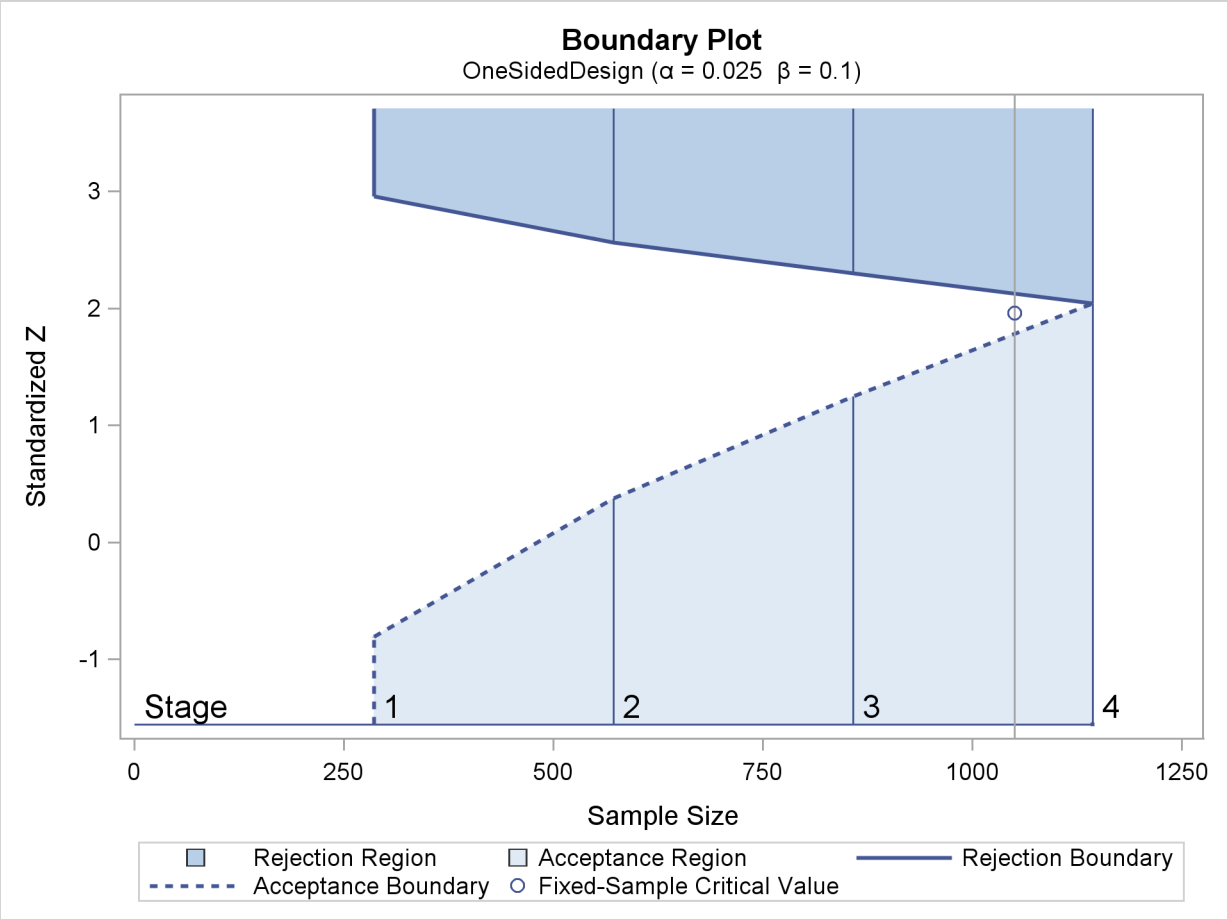


Figure 77.10 corresponds to a one-sided sequential design with early stopping to reject or accept the null hypothesis. For a sequential test with early stopping only to reject the null hypothesis, there are no acceptance boundary values at interim stages. The acceptance boundary value and its associated acceptance region are displayed only at the final stage. Similarly, for a sequential test with early stopping only to accept the null hypothesis, there are no rejection boundary values at interim stages. The rejection boundary value and its associated rejection region are displayed only at the final stage.

For a two-sided design, both the lower and upper boundaries are created. For a design with early stopping to reject the null hypothesis, α boundaries are created. Similarly, for a design with early

stopping to accept the null hypothesis, β boundaries are created. For a design with early stopping to accept or reject the null hypothesis, both the α and β boundaries are created.

For example, Figure 77.11 shows the boundary plot for a two-sided test.

Figure 77.11 Boundary Plot for Two-Sided Test

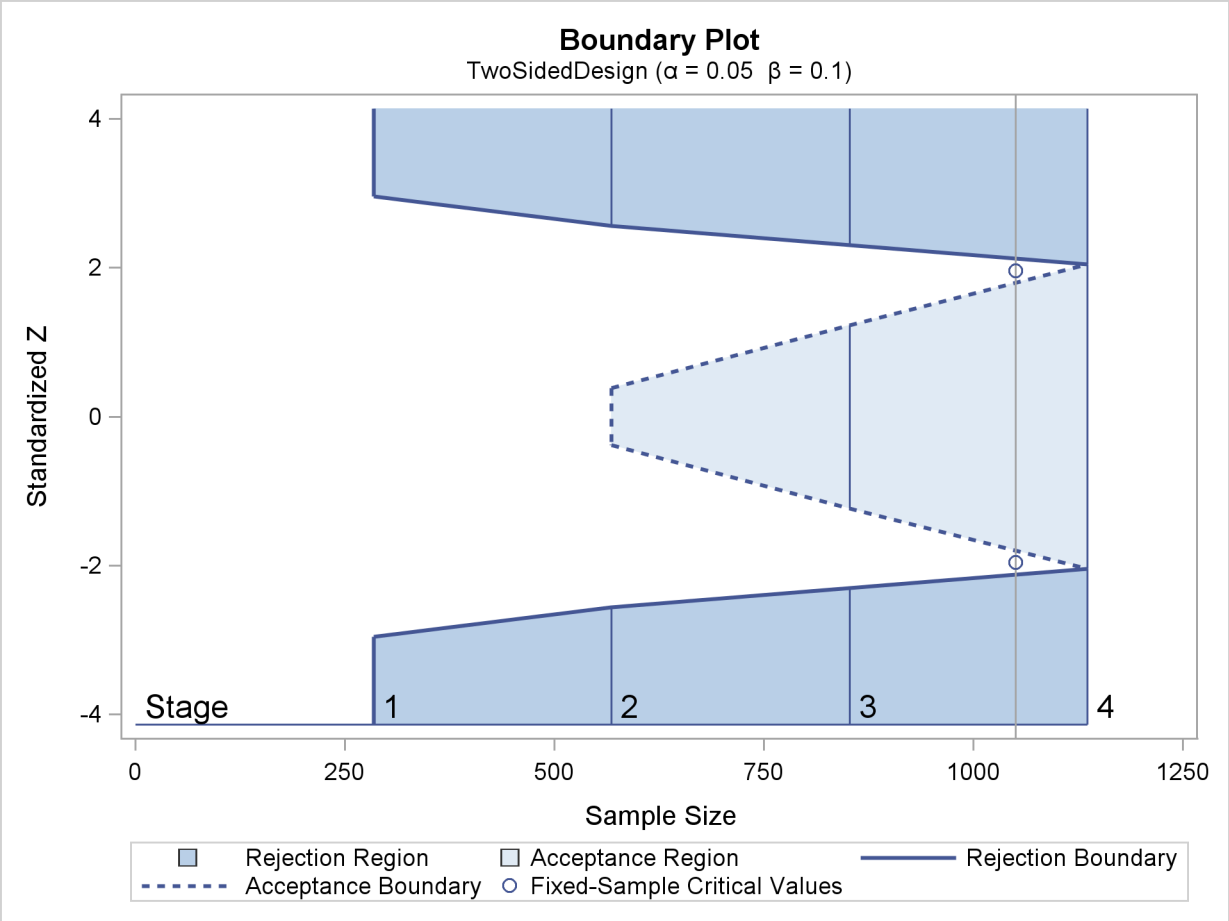


Figure 77.11 corresponds to a two-sided sequential design with early stopping to reject or accept the null hypothesis. For a sequential test with early stopping only to reject the null hypothesis, there are no acceptance boundary values at interim stages. The acceptance boundary value and its associated acceptance region are displayed only at the final stage. Similarly, for a sequential test with early stopping only to accept the null hypothesis, there are no rejection boundary values at interim stages. The rejection boundary value and its associated rejection region are displayed only at the final stage.

Type I and Type II Errors

The Type I error is the error of rejecting the null hypothesis when the null hypothesis is correct, and the Type II error is the error of not rejecting the null hypothesis when the null hypothesis is incorrect. The level of significance α is the probability of making a Type I error. The Type II error depends

on the hypothetical reference of the alternative hypothesis, and the Type II error probability β is defined as the probability of not rejecting the null hypothesis when a specific alternative reference is true. The power $1 - \beta$ is then defined as the probability of rejecting the null hypothesis at the alternative reference.

In a sequential design, if the maximum information and alternative reference are not both specified, the critical values are created such that both the specified Type I and the specified Type II error probability levels are maintained in the design. Otherwise, the critical values are created such that either the specified Type I error probability or the specified Type II error probability is maintained.

One-Sided Tests

For a K-stage group sequential design with an upper alternative hypothesis $H_1 : \theta = \theta_1$ and early stopping to reject or accept the null hypothesis $H_0 : \theta = 0$, the boundaries contain the upper α critical values a_k and upper β critical values $b_k, k = 1, 2, \ldots, K$. At each interim stage, $b_k < a_k$, the null hypothesis H_0 is rejected if the observed statistic $z_k \geq a_k$, H_0 is accepted if $z_k < b_k$, or the process is continued to the next stage if $b_k \leq z_k < a_k$. At the final stage $b_K = a_K$, the hypothesis is either rejected or accepted.

The overall Type I error probability α is given by

$$\alpha = \sum_{k=1}^{K} \alpha_k$$

where α_k is the α spending at stage k. That is, at stage 1,

$$\alpha_1 = P_{\theta=0}(z_1 \geq a_1)$$

At a subsequent stage k,

$$\alpha_k = P_{\theta=0}(b_j \leq z_j < a_j, j = 1, 2, \ldots, k-1, z_k \geq a_k)$$

Similarly, the Type II error probability

$$\beta = \sum_{k=1}^{K} \beta_k$$

where β_k is the β spending at stage k. That is, at stage 1,

$$\beta_1 = P_{\theta=\theta_1}(z_1 < b_1)$$

At a subsequent stage k,

$$\beta_k = P_{\theta=\theta_1}(b_j \leq z_j < a_j, j = 1, 2, \ldots, k-1, z_k < b_k)$$

With an upper alternative hypothesis $H_1 : \theta = \theta_1 > 0$, the power $1 - \beta$ is the probability of rejecting the null hypothesis for the upper alternative.

$$1 - \beta = 1 - \sum_{k=1}^{K} \beta_k = \sum_{k=1}^{K} P_{\theta=\theta_1}(b_j \le z_j < a_j, j = 1, 2, \ldots, k-1, z_k \ge a_k)$$

For a design with early stopping to reject H_0 only, the interim upper β critical values are set to $-\infty$, $b_k = -\infty, k = 1, 2, \ldots, K-1$, and $\beta = \beta_K$. For a design with early stopping to accept H_0 only, the interim upper α critical values are set to ∞, $a_k = \infty, k = 1, 2, \ldots, K-1$, and $\alpha = \alpha_K$.

Similarly, the Type I and Type II error probabilities for a K-stage design with a lower alternative hypothesis $H_0 : \theta = -\theta_1$ can also be derived.

Two-Sided Tests

For a K-stage group sequential design with two-sided alternative hypotheses $H_{1u} : \theta = \theta_{1u}$ and $H_{1l} : \theta = \theta_{1l}$, and early stopping to reject or accept the null hypothesis $H_0 : \theta = 0$, the boundaries contain the upper α critical values a_k, upper β critical values b_k, lower β critical values $\_b_k$, and lower α critical values $\_a_k$, $k = 1, 2, \ldots, K$. At each interim stage, $\_a_k < \_b_k \le b_k < a_k$, the null hypothesis H_0 is rejected if the observed statistic $z_k \ge a_k$ or $z_k \le \_a_k$, H_0 is accepted if $\_b_k < z_k < b_k$, or the process is continued to the next stage if $b_k \le z_k < a_k$ or $\_a_k < z_k \le \_b_k$. At the final stage $b_K = a_K$ and $\_b_K = \_a_K$, the hypothesis is either rejected or accepted.

The overall upper Type I error probability α_u is given by

$$\alpha_u = \sum_{k=1}^{K} \alpha_{uk}$$

where α_{uk} is the α spending at stage k for the upper alternative. That is, at stage 1,

$$\alpha_{u1} = P_{\theta=0}(z_1 \ge a_1)$$

At a subsequent stage k,

$$\alpha_{uk} = P_{\theta=0}(\_a_j < z_j \le \_b_j \text{ or } b_j \le z_j < a_j, j = 1, 2, \ldots, k-1, z_k \ge a_k)$$

Similarly, the overall lower Type I error probability α_l can also be derived, and the overall Type I error probability $\alpha = \alpha_l + \alpha_u$.

The overall upper Type II error probability β_u is given by

$$\beta_u = \sum_{k=1}^{K} \beta_{uk}$$

where β_{uk} is the upper β spending at stage k. That is, at stage 1,

$$\beta_{u1} = P_{\theta=\theta_{1u}}(z_1 < \_a_1 \text{ or } \_b_1 < z_1 < b_1)$$

At a subsequent stage k,

$$\beta_{uk} = P_{\theta=\theta_{1u}}(\_a_j < z_j \le \_b_j \text{ or } b_j \le z_j < a_j,\ j = 1, 2, \ldots, k-1,\ z_k < \_a_k \text{ or } \_b_k < z_k < b_k)$$

With an upper alternative hypothesis $H_1 : \theta = \theta_{1u} > 0$, the power $1 - \beta_u$ is the probability of rejecting the null hypothesis for the upper alternative:

$$1 - \beta_u = 1 - \sum_{k=1}^{K} \beta_{uk}$$

which is

$$P_{\theta=\theta_{1u}}(\_a_j < z_j \le \_b_j \text{ or } b_j \le z_j < a_j,\ j = 1, 2, \ldots, k-1,\ z_k \ge a_k)$$

The overall lower Type II error probability β_l and power $1 - \beta_l$ can be similarly derived.

For a design with early stopping only to reject H_0, both the interim lower and upper β critical values are set to missing, $k = 1, 2, \ldots, K-1$, and $\beta_{lK} = \beta_l$, $\beta_{uK} = \beta_u$. For a design with early stopping only to accept H_0, the interim upper α critical values are set to ∞, $a_{uk} = \infty$, and the interim lower α critical values are set to $-\infty$, $a_{lk} = -\infty$, $k = 1, 2, \ldots, K-1$, and $\alpha_{uK} = \alpha_u$, $\alpha_{lK} = \alpha_l$.

Unified Family Methods

Unified family methods (Kittelson and Emerson 1999) derive boundary values with a specified boundary shape. For example, Pocock's method (Pocock 1977) derives equal boundary values for all stages in the standardized Z scale. In addition to Pocock's method, the unified family methods include the O'Brien-Fleming, power family, and unified family triangular methods.

The boundary values at each stage depend on the information fractions

$$\Pi_k = \frac{I_k}{I_X}$$

where I_k is the information available at stage k and I_X is the maximum information, the information available at the end of the trial if the trial does not stop early.

Boundary Values in Standardized *Z* Scale

With the unified family method, the boundary values for the upper α boundary $Z_{\alpha u}$, upper β boundary $Z_{\beta u}$, lower β boundary $Z_{\beta l}$, and lower α boundary $Z_{\alpha l}$, using the standardized normal scale, are given by the following:

- $Z_{\alpha u}(\Pi_k) = f_{\alpha u}(\Pi_k)\, C_{\alpha u}$
- $Z_{\beta u}(\Pi_k) = \theta_{1u}\, I_k^{\frac{1}{2}} - f_{\beta u}(\Pi_k)\, C_{\beta u}$

- $Z_{\beta l}(\Pi_k) = \theta_{1l}\, I_k^{\frac{1}{2}} + f_{\beta l}(\Pi_k)\, C_{\beta l}$
- $Z_{\alpha l}(\Pi_k) = -f_{\alpha l}(\Pi_k)\, C_{\alpha l}$

where $\theta_{1l}(< 0)$ and $\theta_{1u}(> 0)$ are the lower and upper alternative references, $f_{\alpha l}(\Pi_k)$, $f_{\beta l}(\Pi_k)$, $f_{\beta u}(\Pi_k)$, and $f_{\alpha u}(\Pi_k)$ are the specified shape functions, and $C_{\alpha l}$, $C_{\beta l}$, $C_{\beta u}$, and $C_{\alpha u}$ are the critical values derived to achieve the specified α and β levels.

If a derived lower β boundary value $Z_{\beta l}(\Pi_k)$ is greater than its corresponding upper β boundary value $Z_{\beta u}(\Pi_k)$, then both values are set to missing.

Note that the drift parameters $d_l = \theta_{1l}\sqrt{I_X}$ and $d_u = \theta_{1u}\sqrt{I_X}$ are derived in the SEQDESIGN procedure. The boundary values in standardized Z scale can be derived without specifying the maximum information and alternative reference.

Shape Parameters

The shape function in the SEQDESIGN procedure is given by

$$f(\Pi_k) = f(\Pi_k; \tau, \rho) = \tau\, \Pi_k^{\frac{1}{2}} + \Pi_k^{-\rho} = \Pi_k^{\frac{1}{2}}\, (\tau + \Pi_k^{-(\rho+\frac{1}{2})})$$

where the parameters $\rho \geq 0$ and $0 \leq \tau \leq 2\rho$. can be specified for each boundary separately.

The parameters τ and ρ determine the shape of the boundaries. Special cases of the unified family methods also include power family methods and triangular methods. Table 77.6 summarizes the corresponding parameter values in the unified family for these methods.

Table 77.6 Parameters in the Unified Family for Various Methods

Method	Option	Unified Family	
		Rho	Tau
Pocock	POC	0	0
O'Brien-Fleming	OBF	0.5	0
Power family	POW (RHO=ρ)	ρ	0
Triangular	TRI (TAU=τ)	0.5	τ

Note that the power parameter $\rho = 1/2 - \Delta = \rho^* - 1/2$, where Δ is the power parameter used in Jennison and Turnbull (2000) and Wang and Tsiatis (1987) and ρ^* is the power parameter used in Kittelson and Emerson (1999).

Also note that instead of the three parameters used in the unified family methods by Kittelson and Emerson (1999), only two parameters are used in the SEQDESIGN procedure. The other parameter is fixed at zero.

Boundary Values in MLE Scale

If the maximum information is available, the boundary values derived from a unified family method can also be displayed in the MLE scale:

- $\theta_{\alpha u}(\Pi_k) = I_k^{-\frac{1}{2}} f_{\alpha u}(\Pi_k) C_{\alpha u}$
- $\theta_{\beta u}(\Pi_k) = \theta_{1u} - I_k^{-\frac{1}{2}} f_{\beta u}(\Pi_k) C_{\beta u}$
- $\theta_{\beta l}(\Pi_k) = \theta_{1l} + I_k^{-\frac{1}{2}} f_{\beta l}(\Pi_k) C_{\beta l}$
- $\theta_{\alpha l}(\Pi_k) = -I_k^{-\frac{1}{2}} f_{\alpha l}(\Pi_k) C_{\alpha l}$

These MLE scale boundary values are computed by multiplying $I_k^{-\frac{1}{2}}$ by the standardized Z scale boundary values at stage k.

Boundary Values in Score Scale

If the maximum information is available, the boundary values derived from a unified family method can also be displayed in the score scale:

- $S_{\alpha u}(\Pi_k) = I_k^{\frac{1}{2}} f_{\alpha u}(\Pi_k) C_{\alpha u}$
- $S_{\beta u}(\Pi_k) = \theta_{1u} I_k - I_k^{\frac{1}{2}} f_{\beta u}(\Pi_k) C_{\beta u}$
- $S_{\beta l}(\Pi_k) = \theta_{1l} I_k + I_k^{\frac{1}{2}} f_{\beta l}(\Pi_k) C_{\beta l}$
- $S_{\alpha l}(\Pi_k) = -I_k^{\frac{1}{2}} f_{\alpha l}(\Pi_k) C_{\alpha l}$

These MLE scale boundary values are computed by multiplying $I_k^{\frac{1}{2}}$ by the standardized Z scale boundary values at stage k.

Boundary Values in *p*-Value Scale

For a design with a lower alternative or a two-sided alternative, the p-value scale boundary values are the cumulative normal distribution function values of the standardized Z boundary values:

- $P_{\alpha u}(\Pi_k) = \Phi(Z_{\alpha u}(\Pi_k))$
- $P_{\beta u}(\Pi_k) = \Phi(Z_{\beta u}(\Pi_k))$
- $P_{\beta l}(\Pi_k) = \Phi(Z_{\beta l}(\Pi_k))$
- $P_{\alpha l}(\Pi_k) = \Phi(Z_{\alpha l}(\Pi_k))$

These nominal p-values are the one-sided fixed-sample p-values under the null hypothesis with a lower alternative.

For a one-sided design with an upper alternative, the p-value scale boundary values are the one-sided fixed-sample p-values under the null hypothesis with an upper alternative:

- $P_{\alpha u}(\Pi_k) = 1 - \Phi(Z_{\alpha u}(\Pi_k))$
- $P_{\beta u}(\Pi_k) = 1 - \Phi(Z_{\beta u}(\Pi_k))$

Pocock's Method

The shape function for Pocock's method (Pocock 1977) is given by

$$f(\Pi_k) = 1$$

The resulting boundary values for a two-sided design with an early stopping to reject the null hypothesis $H_0 : \theta = 0$ are as follows:

- $Z_{\alpha u}(\Pi_k) = C_{\alpha u}$
- $Z_{\alpha l}(\Pi_k) = -C_{\alpha l}$

That is, the rejection boundary values are constant over all stages of different information levels in the standardized Z scale.

Note that compared with other designs, Pocock's design tends to stop the trials early with a larger p-value. For a new treatment, Pocock's design to stop a trial early with a large p-value might not be persuasive enough to make a new treatment widely accepted (Pocock and White 1999). A Pocock design is illustrated in Example 77.3.

O'Brien-Fleming Method

The shape function for the O'Brien-Fleming method (O'Brien and Fleming 1979) is given by

$$f(\Pi_k) = \Pi_k^{-\frac{1}{2}}$$

The resulting boundary values for a two-sided design with early stopping to reject the null hypothesis $H_0 : \theta = 0$ are as follows:

- $Z_{\alpha u}(\Pi_k) = \Pi_k^{-\frac{1}{2}} C_{\alpha u}$
- $Z_{\alpha l}(\Pi_k) = -\Pi_k^{-\frac{1}{2}} C_{\alpha l}$

That is, the rejection boundaries are inversely proportional to the square root of the information levels in the standardized Z scale.

In the score scale, these boundaries can be displayed as follows:

- $S_{\alpha u}(\Pi_k) = C_{\alpha u}\, I_X^{\frac{1}{2}}$
- $S_{\alpha l}(\Pi_k) = -C_{\alpha l}\, I_X^{\frac{1}{2}}$

which are constants over all stages in the score scale. An O'Brien-Fleming design is illustrated in Example 77.2.

Power Family Method

The shape function for a power family method (Wang and Tsiatis 1987; Emerson and Fleming 1989; Pampallona and Tsiatis 1994) is given by

$$f(\Pi_k) = \Pi_k^{-\rho}$$

The resulting boundary values for a two-sided design with early stopping to reject the null hypothesis $H_0 : \theta = 0$ are as follows:

- $Z_{\alpha u}(\Pi_k) = \Pi_k^{-\rho}\, C_{\alpha u}$
- $Z_{\alpha l}(\Pi_k) = -\Pi_k^{-\rho}\, C_{\alpha l}$

The rejection boundaries depend on the power parameter ρ. The power family includes the Pocock and O'Brien-Fleming methods, and the power parameter is used to allow continuous movement between these two methods. A power family design is illustrated in Example 77.5.

Triangular Method

The shape function for a triangular method (Kittelson and Emerson 1999) in the unified family is given by

$$f(\Pi_k) = \Pi_k^{-\frac{1}{2}} + \tau\, \Pi_k^{\frac{1}{2}}$$

The resulting boundary values for a two-sided design with early stopping to reject the null hypothesis $H_0 : \theta = 0$ are as follows:

- $Z_{\alpha u}(\Pi_k) = (\Pi_k^{-\frac{1}{2}} + \tau\, \Pi_k^{\frac{1}{2}})\, C_{\alpha u} = C_{\alpha u}\, \Pi_k^{-\frac{1}{2}}\, (1 + \tau\, \Pi_k)$
- $Z_{\alpha l}(\Pi_k) = -(\Pi_k^{-\frac{1}{2}} + \tau\, \Pi_k^{\frac{1}{2}})\, C_{\alpha l} = -C_{\alpha l}\, \Pi_k^{-\frac{1}{2}}\, (1 + \tau\, \Pi_k)$

In the score scale, these boundaries are as follows:

- $S_{\alpha u}(\Pi_k) = C_{\alpha u}\, I_X^{\frac{1}{2}}\,(1 + \tau\,\Pi_k) = C_{\alpha u}\, I_X^{\frac{1}{2}} + C_{\alpha u}\tau I_X^{-\frac{1}{2}}\, I_k$
- $S_{\alpha l}(\Pi_k) = -C_{\alpha l}\, I_X^{\frac{1}{2}}\,(1 + \tau\,\Pi_k) = -C_{\alpha l}\, I_X^{\frac{1}{2}} - C_{\alpha l}\tau I_X^{-\frac{1}{2}}\, I_k$

Thus, in the score scale, the boundary function is a linear function of the information I_k. With these straight-line boundaries, a triangular method for a one-sided trial with early stopping to reject or accept the null hypothesis produces a triangular continuation region. Similarly, for a two-sided design, the continuation region is a union of two separate triangular regions. A triangular method is illustrated in Example 77.6.

Haybittle-Peto Method

The Haybittle-Peto method (Haybittle 1971; Peto et al. 1976) uses a value of 3 for the critical values in interim stages, so that the critical value at the final stage is close to the original design without interim monitoring.

In the SEQDESIGN procedure, the Haybittle-Peto method has been generalized to allow for different boundary values at different stages. That is, with the standardized normal scale, the boundary values are given by the following:

- $Z_{\alpha u}(\Pi_k) = z_{\alpha uk}$
- $Z_{\beta u}(\Pi_k) = \theta_{1u}\, I_k^{\frac{1}{2}} - z_{\beta uk}$
- $Z_{\beta l}(\Pi_k) = \theta_{1l}\, I_k^{\frac{1}{2}} + z_{\beta lk}$
- $Z_{\alpha l}(\Pi_k) = -z_{\alpha lk}$

where θ_{1l} and θ_{1u} are the lower and upper alternative references and the boundary values $z_{\alpha uk}$, $z_{\beta uk}$, $z_{\beta lk}$, and $z_{\alpha lk}$ are specified either explicitly with the HP(Z= *numbers*) option or implicitly with the HP(PVALUE= *numbers*) option. The HP(PVALUE= *numbers*) option specifies the nominal p-values p_k for the corresponding boundary values z_k:

$$z_k = \Phi^{-1}(1 - p_k)$$

The Haybittle-Peto method is illustrated in Example 77.5.

Whitehead Methods

The Whitehead methods (Whitehead and Stratton 1983; Whitehead 1997, 2001) derive boundary values by adjusting the boundary values generated from continuous monitoring. With continuous monitoring, the boundary values are on a straight line in the score scale for each boundary. For a group sequential design, the boundary values at an interim stage k depend on the information fractions

$$\Pi_k = \frac{I_k}{I_X}$$

where I_k is the information available at stage k and I_X is the maximum information, the information available at the end of the trial if the trial does not stop early.

One-Sided Symmetric Designs

A one-sided symmetric design is a one-sided design with identical Type I and Type II error probabilities. For a one-sided symmetric design with an upper alternative, $\alpha_u = \beta_u$, the boundary values in the score scale from continuous monitoring are as follows:

- $S_{\alpha u}(\Pi_k) = C_u\theta_u^{-1} + \tau_u\theta_u I_k$
- $S_{\beta u}(\Pi_k) = \theta_u\, I_k - (C_u\theta_u^{-1} - \tau_u\theta_u I_k)$

where θ_u is the upper alternative reference, τ_u is a specified constant for the slope, $0 \le \tau_u < \frac{1}{2}$, and C_u is a constant, fixed for STOP=BOTH and derived for STOP=ACCEPT and STOP=REJECT.

The upper β boundary value can also be expressed as

- $S_{\beta u}(\Pi_k) = -C_u\theta_u^{-1} + (1-\tau_u)\theta_u I_k$

Thus, these straight-line boundaries form a triangle in the score statistic scale.

To adjust for the nature of discrete monitoring, the group sequential boundary values are given by the following:

- $S_{\alpha u}(\Pi_k) = C_u\theta_u^{-1} + \tau_u\theta_u I_k - g_k$
- $S_{\beta u}(\Pi_k) = -C_u\theta_u^{-1} + (1-\tau_u)\theta_u I_k + g_k$

where $g_1 = 0.583\sqrt{I_1}$ and $g_k = 0.583\sqrt{I_k - I_{(k-1)}}$, $k > 1$ are the adjustments.

Note that with the adjustment g_k, the resulting boundaries form a Christmas tree shape within the original triangle and are referred to as the Christmas tree boundaries (Whitehead 1997, p. 73).

One-Sided Asymmetric Designs

For a one-sided asymmetric design with an upper alternative, $\alpha_u \neq \beta_u$, the boundary values computed using the score scale, are given by the following:

- $S_{\alpha u}(\Pi_k) = C_u \tilde{\theta}_u^{-1} + \tau_u \tilde{\theta}_u I_k - g_k$
- $S_{\beta u}(\Pi_k) = -C_u \tilde{\theta}_u^{-1} + (1-\tau_u) \tilde{\theta}_u I_k + g_k$

where $\tilde{\theta_u}$ is the modified alternative reference

$$\tilde{\theta_u} = \frac{2\Phi^{-1}(1-\alpha_u)}{\Phi^{-1}(1-\alpha_u) + \Phi^{-1}(1-\beta_u)}\,\theta_u$$

The modified alternative reference $\tilde{\theta}_u = \theta_u$ if $\alpha_u = \beta_u$.

For a design with early stopping to reject or accept the null hypothesis, $S_{\alpha u}(1) = S_{\beta u}(1)$, the boundary values at the final stage are equal. The modified drift parameter $\tilde{d_u}$ is given by

$$\tilde{d_u} = \tilde{\theta}_u \sqrt{I_X} = \frac{1}{1-2\tau_u}\left(\sqrt{{h_K}^2 + 2C_u(1-2\tau_u)} - h_K\right)$$

where $h_K = g_K\, {I_X}^{-\frac{1}{2}} = 0.583\sqrt{1-\Pi_{(K-1)}}$.

A one-sided Whitehead design with early stopping to reject or accept the null hypothesis is illustrated in Example 77.7.

Two-Sided Designs

The boundary values for a two-sided design are generated by combining boundary values from two one-sided designs. With the STOP=BOTH option, this produces a double triangular design (Whitehead 1997, p. 98).

The boundary values for a two-sided design, using the score scale, are then given by the following:

- $S_{\alpha u}(\Pi_k) = C_u \tilde{\theta}_u^{-1} + \tau_u \tilde{\theta}_u I_k - g_k$
- $S_{\beta u}(\Pi_k) = -C_u \tilde{\theta}_u^{-1} + (1-\tau_u) \tilde{\theta}_u I_k + g_k$
- $S_{\beta l}(\Pi_k) = -C_l \tilde{\theta}_l^{-1} + (1-\tau_l) \tilde{\theta}_l I_k - g_k$
- $S_{\alpha l}(\Pi_k) = C_l \tilde{\theta}_l^{-1} + \tau_l \tilde{\theta}_l I_k + g_k$

where the modified alternative references are

$$\tilde{\theta_u} = \frac{2\Phi^{-1}(1-\alpha_u)}{\Phi^{-1}(1-\alpha_u) + \Phi^{-1}(1-\beta_u)}\,\theta_u$$

$$\tilde{\theta}_l = \frac{2\Phi^{-1}(1-\alpha_l)}{\Phi^{-1}(1-\alpha_l)+\Phi^{-1}(1-\beta_l)}\,\theta_l$$

The modified alternative reference $\tilde{\theta}_u = \theta_u$ if $\alpha_u = \beta_u$ and $\tilde{\theta}_l = \theta_l$ if $\alpha_l = \beta_l$.

For a design with early stopping to reject or accept the null hypothesis, the two upper boundary values at the final stage are identical and the two lower boundary values at the final stage are identical. That is, $S_{\alpha l}(1) = S_{\beta l}(1)$ and $S_{\alpha u}(1) = S_{\beta u}(1)$. These modified drift parameters are then given by

$$\tilde{d}_l = \tilde{\theta}_l\sqrt{I_X} = \frac{1}{1-2\tau_l}\left(\sqrt{{h_K}^2 + 2C_l(1-2\tau_l)} - h_K\right)$$

$$\tilde{d}_u = \tilde{\theta}_u\sqrt{I_X} = \frac{1}{1-2\tau_u}\left(\sqrt{{h_K}^2 + 2C_u(1-2\tau_u)} - h_K\right)$$

where $h_K = g_K I_X^{-\frac{1}{2}} = 0.583\sqrt{1-\Pi_{(K-1)}}$.

For a design with early stopping to reject the null hypothesis, or a design with early stopping to accept the null hypothesis, you can specify the slope parameters τ_u and τ_l in the TAU= option, and then the intercept parameters C_u and C_l, and the resulting boundary values are derived. If both the maximum information and alternative references are specified, the procedure derives C_u and C_l by maintaining either the overall α levels (BOUNDARYKEY=ALPHA) or the overall β levels (BOUNDARYKEY=BETA). If the maximum information and alternative reference are not both specified, the procedure derives the boundary values C_u and C_l by maintaining both the overall α and overall β levels.

For a design with early stopping to reject or accept the null hypothesis (STOP=BOTH), Whitehead's triangular test uses $\tau_u = \tau_l = 0.25$ and compute $C_u = -2\log(2\alpha_u)$ and $C_l = -2\log(2\alpha_l)$ for the boundary values. If the maximum information and alternative reference are both specified, the BOUNDARYKEY=ALPHA option uses the specified α values to compute the β values and boundary values. The final-stage boundary values are modified to maintain the overall α levels if they exist. Similarly, the BOUNDARYKEY=BETA option uses the specified β values to compute the α values and boundary values. The final-stage boundary values are modified to maintain the overall β levels if they exist.

If the maximum information and alternative reference are not both specified, the specified α and β values are used to derive boundary values. The BOUNDARYKEY=NONE option uses these boundary values without adjustment. The BOUNDARYKEY=ALPHA option modifies the final-stage boundary values to maintain the overall α levels if they exist. Similarly, the BOUNDARYKEY=BETA option modifies the final-stage boundary values to maintain the overall β levels if they exist.

Two-sided Whitehead designs with early stopping to reject the null hypothesis are illustrated in Example 77.9.

Applicable Boundary Keys

Table 77.7 lists applicable boundary keys for a design that uses Whitehead methods.

Table 77.7 Applicable Boundary Keys for Whitehead Methods

	Specified Parameters		Boundary Keys			
Early Stopping	**(Alt Ref – Max Info)**	**Tau**	**Alpha**	**Beta**	**None**	**Both**
Reject H_0	X	X	X	X		
Accept H_0	X	X	X	X		
Reject/Accept H_0	X	0.25	X	X		
Reject H_0		X				X
Accept H_0		X				X
Reject/Accept H_0		0.25	X	X	X	

Note that the symbol "X" under "(Alt Ref – Max Info)" indicates that both alternative reference and maximum information are specified.

For a design with early stopping to reject the null hypothesis, or a design with early stopping to accept the null hypothesis, you can specify the slope parameter τ_u in the TAU= option, and then the intercept parameter C_u and the resulting boundary values are derived. If both the maximum information and alternative reference are specified, the procedure derives C_u by maintaining either the overall α levels (BOUNDARYKEY=ALPHA) or the overall β levels (BOUNDARYKEY=BETA). If the maximum information and alternative reference are not both specified, the procedure derives the boundary values and C_u by maintaining both the overall α and overall β levels.

For a design with early stopping to reject or accept the null hypothesis (STOP=BOTH), Whitehead's triangular test uses $\tau_u = 0.25$ and solves $C_u = 2\log(\frac{1}{2\alpha_u})$ for the boundary values. If the maximum information and alternative reference are both specified, the BOUNDARYKEY=ALPHA option uses the specified α value to compute the β value and boundary values. The final-stage boundary value is modified to maintain the overall α level if it exists. Similarly, the BOUNDARYKEY=BETA option uses the specified β value to compute the α value and boundary values. The final-stage boundary value is modified to maintain the overall β level if it exists.

If the maximum information and alternative reference are not both specified, the specified α and β values are used to derive boundary values. The BOUNDARYKEY=NONE option uses these boundary values without adjustment. The BOUNDARYKEY=ALPHA option modifies the final-stage boundary value to maintain the overall α level if it exists. Similarly, the BOUNDARYKEY=BETA option modifies the final-stage boundary value to maintain the overall β level if it exists.

Error Spending Methods

For each sequential design, the α and β errors spent at each stage can be computed from the boundary values. For example, for a K-stage design with an upper alternative hypothesis $H_1 : \theta = \theta_1$ and early stopping to reject the null hypothesis $H_0 : \theta = 0$, the boundary values in a standardized Z

scale are the upper α critical values $a_k, k = 1, 2, \ldots, K$. At each interim stage, the null hypothesis H_0 is rejected if the observed standardized Z statistic $z_k \geq a_k$. Otherwise, the process continues to the next stage. At the final stage, the hypothesis is rejected if $z_K \geq a_K$. Otherwise, the null hypothesis is accepted.

The boundary values a_k are derived such that the overall Type I error probability

$$\alpha = \sum_{k=1}^{K} \alpha_k$$

where α_k is the α spending at stage k. That is, at stage 1,

$$\alpha_1 = P_{\theta=0}\,(\,z_1 \geq a_1\,)$$

At a subsequent stage k,

$$\alpha_k = P_{\theta=0}\,(\,z_j < a_j, j = 1, 2, \ldots, k-1,\ z_k \geq a_k\,)$$

Since each design can be uniquely identified by the α and β errors spent at each stage, a design can then be derived by specifying the α and β errors to be used at each stage. The error spending method (Lan and DeMets 1983) distributes the error to be used at each stage and then derives the boundary values. Numerous forms of the error spending function are available. Kim and DeMets (1987) examine the functions $f(t) = t$, $f(t) = t^{\frac{3}{2}}$, and $f(t) = t^2$, where t is the information fraction. Jennison and Turnbull (2000, p. 148) generalize these functions to the power functions $f(t; \rho) = t^\rho, \rho > 0$.

The ERRFUNCPOC option uses the cumulative error spending function (Lan and DeMets 1983)

$$E(t) = \begin{cases} 1 & \text{if } t \geq 1 \\ \log(\,1 + (e-1)t) & \text{if } 0 < t < 1 \\ 0 & \text{otherwise} \end{cases}$$

With a specified error of α or β, the cumulative error spending at stage k is $\alpha\, E(\Pi_k)$ or $\beta\, E(\Pi_k)$, where $\Pi_k = I_k / I_X$ is the information fraction at stage k. The method produces boundaries similar to those produced with Pocock's method.

The ERRFUNCOBF option uses the cumulative error spending function (Lan and DeMets 1983)

$$E(t; a) = \begin{cases} 1 & \text{if } t \geq 1 \\ \frac{1}{a}\, 2\,(1 - \Phi(\frac{z_{(1-a/2)}}{\sqrt{t}})) & \text{if } 0 < t < 1 \\ 0 & \text{otherwise} \end{cases}$$

where a is either α for the α spending function or β for the β spending function. That is, with a specified error of α or β, the cumulative error spending at stage k is $\alpha\, E(\Pi_k; \alpha)$ or $\beta\, E(\Pi_k; \beta)$. The method produces boundaries similar to those produced with the O'Brien-Fleming method.

The ERRFUNCGAMMA option uses the gamma cumulative error spending function (Hwang, Shih, and DeCani 1990)

$$E(t;\gamma) = \begin{cases} 1 & \text{if } t \geq 1 \\ \frac{1-e^{-\gamma t}}{1-e^{-\gamma}} & \text{if } 0 < t < 1, \gamma \neq 0 \\ t & \text{if } 0 < t < 1, \gamma = 0 \\ 0 & \text{otherwise} \end{cases}$$

where γ is the parameter γ specified in the GAMMA= option. That is, with a specified error of α or β, the cumulative error spending at stage k is $\alpha\, E(\Pi_k;\gamma)$ or $\beta\, E(\Pi_k;\gamma)$.

The ERRFUNCPOW option uses the cumulative error spending function (Jennison and Turnbull 2000, p. 148)

$$E(t;\rho) = \begin{cases} 1 & \text{if } t \geq 1 \\ t^{\rho} & \text{if } 0 < t < 1 \\ 0 & \text{otherwise} \end{cases}$$

where ρ is the power parameter specified in the RHO= option. That is, with a specified error of α or β, the cumulative error spending at stage k is $\alpha\, E(\Pi_k;\rho)$ or $\beta\, E(\Pi_k;\rho)$.

Error spending methods derive boundary values at each stage sequentially and require much more computation than other types of methods for group sequential trials with a large number of stages, especially for a two-sided asymmetric design with early stopping to accept H_0, or to reject or accept H_0.

Note that for a two-sided design with the STOP=BOTH or STOP=ACCEPT option, at each interim stage, the SEQDESIGN procedure first produces the lower and upper β boundary values based on the one-sided β spending. If the lower β boundary value is greater than or equal to its corresponding upper β boundary value, there is no early stopping to accept the null hypothesis at this stage, and the corresponding β spending is distributed proportionally to the remaining stages.

For the error spending functions not available in the SEQDESIGN procedure, you can first compute the corresponding error spending at each stage explicitly, then use the SEQDESIGN procedure with the ERRSPEND= option to specify these errors directly.

For example, if the information levels are equally spaced in a five-stage design, the option ERRFUNCPOW (RHO=2) produces relative cumulative errors of $(1/5)^2$, $(2/5)^2$, $(3/5)^2$, $(4/5)^2$, and 1. This is equivalent to using the option ERRSPEND (1 4 9 16 25).

A one-sided error spending design is illustrated in Example 77.8 and a two-sided asymmetric error spending design is illustrated in Example 77.11.

Boundary Adjustments for Overlapping Lower and Upper β Boundaries

For the fixed boundary shape methods and Whitehead methods, the boundary values for all stages are derived simultaneously for each boundary. For a two-sided design with STOP=ACCEPT or

STOP=BOTH, simultaneous derivation might result in overlapping of the lower and upper β boundaries. That is, at an interim stage k, the lower β boundary value might be greater than its corresponding upper β boundary value. In this case, these two β boundary values are set to missing and the design does not stop at stage k to accept the null hypothesis (Jennison and Turnbull 2000, p. 113).

For the error spending methods, the boundary values are derived sequentially for the stages. For a two-sided design with STOP=ACCEPT or STOP=BOTH, a small β spending at an interim stage might result in overlapping of the lower and upper β boundaries for the two corresponding one-sided tests. Specifically, this form of overlapping occurs at an interim stage k if the upper β boundary value derived from the one-sided test for the upper alternative is less than the lower β boundary value derived from the one-sided test for the lower alternative (Kittelson and Emerson 1999, pp. 881–882; Rudser and Emerson 2007, p. 6). You can use the BETAOVERLAP= option to specify how this type of overlapping is to be handled.

If BETAOVERLAP=ADJUST (which is the default) is specified, the procedure derives the boundary values for the two-sided design and then checks for overlapping of the two one-sided β boundaries at interim stages. If overlapping occurs at a particular stage, the β boundary values for the two-sided design are set to missing (so the trial does not stop to accept the null hypothesis at this stage), and the β spending values at subsequent stages are adjusted proportionally as follows.

If the β boundary values are set to missing at stage k in a K-stage trial, the adjusted β spending value at stage k, e'_k, is updated for these missing β boundary values, and then the β spending values at subsequent stages are adjusted proportionally by

$$e'_j = e'_k + \frac{e_j - e_k}{e_K - e_k}\,(e_K - e'_k)$$

for $j = k+1, \ldots, K$, where e_j and e'_j are the cumulative β spending values st stage j before and after the adjustment, respectively.

After all these adjusted β spending values are computed, the boundary values are then further modified for these adjusted β spending values.

If you specify BETAOVERLAP=NOADJUST, no adjustment is made when overlapping of one-sided β boundaries occurs. The BETAOVERLAP= option is illustrated in Example 77.10.

Specified and Derived Parameters

In the SEQDESIGN procedure, the type of alternative hypothesis (ALT= option) and the condition for early stopping (STOP= option) must be specified for each sequential design. The drift parameters are derived for each design specified. Other parameters, such as Type I error probability α, Type II error probability β, the alternative reference θ_1, and maximum information are either specified or derived in the SEQDESIGN procedure.

Table 77.8 summarizes the available combinations for the specified and derived parameters in the SEQDESIGN procedure.

Table 77.8 Specified and Derived Parameters in the SEQDESIGN Procedure

Specified Parameters				Derived Parameters				
Alt Ref	**Max Info**	**Alpha**	**Beta**	**Alt Ref**	**Max Info**	**Alpha**	**Beta**	**Drift**
Z	X	X					X	X
Z	X		X			X		X
Z	X					X	X	X
Z		X	X		X			X
	X	X	X	X				X
		X	X					X

The symbol "X" indicates that the parameter is either specified or derived in the design and the symbol "Z" indicates that the alternative reference is either specified explicitly with the ALTREF= option or derived from the SAMPLESIZE statement. The drift parameter is always derived in the SEQDESIGN procedure.

For example, if the ALTREF= option is specified without the MAXINFO= option being specified, then the maximum information is derived in the SEQDESIGN procedure with the specified α and β, as illustrated in Example 77.5.

The drift parameter is the standardized reference difference at the final stage. For a design, the drift parameter is

$$d_l = \theta_{1l}\sqrt{I_X}$$

if it has a lower alternative, and

$$d_u = \theta_{1u}\sqrt{I_X}$$

if it has an upper alternative, where I_X is the maximum information and θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively.

If the alternative reference and the maximum information are not both specified, then the specified α and β are used to derive the drift parameter. Then if either the alternative reference or the maximum information is specified, the other is derived from the drift parameter.

If both the alternative reference and the maximum information are specified, then either the α error or the β error is derived in the procedure. However, for a Haybittle-Peto method with the BOUNDARYKEY=NONE option, both α and β errors are derived from the completely specified boundary values.

