THE POWER TO KNOW.

SAS/STAT® 9.3 User's Guide

The SEQTEST Procedure

(Chapter)

SAS® Documentation

This document is an individual chapter from *SAS/STAT® 9.3 User's Guide*.

The correct bibliographic citation for the complete manual is as follows: SAS Institute Inc. 2011. *SAS/STAT® 9.3 User's Guide*. Cary, NC: SAS Institute Inc.

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

1st electronic book, July 2011

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 81

The SEQTEST Procedure

Contents

Overview: SEQTEST Procedure

The purpose of the SEQTEST procedure is to perform interim analyses for clinical trials. Clinical trials are experiments on human beings 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 analyses of the data. The protocol contains information such as a null hypothesis and an alternative hypothesis, a test statistic, the probability α of a Type I error (incorrectly rejecting the null hypothesis), the probability β of a Type II error (incorrectly accepting the null hypothesis), the sample size needed to attain a specified power (probability of correctly rejecting the null hypothesis) of $1 - \beta$ at an alternative reference, and critical values that are associated with the test statistic for hypothesis testing.

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 data safety monitoring boards or data monitoring 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 rare situations, the board or committee might even 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 formal completion of the trial while maintaining the specified overall Type I and Type II error probability levels.

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. If a group sequential trial stops early, then it usually requires fewer participants than a corresponding fixed-sample trial.

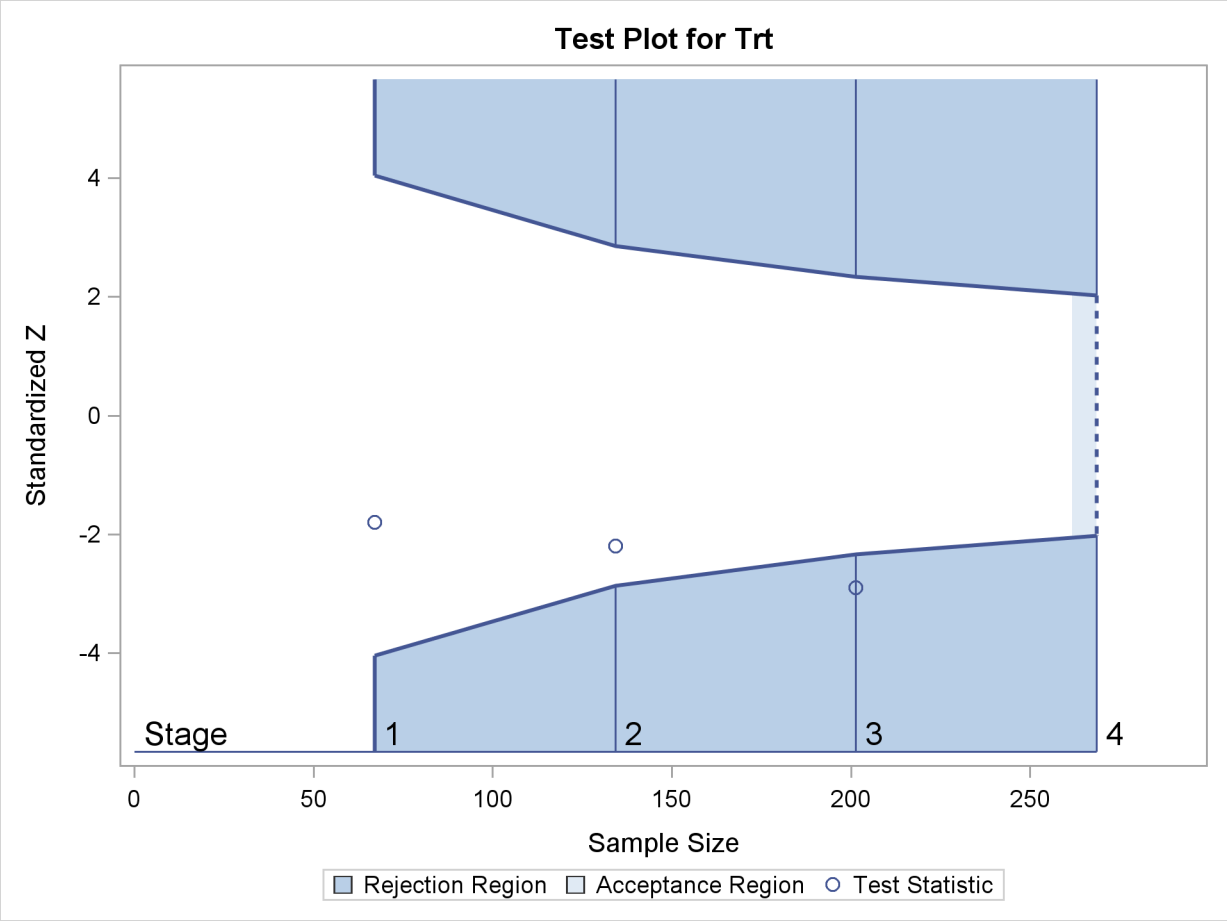
Thus, in most cases, if a group sequential trial stops early for safety of the new treatment, fewer patients will be exposed to the new treatment than in the fixed-sample trial. Also, if a trial stops early for efficacy of the new treatment, the new treatment will be available sooner than it would be in a fixed-sample trial. Furthermore, if a trial stops early, this 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 its corresponding critical values at the stage, 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 to maintain the overall α level, the overall β level, or both the overall α and β levels.

Figure 81.1 shows a two-sided symmetric group sequential trial that stops early to reject the null hypothesis that the parameter Trt is zero.

Figure 81.1 Sequential Plot for Two-Sided Test



The trial has four stages, which are indicated by vertical lines with accompanying stage numbers. With early stopping to reject the null hypothesis, the lower rejection boundary is constructed by connecting the lower critical values (boundary 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 boundary values and test statistics on the standardized Z scale.

At each interim stage, if the standardized Z test statistic falls into a rejection region (the darker shaded areas in Figure 81.1), the trial stops and the null hypothesis is rejected. Otherwise, the trial continues to the next stage. At the final stage (stage 4), the trial is rejected if Z falls into a rejection region. Otherwise, the trial is accepted. In Figure 81.1, 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:

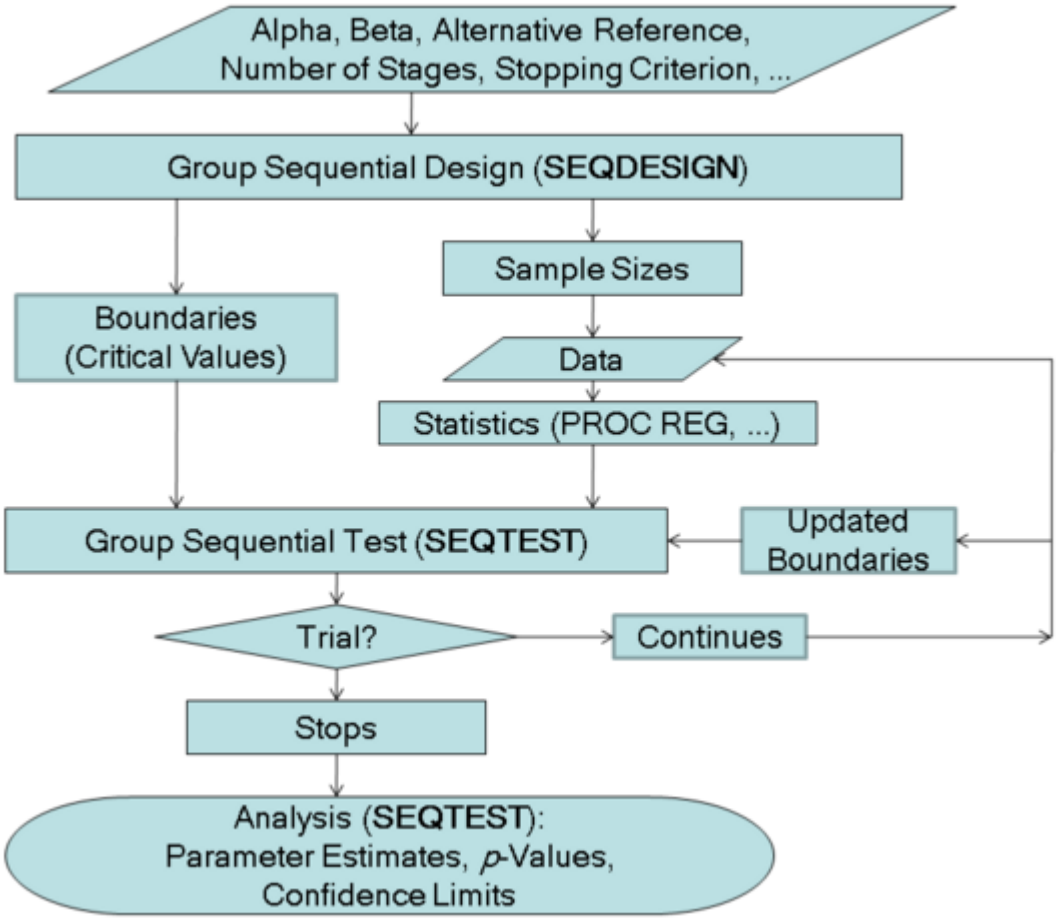
1. You specify the statistical details of the design, including the null and alternative hypotheses, a test statistic for the hypothesis test, the Type I and II error probabilities, a stopping criterion, the total number of stages, and the relative information level at each stage.
2. You compute the 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 companion SEQDESIGN procedure at Step 2 to compute the boundary values and required sample sizes for the trial. You use the 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 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.

Note that for some clinical trials, the information levels are derived from statistics based on individuals specified in the design plan and might not reach the target maximum information level. For example, if an estimate of the variance is used to compute the required sample size for a group sequential trial, the computed variance at each stage might not be the same as the estimated variance. Thus, instead of specifying the number of individuals in the protocol, the information level can be specified. You can then adjust the sample sizes with the updated variance estimates at interim stages to achieve the target maximum information level for the trial (Jennison and Turnbull 2000, p. 295).

The flowchart in Figure 81.2 summarizes the steps in a typical group sequential trial and the relevant SAS procedures.

Figure 81.2 Group Sequential Trial



Features of the SEQTEST Procedure

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 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. At each subsequent stage, the boundary values are derived by using the test information tables created by the SEQTEST procedure at the previous stage.

If the observed information level does not match the corresponding information level in the BOUNDARY= data set, the SEQTEST procedure modifies the boundary values to adjust for new information levels at the current and subsequent stages. See the section "Boundary Adjustments for Information Levels" on page 6895 for a detailed description of these boundary adjustments.

Either you can specify the test statistic and its information level in the DATA= input data set, or you can specify the test statistic and its associated standard error in the PARMS= input data set. With the PARMS=

input data set, the information level for the test statistic is computed from its standard error. See the section "Input Data Sets" on page 6889 for a detailed description of these input data sets.

At the end of a trial, the parameter estimate is computed. The median unbiased estimate, confidence limits, and p-value depend on the specified sample space ordering. A sample space ordering specifies the ordering for test statistics that result in the stopping of a trial. That is, for all the statistics in the rejection region and in acceptance region, the SEQTEST procedure provides three different sample space orderings: the stagewise ordering uses counterclockwise ordering around the continuation region, the LR ordering uses the distance between the observed Z statistic z and its hypothetical value, and the MLE ordering uses the observed maximum likelihood estimate. See the section "Available Sample Space Orderings in a Sequential Test" on page 6902 for a detailed description of these orderings.

Output from the SEQTEST Procedure

In addition to the adjusted boundary values and test results for the group sequential trial, the SEQTEST procedure also computes the following quantities:

- 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
- conditional power given the most recently observed statistic under specified hypothetical references
- predictive power given the most recently observed statistic
- repeated confidence intervals for the parameter from the observed statistic at each stage
- parameter estimate, p-value for hypothesis testing, and median and confidence limits for the parameter at the conclusion of a sequential trial

Getting Started: SEQTEST Procedure

The following example illustrates a clinical study 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 level: an experimental group given the new drug and a placebo control group. Suppose the changes in LDL level after the treatment for patients 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 $H_0 : \theta = \mu_e - \mu_c = 0$. Also suppose that the alternative reference $\theta = -10$ is the clinically meaningful difference that the trial should detect with a high probability (power), and 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=-10.0;
  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.

In the DESIGN statement, the label `TwoSidedOBrienFleming` identifies the design in the output tables. By default (or equivalently if you specify ALT=TWOSIDED and STOP=REJECT), the design has a two-sided alternative hypothesis with early stopping in the interim stages only to reject 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 $H_0 : \theta = 0$. By default (or equivalently if you specify ALPHA=0.05 and BETA=0.10), the design has a Type I error probability $\alpha = 0.05$, and 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 (rejection) 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 81.5.

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. 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 that are 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 that are derived from the boundary information stored in the test information table that was 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 (or equivalently if you specify INFO=EQUAL), 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" table in Figure 81.3 displays design specifications and three derived statistics: 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). 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). When you specify an alternative reference (in this case, ALTREF=−10), the actual maximum information 0.1074 is also computed. 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 81.3 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 "Boundary Information" table in Figure 81.4 displays the information level, the lower and upper alternative references, and the lower and upper boundary values at each stage. By default (or equivalently if you specify INFO=EQUAL), the SEQDESIGN procedure uses equally spaced information levels for all stages.

Figure 81.4 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Information Level N	Alternative Reference Lower	Alternative Reference 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

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

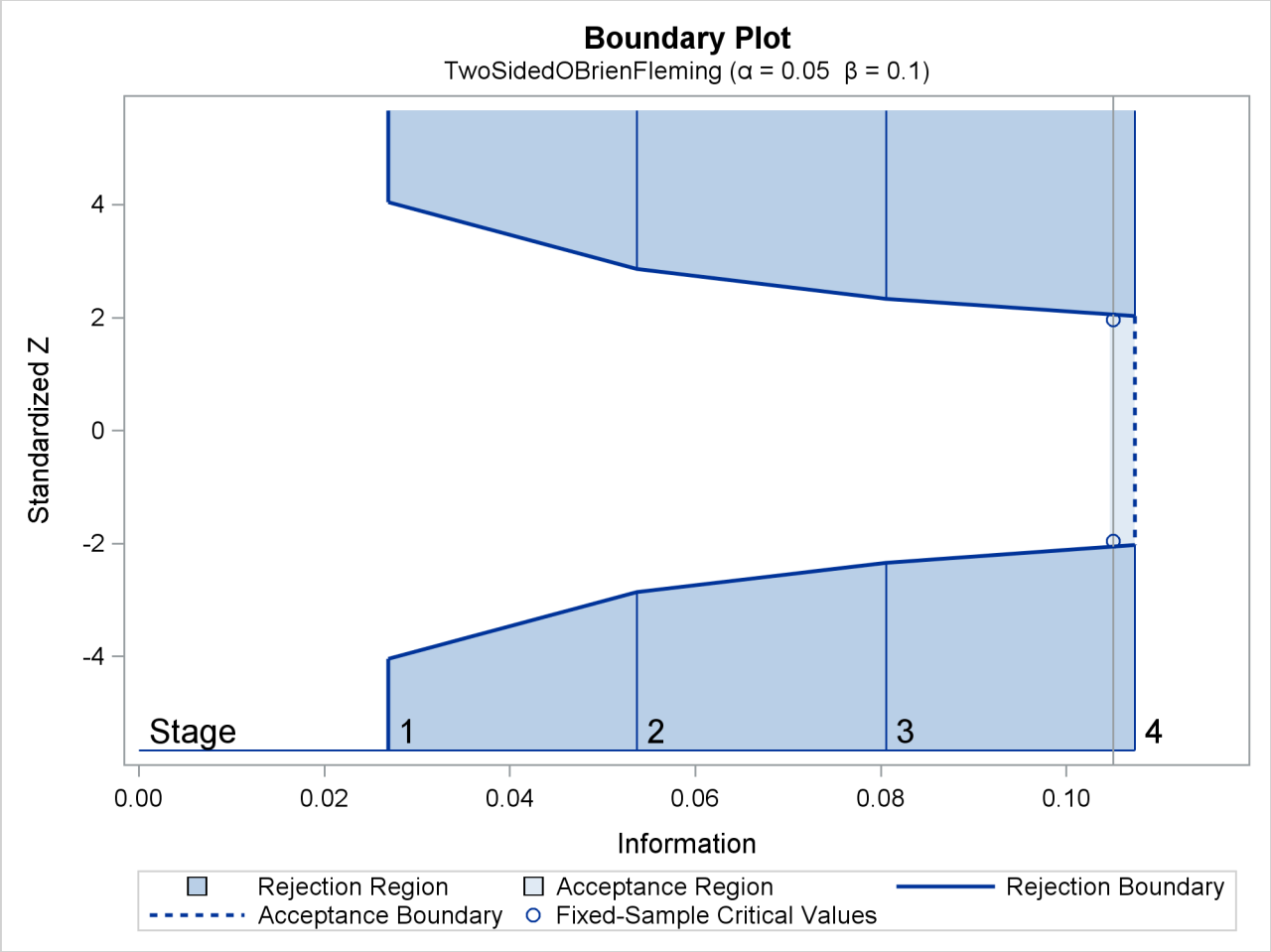
The information proportion is the proportion of maximum information available at each stage. By default (or equivalently if you specify BOUNDARYSCALE=STDZ), the alternative references and boundary values are displayed with the standardized Z statistic scale. The alternative reference in the standardized 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, $k = 1, 2, 3, 4$.

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. Consequently, the SEQTEST procedure modifies the boundary values stored in the BND_LDL data set to adjust for these new information levels.

If ODS Graphics is enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Figure 81.5.

Figure 81.5 O'Brien-Fleming Boundary Plot



This boundary plot displays the boundary values in the "Boundary Information" table in Figure 81.4. The stages are indicated by vertical lines with accompanying stage numbers. The horizontal axis indicates the information levels for the stages. 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. The 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.

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" in the chapter "The SEQDESIGN Procedure" for a detailed description of how these required sample sizes are calculated.

The "Sample Size Summary" table in Figure 81.6 displays the parameters for the sample size computation and the resulting maximum and expected sample sizes.

Figure 81.6 Required 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

With the derived maximum information 0.1074 and the specified MODEL=TWOSAMPLEMEAN (STDDEV=20) option in the SAMPLESIZE statement, the total sample size in each group is

$$N_a = N_b = 2\sigma^2 I_X = 2 \times 20^2 \times 0.1074 = 85.92$$

The "Sample Sizes (N)" table in Figure 81.7 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." By default (or equivalently if you specify WEIGHT=1 in the MODEL=TWOSAMPLEMEAN option), the sample sizes for the two groups are equal for the two-sample test.

Figure 81.7 Required 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, each of the two groups needs 22, 43, 65, and 86 patients for the four stages, respectively.

Suppose that 22 patients are available in each group at stage 1 and that their measurements for LDL are saved in the data set LDL_1. Figure 81.8 lists the first 10 observations in the data set LDL_1.

Figure 81.8 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Trt	Ldl
1	0	33.33
2	1	-14.89
3	0	15.30
4	1	4.71
5	0	26.89
6	1	-48.74
7	0	-39.35
8	1	-8.13
9	0	-8.22
10	1	12.35

The variable Trt is an indicator variable with value 1 for patients in the treatment group and value 0 for patients in the placebo control group. The variable Ldl is the LDL level of these patients.

The following statements use the REG procedure to estimate the mean treatment difference and its associated standard error at stage 1:

```
proc reg data=LDL_1;
   model Ldl=Trt;
ods output ParameterEstimates=Parms_LDL1;
run;
```

The following statements create the data set for the mean treatment difference and its associated standard error as a PARMS= data set, which will subsequently serve as an input data set for PROC SEQTEST. Note that all of the variables are required for a PARMS= data set, as described in the section "PARMS < (TESTVAR= *variable*) > = SAS Data Set" on page 6891.

```
data Parms_LDL1;
   set Parms_LDL1;
   if Variable='Trt';
   _Scale_='MLE';
   _Stage_= 1;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_LDL1;
   title 'Statistics Computed at Stage 1';
run;
```

Figure 81.9 displays the statistics computed at stage 1.

Figure 81.9 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Trt	-2.52591	5.68572	MLE	1

Since the sample sizes derived are based on the estimated variance at the designing phase, the information level that corresponds to the test statistic at stage 1 is estimated by

$$I_1 = \frac{1}{s_1^2} = \frac{1}{5.686^2} = 0.0309$$

where s_1 is the standard error of the treatment estimate.

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_LDL
             Parms(Testvar=Trt)=Parms_LDL1
             infoadj=prop
             ;
ods output Test=Test_LDL1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The PARMS=PARMS_LDL1 option specifies the input data set PARMS_LDL1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=TRT option identifies the test variable TRT in the data set.

By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the maximum information and the Type I error level are maintained. Furthermore, with the INFOADJ=PROP option (which is the default), the information levels at future interim stages (2 and 3) are adjusted proportionally from the levels provided in the BOUNDARY= data set.

The ODS OUTPUT statement with the TEST=TEST_LDL1 option creates an output data set named TEST_LDL1 which contains the updated boundary information for the test at stage 1, and the boundary information that is needed for the group sequential test at the next stage. See the section "Boundary Adjustments for Information Levels" on page 6895 for details.

The "Design Information" table in Figure 81.10 displays design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are adjusted for the updated information levels to maintain the Type I α level, and the maximum information remains the same as in the BOUNDARY= data set. But the derived Type II error probability β and power $1 - \beta$ are slightly different with new information levels. With the updated power $1 - \beta$, the corresponding fixed-sample design is also updated.

Figure 81.10 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_LDL
Data Set	WORK.PARMS_LDL1
Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Number of Stages	4
Alpha	0.05
Beta	0.10074
Power	0.89926
Max Information (Percent of Fixed Sample)	102.4815
Max Information	0.10740291
Null Ref ASN (Percent of Fixed Sample)	101.7765
Alt Ref ASN (Percent of Fixed Sample)	75.4928

The "Test Information" table in Figure 81.11 displays the boundary values for the test statistic. By default (or equivalently if you specify BOUNDARYSCALE=STDZ), these statistics are displayed with the standardized Z scale. With the INFOADJ=PROP option (which is the default), information levels at future interim stages are derived proportionally from the corresponding levels provided in the BOUNDARY= data set.

Figure 81.11 Sequential Tests

Test Information (Standardized Z Scale)
Null Reference = 0

	--Information Level-		------Alternative----- Reference------		----Boundary Values--- ---Lower--	---Upper--
Stage	Proportion	Actual	Lower	Upper	Alpha	Alpha
1	0.2880	0.030934	-1.75879	1.75879	-3.39532	3.39532
2	0.5253	0.056423	-2.37536	2.37536	-2.77374	2.77374
3	0.7627	0.081913	-2.86205	2.86205	-2.32412	2.32412
4	1.0000	0.107403	-3.27724	3.27724	-2.03147	2.03147

Test Information (Standardized Z Scale)
Null Reference = 0

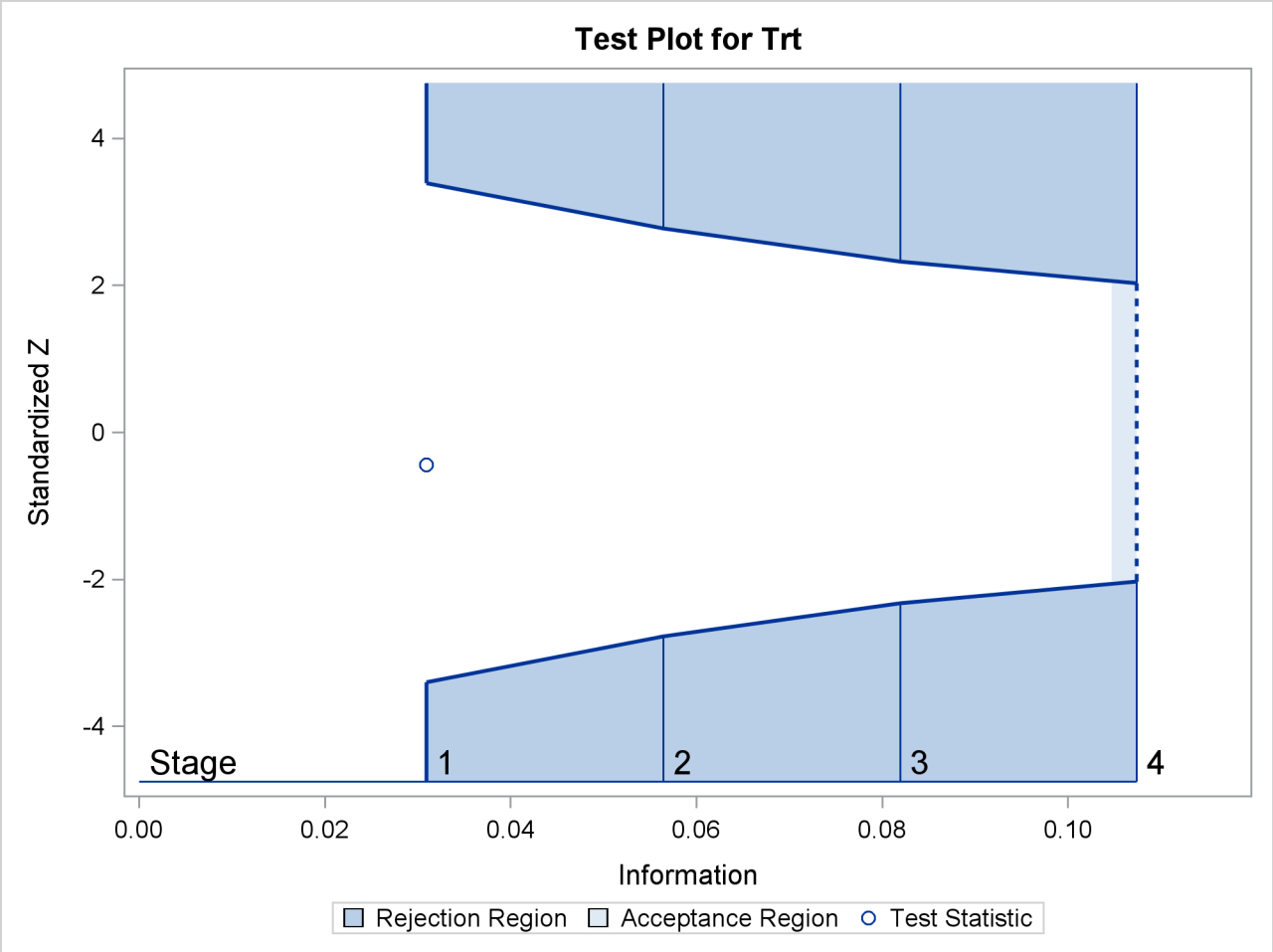
	----------Test---------- ----------Trt----------	
Stage	Estimate	Action
1	-0.44426	Continue
2	.	
3	.	
4	.	

At stage 1, the standardized Z statistic -0.44426 is between the lower and upper α boundary values, and so the trial continues to the next stage. With the observed information level at stage 1, $I_1 = 0.0309$ (which is not substantially different from the target information level at stage 1), the trial continues to the next stage without adjustment of the sample size according to the study plan.

If an observed information level is different from its target level at an interim stage, the sample sizes at future stages can be adjusted to achieve the target maximum information level according to the study plan. That is, a study plan might modify the final sample size to achieve the target maximum information level if the observed information level is different from its target level by a specified amount at the interim stage. For example, if the variance estimate is used to compute the required sample size of a two-sample Z test for mean difference, the study plan might use the current variance estimate to update the required sample size for the trial (Jennison and Turnbull 2000, p. 295). See the section "Applicable Two-Sample Tests and Sample Size Computation" in "The SEQDESIGN Procedure" for a description of how to compute the sample size from the variance estimate.

If ODS Graphics is enabled, a detailed test plot with the rejection and acceptance regions is displayed, as shown in Figure 81.12. This plot displays the boundary values in the "Test Information" table in Figure 81.11. The stages are indicated by vertical lines with accompanying stage numbers. The horizontal axis indicates the information levels for the stages. As expected, the test statistic is in the continuation region between the lower and upper α boundaries.

Figure 81.12 Sequential Test Plot



The following statements use the REG procedure with the data available at the first two stages to estimate the mean treatment difference and its associated standard error at stage 2:

```
proc reg data=LDL_2;
   model Ldl=Trt;
ods output ParameterEstimates=Parms_LDL2;
run;
```

The following statements create and display (in Figure 81.13) the data set for the mean treatment difference and its associated standard error:

```
data Parms_LDL2;
   set Parms_LDL2;
   if Variable='Trt';
   _Scale_='MLE';
   _Stage_= 2;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_LDL2;
   title 'Statistics Computed at Stage 2';
run;
```

Figure 81.13 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Trt	-8.37628	4.24405	MLE	2

Using the standard error for the treatment estimate available at stage 2, the information level that corresponds to the test statistic at stage 2 is estimated by

$$I_2 = \frac{1}{s_2^2} = \frac{1}{4.244^2} = 0.0555$$

where s_2 is the standard error of the treatment estimate at stage 2.

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
proc seqtest Boundary=Test_LDL1
             Parms(Testvar=Trt)=Parms_LDL2
             infoadj=prop
             ;
ods output Test=Test_LDL2;
run;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=TEST_LDL2 option creates an output data set named TEST_LDL2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Test Information" table in Figure 81.14 displays the boundary values for the test statistic with the default standardized Z scale

Figure 81.14 Sequential Tests

The SEQTEST Procedure

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	--Information Level- Proportion	Actual	------Alternative----- -------Reference------ Lower	Upper	----Boundary Values--- ---Lower-- Alpha	---Upper-- Alpha
1	0.2880	0.030934	-1.75879	1.75879	-3.39532	3.39532
2	0.5169	0.055519	-2.35624	2.35624	-2.78456	2.78456
3	0.7585	0.081461	-2.85413	2.85413	-2.32908	2.32908
4	1.0000	0.107403	-3.27724	3.27724	-2.03097	2.03097

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	----------Test---------- ----------Trt---------- Estimate	Action
1	-0.44426	Continue
2	-1.97365	Continue
3	.	
4	.	

At stage 2, the standardized test statistic, $z = -8.37628/4.24405 = -1.97365$, is between its corresponding lower and upper α boundary values. Therefore, the trial continues to the next stage.

The following statements use the REG procedure with the data available at the first three stages to estimate the mean treatment difference and its associated standard error at stage 3:

```
proc reg data=LDL_3;
   model Ldl=Trt;
ods output ParameterEstimates=Parms_LDL3;
run;
```

The following statements create and display (in Figure 81.15) the data set for the mean treatment difference and its associated standard error:

```
data Parms_LDL3;
   set Parms_LDL3;
   if Variable='Trt';
   _Scale_='MLE';
   _Stage_= 3;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_LDL3;
   title 'Statistics Computed at Stage 3';
run;
```

Figure 81.15 Statistics Computed at Stage 3

Statistics Computed at Stage 3

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Trt	-9.21369	3.42149	MLE	3

The following statements invoke the SEQTEST procedure to test for early stopping at stage 3:

```
ods graphics on;
proc seqtest Boundary=Test_LDL2
             Parms(Testvar=Trt)=Parms_LDL3
             infoadj=prop
             ;
ods output Test=Test_LDL3;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 3, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 3, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=TEST_LDL3 option creates an output data set named TEST_LDL3 which contains the updated boundary information for the test at stage 3. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Test Information" table in Figure 81.16 displays the boundary values for the test statistic with the default standardized Z scale.

Figure 81.16 Sequential Tests

The SEQTEST Procedure

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	--Information Level- Proportion	Actual	------Alternative----- -------Reference------ Lower	Upper	----Boundary Values--- ---Lower-- Alpha	---Upper-- Alpha
1	0.2880	0.030934	-1.75879	1.75879	-3.39532	3.39532
2	0.5169	0.055519	-2.35624	2.35624	-2.78456	2.78456
3	0.7953	0.085422	-2.92271	2.92271	-2.25480	2.25480
4	1.0000	0.107403	-3.27724	3.27724	-2.04573	2.04573

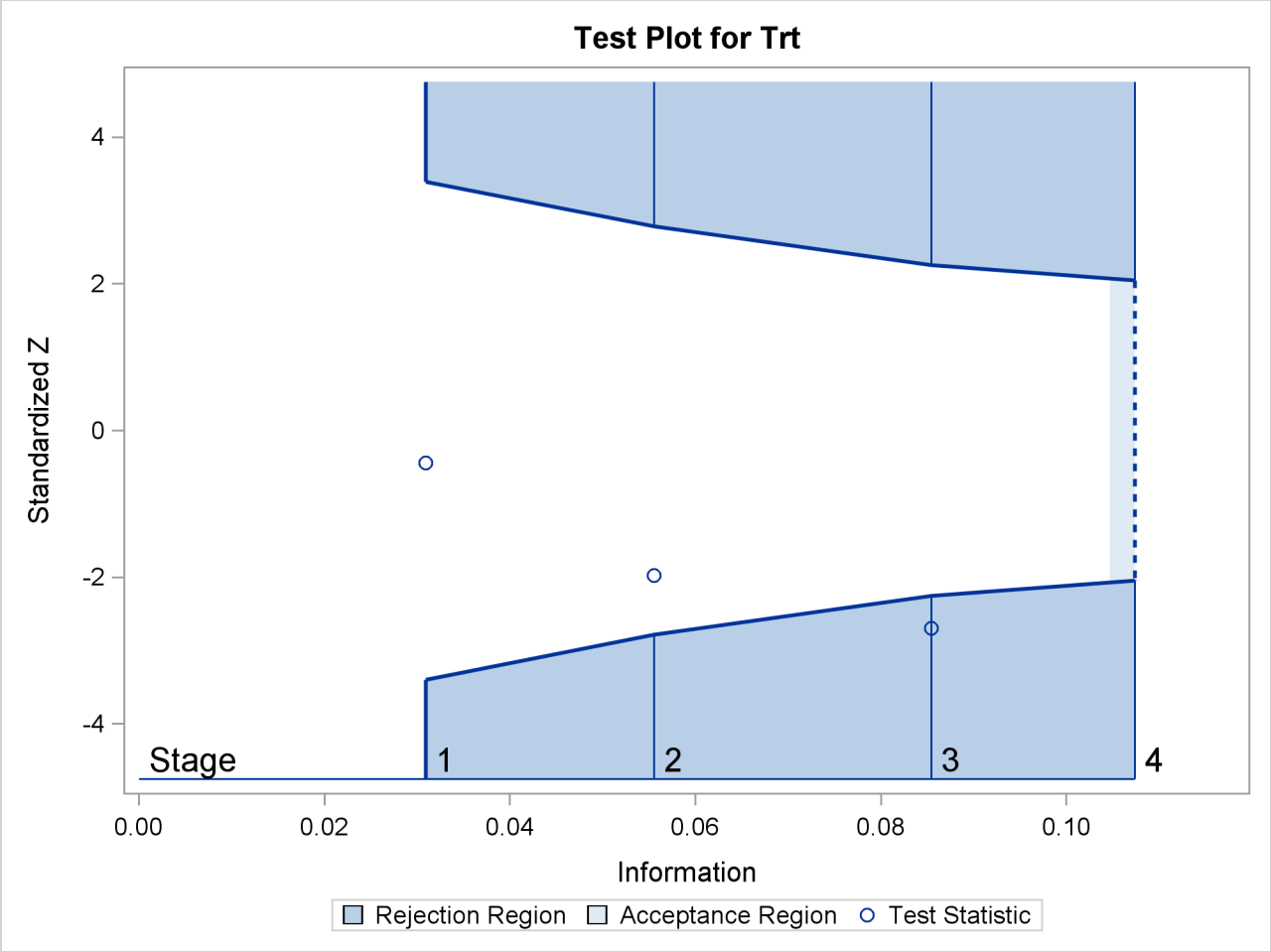
Test Information (Standardized Z Scale)
Null Reference = 0

Stage	----------Test---------- ----------Trt---------- Estimate	Action
1	-0.44426	Continue
2	-1.97365	Continue
3	-2.69289	Reject Null
4	.	

The sequential test stops at stage 3 to reject the null hypothesis for the lower alternative because the test statistic -2.69289 is less than the corresponding upper α boundary -2.25480. That is, the test demonstrates significant beneficial effect for the new drug.

The "Test Plot" displays boundary values for the design and the test statistic at the first three stages, as shown in Figure 81.17. It shows that the test statistic is in the "Rejection Region" below the lower α boundary at stage 3.

Figure 81.17 Sequential Test Plot



When a trial stops, the "Parameter Estimates" table in Figure 81.18 displays the stopping stage, parameter estimate, unbiased median estimate, confidence limits, and p-value under the null hypothesis $H_0 : \theta = 0$. As expected, the p-value 0.0108 is significant at the two-sided α level, $\alpha = 0.05$, and the confidence interval does not contain the value zero. The p-value, unbiased median estimate, and confidence limits depend on the ordering of the sample space (k, z), where k is the stage number and z is the standardized Z statistic. See the section "Analysis after a Sequential Test" on page 6901 for a detailed description of these statistics.

Figure 81.18 Parameter Estimates

Parameter Estimates
Stagewise Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate
Trt	3	-9.213692	0.0108	-9.022891

Parameter Estimates
Stagewise Ordering

Parameter	95% Confidence Limits	
Trt	-15.79845	-2.13138

Syntax: SEQTEST Procedure

The following PROC SEQTEST statement is required for the SEQTEST procedure:

PROC SEQTEST < *options* > ;

PROC SEQTEST Statement

Table 81.1 summarizes the options in the PROC SEQTEST statement.

Table 81.1 Summary of PROC SEQTEST Options

Option	Description
Input Data Sets	
BOUNDARY=	Specifies the data set for boundary information
DATA=	Specifies the data set for parameter estimates and information levels
PARMS=	Specifies the data set for parameter estimates and standard errors
Boundaries	
BETAOVERLAP=	Checks for overlapping of the lower and upper β boundaries at the current and subsequent interim stages in a two-sided design
BOUNDARYKEY=	Specifies the boundary key to maintain Type I and II error probability levels
BOUNDARYSCALE=	Specifies the boundary scale
ERRSPENDADJ=	Specifies error spending methods for boundary adjustments
ERRSPENDMIN=	Specifies minimum error spending values for the boundaries
INFOADJ=	Specifies whether information levels at future interim stages should be adjusted

Table 81.1 *continued*

Option	Description
NSTAGES=	Specifies the number of stages
Test Statistics	
DATA(TESTVAR=)	Specifies the test variable in the DATA= data set
PARMS(TESTVAR=)	Specifies the test variable in the PARMS= data set
***p*-Values and Confidence Intervals**	
CIALPHA=	Specifies the significance levels for the confidence interval
CITYPE=	Specifies the types of confidence interval
ORDER=	Specifies the ordering of the sample space used to derive the *p*-values and confidence limits
Table Output	
CONDPOWER	Displays conditional powers
ERRSPEND	Displays the cumulative error spending at each stage
PREDPOWER	Displays the predictive powers
PSS	Displays the powers and expected sample sizes
RCI	Displays the repeated confidence intervals
STOPPROB	Displays the expected cumulative stopping probabilities
Graphics Output	
PLOTS=ASN	Displays the expected sample numbers plot
PLOTS=CONDPOWER	Displays the conditional powers plot
PLOTS=ERRSPEND	Displays the error spending plot
PLOTS=POWER	Displays the powers plot
PLOTS=RCI	Displays the repeated confidence intervals plot
PLOTS=TEST	Displays the boundary plot with test statistics

The BOUNDARY= option provides the information for the design and is required in the PROC SEQTEST statement. By default, the SEQTEST procedure displays tables of design information and test information. If ODS Graphics is enabled, the procedure also displays a sequential test plot.

The following options can be used in the PROC SEQTEST statement. They are listed in alphabetical order.

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 at the current and subsequent interim stages. This option applies to two-sided designs with early stopping to accept H_0, or to either accept or reject H_0. 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 at the current and subsequent interim stages. 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 6898.

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

BOUNDARY=*SAS-data-set*

names the required SAS data set that contains the design boundary information. At stage 1, the data set is usually created from the "Boundary Information" table created by the SEQDESIGN procedure. At each subsequent stage, the data set is usually created from the "Test Information" table created by the SEQTEST procedure at the previous stage. The data set includes the variables _Scale_ for the boundary scale, _Stop_ for the stopping criterion, and _ALT_ for the type of alternative hypothesis. It also includes _Stage_ for the stage number, Info_Prop for the information proportion, and a set of the boundary variables from Bound_LA, Bound_LB, Bound_UB, and Bound_UA for boundary values at each stage.

The data set might also include _Info_ for the actual information level, NObs for the number of observation, and Events for the number of events required at each stage.

BOUNDARYKEY=ALPHA | BETA | BOTH

specifies the boundary key to be maintained in the boundary adjustments. The BOUNDARYKEY=ALPHA option maintains the Type I α level and derives the Type II error probability, and the BOUNDARYKEY=BETA option maintains the Type II β level and derives the Type I error probability. The BOUNDARYKEY=BOTH option maintains both α and β levels simultaneously by deriving a new maximum information. The default is BOUNDARYKEY=ALPHA.

BOUNDARYSCALE=MLE | SCORE | STDZ | PVALUE

BSCALE=MLE | SCORE | STDZ | PVALUE

specifies the boundary scale to be displayed in the output boundary table and plot. The BOUNDARYSCALE=MLE, BOUNDARYSCALE=SCORE, BOUNDARYSCALE=STDZ, and BOUNDARYSCALE=PVALUE options correspond to the boundary with the maximum likelihood estimator scale, score statistic scale, standardized normal Z scale, and p-value scale, respectively. The default is BOUNDARYSCALE=STDZ.

With the BOUNDARYSCALE=MLE or BOUNDARYSCALE=SCORE option, either the MAXINFO= option must be specified or the _Info_ variable must be in the BOUNDARY= data set to provide the necessary information level at each stage to derive the boundary values. Usually, these values are obtained from analysis output in SAS procedures.

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.

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

specifies the significance levels for the confidence interval, where $0 < \alpha < 1$, $0 < \alpha_l < 0.5$, and $0 < \alpha_u < 0.5$. The default is CIALPHA= 0.05.

For a lower confidence interval (CITYPE=LOWER), the CIALPHA=α option produces a $(1 - \alpha)$ lower confidence interval. For an upper confidence interval (CITYPE=UPPER), the CIALPHA=α option produces a $(1 - \alpha)$ upper confidence interval. The LOWER= and UPPER= suboptions are applicable only for a two-sided confidence interval (CITYPE=TWOSIDED). The LOWER= suboption specifies the lower significance level α_l and the upper significance level $\alpha_u = 1 - \alpha_l$. The UPPER= suboption specifies the upper significance level α_u and the lower significance level $\alpha_l = 1 - \alpha_u$. If both LOWER= and UPPER= suboptions are not specified, $\alpha_l = \alpha_u = \alpha/2$. The significance levels α_l and α_u are then used for the $(1 - \alpha_l)$ lower confidence limit and $(1 - \alpha_u)$ upper confidence limit, respectively.

CITYPE=LOWER | UPPER | TWOSIDED

specifies the type of confidence interval. The CITYPE=LOWER, CITYPE=UPPER, and CITYPE=TWOSIDED options correspond to the lower confidence interval, upper confidence interval, and two-sided confidence interval, respectively. The default is CITYPE=LOWER for the design with an upper alternative, CITYPE=UPPER for the design with a lower alternative, and CITYPE=TWOSIDED for the design with a two-sided alternative.

DATA < (TESTVAR=*variable***) >=***SAS-data-set*

names the SAS data set that contains the test statistic and its associated information level for the stage. The data set includes the stage variable _Stage_ and a variable to identify or derive the information level: _Info_ for the information level, NObs for the number of observation, or Events for the number of events. If the information level that corresponds to the test statistic is not available, the information level derived in the BOUNDARY= data set is used.

If the TESTVAR= option is specified, the data set also includes the test variable specified in the TESTVAR= option and the scale variable _Scale_ for the test statistic. Usually, these test variable values are obtained from analysis output in SAS procedures.

ERRSPENDADJ=*method*

ERRSPENDADJ(*boundary***)=***method*

BOUNDARYADJ=*method*

BOUNDARYADJ(*boundary***)=***method*

specifies methods to compute the error spending values at the current and future interim stages for the boundaries. This option is applicable only if the observed information level at the current stage does not match the value provided in the BOUNDARY= data set. These error spending values are then used to derive the updated boundary values. The default is ERRSPENDADJ=ERRLINE. Note that the information levels at future interim stages are determined by the INFOADJ= option.

The following options specify available error spending methods for boundary adjustment:

NONE

specifies that the cumulative error spending at each interim stage not be changed, even if the corresponding information level has been changed.

ERRLINE

specifies the linear interpolation method for the adjustment.

ERRFUNCGAMMA < (GAMMA= γ **) >**

specifies the gamma function method for the adjustment. The GAMMA= suboption specifies the γ parameter in the function, where $\gamma \leq 3$. The default is GAMMA=-2.

ERRFUNCOBF

specifies the approximate O'Brien-Fleming cumulative error spending function for the adjustment.

ERRFUNCPOC

specifies the approximate Pocock cumulative error spending function for the adjustment.

ERRFUNCPOW < (RHO= ρ **) >**

specifies the power function method for the adjustment. The RHO= suboption specifies the power parameter ρ in the function, where $\rho \geq 0.25$. The default is RHO=2.

See the section "Boundary Adjustments for Information Levels" on page 6895 for a detailed description of the available error spending methods for boundary adjustment in the SEQTEST procedure.

If an error spending method for boundary adjustments is used for all boundaries in a group sequential test, you can use the ERRSPENDADJ=*method* option to specify the method. Otherwise, you can use the following ERRSPENDADJ(*boundary*)=*method* options to specify different methods for the boundaries.

ERRSPENDADJ(ALPHA)=*method*

ERRSPENDADJ(REJECT)=*method*

BOUNDARYADJ(ALPHA)=*method*

BOUNDARYADJ(REJECT)=*method*

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

ERRSPENDADJ(LOWERALPHA)=*method*

ERRSPENDADJ(LOWERREJECT)=*method*

BOUNDARYADJ(LOWERALPHA)=*method*

BOUNDARYADJ(LOWERREJECT)=*method*

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

ERRSPENDADJ(UPPERALPHA)=*method*

ERRSPENDADJ(UPPERREJECT)=*method*

BOUNDARYADJ(UPPERALPHA)=*method*

BOUNDARYADJ(UPPERREJECT)=*method*

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

ERRSPENDADJ(BETA)=*method*

ERRSPENDADJ(ACCEPT)=*method*

BOUNDARYADJ(BETA)=*method*

BOUNDARYADJ(ACCEPT)=*method*

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

ERRSPENDADJ(LOWERBETA)=*method*

ERRSPENDADJ(LOWERACCEPT)=*method*

BOUNDARYADJ(LOWERBETA)=*method*

BOUNDARYADJ(LOWERACCEPT)=*method*

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

ERRSPENDADJ(UPPERBETA)=*method*

ERRSPENDADJ(UPPERACCEPT)=*method*

BOUNDARYADJ(UPPERBETA)=*method*

BOUNDARYADJ(UPPERACCEPT)=*method*

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

ERRSPENDMIN=*numbers*

ERRSPENDMIN(*boundary***)=***numbers*

specifies the minimum error spending values at the current observed and future interim stages for the boundaries specified in the BOUNDARYKEY= option. The default is ERRSPENDMIN=0.

If a set of numbers is used for each boundary in the design, you can use the ERRSPENDMIN=*numbers* option. Otherwise, you can use the following ERRSPENDMIN(*boundary*)=*numbers* options to specify different sets of minimum error spending values for the boundaries. For a boundary, the error spending value at stage 1 is identical to its nominal p-value.

ERRSPENDMIN(ALPHA)=*numbers*

ERRSPENDMIN(REJECT)=*numbers*

specifies the minimum error spending values for the α boundary of a one-sided design or the lower and upper α boundaries of a two-sided design.

ERRSPENDMIN(LOWERALPHA)=*numbers*

ERRSPENDMIN(LOWERREJECT)=*numbers*

specifies the minimum error spending values for the lower α boundary of a two-sided design.

ERRSPENDMIN(UPPERALPHA)=*numbers*

ERRSPENDMIN(UPPERREJECT)=*numbers*

specifies the minimum error spending values for the upper α boundary of a two-sided design.

ERRSPENDMIN(BETA)=*numbers*

ERRSPENDMIN(ACCEPT)=*numbers*

specifies the minimum error spending values for the β boundary of a one-sided design or the lower and upper β boundaries of a two-sided design.

ERRSPENDMIN(LOWERBETA)=*numbers*

ERRSPENDMIN(LOWERACCEPT)=*numbers*

specifies the minimum error spending values for the lower β boundary of a two-sided design.

ERRSPENDMIN(UPPERBETA)=*numbers*

ERRSPENDMIN(UPPERACCEPT)=*numbers*

specifies the minimum error spending values for the upper β boundary of a two-sided design.

INFOADJ=NONE | PROP

specifies whether information levels at future interim stages are to be adjusted. If you specify INFOADJ=NONE, no adjustment is made, and the information levels are preserved at the levels provided in the BOUNDARY= data set. If you specify INFOADJ=PROP (which is the default), the information levels are adjusted proportionally from the levels provided in the BOUNDARY= data set. The section "Information Level Adjustments at Future Stages" on page 6895 describes how the adjustments are computed.

Note that if you specify BOUNDARYKEY=BOTH, the INFOADJ=NONE option is not applicable, and the INFOADJ=PROP option is used to adjusted the information levels at future stages proportionally from the levels provided in the BOUNDARY= data set to maintain both α and β levels.

NSTAGES=*number*

specifies the number of stages for the clinical trial. The default is the number derived from the BOUNDARY= data set.

The specified NSTAGES= number might or might not be the same as the number derived in the BOUNDARY= data set. You can use the NSTAGES= option to set the next stage as the final stage to compute the conditional power, as described in the section "Conditional Power Approach" on page 6899.

ORDER=LR | MLE | STAGEWISE

specifies the ordering of the sample space (k, z), where k is the stage number and z is the observed standardized Z statistic. The ordering is used to derive the p-values for the observed (k, z) statistic and to create unbiased median estimate and confidence limits from the statistic. The ORDER=LR option specifies the LR ordering that compares the distances between observed standardized Z statistics and their corresponding hypothetical values, the ORDER=MLE option specifies the MLE ordering that compares values in the MLE scale, and the ORDER=STAGEWISE specifies the stagewise ordering that uses counterclockwise ordering around the continuation region. The default is ORDER=STAGEWISE. See the section "Available Sample Space Orderings in a Sequential Test" on page 6902 for a detailed description of these sample space orderings.

PARMS < (TESTVAR=*variable***) > =***SAS-data-set*

names the SAS data set that contains the parameter estimate and its associated standard error for the stage. The data set includes the stage variable _Stage_, the test statistic Estimate, the standard error of the estimate StdErr, and the test statistic scale variable _Scale_. The standard error is are used to derive the information level. If the standard error is not available, the information level derived in the BOUNDARY= data set is used.

The data set also includes the variable Parameter, Effect, Variable, or Parm that contains the test variable specified in the TESTVAR= option. Usually, these test variable values are obtained from analysis output in SAS procedures.

Table Output Options

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

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

displays conditional powers given the most recently observed statistic under specified hypothetical references, where the numbers $c_i \geq 0$. In the SEQTEST procedure, the conditional power is the probability that the test statistic at the final stage would exceed the rejection critical value given the observed statistic.

If interim stages exist between the current stage and the final stage, the conditional power is not the conditional probability to reject the null hypothesis H_0. In this case, you can set the next stage as the final stage, and the conditional power is the conditional probability to reject H_0.

For a one-sided test, the powers are derived under the hypothetical references $\theta = \hat{\theta}$ and $\theta = c_i\theta_1$, where $\hat{\theta}$ is the observed statistic, θ_1 is the alternative reference, and c_i are the values specified in the CREF= option. For a two-sided test, the powers are derived under hypothetical references $\theta = \hat{\theta}$,

$\theta = c_i\theta_{1l}$, and $\theta = c_i\theta_{1u}$, where θ_{1l} is the lower alternative reference and θ_{1u} is the upper alternative reference. The default is CREF= 0 0.5 1.0 1.5.

ERRSPEND

displays the error spending at each stage for each sequential boundary.

PREDPOWER

displays predictive powers given the most recently observed statistic. The predictive power is the posterior probability that the test statistic at the final stage would exceed the rejection critical value given the observed statistic and a prior distribution of the hypothetical reference. A noninformative prior is used in the procedure.

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 with the null reference $\theta_0 = 0$, 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.

RCI

displays repeated confidence intervals for the parameter from the observed statistic at each stage. Repeated confidence intervals include both rejection and acceptance confidence intervals.

With the STOP=REJECT or STOP=BOTH option, rejection confidence limits can be derived, and the null hypothesis $H_0 : \theta = 0$ is rejected if the lower rejection confidence limit is greater than 0 or the upper rejection confidence limit is less than 0.

With the STOP=ACCEPT or STOP=BOTH option, acceptance confidence limits can be derived, and the null hypothesis is accepted with alternative hypotheses $H_{1l} : \theta = \theta_{1l}$ and $H_{1u} : \theta = \theta_{1u}$ if the upper acceptance confidence limit is less than θ_{1u} and the lower acceptance confidence limit is greater than θ_{1l}.

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

The following options can be used in the PROC SEQTEST statement to display plots 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=TEST. The global plot option ONLY suppresses the default plots and displays only plots specifically requested.

ODS Graphics must be enabled before requesting plots. For example:

```
ods graphics on;
proc seqtest Boundary=Bnd_LDL
             Parms(Testvar=Trt)=Parms_LDL1
             Plots=(test errspend);
run;
ods graphics off;
```

For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 609 in Chapter 21, "Statistical Graphics Using ODS."

The plot request options include the following.

ALL

produces all appropriate plots.

ASN < (CREF=*numbers*) >

displays a plot of the average sample numbers (expected sample sizes for nonsurvival data or expected number of events for survival data) under various hypothetical references, where the numbers $c_i \geq 0$.

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.

CONDPOWER < (CREF=*numbers*) >

displays a plot of conditional powers given the most recently observed statistic under specified hypothetical references, where the numbers $c_i \geq 0$. In the SEQTEST procedure, the conditional power is the probability that the test statistic at the final stage would exceed the rejection critical value given the observed statistic.

For a one-sided test, the powers are derived under hypothetical references $\theta = \hat{\theta}$ and $\theta = c_i\theta_1$, where $\hat{\theta}$ is the observed statistic, θ_1 is the alternative reference, and c_i are the values specified in the CREF= option. For a two-sided test, the powers are derived under hypothetical references $\theta = \hat{\theta}$, $\theta = c_i\theta_{1l}$, and $\theta = c_i\theta_{1u}$, where θ_{1l} is the lower alternative reference and θ_{1u} is the upper alternative reference. The default is CREF= 0 to 1.5 by 0.01.

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.

RCI

displays a plot of repeated confidence intervals. Repeated confidence intervals include both rejection and acceptance confidence intervals.

With the STOP=REJECT or STOP=BOTH option, rejection confidence limits can be derived and the null hypothesis $H_0 : \theta = 0$ is rejected if the lower rejection confidence limit is greater than 0 or the upper rejection confidence limit is less than 0.

With the STOP=ACCEPT or STOP=BOTH option, acceptance confidence limits can be derived and the null hypothesis is accepted with alternative hypotheses $H_{1l} : \theta = \theta_{1l}$ and $H_{1u} : \theta = \theta_{1u}$ if the upper acceptance confidence limit is less than θ_{1u} and the lower acceptance confidence limit is greater than θ_{1l}.

TEST < (HSCALE=INFO | SAMPLESIZE) >

displays a plot of the sequential boundaries and test variables. Either the information level (HSCALE=INFO) or the sample size (HSCALE=SAMPLESIZE) is displayed on the horizontal axis. The HSCALE=SAMPLESIZE option is applicable only if the sample size information is available in both the input BOUNDARY= data set and input DATA= data set. The stage number for each stage is displayed inside the plot. The default is HSCALE=INFO.

Details: SEQTEST Procedure

Input Data Sets

The BOUNDARY= data set option is required, and if neither the DATA= nor the PARMS= data set option is specified, the procedure derives statistics such as Type I and Type II error probabilities from the BOUNDARY= data set. The resulting boundaries are displayed with the scale specified in the BOUNDARYSCALE= option.

BOUNDARY= SAS Data Set

The BOUNDARY= data set provides the boundary information for the sequential test. At stage 1, the data set is usually created with an ODS OUTPUT statement from the "Boundary Information" table created by the SEQDESIGN procedure. At each subsequent stage, the data set is usually created with an ODS OUTPUT statement from the "Test Information" table that was created by the SEQTEST procedure at the previous stage. See the section "Getting Started: SEQTEST Procedure" on page 6864 for an illustration of the BOUNDARY= data set option.

The BOUNDARY= data set contains the following variables:

- _Scale_, the boundary scale, with the value MLE for the maximum likelihood estimate, STDZ for the standardized Z, SCORE for the score statistic, or PVALUE for the nominal p-value. Note that for a two-sided design, the nominal p-value is the one-sided fixed-sample p-value under the null hypothesis with a lower alternative hypothesis.
- _Stop_, the stopping criterion, with the value REJECT for rejecting the null hypothesis H_0, ACCEPT for accepting H_0, or BOTH for both rejecting and accepting H_0
- _ALT_, the type of alternative hypothesis, with the value UPPER for an upper alternative, LOWER for a lower alternative, or TWOSIDED for a two-sided alternative
- _Stage_, the stage number
- the boundary variables, a subset of Bound_LA for lower α boundary, Bound_LB for lower β boundary, Bound_UB for upper β boundary, and Bound_UA for upper α boundary
- AltRef_L, the lower alternative reference, if ALT=LOWER or ALT=TWOSIDED
- AltRef_U, the upper alternative reference, if ALT=UPPER or ALT=TWOSIDED
- _InfoProp_, the information proportion at each stage

Optionally, the BOUNDARY= data set also contains the following variables:

- _Info_, the information level at each stage
- NObs, the required number of observations for nonsurvival data at each stage
- Events, the required number of events for survival data at each stage
- Parameter, the variable specified in the DATA(TESTVAR=) or PARMS(TESTVAR=) option
- Estimate, the parameter estimate

If the BOUNDARY= data set contains the variable Parameter for the test variable that is specified in the TESTVAR= option, and the variable Estimate for the test statistics, then these test statistics are also displayed in the output test information table and output test plot.

DATA < (TESTVAR= *variable*) > = SAS Data Set

The DATA= data set provides the test variable information for the current stage of the trial. Such data sets are usually created with an ODS OUTPUT statement by using a procedure such as PROC MEANS. See "Example 81.4: Testing a Binomial Proportion" on page 6957 for an illustration of the DATA= data set option.

The DATA= data set includes the following variables:

- _Stage_, the stage number
- _Scale_, the scale for the test statistic, with the value MLE for the maximum likelihood estimate, STDZ for the standardized Z, SCORE for the score statistic, or PVALUE for the nominal p-value
- _Info_, the information level
- NObs, the number of observations for nonsurvival data at each stage
- Events, the number of events for survival data at each stage
- test variable, specified in the TESTVAR= option, contains the test variable value in the scale specified in the _Scale_ variable

With the specified DATA= data set, the procedure derives boundary values from the information levels in the _Info_ variable. If the data set does not include the _Info_ variable, then the information levels are derived from the NObs or Events variable in the DATA= data set if the variable is also in the input BOUNDARY= data set. That is, the information level at stage k is computed as $I_k^* = I_k \times (n_k^*/n_k)$, where I_k and n_k are the information level and sample size at stage k in the BOUNDARY= data set and n_k^* is the sample size at stage k in the DATA= data set. Otherwise, the information levels from the BOUNDARY= data set are used.

If the TESTVAR= option is specified, the DATA= data set must also include the test variable for the test statistic and _Scale_ variable for the corresponding scale. Note that for a two-sided design, the nominal p-value is the one-sided fixed-sample p-value under the null hypothesis with a lower alternative hypothesis.

PARMS < (TESTVAR= *variable*) > = SAS Data Set

The PARMS= data set provides a parameter estimate and associated standard error for the current stage of the trial. Such data sets are usually created with an ODS OUTPUT statement by using procedures such as the GENMOD, GLM, LOGISTIC, and REG procedures. See the section "Getting Started: SEQTEST Procedure" on page 6864 for an illustration of the PARMS= data set option.

The PARMS= data set includes the following variables:

- _Stage_, the stage number
- _Scale_, the scale for the test statistic, with the value MLE for the maximum likelihood estimate, STDZ for the standardized Z, SCORE for the score statistic, or PVALUE for the nominal p-value
- Parameter, Effect, Variable, or Parm, which contains the variable specified in the TESTVAR= option
- Estimate, the parameter estimate
- StdErr, standard error of the parameter estimate

With the specified PARMS= data set, the information level is derived from the StdErr variable. For a score statistic, the information level I_k is the variance of the statistic, $\hat{s}_k^2$, where $\hat{s}_k$ is the standard error in the StdErr variable. Otherwise, the information level is the inverse of the variance of the statistic, $\hat{s}_k^{-2}$. If the data set does not include the StdErr variable, the information levels derived from the BOUNDARY= data set are used.

If the TESTVAR= option is specified, the PARMS= data set also includes the variable Parameter, Effect, Variable, or Parm for the test variable, Estimate for the test statistic, and _Scale_ variable for the corresponding scale. Note that for a two-sided design, the nominal p-value is the one-sided fixed-sample p-value under the null hypothesis with a lower alternative hypothesis.

Boundary Variables

The boundaries created in group sequential trials depend on the type of the alternative hypothesis and the early stopping criterion. Table 81.2 shows the boundaries created with various design specifications.

Table 81.2 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 different boundaries can be generated in a group sequential design:

- the upper α boundary, used to reject the null hypothesis in favor of an upper alternative hypothesis
- the upper β boundary, used to accept the null hypothesis with an upper alternative hypothesis
- the lower β boundary, used to accept the null hypothesis with a lower alternative hypothesis
- the lower α boundary, used to reject the null hypothesis in favor of a lower alternative hypothesis

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

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 81.19 shows the boundary plot for a one-sided test with an upper alternative.

Figure 81.19 Boundary Plot for One-Sided Test

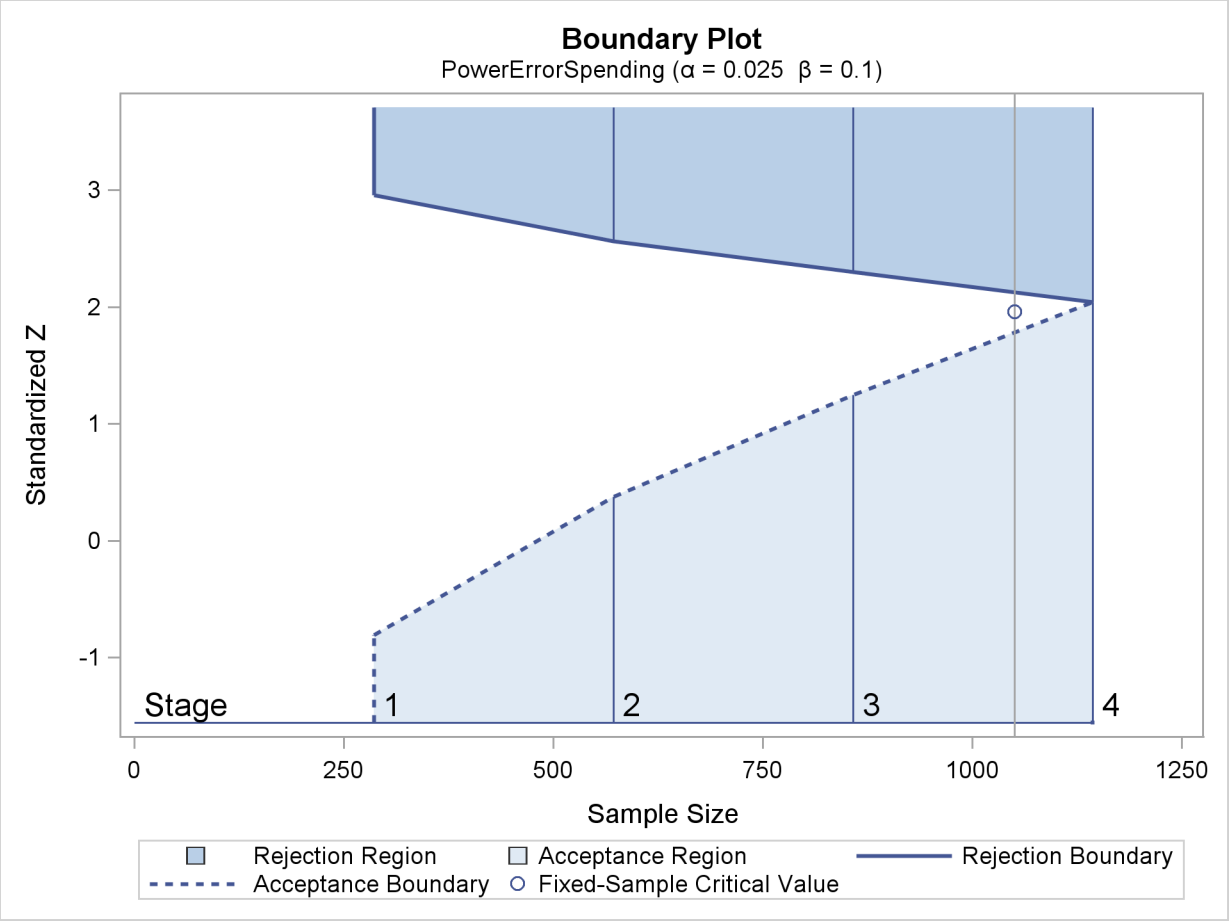


Figure 81.19 corresponds to a one-sided sequential design with early stopping to reject or accept the null hypothesis. For a sequential test with early stopping only to reject the null hypothesis, there are no acceptance boundary values at interim stages. The acceptance boundary value and its associated 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 81.20 shows the boundary plot for a two-sided test.

Figure 81.20 Boundary Plot for Two-Sided Test

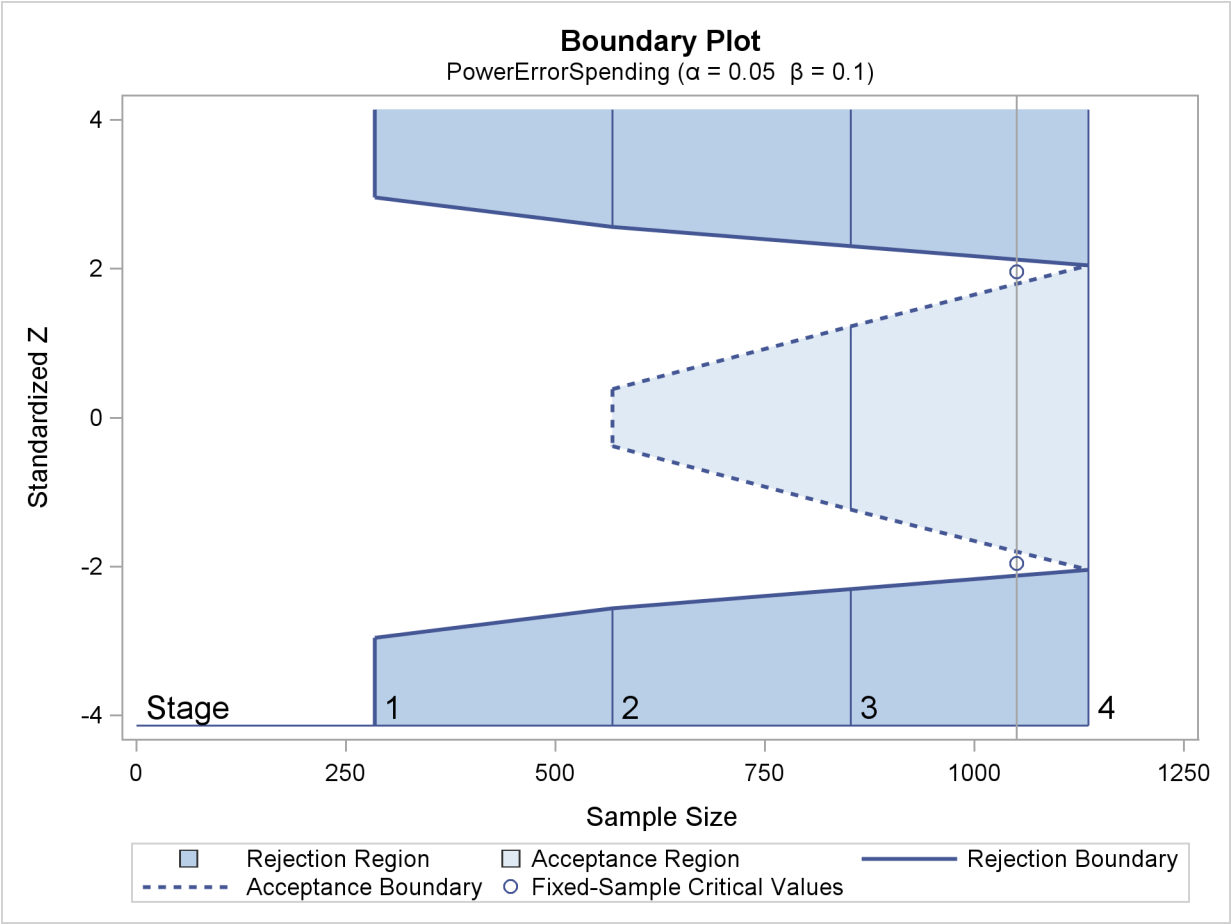


Figure 81.20 corresponds to a two-sided sequential design with early stopping to reject or accept the null hypothesis. For a sequential test with early stopping only to reject the null hypothesis, there are no acceptance boundary values at interim stages. The acceptance boundary value and its associated 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.

Information Level Adjustments at Future Stages

In a group sequential clinical trial, the information level for the observed test statistic at the current stage generally does not match the corresponding information level in the BOUNDARY= data set. By default (or equivalently if you specify INFOADJ=PROP), the SEQTEST procedure accommodates the observed information level by adjusting the information levels at future interim stages. The adjustment of information levels depends on the boundary key to be maintained in the boundary adjustments, which in turn is determined by the BOUNDARYKEY= option.

If you specify BOUNDARYKEY=ALPHA (which is the default) or BOUNDARYKEY=BETA, the maximum information level (the information level at the final stage) provided in the BOUNDARY= data set is maintained. In this case, if an observed information level at the current stage is different from the level provided in the BOUNDARY= data set, you can use the INFOADJ= option to determine whether the information levels at subsequent interim stages are to be adjusted. Specifying INFOADJ=NONE preserves the levels provided in the BOUNDARY= data set without adjustment. Specifying INFOADJ=PROP proportionally adjusts the levels provided in the BOUNDARY= data set as follows.

Denote the information level at stage k for the K-stage design that is stored in the BOUNDARY= data set by I_k, $k = 1, 2, \ldots, K$. Also denote the information level that corresponds to the test statistic at an interim stage k_0 by I'_{k_0}, $1 \leq k_0 \leq (K-1)$. Then for the updated design, the information level at stage k, $k = k_0 + 1, \ldots, (K-1)$, is computed as

$$I'_k = I'_{k_0} + (I_K - I'_{k_0}) \frac{I_k - I_{k_0}}{I_K - I_{k_0}}$$

Note that if $I'_{k_0} \geq I_K$, the information level at stage k_0 reaches the maximum information level in the design, the trial stops at stage k_0, and no future information levels are derived.

If you specify BOUNDARYKEY=BOTH, the maximum information level for the trial is not necessarily the same as the maximum information level saved in the BOUNDARY= data set. In this case, the INFOADJ=NONE option is not applicable, and the INFOADJ=PROP option is used to proportionally adjust the information levels at future interim stages with the updated maximum information I'_K. That is, with an updated I'_K, the information level at a future interim stage k is computed as

$$I'_k = I'_{k_0} + (I'_K - I'_{k_0}) \frac{I_k - I_{k_0}}{I'_K - I_{k_0}}$$

Boundary Adjustments for Information Levels

In a group sequential clinical trial, if the information level for the observed test statistic does not match the corresponding information level in the BOUNDARY= data set, the INFOADJ=PROP option (which is the default) can be used to modify information levels at future stages to accommodate this observed information level. With the adjusted information levels, the ERRSPENDADJ= option provides various methods to compute error spending values at the current and future interim stages. These error spending values are then used to derive boundary values in the SEQTEST procedure. See the section "Error Spending Methods" in the chapter "The SEQDESIGN Procedure" for a detailed description of how to use these error spending values to derive boundary values.

The ERRSPENDADJ=NONE option keeps the error spending the same at each stage. The ERRSPENDADJ=ERRLINE option uses a linear interpolation on the cumulative error spending in the design stored in the BOUNDARY= data set to derive the error spending for each unmatched information level (Kittelson and Emerson 1999, p. 882). That is, the cumulative error spending for an information level I is computed as

$$e(I) = \begin{cases} e_1 \left(\frac{I}{I_1}\right) & \text{if } I < I_1 \\ e_j + (\alpha_{j+1} - \alpha_j) \left(\frac{I - I_j}{I_{j+1} - I_j}\right) & \text{if } I_j \le I < I_{j+1} \\ e_K & \text{if } I \ge I_K \end{cases}$$

where $e_1, e_2, \ldots, e_K$ are the cumulative errors at the K stages of the design that is stored in the BOUNDARY= data set.

The ERRSPENDADJ=ERRFUNCPOC option uses Pocock-type cumulative error spending function (Lan and DeMets 1983):

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

With an error level of α or β, the cumulative error spending for an information level I is $e(I) = \alpha\, E(I/I_K)$ or $e(I) = \beta\, E(I/I_K)$.

The ERRSPENDADJ=ERRFUNCOBF option uses O'Brien-Fleming-type cumulative error spending function (Lan and DeMets 1983):

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

where a is either α for the α spending function or β for the β spending function, and Φ is the cumulative distribution function of the standardized Z statistic. That is, with an error level of α or β, the cumulative error spending for an information level I is $e(I) = \alpha\, E(I/I_K; \alpha)$ or $e(I) = \beta\, E(I/I_K; \beta)$.

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

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

where γ is the parameter γ specified in the GAMMA= option. That is, with an error level of α or β, the cumulative error spending for an information level I is $e(I) = \alpha\, E(I/I_K; \gamma)$ or $e(I) = \beta\, E(I/I_K; \gamma)$.

The ERRSPENDADJ=ERRFUNCPOW option uses power cumulative error spending function (Jennison and Turnbull 2000, p. 148):

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

where ρ is the power parameter specified in the RHO= suboption. With an error level of α or β, the cumulative error spending for an information level I is $e(I) = \alpha\, E(I/I_K; \rho)$ or $e(I) = \beta\, E(I/I_K; \rho)$.

If the BOUNDARYKEY=BOTH option is specified, the maximum information required for the trial might not be the same as the maximum information level stored in the BOUNDARY= data set. In this case, the information levels at future stages are adjusted proportionally, and the same error spending values that were computed based on the maximum information level stored in the BOUNDARY= data set are used to derive boundary values for the trial.

If an error spending function is used to create boundaries for the design in the SEQDESIGN procedure, then in order to better maintain the design features throughout the group sequential trial, the same error spending function to create boundaries for the design in the SEQDESIGN procedure should be used to modify boundaries in the SEQTEST procedure at each subsequent stage.

Boundary Adjustments for Minimum Error Spending

In a group sequential clinical trial, boundary values created from a design such as an O'Brien-Fleming design might be too conservative in early stages. Thus the trial is unlikely to stop in early stages. Lan and Demets (1983, p. 662) suggest truncating boundary values to a number such as 3.5 for the trial to have a reasonable probability of stopping at early stages. Instead of truncating boundary values by a specified number, the ERRSPENDMIN= option provides individual minimum error spending at each interim stage to stop the trial early.

For a K-stage trial, denote the derived cumulative error spending at stage k after adjusting for information levels by $e_k, k = 1, 2, \ldots, K$. Also denote the specified minimum error spending at interim stage k by $\epsilon_k, k = 1, 2, \ldots, K-1$. Then the cumulative error spending at stage 1 is $e'_1 = \max(e_1,\ \epsilon_1)$. If $e_1 < e'_1$, the error spending values at subsequent interim stages are adjusted proportionally by

$$e'_j = e'_1 + \frac{e_j - e_1}{e_K - e_1}\,(e_K - e'_1)$$

for $j = 2, \ldots, K-1$.

The process is repeated at each subsequent interim stage. That is, at stage $k, k = 2, \ldots, K-1$, denote the updated cumulative β spending at stage j by e_j, $j = k, k+1, \ldots, K$. Then the cumulative error spending at stage k is $e'_k = \max(e_k,\ e'_{k-1} + \epsilon_k)$. If $e_k < e'_k$, the error spending values at subsequent interim stages are adjusted proportionally by

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

for $j = k+1, \ldots, K-1$.

Note that the ERRSPENDMIN= option is applicable only to the boundaries specified in the BOUNDARYKEY= option. That is, the ERRSPENDMIN= option is applicable to the α boundaries with BOUNDARYKEY=ALPHA or BOUNDARYKEY=BOTH, and it is applicable to the β boundaries with BOUNDARYKEY=BETA or BOUNDARYKEY=BOTH.

Boundary Adjustments for Overlapping Lower and Upper β Boundaries

In the SEQTEST procedure, the α and β spending values at the stages are used to derive the boundary values for the trial. For a two-sided design with early stopping to accept H_0, or to either reject or accept H_0, a zero β spending at an interim stage sets the β boundary values to missing. A small β spending at the current or subsequent 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 that is derived from the one-sided test for the upper alternative is less than the lower β boundary value that is 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 the current and subsequent 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 cumulative β spending values at 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.

Stochastic Curtailment

Lan, Simon, and Halperin (1982) introduce stochastic curtailment to stop a trial if, given current data, it is likely to predict the outcome of the trial with high probability. That is, a trial can be stopped to reject the null hypothesis H_0 if, given current data in the analyses, the conditional probability of rejecting H_0 under H_0 at the end of the trial is greater than γ, where the constant γ should be between 0.5 and 1 and values of 0.8 or 0.9 are recommended (Jennison and Turnbull 2000, p. 206). Similarly, a trial can be stopped to accept the null hypothesis H_0 if, given current data in the analyses, the conditional probability of rejecting H_0 under the alternative hypothesis H_1 at the end of the trial is less than γ.

The following two approaches for stochastic curtailment are available in the SEQTEST procedures: conditional power approach and predictive power approach. For each approach, the derived group sequential test is used as the reference test for rejection.

Conditional Power Approach

In the SEQTEST procedure, the conditional power at an interim stage k is the probability that the test statistic at the final stage (stage K) would exceed the rejection critical value (Cui, Hung, and Wang 1999, p. 854; Emerson, Kittelson, and Gillen 2005, p. 13). If there exist interim stages between the kth stage and the final stage, $k < K - 1$, the conditional power is not the conditional probability to reject the null hypothesis H_0. In this case, you can set the next stage as the final stage, and the conditional power is the conditional probability to reject H_0.

The conditional distribution of Z_K given the observed statistic z_k at the kth stage and the hypothetical reference θ is

$$Z_K \,|\, (z_k, \theta) \;\sim\; N\left(z_k \;\Pi_k^{\frac{1}{2}} + \theta \; I_X^{\frac{1}{2}} \left(1 - \Pi_k\right),\; 1 - \Pi_k \right)$$

where $\Pi_k = I_k / I_X$ is the fraction of information at the kth stage.

The power for the upper alternative, prob($Z_K > a_K \,|\, z_k, \theta$), is then given by

$$p_{ku}(\theta) = \Phi\left((1 - \Pi_k)^{-\frac{1}{2}} \left(z_k \;\Pi_k^{\frac{1}{2}} - a_K\right) \;+\; \theta \; I_X^{\frac{1}{2}} \left(1 - \Pi_k\right)^{\frac{1}{2}} \right)$$

where Φ is the cumulative distribution function of the standardized Z statistic and a_K is the upper critical value at the final stage.

Similarly, the power for the lower alternative, prob($Z_K < a\_K \,|\, z_k, \theta$), is

$$p_{kl}(\theta) = 1 - \Phi\left((1 - \Pi_k)^{-\frac{1}{2}} \left(z_k \;\Pi_k^{\frac{1}{2}} - a\_K\right) \;+\; \theta \; I_X^{\frac{1}{2}} \left(1 - \Pi_k\right)^{\frac{1}{2}} \right)$$

where a_K is the lower critical value at the final stage.

A special case of the conditional power is the futility index (Ware, Muller, and Braunwald, 1985). It is one minus the conditional power under $H_1 : \theta = \theta_1$:

$$1 - p_{ku}(\theta_1) \;\text{ or }\; 1 - p_{kl}(\theta_1)$$

That is, it is the probability of accepting the null hypothesis under the alternative hypothesis given current data. A high futility index indicates a small probability of success (rejecting H_0) given the current data.

If $\theta = \hat{\theta}_k = z_k I_k^{-\frac{1}{2}}$, the maximum likelihood estimate at stage k, the powers for the upper and lower alternatives can be simplified:

$$p_{ku}(\theta) = \Phi\left((1 - \Pi_k)^{-\frac{1}{2}} \; (z_k \; \Pi_k^{-\frac{1}{2}} - a_K) \right)$$

$$p_{kl}(\theta) = 1 - \Phi\left((1 - \Pi_k)^{-\frac{1}{2}} \; (z_k \; \Pi_k^{-\frac{1}{2}} - a\_K) \right)$$

Predictive Power Approach

The conditional power depends on the specified reference θ, which might be supported by the current data (Jennison and Turnbull 2000, p. 210). An alternative is to use the predictive power (Herson 1979), which is

a weighted average of the conditional power over values of θ. Without prior knowledge about θ, then with $\hat{\theta} = z_k/\sqrt{I_k}$, the maximum likelihood estimate at stage k, the posterior distribution for θ (Jennison and Turnbull 2000, p. 211) is

$$\theta \mid Z_K \sim N\left(\frac{z_k}{\sqrt{I_k}},\ \frac{1}{I_k}\right)$$

Thus, the predictive power at stage k for the upper and lower alternatives can be derived as

$$p_{ku} = 1 - \Phi\left((1-\Pi_k)^{-\frac{1}{2}}\ (a_K\ \Pi_k^{\frac{1}{2}} - z_k)\right)$$

$$p_{kl} = \Phi\left((1-\Pi_k)^{-\frac{1}{2}}\ (a_{\_K}\ \Pi_k^{\frac{1}{2}} - z_k)\right)$$

where a_K and $a_{\_K}$ are the upper and lower critical values at the final stage.

Repeated Confidence Intervals

In a group sequential test, repeated confidence intervals for a parameter θ are defined as a sequence of intervals $(\hat{\theta}_{kl}, \hat{\theta}_{ku}), k = 1, 2, \ldots, K$, for which a simultaneous coverage probability is maintained (Jennison and Turnbull 2000, p. 189). That is, a $(1-\alpha)$ sequence of repeated confidence intervals has

$$\text{Prob}(\hat{\theta}_{kl} \le \theta \le \hat{\theta}_{ku}) = 1 - \alpha$$

These confidence limits $\hat{\theta}_{kl}$ and $\hat{\theta}_{ku}$ can be created from observed statistic and boundary values at each stage.

Two-Sided Repeated Confidence Intervals

Two sequences of repeated confidence intervals can be derived for a two-sided test. One is a $(1-\alpha_l-\alpha_u)$ rejection repeated confidence intervals $(\hat{\theta}_{kl}(\alpha), \hat{\theta}_{ku}(\alpha)), k = 1, 2, \ldots, K$, and the other is a $(1-\beta_l-\beta_u)$ acceptance repeated confidence intervals $(\hat{\theta}_{kl}(\beta), \hat{\theta}_{ku}(\beta)), k = 1, 2, \ldots, K$, where α_l and α_u are the lower and upper Type I error probabilities for the test and β_l and β_u are the lower and upper Type II error probabilities for the test (Jennison and Turnbull 2000, p. 196).

The rejection lower and upper repeated confidence limits at stage k are

$$\hat{\theta}_{kl}(\alpha) = \hat{\theta}_k - \frac{a_k}{\sqrt{I_k}} \qquad \hat{\theta}_{ku}(\alpha) = \hat{\theta}_k - \frac{a_{\_k}}{\sqrt{I_k}}$$

The hypothesis is rejected for upper alternative if the lower limit $\hat{\theta}_{kl}(\alpha) > \theta_{0u}$ and is rejected for lower alternative if the upper limit $\hat{\theta}_{ku}(\alpha) < \theta_{0l}$. That is, the hypothesis is rejected if both θ_{0l} and θ_{0u} are not in a rejection repeated confidence interval $(\hat{\theta}_{kl}(\alpha), \hat{\theta}_{ku}(\alpha))$.

The acceptance lower and upper repeated confidence limits at stage k are

$$\hat{\theta}_{kl}(\beta) = \hat{\theta}_k + \left(\theta_{1l} - \frac{b_{\_k}}{\sqrt{I_k}}\right) \qquad \hat{\theta}_{ku}(\beta) = \hat{\theta}_k + \left(\theta_{1u} - \frac{b_k}{\sqrt{I_k}}\right)$$

The hypothesis is accepted if the lower limit $\hat{\theta}_{kl}(\beta) > \theta_{1l}$ and the upper limit $\hat{\theta}_{ku}(\beta) < \theta_{1u}$. That is, a repeated confidence interval is contained in the interval $(\theta_{1l}, \theta_{1u})$.

One-Sided Repeated Confidence Intervals

Like the two-sided repeated confidence intervals, two sequences of repeated confidence intervals can be derived for a one-sided test. Suppose the one-sided test has an upper alternative θ_{1u}. Then one sequence of repeated confidence intervals is a $(1-\alpha_u)$ rejection repeated confidence intervals $(\hat{\theta}_{kl}(\alpha), \infty)$, $k = 1, 2, \ldots, K$, and the other is a $(1-\beta_u)$ acceptance repeated confidence intervals $(-\infty, \hat{\theta}_{ku}(\beta))$, $k = 1, 2, \ldots, K$, where α_u and β_u are the upper Type I and Type II error probabilities for the test. Thus, a sequence of repeated confidence intervals with confidence level greater than or equal to $(1-\alpha_u-\beta_u)$ is given by $(\hat{\theta}_{kl}(\alpha), \hat{\theta}_{ku}(\beta))$.

The rejection lower repeated confidence limit and the acceptance upper repeated confidence limit at stage k are

$$\hat{\theta}_{kl}(\alpha) = \hat{\theta}_k - \left(\frac{a_k}{\sqrt{I_k}} - \theta_{0u}\right) \qquad \hat{\theta}_{ku}(\beta) = \hat{\theta}_k + \left(\theta_{1u} - \frac{b_k}{\sqrt{I_k}}\right)$$

The hypothesis is rejected if the lower limit $\hat{\theta}_{kl}(\alpha) > \theta_{0u}$. and it is accepted if the upper limit $\hat{\theta}_{ku}(\beta) < \theta_{1u}$.

Analysis after a Sequential Test

At the end of a trial, the hypothesis is either rejected or accepted. But the p-value, median, and confidence limits depend on the ordering the sample space (k, z), where k is the stage number and z is the standardized Z statistic.

Following the notations used in Jennison and Turnbull (2000, pp. 179–180), $(k', z') \succ (k, z)$ if (k', z') has a higher order or more extreme than (k, z). Then for a given ordering, the p-value, median, and confidence limits associated with the observed statistics (k, z) can be derived.

p-value

With the observed pair of statistics (k_0, z_0) when the trial is stopped, a one-sided upper p-value is computed as

$$\text{Prob}\{\,(k, z) \succeq (k_0, z_0)\,\}$$

A one-sided lower p-value is computed as

$$\text{Prob}\{\,(k, z) \preceq (k_0, z_0)\,\}$$

A two-sided p-value is twice the smaller of the lower and upper p-values.

Median Unbiased Estimate

With the observed pair (k_0, z_0), a median unbiased estimate θ_m is computed from

$$\text{Prob}\{\,(k,z) \succeq (k_0,z_0)\,|\,\theta_m\,\} = 0.50$$

Confidence Limits

With the observed pair (k_0, z_0), a lower $(1-\alpha_l)$ confidence limit for θ, θ_l, is computed from

$$\text{Prob}\{\,(k,z) \succeq (k_0,z_0)\,|\,\theta_l\,\} = \alpha_l$$

Similarly, an upper $(1-\alpha_u)$ confidence limit for θ, θ_u, is computed from

$$\text{Prob}\{\,(k,z) \preceq (k_0,z_0)\,|\,\theta_u\,\} = \alpha_u$$

Available Sample Space Orderings in a Sequential Test

At the end of a trial, the hypothesis is either rejected or accepted. Denote the stage number and the statistic at the end of a trial by a pair of statistics (k, z), where k is the stage number and z is the standardized Z statistic. Then an ordering on the sample space (k, z) is needed to derive the p-value, median, and confidence limits associated with the observed statistics (k^*, z^*).

The SEQTEST procedure provides the stagewise, LR, and MLE orderings. Refer to Jennison and Turnbull (2000 pp. 179–187) for a detailed description and comparison of these orderings.

Stagewise Ordering

If the continuation regions of a design are intervals, the stagewise ordering (Fairbanks and Madsen 1982; Tsiatis, Rosner, and Mehta 1984; Jennison and Turnbull 2000, pp. 179–180) uses counter-clockwise ordering around the continuation region to compute the p-value, unbiased median estimate, and confidence limits. This ordering depends on the stopping region, stopping stage, and standardized statistic at the stopping stage. But it does not depend on information levels beyond the observed stage. For a one-sided design with an upper alternative, $(k', z') \succ (k, z)$ if one of the following criteria holds:

- $k' = k$ and $z' > z$
- $k' < k$ and $z' \geq a_{k'}$, the upper α boundary at stage k'
- $k' > k$ and $z < b_k$, the upper β boundary at stage k

Similar criteria can be derived for a one-sided design with a lower alternative.

For a two-sided design with early stopping to reject the null hypothesis, $(k', z') \succ (k, z)$ if one of the following criteria holds:

- $k' = k$ and $z' > z$
- $k' < k$ and $z' \geq a_{k'}$, the upper α boundary at stage k'
- $k' > k$ and $z \leq \_a_k$, the lower α boundary at stage k

Note that the stagewise ordering is not applicable for two-sided designs with early stopping to accept H_0 or to either accept or reject H_0, which might have two disjoint continuous intervals at each interim stage.

For a two-sided design with early stopping either to reject or to accept the null hypothesis, $(k', z') \succ (k, z)$ if one of the following criteria holds:

- $z' \geq a_{k'}$ and $z < b_k$
- $z' > \_b_{k'}$ and $z \leq \_a_k$

That is, each value in the continuation region is less extreme than each value in the upper rejection region and more extreme than each value in the lower rejection region. Then, combining with the ordering defined for a two-sided design with early stopping to reject the null hypothesis, the p-value, median, and confidence limits can be derived for the observed statistics in the lower or upper rejection region.

Thus, if the stagewise ordering is specified in the SEQTEST procedure for a two-sided design with early stopping to either reject or accept the null hypothesis, the stagewise ordering is used to derive these statistics only if the observed statistics is in the lower or upper rejection region. Otherwise, the LR ordering is used.

LR Ordering

The LR ordering (Chang 1989) depends on the observed standardized Z statistic z, information levels, and a specified hypothetical reference. For the LR ordering under a given hypothesis $H : \theta = \theta_g$, $(k', z') \succ (k, z)$ if

$$(z' - \theta_g \sqrt{I_{k'}}) \; > \; (z - \theta_g \sqrt{I_k})$$

Under the null hypothesis $H_0 : \theta = 0$, it reduces to

$$z' > z$$

and can be used to derive statistics under H_0, such as p-values.

The LR ordering is applicable to all designs if all information levels are available. But depending on the boundary shape, some observed statistics (k, z) in the rejection region might be less extreme than the statistics in the acceptance region. That is, the p-value for observed statistics in the rejection region might be greater than the significance level.

MLE Ordering

The MLE ordering (Emerson and Fleming 1990) depends only on the observed maximum likelihood estimate. $(k', z') \succ (k, z)$ if

$$\frac{z'}{\sqrt{I_{k'}}} > \frac{z}{\sqrt{I_k}}$$

The MLE ordering is applicable to all designs if all information levels are available.

Applicable Tests and Sample Size Computation

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

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

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

If the data are not from a normal distribution such as 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 clinical trial, the sample size required depends on the Type I error probability α, reference improvement θ_1, power $1-\beta$, and variance of the response variable. Given a 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

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

where the parameter θ depends on the test specified in the clinical trial. For example, if you are comparing two binomial populations $H_0 : \theta = 0$, then $\theta = p_t - p_c$ is the difference between two proportions if the proportion difference statistic is used, and $\theta = \log\left(\frac{p_t(1-p_c)}{p_c(1-p_t)}\right)$, 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's information for θ.

The resulting statistic $\hat{\theta}$ corresponds to the MLE 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 scale (BOUNDARYSCALE=SCORE).

Alternatively, if the score statistic is derived, it can also be used as the test statistic and its asymptotic variance is given by Fisher's information.

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

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

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.

With a specified test statistic, the required sample sizes at the stages can be computed. These tests include commonly used tests for normal means, binomial proportions, and survival distributions. See the section "Sample Size Computation" in "The SEQDESIGN Procedure" for a description of these tests.

Table Output

The SEQTEST procedure displays the "Design Information" and "Test Information" tables by default.

Conditional Power

The "Conditional Power Information" table displays the following information under a hypothetical reference:

- stopping stage
- MLE, observed maximum likelihood estimate
- conditional power under the hypothetical reference

For a one-sided test, the power are derived under hypothetical references $\theta = \hat{\theta}$ and $\theta = c_i\theta_1$, where $\hat{\theta}$ is the observed statistic, θ_1 is the alternative reference, and c_i are the values specified in the CREF= option. For a two-sided test, the power are derived under the hypothetical references $\theta = \hat{\theta}$, $\theta = c_i\theta_{1l}$, and $\theta = c_i\theta_{1u}$, where θ_{1l} is the lower alternative reference and θ_{1u} is the upper alternative reference. The default is CREF= 0 0.5 1.0 1.5.

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 size 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 size required under the alternative reference for the group sequential design in percentage of the corresponding fixed-sample design.

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

Parameter Estimates

The "Parameter Estimates" table displays the following information at the conclusion of a sequential trial:

- stopping stage
- parameter estimate
- median and confidence limits based on the specified ordering
- p-value for the hypothesis H_0 based on the specified ordering

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.

Predictive Power

The "Predictive Power Information" table displays the following information:

- stopping stage
- MLE, observed maximum likelihood estimate
- predictive power

Repeated Confidence Intervals

The "Repeated Confidence Intervals" table displays the following information for the observed statistic at each stage:

- information level
- parameter estimate
- rejection confidence limits. The null hypothesis is rejected for the upper alternative if the lower rejection confidence limit is greater than the null parameter value. Similarly, the null hypothesis is rejected for the lower alternative if the upper rejection confidence limit is less than the null parameter value.
- acceptance confidence limits. The upper alternative hypothesis is rejected if the upper acceptance confidence limit is less than the upper alternative value. Similarly, the lower alternative hypothesis is rejected if the lower acceptance confidence limit is greater than the upper alternative value. For a two-sided design, if both upper and lower alternative hypothesis are rejected, the null hypothesis is accepted.

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 $k_0 + d$ and is derived from the expected information level

$$\sum_{k=1}^{K} p_k\, I_k = I_{k_0} + d\,(I_{(k_0+1)} - I_{k_0})$$

where p_k is the stopping probability at stage k and $0 \leq d < 1$.

For equally spaced information levels, the expected stopping stage is reduced to the weighted average

$$\sum_{k=1}^{K} p_k\, k$$

Test Information

The "Test Information" table displays the following information at each stage:

- proportion of information
- actual information level, if the maximum information is available from the input BOUNDARY= data set
- 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 test 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.

If the test variable is specified, the table also displays the following:

- test statistic
- resulting action of test statistic: continue to the next stage, accept the null hypothesis H_0, or reject H_0

ODS Table Names

PROC SEQTEST 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 81.3. For more information about ODS, see Chapter 20, "Using the Output Delivery System."

Table 81.3 ODS Tables Produced by PROC SEQTEST

ODS Table Name	Description	Option
CondPower	Conditional power	CONDPOWER
Design	Design information	
ErrSpend	Error spending	ERRSPEND
ParameterEstimates	Parameter estimates	DATA(TESTVAR=) or PARMS(TESTVAR=)
PowerSampleSize	Power and expected sample sizes	PSS
PredPower	Predictive power	PREDPOWER
RepeatedCI	Repeated confidence intervals	RCI
StopProb	Stopping probabilities	STOPPROB
Test	Test statistics and boundary values	

Graphics Output

This section describes the use of ODS for creating graphics with the SEQTEST procedure. To request these graphs, ODS Graphics must be enabled and you must specify the associated graphics options in the PROC SEQTEST statement. For more information about ODS Graphics, 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.

If the trial stops after the sequential test, the hypothetical reference corresponding to the test statistic is also indicated in the plot.

Conditional Power Plot

The PLOTS=CONDPOWER option displays the conditional powers given the observed statistic under various hypothetical references. These powers are connected and are displayed in the "Conditional Power Plot" graph.

For a one-sided test, the power are derived under the hypothetical references $\theta = \hat{\theta}$ and $\theta = c_i\theta_1$, where $\hat{\theta}$ is the observed statistic, θ_1 is the alternative reference, and c_i are the values specified in the CREF= option. The horizontal axis displays these c_i values for hypothetical references.

For a two-sided test, the power are derived under hypothetical references $\theta = \hat{\theta}$, $\theta = c_i\theta_{1l}$, and $\theta = c_i\theta_{1u}$, where θ_{1l} is the lower alternative reference and θ_{1u} is the upper alternative reference. The horizontal axis displays $-c_i$ values for hypothetical references $\theta = c_i\theta_{1l}$ and c_i values for hypothetical references $\theta = c_i\theta_{1u}$.

If the trial stops after the sequential test, the hypothetical reference corresponding to the test statistic is also indicated in the plot.

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.

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.

If the trial stops after the sequential test, the hypothetical reference corresponding to the test statistic is also indicated in the plot.

Repeated Confidence Intervals Plot

The PLOTS=RCI option displays repeated confidence intervals at each stage given the observed statistic at that stage. These repeated confidence intervals are displayed in the "Repeated Confidence Intervals Plot" graph.

Sequential Test Plot

The PLOTS=TEST option displays boundary values and test statistics in the "Test Plot" graph. The boundary values are connected for each boundary, and both the stage number and the information level at each stage are displayed. The legend table identifies the acceptance and rejection regions in the plot.

ODS Graphics

Statistical procedures use ODS Graphics to create graphs as part of their output. ODS Graphics is described in detail in Chapter 21, "Statistical Graphics Using ODS."

Before you create graphs, ODS Graphics must be enabled (for example, with the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 609 in Chapter 21, "Statistical Graphics Using ODS."

The overall appearance of graphs is controlled by ODS styles. Styles and other aspects of using ODS Graphics are discussed in the section "A Primer on ODS Statistical Graphics" on page 608 in Chapter 21, "Statistical Graphics Using ODS."

PROC SEQTEST assigns a name to each graph it creates. You can use these names to reference the graphs when using ODS. To request these graphs, ODS Graphics must be enabled and you must specify the options indicated in Table 81.4.

Table 81.4 Graphs Produced by PROC SEQTEST

ODS Graph Name	Plot Description	Option
AsnPlot	Average sample numbers	PLOTS=ASN
CondPowerPlot	Conditional power curves	PLOTS=CONDPOWER
ErrSpendPlot	Error spending	PLOTS=ERRSPEND
PowerPlot	Power curves	PLOTS=POWER
RepeatedCIPlot	Repeated confidence intervals	PLOTS=RCI
TestPlot	Boundary values and test statistics	PLOTS=TEST

Acknowledgments

In addition to being shaped by the research literature listed in the section "References" on page 7028, 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.

Examples: SEQTEST Procedure

The following examples perform group sequential tests with various designs and test statistics.

Four statistic scales are available for the input boundary values, the input test statistic, and the displayed test information in the SEQTEST procedure. These are the maximum likelihood estimator scale, score statistic scale, standardized normal Z scale, and p-value scale. There is a unique one-to-one transformation between any two of the scales, and you can use different scales for the input boundary values and input test statistic. These boundary values and test statistic are displayed with the scale specified in the BOUNDARYSCALE= option in the SEQTEST procedure.

Example 81.1: Testing the Difference between Two Proportions

This example demonstrates group sequential tests that use an O'Brien-Fleming group sequential design. A clinic is studying the effect of vitamin C supplements in treating flu symptoms. The study consists of patients in the clinic who have exhibited the first sign of flu symptoms within the last 24 hours. These patients are randomly assigned to either the control group (which receives 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 patient are recorded.

Suppose that you know from past experience that flu symptoms disappear in five days for 60% of patients who experience flu symptoms. The clinic would like to detect a 75% symptom disappearance with a high

probability. A test that compares the proportions directly specifies the null hypothesis $H_0 : \theta = p_t - p_c = 0$ with a one-sided alternative $H_1 : \theta > 0$ and a power of 0.90 at $H_1 : \theta = 0.15$, where p_t and p_c are the proportions of symptom disappearance in the treatment group and control group, respectively.

The following statements invoke the SEQDESIGN procedure and request a four-stage group sequential design by using an O'Brien-Fleming method for normally distributed data. The design uses a one-sided alternative hypothesis with early stopping either to accept or reject the null hypothesis H_0. The BOUNDARYSCALE=MLE option uses the MLE scale to display statistics in the boundary table and boundary plots.

