## Chapter 18 Introduction to Power and Sample Size Analysis

Contents
Overview ..... 365
Coverage of Statistical Analyses ..... 366
Statistical Background ..... 367
Hypothesis Testing, Power, and Confidence Interval Precision ..... 367
Standard Hypothesis Tests ..... 367
Equivalence and Noninferiority ..... 367
Confidence Interval Precision ..... 368
Computing Power and Sample Size ..... 368
Power and Study Planning ..... 369
Components of Study Planning ..... 369
Effect Size ..... 370
Uncertainty and Sensitivity Analysis ..... 371
SAS/STAT Tools for Power and Sample Size Analysis ..... 371
Basic Graphs (POWER, GLMPOWER, Power and Sample Size Application) ..... 372
Highly Customized Graphs (POWER, GLMPOWER) ..... 374
Formatted Tables (\%POWTABLE Macro) ..... 375
Narratives and Graphical User Interface (Power and Sample Size Application) ..... 377
Customized Power Formulas (DATA Step) ..... 379
Empirical Power Simulation (DATA Step, SAS/STAT Software) ..... 380
References ..... 381

## Overview

Power and sample size analysis optimizes the resource usage and design of a study, improving chances of conclusive results with maximum efficiency. The standard statistical testing paradigm implicitly assumes that Type I errors (mistakenly concluding significance when there is no true effect) are more costly than Type II errors (missing a truly significant result). This may be appropriate for your situation, or the relative costs of the two types of error may be reversed. For example, in screening experiments for drug development, it is often less damaging to carry a few false positives forward for follow-up testing than to miss potential leads. Power and sample size analysis can help you achieve your desired balance between Type I and Type II errors. With optimal designs and sample sizes, you can improve your chances of detecting effects that might otherwise have been ignored, save money and time, and perhaps minimize risks to subjects.

Relevant tools in SAS/STAT software for power and sample size analysis include the following:

- the GLMPOWER procedure
- the POWER procedure
- the Power and Sample Size application
- the \%POWTABLE macro
- various procedures, statements, and functions in Base SAS and SAS/STAT for developing customized formulas and simulations

These tools, discussed in detail in the section "SAS/STAT Tools for Power and Sample Size Analysis" on page 371 , deal exclusively with prospective analysis-that is, planning for a future study. This is in contrast to retrospective analysis for a past study, which is not supported by the main tools. Although retrospective analysis is more convenient to perform, it is often uninformative or misleading, especially when power is computed directly based on observed data.

The goals of prospective power and sample size analysis include the following:

- determining the sample size required to get a significant result with adequate probability (power)
- characterizing the power of a study to detect a meaningful effect
- computing the probability of achieving the desired precision of a confidence interval, or the sample size required to ensure this probability
- conducting what-if analyses to assess how sensitive the power or required sample size is to other factors

The phrase power analysis is used for the remainder of this document as a shorthand to represent any or all of these goals. For more information about the GLMPOWER procedure, see Chapter 47, "The GLMPOWER Procedure." For more information about the POWER procedure, see Chapter 77, "The POWER Procedure." For more information about the Power and Sample Size application, see Chapter 78, "The Power and Sample Size Application."

## Coverage of Statistical Analyses

The GLMPOWER procedure covers power analysis for Type III $F$ tests and contrasts of fixed effects in univariate and multivariate linear models. For univariate models, you can specify covariates, which can be continuous or categorical. For multivariate models, you can choose among Wilks' likelihood ratio, Hotelling-Lawley trace, and Pillai's trace $F$ tests for multivariate analysis of variance (MANOVA) and among uncorrected, Greenhouse-Geisser, Huynh-Feldt, and Box conservative $F$ tests for the univariate approach to repeated measures. Tests and contrasts that involve random effects are not supported.

The POWER procedure covers power analysis for the following:

- $t$ tests, equivalence tests, and confidence intervals for means
- tests, equivalence tests, and confidence intervals for binomial proportions
- multiple regression
- tests of correlation and partial correlation
- one-way analysis of variance
- rank tests for comparing two survival curves
- logistic regression with binary response
- Wilcoxon Mann-Whitney rank-sum test

The Power and Sample Size application covers a large subset of the analyses in the GLMPOWER and POWER procedures.

## Statistical Background

## Hypothesis Testing, Power, and Confidence Interval Precision

## Standard Hypothesis Tests

In statistical hypothesis testing, you typically express the belief that some effect exists in a population by specifying an alternative hypothesis $H_{1}$. You state a null hypothesis $H_{0}$ as the assertion that the effect does not exist and attempt to gather evidence to reject $H_{0}$ in favor of $H_{1}$. Evidence is gathered in the form of sample data, and a statistical test is used to assess $H_{0}$. If $H_{0}$ is rejected but there really is no effect, this is called a Type I error. The probability of a Type I error is usually designated "alpha" or $\alpha$, and statistical tests are designed to ensure that $\alpha$ is suitably small (for example, less than 0.05 ).
If there is an effect in the population but $H_{0}$ is not rejected in the statistical test, then a Type II error has been committed. The probability of a Type II error is usually designated "beta" or $\beta$. The probability $1-\beta$ of avoiding a Type II error (that is, correctly rejecting $H_{0}$ and achieving statistical significance) is called the power of the test.
Most, but not all, of the power analyses in the GLMPOWER and POWER procedures are based on such standard hypothesis tests.

## Equivalence and Noninferiority

Whereas the standard two-sided hypothesis test for a parameter $\mu$ (such as a mean difference) aims to demonstrate that it is significantly different than a null value $\mu_{0}$ :

$$
\begin{aligned}
& H_{0}: \mu=\mu_{0} \\
& H_{1}: \mu \neq \mu_{0}
\end{aligned}
$$

an equivalence test instead aims to demonstrate that it is significantly similar to some value, expressed in terms of a range $\theta_{L}, \theta_{U}$ around that value:

$$
\begin{aligned}
& H_{0}: \mu<\theta_{L} \text { or } \mu>\theta_{U} \\
& H_{1}: \theta_{L} \leq \mu \leq \theta_{U}
\end{aligned}
$$

Whereas the standard one-sided hypothesis test for $\mu$ (say, the upper one-sided test) aims to demonstrate that it is significantly greater than $\mu_{0}$ :

$$
\begin{aligned}
& H_{0}: \mu \leq \mu_{0} \\
& H_{1}: \mu>\mu_{0}
\end{aligned}
$$

a corresponding noninferiority test aims to demonstrate that it is not significantly less than $\mu_{0}$, expressed in terms of a margin $\delta>0$ :

$$
\begin{aligned}
& H_{0}: \mu \leq \mu_{0}-\delta \\
& H_{1}: \mu>\mu_{0}-\delta
\end{aligned}
$$

Corresponding forms of these hypotheses with the inequalities reversed apply to lower one-sided noninferiority tests (sometimes called nonsuperiority tests).

The POWER procedure performs power analyses for equivalence tests for one-sample, paired, and twosample tests of normal and lognormal mean differences and ratios. It also supports noninferiority tests for a variety of analyses of means, proportions, and correlation, both directly (with a MARGIN= option representing $\delta$ ) and indirectly (with an option for a custom null value representing the sum or difference of $\mu_{0}$ and $\delta$ ).

## Confidence Interval Precision

An analysis of confidence interval precision is analogous to a traditional power analysis, with CI Half-Width taking the place of effect size and Prob(Width) taking the place of power. The CI Half-Width is the margin of error associated with the confidence interval, the distance between the point estimate and an endpoint. The $\operatorname{Prob}($ Width $)$ is the probability of obtaining a confidence interval with at most a target half-width.

The POWER procedure performs confidence interval precision analyses for $t$-based confidence intervals for one-sample, paired, and two-sample designs, and for several varieties of confidence intervals for a binomial proportion.

## Computing Power and Sample Size

For some statistical models and tests, power analysis calculations are exact-that is, they are based on a mathematically accurate formula that expresses power in terms of the other components. Such formulas typically involve either enumeration or noncentral versions of the distribution of the test statistic.

When a power computation is based on a noncentral $t, F$, or chi-square distribution, the noncentrality parameter generally has the same form as the test statistic, with the conjectured population parameters in place of their corresponding estimators.

For example, the test statistic for a two-sample $t$ test is computed as follows:

$$
t=N^{\frac{1}{2}}\left(w_{1} w_{2}\right)^{\frac{1}{2}}\left(\frac{\bar{x}_{2}-\bar{x}_{1}-\mu_{0}}{s_{p}}\right)
$$

where $N$ is the total sample size, $w_{1}$ and $w_{2}$ are the group allocation weights, $\bar{x}_{1}$ and $\bar{x}_{2}$ are the sample means, $\mu_{0}$ is the null mean difference, and $s_{p}$ is the pooled standard deviation. Under the null hypothesis, the
statistic $F=t^{2}$ is distributed as $F(1, N-2)$. In general, $F$ has a noncentral $F$ distribution $F\left(1, N-2, \delta^{2}\right)$ where

$$
\delta=N^{\frac{1}{2}}\left(w_{1} w_{2}\right)^{\frac{1}{2}}\left(\frac{\mu_{\mathrm{diff}}-\mu_{0}}{\sigma}\right)
$$

and $\mu_{\text {diff }}$ and $\sigma$ are the (unknown) true mean difference and common group standard deviation, respectively. Note that the square-root noncentrality $\delta$ is exactly the same as the $t$ statistic except that the estimators of mean difference and standard deviation are replaced by their corresponding true population values.

The power for the two-sided two-sample $t$ test with significance level $\alpha$ is computed as

$$
P\left(F \geq F_{1-\alpha}(1, N-2)\right)
$$

where $F$ is distributed as $F\left(1, N-2, \delta^{2}\right)$ and $F_{1-\alpha}(1, N-2)$ is the $100(1-\alpha) \%$ quantile of the central $F$ distribution with 1 and $N-2$ degrees of freedom. See the section "Customized Power Formulas (DATA Step)" on page 379 for an example of the implementation of this formula in the DATA step.

In the absence of exact mathematical results, approximate formulas can sometimes be used. When neither exact power computations nor reasonable approximations are possible, simulation provides an increasingly viable alternative. You specify values for model parameters and use them to randomly generate a large number of hypothetical data sets. Applying the statistical test to each data set, you estimate power with the percentage of times the null hypothesis is rejected. While the simulation approach is computationally intensive, faster computing makes this less of an issue. A simulation-based power analysis is always a valid option, and, with a large number of data set replications, it can often be more accurate than approximations. See the section "Empirical Power Simulation (DATA Step, SAS/STAT Software)" on page 380 for an example of an empirical power simulation.

Sample size is usually computed by iterative numerical methods because it often cannot be expressed in closed form as a function of the other parameters. Sample size tends to appear in both a noncentrality parameter and a degrees of freedom term for the critical value.

## Power and Study Planning

Power analysis is most effective when performed at the study planning stage, and as such it encourages early collaboration between researcher and statistician. It also focuses attention on effect sizes and variability in the underlying scientific process, concepts that both researcher and statistician should consider carefully at this stage.

There are many factors involved in a power analysis, such as the research objective, design, data analysis method, power, sample size, Type I error, variability, and effect size. By performing a power analysis, you can learn about the relationships between these factors, optimizing those that are under your control and exploring the implications of those that are fixed or unknown.

## Components of Study Planning

Even when the research questions and study design seem straightforward, the ensuing power analysis can seem technically daunting. It is often helpful to break the process down into five components:

- Study Design: What is the structure of the planned design? This must be clearly and completely specified. What groups and treatments ("cells" and "factors" of the design) are going to be assessed, and what will be the relative sizes of those cells? How is each case going to be studied-that is, what is the primary outcome measure ("dependent variable")? Will covariates be measured and included in the statistical model?
- Scenario Model: What are your beliefs about patterns in the data? Imagine that you had unlimited time and resources to execute the study design, so that you could gather an "infinite data set." Characterize that infinite data set as best you can using a mathematical model, realizing that it will be a simplification of reality. Alternatively, as is common with complex linear models, you may decide to construct an "exemplary" data set that mimics the infinite data set. However you do this, your scenario model should capture the key features of the study design and the main relationships among the primary outcome variables and study factors.
- Effects and Variability: What exactly are the "signals and noises" in the patterns you suspect? Set specific values for the parameters of your scenario model, keeping at most one unspecified. It is often enlightening to consider a variety of realistic possibilities for the key values by performing a sensitivity analysis, to explore the consequences of competing views on what the infinite data set might look like.
- Statistical Method: How will you cast your model in statistical terms and conduct the eventual data analysis? Define the statistical models and procedures that will be used to embody the study design and estimate and/or test the effects central to the research question. What significance levels will be used? Will one- or two-sided tests be used?
- Aim of Assessment: Finally, what needs to be determined in the power analysis? Most often you want to examine the statistical powers obtained across the various scenarios for the effects, variability, alternative varieties of the statistical procedures to be used, and the feasible total sample sizes. Sometimes the goal is to find sample size values that provide given levels of power, say $85 \%, 90 \%$, or $95 \%$.


## Effect Size

There is some confusion in practice about how to postulate the effect size. One alternative is to specify the effect size that represents minimal clinical significance; then the result of the power analysis reveals the chances of detecting a minimally meaningful effect size. Often this minimal effect size is so small that it requires excessive resources to detect. Another alternative is to make an educated guess of the true underlying effect size. Then the power analysis determines the chance of detecting the effect size that is believed to be true. The choice is ultimately determined by the research goals. Finally, you can specify a collection of possible values, perhaps spanning the range between minimally meaningful effects and larger surmised effects.

You can arrive at values for required quantities in a power analysis, such as effect sizes and measures of variability, in many different ways. For example, you can use pilot data, results of previous studies reported in literature, educated guesses derived from theory, or educated guesses derived from partial data (a small sample or even just quantiles).

## Uncertainty and Sensitivity Analysis

Uncertainty is a fact of life in any power analysis, because at least some of the numbers used are best guesses of unknown values. The result of a power calculation, whether it be achieved power or required sample size or something else, serves only as a point estimate, conditional on the conjectured values of the other components. It is not feasible in general to quantify the variability involved in using educated guesses or undocumented results to specify these components. If observed data are used, relevant adjustments for variability in the data tend to be problematic in the sense of producing confidence intervals for power that are too wide for practical use. But there is a useful way for you to characterize the uncertainty in your power analysis, and also discover the extent to which statistical power is affected by each component. You can posit a reasonable range for each input component, vary each one within its range, and observe the variety of results in the form of tables or graphs.

## SAS/STAT Tools for Power and Sample Size Analysis

This section demonstrates how you can use the different SAS power analysis tools mentioned in the section "Overview" on page 365 to generate graphs, tables, and narratives; implement your own power formulas; and simulate empirical power.

Suppose you want to compute the power of a two-sample $t$ test. You conjecture that the mean difference is between 5 and 6 and that the common group standard deviation is between 12 and 18 . You plan to use a significance level between 0.05 and 0.1 and a sample size between 100 and 200. The following SAS statements use the POWER procedure to compute the power for these scenarios:

```
proc power;
    twosamplemeans test=diff
        meandiff = 5 6
        stddev = 12 18
        alpha = 0.05 0.1
        ntotal = 100 200
        power = .;
run;
```

Figure 18.1 shows the results. Depending on the plausibility of the various combinations of input parameter values, the power ranges between 0.379 and 0.970 .

Figure 18.1 PROC POWER Tabular Output
The POWER Procedure
Two-Sample t Test for Mean Difference

| Computed Power |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Alpha | Mean Diff |  | $\underset{\text { Total }}{\mathrm{N}}$ | Power |
| 1 | 0.05 | 5 | 12 | 100 | 0.541 |
| 2 | 0.05 | 5 | 12 | 200 | 0.834 |
| 3 | 0.05 | 5 | 18 | 100 | 0.280 |
| 4 | 0.05 | 5 | 18 | 200 | 0.498 |
| 5 | 0.05 | 6 | 12 | 100 | 0.697 |
| 6 | 0.05 | 6 | 12 | 200 | 0.940 |
| 7 | 0.05 | 6 | 18 | 100 | 0.379 |
| 8 | 0.05 | 6 | 18 | 200 | 0.650 |
| 9 | 0.10 | 5 | 12 | 100 | 0.664 |
| 10 | 0.10 | 5 | 12 | 200 | 0.902 |
| 11 | 0.10 | 5 | 18 | 100 | 0.397 |
| 12 | 0.10 | 5 | 18 | 200 | 0.623 |
| 13 | 0.10 | 6 | 12 | 100 | 0.799 |
| 14 | 0.10 | 6 | 12 | 200 | 0.970 |
| 15 | 0.10 | 6 | 18 | 100 | 0.505 |
| 16 | 0.10 | 6 | 18 | 200 | 0.759 |

The following seven sections illustrate additional ways of displaying these results using the different SAS tools.

## Basic Graphs (POWER, GLMPOWER, Power and Sample Size Application)

If you include a PLOT statement, the GLMPOWER and POWER procedures produce standard power curves, which represent any multivalued input parameters with varying line styles, symbols, colors, and/or panels. The Power and Sample Size application also has an option to produce power curves. If ODS Graphics is enabled, then graphs are created using ODS Graphics; otherwise, traditional graphs are produced.

To display default power curves for the preceding PROC POWER call, add the PLOT statement with no arguments as follows:

```
ods graphics on;
proc power plotonly;
    twosamplemeans test=diff
        meandiff = 5 6
        stddev = 12 18
        alpha = 0.05 0.1
        ntotal = 100 200
        power = .;
    plot;
run;
ods graphics off;
```

The ODS GRAPHICS ON statement enables ODS Graphics.
Figure 18.2 shows the results. Note that the line style varies by the significance level $\alpha$, the symbol varies by the mean difference, and the panel varies by standard deviation.

Figure 18.2 PROC POWER Default Graphical Output


Figure 18.2 continued


## Highly Customized Graphs (POWER, GLMPOWER)

Example 77.8 of Chapter 77, "The POWER Procedure," demonstrates various ways you can modify and enhance plots created in the GLMPOWER or POWER procedures:

- assigning analysis parameters to axes
- fine-tuning a sample size axis
- adding reference lines
- linking plot features to analysis parameters
- choosing key (legend) styles
- modifying symbol locations

For example, replace the default PLOT statement with the following statement to modify the graphical results in Figure 18.2 to lower the minimum sample size to 60 , show a reference line at power= 0.9 with
corresponding sample size values, distinguish standard deviation by color instead of panel, and swap the roles of $\alpha$ and mean difference:

```
plot
    min=60
    yopts=(ref=0.9 crossref=yes)
    vary(color by stddev, linestyle by meandiff, symbol by alpha);
```

Figure 18.3 shows the results. The plot reveals that only the scenarios with the largest mean difference and smallest standard deviation achieve a power of at least 0.9 for this sample size range.

Figure 18.3 PROC POWER Customized Graphical Output


## Formatted Tables (\%POWTABLE Macro)

The \%POWTABLE macro renders the output of the POWER and GLMPOWER procedures in rectangular form, and it optionally produces simplified results using weighted means across chosen variables. PROC REPORT and the Output Delivery System (ODS) are used to generate the tables. Base SAS and SAS/STAT 9.1 or higher versions are required.

You can run the \%POWTABLE macro for the output in Figure 18.1 to display the results in a form more suitable for quickly discerning relationships among parameters. First use the ODS OUTPUT statement to assign the "Output" table produced by the POWER procedure to a data set as follows:

```
ods output output=powdata;
```

Next, specify the same PROC POWER statements that generate Figure 18.1. Finally, use the \%POWTABLE macro to assign analysis parameters to table dimensions. To create a table of computed power values with mean difference assigned to rows, sample size and $\alpha$ assigned to columns, and standard deviation assigned to "panels" (rendered by default as rows separated by blank lines), specify the following statements:

```
%powtable ( Data = powdata,
    Entries = power,
    Rows = meandiff,
    Cols = ntotal alpha,
    Panels = stddev )
```

Figure 18.4 shows the results.
Figure 18.4 \%POWTABLE Macro Output

| The POWTABLE Macro |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Entries are Power |  |  |  |  |  |
|  |  | N Total |  |  |  |
|  |  | $\begin{gathered} 100 \\ \text { Alpha } \end{gathered}$ |  | $\begin{gathered} 200 \\ \text { Alpha } \end{gathered}$ |  |
|  |  |  |  |  |  |
|  |  | 0.05 | 0.10 | 0.05 | 0.10 |
| Std Mean Dev Diff |  |  |  |  |  |
| -- | -- | -- | -- | -- | -- |
| 12 |  | 50.541 | 0.664 | 0.834 | 0.902 |
|  |  | 60.697 | 0.799 | 0.940 | 0.970 |
| 18 |  | 50.280 | 0.397 | 0.498 | 0.623 |
|  |  | 60.379 | 0.505 | 0.650 | 0.759 |

## Narratives and Graphical User Interface (Power and Sample Size Application)

The Power and Sample Size application produces narratives for the results. Narratives are descriptions of the input parameters and a statement about the computed power or sample size.