For a nonsymmetric two-sided design with different lower and upper specifications (such as different lower and upper α errors, β errors, or alternative references in absolute values), the derived lower and upper boundaries are not symmetric. If the alternative references θ_{1l} and θ_{1u} are not both specified, then the SEQDESIGN procedure assumes symmetric alternative references, $\theta_1 = \theta_{1u} = -\theta_{1l}$, for the computation of the boundary values.

Applicable Boundary Keys

In the SEQDESIGN procedure, the BOUNDARYKEY= option in the DESIGN statement specifies the types of errors to be maintained for the design. Table 77.9 lists applicable boundary keys for designs that use unified family and Haybittle-Peto methods, designs that use error spending methods, and designs that use the Haybittle-Peto method only.

Table 77.9 Applicable Boundary Keys for Designs without Whitehead Methods

Method	Specified Parameters (Alt Ref – Max Info)	Boundary Keys Alpha	Beta	None	Both
Unified	X	X	X		
Unified/Haybittle-Peto	X	X	X		
Error spending	X	X	X		
Haybittle-Peto	X	X	X	X	
Unified/Haybittle-Peto					X
Error spending					X
Haybittle-Peto					X

Note that the symbol "X" under "(Alt Ref – Max Info)" indicates that both the alternative reference and maximum information are specified, and the method "Unified/Haybittle-Peto" indicates that both the unified method and the Haybittle-Peto method are used in the same design.

If the ALTREF= and MAXINFO= options are both specified, then Type I and Type II error probability levels cannot be met simultaneously if both error probabilities are specified. The BOUNDARYKEY=ALPHA option maintains the Type I error probability level α and derives Type II error probability β. The BOUNDARYKEY=BETA option maintains the Type II error probability level β and derives Type I error probability α.

If the Haybittle-Peto method is used for all boundaries, the BOUNDARYKEY=NONE option uses the specified α boundary value (STOP=REJECT or STOP=BOTH) and the specified β boundary value (STOP=ACCEPT) at the final stage for the design.

If the ALTREF= and MAXINFO= options are not both specified, the BOUNDARYKEY=BOTH option derives boundary values that maintain both Type I and Type II error probability levels.

Table 77.10 lists applicable boundary keys for a design that uses Whitehead methods.

Table 77.10 Applicable Boundary Keys for Whitehead Methods

Early Stopping	Specified Parameters (Alt Ref – Max Info)	Tau	Boundary Keys Alpha	Beta	None	Both
Reject H_0	X	X	X	X		
Accept H_0	X	X	X	X		
Reject/Accept H_0	X	0.25	X	X		
Reject H_0		X				X
Accept H_0		X				X
Reject/Accept H_0		0.25	X	X	X	

Note that the symbol "X" under "(Alt Ref – Max Info)" indicates that both alternative reference and maximum information are specified.

If the ALTREF= and MAXINFO= options are both specified, then Type I and Type II error probability levels cannot be achieved simultaneously if both are specified. the BOUNDARYKEY=ALPHA option maintains the Type I error probability level α and derives Type II error probability β. The BOUNDARYKEY=BETA option maintains the Type II error probability level β and derives the Type I error probability α.

If the ALTREF= and MAXINFO= options are not both specified, then for a design with the STOP=REJECT or STOP=ACCEPT option, the BOUNDARYKEY=BOTH option derives boundary values that maintain both Type I and Type II error probability levels. If the STOP=BOTH option is specified, Whitehead's triangular method produces boundaries with approximate Type I and Type II error probabilities. The BOUNDARYKEY=NONE option specifies no adjustment to these boundaries. The BOUNDARYKEY=ALPHA and BOUNDARYKEY=BETA options maintain the Type I error probability level α and Type II error probability level β, respectively, by adjusting boundary values at the final stage.

Sample Size Computation

The SEQDESIGN procedure assumes that the data are from a multivariate normal distribution and the sequence of the standardized test statistics $\{Z_1, Z_2, \ldots, Z_K\}$ has the following canonical joint distribution:

- $(Z_1, Z_2, \ldots, Z_K)$ is multivariate normal
- $Z_k \sim N\left(\theta\sqrt{I_k}, 1\right)$
- $\mathrm{Cov}(Z_{k_1}, Z_{k_2}) = \sqrt{I_{k_1}/I_{k_2}} \quad , \quad 1 \le k_1 \le k_2 \le K$

where K is the total number of stages and I_k is the information available at stage k.

If the test statistic is computed from the data that are not from a normal distribution, such as a binomial distribution, then it is assumed that the test statistic is computed from a large sample such that the statistic has an approximately normal distribution.

In a typical clinical trial, the sample size required depends on the Type I error probability level α, alternative reference θ_1, power $1-\beta$, and variance of the response variable. Given a one-sided null hypothesis $H_0 : \theta = 0$ with an upper alternative hypothesis $H_1 : \theta = \theta_1$, the information required for a fixed-sample test is given by

$$I_0 = \frac{(\Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta))^2}{\theta_1^2}$$

The parameter θ and the subsequent alternative reference θ_1 depend on the test specified in the clinical trial. For example, suppose you are comparing two binomial populations $p_a = p_b$; then $\theta = p_a - p_b$ is the difference between two proportions if the proportion difference statistic is used,

and $\theta = \log\left(\frac{p_a(1-p_b)}{p_b(1-p_a)}\right)$, the log odds ratio for the two proportions if the log odds ratio statistic is used.

If the maximum likelihood estimate $\hat{\theta}$ from the likelihood function can be derived, then the asymptotic variance for $\hat{\theta}$ is $\mathrm{Var}(\hat{\theta}) = 1/I$, where I is Fisher information for θ. The resulting statistic $\hat{\theta}$ corresponds to the MLE statistic scale as specified in the BOUNDARYSCALE=MLE option in the PROC SEQDESIGN statement, $\hat{\theta}\sqrt{I}$ corresponds to the standardized Z scale (BOUNDARYSCALE=STDZ), and $\hat{\theta}\, I$ corresponds to the score statistic scale (BOUNDARYSCALE=SCORE).

Alternatively, if the score statistic S is derived in a statistical procedure, it can be used as the test statistic and its asymptotic variance is given by Fisher information, I. In this case, $S/\sqrt{I}$ corresponds to the standardized Z scale and S/I corresponds to the MLE statistic scale.

For a group sequential trial, the maximum information I_X is derived in the SEQDESIGN procedure with the specified α, β, and θ_1. With the maximum information

$$I_X = \frac{1}{\mathrm{Var}(\hat{\theta})}$$

the sample size required for a specified test statistic in the trial can be evaluated or estimated from the known or estimated variance of the response variable. Note that different designs might produce different maximum information levels for the same hypothesis, and this in turn might require a different number of observations for the trial.

If each observation in the data set provides one unit of information in a hypothesis testing, such as a one-sample test for the mean, the required sample size for the sequential design can be derived from the maximum information. However, for a survival analysis, an individual in the survival time data might provide only partial information because of censoring. In this case, the required number of events can be derived from the maximum information. With addition accrual information, the sample size can also be computed.

The SEQDESIGN procedure provides sample size computation for some one-sample and two-sample tests in the SAMPLESIZE statement. It also provides sample size computation for tests of a parameter in regression models such as normal regression, logistic regression, and proportional hazards regression. In addition, the procedure can also compute the required sample size or number of events from the corresponding number in the fixed-sample design.

Table 77.11 lists the options available in the SAMPLESIZE statement.

Table 77.11 SAMPLESIZE Statement Options

Option	Description
Fixed-Sample Models	
INPUTNOBS	specifies sample size for fixed-sample design
INPUTNEVENTS	specifies number of events for fixed-sample design
One-Sample Models	
ONESAMPLEMEAN	specifies one-sample Z test for mean
ONESAMPLEFREQ	specifies one-sample test for binomial proportion

Table 77.11 *continued*

Option	Description
Two-Sample Models	
TWOSAMPLEMEAN	specifies two-sample Z test for mean difference
TWOSAMPLEFREQ	specifies two-sample test for binomial proportions
TWOSAMPLESURVIVAL	specifies log-rank test for two survival distributions
Regression Models	
REG	specifies test for a regression parameter
LOGISTIC	specifies test for a logistic regression parameter
PHREG	specifies test for a proportional hazards regression parameter

The MODEL=INPUTNOBS and MODEL=INPUTNEVENTS options are described next, and the remaining options are described in the next three sections.

Input Sample Size for Fixed-Sample Design

The MODEL=INPUTNOBS option derives the sample size required for a group sequential trial from the sample size n_0 for the corresponding fixed-sample design. With the N= n_0 option specifying the sample size n_0 for a fixed-sample design, the sample size required for a group sequential trial is then computed as

$$N_X = \frac{I_X}{I_0}\, n_0$$

where I_X is the maximum information for the group sequential design and I_0 is the information for the corresponding fixed-sample design. The information ratio between I_X and I_0 is derived in the SEQDESIGN procedure.

The SAMPLE=ONE option specifies a one-sample test, and the SAMPLE=TWO option specifies a two-sample test. For a two-sample test, the WEIGHT= option specifies the sample size allocation weights for the two groups.

Input Number of Events for Fixed-Sample Design

The MODEL=INPUTNOBS option derives the number of events required for a group sequential trial from the number of events d_0 for the corresponding fixed-sample design. With the D= d_0 option specifies the number of events d_0 for a fixed-sample survival analysis, the number of events required for a group sequential trial is then computed as

$$d_X = \frac{I_X}{I_0}\, d_0$$

where I_X is the maximum information for the group sequential design and I_0 is the information for the corresponding fixed-sample design. The information ratio between I_X and I_0 is derived in the SEQDESIGN procedure.

The SAMPLE=ONE option specifies a one-sample test, and the SAMPLE=TWO option specifies a two-sample test. For a two-sample test, the WEIGHT= option specifies the sample size allocation weights for the two groups.

With the computed number of events d_X for a group sequential survival design, the required total sample size and sample size at each stage can be derived with specifications of hazard rates, accrual rate, and accrual time.

For a study group, if the hazard rate h is constant, corresponding to an exponential survival distribution, and the individual accrual is uniform in the accrual time T_a with a constant accrual rate r_a, Kim and Tsiatis (1990, pp. 83–84) show that the expected number of events by time t is given by

$$D_h(t) = \begin{cases} r_a \left(t - \frac{1-e^{-ht}}{h} \right) & \text{if } t \le T_a \\ r_a \left(T_a - \frac{e^{-ht}}{h}(e^{hT_a} - 1) \right) & \text{if } t > T_a \end{cases}$$

For a one-sample design, such as a proportional hazards regression, the expected number of events by time t is $E(t) = D_h(t)$, where h is the hazard rate for the group. For a two-sample design, such as a log-rank test for two survival distributions, the expected number of events by time t is

$$E(t) = \frac{R}{R+1} D_{h_a}(t) + \frac{1}{R+1} D_{h_b}(t)$$

where h_a and h_b are hazard rates in groups A and B, respectively, and R is the ratio of the sample size allocation weights w_a/w_b.

If the accrual rate r_a is specified without the accrual time T_a, follow-up time T_f, and total study time $T = T_a + T_f$, the SEQDESIGN procedure computes the minimum and maximum accrual times from the following equation, as described in Kim and Tsiatis (1990, p. 85):

$$\frac{d_X}{r_a} \le T_a \le E^{-1}(d_X)$$

If the accrual rate r_a is specified with one of the three time parameters—the accrual time, follow-up time, and total study time—then the other two time parameters are computed in the SEQDESIGN procedure. Similarly, if the accrual rate r_a is not specified, but two of the three time parameters are specified, then the accrual rate is derived in the SEQDESIGN procedure.

With the accrual rate r_a and the accrual time T_a, the total sample size is

$$N_X = r_a \; T_a$$

At each stage k, the number of events is given by

$$d_k = \frac{I_k}{I_X} \; d_X$$

The corresponding time T_k can be derived from the equation for the expected number of events, $E(t) = d_k$, and the resulting sample size is computed as

$$N_k = r_a \; T_k$$

The following three sections describe examples of test statistics with their resulting information levels, which can then be used to derive the required sample size. The maximum likelihood estimators are used for all tests except to compare two survival distributions with a log-rank test, where a score statistic is used.

Applicable One-Sample Tests and Sample Size Computation

The SEQDESIGN procedure provides sample size computation for two one-sample tests: normal mean and binomial proportion. The required sample size depends on the variance of the response variable—that is, the sample proportion for a binomial proportion test.

In a typical clinical trial, a hypothesis is designed to reject, not accept, the null hypothesis to show the evidence for the alternative hypothesis. Thus, in most cases, the proportion under the alternative hypothesis is used to derive the required sample size. For a test of the binomial proportion, the REF=NULLPROP and REF=PROP options use proportions under the null and alternative hypotheses, respectively.

Test for a Normal Mean

The MODEL=ONESAMPLEMEAN option in the SAMPLESIZE statement derives the sample size required to test a normal mean by using the sample mean statistic for the null hypothesis $\mu = \mu_0$. At stage k, the sample mean is computed as

$$\overline{y}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} y_{kj}$$

where y_{kj} is the value of the jth observation available in the kth stage and N_k is the cumulative sample size at stage k.

An equivalent hypothesis is $H_0 : \theta = 0$, where $\theta = \mu - \mu_0$.

The MLE statistic for θ,

$$\hat{\theta}_k = \overline{y}_k - \mu_0 \sim N\left(\theta,\ {I_k}^{-1}\right)$$

where the information

$$I_k = \frac{1}{\mathrm{Var}(\hat{\theta})} = \frac{1}{\mathrm{Var}(\overline{y}_k)} = \frac{N_k}{\sigma^2}$$

is the inverse of the variance.

That is, the standardized statistic

$$Z_k = \hat{\theta}_k \sqrt{I_k} = (\overline{y}_k - \mu_0)\sqrt{I_k} \sim N\left(\theta\sqrt{I_k},\ 1\right)$$

Thus, to test the hypothesis $H_0 : \theta = 0$ against a two-sided alternative $H_1 : \theta = \theta_1$, H_0 is rejected at stage k if the statistic Z_k is less than or equal to the lower α boundary value or if Z_k is greater than or equal to the upper α boundary value at stage k.

If the variance σ^2 is unknown, the sample variance can be used if it is assumed that the sample variance is computed from a large sample such that the test statistic has an approximately normal distribution.

The maximum information is needed to derive the required sample size. If the maximum information is not specified or derived with the ALTREF= option in the procedure, the MEAN=θ_1 option in the SAMPLESIZE statement is used to specify the alternative reference and thus to derive the maximum information.

In the SEQDESIGN procedure, the computed total sample size

$$N_K = \sigma^2 \; I_X$$

where I_X is the maximum information and σ is the specified standard deviation. With an available maximum information, you can specify the MODEL=ONESAMPLEMEAN(STDDEV= σ) option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage. A procedure such as PROC MEANS can be used to derive a one-sample Z test for a normal mean.

Test for a Binomial Proportion

The MODEL=ONESAMPLEFREQ option in the SAMPLESIZE statement derives the sample size required to test a binomial proportion by using the null hypothesis $p = p_0$, where p is the proportion of a binomial population. At stage k, the MLE for p is computed as

$$\hat{p}_k = \frac{1}{N_k} \sum_{j=1}^{N_k} y_{kj}$$

where y_{kj} is the value of the jth observation available in the kth stage and N_k is the cumulative sample size at stage k.

An equivalent hypothesis is $H_0 : \theta = 0$, where $\theta = p - p_0$. If p_0 is not close to 0 or 1, then for a large sample, $\hat{\theta}_k = \hat{p}_k - p_0$ has an approximately normal distribution

$$\hat{\theta}_k \sim N\left(\theta, \; I_k^{-1}\right)$$

where the information $I_k = (p\,(1-p) \,/\, N_k)^{-1}$ is the inverse of the variance Var($\hat{\theta}$).

Then the standardized statistic

$$Z_k = \hat{\theta}_k \sqrt{I_k} \sim N\left(\theta \sqrt{I_k}, \; 1\right)$$

In practice, the estimated sample proportion $\hat{p}$ at stage k can be used to derive the information I_k and test statistic Z_k. Thus, to test the hypothesis H_0 against an upper alternative $H_1 : \theta = \theta_1 > 0$, H_0 is rejected at stage k if the statistic Z_k is greater than or equal to the upper α boundary at stage k.

The maximum information I_X is needed to derive the required sample size. If the maximum information is not specified or derived with the ALTREF= option in the procedure, the PROP= option in the SAMPLESIZE statement is used to specify the alternative reference and to derive the maximum information for the sample size calculation.

It is assumed that the sample size is sufficiently large such that the test statistic has an approximately normal distribution. With the hypotheses $H_0 : p = p_0$ and $H_1 : p = p_1$, the SEQDESIGN procedure derives the total sample size

$$N_X = p^* (1 - p^*) I_X$$

where $p^* = p_0$ if REF=NULLPROP is specified. Otherwise, $p^* = p_1$.

If the PROP= option in the SAMPLESIZE statement is not specified, then the alternative reference θ_1 derived in the SEQDESIGN procedure is used to compute $p_1 = p_0 + \theta_1$.

The ALTREF= option in the PROC statement can be used to specify θ_1. Otherwise, the PROP= option in the SAMPLESIZE statement must be specified.

For example, with $H_0 : p = 0.5$, $H_1 : p = 0.6$ and the default REF=PROP,

$$N_K = p^*(1 - p^*) I_X = (0.6 \times 0.4) I_X = 0.24 I_X$$

You can specify the MODEL=ONESAMPLEFREQ option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage. A procedure such as PROC GENMOD with the default DIST=NORMAL option in the MODEL statement can be used to derive the Z test for a binomial proportion.

Applicable Two-Sample Tests and Sample Size Computation

The SEQDESIGN procedure provides sample size computation for two-sample tests: the test for the difference between two normal means, tests for binomial proportions, and the log-rank test for two survival distributions. These tests for binomial proportions include the test for the difference between two binomial proportions, the log odds ratio test for binomial proportions, and the log relative risk test for binomial proportions,

For a test of difference between two sample means, the required sample size depends on the assumed sample variances. Similarly, for a test of two-sample proportions, the required sample size depends on the assumed sample proportions. For a log-rank test of two survival distributions, the required sample size depends on the assumed sample hazard rates, accrual rate, and accrual time.

If the REF=NULLPROP or REF=NULLHAZARD option is specified, the proportions or hazard rates under the null hypothesis are used to derive the required sample size or number of events. Otherwise, the default REF=PROP option in the TWOSAMPLEFREQ option or the REF=HAZARD option in the TWOSAMPLESURVIVAL option uses proportions or hazard rates under the alternative hypothesis to derive the required sample size or number of events.

Test for the Difference between Two Normal Means

The MODEL=TWOSAMPLEMEAN option in the SAMPLESIZE statement derives the sample size required to test the difference between the means of two normal populations μ_a and μ_b by using the null hypothesis $H_0 : \theta = 0$, where $\theta = \mu_a - \mu_b$.

At stage k, the MLE for θ is computed as

$$\hat{\theta}_k = \overline{y}_{ak} - \overline{y}_{bk} = \frac{1}{N_{ak}} \sum_{j=1}^{N_{ak}} y_{akj} - \frac{1}{N_{bk}} \sum_{j=1}^{N_{bk}} y_{bkj}$$

where y_{akj} and y_{bkj} are the values of the jth observation available in the kth stage groups A and B, respectively, and N_{ak} and N_{bk} are the cumulative sample sizes at stage k for these two groups.

The statistic $\hat{\theta}_k$ has a normal distribution

$$\hat{\theta}_k \sim N\left(\theta,\ {I_k}^{-1}\right)$$

where the information I_k is the inverse of the variance $\text{Var}(\hat{\theta}_k) = \sigma_a^2/N_{ak} + \sigma_b^2/N_{bk}$.

Then the standardized statistic

$$Z_k = \hat{\theta}_k \sqrt{I_k} \sim N\left(\theta \sqrt{I_k},\ 1\right)$$

Thus, to test the hypothesis $H_0 : \theta = 0$ against an upper alternative $H_1 : \theta = \theta_1, \theta_1 > 0$, H_0 is rejected at stage k if the statistic $Z_k \geq a_k$, the upper α boundary for the standardized Z statistic at stage k.

If the variances σ_a^2 and σ_b^2 are unknown, the sample variances can be used to derive the information I_k if it is assumed that each sample variance is computed from a large sample such that the test statistic has an approximately normal distribution.

The maximum information is needed to derive the required sample size. If the maximum information is not specified or derived in the procedure, the alternative reference θ_1^* specified in the MEANDIFF option is used to derive the maximum information.

Note that in order to derive the sample sizes N_{ak} and N_{bk} uniquely from the information, $N_{ak} = R\, N_{bk}$ is assumed for $k = 1, 2, \ldots, K$, where $R = w_a/w_b$ is the constant allocation ratio computed from the WEIGHT=w_a w_b option in the SAMPLESIZE statement.

In PROC SEQDESIGN, the computed total sample sizes for the two groups are

$$N_{aK} = (\sigma_a^2 + R\,\sigma_b^2)\, I_X = R\,(\frac{\sigma_a^2}{R} + \sigma_b^2)\, I_X$$

$$N_{bK} = (\frac{\sigma_a^2}{R} + \sigma_b^2)\, I_X$$

where I_X is the maximum information derived in the SEQDESIGN procedure, R is the constant allocation ratio, and σ_a and σ_b are the specified standard deviations.

For $R = 1$, the two sample sizes are equal, then

$$N_{aK} = N_{bK} = \frac{N_K}{2} = (\sigma_a^2 + \sigma_b^2)\, I_X$$

If the variances from the two groups are equal, $\sigma_a^2 = \sigma_b^2 = \sigma^2$, then the total sample sizes for the two groups are

$$N_{aK} = (1 + R)\,\sigma^2\, I_X$$

$$N_{bK} = (1 + \frac{1}{R})\,\sigma^2\, I_X$$

and the total sample size is

$$N_X = N_{aK} + N_{bK} = \frac{(R+1)^2}{R}\,\sigma^2\, I_X$$

Furthermore, for $R = 1$, the two sample sizes are equal, then

$$N_{aK} = N_{bK} = \frac{N_X}{2} = 2\,\sigma^2\, I_X$$

With an available maximum information, you can specify the MODEL=TWOSAMPLEMEAN(WEIGHT= R STDDEV= σ_a σ_b) option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage. A procedure such as PROC GLM can be used to derive the two-sample Z test for the mean difference.

Test for the Difference between Two Binomial Proportions

The MODEL=TWOSAMPLEFREQ(TEST=PROP) option in the SAMPLESIZE statement derives the sample size required to test the difference between two binomial populations with $H_0 : \theta = 0$, where $\theta = p_a - p_b$. At stage k, the MLE for θ is

$$\hat{\theta}_k = \hat{p}_{ak} - \hat{p}_{bk} = \frac{1}{N_{ak}} \sum_{j=1}^{N_{ak}} y_{akj} - \frac{1}{N_{bk}} \sum_{j=1}^{N_{bk}} y_{bkj}$$

where y_{akj} and y_{bkj} are the values of the jth observation available in the kth stage for groups A and B, respectively, and N_{ak} and N_{bk} are the cumulative sample sizes at stage k for these two groups.

For sufficiently large sample sizes N_{ak} and N_{bk}, the statistic $\hat{\theta}_k$ has an approximate normal distribution

$$\hat{\theta}_k \sim N\left(\theta,\, {I_k}^{-1}\right)$$

where the information is the inverse of the variance

$$\mathrm{Var}(\hat{\theta}_k) = \frac{p_a\,(1-p_a)}{N_{ak}} + \frac{p_b\,(1-p_b)}{N_{bk}}$$

Thus, the standardized statistic

$$Z_k = \hat{\theta}_k \sqrt{I_k} \sim N\left(\theta\sqrt{I_k},\, 1\right)$$

In practice, $p_a = \hat{p}_a$ and $p_b = \hat{p}_b$, the estimated sample proportions for groups A and B, respectively, at stage k, can be used to derive the information I_k and the test statistic Z_k. Thus, to test the hypothesis H_0 against an upper alternative $H_1 : \theta > 0$, H_0 is rejected at stage k if the statistic $Z_k \geq a_k$, the upper α boundary for the standardized Z statistic at stage k.

The maximum information I_X is needed to derive the required sample size. If the maximum information is not specified or derived with the ALTREF= option in the procedure, the PROP= option in the SAMPLESIZE statement is used to provide proportions under the alternative hypothesis for the alternative reference and then to derive the maximum information.

The proportions in the two groups are needed to derive the sample size. Also, in order to derive the sample sizes N_{ak} and N_{bk} uniquely from the information, $N_{ak} = R\, N_{bk}$ is assumed for $k = 1, 2, \ldots, K$, where $R = w_a/w_b$ is the constant allocation ratio computed from the WEIGHT=w_a w_b option in the SAMPLESIZE statement. Then

$$I_X = \left(\frac{p_a\,(1-p_a)}{N_{aK}} + \frac{p_b\,(1-p_b)}{N_{bK}} \right)^{-1} = \frac{N_{aK}}{p_a(1-p_a) + R\,p_b(1-p_b)}$$

In PROC SEQDESIGN, the total sample sizes in the two groups are computed as

$$N_{aK} = \left(p_a^*\,(1-p_a^*) + R\,p_b^*\,(1-p_b^*) \right)\, I_X$$

$$N_{bK} = \frac{1}{R}\, N_{aK}$$

where $R = w_a/w_b$ is the constant allocation ratio, and p_a^* and p_b^* are proportions specified with the REF= option:

- REF=NULLPROP uses proportions under H_0: $p_a^* = p_{0a}$, $p_b^* = p_{0b}$
- REF=AVGNULLPROP uses the average proportion under H_0: $p_a^* = p_b^* = (Rp_{0a} + p_{0b})/(R+1)$
- REF=PROP uses proportions under H_1: $p_a^* = p_{1a}$, $p_b^* = p_{1b}$
- REF=AVGPROP uses the average proportion under $H_1 : p_a^* = p_b^* = (Rp_{1a} + p_{1b})/(R+1)$

The total sample size is given by

$$N_X = N_{aK} + N_{bK} = (R+1)\left(\frac{1}{R}\, p_a^*\,(1-p_a^*) + p_b^*\,(1-p_b^*) \right)\, I_X$$

For $R = 1$, the two sample sizes are equal,

$$N_{aK} = N_{bK} = \frac{N_X}{2} = \left(p_a^*\,(1-p_a^*) + p_b^*\,(1-p_b^*) \right)\, I_X$$

You can specify the MODEL=TWOSAMPLEFREQ(TEST=PROP WEIGHT=R) option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage. A procedure such as PROC GENMOD with the default DIST=NORMAL option in the MODEL statement can be used to derive the two-sample Z test for proportion difference.

Test for Two Binomial Proportions with a Log Odds Ratio Statistic

The MODEL=TWOSAMPLEFREQ(TEST=LOGOR) option in the SAMPLESIZE statement derives the sample size required to test two binomial proportions by using a log odds ratio statistic. The odds ratio is the ratio of the odds in one group to the odds in the other group, and the log odds ratio is the logarithm of the odds ratio

$$\theta = \log\left(\frac{p_a/(1-p_a)}{p_b/(1-p_b)}\right) = \log\left(\frac{p_a(1-p_b)}{p_b(1-p_a)}\right)$$

The hypothesis of no difference between two proportions, $p_a = p_b$, can be tested through the null hypothesis $H_0 : \theta = 0$, where θ is the log odds ratio. For example, with $H_0 : p_a = p_b = 0.6$ and $H_1 : p_a = 0.8$, $p_b = 0.6$, it corresponds to the equivalent hypothesis $H_0 : \theta = 0$ and $H_1 : \theta = \log\left(\frac{0.8(1-0.6)}{0.6(1-0.8)}\right) = \log(8/3) = 0.98083$.

The maximum likelihood estimate of θ is given by

$$\hat{\theta} = \log\left(\frac{\hat{p}_a(1-\hat{p}_b)}{\hat{p}_b(1-\hat{p}_a)}\right)$$

with an asymptotic variance

$$\mathrm{Var}(\hat{\theta}) = I^{-1} = \frac{1}{N_a\, p_a(1-p_a)} + \frac{1}{N_b\, p_b(1-p_b)}$$

where I is the information (Diggle et al. 2002, pp. 341–342). That is, the standardized statistic

$$Z_k = \hat{\theta}_k \sqrt{I_k} \sim N\left(\theta\sqrt{I_k},\, 1\right)$$

In practice, $p_a = \hat{p}_a$ and $p_b = \hat{p}_b$, the estimated sample proportions for groups A and B, respectively, at stage k, can be used to derive the information I_k and the test statistic $Z_k = \hat{\theta}_k\sqrt{I_k}$ if the two sample sizes N_a and N_b are sufficiently large such that the test statistic has an approximately normal distribution.

The maximum information I_X is needed to derive the required sample size. If the maximum information is not specified or derived with the ALTREF= option in the procedure, the PROP= option in the SAMPLESIZE statement is used to provide proportions under the alternative hypothesis for the alternative reference and then to derive the maximum information.

In order to derive the sample sizes N_{ak} and N_{bk} uniquely from the information, $N_{ak} = R\, N_{bk}$ is assumed for $k = 1, 2, \ldots, K$, where $R = w_a/w_b$ is the constant allocation ratio computed from the WEIGHT=w_a w_b option in the SAMPLESIZE statement. Then with

$$I_X = N_{bK}\left(\frac{1}{R\, p_a(1-p_a)} + \frac{1}{p_b(1-p_b)}\right)^{-1}$$

the sample size can be computed.

In PROC SEQDESIGN, the total sample sizes in the two groups are computed as

$$N_{bK} = I_X \left(\frac{1}{R\, p_a^*(1-p_a^*)} + \frac{1}{p_b^*(1-p_b^*)} \right)$$

$$N_{aK} = R\, N_{bK}$$

where $R = w_a/w_b$ is the constant allocation ratio, and p_a^* and p_b^* are proportions specified with the REF= option:

- REF=NULLPROP uses proportions under H_0: $p_a^* = p_{0a}$, $p_b^* = p_{0b}$
- REF=AVGNULLPROP uses the average proportion under H_0: $p_a^* = p_b^* = (Rp_{0a} + p_{0b})/(R+1)$
- REF=PROP uses proportions under H_1: $p_a^* = p_{1a}$, $p_b^* = p_{1b}$
- REF=AVGPROP uses the average proportion under H_1 : $p_a^* = p_b^* = (Rp_{1a} + p_{1b})/(R+1)$

You can specify the MODEL=TWOSAMPLEFREQ(TEST=LOGOR WEIGHT=R) option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage. A procedure such as PROC LOGISTIC can be used to derive the log odds ratio statistic.

Test for Two Binomial Proportions with a Log Relative Risk Statistic

The MODEL=TWOSAMPLEFREQ(TEST=LOGRR) option in the SAMPLESIZE statement derives the sample size required to test two binomial proportions by using a log relative risk statistic. The relative risk is the ratio of the proportion in one group to the proportion in the other group. The log relative risk statistic is the logarithm of the relative risk

$$\theta = \log\left(\frac{p_a}{p_b}\right)$$

The hypothesis of no difference between two proportions, $p_a = p_b$, can be tested through the null hypothesis $H_0 : \theta = 0$. For example, with $H_0 : p_a = p_b = 0.6$ and $H_1 : p_a = 0.8$, $p_b = 0.6$, it corresponds to the equivalent hypothesis $H_0 : \theta = 0$ and $H_1 : \theta = \log\left(\frac{0.8}{0.6}\right) = \log(4/3) = 0.28768$.

The maximum likelihood estimate of θ is given by

$$\hat{\theta} = \log\left(\frac{\hat{p}_a}{\hat{p}_b}\right)$$

with an asymptotic variance

$$I^{-1} = \frac{1-p_a}{N_a\, p_a} + \frac{1-p_b}{N_b\, p_b}$$

where I is the information (Chow and Liu 1998, p. 329).

In practice, $p_a = \hat{p}_a$ and $p_b = \hat{p}_b$, the estimated sample proportions for groups A and B, respectively, at stage k, are used to derive the information I_k and the test statistic $Z_k = \hat{\theta}_k \sqrt{I_k}$.

The maximum information I_X and proportions p_a and p_b are needed to derive the required sample size. If the maximum information is not specified or derived with the ALTREF= option in the procedure, the PROP= option in the SAMPLESIZE statement is used to provide proportions under the alternative hypothesis for the alternative reference and then to derive the maximum information.

Note that in order to derive the sample sizes N_{ak} and N_{bk} uniquely from the information, $N_{ak} = R\, N_{bk}$ is assumed for $k = 1, 2, \ldots, K$, where $R = w_a/w_b$ is the constant allocation ratio computed from the WEIGHT=w_a w_b option in the SAMPLESIZE statement. Then the sample size can be computed from

$$I_X = N_{bK} \left(\frac{1 - p_a}{R\, p_a} + \frac{1 - p_b}{p_b} \right)^{-1}$$

In PROC SEQDESIGN, the computed sample sizes in the two groups are

$$N_{bK} = I_X \left(\frac{1 - p_a^*}{R\, p_a^*} + \frac{1 - p_b^*}{p_b^*} \right)$$

$$N_{aK} = R\, N_{bK}$$

where $R = w_a/w_b$ is the constant allocation ratio, and p_a^* and p_b^* are proportions specified with the REF= option:

- REF=NULLPROP uses proportions under H_0: $p_a^* = p_{0a}$, $p_b^* = p_{0b}$
- REF=AVGNULLPROP uses the average proportion under H_0: $p_a^* = p_b^* = (Rp_{0a} + p_{0b})/(R+1)$
- REF=PROP uses proportions under H_1: $p_a^* = p_{1a}$, $p_b^* = p_{1b}$
- REF=AVGPROP uses the average proportion under H_1 : $p_a^* = p_b^* = (Rp_{1a} + p_{1b})/(R+1)$

You can specify the MODEL=TWOSAMPLEFREQ(TEST=LOGRR WEIGHT=R) option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage. A procedure such as PROC LOGISTIC can be used to derive the log relative risk statistic.

Test for Two Survival Distributions with a Log-Rank Test

The MODEL=TWOSAMPLESURV option in the SAMPLESIZE statement derives the number of events required for a log-rank test of two survival distributions. The analysis of survival data involves the survival times for both censored and uncensored data. A noncensored survival time is the time from treatment to an event such as remission or relapse for an individual. A censored survival time is the time from treatment to the time of analysis for an individual surviving at that time, and the status is unknown beyond that time.

Let T be the random variable of the survival time. Then the survival function

$$S(t) = \Pr(T > t)$$

is the probability that an individual from the population has a survival time that exceeds t. And the hazard function is given by

$$h(t) = \frac{f(t)}{S(t)}$$

where $f(t)$ is the density function of T.

The hazard functions can be used to test the equality of two survival distributions $S_a(t) = S_b(t)$ with the null hypothesis $H_0 : h_a(t) = h_b(t), t > 0$, where $S_a(t)$ and $S_b(t)$ are survival functions for groups A and B, respectively, and $h_a(t)$ and $h_b(t)$ are the corresponding hazard functions.

If the two hazards are proportional, $h_a(t) = \lambda\, h_b(t)$, where λ is a constant, then an equivalent null hypothesis is

$$H_0 : \lambda = \frac{h_a(t)}{h_b(t)} = 1$$

Alternatively, another equivalent null hypothesis is given by

$$H_0 : \theta = -\log(\lambda) = 0$$

Suppose that the hazard rate h is a constant. Then with a specified median survival time T_m, the hazard rate can be derived from the equation

$$e^{-h\,T_m} = \frac{1}{2}$$

Denote the distinct event times at stage k as $\tau_{kj}, j = 1, 2, \ldots, t_k$, where t_k is the total number of distinct event times. Then the score statistic is the log-rank statistic (Jennison and Turnbull 2000, pp. 259–261; Whitehead 1997, pp. 36–39)

$$S_k = \sum_{j=1}^{t_k} (d_{akj} - e_{akj})$$

where d_{akj} is the number of events from group A and e_{akj} is the number of expected events from A. The number of expected events from A is computed as

$$e_{akj} = d_{kj} \frac{r_{akj}}{r_{kj}}$$

where d_{kj} is the number of events from both groups, r_{akj} is the number of individuals from the treatment group who survived up to time τ_{kj}, and r_{kj} is the number of individuals from both groups who survived up to time τ_{kj}.

If the number of events d_{kj} is small relative to r_{kj}, the number of individuals survived up to time τ_{kj}, then with a sufficiently large sample size, S_k has an approximately normal distribution

$$S_k \sim N(\theta\, I_k,\ I_k)$$

where the variance of S_k is the estimated information

$$I_k = \sum_{j=1}^{t_k} \frac{r_{akj}\, r_{bkj}\, d_{kj}}{r_{kj}^2}$$

In order to derive the number of events from the information I_k, $N_{ak} = R\, N_{bk}$ is assumed for $k = 1, 2, \ldots, K$, where $R = w_a/w_b$ is the constant allocation ratio computed from the WEIGHT=w_a w_b option in the SAMPLESIZE statement.

The maximum information I_X is needed to derive the required sample size. If the maximum information is specified or derived with the ALTREF= option in the procedure, the HAZARD=, MEDSURVTIME=, and HAZARDRATIO= options are not applicable. Otherwise, the HAZARD=, MEDSURVTIME=, or HAZARDRATIO= option is used to compute the alternative reference and then to derive the maximum information for the sample size calculation.

With $N_{aK} = R\, N_{bK}$, if the number of events is few relative to the number of individuals who survived, then $r_{aKj} \approx R\, r_{bKj}$, and

$$I_X \approx \sum_{j=1}^{t_K} \frac{R}{(R+1)^2} d_{Kj} = \frac{R}{(R+1)^2}\, D_X$$

where D_X is the total number of events.

Thus, the required total number of events

$$D_X = \frac{(R+1)^2}{R}\, I_X$$

For a study group, if the hazard rate is constant, corresponding to an exponential survival distribution, and the individual accrual is uniform in the accrual time T_a with a constant accrual rate r_a, then the required total sample size and sample size at each stage can be derived. See the section "Input Number of Events for Fixed-Sample Design" on page 5856 for a detailed description of the sample size computation that uses hazard rates, accrual rate, and accrual time.

You can specify the MODEL=TWOSAMPLESURVIVAL option in the SAMPLESIZE statement to compute the required total number of events and individual number of events at each stage. With the specifications of hazard rates, accrual rate, and accrual time, the required total sample size and individual sample size at each stage can also be derived. If the REF=NULLHAZARD option is specified, the hazard rates under the null hypothesis, h_{0a} and h_{0b}, are used in the sample size computation. Otherwise, the hazard rates under the alternative hypothesis, h_{1a} and h_{1b}, are used. A procedure such as PROC LIFETEST can be used to derive the log-rank statistic.

Applicable Regression Parameter Tests and Sample Size Computation

The SEQDESIGN procedure provides sample size computation for tests of a regression parameter in three regression models: normal regression, logistic regression, and proportional hazards regression.

To test a parameter β_1 in a regression model, the variance of the parameter estimate $\hat{\beta}_1$ is needed for the sample size computation. In a simple regression model with one covariate X1, the variance of $\hat{\beta}_1$ is inversely related to the variance of X1, σ_x^2. That is,

$$\text{Var}(\hat{\beta}_1) \propto \frac{1}{N\,\sigma_x^2}$$

for the normal regression and logistic regression models, where N is the sample size, and

$$\text{Var}(\hat{\beta}_1) \propto \frac{1}{D\,\sigma_x^2}$$

for the proportional hazards regression model, where D is the number of events.

For a regression model with more than one covariate, the variance of $\hat{\beta}_1$ for the normal regression and logistic regression models is inversely related to the variance of X1 after adjusting for other covariates. That is,

$$\text{Var}(\hat{\beta}_1) \propto \frac{1}{N\,(1-r_x^2)\,\sigma_x^2}$$

where $\hat{\beta}_1$ is the estimate of the parameter β_1 in the model and r_x^2 is the R square from the regression of X1 on other covariates—that is, the proportion of the variance σ_x^2 explained by these covariates.

Similarly, for a proportional hazards regression model,

$$\text{Var}(\hat{\beta}_1) \propto \frac{1}{D\,(1-r_x^2)\,\sigma_x^2}$$

Thus, with the derived maximum information, the required sample size or number of events can also be computed for the testing of a parameter in a regression model with covariates.

Test for a Parameter in the Regression Model

The MODEL=REG option in the SAMPLESIZE statement derives the sample size required for a Z test of a normal regression. For a normal linear regression model, the response variable is normally distributed with the mean equal to a linear function of the explanatory variables and the constant variance σ^2.

The normal linear model is

$$\mathbf{y} \sim N\left(\mathbf{X}\boldsymbol{\beta},\, \sigma_y^2\,\mathbf{I}_{(N)}\right)$$

where $\mathbf{Y}_{(N\times 1)}$ is the vector of the N observed responses, $\mathbf{X}_{(N\times p)}$ is the design matrix for these N observations, $\boldsymbol{\beta}_{(p\times 1)}$ is the parameter vector, and $\mathbf{I}_{(N)}$ is the $(N \times N)$ identity matrix.

The least squares estimate is

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{Y}$$

and is normally distributed with mean $\boldsymbol{\beta}$ and variance

$$\text{Var}(\hat{\boldsymbol{\beta}}) = \sigma_y^2(\mathbf{X}'\mathbf{X})^{-1}$$

For a model with only one covariate X1,

$$\hat{\beta}_1 \sim N\left(\beta_1,\ \text{Var}(\hat{\beta}_1)\right)$$

where the variance

$$\text{Var}(\hat{\beta}_1) = I_{\beta_1}^{-1} = \sigma_y^2\ \frac{1}{N\,\sigma_x^2}$$

Thus, with the derived maximum information $I_X = I_{\beta_1}$, the required sample size is given by

$$N = I_X\ \frac{\sigma_y^2}{\sigma_x^2}$$

For a normal linear model with more than one covariate, the variance of a single parameter β_1 is

$$\text{Var}(\hat{\beta}_1) = \sigma_y^2\ (\mathbf{X}'\mathbf{X})^{-1}_{(11)} = \sigma_y^2\ \frac{1}{N\,\sigma_x^2\,(1-r_x^2)}$$

where $(\mathbf{X}'\mathbf{X})^{-1}_{(11)}$ is the diagonal element of the $(\mathbf{X}'\mathbf{X})^{-1}$ matrix corresponding to the parameter β_1, σ_x^2 is the variance of the variable X1, and r_x^2 is the proportion of variance of X1 explained by other covariates. The value $\sigma_x^2\,(1-r_x^2)$ represents the variance of X1 after adjusting for all other covariates.

Thus, with the derived maximum information I_X, the required sample size is

$$N = I_X\ \frac{\sigma_y^2}{(1-r_x^2)\,\sigma_x^2}$$

In the SEQDESIGN procedure, you can specify the MODEL=REG(VARIANCE=σ_y^2 XVARIANCE=σ_x^2 XRSQUARE=r_x^2) option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage. A SAS procedure such as PROC REG can be used to compute the parameter estimate and its standard error at each stage.

Test for a Parameter in the Logistic Regression Model

The MODEL=LOGISTIC option in the SAMPLESIZE statement derives the sample size required for a Z test of a logistic regression parameter. The linear logistic model has the form

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \mathbf{x}\boldsymbol{\beta}$$

where p is the response probability to be modeled and $\boldsymbol{\beta}$ is a vector of parameters.

Following the derivation in the section "Test for a Parameter in the Regression Model" on page 5869, the required sample size for testing a parameter in $\boldsymbol{\beta}$ is given by

$$N = I_X\ \frac{\sigma_y^2}{(1-r_x^2)\,\sigma_x^2}$$

With the variance of the logit response, $\sigma_y^2 = 1/(p(1-p))$,

$$N = I_X \frac{1}{p(1-p)} \frac{1}{(1-r_x^2)\,\sigma_x^2}$$

where σ_x^2 is the variance of X and r_x^2 is the proportion of variance explained by other covariates.

In the SEQDESIGN procedure, you can specify the MODEL=LOGISTIC(PROP=p XVARIANCE=σ_x^2 XRSQUARE=r_x^2) option in the SAMPLESIZE statement to compute the required total sample size and individual sample size at each stage.

A SAS procedure such as PROC LOGISTIC can be used to compute the parameter estimate and its standard error at each stage.

Test for a Parameter in the Proportional Hazards Regression Model

The MODEL=PHREG option in the SAMPLESIZE statement derives the number of events required for a Z test of a proportional hazards regression parameter. For analyses of survival data, Cox's semiparametric model is often used to examine the effect of explanatory variables on hazard rates. The survival time of each observation in the population is assumed to follow its own hazard function, $h_i(t)$, expressed as

$$h_i(t) = h(t; \mathbf{X}_i) = h_0(t)\exp(\mathbf{X}_i'\boldsymbol{\beta})$$

where $h_0(t)$ is an arbitrary and unspecified baseline hazard function, $\mathbf{x}_i$ is the vector of explanatory variables for the ith individual, and $\boldsymbol{\beta}$ is the vector of regression parameters associated with the explanatory variables.

Hsieh and Lavori (2000, p. 553) show that the required number of events for testing a parameter in $\boldsymbol{\beta}$, β_1, associated with the variable X1 is given by

$$D_X = I_X \frac{1}{(1-r_x^2)\,\sigma_x^2}$$

where σ_x^2 is the variance of X1 and r_x^2 is the proportion of variance of X1 explained by other covariates.

In the SEQDESIGN procedure, you can specify the MODEL=PHREG(XVARIANCE=σ_x^2 XRSQUARE=r_x^2) option in the SAMPLESIZE statement to compute the required number of events and individual number of events at each stage.

A SAS procedure such as PROC PHREG can be used to compute the parameter estimate and its standard error at each stage.

Note that for a two-sample test, X1 is an indicator variable and is the only covariate in the model. Thus, if the two sample sizes are equal, then the variance $\sigma_x^2 = 1/4$ and the required number of events for testing the parameter β_1 is given by

$$D_X = I_X \frac{1}{\sigma_x^2} = 4\,I_X$$

See the section "Input Number of Events for Fixed-Sample Design" on page 5856 for a detailed description of the sample size computation that uses hazard rates, accrual rate, and accrual time.

Aspects of Group Sequential Designs

This section summarizes various aspects of group sequential designs that are encountered in applications of the SEQDESIGN procedure. Features are illustrated through two-sided designs with $\alpha = 0.05$ and $\beta = 0.10$. The null hypothesis $H_0 : \theta = 0$ and an alternative reference $\theta_1 = \pm 0.25$ are used for the designs with early stopping only to reject the null hypothesis.

Canonical Joint Distribution

The SEQDESIGN procedure assumes that with a total number of stages K, the sequence of the standardized test statistics $\{Z_1, Z_2, \ldots, Z_K\}$ has the canonical joint distribution with information levels $\{I_1, I_2, \ldots, I_K\}$ for the parameter θ (Jennison and Turnbull 2000, p. 49):

- $(Z_1, Z_2, \ldots, Z_K)$ is multivariate normal
- $Z_k \sim N\left(\theta\sqrt{I_k},\ 1\right),\ k = 1, 2, \ldots, K$
- $\mathrm{Cov}(Z_{k_1}, Z_{k_2}) = \sqrt{(I_{k_1}/I_{k_2})},\quad 1 \le k_1 \le k_2 \le K$

Normality Assumption

The SEQDESIGN procedure derives the boundary values by assuming that the sequence of the standardized test statistics $\{Z_1, Z_2, \ldots, Z_K\}$ has the canonical joint distribution with information levels $\{I_1, I_2, \ldots, I_K\}$ for the parameter θ. If the test statistic Z_k does not have a normal distribution, it is assumed that the test statistic is computed from a large sample, so that the resulting statistic has an approximately normal distribution.