```
ods graphics on;
proc seqdesign altref=0.15
               boundaryscale=mle
               ;
   OBrienFleming: design method=obf
                         nstages=4
                         alt=upper
                         stop=both
                         alpha=0.025
                         ;
   samplesize model=twosamplefreq(nullprop=0.6 test=prop);
ods output Boundary=Bnd_Count;
run;
ods graphics off;
```

The ODS OUTPUT statement with the BOUNDARY=BND_COUNT option creates an output data set named BND_COUNT which contains the resulting boundary information for the subsequent sequential tests.

The "Design Information" table in Output 81.1.1 displays design specifications. With the specified alternative hypothesis $H_1 : \theta = 0.15$, the maximum information is derived to achieve a power of 0.90 at H_1. The derived fixed-sample information ratio 1.0767 is the maximum information needed for a group sequential design relative to its corresponding fixed-sample design.

Output 81.1.1 O'Brien-Fleming Design Information

The SEQDESIGN Procedure
Design: OBrienFleming

Design Information

Statistic Distribution	Normal
Boundary Scale	MLE
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 "Boundary Information" table in Output 81.1.2 displays the information level, alternative reference, and boundary values at each stage. With the BOUNDARYSCALE=MLE option, the SEQDESIGN procedure displays the output boundaries with the maximum likelihood estimator scale.

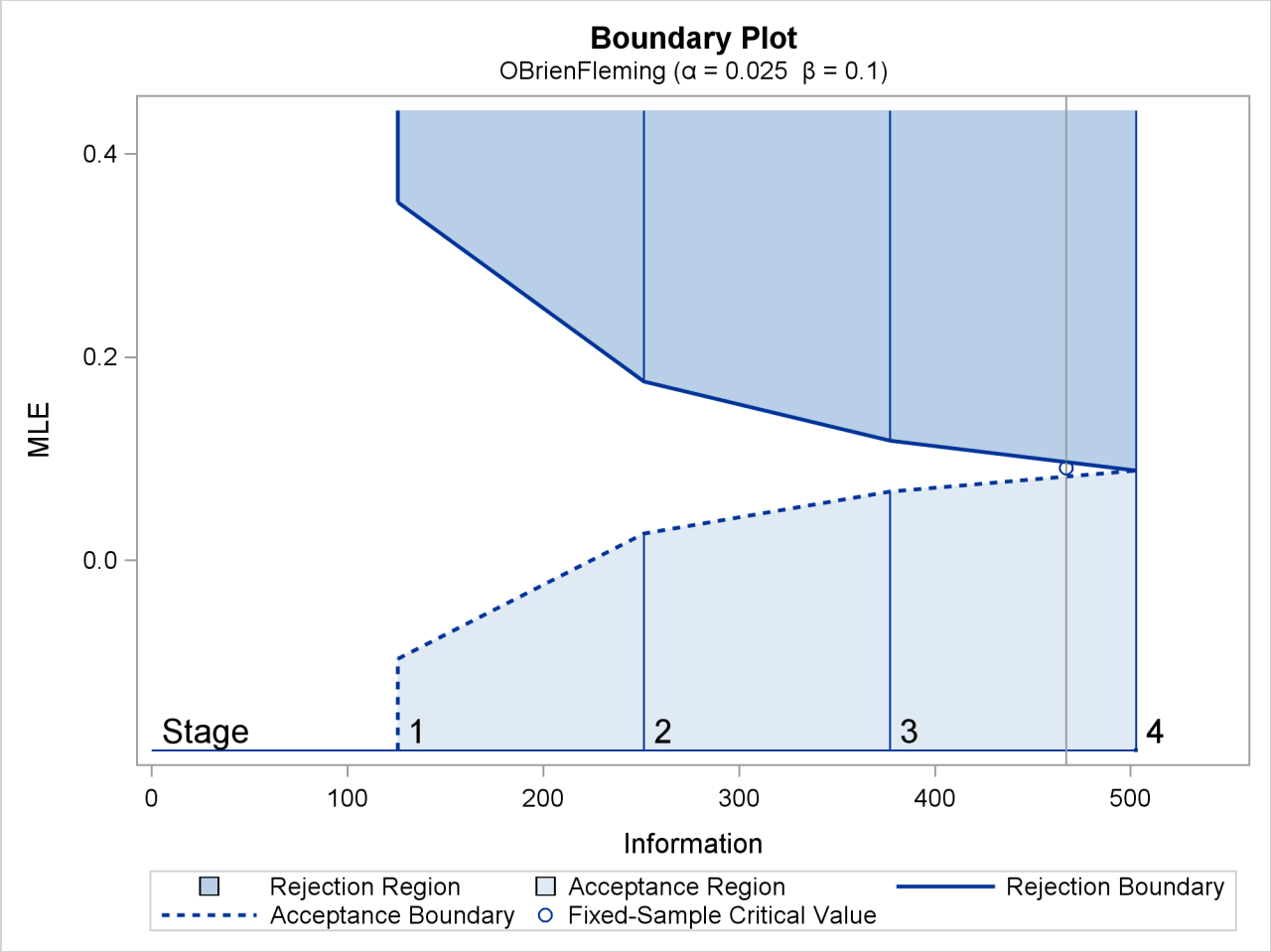
Output 81.1.2 O'Brien-Fleming Boundary Information

Boundary Information (MLE Scale)
Null Reference = 0

	Information Level			Alternative Reference	Boundary Values	
					Upper	
Stage	Proportion	Actual	N	Upper	Beta	Alpha
1	0.2500	125.7086	107.4808	0.15000	-0.09709	0.35291
2	0.5000	251.4171	214.9617	0.15000	0.02645	0.17645
3	0.7500	377.1257	322.4425	0.15000	0.06764	0.11764
4	1.0000	502.8343	429.9233	0.15000	0.08823	0.08823

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.1.3. The horizontal axis indicates the information levels for the design. The stages are indicated by vertical lines with accompanying stage numbers. If the test statistic at a stage is in a rejection region, the trial stops and the hypothesis is rejected. If the 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 or an acceptance region, the trial continues to the next stage.

Output 81.1.3 O'Brien-Fleming Boundary Plot



The boundary plot also displays the information level and critical value for the corresponding fixed-sample design. The solid and dashed lines at the fixed-sample information level correspond to the rejection and acceptance lines, respectively.

With the SAMPLESIZE statement, the maximum information is used to derive the required sample size for the study. The "Sample Size Summary" table in Output 81.1.4 displays parameters for the sample size computation.

Output 81.1.4 Required 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

With the derived maximum information and the specified MODEL= option in the SAMPLESIZE statement, the total sample size in each group for testing the difference between two proportions under the alternative hypothesis is

$$N_1 = N_2 = \left(p_{1c}\left(1 - p_{1c}\right) + p_{1t}\left(1 - p_{1t}\right)\right) I_X$$

where $p_{1c} = 0.6$ and $p_{1t} = p_{1c} + \theta_1 = 0.75$. By default (or equivalently if you specify REF=PROP in the MODEL=TWOSAMPLEFREQ option), the required sample sizes are computed under the alternative hypothesis. See the section "Test for the Difference between Two Binomial Proportions" in the chapter "The SEQDESIGN Procedure" for a description of these parameters.

The "Sample Sizes (N)" table in Output 81.1.5 displays the required sample sizes at each stage, in both fractional and integer numbers. The derived sample sizes under the heading Fractional N which correspond to the design are not integers. These sample sizes are rounded up to integers under the heading Ceiling N. In practice, integer sample sizes are used, and the information levels increase slightly. Thus, 54, 108, 162, and 215 patients are needed in each group for the four stages, respectively.

Output 81.1.5 Required 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

Suppose the trial follows the study plan, and 54 patients are available in each group at stage 1. The data set count_1 contains these 108 patients. Output 81.1.6 lists the first 10 observations of the data set.

Output 81.1.6 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Trt	Resp
1	0	0
2	1	1
3	0	1
4	1	0
5	0	0
6	1	0
7	0	0
8	1	1
9	0	1
10	1	0

The Trt variable is a grouping variable with value 0 for a patient in the placebo control group and value 1 for a patient in the treatment group who is given vitamin C supplements. The Resp variable is an indicator variable with value 1 for a patient without flu symptoms after five days and value 0 for a patient with flu symptoms after five days.

The following statements use the GENMOD procedure to estimate the treatment effect at stage 1:

```
proc genmod data=count_1;
   model Resp= Trt;
ods output ParameterEstimates=Parms_Count1;
run;
```

Output 81.1.7 displays the treatment effect at stage 1.

Output 81.1.7 Stage 1 Treatment Difference

The GENMOD Procedure

Analysis Of Maximum Likelihood Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Wald Chi-Square	Pr > ChiSq
Intercept	1	0.6296	0.0627	0.5066	0.7526	100.68	<.0001
Trt	1	0.1111	0.0887	-0.0628	0.2850	1.57	0.2105
Scale	1	0.4611	0.0314	0.4035	0.5269		

NOTE: The scale parameter was estimated by maximum likelihood.

The test statistic is $\hat{\theta}_1 = \hat{p}_t - \hat{p}_c = 0.1111$, and its associated standard error is

$$\sqrt{\mathrm{Var}(\hat{\theta}_1)} = \sqrt{\frac{\hat{p}_c(1-\hat{p}_c)}{54} + \frac{\hat{p}_t(1-\hat{p}_t)}{54}} = 0.0887$$

The following statements create and display (in Output 81.1.8) the data set that contains the parameter estimate at stage 1, $\hat{\theta}_1 = 0.1111$, and its associated standard error $\sqrt{\text{Var}(\hat{\theta}_1)} = 0.0887$ which are used in the SEQTEST procedure:

```
data Parms_Count1;
   set Parms_Count1;
   if Parameter='Trt';
   _Scale_='MLE';
   _Stage_= 1;
   keep _Scale_ _Stage_ Parameter Estimate StdErr;
run;

proc print data=Parms_Count1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.1.8 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Parameter	Estimate	StdErr	_Scale_	_Stage_
1	Trt	0.1111	0.0887	MLE	1

The initial required sample sizes are derived with the proportions $p_c = 0.6$ and $p_t = 0.75$. If the observed proportions are different from these assumed values, or if the number of available patients is different from the study plan in one of the stages, then the information level that corresponds to the test statistic is estimated from

$$I_k = \frac{1}{\text{Var}(\hat{\theta}_k)}$$

The following statements invoke the SEQTEST procedure and test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Count
             Parms(Testvar=Trt)=Parms_Count1
             infoadj=none
             errspendmin=0.001
             boundaryscale=mle
             errspend
             plots=errspend
             ;
ods output Test=Test_Count1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The PARMS=PARMS_COUNT1 option specifies the input data set PARMS_COUNT1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=TRT option identifies the test variable TRT in the data set. The INFOADJ=NONE option maintains the information levels at future interim stages (2 and 3) as provided in the

BOUNDARY= data set. The BOUNDARYSCALE=MLE option displays the output boundaries in terms of the MLE scale.

The O'Brien-Fleming design is conservative in early stages and might not be desirable in a clinical trial. The ERRSPENDMIN=0.001 option specifies the minimum error spending at each stage to be 0.001, and it might increase the corresponding nominal p-value in early stages for the trial. The BOUNDARYSCALE=MLE option uses the MLE scale to display test statistics in the boundary table and boundary plots.

The ODS OUTPUT statement with the TEST=TEST_COUNT1 option creates an output data set named TEST_COUNT1 which contains the updated boundary information for the test at stage 1. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.1.9 displays design specifications. The derived statistics, such as the overall α and β levels, are derived from the specified maximum information and boundary values in the BOUNDARY= data set. Note that with a minor change in the information level at stage 1, the power also changes slightly from the design provided in the BOUNDARY= data set.

Output 81.1.9 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_COUNT
Data Set	WORK.PARMS_COUNT1
Statistic Distribution	Normal
Boundary Scale	MLE
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Number of Stages	4
Alpha	0.025
Beta	0.10147
Power	0.89853
Max Information (Percent of Fixed Sample)	108.2301
Max Information	502.834283
Null Ref ASN (Percent of Fixed Sample)	61.09917
Alt Ref ASN (Percent of Fixed Sample)	73.9745

With the ERRSPEND option, the "Error Spending Information" table in Output 81.1.10 displays cumulative error spending at each stage for each boundary. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the Type I error level $\alpha = 0.025$ is maintained. Furthermore, with the ERRSPENDMIN=0.001 option, the α spending at each stage is greater than or equal to 0.001.

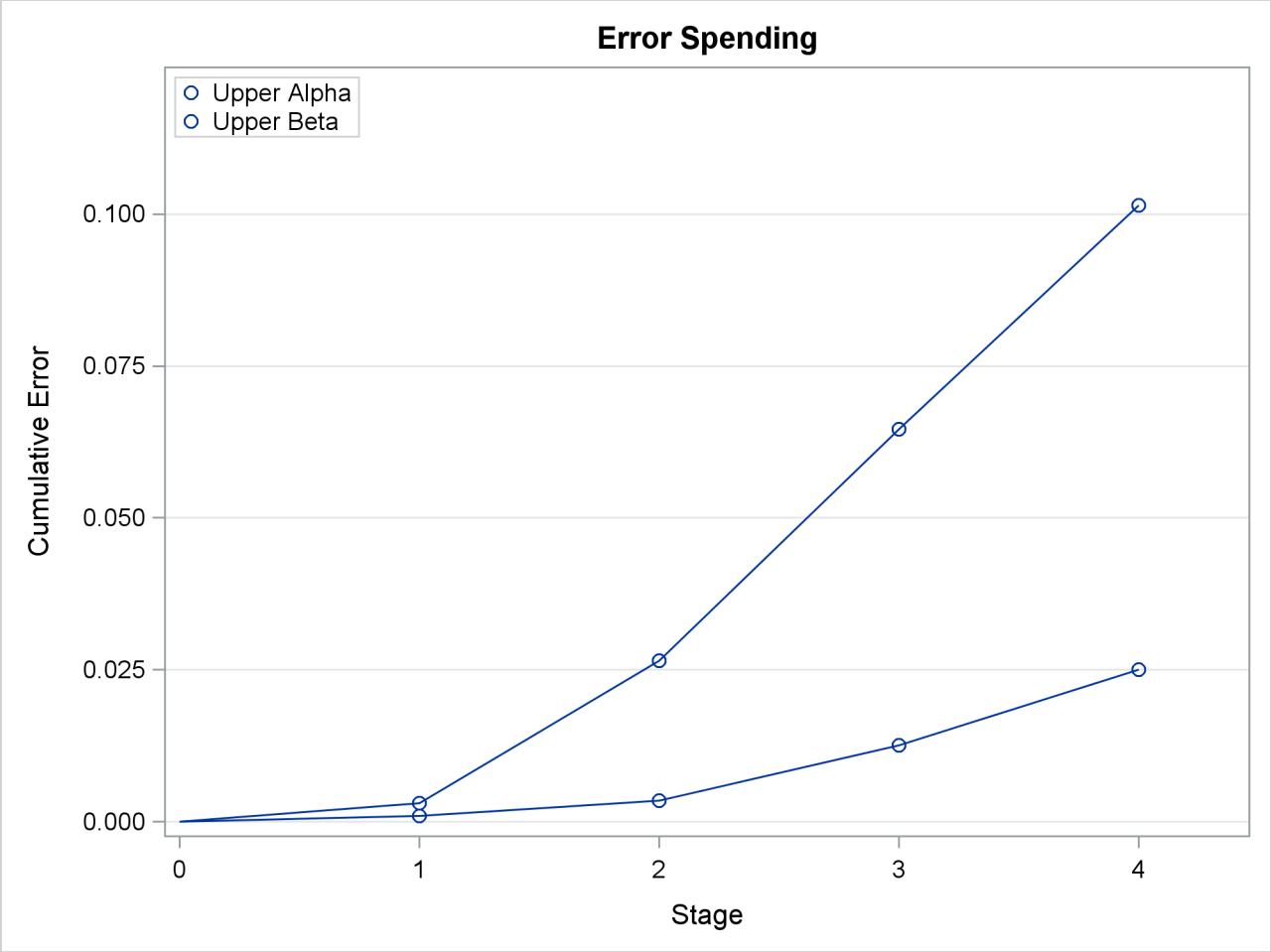
Output 81.1.10 Error Spending Information

Error Spending Information

Stage	---Information Level-- Proportion	Actual	-Cumulative Error Spending- ----------Upper---------- Beta	Alpha
1	0.2525	126.9871	0.00308	0.00100
2	0.5000	251.4171	0.02653	0.00343
3	0.7500	377.1257	0.06456	0.01254
4	1.0000	502.8343	0.10147	0.02500

With the PLOTS=ERRSPEND option, the procedure displays a plot of error spending for each boundary, as shown in Output 81.1.11. The error spending values in the "Error Spending Information" table in Output 81.1.10 are displayed in the plot.

Output 81.1.11 Error Spending Plot



The "Test Information" table in Output 81.1.12 displays the boundary values for the design, test statistic, and resulting action at each stage. With the BOUNDARYSCALE=MLE option, the maximum likelihood

estimator scale is used for the test statistic and boundary values. The table shows that the test statistic 0.1111 is between the upper α and β boundaries, so the trial continues to the next stage.

Output 81.1.12 Sequential Test

Test Information (MLE Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper	Boundary Values Upper Beta	Boundary Values Upper Alpha
1	0.2525	126.9871	0.15000	-0.09306	0.27423
2	0.5000	251.4171	0.15000	0.02674	0.17527
3	0.7500	377.1257	0.15000	0.06805	0.11792
4	1.0000	502.8343	0.15000	0.08875	0.08875

Test Information (MLE Scale)
Null Reference = 0

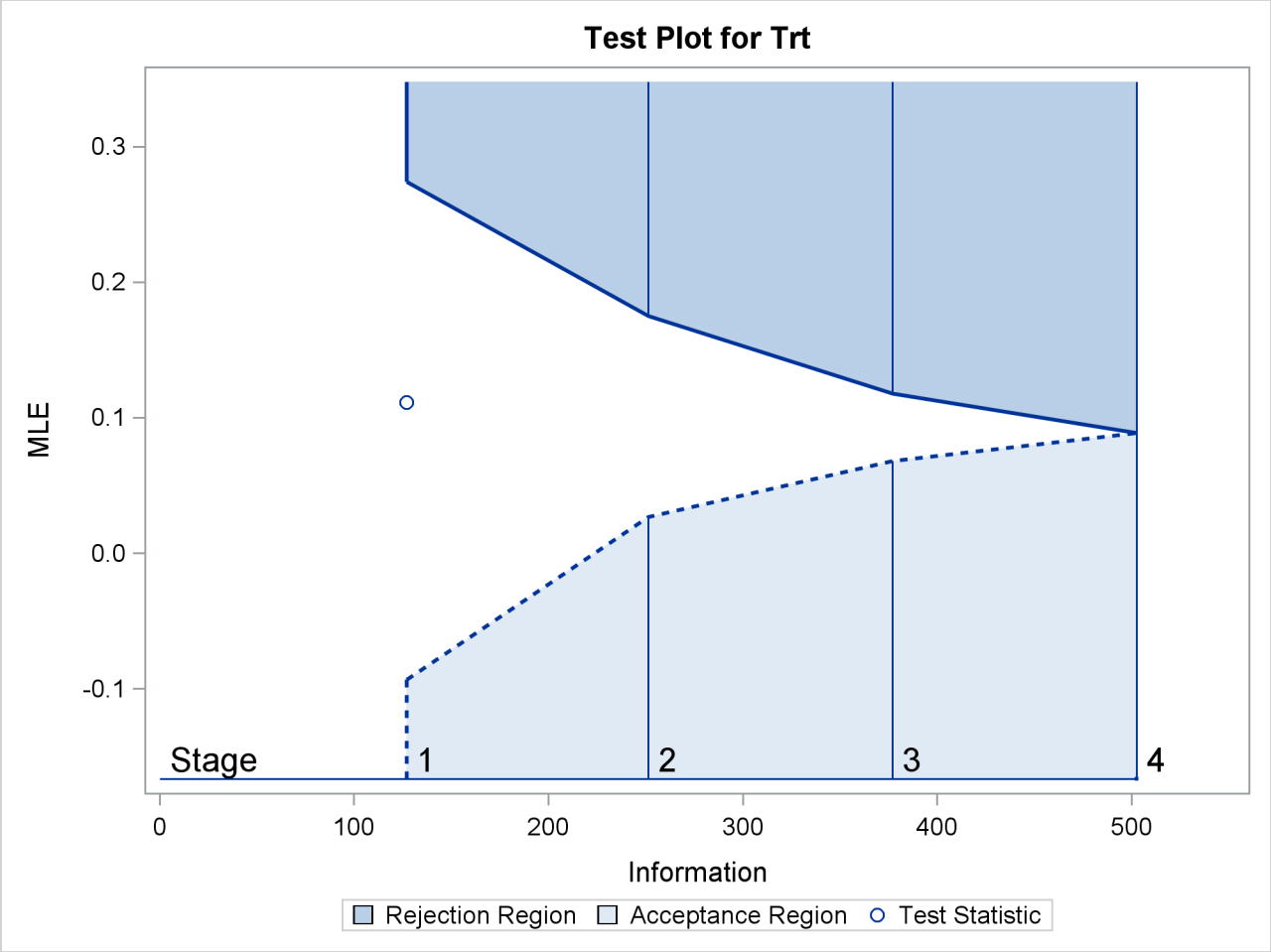
Stage	Test Trt Estimate	Test Trt Action
1	0.11111	Continue
2	.	
3	.	
4	.	

The information level at stage 1 is derived from the standard error,

$$I_1 = \frac{1}{\mathrm{Var}(\hat{\theta}_1)} = \frac{1}{(\mathrm{s.e.}(\hat{\theta}_1))^2} = 126.987$$

By default (or equivalently if you specify PLOTS=TEST), the "Test Plot" graph displays boundary values of the design and the test statistic at stage 1, as shown in Output 81.1.13. It also shows that the observed statistic is in the continuation region.

Output 81.1.13 Sequential Test Plot



The observed information level at stage 1, $I_1 = 126.987$, is slightly larger than the target information level at the design. If an observed information level in the study is substantially different from its target level in the design, then the sample sizes should be adjusted in the subsequent stages to achieve the target information levels.

Suppose the trial continues to the next stage, and 108 patients are available in each group at stage 2. The data set COUNT_2 contains these 216 patients.

The following statements use the GENMOD procedure to estimate the treatment effect at stage 2:

```
proc genmod data=Count_2;
   model Resp= Trt;
ods output ParameterEstimates=Parms_Count2;
run;
```

The following statements create the parameter estimate at stage 2, $\hat{\theta}_2 = \hat{p}_2 - \hat{p}_1 = 0.1759$, and its associated standard error $\sqrt{\text{Var}(\hat{\theta}_2)} = 0.0623$ into a test data set:

```
data Parms_Count2;
   set Parms_Count2;
   if Parameter='Trt';
   _Scale_='MLE';
   _Stage_= 2;
   keep _Scale_ _Stage_ Parameter Estimate StdErr;
run;

proc print data=Parms_Count2;
   title 'Statistics Computed at Stage 2';
run;
```

Output 81.1.14 displays the test statistics at stage 2.

Output 81.1.14 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	Parameter	Estimate	StdErr	_Scale_	_Stage_
1	Trt	0.1759	0.0623	MLE	2

The following statements invoke the SEQTEST procedure and test for early stopping at stage 2:

```
ods graphics on;
proc seqtest Boundary=Test_Count1
             Parms(Testvar=Trt)=Parms_Count2
             infoadj=none
             boundaryscale=mle
             ;
ods output Test=Test_Count2;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=TEST_COUNT2 option creates an output data set named TEST_COUNT2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Test Information" table in Output 81.1.15 displays the boundary values, test statistic, and resulting action at each stage. The table shows that the test statistic 0.17593 is larger than the corresponding upper alpha boundary value, so the trial stops to reject the hypothesis.

Output 81.1.15 Sequential Test

The SEQTEST Procedure

Test Information (MLE Scale)
Null Reference = 0

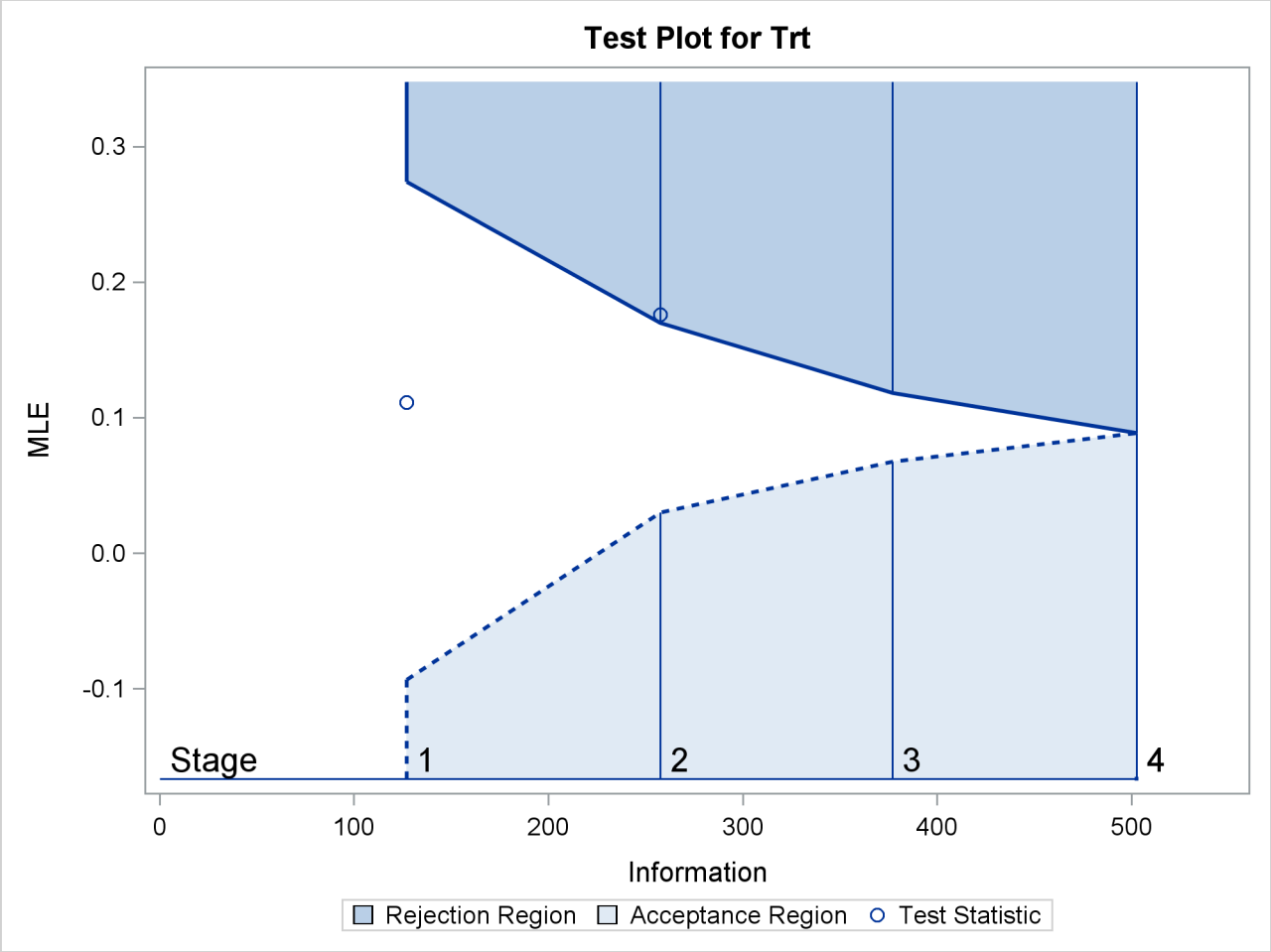
Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper	Boundary Values Upper Beta	Boundary Values Upper Alpha
1	0.2525	126.9871	0.15000	-0.09306	0.27423
2	0.5122	257.5571	0.15000	0.03019	0.17001
3	0.7500	377.1257	0.15000	0.06783	0.11826
4	1.0000	502.8343	0.15000	0.08878	0.08878

Test Information (MLE Scale)
Null Reference = 0

Stage	Test Trt Estimate	Test Trt Action
1	0.11111	Continue
2	0.17593	Reject Null
3	.	
4	.	

With ODS Graphics enabled, the "Test Plot" is displayed, as shown in Output 81.1.16. The plot displays boundary values of the design and the test statistics at the first two stages. As expected, the test statistic at stage 2 is in the "Upper Rejection Region" above the upper alpha boundary.

Output 81.1.16 Sequential Test Plot



After a trial is stopped, the "Parameter Estimates" table in Output 81.1.17 displays the stopping stage and the maximum likelihood estimate of the parameter. It also displays the p-value, median estimate, and confidence limits for the parameter that correspond to the observed statistic by using the specified sample space ordering.

Output 81.1.17 Parameter Estimates

Parameter Estimates
Stagewise Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate	Lower 95% CL
Trt	2	0.175926	0.0031	0.174462	0.07059

The MLE statistic at the stopping stage is the maximum likelihood estimate of the parameter and is biased. The computation of p-value, unbiased median estimate, and confidence limits depends on the ordering of the sample space (k, z), where k is the stage number and z is the observed standardized Z statistic. By default

(or equivalently if you specify ORDER=STAGEWISE), the stagewise ordering that uses counterclockwise ordering around the continuation region is used to compute the p-value, unbiased median estimate, and confidence limits. As expected, the p-value is less than 0.025, and the confidence interval does not contain the null reference zero. With the stagewise ordering, the p-value is computed as

$$P_{\theta=0}\left(Z_1 > a_1\right) + P_{\theta=0}\left(Z_2 > z_2 \mid b_1 < Z_1 < a_1\right)$$

where z_2 is the observed standardized Z statistic at stage 2, Z_1 is the standardized normal variate at stage 1, Z_2 is the standardized normal variate at stage 2, and a_1 and b_1 are the stage 1 upper rejection and acceptance boundary values, respectively.

See the section "Available Sample Space Orderings in a Sequential Test" on page 6902 for a detailed description of the stagewise ordering.

Example 81.2: Testing an Effect in a Regression Model

This example demonstrates a two-sided group sequential test that uses an error spending design with early stopping to reject the null hypothesis. A study is conducted to examine the effects of Age (years), Weight (kg), RunTime (time in minutes to run 1.5 miles), RunPulse (heart rate while running), and MaxPulse (maximum heart rate recorded while running) on Oxygen (oxygen intake rate, ml per kg body weight per minute). The primary interest is whether oxygen intake rate is associated with weight.

The hypothesis is tested using the following linear model:

$$\text{Oxygen} = \text{Age} + \text{Weight} + \text{RunTime} + \text{RunPulse} + \text{MaxPulse}$$

The null hypothesis is $H_0 : \beta_w = 0$, where β_w is the regression parameter for the variable Weight. Suppose that $\beta_w = 0.10$ is the reference improvement that should be detected at a 0.90 level. Then the maximum information I_X can be derived in the SEQDESIGN procedure.

Following the derivations in the section "Test for a Parameter in the Regression Model" in the chapter "The SEQDESIGN Procedure," the required sample size can be derived from

$$N = I_X \, \frac{\sigma_y^2}{\left(1 - r_x^2\right)\sigma_x^2}$$

where σ_y^2 is the variance of the response variable in the regression model, r_x^2 is the proportion of variance of Weight explained by other covariates, and σ_x^2 is the variance of Weight.

Further suppose that from past experience, $\sigma_y^2 = 5$, $r_x^2 = 0.10$, and $\sigma_x^2 = 64$. Then the required sample size can be derived using the SAMPLESIZE statement in the SEQDESIGN procedure.

The following statements invoke the SEQDESIGN procedure and request a three-stage group sequential design for normally distributed data to test the null hypothesis of a regression parameter $H_0 : \beta_w = 0$ against the alternative $H_1 : \beta_w \neq 0$:

```
ods graphics on;
proc seqdesign altref=0.10;
   OBFErrorFunction: design method=errfuncobf
                            nstages=3
                            info=cum(2 3 4)
                            ;
   samplesize model=reg(variance=5 xvariance=64 xrsquare=0.10);
ods output Boundary=Bnd_Fit;
run;
ods graphics off;
```

By default (or equivalently if you specify ALPHA=0.05 and BETA=0.10), the procedure uses a Type I error probability 0.05 and a Type II error probability 0.10. The ALTREF=0.10 option specifies a power of $1 - \beta = 0.90$ at the alternative hypothesis $H_1 : \beta_w = \pm 0.10$. The INFO=CUM(2 3 4) option specifies that the study perform the first interim analysis with information proportion $2/4 = 0.5$—that is, after half of the total observations are collected.

The ODS OUTPUT statement with the BOUNDARY=BND_FIT option creates an output data set named BND_FIT which contains the resulting boundary information for the subsequent sequential tests.

The "Design Information" table in Output 81.2.1 displays design specifications and derived statistics. Since the alternative reference is specified, the maximum information is derived.

Output 81.2.1 Design Information

The SEQDESIGN Procedure
Design: OBFErrorFunction

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	0.1
Number of Stages	3
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	101.8276
Max Information	1069.948
Null Ref ASN (Percent of Fixed Sample)	101.2587
Alt Ref ASN (Percent of Fixed Sample)	77.81586

The "Boundary Information" table in Output 81.2.2 displays information level, alternative reference, and boundary values at each stage.

Output 81.2.2 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

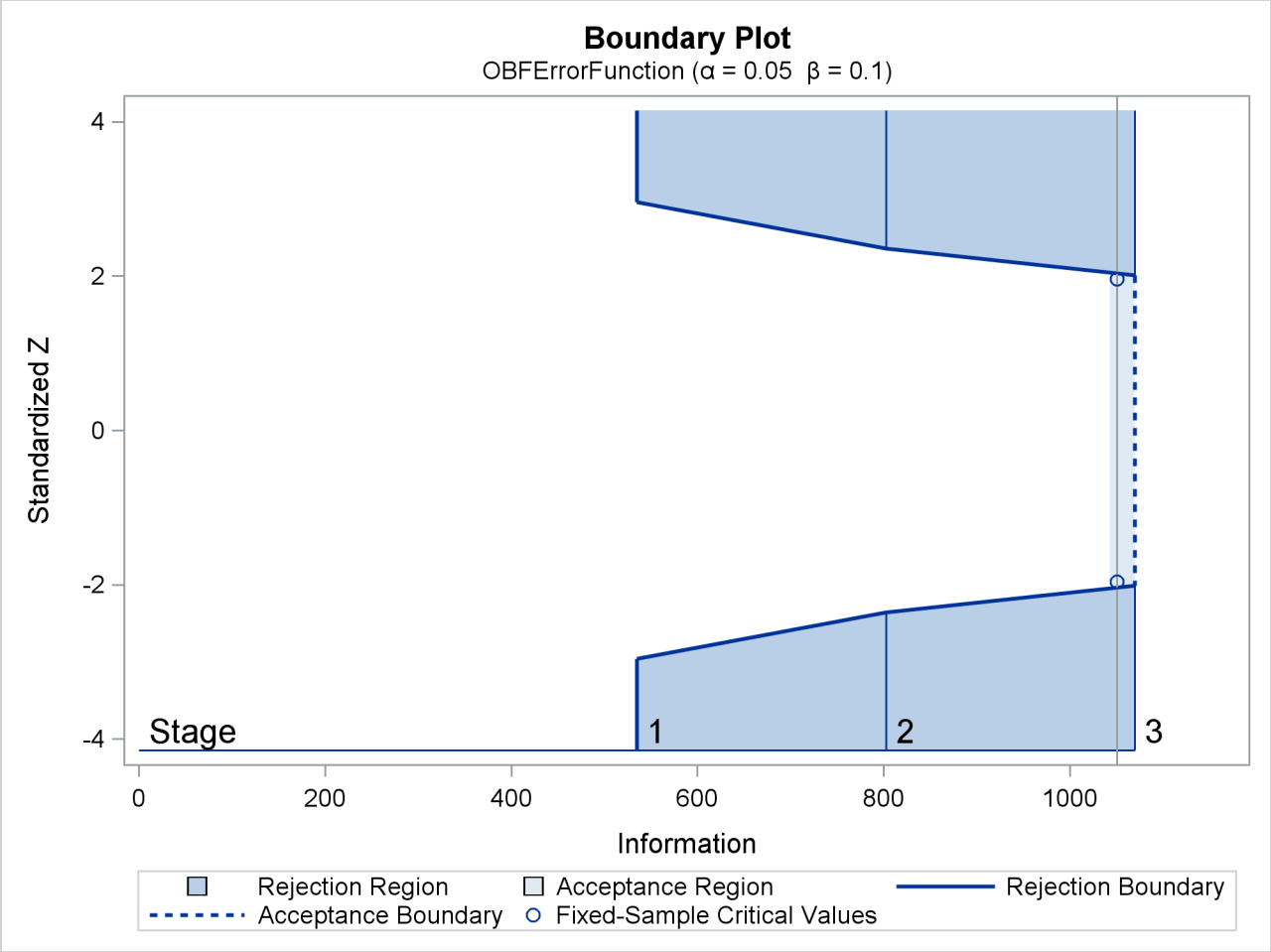
Stage	Information Level Proportion	Information Level Actual	Information Level N	Alternative Reference Lower	Alternative Reference Upper
1	0.5000	534.9738	46.43869	-2.31295	2.31295
2	0.7500	802.4606	69.65804	-2.83277	2.83277
3	1.0000	1069.948	92.87739	-3.27101	3.27101

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Boundary Values Lower Alpha	Boundary Values Upper Alpha
1	-2.96259	2.96259
2	-2.35902	2.35902
3	-2.01409	2.01409

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.2.3. The boundary plot also displays the information level and critical value for the corresponding fixed-sample design. This design has characteristics of an O'Brien-Fleming design; the probability for early stopping is low, and the maximum information and critical values at the final stage are similar to those of the corresponding fixed-sample design.

Output 81.2.3 Boundary Plot



With the MODEL=REG option in the SAMPLESIZE statement, the "Sample Size Summary" table in Output 81.2.4 displays the parameters for the sample size computation.

Output 81.2.4 Required Sample Size Summary

Sample Size Summary

Test	Reg Parameter
Parameter	0.1
Variance	5
X Variance	64
R Square (X)	0.1
Max Sample Size	92.87739
Expected Sample Size (Null Ref)	92.35845
Expected Sample Size (Alt Ref)	70.97617

The "Sample Sizes" table in Output 81.2.5 displays the required sample sizes for the group sequential clinical trial.

Output 81.2.5 Required Sample Sizes

Sample Sizes (N)
Z Test for Regression Parameter

	Fractional N		Ceiling N	
Stage	N	Information	N	Information
1	46.44	535.0	47	541.4
2	69.66	802.5	70	806.4
3	92.88	1069.9	93	1071.4

Thus, 47, 70, and 93 individuals are needed in stages 1, 2, and 3, respectively. Since the sample sizes are derived from estimated values of σ_y^2, r_x^2, and σ_x^2, the actual information levels might not achieve the target information levels. Thus, instead of specifying sample sizes in the protocol, you can specify the maximum information levels. Then if an actual information level is much less than the target level, you can increase the sample sizes for the remaining stages to achieve the desired information levels and power.

Suppose that 47 individuals are available at stage 1. Output 81.2.6 lists the first 10 observations of the trial data.

Output 81.2.6 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Oxygen	Age	Weight	RunTime	Run Pulse	Max Pulse
1	54.5521	44	87.7676	11.6949	178.435	181.607
2	52.2821	40	75.4853	9.8872	184.433	183.667
3	62.1871	44	89.0638	8.7950	155.540	167.108
4	65.3269	42	67.7310	8.4577	162.926	173.877
5	59.9809	37	93.1902	9.3228	179.033	180.144
6	52.5588	47	75.9044	12.0385	177.753	175.033
7	51.7838	40	73.5422	11.6607	175.838	178.140
8	57.0024	43	81.2861	11.2219	160.963	171.770
9	48.0775	44	85.2290	13.1789	173.722	176.548
10	68.3357	38	80.2490	8.5066	171.824	184.011

The following statements use the REG procedure to estimate the slope β_w and its associated standard error at stage 1:

```
proc reg data=Fit_1;
   model Oxygen=Age Weight RunTime RunPulse MaxPulse;
   ods output ParameterEstimates=Parms_Fit1;
run;
```

The following statements create and display (in Output 81.2.7) the input data set that contains slope β_w and its associated standard error for the SEQTEST procedure:

```
data Parms_Fit1;
   set Parms_Fit1;
   if Variable='Weight';
   _Scale_='MLE';
   _Stage_= 1;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_Fit1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.2.7 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Weight	0.03772	0.04345	MLE	1

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Fit
             Parms(Testvar=Weight)=Parms_Fit1
             infoadj=prop
             errspendadj=errfuncobf
             order=lr
             stopprob
             ;
ods output Test=Test_Fit1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1 (this data set was generated in the SEQDESIGN procedure). Recall that these boundary values were derived for the information levels specified with the INFO=CUM(2 3 4) option in the SEQDESIGN procedure. The PARMS=PARMS_FIT1 option specifies the input data set PARMS_FIT1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=WEIGHT option identifies the test variable WEIGHT in the data set.

If the computed information level for stage 1 is not the same as the value provided in the BOUNDARY= data set, the INFOADJ=PROP option (which is the default) proportionally adjusts the information levels at future interim stages from the levels provided in the BOUNDARY= data set. The ORDER=LR option uses the LR ordering to derive the p-value, the unbiased median estimate, and the confidence limits for the regression slope estimate. The ERRSPENDADJ=ERRFUNCOBF option adjusts the boundaries with the updated error spending values generated from an O'Brien-Fleming-type cumulative error spending function.

The ODS OUTPUT statement with the TEST=TEST_FIT1 option creates an output data set named TEST_FIT1 which contains the updated boundary information for the test at stage 1. The adjustment is needed because the observed information level is different from the information level in the BOUNDARY= data set. The data set TEST_FIT1 also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.2.8 displays the design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are modified for the new information levels to maintain the Type I α level. The maximum information remains the same as in the BOUNDARY= data set, but the derived Type II error probability β and power $1-\beta$ are different because of the new information level.

Output 81.2.8 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_FIT
Data Set	WORK.PARMS_FIT1
Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Number of Stages	3
Alpha	0.05
Beta	0.09994
Power	0.90006
Max Information (Percent of Fixed Sample)	101.8057
Max Information	1069.94751
Null Ref ASN (Percent of Fixed Sample)	101.2416
Alt Ref ASN (Percent of Fixed Sample)	77.87607

With the STOPPROB option, the "Expected Cumulative Stopping Probabilities" table in Output 81.2.9 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 $c_i = 0, 0.5, 1, 1.5$ by default. You can specify other values for c_i with the CREF= option.

Output 81.2.9 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	2.978	Reject Null	0.00289	0.01906	0.05000
0.5000	2.792	Reject Null	0.03373	0.17443	0.36566
1.0000	2.069	Reject Null	0.24884	0.68206	0.90006
1.5000	1.348	Reject Null	0.68172	0.97032	0.99820

The "Test Information" table in Output 81.2.10 displays the boundary values for the test statistic. By default (or equivalently if you specify BOUNDARYSCALE=STDZ), these statistics are displayed with the standardized Z scale. The information level at stage 1 is derived from the standard error s_1 in the PARMS= data set,

$$I_1 = \frac{1}{s_1^2} = \frac{1}{0.043453^2} = 529.62$$

Output 81.2.10 Sequential Tests

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	--Information Level- Proportion	Actual	------Alternative----- -------Reference------ Lower	Upper	----Boundary Values--- ---Lower-- Alpha	---Upper-- Alpha
1	0.4950	529.6232	-2.30135	2.30135	-2.97951	2.97951
2	0.7475	799.7853	-2.82805	2.82805	-2.36291	2.36291
3	1.0000	1069.948	-3.27101	3.27101	-2.01336	2.01336

Test Information (Standardized Z Scale)
Null Reference = 0

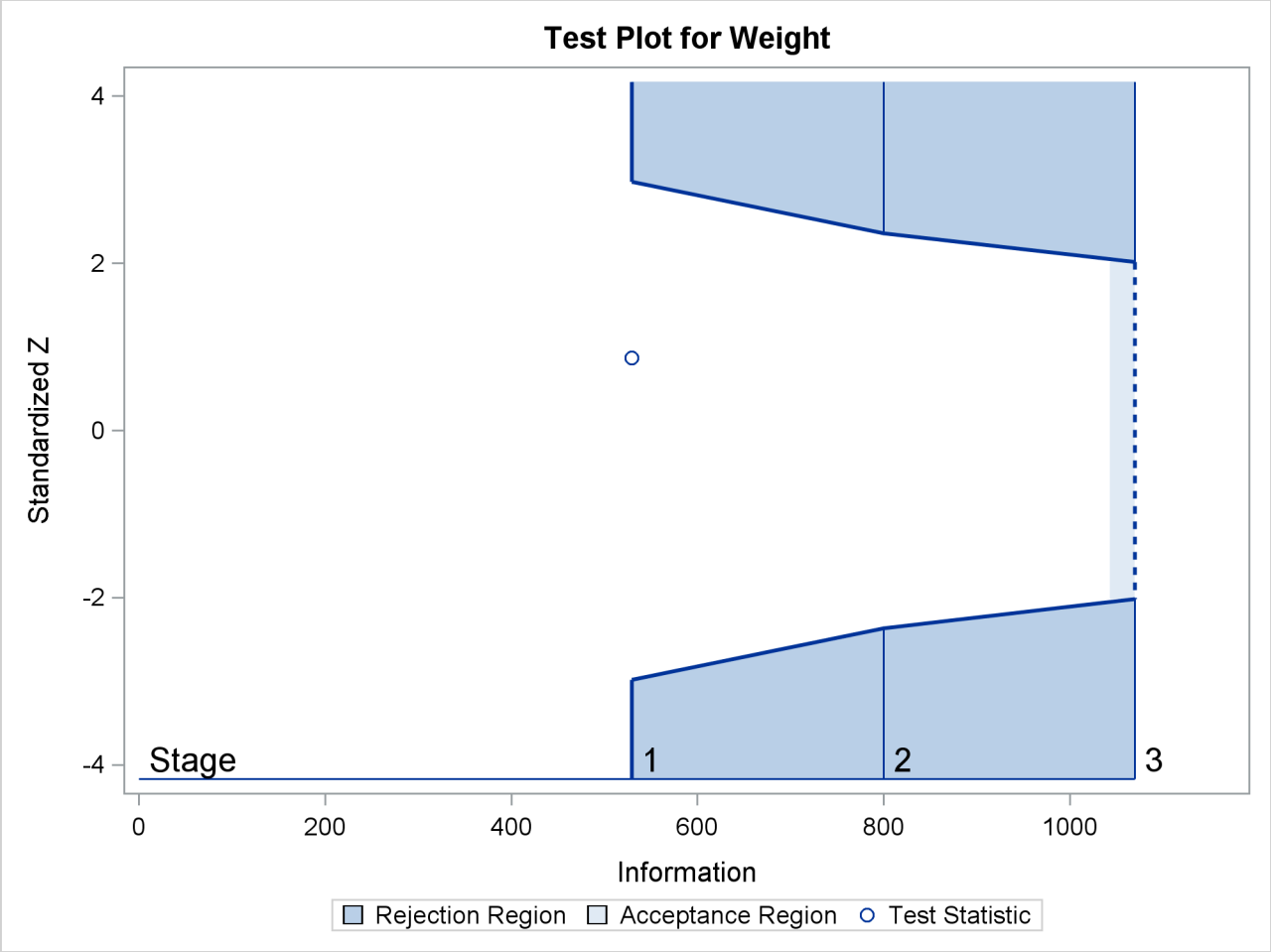
Stage	----------Test---------- ----------Weight--------- Estimate	Action
1	0.86798	Continue
2	.	
3	.	

With the INFOADJ=PROP option (which is the default), the information level at stage 2 is derived proportionally from the observed information at stage 1 and the information levels in the BOUNDARY= data set. See the section "Boundary Adjustments for Information Levels" on page 6895 for details about how the adjusted information levels are computed.

At stage 1, the standardized Z statistic 0.86798 is between the lower α boundary -2.97951 and the upper α boundary 2.97951, so the trial continues to the next stage.

With ODS Graphics enabled, a boundary plot with test statistics is displayed, as shown in Output 81.2.11. As expected, the test statistic is in the continuation region between the lower and upper α boundaries.

Output 81.2.11 Sequential Test Plot



The following statements use the REG procedure to estimate the slope β_w and its associated standard error at stage 2:

```
proc reg data=Fit_2;
   model Oxygen=Age Weight RunTime RunPulse MaxPulse;
   ods output ParameterEstimates=Parms_Fit2;
run;
```

Note that the data set Fit_2 contains both the data from stage 1 and the data from stage 2.

The following statements create and display (in Output 81.2.12) the input data set that contains slope β_w and its associated standard error at stage 2 for the SEQTEST procedure:

```
data Parms_Fit2;
   set Parms_Fit2;
   if Variable='Weight';
   _Scale_='MLE';
   _Stage_= 2;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;
```

```
proc print data=Parms_Fit2;
   title 'Statistics Computed at Stage 2';
run;
```

Output 81.2.12 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Weight	0.02932	0.03520	MLE	2

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
proc seqtest Boundary=Test_Fit1
             Parms(Testvar=Weight)=Parms_Fit2
             infoadj=prop
             errspendadj=errfuncobf
             order=lr
             ;
ods output Test=Test_Fit2;
run;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=TEST_FIT2 option creates an output data set named TEST_FIT2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

Since the data set PARMS_FIT2 does not contain the test information at stage 1, the information level at stage 1 in the TEST_FIT1 data set is used to generate boundary values for the test at stage 2.

Following the process at stage 1, the slope estimate is also between its corresponding lower and upper α boundary values, so the trial continues to the next stage.

The following statements use the REG procedure to estimate the slope β_w and its associated standard error at the final stage:

```
proc reg data=Fit_3;
   model Oxygen=Age Weight RunTime RunPulse MaxPulse;
   ods output ParameterEstimates=Parms_Fit3;
run;
```

The following statements create the input data set that contains slope β_w and its associated standard error at stage 3 for the SEQTEST procedure:

```
data Parms_Fit3;
   set Parms_Fit3;
   if Variable='Weight';
   _Scale_='MLE';
   _Stage_= 3;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;
```

The following statements print (in Output 81.2.13) the test statistics at stage 3:

```
proc print data=Parms_Fit3;
   title 'Statistics Computed at Stage 3';
run;
```

Output 81.2.13 Statistics Computed at Stage 3

Statistics Computed at Stage 3

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Weight	0.02189	0.03028	MLE	3

The following statements invoke the SEQTEST procedure to test the hypothesis:

```
ods graphics on;
proc seqtest Boundary=Test_Fit2
             Parms(testvar=Weight)=Parms_Fit3
             errspendadj=errfuncobf
             order=lr
             pss
             plots=(asn power)
             ;
ods output Test=Test_Fit3;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 3, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 3, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=TEST_FIT3 option creates an output data set named TEST_FIT3 which contains the updated boundary information for the test at stage 3.

The "Design Information" table in Output 81.2.14 displays design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are modified for the new information levels to maintain the Type I α level.

Output 81.2.14 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.TEST_FIT2
Data Set	WORK.PARMS_FIT3
Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Number of Stages	3
Alpha	0.05
Beta	0.09514
Power	0.90486
Max Information (Percent of Fixed Sample)	102.0102
Max Information	1090.63724
Null Ref ASN (Percent of Fixed Sample)	101.4122
Alt Ref ASN (Percent of Fixed Sample)	77.22139

The maximum information is derived from the standard error associated with the slope estimate at the final stage and is larger than the target level. The derived Type II error probability β and power $1-\beta$ are different because of the new information levels.

With the PSS option, the "Power and Expected Sample Sizes" table in Output 81.2.15 displays powers and expected mean sample sizes under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and $c_i = 0, 0.5, 1, 1.5$ by default. You can specify the c_i values with the CREF= option.

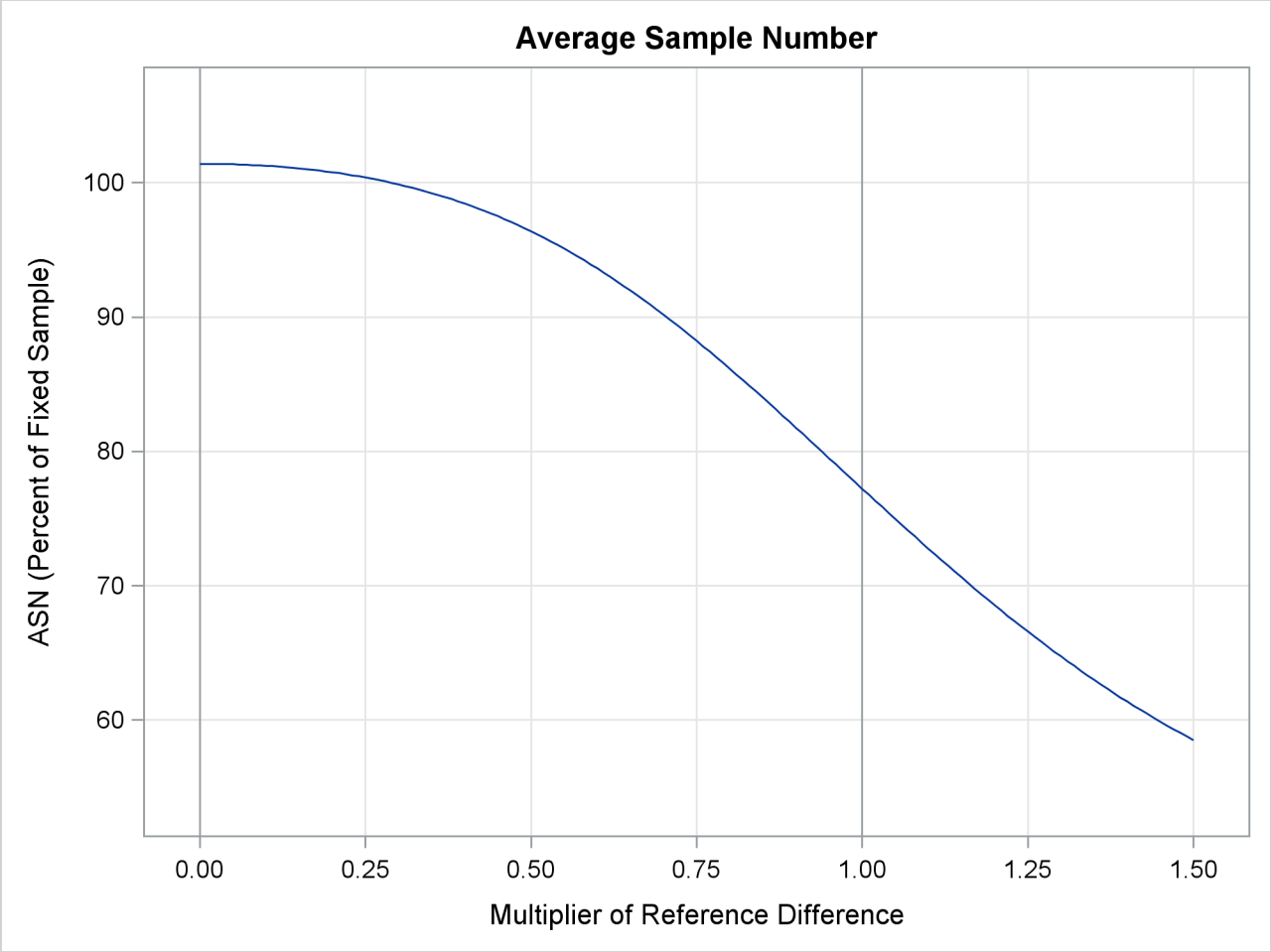
Output 81.2.15 Power and Expected Sample Size Information

Powers and Expected Sample Sizes
Reference = CRef * (Alt Reference)

CRef	Power	-Sample Size- Percent Fixed-Sample
0.0000	0.02500	101.4122
0.5000	0.37046	96.3754
1.0000	0.90486	77.2214
1.5000	0.99844	58.5301

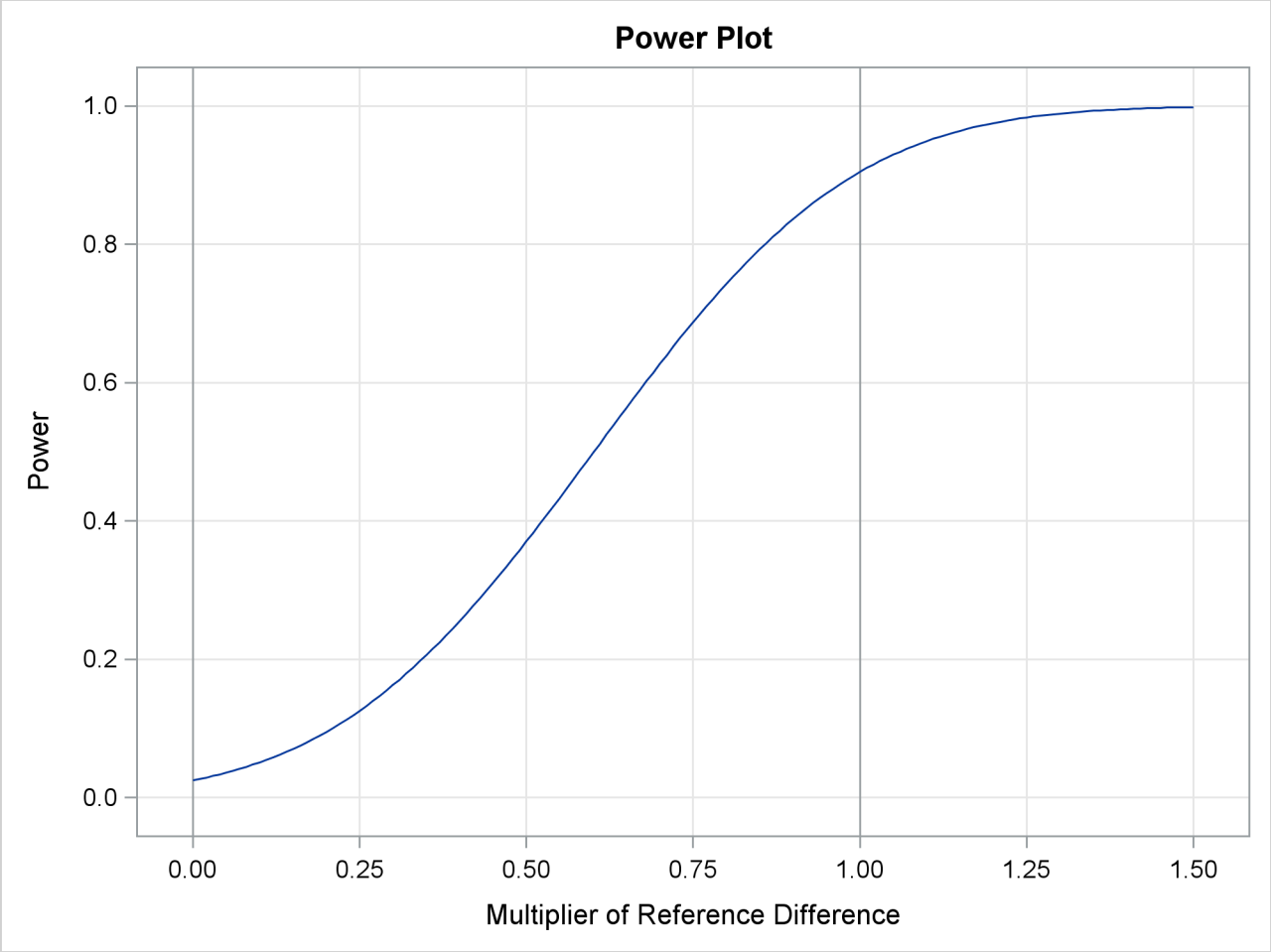
With the PLOTS=ASN option, the procedure displays a plot of expected sample sizes under various hypothetical references, as shown in Output 81.2.16. By default, expected sample sizes under the hypotheses $\theta = c_i\,\theta_1$, $c_i = 0, 0.01, 0.02, \ldots, 1.50$, are displayed, where θ_1 is the alternative reference.

Output 81.2.16 ASN Plot



With the PLOTS=POWER option, the procedure displays a plot of powers under various hypothetical references, as shown in Output 81.2.17. By default, powers under hypothetical references $\theta = c_i\,\theta_1$ are displayed, where $c_i = 0, 0.01, 0.02, \ldots, 1.50$ by default. You can specify c_i values with the CREF= option. The c_i values are displayed on the horizontal axis.

Output 81.2.17 Power Plot



Under the null hypothesis, $c_i = 0$, the power is 0.025, which is the upper Type I error probability. Under the alternative hypothesis, $c_i = 1$, the power is 0.90486, which is one minus the Type II error probability.

The "Test Information" table in Output 81.2.18 displays the boundary values for the test statistic with the default standardized Z scale. The information level at the current stage is derived from the standard error for the current stage in the PARMS= data set. At stage 3, the standardized slope estimate 0.72284 is still between the lower and upper α boundary values. Since it is the final stage, the trial stops to accept the null hypothesis that the variable Weight has no effect on the oxygen intake rate after adjusting for other covariates.

Output 81.2.18 Sequential Tests

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	--Information Level- Proportion	Actual	------Alternative----- -------Reference------ Lower	Upper	----Boundary Values--- ---Lower-- Alpha	---Upper-- Alpha
1	0.4856	529.6232	-2.30135	2.30135	-2.97951	2.97951
2	0.7401	807.1954	-2.84112	2.84112	-2.34945	2.34945
3	1.0000	1090.637	-3.30248	3.30248	-2.01885	2.01885

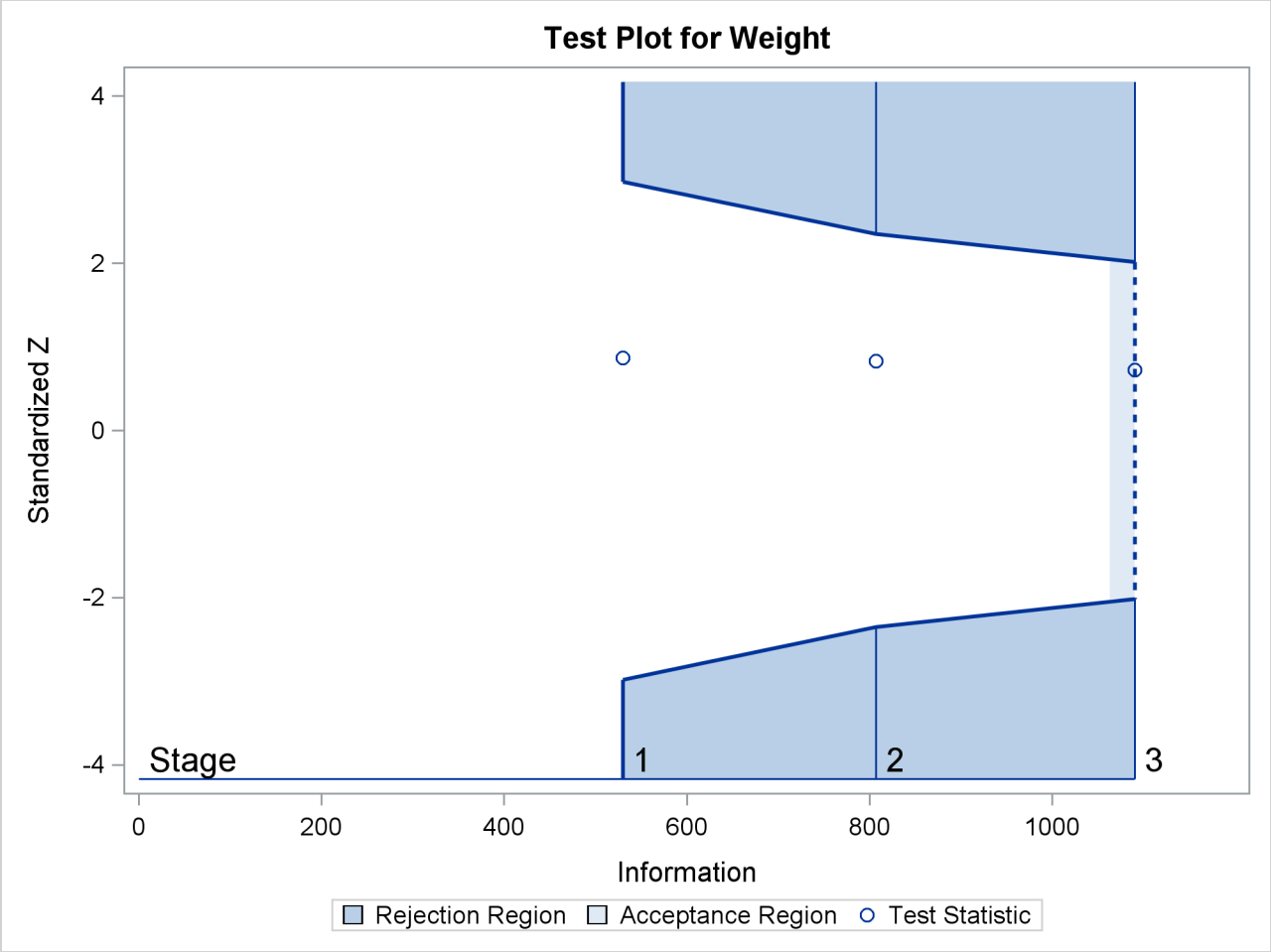
Test Information (Standardized Z Scale)
Null Reference = 0

Stage	----------Test---------- ----------Weight--------- Estimate	Action
1	0.86798	Continue
2	0.83305	Continue
3	0.72284	Accept Null

Since the data set FIT_3 contains the test information only at stage 3, the information levels at previous stages in the TEST_FIT2 data set are used to generate boundary values for the test.

With ODS Graphics enabled, a boundary plot with test statistics is displayed, as shown in Output 81.2.19. As expected, the test statistic is in the acceptance region between the lower and upper α boundaries at the final stage.

Output 81.2.19 Sequential Test Plot



After a trial is stopped, the "Parameter Estimates" table in Output 81.2.20 displays the stopping stage, parameter estimate, unbiased median estimate, confidence limits, and the p-value under the null hypothesis $H_0 : \beta_w = 0$.

Output 81.2.20 Parameter Estimates

Parameter Estimates
LR Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate
Weight	3	0.021888	0.4699	0.021884

Parameter Estimates
LR Ordering

Parameter	95% Confidence Limits	
Weight	-0.03747	0.08123

As expected, the p-value 0.4699 is not significant at the $\alpha = 0.05$ level, and the confidence interval does contain the value zero. The p-value, unbiased median estimate, and confidence limits depend on the ordering of the sample space (k, z), where k is the stage number and z is the standardized Z statistic. With the specified LR ordering, the p-values are computed with the ordering $(k', z') \succ (k, z)$ if $z' > z$. See the section "Available Sample Space Orderings in a Sequential Test" on page 6902 for a detailed description of the LR ordering.

Example 81.3: Testing an Effect with Early Stopping to Accept H_0

This example demonstrates a two-sided group sequential test that uses an error spending design with early stopping to accept the null hypothesis H_0. The example is similar to Example 81.2 but with early stopping to accept H_0.

A study is conducted to examine the effects of Age (years), Weight (kg), RunTime (time in minutes to run 1.5 miles), RunPulse (heart rate while running), and MaxPulse (maximum heart rate recorded while running) on Oxygen (oxygen intake rate, ml per kg body weight per minute). The primary interest is whether oxygen intake rate is associated with weight.

The hypothesis is tested using the following linear model:

$$\text{Oxygen} = \text{Age} + \text{Weight} + \text{RunTime} + \text{RunPulse} + \text{MaxPulse}$$

The null hypothesis is $H_0 : \beta_w = 0$, where β_w is the regression parameter for the variable Weight. Suppose that $\beta_w = 0.10$ is the reference improvement that should be detected at a 0.90 level. Then the maximum information I_X can be derived in the SEQDESIGN procedure.

Following the derivations in the section "Test for a Parameter in the Regression Model" in the chapter "The SEQDESIGN Procedure," the required sample size can be derived from

$$N = I_X \frac{\sigma_y^2}{(1 - r_x^2)\,\sigma_x^2}$$

where σ_y^2 is the variance of the response variable in the regression model, r_x^2 is the proportion of variance of Weight explained by other covariates, and σ_x^2 is the variance of Weight.

Further suppose that from past experience, $\sigma_y^2 = 5$, $r_x^2 = 0.10$, and $\sigma_x^2 = 64$. Then the required sample size can be derived using the SAMPLESIZE statement in the SEQDESIGN procedure.

The following statements invoke the SEQDESIGN procedure and request a three-stage group sequential design for normally distributed data to test the null hypothesis of a regression parameter $H_0 : \beta_w = 0$ against the alternative $H_1 : \beta_w \neq 0$:

```
ods graphics on;
proc seqdesign altref=0.10;
   OBFErrorFunction: design method=errfuncgamma
                            stop=accept
                            nstages=3
                            info=cum(2 3 4);
```

```
   samplesize model=reg( variance=5 xvariance=64 xrsquare=0.10);
ods output Boundary=Bnd_Fit;
run;
ods graphics off;
```

By default (or equivalently if you specify ALPHA=0.05 and BETA=0.10), the procedure uses a Type I error probability 0.05 and a Type II error probability 0.10. The ALTREF=0.10 option specifies a power of $1-\beta = 0.90$ at the alternative hypothesis $H_1 : \beta_w = \pm 0.10$. The INFO=CUM(2 3 4) option specifies that the study perform the first interim analysis with information proportion $2/4 = 0.5$—that is, after half of the total observations are collected.

The ODS OUTPUT statement with the BOUNDARY=BND_FIT option creates an output data set named BND_FIT which contains the resulting boundary information for the subsequent sequential tests.

The "Design Information" table in Output 81.3.1 displays design specifications and derived statistics. Since the alternative reference is specified, the maximum information is derived.

Output 81.3.1 Error Spending Design Information

The SEQDESIGN Procedure
Design: OBFErrorFunction

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	0.1
Number of Stages	3
Alpha	0.05
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	103.9245
Max Information	1091.972
Null Ref ASN (Percent of Fixed Sample)	75.00521
Alt Ref ASN (Percent of Fixed Sample)	101.8099

The "Boundary Information" table in Output 81.3.2 displays information level, alternative reference, and boundary values at each stage.

Output 81.3.2 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

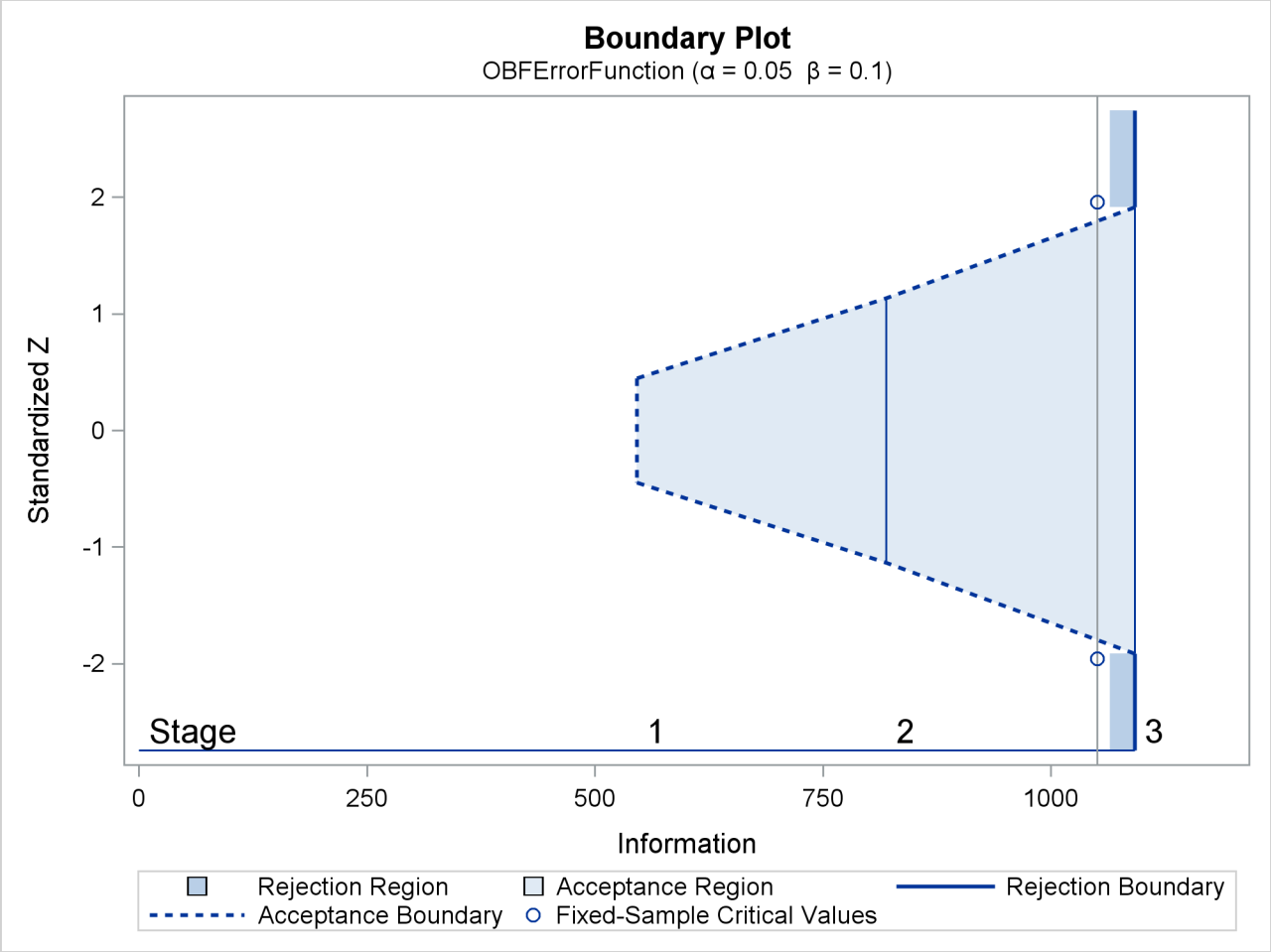
Stage	Information Level Proportion	Information Level Actual	Information Level N	Alternative Reference Lower	Alternative Reference Upper
1	0.5000	545.9862	47.39463	-2.33663	2.33663
2	0.7500	818.9792	71.09195	-2.86178	2.86178
3	1.0000	1091.972	94.78926	-3.30450	3.30450

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Boundary Values Lower Beta	Boundary Values Upper Beta
1	-0.44937	0.44937
2	-1.13583	1.13583
3	-1.91428	1.91428

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.3.3. The boundary plot also displays the information level and critical value for the corresponding fixed-sample design.

Output 81.3.3 Boundary Plot



With the MODEL=REG option in the SAMPLESIZE statement, the "Sample Size Summary" table in Output 81.3.4 displays the parameters for the sample size computation.

Output 81.3.4 Required Sample Size Summary

Sample Size Summary

Test	Reg Parameter
Parameter	0.1
Variance	5
X Variance	64
R Square (X)	0.1
Max Sample Size	94.78926
Expected Sample Size (Null Ref)	68.41207
Expected Sample Size (Alt Ref)	92.86057

The "Sample Sizes" table in Output 81.3.5 displays the required sample sizes for the group sequential clinical trial.

Output 81.3.5 Required Sample Sizes

Sample Sizes (N)
Z Test for Regression Parameter

	Fractional N		Ceiling N	
Stage	N	Information	N	Information
1	47.39	546.0	48	553.0
2	71.09	819.0	72	829.4
3	94.79	1092.0	95	1094.4

Thus, 48, 72, and 95 individuals are needed in stages 1, 2, and 3, respectively. Since the sample sizes are derived from estimated values of σ_y^2, r_x^2, and σ_x^2, the actual information levels might not achieve the target information levels. Thus, instead of specifying sample sizes in the protocol, you can specify the maximum information levels. Then if an actual information level is much less than the target level, you can increase the sample sizes for the remaining stages to achieve the desired information levels and power.

Suppose that 48 individuals are available at stage 1. Output 81.3.6 lists the first 10 observations of the trial data.

Output 81.3.6 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Oxygen	Age	Weight	RunTime	Run Pulse	Max Pulse
1	54.5521	44	87.7676	11.6949	178.435	181.607
2	52.2821	40	75.4853	9.8872	184.433	183.667
3	62.1871	44	89.0638	8.7950	155.540	167.108
4	65.3269	42	67.7310	8.4577	162.926	173.877
5	59.9809	37	93.1902	9.3228	179.033	180.144
6	52.5588	47	75.9044	12.0385	177.753	175.033
7	51.7838	40	73.5422	11.6607	175.838	178.140
8	57.0024	43	81.2861	11.2219	160.963	171.770
9	48.0775	44	85.2290	13.1789	173.722	176.548
10	68.3357	38	80.2490	8.5066	171.824	184.011

The following statements use the REG procedure to estimate the slope β_w and its associated standard error at stage 1:

```
proc reg data=Fit_1;
   model Oxygen=Age Weight RunTime RunPulse MaxPulse;
   ods output ParameterEstimates=Parms_Fit1;
run;
```

The following statements create and display (in Output 81.3.7) the input data set that contains slope β_w and its associated standard error for the SEQTEST procedure:

```
data Parms_Fit1;
   set Parms_Fit1;
   if Variable='Weight';
   _Scale_='MLE';
   _Stage_= 1;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_Fit1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.3.7 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Weight	0.04660	0.04308	MLE	1

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Fit
             Parms(testvar=Weight)=Parms_Fit1
             infoadj=none
             errspendadj=errfuncgamma
             stopprob
             order=lr
             ;
ods output Test=Test_Fit1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The PARMS=PARMS_FIT1 option specifies the input data set PARMS_FIT1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=WEIGHT option identifies the test variable WEIGHT in the data set. The INFOADJ=NONE option maintains the information level for stage 2 at the value provided in the BOUNDARY= data set.