For example, the Power and Sample Size application creates the following narrative for the scenario corresponding to the first row in Figure 18.1:
"For a two-sample pooled $t$ test of a normal mean difference with a two-sided significance level of 0.05 , assuming a common standard deviation of 12 , a total sample size of 100 assuming a balanced design has a power of 0.541 to detect a mean difference of 5 ."

The Power and Sample Size application also provides multiple input parameter options, stores the results in a project format, displays power curves, and shows the SAS log and SAS code. You can access each project to review the results or to edit your input parameters and produce another analysis.

Where appropriate, several alternate ways of entering values for certain parameters are offered. For example, in the two-sample $t$ test analysis, sample sizes can be entered in any of several parameterizations:

- total sample size in a balanced design
- sample size per group in a balanced design
- total sample size and group allocation weights
- groupwise sample sizes

See Figure 18.5 for an illustration of the application, showing the sample size input page for a two-sample $t$ test.

Figure 18.5 Power and Sample Size Application


## Customized Power Formulas (DATA Step)

If you want to perform a power computation for an analysis that is not currently supported directly in SAS/STAT tools, and you have a power formula, then you can program the formula in the DATA step.

For purposes of illustration, here is the power formula in the section "Computing Power and Sample Size" on page 368 implemented in the DATA step to compute power for the $t$ test example:

```
data tpow;
    do meandiff = 5, 6;
        do stddev = 12, 18;
            do alpha = 0.05, 0.1;
                do ntotal = 100, 200;
                ncp = ntotal * 0.5 * 0.5 * meandiff**2 / stddev**2;
                critval = finv(1-alpha, 1, ntotal-2, 0);
                power = sdf('f', critval, 1, ntotal-2, ncp);
                output;
                end;
            end;
        end;
    end;
run;
proc print data=tpow;
run;
```

The output is shown in Figure 18.6.
Figure 18.6 Customized Power Formula (DATA Step)

| Obs meandiff stddev alpha ntotal |  |  |  |  | ncp | critval | power |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | 5 | 12 | 0.05 | 100 | 4.3403 | 3.93811 | 0.54102 |
| 2 | 5 | 12 | 0.05 | 200 | 8.6806 | 3.88885 | 0.83447 |
| 3 | 5 | 12 | 0.10 | 100 | 4.3403 | 2.75743 | 0.66434 |
| 4 | 5 | 12 | 0.10 | 200 | 8.6806 | 2.73104 | 0.90171 |
| 5 | 5 | 18 | 0.05 | 100 | 1.9290 | 3.93811 | 0.27981 |
| 6 | 5 | 18 | 0.05 | 200 | 3.8580 | 3.88885 | 0.49793 |
| 7 | 5 | 18 | 0.10 | 100 | 1.9290 | 2.75743 | 0.39654 |
| 8 | 5 | 18 | 0.10 | 200 | 3.8580 | 2.73104 | 0.62287 |
| 9 | 6 | 12 | 0.05 | 100 | 6.2500 | 3.93811 | 0.69689 |
| 10 | 6 | 12 | 0.05 | 200 | 12.5000 | 3.88885 | 0.94043 |
| 11 | 6 | 12 | 0.10 | 100 | 6.2500 | 2.75743 | 0.79895 |
| 12 | 6 | 12 | 0.10 | 200 | 12.5000 | 2.73104 | 0.96985 |
| 13 | 6 | 18 | 0.05 | 100 | 2.7778 | 3.93811 | 0.37857 |
| 14 | 6 | 18 | 0.05 | 200 | 5.5556 | 3.88885 | 0.65012 |
| 15 | 6 | 18 | 0.10 | 100 | 2.7778 | 2.75743 | 0.50459 |
| 16 | 6 | 18 | 0.10 | 200 | 5.5556 | 2.73104 | 0.75935 |

## Empirical Power Simulation (DATA Step, SAS/STAT Software)

You can obtain a highly accurate power estimate by simulating the power empirically. You need to use this approach for analyses that are not supported directly in SAS/STAT tools and for which you lack a power formula. But the simulation approach is also a viable alternative to existing power approximations. A high number of simulations will yield a more accurate estimate than a non-exact power approximation.

Although exact power computations for the two-sample $t$ test are supported in several of the SAS/STAT tools, suppose for purposes of illustration that you want to simulate power for the continuing $t$ test example. This section describes how you can use the DATA step and SAS/STAT software to do this.

The simulation involves generating a large number of data sets according to the distributions defined by the power analysis input parameters, computing the relevant $p$-value for each data set, and then estimating the power as the proportion of times that the $p$-value is significant.

The following statements compute a power estimate along with a $95 \%$ confidence interval for power for the first scenario in the two-sample $t$ test example, with 10,000 simulations:

```
%let meandiff = 5;
%let stddev = 12;
%let alpha = 0.05;
%let ntotal = 100;
%let nsim = 10000;
data simdata;
    call streaminit(123);
    do isim = 1 to &nsim;
        do i = 1 to floor(&ntotal/2);
            group = 1;
            y = rand('normal', 0 , &stddev);
            output;
            group = 2;
            y = rand('normal', &meandiff, &stddev);
            output;
        end;
    end;
run;
ods listing close;
proc ttest data=simdata;
    ods output ttests=tests;
    by isim;
    class group;
    var y;
run;
ods listing;
data tests;
    set tests;
    where method="Pooled";
    issig = probt < &alpha;
run;
```

```
proc freq data=tests;
    ods select binomial;
    tables issig / binomial(level='1');
run;
```

First the DATA step is used to randomly generate $n \operatorname{sim}=10,000$ data sets based on the meandiff, stddev, and ntotal parameters and the normal distribution, consistent with the assumptions underlying the two-sample $t$ test. These data sets are contained in a large SAS data set called simdata indexed by the variable isim.

The CALL STREAMINIT(123) statement initializes the random number generator with a specific sequence and ensures repeatable results for purposes of this example. (NOTE: Skip this step when you are performing actual power simulations.)

The TTEST procedure is run using isim as a BY variable, with the ODS LISTING CLOSE statement to suppress output. The ODS OUTPUT statement saves the "TTests" table to a data set called tests. The $p$-values are contained in a column called probt.

The subsequent DATA step defines a variable called issig to flag the significant $p$-values.
Finally, the FREQ procedure computes the empirical power estimate as the estimate of $\mathrm{P}(\mathrm{issig}=1)$ and provides approximate and exact confidence intervals for this estimate.

Figure 18.7 shows the results. The estimated power is 0.5388 with $95 \%$ confidence interval ( $0.5290,0.5486$ ). Note that the exact power of 0.541 shown in the first row in Figure 18.1 is contained within this tight confidence interval.

Figure 18.7 Simulated Power (DATA Step, SAS/STAT Software)
The FREQ Procedure

| Binomial Proportion <br> issig = 1 |  |
| :--- | ---: |
| Proportion | 0.5388 |
| ASE | 0.0050 |
| 95\% Lower Conf Limit | 0.5290 |
| 95\% Upper Conf Limit | 0.5486 |
|  |  |
| Exact Conf Limits |  |
| 95\% Lower Conf Limit | 0.5290 |
| 95\% Upper Conf Limit | 0.5486 |

## References

Castelloe, J. M. (2000), "Sample Size Computations and Power Analysis with the SAS System," in Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc.

Castelloe, J. M. and O'Brien, R. G. (2001), "Power and Sample Size Determination for Linear Models," in Proceedings of the Twenty-Sixth Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc.

Lenth, R. V. (2001), "Some Practical Guidelines for Effective Sample Size Determination," American Statistician, 55, 187-193.

Muller, K. E. and Benignus, V. A. (1992), "Increasing Scientific Power with Statistical Power," Neurotoxicology and Teratology, 14, 211-219.

O’Brien, R. G. and Castelloe, J. M. (2007), "Sample-Size Analysis for Traditional Hypothesis Testing: Concepts and Issues," in A. Dmitrienko, C. Chuang-Stein, and R. D'Agostino, eds., Pharmaceutical Statistics Using SAS: A Practical Guide, 237-271, Cary, NC: SAS Institute Inc.

O’Brien, R. G. and Muller, K. E. (1993), "Unified Power Analysis for $t$-Tests through Multivariate Hypotheses," in L. K. Edwards, ed., Applied Analysis of Variance in Behavioral Science, 297-344, New York: Marcel Dekker.

## Chapter 47 The GLMPOWER Procedure

## Contents

Overview: GLMPOWER Procedure ..... 3598
Getting Started: GLMPOWER Procedure ..... 3599
Simple Two-Way ANOVA ..... 3599
Incorporating Contrasts, Unbalanced Designs, and Multiple Means Scenarios ..... 3603
Syntax: GLMPOWER Procedure ..... 3605
PROC GLMPOWER Statement ..... 3606
BY Statement ..... 3607
CLASS Statement ..... 3608
CONTRAST Statement ..... 3608
MANOVA Statement ..... 3609
MODEL Statement ..... 3611
PLOT Statement ..... 3611
POWER Statement ..... 3616
REPEATED Statement ..... 3623
WEIGHT Statement ..... 3626
Details: GLMPOWER Procedure ..... 3626
Specifying Value Lists in the POWER Statement ..... 3626
Number-Lists ..... 3627
Name-Lists ..... 3627
Keyword-Lists ..... 3627
Sample Size Adjustment Options ..... 3627
Error and Information Output ..... 3628
Displayed Output ..... 3629
ODS Table Names ..... 3630
Computational Methods and Formulas ..... 3630
Contrasts in Fixed-Effect Univariate Models ..... 3630
Adjustments for Covariates in Univariate Models ..... 3632
Contrasts in Fixed-Effect Multivariate Models ..... 3633
ODS Graphics ..... 3641
Examples: GLMPOWER Procedure ..... 3641
Example 47.1: One-Way ANOVA ..... 3641
Example 47.2: Two-Way ANOVA with Covariate ..... 3647
Example 47.3: Repeated Measures ANOVA ..... 3653
References ..... 3657

## Overview: GLMPOWER Procedure

Power and sample size analysis optimizes the resource usage and design of a study, improving chances of conclusive results with maximum efficiency. The GLMPOWER procedure performs prospective power and sample size analysis for linear models, with a variety of goals:

- determining the sample size required to get a significant result with adequate probability (power)
- characterizing the power of a study to detect a meaningful effect
- conducting what-if analyses to assess sensitivity of the power or required sample size to other factors

Here prospective indicates that the analysis pertains to planning for a future study. This is in contrast to retrospective analysis for a past study, which is not supported by this procedure.

The statistical analyses that are covered include Type III $F$ tests and contrasts of fixed effects in univariate and multivariate linear models. For univariate models, you can specify covariates, which can be continuous or categorical. For multivariate models, you can choose among Wilks' likelihood ratio, Hotelling-Lawley trace, and Pillai's trace $F$ tests for multivariate analysis of variance (MANOVA) and among uncorrected, Greenhouse-Geisser, Huynh-Feldt, and Box conservative $F$ tests for the univariate approach to repeated measures. Tests and contrasts that involve random effects are not supported. For power and sample size analyses in a variety of other statistical situations, see Chapter 77, "The POWER Procedure."

Input for PROC GLMPOWER includes the following components, which are considered in study planning:

- design (including subject profiles and their allocation weights)
- statistical model and test
- between-subject contrasts of class effects
- within-subject contrasts (for multivariate models)
- significance level (alpha)
- surmised response means for subject profiles (often called "cell means")
- surmised variability (and correlation for multivariate models)
- power
- sample size

In order to identify power or sample size as the result parameter, you designate it by a missing value in the input. The procedure calculates this result value over one or more scenarios of input values for all other components.

You specify the design and the cell means by using an exemplary data set, a data set of artificial values that is constructed to represent the intended sampling design and the surmised response means in the underlying population. You specify the model and between-subject contrasts by using MODEL and CONTRAST statements similar to those in the GLM, ANOVA, and MIXED procedures. For multivariate models, you
specify the within-subject contrasts by using MANOVA and REPEATED statements similar to those in the GLM and MIXED procedures. You specify the remaining parameters by using the POWER statement, which is similar to analysis statements in the POWER procedure.

In addition to tabular results, PROC GLMPOWER produces graphs. You can produce the most common types of plots easily with default settings and use a variety of options for more customized graphics. For example, you can control the choice of axis variables, axis ranges, number of plotted points, mapping of graphical features (such as color, line style, symbol, and panel) to analysis parameters, and legend appearance.

If ODS Graphics is enabled, then PROC GLMPOWER uses ODS Graphics to create graphs; otherwise, traditional graphs are produced.

For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 606 in Chapter 21, "Statistical Graphics Using ODS."

For specific information about the statistical graphics and options available with the GLMPOWER procedure, see the PLOT statement and the section "ODS Graphics" on page 3641.

The GLMPOWER procedure is one of several tools available in SAS/STAT software for power and sample size analysis. PROC POWER covers a variety of other analyses such as $t$ tests, equivalence tests, confidence intervals, binomial proportions, multiple regression, one-way ANOVA, survival analysis, logistic regression, and the Wilcoxon rank-sum test. The Power and Sample Size application provides a user interface and implements many of the analyses supported in the procedures. See Chapter 77, "The POWER Procedure," and Chapter 78, "The Power and Sample Size Application," for details.

The following sections of this chapter describe how to use PROC GLMPOWER and discuss the underlying statistical methodology. The section "Getting Started: GLMPOWER Procedure" on page 3599 introduces PROC GLMPOWER with examples of power computation for a two-way analysis of variance. The section "Syntax: GLMPOWER Procedure" on page 3605 describes the syntax of the procedure. The section "Details: GLMPOWER Procedure" on page 3626 summarizes the methods employed by PROC GLMPOWER and provides details on several special topics. The section "Examples: GLMPOWER Procedure" on page 3641 illustrates the use of the GLMPOWER procedure with several applications.

For an overview of methodology and SAS tools for power and sample size analysis, see Chapter 18, "Introduction to Power and Sample Size Analysis." For more discussion and examples for linear models, see Castelloe and O'Brien (2001); O'Brien and Shieh (1992); Muller and Benignus (1992); O'Brien and Muller (1993). For additional discussion of general power and sample size concepts, see O'Brien and Castelloe (2007); Castelloe (2000); Muller and Benignus (1992); Lenth (2001).

## Getting Started: GLMPOWER Procedure

## Simple Two-Way ANOVA

This example demonstrates how to use PROC GLMPOWER to compute and plot power for each effect test in a two-way analysis of variance (ANOVA).

Suppose you are planning an experiment to study the effect of light exposure at three levels on the growth of two varieties of flowers. The planned data analysis is a two-way ANOVA with flower height (measured
at two weeks) as the response and a model consisting of the effects of light exposure, flower variety, and their interaction. You want to calculate the power of each effect test for a balanced design with a total of 60 specimens ( 10 for each combination of exposure and variety) with $\alpha=0.05$ for each test.

As a first step, create an exemplary data set describing your conjectures about the underlying population means. You believe that the mean flower height for each combination of variety and exposure level (that is, for each design profile, or for each cell in the design) roughly follows Table 47.1.

Table 47.1 Mean Flower Height (in cm) by Variety and Exposure

|  | Exposure |  |  |
| :---: | :---: | :---: | :---: |
| Variety | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ |
| 1 | 14 | 16 | 21 |
| 2 | 10 | 15 | 16 |

The following statements create a data set named Exemplary containing these cell means.

```
data Exemplary;
    do Variety = 1 to 2;
        do Exposure = 1 to 3;
            input Height @@;
            output;
        end;
    end;
    datalines;
        14 16 21
        10 15 16
;
```

You also conjecture that the error standard deviation is about 5 cm .
Use the DATA= option in the PROC GLMPOWER statement to specify Exemplary as the exemplary data set. Identify the classification variables (Variety and Exposure) by using the CLASS statement. Specify the model by using the MODEL statement. Use the POWER statement to specify power as the result parameter and provide values for the other analysis parameters, error standard deviation and total sample size. The following SAS statements perform the power analysis:

```
proc glmpower data=Exemplary;
    class Variety Exposure;
    model Height = Variety | Exposure;
    power
        stddev = 5
        ntotal = 60
        power = .;
run;
```

The MODEL statement defines the full model including both main effects and the interaction. The POWER= option in the POWER statement identifies power as the result parameter with a missing value (POWER=.). The STDDEV= option specifies an error standard deviation of 5, and the NTOTAL= option specifies a total sample size of 60. The default value for the ALPHA $=$ option sets the significance level to $\alpha=0.05$.

Figure 47.1 shows the output.

Figure 47.1 Sample Size Analysis for Two-Way ANOVA
The GLMPOWER Procedure

| Fixed Scenario Elements |  |  |  |
| :--- | ---: | ---: | :---: |
| Dependent Variable | Height |  |  |
| Error Standard Deviation | 5 |  |  |
| Total Sample Size | 60 |  |  |
| Alpha | 0.05 |  |  |
| Error Degrees of Freedom | 54 |  |  |
| Computed Power |  |  |  |
| Test |  |  |  |
| Index | Source | DF Power |  |
| $\mathbf{1}$ | Variety | 1 |  |
| 2 | Exposure | 2 |  |
| 3 | 0.718 |  |  |
| $\mathbf{3}$ | Variety*Exposure | 2 |  |

The power is about 0.72 for the test of the Variety effect. In other words, there is a probability of 0.72 that the test of the Variety effect will produce a significant result (given the assumptions for the means and error standard deviation). The power is 0.96 for the test of the Exposure effect and 0.19 for the interaction test.

Now, suppose you want to account for some of your uncertainty in conjecturing the true error standard deviation by evaluating the power at reasonable low and high values, 4 and 6.5 . You also want to plot power for sample sizes between 30 and 90 . The following statements perform the analysis:

```
ods graphics on;
proc glmpower data=Exemplary;
    class Variety Exposure;
    model Height = Variety | Exposure;
    power
            stddev = 4 6.5
            ntotal = 60
            power = .;
    plot x=n min=30 max=90;
run;
ods graphics off;
```

The PLOT statement with the $\mathrm{X}=\mathrm{N}$ option requests a plot with sample size on the X axis. (The result parameter-in this case, power-is always plotted on the other axis.) The MIN= and MAX=options in the PLOT statement specify the sample size range. The ODS GRAPHICS ON statement enables ODS Graphics.
Figure 47.2 shows the output, and Figure 47.3 shows the plot.

Figure 47.2 Sample Size Analysis for Two-Way ANOVA with Input Ranges

## The GLMPOWER Procedure

| Fixed Scenario Elements |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Dependent Variable |  |  | Height |  |
| Total Sample Size |  |  | 60 |  |
| Alpha |  |  | 0.05 |  |
| Error Degrees of Freedom |  |  | 54 |  |
| Computed Power |  |  |  |  |
| Index | $x$ Source | $\begin{aligned} & \text { Std } \\ & \text { Dev } \end{aligned}$ | Test DF | Power |
| 1 | 1 Variety | 4.0 | 1 | 0.887 |
| 2 | 2 Variety | 6.5 | 1 | 0.496 |
| 3 | 3 Exposure | 4.0 | 2 | 0.996 |
| 4 | 4 Exposure | 6.5 | 2 | 0.793 |
|  | 5 Variety*Exposure | 4.0 | 2 | 0.280 |
|  | 6 Variety*Exposure | 6.5 | 2 | 0.130 |

Figure 47.2 reveals that the power ranges from about 0.130 to 0.996 for the different effect tests and scenarios for standard deviation, with a sample size of 60 . In Figure 47.3, the line style identifies the effect test, and the plotting symbol identifies the standard deviation. The locations of the plotting symbols identify actual computed powers; the curves are linear interpolations of these points. Note that the computed points in the plot occur at sample size multiples of 6 , because there are 6 cells in the design (and by default, sample sizes are rounded to produce integer cell sizes).