Number of Stages

For group sequential trials with fixed significance level α, power $1 - \beta$, and alternative reference, if the number of stages is increased, the required maximum information is also increased, but the average sample number under the alternative hypothesis is likely to decrease.

For example, with O'Brien-Fleming designs specified in the section, the maximum information increases from 168.12 for a fixed-sample design to 169.32 for a two-stage design, 172.57 for a five-stage design, and then 174.42 for a ten-stage design. In the mean time, the average sample number (as a percentage of fixed-sample) under the alternative hypothesis decreases from 100 for a fixed-sample design to 85.11 for a two-stage design, 75.03 for a five-stage design, and then 71.80 for a ten-stage design. The reduction in average sample number decreases as the number of stages increases. Thus there seems to be little to gain from choosing a design with more than five stages (Pocock 1982, p. 155).

Alternative Reference

The alternative reference θ_1 is the hypothetical reference under the alternative hypothesis at which the power is computed. It is a treatment value that the investigators would hope to detect with high probability (Jennison and Turnbull 2000, p. 21).

For a group sequential design with specified parameters such as α and β errors, the drift parameter $\theta_1\sqrt{I_X}$ is always derived in the SEQDESIGN procedure. Thus, with a smaller alternative reference θ_1, a larger maximum information level I_X is needed. That is, in order to detect a smaller difference with the same high power, a larger sample size is required.

Maximum Information

In a clinical trial, the amount of information about an unknown parameter available from the data can be measured by the Fisher information, the variance of the score statistic. The maximum information is the information level needed at the final stage of the group sequential trial if the trial does not stop at an interim stage. For a group sequential design, the maximum information can be derived with the specified alternative reference.

The maximum information is proportional to the sample size or number of events required for the design. Thus, it can also be used to compare different designs. Generally, a design with a larger probability to stop the trial early tends to have a larger maximum information. For example, with four-stage designs specified in the section, the Pocock method has a maximum information of 198.91 and the O'Brien-Fleming method has a maximum information of 171.85, indicating a much larger information level required for the Pocock method.

Drift Parameter

The drift parameter $\theta_1\sqrt{I_X}$ is derived for each design in the SEQDESIGN procedure, where θ_1 is the alternative reference. It is proportional to the square root of maximum information required for the design and can be used to compare maximum information for different designs with the same alternative reference. For example, with the four-stage designs specified in the section, the Pocock method has a drift parameter 3.526 and the O'Brien-Fleming method has a drift parameter 3.277, indicating that a larger maximum information level is required for the Pocock method than for the O'Brien-Fleming method.

Average Sample Number

The average sample number is the expected sample size (for nonsurvival data) or expected number of events (for survival data) of the design under a specific hypothetical reference. The percent average sample numbers with respect to the corresponding fixed-sample design are displayed in the SEQDESIGN procedure.

The design that requires a larger maximum information level tends to have a smaller average sample number under the alternative hypothesis. For example, with the four-stage designs specified in the section, the Pocock design has a maximum information of 198.91 and an average sample number

(in percentage of fixed-sample design) of 69.748 under the alternative hypothesis, and the O'Brien-Fleming design has a maximum information of 171.85 and an average sample number of 76.740.

Sample Size

The maximum information for the sequential design expressed as a percentage of its corresponding fixed-sample information is derived in the SEQDESIGN procedure. The sample size or number of events needed for a group sequential trial is computed by multiplying the sample size or number of events for the corresponding fixed-sample design by the derived percentage.

If the sample size or number of events for the fixed-sample design is available, you can use the MODEL=INPUTNOBS or MODEL=INPUTNEVENTS option in the SAMPLESIZE statement to derive the sample size or number of events needed at each stage. Otherwise, with the specified or derived maximum information, you can use the MODEL= option in the SAMPLESIZE statement to specify a hypothesis test and then to derive the sample size or number of events needed at each stage. See the section "Sample Size Computation" on page 5854 for the sample size computation for commonly used tests.

Summary of Methods in Group Sequential Designs

There are three different types of methods available in the SEQDESIGN procedure: fixed boundary shape methods for specified boundary shape, Whitehead methods for boundaries from continuous monitoring, and error spending methods for specified error spending at each stage.

The fixed boundary shape methods include unified family methods and Haybittle-Peto methods. The unified family methods include Pocock, O'Brien-Fleming, power family, and triangular methods.

Pocock Method

Pocock derives the constant boundary on the standardized Z scale to demonstrate the sequential design while maintaining the overall α and β levels (Pocock 1977). The resulting boundary tends to stop the trials early with a larger p-value. This boundary is commonly called a Pocock boundary, but Pocock himself does not advocate these boundary values for stopping a trial early to reject the null hypothesis, because large p-values might not be persuasive enough (Pocock and White 1999). Also, the nominal p-value at the final stage is much smaller than the overall p-value of the design. That is, the trial might stop at the final stage with a small nominal p-value, but the test is not rejected, which might not be easy to justify.

O'Brien-Fleming Method

O'Brien-Fleming boundary values are inversely proportional to the square root of information levels on the standardized Z scale (O'Brien and Fleming 1979). The O'Brien-Fleming boundary is

conservative in the early stages and tends to stop the trials early only with a small p-value. But the nominal value at the final stage is close to the overall p-value of the design.

Power Family Method

The power family method (Wang and Tsiatis 1987; Emerson and Fleming 1989; Pampallona and Tsiatis 1994) generalizes the Pocock and O'Brien-Fleming methods with a power parameter to allow continuous movement between the Pocock and O'Brien-Fleming methods. The power parameter is $\rho = 0$ for the Pocock method and $\rho = 0.5$ for the O'Brien-Fleming method.

Triangular Method

The unified family triangular method (Kittelson and Emerson 1999) contains straight-line boundaries on the score scale. For a one-sided trial with early stopping either to reject and to accept the null hypothesis, the method produces a triangular continuation region. The boundary shape is specified with the slope parameter τ.

Unified Family Method

The unified family method (Kittelson and Emerson 1999) extends power family methods to incorporate the triangular method, which contains straight-line boundaries on the score scale.

Haybittle-Peto Method

The Haybittle-Peto method (Haybittle 1971; Peto et al. 1976) uses a Z value of 3 for the critical values in interim stages and derives the critical value at the final stage. With this method, the final-stage critical value is close to the original design without interim monitoring. The SEQDESIGN procedure extends this method further to allow for different Z or nominal p-values for the boundaries.

Whitehead Method

Whitehead methods (Whitehead and Stratton 1983; Whitehead 1997, 2001) derive the boundary values by adapting the continuous monitoring tests to the discrete monitoring of group sequential tests. With early stopping to reject or accept the null hypothesis in a one-sided test, the derived continuation region has a triangular shape on the score-scaled boundaries. Only elementary calculations are needed to derive the boundary values in Whitehead's triangular methods. The resulting Type I error probability and power are extremely close but differ slightly from the specified values due to the approximations used in deriving the tests (Jennison and Turnbull 2000, p. 106). The SEQDESIGN procedure provides the BOUNDARYKEY= option to adjust the boundary value at the final stage for the exact Type I or Type II error probability levels.

Error Spending Method

The error spending method uses the specified α and β errors to be used at each stage of the design to derive the boundary values.

Error Spending Function Method

The error spending function method uses the error spending function to compute the α and β errors to be used at each stage of the design and then to derive the boundary values for these errors. The following four error spending functions are available in the SEQDESIGN procedure:

- The Pocock-type error spending function (Lan and DeMets 1983) produces boundaries similar to those produced with Pocock's method.
- The O'Brien-Fleming-type error spending function (Lan and DeMets 1983) produces boundaries similar to those produced with the O'Brien-Fleming method.
- The gamma error spending function (Hwang, Shih, and DeCani 1990) specifies a gamma cumulative error spending function indexed by the gamma parameter γ. The boundaries created with $\gamma = 1$ are similar to the boundaries from the Pocock method, and the boundaries created with $\gamma = -4$ or $\gamma = -5$ are similar to the boundaries from the O'Brien-Fleming method.
- The power error spending function (Jennison and Turnbull 2000, p. 148) specifies a power cumulative error spending function indexed by the power parameter ρ. The boundaries created with $\rho = 1$ are similar to the boundaries from the Pocock method, and the boundaries created with $\rho = 3$ are similar to the boundaries from the O'Brien-Fleming method.

Table Output

For each design, the SEQDESIGN procedure displays the "Design Information," "Method Information," and "Boundary Information" tables by default.

Boundary Information

The "Boundary Information" table displays the following information at each stage:

- proportion of information
- actual information level, if the maximum information is either specified or derived
- alternative references with the specified statistic scale. If a p-value scale is specified, the standardized Z scale is used.
- boundary values with the specified statistic scale to reject or accept the null hypothesis

Note that implicitly, the boundary information table also contains variables for the boundary scale, stopping criterion, and type of alternative hypothesis. That is, if an ODS statement is used to save the table, the data set also contains the variables _Scale_ for the boundary scale, _Stop_ for the stopping criterion, and _ALT_ for the type of alternative hypothesis.

Design Information

The "Design Information" table displays the design specifications and derived statistics. The derived Max Information (Percent Fixed-Sample) is the maximum information for the sequential design in percentage of the corresponding fixed-sample information.

The Null Ref ASN (Percent Fixed-Sample) is the average sample number (expected sample size for nonsurvival data or expected number of events for survival data) required under the null hypothesis for the group sequential design in percentage of the corresponding fixed-sample design. Similarly, the Alt Ref ASN (Percent Fixed-Sample) is the average sample number required under the alternative reference for the group sequential design in percentage of the corresponding fixed-sample design.

If both the maximum information (MAXINFO= option) and the alternative reference θ_1 (ALTREF= option) are specified, then either the ALPHA= option is used to derive the Type II error probability β (BOUNDARYKEY=ALPHA) or the BETA= option is used to derive the Type I error probability α (BOUNDARYKEY=BETA).

Error Spending Information

The "Error Spending Information" table displays the following information at each stage:

- proportion of information
- actual information level, if the maximum information is either specified or derived
- cumulative error spending for each boundary

Method Information

The "Method Information" table displays detailed method information for the design. For each boundary, it displays the following:

- the group sequential method used
- the α or β errors
- the specified parameter ρ, if an error spending function is used
- the specified parameters ρ and τ with the derived critical value C, if a unified family method is used

- the alternative reference θ_1, if either the ALTREF= or the MAXINFO= option is specified
- the derived drift parameter, $\theta_1\sqrt{I_X}$, where I_X is the maximum information and θ_1 is the alternative reference

Note that the alternative references are displayed with the MLE scale in the "Method Information" table. In contrast, the alternative references in the "Boundary Information" table are displayed with the specified statistic scale (if the p-value scale is not specified) or the standardized Z scale (if the p-value scale is specified).

Powers and Expected Sample Sizes

The "Powers and Expected Sample Sizes" table displays the following information under each of the specified hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and c_i are values specified in the CREF= option.

- coefficient c_i for the hypothetical references. The value $c_i = 0$ corresponds to the null hypothesis and $c_i = 1$ corresponds to the alternative hypothesis
- power
- expected sample size, as percentage of fixed-sample size

For a one-sided design, the power and expected sample sizes under the hypothetical references $\theta = c_i\theta_1$ are displayed.

For a two-sided symmetric design, the power and expected sample sizes under each of the hypothetical references $\theta = c_i\theta_{1u}$ are displayed, where θ_{1u} is the upper alternative reference.

For a two-sided asymmetric design, the power and expected sample sizes under each of the hypothetical references $\theta = c_i\theta_{1l}$ and $\theta = c_i\theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively.

For a two-sided design, the power is the probability of correctly rejecting the null hypothesis for the correct alternative. Thus, under the null hypothesis, the displayed power corresponds to a one-sided Type I error probability level—that is, the lower α level or the upper α level.

The expected sample size as a percentage of the corresponding fixed-sample design is

$$100 \times \frac{\sum_{k=1}^{K} p_k\, I_k}{I_0}$$

where p_k is the stopping probability at stage k, $\sum_{k=1}^{K} p_k\, I_k$ is the expected information level, and I_0 is the information level for the fixed-sample design.

Sample Size Summary

When you use the SAMPLESIZE statement with the SEQDESIGN procedure, the "Sample Size Summary" table displays parameters for the sample size computation. It also displays the expected sample sizes or numbers of events for the model under both the null and alternative hypotheses.

The expected sample size is the average sample size

$$\frac{\sum_{k=1}^{K} p_k I_k}{I_0} N_0$$

where p_k is the stopping probability at stage k, $\sum_{k=1}^{K} p_k I_k$ is the expected information level, I_0 is the information level for the fixed-sample design, and N_0 is the sample size for the fixed-sample design.

The expected number of events is the average number of events

$$\frac{\sum_{k=1}^{K} p_k I_k}{I_0} D_0$$

where D_0 is the fixed-sample number of events for the model.

Sample Size Information

The "Sample Sizes (N)" table displays the required sample sizes and information levels at each stage, in both fractional and integer numbers. The derived fractional sample sizes are under the heading "Fractional N." These sample sizes are rounded up to integers under the heading "Ceiling N." The matched integer sample sizes are also displayed for two-sample tests.

The "Required Number of Events (D)" table displays the required number of events required and information level at each stage.

The "Number of Events (D) and Sample Sizes (N)" table displays the number of events and sample size required at each stage with the study time. The derived times under the heading "Fractional Time" are not integers. These times are rounded up to integers under the heading "Ceiling Time."

Stopping Probabilities

The "Expected Cumulative Stopping Probabilities" table displays the following information under each of the specified hypothetical references $\theta = c_i \theta_1$, where c_i are values specified in the CREF= option, and θ_1 is the alternative reference:

- coefficient c_i for the hypothetical references. The value $c_i = 0$ corresponds to the null hypothesis, and $c_i = 1$ corresponds to the alternative hypothesis
- expected stopping stage
- source of the stopping probability: reject H_0 (with STOP=REJECT or STOP=BOTH), accept H_0 (with STOP=ACCEPT or STOP=BOTH), or either reject or accept H_0 (with STOP=BOTH)
- expected cumulative stopping probabilities at each stage

For a one-sided design, the expected cumulative stopping probabilities under the hypothetical references $\theta = c_i \theta_1$ are displayed.

For a two-sided design, the expected cumulative stopping probabilities under each of the hypothetical references $\theta = c_i\theta_{1l}$ and $\theta = c_i\theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively.

Note that for a symmetric two-sided design, only the expected cumulative stopping probabilities under the hypothetical references $\theta = c_i\theta_{1u}$ are derived.

The expected stopping stage is given by $k_0 + d$, where the integer k_0 and the fraction d $(0 \leq d < 1)$ are derived from the expected information level equation

$$\sum_{k=1}^{K} p_k\, I_k = I_{k_0} + d\,(I_{(k_0+1)} - I_{k_0})$$

where p_k is the stopping probability at stage k.

For equally spaced information levels, the expected stopping stage is reduced to the weighted average

$$\sum_{k=1}^{K} p_k\, k$$

ODS Table Names

PROC SEQDESIGN assigns a name to each table it creates. You must use these names to reference tables when using the Output Delivery System (ODS). These names are listed in Table 77.12. For more information about ODS, see Chapter 20, "Using the Output Delivery System."

Table 77.12 ODS Tables Produced by PROC SEQDESIGN

ODS Table Name	Description	Statement	Option
Boundary	Boundary values		
Design	Design information		
ErrSpend	Error spending		ERRSPEND
Method	Method information		
PowerSampleSize	Power and expected sample size		PSS
SampleSize	Derived sample sizes	SAMPLESIZE	
SampleSizeSummary	Sample size summary	SAMPLESIZE	
StopProb	Stopping probabilities		STOPPROB

Graphics Output

This section describes the use of ODS for creating graphics with the SEQDESIGN procedure. To request these graphs, you must specify the ODS GRAPHICS ON statement in addition to the associated graphics options in the PROC SEQDESIGN statement. Except for the PLOTS=BOUNDARY

option, where a detailed boundary plot is generated for each design separately, each option produces a plot for all designs together. For more information about the ODS GRAPHICS statement, see Chapter 21, "Statistical Graphics Using ODS."

Sequential ASN Plot

The PLOTS=ASN option displays the average sample numbers (expected sample sizes for nonsurvival data or expected numbers of events for survival data) under various hypothetical references. The average sample numbers are connected for each design, and these connected curves for all designs are displayed in the "Sequential ASN Plot" graph.

For a one-sided design, average sample numbers under the hypothetical references $\theta = c_i\theta_1$ are displayed, where c_i are the values specified in the CREF= option and θ_1 is the alternative reference. The horizontal axis displays the c_i values of these hypothetical references.

For a two-sided design, average sample numbers under each of the hypothetical references $\theta = c_i\theta_{1l}$ and $\theta = c_i\theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively. The horizontal axis displays $-c_i$ values for lower hypothetical references $\theta = c_i\theta_{1l}$ and c_i values for upper hypothetical references $\theta = c_i\theta_{1u}$.

Note that for a symmetric two-sided design, only average sample numbers under the hypothetical references $\theta = c_i\theta_{1u}$ are derived.

Sequential Boundary Plot

The PLOTS=BOUNDARY option displays boundary values and the acceptance and rejection regions at each stage for each design separately in the "Detailed Boundary Information" graph. The BOUNDARYSCALE= option is used to specify the scale of the boundaries on the vertical axis. The keywords MLE, SCORE, STDZ, and PVALUE in the BOUNDARYSCALE= option correspond to the boundary with the MLE scale, score statistic scale, standardized normal Z scale, and p-value scale, respectively.

The stage numbers are displayed on the horizontal axis. In addition, the HSCALE= option in the PLOTS=BOUNDARY option can be used to specify the scale on the horizontal axis. The keywords INFO and SAMPLESIZE in the HSCALE= option correspond to the information levels and sample sizes, respectively.

Combined Sequential Boundary Plot

The PLOTS=COMBINEDBOUNDARY option displays boundary values. The boundary values are connected for each boundary in each design, and these connected curves for all designs are displayed in the "Sequential Boundary Information" graph. The BOUNDARYSCALE= option is used to specify the scale of the boundaries on the vertical axis. The keywords MLE, SCORE, STDZ, and PVALUE in the BOUNDARYSCALE= option correspond to the boundary with the MLE scale, score statistic scale, standardized normal Z scale, and p-value scale, respectively.

The HSCALE= option in the PLOTS=COMBINEDBOUNDARY option can be used to specify the scale on the horizontal axis. The keywords INFO, SAMPLESIZE, and STAGE in the HSCALE= option correspond to the information levels, sample sizes, and stage numbers, respectively.

Sequential Error Spending Plot

The PLOTS=ERRSPEND option displays the cumulative error spending at each stage on each boundary in the "Sequential Error Spending Plot" graph. A legend table uses the design labels to identify the curves for the corresponding design in the plot. Another legend table uses symbols to identify boundaries in the plot.

Sequential Power Plot

The PLOTS=POWER option displays the powers under various hypothetical references. The powers are connected for each design, and these connected curves for all designs are displayed in the "Sequential Power Plot" graph.

For a one-sided design, powers under hypothetical references $\theta = c_i\theta_1$ are displayed, where c_i are the values specified in the CREF= option and θ_1 is the alternative reference. The horizontal axis displays the c_i values of these hypothetical references.

For a two-sided design, powers under hypothetical references $\theta = c_i\theta_{1l}$ and $\theta = c_i\theta_{1u}$ are displayed, where θ_{1l} and θ_{1u} are the lower and upper alternative references, respectively. The horizontal axis displays $-c_i$ values for lower hypothetical references $\theta = c_i\theta_{1l}$ and c_i values for upper hypothetical references $\theta = c_i\theta_{1u}$.

Note that for a symmetric two-sided design, only powers under hypothetical references $\theta = c_i\theta_{1u}$ are derived.

ODS Graphics

PROC SEQDESIGN assigns a name to each graph it creates. You can use these names to reference the graphs when using ODS. The names are listed in Table 77.13.

To request these graphs, you must specify the ODS GRAPHICS ON statement in addition to the options indicated in Table 77.13. For more information about the ODS GRAPHICS statement, see Chapter 21, "Statistical Graphics Using ODS."

Table 77.13 ODS Graphics Produced by PROC SEQDESIGN

ODS Graph Name	Plot Description	Option
ASNPlot	Average sample numbers	PLOTS=ASN
BoundaryPlot	Detailed boundary values	PLOTS=BOUNDARY
CombinedBoundaryPlot	Boundary values	PLOTS=COMBINEDBOUNDARY
ErrSpendPlot	Error spending	PLOTS=ERRSPEND
PowerPlot	Power curves	PLOTS=POWER

Examples: SEQDESIGN Procedure

The following examples demonstrate the usage of group sequential methods. Example 77.1 uses the NSTAGES=1 option to derive boundaries of critical values for a fixed-sample design. The remaining examples use different methods to create boundaries for various group sequential designs.

Example 77.1: Creating Fixed-Sample Designs

This example demonstrates a one-sided fixed-sample design and a two-sided fixed-sample design. The following statements request a fixed-sample design with an upper alternative:

```
ods graphics on;
proc seqdesign pss
               ;
   OneSidedFixedSample: design nstages=1
                               alt=upper
                               alpha=0.025 beta=0.10
                               ;
   samplesize model=onesamplemean(mean=0.25);
run;
ods graphics off;
```

In the DESIGN statement, the label `OneSidedFixedSample` identifies the design in the output tables. The NSTAGES=1 option specifies that the design has only one stage; this corresponds to a fixed-sample design. In the SEQDESIGN procedure, the null hypothesis for the design is $H_0 : \theta = 0$ and the ALT=UPPER option specifies an upper alternative hypothesis $H_1 : \theta = \theta_1 > 0$. The MEAN=0.25 option in the SAMPLESIZE statement specifies the upper alternative reference $\theta_1 = 0.25$.

The options ALPHA=0.025 and BETA=0.10 specify the Type I error probability level $\alpha = 0.025$ and the Type II error probability level $\beta = 0.10$. That is, the design has a power $1 - \beta = 0.90$ at $\theta_1 = 0.25$.

The "Design Information" table in Output 77.1.1 displays design specifications and the derived statistics such as power. As expected, the derived statistics such as maximum information and average sample number (in percentage of its corresponding fixed-sample information) are 100 for the fixed-sample design (NSTAGES=1). Also, for a fixed-sample design, the STOP= and METHOD= options in the DESIGN statement are not applicable.

Output 77.1.1 One-Sided Fixed-Sample Design Information

The SEQDESIGN Procedure
Design: OneSidedFixedSample

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Alternative Reference	0.25
Number of Stages	1
Alpha	0.025
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	100
Max Information	168.1188
Null Ref ASN (Percent of Fixed Sample)	100
Alt Ref ASN (Percent of Fixed Sample)	100

The "Method Information" table in Output 77.1.2 displays the α and β error levels. It also displays the derived drift parameter, which is the standardized reference improvement, $\theta_1\sqrt{I_0}$, where θ_1 is the alternative reference and I_0 is the maximum information for the design. If either θ_1 or I_0 is specified, the other statistic is derived in the SEQDESIGN procedure. For a fixed-sample design,

$$\theta_1\sqrt{I_0} = \Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta) = \Phi^{-1}(0.975) + \Phi^{-1}(0.90) = 3.2415$$

Output 77.1.2 Method Information

Method Information

Boundary	Alpha	Beta	Alternative Reference	Drift
Upper Alpha	0.02500	0.10000	0.25	3.241516

The "Boundary Information" table in Output 77.1.3 displays information level, alternative reference, and boundary value at each stage. The information proportion indicates the proportion of maximum information available at the stage. With only one stage for a fixed-sample design, the proportion is 1. With the SAMPLESIZE statement, the required sample size N is also displayed under the heading "Information Level."

Output 77.1.3 Boundary Information

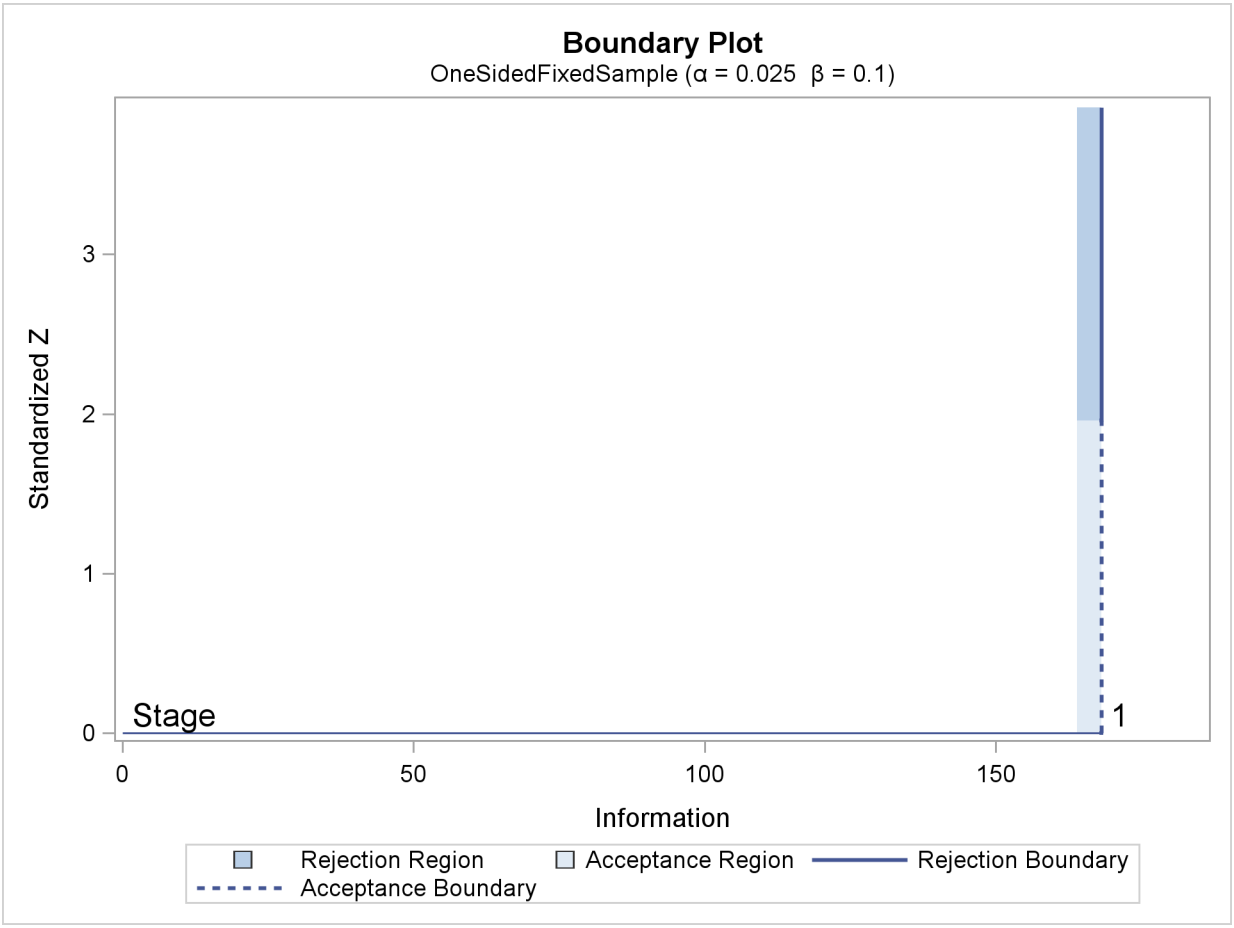
Boundary Information (Standardized Z Scale)
Null Reference = 0

	--------Information Level-------			-Alternative- --Reference--	-Boundary Values- ------Upper------
Stage	Proportion	Actual	N	Upper	Alpha
1	1.0000	168.1188	168.1188	3.24152	1.95996

With the default BOUNDARYSCALE=STDZ option, output alternative references and boundaries are displayed with the standardized normal Z scale. The alternative reference on the standardized Z scale at stage 1 is given by $\theta_1\sqrt{I_1}$, where I_1 is the information level at stage 1. With a boundary value 1.96, the hypothesis of $\theta = 0$ is rejected if the standardized normal statistic $Z \geq 1.96$.

If the ODS GRAPHICS ON statement is specified, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.1.4. The boundary values in the "Boundary Information" table in Figure 77.1.3 are displayed in the plot.

Output 77.1.4 Boundary Plot



The "Sample Size Summary" table in Output 77.1.5 displays parameters for the sample size computation of the test for a normal mean.

Output 77.1.5 Sample Size Summary

Sample Size Summary

Test	One-Sample Mean
Mean	0.25
Standard Deviation	1
Max Sample Size	168.1188
Expected Sample Size (Null Ref)	168.1188
Expected Sample Size (Alt Ref)	168.1188

The "Sample Sizes (N)" table in Output 77.1.6 displays the derived sample sizes, in both fractional and integer numbers. With the resulting integer sample sizes, the corresponding information level is slightly larger than the level from the design. This can increase the power slightly if the integer sample size is used in the trial.

Output 77.1.6 Derived Sample Sizes

Sample Sizes (N)
One-Sample Z Test for Mean

	Fractional N		Ceiling N	
Stage	N	Information	N	Information
1	168.12	168.1	169	169.0

The following statements request a two-sided fixed-sample design with a specified alternative reference:

```
ods graphics on;
proc seqdesign altref=1.2
               pss
               ;
   TwoSidedFixedSample: design nstages=1
                               alt=twosided
                               alpha=0.05 beta=0.10
                               ;
   samplesize model=twosamplemean(stddev=2 weight=2);
run;
ods graphics off;
```

In the SEQDESIGN procedure, the null hypothesis for the design is $H_0 : \theta = 0$. The ALT=TWOSIDED option specifies a two-sided alternative hypothesis $H_1 : \theta = \theta_1 \neq 0$. The ALTREF=1.2 option in the PROC SEQDESIGN statement specifies the alternative reference $\theta_1 = \pm 1.2$.

The ALPHA=0.05 option specifies the two-sided Type I error probability level $\alpha = 0.05$. That is, the lower and upper Type I error probabilities $\alpha_l = \alpha_u = 0.025$. The BETA=0.10 option specifies the Type II error probability level $\beta = 0.10$, and the design has a power $1 - \beta = 0.90$ at the alternative reference $\theta_1 = \pm 1.2$.

The "Design Information" table in Output 77.1.7 displays design specifications and the derived power. With a specified alternative reference, the maximum information is derived.

Output 77.1.7 Two-Sided Fixed-Sample Design Information

The SEQDESIGN Procedure
Design: TwoSidedFixedSample

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Alternative Reference	1.2
Number of Stages	1
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	100
Max Information	7.296822
Null Ref ASN (Percent of Fixed Sample)	100
Alt Ref ASN (Percent of Fixed Sample)	100

The "Method Information" table in Output 77.1.8 displays the α and β errors, alternative references, and drift parameter. For a fixed-sample design, the derived drift parameter

$$\theta_1 \sqrt{I_0} = \Phi^{-1}(1 - \frac{\alpha}{2}) + \Phi^{-1}(1 - \beta) = \Phi^{-1}(0.975) + \Phi^{-1}(0.90) = 3.2415$$

Output 77.1.8 Method Information

Method Information

Boundary	Alpha	Beta	Alternative Reference	Drift
Upper Alpha	0.02500	0.10000	1.2	3.241516
Lower Alpha	0.02500	0.10000	-1.2	-3.24152

With a specified alternative reference $\theta_1 = 1.2$, the maximum information

$$I_0 = \left(\frac{3.2415}{1.2}\right)^2 = 7.2968$$

The default "Boundary Information" table in Output 77.1.9 displays information level, alternative reference, and boundary values. With the default BOUNDARYSCALE=STDZ option, alternative reference and boundary values are displayed with the standardized normal Z scale. Thus, the standardized alternative references $\pm\theta_1\sqrt{I_0}$ are displayed.

Output 77.1.9 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Information Level			Alternative Reference	
Stage	Proportion	Actual	N	Lower	Upper
1	1.0000	7.296822	131.3428	-3.24152	3.24152

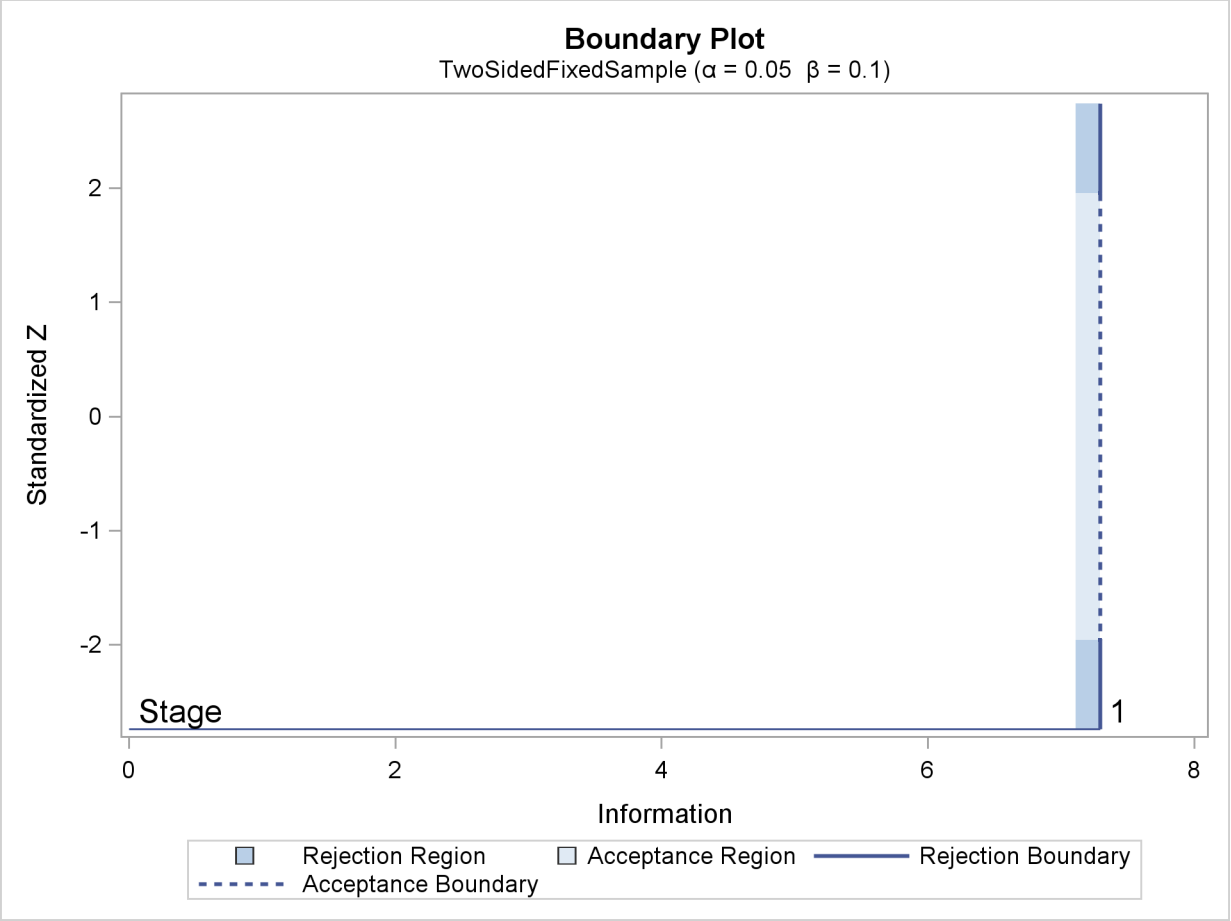
Boundary Information (Standardized Z Scale)
Null Reference = 0

	Boundary Values	
	Lower	Upper
Stage	Alpha	Alpha
1	-1.95996	1.95996

With boundary values of -1.96 and 1.96, the hypothesis of $\theta = 0$ is rejected if the standardized normal statistic $Z \geq 1.96$ or $Z \leq -1.96$.

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.1.10 . The boundary values in the "Boundary Information" table in Figure 77.1.9 are displayed in the plot.

Output 77.1.10 Boundary Plot



The "Sample Size Summary" table in Output 77.1.11 displays parameters for the sample size computation of the test for a normal mean.

Output 77.1.11 Sample Size Summary

Sample Size Summary

Test	Two-Sample Means
Mean Difference	1.2
Standard Deviation	2
Max Sample Size	131.3428
Expected Sample Size (Null Ref)	131.3428
Expected Sample Size (Alt Ref)	131.3428
Weight (Group A)	2
Weight (Group B)	1

The "Sample Sizes (N)" table in Output 77.1.12 displays the derived sample sizes, in both fractional and integer numbers. With the WEIGHT=2 option, the allocation ratio is 2 for the first group and 1 for the second group. With the resulting integer sample sizes, the corresponding information level is slightly larger than the level from the design. This can increase the power slightly if the integer sample size is used in the trial.

Output 77.1.12 Derived Sample Sizes

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Fractional N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	131.34	87.56	43.78	7.2968

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Ceiling N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	132	88	44	7.3333

Example 77.2: Creating a One-Sided O'Brien-Fleming Design

This example demonstrates a group sequential design for a clinical study. A clinic is conducting a study on the effect of vitamin C supplements in treating flu symptoms. The study groups consist of patients in the clinic with their first sign of flu symptoms within the last 24 hours. These individuals are randomly assigned to either the control group, which receives the placebo pills, or the treatment group, which receives large doses of vitamin C supplements. At the end of a five-day period, the flu symptoms of each individual are recorded.

Suppose that from past experience, 60% of individuals experiencing flu symptoms have the symptoms disappeared within five days. The clinic wants to detect a 75% symptoms disappearance with a high probability in the trial. A test that compares the proportions directly is to specify a null hypothesis $H_0 : \theta = p_a - p_b = 0$ with a Type I error probability level $\alpha = 0.025$, where p_a and p_b are the proportions of symptoms' disappearance in the treatment group and control group, respectively. A one-sided alternative $H_1 : \theta > 0$ is also specified with a power of $1 - \beta = 0.90$ at $H_1 : \theta = 0.15$.

For a one-sided fixed-sample design, the critical value for the standardized Z test statistic is given by $C_\alpha = \Phi^{-1}(1-\alpha) = 1.96$. That is, at the end of study, if the test statistic $z \geq C_\alpha$, then the null hypothesis is rejected and the efficacy of vitamin C supplements is declared. Otherwise, the null hypothesis is not rejected and the effect of vitamin C supplements is not significant.

To achieve a $1 - \beta = 0.90$ power at $H_1 : \theta = 0.15$ for a fixed-sample design, the information required is given by

$$I_0 = \frac{(\Phi^{-1}(1-\alpha) + \Phi^{-1}(1-\beta))^2}{0.15^2} = \frac{(1.96 + 1.28155)^2}{0.0225} = 466.99$$

With an equal sample size on the treatment and control groups, $N_a = N_b$, the sample size required for each group under H_1 is computed from the information I_0:

$$N_a = N_b = \left(p_{1a}\left(1 - p_{1a}\right) + p_{1b}\left(1 - p_{1b}\right)\right) I_0$$

where $p_{1a} = 0.75$ and $p_{1b} = 0.60$ are proportions in the treatment and control groups under H_1. That is,

$$N_a = N_b = (0.75 \times 0.25 + 0.6 \times 0.4) \times 466.99 = 199.64$$

Thus, 200 individuals are required for each group in the fixed-sample study. See the section "Test for the Difference between Two Binomial Proportions" on page 5862 for a detailed derivation of these required sample sizes.

Instead of a fixed-sample design for the trial, a group sequential design is used to stop the trial early for ethical concerns of possible harm or an unexpected strong efficacy outcome of the new drug. It can also save time and resources in the process. The following statements invoke the SEQDESIGN procedure and request a four-stage group sequential design that uses an O'Brien-Fleming method for normally distributed statistics. The design uses a one-sided alternative hypothesis H_1 with early stopping to reject or accept H_0.

```
ods graphics on;
proc seqdesign altref=0.15
               ;
   OneSidedOBrienFleming: design nstages=4
                          method=obf
                          alt=upper   stop=both
                          alpha=0.025 beta=0.10
                          ;
   samplesize model=twosamplefreq(nullprop=0.6 test=prop);
ods output Boundary=Bnd_Prop;
run;
ods graphics off;
```

In a sequential design, a hypothesis can be rejected, accepted, or continued to the next time point at each interim stage. The STOP=BOTH option specifies early stopping to reject or accept the null hypothesis. The "Design Information," "Method Information," and "Boundary Information" tables are displayed by default.

The "Design Information" table in Output 77.2.1 displays design specifications and derived statistics such as power and maximum information. With a specified alternative reference, ALTREF=0.15, the maximum information I_X is derived.

Output 77.2.1 Design Information

The SEQDESIGN Procedure
Design: OneSidedOBrienFleming

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Method	O'Brien-Fleming
Boundary Key	Both
Alternative Reference	0.15
Number of Stages	4
Alpha	0.025
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	107.6741
Max Information	502.8343
Null Ref ASN (Percent of Fixed Sample)	61.12891
Alt Ref ASN (Percent of Fixed Sample)	75.89782

The Max Information (Percent Fixed-Sample) is the ratio in percentage between the maximum information for the group sequential design and the information required for a corresponding fixed-sample design:

$$100 \times \frac{I_X}{I_0} = 100 \times \frac{502.83}{466.99} = 107.67$$

That is, if the group sequential trial does not stop at any interim stages, the information needed is 7.67% more than is needed for the corresponding fixed-sample design. For a two-sample test for binomial proportions, the information is proportional to the sample size. Thus, 7.67% more observations are needed for the group sequential trial.

The Null Ref ASN (Percent Fixed-Sample) is the ratio in percentage between the expected sample size required under the null hypothesis for the group sequential design and the sample size required for the corresponding fixed-sample design. With a ratio of 61.1%, the expected sample size for the group sequential trial under the null hypothesis is 61.1% of the sample size in the corresponding fixed-sample design.

Similarly, the Alt Ref ASN (Percent Fixed-Sample) is the ratio in percentage between the expected sample size required under the alternative hypothesis for the group sequential design and the sample size required for the corresponding fixed-sample design. With a ratio of 75.9%, the expected sample size for the group sequential trial under the alternative hypothesis is 75.9% of the sample size in the corresponding fixed-sample design.

For a one-sided design with an upper alternative and early stopping to reject or accept the null hypothesis, upper α and β boundaries are created. The "Method Information" table in Output 77.2.2 displays the Type I error probability α, the Type II error probability β, and the derived drift parameter. The drift parameter is the standardized reference improvement between the alternative and null hypotheses at the final stage. It is also the standardized alternative reference at the final stage if the null reference is zero.

Output 77.2.2 Method Information

Method Information

Boundary	Method	Alpha	Beta	Unified Family: Rho	Unified Family: Tau	Unified Family: C
Upper Alpha	O'Brien-Fleming	0.02500	.	0.5	0	1.9784
Upper Beta	O'Brien-Fleming	.	0.10000	0.5	0	1.3852

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.15	3.363595
Upper Beta	0.15	3.363595

With the METHOD=OBF option, the O'Brien-Fleming method is used for each boundary. The O'Brien-Fleming method is one of the unified family methods, and the "Method Information" table displays the corresponding parameter ρ in the unified family method. The table also displays the critical values $C_\alpha = 1.9784$ for the α boundary and $C_\beta = 1.3852$ for the β boundary. These critical values are used to create the boundary values. With the default BOUNDARYSCALE=STDZ option, the boundaries are displayed with the standardized Z statistic scale.

The "Boundary Information" table in Output 77.2.3 displays information level, alternative reference, and boundary values at each stage. The default BOUNDARYSCALE=STDZ option specifies that the standardized Z scale be used to display the alternative references and boundary values. The resulting standardized alternative reference at stage k is given by $\theta_1 \sqrt{I_k}$, where θ_1 is the alternative reference and I_k is the information level at stage k, $k = 1, 2, 3, 4$.

Output 77.2.3 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Information Level: Proportion	Information Level: Actual	Information Level: N	Alternative Reference: Upper	Boundary Values: Upper: Beta	Boundary Values: Upper: Alpha
1	0.2500	125.7086	107.4808	1.68180	-1.08860	3.95679
2	0.5000	251.4171	214.9617	2.37842	0.41946	2.79788
3	0.7500	377.1257	322.4425	2.91296	1.31347	2.28446
4	1.0000	502.8343	429.9233	3.36360	1.97840	1.97840

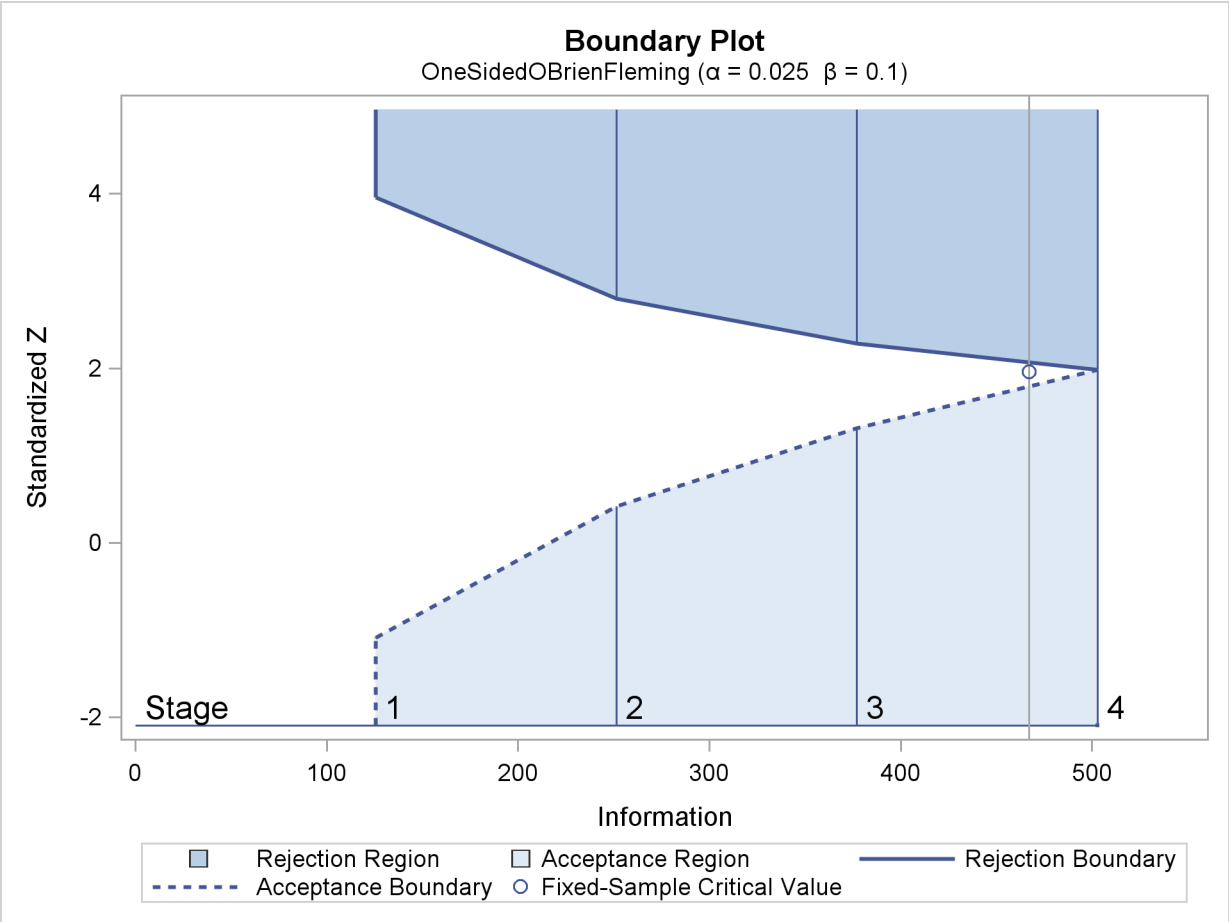
By default, equally spaced information levels are used. An information proportion is the proportion of maximum information available at each stage. With the derived maximum information, the actual information level at each stage is also displayed. With the SAMPLESIZE statement, the required sample size N is also displayed under the heading "Information Level."