The ORDER=LR option uses the LR ordering to derive the p-value, the unbiased median estimate, and the confidence limits for the regression slope estimate. The ERRSPENDADJ=ERRFUNCGAMMA option adjusts the boundaries with the updated error spending values generated from a gamma cumulative error spending function.

The ODS OUTPUT statement with the TEST=TEST_FIT1 option creates an output data set named TEST_FIT1 which contains the updated boundary information for the test at stage 1. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.3.8 displays the design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are modified for the new informa-

tion levels to maintain the Type I α level. The maximum information remains the same as in the BOUNDARY= data set, but the derived Type II error probability β and power $1-\beta$ are different because of the new information level.

Output 81.3.8 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_FIT
Data Set	WORK.PARMS_FIT1
Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Accept Null
Number of Stages	3
Alpha	0.05
Beta	0.10007
Power	0.89993
Max Information (Percent of Fixed Sample)	103.9498
Max Information	1091.97232
Null Ref ASN (Percent of Fixed Sample)	75.15846
Alt Ref ASN (Percent of Fixed Sample)	101.8296

With the STOPPROB option, the "Expected Cumulative Stopping Probabilities" table in Output 81.3.9 displays the expected stopping stage and the cumulative stopping probability of accepting the null hypothesis at each stage under various hypothetical references $\theta = c_i\theta_1$, where θ_1 is the alternative reference and $c_i = 0, 0.5, 1, 1.5$ by default. You can specify other values for c_i with the CREF= option.

Output 81.3.9 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	1.895	Accept Null	0.33304	0.76607	0.95000
0.5000	2.409	Accept Null	0.17680	0.40947	0.62828
1.0000	2.918	Accept Null	0.02636	0.05453	0.10007
1.5000	2.997	Accept Null	0.00109	0.00166	0.00242

The "Test Information" table in Output 81.3.10 displays the boundary values for the test statistic. By default (or equivalently if you specify BOUNDARYSCALE=STDZ), these statistics are displayed with the standardized Z scale. The information level at stage 1 is derived from the standard error s_1 in the PARMS= data set,

$$I_1 = \frac{1}{s_1^2} = \frac{1}{0.04308^2} = 538.8$$

Output 81.3.10 Sequential Tests

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Alternative Reference Lower	Alternative Reference Upper
1	0.4934	538.7887	-2.32118	2.32118
2	0.7500	818.9792	-2.86178	2.86178
3	1.0000	1091.972	-3.30450	3.30450

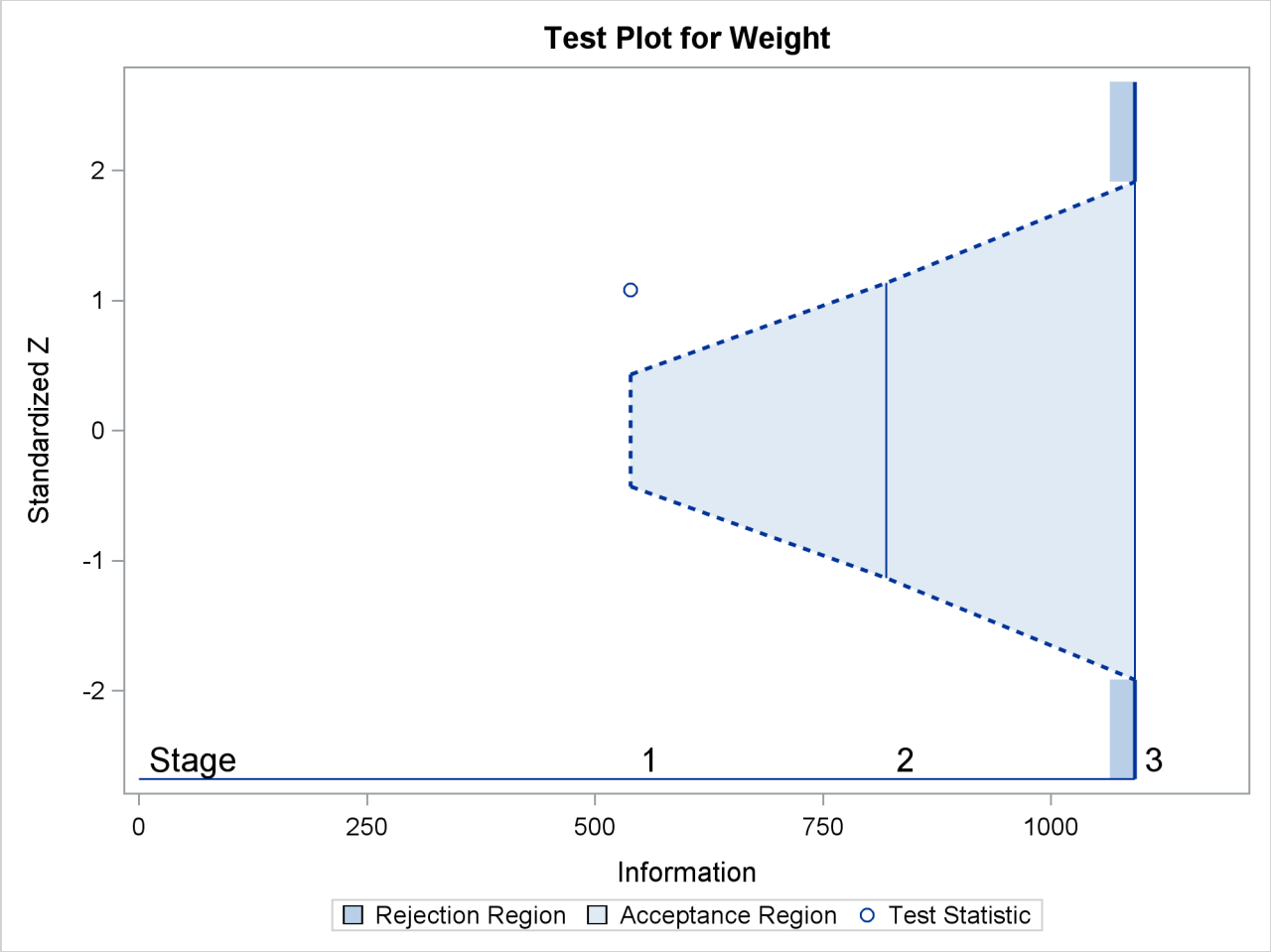
Test Information (Standardized Z Scale)
Null Reference = 0

Stage	Boundary Values Lower Beta	Boundary Values Upper Beta	Test Weight Estimate	Test Weight Action
1	-0.43033	0.43033	1.08174	Continue
2	-1.13623	1.13623	.	
3	-1.91431	1.91431	.	

At stage 1, the standardized Z statistic 1.08174 is greater than the upper β boundary 0.43033, so the trial continues to the next stage.

With ODS Graphics enabled, a boundary plot with test statistics is displayed, as shown in Output 81.3.11. As expected, the test statistic is in the continuation region.

Output 81.3.11 Sequential Test Plot



The following statements use the REG procedure to estimate the slope β_w and its associated standard error at stage 2:

```
proc reg data=Fit_2;
   model Oxygen=Age Weight RunTime RunPulse MaxPulse;
   ods output ParameterEstimates=Parms_Fit2;
run;
```

Note that the data set Fit_2 contains both the data from stage 1 and the data from stage 2,

The following statements create and display (in Output 81.3.12) the input data set that contains slope β_w and its associated standard error at stage 2 for the SEQTEST procedure:

```
data Parms_Fit2;
   set Parms_Fit2;
   if Variable='Weight';
   _Scale_='MLE';
   _Stage_= 2;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;
```

```
proc print data=Parms_Fit2;
   title 'Statistics Computed at Stage 2';
run;
```

Output 81.3.12 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	Weight	0.02925	0.03490	MLE	2

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
ods graphics on;
proc seqtest Boundary=Test_Fit1
             Parms(testvar=Weight)=Parms_Fit2
             errspendadj=errfuncgamma
             order=lr
             pss
             plots=(asn power)
             ;
ods output Test=Test_Fit2;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set.

Since the data set PARMS_FIT2 does not contain the test information at stage 1, the information level at stage 1 in the TEST_FIT1 data set is used to generate boundary values for the test.

The ORDER=LR option uses the LR ordering to derive the p-value, unbiased median estimate, and confidence limits for the regression slope estimate.

The ODS OUTPUT statement with the TEST=TEST_FIT2 option creates an output data set named TEST_FIT2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.3.13 displays design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are modified for the new information levels to maintain the Type I α level.

Output 81.3.13 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.TEST_FIT1
Data Set	WORK.PARMS_FIT2
Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Accept Null
Number of Stages	3
Alpha	0.05
Beta	0.10009
Power	0.89991
Max Information (Percent of Fixed Sample)	103.9566
Max Information	1091.97232
Null Ref ASN (Percent of Fixed Sample)	75.18254
Alt Ref ASN (Percent of Fixed Sample)	101.8349

The derived Type II error probability β and power $1-\beta$ are different because of the new information levels.

With the PSS option, the "Power and Expected Sample Sizes" table in Output 81.3.14 displays powers and expected mean 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.

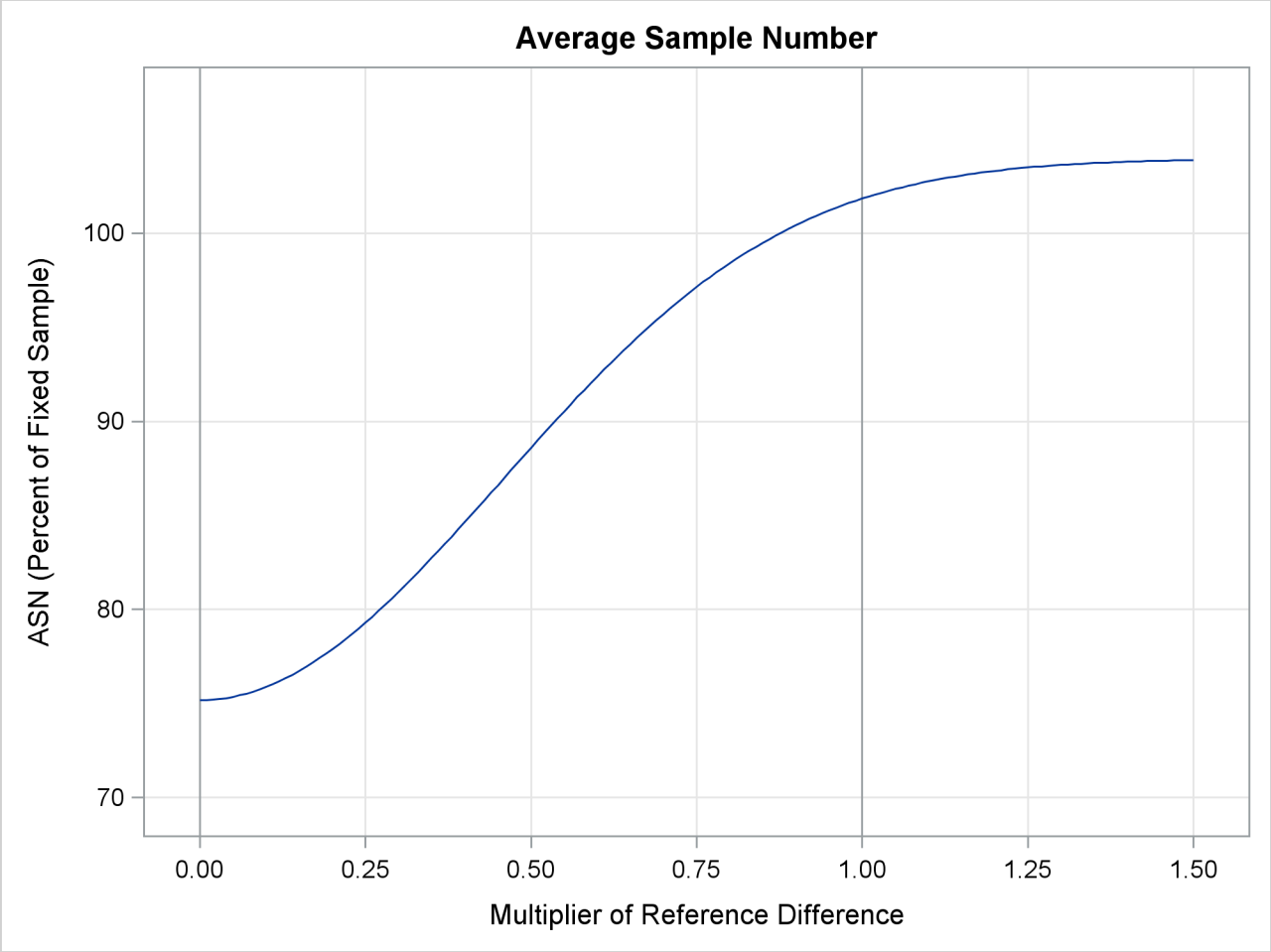
Output 81.3.14 Power and Expected Sample Size Information

Powers and Expected Sample Sizes
Reference = CRef * (Alt Reference)

CRef	Power	-Sample Size- Percent Fixed-Sample
0.0000	0.02500	75.1825
0.5000	0.37154	88.5975
1.0000	0.89991	101.8349
1.5000	0.99758	103.8843

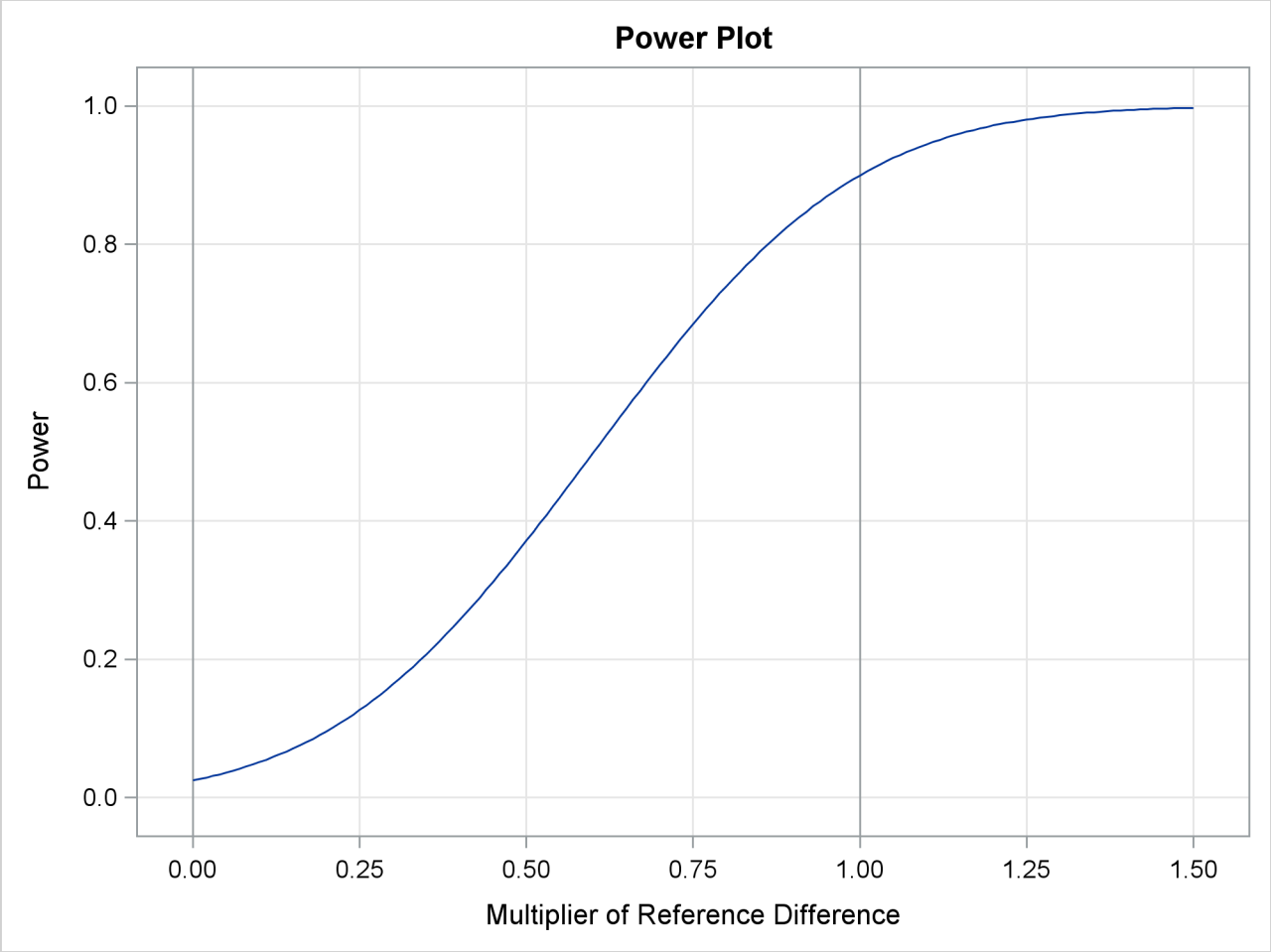
With the PLOTS=ASN option, the procedure displays a plot of expected sample sizes under various hypothetical references, as shown in Output 81.3.15. By default, expected sample sizes under the hypotheses $\theta = c_i\,\theta_1$, $c_i = 0, 0.01, 0.02, \ldots, 1.50$, are displayed, where θ_1 is the alternative reference.

Output 81.3.15 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 81.3.16. By default, powers under hypothetical references $\theta = c_i\,\theta_1$ are displayed, where $c_i = 0, 0.01, 0.02, \ldots, 1.50$ by default. You can specify c_i values with the CREF= option. The c_i values are displayed on the horizontal axis.

Output 81.3.16 Power Plot



Under the null hypothesis, $c_i = 0$, the power is 0.025, which is the upper Type I error probability. Under the alternative hypothesis, $c_i = 1$, the power is 0.89991, which is one minus the Type II error probability, as displayed in the "Design Information" table in Output 81.3.13.

The "Test Information" table in Output 81.3.17 displays the boundary values for the test statistic with the default standardized Z scale. At stage 2, the standardized slope estimate 0.83805 is between the lower and upper β boundary values. The trial stops to accept the null hypothesis that the variable Weight has no effect on the oxygen intake rate after adjusting for other covariates.

Output 81.3.17 Sequential Tests

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Alternative Reference Lower	Alternative Reference Upper
1	0.4934	538.7887	-2.32118	2.32118
2	0.7517	820.8509	-2.86505	2.86505
3	1.0000	1091.972	-3.30450	3.30450

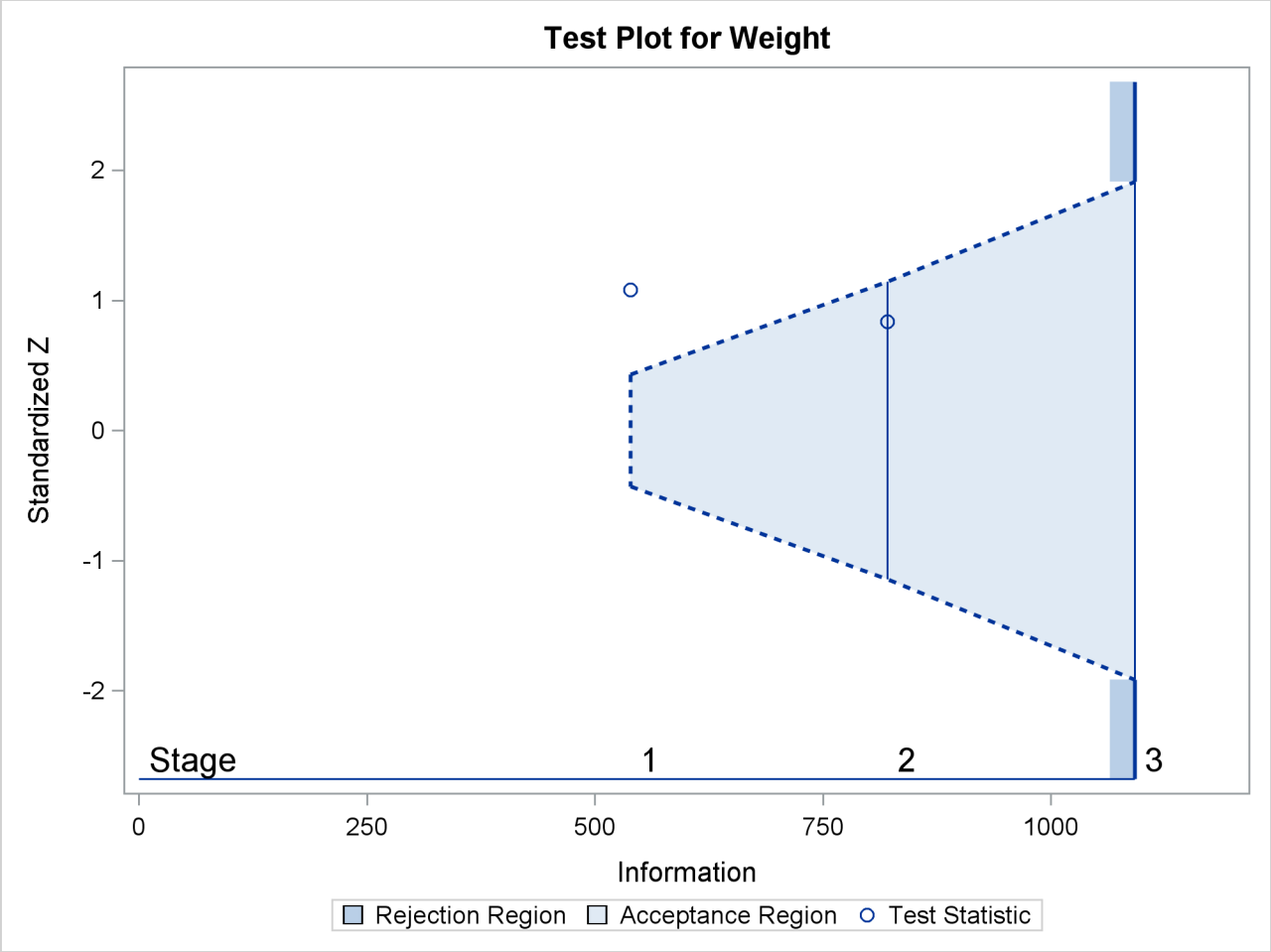
Test Information (Standardized Z Scale)
Null Reference = 0

Stage	Boundary Values Lower Beta	Boundary Values Upper Beta	Test Weight Estimate	Test Weight Action
1	-0.43033	0.43033	1.08174	Continue
2	-1.14239	1.14239	0.83805	Accept Null
3	-1.91408	1.91408	.	

Since the data set PARMS_FIT2 contains the test information only at stage 2, the information level at stage 1 in the TEST_FIT1 data set is used to generate boundary values for the test.

With ODS Graphics enabled, a boundary plot with test statistics is displayed, as shown in Output 81.3.18. As expected, the test statistic is in the acceptance region between the lower and upper α boundaries at the final stage.

Output 81.3.18 Sequential Test Plot



After a trial is stopped, the "Parameter Estimates" table in Output 81.3.19 displays the stopping stage, parameter estimate, unbiased median estimate, confidence limits, and the p-value under the null hypothesis $H_0 : \beta_w = 0$. As expected, the p-value 0.3056 is not significant at the $\alpha = 0.05$ level, and the confidence interval does contain the value zero. The p-value, unbiased median estimate, and confidence limits depend on the ordering of the sample space (k, z), where k is the stage number and z is the standardized Z statistic. With the specified LR ordering, the p-values are computed with the ordering $(k', z') \succ (k, z)$ if $z' > z$. See the section "Available Sample Space Orderings in a Sequential Test" on page 6902 for a detailed description of the LR ordering.

Output 81.3.19 Parameter Estimates

Parameter Estimates
LR Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate
Weight	2	0.029251	0.3056	0.037080

Parameter Estimates
LR Ordering

Parameter	95% Confidence Limits	
Weight	-0.03368	0.10532

Example 81.4: Testing a Binomial Proportion

This example tests a binomial proportion by using a four-stage group sequential design. Suppose a supermarket is developing a new store-brand coffee. From past studies, the positive response for the current store-brand coffee from customers is around 60%. The store is interested in whether the new brand has a better positive response than the current brand.

A power family method is used for the group sequential trial with the null hypothesis $H_0 : p = p_0 = 0.60$ and a one-sided upper alternative with a power of 0.80 at $H_1 : p = 0.70$. To accommodate the zero null reference that is assumed in the SEQDESIGN procedure, an equivalent hypothesis $H_0 : \theta = 0$ with $H_1 : \theta = 0.10$ is used, where $\theta = p - p_0$. The following statements request a power family method with early stopping to reject the null hypothesis:

```
ods graphics on;
proc seqdesign altref=0.10
               boundaryscale=mle
               ;
   PowerFamily: design method=pow
                       nstages=4
                       alt=upper
                       beta=0.20
                       ;
   samplesize model=onesamplefreq( nullprop=0.6);
ods output Boundary=Bnd_Prop;
run;
ods graphics off;
```

The NULLPROP= option in the SAMPLESIZE statement specifies $p_0 = 0.60$ for the sample size computation. The ODS OUTPUT statement with the BOUNDARY=BND_PROP option creates an output data set named BND_PROP which contains the resulting boundary information for the subsequent sequential tests.

With the BOUNDARYSCALE=MLE option, the procedure displays the output boundaries in terms of the maximum likelihood estimates. The "Design Information" table in Output 81.4.1 displays design specifica-

tions and derived statistics. With the specified alternative reference $\theta_1 = p_1 - p_0 = 0.7 - 0.6 = 0.1$, the maximum information 670.38 is also derived.

Output 81.4.1 Design Information

The SEQDESIGN Procedure
Design: PowerFamily

Design Information

Statistic Distribution	Normal
Boundary Scale	MLE
Alternative Hypothesis	Upper
Early Stop	Reject Null
Method	Power Family
Boundary Key	Both
Alternative Reference	0.1
Number of Stages	4
Alpha	0.05
Beta	0.2
Power	0.8
Max Information (Percent of Fixed Sample)	108.4306
Max Information	670.3782
Null Ref ASN (Percent of Fixed Sample)	106.9276
Alt Ref ASN (Percent of Fixed Sample)	78.51072

The "Boundary Information" table in Output 81.4.2 displays the information level, alternative reference, and boundary values at each stage. With the STOP=REJECT option, only the rejection boundary values are displayed.

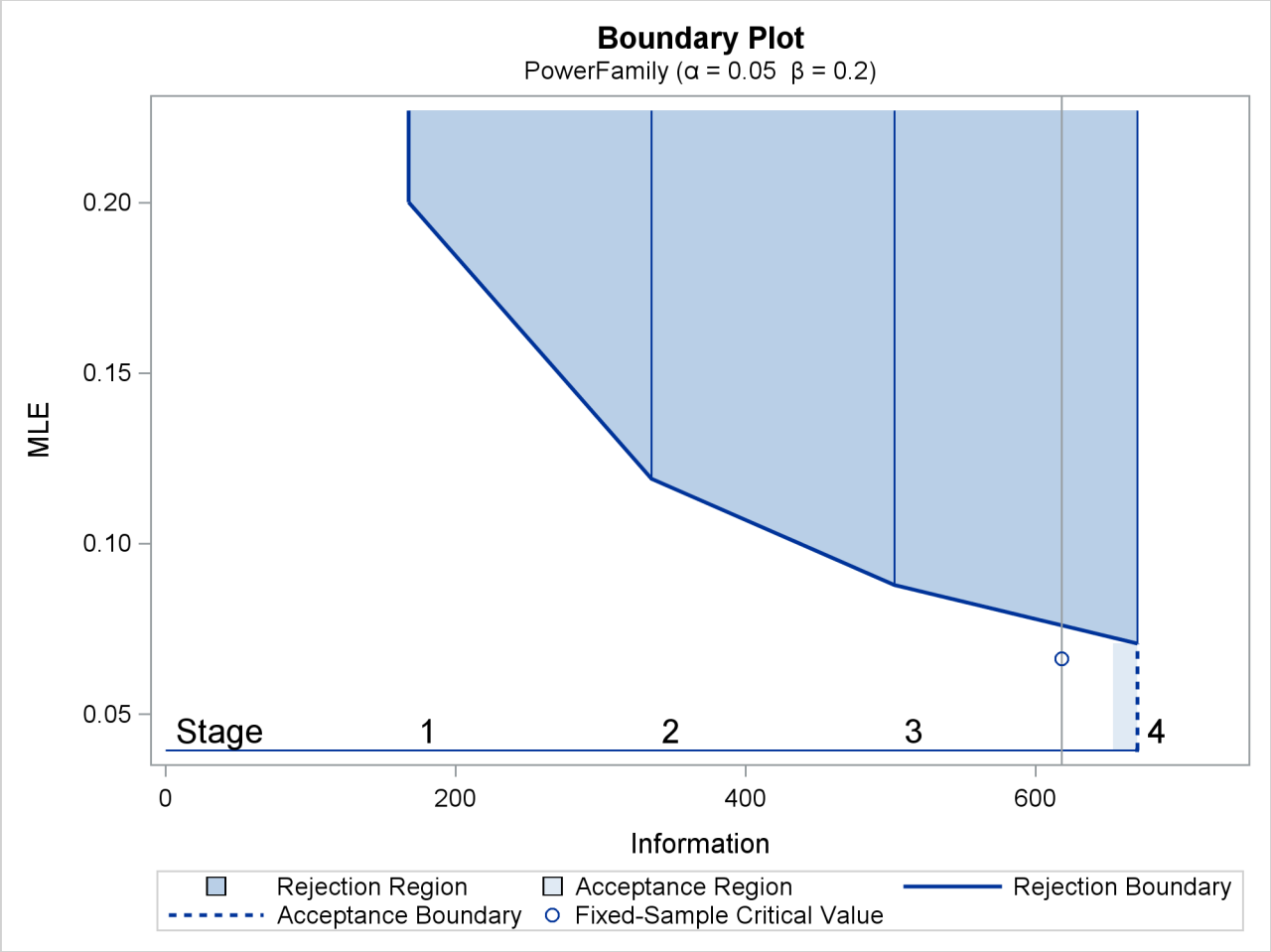
Output 81.4.2 Boundary Information

Boundary Information (MLE Scale)
Null Reference = 0

	Information Level			Alternative Reference	Boundary Values
Stage	Proportion	Actual	N	Upper	Upper Alpha
1	0.2500	167.5945	35.19485	0.10000	0.20018
2	0.5000	335.1891	70.38971	0.10000	0.11903
3	0.7500	502.7836	105.5846	0.10000	0.08782
4	1.0000	670.3782	140.7794	0.10000	0.07077

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.4.3.

Output 81.4.3 Boundary Plot



With the MODEL=ONESAMPLEFREQ option in the SAMPLESIZE statement, the "Sample Size Summary" table in Output 81.4.4 displays the parameters for the sample size computation.

Output 81.4.4 Required Sample Size Summary

Sample Size Summary

Test	One-Sample Proportion
Null Proportion	0.6
Proportion	0.7
Test Statistic	Z for Proportion
Reference Proportion	Alt Ref
Max Sample Size	140.7794
Expected Sample Size (Null Ref)	138.828
Expected Sample Size (Alt Ref)	101.9333

The "Sample Sizes" table in Output 81.4.5 displays the required sample sizes for the group sequential clinical trial.

Output 81.4.5 Required Sample Sizes

Sample Sizes (N)
One-Sample Z Test for Proportion

	Fractional N		Ceiling N	
Stage	N	Information	N	Information
1	35.19	167.6	36	171.4
2	70.39	335.2	71	338.1
3	105.58	502.8	106	504.8
4	140.78	670.4	141	671.4

Thus, 36 customers are needed at stage 1, and 35 new customers are needed at each of the remaining stages. Suppose that 36 customers are available at stage 1. Output 81.4.6 lists the 10 observations in the data set count_1.

Output 81.4.6 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Resp
1	1
2	1
3	0
4	0
5	1
6	1
7	0
8	1
9	1
10	1

The Resp variable is an indicator variable with a value of 1 for a customer with a positive response and a value of 0 for a customer without a positive response.

The following statements use the MEANS procedure to compute the mean response at stage 1:

```
proc means data=Prop_1;
   var Resp;
   ods output Summary=Data_Prop1;
run;
```

The following statements create and display (in Output 81.4.7) the data set for the centered mean positive response, $\hat{p} - p_0$:

```
data Data_Prop1;
   set Data_Prop1;
   _Scale_='MLE';
   _Stage_= 1;
   NObs= Resp_N;
   PDiff= Resp_Mean - 0.6;
   keep _Scale_ _Stage_ NObs PDiff;
run;
proc print data=Data_Prop1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.4.7 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	_Scale_	_Stage_	NObs	PDiff
1	MLE	1	36	-0.016667

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Prop
             Data(Testvar=PDiff)=Data_Prop1
             infoadj=prop
             boundarykey=both
             boundaryscale=mle
             ;
ods output Test=Test_Prop1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The DATA=DATA_PROP1 option specifies the input data set DATA_PROP1 that contains the test statistic and its associated sample size at stage 1, and the TESTVAR=PDIFF option identifies the test variable PDIFF in the data set.

If the computed information level for stage 1 is not the same as the value provided in the BOUNDARY= data set, the INFOADJ=PROP option (which is the default) proportionally adjusts the information levels at future interim stages from the levels provided in the BOUNDARY= data set. The BOUNDARYKEY=BOTH option maintains both the α and β levels. The BOUNDARYSCALE=MLE option displays the output boundaries in terms of the MLE scale.

The ODS OUTPUT statement with the TEST=TEST_PROP1 option creates an output data set named TEST_PROP1 which contains the updated boundary information for the test at stage 1. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.4.8 displays design specifications. With the specified BOUNDARYKEY=BOTH option, the information levels and boundary values at future stages are modified to maintain both the α and β levels.

Output 81.4.8 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_PROP
Data Set	WORK.DATA_PROP1
Statistic Distribution	Normal
Boundary Scale	MLE
Alternative Hypothesis	Upper
Early Stop	Reject Null
Number of Stages	4
Alpha	0.05
Beta	0.2
Power	0.8
Max Information (Percent of Fixed Sample)	108.4795
Max Information	670.680662
Null Ref ASN (Percent of Fixed Sample)	106.9693
Alt Ref ASN (Percent of Fixed Sample)	78.44835

The "Test Information" table in Output 81.4.9 displays the boundary values for the test statistic with the specified MLE scale.

Output 81.4.9 Sequential Tests

Test Information (MLE Scale)
Null Reference = 0

	Information Level			Alternative Reference	Boundary Values Upper
Stage	Proportion	Actual	N	Upper	Alpha
1	0.2556	171.4286	35.98376	0.10000	0.19638
2	0.5037	337.8459	70.91565	0.10000	0.11843
3	0.7519	504.2633	105.8475	0.10000	0.08770
4	1.0000	670.6807	140.7794	0.10000	0.07080

Test Information (MLE Scale)
Null Reference = 0

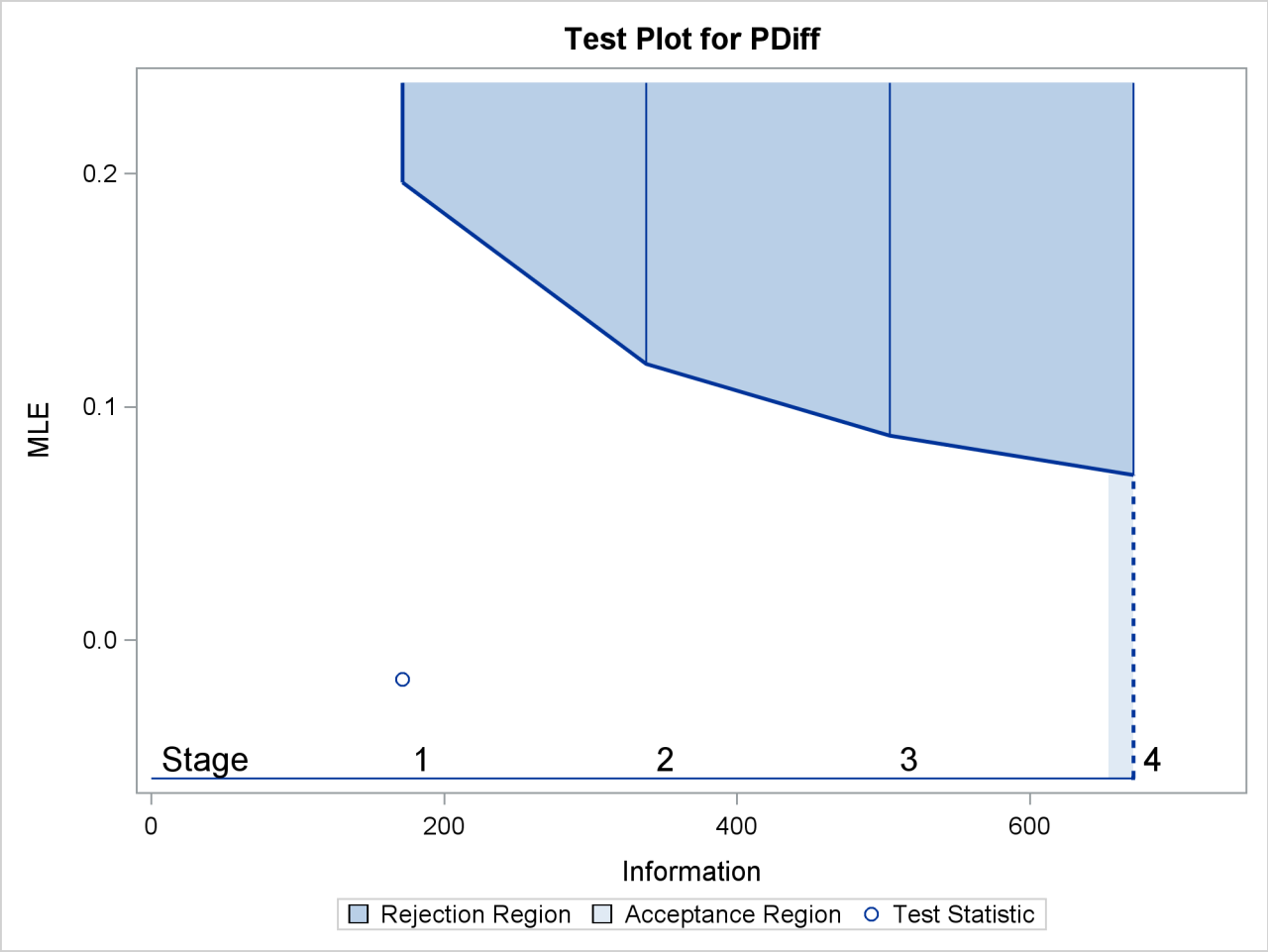
	Test PDiff	
Stage	Estimate	Action
1	-0.01667	Continue
2	.	
3	.	
4	.	

The information level at stage 1 is computed as $I_1^* = I_1 \times (n_1^*/n_1)$, where I_1 and n_1 are the information level and sample size at stage 1 in the BOUNDARY= data set, and $n_1^* = 36$ is the available sample size at stage 1.

With the INFOADJ=PROP option (which is the default), the information levels at interim stages 2 and 3 are derived proportionally from the information levels in the BOUNDARY= data set. At stage 1, the statistic $\hat{\theta} = \hat{p} - p_0 = 0.58333 - 0.6 = -0.01667$ is less than the upper α boundary value 0.19638, so the trial continues to the next stage.

With ODS Graphics enabled, a boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.4.10. As expected, the test statistic is in the continuation region.

Output 81.4.10 Sequential Test Plot



The following statements use the MEANS procedure to compute the mean response at stage 2:

```
proc means data=Prop_2;
   var Resp;
   ods output Summary=Data_Prop2;
run;
```

The following statements create and display (in Output 81.4.11) the data set for the centered mean positive response ($\hat{p} - p_0$) at stage 2:

```
data Data_Prop2;
   set Data_Prop2;
   _Scale_='MLE';
   _Stage_= 2;
   NObs= Resp_N;
   PDiff= Resp_Mean - 0.6;
   keep _Scale_ _Stage_ NObs PDiff;
run;

proc print data=Data_Prop2;
   title 'Statistics Computed at Stage 2';
run;
```

Output 81.4.11 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	_Scale_	_Stage_	NObs	PDiff
1	MLE	2	71	-0.064789

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
ods graphics on;
proc seqtest Boundary=Test_Prop1
             Data(Testvar=PDiff)=Data_Prop2
             infoadj=prop
             boundarykey=both
             boundaryscale=mle
             condpower(cref=1)
             predpower
             plots=condpower
             ;
ods output test=Test_Prop2;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The DATA= option specifies the input data set that contains the test statistic and its associated sample size at stage 2, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=TEST_PROP2 option creates an output data set named TEST_PROP2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The CONDPOWER(CREF=1) option requests the conditional power with the observed statistic under the alternative hypothesis, in addition to the conditional power under the hypothetical reference $\theta = \hat{\theta}$, the MLE estimate. The PREDPOWER option requests the noninformative predictive power with the observed statistic.

The "Test Information" table in Output 81.4.12 displays the boundary values for the test statistic with the specified MLE scale. The test statistic $\hat{\theta} = -0.06479$ is less than the corresponding upper α boundary 0.11831, so the sequential test does not stop at stage 2 to reject the null hypothesis.

Output 81.4.12 Sequential Tests

The SEQTEST Procedure

Test Information (MLE Scale)
Null Reference = 0

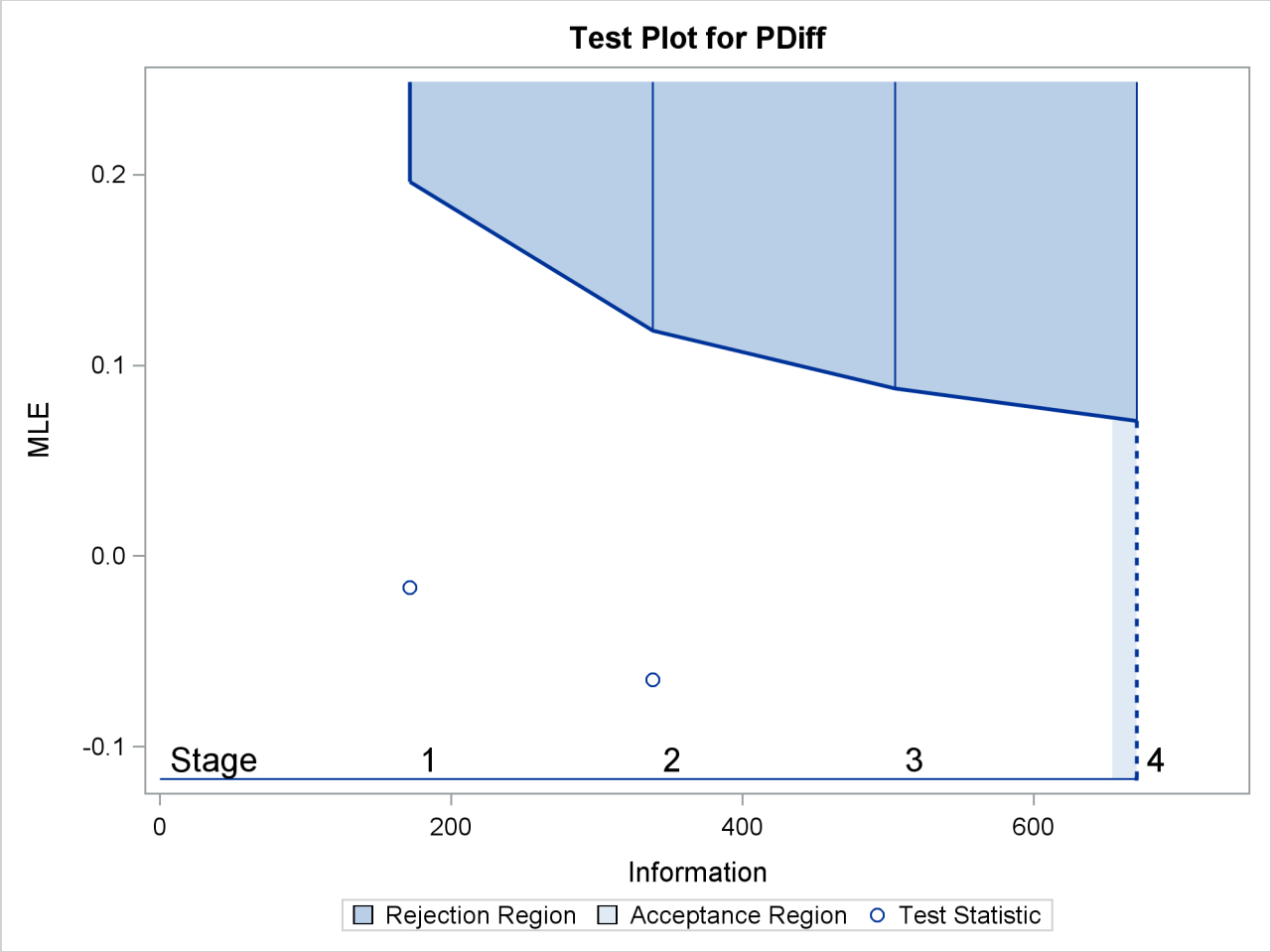
	Information Level			Alternative Reference	Boundary Values Upper
Stage	Proportion	Actual	N	Upper	Alpha
1	0.2556	171.4286	35.98223	0.10000	0.19638
2	0.5043	338.2478	70.99698	0.10000	0.11831
3	0.7522	504.4785	105.8882	0.10000	0.08767
4	1.0000	670.7092	140.7794	0.10000	0.07081

Test Information (MLE Scale)
Null Reference = 0

	Test PDiff	
Stage	Estimate	Action
1	-0.01667	Continue
2	-0.06479	Continue
3	.	
4	.	

With ODS Graphics enabled, the "Test Plot" displays boundary values of the design and the test statistic, as shown in Output 81.4.13. It also shows that the test statistic is in the "Continuation Region" below the upper α boundary value at stage 2.

Output 81.4.13 Sequential Test Plot



The predictive power is the probability to reject the null hypothesis under the posterior distribution with a noninformative prior given the observed statistic $\hat{\theta} = -0.06479$. The "Predictive Power Information" table in Output 81.4.14 indicates that the predictive power at $\hat{\theta} = -0.06479$ is 0.0002.

Output 81.4.14 Predictive Power

Predictive Power Information

Stopping Stage	MLE	Predictive Power
2	-0.06479	0.00020

The "Conditional Power Information" table in Output 81.4.15 displays conditional powers given the observed statistic under hypothetical references $\theta = \hat{\theta}$, the maximum likelihood estimate, and $\theta = \theta_1$. The constant c under CRef for the MLE is derived from $\hat{\theta} = c\theta_1$; that is, $c = \hat{\theta}/\theta_1 = -0.06479/0.1 = -0.6479$.

Output 81.4.15 Conditional Power

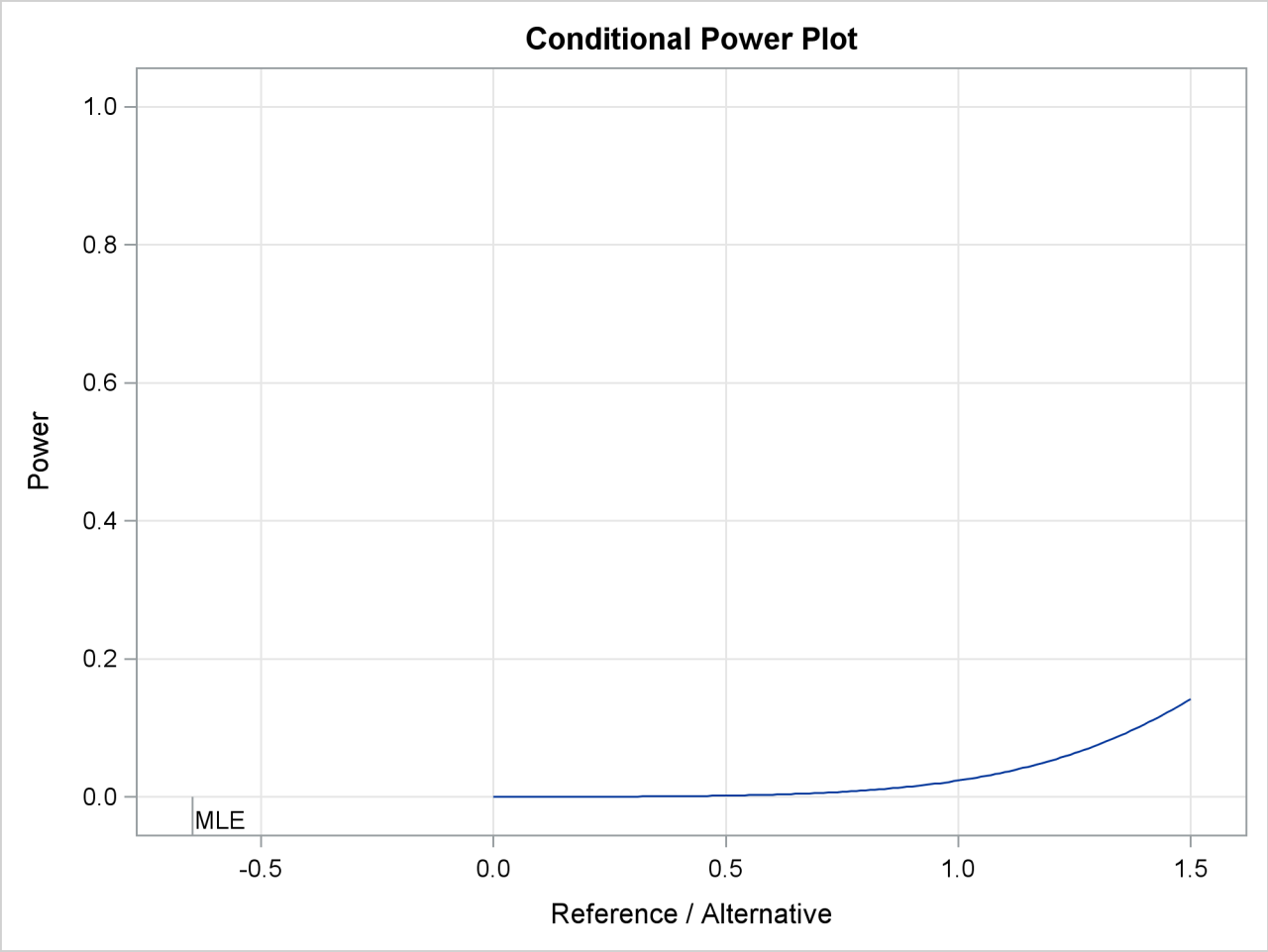
Conditional Power Information
Reference = CRef * (Alt Reference)

Stopping Stage	MLE	Reference: Ref	Reference: CRef	Conditional Power
2	−0.06479	MLE	−0.6479	0.00000
2	−0.06479	Alternative	1.0000	0.02368

The conditional power is the probability of rejecting the null hypothesis under these hypothetical references given the observed statistic $\hat{\theta} = -0.06479$. The table in Output 81.4.15 shows a weak conditional power of 0.02368 under the alternative hypothesis.

The "Conditional Power Plot" displays conditional powers given the observed statistic under various hypothetical references, as shown in Output 81.4.16. These references include $\theta = \hat{\theta}$, the maximum likelihood estimate, and $\theta = c_i\theta_1$, where θ_1 is the alternative reference and $c_i = 0, 0.01, \ldots, 1.50$ are constants that are specified in the CREF= option. Output 81.4.16 shows that the conditional power increases as c_i increases.

Output 81.4.16 Conditional Power Plot



With a predictive power 0.0002 and a conditional power of 0.02368 under H_1, the supermarket decides to stop the trial and accept the null hypothesis. That is, the positive response for the new store-brand coffee is not better than that for the current store-brand coffee.

Output 81.4.17 Predictive Power

Predictive Power Information

Stopping Stage	MLE	Predictive Power
2	-0.06479	0.00020

In the SEQTEST procedure, the conditional probability at an interim stage k is the probability that the test statistic at the final stage (stage 4) would exceed the rejection critical value. Since an interim stage exists between the current stage (stage 2) and the final stage, the conditional power is not the conditional probability to reject the null hypothesis H_0.

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2. The NSTAGES=3 option sets the next stage as the final stage (stage 3), and the BOUNDARYKEY=BOTH option derives the information level at stage 3 that maintain both Type I and Type II error probability levels. The CONDPOWER(CREF=1) option requests the conditional power with the observed statistic under the alternative hypothesis, in addition to the conditional power under the hypothetical reference $\theta = \hat{\theta}$, the MLE estimate.

```
proc seqtest Boundary=Test_Prop1
             Data(Testvar=PDiff)=Data_Prop2
             nstages=3
             boundarykey=both
             boundaryscale=mle
             condpower(cref=1)
             ;
run;
```

The "Test Information" table in Output 81.4.18 displays the boundary values for the test statistic with the specified MLE scale, assuming that the next stage is the final stage.

Output 81.4.18 Sequential Tests

The SEQTEST Procedure

Test Information (MLE Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Information Level N	Alternative Reference Upper	Boundary Values Upper Alpha
1	0.2645	171.4286	37.23405	0.10000	0.19638
2	0.5219	338.2478	73.46696	0.10000	0.11831
3	1.0000	648.1598	140.7794	0.10000	0.06831

Test Information (MLE Scale)
Null Reference = 0

Test PDiff

Stage	Estimate	Action
1	-0.01667	Continue
2	-0.06479	Continue
3	.	

The "Conditional Power Information" table in Output 81.4.19 displays conditional powers given the observed statistic, assuming that the next stage is the final stage.

Output 81.4.19 Conditional Power

Conditional Power Information
Reference = CRef * (Alt Reference)

Stopping Stage	MLE	Reference Ref	Reference CRef	Conditional Power
2	-0.06479	MLE	-0.6479	0.00000
2	-0.06479	Alternative	1.0000	0.02278

The conditional power is the probability of rejecting the null hypothesis under these hypothetical references given the observed statistic $\hat{\theta} = -0.06479$. The table in Output 81.4.19 also shows a weak conditional power of 0.02278 under the alternative hypothesis.

Example 81.5: Comparing Two Proportions with a Log Odds Ratio Test

This example compares two binomial proportions by using a log odds ratio statistic in a five-stage group sequential test. A clinic is studying the effect of vitamin C supplements in treating flu symptoms. The study consists of patients in the clinic who exhibit the first sign of flu symptoms within the last 24 hours. These patients are randomly assigned to either the control group (which receives 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 patient are recorded.

Suppose that you know from past experience that flu symptoms disappear in five days for 60% of patients who experience flu symptoms. The clinic would like to detect a 70% symptom disappearance with a high probability. A test that compares the proportions directly specifies the null hypothesis $H_0 : \theta = p_t - p_c = 0$ with a one-sided alternative $H_1 : \theta > 0$ and a power of 0.90 at $H_1 : \theta = 0.10$, where p_t and p_c are the proportions of symptom disappearance in the treatment group and control group, respectively. An alternative trial tests an equivalent hypothesis by using the log odds ratio statistics:

$$\theta = \log\left(\frac{(\frac{p_t}{1-p_t})}{(\frac{p_c}{1-p_c})}\right)$$

Then the null hypothesis is $H_0 : \theta = \theta_0 = 0$ and the alternative hypothesis is

$$H_1 : \theta = \theta_1 = \log\left(\frac{(\frac{0.70}{0.30})}{(\frac{0.6}{0.4})}\right) = 0.441833$$

The following statements invoke the SEQDESIGN procedure and request a five-stage group sequential design by using an error spending function method for normally distributed statistics. The design uses a two-sided alternative hypothesis with early stopping to reject the null hypothesis H_0.

```
ods graphics on;
proc seqdesign altref=0.441833
               boundaryscale=mle
               ;
   OneSidedErrorSpending: design method=errfuncpow
                                nstages=5
                                alt=upper
                                stop=accept
                                alpha=0.025;
   samplesize model=twosamplefreq( nullprop=0.6 test=logor);
ods output Boundary=Bnd_CSup;
run;
ods graphics off;
```

The ODS OUTPUT statement with the BOUNDARY=BND_CSUP option creates an output data set named BND_CSUP which contains the resulting boundary information for the subsequent sequential tests.

The "Design Information" table in Output 81.5.1 displays design specifications and derived statistics. With the specified alternative reference, the maximum information 56.30934 is derived.

Output 81.5.1 Design Information

The SEQDESIGN Procedure
Design: OneSidedErrorSpending

Design Information

Statistic Distribution	Normal
Boundary Scale	MLE
Alternative Hypothesis	Upper
Early Stop	Accept Null
Method	Error Spending
Boundary Key	Both
Alternative Reference	0.441833
Number of Stages	5
Alpha	0.025
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	104.6166
Max Information	56.30934
Null Ref ASN (Percent of Fixed Sample)	57.21399
Alt Ref ASN (Percent of Fixed Sample)	102.1058

The "Boundary Information" table in Output 81.5.2 displays information level, alternative reference, and boundary values at each stage. With the specified BOUNDARYSCALE=MLE option, the procedure displays the output boundaries in terms of the MLE scale.

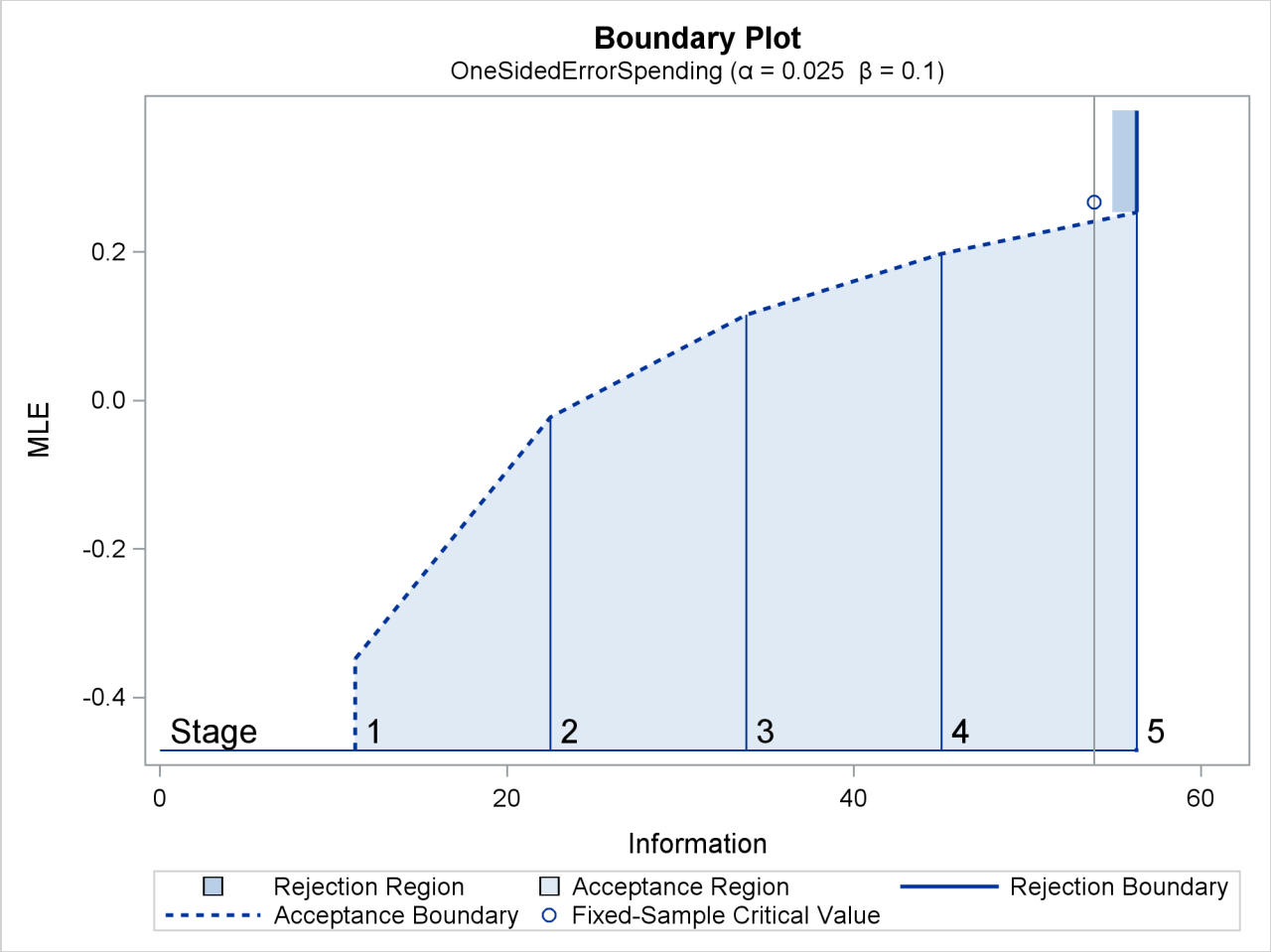
Output 81.5.2 Boundary Information

Boundary Information (MLE Scale)
Null Reference = 0

	Information Level			Alternative Reference	Boundary Values
Stage	Proportion	Actual	N	Upper	Upper Beta
1	0.2000	11.26187	201.1048	0.44183	-0.34844
2	0.4000	22.52374	402.2096	0.44183	-0.02262
3	0.6000	33.7856	603.3144	0.44183	0.11527
4	0.8000	45.04747	804.4192	0.44183	0.19708
5	1.0000	56.30934	1005.524	0.44183	0.25345

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.5.3.

Output 81.5.3 Boundary Plot



With the SAMPLESIZE statement, the "Sample Size Summary" table in Output 81.5.4 displays the parameters for the sample size computation.

Output 81.5.4 Sample Size Summary

Sample Size Summary

Test	Two-Sample Proportions
Null Proportion	0.6
Proportion (Group A)	0.7
Test Statistic	Log Odds Ratio
Reference Proportions	Alt Ref
Max Sample Size	1005.524
Expected Sample Size (Null Ref)	549.9132
Expected Sample Size (Alt Ref)	981.3914

The "Sample Sizes" table in Output 81.5.5 displays the required sample sizes for the group sequential clinical trial.

Output 81.5.5 Required Sample Sizes

Sample Sizes (N)
Two-Sample Log Odds Ratio Test for Proportion Difference

	Fractional N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	201.10	100.55	100.55	11.2619
2	402.21	201.10	201.10	22.5237
3	603.31	301.66	301.66	33.7856
4	804.42	402.21	402.21	45.0475
5	1005.52	502.76	502.76	56.3093

Sample Sizes (N)
Two-Sample Log Odds Ratio Test for Proportion Difference

	Ceiling N			
Stage	N	N(Grp 1)	N(Grp 2)	Information
1	202	101	101	11.3120
2	404	202	202	22.6240
3	604	302	302	33.8240
4	806	403	403	45.1360
5	1006	503	503	56.3360

Thus, 101 new patients are needed in each group at stages 1, 2, and 4, and 100 new patients are needed in each group at stages 3 and 5. Suppose that 101 patients are available in each group at stage 1. Output 81.5.6 lists the 10 observations in the data set count_1.

Output 81.5.6 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	TrtGrp	Resp
1	Control	1
2	C_Sup	0
3	Control	0
4	C_Sup	1
5	Control	1
6	C_Sup	1
7	Control	1
8	C_Sup	0
9	Control	0
10	C_Sup	1

The TrtGrp variable is a grouping variable with the value Control for a patient in the placebo control group and the value C_Sup for a patient in the treatment group who receives vitamin C supplements. The Resp variable is an indicator variable with the value 1 for a patient without flu symptoms after five days and the value 0 for a patient with flu symptoms after five days.

The following statements use the LOGISTIC procedure to compute the log odds ratio statistic and its associated standard error at stage 1:

```
proc logistic data=CSup_1 descending;
   class TrtGrp / param=ref;
   model Resp= TrtGrp;
   ods output ParameterEstimates=Parms_CSup1;
run;
```

The DESCENDING option is used to reverse the order for the response levels, so the LOGISTIC procedure is modeling the probability that Resp = 1.

The following statements create and display (in Output 81.5.7) the data set for the log odds ratio statistic and its associated standard error:

```
data Parms_CSup1;
   set Parms_CSup1;
   if Variable='TrtGrp' and ClassVal0='C_Sup';
   _Scale_='MLE';
   _Stage_= 1;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_CSup1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.5.7 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	TrtGrp	0.3247	0.2856	MLE	1

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_CSup
             Parms(Testvar=TrtGrp)=Parms_CSup1
             infoadj=prop
             errspendadj=errfuncpow
             boundarykey=both
             boundaryscale=mle
             ;
ods output test=Test_CSup1; run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The PARMS=PARMS_CSUP1 option specifies the input data set PARMS_CSUP1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=TRTGRP option identifies the test variable TRTGRP in the data set.

If the computed information level for stage 1 is not the same as the value provided in the BOUNDARY= data set, the INFOADJ=PROP option (which is the default) proportionally adjusts the information levels at future interim stages from the levels provided in the BOUNDARY= data set. The ERRSPENDADJ=ERRFUNCPOW option adjusts the boundaries with the updated error spending values generated from the power error spending function. The BOUNDARYKEY=BOTH option maintains both the α and β levels. The BOUNDARYSCALE=MLE option displays the output boundaries in terms of the MLE scale.

The ODS OUTPUT statement with the TEST=TEST_CSUP1 option creates an output data set named TEST_CSUP1 which contains the updated boundary information for the test at stage 1. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.5.8 displays design specifications. With the specified BOUNDARYKEY=BOTH option, the information levels and boundary values at future stages are modified to maintain both the α and β levels.

Output 81.5.8 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_CSUP
Data Set	WORK.PARMS_CSUP1
Statistic Distribution	Normal
Boundary Scale	MLE
Alternative Hypothesis	Upper
Early Stop	Accept Null
Number of Stages	5
Alpha	0.025
Beta	0.1
Power	0.9
Max Information (Percent of Fixed Sample)	104.6673
Max Information	56.3361718
Null Ref ASN (Percent of Fixed Sample)	57.02894
Alt Ref ASN (Percent of Fixed Sample)	102.1369

The "Test Information" table in Output 81.5.9 displays the boundary values for the test statistic with the specified MLE scale. With the INFOADJ=PROP option (which is the default), the information levels at future interim stages are derived proportionally from the observed information at stage 1 and the information levels in the BOUNDARY= data set.

Since the information level at stage 1 is derived from the PARMS= data set and other information levels are not specified, equal increments are used at remaining stages. At stage 1, the MLE statistic 0.32474 is greater than the corresponding upper β boundary value -0.29906, so the sequential test continues to the next stage.

Output 81.5.9 Sequential Tests

Test Information (MLE Scale)
Null Reference = 0

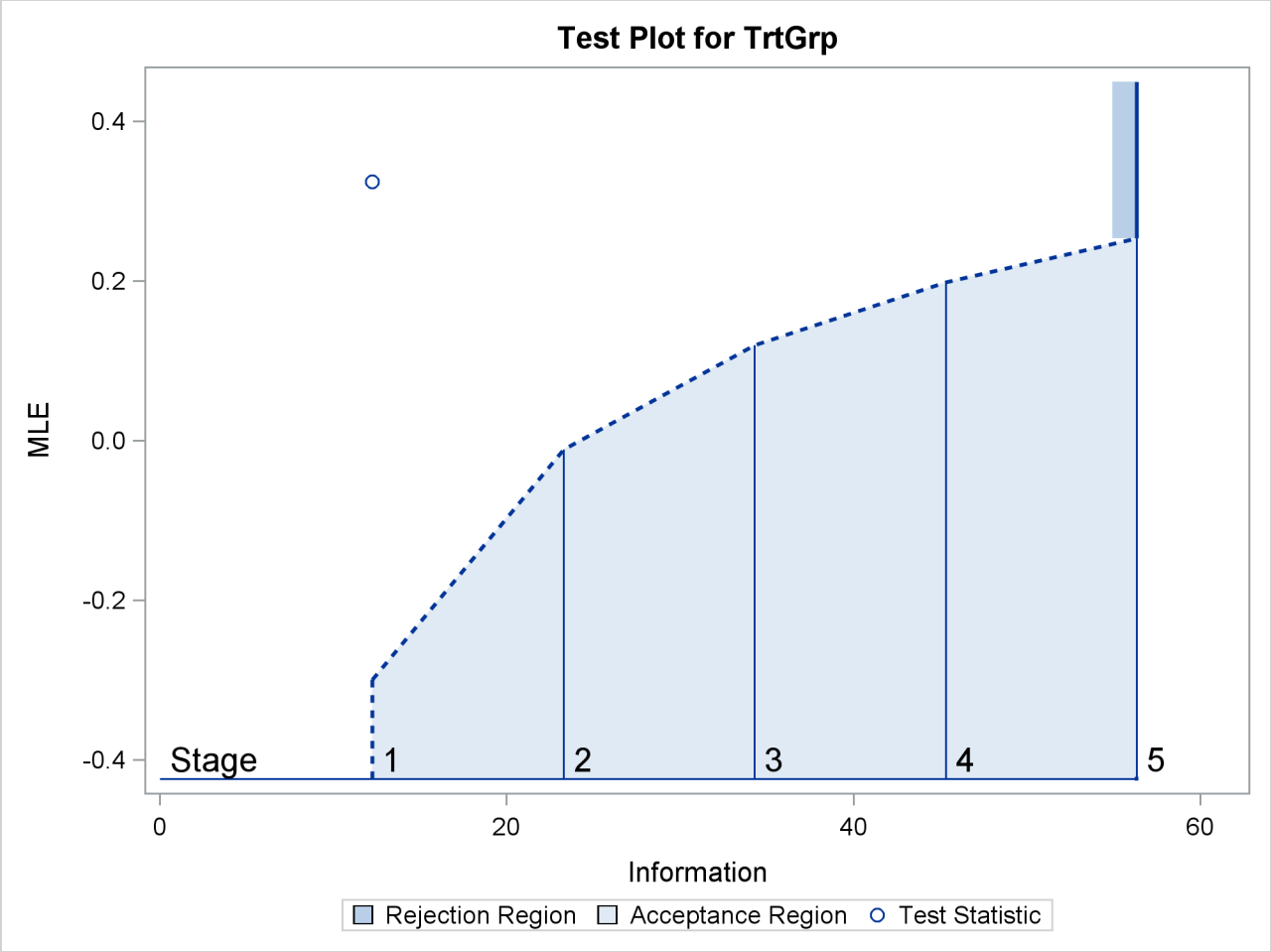
Stage	---Information Level-- Proportion	Actual	-Alternative- --Reference-- Upper
1	0.2176	12.26014	0.44183
2	0.4132	23.27914	0.44183
3	0.6088	34.29815	0.44183
4	0.8044	45.31716	0.44183
5	1.0000	56.33617	0.44183

Test Information (MLE Scale)
Null Reference = 0

Stage	-Boundary Values- ------Upper------ Beta	----------Test---------- ----------TrtGrp--------- Estimate	Action
1	-0.29906	0.32474	Continue
2	-0.01067	.	
3	0.11942	.	
4	0.19829	.	
5	0.25325	.	

With ODS Graphics enabled, a boundary plot with the boundary values and test statistics is displayed, as shown in Output 81.5.10. As expected, the test statistic is in the continuation region.

Output 81.5.10 Sequential Test Plot



The following statements use the LOGISTIC procedure to compute the log odds ratio statistic and its associated standard error at stage 2:

```
proc logistic data=CSup_2 descending;
   class TrtGrp / param=ref;
   model Resp= TrtGrp;
   ods output ParameterEstimates=Parms_CSup2;
run;
```

The following statements create and display (in Output 81.5.11) the data set for the mean positive response and its associated standard error at stage 2:

```
data Parms_CSup2;
   set Parms_CSup2;
   if Variable='TrtGrp' and ClassVal0='C_Sup';
   _Scale_='MLE';
   _Stage_= 2;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;
proc print data=Parms_CSup2;
   title 'Statistics Computed at Stage 2';
run;
```

Output 81.5.11 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	TrtGrp	0.2356	0.2073	MLE	2

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
proc seqtest Boundary=Test_CSup1
             Parms( testvar=TrtGrp)=Parms_CSup2
             infoadj=prop
             errspendadj=errfuncpow
             boundarykey=both
             boundaryscale=mle
             ;
ods output Test=Test_CSup2;
run;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=CSUP_LDL2 option creates an output data set named CSUP_LDL2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Test Information" table in Output 81.5.12 displays the boundary values for the test statistic with the specified MLE scale. The test statistic 0.2356 is greater than the corresponding upper β boundary value -0.01068, so the sequential test continues to the next stage.

Output 81.5.12 Sequential Tests

The SEQTEST Procedure

Test Information (MLE Scale)
Null Reference = 0

Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper
1	0.2176	12.26014	0.44183
2	0.4132	23.27916	0.44183
3	0.6088	34.29799	0.44183
4	0.8044	45.31681	0.44183
5	1.0000	56.33563	0.44183

Test Information (MLE Scale)
Null Reference = 0

Stage	Boundary Values Upper Beta	Test TrtGrp Estimate	Test TrtGrp Action
1	-0.29906	0.32474	Continue
2	-0.01068	0.23560	Continue
3	0.11942	.	
4	0.19829	.	
5	0.25325	.	

Similar results are found at stages 3 and stage 4, so the trial continues to the final stage. The following statements use the LOGISTIC procedure to compute the log odds ratio statistic and its associated standard error at stage 5:

```
proc logistic data=CSup_5 descending;
   class TrtGrp / param=ref;
   model Resp= TrtGrp;
   ods output ParameterEstimates=Parms_CSup5;
run;
```

The following statements create and display (in Output 81.5.13) the data set for the log odds ratio statistic and its associated standard error at stage 5:

```
data Parms_CSup5;
   set Parms_CSup5;
   if Variable='TrtGrp' and ClassVal0='C_Sup';
   _Scale_='MLE';
   _Stage_= 5;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_CSup5;
   title 'Statistics Computed at Stage 5';
run;
```

Output 81.5.13 Statistics Computed at Stage 5

Statistics Computed at Stage 5

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	TrtGrp	0.2043	0.1334	MLE	5

The following statements invoke the SEQTEST procedure to test for the hypothesis at stage 5:

```
ods graphics on;
proc seqtest Boundary=Test_CSup4
             Parms( testvar=TrtGrp)=Parms_CSup5
             errspendadj=errfuncpow
             boundaryscale=mle
             cialpha=.025
             rci
             plots=rci
             ;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 5, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 5, and the TESTVAR= option identifies the test variable in the data set. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary value at stage 5 is derived to maintain the α level.

The "Test Information" table in Output 81.5.14 displays the boundary values for the test statistic with the specified MLE scale. The test statistic 0.2043 is less than the corresponding upper β boundary 0.25375, so the sequential test stops to accept the null hypothesis. That is, there is no reduction in duration of symptoms for the group receiving vitamin C supplements.

Output 81.5.14 Sequential Tests

The SEQTEST Procedure

Test Information (MLE Scale)
Null Reference = 0

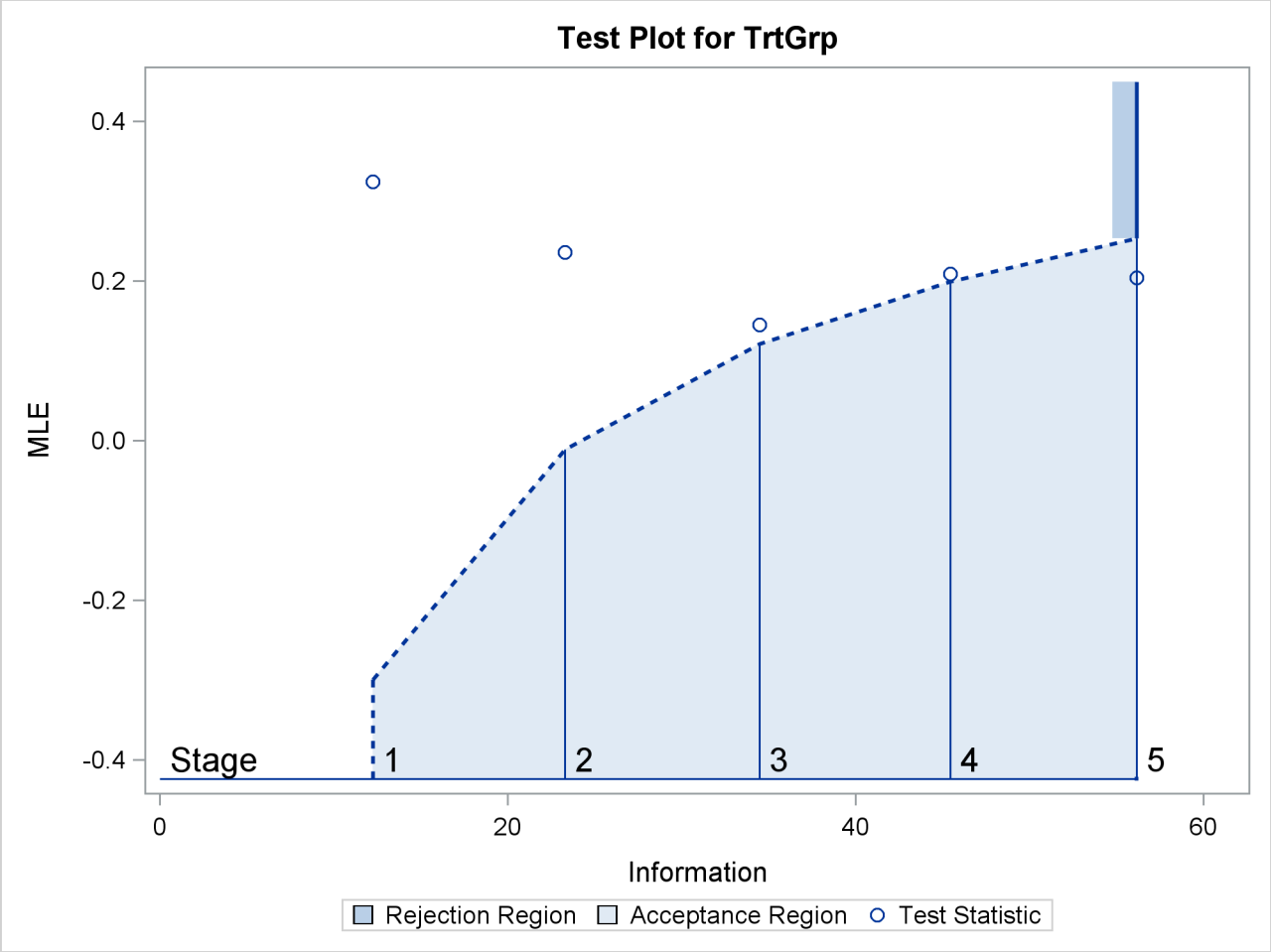
Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper
1	0.2183	12.26014	0.44183
2	0.4145	23.27916	0.44183
3	0.6141	34.48793	0.44183
4	0.8092	45.44685	0.44183
5	1.0000	56.16068	0.44183

Test Information (MLE Scale)
Null Reference = 0

Stage	Boundary Values Upper Beta	Test TrtGrp Estimate	Test TrtGrp Action
1	-0.29906	0.32474	Continue
2	-0.01068	0.23560	Continue
3	0.12134	0.14482	Continue
4	0.19899	0.20855	Continue
5	0.25375	0.20430	Accept Null

The "Test Plot" displays boundary values of the design and the test statistics, as shown in Output 81.5.15. It also shows that the test statistic is in the "Acceptance Region" at the final stage.

Output 81.5.15 Sequential Test Plot



After a trial is stopped, the "Parameter Estimates" table in Output 81.5.16 displays the stopping stage, parameter estimate, unbiased median estimate, confidence limits, and the p-value under the null hypothesis $H_0 : \theta = 0$. As expected, the p-value 0.0456 is not significant at $\alpha = 0.025$ level and the lower 97.5% confidence limit is less than the value $\theta_0 = 0$. The p-value, unbiased median estimate, and confidence limits depend on the ordering of the sample space (k, z), where k is the stage number and z is the standardized Z statistic.

Output 81.5.16 Parameter Estimates

Parameter Estimates
Stagewise Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate	Lower 97.5% CL
TrtGrp	5	0.204303	0.0456	0.234494	-0.03712

Since the test is accepted at stage 5, the p-value computed by using the default stagewise ordering can be expressed as

$$\alpha_u = P_{\theta=0}\left(z_5 < Z_5 \mid b_k < Z_k, k < 5\right)$$

where $z_5 = 1.53105$ is the test statistic at stage 5, Z_k is a standardized normal variate at stage k, and b_k is the upper β boundary value in the standardized Z scale at stage $k, k = 1, 2, \ldots, 5$.

With the RCI option, the "Repeated Confidence Intervals" table in Output 81.5.17 displays repeated confidence intervals for the parameter. For a one-sided test with an upper alternative hypothesis, since the upper acceptance repeated confidence limit 0.3924 at the final stage is less than the alternative reference 0.441833, the null hypothesis is accepted.

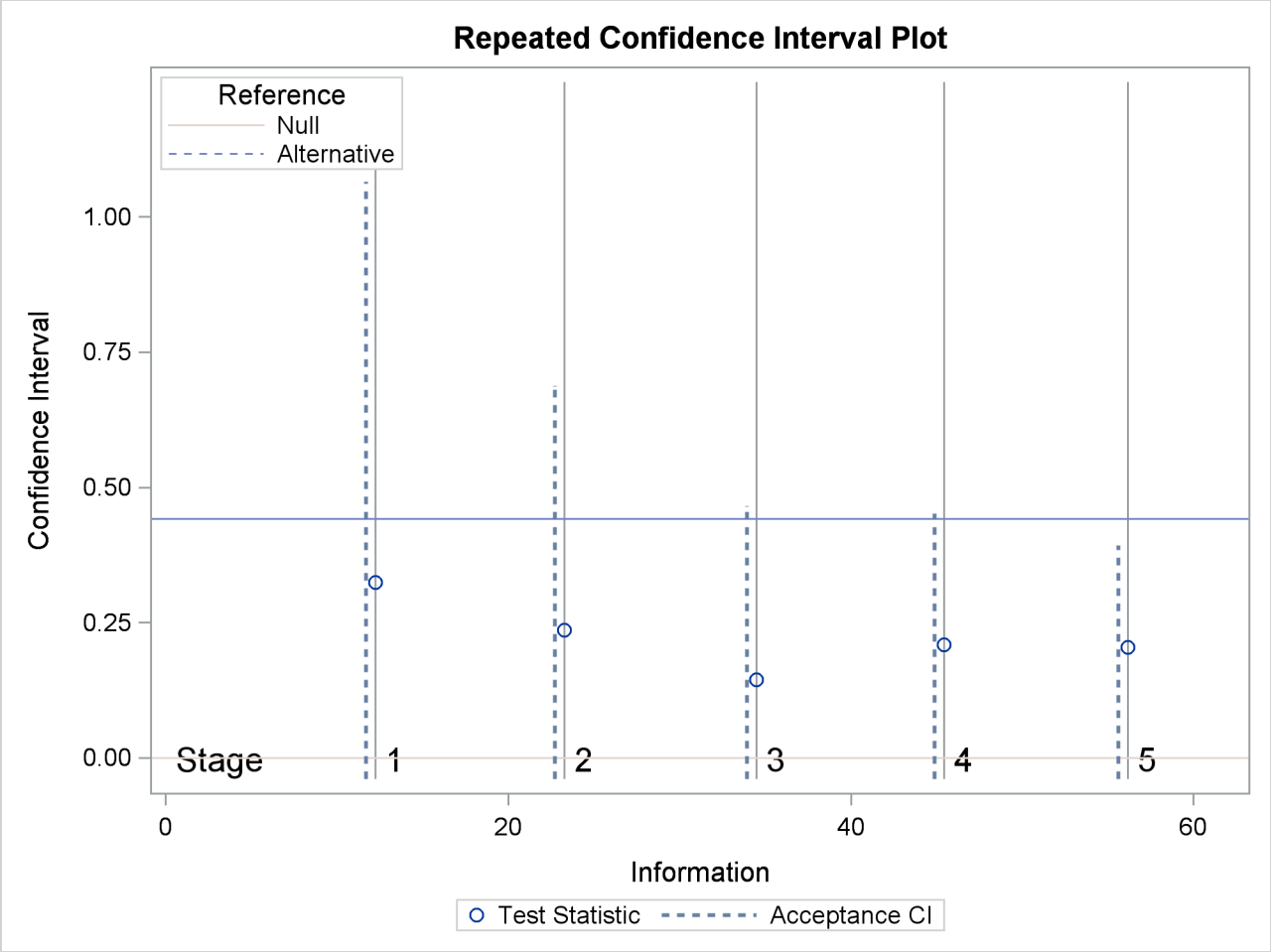
Output 81.5.17 Repeated Confidence Intervals

Repeated Confidence Intervals

Stage	Information Level	Parameter Estimate	-Acceptance Boundary- Upper 89.94% CL
1	12.2601	0.32474	1.0656
2	23.2792	0.23560	0.6881
3	34.4879	0.14482	0.4653
4	45.4468	0.20855	0.4514
5	56.1607	0.20430	0.3924

With the PLOTS=RCI option, the "Repeated Confidence Intervals Plot" displays repeated confidence intervals for the parameter, as shown in Output 81.5.18. It shows that the upper acceptance repeated confidence limit at the final stage is less than the alternative reference 0.441833. This implies that the study accepts the null hypothesis at the final stage.

Output 81.5.18 Repeated Confidence Intervals Plot



Example 81.6: Comparing Two Survival Distributions with a Log-Rank Test

This example requests a log-rank test that compares two survival distributions for the treatment effect (Jennison and Turnbull 2000, pp. 77–79; Whitehead 1997, pp. 36–39).

A clinic is studying the effect of a new cancer treatment. The study consists of mice exposed to a carcinogen and randomized to either the control group or the treatment group. The event of interest is the 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" in the chapter "The SEQDESIGN Procedure," the hypothesis $H_0 : \theta = -\log(\lambda) = 0$ with an alternative hypothesis $H_1 : \theta = \theta_1 > 0$ can be used, where λ is the hazard ratio between the treatment group and control group.

Suppose that from past experience, the median survival time for the control group is $t_0 = 20$ weeks, and the study would like to detect a $t_1 = 40$ weeks 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 $h_0 = 0.03466$ and $h_1 = 0.01733$, the hazard ratio $\lambda_1 = h_1/h_0 = 1/2$ and the alternative hypothesis is

$$\theta_1 = -\log(\lambda_1) = -\log(\frac{1}{2}) = 0.69315$$

The following statements invoke the SEQDESIGN procedure and request a four-stage group sequential design for normally distributed data. The design uses a one-sided alternative hypothesis with early stopping to reject and to accept the null hypothesis H_0. Whitehead's triangular method is used to derive the boundaries.