Figure 47.3 Plot of Power versus Sample Size for Two-Way ANOVA with Input Ranges


## Incorporating Contrasts, Unbalanced Designs, and Multiple Means Scenarios

Suppose you want to compute power for the two-way ANOVA described in the section "Simple Two-Way ANOVA" on page 3599 , but you want to additionally perform the following tasks:

- try an unbalanced sample size allocation with respect to Exposure, using twice as many samples for levels 2 and 3 as for level 1
- consider an additional, less optimistic scenario for the cell means, shown in Table 47.2
- test a contrast of Exposure comparing levels 1 and 3

Table 47.2 Additional Cell Means Scenario

|  | Exposure |  |  |
| :---: | :---: | :---: | :---: |
| Variety | $\mathbf{1}$ | $\mathbf{2}$ | $\mathbf{3}$ |
| 1 | 15 | 16 | 20 |
| 2 | 11 | 14 | 15 |

To specify the unbalanced design and the additional cell means scenario, you can add two new variables to the exemplary data set (Weight for the sample size weights, and HeightNew for the new cell means scenario). Change the name of the original cell means scenario to HeightOrig. The following statements define the exemplary data set:

```
data Exemplary;
    input Variety $ Exposure $ HeightOrig HeightNew Weight;
    datalines;
\begin{tabular}{lllll}
1 & 1 & 14 & 15 & 1 \\
1 & 2 & 16 & 16 & 2 \\
1 & 3 & 21 & 20 & 2 \\
2 & 1 & 10 & 11 & 1 \\
2 & 2 & 15 & 14 & 2 \\
2 & 3 & 16 & 15 & 2
\end{tabular}
;
```

In PROC GLMPOWER, specify the name of the weight variable by using the WEIGHT statement, and specify the name of the cell means variables as dependent variables in the MODEL statement. Use the CONTRAST statement to specify the contrast as you would in PROC GLM. The following statements perform the sample size analysis.

```
proc glmpower data=Exemplary;
    class Variety Exposure;
    model HeightOrig HeightNew = Variety | Exposure;
    weight Weight;
    contrast 'Exposure=1 vs Exposure=3' Exposure 1 0 -1;
    power
        stddev = 5
        ntotal = 60
        power = .;
run;
```

Figure 47.4 shows the output.
Figure 47.4 Sample Size Analysis for More Complex Two-Way ANOVA

## The GLMPOWER Procedure

| Fixed Scenario Elements |  |
| :--- | ---: |
| Weight Variable | Weight |
| Error Standard Deviation | 5 |
| Total Sample Size | 60 |
| Alpha | 0.05 |
| Error Degrees of Freedom | 54 |

Figure 47.4 continued

| Computed Power |  |  |  |  |  |
| ---: | :--- | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  |  |
| Index | Dependent | Type | Source | DF | Power |
| $\mathbf{1}$ | HeightOrig | Effect | Variety | 1 | 0.672 |
| $\mathbf{2}$ | HeightOrig | Effect | Exposure | 2 | 0.911 |
| $\mathbf{3}$ | HeightOrig | Effect | Variety*Exposure | 2 | 0.217 |
| $\mathbf{4}$ | HeightOrig | Contrast | Exposure=1 vs Exposure=3 | 1 | 0.951 |
| $\mathbf{5}$ | HeightNew | Effect | Variety | 1 | 0.754 |
| $\mathbf{6}$ | HeightNew | Effect | Exposure | 2 | 0.633 |
| $\mathbf{7}$ | HeightNew | Effect | Variety*Exposure | 2 | 0.137 |
| $\mathbf{8}$ | HeightNew | Contrast | Exposure=1 vs Exposure=3 | $\mathbf{1}$ | 0.705 |

The power of the contrast of Exposure levels 1 and 3 is about 0.95 for the original cell means scenario (HeightOrig) and only 0.71 for the new one (HeightNew). The power is higher for the test of Variety, but lower for the tests of Exposure and of Variety*Exposure for the new cell means scenario compared to the original one. Note also for the HeightOrig scenario that the power for the unbalanced design (Figure 47.4) compared to the balanced design (Figure 47.1) is slightly lower for the tests of Variety and Exposure, but slightly higher for the test of Variety*Exposure.

## Syntax: GLMPOWER Procedure

The following statements are available in the GLMPOWER procedure:

```
PROC GLMPOWER < options> ;
    BY variables;
    CLASS variables;
    CONTRAST 'label' effect values < . . effect values> </ options> ;
    MANOVA 'label' < test-options > < / detail-options > ;
    MODEL dependent-variables = independent-effects ;
    PLOT < plot-options> </ graph-options> ;
    POWER < options > ;
    REPEATED factor-specification;
    WEIGHT variable;
```

The PROC GLMPOWER statement, the MODEL statement, and the POWER statement are required. If your model contains classification effects, the classification variables must be listed in a CLASS statement, and the CLASS statement must appear before the MODEL statement. In addition, CONTRAST and POWER statements must appear after the MODEL statement. PLOT statements must appear after the POWER statement that defines the analysis for the plot.

If you specify one or more MANOVA or REPEATED statements, then the model is assumed to be multivariate. Otherwise, a univariate model is assumed, in which case multiple dependent variables represent cell means scenarios for a single response.

You can use multiple CONTRAST, MANOVA, REPEATED, POWER, and PLOT statements. Each CONTRAST statement defines a separate between-subject contrast. Each MANOVA or REPEATED statement
defines a separate within-subject contrast for a multivariate model. Each POWER statement produces a separate analysis and uses the information that is contained in the CLASS, MODEL, WEIGHT, CONTRAST, MANOVA, and REPEATED statements. Each PLOT statement refers to the previous POWER statement and generates a separate graph (or set of graphs).
Table 47.3 summarizes the basic functions of each statement in PROC GLMPOWER. The syntax of each statement in Table 47.3 is described in the following pages.

Table 47.3 Statements in the GLMPOWER Procedure

| Statement | Description |
| :--- | :--- |
| PROC GLMPOWER | Invokes procedure and specifies exemplary data set <br> BY <br> CLASS <br> analysis |
| CONTRAST | Declares classification variables |
| MANOVA | Defines between-subject linear tests of model <br> parameters |
| MODEL | Defines within-subject linear tests of model param- <br> eters for multivariate models, in terms of contrast <br> matrix coefficients |
| PLOT | Defines model and specifies dependent variables; <br> for univariate models, multiple dependent variables <br> represent cell means scenarios for a single response |
| POWER | Displays graphs for preceding POWER statement |
| REPEATED | Identifies parameter to solve for and provides one <br> or more scenarios for values of other analysis <br> parameters |
| WEIGHT | Defines within-subject linear tests of model param- <br> eters for multivariate models, in terms of common <br> repeated measures transformations of the dependent <br> variables |

## PROC GLMPOWER Statement

PROC GLMPOWER < options > ;
The PROC GLMPOWER statement invokes the GLMPOWER procedure. You can specify the following options.

## DATA=SAS-data-set

names a SAS data set to be used as the exemplary data set, which is an artificial data set constructed to represent the intended sampling design and the conjectured response means for the underlying population.

## ORDER=DATA | FORMATTED | FREQ | INTERNAL

specifies the sort order for the levels of the classification variables (which are specified in the CLASS statement).

This option applies to the levels for all classification variables, except when you use the (default) ORDER=FORMATTED option with numeric classification variables that have no explicit format. In that case, the levels of such variables are ordered by their internal value.

The ORDER= option can take the following values:

| Value of ORDER= | Levels Sorted By |
| :--- | :--- |
| DATA | Order of appearance in the input data set |
| FORMATTED | External formatted value, except for numeric variables with <br> no explicit format, which are sorted by their unformatted <br> (internal) value |
| FREQ | Descending frequency count; levels with the most observa- <br> tions come first in the order |
| INTERNAL | Unformatted value |

By default, ORDER=FORMATTED. For ORDER=FORMATTED and ORDER=INTERNAL, the sort order is machine-dependent.

For more information about sort order, see the chapter on the SORT procedure in the Base SAS Procedures Guide and the discussion of BY-group processing in SAS Language Reference: Concepts.

## PLOTONLY

specifies that only graphical results from the PLOT statement be produced.

## BY Statement

## BY variables ;

You can specify a BY statement with PROC GLMPOWER to obtain separate analyses of observations in groups that are defined by the BY variables. When a BY statement appears, the procedure expects the input data set to be sorted in order of the BY variables. If you specify more than one BY statement, only the last one specified is used.

If your input data set is not sorted in ascending order, use one of the following alternatives:

- Sort the data by using the SORT procedure with a similar BY statement.
- Specify the NOTSORTED or DESCENDING option in the BY statement for the GLMPOWER procedure. The NOTSORTED option does not mean that the data are unsorted but rather that the
data are arranged in groups (according to values of the BY variables) and that these groups are not necessarily in alphabetical or increasing numeric order.
- Create an index on the BY variables by using the DATASETS procedure (in Base SAS software).

Because sorting the data changes the order in which PROC GLMPOWER reads observations, the sort order for the levels of the classification variables might be affected if you have also specified ORDER=DATA in the PROC GLMPOWER statement. This, in turn, affects specifications in CONTRAST statements.

For more information about BY-group processing, see the discussion in SAS Language Reference: Concepts. For more information about the DATASETS procedure, see the discussion in the Base SAS Procedures Guide.

## CLASS Statement

## CLASS variables ;

The CLASS statement names the classification variables to be used in the analysis. If you use the CLASS statement, it must appear before the MODEL statement.

Classification variables can be either character or numeric. By default, class levels are determined from the entire set of formatted values of the CLASS variables.

## CONTRAST Statement

CONTRAST 'label' effect values < . . . effect values> </ options> ;
The CONTRAST statement enables you to define custom Type III hypothesis tests by specifying an $\mathbf{L}$ vector or matrix for testing either the hypothesis $\mathbf{L} \boldsymbol{\beta}=0$ (for univariate models) or the hypothesis $\mathbf{L B M}=0$ (for multivariate models). The $\mathbf{L}$ matrix consists of one or more between-subject contrasts.

To use this feature, you must be familiar with the details of the model parameterization that PROC GLM uses. For more information, see the section "Parameterization of PROC GLM Models" on page 3456 in Chapter 45, "The GLM Procedure." All the elements of the $\mathbf{L}$ matrix can be given, or if only certain portions of the $\mathbf{L}$ matrix are given, PROC GLMPOWER constructs the remaining elements from the context (in a manner similar to that in rule 4 in the section "Construction of Least Squares Means" on page 3489 in Chapter 45, "The GLM Procedure").

There is no limit to the number of CONTRAST statements that you can specify, but they must appear after the MODEL statement. Each power analysis includes tests for all CONTRAST statements.
You can specify the following arguments:
label identifies the contrast in the output. A label is required for every contrast that is specified. Labels must be enclosed in single or double quotation marks.
effect identifies an effect that appears in the MODEL statement, or the INTERCEPT effect. You do not need to include all effects that appear in the MODEL statement.
values are constants that are elements of the $\mathbf{L}$ matrix associated with the effect.
You can specify the following option in the CONTRAST statement after a slash (/):

## SINGULAR=number

tunes the estimability checking. If $\operatorname{ABS}(\mathbf{L}-\mathbf{L H})>C \times$ number for any row in the contrast, then $\mathbf{L}$ is declared nonestimable. $\mathbf{H}$ is the $\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{X}^{\prime} \mathbf{X}$ matrix, and $C$ is $\operatorname{ABS}(\mathbf{L})$ except for rows where $\mathbf{L}$ is zero, and then it is 1 . The default value for the SINGULAR $=$ option is $10^{-4}$. Values for the SINGULAR= option must be between 0 and 1 .

As stated previously, the CONTRAST statement enables you to define custom hypothesis tests. If the hypothesis is testable in a univariate model, then the hypothesis sum of squares, $\operatorname{SS}\left(H_{0}: \mathbf{L} \boldsymbol{\beta}=0\right)$, is computed as

$$
(\mathbf{L b})^{\prime}\left(\mathbf{L}\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{L}^{\prime}\right)^{-1}(\mathbf{L b})
$$

where $\mathbf{b}=\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{X}^{\prime} \mathbf{y}$.
For testable hypotheses in a multivariate model, the usual multivariate tests are defined by using

$$
\mathbf{H}=\mathbf{M}^{\prime}(\mathbf{L B})^{\prime}\left(\mathbf{L}\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{L}^{\prime}\right)^{-1}(\mathbf{L B}) \mathbf{M}
$$

where $\mathbf{B}=\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{X}^{\prime} \mathbf{Y}$ and $\mathbf{Y}$ is the matrix of multivariate responses or dependent variables.
The degrees of freedom associated with the hypothesis are equal to the row rank of $\mathbf{L}$. The sum of squares computed in this situation is equivalent to the sum of squares computed using an $\mathbf{L}$ matrix with any row deleted that is a linear combination of previous rows.

Multiple-degrees-of-freedom hypotheses can be specified by separating the rows of the $\mathbf{L}$ matrix with commas.

## MANOVA Statement

## MANOVA 'label' < test-options> </detail-options>;

If the MODEL statement includes more than one dependent variable, you can indicate a multivariate model and define transformations of dependent variables by using the MANOVA statement.

The MANOVA statement enables you to define custom Type III hypothesis tests by specifying an $\mathbf{M}$ vector or matrix for testing the hypothesis $\mathbf{L} \boldsymbol{\beta} \mathbf{M}=0$. The $\mathbf{L}$ matrix consists of one or more between-subject contrasts that involve the model effects, and the M matrix consists of one or more within-subject contrasts.

To use this feature, you must be familiar with the details of multivariate model and contrast parameterizations that are used in PROC GLM. For more information, see the sections "Multivariate Analysis of Variance" on page 3492 and "Repeated Measures Analysis of Variance" on page 3493 in Chapter 45, "The GLM Procedure." For information about the power and sample size computational methods and formulas, see the section "Contrasts in Fixed-Effect Multivariate Models" on page 3633.

You can use either the MANOVA statement or the REPEATED statement with any of the tests for multivariate models that are supported in the MTEST= option in the POWER statement. For handling repeated measures on the same experimental unit, you would usually use the REPEATED statement instead of the MANOVA statement. But you can use the MANOVA statement in repeated measures situations, in addition to situations where you have clusters or multiple outcome variables. The differences between the MANOVA and REPEATED statements are as follows:

- You can use the MANOVA statement to construct any M matrix, but you must specify the coefficients explicitly (except for the default identity matrix).
- You can use the REPEATED statement to specify commonly used contrasts by using keywords rather than coefficients, but you are limited to only those forms of the $\mathbf{M}$ matrix.

There is no limit to the number of MANOVA statements that you can specify. Each power analysis includes tests for all MANOVA statements.

The label identifies the dependent variable transformation in the output. The label serves the same purpose as the factor-name in the REPEATED statement, enabling you to use the MANOVA statement for tests of within-subject effects and within-subject-by-between-subject interactions. A label is required for every transformation that is specified. Labels must be enclosed in single or double quotation marks.

## Test Options

You can specify the following test-option in the MANOVA statement:
$\mathbf{M}=$ equation, . . . , equation | (row-of-matrix, . . . , row-of-matrix)
specifies a transformation matrix for the dependent variables that are listed in the MODEL statement.
The equations in the $M=$ specification are of the form

$$
\begin{array}{rc}
c_{1} \times \text { dependent-variable } & \pm c_{2} \times \text { dependent-variable } \\
\cdots & \pm \quad c_{n} \times \text { dependent-variable }
\end{array}
$$

where the $c_{i}$ values are coefficients of the various dependent-variables. If the value of a given $c_{i}$ is 1 , it can be omitted; in other words, $1 \times Y$ is the same as $Y$. Equations should involve two or more dependent variables.

Alternatively, you can input the transformation matrix directly by entering the elements of the matrix, using commas to separate the rows and parentheses to surround the matrix. When you use this alternate form of input, the number of elements in each row must equal the number of dependent variables. Although these combinations actually represent the columns of the $\mathbf{M}$ matrix, they are displayed by rows.

When you include an $M=$ specification, the tests are based on the variables that are defined by the equations in the specification, not the original dependent variables. If you omit the $\mathrm{M}=$ option, the tests are based on the original dependent variables in the MODEL statement. Omitting the $\mathrm{M}=$ option is equivalent to specifying an identity matrix, as in the following example, which assumes three dependent variables:

```
MANOVA 'Identity' M=(1 0 0,
    0 1 0,
    0 0 1);
```

For more examples of the $\mathrm{M}=$ option, see the section "Examples" on page 3431 in Chapter 45, "The GLM Procedure." The syntax and functionality of the $M=$ option in PROC GLM are the same as in PROC GLMPOWER.

## Detail Options

You can specify the following detail-option in the MANOVA statement after a slash (/):
ORTH
requests that the transformation matrix in the $\mathrm{M}=$ specification of the MANOVA statement be orthogonalized by rows before the analysis.

## MODEL Statement

MODEL dependent-variables = independent-effects ;
The MODEL statement names the dependent variables and independent effects. If one or more MANOVA or REPEATED statements are specified, then multiple dependent variables define a multivariate model. In the absence of the MANOVA and REPEATED statements, a univariate model is assumed, and multiple dependent variables represent different scenarios for the cell means.

The independent-effects can involve classification variables, continuous variables, or both. You can include main effects and interactions by using the effects notation of PROC GLM; for more information, see the section "Specification of Effects" on page 3453 in Chapter 45, "The GLM Procedure." For any model effect that involves classification variables (main effects and interactions), the number of levels cannot exceed 32,767 . If no independent effects are specified, only an intercept term is fit. You can specify only one MODEL statement, and it must appear before the POWER statement if the EFFECTS= option is specified in the POWER statement.

For a univariate model, you can account for covariates without specifying them explicitly in the model by using the NCOVARIATES $=$ option and either the CORRXY= or PROPVARREDUCTION= option in the POWER statement. For a multivariate model, you must explicitly specify any covariates in the MODEL statement.

The values of dependent variables in the exemplary data set (the data set named by the DATA= option in the PROC GLMPOWER statement) are surmised response means across subject profiles. For a univariate model, multiple dependent variables correspond to multiple scenarios for these cell means.

The MODEL statement is required. You can specify only one MODEL statement.

## PLOT Statement

PLOT < plot-options> </ graph-options> ;
The PLOT statement produces a graph or set of graphs for the sample size analysis defined by the previous POWER statement. The plot-options define the plot characteristics, and the graph-options are like those in SAS/GRAPH software. If ODS Graphics is enabled, then the PLOT statement uses ODS Graphics to create graphs. For example:

```
ods graphics on;
proc glmpower data=Exemplary;
    class Variety Exposure;
    model Height = Variety | Exposure;
    power
        stddev = 4 6.5
        ntotal = 60
        power = .;
    plot x=n min=30 max=90;
run;
ods graphics off;
```

Otherwise, traditional graphics are produced. For example:

```
ods graphics off;
proc glmpower data=Exemplary;
    class Variety Exposure;
    model Height = Variety | Exposure;
    power
        stddev = 4 6.5
        ntotal = 60
        power = .;
    plot x=n min=30 max=90;
run;
```

For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 606 in Chapter 21, "Statistical Graphics Using ODS."

Table 47.4 summarizes the options available in the PLOT statement.

Table 47.4 PLOT Statement Options

| Option | Description |
| :--- | :--- |
| Plot Options |  |
| INTERPOL $=$ | Specifies the type of curve to draw |
| KEY $=$ | Specifies the style of key for the plot |
| MARKERS $=$ | Specifies the locations for plotting symbols |
| MAX $=$ | Specifies the maximum of the range of values |
| MIN $=$ | Specifies the minimum of the range of values |
| NPOINTS $=$ | Specifies the number of values |
| STEP $=$ | Specifies the increment between values |
| VARY | Specifies how plot features should be linked to varying analysis parameters |
| X= | Specifies a plot with the requested type of parameter on the X axis |
| XOPTS $=$ | Specifies plot characteristics pertaining to the X axis |
| Y= | Specifies a plot with the requested type of parameter on the Y axis |
| YOPTS $=$ | Specifies plot characteristics pertaining to the Y axis |
| Graph Options |  |
| DESCRIPTION= $=$ | Specifies a descriptive string |
| NAME $=$ | Specifies a name for the catalog entry for the plot |

## Options

You can specify the following plot-options in the PLOT statement.

## INTERPOL=JOIN | NONE

specifies the type of curve to draw through the computed points. The INTERPOL=JOIN option connects computed points with straight lines. The INTERPOL=NONE option leaves computed points unconnected.

## KEY=BYCURVE < (bycurve-options) >

KEY=BYFEATURE < (byfeature-options) >
KEY=ONCURVES
specifies the style of key (or "legend") for the plot. The default is KEY=BYFEATURE, which specifies a key with a column of entries for each plot feature (line style, color, and/or symbol). Each entry shows the mapping between a value of the feature and the value(s) of the analysis parameter(s) linked to that feature. The KEY=BYCURVE option specifies a key with each row identifying a distinct curve in the plot. The KEY=ONCURVES option places a curve-specific label adjacent to each curve.

You can specify the following byfeature-options in parentheses after the KEY=BYCURVE option.

## NUMBERS=OFF | ON

specifies how the key should identify curves. If NUMBERS=OFF, then the key includes symbol, color, and line style samples to identify the curves. If NUMBERS $=O N$, then the key includes numbers matching numeric labels placed adjacent to the curves. The default is NUMBERS=ON.

## POS=BOTTOM | INSET

specifies the position of the key. The POS=BOTTOM option places the key below the X axis. The POS=INSET option places the key inside the plotting region and attempts to choose the least crowded corner. The default is POS=BOTTOM.

You can specify the following byfeature-options in parentheses after KEY=BYFEATURE option.

## POS=BOTTOM | INSET

specifies the position of the key. The POS=BOTTOM option places the key below the X axis. The POS=INSET option places the key inside the plotting region and attempts to choose the least crowded corner. The default is POS=BOTTOM.

## MARKERS=ANALYSIS | COMPUTED | NICE | NONE

specifies the locations for plotting symbols.
The MARKERS=ANALYSIS option places plotting symbols at locations corresponding to the values of the relevant input parameter from the POWER statement preceding the PLOT statement.
The MARKERS=COMPUTED option (the default) places plotting symbols at the locations of actual computed points from the sample size analysis.

The MARKERS=NICE option places plotting symbols at tick mark locations (corresponding to the argument axis).

The MARKERS=NONE option disables plotting symbols.

## MAX=number | DATAMAX

specifies the maximum of the range of values for the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). The default is DATAMAX, which specifies the maximum value that occurs for this parameter in the POWER statement that precedes the PLOT statement.