At each interim stage, if the standardized Z test statistic is larger than or equal to the corresponding upper α boundary, then the hypothesis $H_0 : \theta = 0$ is rejected. If the test statistic is less than the corresponding upper β boundary, then the trial is stopped and the hypothesis H_0 is accepted. Otherwise, the process continues to the next stage. At the final stage, stage 4, the trial stops and the hypothesis H_0 is rejected if the standardized Z statistic $Z_4 \geq 1.9784$. Otherwise, the trial is accepted.

The ODS OUTPUT statement with the BOUNDARY=BND_PROP option creates an output data set that contains the resulting boundary information. After the actual data from the clinical trial are collected and analyzed at each stage with a procedure such as PROC GENMOD, the SEQTEST procedure is used to test the resulting statistics at stage 1 with the boundary information stored in the BOUND_PROP data set.

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.2.4.

Output 77.2.4 Boundary Plot



The horizontal axis indicates the information levels for the design. The stages are indicated by vertical lines with accompanying stage numbers. If at any stage a test statistic is in a rejection region, the trial stops and the hypothesis is rejected. If a test statistic is in an acceptance region, then the trial also stops and the hypothesis is accepted. If the statistic is not in a rejection region

or an acceptance region, the trial continues to the next stage. The boundary plot also displays the information level and critical value for the corresponding fixed-sample design.

The SEQDESIGN procedure derives the drift parameter $\theta_1\sqrt{I_X}$, where θ_1 is the alternative reference and I_X is the maximum information. If either θ_1 or I_X is specified, the other can be derived. With the SAMPLESIZE statement, the maximum information is used to compute the required sample size for the study.

The "Sample Size Summary" table in Output 77.2.5 displays parameters for the sample size computation. With the MODEL=TWOSAMPLEFREQ(NULLPROP=0.6 TEST=PROP) option in the SAMPLESIZE statement, the total sample size in each group for testing the difference between two proportions is computed. With the default REF=PROP option, the required sample sizes are computed under the alternative hypothesis. That is,

$$N_a = N_b = \left(p_{1a}\left(1 - p_{1a}\right) + p_{1b}\left(1 - p_{1b}\right)\right) I_X$$

where $p_{1b} = 0.60$ and $p_{1a} = p_{1b} + \theta_1 = 0.75$ are the proportions in the control and treatment groups, respectively, under the alternative hypothesis H_1. See the section "Test for the Difference between Two Binomial Proportions" on page 5862 for a detailed description of these parameters.

Output 77.2.5 Sample Size Summary

Sample Size Summary

Test	Two-Sample Proportions
Null Proportion	0.6
Proportion (Group A)	0.75
Test Statistic	Z for Proportion
Reference Proportions	Alt Ref
Max Sample Size	429.9233
Expected Sample Size (Null Ref)	244.0768
Expected Sample Size (Alt Ref)	303.0464

The "Sample Sizes (N)" table in Output 77.2.6 displays the required sample sizes at each stage, in both fractional and integer numbers. The derived fractional sample sizes are under the heading "Fractional N." These sample sizes are rounded up to integers under the heading "Ceiling N." In practice, integer sample sizes are used, and the resulting information levels increase slightly. Thus, 54, 108, 162, and 215 individuals are needed in each group for the four stages, respectively.

Output 77.2.6 Derived Sample Sizes

Sample Sizes (N)
Two-Sample Z Test for Proportion Difference

	Fractional N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	107.48	53.74	53.74	125.7
2	214.96	107.48	107.48	251.4
3	322.44	161.22	161.22	377.1
4	429.92	214.96	214.96	502.8

Sample Sizes (N)
Two-Sample Z Test for Proportion Difference

	Ceiling N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	108	54	54	126.3
2	216	108	108	252.6
3	324	162	162	378.9
4	430	215	215	502.9

Example 77.3: Creating Two-Sided Pocock and O'Brien-Fleming Designs

This example requests two 4-stage group sequential designs for normally distributed statistics with equally spaced information levels at all stages. One design uses Pocock's method and the other uses the O'Brien-Fleming method. With the default BOUNDARYSCALE=STDZ option, the output boundaries are displayed with the standardized normal Z scale.

In the following statements, the default ALT=TWOSIDED option specifies a null hypothesis with a two-sided alternative, and the default STOP=REJECT option indicates an early stop to reject the null hypothesis H_0:

```
proc seqdesign altref=0.4
               pss
               stopprob
               errspend
               ;
   TwoSidedPocock:         design nstages=4 method=poc;
   TwoSidedOBrienFleming: design nstages=4 method=obf;
   samplesize model=twosamplemean(stddev=0.8 weight=2);
run;
```

The "Design Information" table in Output 77.3.1 displays design specifications and derived statistics for the Pocock's design. With the specified ALTREF= option, the maximum information $I_X = 77.6984$ is also derived.

Output 77.3.1 Pocock Design Information

The SEQDESIGN Procedure
Design: TwoSidedPocock

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Method	Pocock
Boundary Key	Both
Alternative Reference	0.4
Number of Stages	4
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	118.3143
Max Information	77.69844
Null Ref ASN (Percent of Fixed Sample)	115.6074
Alt Ref ASN (Percent of Fixed Sample)	69.74805

With the corresponding fixed-sample information

$$I_0 = \frac{(\Phi^{-1}(1-\alpha/2) + \Phi^{-1}(1-\beta))^2}{0.4^2} = \frac{(1.96 + 1.28155)^2}{0.16} = 65.6728$$

the fixed-sample information ratio is $77.6984/65.6728 = 1.1831$.

For a two-sided design with early stopping to reject the null hypothesis, lower and upper α boundaries are created. The "Method Information" table in Output 77.3.2 displays the α and β errors, alternative references, and derived drift parameters, which are the standardized alternative references at the final stage.

Output 77.3.2 Method Information

Method Information

Boundary	Method	Alpha	Beta	-------Unified Family------ Rho	Tau	C
Upper Alpha	Pocock	0.02500	0.10000	0	0	2.36129
Lower Alpha	Pocock	0.02500	0.10000	0	0	2.36129

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.4	3.525869
Lower Alpha	-0.4	-3.52587

With the METHOD=POC option, the Pocock method is used for each boundary. The Pocock method is one of the unified family methods, and the table also displays its corresponding parameters $\rho = 0$ as a unified family method and the derived parameters $C = 2.3613$ for the boundary values.

With the PSS option, the "Power and Expected Sample Sizes" table in Output 77.3.3 displays powers and expected sample sizes under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and c_i are values specified in the CREF= option. By default, $c_i = 0, 0.5, 1.0, 1.5$.

Output 77.3.3 Power and Expected Sample Size Information

Powers and Expected Sample Sizes
Reference = CRef * (Alt Reference)

CRef	Power	-Sample Size- Percent Fixed-Sample
0.0000	0.02500	115.6074
0.5000	0.34252	104.0615
1.0000	0.90000	69.7480
1.5000	0.99869	43.6600

Note that at $c_i = 0$, the null reference $\theta = 0$, and the power 0.025 corresponds to the one-sided Type I error probability 0.025. At $c_i = 1$, $\theta = \theta_1$, the power 0.9 is the power of the design. The expected sample sizes are displayed in a percentage scale to its corresponding fixed-sample size design. With the specified SAMPLESIZE statement, the expected sample sizes for the specified model in the SAMPLESIZE statement are also displayed.

With the STOPPROB option, the "Expected Cumulative Stopping Probabilities" table in Output 77.3.4 displays the expected cumulative stopping stage and cumulative stopping probability to reject the null hypothesis H_0 at each stage under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and c_i are values specified in the CREF= option. By default, $c_i = 0, 0.5, 1.0, 1.5$.

Output 77.3.4 Stopping Probabilities

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

CRef	Expected Stopping Stage	Source	--------Stopping Probabilities-------- Stage_1	Stage_2	Stage_3	Stage_4
0.0000	3.908	Reject Null	0.01821	0.03155	0.04176	0.05000
0.5000	3.518	Reject Null	0.07005	0.15939	0.25242	0.34327
1.0000	2.358	Reject Null	0.27482	0.58074	0.78638	0.90002
1.5000	1.476	Reject Null	0.61145	0.92348	0.98900	0.99869

Note that at $c_i = 0$, the cumulative stopping probability to reject H_0 at the final stage is the overall Type I error probability 0.05. At $c_i = 1$, the alternative hypothesis $H_1 : \theta = \theta_1$, the cumulative

stopping probability to reject H_0 includes both the probability in the lower rejection region and the probability in the upper rejection region. This stopping probability to reject H_0 at the final stage, 0.90002, is slightly greater than the power $1 - \beta = 0.90$, which corresponds to the cumulative stopping probability in the upper rejection region only. See the section "Type I and Type II Errors" on page 5836 for a detailed description of the Type II error probability β.

The "Boundary Information" table in Output 77.3.5 displays the information level, alternative references, and boundary values at each stage. The default BOUNDARYSCALE=STDZ option specifies that the standardized Z scale be used to display the alternative references and boundary values. The resulting standardized alternative reference at stage k is given by $\theta_1 \sqrt{I_k}$, where θ_1 is the alternative reference and I_k is the information level at stage k, $k = 1, 2, 3, 4$.

Output 77.3.5 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Information Level			Alternative Reference	
Stage	Proportion	Actual	N	Lower	Upper
1	0.2500	19.42461	55.94288	-1.76293	1.76293
2	0.5000	38.84922	111.8858	-2.49317	2.49317
3	0.7500	58.27383	167.8286	-3.05349	3.05349
4	1.0000	77.69844	223.7715	-3.52587	3.52587

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Boundary Values	
	Lower	Upper
Stage	Alpha	Alpha
1	-2.36129	2.36129
2	-2.36129	2.36129
3	-2.36129	2.36129
4	-2.36129	2.36129

By default, equally spaced information levels are used, and the procedure displays the output boundaries in terms of standardized statistics (BOUNDARYSCALE=STDZ). With the SAMPLESIZE statement, the required sample size N is also displayed under the heading "Information Level." With the Pocock method, the standardized Z boundary values are identical at all stages for each α boundary.

At each interim stage, the hypothesis of $H_0 : \theta = 0$ is rejected if the standardized normal test statistic $z \leq -2.36129$, the lower α boundary, or $z \geq 2.36129$, the upper α boundary. Otherwise, the trial continues to the next stage. At the final stage, stage 4, the trial stops and the hypothesis is rejected if the test statistic $|z_4| \geq 2.36129$. Otherwise, the hypothesis is accepted.

The "Error Spending Information" in Output 77.3.6 displays cumulative error spending at each stage for each boundary. It shows that more α errors are used in early stages than in later stages.

Output 77.3.6 Error Spending Information

Error Spending Information

Stage	-Information Level- Proportion	--------Cumulative Error Spending------- -------Lower------ Alpha	Beta	-------Upper------ Beta	Alpha
1	0.2500	0.00911	0.00002	0.00002	0.00911
2	0.5000	0.01577	0.00002	0.00002	0.01577
3	0.7500	0.02088	0.00002	0.00002	0.02088
4	1.0000	0.02500	0.10000	0.10000	0.02500

The "Sample Size Summary" table in Output 77.3.7 displays the specified parameters for the sample size computation of the two-sample test for mean difference.

Output 77.3.7 Sample Size Summary

Sample Size Summary

Test	Two-Sample Means
Mean Difference	0.4
Standard Deviation	0.8
Max Sample Size	223.7715
Expected Sample Size (Null Ref)	218.652
Expected Sample Size (Alt Ref)	131.9167
Weight (Group A)	2
Weight (Group B)	1

The "Sample Sizes (N)" table in Output 77.3.8 displays the derived sample sizes at each stage, in both fractional and integer numbers. With the WEIGHT=2 option, the allocation ratio is 2 for the first group and 1 for the second group. See the section "Test for the Difference between Two Normal Means" on page 5860 for the derivation of these sample sizes. With the fixed-sample information ratio 1.1831, the derived sample sizes in fractional numbers are derived by multiplying 1.1831 by the corresponding sample sizes in the fixed-sample design.

Output 77.3.8 Sample Sizes

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Fractional N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	55.94	37.30	18.65	19.4246
2	111.89	74.59	37.30	38.8492
3	167.83	111.89	55.94	58.2738
4	223.77	149.18	74.59	77.6984

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Ceiling N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	57	38	19	19.7917
2	113	75	38	39.4082
3	168	112	56	58.3333
4	225	150	75	78.1250

These fractional sample sizes are rounded up to integers under the heading "Ceiling N." When the resulting integer sample sizes are used, the corresponding information levels are slightly larger than the levels from the design. This can increase the power slightly if a trial uses these integer sample sizes.

Note that compared with other designs, a Pocock design can stop the trial early with a larger p-value. However, this might not be persuasive enough to make a new treatment widely accepted (Pocock and White, 1999).

The "Design Information" table in Output 77.3.9 displays design specifications and the derived statistics for the O'Brien-Fleming design. With the specified ALTREF= option, the maximum information $I_X = 67.1268$ is derived.

Output 77.3.9 O'Brien-Fleming Design Information

The SEQDESIGN Procedure
Design: TwoSidedOBrienFleming

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Method	O'Brien-Fleming
Boundary Key	Both
Alternative Reference	0.4
Number of Stages	4
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	102.2163
Max Information	67.12682
Null Ref ASN (Percent of Fixed Sample)	101.5728
Alt Ref ASN (Percent of Fixed Sample)	76.7397

With the corresponding fixed-sample information

$$I_0 = \frac{(\Phi^{-1}(1-\alpha/2) + \Phi^{-1}(1-\beta))^2}{0.4^2} = \frac{(1.96 + 1.28155)^2}{0.16} = 65.6728$$

the fixed-sample information ratio is $67.1268/65.6728 = 1.022$. That is, the maximum information for the O'Brien-Fleming design is only 2.2% more than for the corresponding fixed-sample design.

The "Method Information" table in Output 77.3.10 displays the Type I α level and Type II β level. It also displays the derived drift parameter $\theta_1\sqrt{I_X}$, which is the standardized alternative reference at the final stage.

Output 77.3.10 Method Information

Method Information

				Unified Family		
Boundary	Method	Alpha	Beta	Rho	Tau	C
Upper Alpha	O'Brien-Fleming	0.02500	0.10000	0.5	0	2.02429
Lower Alpha	O'Brien-Fleming	0.02500	0.10000	0.5	0	2.02429

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.4	3.277238
Lower Alpha	-0.4	-3.27724

With the METHOD=OBF option, the O'Brien-Fleming method is used for each boundary. The O'Brien-Fleming method is one of the unified family methods, and the table also displays its corresponding parameters $\rho = 0.5$ as a unified family method and the derived parameter $C = 2.0243$ for the boundary values.

With the PSS option, the "Power and Expected Sample Sizes" table in Output 77.3.11 displays powers and expected sample sizes under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and c_i are values specified in the CREF= option.

Output 77.3.11 Power and Expected Sample Size Information

Powers and Expected Sample Sizes
Reference = CRef * (Alt Reference)

CRef	Power	-Sample Size- Percent Fixed-Sample
0.0000	0.02500	101.5728
0.5000	0.36495	96.3684
1.0000	0.90000	76.7397
1.5000	0.99821	57.2590

Compared with the corresponding Pocock design, the O'Brien-Fleming design has a smaller maximum sample size, and smaller expected sample sizes under hypothetical references $\theta = 0$ and $\theta = 0.5\,\theta_1$, but larger expected sample sizes under hypothetical references $\theta = \theta_1$ and $\theta = 1.5\,\theta_1$.

With the STOPPROB option, the "Expected Cumulative Stopping Probabilities" table in Output 77.3.12 displays the expected stopping stage and cumulative stopping probability to reject the null hypothesis at each stage under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and c_i are values specified in the CREF= option.

Output 77.3.12 Stopping Probabilities

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

CRef	Expected Stopping Stage	Source	--------Stopping Probabilities------- Stage_1	Stage_2	Stage_3	Stage_4
0.0000	3.975	Reject Null	0.00005	0.00422	0.02091	0.05000
0.5000	3.771	Reject Null	0.00062	0.04430	0.18392	0.36515
1.0000	3.003	Reject Null	0.00798	0.29296	0.69603	0.90000
1.5000	2.241	Reject Null	0.05584	0.73031	0.97315	0.99821

Compared with the corresponding Pocock design, the O'Brien-Fleming design has smaller stopping probabilities in early stages under each hypothetical reference.

The "Boundary Information" table in Output 77.3.13 displays the boundary values for the design that uses the O'Brien-Fleming method. Compared with the Pocock method, the standardized statistics α boundary values derived from the O'Brien-Fleming method in absolute values are larger in

early stages and smaller in later stages. This makes the O'Brien-Fleming design less likely to reject the null hypothesis in early stages than the Pocock design. With the derived parameter $C = 2.0243$ for the α boundary, the α boundaries at stage j are computed as $C\sqrt{4/j}$, $j = 1, \ldots, 4$.

Output 77.3.13 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Information Level			Alternative Reference	
Stage	Proportion	Actual	N	Lower	Upper
1	0.2500	16.7817	48.33131	-1.63862	1.63862
2	0.5000	33.56341	96.66262	-2.31736	2.31736
3	0.7500	50.34511	144.9939	-2.83817	2.83817
4	1.0000	67.12682	193.3252	-3.27724	3.27724

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Boundary Values	
	Lower	Upper
Stage	Alpha	Alpha
1	-4.04859	4.04859
2	-2.86278	2.86278
3	-2.33745	2.33745
4	-2.02429	2.02429

The "Error Spending Information" in Output 77.3.14 displays cumulative error spending at each stage for each boundary. With smaller α spending in early stages for the O'Brien-Fleming method, it also indicates that the O'Brien-Fleming design is less likely to reject the null hypothesis in early stages than the Pocock design.

Output 77.3.14 Error Spending Information

Error Spending Information

	Information Level	Cumulative Error Spending			
		Lower		Upper	
Stage	Proportion	Alpha	Beta	Beta	Alpha
1	0.2500	0.00003	0.00000	0.00000	0.00003
2	0.5000	0.00211	0.00000	0.00000	0.00211
3	0.7500	0.01046	0.00000	0.00000	0.01046
4	1.0000	0.02500	0.10000	0.10000	0.02500

The "Sample Size Summary" table in Output 77.3.15 displays the specified parameters for the sample size computation of the two-sample test for mean difference.

Output 77.3.15 Sample Size Summary

Sample Size Summary

Test	Two-Sample Means
Mean Difference	0.4
Standard Deviation	0.8
Max Sample Size	193.3252
Expected Sample Size (Null Ref)	192.1081
Expected Sample Size (Alt Ref)	145.1404
Weight (Group A)	2
Weight (Group B)	1

The "Sample Sizes (N)" table in Output 77.3.16 displays the derived sample sizes at each stage, in both fractional and integer numbers. With the fixed-sample information ratio 1.0222, the required sample sizes in fractional numbers are derived by multiplying 1.0222 by the corresponding sample sizes in the fixed-sample design.

Output 77.3.16 Derived Sample Sizes

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Fractional N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	48.33	32.22	16.11	16.7817
2	96.66	64.44	32.22	33.5634
3	144.99	96.66	48.33	50.3451
4	193.33	128.88	64.44	67.1268

Sample Sizes (N)
Two-Sample Z Test for Mean Difference

	Ceiling N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	50	33	17	17.5313
2	98	65	33	34.1996
3	146	97	49	50.8669
4	194	129	65	67.5338

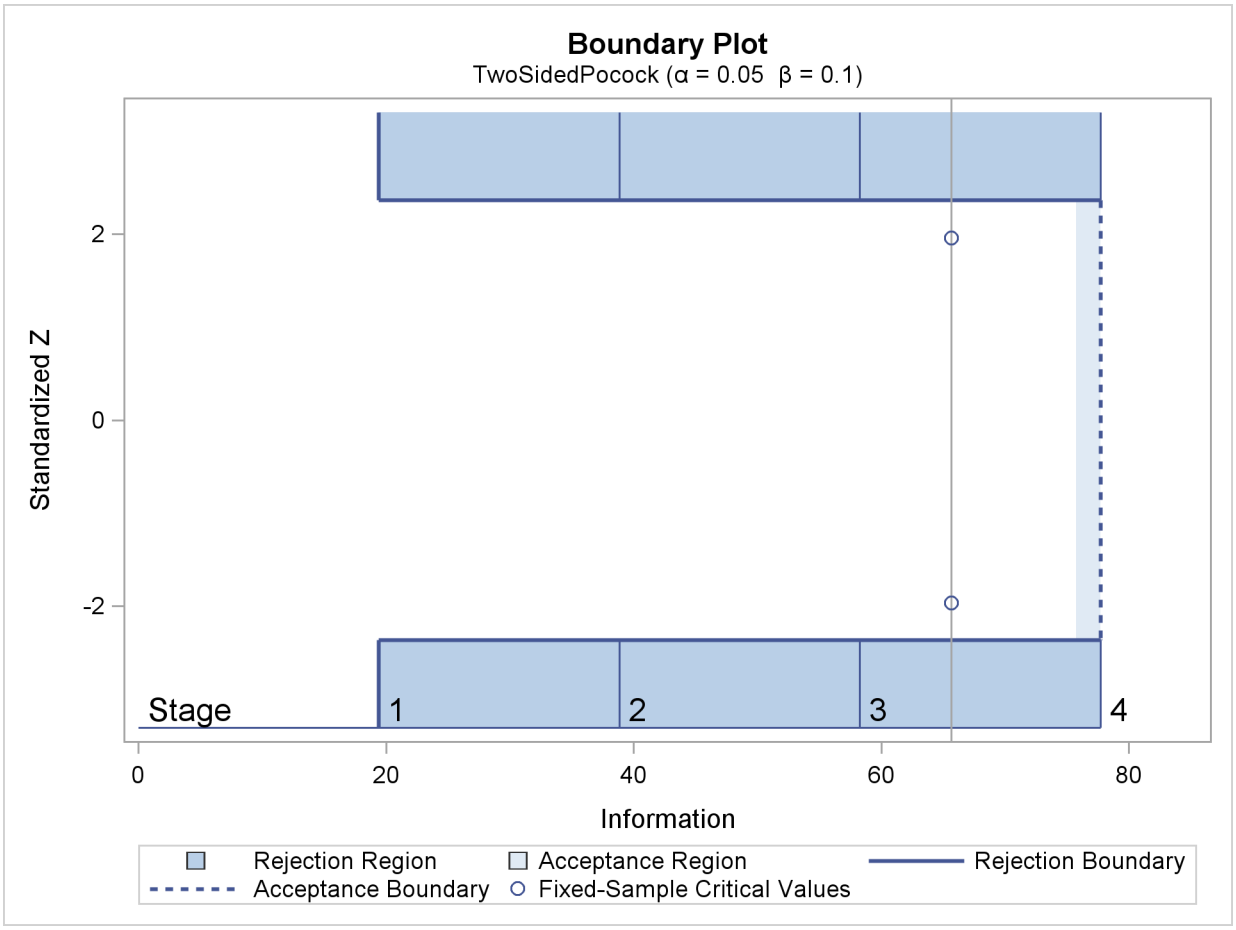
Example 77.4: Generating Graphics Display for Sequential Designs

This example creates the same group sequential design as in Example 77.3 and creates graphics by using ODS Graphics. The following statements request all available graphs in the SEQDESIGN procedure:

```
ods graphics on;
proc seqdesign altref=0.4
               plots=all
               ;
   TwoSidedPocock:         design nstages=4 method=poc;
   TwoSidedOBrienFleming: design nstages=4 method=obf;
run;
ods graphics off;
```

With the PLOTS=ALL option, a detailed boundary plot with the rejection region and acceptance region is displayed for the Pocock design, as shown in Output 77.4.1. With the default STOP=REJECT option, the rejection boundaries are also generated at interim stages.

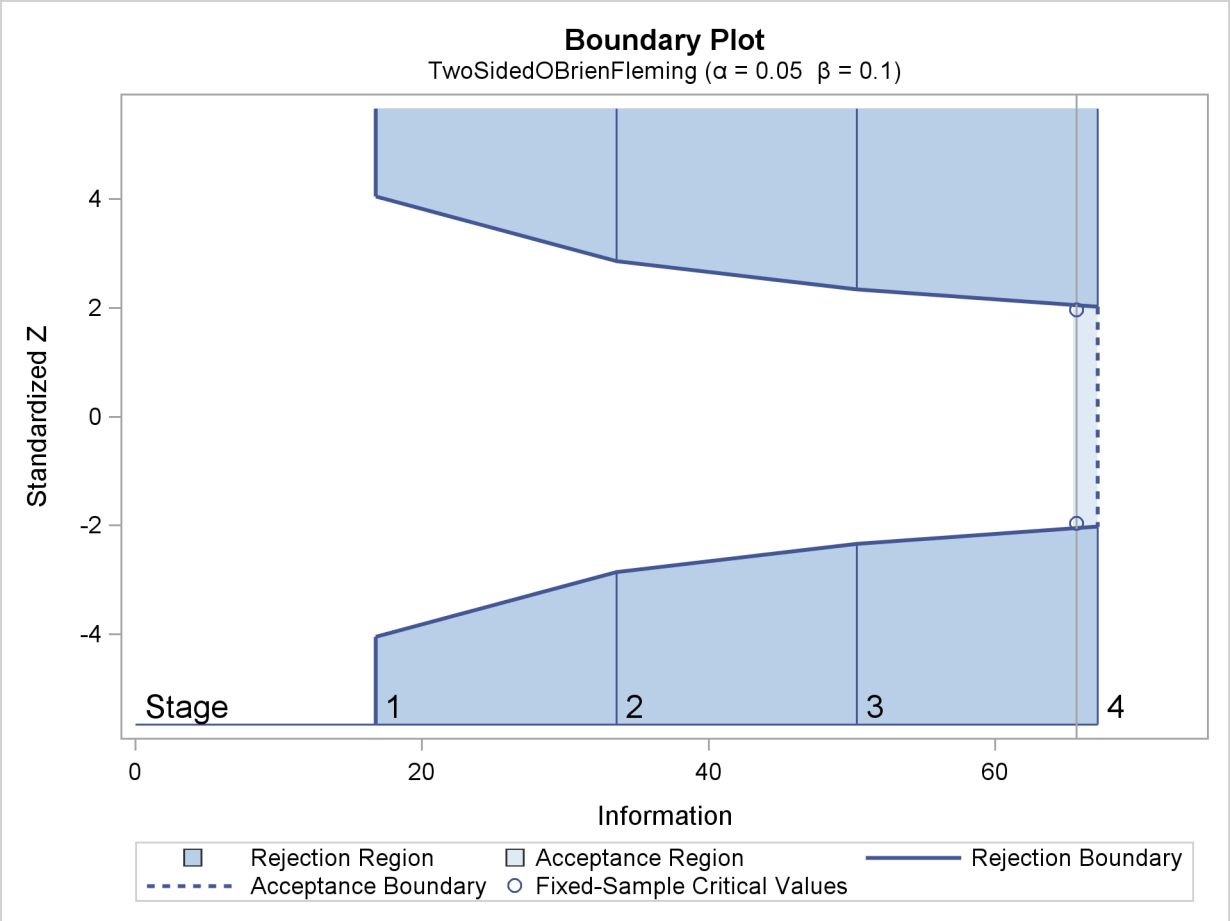
Output 77.4.1 Pocock Boundary Plot



The plot shows identical boundary values in each boundary in the standardized Z scale for the Pocock design. The information level and critical value for the corresponding fixed-sample design are also displayed.

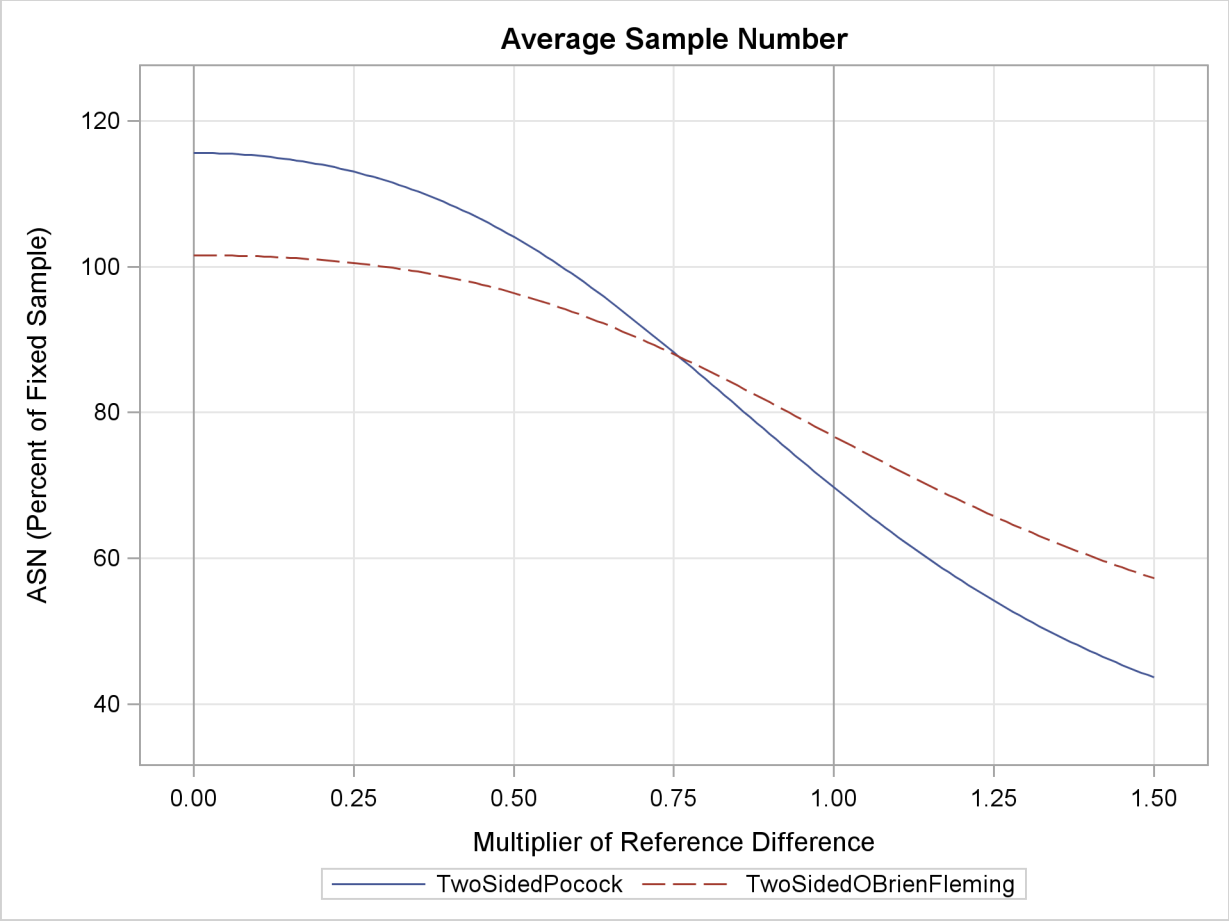
With the PLOTS=ALL option, a detailed boundary plot with the rejection region and acceptance region is also displayed for the O'Brien-Fleming design, as shown in Output 77.4.2. The plot shows that the rejection boundary values are decreasing as the trial advances in the standardized Z scale.

Output 77.4.2 O'Brien-Fleming Boundary Plot



With the PLOTS=ALL option, the procedure displays a plot of average sample numbers (expected sample sizes for nonsurvival data or expected numbers of events for survival data) under various hypothetical references for all designs simultaneously, as shown in Output 77.4.3. By default, the option CREF= $0, 0.01, 0.02, \ldots, 1.50$ and expected sample sizes under the hypothetical references $\theta = c_i\,\theta_1$ are displayed, where c_i are values specified in the CREF= option. These CREF= values are displayed on the horizontal axis.

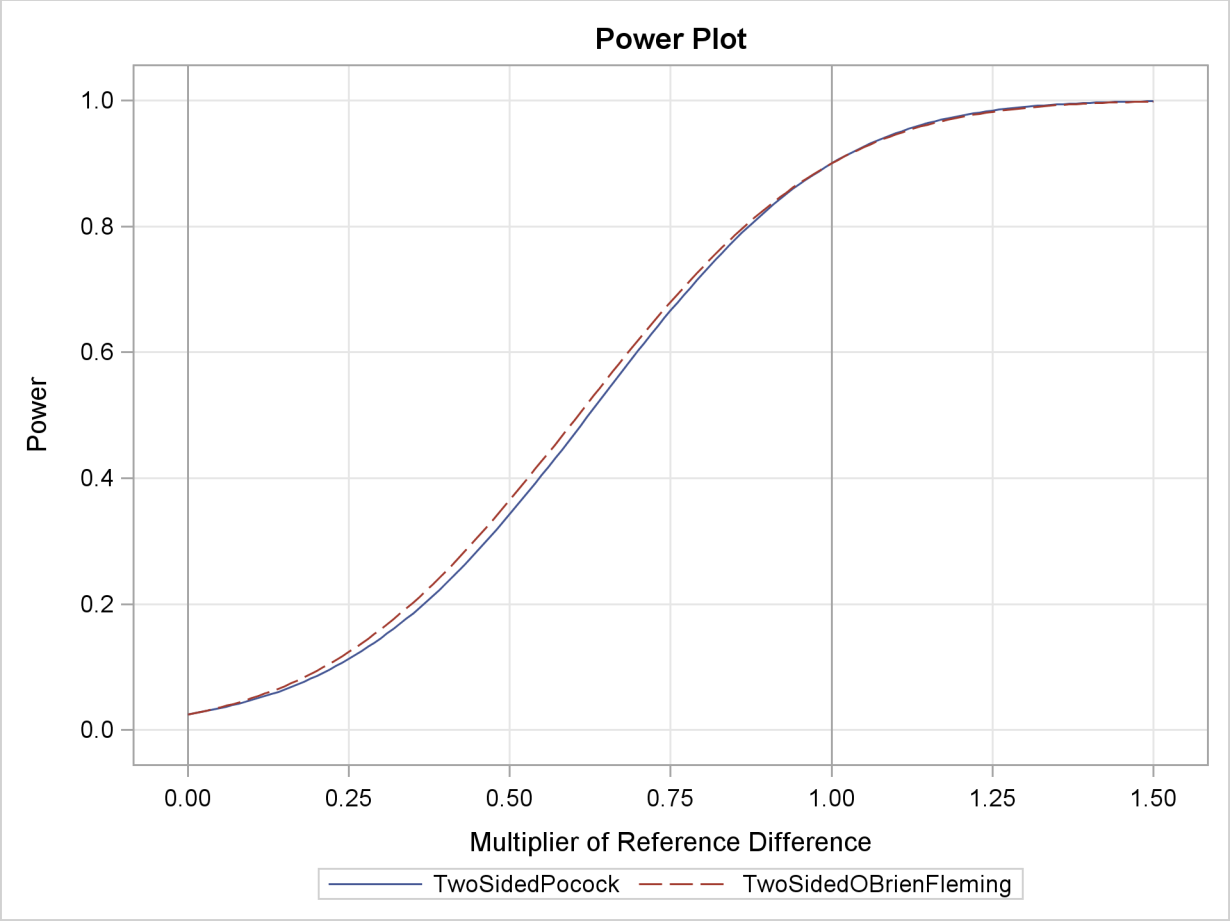
Output 77.4.3 ASN Plot



The plot shows that the Pocock design has a larger expected sample size than the O'Brien-Fleming design under the null hypothesis ($c_i = 0$) and has a smaller expected sample size under the alternative hypothesis ($c_i = 1$).

With the PLOTS=ALL option, the procedure displays a plot of the power curves under various hypothetical references for all designs simultaneously, as shown in Output 77.4.4. By default, the option CREF= $0, 0.01, 0.02, \ldots, 1.50$ and powers under hypothetical references $\theta = c_i\,\theta_1$ are displayed, where c_i are values specified in the CREF= option. These CREF= values are displayed on the horizontal axis.

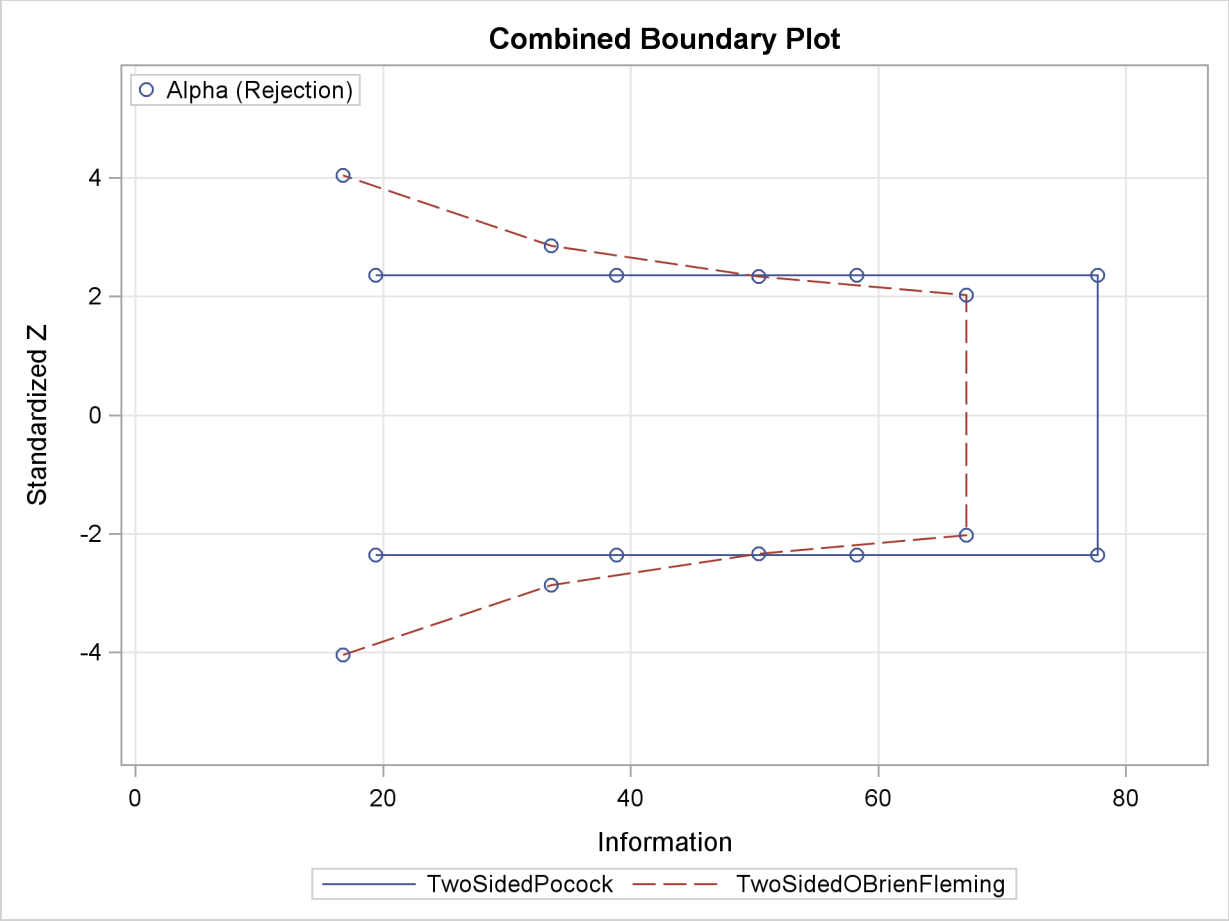
Output 77.4.4 Power Plot



Under the null hypothesis, $c_i = 0$, the power is 0.025, the upper Type I error probability. Under the alternative hypothesis, $c_i = 1$, the power is 0.9, one minus the Type II error probability. The plot shows only minor difference between the two designs.

With the PLOTS=ALL option, the procedure displays a plot of sequential boundaries for all designs simultaneously, as shown in Output 77.4.5. By default, the information levels are used on the horizontal axis.

Output 77.4.5 Combined Boundary Plot



The plot shows that the α boundary values (in absolute value) created from the O'Brien-Fleming method are greater in early stages and smaller in later stages than the boundary values from the Pocock method. The plot also shows that the information level in the Pocock design is larger than the corresponding level in the O'Brien-Fleming design at each stage.

With the PLOTS=ALL option, the procedure displays a plot of cumulative error spends for all boundaries in the designs simultaneously, as shown in Output 77.4.6. With a symmetric two-sided design, cumulative error spending is displayed only for the upper α boundary. The plot shows that for the upper α boundary, the O'Brien-Fleming method spends fewer errors in early stages and more errors in later stages than the corresponding Pocock method.

Output 77.4.6 Error Spending Plot

Example 77.5: Creating Designs Using Haybittle-Peto Methods

This example requests two 3-stage group sequential designs for normally distributed statistics. Each design uses a Haybittle-Peto method with a two-sided alternative and early stopping to reject the hypothesis. One design uses the specified interim boundary Z values and derives the final-stage boundary value for the specified α and β errors. The other design uses the specified boundary Z values and derives the overall α and β errors.

The following statements specify the interim boundary Z values and derive the final-stage boundary value for the specified $\alpha = 0.05$ and $\beta = 0.10$:

```
ods graphics on;
proc seqdesign altref=0.25
               errspend
               stopprob
               plots=errspend
               ;
   OneSidedPeto: design nstages=3
                 method=peto( z=3)
```

```
                    alt=upper   stop=reject
                    alpha=0.05  beta=0.10;
run;
ods graphics off;
```

The "Design Information" table in Output 77.5.1 displays design specifications and maximum information in percentage of its corresponding fixed-sample design.

Output 77.5.1 Haybittle-Peto Design Information

The SEQDESIGN Procedure
Design: OneSidedPeto

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Early Stop	Reject Null
Method	Haybittle-Peto
Boundary Key	Both
Alternative Reference	0.25
Number of Stages	3
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	100.2466
Max Information	137.3592
Null Ref ASN (Percent of Fixed Sample)	100.1192
Alt Ref ASN (Percent of Fixed Sample)	87.35

The "Method Information" table in Output 77.5.2 displays the α and β errors and the derived drift parameter, which is the standardized alternative reference at the final stage.

Output 77.5.2 Method Information

Method Information

Boundary	Method	Alpha	Beta	Alternative Reference	Drift
Upper Alpha	Haybittle-Peto	0.05000	0.10000	0.25	2.930009

With the STOPPROB option, the "Expected Cumulative Stopping Probabilities" table in Output 77.5.3 displays the expected stopping stage and cumulative stopping probability to reject the null hypothesis at each stage under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and $c_i = 0, 0.5, 1, 1.5$ are the default values in the CREF= option.

Output 77.5.3 Stopping Probabilities

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

CRef	Expected Stopping Stage	Source	----Stopping Probabilities--- Stage_1	Stage_2	Stage_3
0.0000	2.996	Reject Null	0.00135	0.00246	0.05000
0.5000	2.941	Reject Null	0.01561	0.04372	0.42762
1.0000	2.614	Reject Null	0.09538	0.29057	0.90000
1.5000	1.944	Reject Null	0.32185	0.73442	0.99698

The "Boundary Information" table in Output 77.5.4 displays information level, alternative references, and boundary values. The default BOUNDARYSCALE=STDZ option specifies that the standardized Z scale be used to display the alternative references and boundary values. The resulting standardized alternative reference at stage k is given by $\theta_1\sqrt{I_k}$, where θ_1 is the alternative reference and I_k is the information level at stage k, $k = 1, 2, 3$.

Output 77.5.4 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	---Information Level-- Proportion	Actual	-Alternative- --Reference-- Upper	-Boundary Values- ------Upper------ Alpha
1	0.3333	45.7864	1.69164	3.00000
2	0.6667	91.57281	2.39234	3.00000
3	1.0000	137.3592	2.93001	1.65042

At each interim stage, if the standardized statistic $z \geq 3$, the trial is stopped and the null hypothesis is rejected. If the statistic $z < 3$, the trial continues to the next stage. At the final stage, the null hypothesis is rejected if the statistic $z_3 > 1.65$. Otherwise, the hypothesis is accepted. Note that the boundary values at the final stage, 1.65, are close to the critical values 1.645 in the corresponding fixed-sample design.

The "Error Spending Information" in Output 77.5.5 displays cumulative error spending at each stage for each boundary. The stage 1 α spending 0.00135 corresponds to the one-sided p-value for a standardized Z statistic, $Z > 3$.