```
ods graphics on;
proc seqdesign altref=0.69315
               boundaryscale=score
               ;
   OneSidedWhitehead: design method=whitehead
                             nstages=4
                             boundarykey=alpha
                             alt=upper  stop=both
                             beta=0.20;
   samplesize model=twosamplesurvival
                    ( nullhazard=0.03466
                      accrate=10);
run;
ods graphics off;
```

A Whitehead method creates boundaries that approximately satisfy the Type I and Type II error probability level specification. The BOUNDARYKEY=ALPHA option is used to adjust the boundary value at the last stage and to meet the specified Type I probability level.

The specified ACCRATE=10 option indicates that 10 mice will be accrued each week and the resulting minimum and maximum accrual times are displayed. With the BOUNDARYSCALE=SCORE option, the procedure displays the output boundaries with the score statistics.

The "Design Information" table in Output 81.6.1 displays design specifications and derived statistics.

Output 81.6.1 Design Information

The SEQDESIGN Procedure
Design: OneSidedWhitehead

Design Information

Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Method	Whitehead
Boundary Key	Alpha
Alternative Reference	0.69315
Number of Stages	4
Alpha	0.05
Beta	0.20044
Power	0.79956
Max Information (Percent of Fixed Sample)	129.9894
Max Information	16.70624
Null Ref ASN (Percent of Fixed Sample)	62.6302
Alt Ref ASN (Percent of Fixed Sample)	74.00064

The "Boundary Information" table in Output 81.6.2 displays the information level, alternative reference, and boundary values at each stage.

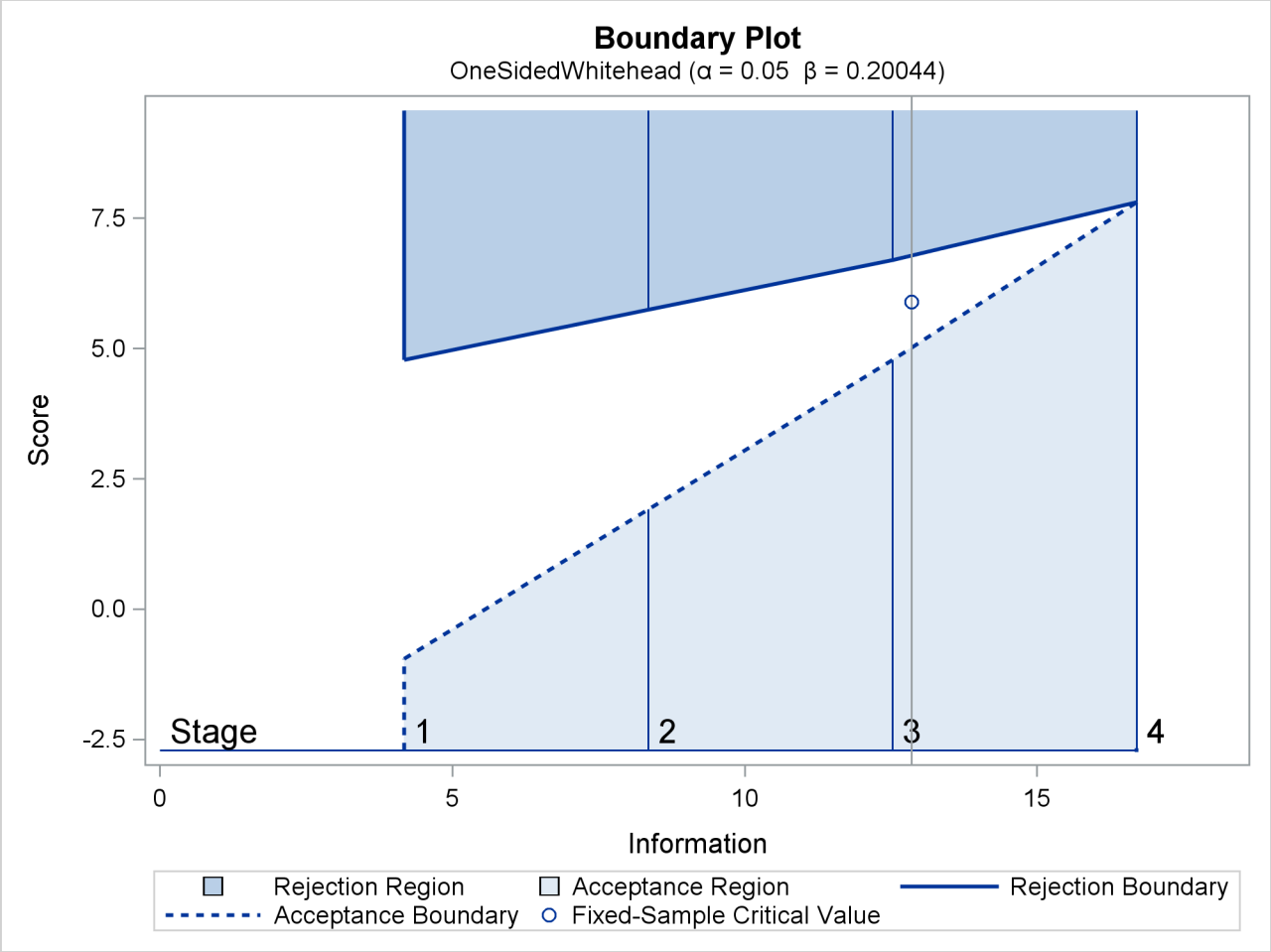
Output 81.6.2 Boundary Information

Boundary Information (Score Scale)
Null Reference = 0

	Information Level			Alternative Reference	Boundary Values	
					Upper	
Stage	Proportion	Actual	Events	Upper	Beta	Alpha
1	0.2500	4.176561	16.70624	2.89498	-0.95755	4.78773
2	0.5000	8.353122	33.41249	5.78997	1.91509	5.74527
3	0.7500	12.52968	50.11873	8.68495	4.78773	6.70282
4	1.0000	16.70624	66.82498	11.57993	7.81296	7.81296

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.6.3.

Output 81.6.3 Boundary Plot



With the MODEL=TWOSAMPLESURVIVAL option in the SAMPLESIZE statement, the "Sample Size Summary" table in Output 81.6.4 displays the parameters for the sample size computation.

Output 81.6.4 Required Sample Size Summary

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	0.499999
log(Hazard Ratio)	-0.69315
Reference Hazards	Alt Ref
Accrual Rate	10
Min Accrual Time	6.682498
Min Sample Size	66.82498
Max Accrual Time	25.401
Max Sample Size	254.01
Max Number of Events	66.82498

With a minimum accrual time of 6.6825 weeks and a maximum accrual time of 25.401 weeks, an accrual time of 20 weeks is used in the study.

The "Numbers of Events" table in Output 81.6.5 displays the required number of events for the group sequential clinical trial.

Output 81.6.5 Required Numbers of Events

Numbers of Events (D)
Two-Sample Log-Rank Test

Stage	D	Information
1	16.71	4.1766
2	33.41	8.3531
3	50.12	12.5297
4	66.82	16.7062

The following statements invoke the SEQDESIGN procedure and provide more detailed sample size information:

```
proc seqdesign altref=0.69315
               boundaryscale=score
               ;
   OneSidedWhitehead: design method=whitehead
                             nstages=4
                             boundarykey=alpha
                             alt=upper
                             stop=both
                             beta=0.20;
   samplesize model=twosamplesurvival
                     ( nullhazard=0.03466
                       accrate=10 acctime=20);
ods output Boundary=Bnd_Surv;
run;
```

The ODS OUTPUT statement with the BOUNDARY=BND_SURV option creates an output data set named BND_SURV which contains the resulting boundary information for the subsequent sequential tests.

With an accrual time of 20 weeks, the "Sample Size Summary" table in Output 81.6.6 displays the follow-up time for the trial.

Output 81.6.6 Required Sample Size Summary

The SEQDESIGN Procedure
Design: OneSidedWhitehead

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	0.499999
log(Hazard Ratio)	-0.69315
Reference Hazards	Alt Ref
Accrual Rate	10
Accrual Time	20
Follow-up Time	6.47422
Total Time	26.47422
Max Number of Events	66.82498
Max Sample Size	200
Expected Sample Size (Null Ref)	161.5937
Expected Sample Size (Alt Ref)	172.4689

The "Numbers of Events and Sample Sizes" table in Output 81.6.7 displays the required sample sizes for the group sequential clinical trial.

Output 81.6.7 Numbers 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	11.9866	119.87	59.93	59.93
2	33.41	11.84	21.57	17.3584	173.58	86.79	86.79
3	50.12	18.01	32.11	21.7479	200.00	100.00	100.00
4	66.82	24.46	42.37	26.4742	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	8.3531	35.73	12.68	23.04	18	180.00	90.00
3	12.5297	51.07	18.37	32.70	22	200.00	100.00
4	16.7062	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	17.1377

Thus the study will perform three interim analyses after 12, 18, and 22 weeks and a final analysis after 27 weeks if the study does not stop at any of the interim analyses.

Note that the SEQDESIGN procedure does not compute numbers of events or sample sizes for all statistical models. If the number of events or sample size for a fixed-sample design is available, then the MODEL=INPUTNEVENTS or MODEL=INPUTNOBS option can be used to input fixed-sample information. For example, with a required fixed-sample number of events 51.4073, the following SAMPLESIZE statement can be used to produce the same sample size results:

```
samplesize model=inputnevents
                ( d=51.47 sample=two
                  hazard=0.03466 0.01733
                  accrate=10 acctime=20);
```

Suppose that 120 mice are available after week 12 for the first interim analysis. Output 81.6.8 lists the 10 observations in the data set weeks_1.

Output 81.6.8 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Trt Gp	Event	Weeks
1	0	0	11
2	1	0	11
3	0	0	11
4	1	0	11
5	0	1	6
6	1	0	11
7	0	0	11
8	1	0	11
9	0	1	9
10	1	0	11

The TrtGp variable is a grouping variable with the value 0 for a mouse in the placebo control group and the value 1 for a mouse in the treatment group. The Weeks variable is the survival time variable measured in weeks and the Event variable is the censoring variable with the value 0 indicating censoring. That is, the values of Weeks are considered censored if the corresponding values of Event are 0; otherwise, they are considered as event times.

The following statements use the LIFETEST procedure to estimate the log-rank statistic at stage 1:

```
proc lifetest data=Surv_1;
   time Weeks*Event(0);
   test TrtGp;
ods output logunichisq=Parms_Surv1;
run;
```

The following statements create and display (in Output 81.6.9) the data set for the log-rank statistic and its associated standard error:

```
data Parms_Surv1;
   set Parms_Surv1(rename=(Statistic=Estimate));
   if Variable='TrtGp';
   _Scale_='Score';
   _Stage_= 1;
   keep Variable _Scale_ _Stage_ StdErr Estimate;
run;

proc print data=Parms_Surv1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.6.9 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	TrtGp	3.2004	1.9979	Score	1

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Surv
             Parms(Testvar=TrtGp)=Parms_Surv1
             infoadj=none
             boundaryscale=score
             ;
ods output Test=Test_Surv1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The PARMS=PARMS_SURV1 option specifies the input data set PARMS_SURV1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=TRTGP option identifies the test variable TRTGP in the data set. The INFOADJ=NONE option maintains the information levels for future interim stages (2 and 3) at the values provided in the BOUNDARY= data set.

The ODS OUTPUT statement with the TEST=TEST_SURV1 option creates an output data set named TEST_SURV1 which contains the updated boundary information for the test at stage 1. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.6.10 displays design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the maximum information and the Type I error level α are preserved. Since the computed information level at stage 1 is not the same as the value provided in the BOUNDARY= data set, the power has been modified.

Output 81.6.10 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_SURV
Data Set	WORK.PARMS_SURV1
Statistic Distribution	Normal
Boundary Scale	Score
Alternative Hypothesis	Upper
Early Stop	Accept/Reject Null
Number of Stages	4
Alpha	0.05
Beta	0.20055
Power	0.79945
Max Information (Percent of Fixed Sample)	130.0335
Max Information	16.7062448
Null Ref ASN (Percent of Fixed Sample)	62.80855
Alt Ref ASN (Percent of Fixed Sample)	74.19155

The "Test Information" table in Output 81.6.11 displays the boundary values for the test statistic with the SCORE statistic scale. Since only the information level at stage 1 is specified in the DATA= data set, the information levels at subsequent stages are derived proportionally from the corresponding information

levels provided in the BOUNDARY= data set. At stage 1, the score statistic 3.2004 is between the upper β boundary value -1.0386 and the upper α boundary value 4.7142, so the trial continues to the next stage.

Output 81.6.11 Sequential Tests

Test Information (Score Scale)
Null Reference = 0

Stage	---Information Level-- Proportion	Actual	-Alternative- --Reference-- Upper	-----Boundary Values---- ----------Upper--------- Beta	Alpha
1	0.2389	3.991698	2.76685	-1.03860	4.71422
2	0.5000	8.353122	5.78997	1.91799	5.73971
3	0.7500	12.52968	8.68495	4.78802	6.70284
4	1.0000	16.70624	11.57993	7.81346	7.81346

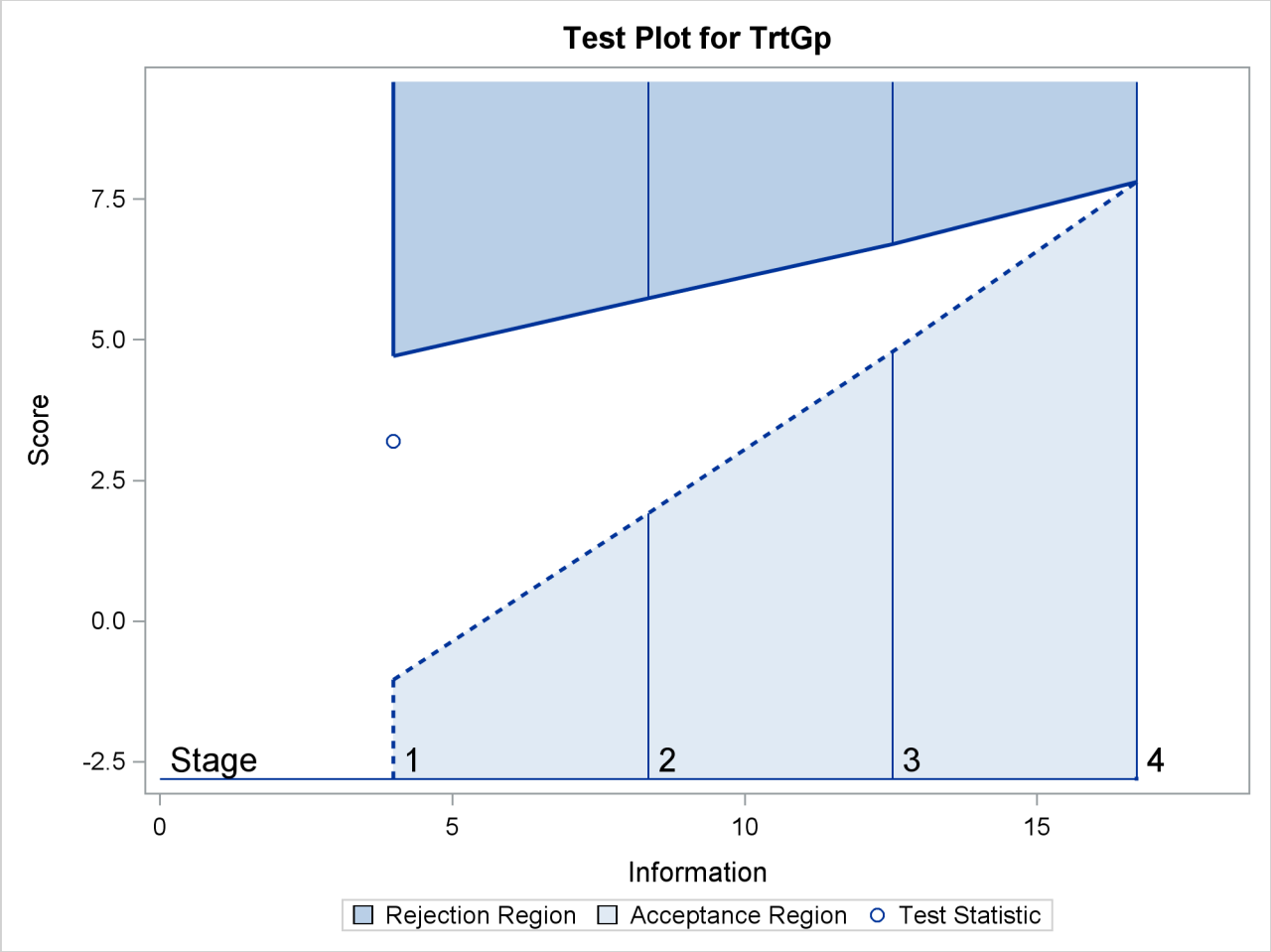
Test Information (Score Scale)
Null Reference = 0

Stage	----------Test---------- ----------TrtGp---------- Estimate	Action
1	3.20040	Continue
2	.	
3	.	
4	.	

Note that the observed information level 3.9917 corresponds to a proportion of 0.2389 in information level. If the observed information level is much smaller than the target proportion of 0.25, then you need to increase the accrual rate, accrual time, or follow-up time to achieve the target maximum information level for the trial. Scharfstein and Tsiatis (1998) use the simulation and bootstrap methods to modify the trial at interim stages to achieve the target maximum information level. These modifications should be specified in the study protocol or study plan before the study starts.

With ODS Graphics enabled, a boundary plot with test statistics is displayed, as shown in Output 81.6.12. As expected, the test statistic is in the continuation region between the upper β and α boundary values.

Output 81.6.12 Sequential Test Plot



Note that the input DATA= option can also be used for the test statistics. For example, the following statements create and display (in Output 81.6.13) the data set for the log-rank statistic and its associated standard error after the LIFETEST procedure. Since the log-rank statistic is a score statistic, the corresponding information level is the variance of the statistic.

```
proc lifetest data=Surv_1;
   time Weeks*Event(0);
   test TrtGp;
ods output logunichisq=Parms_Surv1a;
run;

data Parms_Surv1a;
   set Parms_Surv1a(rename=(Statistic=TrtGp));
   keep _Scale_ _Stage_ _Info_ TrtGp;
   _Scale_='Score';
   _Stage_= 1;
   _Info_= StdErr * StdErr;
   if Variable='TrtGp';
run;
```

```
proc print data=Parms_Surv1a;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.6.13 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	TrtGp	_Scale_	_Stage_	_Info_
1	3.2004	Score	1	3.99170

The following statements can be used to invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Surv
             Data(Testvar=TrtGp)=Parms_Surv1a
             infoadj=none
             boundaryscale=score
             ;
ods output Test=Test_Surv1;
run;
ods graphics off;
```

The following statements use the LIFETEST procedure to compute the log-rank statistic and its associated standard error at stage 2:

```
proc lifetest data=Surv_2;
   time Weeks*Event(0);
   test TrtGp;
ods output logunichisq=Parms_Surv2;
run;
```

The following statements create and display (in Output 81.6.14) the data set for the log-rank statistic and its associated standard error for each of the first two stages:

```
data Parms_Surv2;
   set Parms_Surv2 (rename=(Statistic=Estimate));
   if Variable='TrtGp';
   _Scale_='Score';
   _Stage_= 2;
   keep Variable _Scale_ _Stage_ StdErr Estimate;
run;

proc print data=Parms_Surv2;
   title 'Statistics Computed at Stage 2';
run;
```

Output 81.6.14 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	TrtGp	7.3136	2.9489	Score	2

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
ods graphics on;
proc seqtest Boundary=Test_Surv1
             Parms(Testvar=TrtGp)=Parms_Surv2
             infoadj=none
             boundaryscale=score
             citype=lower
             ;
ods output Test=Test_Surv2;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set. The INFOADJ=NONE option maintains the information level for stage 3 at the value provided in the BOUNDARY= data set.

The ODS OUTPUT statement with the TEST=TEST_SURV2 option creates an output data set named TEST_SURV2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage if the trial does not stop at the current stage.

The "Test Information" table in Output 81.6.15 displays the boundary values for the test statistic. The test statistic 7.31365 is larger than the corresponding upper α boundary 5.79886, so the study stops and rejects the null hypothesis. That is, there is evidence of reduction in hazard rate for the new treatment.

Output 81.6.15 Sequential Tests

The SEQTEST Procedure

Test Information (Score Scale)
Null Reference = 0

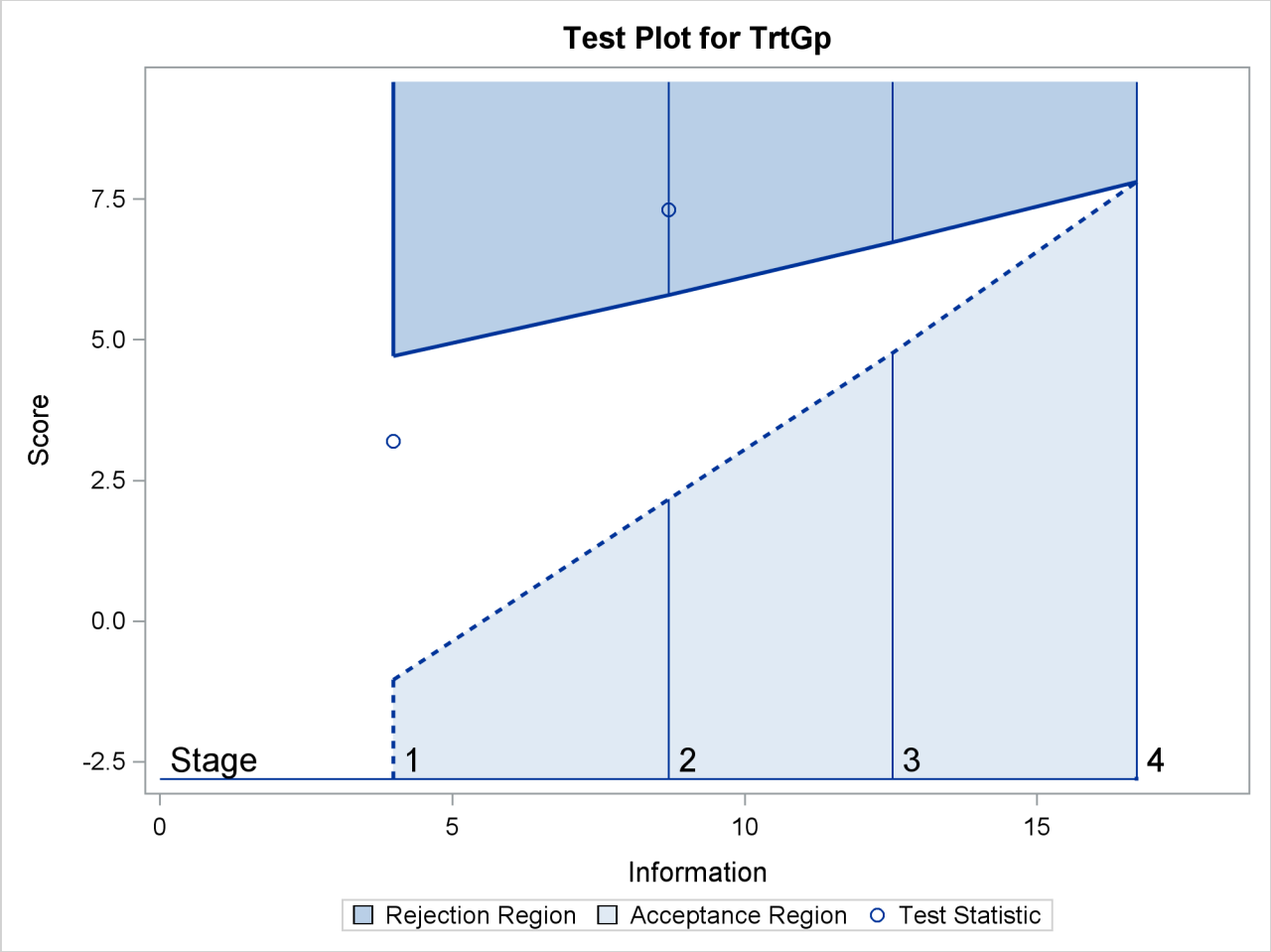
Stage	Information Level Proportion	Information Level Actual	Alternative Reference Upper	Boundary Values Upper Beta	Boundary Values Upper Alpha
1	0.2389	3.991698	2.76685	-1.03860	4.71422
2	0.5205	8.696125	6.02772	2.17045	5.79332
3	0.7500	12.52968	8.68495	4.76306	6.72915
4	1.0000	16.70624	11.57993	7.81287	7.81287

Test Information (Score Scale)
Null Reference = 0

	Test TrtGp	
Stage	Estimate	Action
1	3.20040	Continue
2	7.31365	Reject Null
3	.	
4	.	

With ODS Graphics enabled, the "Test Plot" displays boundary values of the design and the test statistic at the first two stages, as shown in Output 81.6.16. It also shows that the test statistic is in the "Rejection Region" above the upper α boundary value at stage 2.

Output 81.6.16 Sequential Test Plot



After the stopping of a trial, the "Parameter Estimates" table in Output 81.6.17 displays the stopping stage, parameter estimate, unbiased median estimate, confidence limits, and p-value under the null hypothesis $H_0 : \theta = 0$.

Output 81.6.17 Parameter Estimates

Parameter Estimates
Stagewise Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate	Lower 95% CL
TrtGp	2	0.841024	0.0139	0.810329	0.21615

As expected, the p-value 0.0139 is significant at the $\alpha = 0.05$ level and the lower 95% confidence limit is larger than $\theta_0 = 0$. The p-value, unbiased median estimate, and lower confidence limit depend on the ordering of the sample space (k, z), where k is the stage number and z is the standardized Z statistic. With

the specified stagewise ordering, the p-value is $p_1 + p_2$, where p_1 is the α spending at stage 1,

$$p_1 = P_{\theta=0}(Z_1\sqrt{I_1} \geq 4.71422) = 0.00915$$

$$p_2 = P_{\theta=0}(Z_2\sqrt{I_2} \geq 7.31365 \mid -1.04069 < Z_1\sqrt{I_1} < 4.71422)$$

where Z_k is a standardized normal variate and I_k is the information level at stage k for $k = 1, 2$.

Example 81.7: Testing an Effect in a Proportional Hazards Regression Model

This example compares two survival distributions for the treatment effect. The example uses a power family method to generate two-sided asymmetric boundaries and then uses a proportional hazards regression model to test the hypothesis with a covariate.

A clinic is conducting a clinical study for the effect of a new cancer treatment. The study consists of mice exposed to a carcinogen and randomized to either the control group or the treatment group. The event of interest is the death from cancer induced by the carcinogen, and the response is the time from randomization to death.

Consider the proportional hazards regression model

$$h(t;\ \text{TrtGp},\ \text{Wgt}) = h_0(t)\ \exp(\beta_g\ \text{TrtGp}\ +\ \beta_w \text{Wgt})$$

where $h_0(t)$ is an arbitrary and unspecified baseline hazard function, TrtGp is the grouping variable for the two groups, Wgt is the initial weight of the mice, and β_g and β_w are the regression parameters associated with the variables TrtGp and Wgt, respectively. The grouping variable has the value 0 for each mouse in the control group and the value 1 for each mouse in the treatment group.

The hypothesis $H_0 : \beta_g = 0$ with an alternative hypothesis $H_1 : \beta_g \neq 0$ is used for the study.

Suppose that from past experience, the median survival time for the control group is $t_0 = 20$ weeks. The study would like to detect a $t_1 = 40$ weeks 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 the hazard rates $h_0 = 0.03466$ and $h_1 = 0.01733$, the hazard ratio $\exp(\beta_g) = h_1/h_0 = 1/2$ and the alternative hypothesis

$$\beta_{g1} = \log(\frac{1}{2}) = -0.69315$$

Following the derivations in the section "Test for a Parameter in the Proportional Hazards Regression Model" in the chapter "The SEQDESIGN Procedure," the required number of events for testing a parameter in $\boldsymbol{\beta}$ is given by

$$D_X = I_X \frac{1}{(1 - r_x^2)\ \sigma_x^2}$$

where σ_x^2 is the variance of TrtGp and r_x^2 is the proportion of variance of TrtGp explained by the variable Wgt.

If the two groups have the same number of mice in the study, then the MLE of the variance is $\hat{\sigma}_x^2 = 0.25$. Further, if $r_x^2 = 0.10$, then you can specify the MODEL=PHREG(XVARIANCE=0.25 XRSQUARE=0.10) option in the SAMPLESIZE statement in the SEQDESIGN procedure to compute the required number of events and the individual number of events at each stage.

The following statements invoke the SEQDESIGN procedure and request a four-stage group sequential design for normally distributed data. The design uses a two-sided alternative hypothesis with early stopping to reject the null hypothesis H_0. A power family method is used to derive the boundaries.

```
ods graphics on;
proc seqdesign altref=0.69315;
   TwoSidedPowerFamily: design method=pow
                               nstages=4
                               alpha=0.075(lower=0.025)
                               beta=0.20;
   samplesize model=phreg( xvariance=0.25 xrsquare=0.10
                           hazard=0.02451 accrate=10);
run;
ods graphics off;
```

The ALPHA=0.075(LOWER=0.025) option specifies a lower α level 0.025 for the lower rejection boundary and an upper α level $0.05 = 0.075 - 0.025$ for the upper rejection boundary. The geometric average hazard $\sqrt{h_0 \times h_1} = \sqrt{0.03466 \times 0.01733} = 0.02451$ is used in the HAZARD= option in the SAMPLESIZE statement to compute the required sample size. The specified ACCRATE=10 option indicates that 10 mice will be accrued each week and the resulting minimum and maximum accrual times will be displayed.

The "Design Information" table in Output 81.7.1 displays the design specifications and the derived statistics.

Output 81.7.1 Design Information

The SEQDESIGN Procedure
Design: TwoSidedPowerFamily

Design Information

Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Method	Power Family
Boundary Key	Both
Alternative Reference	0.69315
Number of Stages	4
Alpha	0.075
Alpha (Lower)	0.025
Alpha (Upper)	0.05
Beta (Lower)	0.2
Beta (Upper)	0.12764
Power (Lower)	0.8
Power (Upper)	0.87236
Max Information (Percent of Fixed Sample)	106.468
Max Information	17.39288
Null Ref ASN (Percent of Fixed Sample)	104.3691
Lower Alt Ref ASN (Number of Events)	58.04014
Upper Alt Ref ASN (Number of Events)	52.05395

The "Boundary Information" table in Output 81.7.2 displays the information level, alternative reference, and boundary values at each stage. By default (or equivalently if you specify BOUNDARYSCALE=STDZ), the procedure displays the output boundaries with the standardized Z statistic.

Output 81.7.2 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

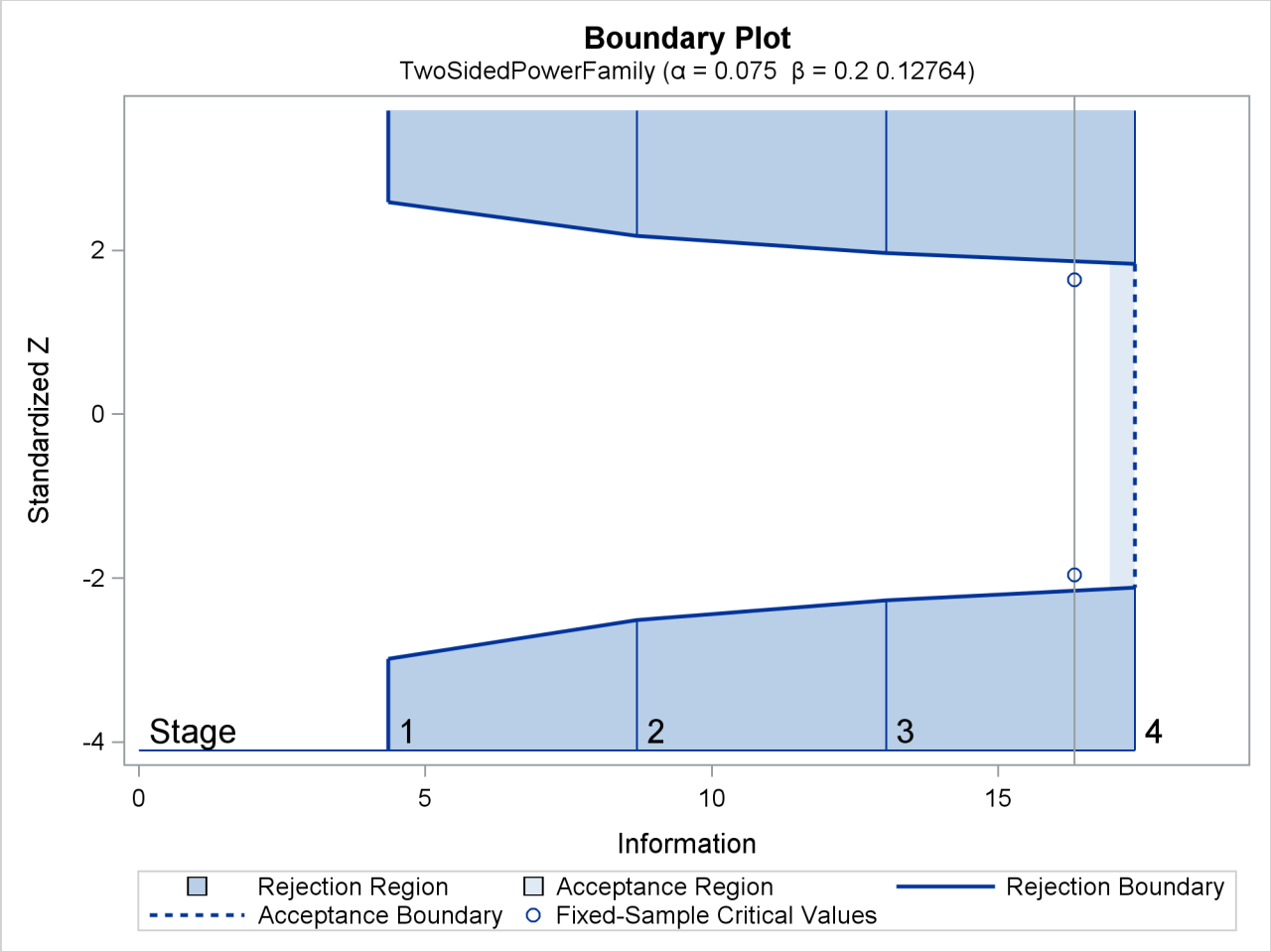
Stage	Information Level Proportion	Information Level Actual	Information Level Events	Alternative Reference Lower	Alternative Reference Upper
1	0.2500	4.348221	19.32543	-1.44538	1.44538
2	0.5000	8.696441	38.65085	-2.04408	2.04408
3	0.7500	13.04466	57.97628	-2.50348	2.50348
4	1.0000	17.39288	77.3017	-2.89077	2.89077

Boundary Information (Standardized Z Scale)
Null Reference = 0

Stage	Boundary Values Lower Alpha	Boundary Values Upper Alpha
1	-2.98871	2.59149
2	-2.51320	2.17917
3	-2.27093	1.96910
4	-2.11334	1.83246

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.7.3.

Output 81.7.3 Boundary Plot



With the MODEL=PHREG option in the SAMPLESIZE statement, the "Sample Size Summary" table in Output 81.7.4 displays the parameters used in the sample size computation for the proportional hazards regression model.

Output 81.7.4 Required Sample Size Summary

Sample Size Summary

Test	PH Reg Parameter
Parameter	0.69315
X Variance	0.25
R Square (X)	0.1
Hazard Rate	0.02451
Accrual Rate	10
Min Accrual Time	7.73017
Min Sample Size	77.3017
Max Accrual Time	27.97872
Max Sample Size	279.7872
Max Number of Events	77.3017

With a minimum accrual time of 7.73 weeks and maximum accrual time of 27.98 weeks, an accrual time of 20 weeks is used in the study. The "Numbers of Events" table in Output 81.7.5 displays the required numbers of events for the group sequential clinical trial.

Output 81.7.5 Required Sample Sizes

Numbers of Events (D)
Z Test for PH Regression Parameter

Stage	D	Information
1	19.33	4.3482
2	38.65	8.6964
3	57.98	13.0447
4	77.30	17.3929

The following statements invoke the SEQDESIGN procedure and provide more detailed sample size information with a 20-week accrual time:

```
proc seqdesign altref=0.69315;
   TwoSidedPowerFamily: design method=pow
                               nstages=4
                               alpha=0.075(lower=0.025)
                               beta=0.20;
   samplesize model=phreg( xvariance=0.25 xrsquare=0.10
                           hazard=0.02451
                           accrate=10 acctime=20);
ods output Boundary=Bnd_Time;
run;
```

The ODS OUTPUT statement with the BOUNDARY=BND_TIME option creates an output data set named BND_TIME which contains the resulting boundary information for the subsequent sequential tests.

With an accrual time of 20 weeks, the "Sample Size Summary" table in Output 81.7.6 displays the follow-up time for the trial.

Output 81.7.6 Sample Size Summary

The SEQDESIGN Procedure
Design: TwoSidedPowerFamily

Sample Size Summary

Test	PH Reg Parameter
Parameter	0.69315
X Variance	0.25
R Square (X)	0.1
Hazard Rate	0.02451
Accrual Rate	10
Accrual Time	20
Follow-up Time	10.34195
Total Time	30.34195
Max Number of Events	77.3017
Max Sample Size	200
Expected Sample Size (Null Ref)	199.4282
Expected Sample Size (Alt Ref)	188.6561

The "Numbers of Events and Sample Sizes" table in Output 81.7.7 displays the required sample sizes for the group sequential clinical trial.

Output 81.7.7 Numbers of Events and Sample Sizes

Numbers of Events (D) and Sample Sizes (N)
Z Test for PH Regression Parameter

	Fractional Time			
Stage	D	Time	N	Information
1	19.33	13.2362	132.36	4.3482
2	38.65	19.1466	191.47	8.6964
3	57.98	24.3744	200.00	13.0447
4	77.30	30.3420	200.00	17.3929

Numbers of Events (D) and Sample Sizes (N)
Z Test for PH Regression Parameter

	Ceiling Time			
Stage	D	Time	N	Information
1	21.49	14	140.00	4.8359
2	41.90	20	200.00	9.4281
3	60.14	25	200.00	13.5309
4	79.26	31	200.00	17.8346

Thus, the study will perform three interim analyses after 14, 20, and 25 weeks and a final analysis after 31 weeks if the study does not stop at any of the interim analyses.

Suppose 140 mice are available for the first interim analysis after week 14. Output 81.7.8 lists the first 10 observations in the data set weeks_1.

Output 81.7.8 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Trt Gp	Event	Wgt	Weeks
1	0	0	22.1659	12
2	1	0	28.4458	12
3	0	0	26.2857	12
4	1	0	25.0283	12
5	0	0	21.5114	12
6	1	0	23.2240	12
7	0	1	22.6845	6
8	1	0	27.9292	12
9	0	0	22.5514	12
10	1	1	27.3793	11

The TrtGp variable is a grouping variable with the value 0 for a mouse in the placebo control group and the value 1 for a mouse in the treatment group.

The Weeks variable is the survival time variable measured in weeks and the Event variable is the censoring variable with the value 0 indicating censoring. That is, the values of Weeks are considered censored if the corresponding values of Event are 0; otherwise, they are considered as event times.

The following statements use the PHREG procedure to estimate the treatment effect after adjusting for the Wgt variable at stage 1:

```
proc phreg data=Time_1;
   model Weeks*Event(0)= TrtGp Wgt;
ods output parameterestimates=Parms_Time1;
run;
```

The following statements create and display (in Output 81.7.9) the data set for the treatment effect MLE statistic and its associated standard error. Note that for a MLE statistic, the inverse of the variance of the statistic is the information.

```
data Parms_Time1;
   set Parms_Time1;
   if Parameter='TrtGp';
   _Scale_='MLE';
   _Stage_= 1;
   keep _Scale_ _Stage_ Parameter Estimate StdErr;
run;

proc print data=Parms_Time1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.7.9 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Parameter	Estimate	StdErr	_Scale_	_Stage_
1	TrtGp	0.00836	0.46588	MLE	1

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Time
             Parms(Testvar=TrtGp)=Parms_Time1
             infoadj=prop
             order=lr
             ;
ods output Test=Test_Time1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The PARMS=PARMS_TIME1 option specifies the input data set PARMS_TIME1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=TRTGP option identifies the test variable TRTGP in the data set.

If the computed information level for stage 1 is not the same as the value provided in the BOUNDARY= data set, the INFOADJ=PROP option (which is the default) proportionally adjusts the information levels at future interim stages from the levels provided in the BOUNDARY= data set. The ORDER=LR option uses the LR ordering to derive the p-value, the unbiased median estimate, and the confidence limits for the regression slope estimate.

The ODS OUTPUT statement with the TEST=TEST_TIME1 option creates an output data set named TEST_TIME1 which contains the updated boundary information for the test at stage 1. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.7.10 displays design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are modified for the new information levels to maintain the Type I α level. The maximum information and the power have been modified for the new information levels.

Output 81.7.10 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_TIME
Data Set	WORK.PARMS_TIME1
Statistic Distribution	Normal
Boundary Scale	Standardized Z
Alternative Hypothesis	Two-Sided
Early Stop	Reject Null
Number of Stages	4
Alpha	0.075
Alpha (Lower)	0.025
Alpha (Upper)	0.05
Beta (Lower)	0.20048
Beta (Upper)	0.12795
Power (Lower)	0.79952
Power (Upper)	0.87205
Max Information (Percent of Fixed Sample)	106.5982
Max Information	17.3928828
Null Ref ASN (Percent of Fixed Sample)	104.4715
Lower Alt Ref ASN (Percent of Fixed Sample)	79.7886
Upper Alt Ref ASN (Percent of Fixed Sample)	71.53877

The "Test Information" table in Output 81.7.11 displays the boundary values for the test statistic with the MLE statistic scale.

Output 81.7.11 Sequential Tests

Test Information (Standardized Z Scale)
Null Reference = 0

	--Information Level-		------Alternative----- Reference------		----Boundary Values--- ---Lower--	---Upper--
Stage	Proportion	Actual	Lower	Upper	Alpha	Alpha
1	0.2649	4.607347	-1.48783	1.48783	-2.92457	2.54086
2	0.5099	8.869192	-2.06428	2.06428	-2.50505	2.17290
3	0.7550	13.13104	-2.51175	2.51175	-2.27093	1.96941
4	1.0000	17.39288	-2.89077	2.89077	-2.11635	1.83531

Test Information (Standardized Z Scale)
Null Reference = 0

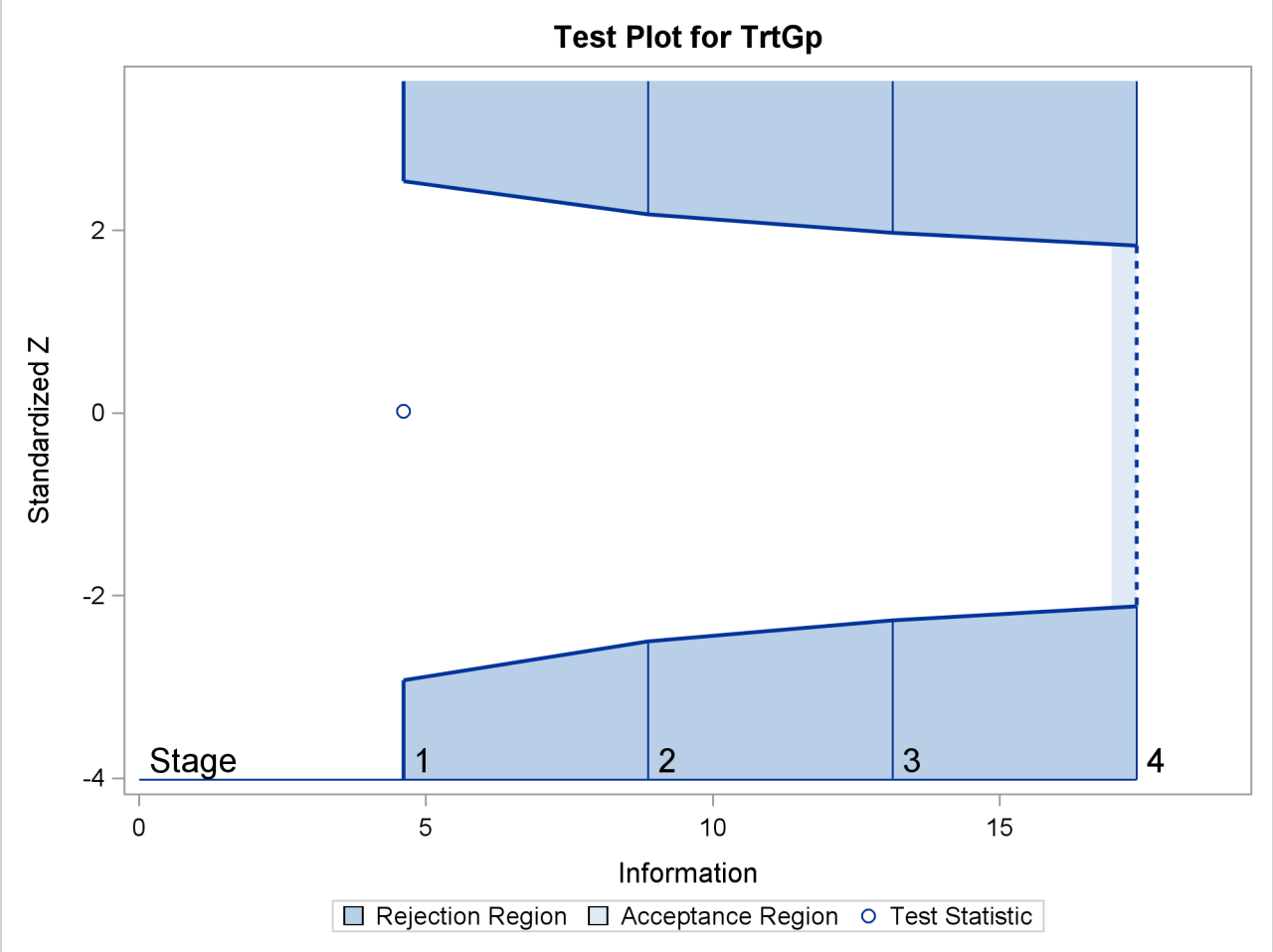
	----------Test---------- ----------TrtGp----------	
Stage	Estimate	Action
1	0.01795	Continue
2	.	
3	.	
4	.	

With the INFOADJ=PROP option (which is the default), the information levels at interim stages 2 and 3 are derived proportionally from the information levels in the BOUNDARY= data set. At stage 1, the standardized Z statistic 0.01795 is between the lower and upper α boundary values of -2.92457 and 2.54086, so the trial continues to the next stage.

Note that the observed information level 4.6073 corresponds to a proportion of 0.2649 in the information level. If the observed information level is much larger than the target proportion of 0.25, then you can decrease the accrual rate, accrual time, or follow-up time to achieve target information levels for subsequent stages. These modifications should be specified in the study plan before the study begins.

With ODS Graphics enabled, a boundary plot with test statistics is displayed, as shown in Output 81.7.12. As expected, the test statistic is in the continuation region between the lower and upper α boundary values.

Output 81.7.12 Sequential Test Plot



The following statements use the PHREG procedure to compute the MLE statistic and its associated standard error at stage 2:

```
proc phreg data=Time_2;
   model Weeks*Event(0)= TrtGp Wgt;
ods output parameterestimates= Parms_Time2;
run;
```

The following statements create the data set for the MLE statistic and its associated standard error at stage 2:

```
data Parms_Time2;
   set Parms_Time2;
   if Parameter='TrtGp';
   _Scale_='MLE';
   _Stage_= 2;
   keep _Scale_ _Stage_ Parameter Estimate StdErr;
run;
```

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
proc seqtest Boundary=Test_Time1
             Parms(Testvar=TrtGp)=Parms_Time2
             infoadj=prop
             order=lr
             ;
ods output Test=Test_Time2;
run;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set.

The ODS OUTPUT statement with the TEST=TEST_TIME2 option creates an output data set named TEST_TIME2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Test Information" table in Output 81.7.13 displays the boundary values for the test statistic with the MLE statistic scale. At stage 2, the standardized Z statistic -0.43552 is between the lower α and upper boundary values, -2.47689 and 2.14819, respectively, so the trial continues to the next stage.

Output 81.7.13 Sequential Tests

The SEQTEST Procedure

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	--Information Level- Proportion	Actual	------Alternative----- -------Reference------ Lower	Upper	----Boundary Values--- ---Lower-- Alpha	---Upper-- Alpha
1	0.2649	4.607347	-1.48783	1.48783	-2.92457	2.54086
2	0.5251	9.132918	-2.09475	2.09475	-2.47689	2.14819
3	0.7625	13.2629	-2.52433	2.52433	-2.26878	1.96770
4	1.0000	17.39288	-2.89077	2.89077	-2.12017	1.83880

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	----------Test---------- ----------TrtGp---------- Estimate	Action
1	0.01795	Continue
2	-0.43552	Continue
3	.	
4	.	

Since the data set PARMS_Time2 contains the test information only at stage 2, the information level at stage 1 in the TEST_Time1 data set is used to generate boundary values for the test.

Similarly, the test statistic at stage 3 is also between its corresponding lower and upper α boundary values. The trial continues to the next stage.

The following statements use the PHREG procedure to compute the MLE statistic and its associated standard error at the final stage:

```
proc phreg data=Time_4;
   model Weeks*Event(0)= TrtGp Wgt;
ods output parameterestimates= Parms_Time4;
run;
```

The following statements create and display (in Output 81.7.14) the data set for the MLE statistic and its associated standard error at each stage of the study:

```
data Parms_Time4;
   set Parms_Time4;
   if Parameter='TrtGp';
   _Scale_='MLE';
   _Stage_= 4;
   keep _Scale_ _Stage_ Parameter Estimate StdErr;
run;
proc print data=Parms_Time4;
   title 'Statistics Computed at Stage 4';
run;
```

Output 81.7.14 Statistics Computed at Stage 4

Statistics Computed at Stage 4

Obs	Parameter	Estimate	StdErr	_Scale_	_Stage_
1	TrtGp	-0.04451	0.23971	MLE	4

The following statements invoke the SEQTEST procedure to test the hypothesis at stage 4:

```
ods graphics on;
proc seqtest Boundary=Test_Time3
             Parms(Testvar=TrtGp)=Parms_Time4
             order=lr
             ;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 4, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 4, and the TESTVAR= option identifies the test variable in the data set.