## MIN=number | DATAMIN

specifies the minimum of the range of values for the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). The default is DATAMIN, which specifies the minimum value that occurs for this parameter in the POWER statement that precedes the PLOT statement.

## NPOINTS=number

## NPTS=number

specifies the number of values for the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). You cannot use the NPOINTS $=$ and STEP= options simultaneously. The default value for typical situations is 20 .

## STEP=number

specifies the increment between values of the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). You cannot use the STEP= and NPOINTS= options simultaneously. By default, the NPOINTS= option is used instead of the STEP= option.

VARY ( feature < BY parameter-list $><, \ldots$, feature < BY parameter-list >> )
specifies how plot features should be linked to varying analysis parameters. Available features are COLOR, LINESTYLE, PANEL, and SYMBOL. A "panel" refers to a separate plot with a heading identifying the subset of values represented in the plot.

The parameter-list is a list of one or more names, separated by spaces. Each name must match the name of an analysis option used in the POWER statement preceding the PLOT statement, or one of the following keywords:

- SOURCE, which represents the model effects and contrasts in a univariate model and the betweensubject effects and contrasts in a multivariate model
- DEPENDENT, which represents the cell means scenarios in a univariate model or the dependent variable transformations in a multivariate model

If the name represents an analysis option that is specified in the POWER statement, then it must be the primary name for the analysis option-that is, the one that is listed first in the syntax description.

If you omit the $<$ BY parameter-list $>$ portion for a feature, then one or more multivalued parameters from the analysis are automatically selected for you.

## X=N | POWER

specifies a plot with the requested type of parameter on the $X$ axis and the parameter being solved for on the $Y$ axis. When $X=N$, sample size is assigned to the $X$ axis. When $X=P O W E R$, power is assigned to the X axis. You cannot use the $\mathrm{X}=$ and $\mathrm{Y}=$ options simultaneously. The default is $\mathrm{X}=\mathrm{POWER}$, unless the result parameter is power, in which case the default is $X=N$.

## XOPTS= ( $x$-options )

specifies plot characteristics pertaining to the X axis.
You can specify the following $x$-options in parentheses.

## CROSSREF=NO | YES

specifies whether the reference lines defined by the $\mathrm{REF}=x$-option should be crossed with a reference line on the Y axis that indicates the solution point on the curve.

## REF=number-list

specifies locations for reference lines extending from the X axis across the entire plotting region. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## Y=N | POWER

specifies a plot with the requested type of parameter on the Y axis and the parameter being solved for on the X axis. When $\mathrm{Y}=\mathrm{N}$, sample size is assigned to the Y axis. When $\mathrm{Y}=\mathrm{POWER}$, power is assigned to the Y axis. You cannot use the $\mathrm{Y}=$ and $\mathrm{X}=$ options simultaneously. By default, the $\mathrm{X}=$ option is used instead of the $\mathrm{Y}=$ option.

## YOPTS=( $y$-options )

specifies plot characteristics pertaining to the Y axis.
You can specify the following $y$-options in parentheses.

## CROSSREF=NO | YES

specifies whether the reference lines defined by the REF=y-option should be crossed with a reference line on the X axis that indicates the solution point on the curve.

## REF=number-list

specifies locations for reference lines extending from the Y axis across the entire plotting region. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

You can specify the following graph-options in the PLOT statement after a slash (/).

## DESCRIPTION=’string'

specifies a descriptive string of up to 40 characters that appears in the "Description" field of the graphics catalog. The description does not appear on the plots. By default, PROC GLMPOWER assigns a description either of the form " $Y$ versus $X$ " (for a single-panel plot) or of the form " $Y$ versus $X(S)$," where $Y$ is the parameter on the Y axis, $X$ is the parameter on the X axis, and $S$ is a description of the subset represented on the current panel of a multipanel plot.

## NAME='string '

specifies a name of up to eight characters for the catalog entry for the plot. The default name is PLOTn, where $n$ is the number of the plot statement within the current invocation of PROC GLMPOWER. If the name duplicates the name of an existing entry, SAS/GRAPH software adds a number to the duplicate name to create a unique entry-for example, PLOT11 and PLOT12 for the second and third panels of a multipanel plot generated in the first PLOT statement in an invocation of PROC GLMPOWER.

## POWER Statement

```
POWER < options > ;
```

The POWER statement performs power and sample size analyses for the Type III $F$ tests that are specified in the MTEST= option, for each effect in the model that is defined by the MODEL statement, and for the contrasts that are defined by all CONTRAST, MANOVA, and REPEATED statements. The POWER statement must appear after the MODEL statement if the EFFECTS= option is used in the POWER statement.

For information about the power and sample size computational methods and formulas, see the section "Computational Methods and Formulas" on page 3630.

## Summary of Options

Table 47.5 summarizes the options available in the POWER statement.

Table 47.5 POWER Statement Options

| Option | Description |
| :---: | :---: |
| Specify test statistic |  |
| MTEST= | Specifies the test statistic for a multivariate model |
| UEPSDEF= | Specifies the form of the Huynh-Feldt epsilon for MTEST=HF |
| Specify analysis information |  |
| ALPHA= | Specifies the level of significance of each test |
| EFFECTS | Specifies the model effects |
| Specify covariates for a univariate model |  |
| CORRXY= | Specifies multiple correlation ( $\rho$ ) between covariates and response |
| NCOVARIATES= | Specifies additional degrees of freedom due to covariates |
| PROPVARREDUCTION= | Specifies proportional variance reduction (r) due to covariates |
| Specify variability |  |
| CORRMAT= | Specifies the correlation matrix of the dependent variables in a multivariate model |
| CORRS $=$ | Specifies the correlations among the dependent variables in a multivariate model |
| COVMAT= | Specifies the covariance matrix of the dependent variables in a multivariate model |
| MATRIX= | Defines a matrix or vector |
| SQRTVAR= | Specifies the vector of error standard deviations for each dependent variable in a multivariate model |
| STDDEV= | Specifies the common error standard deviation |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NTOTAL= | Specifies the sample size |
| Specify power |  |
| POWER= | Specifies power |
| Choose computational method |  |
| METHOD= | Specifies the computational method for multivariate tests |

Table 47.5 continued

| Option | Description |
| :--- | :--- |
| Control ordering in output |  |
| DEPENDENT | Specifies the location of the Dependent or Transformation column in the <br> output |
| OUTPUTORDER $=$ | Controls ordering in output |

Table 47.6 summarizes the valid result parameters.

Table 47.6 Summary of Result Parameters in the POWER
Statement

| Solve for | Syntax |
| :--- | :--- |
| Power | POWER $=$. |
| Sample size | NTOTAL $=$. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of each test. The default is 0.05 , corresponding to the usual $0.05 \times$ $100 \%=5 \%$ level of significance. Note that this is a test-wise significance level with the same value for all tests, not incorporating any corrections for multiple testing. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

CORRMAT=name-list
specifies the correlation matrix of the dependent variables in a multivariate model, by using labels that are specified in the MATRIX= option. The corresponding matrices that are defined in the MATRIX= option must have either a lower triangular form that includes the diagonal of 1's or a linear exponent autoregressive (LEAR) correlation structure. The matrix must be positive definite. You can use the CORRMAT = option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements. For information about specifying the name-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## CORRS=name-list

specifies the correlations among the dependent variables in a multivariate model, by using labels that are specified in the MATRIX= option. The corresponding matrices that are defined in the MATRIX= option must have either a lower triangular form that excludes the diagonal of 1's or a LEAR correlation structure. The matrix must be positive definite. You can use the CORRS= option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements. For information about specifying the name-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

COVMAT=name-list
specifies the covariance matrix of the dependent variables in a multivariate model, by using labels that are specified in the MATRIX= option. The corresponding matrices that are defined in the MATRIX= option must have a lower triangular form that includes the diagonal of error variances. The matrix must be positive definite. You can use the COVMAT= option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements. For information about specifying the name-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626 .

## CORRXY=number-list

specifies the multiple correlation $(\rho)$ between all covariates and the response for a univariate model. The error standard deviation that is given by the $\mathrm{STDDEV}=$ option is consequently reduced by multiplying it by a factor of $\left(1-\rho^{2}\right)^{\frac{1}{2}}$, provided that the number of covariates (as determined by the NCOVARIATES= option) is greater than 0 . You cannot use the CORRXY= and PROPVARREDUCTION= options simultaneously. You cannot use the CORRXY= option when you have a multivariate model-that is, in the presence of a MANOVA or REPEATED statement. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## DEPENDENT

specifies the location of the Dependent column (for a univariate model) or the Transformation column (for a multivariate model) in the output when you specify the OUTPUTORDER=REVERSE option or OUTPUTORDER=SYNTAX option, according to its relative position in the POWER statement.

## EFFECTS <= < ( effect . . . effect ) >>

specifies the model effects to include in the power analysis. By default, or if the EFFECTS keyword is specified without the equal sign (=), all model effects are included. Specify EFFECTS=() to exclude all model effect tests from the power analysis. You can include main effects and interactions by using the effects notation of PROC GLM; see the section "Specification of Effects" on page 3453 in Chapter 45, "The GLM Procedure" for further details. The MODEL statement must appear before the POWER statement if the EFFECTS option is used.

## MATRIX('label')=matrix-specification

defines a matrix or vector that you can use along with the CORRMAT=, CORRS $=$, COVMAT=, and SQRTVAR= options when you have a multivariate model.

The matrix-specification can have one of the following three forms:

## 1. Raw values

(values)
specifies the values of a matrix or vector in one of the following forms:

- a matrix in lower triangular form, for use with the CORRMAT $=$ or COVMAT $=$ option
- a matrix in strictly lower triangular form, for use with the CORRS= option
- a vector, for use with the SQRTVAR= option

A matrix in lower triangular form contains the diagonal the and values below it. For example, you can represent a $3 \times 3$ correlation matrix for use with the CORRMAT= option as follows:

```
MATRIX ('MyCorrMat') = ( 1
    0.5 1
    0.2 0.5 1)
```

You can represent the same correlation matrix in strictly lower triangular form for use with the CORRS= option as follows:

```
MATRIX ('MyCorrs') = (0.5
    0.2 0.5)
```

An example of a vector for use with the SQRTVAR= option is as follows:

```
MATRIX ('MySqrtVar') = (3.2 4.5 3.7)
```


## 2. Linear exponent autoregressive (LEAR) correlation structure

LEAR (base-corr, corr-decay <, nlevels <, level-values >> )
specifies a LEAR correlation structure for use with the CORRMAT $=$ or CORRS $=$ option. The LEAR structure is useful for characterizing exponentially decaying within-subject correlation that decays at a rate slower or faster than $\operatorname{AR}(1)$. Special cases include compound symmetry, first-order autoregressive (AR(1)), and first-order moving average correlation structures.

The LEAR correlation structure is related to the spatial covariance structures in PROC MIXED and is discussed in Simpson et al. (2010).

The base-corr $(\rho)$ is the correlation between variables whose level-values are one unit apart, and it must satisfy $0 \leq \rho<1$. The corr-decay ( $\delta$ ) is the correlation decay rate, which must be nonnegative. The default value for the number of levels, nlevels ( $n$ ), is the number of dependent variables. The $n$ level-values, denoted $\left\{l_{1}, \ldots, l_{n}\right\}$, must be distinct. The default level-values are $\{1, \ldots, n\}$. Let $d_{j k}$ denote the distance between levels $j$ and $k,\left(d_{j k}=\left|l_{j}-l_{k}\right|\right)$. Let $d_{\min }=\min \left(d_{j k}: j \neq k\right)$ and $d_{\max }=\max \left(d_{j k}: j \neq k\right)$. The $(i, j)$ th element of the correlation matrix according to the LEAR model is defined as

$$
\rho_{j k}= \begin{cases}1 & \text { if } j=k \\ \rho_{\min }+\delta\left[\left(d_{j k}-d_{\min }\right) /\left(d_{\max }-d_{\min }\right)\right] & \text { if } j \neq k \text { and } d_{\min } \neq d_{\max } \\ \rho & \text { if } j \neq k \text { and } d_{\min }=d_{\max }\end{cases}
$$

Compound symmetry is the special case $\delta=0$. $\operatorname{AR}(1)$ is the special case $\delta=d_{\max }-d_{\min }$. As $\delta \rightarrow \infty$, the model approaches the first-order moving average model.

## 3. Kronecker product

'matrix-name’ @ 'matrix-name’<@ ...@ 'matrix-name’>
specifies a direct (Kronecker) product of two or more matrices for use with the CORRMAT=, CORRS=, or COVMAT= option. This form is useful when you have more than one type of distinction among dependent variables. For example, suppose you have a three-level repeated measurement factor with correlation that is 0.4 for neighboring measurements and that decays slightly more slowly than $\operatorname{AR}(1)$
across more distant measurements. You also have four clusters that you believe satisfy compound symmetry with a correlation of 0.3 . Your level values are the same as the default for the LEAR model. You can specify this correlation structure as follows:

```
MATRIX ('RepMeasures') = LEAR (0.4, 1.5, 3)
MATRIX ('Clusters') = LEAR (0.3, 0, 4)
MATRIX ('FullCorr') = 'RepMeasures' @ 'Clusters'
CORRMAT = 'FullCorr'
```

You can use the MATRIX option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements.

## METHOD=MULLERPETERSON | MP

## METHOD=OBRIENSHIEH | OS

specifies the power computation method for the multivariate tests (MTEST=HLT, MTEST=PT, and MTEST=WILKS). METHOD=OBRIENSHIEH (the default) is based on O'Brien and Shieh (1992), and METHOD=MULLERPETERSON is based on Muller and Peterson (1984). For information about the associated power and sample size computational methods and formulas, see the section "Multivariate Tests" on page 3635.

If the dependent variable transformation consists of a single contrast ( $r_{M}=1$ ), then the two methods are identical and compute exact power. If $r_{M}>1$ but the model effect or between-subject contrast has only one degree of freedom ( $r_{L}=1$ ), then METHOD=OBRIENSHIEH computes exact results and METHOD=MULLERPETERSON computes approximate results. If $r_{M}>1$ and $r_{L}>1$, then both methods compute approximate results.
You can use the METHOD= option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements.

## MTEST=test-list

specifies the form of the $F$ test for a multivariate model. Seven keywords are available, as discussed in the following paragraphs: BOX, GG, HF, HLT, PT, UNCORR, and WILKS. For information about specifying the keyword-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

Three of these tests are multivariate, corresponding to the default MSTAT=FAPPROX option in the MANOVA and REPEATED statements in PROC GLM:

- MTEST=HLT (the default) is the Hotelling-Lawley trace
- MTEST=PT is Pillai's trace
- MTEST=WILKS is Wilks' lambda

For more information about these multivariate tests, see the section "Multivariate Tests" on page 90 in Chapter 4, "Introduction to Regression Procedures." For information about the associated power and sample size computational methods and formulas, see the section "Multivariate Tests" on page 3635.
The other four tests are univariate, corresponding to the univariate approach to repeated measures in the REPEATED statement in PROC GLM:

- MTEST=UNCORR is the uncorrected univariate $F$ test, assuming sphericity $(\varepsilon=1)$.
- MTEST=GG is the $F$ test with the Greenhouse-Geisser adjustment, estimating $\varepsilon$ by its maximum likelihood estimate $\hat{\varepsilon}$.
- MTEST=HF is the $F$ test with the Huynh-Feldt adjustment as specified by the UEPSDEF= option, using an approximately unbiased estimate $\tilde{\varepsilon}$.
- MTEST=BOX is the $F$ test with Box's conservative adjustment, estimating sphericity by its smallest possible value $1 / r_{M}$, where $r_{M}$ is the number of within-subject contrasts in the dependent variable transformation.

For more information about these univariate tests, see the section "Hypothesis Testing in Repeated Measures Analysis" on page 3496 in Chapter 45, "The GLM Procedure." For information about the associated power and sample size computational methods and formulas, see the section "Univariate Tests" on page 3638.

These tests are all of the form $\mathbf{L} \boldsymbol{\beta} \mathbf{M}=0$, where $\mathbf{L}$ is a between-subject contrast, $\boldsymbol{\beta}$ is the matrix of model parameters, and $\mathbf{M}$ is a within-subject contrast.

You can use the MTEST= option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements.

## NCOVARIATES=number-list

NCOVARIATE=number-list

## NCOVS=number-list

## NCOV=number-list

specifies the number of additional degrees of freedom to accommodate covariate effects-both class and continuous-not listed in the MODEL statement, for a univariate model. The error degrees of freedom are consequently reduced by the value of the NCOVARIATES $=$ option, and the error standard deviation (whose unadjusted value is provided with the STDDEV= option) is reduced according to the value of the CORRXY= or PROPVARREDUCTION= option. You cannot use the NCOVARIATES= option when you have a multivariate model-that is, in the presence of a MANOVA or REPEATED statement. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 3627 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). The term "sample size" here refers to the number of independent sampling units. Values for the sample size for a univariate model must be no smaller than the model degrees of freedom (counting the covariates if present). The minimum required sample size for a multivariate model depends on the analysis and computational method; for more information, see the section "Contrasts in Fixed-Effect Multivariate Models" on page 3633. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- DEPENDENT
- EFFECTS
- weight variable (from the WEIGHT statement)
- ALPHA=
- NCOVARIATES=
- CORRXY=
- PROPVARREDUCTION=
- STDDEV=
- NTOTAL=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the POWER statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the POWER statement.

## POWER=number-list

specifies the desired power of each test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability (for example, 0.9) rather than a percentage. Note that this is a test-wise power with the same value for all tests, without any correction for multiple testing. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## PROPVARREDUCTION=number-list

## PVRED=number-list

specifies the proportional reduction $(r)$ in total R square incurred by the covariates-in other words, the amount of additional variation explained by the covariates-for a univariate model. The error standard deviation that is given by the STDDEV= option is consequently reduced by multiplying it by a factor of $(1-r)^{\frac{1}{2}}$, provided that the number of covariates (as determined by the NCOVARIATES $=$ option) is greater than 0 . You cannot use the PROPVARREDUCTION= and CORRXY= options simultaneously. You cannot use the PROPVARREDUCTION= option when you have a multivariate model-that is, in the presence of a MANOVA or REPEATED statement. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## SQRTVAR=name-list

specifies the vector of standard deviations-that is, the square roots of the variances-of the dependent variables in a multivariate model, by using labels that are specified using the MATRIX= option. The standard deviation values must be positive. You can use the SQRTVAR=option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements. For information about specifying the name-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## STDDEV=number-list

specifies the error standard deviation, or root MSE. For a multivariate model, each value in the numberlist is taken to be a common value for all dependent variables. If covariates are specified by using the NCOVARIATES $=$ option, then the STDDEV= option denotes the error standard deviation before accounting for these covariates. For information about specifying the number-list, see the section "Specifying Value Lists in the POWER Statement" on page 3626.

## UEPSDEF=unbiased-epsilon-definition

specifies the type of adjustment for MTEST=HF. The default is UEPSDEF=HFL, corresponding to the corrected form of the Huynh-Feldt adjustment (Huynh and Feldt 1976; Lecoutre 1991). Other alternatives are UEPSDEF=HF; the uncorrected Huynh-Feldt adjustment; and UEPSDEF=CM, the adjustment of Chi et al. (2012). For more information about these adjustments, see the section "Hypothesis Testing in Repeated Measures Analysis" on page 3496 in Chapter 45, "The GLM Procedure." You can use the UEPSDEF= option only when you have a multivariate model-that is, in the presence of one or more MANOVA or REPEATED statements. For information about the associated power and sample size computational methods and formulas, see the section "Univariate Tests" on page 3638.

## Restrictions on Option Combinations

To specify the variability in a multivariate model, choose one of the following parameterizations:

- covariance matrix (using the MATRIX= and COVMAT= options)
- standard deviations and correlations (using the MATRIX=, SQRTVAR=, and CORRS= options)
- common standard deviation and correlations (using the $\operatorname{STDDEV}=$, MATRIX $=$, and CORRS $=$ options)
- standard deviations and correlation matrix (using the MATRIX=, SQRTVAR=, and CORRMAT= options)
- common standard deviation and correlation matrix (using the $\operatorname{STDDEV}=$, MATRIX $=$, and CORRMAT= options)

For the relationship between covariates and response in a univariate model, specify either the multiple correlation (by using the CORRXY= option) or the proportional reduction in total R square (by using the PROPVARREDUCTION= option).

## REPEATED Statement

REPEATED factor-specification ;
If the MODEL statement includes more than one dependent variable, you can indicate a multivariate model and define transformations of dependent variables by using the REPEATED statement.

The REPEATED statement enables you to define custom Type III hypothesis tests by choosing from among several transformations of the dependent variables: contrast, Helmert, identity, mean, polynomial, and profile. You can specify a transformation for each repeated factor (often called a within-subject factor), and each combination of repeated factors produces an $\mathbf{M}$ vector or matrix for testing the hypothesis $\mathbf{L} \boldsymbol{\beta} \mathbf{M}=0$. The $\mathbf{L}$ matrix consists of one or more between-subject contrasts that involve the model effects, and the M matrix
consists of one or more within-subject contrasts that involve the repeated factors. There is no limit to the number of repeated factors that you can specify.