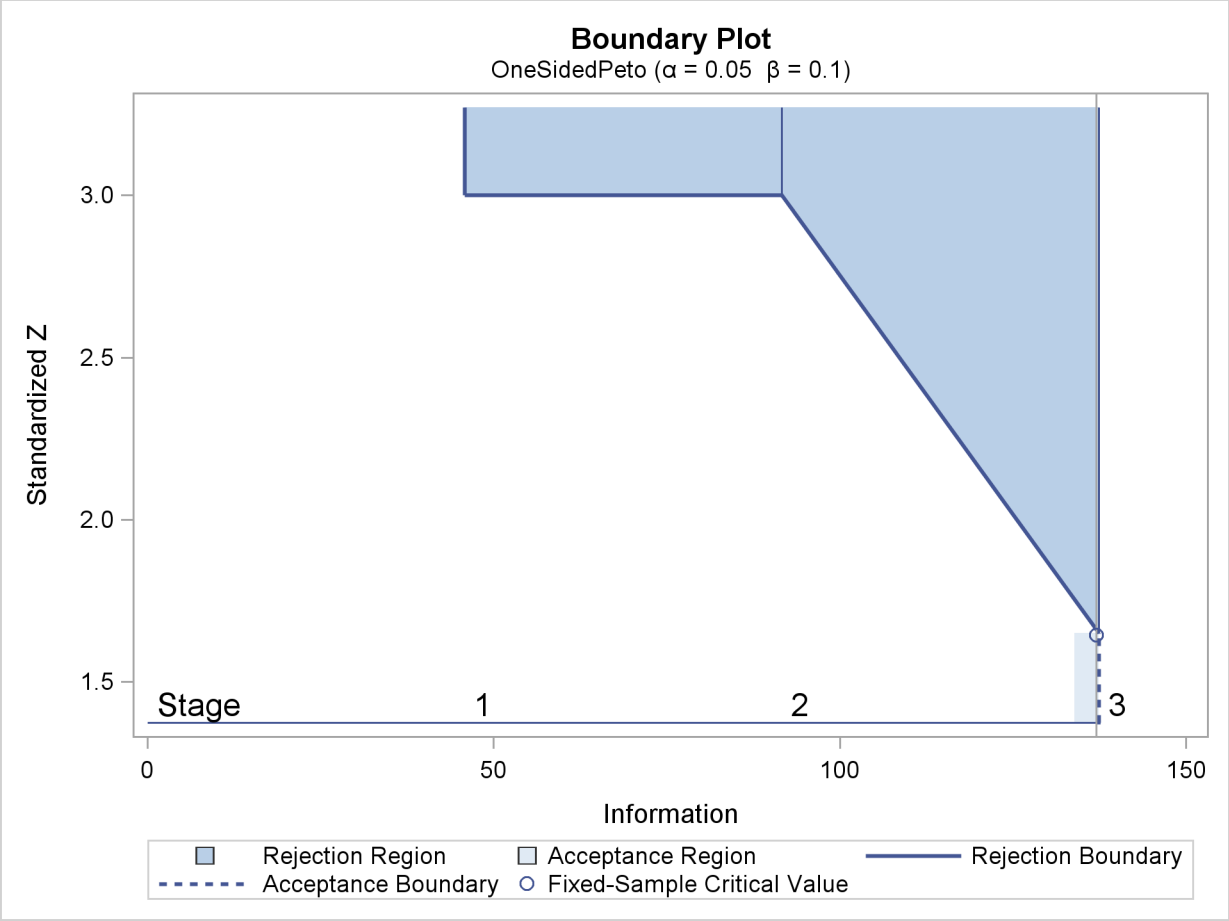
Output 77.5.5 Error Spending Information

Error Spending Information

Stage	-Information Level- Proportion	-Cumulative Error Spending- -----Upper----- Beta	Alpha
1	0.3333	0.00000	0.00135
2	0.6667	0.00000	0.00246
3	1.0000	0.10000	0.05000

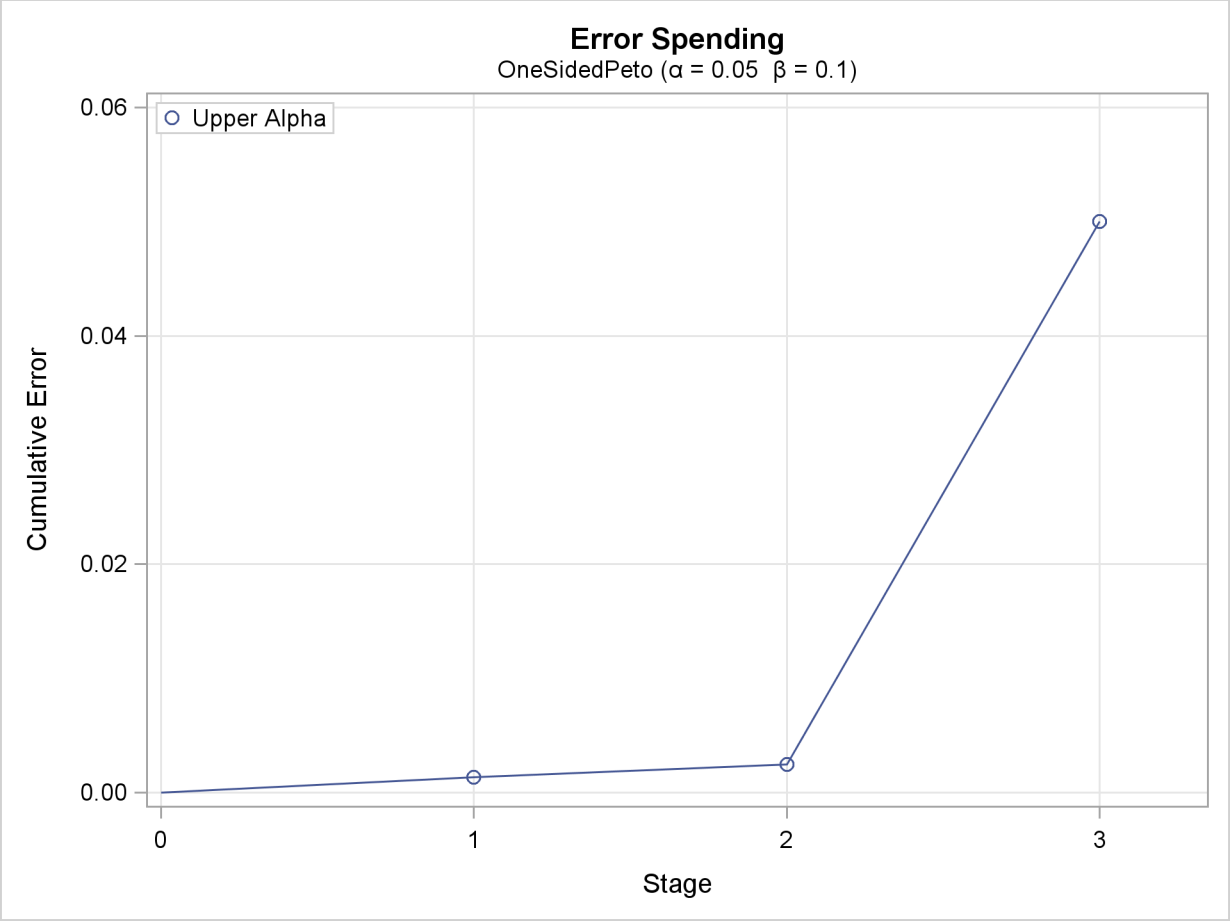
With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.5.6. With the STOP=REJECT option, the interim rejection boundaries are displayed.

Output 77.5.6 Boundary Plot



With the PLOTS=ERRSPEND option, the procedure displays a plot of error spending for each boundary, as shown in Output 77.5.7. The error spending values in the "Error Spending Information" in Output 77.5.4 are displayed in the plot. As expected, the error spending at each of the first two stages is small, with the standardized Z boundary value 3.

Output 77.5.7 Error Spending Plot



The following statements specify the boundary Z values and derive the α and β errors from these completely specified boundary values:

```
ods graphics on;
proc seqdesign altref=0.25
               maxinfo=200
               errspend
               stopprob
               plots=errspend
               ;
   OneSidedPeto: design nstages=3
                 method=peto(z=3 2.5 2)
                 alt=upper  stop=reject
                 boundarykey=none
                 ;
run;
ods graphics off;
```

The "Design Information" table in Output 77.5.8 displays design specifications and derived α and β error levels.

Output 77.5.8 Design Information

The SEQDESIGN Procedure
Design: OneSidedPeto

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Early Stop	Reject Null
Method	Haybittle-Peto
Boundary Key	None
Alternative Reference	0.25
Number of Stages	3
Alpha	0.02532
Beta	0.06035
Power	0.93965
Max Information (Percent of Fixed Sample)	101.6769
Max Information	200
Null Ref ASN (Percent of Fixed Sample)	101.3933
Alt Ref ASN (Percent of Fixed Sample)	73.74031

The "Method Information" table in Output 77.5.9 displays the α and β errors and the derived drift parameter for each boundary.

Output 77.5.9 Method Information

Method Information

Boundary	Method	Alpha	Beta	Alternative Reference	Drift
Upper Alpha	Haybittle-Peto	0.02532	0.06035	0.25	3.535534

With the STOPPROB option, the "Expected Cumulative Stopping Probabilities" table in Output 77.5.10 displays the expected stopping stage and cumulative stopping probability to reject the null hypothesis at each stage under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and $c_i = 0, 0.5, 1, 1.5$ are the default values in the CREF= option.

Output 77.5.10 Stopping Probabilities

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

CRef	Expected Stopping Stage	Source	----Stopping Probabilities--- Stage_1	Stage_2	Stage_3
0.0000	2.992	Reject Null	0.00135	0.00702	0.02532
0.5000	2.826	Reject Null	0.02389	0.15030	0.41775
1.0000	2.176	Reject Null	0.16884	0.65544	0.93965
1.5000	1.508	Reject Null	0.52466	0.96708	0.99954

The "Boundary Information" table in Output 77.5.11 displays information level, alternative references, and boundary values.

Output 77.5.11 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper	Boundary Values Upper Alpha
1	0.3333	66.66667	2.04124	3.00000
2	0.6667	133.3333	2.88675	2.50000
3	1.0000	200	3.53553	2.00000

The "Error Spending Information" in Output 77.5.12 displays cumulative error spending at each stage for each boundary. The first-stage α spending 0.00135 corresponds to the one-sided p-value for a standardized Z statistic, $Z > 3$.

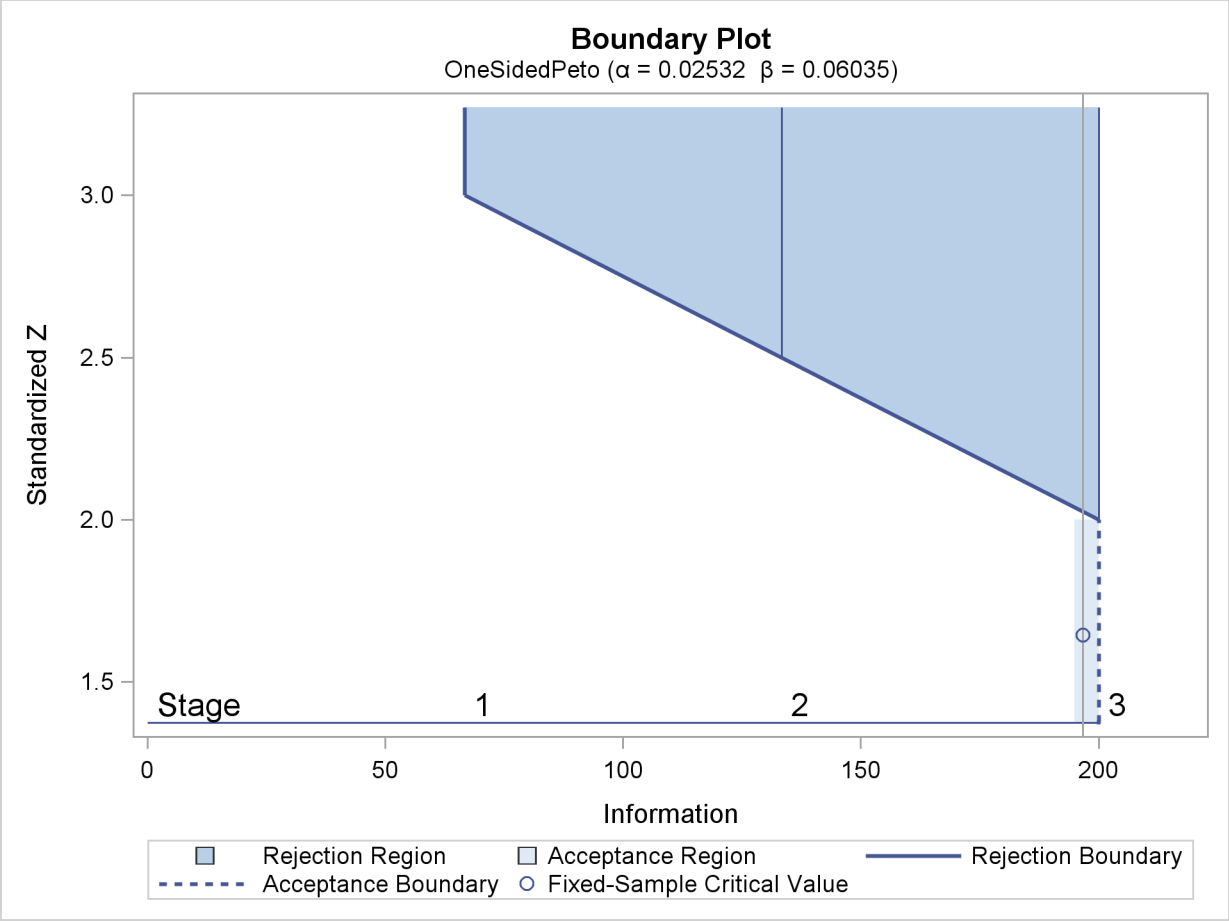
Output 77.5.12 Error Spending Information

Error Spending Information

Stage	Information Level Proportion	Cumulative Error Spending Upper Beta	Cumulative Error Spending Upper Alpha
1	0.3333	0.00000	0.00135
2	0.6667	0.00000	0.00702
3	1.0000	0.06035	0.02532

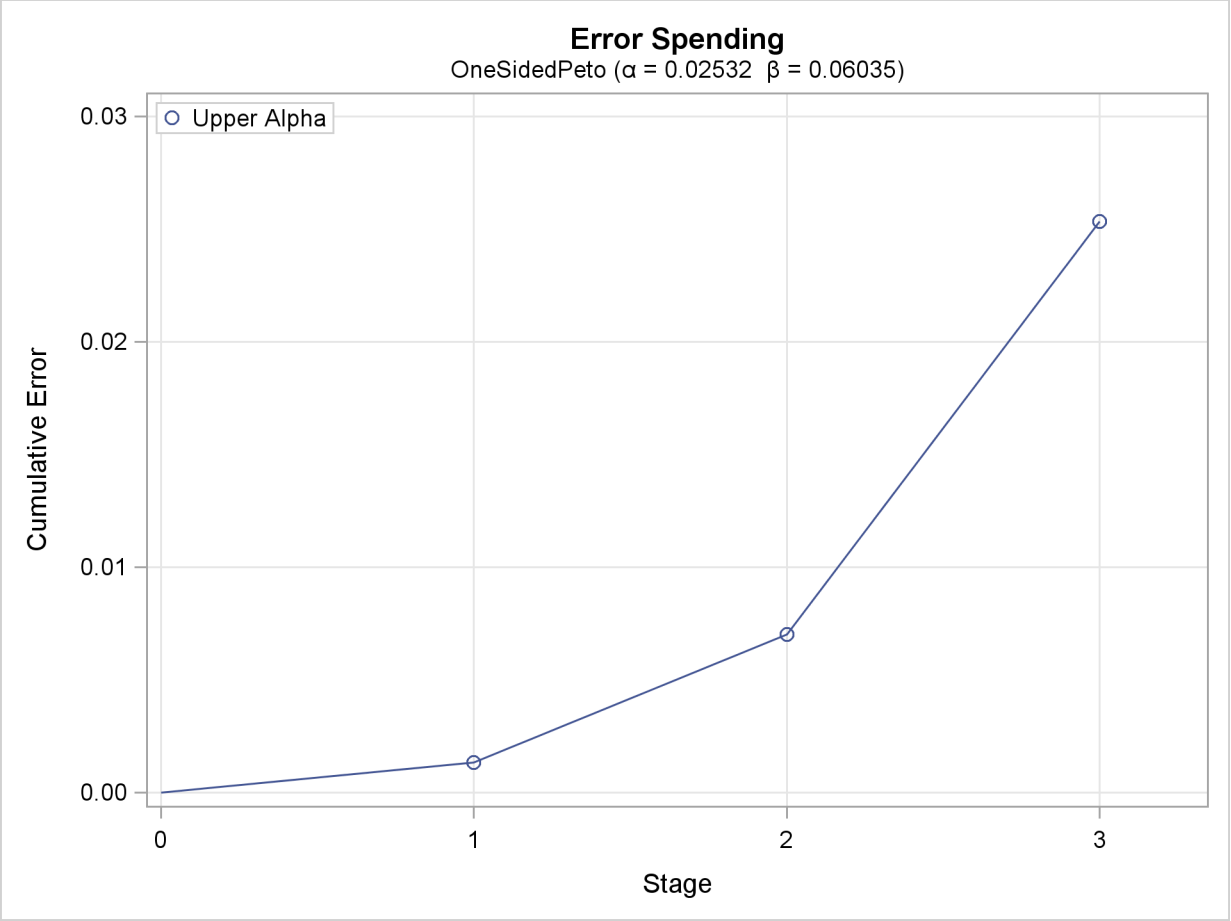
With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.5.13. With the STOP=REJECT option, the interim rejection boundaries are displayed.

Output 77.5.13 Boundary Plot



With the PLOTS=ERRSPEND option, the procedure displays a plot of error spending for each boundary, as shown in Output 77.5.14. The error spending values in the "Error Spending Information" table in Output 77.5.10 are displayed in the plot.

Output 77.5.14 Error Spending Plot



Example 77.6: Creating Designs with Various Stopping Criteria

This example requests three 5-stage group sequential designs for normally distributed statistics. Each design uses a triangular method with the specified one-sided upper alternative reference $\theta_1 = 0.2$. The resulting boundary values are displayed with the score scale. Note that these unified family triangular designs are different from Whitehead's triangular designs.

The following statements request three designs with different stopping criterion:

```
ods graphics on;
proc seqdesign altref=0.2
               bscale=score
               errspend
               plots=(combinedboundary errspend(hscale=info))
               ;
   StopToRejectAccept: design nstages=5 method=tri alt=upper stop=both;
   StopToReject:       design nstages=5 method=tri alt=upper stop=reject;
   StopToAccept:       design nstages=5 method=tri alt=upper stop=accept;
run;
ods graphics off;
```

The first design has early stopping to reject or accept the null hypothesis H_0.

The "Design Information" table in Output 77.6.1 displays design specifications and derived statistics. With the specified alternative reference, the maximum information is derived.

Output 77.6.1 Triangular Design Information

The SEQDESIGN Procedure
Design: StopToRejectAccept

Design Information

Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Method	Triangular
Boundary Key	Both
Alternative Reference	0.2
Number of Stages	5
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	140.0293
Max Information	299.797
Null Ref ASN (Percent of Fixed Sample)	59.11973
Alt Ref ASN (Percent of Fixed Sample)	66.94909

The "Method Information" table in Output 77.6.2 displays the α and β errors and the derived drift parameter, which is the standardized alternative reference at the final stage. The table also shows the corresponding parameters for a triangular method as a unified family method.

Output 77.6.2 Method Information

Method Information

				Unified Family		
Boundary	Method	Alpha	Beta	Rho	Tau	C
Upper Alpha	Triangular	0.05000	.	0.5	1	0.94394
Upper Beta	Triangular	.	0.10000	0.5	1	0.78753

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.2	3.46293
Upper Beta	0.2	3.46293

The "Boundary Information" table in Output 77.6.3 displays information level, alternative reference, and boundary values. With the specified BOUNDARYSCALE=SCORE option, the alternative reference and boundary values are displayed in the score statistic scale. With a score scale, the alternative reference is $\theta_1 I_k$, where θ_1 is the specified alternative reference and I_k is the information level at stage k, $k = 1, 2, \ldots, 5$.

Output 77.6.3 Boundary Information

Boundary Information (Score Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper	Boundary Values Upper Beta	Boundary Values Upper Alpha
1	0.2000	59.9594	11.99188	-4.37102	19.61274
2	0.4000	119.9188	23.98376	4.89371	22.88154
3	0.6000	179.8782	35.97564	14.15845	26.15033
4	0.8000	239.8376	47.96752	23.42318	29.41912
5	1.0000	299.797	59.95940	32.68791	32.68791

The "Error Spending Information" table in Output 77.6.4 displays cumulative error spending at each stage for each boundary.

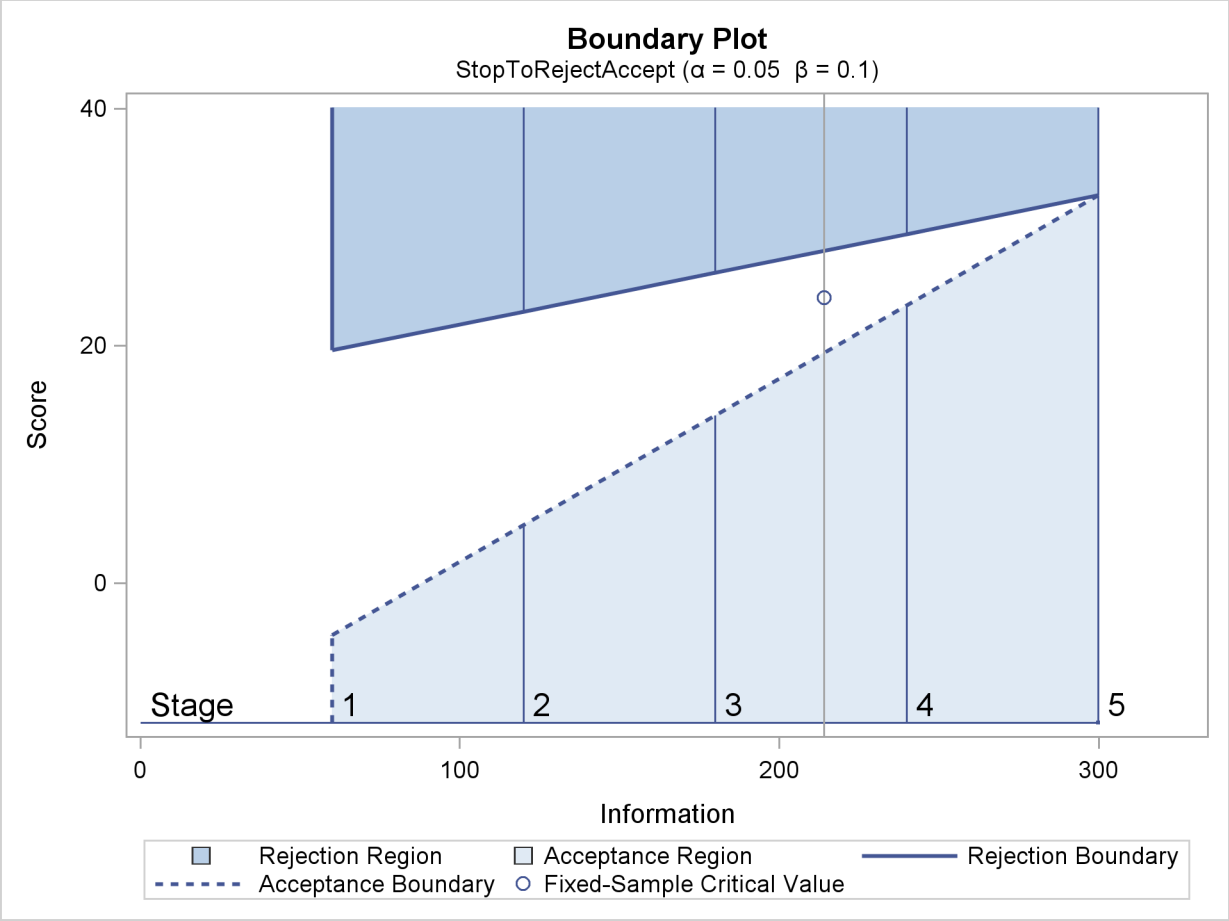
Output 77.6.4 Error Spending Information

Error Spending Information

Stage	Information Level Proportion	Cumulative Error Spending Upper Beta	Cumulative Error Spending Upper Alpha
1	0.2000	0.01729	0.00566
2	0.4000	0.04927	0.02138
3	0.6000	0.07611	0.03643
4	0.8000	0.09357	0.04641
5	1.0000	0.10000	0.05000

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.6.5. With the STOP=BOTH option, both the acceptance and rejection boundaries at interim stages are displayed. With the score scale, the acceptance and rejection boundaries are straight lines and form a triangular-shape continuation region.

Output 77.6.5 Boundary Plot with Score Statistics



The second design has early stopping only to reject the null hypothesis H_0.

The "Design Information" table in Output 77.6.6 displays design specifications and derived statistics. With the specified alternative reference, the maximum information is derived.

Output 77.6.6 Triangular Design Information

The SEQDESIGN Procedure
Design: StopToReject

Design Information

Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Reject Null
Method	Triangular
Boundary Key	Both
Alternative Reference	0.2
Number of Stages	5
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	113.4443
Max Information	242.8799
Null Ref ASN (Percent of Fixed Sample)	111.3399
Alt Ref ASN (Percent of Fixed Sample)	67.41968

The "Method Information" table in Output 77.6.7 displays the α and β errors and the derived drift parameter. The table also shows the corresponding parameters for a triangular method as a unified family method.

Output 77.6.7 Method Information

Method Information

				Unified Family		
Boundary	Method	Alpha	Beta	Rho	Tau	C
Upper Alpha	Triangular	0.05000	0.10000	0.5	1	0.9833

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.2	3.116921

The "Boundary Information" table in Output 77.6.8 displays information level, alternative reference, and boundary values. With the specified BOUNDARYSCALE=SCORE option, the alternative reference and boundary values are displayed in the score statistic scale.

Output 77.6.8 Boundary Information

Boundary Information (Score Scale)
Null Reference = 0

Stage	---Information Level-- Proportion	Actual	-Alternative- --Reference-- Upper	-Boundary Values- ------Upper------ Alpha
1	0.2000	48.57597	9.71519	18.38919
2	0.4000	97.15194	19.43039	21.45405
3	0.6000	145.7279	29.14558	24.51891
4	0.8000	194.3039	38.86078	27.58378
5	1.0000	242.8799	48.57597	30.64864

The "Error Spending Information" table in Output 77.6.9 displays cumulative error spending at each stage for each boundary.

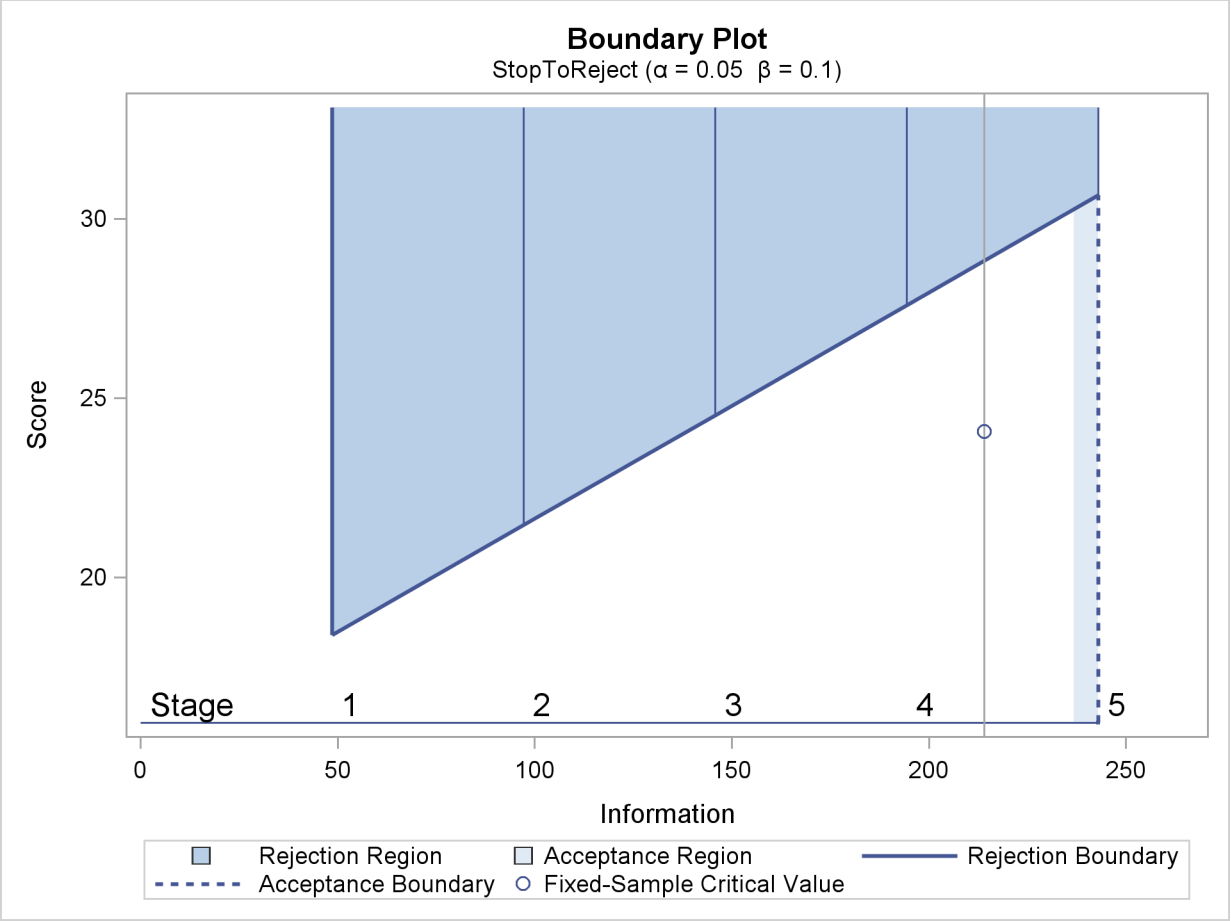
Output 77.6.9 Error Spending Information

Error Spending Information

Stage	-Information Level- Proportion	-Cumulative Error Spending- ----------Upper---------- Beta	Alpha
1	0.2000	0.00000	0.00416
2	0.4000	0.00000	0.01705
3	0.6000	0.00000	0.03027
4	0.8000	0.00000	0.04127
5	1.0000	0.10000	0.05000

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.6.10. For a triangular design, these rejection boundaries form a straight line with the score scale.

Output 77.6.10 Boundary Plot with Score Statistics



The third design has early stopping to accept the null hypothesis H_0.

The "Design Information" table in Output 77.6.11 displays design specifications and derived statistics. With the specified alternative reference, the maximum information is derived.

Output 77.6.11 Triangular Design Information

The SEQDESIGN Procedure
Design: StopToAccept

Design Information

Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Accept Null
Method	Triangular
Boundary Key	Both
Alternative Reference	0.2
Number of Stages	5
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	114.9925
Max Information	246.1945
Null Ref ASN (Percent of Fixed Sample)	57.83208
Alt Ref ASN (Percent of Fixed Sample)	110.2477

The "Method Information" table in Output 77.6.12 displays the α and β errors and the derived drift parameter. The table also shows the corresponding parameters for a triangular method as a unified family method.

Output 77.6.12 Method Information

Method Information

				Unified Family		
Boundary	Method	Alpha	Beta	Rho	Tau	C
Upper Beta	Triangular	0.05000	0.10000	0.5	1	0.82154

Method Information

Boundary	Alternative Reference	Drift
Upper Beta	0.2	3.138117

The "Boundary Information" table in Output 77.6.13 displays information level, alternative reference, and boundary values. With the specified BOUNDARYSCALE=SCORE option, the alternative reference and boundary values are displayed in the score statistic scale.

Output 77.6.13 Boundary Information

Boundary Information (Score Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper	Boundary Values Upper Beta
1	0.2000	49.2389	9.84778	-5.62074
2	0.4000	98.4778	19.69556	1.64895
3	0.6000	147.7167	29.54334	8.91865
4	0.8000	196.9556	39.39112	16.18834
5	1.0000	246.1945	49.23890	23.45803

The "Error Spending Information" table in Output 77.6.14 displays cumulative error spending at each stage for each boundary.

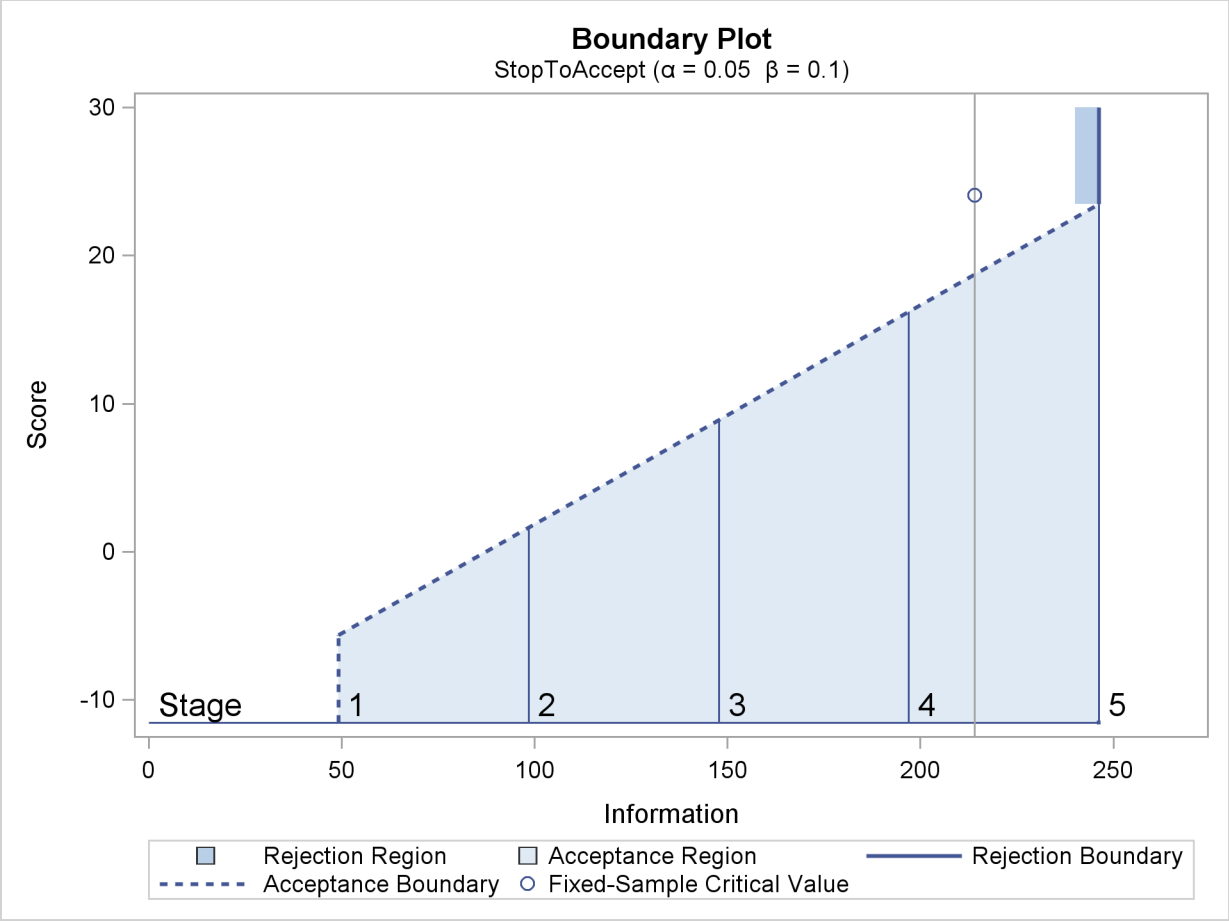
Output 77.6.14 Error Spending Information

Error Spending Information

Stage	Information Level Proportion	Cumulative Error Spending Upper Beta	Cumulative Error Spending Upper Alpha
1	0.2000	0.01375	0.00000
2	0.4000	0.04149	0.00000
3	0.6000	0.06594	0.00000
4	0.8000	0.08513	0.00000
5	1.0000	0.10000	0.05000

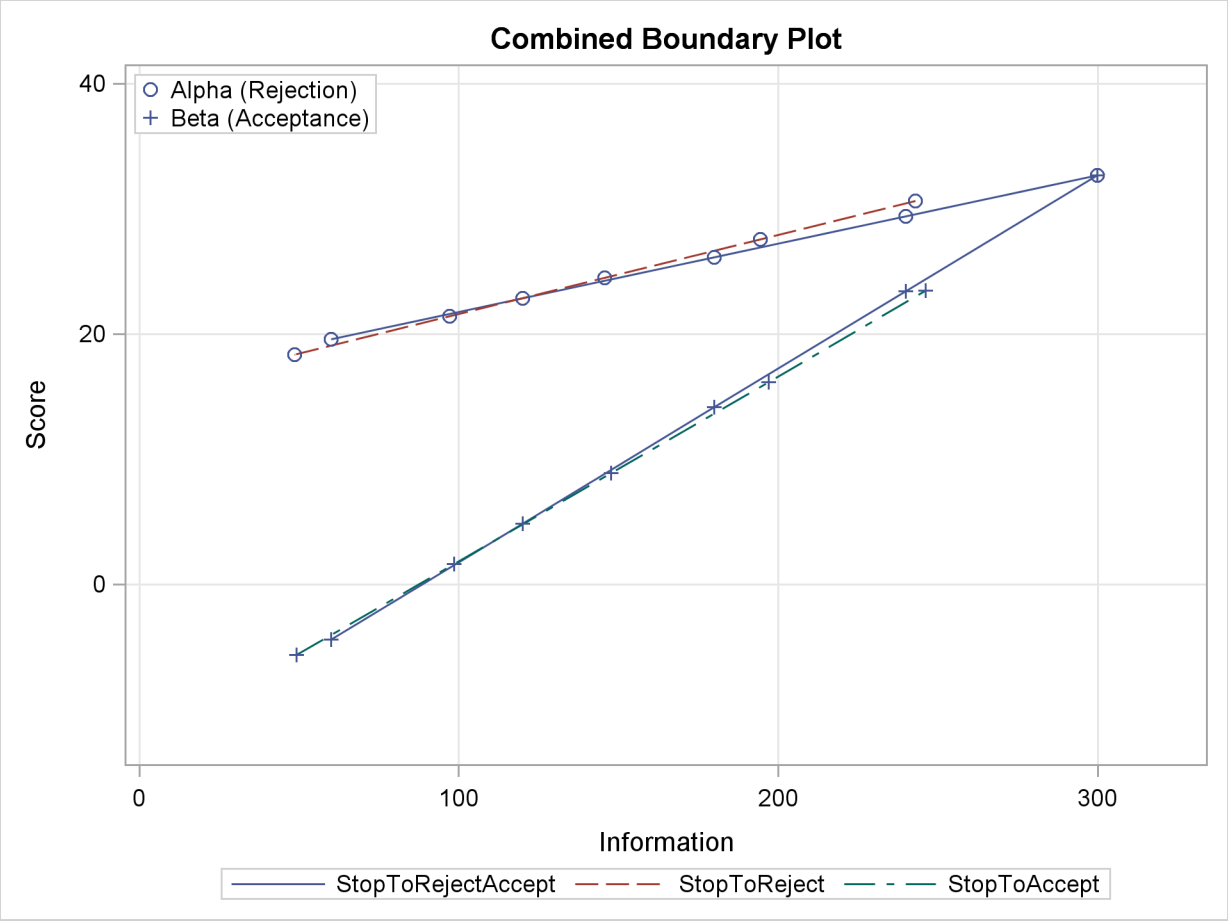
With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.6.15. For a triangular design, these rejection boundaries form a straight line with the score scale.

Output 77.6.15 Boundary Plot with Score Scale



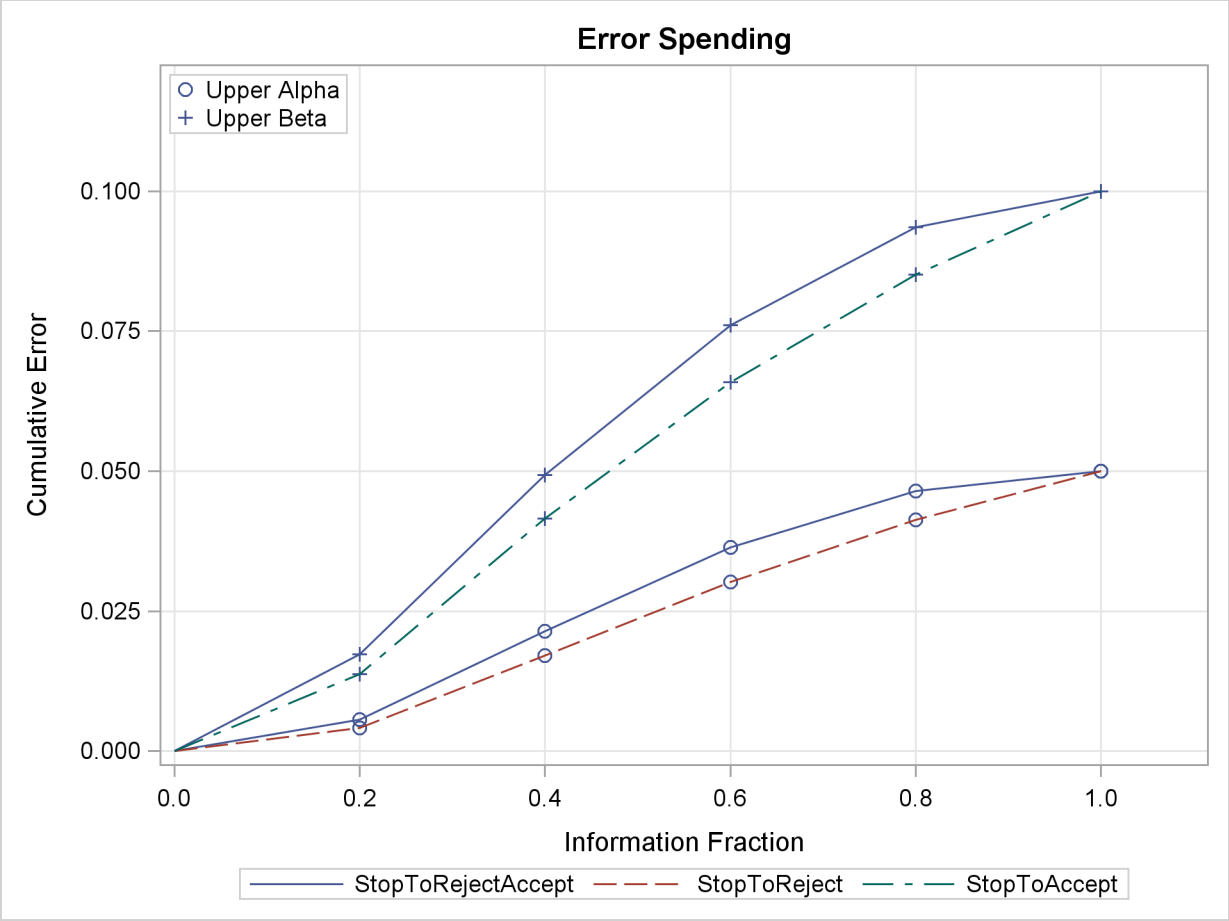
With the PLOTS=COMBINEDBOUNDARY option, a plot of the resulting sequential boundaries for all designs is displayed, as shown in Output 77.6.16. The plot shows that the design with early stopping to reject and to accept H_0 has larger maximum information than the other two designs.

Output 77.6.16 Combined Boundary Plot with Score Scale



With the PLOTS=ERRSPEND(HSCALE=INFO) option, the error spending plot is displayed with the information level on the horizontal axis, as shown in Output 77.6.17. The design with early stopping to reject or accept the null hypothesis H_0 has larger α spending and larger β spending in early stages than the other two designs.

Output 77.6.17 Error Spending Plot



Example 77.7: Creating Whitehead's Triangular Designs

This example requests three 4-stage Whitehead's triangular designs for normally distributed statistics. Each design has a one-sided alternative hypothesis with early stopping to reject or accept the null hypothesis H_0. Note that Whitehead's triangular designs are different from unified family triangular designs.

Suppose that a clinic is conducting a study of the effect of a new cancer treatment. The study consists of exposing mice to a carcinogen and randomly assigning them to either the control group or the treatment group. The event of interest is death from cancer induced by the carcinogen, and the response is the time from randomization to death.

Following the derivations in the section "Test for Two Survival Distributions with a Log-Rank Test" on page 5866, the hypothesis $H_0 : \theta = -\log(\lambda) = 0$ with an alternative hypothesis $H_1 : \theta = \theta_1 > 0$ is used, where λ is the hazard ratio between the treatment group and the control group.

Also suppose that from past experience, the median survival time for the control group is $t_0 = 20$ days, and the study wants to detect a $t_1 = 40$ days' median survival time with a 80% power in the trial. Assuming exponential survival functions for the two groups, the hazard rates can be computed

from

$$S_j(t_j) = e^{-h_j t_j} = \frac{1}{2}$$

where $j = 0, 1$.

Thus, with $h_0 = 0.0346574$ and $h_1 = 0.0173287$, the hazard ratio $\lambda_1 = h_1/h_0 = 1/2$, and the alternative reference is

$$\theta_1 = -\log(\lambda_1) = -\log(\frac{1}{2}) = 0.693147$$

The following statements invoke the SEQDESIGN procedure and specify three Whitehead's triangular designs:

```
ods graphics on;
proc seqdesign altref=0.693147
               bscale=score
               plots=combinedboundary
               ;
   BoundaryKeyNone:  design nstages=4
                            method=whitehead
                            boundarykey=none
                            alt=upper   stop=both
                            alpha=0.05  beta=0.20
                            ;
   BoundaryKeyAlpha: design nstages=4
                            method=whitehead
                            boundarykey=alpha
                            alt=upper   stop=both
                            alpha=0.05  beta=0.20
                            ;
   BoundaryKeyBeta:  design nstages=4
                            method=whitehead
                            boundarykey=beta
                            alt=upper   stop=both
                            alpha=0.05  beta=0.20
                            ;
run;
ods graphics off;
```

Whitehead methods with early stopping to reject or accept the null hypothesis create boundaries that approximately satisfy the Type I and Type II error probability specification. The BOUNDARYKEY=NONE option specifies no adjustment to the boundary value at the final stage to maintain either a Type I or a Type II error probability level.

The "Design Information" table in Output 77.7.1 displays design specifications and maximum information. Note that with the BOUNDARYKEY=NONE option, the derived errors $\alpha = 0.05071$ and $\beta = 0.19771$ are not the same as the specified errors $\alpha = 0.05$ and $\beta = 0.20$.

Output 77.7.1 Whitehead Design Information

The SEQDESIGN Procedure
Design: BoundaryKeyNone

Design Information

Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Method	Whitehead
Boundary Key	None
Alternative Reference	0.693147
Number of Stages	4
Alpha	0.05071
Beta	0.19771
Power	0.80229
Max Information (Percent of Fixed Sample)	129.6815
Max Information	16.70639
Null Ref ASN (Percent of Fixed Sample)	62.48184
Alt Ref ASN (Percent of Fixed Sample)	73.82535

The "Method Information" table in Output 77.7.2 displays the derived α and β errors and the derived drift parameter. The derived errors $\alpha = 0.05071$ and $\beta = 0.19771$ are not exactly the same as the specified errors $\alpha = 0.05$ and $\beta = 0.20$ with the BOUNDARYKEY=NONE option.

Output 77.7.2 Method Information

Method Information

				Whitehead	
Boundary	Method	Alpha	Beta	Tau	C
Upper Alpha	Whitehead	0.05071	.	0.25	4.60517
Upper Beta	Whitehead	.	0.19771	0.25	4.60517

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.693147	2.833131
Upper Beta	0.693147	2.833131

The "Boundary Information" table in Output 77.7.3 displays information level, alternative reference, and boundary values. With the specified BOUNDARYSCALE=SCORE option, the alternative reference and boundary values are displayed with the score statistics scale.