The "Test Information" table in Output 81.7.15 displays the boundary values for the test statistic. The standardized test statistic -0.1857 is between the lower and upper α boundary values of -2.10447 and 1.82112, respectively, so the study stops and accepts the null hypothesis. That is, there is no evidence of reduction in hazard rate for the new treatment.

Output 81.7.15 Sequential Tests

The SEQTEST Procedure

Test Information (Standardized Z Scale)
Null Reference = 0

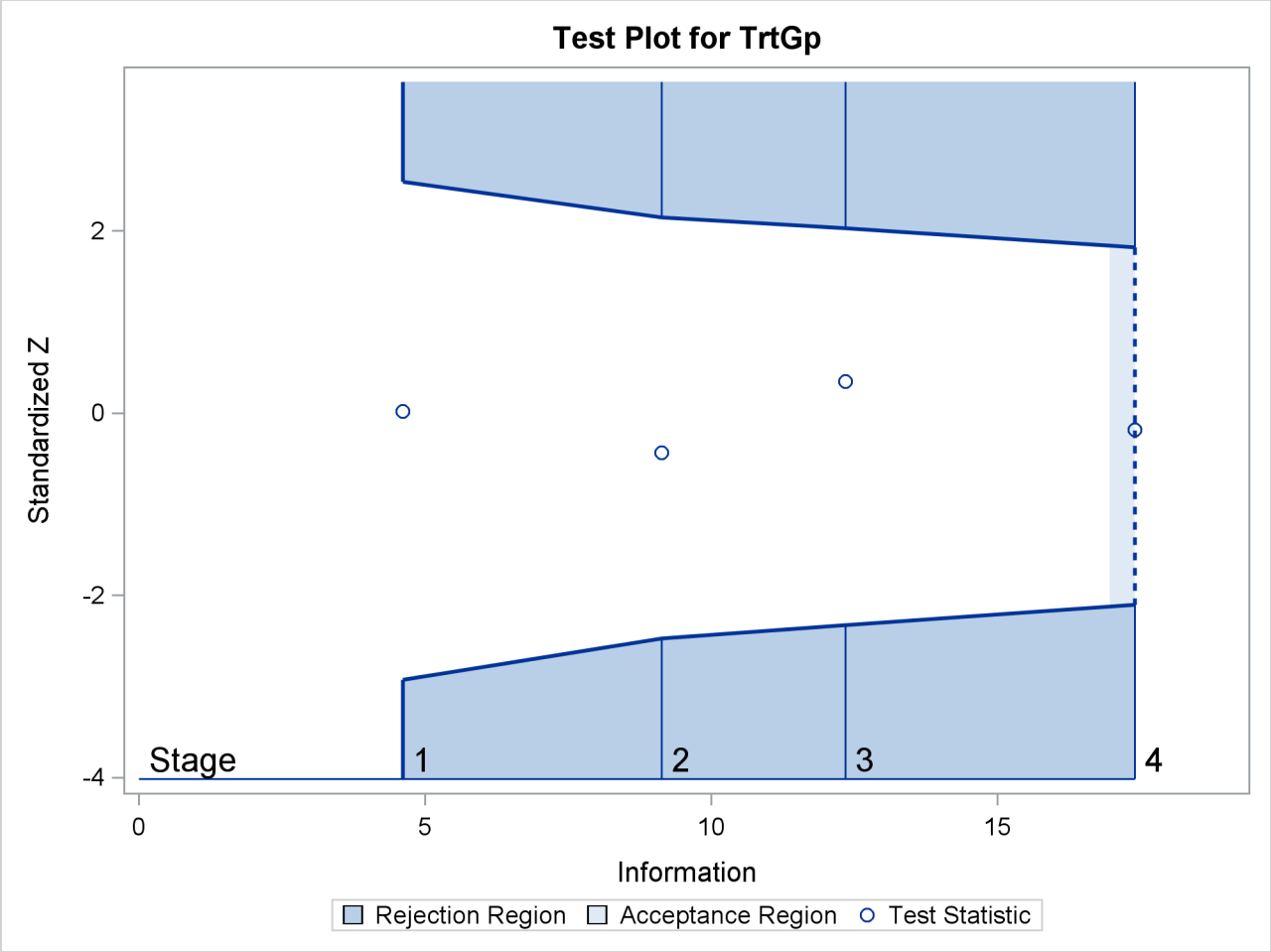
Stage	--Information Level- Proportion	Actual	------Alternative----- -------Reference------ Lower	Upper	----Boundary Values--- ---Lower-- Alpha	---Upper-- Alpha
1	0.2647	4.607347	-1.48783	1.48783	-2.92457	2.54086
2	0.5248	9.132918	-2.09475	2.09475	-2.47689	2.14819
3	0.7095	12.34753	-2.43566	2.43566	-2.32705	2.02634
4	1.0000	17.40274	-2.89159	2.89159	-2.10447	1.82112

Test Information (Standardized Z Scale)
Null Reference = 0

Stage	----------Test---------- ----------TrtGp---------- Estimate	Action
1	0.01795	Continue
2	-0.43552	Continue
3	0.34864	Continue
4	-0.18570	Accept Null

The "Test Plot" displays boundary values of the design and the test statistic at the first two stages, as shown in Output 81.7.16. It also shows that the test statistic is in the "Acceptance Region" between the lower and upper α boundary values at stage 4.

Output 81.7.16 Sequential Test Plot



After the stopping of a trial, the "Parameter Estimates" table in Output 81.7.17 displays the stopping stage, parameter estimate, unbiased median estimate, confidence limits, and p-value under the null hypothesis $H_0 : \theta = 0$.

Output 81.7.17 Parameter Estimates

Parameter Estimates
LR Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate
TrtGp	4	-0.044514	0.8525	-0.044577

Parameter Estimates
LR Ordering

Parameter	95% Confidence Limits	
TrtGp	-0.51461	0.42538

As expected, the two-sided p-value 0.8525 is not significant at the lower $\alpha = 0.025$ level and the upper $\alpha = 0.05$ level, and the two-sided 95% confidence interval contains the null value zero. The p-value, unbiased median estimate, and lower confidence limit depend on the ordering of the sample space (k, z), where k is the stage number and z is the standardized Z statistic. With the specified LR ordering, the two-sided p-value is derived from the one-sided p-value

$$p_u = \sum_{k=1}^{4} P_{\theta=0}\left(Z_k \geq z_4 \mid \_a_{k'} < Z_{k'} < a_{k'}, k' < k\right)$$

where $z_4 = -0.1857$ is the observed test statistic at stage 4, Z_k is a standardized normal variate at stage k, and $\_a_{k'}$ and $a_{k'}$ are the stage k lower and upper rejection boundary values, respectively.

Thus,

$$p_u = \alpha_u + P_{\theta=0}\left(z_4 \leq Z_4 < a_4 \mid \_a_{k'} < Z_{k'} < a_{k'}, k' < 4\right)$$

where $\alpha_u = 0.05$ is the upper α level and $a_4 = 1.82112$.

Since $P_{\theta=0}\left(z_4 \leq Z_4 \leq a_4 \mid \_a_{k'} < Z_{k'} < a_{k'}, k' < 4\right) = 0.52374$, $p_u = 0.05 + 0.52374 = 0.57374$, which is greater than 0.50. Thus, the two-sided p-value is given by $2 \times (1.0 - p_u) = 0.8525$.

Example 81.8: Testing an Effect in a Logistic Regression Model

This example requests a two-sided test for the dose effect in a dose-response model (Whitehead 1997, pp. 262–263). Consider the logistic regression model

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 \, \text{LDose}$$

where $p = \text{Prob}(\text{Resp} = 1 \mid \text{LDose})$ is the response probability to be modeled for the binary response Resp and LDose = log(Dose +1) is the covariate. The dose levels are 0 for the control group, and they are 1, 3, and 6 for the three treatment groups.

Following the derivations in the section "Test for a Parameter in the Logistic Regression Model" in the chapter "The SEQDESIGN Procedure," the required sample size can be derived from

$$N = I_X \, \frac{\sigma_y^2}{(1 - r_x^2)\,\sigma_x^2}$$

where σ_y^2 is the variance of the response variable in the logistic regression model, r_x^2 is the proportion of variance of LDose explained by other covariates, and σ_x^2 is the variance of LDose.

Since LDose is the only covariate in the model, $r_x^2 = 0$. For a logistic model, the variance σ^2 can be estimated by

$$\sigma_y^2 = \frac{1}{\hat{p}(1-\hat{p})}$$

where $\hat{p}$ is the estimated probability of the response variable Resp. Thus, the sample size can be computed as

$$N = I_X \, \frac{1}{p(1-p)} \, \frac{1}{\sigma_x^2}$$

The null hypothesis $H_0 : \beta_1 = 0$ corresponds to no treatment effect. Suppose that the alternative hypothesis $H_1 : \beta_1 = 0.5$ is the reference improvement that should be detected at a 0.90 level.

Note that $\beta_1 = 0.5$ corresponds to an odds ratio of 2 between the treatment group with dose level 3 and the control group. The log odds ratio between the two groups is

$$\log\left(\frac{p_t(1-p_c)}{(1-p_t)p_0}\right) = \log\left(\frac{p_t}{1-p_t}\right) - \log\left(\frac{p_c}{1-p_c}\right)$$

which corresponds to

$$(\beta_0 + \beta_1 \log(3+1)) \; - \; (\beta_0 + \beta_1 \log(1)) = \beta_1 \log(4) = \log(2)$$

If the same number of patients are assigned in each of the four groups, then the MLE of the variance of LDose is $\hat{\sigma}_x^2 = 0.5345$. Further, if the response rate is 0.40, then the required sample size can be derived using the SAMPLESIZE statement in the SEQDESIGN procedure.

The following statements invoke the SEQDESIGN procedure and request a three-stage group sequential design for normally distributed data. The design has a null hypothesis of no treatment effect $H_0 : \beta_1 = 0$ with early stopping to reject the null hypothesis with a two-sided alternative hypothesis $H_1 : \beta_1 = \pm 0.5$.

```
ods graphics on;
proc seqdesign altref=0.5;
   TwoSidedErrorSpending: design method=errfuncpow
                                method(loweralpha)=errfuncpow(rho=1)
                                method(upperalpha)=errfuncpow(rho=3)
                                nstages=3
                                stop=both;
   samplesize model=logistic( prop=0.4 xvariance=0.5345);
ods output Boundary=Bnd_Dose;
run;
ods graphics off;
```

The ODS OUTPUT statement with the BOUNDARY=BND_DOSE option creates an output data set named BND_DOSE which contains the resulting boundary information for the subsequent sequential tests.

The "Design Information" table in Output 81.8.1 displays design specifications and derived statistics. Since the alternative reference is specified, the maximum information 47.22445 is derived.

Output 81.8.1 Error Spending Design Information

The SEQDESIGN Procedure
Design: TwoSidedErrorSpending

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	0.5
Number of Stages	3
Alpha	0.05
Beta (Lower)	0.1
Beta (Upper)	0.07871
Power (Lower)	0.9
Power (Upper)	0.92129
Max Information (Percent of Fixed Sample)	103.7223
Max Information	47.22445
Null Ref ASN (Percent of Fixed Sample)	79.47628
Lower Alt Ref ASN (Sample Size)	234.1646
Upper Alt Ref ASN (Sample Size)	270.4058

The "Boundary Information" table in Output 81.8.2 displays the information level, alternative reference, and boundary values at each stage. By default (or equivalently if you specify BOUNDARYSCALE=STDZ), the boundary values are displayed with the standardized Z statistic scale.

Output 81.8.2 Boundary Information

Boundary Information (Standardized Z Scale)
Null Reference = 0

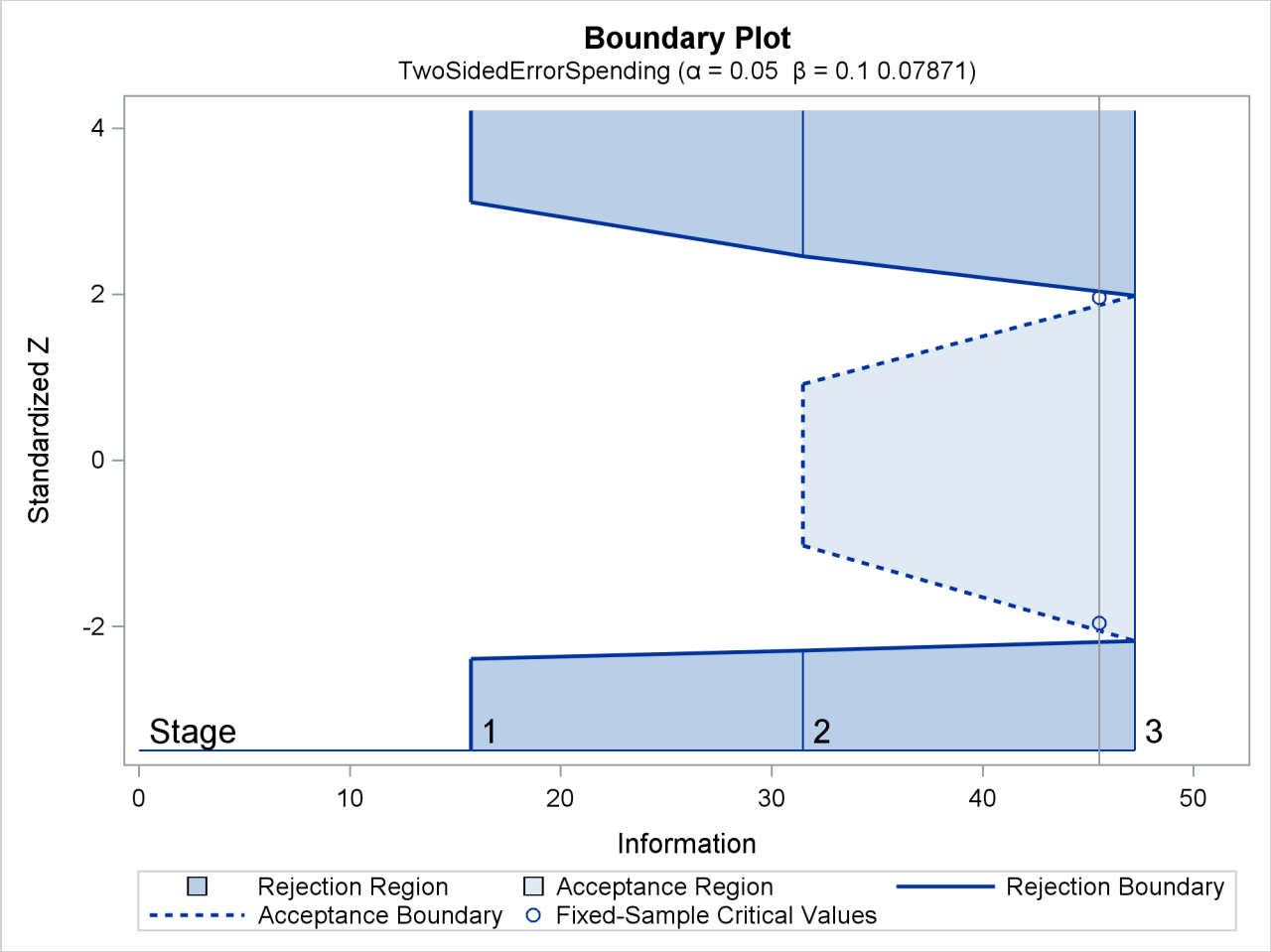
	Information Level			Alternative Reference	
Stage	Proportion	Actual	N	Lower	Upper
1	0.3333	15.74148	122.7119	-1.98378	1.98378
2	0.6667	31.48297	245.4238	-2.80548	2.80548
3	1.0000	47.22445	368.1357	-3.43600	3.43600

Boundary Information (Standardized Z Scale)
Null Reference = 0

	Boundary Values			
	Lower		Upper	
Stage	Alpha	Beta	Beta	Alpha
1	-2.39398	.	.	3.11302
2	-2.29380	-1.02812	0.91855	2.46193
3	-2.17479	-2.17479	1.98311	1.98311

With ODS Graphics enabled, a detailed boundary plot with the rejection and acceptance regions is displayed, as shown in Output 81.8.3.

Output 81.8.3 Boundary Plot



With the SAMPLESIZE statement, the "Sample Size Summary" table in Output 81.8.4 displays the parameters for the sample size computation.

Output 81.8.4 Required Sample Size Summary

Sample Size Summary

Test	Logistic Reg Parameter
Parameter	0.5
Proportion	0.4
X Variance	0.5345
R Square (X)	0
Max Sample Size	368.1357
Expected Sample Size (Null Ref)	282.0807
Expected Sample Size (Alt Ref)	270.4058

The "Sample Sizes" table in Output 81.8.5 displays the required sample sizes for the group sequential clinical trial.

Output 81.8.5 Required Sample Sizes

Sample Sizes (N)
Z Test for Logistic Regression Parameter

	Fractional N		Ceiling N	
Stage	N	Information	N	Information
1	122.71	15.7415	123	15.7784
2	245.42	31.4830	246	31.5569
3	368.14	47.2245	369	47.3353

That is, 123 new patients are needed in each stage and the number is rounded up to 124 for each stage to have a multiple of four for the four dose levels in the trial. Note that since the sample sizes are derived from an estimated response probability and are rounded up, the actual information levels might not match the corresponding target information levels.

Output 81.8.6 lists the first 10 observations of the trial data.

Output 81.8.6 Clinical Trial Data

First 10 Obs in the Trial Data

Obs	Resp	Dose	LDose
1	0	0	0.00000
2	0	1	0.69315
3	1	3	1.38629
4	1	6	1.94591
5	1	0	0.00000
6	1	1	0.69315
7	1	3	1.38629
8	1	6	1.94591
9	0	0	0.00000
10	0	1	0.69315

The following statements use the LOGISTIC procedure to estimate the slope β_1 and its associated standard error at stage 1:

```
proc logistic data=Dose_1;
   model Resp(event='1')= LDose;
   ods output ParameterEstimates=Parms_Dose1;
run;
```

The following statements create and display (in Output 81.8.7) the input data set that contains slope β_1 and its associated standard error for the SEQTEST procedure:

```
data Parms_Dose1;
   set Parms_Dose1;
   if Variable='LDose';
   _Scale_='MLE';
   _Stage_= 1;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_Dose1;
   title 'Statistics Computed at Stage 1';
run;
```

Output 81.8.7 Statistics Computed at Stage 1

Statistics Computed at Stage 1

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	LDose	0.5741	0.2544	MLE	1

The following statements invoke the SEQTEST procedure to test for early stopping at stage 1:

```
ods graphics on;
proc seqtest Boundary=Bnd_Dose
             Parms(testvar=LDose)=Parms_Dose1
             infoadj=prop
             order=mle
             boundaryscale=mle
             ;
ods output Test=Test_Dose1;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 1, which was generated in the SEQDESIGN procedure. The PARMS=PARMS_DOSE1 option specifies the input data set PARMS_DOSE1 that contains the test statistic and its associated standard error at stage 1, and the TESTVAR=LDOSE option identifies the test variable LDOSE in the data set.

If the computed information level for stage 1 is not the same as the value provided in the BOUNDARY= data set, the INFOADJ=PROP option (which is the default) proportionally adjusts the information levels at future interim stages from the levels provided in the BOUNDARY= data set. The ORDER=MLE option uses the MLE ordering to derive the p-value, the unbiased median estimate, and the confidence limits for the regression slope estimate.

The ODS OUTPUT statement with the TEST=TEST_DOSE1 option creates an output data set named TEST_DOSE1 which contains the updated boundary information for the test at stage 1. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.8.8 displays design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are modified for the new information

levels to maintain the Type I α level. The maximum information remains the same as the design stored in the BOUNDARY= data set, but the derived Type II error probability β and power $1-\beta$ are different because of the new information levels.

Output 81.8.8 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.BND_DOSE
Data Set	WORK.PARMS_DOSE1
Statistic Distribution	Normal
Boundary Scale	MLE
Alternative Hypothesis	Two-Sided
Early Stop	Accept/Reject Null
Number of Stages	3
Alpha	0.05
Beta (Lower)	0.09992
Beta (Upper)	0.07871
Power (Lower)	0.90008
Power (Upper)	0.92129
Max Information (Percent of Fixed Sample)	103.7231
Max Information	47.2244524
Null Ref ASN (Percent of Fixed Sample)	79.45049
Lower Alt Ref ASN (Percent of Fixed Sample)	66.05269
Upper Alt Ref ASN (Percent of Fixed Sample)	76.24189

The "Test Information" table in Output 81.8.9 displays the boundary values for the test statistic with the specified MLE scale.

Output 81.8.9 Sequential Tests

Test Information (MLE Scale)
Null Reference = 0

	Information Level		Alternative Reference	
Stage	Proportion	Actual	Lower	Upper
1	0.3272	15.45062	-0.50000	0.50000
2	0.6636	31.33753	-0.50000	0.50000
3	1.0000	47.22445	-0.50000	0.50000

Test Information (MLE Scale)
Null Reference = 0

	Boundary Values				Test	
	Lower		Upper		LDose	
Stage	Alpha	Beta	Beta	Alpha	Estimate	Action
1	-0.61078	.	.	0.79337	0.57409	Continue
2	-0.40974	-0.18169	0.16222	0.44033	.	
3	-0.31633	-0.31633	0.28860	0.28860	.	

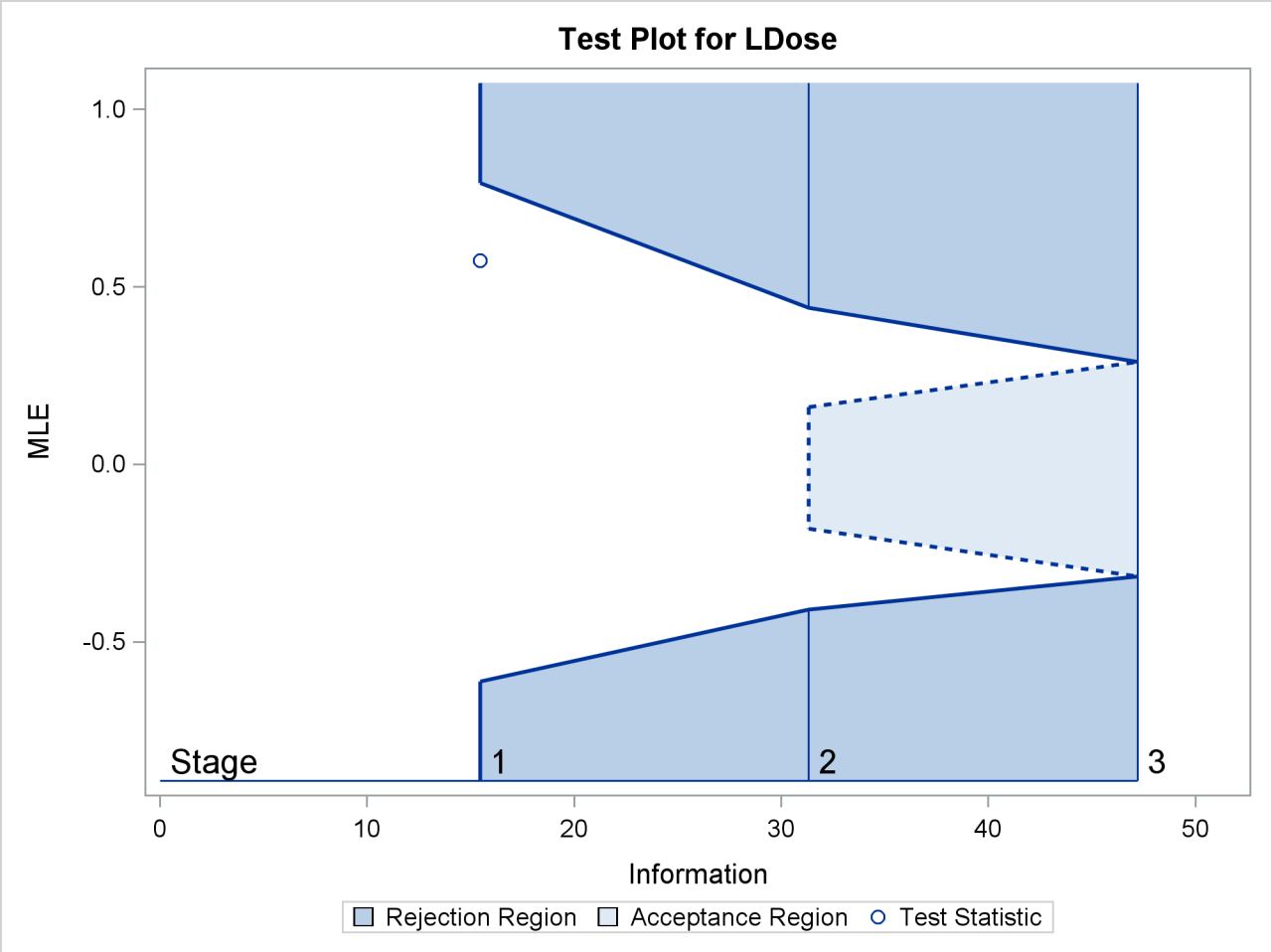
The information level at stage 1 is derived from the standard error s_1 in the PARMS= data set,

$$I_1 = \frac{1}{s_1^2} = \frac{1}{0.2544^2} = 15.45$$

With the INFOADJ=PROP option (which is the default), the information level at stage 2 is derived proportionally from the observed information at stage 1 and the information levels in the BOUNDARY= data set. At stage 1, the β boundary values are missing and there is no early stopping to accept H_0. The MLE statistic 0.57409 is between the lower and upper α boundaries, so the trial continues to the next stage.

With ODS Graphics enabled, a boundary plot with the boundary values and test statistics is displayed, as shown in Output 81.8.10. As expected, the test statistic is in the continuation region below the upper alpha boundary.

Output 81.8.10 Sequential Test Plot



The following statements use the LOGISTIC procedure to estimate the slope β_1 and its associated standard error at stage 2:

```
proc logistic data=dose_2;
   model Resp(event='1')=LDose;
   ods output ParameterEstimates=Parms_Dose2;
run;
```

The following statements create and display (in Output 81.8.11) the input data set that contains slope β_1 and its associated standard error at stage 2 for the SEQTEST procedure:

```
data Parms_Dose2;
   set Parms_Dose2;
   if Variable='LDose';
   _Scale_='MLE';
   _Stage_= 2;
   keep _Scale_ _Stage_ Variable Estimate StdErr;
run;

proc print data=Parms_Dose2;
   title 'Statistics Computed at Stage 2';
run;
```

Output 81.8.11 Statistics Computed at Stage 2

Statistics Computed at Stage 2

Obs	Variable	Estimate	StdErr	_Scale_	_Stage_
1	LDose	0.5213	0.1788	MLE	2

The following statements invoke the SEQTEST procedure to test for early stopping at stage 2:

```
ods graphics on;
proc seqtest Boundary=Test_Dose1
             Parms(Testvar=LDose)=Parms_Dose2
             infoadj=prop
             order=mle
             boundaryscale=mle
             rci
             plots=rci
             ;
ods output Test=Test_Dose2;
run;
ods graphics off;
```

The BOUNDARY= option specifies the input data set that provides the boundary information for the trial at stage 2, which was generated by the SEQTEST procedure at the previous stage. The PARMS= option specifies the input data set that contains the test statistic and its associated standard error at stage 2, and the TESTVAR= option identifies the test variable in the data set.

The ORDER=MLE option uses the MLE ordering to derive the p-value, unbiased median estimate, and confidence limits for the regression slope estimate.

The ODS OUTPUT statement with the TEST=TEST_DOSE2 option creates an output data set named TEST_DOSE2 which contains the updated boundary information for the test at stage 2. The data set also provides the boundary information that is needed for the group sequential test at the next stage.

The "Design Information" table in Output 81.8.12 displays design specifications. By default (or equivalently if you specify BOUNDARYKEY=ALPHA), the boundary values are modified for the new information levels to maintain the Type I α level.

Output 81.8.12 Design Information

The SEQTEST Procedure

Design Information

BOUNDARY Data Set	WORK.TEST_DOSE1
Data Set	WORK.PARMS_DOSE2
Statistic Distribution	Normal
Boundary Scale	MLE
Alternative Hypothesis	Two-Sided
Early Stop	Accept/Reject Null
Number of Stages	3
Alpha	0.05
Beta (Lower)	0.0999
Beta (Upper)	0.07871
Power (Lower)	0.9001
Power (Upper)	0.92129
Max Information (Percent of Fixed Sample)	103.7227
Max Information	47.2244524
Null Ref ASN (Percent of Fixed Sample)	79.44086
Lower Alt Ref ASN (Percent of Fixed Sample)	66.04641
Upper Alt Ref ASN (Percent of Fixed Sample)	76.22739

The information is derived from the standard error associated with the slope estimate at the final stage and is larger than the target level. The derived Type II error probability β and power $1-\beta$ are different because of the new information levels.

The "Test Information" table in Output 81.8.13 displays the boundary values for the test statistic with the specified MLE scale. The information levels are derived from the standard errors in the PARMS= data set. At stage 2, the slope estimate 0.52128 is larger than 0.44091, the upper α boundary value, the trial stops to reject the null hypothesis of no treatment effect.

Output 81.8.13 Sequential Tests

Test Information (MLE Scale)
Null Reference = 0

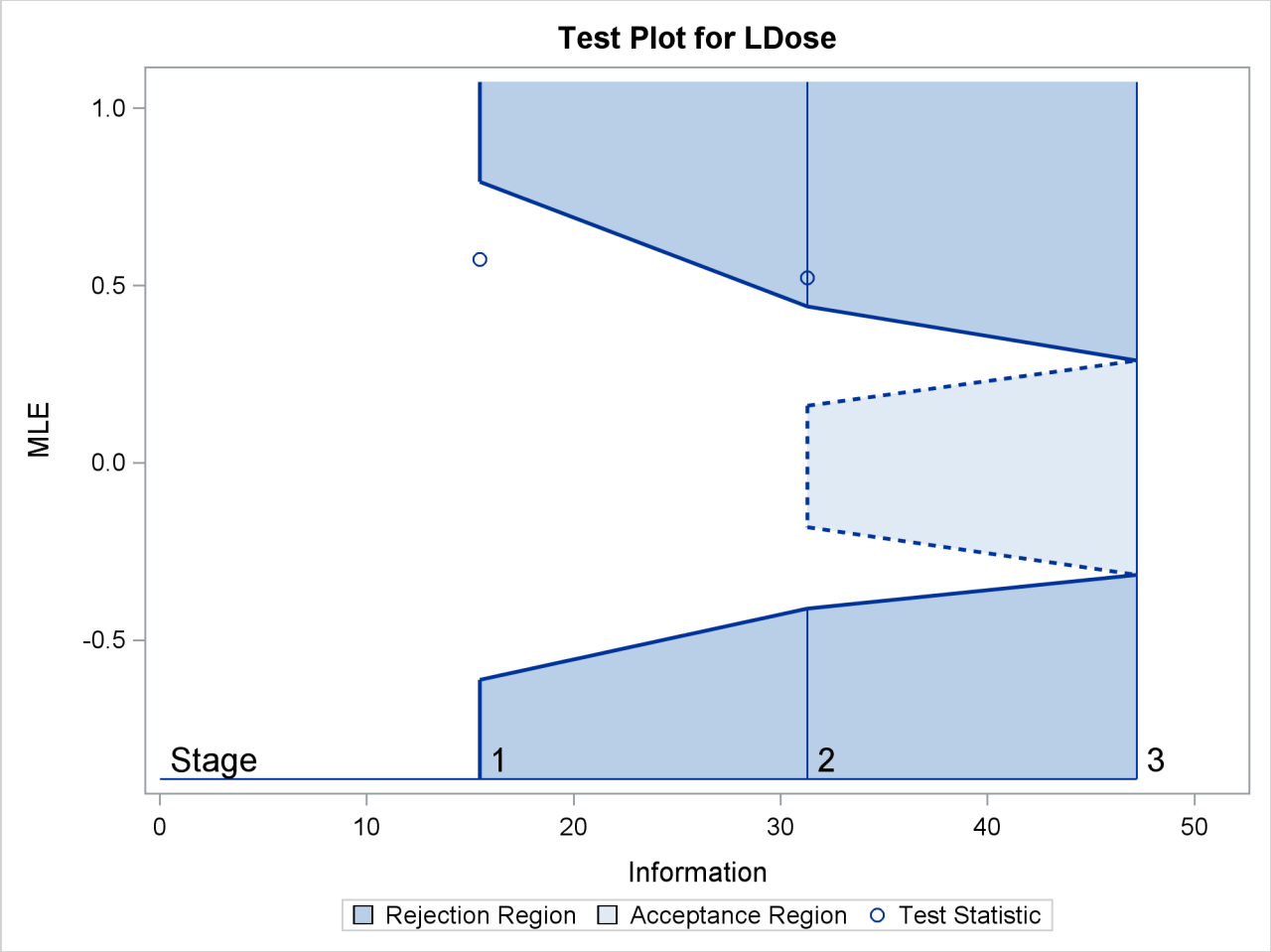
Stage	---Information Level--- Proportion	Actual	-------Alternative------ --------Reference------- Lower	Upper
1	0.3272	15.45062	−0.50000	0.50000
2	0.6624	31.28346	−0.50000	0.50000
3	1.0000	47.22445	−0.50000	0.50000

Test Information (MLE Scale)
Null Reference = 0

	----------Boundary Values---------- ---------Lower---------		---------Upper---------		----------Test--------- ---------LDose---------	
Stage	Alpha	Beta	Beta	Alpha	Estimate	Action
1	−0.61078	.	.	0.79337	0.57409	Continue
2	−0.41028	−0.18112	0.16166	0.44091	0.52128	Reject Null
3	−0.31628	−0.31628	0.28861	0.28861	.	

With ODS Graphics enabled, a boundary plot with the boundary values and test statistics is displayed, as shown in Output 81.8.14. As expected, the test statistic is above the upper α boundary in the upper rejection region at stage 2.

Output 81.8.14 Sequential Test Plot



After a trial is stopped, the "Parameter Estimates" table in Output 81.8.15 displays the stopping stage, parameter estimate, unbiased median estimate, confidence limits, and the p-value under the null hypothesis $H_0 : \beta_1 = 0$.

Output 81.8.15 Parameter Estimates

Parameter Estimates
MLE Ordering

Parameter	Stopping Stage	MLE	p-Value for H0:Parm=0	Median Estimate
LDose	2	0.521275	0.0050	0.502647

Parameter Estimates
MLE Ordering

Parameter	95% Confidence Limits	
LDose	0.15745	0.85154

With the ORDER=MLE option, the MLE ordering is used to compute the p-value, unbiased median estimate, and confidence limits. As expected, the p-value 0.005 is significant at the $\alpha = 0.05$ level and the confidence interval does not contain the null reference zero.

With the RCI option, the "Repeated Confidence Intervals" table in Output 81.8.16 displays repeated confidence intervals for the parameter. For a two-sided test, since the rejection lower repeated confidence limit 0.0804 is greater than the null reference zero, the trial is stopped to reject the hypothesis.

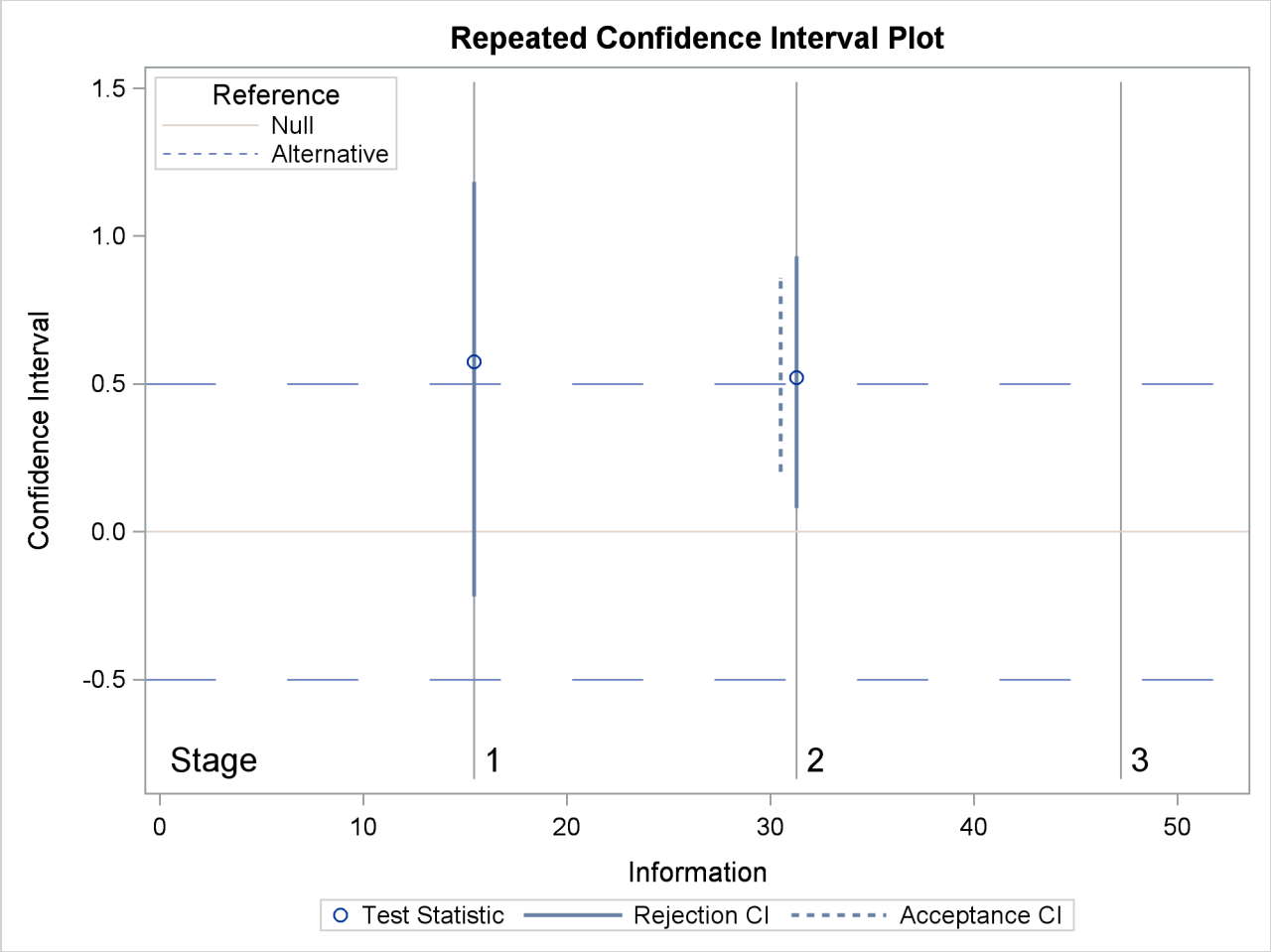
Output 81.8.16 Repeated Confidence Intervals

Repeated Confidence Intervals

Stage	Information Level	Parameter Estimate	-Rejection Boundary- Lower 97.5% Repeated CL	-Rejection Boundary- Upper 97.5% Repeated CL	-Acceptance Boundary- Lower 90.01% Repeated CL	-Acceptance Boundary- Upper 92.13% Repeated CL
1	15.4506	0.57409	-0.2193	1.1849	.	.
2	31.2835	0.52128	0.0804	0.9316	0.2024	0.8596

With the PLOTS=RCI option, the "Repeated Confidence Intervals Plot" displays repeated confidence intervals for the parameter, as shown in Output 81.8.17. It shows that the null reference zero is inside the rejection repeated confidence interval at stage 1 but outside the rejection repeated confidence interval at stage 2. This implies that the study stops at stage 2 to reject the hypothesis.

Output 81.8.17 Repeated Confidence Intervals Plot



Note that the hypothesis is accepted if at any stage, the acceptance repeated confidence interval falls within the interval $(-0.5, 0.5)$ of the alternative references.

References

Chang, M. N. (1989), "Confidence Intervals for a Normal Mean Following a Group Sequential Test," *Biometrics*, 45, 247–254.

Chang, M. N., Gould, A. L., and Snapinn, S. M. (1995), "P-values for Group Sequential Testing," *Biometrika*, 82, 650–654.

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, E. R. (1972), "Regression Models and Life-Tables (with Discussion)," *Journal of the Royal Statistical Society, B*, 34, 187–220.

Cui, L., Hung, H. M. J., and Wang, S. (1999), "Modification of Sample Size in Group Sequential Clinical Trials," *Biometrics*, 55, 853–857.

DeMets, D. L., Furberg, C. D., and Friedman, L. M. (2006), *Data Monitoring in Clinical Trials,* New York: Springer.

Dmitrienko, A., Molenberghs, G., Chuang-Stein, C., and Offen, W. (2005), *Analysis of Clinical Trials Using SAS: A Practical Guide,* Cary, NC: SAS Institute Inc.

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. and Fleming, T. R. (1990), "Parameter Estimation Following Group Sequential Hypothesis Testing," *Biometrika*, 77, 875–892.

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*.

Fairbanks, K. and Madsen, R. (1982), "P Values for Tests Using a Repeated Significance Test Design," *Biometrika*, 69, 69–74.

Food and Drug Administration (1998), "E9: Statistical Principles for Clinical Trials," *Federal Register*, 63 (179), 49583–49598.

Herson, J. (1979), "Predictive Probability Early Termination for Phase II Clinical Trials," *Biometrics*, 35, 775–783.

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. (2000), *Group Sequential Methods with Applications to Clinical Trials,* New York: Chapman & Hall.

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.

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.

Proschan, M. A., Lan, K. K. G., and Wittes, J. T. (2006), *Statistical Monitoring of Clinical Trials,* New York: Springer.

Rosner, G. L. and Tsiatis, A. A. (1988), "Exact Confidence Intervals Following a Group Sequential Trial: A Comparison of Methods," *Biometrika*, 75, 723–729.

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.

Scharfstein, D. O. and Tsiatis, A. A. (1998), "The Use of Simulation and Bootstrap in Information-Based Group Sequential Studies," *Statistics in Medicine*, 17, 75–87.

Tsiatis, A. A. and Mehta C. R. (2003), "On the Inefficiency of the Adaptive Design for Monitoring Clinical Trials," *Biometrika*, 90, 367–378.

Tsiatis, A. A., Rosner G. L., and Mehta C. R. (1984), "Exact Confidence Intervals Following a Group Sequential Test," *Biometrics*, 40, 797–803.

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.

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**.