Usually, the variables on the left side of the equation in the MODEL statement represent one repeated response variable. This does not mean that you are limited to listing only one factor in the REPEATED statement. For example, one repeated response variable (wellness rating) might be measured six times (implying variables Y 1 to Y 6 on the left side of the equal sign in the MODEL statement), with the associated within-subject factors rater and time (implying two factors listed in the REPEATED statement). However, designs that have two or more repeated response variables can be handled by using the IDENTITY transformation.

To use this feature, you must be familiar with the details of multivariate model and contrast parameterizations that PROC GLM uses. For more information, see the sections "Repeated Measures Analysis of Variance" on page 3493 and "Multivariate Analysis of Variance" on page 3492 in Chapter 45, "The GLM Procedure." For information about the power and sample size computational methods and formulas, see the section "Contrasts in Fixed-Effect Multivariate Models" on page 3633.

If you specify one or more REPEATED statements, then a "Mean(Dep)" transformation is added to the power analysis. This transformation is the mean of the dependent variables, the same transformation that is used implicitly in the "Tests of Hypotheses for Between Subjects Effects" table in PROC GLM. In addition, the Intercept model effect is included in the power analysis. If the REPEATED statement is not specified, then tests that involve the Intercept are excluded from the power analysis.

You can use either the REPEATED statement or the MANOVA statement along with any of the tests for multivariate models that are supported in the MTEST= option in the POWER statement. The REPEATED statement is usually used for handling repeated measurements on the same experimental unit, but you can also use the REPEATED statement for other situations, such as clusters or multiple outcome variables. The differences between the REPEATED and MANOVA statements are as follows:

- You can use the REPEATED statement to specify commonly used contrasts by using keywords rather than coefficients, but you are limited to only those forms of the $\mathbf{M}$ matrix.
- You can use the MANOVA statement to construct any M matrix, but you must specify the coefficients explicitly (except for the default identity matrix).

There is no limit to the number of REPEATED statements that you can specify. Each power analysis includes tests for all REPEATED statements and also (if you specify at least one REPEATED statement) the extra "Mean(Dep)" transformation that was previously mentioned.

The simplest form of the REPEATED statement requires only a factor-name. When you have two or more repeated factors, you must specify the factor-name and number of levels (levels) for each factor. Optionally, you can specify the actual values for the levels (level-values) and a transformation that defines single-degree-of-freedom contrasts. When you specify more than one within-subject factor, the factor-names (and associated level and transformation information) must be separated by a comma in the REPEATED statement.

The factor-specification for the REPEATED statement can include any number of individual factor specifications, separated by commas, of the following form:

> factor-name levels < (level-values) > <transformation>
where
factor-name
levels gives the number of levels associated with the factor being defined. When there is only one within-subject factor, the number of levels is equal to the number of dependent variables. In this case, levels is optional. When more than one within-subject factor is defined, however, levels is required, and the product of the number of levels of all the factors must equal the number of dependent variables in the MODEL statement.
(level-values)
names a factor to be associated with the dependent variables. The name should not be the same as any variable name that already exists in the data set being analyzed and should conform to the usual conventions of SAS variable names.

When you specify more than one factor, list the dependent variables in the MODEL statement so that the within-subject factors that you define in the REPEATED statement are nested; that is, the first factor you define in the REPEATED statement should be the one whose values change least frequently.
gives values that correspond to levels of a repeated measures factor. These values are used as spacings for constructing orthogonal polynomial contrasts if you specify a POLYNOMIAL transformation. The number of values that you specify must correspond to the number of levels for that factor in the REPEATED statement. Enclose the level-values in parentheses.

The following transformation keywords define single-degree-of-freedom contrasts for factors that you specify in the REPEATED statement. Because the number of contrasts that are generated is always one less than the number of levels of the factor, you have some control over which contrast is omitted from the analysis by which transformation you select. The only exception is the IDENTITY transformation, which is not composed of contrasts and has the same degrees of freedom as the factor has levels. By default, PROC GLMPOWER uses the CONTRAST transformation.

## CONTRAST<(ordinal-reference-level) >

generates contrasts between levels of the factor and a reference level. By default, PROC GLMPOWER uses the last level as the reference level; you can optionally specify a reference level in parentheses after the keyword CONTRAST. The reference level corresponds to the ordinal value of the level rather than the level value that is specified. For example, to generate contrasts between the first level of a factor and the other levels, specify CONTRAST(1).

## HELMERT

generates contrasts between each level of the factor and the mean of subsequent levels.

## IDENTITY

generates an identity transformation that corresponds to the associated factor. This transformation is not composed of contrasts; it has $n$ degrees of freedom for an $n$-level factor, instead of $n-1$ degrees of freedom.

## MEAN < (ordinal-reference-level) >

generates contrasts between levels of the factor and the mean of all other levels of the factor. Specifying a reference level eliminates the contrast between that level and the mean. When no reference level is specified, the contrast that involves the last level is omitted. For an example, see the CONTRAST transformation.

## POLYNOMIAL

generates orthogonal polynomial contrasts. Level values, if provided, are used as spacings in the construction of the polynomials; otherwise, equal spacing is assumed.

## PROFILE

generates contrasts between adjacent levels of the factor.

## Examples

When you specify more than one factor, list the dependent variables in the MODEL statement so that the within-subject factors that you define in the REPEATED statement are nested; that is, the first factor you define in the REPEATED statement should be the one whose values change least frequently. For example, assume that two raters submit a wellness rating at each of three times, for a total of six dependent variables for each subject. Consider the following statements:

```
proc glm;
    class treatment;
    model Y1-Y6 = treatment;
    repeated rater 2, time 3;
run;
```

The variables are listed in the MODEL statement as Y1 through Y6, so the REPEATED statement in the preceding statements implies the following structure:

|  | Dependent Variables |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
|  | Y 1 | Y 2 | Y 3 | Y 4 | Y 6 |  |
| Value of rater | 1 | 1 | 1 | 2 | 2 | 2 |
| Value of time | 1 | 2 | 3 | 1 | 2 | 3 |

## WEIGHT Statement

## WEIGHT variable ;

The WEIGHT statement names a variable that provides a profile weight ("cell weight") for each observation in the exemplary data set specified by the DATA= option in the PROC GLMPOWER statement.
If the WEIGHT statement is not used, then a balanced design is assumed with default cell weights of 1 .

## Details: GLMPOWER Procedure

## Specifying Value Lists in the POWER Statement

To specify one or more scenarios for an analysis parameter (or set of parameters) in the POWER statement, you provide a list of values for the option that corresponds to the parameter(s). To identify the parameter you want to solve for, you place a missing value in the appropriate list.

There are three basic types of such lists: number-lists, name-lists, and keyword-lists. Scenarios for scalarvalued parameters, such as power, are represented by a number-list. Scenarios for named parameters, such as correlation matrices, are represented by a name-list. Some parameters, such as the test statistic for a multivariate model, have values that are represented by one or more keywords in a keyword-list.

## Number-Lists

A number-list can be one of two things: a series of one or more numbers expressed in the form of one or more DOLISTs, or a missing value indicator (. ).

The DOLIST format is the same as in the DATA step. For example, you can specify four scenarios (30,50, 70, and 100) for a total sample size in either of the following ways:

```
NTOTAL = 30 50 70 100
NTOTAL = 30 to 70 by 20 100
```

A missing value identifies a parameter as the result parameter; it is valid only with options representing parameters you can solve for in a given analysis. For example, you can request a solution for NTOTAL as follows:

```
NTOTAL = .
```


## Name-Lists

A name-list is a list of one or more names that are enclosed in single or double quotation marks and separated by spaces. For example, you can specify two scenarios for the correlation matrix in a multivariate model as follows:

```
CORRMAT = "Corr A" "Corr B"
```


## Keyword-Lists

A keyword-list is a list of one or more keywords, separated by spaces. For example, you can specify both the multivariate Hotelling-Lawley trace and uncorrected univariate $F$ test for a multivariate model as follows:

```
MTEST = HLT UNCORR
```


## Sample Size Adjustment Options

By default, PROC GLMPOWER rounds sample sizes conservatively (down in the input, up in the output) so that all total sizes and sample sizes for individual design profiles are integers. This is generally considered conservative because it selects the closest realistic design providing at most the power of the (possibly fractional) input or mathematically optimized design. In addition, all design profile sizes are adjusted to be multiples of their corresponding weights. If a design profile is present more than once in the exemplary data set, then the weights for that design profile are summed. For example, if a particular design profile is present twice in the exemplary data set with weight values 2 and 6 , then all sample sizes for this design profile become multiples of $2+6=8$.

With the NFRACTIONAL option, sample size input is not rounded, and sample size output is reported in two versions, a raw "fractional" version and a "ceiling" version rounded up to the nearest integer.
Whenever an input sample size is adjusted, both the original ("nominal") and adjusted ("actual") sample sizes are reported. Whenever computed output sample sizes are adjusted, both the original input ("nominal") power and the achieved ("actual") power at the adjusted sample size are reported.

## Error and Information Output

The Error column in the main output table explains reasons for missing results and flags numerical results that are bounds rather than exact answers.

The Info column provides further information about Error entries, warnings about any boundary conditions detected, and notes about any adjustments to input. Note that the Info column is hidden by default in the main output. You can view it by using the ODS OUTPUT statement to save the output as a data set and the PRINT procedure. For example, the following SAS statements print both the Error and Info columns for a power computation in a one-way ANOVA:

```
data MyExemp;
    input A $ Y1 Y2;
    datalines;
            1 10 11
            2 12 11
            3 15 11
;
proc glmpower data=MyExemp;
    class A;
    model Y1 Y2 = A;
    power
        stddev = 2
        ntotal = 3 10
        power = .;
    ods output output=Power;
run;
proc print noobs data=Power;
    var NominalNTotal NTotal Dependent Power Error Info;
run;
```

The output is shown in Figure 47.5.
Figure 47.5 Error and Information Columns

| NominalNTotal | NTotal | Dependent | Power | Error | Info |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 3 | 3 | Y1 |  | Invalid input | Error DF=0 |
| 10 |  | Y1 | 0.557 |  | Input N adjusted |
| 3 | 3 | Y2 |  | Invalid input | Error DF=0 / No effect |
| 10 | 9 | Y2 | 0.050 |  | Input N adjusted/ No effect |

The sample size of 3 specified with the NTOTAL= option causes an "Invalid input" message in the Error column and an "Error DF=0" message in the Info column, because a sample size of 3 is so small that there
are no degrees of freedom left for the error term. The sample size of 10 causes an "Input N adjusted" message in the Info column, because it is rounded down to 9 to produce integer group sizes of 3 per cell. The cell means scenario represented by the dependent variable Y2 causes a "No effect" message to appear in the Info column, because the means in this scenario are all equal.

## Displayed Output

If you use the PLOTONLY option in the PROC GLMPOWER statement, the procedure displays only graphical output. Otherwise, the displayed output of the GLMPOWER procedure includes the following:

- the "Fixed Scenario Elements" table, which shows all applicable single-valued analysis parameters, in the following order: the dependent variable that represents the cell means scenario (for a univariate model) or the dependent variable transformation (for a multivariate model), the source of the test (that is, the model effect or between-subject contrast), the weight variable, parameters that are input explicitly, parameters that are supplied with defaults, and ancillary results
- an output table that shows the following when applicable (in order): the index of the scenario, the dependent variable that represents the cell means scenario (for a univariate model) or the dependent variable transformation (for a multivariate model), the type of the test, the source of the test (that is, the model effect or between-subject contrast), all multivalued input, ancillary results, the primary computed result, and error descriptions
- plots (if requested)

The exception to these ordering conventions is that the DEPENDENT and EFFECTS= options can be used along with the OUTPUTORDER=SYNTAX or OUTPUTORDER=REVERSE option in the POWER statement to specify the relative location of the output for dependent variable and type and source of test.
Ancillary results include the following:

- Actual Power, the achieved power, if it differs from the input (Nominal) power value
- Actual Alpha, the achieved significance level, if it differs from the input (Nominal) alpha value
- fractional sample size, if the NFRACTIONAL option is used in the POWER statement
- test or numerator degrees of freedom in the test's critical value
- error or denominator degrees of freedom in the test's critical value
- Effect, the combination of the within-subject Transformation contrast and between-subject Source test or contrast in a multivariate model

If sample size is the result parameter and the NFRACTIONAL option is used in the POWER statement, then both "Fractional" and "Ceiling" sample size results are displayed. Fractional sample sizes correspond to the "Nominal" values of power. Ceiling sample sizes are simply the fractional sample sizes rounded up to the nearest integer; they correspond to "Actual" values of power.

The noncentrality parameter is computed and stored in a hidden column called Noncentrality in the "Output" table. If a univariate test for a multivariate model is specified (that is, one of MTEST=BOX, MTEST=GG, MTEST=HF, or MTEST=UNCORR), then the numerator and denominator degrees of freedom that are used in the noncentral $F$ approximation of the test statistic distribution are computed and stored in hidden columns called NumNCDF and DenNCDF, respectively, in the "Output" table. These are the only tests for which the degrees of freedom in the noncentral $F$ approximation of the test statistic are different from those in the critical value.

## ODS Table Names

PROC GLMPOWER assigns a name to each table that it creates. You can use these names to reference the table when using the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in Table 47.7. For more information about ODS, see Chapter 20, "Using the Output Delivery System."

Table 47.7 ODS Tables Produced by PROC GLMPOWER

| ODS Table Name | Description | Statement |
| :--- | :--- | :--- |
| FixedElements | Factoid with single-valued analysis parameters | Default |
| Output | All input and computed analysis parameters, error <br> messages, and information messages for each scenario | Default |
| PlotContent | Data contained in plots, including analysis parameters <br> and indices identifying plot features. (NOTE: This | PLOT |
|  | table is saved as a data set and not displayed in PROC |  |
|  | GLMPOWER output.) |  |

## Computational Methods and Formulas

This section describes the approaches that PROC GLMPOWER uses to compute power and sample size.

## Contrasts in Fixed-Effect Univariate Models

The univariate linear model has the form

$$
\mathrm{y}=\mathrm{X} \boldsymbol{\beta}+\boldsymbol{\epsilon}
$$

where $\mathbf{y}$ is the $N \times 1$ vector of responses, $\mathbf{X}$ is the $N \times k$ design matrix, $\boldsymbol{\beta}$ is the $k \times 1$ vector of model parameters corresponding to the columns of $\mathbf{X}$, and $\boldsymbol{\epsilon}$ is an $N \times 1$ vector of errors with

$$
\begin{equation*}
\epsilon_{1}, \ldots, \epsilon_{N} \sim \mathrm{~N}\left(0, \sigma^{2}\right) \tag{iid}
\end{equation*}
$$

In PROC GLMPOWER, the model parameters $\boldsymbol{\beta}$ are not specified directly, but rather indirectly as $\mathbf{y}^{\star}$, which represents either conjectured response means or typical response values for each design profile. The $\mathrm{y}^{\star}$ values are manifested as the dependent variable in the MODEL statement. The vector $\boldsymbol{\beta}$ is obtained from $\mathbf{y}^{\star}$ according to the least squares equation,

$$
\boldsymbol{\beta}=\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{X}^{\prime} \mathbf{y}^{\star}
$$

Note that, in general, there is not a one-to-one mapping between $\mathbf{y}^{\star}$ and $\boldsymbol{\beta}$. Many different scenarios for $\mathbf{y}^{\star}$ might lead to the same $\boldsymbol{\beta}$. If you specify $\mathbf{y}^{\star}$ with the intention of representing cell means, keep in mind that PROC GLMPOWER allows scenarios that are not valid cell means according to the model that is specified in the MODEL statement. For example, if $\mathbf{y}^{\star}$ exhibits an interaction effect but the corresponding interaction term is left out of the model, then the cell means ( $\mathbf{X} \boldsymbol{\beta}$ ) that are derived from $\boldsymbol{\beta}$ differ from $\mathbf{y}^{\star}$. In particular, the cell means that are derived in this way are the projection of $\mathbf{y}^{\star}$ onto the model space.
It is convenient in power analysis to parameterize the design matrix $\mathbf{X}$ in three parts, $\{\ddot{\mathbf{X}}, \mathbf{w}, N\}$, defined as follows:

1. The $q \times k$ essence design matrix $\ddot{\mathbf{X}}$ is the collection of unique rows of $\mathbf{X}$. Its rows are sometimes referred to as "design profiles." Here, $q \leq N$ is defined simply as the number of unique rows of $\mathbf{X}$.
2. The $q \times 1$ weight vector $\mathbf{w}$ reveals the relative proportions of design profiles, and $\mathbf{W}=\operatorname{diag}(\mathbf{w})$. Row $i$ of $\ddot{\mathbf{X}}$ is to be included in the design $w_{i}$ times for every $w_{j}$ times that row $j$ is included. The weights are assumed to be standardized (that is, they sum up to 1 ).
3. The total sample size is $N$. This is the number of rows in $\mathbf{X}$. If you gather $N w_{i}=n_{i}$ copies of the $i$ th row of $\ddot{\mathbf{X}}$, for $i=1, \ldots, q$, then you end up with $\mathbf{X}$.

The preceding quantities are derived from PROC GLMPOWER syntax as follows:

- Values for $\ddot{\mathbf{X}}, \mathbf{y}^{\star}$, and $\mathbf{w}$ are specified in the exemplary data set (from using the DATA= option in the PROC GLMPOWER statement), and the corresponding variables are identified in the CLASS, MODEL, and WEIGHT statements.
- $N$ is specified in the NTOTAL= option in the POWER statement.

It is useful to express the crossproduct matrix $\mathbf{X}^{\prime} \mathbf{X}$ in terms of these three parts,

$$
\mathbf{X}^{\prime} \mathbf{X}=N \ddot{\mathbf{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}
$$

because this expression factors out the portion $(N)$ that depends on sample size and the portion ( $\ddot{\mathbf{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}$ ) that depends only on the design structure.

A general linear hypothesis for the univariate model has the form

$$
\begin{aligned}
& H_{0}: \mathbf{L} \boldsymbol{\beta}=\boldsymbol{\theta}_{0} \\
& H_{A}: \mathbf{L} \boldsymbol{\beta} \neq \boldsymbol{\theta}_{0}
\end{aligned}
$$

where $\mathbf{L}$ is an $1 \times k$ contrast matrix with rank $r_{L}$ and $\boldsymbol{\theta}_{0}$ is the null value (usually just a vector of zeros).
Note that model effect tests are just contrasts that use special forms of $\mathbf{L}$. Thus, this scheme covers both effect tests (which are specified in the MODEL statement and the EFFECTS = option in the POWER statement) and custom contrasts (which are specified in the CONTRAST statement).

The model degrees of freedom $\mathrm{DF}_{\mathrm{M}}$ are equal to the rank of $\mathbf{X}$, denoted $r_{X}$. The error degrees of freedom $\mathrm{DF}_{\mathrm{E}}$ are equal to $N-r_{X}$. The sample size $N$ must be at least $\mathrm{DF}_{\mathrm{M}}$ plus the number of covariates.

The test statistic is

$$
F=\frac{\left(\frac{\mathrm{SS}_{\mathrm{H}}}{r_{\mathrm{L}}}\right)}{\hat{\sigma}^{2}}
$$

where

$$
\begin{aligned}
\mathrm{SS}_{\mathrm{H}} & =\frac{1}{N}\left(\mathbf{L} \hat{\boldsymbol{\beta}}-\boldsymbol{\theta}_{0}\right)^{\prime}\left(\mathbf{L}\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{L}^{\prime}\right)^{-1}\left(\mathbf{L} \hat{\boldsymbol{\beta}}-\boldsymbol{\theta}_{0}\right) \\
\hat{\boldsymbol{\beta}} & =\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{X}^{\prime} \mathbf{y} \\
\hat{\sigma}^{2} & =\frac{1}{\mathrm{DF}_{\mathrm{E}}}(\mathbf{y}-\mathbf{X} \hat{\boldsymbol{\beta}})^{\prime}(\mathbf{y}-\mathbf{X} \hat{\boldsymbol{\beta}})
\end{aligned}
$$

Under $H_{0}, F \sim F\left(r_{L}, \mathrm{DF}_{\mathrm{E}}\right)$. Under $H_{A}, F$ is distributed as $F\left(r_{L}, \mathrm{DF}_{\mathrm{E}}, \lambda\right)$ with noncentrality

$$
\lambda=N\left(\mathbf{L} \boldsymbol{\beta}-\boldsymbol{\theta}_{0}\right)^{\prime}\left(\mathbf{L}\left(\ddot{\mathbf{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}\right)^{-1} \mathbf{L}^{\prime}\right)^{-1}\left(\mathbf{L} \boldsymbol{\beta}-\boldsymbol{\theta}_{0}\right) \sigma^{-2}
$$

The value of $\sigma$ is specified in the STDDEV $=$ option in the POWER statement.
Muller and Peterson (1984) give the exact power of the test as

$$
\text { power }=P\left(F\left(r_{L}, \mathrm{DF}_{\mathrm{E}}, \lambda\right) \geq F_{1-\alpha}\left(r_{L}, \mathrm{DF}_{\mathrm{E}}\right)\right)
$$

The value of $\alpha$ is specified in the ALPHA= option in the POWER statement.
Sample size is computed by inverting the power equation.
See Muller and Benignus (1992) and O'Brien and Shieh (1992) for additional discussion.