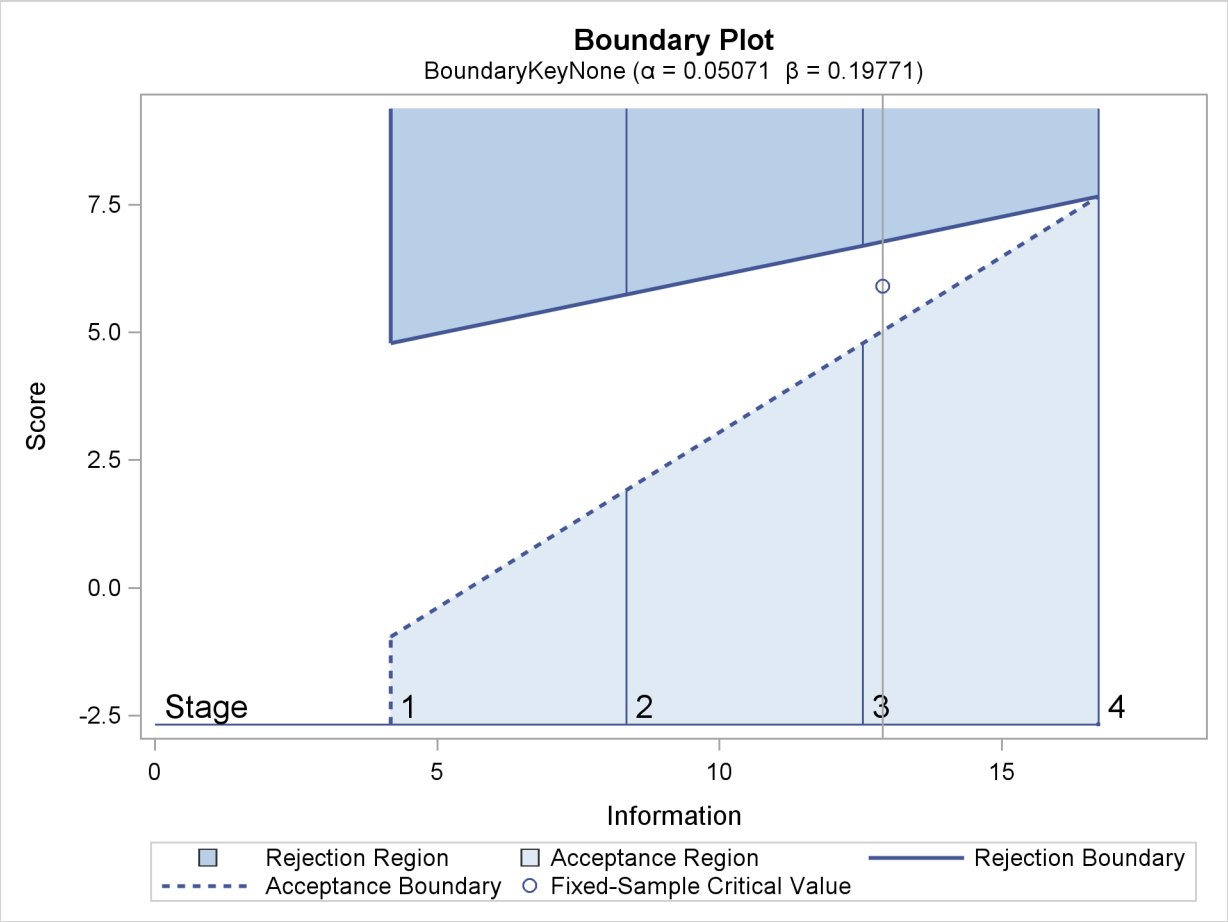
Output 77.7.3 Boundary Information

Boundary Information (Score Scale)
Null Reference = 0

Stage	---Information Level-- Proportion	Actual	-Alternative- --Reference-- Upper	-----Boundary Values---- ----------Upper---------- Beta	Alpha
1	0.2500	4.176597	2.89500	-0.95755	4.78775
2	0.5000	8.353195	5.78999	1.91510	5.74530
3	0.7500	12.52979	8.68499	4.78775	6.70285
4	1.0000	16.70639	11.57998	7.66039	7.66039

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.7.4.

Output 77.7.4 Boundary Plot



The second design uses the BOUNDARYKEY=ALPHA option to adjust the boundary value at the final stage to maintain the Type I error probability level.

The "Design Information" table in Output 77.7.5 displays design specifications and the derived maximum information. Note that with the BOUNDARYKEY=ALPHA option, the specified Type I error probability $\alpha = 0.05$ is maintained.

Output 77.7.5 Whitehead Design Information

The SEQDESIGN Procedure
Design: BoundaryKeyAlpha

Design Information

Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Method	Whitehead
Boundary Key	Alpha
Alternative Reference	0.693147
Number of Stages	4
Alpha	0.05
Beta	0.20044
Power	0.79956
Max Information (Percent of Fixed Sample)	129.9894
Max Information	16.70639
Null Ref ASN (Percent of Fixed Sample)	62.6302
Alt Ref ASN (Percent of Fixed Sample)	74.00064

The "Method Information" table in Output 77.7.6 displays the specified and derived α and β errors and the derived drift parameter. The derived Type I error probability is the same as the specified $\alpha = 0.05$ and the derived Type II error probability $\beta = 0.20044$ is not the same as the specified $\beta = 0.20$ with the BOUNDARYKEY=ALPHA option.

Output 77.7.6 Method Information

Method Information

				------Whitehead-----	
Boundary	Method	Alpha	Beta	Tau	C
Upper Alpha	Whitehead	0.05000	.	0.25	4.60517
Upper Beta	Whitehead	.	0.20044	0.25	4.60517

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.693147	2.833131
Upper Beta	0.693147	2.833131

The "Boundary Information" table in Output 77.7.7 displays information level, alternative reference, and boundary values.

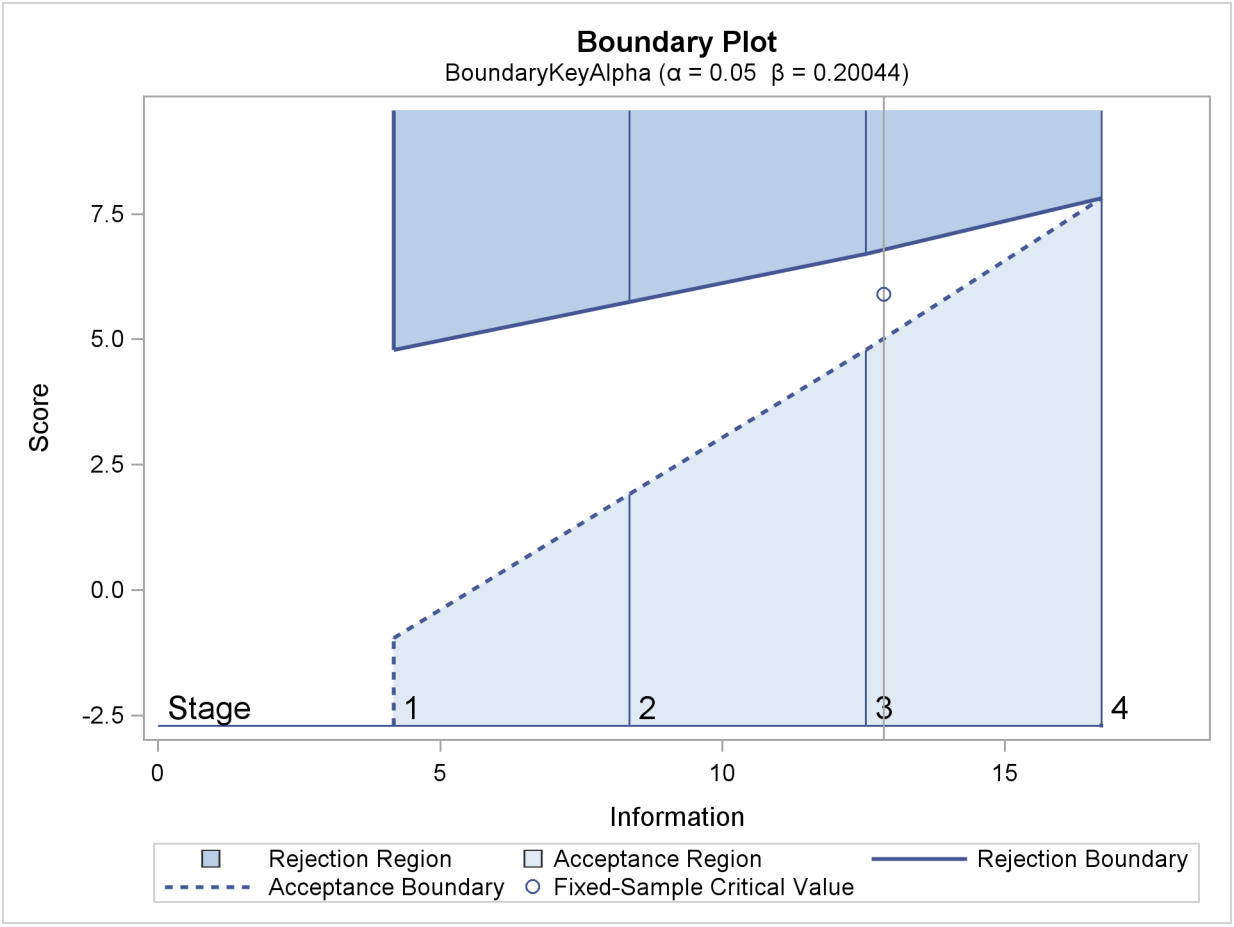
Output 77.7.7 Boundary Information

Boundary Information (Score Scale)
Null Reference = 0

Stage	---Information Level-- Proportion	Actual	-Alternative- --Reference-- Upper	-----Boundary Values---- ----------Upper--------- Beta	Alpha
1	0.2500	4.176597	2.89500	-0.95755	4.78775
2	0.5000	8.353195	5.78999	1.91510	5.74530
3	0.7500	12.52979	8.68499	4.78775	6.70285
4	1.0000	16.70639	11.57998	7.81300	7.81300

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.7.8.

Output 77.7.8 Boundary Plot



The third design specifies the BOUNDARYKEY=BETA option to derive the boundary values to maintain the Type II error probability level β.

The "Design Information" table in Output 77.7.9 displays design specifications and the derived maximum information. Note that with the BOUNDARYKEY=BETA option, the specified Type II error probability $\beta = 0.20$ is maintained.

Output 77.7.9 Whitehead Design Information

The SEQDESIGN Procedure
Design: BoundaryKeyBeta

Design Information

Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Method	Whitehead
Boundary Key	Beta
Alternative Reference	0.693147
Number of Stages	4
Alpha	0.05011
Beta	0.2
Power	0.8
Max Information (Percent of Fixed Sample)	129.9364
Max Information	16.70639
Null Ref ASN (Percent of Fixed Sample)	62.60462
Alt Ref ASN (Percent of Fixed Sample)	73.97042

The "Method Information" table in Output 77.7.10 displays the α and β errors and the derived drift parameter. The derived Type II error probability is the same as the specified $\beta = 0.20$ and the derived Type I error probability $\alpha = 0.05011$ is not the same as the specified $\alpha = 0.05$ with the BOUNDARYKEY=BETA option.

Output 77.7.10 Method Information

Method Information

				------Whitehead-----	
Boundary	Method	Alpha	Beta	Tau	C
Upper Alpha	Whitehead	0.05011	.	0.25	4.60517
Upper Beta	Whitehead	.	0.20000	0.25	4.60517

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.693147	2.833131
Upper Beta	0.693147	2.833131

The "Boundary Information" table in Output 77.7.11 displays information level, alternative reference, and boundary values.

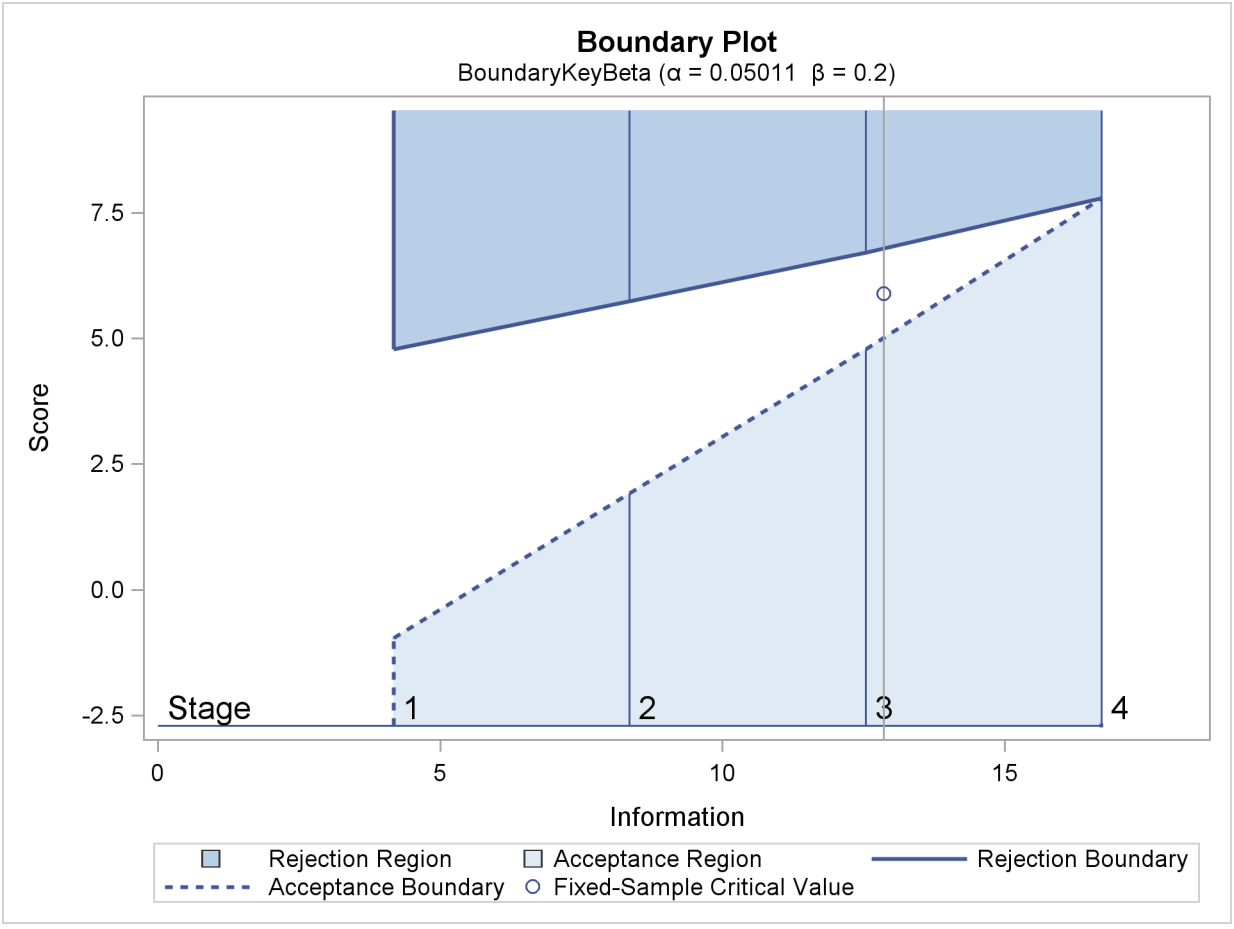
Output 77.7.11 Boundary Information

Boundary Information (Score Scale)
Null Reference = 0

	---Information Level--		-Alternative- --Reference--	-----Boundary Values---- ----------Upper---------	
Stage	Proportion	Actual	Upper	Beta	Alpha
1	0.2500	4.176597	2.89500	-0.95755	4.78775
2	0.5000	8.353195	5.78999	1.91510	5.74530
3	0.7500	12.52979	8.68499	4.78775	6.70285
4	1.0000	16.70639	11.57998	7.78899	7.78899

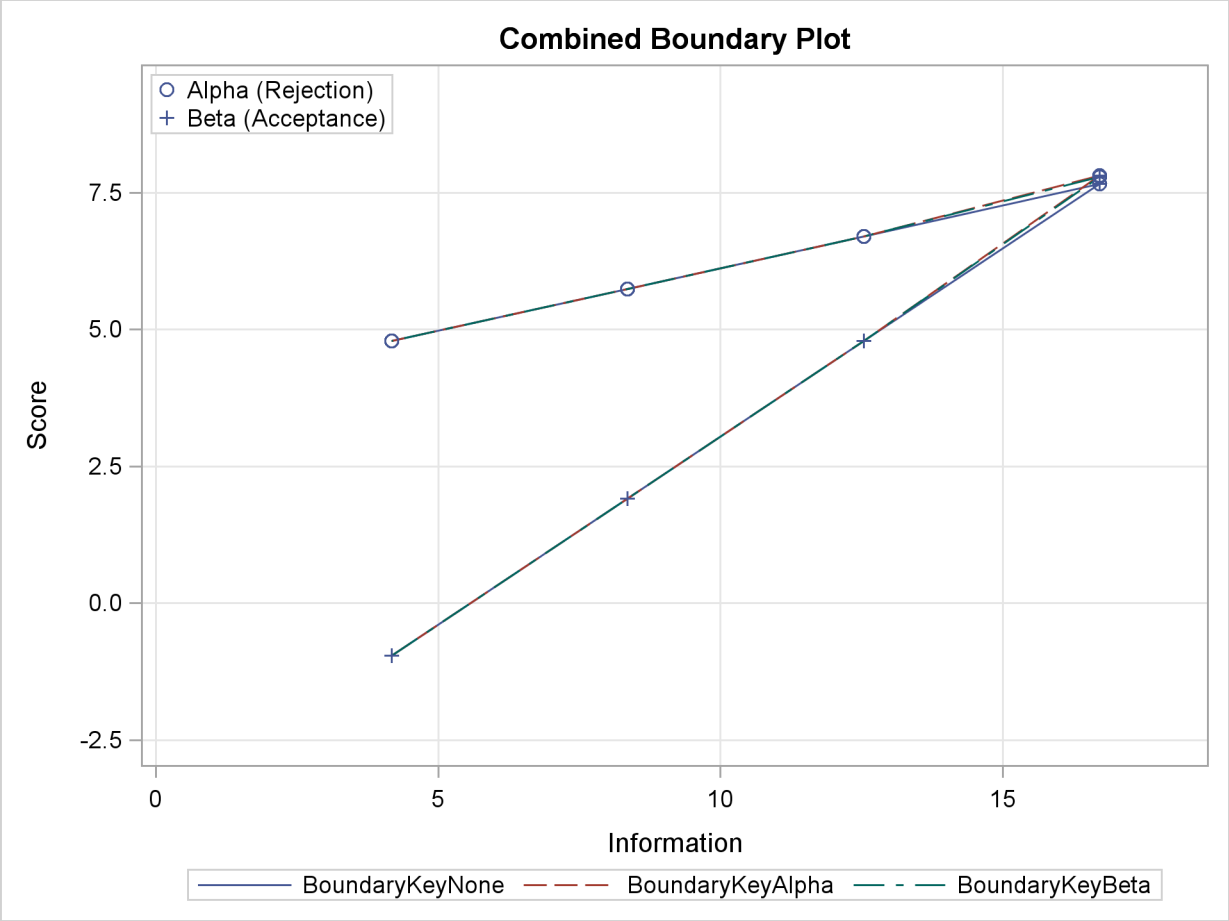
With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.7.12.

Output 77.7.12 Boundary Plot



With the PLOTS=COMBINEDBOUNDARY option, a combined plot of group sequential boundaries for all designs is displayed, as shown in Output 77.7.13. It shows that three designs are similar, with a slightly smaller boundary value at the final stage for the design with the BOUNDARYKEY=NONE option.

Output 77.7.13 Combined Boundary Plot



The following statements invoke the SEQDESIGN procedure and specify the SAMPLESIZE statement to derive required sample sizes for a log-rank test comparing two survival distributions for the treatment effect (Jennison and Turnbull 2000 pp. 77–79; Whitehead 1997, pp. 36–39):

```
proc seqdesign altref=0.693147
               bscale=score
               ;
  BoundaryKeyAlpha: design nstages=4
                           method=whitehead
                           boundarykey=alpha
                           alt=upper   stop=both
                           alpha=0.05  beta=0.20
                           ;
  samplesize model=twosamplesurvival
                  ( nullhazard=0.03466 accrate=10);
run;
```

The design is identical to the previous design with the BOUNDARYKEY=ALPHA option except with the addition of the sample size computation.

The "Sample Size Summary" table in Output 77.7.14 displays parameters for the sample size computation. Since the ACCTIME= option is not specified for the accrual time, the minimum and maximum accrual times are derived for the specified accrual rate.

Output 77.7.14 Sample Size Summary

The SEQDESIGN Procedure
Design: BoundaryKeyAlpha

Sample Size Summary

Test	Two-Sample Survival
Null Hazard Rate	0.03466
Hazard Rate (Group A)	0.01733
Hazard Rate (Group B)	0.03466
Hazard Ratio	0.5
log(Hazard Ratio)	-0.69315
Reference Hazards	Alt Ref
Accrual Rate	10
Min Accrual Time	6.682556
Min Sample Size	66.82556
Max Accrual Time	25.40111
Max Sample Size	254.0111
Max Number of Events	66.82556

If the ACCTIME=20 option is specified in the SAMPLESIZE statement, the "Sample Size Summary" table in Output 77.7.15 also displays the follow-up time and maximum sample size with the specified accrual time.

Output 77.7.15 Sample Size Summary

The SEQDESIGN Procedure
Design: WhiteheadKeyAlpha

Sample Size Summary

Test	Two-Sample Survival
Null Hazard Rate	0.03466
Hazard Rate (Group A)	0.01733
Hazard Rate (Group B)	0.03466
Hazard Ratio	0.5
log(Hazard Ratio)	-0.69315
Reference Hazards	Alt Ref
Accrual Rate	10
Accrual Time	20
Follow-up Time	6.474376
Total Time	26.47438
Max Number of Events	66.82556
Max Sample Size	200
Expected Sample Size (Null Ref)	161.5941
Expected Sample Size (Alt Ref)	172.4693

The "Number of Events (D) and Sample Sizes (N)" table in Output 77.7.16 displays the required time at each stage, in both fractional and integer numbers. The derived times under the heading "Fractional Time" are not integers. These times are rounded up to integers under the heading "Ceiling Time." The table also displays the numbers of events and sample sizes at each stage.

Output 77.7.16 Number of Events and Sample Sizes

Numbers of Events (D) and Sample Sizes (N)
Two-Sample Log-Rank Test

	Fractional Time						
Stage	D	D(Grp 1)	D(Grp 2)	Time	N	N(Grp 1)	N(Grp 2)
1	16.71	5.82	10.89	11.9867	119.87	59.93	59.93
2	33.41	11.84	21.57	17.3585	173.58	86.79	86.79
3	50.12	18.01	32.11	21.7480	200.00	100.00	100.00
4	66.83	24.46	42.37	26.4744	200.00	100.00	100.00

Numbers of Events (D) and Sample Sizes (N)
Two-Sample Log-Rank Test

	Fractional Time	Ceiling Time					
Stage	Information	D	D(Grp 1)	D(Grp 2)	Time	N	N(Grp 1)
1	4.1766	16.74	5.83	10.91	12	120.00	60.00
2	8.3532	35.73	12.68	23.04	18	180.00	90.00
3	12.5298	51.07	18.37	32.70	22	200.00	100.00
4	16.7064	68.55	25.14	43.41	27	200.00	100.00

Numbers of Events (D) and Sample Sizes (N)
Two-Sample Log-Rank Test

	Ceiling Time	
Stage	N(Grp 2)	Information
1	60.00	4.1854
2	90.00	8.9322
3	100.00	12.7667
4	100.00	17.1378

Example 77.8: Creating a One-Sided Error Spending Design

This example requests a five-stage, one-sided group sequential design for normally distributed statistics. The design uses an O'Brien-Fleming-type error spending function for the α boundary and a Pocock-type error spending function for the β boundary. The following statements request a one-sided design by using different α and β spending functions:

```
ods graphics on;
proc seqdesign altref=0.2  errspend
               pss(cref=0 0.5 1)
               stopprob(cref=0 0.5 1)
```

```
                plots=(asn power errspend)
                ;
   OneSidedErrorSpending: design nstages=5
                          method(alpha)=errfuncobf
                          method(beta)=errfuncpoc
                          alt=upper  stop=both
                          alpha=0.025
                          ;
run;
ods graphics off;
```

The "Design Information" table in Output 77.8.1 displays design specifications and the derived statistics. With the specified alternative reference, the maximum information is derived.

Output 77.8.1 Error Spending Method Design Information

The SEQDESIGN Procedure
Design: OneSidedErrorSpending

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Method	Error Spending
Boundary Key	Both
Alternative Reference	0.2
Number of Stages	5
Alpha	0.025
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	119.4278
Max Information	313.7196
Null Ref ASN (Percent of Fixed Sample)	50.35408
Alt Ref ASN (Percent of Fixed Sample)	78.77223

The "Method Information" table in Output 77.8.2 displays the α and β errors, alternative reference, and derived drift parameter, which is the standardized alternative reference at the final stage.

Output 77.8.2 Method Information

Method Information

Boundary	Method	Alpha	Beta	----Error Spending---- Function
Upper Alpha	Error Spending	0.02500	.	Approx O'Brien-Fleming
Upper Beta	Error Spending	.	0.10000	Approx Pocock

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	0.2	3.542426
Upper Beta	0.2	3.542426

With the STOPPROB option, the "Expected Cumulative Stopping Probabilities" table in Output 77.8.3 displays the expected stopping stage and cumulative stopping probability to reject the null hypothesis at each stage under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and c_i are values specified in the CREF= option.

Output 77.8.3 Stopping Probabilities

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

CRef	Expected Stopping Stage	Source
0.0000	2.108	Reject Null
0.0000	2.108	Accept Null
0.0000	2.108	Total
0.5000	3.296	Reject Null
0.5000	3.296	Accept Null
0.5000	3.296	Total
1.0000	3.298	Reject Null
1.0000	3.298	Accept Null
1.0000	3.298	Total

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

	Stopping Probabilities				
CRef	Stage_1	Stage_2	Stage_3	Stage_4	Stage_5
0.0000	0.00000	0.00039	0.00381	0.01221	0.02500
0.0000	0.38080	0.69133	0.86162	0.94170	0.97500
0.0000	0.38080	0.69173	0.86543	0.95391	1.00000
0.5000	0.00002	0.01265	0.09650	0.24465	0.38724
0.5000	0.13665	0.28063	0.41080	0.52230	0.61276
0.5000	0.13667	0.29328	0.50730	0.76695	1.00000
1.0000	0.00050	0.13209	0.52642	0.80390	0.90000
1.0000	0.02954	0.05231	0.07085	0.08648	0.10000
1.0000	0.03004	0.18440	0.59728	0.89039	1.00000

With the PSS option, the "Power and Expected Sample Sizes" table in Output 77.8.4 displays powers and expected sample sizes under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and $c_i = 0, 0.5, 1, 1.5$ are the default values in the CREF= option.

Output 77.8.4 Power and Expected Sample Size Information

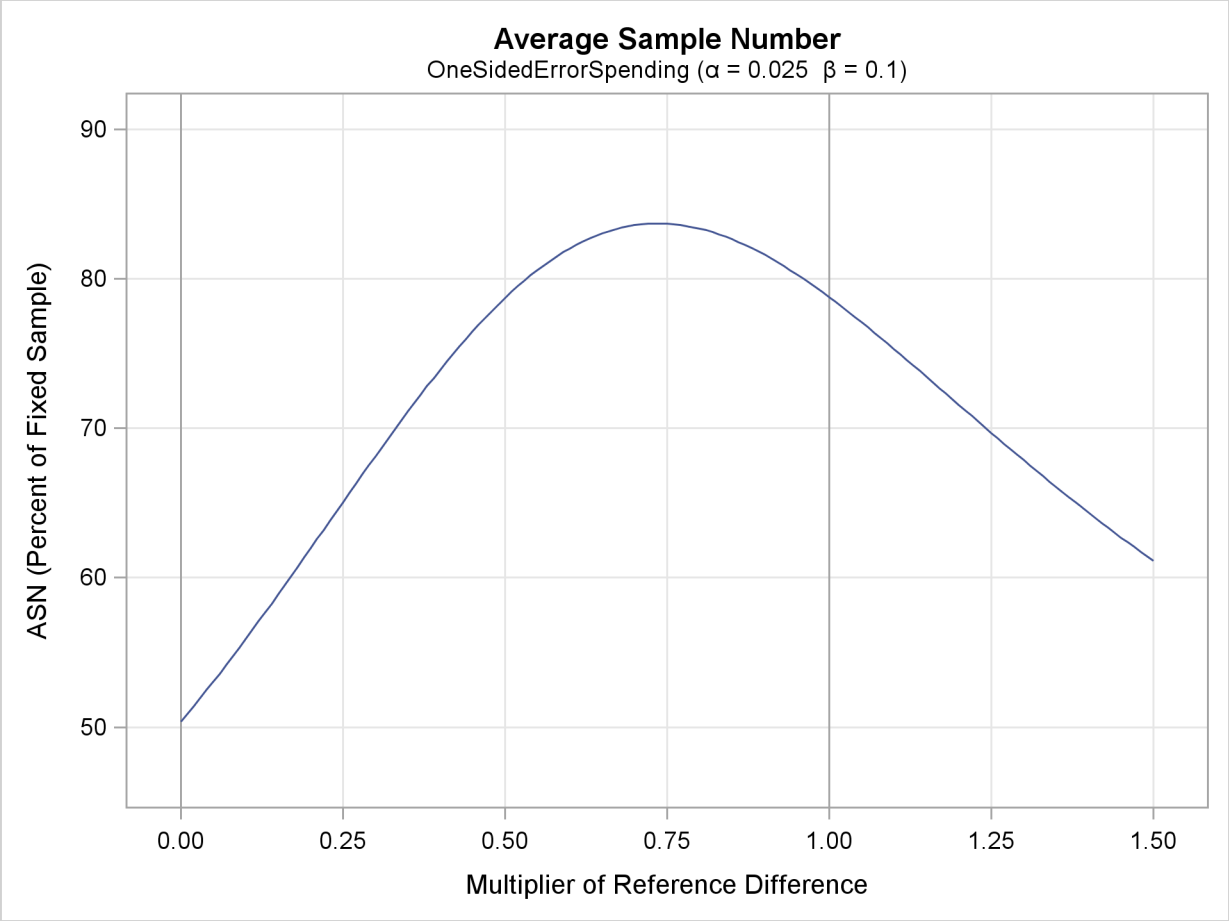
Powers and Expected Sample Sizes
Reference = CRef * (Alt Reference)

CRef	Power	-Sample Size- Percent Fixed-Sample
0.0000	0.02500	50.3541
0.5000	0.38724	78.7219
1.0000	0.90000	78.7722

With the PLOTS=ASN option, the procedure displays a plot of expected sample sizes under various hypothetical references, as shown in Output 77.8.5. By default, expected sample sizes under the

hypotheses $\theta = c_i\,\theta_1$, $c_i = 0, 0.01, 0.02, \ldots, 1.50$, are displayed, where θ_1 is the alternative reference.

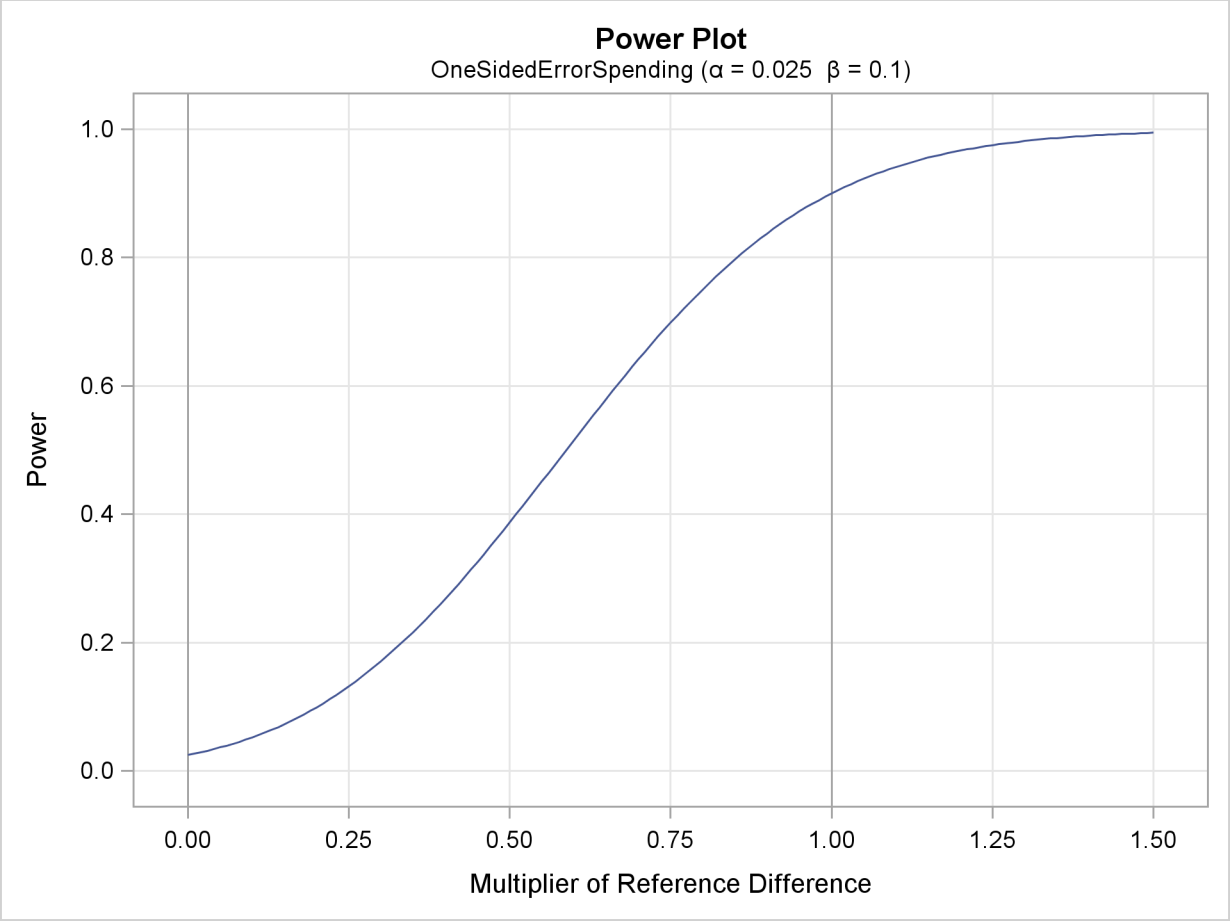
Output 77.8.5 ASN Plot



With the PLOTS=POWER option, the procedure displays a plot of the power curves under various hypothetical references for all designs simultaneously, as shown in Output 77.8.6. By default, the option CREF= $0, 0.01, 0.02, \ldots, 1.50$ and powers under hypothetical references $\theta = c_i\,\theta_1$ are displayed, where c_i are values specified in the CREF= option. These CREF= values are displayed on the horizontal axis.

Under the null hypothesis, $c_i = 0$, the power is 0.025, the upper Type I error probability. Under the alternative hypothesis, $c_i = 1$, the power is 0.9, one minus the Type II error probability. The plot shows only minor difference between the two designs.

Output 77.8.6 Power Plot



The "Boundary Information" table in Output 77.8.7 displays information level, alternative reference, and boundary values. By default, the alternative reference and boundary values are displayed with the default standardized Z scale. That is, the resulting standardized alternative reference at stage k is given by $\theta_1\sqrt{I_k}$, where θ_1 is the specified alternative reference and I_k is the information level at stage k, $k = 1, 2, \ldots, 5$.

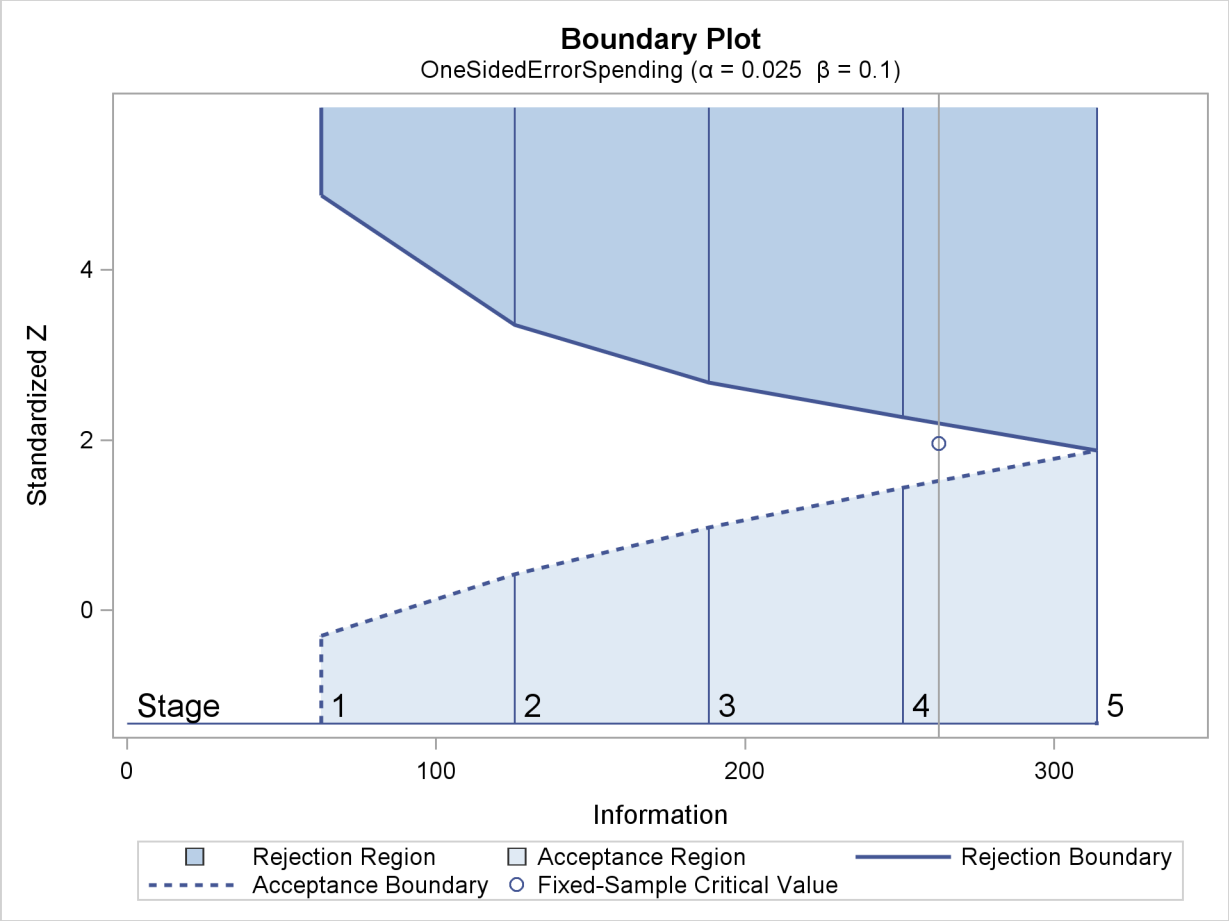
Output 77.8.7 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

	---Information Level--		-Alternative- --Reference--	-----Boundary Values---- ----------Upper---------	
Stage	Proportion	Actual	Upper	Beta	Alpha
1	0.2000	62.74393	1.58422	-0.30338	4.87688
2	0.4000	125.4879	2.24043	0.41667	3.35706
3	0.6000	188.2318	2.74395	0.97165	2.67766
4	0.8000	250.9757	3.16844	1.43627	2.26535
5	1.0000	313.7196	3.54243	1.87522	1.87522

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.8.8. This plot displays the boundary values in the "Boundary Information" table in Output 77.8.7.

Output 77.8.8 Boundary Plot



The "Error Spending Information" table in Output 77.8.9 displays cumulative error spending at each stage for each boundary.

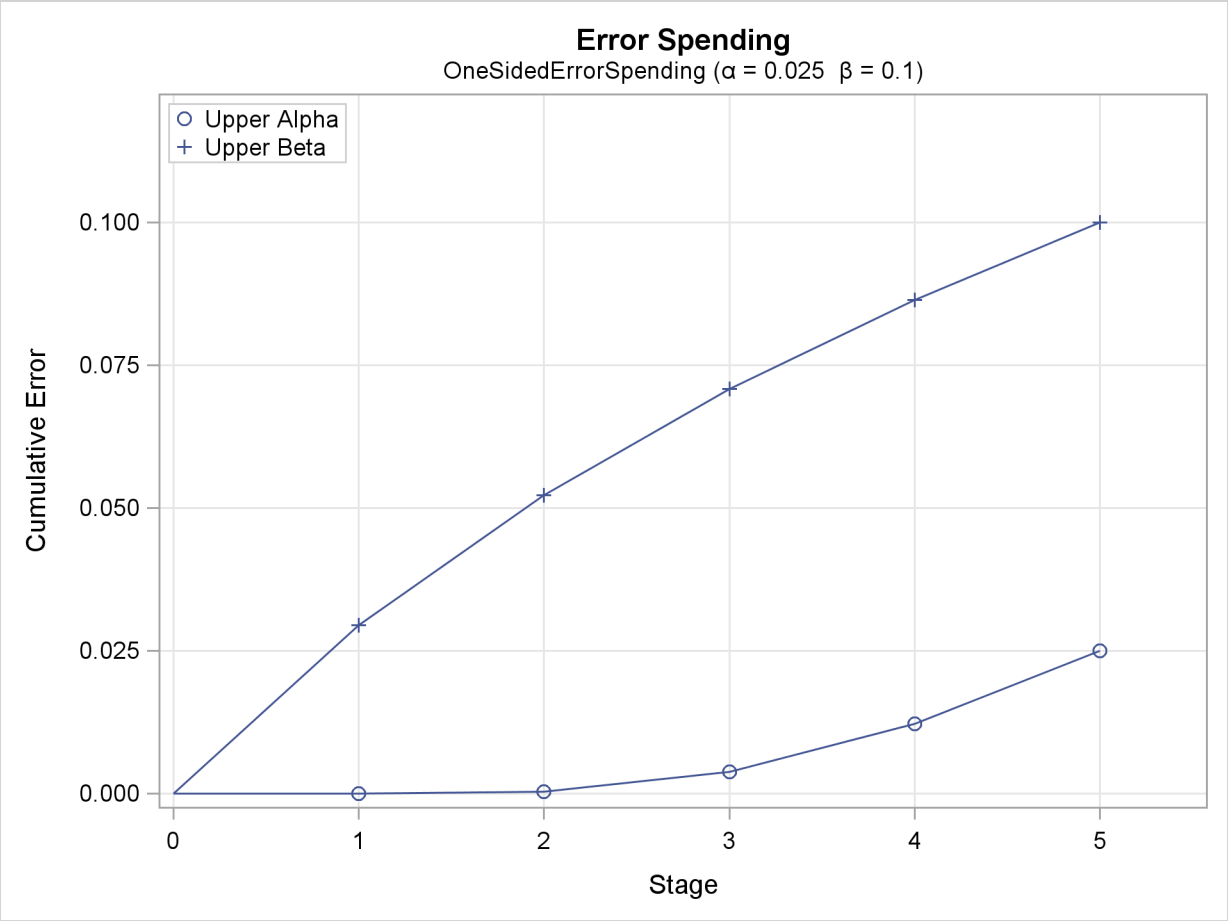
Output 77.8.9 Error Spending Information

Error Spending Information

Stage	-Information Level- Proportion	-Cumulative Error Spending- ----------Upper---------- Beta	Alpha
1	0.2000	0.02954	0.00000
2	0.4000	0.05231	0.00039
3	0.6000	0.07085	0.00381
4	0.8000	0.08648	0.01221
5	1.0000	0.10000	0.02500

With the PLOTS=ERRSPEND option, the procedure displays a plot of error spending for each boundary, as shown in Output 77.8.10. This plot displays the cumulative error spending at each stage in the "Error Spending Information" table in Output 77.8.9. The O'Brien-Fleming-type α spending function is conservative in early stages because it uses much less at early stages than in the later stages. In contrast, the Pocock-type β spending function uses more at early stages than in the later stages.

Output 77.8.10 Error Spending Plot



Example 77.9: Creating Designs with Various Number of Stages

This example requests three group sequential designs for normally distributed statistics. Each design uses the power family error spending function with the default power parameter $\rho = 2$. The specified error spending method is between the approximated Pocock method ($\rho = 1$) and the approximated O'Brien-Fleming method ($\rho = 3$) (Jennison and Turnbull 1999, p. 148). The three designs are identical except for the specified number of stages. The following statements request the group sequential design with the default standardized Z scale for the boundary values:

```
ods graphics on;
proc seqdesign plots=( asn
                       power
                       combinedboundary
                       errspend(hscale=info)
                       )
                ;
   TwoStageDesign:  design nstages=2
                    method=errfuncpow
                    alt=upper  stop=reject
                    ;
   FiveStageDesign: design nstages=5
                    method=errfuncpow
                    alt=upper  stop=reject
                    ;
   TenStageDesign:  design nstages=10
                    method=errfuncpow
                    alt=upper  stop=reject
                    ;
run;
ods graphics off;
```

The "Design Information" table in Output 77.9.1 displays design information for the two-stage design.

Output 77.9.1 Design Information

The SEQDESIGN Procedure
Design: TwoStageDesign

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Early Stop	Reject Null
Method	Error Spending
Boundary Key	Both
Number of Stages	2
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	102.4167
Null Ref ASN (Percent of Fixed Sample)	101.7766
Alt Ref ASN (Percent of Fixed Sample)	79.81021

The "Boundary Information" table in Output 77.9.2 displays the information level, alternative reference, and boundary values with the default standardized normal Z scale. The resulting standardized alternative reference at stage k is given by $\theta_1\sqrt{I_k}$, where θ_1 is the alternative reference and I_k is the information level at stage k, $k = 1, 2$.

Output 77.9.2 Boundary Information in *Z* Scale

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	-Information Level- Proportion	-Alternative- --Reference-- Upper	-Boundary Values- ------Upper------ Alpha
1	0.5000	2.09414	2.24140
2	1.0000	2.96156	1.69970

The "Design Information" table in Output 77.9.3 displays design information for the five-stage design. Compared with the two-stage design in Output 77.9.1, the maximum information increases from 102.42 to 105.62, and the average sample number under the alternative reference (Alt Ref ASN) decreases from 79.81 to 69.64.

Output 77.9.3 Design Information

The SEQDESIGN Procedure
Design: FiveStageDesign

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Early Stop	Reject Null
Method	Error Spending
Boundary Key	Both
Number of Stages	5
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	105.6235
Null Ref ASN (Percent of Fixed Sample)	104.356
Alt Ref ASN (Percent of Fixed Sample)	69.64322

The "Boundary Information" table in Output 77.9.4 displays the information level, alternative reference, and boundary values with the default standardized normal Z scale.

Output 77.9.4 Boundary Information in Z Scale

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	-Information Level- Proportion	-Alternative- --Reference-- Upper	-Boundary Values- ------Upper------ Alpha
1	0.2000	1.34502	2.87816
2	0.4000	1.90215	2.47023
3	0.6000	2.32965	2.20095
4	0.8000	2.69005	1.98182
5	1.0000	3.00756	1.79024

The "Design Information" table in Output 77.9.5 displays design information for the ten-stage design. Compared with the five-stage design in Output 77.9.3, the maximum information increases further from 105.62 to 107.26 and under the alternative reference, the average sample number decreases further from 69.64 to 66.36.