## Adjustments for Covariates in Univariate Models

If you specify covariates in a univariate model (whether continuous or categorical), then two adjustments are made in order to compute approximate power in the presence of the covariates. Let $n_{v}$ denote the number of covariates (counting dummy variables for categorical covariates individually) as specified in the NCOVARIATES $=$ option in the POWER statement. In other words, $n_{v}$ is the total degrees of freedom used by the covariates. The adjustments are as follows:

1. The error degrees of freedom decrease by $n_{\nu}$.
2. The error standard deviation $\sigma$ shrinks by a factor of $\left(1-\rho^{2}\right)^{\frac{1}{2}}$ (if the CORRXY $=$ option is used to specify the correlation $\rho$ between covariates and response) or $(1-r)^{\frac{1}{2}}$ (if the PROPVARREDUCTION= option is used to specify the proportional reduction in total $R^{2}$ incurred by the covariates). Let $\sigma^{\star}$ represent the updated value of $\sigma$.

As a result of these changes, the power is computed as

$$
\text { power }=P\left(F\left(r_{L}, \mathrm{DF}_{\mathrm{E}}-n_{v}, \lambda^{\star}\right) \geq F_{1-\alpha}\left(r_{L}, N-r_{x}-n_{\nu}\right)\right)
$$

where $\lambda^{\star}$ is calculated using $\sigma^{\star}$ rather than $\sigma$ :

$$
\lambda^{\star}=N\left(\mathbf{L} \boldsymbol{\beta}-\boldsymbol{\theta}_{0}\right)^{\prime}\left(\mathbf{L}\left(\ddot{\mathbf{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}\right)^{-1} \mathbf{L}^{\prime}\right)^{-1}\left(\mathbf{L} \boldsymbol{\beta}-\boldsymbol{\theta}_{0}\right)\left(\sigma^{\star}\right)^{-2}
$$

## Contrasts in Fixed-Effect Multivariate Models

The multivariate model has the form

$$
\mathbf{Y}=\mathbf{X} \boldsymbol{\beta}+\boldsymbol{\epsilon}
$$

where $\mathbf{Y}$ is the $N \times p$ vector of responses, for $p>1 ; \mathbf{X}$ is the $N \times k$ design matrix; $\boldsymbol{\beta}$ is the $k \times p$ matrix of model parameters that correspond to the columns of $\mathbf{X}$ and $\mathbf{Y}$; and $\boldsymbol{\epsilon}$ is an $N \times p$ vector of errors, where

$$
\epsilon_{1}, \ldots, \epsilon_{N} \sim \mathrm{~N}(0, \boldsymbol{\Sigma}) \quad \text { (iid) }
$$

In PROC GLMPOWER, the model parameters $\boldsymbol{\beta}$ are not specified directly, but rather indirectly as $\mathbf{Y}^{\star}$, which represents either conjectured response means or typical response values for each design profile. The $\mathbf{Y}^{\star}$ values are manifested as the collection of dependent variables in the MODEL statement. The matrix $\boldsymbol{\beta}$ is obtained from $\mathbf{Y}^{\star}$ according to the least squares equation,

$$
\boldsymbol{\beta}=\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{X}^{\prime} \mathbf{Y}^{\star}
$$

Note that, in general, there is not a one-to-one mapping between $\mathbf{Y}^{\star}$ and $\boldsymbol{\beta}$. Many different scenarios for $\mathbf{Y}^{\star}$ might lead to the same $\boldsymbol{\beta}$. If you specify $\mathbf{Y}^{\star}$ with the intention of representing cell means, keep in mind that PROC GLMPOWER allows scenarios that are not valid cell means according to the model that is specified in the MODEL statement. For example, if $\mathbf{Y}^{\star}$ exhibits an interaction effect but the corresponding interaction term is left out of the model, then the cell means ( $\mathbf{X} \boldsymbol{\beta}$ ) that are derived from $\boldsymbol{\beta}$ differ from $\mathbf{Y}^{\star}$. In particular, the cell means that are derived in this way are the projection of $\mathbf{Y}^{\star}$ onto the model space.

It is convenient in power analysis to parameterize the design matrix $\mathbf{X}$ in three parts, $\{\ddot{\mathbf{X}}, \mathbf{W}, N\}$, defined as follows:

1. The $q \times k$ essence design matrix $\ddot{\mathbf{X}}$ is the collection of unique rows of $\mathbf{X}$. Its rows are sometimes referred to as "design profiles." Here, $q \leq N$ is defined simply as the number of unique rows of $\mathbf{X}$.
2. The $q \times 1$ weight vector $\mathbf{w}$ reveals the relative proportions of design profiles, and $\mathbf{W}=\operatorname{diag}(\mathbf{w})$. Row $i$ of $\ddot{\mathbf{X}}$ is to be included in the design $w_{i}$ times for every $w_{j}$ times that row $j$ is included. The weights are assumed to be standardized (that is, they sum up to 1 ).
3. The total sample size is $N$. This is the number of rows in $\mathbf{X}$. If you gather $N w_{i}=n_{i}$ copies of the $i$ th row of $\ddot{\mathbf{X}}$, for $i=1, \ldots, q$, then you end up with $\mathbf{X}$.

The preceding quantities are derived from PROC GLMPOWER syntax as follows:

- Values for $\ddot{\mathbf{X}}, \mathbf{Y}^{\star}$, and $\mathbf{w}$ are specified in the exemplary data set (from using the DATA= option in the PROC GLMPOWER statement), and the corresponding variables are identified in the CLASS, MODEL, and WEIGHT statements.
- $N$ is specified in the NTOTAL= option in the POWER statement.

It is useful to express the crossproduct matrix $\mathbf{X}^{\prime} \mathbf{X}$ in terms of these three parts,

$$
\mathbf{X}^{\prime} \mathbf{X}=N \ddot{\mathbf{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}
$$

because this expression factors out the portion $(N)$ that depends on sample size and the portion ( $\ddot{\mathbf{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}$ ) that depends only on the design structure.

A general linear hypothesis for the univariate model has the form

$$
\begin{aligned}
& H_{0}: \mathbf{L} \boldsymbol{\beta} \mathbf{M}=\boldsymbol{\theta}_{0} \\
& H_{A}: \mathbf{L} \boldsymbol{\beta} \mathbf{M} \neq \boldsymbol{\theta}_{0}
\end{aligned}
$$

where $\mathbf{L}$ is an $l \times k$ between-subject contrast matrix with $\operatorname{rank} r_{L}, \mathbf{M}$ is a $p \times m$ within-subject contrast matrix with rank $r_{M}$, and $\boldsymbol{\theta}_{0}$ is an $1 \times m$ null contrast matrix (usually just a matrix of zeros).

Note that model effect tests are just between-subject contrasts that use special forms of $\mathbf{L}$, combined with an $\mathbf{M}$ that is the $p \times 1$ mean transformation vector of the dependent variables (a vector of values all equal to $1 / p$ ). Thus, this scheme covers both effect tests (which are specified in the MODEL statement and the EFFECTS = option in the POWER statement) and custom between-subject contrasts (which are specified in the CONTRAST statement).

The $\mathbf{M}$ matrix is often referred to as the dependent variable transformation and is specified in the MANOVA or REPEATED statement.

The model degrees of freedom $\mathrm{DF}_{\mathrm{M}}$ are equal to the rank of $\mathbf{X}$, denoted $r_{X}$. The error degrees of freedom $\mathrm{DF}_{\mathrm{E}}$ are equal to $N-r_{X}$.

The hypothesis sum of squares $\mathrm{SS}_{\mathrm{H}}$ in the univariate model generalizes to the hypothesis SSCP matrix in the multivariate model,

$$
\mathbf{H}=\left(\mathbf{L} \hat{\boldsymbol{\beta}} \mathrm{M}-\boldsymbol{\theta}_{0}\right)^{\prime}\left(\mathbf{L}\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-1} \mathbf{L}^{\prime}\right)^{-1}\left(\mathbf{L} \hat{\boldsymbol{\beta}} \mathrm{M}-\boldsymbol{\theta}_{0}\right)
$$

The error sum of squares $\hat{\sigma}^{2}\left(N-r_{X}\right)$ in the univariate model generalizes to the error SSCP matrix in the multivariate model,

$$
\mathbf{E}=\left(N-r_{X}\right) \mathbf{M}^{\prime} \hat{\mathbf{\Sigma}} \mathbf{M}
$$

where

$$
\hat{\boldsymbol{\Sigma}}=(\mathbf{Y}-\mathbf{X} \hat{\boldsymbol{\beta}})^{\prime}(\mathbf{Y}-\mathbf{X} \hat{\boldsymbol{\beta}}) /\left(N-r_{X}\right)
$$

and

$$
\hat{\boldsymbol{\beta}}=\left(\mathbf{X}^{\prime} \mathbf{X}\right)^{-} \mathbf{X}^{\prime} \mathbf{Y}
$$

The population counterpart of $\mathbf{H} / N$ is

$$
\mathbf{H}^{\star}=\left(\mathbf{L} \boldsymbol{\beta} \mathbf{M}-\boldsymbol{\theta}_{0}\right)^{\prime}\left(\mathbf{L}\left(\ddot{\mathbf{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}\right)^{-} \mathbf{L}^{\prime}\right)^{-1}\left(\mathbf{L} \boldsymbol{\beta} \mathbf{M}-\boldsymbol{\theta}_{0}\right)
$$

and the population counterpart of $\mathbf{E} / N$ is

$$
\mathbf{E}^{\star}=\mathbf{M}^{\prime} \boldsymbol{\Sigma} \mathbf{M}
$$

The elements of $\Sigma$ are specified in the MATRIX= and STDDEV= options and identified in the CORRMAT=, CORRS $=$, COVMAT $=$, and SQRTVAR = options in the POWER statement.

The power and sample size computations for all the tests that are supported in the MTEST= option in the POWER statement are based on $\mathbf{H}^{\star}$ and $\mathbf{E}^{\star}$. The following two subsections cover the computational methods and formulas for the multivariate and univariate tests that are supported in the MTEST= and UEPSDEF= options in the POWER statement.

## Multivariate Tests

Power computations for multivariate tests are based on O'Brien and Shieh (1992) (for METHOD=OBRIENSHIEH) and Muller and Peterson (1984) (for METHOD=MULLERPETERSON).

Let $s=\min \left(r_{L}, r_{M}\right)$, the smaller of the between-subject and within-subject contrast degrees of freedom. Critical value computations assume that under $H_{0}$, the test statistic $F$ is distributed as $F\left(r_{L} r_{M}, \nu_{2}\right)$, where $\nu_{2}=\left(N-r_{X}\right)-r_{M}+1$ if $s=1$ but depends on the choice of test if $s>1$. Power computations assume that under $H_{A}, F$ is distributed as $F\left(r_{L} r_{M}, \nu_{2}, \lambda\right)$, where the noncentrality $\lambda$ depends on $r_{L}, r_{M}$, the choice of test, and the power computation method.

Formulas for the test statistic $F$, denominator degrees of freedom $\nu_{2}$, and noncentrality $\lambda$ for all combinations of dimensions, tests, and methods are given in the following subsections.

The power in each case is computed as

$$
\text { power }=P\left(F\left(r_{L} r_{M}, \nu_{2}, \lambda\right) \geq F_{1-\alpha}\left(r_{L} r_{M}, \nu_{2}\right)\right)
$$

Computed power is exact for some cases and approximate for others. Sample size is computed by inverting the power equation.

Let $\boldsymbol{\Delta}=\mathbf{E}^{-1} \mathbf{H}$, and define $\boldsymbol{\phi}$ as the $s \times 1$ vector of ordered positive eigenvalues of $\boldsymbol{\Delta}, \boldsymbol{\phi}=\left\{\phi_{1}, \ldots, \phi_{s}\right\}$, where $\phi_{1} \geq \cdots \geq \phi_{s}>0$. The population equivalent is

$$
\begin{aligned}
\Delta^{\star} & =\mathbf{E}^{\star-1} \mathbf{H}^{\star} \\
& =\left(\mathbf{M}^{\prime} \boldsymbol{\Sigma} M\right)^{-1}\left(\mathbf{L} \boldsymbol{\beta} \mathbf{M}-\boldsymbol{\theta}_{0}\right)^{\prime}\left(\mathbf{L}\left(\ddot{\mathrm{X}}^{\prime} \mathbf{W} \ddot{\mathbf{X}}\right)^{-1} \mathbf{L}^{\prime}\right)^{-1}\left(\mathbf{L} \boldsymbol{\beta} \mathbf{M}-\boldsymbol{\theta}_{0}\right)
\end{aligned}
$$

where $\phi^{\star}$ is the $s \times 1$ vector of ordered positive eigenvalues of $\Delta^{\star}, \phi^{\star}=\left\{\phi_{1}^{\star}, \ldots, \phi_{s}^{\star}\right\}$ for $\phi_{1}^{\star} \geq \cdots \geq$ $\phi_{s}^{\star}>0$.

## Case 1: $s=1$

When $s=1$, all three multivariate tests (MTEST=HLT, MTEST=PT, and MTEST=WILKS) are equivalent. The test statistic is $F=\phi_{1} \nu_{2} /\left(r_{L} r_{M}\right)$, where $\nu_{2}=\left(N-r_{X}\right)-r_{M}+1$.

When the dependent variable transformation has a single degree of freedom ( $r_{M}=1$ ), METHOD=OBRIENSHIEH and METHOD=MULLERPETERSON are the same, computing exact power by using noncentrality $\lambda=N \Delta^{\star}$. The sample size must satisfy $N \geq r_{X}+1$.

When the dependent variable transformation has more than one degree of freedom but the between-subject contrast has a single degree of freedom $\left(r_{M}>1, r_{L}=1\right)$, METHOD=OBRIENSHIEH computes exact power by using noncentrality $\lambda=N \phi_{1}^{\star}$, and METHOD=MULLERPETERSON computes approximate power by using

$$
\lambda=\frac{\left(N-r_{X}\right)-r_{M}+1}{\left(N-r_{X}\right)} N \phi_{1}^{\star}
$$

The sample size must satisfy $N \geq r_{X}+r_{M}$.
Case 2: $s>1$
When both the dependent variable transformation and the between-subject contrast have more than one degree of freedom $(s>1)$, METHOD=OBRIENSHIEH computes the noncentrality as $\lambda=N \lambda^{\star}$, where $\lambda^{\star}$ is the primary noncentrality. The form of $\lambda^{\star}$ depends on the choice of test statistic.

METHOD $=$ MULLERPETERSON computes the noncentrality as $\lambda=\nu_{2} \lambda^{(\mathrm{MP})^{\star}}$, where $\lambda^{(\mathrm{MP})^{\star}}$ has the same form as $\lambda^{\star}$ except that $\phi^{\star}$ is replaced by

$$
\boldsymbol{\phi}^{(\mathrm{MP})^{\star}}=\frac{N}{\left(N-r_{X}\right)} \boldsymbol{\phi}^{\star}
$$

Computed power is approximate for both methods when $s>1$.

## Hotelling-Lawley Trace (MTEST=HLT) When $s>1$

If $N>r_{X}+r_{M}+1$, then the denominator degrees of freedom for the Hotelling-Lawley trace are $\nu_{2}=\nu_{2 a}$,

$$
\nu_{2 a}=4+\left(r_{L} r_{M}+2\right) g
$$

where

$$
g=\frac{\left(N-r_{X}\right)^{2}-\left(N-r_{X}\right)\left(2 r_{M}+3\right)+r_{M}\left(r_{M}+3\right)}{\left(N-r_{X}\right)\left(r_{L}+r_{M}+1\right)-\left(r_{L}+2 r_{M}+r_{M}^{2}-1\right)}
$$

which is the same as $v_{2}^{\left(T_{2}\right)}$ in O'Brien and Shieh (1992) and is due to McKeon (1974).
If $N \leq r_{X}+r_{M}+1$, then $\nu_{2}=\nu_{2 b}$,

$$
v_{2 b}=s\left(\left(N-r_{X}\right)-r_{M}-1\right)+2
$$

which is the same as both $v_{2}^{\left(T_{1}\right)}$ in O'Brien and Shieh (1992) and $\nu_{2}$ in Muller and Peterson (1984) and is due to Pillai and Samson (1959).

The primary noncentrality is

$$
\lambda^{\star}=\sum_{i=1}^{s} \phi_{i}^{\star}
$$

The sample size must satisfy

$$
N \geq r_{X}+r_{M}+1-1 / s
$$

If $N>r_{X}+r_{M}+1$, then the test statistic is

$$
F=\frac{U / v_{1}}{c / v_{2 a}}
$$

where

$$
\begin{aligned}
U & =\operatorname{trace}\left(\mathbf{E}^{-1} \mathbf{H}\right) \\
& =\sum_{i=1}^{s} \phi_{i}
\end{aligned}
$$

and

$$
c=\frac{2+\left(r_{L} r_{M}+2\right) g}{N-r_{X}-r_{M}-1}
$$

If $N \leq r_{X}+r_{M}+1$, then the test statistic is

$$
F=\frac{U / v_{1}}{s / v_{2 b}}
$$

## Pillai's Trace (MTEST=PT) When $s>1$

The denominator degrees of freedom for Pillai's trace are

$$
\nu_{2}=s\left(\left(N-r_{X}\right)+s-r_{M}\right)
$$

The primary noncentrality is

$$
\lambda^{\star}=s\left(\frac{\sum_{i=1}^{s} \frac{\phi_{i}^{\star}}{1+\phi_{i}^{\star}}}{s-\sum_{i=1}^{s} \frac{\phi_{i}^{\star}}{1+\phi_{i}^{\star}}}\right)
$$

The sample size must satisfy

$$
N \geq r_{X}+r_{M}+1 / s-s
$$

The test statistic is

$$
F=\frac{V / \nu_{1}}{(s-V) / \nu_{2}}
$$

where

$$
\begin{aligned}
V & =\operatorname{trace}\left(\mathbf{H}(\mathbf{H}+\mathbf{E})^{-1}\right) \\
& =\sum_{i=1}^{s} \frac{\phi_{i}}{1+\phi_{i}}
\end{aligned}
$$

## Wilks' Lambda (MTEST=WILKS) When $s>1$

The denominator degrees of freedom for Wilks' lambda are

$$
\nu_{2}=t\left[\left(N-r_{X}\right)-0.5\left(r_{M}-r_{L}+1\right)\right]-0.5\left(r_{L} r_{M}-2\right)
$$

where

$$
t= \begin{cases}1 & \text { if } r_{L} r_{M} \leq 3 \\ {\left[\frac{\left(r_{L} r_{M}\right)^{2}-4}{r_{L}^{2}+r_{M}^{2}-5}\right]^{\frac{1}{2}}} & \text { if } r_{L} r_{M} \geq 4\end{cases}
$$

The primary noncentrality is

$$
\lambda^{\star}=t\left[\left(\prod_{i=1}^{s}\left[\left(1+\phi_{i}^{\star}\right)^{-1}\right]\right)^{-\frac{1}{t}}-1\right]
$$

The sample size must satisfy

$$
N \geq\left(1+0.5\left(r_{L} r_{M}-2\right)\right) / t+r_{X}+\left(r_{M}-r_{L}+1\right) / 2
$$

The test statistic is

$$
F=\frac{\left(1-\Lambda^{1 / t}\right) / \nu_{1}}{\Lambda^{1 / t} / \nu_{2}}
$$

where

$$
\begin{aligned}
\Lambda & =\operatorname{det}(\mathbf{E}) / \operatorname{det}(\mathbf{H}+\mathbf{E}) \\
& =\prod_{i=1}^{s}\left[\left(1+\phi_{i}\right)^{-1}\right]
\end{aligned}
$$

## Univariate Tests

Power computations for univariate tests are based on Muller et al. (2007) and Muller and Barton (1989).
The test statistic is

$$
F=\frac{\operatorname{trace}(\mathbf{H}) / r_{L}}{\operatorname{trace}(\mathbf{E}) /\left(N-r_{X}\right)}
$$

Critical value computations assume that under $H_{0}, F$ is distributed as $F\left(\nu_{1}, \nu_{2}\right)$, where $\nu_{1}$ and $\nu_{2}$ depend on the choice of test.

The four tests for the univariate approach to repeated measures differ in their assumptions about the sphericity $\varepsilon$ of $\mathbf{E}^{\star}$,

$$
\varepsilon=\frac{\operatorname{trace}^{2}\left(\mathbf{E}^{\star}\right)}{r_{M} \operatorname{trace}\left(\mathbf{E}^{\star 2}\right)}
$$

Power computations assume that under $H_{A}, F$ is distributed as $F\left(v_{1}^{\star}, v_{2}^{\star}, \lambda\right)$.
Formulas for $\nu_{1}$ and $\nu_{2}$ for each test and formulas for $\nu_{1}^{\star}, \nu_{2}^{\star}$, and $\lambda$ are given in the following subsections. The power in each case is approximated as

$$
\text { power }=P\left(F\left(v_{1}^{\star}, \nu_{2}^{\star}, \lambda\right) \geq F_{1-\alpha}\left(\nu_{1}, \nu_{2}\right)\right)
$$

Sample size is computed by inverting the power equation.
The sample size must be large enough to yield $\nu_{1}>0, \nu_{1}^{\star}>0, \nu_{2} \geq 1$, and $\nu_{2}^{\star} \geq 1$.
Because these univariate tests are biased, the achieved significance level might differ from the nominal significance level. The actual alpha is computed in the same way as the power, except that the noncentrality parameter $\lambda$ is set to 0 .
Define $\phi^{(\mathrm{E})}$ as the vector of ordered eigenvalues of $\mathbf{E}^{\star}, \boldsymbol{\phi}^{(\mathrm{E})}=\left\{\phi_{1}^{(\mathrm{E})}, \ldots, \phi_{r_{M}}^{(\mathrm{E})}\right\}$, where $\phi_{1}^{(\mathrm{E})} \geq \cdots \geq \phi_{r_{M}}^{(\mathrm{E})}$, and define $\boldsymbol{\gamma}_{j}^{(\mathrm{E})}$ as the $j$ th eigenvector of $\mathbf{E}^{\star}$. Critical values and power computations are based on the following intermediate parameters:

$$
\omega_{* j}=N\left(\boldsymbol{\gamma}_{j}^{(\mathrm{E})}\right)^{\prime} \mathbf{H}^{\star} \boldsymbol{\gamma}_{j}^{(\mathrm{E})} / \phi_{j}^{(\mathrm{E})}
$$

$$
\begin{aligned}
& S_{t 1}=\sum_{j=1}^{r_{M}} \phi_{j}^{(\mathrm{E})} \\
& S_{t 2}=\sum_{j=1}^{r_{M}} \phi_{j}^{(\mathrm{E})} \omega_{* j} \\
& S_{t 3}=\sum_{j=1}^{r_{M}}\left(\phi_{j}^{(\mathrm{E})}\right)^{2} \\
& S_{t 4}=\sum_{j=1}^{r_{M}}\left(\phi_{j}^{(\mathrm{E})}\right)^{2} \omega_{* j} \\
& R_{* 1}=\frac{r_{L} S_{t 3}+2 S_{t 4}}{r_{L} S_{t 1}+2 S_{t 2}} \\
& R_{* 2}=\frac{S_{t 3}}{S_{t 1}} \\
& \mathrm{E}\left(t_{1}\right)=2\left(N-r_{X}\right) S_{t 3}+\left(N-r_{X}\right)^{2} S_{t 1}^{2} \\
& \mathrm{E}\left(t_{2}\right)=\left(N-r_{X}\right)\left(\left(N-r_{X}\right)+2\right) S_{t 3}+2\left(N-r_{X}\right) \sum_{j_{1}=2}^{r_{M}} \sum_{j_{2}=1}^{j_{1}-1} \phi_{j_{1}}^{(\mathrm{E})} \phi_{j_{2}}^{(\mathrm{E})}
\end{aligned}
$$

The degrees of freedom and noncentrality in the noncentral $F$ approximation of the test statistic are computed as follows:

$$
\begin{aligned}
& \nu_{1}^{\star}=\frac{r_{L} S_{t 1}}{R_{* 1}} \\
& \nu_{2}^{\star}=\frac{\left(N-r_{X}\right) S_{t 1}}{R_{* 2}} \\
& \lambda=\frac{S_{t 2}}{R_{* 1}}
\end{aligned}
$$

## Uncorrected Test

The uncorrected test assumes sphericity $\varepsilon=1$, in which case the null $F$ distribution is exact, with the following degrees of freedom:

$$
\begin{aligned}
& \nu_{1}=r_{L} r_{M} \\
& \nu_{2}=r_{M}\left(N-r_{X}\right)
\end{aligned}
$$

## Greenhouse-Geisser Adjustment (MTEST=UNCORR)

The Greenhouse-Geisser adjustment to the uncorrected test reduces degrees of freedom by the MLE $\hat{\varepsilon}$ of the sphericity,

$$
\hat{\varepsilon}=\frac{\operatorname{trace}^{2}(\mathbf{E})}{r_{M} \operatorname{trace}\left(\mathbf{E}^{2}\right)}
$$

An approximation for the expected value of $\hat{\varepsilon}$ is used to compute the degrees of freedom for the null $F$ distribution,

$$
\begin{aligned}
& \nu_{1}=r_{L} r_{M} \mathrm{E}(\hat{\varepsilon}) \\
& \nu_{2}=r_{M}\left(N-r_{X}\right) \mathrm{E}(\hat{\varepsilon})
\end{aligned}
$$

where

$$
\mathrm{E}(\hat{\varepsilon})=\frac{\mathrm{E}\left(t_{1}\right)}{r_{M} \mathrm{E}\left(t_{2}\right)}
$$

## Huynh-Feldt Adjustments (MTEST=HF)

The Huynh-Feldt adjustment reduces degrees of freedom by a nearly unbiased estimate $\tilde{\varepsilon}$ of the sphericity,

$$
\tilde{\varepsilon}= \begin{cases}\frac{N r_{M} \hat{\varepsilon}-2}{r_{M}\left[\left(N-r_{X}\right)-r_{M} \hat{\hat{\varepsilon}}\right]} & \text { if UEPSDEF=HF } \\ \frac{\left.N-r_{X}+1\right) r_{M} \hat{\varepsilon}-2}{r_{M}\left[\left(N-r_{X}\right)-r_{M} \hat{\varepsilon}\right]} & \text { if UEPSDEF=HFL } \\ \left(\frac{\left(\nu_{a}-2\right)\left(\nu_{a}-4\right)}{\nu_{a}^{2}}\right)\left(\frac{\left(N-r_{X}+1\right) r_{M} \hat{\varepsilon}-2}{r_{M}\left[\left(N-r_{X}\right)-r_{M} \hat{\varepsilon}\right]}\right) & \text { if UEPSDEF=CM }\end{cases}
$$

where

$$
v_{a}=\left(N-r_{X}-1\right)+\left(N-r_{X}\right)\left(N-r_{X}-1\right) / 2
$$

The value of $\tilde{\varepsilon}$ is truncated if necessary to be at least $1 / r_{M}$ and at most 1 .
An approximation for the expected value of $\tilde{\varepsilon}$ is used to compute the degrees of freedom for the null $F$ distribution,

$$
\begin{aligned}
& \nu_{1}=r_{L} r_{M} \mathrm{E}_{\mathrm{t}}(\tilde{\varepsilon}) \\
& \nu_{2}=r_{M}\left(N-r_{X}\right) \mathrm{E}_{\mathrm{t}}(\tilde{\varepsilon})
\end{aligned}
$$

where

$$
\mathrm{E}_{\mathrm{t}}(\tilde{\varepsilon})=\min \left(\max \left(\mathrm{E}(\tilde{\varepsilon}), 1 / r_{M}\right), 1\right)
$$

and

## Box Conservative Test (MTEST=BOX)

The Box conservative test assumes the worst case for sphericity, $\varepsilon=1 / r_{M}$, leading to the following degrees of freedom for the null $F$ distribution:

$$
\begin{aligned}
& \nu_{1}=r_{L} \\
& \nu_{2}=\left(N-r_{X}\right)
\end{aligned}
$$

## 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, by specifying the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 606 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 605 in Chapter 21, "Statistical Graphics Using ODS."

If ODS Graphics is not enabled, then PROC GLMPOWER creates traditional graphics.
You can reference every graph produced through ODS Graphics with a name. The names of the graphs that PROC GLMPOWER generates are listed in Table 47.8, along with the required statements and options.

Table 47.8 Graphs Produced by PROC GLMPOWER

| ODS Graph Name | Plot Description | Option |
| :--- | :--- | :--- |
| PowerPlot | Plot with power and sample <br> size on the axes | PLOT |
| PowerAbort | Empty plot that shows an error <br> message when a plot could not <br> be produced | PLOT |

## Examples: GLMPOWER Procedure

## Example 47.1: One-Way ANOVA

This example deals with the same situation as in Example 77.1 in Chapter 77, "The POWER Procedure."
Hocking (1985, p. 109) describes a study of the effectiveness of electrolytes in reducing lactic acid buildup for long-distance runners. You are planning a similar study in which you will allocate five different fluids to runners on a 10 -mile course and measure lactic acid buildup immediately after the race. The fluids consist of water and two commercial electrolyte drinks, EZDure and LactoZap, each prepared at two concentrations, low (EZD1 and LZ1) and high (EZD2 and LZ2).

You conjecture that the standard deviation of lactic acid measurements given any particular fluid is about 3.75, and that the expected lactic acid values will correspond roughly to Table 47.9. You are least familiar with the LZ1 drink and hence decide to consider a range of reasonable values for that mean.

Table 47.9 Mean Lactic Acid Buildup by Fluid

| Water | EZD1 | EZD2 | LZ1 | LZ2 |
| :---: | :---: | :---: | :---: | :---: |
| 35.6 | 33.7 | 30.2 | 29 or 28 | 25.9 |

You are interested in four different comparisons, shown in Table 47.10 with appropriate contrast coefficients.
Table 47.10 Planned Comparisons

|  | Contrast Coefficients |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
| Comparison | Water | EZD1 | EZD2 | LZ1 | LZ2 |
| Water versus electrolytes | 4 | -1 | -1 | -1 | -1 |
| EZD versus LZ | 0 | 1 | 1 | -1 | -1 |
| EZD1 versus EZD2 | 0 | 1 | -1 | 0 | 0 |
| LZ1 versus LZ2 | 0 | 0 | 0 | 1 | -1 |

For each of these contrasts you want to determine the sample size required to achieve a power of 0.9 for detecting an effect with magnitude in accord with Table 47.9. You are not yet attempting to choose a single sample size for the study, but rather checking the range of sample sizes needed for individual contrasts. You plan to test each contrast at $\alpha=0.025$. In the interests of reducing costs, you will provide twice as many runners with water as with any of the electrolytes; that is, you will use a sample size weighting scheme of 2:1:1:1:1.
Before calling PROC GLMPOWER, you need to create the exemplary data set to specify means and weights for the design profiles:

```
data Fluids;
    input Fluid $ LacticAcid1 LacticAcid2 CellWgt;
    datalines;
```



```
            EZD1 33.7 33.7 1
            EZD2 30.2 30.2 1
            LZ1 29 28 1
            LZ2 25.9 25.9 1
;
```

The variable LacticAcid1 represents the cell means scenario with the larger LZ1 mean (29), and LacticAcid2 represents the scenario with the smaller LZ1 mean (28). The variable CellWgt contains the sample size allocation weights.

Use the DATA= option in the PROC GLMPOWER statement to specify Fluids as the exemplary data set. The following statements perform the sample size analysis:

```
proc glmpower data=Fluids;
    class Fluid;
    model LacticAcid1 LacticAcid2 = Fluid;
    weight CellWgt;
    contrast "Water vs. others" Fluid -1 -1 -1 -1 4;
```

```
    contrast "EZD vs. LZ" Fluid 1 1 -1 -1 0;
    contrast "EZD1 vs. EZD2" Fluid 1 -1 0 0 0;
    contrast "LZ1 vs. LZ2" Fluid 0 0 1 -1 0;
    power
    stddev = 3.75
    alpha = 0.025
    ntotal = .
    power = 0.9;
run;
```

The CLASS statement identifies Fluid as a classification variable. The MODEL statement specifies the model and the two cell means scenarios LacticAcid1 and LacticAcid2. The WEIGHT statement identifies CellWgt as the weight variable. The CONTRAST statement specifies the contrasts. Since PROC GLMPOWER by default processes class levels in order of formatted values, the contrast coefficients correspond to the following order: EZD1, EZD2, LZ1, LZ2, Water. (Note: You could use the ORDER=DATA option in the PROC GLMPOWER statement to achieve the same ordering as in Table 47.10 instead.) The POWER statement specifies total sample size as the result parameter and provides values for the other analysis parameters (error standard deviation, alpha, and power).

Output 47.1.1 displays the results.
Output 47.1.1 Sample Sizes for One-Way ANOVA Contrasts
The GLMPOWER Procedure

| Fixed Scenario Elements |  |
| :--- | ---: |
| Weight Variable | CellWgt |
| Alpha | 0.025 |
| Error Standard Deviation | 3.75 |
| Nominal Power | 0.9 |


| Computed N Total |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Dependent | Type | Source | Test DF | $\begin{aligned} & \text { Error } \\ & \text { DF } \end{aligned}$ | Actual Power | $\underset{\text { Total }}{\mathrm{N}}$ |
| 1 | LacticAcid1 | Effect | Fluid | 4 | 25 | 0.958 | 30 |
| 2 | LacticAcid1 | Contrast | Water vs. others | 1 | 25 | 0.947 | 30 |
| 3 | LacticAcid1 | Contrast | EZD vs. LZ | 1 | 55 | 0.929 | 60 |
| 4 | LacticAcid1 | Contrast | EZD1 vs. EZD2 | 1 | 169 | 0.901 | 174 |
| 5 | LacticAcid1 | Contrast | LZ1 vs. LZ2 | 1 | 217 | 0.902 | 222 |
| 6 | LacticAcid2 | Effect | Fluid | 4 | 25 | 0.972 | 30 |
| 7 | LacticAcid2 | Contrast | Water vs. others | 1 | 19 | 0.901 | 24 |
| 8 | LacticAcid2 | Contrast | EZD vs. LZ | 1 | 43 | 0.922 | 48 |
| 9 | LacticAcid2 | Contrast | EZD1 vs. EZD2 | 1 | 169 | 0.901 | 174 |
|  | LacticAcid2 | Contrast | LZ1 vs. LZ2 | 1 | 475 | 0.902 | 480 |

The sample sizes range from 24 for the comparison of water versus electrolytes to 480 for the comparison of LZ1 versus LZ2, both assuming the smaller LZ1 mean. The sample size for the latter comparison is relatively large because the small mean difference of $28-25.9=2.1$ is hard to detect. PROC GLMPOWER also includes the effect test for Fluid. Note that, in this case, it is equivalent to TEST=OVERALL_F in the ONEWAYANOVA statement of PROC POWER, since there is only one effect in the model.

The Nominal Power of 0.9 in the "Fixed Scenario Elements" table in Output 47.1.1 represents the input target power, and the Actual Power column in the "Computed N Total" table is the power at the sample size ( N Total) adjusted to achieve the specified sample weighting. Note that all of the sample sizes are rounded up to multiples of 6 to preserve integer group sizes (since the group weights add up to 6). You can use the NFRACTIONAL option in the POWER statement to compute raw fractional sample sizes.

Suppose you want to plot the required sample size for the range of power values from 0.5 to 0.95 . First, define the analysis by specifying the same statements as before, but add the PLOTONLY option to the PROC GLMPOWER statement to disable the nongraphical results. Next, specify the PLOT statement with $\mathrm{X}=\mathrm{POWER}$ to request a plot with power on the X axis. (The result parameter-here sample size-is always plotted on the other axis.) Use the $\mathrm{MIN}=$ and MAX= options in the PLOT statement to specify the power range. The following statements produce the plot:

```
ods graphics on;
proc glmpower data=Fluids plotonly;
    class Fluid;
    model LacticAcid1 LacticAcid2 = Fluid;
    weight CellWgt;
    contrast "Water vs. others" Fluid -1 -1 -1 -1 4;
    contrast "EZD vs. LZ" Fluid 1 1 -1 -1 0;
    contrast "EZD1 vs. EZD2" Fluid 1 -1 0 0 0;
    contrast "LZ1 vs. LZ2" Fluid 0 0 1 -1 0;
    power
        stddev = 3.75
        alpha = 0.025
        ntotal = .
        power = 0.9;
    plot x=power min=. 5 max=.95;
run;
```

Output 47.1.2 Plot of Sample Size versus Power for One-Way ANOVA Contrasts


In Output 47.1.2, the line style identifies the cell means scenario, and the plotting symbol identifies the test. The plotting symbol locations identify actual computed powers; the curves are linear interpolations of these points. The plot shows that the required sample size is highest for the test of LZ1 versus LZ2, which was previously found to require the most resources.

Note that some of the plotted points in Output 47.1.2 are unevenly spaced. This is because the plotted points are the rounded sample size results at their corresponding actual power levels. The range specified with the MIN= and MAX= values in the PLOT statement corresponds to nominal power levels. In some cases, actual power is substantially higher than nominal power. To obtain plots with evenly spaced points (but with fractional sample sizes at the computed points), you can use the NFRACTIONAL option in the POWER statement preceding the PLOT statement.

Finally, suppose you want to plot the power for the range of sample sizes you will likely consider for the study (the range of 24 to 480 that achieves 0.9 power for different comparisons). In the POWER statement, identify power as the result (POWER=.), and specify any total sample size value (say, NTOTAL=100). Specify the PLOT statement with $\mathrm{X}=\mathrm{N}$ to request a plot with sample size on the X axis.

The following statements produce the plot:

```
proc glmpower data=Fluids plotonly;
    class Fluid;
    model LacticAcid1 LacticAcid2 = Fluid;
    weight CellWgt;
    contrast "Water vs. others" Fluid -1 -1 -1 -1 4;
    contrast "EZD vs. LZ" Fluid 1 1 -1 -1 0;
    contrast "EZD1 vs. EZD2" Fluid 1 -1 0 0 0;
    contrast "LZ1 vs. LZ2" Fluid 0 0 1 -1 0;
    power
            stddev = 3.75
            alpha = 0.025
            ntotal = 24
            power = .;
    plot x=n min=24 max=480;
run;
ods graphics off;
```

Note that the value 100 specified with the NTOTAL=100 option is not used. It is overridden in the plot by the MIN = and MAX = options in the PLOT statement, and the PLOTONLY option in the PROC GLMPOWER statement disables nongraphical results. But the NTOTAL= option (along with a value) is still needed in the POWER statement as a placeholder, to identify the desired parameterization for sample size.

See Output 47.1.3 for the plot.

Output 47.1.3 Plot of Power versus Sample Size for One-Way ANOVA Contrasts


Although Output 47.1.2 and Output 47.1.3 surface essentially the same computations for practical power ranges, they each provide a different quick visual assessment. Output 47.1.2 reveals the range of required sample sizes for powers of interest, and Output 47.1.3 reveals the range of achieved powers for sample sizes of interest.

## Example 47.2: Two-Way ANOVA with Covariate

Suppose you can enhance the planned study discussed in Example 47.1 in two ways:

- incorporate results from races at two different altitudes ("high" and "low")
- measure the body mass index of each runner before the race

This is equivalent to adding a second fixed effect and a continuous covariate to your model.
Since lactic acid buildup is more pronounced at higher altitudes, you will include altitude as a factor in the model along with fluid, extending the one-way ANOVA to a two-way ANOVA. In doing so, you expect to
lower the residual standard deviation from about 3.75 to 3.5 (in addition to generalizing the study results). You assume there is negligible interaction between fluid and altitude and plan to use a main-effects-only model. You conjecture that the mean lactic acid buildup follows Table 47.11.

Table 47.11 Mean Lactic Acid Buildup by Fluid and Altitude

|  | Fluid |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Altitude | Water | EZD1 | EZD2 | LZ1 | LZ2 |
| High | 36.9 | 35.0 | 31.5 | 30 | 27.1 |
| Low | 34.3 | 32.4 | 28.9 | 27 | 24.7 |

By including a measurement of body mass index as a covariate in the study, you hope to further reduce the error variability. The extent of this reduction in variability is commonly expressed in two alternative ways: (1) the correlation between the covariates and the response or (2) the proportional reduction in total R square incurred by the covariates. You prefer the former and guess that the correlation between body mass index and lactic acid buildup is between 0.2 and 0.3 . You specify these estimates with the NCOVARIATES $=$ and CORRXY = options in the POWER statement. The covariate is not included in the MODEL statement.
You are interested in the same four fluid comparisons as in Example 47.1, shown in Table 47.10, except this time you want to marginalize over the effect of altitude.

For each of these contrasts, you want to determine the sample size required to achieve a power of 0.9 to detect an effect with magnitude according to Table 47.11. You are not yet attempting to choose a single sample size for the study, but rather checking the range of sample sizes needed by individual contrasts. You plan to test each contrast at $\alpha=0.025$. You will provide twice as many runners with water as with any of the electrolytes, and you predict that you can study approximately two-thirds as many runners at high altitude than at low altitude. The resulting planned sample size weighting scheme is shown in Table 47.12. Since the scheme is only approximate, you use the NFRACTIONAL option in the POWER statement to disable the rounding of sample sizes up to integers satisfying the weights exactly.