Output 77.9.5 Design Information

The SEQDESIGN Procedure
Design: TenStageDesign

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Upper
Early Stop	Reject Null
Method	Error Spending
Boundary Key	Both
Number of Stages	10
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	107.256
Null Ref ASN (Percent of Fixed Sample)	105.7276
Alt Ref ASN (Percent of Fixed Sample)	66.35565

The "Boundary Information" table in Output 77.9.6 displays the information level, alternative reference, and boundary values with the default standardized normal Z scale.

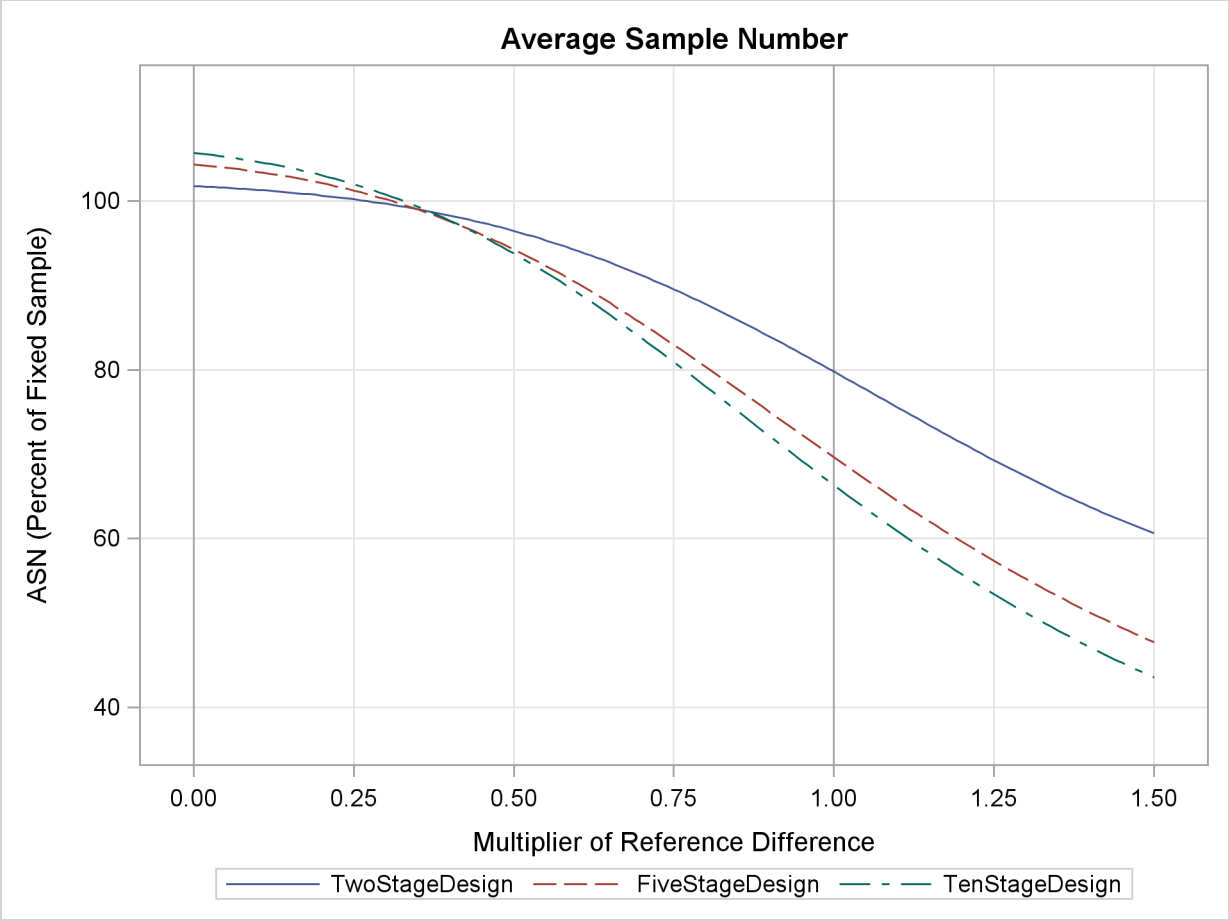
Output 77.9.6 Boundary Information in *Z* Scale

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	-Information Level- Proportion	-Alternative- --Reference-- Upper	-Boundary Values- ------Upper------ Alpha
1	0.1000	0.95840	3.29053
2	0.2000	1.35538	2.94037
3	0.3000	1.65999	2.72115
4	0.4000	1.91679	2.54808
5	0.5000	2.14304	2.40114
6	0.6000	2.34758	2.27127
7	0.7000	2.53568	2.15359
8	0.8000	2.71076	2.04503
9	0.9000	2.87519	1.94355
10	1.0000	3.03072	1.84765

With the PLOTS=ASN option, the procedure displays a plot of average sample numbers under various hypothetical references for all designs simultaneously, as shown in Output 77.9.7. By default, the option CREF= $0, 0.01, 0.02, \ldots, 1.50$ and expected sample sizes under the hypothetical references $\theta = c_i\,\theta_1$ are displayed, where c_i are values specified in the CREF= option. These CREF= values are displayed on the horizontal axis.

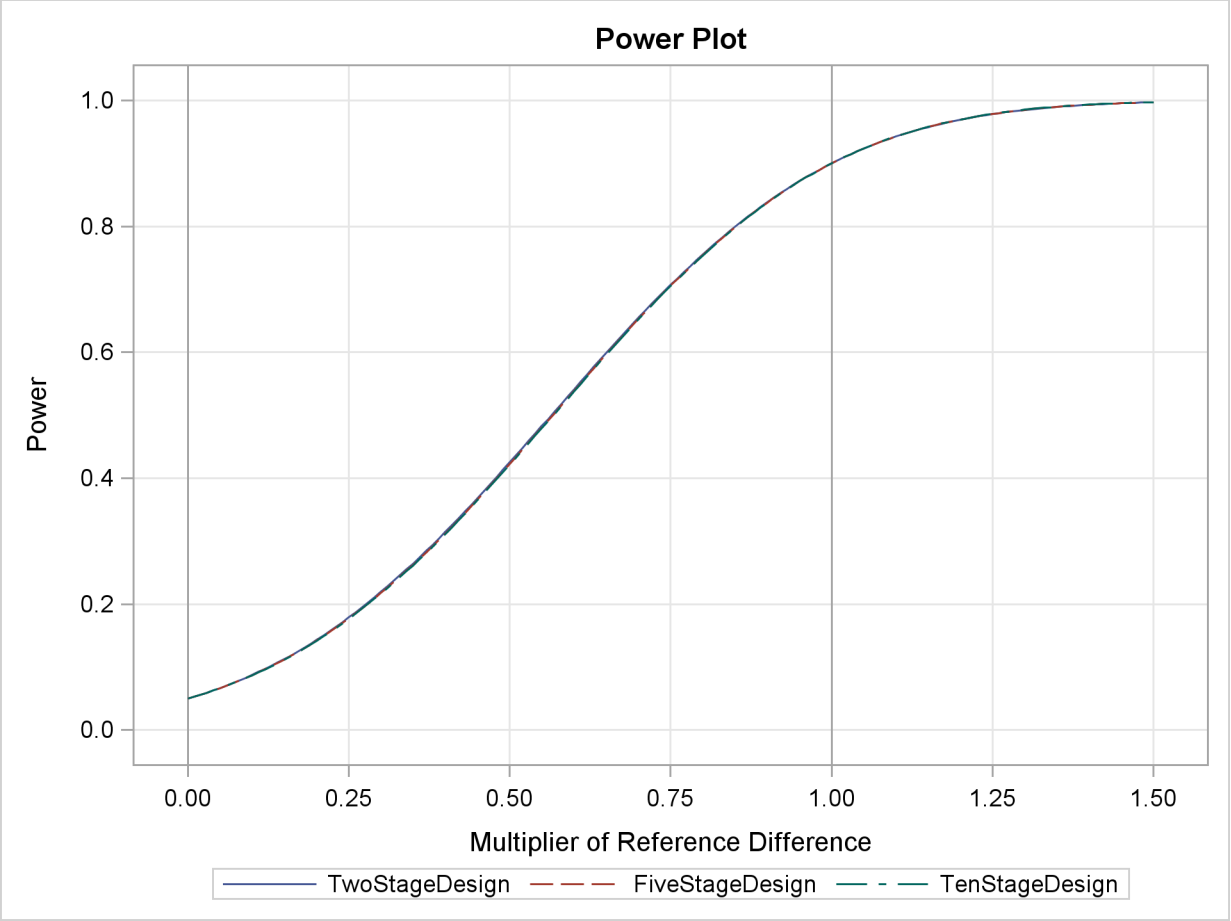
Output 77.9.7 ASN Plot



The plot shows that as the number of stages increases, the average sample number as a percentage of the fixed-sample design increases under the null hypothesis ($c_i = 0$) but decreases under the alternative hypothesis ($c_i = 1$).

With the PLOTS=POWER option, the procedure displays a plot of the power curves under various hypothetical references for all designs simultaneously, as shown in Output 77.9.8. By default, the option CREF= $0, 0.01, 0.02, \ldots, 1.50$ and powers under hypothetical references $\theta = c_i\,\theta_1$ are displayed, where c_i are values specified in the CREF= option. These CREF= values are displayed on the horizontal axis.

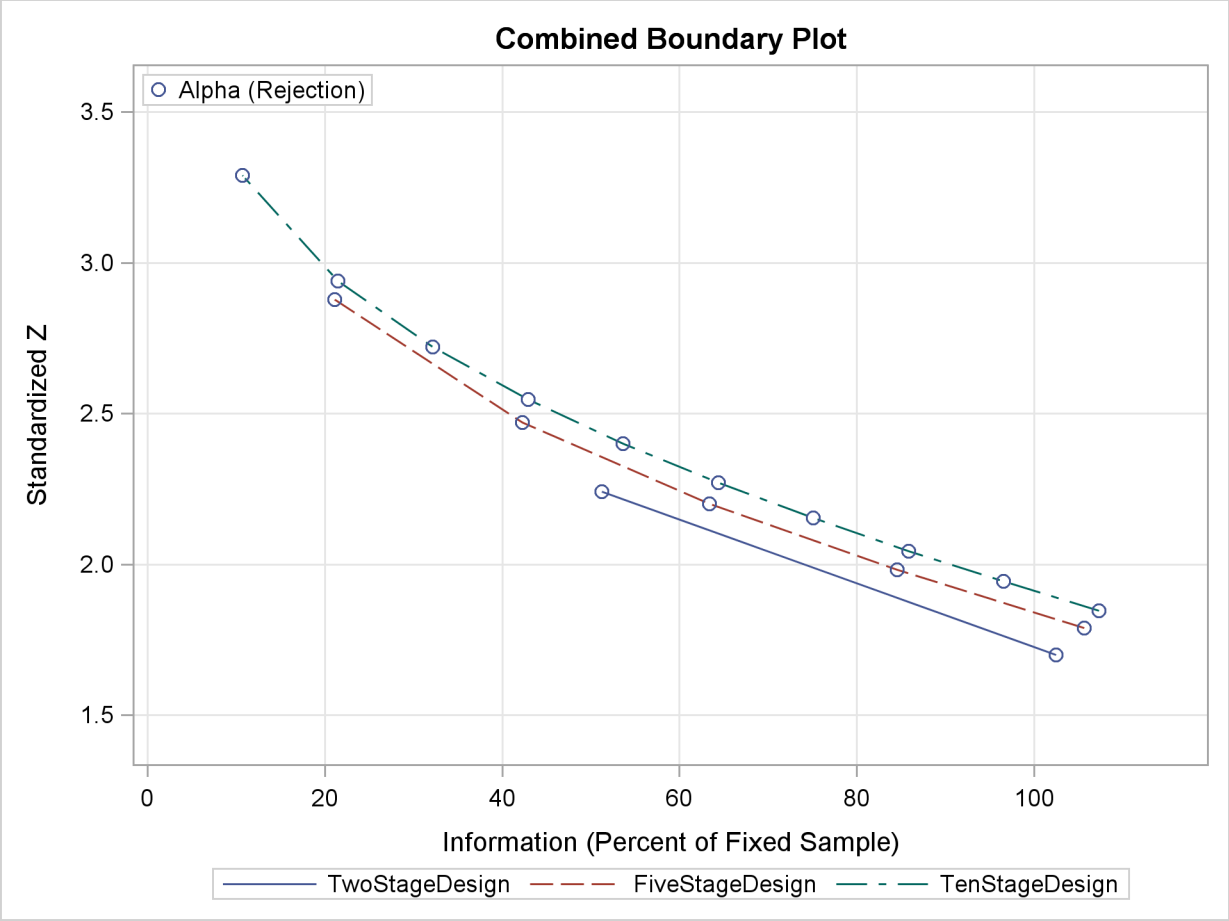
Output 77.9.8 Power Plot



Under the null hypothesis, $c_i = 0$, the power is 0.05, the upper Type I error probability. Under the alternative hypothesis, $c_i = 1$, the power is 0.9, one minus the Type II error probability. The plot shows only minor difference among the three designs.

With the PLOTS=COMBINEDBOUNDARY option, the procedure displays a plot of sequential boundaries for all designs simultaneously, as shown in Output 77.9.9. By default, the information levels are used on the horizontal axis. Since the maximum information is not available for the design, the percent information ratios with respect to the corresponding fixed-sample design are displayed in the plot.

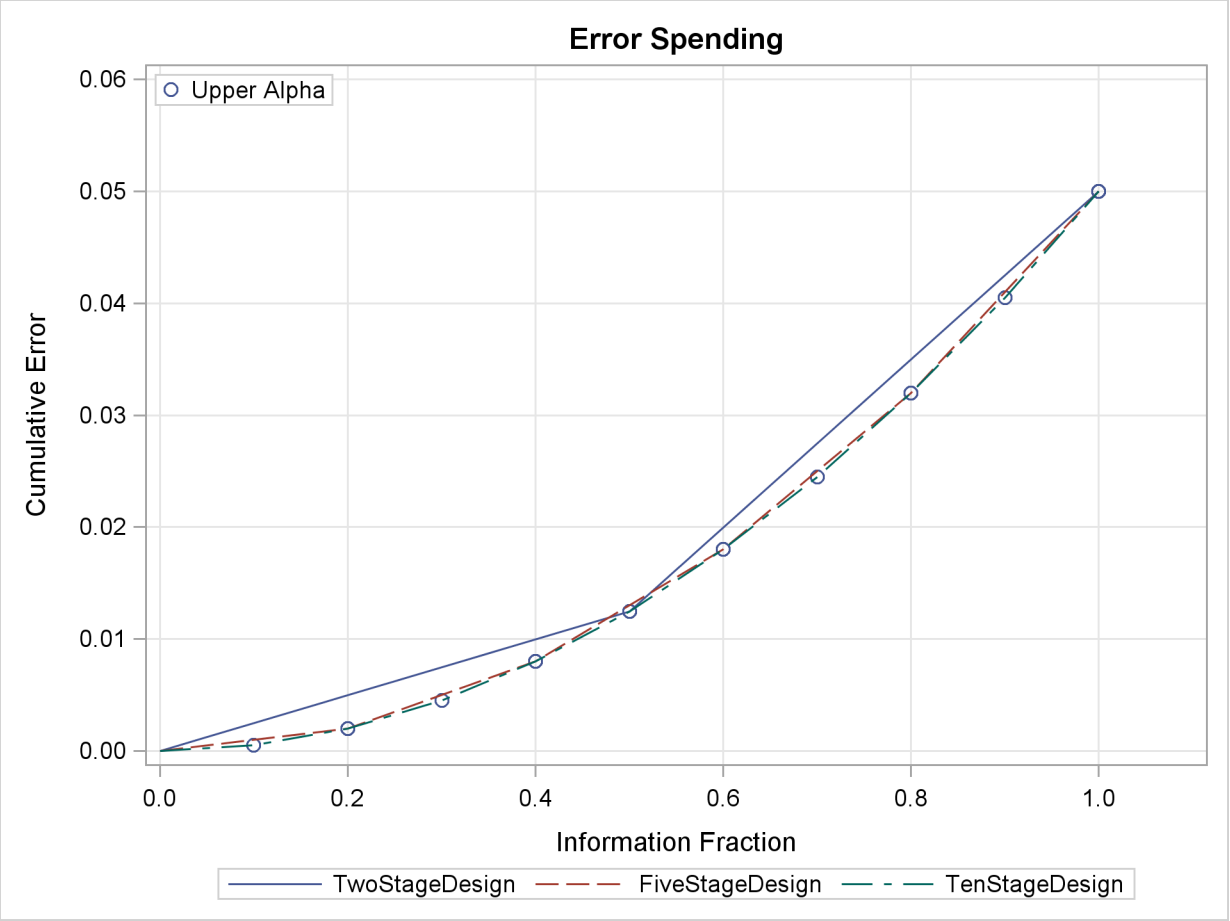
Output 77.9.9 Combined Boundary Plot



The plot shows that as the number of stages increases, the maximum information increases and the α boundary values also increase.

With the PLOTS=ERRSPEND(HSCALE=INFO) option, the procedure displays a plot of cumulative error spends for all boundaries in the designs simultaneously, as shown in Output 77.9.10.

Output 77.9.10 Error Spending Plot



The plot shows similar error spending for these three designs since all three designs are generated from the same power family error spending function.

Example 77.10: Creating Two-Sided Error Spending Designs with and without Overlapping Lower and Upper β Boundaries

This example requests two three-stage group sequential designs for normally distributed statistics. Each design uses a power family error spending function with a specified two-sided alternative hypothesis $H_1 : \theta_1 = \pm 0.2$ and early stopping only to accept the null hypothesis H_0.

The first design uses the BETAOVERLAP=NOADJUST option to derive acceptance boundary values without adjusting for the possible overlapping of the lower and upper β boundaries computed from the two corresponding one-sided tests. The second design uses the BETAOVERLAP=ADJUST option to test the overlapping of the β boundaries at each interim stage based on the two corresponding one-sided tests and then to set the β boundary values at the stage to missing if overlapping occurs at that stage.

The following statements request a two-sided design with the BETAOVERLAP=NOADJUST option:

```
ods graphics on;
proc seqdesign altref=0.2 errspend;
   design nstages=3
          method=errfuncpow
          alt=twosided  stop=accept
          betaoverlap=noadjust
          beta=0.09
          ;
run;
ods graphics off;
```

The "Design Information" table in Output 77.10.1 displays design specifications and the derived statistics for the first design. With the specified alternative reference $\theta_1 = 0.2$, the maximum information is derived.

Output 77.10.1 Design Information

The SEQDESIGN Procedure
Design: Design_1

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Accept Null
Method	Error Spending
Boundary Key	Both
Alternative Reference	0.2
Number of Stages	3
Alpha	0.05
Beta	0.09
Power	0.91
Max Information (Percent of Fixed Sample)	103.8789
Max Information	282.9328
Null Ref ASN (Percent of Fixed Sample)	79.20197
Alt Ref ASN (Percent of Fixed Sample)	102.1476

The "Boundary Information" table in Output 77.10.2 displays the information level, alternative reference, and boundary values. With a specified alternative reference θ_1, the maximum information is derived from the procedure, and the actual information level at each stage is displayed in the table. With the default BOUNDARYSCALE=STDZ option, the alternative reference in the standardized Z scale at stage k is given by $\theta_1\sqrt{I_k}$, where θ_1 is the specified alternative reference and I_k is the information level at stage k, $k = 1, 2, 3$.

Output 77.10.2 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Information Level- Proportion	Information Level- Actual	Alternative Reference Lower	Alternative Reference Upper	Boundary Values Lower Beta	Boundary Values Upper Beta
1	0.3333	94.31094	-1.94228	1.94228	-0.08239	0.08239
2	0.6667	188.6219	-2.74679	2.74679	-0.90351	0.90351
3	1.0000	282.9328	-3.36412	3.36412	-1.92519	1.92519

The "Error Spending Information" table in Output 77.10.3 displays the cumulative error spending at each stage for each boundary.

Output 77.10.3 Error Spending Information

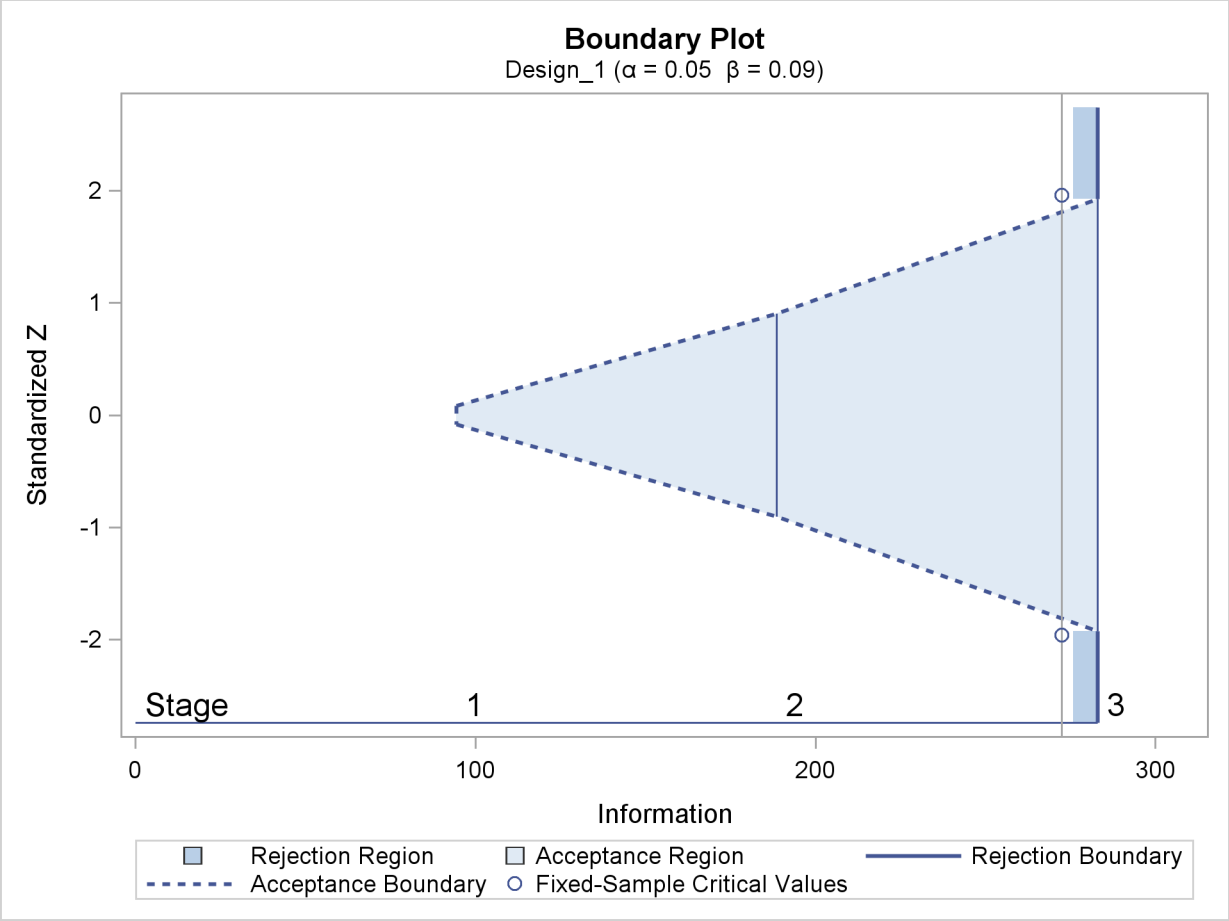
Error Spending Information

Stage	Information Level- Proportion	Cumulative Error Spending Lower Alpha	Cumulative Error Spending Lower Beta	Cumulative Error Spending Upper Beta	Cumulative Error Spending Upper Alpha
1	0.3333	0.00000	0.01000	0.01000	0.00000
2	0.6667	0.00000	0.04000	0.04000	0.00000
3	1.0000	0.02500	0.09000	0.09000	0.02500

With the STOP=ACCEPT option, the design does not stop at interim stages to reject H_0, and the α spending at each interim stage is zero. For the power family error spending function with the default parameter $\rho = 2$, the beta spending at stage 1 is $(1/3)^\rho \beta = (1/3)^2 0.09 = 0.01$, and the cumulative beta spending at stage 2 is $(2/3)^\rho \beta = (2/3)^2 0.09 = 0.04$.

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the acceptance and rejection regions is displayed by default, as shown in Output 77.10.4.

Output 77.10.4 Boundary Plot



The following statements request a two-sided design with the BETAOVERLAP=ADJUST option, which is the default:

```
ods graphics on;
proc seqdesign altref=0.2 errspend;
   design nstages=3
          method=errfuncpow
          alt=twosided
          stop=accept
          betaoverlap=adjust
          beta=0.09
          ;
run;
ods graphics off;
```

With the BETAOVERLAP=ADJUST option, the procedure first derives the usual β boundary values for the two-sided design and then checks for overlapping of the β boundaries for the two corresponding one-sided tests at each stage. If this type of overlapping occurs at a particular stage, the β boundary values for that stage are set to missing, the β spending values at that stage are reset to zero, and the β spending values at subsequent stages are adjusted proportionally.

The boundary values without adjusting for the possible overlapping of the two one-sided β boundaries are identical to the boundary values derived in the first design (with the BETAOVERLAP=NOADJUST option, as shown in Output 77.10.2). At stage 1, the upper β boundary value for the corresponding one-sided test is

$$\theta_1\sqrt{I_1} - \Phi^{-1}(1-\beta_1) = 0.2\sqrt{94.31094} - \Phi^{-1}(0.99) = 1.94228 - 2.32635 = -0.38407$$

where $\theta_1 = 0.2$ is the upper alternative reference, $I_1 = 94.31094$ is the information level at stage 1, and $\beta_1 = 0.01$ is the β spending at stage 1 (as shown in Output 77.10.3).

Similarly, the lower β boundary value for the corresponding one-sided test is computed as 0.38407. Since the upper β boundary value is less than the lower β boundary at stage 1, overlapping occurs, and so the β boundary values for the two-sided design are set to missing at stage 1.

With the β boundary values set to missing at stage 1 and the β spending $\beta'_1 = 0$ the β spending values at subsequent interim stages are adjusted proportionally. In this example, the adjusted β spending at stage 2 is computed as

$$\beta'_2 = \beta'_1 + \frac{\beta_2 - \beta_1}{\beta_3 - \beta_1}(\beta_3 - \beta'_1) = 0 + \frac{0.04 - 0.01}{0.09 - 0.01}\,0.09 = 0.03375$$

where β_k is the cumulative β spending at stage k before the adjustment, $k = 1, 2, 3$.

The "Design Information" table in Output 77.10.5 displays design specifications and derived statistics for the design.

Output 77.10.5 Design Information

The SEQDESIGN Procedure
Design: Design_1

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Accept Null
Method	Error Spending
Boundary Key	Both
Alternative Reference	0.2
Number of Stages	3
Alpha	0.05
Beta	0.09
Power	0.91
Max Information (Percent of Fixed Sample)	101.9388
Max Information	277.649
Null Ref ASN (Percent of Fixed Sample)	80.56408
Alt Ref ASN (Percent of Fixed Sample)	100.792

The "Boundary Information" table in Output 77.10.6 displays the information levels, alternative references, and boundary values.

Output 77.10.6 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Information Level- Proportion	Information Level- Actual	Alternative- Reference- Lower	Alternative- Reference- Upper	Boundary Values- Lower- Beta	Boundary Values- Upper- Beta
1	0.3333	92.54967	-1.92405	1.92405	.	.
2	0.6667	185.0993	-2.72102	2.72102	-0.89469	0.89469
3	1.0000	277.649	-3.33256	3.33256	-1.93494	1.93494

The "Error Spending Information" table in Output 77.10.7 displays the cumulative error spending at each stage for each boundary.

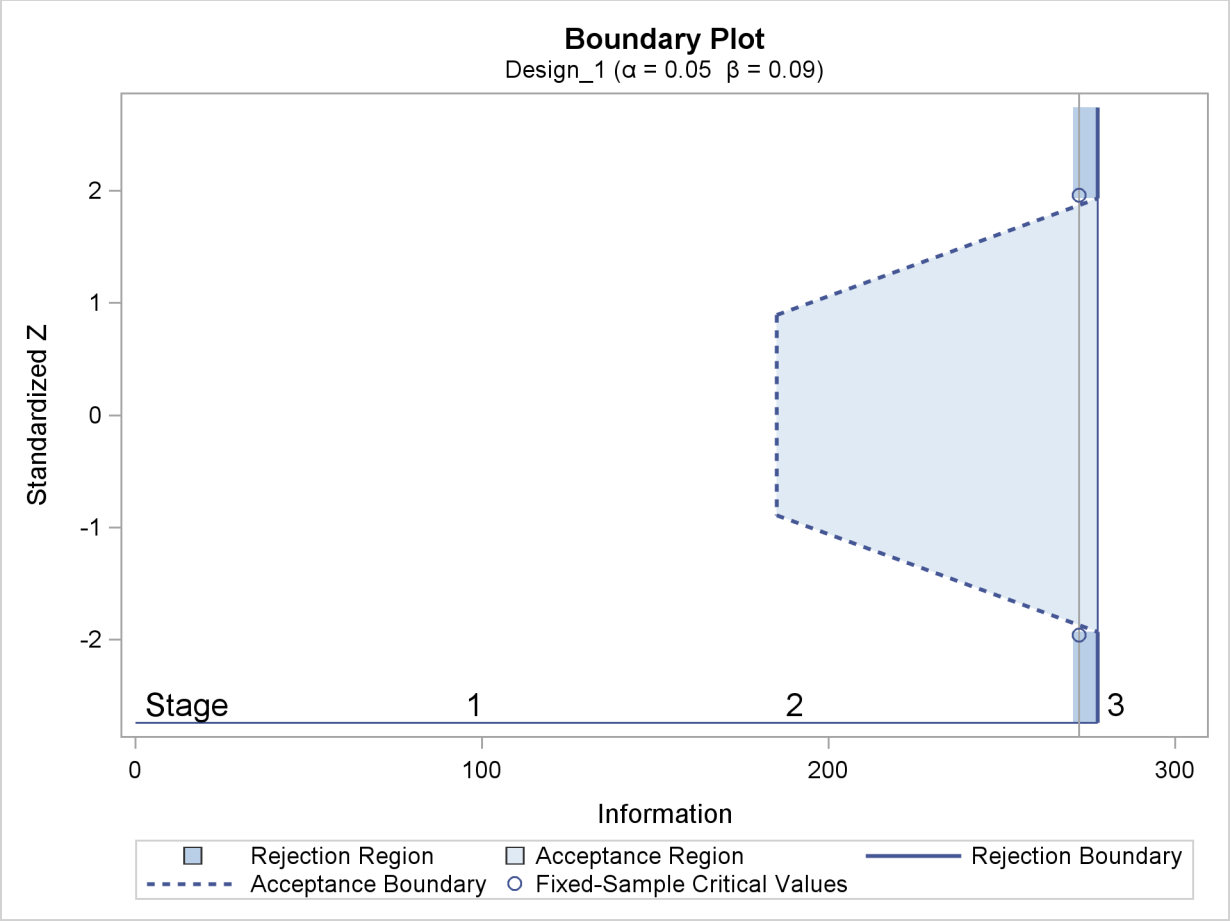
Output 77.10.7 Error Spending Information

Error Spending Information

Stage	Information Level- Proportion	Cumulative Error Spending- Lower- Alpha	Cumulative Error Spending- Lower- Beta	Cumulative Error Spending- Upper- Beta	Cumulative Error Spending- Upper- Alpha
1	0.3333	0.00000	0.00000	0.00000	0.00000
2	0.6667	0.00000	0.03375	0.03375	0.00000
3	1.0000	0.02500	0.09000	0.09000	0.02500

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the acceptance and rejection regions is displayed by default, as shown in Output 77.10.8.

Output 77.10.8 Boundary Plot



Example 77.11: Creating a Two-Sided Asymmetric Error Spending Design with Early Stopping to Reject H_0

This example requests a three-stage two-sided asymmetric group sequential design for normally distributed statistics.

The O'Brien-Fleming boundary can be approximated using a power family error spending function with parameter $\rho = 3$, and the Pocock boundary can be approximated using a power family error spending function with parameter $\rho = 1$ (Jennison and Turnbull 2000, p. 148). The following statements use the power family error spending function to creates a two-sided asymmetric design with early stopping to reject the null hypothesis H_0:

```
ods graphics on;
proc seqdesign altref=1.0
               pss(cref=0 0.5 1)
               stopprob(cref=0 0.5 1)
               errspend
               plots=(asn power errspend)
               ;
```

```
   TwoSidedErrorSpending: design nstages=3
                          method(upperalpha)=errfuncpow(rho=3)
                          method(loweralpha)=errfuncpow(rho=1)
                          info=cum(2 3 4)
                          alt=twosided
                          stop=reject
                          alpha=0.075(upper=0.025)
                          ;
run;
ods graphics off;
```

The design uses power family error spending functions with $\rho = 1$ for the lower α boundary and $\rho = 3$ for the upper α boundary. Thus, the design is conservative in the early stages and tends to stop the trials early only with a small p-value for the upper α boundary. The upper α level 0.025 is specified explicitly, and the lower α level is computed as $0.075 - 0.025 = 0.05$.

The "Design Information" table in Output 77.11.1 displays design specifications and the derived maximum information. Note that in order to attain the same information level for the asymmetric lower and upper boundaries, the derived power at the lower alternative 0.92963 is larger than the default 0.90.

Output 77.11.1 Design Information

The SEQDESIGN Procedure
Design: TwoSidedErrorSpending

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Method	Error Spending
Boundary Key	Both
Alternative Reference	1
Number of Stages	3
Alpha	0.075
Alpha (Lower)	0.05
Alpha (Upper)	0.025
Beta (Lower)	0.07037
Beta (Upper)	0.1
Power (Lower)	0.92963
Power (Upper)	0.9
Max Information (Percent of Fixed Sample)	102.4384
Max Information	10.76365
Null Ref ASN (Percent of Fixed Sample)	100.4877
Lower Alt Ref ASN (Percent of Fixed Sample)	64.8288
Upper Alt Ref ASN (Percent of Fixed Sample)	75.98778

The "Method Information" table in Output 77.11.2 displays the specified α and β error levels and the derived drift parameter. With the same information level used for the asymmetric lower and upper boundaries, only one of the β levels is maintained, and the other is derived to have the level less than or equal to the default level.

Output 77.11.2 Method Information

Method Information

Boundary	Method	Alpha	Beta	----Error Spending---- Function
Upper Alpha	Error Spending	0.02500	0.10000	Power (Rho=3)
Lower Alpha	Error Spending	0.05000	0.07037	Power (Rho=1)

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	1	3.280801
Lower Alpha	-1	-3.2808

With the STOPPROB(CREF=0 0.5 1) option, the "Expected Cumulative Stopping Probabilities" table in Output 77.11.3 displays the expected stopping stage and cumulative stopping probability to reject the null hypothesis H_0 at each stage under hypothetical references $\theta = 0$ (null hypothesis H_0), $\theta = 0.5\,\theta_1$, and $\theta = \theta_1$ (alternative hypothesis H_1), where θ_1 is the alternative reference.

Output 77.11.3 Stopping Probabilities

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

CRef	Ref	Expected Stopping Stage	Source
0.0000	Lower Alt	2.924	Rej Null (Lower Alt)
0.0000	Lower Alt	2.924	Rej Null (Upper Alt)
0.0000	Lower Alt	2.924	Reject Null
0.5000	Lower Alt	2.456	Rej Null (Lower Alt)
0.5000	Lower Alt	2.456	Rej Null (Upper Alt)
0.5000	Lower Alt	2.456	Reject Null
1.0000	Lower Alt	1.531	Rej Null (Lower Alt)
1.0000	Lower Alt	1.531	Rej Null (Upper Alt)
1.0000	Lower Alt	1.531	Reject Null
0.0000	Upper Alt	2.924	Rej Null (Lower Alt)
0.0000	Upper Alt	2.924	Rej Null (Upper Alt)
0.0000	Upper Alt	2.924	Reject Null
0.5000	Upper Alt	2.758	Rej Null (Lower Alt)
0.5000	Upper Alt	2.758	Rej Null (Upper Alt)
0.5000	Upper Alt	2.758	Reject Null
1.0000	Upper Alt	1.967	Rej Null (Lower Alt)
1.0000	Upper Alt	1.967	Rej Null (Upper Alt)
1.0000	Upper Alt	1.967	Reject Null

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

		----Stopping Probabilities---		
CRef	Ref	Stage_1	Stage_2	Stage_3
0.0000	Lower Alt	0.02500	0.03750	0.05000
0.0000	Lower Alt	0.00313	0.01055	0.02500
0.0000	Lower Alt	0.02813	0.04805	0.07500
0.5000	Lower Alt	0.21185	0.33190	0.45370
0.5000	Lower Alt	0.00005	0.00012	0.00021
0.5000	Lower Alt	0.21190	0.33202	0.45391
1.0000	Lower Alt	0.64054	0.82803	0.92963
1.0000	Lower Alt	0.00000	0.00000	0.00000
1.0000	Lower Alt	0.64054	0.82803	0.92963
0.0000	Upper Alt	0.02500	0.03750	0.05000
0.0000	Upper Alt	0.00313	0.01055	0.02500
0.0000	Upper Alt	0.02813	0.04805	0.07500
0.5000	Upper Alt	0.00090	0.00110	0.00120
0.5000	Upper Alt	0.05769	0.18269	0.36458
0.5000	Upper Alt	0.05860	0.18379	0.36578
1.0000	Upper Alt	0.00001	0.00001	0.00001
1.0000	Upper Alt	0.33926	0.69356	0.90000
1.0000	Upper Alt	0.33927	0.69357	0.90001

"Rej Null (Lower Alt)" and "Rej Null (Upper Alt)" under the heading "Source" indicate the probabilities of rejecting the null hypothesis for the lower alternative and for the upper alternative, respectively. "Reject Null" indicates the probability of rejecting the null hypothesis for either the lower or upper alternative.

Note that with the STOP=REJECT option, the cumulative stopping probability of accepting the null hypothesis H_0 at each interim stage is zero and is not displayed.

With the PSS(CREF=0 0.5 1.0) option, the "Power and Expected Sample Sizes" table in Output 77.11.4 displays powers and expected sample sizes under hypothetical references $\theta = 0$ (null hypothesis H_0), $\theta = 0.5\,\theta_1$, and $\theta = \theta_1$ (alternative hypothesis H_1), where θ_1 is the alternative reference. The expected sample sizes are displayed in a percentage scale relative to the corresponding fixed-sample size design.

Output 77.11.4 Power and Expected Sample Size Information

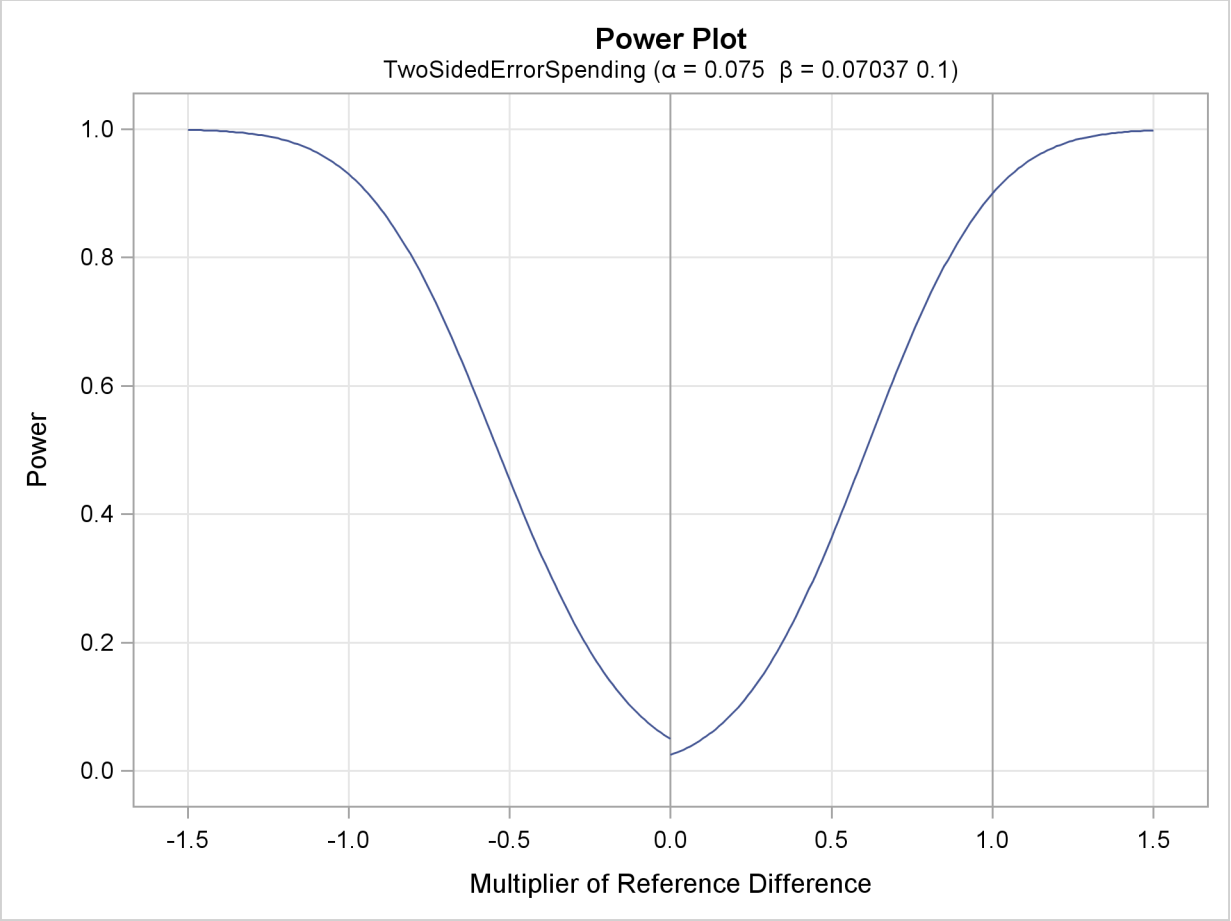
Powers and Expected Sample Sizes
Reference = CRef * (Alt Reference)

CRef	Ref	Power	-Sample Size- Percent Fixed-Sample
0.0000	Lower Alt	0.05000	100.4877
0.5000	Lower Alt	0.45370	88.5090
1.0000	Lower Alt	0.92963	64.8288
0.0000	Upper Alt	0.02500	100.4877
0.5000	Upper Alt	0.36458	96.2309
1.0000	Upper Alt	0.90000	75.9878

Note that at $c_i = 0$, the null reference $\theta = 0$, the power with the lower alternative is the lower α error 0.05, and the power with the upper alternative is the upper α error 0.025. At $c_i = 1$, the alternative reference $\theta = \theta_1$, the power with the upper alternative is the specified power 0.90, and the power with the lower alternative 0.92963 is greater than the specified power 0.90 because the same information level is used for these two asymmetric boundaries.

With the PLOTS=POWER option, the procedure displays a plot of the power curves under various hypothetical references, as shown in Output 77.11.5. By default, powers under the lower hypotheses $\theta = c_i\,\theta_{1l}$ and under the upper hypotheses $\theta = c_i\,\theta_{1u}$ are displayed for a two-sided asymmetric design, where $c_i = 0, 0.01, 0.02, \ldots, 1.50$ and $\theta_{1l} = -1$ and $\theta_{1u} = 1$ are the lower and upper alternative references, respectively.

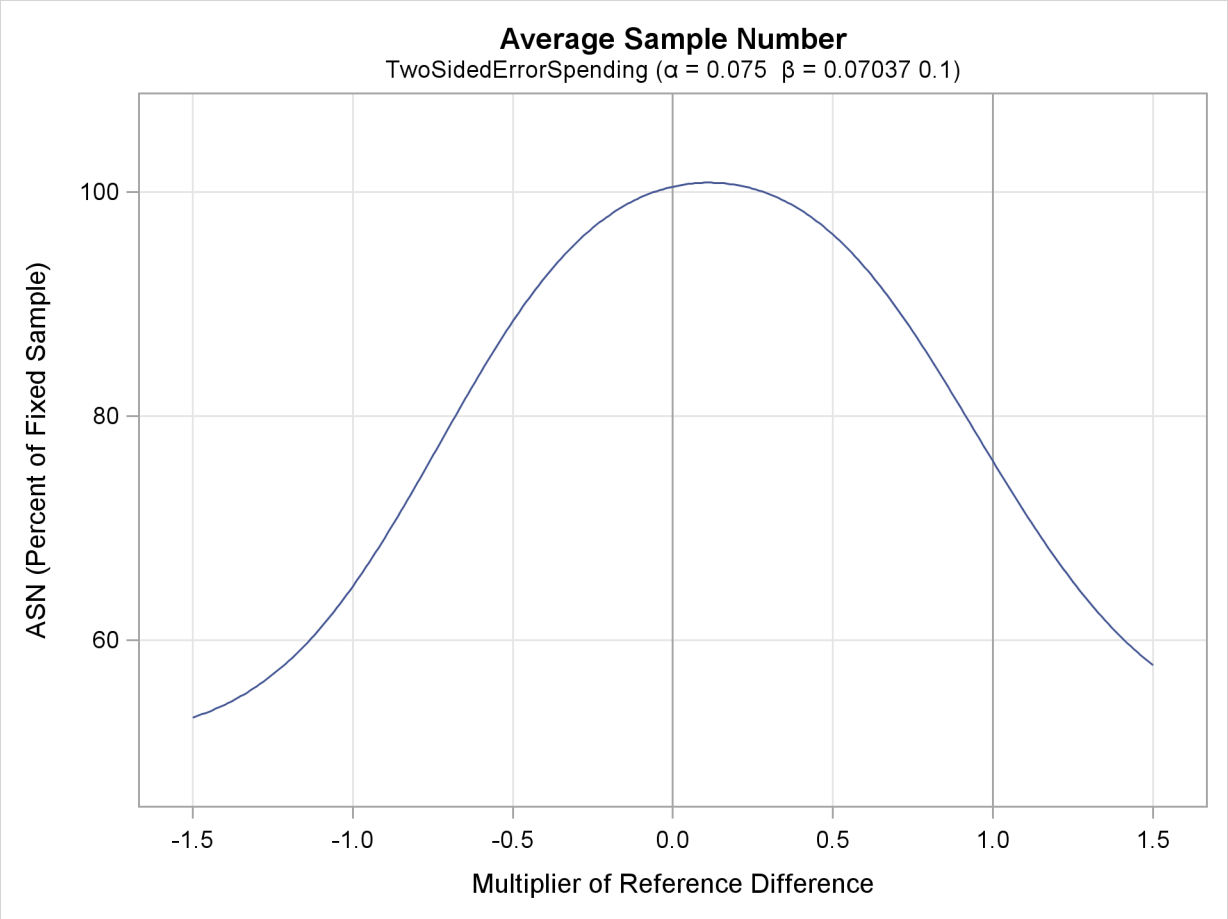
Output 77.11.5 Power Plot



The horizontal axis displays the multiplier of the reference difference. A positive multiplier corresponds to c_i for the upper alternative hypothesis, and a negative multiplier corresponds to $-c_i$ for the lower alternative hypothesis. For lower reference hypotheses, the power is the lower α error 0.05 under the null hypothesis ($c_i = 0$) and is 0.92963 under the alternative hypothesis ($c_i = 1$). For upper reference hypotheses, the power is the upper α error 0.025 under the null hypothesis ($c_i = 0$) and is 0.90 under the alternative hypothesis ($c_i = 1$).

With the PLOTS=ASN option, the procedure displays a plot of expected sample sizes under various hypothetical references, as shown in Output 77.11.6. By default, expected sample sizes under the lower hypotheses $\theta = c_i\,\theta_{1l}$ and under the upper hypotheses $\theta = c_i\,\theta_{1u}$, $c_i = 0, 0.01, 0.02, \ldots, 1.50$, are displayed for a two-sided asymmetric design, where $\theta_{1l} = -1$ and $\theta_{1u} = 1$ are the lower and upper alternative references, respectively.

Output 77.11.6 ASN Plot



The horizontal axis displays the multiplier of the reference difference. A positive multiplier corresponds to c_i for the upper alternative hypothesis and a negative multiplier corresponds to $-c_i$ for the lower alternative hypothesis.

The "Boundary Information" table in Output 77.11.7 displays the information levels, alternative references, and boundary values. The default BOUNDARYSCALE=STDZ option specifies that the standardized Z scale be used to display the alternative references and boundary values. The resulting standardized alternative references at stage k are given by $\pm\theta_1\sqrt{I_k}$, where θ_1 is the specified alternative reference and I_k is the information level at stage k, $k = 1, 2, 3$.

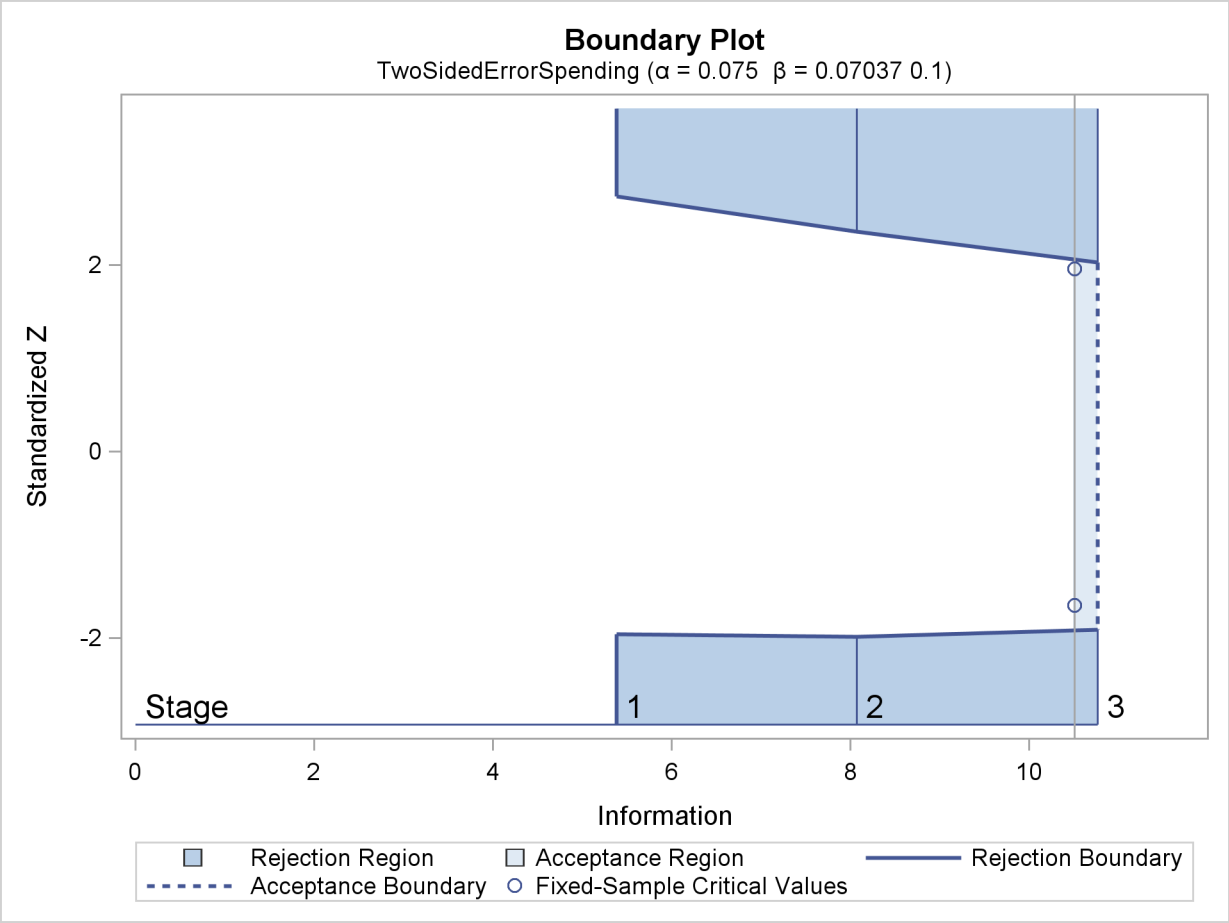
Output 77.11.7 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	--Information Level- Proportion	Actual	------Alternative----- -------Reference------ Lower	Upper	----Boundary Values--- ---Lower-- Alpha	---Upper-- Alpha
1	0.5000	5.381827	-2.31988	2.31988	-1.95996	2.73437
2	0.7500	8.07274	-2.84126	2.84126	-1.98394	2.35681
3	1.0000	10.76365	-3.28080	3.28080	-1.90855	2.02853

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.11.8.

Output 77.11.8 Boundary Plot



The "Error Spending Information" table in Output 77.11.9 displays the cumulative error spending at each stage for each boundary.

Output 77.11.9 Error Spending Information

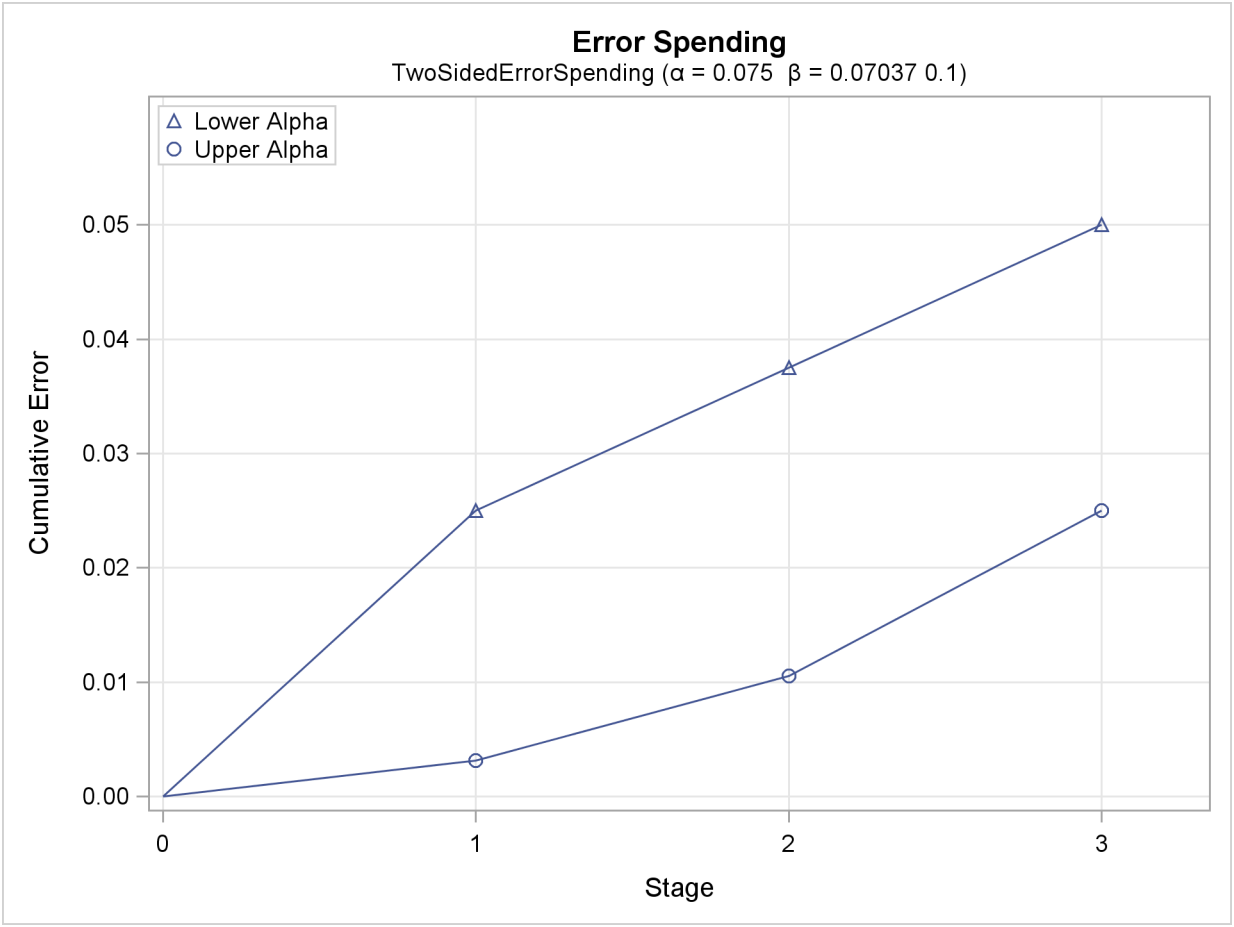
Error Spending Information

Stage	-Information Level- Proportion	--------Cumulative Error Spending------- -------Lower------ Alpha	Beta	-------Upper------ Beta	Alpha
1	0.5000	0.02500	0.00000	0.00001	0.00313
2	0.7500	0.03750	0.00000	0.00001	0.01055
3	1.0000	0.05000	0.07037	0.10000	0.02500

With the STOP=REJECT option, there is no early stopping to accept H_0, and the corresponding β spending at an interim stage is computed from the rejection region. For example, the upper β spending at stage 1 (0.00001) is the probability of rejecting H_0 for the lower alternative under the upper alternative reference.

With the PLOTS=ERRSPEND option, the procedure displays a plot of the cumulative error spending on each boundary at each stage, as shown in Output 77.11.10.

Output 77.11.10 Error Spending Plot



Example 77.12: Creating a Two-Sided Asymmetric Error Spending Design with Early Stopping to Reject or Accept H_0

This example requests a four-stage two-sided asymmetric group sequential design for normally distributed statistics. The O'Brien-Fleming boundary can be approximated by a gamma family error spending function with parameter $\gamma = -4$ or -5, and the Pocock boundary can be approximated with parameter $\gamma = 1$ (Hwang, Shih, and DeCani 1990, p. 1440). The following statements use the gamma error spending function with early stopping to reject or accept the null hypothesis H_0:

```
ods graphics on;
proc seqdesign altref=2
               pss(cref=0 0.5 1)
               stopprob(cref=0 1)
               errspend
               plots=(asn power errspend)
               ;
TwoSidedAsymmetric: design nstages=4
                    method=errfuncgamma(gamma=1)
                    method(beta)=errfuncgamma(gamma=-2)
                    method(upperalpha)=errfuncgamma(gamma=-5)
                    alt=twosided
                    stop=both
                    beta=0.1
                    ;
run;
ods graphics off;
```

The design uses gamma family error spending functions with $\gamma = -5$ for the upper α boundary, $\gamma = 1$ for the lower α boundary, and $\gamma = -2$ for the lower and upper β boundaries.

The "Design Information" table in Output 77.12.1 displays design specifications and the derived maximum information. Note that in order to attain the same information level for the asymmetric lower and upper boundaries, the derived power at the upper alternative 0.93655 is larger than the specified $1 - \beta = 0.90$.

Output 77.12.1 Design Information

The SEQDESIGN Procedure
Design: TwoSidedAsymmetric

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Accept/Reject Null
Method	Error Spending
Boundary Key	Both
Alternative Reference	2
Number of Stages	4
Alpha	0.05
Beta (Lower)	0.1
Beta (Upper)	0.06345
Power (Lower)	0.9
Power (Upper)	0.93655
Max Information (Percent of Fixed Sample)	104.0688
Max Information	3.162386
Null Ref ASN (Percent of Fixed Sample)	74.16654
Lower Alt Ref ASN (Percent of Fixed Sample)	59.10271
Upper Alt Ref ASN (Percent of Fixed Sample)	73.78797

The "Method Information" table in Output 77.11.2 displays the specified α and β error levels and the derived drift parameter. With the same information level used for the asymmetric lower and upper boundaries, only one of the β levels is maintained and the other is derived to have the level less than or equal to the specified level.

Output 77.12.2 Method Information

Method Information

Boundary	Method	Alpha	Beta	----Error Spending---- Function
Upper Alpha	Error Spending	0.02500	.	Gamma (Gamma=-5)
Upper Beta	Error Spending	.	0.06345	Gamma (Gamma=-2)
Lower Beta	Error Spending	.	0.10000	Gamma (Gamma=-2)
Lower Alpha	Error Spending	0.02500	.	Gamma (Gamma=1)

Method Information

Boundary	Alternative Reference	Drift
Upper Alpha	2	3.55662
Upper Beta	2	3.55662
Lower Beta	-2	-3.55662
Lower Alpha	-2	-3.55662

With the STOPPROB(CREF=0 1) option, the "Expected Cumulative Stopping Probabilities" table in Output 77.12.3 displays the expected stopping stage and cumulative stopping probabilities at each stage under the null reference $\theta = 0$ and under the alternative reference $\theta = \theta_1$.

Output 77.12.3 Stopping Probabilities

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

CRef	Ref	Expected Stopping Stage	Source
0.0000	Lower Alt	2.851	Rej Null (Lower Alt)
0.0000	Lower Alt	2.851	Rej Null (Upper Alt)
0.0000	Lower Alt	2.851	Reject Null
0.0000	Lower Alt	2.851	Accept Null
0.0000	Lower Alt	2.851	Total
1.0000	Lower Alt	2.272	Rej Null (Lower Alt)
1.0000	Lower Alt	2.272	Rej Null (Upper Alt)
1.0000	Lower Alt	2.272	Reject Null
1.0000	Lower Alt	2.272	Accept Null
1.0000	Lower Alt	2.272	Total
0.0000	Upper Alt	2.851	Rej Null (Lower Alt)
0.0000	Upper Alt	2.851	Rej Null (Upper Alt)
0.0000	Upper Alt	2.851	Reject Null
0.0000	Upper Alt	2.851	Accept Null
0.0000	Upper Alt	2.851	Total
1.0000	Upper Alt	2.836	Rej Null (Lower Alt)
1.0000	Upper Alt	2.836	Rej Null (Upper Alt)
1.0000	Upper Alt	2.836	Reject Null
1.0000	Upper Alt	2.836	Accept Null
1.0000	Upper Alt	2.836	Total

Expected Cumulative Stopping Probabilities
Reference = CRef * (Alt Reference)

		Stopping Probabilities			
CRef	Ref	Stage_1	Stage_2	Stage_3	Stage_4
0.0000	Lower Alt	0.00875	0.01556	0.02087	0.02500
0.0000	Lower Alt	0.00042	0.00190	0.00704	0.02500
0.0000	Lower Alt	0.00917	0.01746	0.02791	0.05000
0.0000	Lower Alt	0.00000	0.30125	0.79354	0.95000
0.0000	Lower Alt	0.00917	0.31870	0.82145	1.00000
1.0000	Lower Alt	0.27499	0.58934	0.79601	0.90000
1.0000	Lower Alt	0.00000	0.00000	0.00000	0.00000
1.0000	Lower Alt	0.27499	0.58934	0.79601	0.90000
1.0000	Lower Alt	0.00000	0.01863	0.04935	0.10000
1.0000	Lower Alt	0.27499	0.60797	0.84536	1.00000
0.0000	Upper Alt	0.00875	0.01556	0.02087	0.02500
0.0000	Upper Alt	0.00042	0.00190	0.00704	0.02500
0.0000	Upper Alt	0.00917	0.01746	0.02791	0.05000
0.0000	Upper Alt	0.00000	0.30125	0.79354	0.95000
0.0000	Upper Alt	0.00917	0.31870	0.82145	1.00000
1.0000	Upper Alt	0.00002	0.00002	0.00002	0.00002
1.0000	Upper Alt	0.05945	0.33802	0.72323	0.93655
1.0000	Upper Alt	0.05947	0.33804	0.72325	0.93657
1.0000	Upper Alt	0.00000	0.01182	0.03131	0.06343
1.0000	Upper Alt	0.05947	0.34986	0.75456	1.00000

"Rej Null (Lower Alt)" and "Rej Null (Upper Alt)" under the heading "Source" indicate the probabilities of rejecting the null hypothesis for the lower alternative and for the upper alternative, respectively. "Reject Null" indicates the probability of rejecting the null hypothesis for either the lower or upper alternative, "Accept Null" indicates the probability of accepting the null hypothesis, and "Total" indicates the total probability of stopping the trial.

With the PSS(CREF=0 0.5 1.0) option, the "Power and Expected Sample Sizes" table in Output 77.12.4 displays powers and expected sample sizes under hypothetical references $\theta = 0$ (null hypothesis H_0), $\theta = 0.5\,\theta_1$, and $\theta = \theta_1$ (alternative hypothesis H_1), where θ_1 is the alternative reference. The expected sample sizes are displayed in a scale that indicates a percentage of its corresponding fixed-sample size design.

Output 77.12.4 Power and Expected Sample Size Information

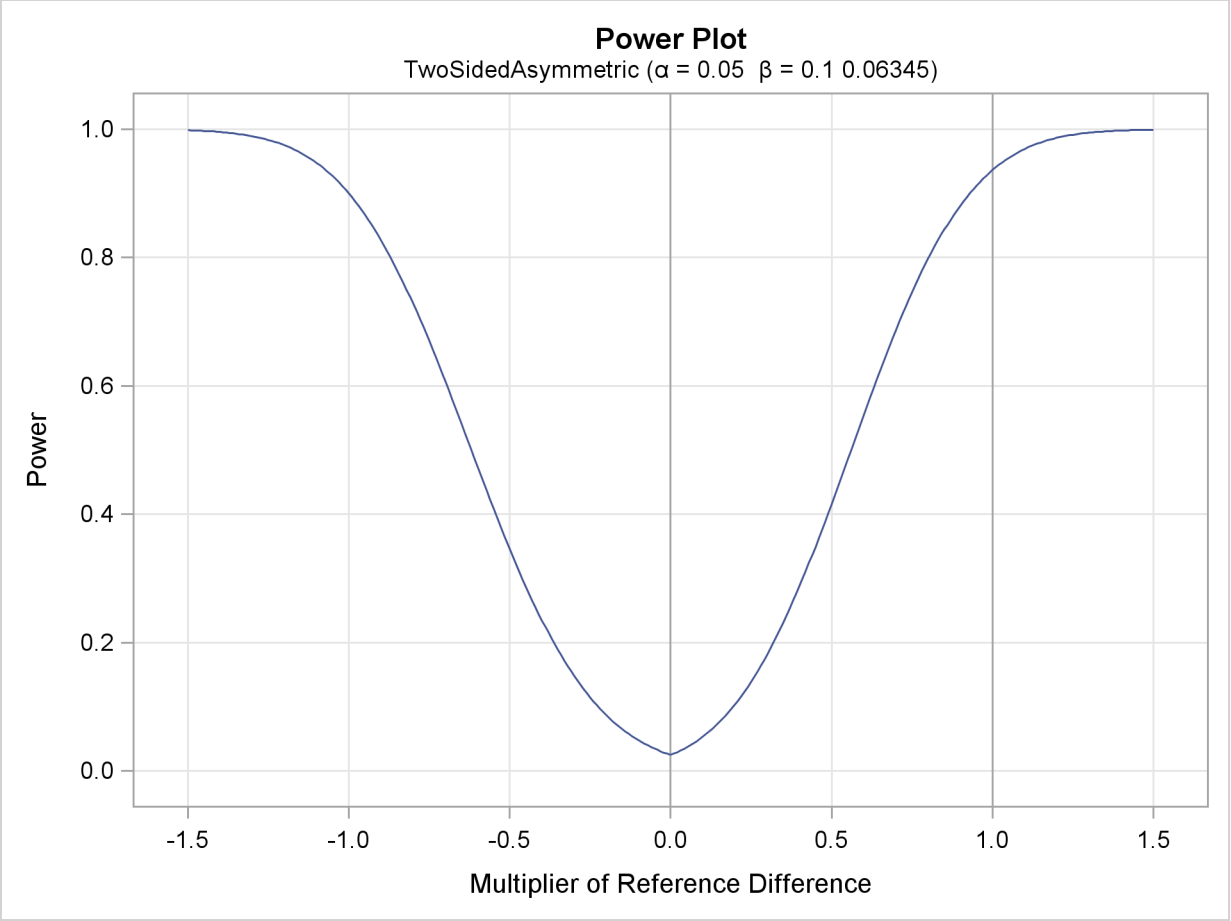
Powers and Expected Sample Sizes
Reference = CRef * (Alt Reference)

CRef	Ref	Power	-Sample Size- Percent Fixed-Sample
0.0000	Lower Alt	0.02500	74.1665
0.5000	Lower Alt	0.34601	75.8425
1.0000	Lower Alt	0.90000	59.1027
0.0000	Upper Alt	0.02500	74.1665
0.5000	Upper Alt	0.41647	85.3976
1.0000	Upper Alt	0.93655	73.7880

Note that at $c_i = 0$, the null reference $\theta = 0$, the power with the lower alternative is the lower α error 0.025, and the power with the upper alternative is the upper α error 0.025. At $c_i = 1$, the alternative reference $\theta = \theta_1$, the power with the lower alternative is the specified power 0.90, and the power with the upper alternative 0.93655 is greater than the specified power 0.90 because the same information level is used for these two asymmetric boundaries.

With the PLOTS=POWER option, the procedure displays a plot of the power curves under various hypothetical references, as shown in Output 77.12.5. By default, powers under the lower hypotheses $\theta = c_i\,\theta_{1l}$ and under the upper hypotheses $\theta = c_i\,\theta_{1u}$, are displayed for a two-sided asymmetric design, where $c_i = 0, 0.01, 0.02, \ldots, 1.50$ and $\theta_{1l} = -1$ and $\theta_{1u} = 1$ are the lower and upper alternative references, respectively.

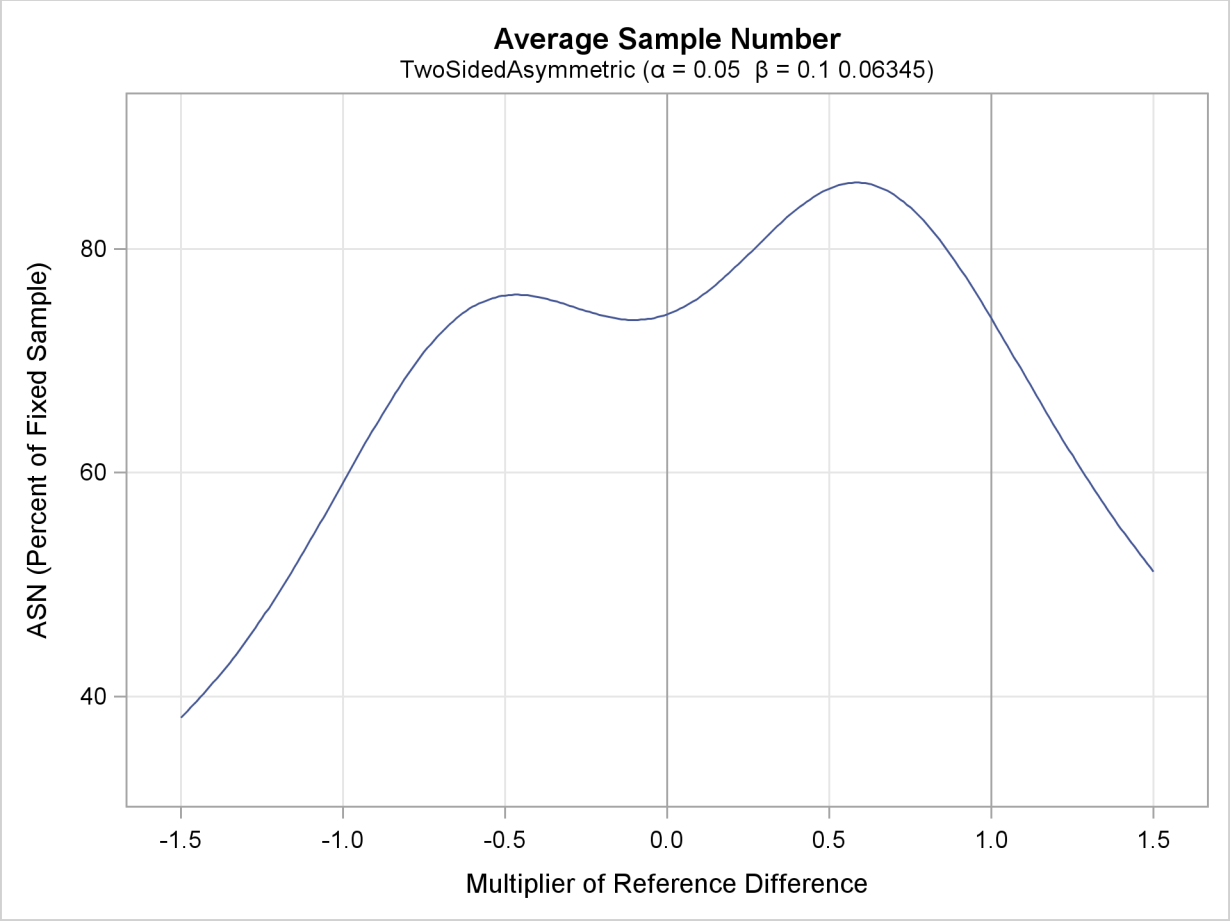
Output 77.12.5 Power Plot



The horizontal axis displays the multiplier of the reference difference. A positive multiplier corresponds to c_i for the upper alternative hypothesis, and a negative multiplier corresponds to $-c_i$ for the lower alternative hypothesis. For lower reference hypotheses, the power is the lower α error 0.025 under the null hypothesis ($c_i = 0$) and is 0.90 under the alternative hypothesis ($c_i = 1$). For upper reference hypotheses, the power is the upper α error 0.025 under the null hypothesis ($c_i = 0$) and is 0.93655 under the alternative hypothesis ($c_i = 1$).

With the PLOTS=ASN option, the procedure displays a plot of expected sample sizes under various hypothetical references, as shown in Output 77.12.6. By default, expected sample sizes under the lower hypotheses $\theta = c_i\,\theta_{1l}$ and under the upper hypotheses $\theta = c_i\,\theta_{1u}$ are displayed for a two-sided asymmetric design, where $c_i = 0, 0.01, 0.02, \ldots, 1.50$ and $\theta_{1l} = -1$ and $\theta_{1u} = 1$ are the lower and upper alternative references, respectively.

Output 77.12.6 ASN Plot



The horizontal axis displays the multiplier of the reference difference. A positive multiplier corresponds to c_i for the upper alternative hypothesis, and a negative multiplier corresponds to $-c_i$ for the lower alternative hypothesis.

By default, or when you specify BETAOVERLAP=ADJUST, the SEQDESIGN procedure first derives boundary values without adjusting for the possible overlapping of the two one-sided β boundaries based on two corresponding one-sided tests. Then the procedure checks for overlapping of the β boundaries at the interim stages. Since the two β boundaries overlap at stage 1, the β boundary values for stage 1 are set to missing, the β spending values at stage 1 are set to zero, and the β spending values at subsequent stages are adjusted proportionally.

The "Boundary Information" table in Output 77.12.7 displays the information levels, alternative references, and boundary values. The default BOUNDARYSCALE=STDZ option specifies that the standardized Z scale be used to display the alternative references and boundary values. The resulting standardized alternative references at stage k is given by $\pm\theta_1\sqrt{I_k}$, where θ_1 is the specified alternative reference and I_k is the information level at stage k, $k = 1, 2, 3, 4$.

Output 77.12.7 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

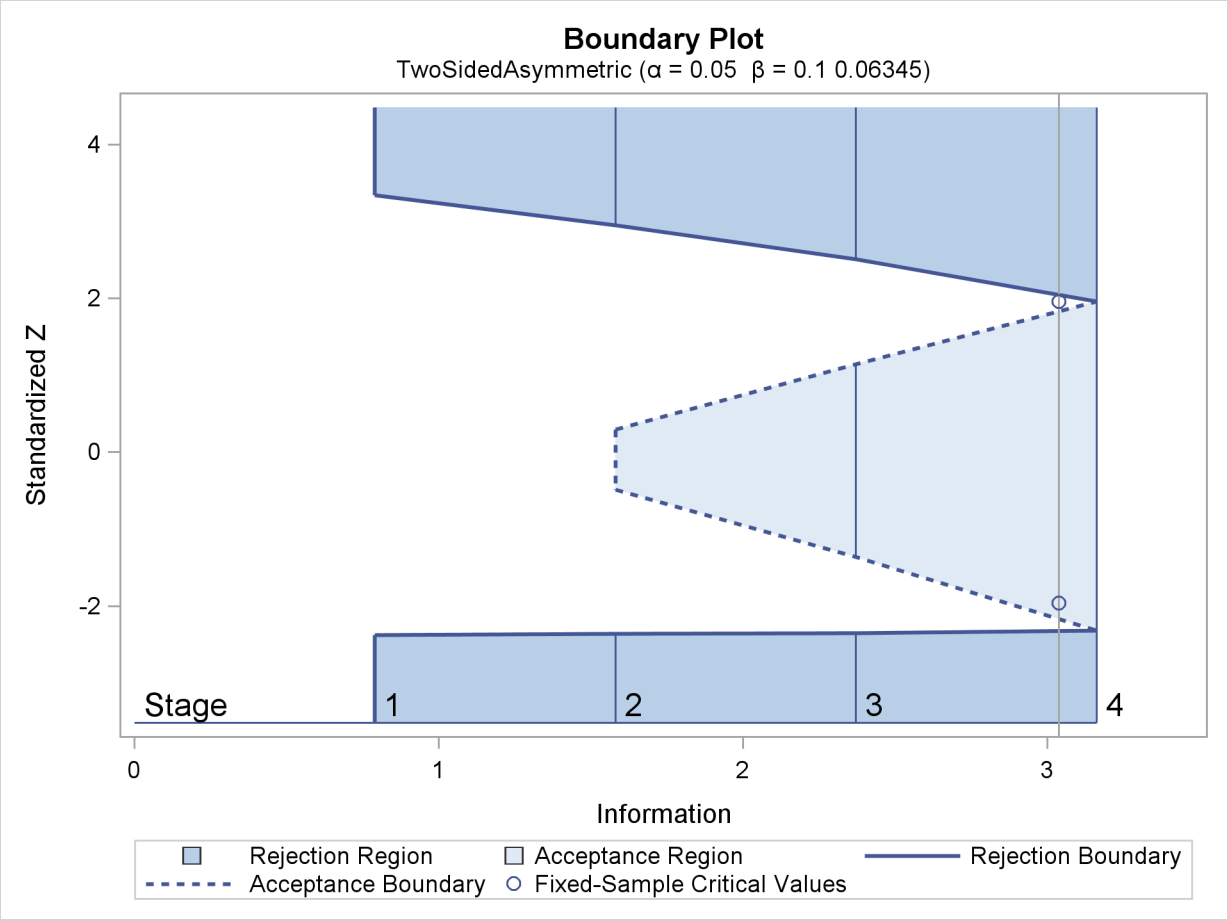
Stage	---Information Level-- Proportion	Actual	-------Alternative------ --------Reference------- Lower	Upper
1	0.2500	0.790597	-1.77831	1.77831
2	0.5000	1.581193	-2.51491	2.51491
3	0.7500	2.37179	-3.08012	3.08012
4	1.0000	3.162386	-3.55662	3.55662

Boundary Information (Standardized Z Scale)
Null Reference = 0

	----------Boundary Values---------- ----------Lower----------		----------Upper----------	
Stage	Alpha	Beta	Beta	Alpha
1	-2.37610	.	.	3.33772
2	-2.35714	-0.48408	0.29400	2.94871
3	-2.34861	-1.36183	1.13898	2.50473
4	-2.32105	-2.32105	1.95675	1.95675

With the specified ODS GRAPHICS ON statement, a detailed boundary plot with the rejection and acceptance regions is displayed by default, as shown in Output 77.12.8.

Output 77.12.8 Boundary Plot



The "Error Spending Information" in Output 77.12.9 displays the cumulative error spending at each stage for each boundary.

Output 77.12.9 Error Spending Information

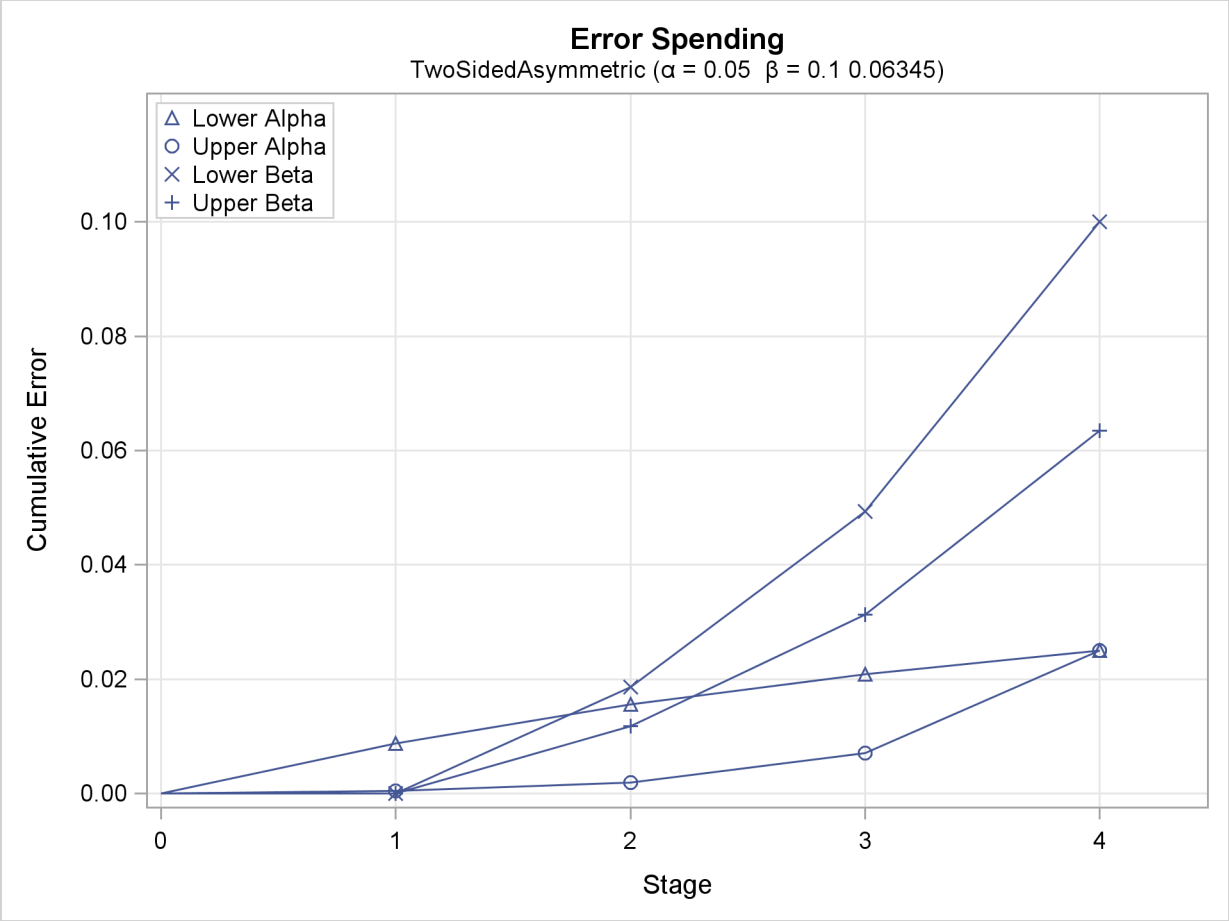
Error Spending Information

		Cumulative Error Spending			
	Information Level	Lower		Upper	
Stage	Proportion	Alpha	Beta	Beta	Alpha
1	0.2500	0.00875	0.00000	0.00002	0.00042
2	0.5000	0.01556	0.01863	0.01184	0.00190
3	0.7500	0.02087	0.04935	0.03132	0.00704
4	1.0000	0.02500	0.10000	0.06345	0.02500

With the β boundary values missing at stage 1, there is no early stopping to accept H_0 at stage 1, and the corresponding β spending at stage 1 is computed from the rejection region. For example, the upper β spending at stage 1 (0.00002) is the probability of rejecting H_0 for the lower alternative under the upper alternative reference.

With the PLOTS=ERRSPEND option, the procedure displays a plot of the cumulative error spending on each boundary at each stage, as shown in Output 77.12.10.

Output 77.12.10 Error Spending Plot



Acknowledgments

In addition to being shaped by the research literature listed in the section "References" on page 5979, the development of the SEQDESIGN and SEQTEST procedures has benefited significantly from the advice and expertise of the following researchers:

- Lu Cui, Eisai Medical Research
- Alex Dmitrienko, Eli Lilly
- Scott Emerson, University of Washington
- Gordon Lan, Johnson & Johnson
- Steve Snapinn, Amgen
- John Whitehead, University of Reading

The time and effort that these researchers have contributed is gratefully acknowledged.

References

Armitage, P., McPherson, C. K., and Rowe, B. C. (1969), "Repeated Significance Test on Accumulating Data," *Journal of the Royal Statistical Society, Series A*, 132, 235–244.

Chow, S. C. and Liu, J. P. (1998), *Design and Analysis of Clinical Trials, Concept and Methodologies,* New York: John Wiley & Sons.

Chow, S. C., Shao, J., and Wang, H. (2003), *Sample Size Calculations in Clinical Research,* Boca Raton, FL: CRC Press.

Cox, D. R. and Hinkley, D. V.(1974), *Theoretical Statistics,* London: Chapman & Hall.

DeMets, D. L., Furberg, C. D., and Friedman, L. M. (2006), *Data Monitoring in Clinical Trials,* New York: Springer.

Diggle, P. J., Heagerty, P., Liang, K. Y., and Zeger, S. L. (2002), *Analysis of Longitudinal Data*, Second Edition, New York: Oxford University Press.

Dmitrienko, A., Molenberghs, G., Chuang-Stein, C., and Offen, W. (2005), *Analysis of Clinical Trials Using SAS: A Practical Guide,* Cary, NC: SAS Institute.

Efron, B. and Hinkley, D. V.(1978), "Assessing the Accuracy of the Maximum Likelihood Estimator: Observed versus Expected Fisher Information," *Biometrika*, 65, 457–483.

Ellenberg, S. S., Fleming, T. R., and DeMets, D. L. (2003), *Data Monitoring Committees in Clinical Trials,* New York: John Wiley & Sons.

Emerson, S. S. (1996), "Statistical Packages for Group Sequential Methods," *The American Statistician*, 50, 183–192.

Emerson, S. S. and Fleming, T. R. (1989), "Symmetric Group Sequential Designs," *Biometrics*, 45, 905–923.

Emerson, S. S., Kittelson, J. M., and Gillen, D. L. (2005), "On the Use of Stochastic Curtailment in Group Sequential Clinical Trials," *UW Biostatistics Working Paper Series*, *http://www.bepress.com/uwbiostat/paper243*.

Food and Drug Administration (1998), "E9: Statistical Principles for Clinical Trials," *Federal Register*, 63 (179), 49583–49598.

Haybittle, J. L. (1971), "Repeated Assessment of Results in Clinical Trials of Cancer Treatment," *Brit. J. Radiology*, 44, 793–797.

Hsieh, F. Y. and Lavori, P. W. (2000), "Sample-Size Calculations for the Cox Proportional Hazards Regression Model with Nonbinary Covariates," *Controlled Clinical Trials*, 21, 552–560.

Hwang, I. K., Shih, W. J., and DeCani, J. S. (1990), "Group Sequential Designs Using a Family of Type I Error Probability Spending Functions," *Statistics in Medicine*, 9, 1439–1445.

Jennison, C. and Turnbull, B. W. (1990), "Statistical Approaches to Interim Monitoring of Medical Trials: A Review and Commentary," *Statistical Science*, 5, 299–317.

Jennison, C. and Turnbull, B. W. (2000), *Group Sequential Methods with Applications to Clinical Trials,* New York: Chapman & Hall.

Kalbfleisch, J. D. and Prentice, R. L. (1980), *The Statistical Analysis of Failure Time Data,* New York: John Wiley & Sons.

Kim, K. and DeMets, D. L. (1987), "Design and Analysis of Group Sequential Tests Based on the Type I Error Spending Rate Function," *Biometrika*, 74, 149–154.

Kim, K. and Tsiatis, A. A. (1990), "Study Duration for Clinical Trials with Survival Response and Early Stopping Rule," *Biometrics*, 46, 81–92.

Kittelson, J. M. and Emerson, S. S. (1999), "A Unifying Family of Group Sequential Test Designs," *Biometrics*, 55, 874–882.

Lan, K. K. G. and DeMets, D. L. (1983), "Discrete Sequential Boundaries for Clinical Trials," *Biometrika*, 70, 659–663.

Lan, K. K. G., Lachin, J. M, and Bautista, O. (2003), "Over-ruling a Group Sequential Boundary: A Stopping Rule versus a Guideline," *Statistics in Medicine*, 22, 3347–3355.

Lan, K. K. G., Simon, R., and Halperin, M. (1982), "Stochastically Curtailed Tests in Long-Term Clinical Trials," *Sequential Analysis*, 1, 207–219.

Lindgren, B. W. (1976), *Statistical Theory,* Third Edition, New York: Macmillan.

McCullagh, P. and Nelder, J. A. (1989), *Generalized Linear Models,* Second Edition, New York: Chapman & Hall/CRC.

Mehta, C. R. and Tsiatis, A. A. (2001), "Flexible Sample Size Considerations under Information Based Interim Monitoring," *Drug Information Journal*, 35, 1095–1112.

O'Brien, P. C. and Fleming, T. R. (1979), "A Multiple Testing Procedure for Clinical Trials," *Biometrics*, 35, 549–556.

O'Neill, R. T. (1994), "Interim Analysis, A Regulatory Perspective on Data Monitoring and Interim Analysis," *Statistics in the Pharmaceutical Industry,* Revised and Expanded Second Edition, ed. C. R. Buncher and J-Y Tsay, New York: Marcel Dekker, 285–290.

Pampallona, S. and Tsiatis, A. A. (1994), "Group Sequential Designs for One-Sided and Two-Sided Hypothesis Testing with Provision for Early Stopping in Favor of the Null Hypothesis," *J. Statist. Planning and Inference*, 42, 19–35.

Peto, R., Pike, M. C., et al. (1976), "Design and Analysis of Randomized Clinical Trials Requiring Prolonged Observation of Each Patient: I. Introduction and Design," *British Journal of Cancer*, 34, 585–612.

Pocock, S. J. (1977), "Group Sequential Methods in the Design and Analysis of Clinical Trials," *Biometrika*, 64, 191–199.

Pocock, S. J. (1982), "Interim Analyses for Randomized Clinical Trials: The Group Sequential Approach," *Biometrics*, 38, 153–162.

Pocock, S. J. and White, I. (1999), "Trials Stopped Early: Too Good to Be True?" *Lancet*, 353, 943–944.

Proschan, M. A., Lan, K. K. G., and Wittes, J. T. (2006), *Statistical Monitoring of Clinical Trials,* New York: Springer.

Rudser, K.D. and Emerson, S.S. (2007), "Implementing Type I & Type II Error Spending for Two-Sided Group Sequential Designs," *Contemporary Clinical Trials*, doi:10.1016/j.cct.2007.09.002.

Schoenfeld, D. A. (1983), "Sample-Size Formula for the Proportional-Hazards Regression Model," *Biometrics*, 39, 499–503.

Senn, S. (1997), *Statistical Issues in Drug Development,* New York: John Wiley & Sons.

Snapinn, S. M. (2000), "Noninferiority Trials," *Current Controlled Trials in Cardiovascular Medicine*, 1, 19–21.

Wang, S. K. and Tsiatis, A. A. (1987), "Approximately Optimal One-Parameter Boundaries for Group Sequential Trials," *Biometrics*, 43, 193–200.

Ware, J. H., Muller, J. E., and Braunwald, E. (1985), "The Futility Index: An Approach to the Cost-Effective Termination of Randomized Clinical Trials" *American Journal of Medicine*, 78, 635–643.

Whitehead, J. (1997), *The Design and Analysis of Sequential Clinical Trials,* Revised Second Edition, Chichester: John Wiley & Sons.

Whitehead, J. (2001), "Use of the Triangular Test in Sequential Clinical Trials," *Handbook of Statistics in Clinical Oncology,* ed. J. Crowley, New York: Marcel Dekker, 211–228.

Whitehead, J. and Jones, D. R. (1979), "The Analysis of Sequential Clinical Trials," *Biometrika*, 66, 443–452.

Whitehead, J. and Stratton, I. (1983), "Group Sequential Clinical Trials with Triangular Continuation Regions," *Biometrics*, 39, 227–236.

Subject Index

Syntax Index

Your Turn

We welcome your feedback.

- If you have comments about this book, please send them to **yourturn@sas.com**. Include the full title and page numbers (if applicable).
- If you have comments about the software, please send them to **suggest@sas.com**.

SAS® Publishing Delivers!

Whether you are new to the work force or an experienced professional, you need to distinguish yourself in this rapidly changing and competitive job market. SAS® Publishing provides you with a wide range of resources to help you set yourself apart. Visit us online at support.sas.com/bookstore.

SAS® Press

Need to learn the basics? Struggling with a programming problem? You'll find the expert answers that you need in example-rich books from SAS Press. Written by experienced SAS professionals from around the world, SAS Press books deliver real-world insights on a broad range of topics for all skill levels.

support.sas.com/saspress

SAS® Documentation

To successfully implement applications using SAS software, companies in every industry and on every continent all turn to the one source for accurate, timely, and reliable information: SAS documentation. We currently produce the following types of reference documentation to improve your work experience:

- Online help that is built into the software.
- Tutorials that are integrated into the product.
- Reference documentation delivered in HTML and PDF – **free** on the Web.
- Hard-copy books.

support.sas.com/publishing

SAS® Publishing News

Subscribe to SAS Publishing News to receive up-to-date information about all new SAS titles, author podcasts, and new Web site features via e-mail. Complete instructions on how to subscribe, as well as access to past issues, are available at our Web site.

support.sas.com/spn

sas | THE POWER TO KNOW®