Table 47.12 Approximate Sample Size Allocation Weights

|  | Fluid |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: |
| Altitude | Water | EZD1 | EZD2 | LZ1 | LZ2 |
| High | 4 | 2 | 2 | 2 | 2 |
| Low | 6 | 3 | 3 | 3 | 3 |

First, you create the exemplary data set to specify means and weights for the design profiles:

```
data Fluids2;
    input Altitude $ Fluid $ LacticAcid CellWgt;
    datalines;
\begin{tabular}{llll} 
High & Water & 36.9 & 4 \\
High & EZD1 & 35.0 & 2 \\
High & EZD2 & 31.5 & 2 \\
High & LZ1 & 30 & 2 \\
High & LZ2 & 27.1 & 2
\end{tabular}
```

| Low | Water | 34.3 | 6 |
| :--- | :--- | :--- | :--- |
| Low | EZD1 | 32.4 | 3 |
| Low | EZD2 | 28.9 | 3 |
| Low | LZ1 | 27 | 3 |
| Low | LZ2 | 24.7 | 3 |

The variables Altitude, Fluid, and LacticAcid specify the factors and cell means in Table 47.11. The variable CellWgt contains the sample size allocation weights in Table 47.12.

Use the DATA= option in the PROC GLMPOWER statement to specify Fluids2 as the exemplary data set. The following statements perform the sample size analysis:

```
proc glmpower data=Fluids2;
    class Altitude Fluid;
    model LacticAcid = Altitude Fluid;
    weight CellWgt;
    contrast "Water vs. others" Fluid -1 -1 -1 -1 4;
    contrast "EZD vs. LZ" Fluid 1 1 -1 -1 0;
    contrast "EZD1 vs. EZD2" Fluid 1 -1 0 0 0;
    contrast "LZ1 vs. LZ2" Fluid 0 0 1 -1 0;
    power
        nfractional
        stddev = 3.5
        ncovariates = 1
        corrxy = 0.2 0.3 0
        alpha = 0.025
        ntotal = .
        power = 0.9;
run;
```

The CLASS statement identifies Altitude and Fluid as classification variables. The MODEL statement specifies the model, and the WEIGHT statement identifies CellWgt as the weight variable. The CONTRAST statement specifies the contrasts in Table 47.10. As in Example 47.1, the order of the contrast coefficients corresponds to the formatted class levels (EZD1, EZD2, LZ1, LZ2, Water). The POWER statement specifies total sample size as the result parameter and provides values for the other analysis parameters. The NCOVARIATES= option specifies the single covariate (body mass index), and the CORRXY= option specifies the two scenarios for its correlation with lactic acid buildup ( 0.2 and 0.3 ). Output 47.2.1 displays the results.

## Output 47.2.1 Sample Sizes for Two-Way ANOVA Contrasts

The GLMPOWER Procedure

| Fixed Scenario Elements |  |
| :--- | ---: |
| Dependent Variable | LacticAcid |
| Weight Variable | CellWgt |
| Alpha | 0.025 |
| Number of Covariates | 1 |
| Std Dev Without Covariate Adjustment | 3.5 |
| Nominal Power | 0.9 |

Output 47.2.1 continued

| Computed Ceiling N Total |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Type | Source | $\begin{gathered} \text { Corr } \\ \text { XY } \end{gathered}$ | Adj Std <br> Dev | Test DF | $\begin{aligned} & \text { Error } \\ & \text { DF } \end{aligned}$ | Fractional N Total | Actual <br> Power | Ceiling Total |
| 1 | Effect | Altitude | 0.2 | 3.43 | 1 | 84 | 90.418451 | 0.902 | 91 |
| 2 | Effect | Altitude | 0.3 | 3.34 | 1 | 79 | 85.862649 | 0.901 | 86 |
| 3 | Effect | Altitude | 0.0 | 3.50 | 1 | 88 | 94.063984 | 0.903 | 95 |
| 4 | Effect | Fluid | 0.2 | 3.43 | 4 | 16 | 22.446173 | 0.912 | 23 |
| 5 | Effect | Fluid | 0.3 | 3.34 | 4 | 15 | 21.687544 | 0.908 | 22 |
| 6 | Effect | Fluid | 0.0 | 3.50 | 4 | 17 | 23.055716 | 0.919 | 24 |
| 7 | Contrast | Water vs. others | 0.2 | 3.43 | 1 | 15 | 21.720195 | 0.905 | 22 |
| 8 | Contrast | Water vs. others | 0.3 | 3.34 | 1 | 14 | 20.848805 | 0.903 | 21 |
| 9 | Contrast | Water vs. others | 0.0 | 3.50 | 1 | 16 | 22.422381 | 0.910 | 23 |
| 10 | Contrast | EZD vs. LZ | 0.2 | 3.43 | 1 | 35 | 41.657424 | 0.903 | 42 |
| 11 | Contrast | EZD vs. LZ | 0.3 | 3.34 | 1 | 33 | 39.674037 | 0.903 | 40 |
| 12 | Contrast | EZD vs. LZ | 0.0 | 3.50 | 1 | 37 | 43.246415 | 0.906 | 44 |
| 13 | Contrast | EZD1 vs. EZD2 | 0.2 | 3.43 | 1 | 139 | 145.613657 | 0.901 | 146 |
| 14 | Contrast | EZD1 vs. EZD2 | 0.3 | 3.34 | 1 | 132 | 138.173983 | 0.902 | 139 |
| 15 | Contrast | EZD1 vs. EZD2 | 0.0 | 3.50 | 1 | 145 | 151.565917 | 0.901 | 152 |
| 16 | Contrast | LZ1 vs. LZ2 | 0.2 | 3.43 | 1 | 268 | 274.055008 | 0.901 | 275 |
| 17 | Contrast | LZ1 vs. LZ2 | 0.3 | 3.34 | 1 | 253 | 259.919126 | 0.900 | 260 |
| 18 | Contrast | LZ1 vs. LZ2 | 0.0 | 3.50 | 1 | 279 | 285.363976 | 0.901 | 286 |

The sample sizes in Output 47.2.1 range from 21 for the comparison of water versus electrolytes (assuming a correlation of 0.3 between body mass and lactic acid buildup) to 275 for the comparison of LZ1 versus LZ2 (assuming a correlation of 0.2). PROC GLMPOWER also includes the effect tests for Altitude and Fluid. Note that the required sample sizes for this study are lower than those for the study in Example 47.1.

Note that the error standard deviation has been reduced from 3.5 to 3.43 (when correlation is 0.2 ) or 3.34 (when correlation is 0.3 ) in the approximation of the effect of the body mass index covariate. The error degrees of freedom has also been automatically adjusted, lowered by 1 (the number of covariates).
Suppose you want to plot the required sample size for the range of power values from 0.5 to 0.95 . First, define the analysis by specifying the same statements as before, but add the PLOTONLY option to the PROC GLMPOWER statement to disable the nongraphical results. Next, specify the PLOT statement with $\mathrm{X}=$ POWER to request a plot with power on the X axis. Sample size is automatically placed on the Y axis. Use the MIN= and MAX= options in the PLOT statement to specify the power range. The following statements produce the plot:

```
ods graphics on;
proc glmpower data=Fluids2 plotonly;
    class Altitude Fluid;
    model LacticAcid = Altitude Fluid;
    weight CellWgt;
    contrast "Water vs. others" Fluid -1 -1 -1 -1 4;
    contrast "EZD vs. LZ" Fluid 1 1 -1 -1 0;
```

```
    contrast "EZD1 vs. EZD2" Fluid 1 -1 0 0 0;
    contrast "LZ1 vs. LZ2" Fluid 0 0 1 -1 0;
power
    nfractional
    stddev = 3.5
    ncovariates = 1
    corrxy = 0.2 0.3 0
    alpha = 0.025
    ntotal = .
        power = 0.9;
        plot x=power min=.5 max=.95;
run;
```

See Output 47.2.2 for the resulting plot.
Output 47.2.2 Plot of Sample Size versus Power for Two-Way ANOVA Contrasts


In Output 47.1.2, the line style identifies the test, and the plotting symbol identifies the scenario for the correlation between covariate and response. The plotting symbol locations identify actual computed powers; the curves are linear interpolations of these points. As in Example 47.1, the required sample size is highest for the test of LZ1 versus LZ2.

Finally, suppose you want to plot the power for the range of sample sizes you will likely consider for the study (the range of 21 to 275 that achieves 0.9 power for different comparisons). In the POWER statement, identify power as the result (POWER=.), and specify NTOTAL=21. Specify the PLOT statement with $X=N$ to request a plot with sample size on the X axis.

The following statements produce the plot:

```
proc glmpower data=Fluids2 plotonly;
    class Altitude Fluid;
    model LacticAcid = Altitude Fluid;
    weight CellWgt;
    contrast "Water vs. others" Fluid -1 -1 -1 -1 4;
    contrast "EZD vs. LZ" Fluid 1 1 -1 -1 0;
    contrast "EZD1 vs. EZD2" Fluid 1 -1 0 0 0;
    contrast "LZ1 vs. LZ2" Fluid 0 0 1 -1 0;
    power
        nfractional
        stddev = 3.5
        ncovariates = 1
        corrxy = 0.2 0.30
        alpha = 0.025
        ntotal = 21
        power = .;
    plot x=n min=21 max=275;
run;
ods graphics off;
```

The MAX=275 option in the PLOT statement sets the maximum sample size value. The MIN= option automatically defaults to the value of 21 from the NTOTAL= option in the POWER statement.

See Output 47.2.3 for the plot.

Output 47.2.3 Plot of Power versus Sample Size for Two-Way ANOVA Contrasts


Although Output 47.2.2 and Output 47.2.3 surface essentially the same computations for practical power ranges, they each provide a different quick visual assessment. Output 47.2.2 reveals the range of required sample sizes for powers of interest, and Output 47.2.3 reveals the range of powers achieved for sample sizes of interest.

## Example 47.3: Repeated Measures ANOVA

Logan, Baron, and Kohout (1995) and Guo et al. (2013) study the effect of a dental intervention on the memory of pain after root canal therapy. The intervention is a sensory focus strategy, in which patients are instructed to pay attention only to the physical sensations in their mouth during the root canal procedure.
Suppose you are interested in the long-term effects of this sensory focus intervention, because avoidance behavior has been shown to build along with memory of pain. You are planning a study to compare sensory focus to standard of care over a period of a year, asking patients to self-report their memory of pain immediately after the procedure and then again at 1 week, 6 months, and 12 months. You use a scale from 0 (no pain remembered) to 5 (maximum pain remembered).

The between-subject factor in your model is treatment, with two levels (sensory focus versus standard of care), and you allocate each treatment equally for a balanced design. The within-subject factor is time, with four levels ( $0,1,26$, and 52 weeks).

You want to determine the number of patients who are needed in order to achieve a power of 0.9 at significance level $\alpha=0.01$ for the test of the interaction between time and treatment, where the contrast over time contains all pairwise comparisons. You also want to generate a plot of power versus sample size that covers the power range of 0.05 to 0.99 .

The default Hotelling-Lawley $F$ test is appropriate for this study, especially because it is the same as the Wald test in PROC MIXED with the DDFM=KR Kenward-Roger degrees-of-freedom method and an unstructured covariance model.

You conjecture that the mean memory of pain for each treatment follows the information in Table 47.13.

Table 47.13 Mean Memory of Pain by Treatment

|  | Time Since Root Canal Therapy |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Treatment | Later on Same Day | $\mathbf{1}$ Week | 6 Months | $\mathbf{1 2}$ Months |
| Sensory Focus | 2.40 | 2.38 | 2.05 | 1.90 |
| Standard of Care | 2.40 | 2.39 | 2.36 | 2.30 |

The following statements create a data set named Pain that is to contain these means over treatment and time:

```
data Pain;
    input Treatment $ PainMem0 PainMem1Wk PainMem6Mo PainMem12Mo;
    datalines;
        SensoryFocus 2.40 2.38 2.05 1.90
        StandardOfCare 2.40 2.39 2.36 2.30
;
```

The variable Treatment specifies the two treatments. The four variables PainMem0, PainMem1Wk, PainMem6Mo, and PainMem12Mo specify the mean memory of pain scores in Table 47.13.

To characterize the variability, you must specify a set of parameters that defines the entire covariance matrix of the residuals. You conjecture that the error standard deviation is the same at all four time points, with a value somewhere between 0.92 and 1.04 , and you account for your uncertainty by including both the lower and upper ends of this range in the sample size analysis. You believe that the correlation has a linear exponent autoregressive (LEAR) structure, with a correlation of about 0.6 between measurements one week apart and a decay rate of about 0.8 over one-week intervals. The correlation matrix that contains these LEAR parameters, rounded to three decimal places, is shown in Table 47.14.

Table 47.14 Conjectured Correlation Matrix

|  | $\mathbf{0}$ | $\mathbf{1}$ | $\mathbf{2 6}$ | $\mathbf{5 2}$ |
| :---: | :---: | :---: | :---: | :---: |
| $\mathbf{0}$ | 1 | 0.6 | 0.491 | 0.399 |
| $\mathbf{1}$ | 0.6 | 1 | 0.495 | 0.402 |
| $\mathbf{2 6}$ | 0.491 | 0.495 | 1 | 0.491 |
| $\mathbf{5 2}$ | 0.399 | 0.402 | 0.491 | 1 |

Use the DATA= option in the PROC GLMPOWER statement to specify Pain as the exemplary data set. Specify the between- and within-subject factors and the model by using the CLASS, MODEL, and REPEATED statements just as you would in PROC GLM for the repeated measures data analysis. Use the POWER statement to indicate sample size as the result parameter and specify the other analysis parameters, and use the PLOT statement to generate the power curves. The following statements perform the sample size analysis:

```
ods graphics on;
proc glmpower data=Pain;
    class Treatment;
    model PainMem0 PainMem1Wk PainMem6Mo PainMem12Mo = Treatment;
    repeated Time contrast;
    power
        mtest = hlt
        alpha = 0.01
        power = . }
        ntotal = .
        stddev = 0.92 1.04
        matrix ("PainCorr") = lear(0.6, 0.8, 4, 0 1 26 52)
        corrmat = "PainCorr";
    plot y=power min=0.05 max=0.99 yopts=(ref=0.9)
        vary (linestyle by stddev, symbol by dependent source);
run;
ods graphics off;
```

The STDDEV= option specifies the two scenarios for the common residual standard deviation, 0.92 and 1.04 . The MATRIX= option defines the LEAR correlation structure, and the CORRMAT= option specifies it as the correlation matrix of the residuals. The Y=POWER option in the PLOT statement requests a plot that has power on the Y axis. (The result parameter-in this case, total sample size-is always plotted on the other axis.) The MIN $=$ and MAX $=$ options in the PLOT statement specify the power range. The YOPTS=(REF=) option adds a reference line at the target power value of 0.9. The VARY option specifies that the line style vary by the residual standard deviation and that the plotting symbol vary by the combination of within-subject and between-subject effects. The ODS GRAPHICS ON statement enables ODS Graphics.

Output 47.3.1 shows the output, and Output 47.3.2 shows the plot.
Output 47.3.1 Sample Size Analysis for Repeated Measures
The GLMPOWER Procedure F Test for Multivariate Model

| Fixed Scenario Elements |  |
| :--- | ---: |
| Wilks/HLT/PT Method | O'Brien-Shieh |
| F Test | Hotelling-Lawley Trace |
| Alpha | 0.01 |
| Correlation Matrix | PainCorr |
| Nominal Power | 0.9 |

Output 47.3.1 continued

| Computed N Total |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Transformation | Source | Std <br> Dev | Effect | Num DF | Den DF | Actual Power | N Total |
| 1 | Time | Intercept | 0.92 | Time | 3 | 176 | 0.900 | 180 |
| 2 | Time | Intercept | 1.04 | Time | 3 | 226 | 0.903 | 230 |
| 3 | Time | Treatment | 0.92 | Time*Treatment | 3 | 346 | 0.901 | 350 |
| 4 | Time | Treatment | 1.04 | Time*Treatment | 3 | 442 | 0.901 | 446 |
| 5 | Mean(Dep) | Intercept | 0.92 | Intercept | 1 | 4 | 0.960 | 6 |
| 6 | Mean(Dep) | Intercept | 1.04 | Intercept | 1 | 4 | 0.907 | 6 |
| 7 | Mean(Dep) | Treatment | 0.92 | Treatment | 1 | 950 | 0.900 | 952 |
| 8 | Mean(Dep) | Treatment | 1.04 | Treatment | 1 | 1214 | 0.900 | 1216 |

Output 47.3.1 reveals that the required sample size to achieve a power of 0.9 for the test of the Time*Treatment interaction is 350 for the error standard deviation of 0.92 and 446 for the error standard deviation of 1.04.

Output 47.3.2 Plot of Power versus Sample Size for Repeated Measures Analysis


## References

Castelloe, J. M. (2000), "Sample Size Computations and Power Analysis with the SAS System," in Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc.

Castelloe, J. M. and O'Brien, R. G. (2001), "Power and Sample Size Determination for Linear Models," in Proceedings of the Twenty-Sixth Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc.

Chi, Y.-Y., Gribbin, M. J., Lamers, Y., Gregory, J. F., III, and Muller, K. E. (2012), "Global Hypothesis Testing for High-Dimensional Repeated Measures Outcomes," Statistics in Medicine, 31, 724-742.

Guo, Y., Logan, H. L., Glueck, D. H., and Muller, K. E. (2013), "Selecting a Sample Size for Studies with Repeated Measures," BMC Medical Research Methodology, 13.

Hocking, R. R. (1985), The Analysis of Linear Models, Monterey, CA: Brooks/Cole.
Huynh, H. and Feldt, L. S. (1976), "Estimation of the Box Correction for Degrees of Freedom from Sample Data in the Randomized Block and Split Plot Designs," Journal of Educational Statistics, 1, 69-82.

Lecoutre, B. (1991), "A Correction for the Epsilon Approximate Test with Repeated Measures Design with Two or More Independent Groups," Journal of Educational Statistics, 16, 371-372.

Lenth, R. V. (2001), "Some Practical Guidelines for Effective Sample Size Determination," American Statistician, 55, 187-193.

Logan, H. L., Baron, R. S., and Kohout, F. (1995), "Sensory Focus as Therapeutic Treatments for Acute Pain," Psychosomatic Medicine, 57, 475-484.

McKeon, J. J. (1974), " $F$ Approximations to the Distribution of Hotelling's $T_{0}^{2}$," Biometrika, 61, 381-383.
Muller, K. E. and Barton, C. N. (1989), "Approximate Power for Repeated-Measures ANOVA Lacking Sphericity," Journal of the American Statistical Association, 84, 549-555, also see "Correction to Approximate Power for Repeated-Measures ANOVA Lacking Sphericity," Journal of the American Statistical Association (1991), 86:255-256.

Muller, K. E. and Benignus, V. A. (1992), "Increasing Scientific Power with Statistical Power," Neurotoxicology and Teratology, 14, 211-219.

Muller, K. E., Edwards, L. J., Simpson, S. L., and Taylor, D. J. (2007), "Statistical Tests with Accurate Size and Power for Balanced Linear Mixed Models," Statistics in Medicine, 26, 3639-3660.

Muller, K. E. and Peterson, B. L. (1984), "Practical Methods for Computing Power in Testing the Multivariate General Linear Hypothesis," Computational Statistics and Data Analysis, 2, 143-158.

O’Brien, R. G. and Castelloe, J. M. (2007), "Sample-Size Analysis for Traditional Hypothesis Testing: Concepts and Issues," in A. Dmitrienko, C. Chuang-Stein, and R. D'Agostino, eds., Pharmaceutical Statistics Using SAS: A Practical Guide, 237-271, Cary, NC: SAS Institute Inc.

O'Brien, R. G. and Muller, K. E. (1993), "Unified Power Analysis for $t$-Tests through Multivariate Hypotheses," in L. K. Edwards, ed., Applied Analysis of Variance in Behavioral Science, 297-344, New York: Marcel Dekker.

O'Brien, R. G. and Shieh, G. (1992), "Pragmatic, Unifying Algorithm Gives Power Probabilities for Common $F$ Tests of the Multivariate General Linear Hypothesis," Poster presented at the American Statistical Association Meetings, Boston, Statistical Computing Section.

Pillai, K. C. S. and Samson, P., Jr. (1959), "On Hotelling's Generalization of $T^{2}$," Biometrika, 46, 160-168.
Simpson, S. L., Edwards, L. J., Muller, K. E., Sen, P. K., and Styner, M. A. (2010), "A Linear Exponent AR(1) Family of Correlation Structures," Statistics in Medicine, 29, 1825-1838.

## Chapter 77 The POWER Procedure

## Contents

Overview: POWER Procedure ..... 6266
Getting Started: POWER Procedure ..... 6268
Computing Power for a One-Sample $t$ Test ..... 6268
Determining Required Sample Size for a Two-Sample $t$ Test ..... 6271
Syntax: POWER Procedure ..... 6274
PROC POWER Statement ..... 6276
LOGISTIC Statement ..... 6276
MULTREG Statement ..... 6283
ONECORR Statement ..... 6288
ONESAMPLEFREQ Statement ..... 6292
ONESAMPLEMEANS Statement ..... 6300
ONEWAYANOVA Statement ..... 6306
PAIREDFREQ Statement ..... 6310
PAIREDMEANS Statement ..... 6317
PLOT Statement ..... 6326
TWOSAMPLEFREQ Statement ..... 6330
TWOSAMPLEMEANS Statement ..... 6337
TWOSAMPLESURVIVAL Statement ..... 6347
TWOSAMPLEWILCOXON Statement ..... 6358
Details: POWER Procedure ..... 6363
Overview of Power Concepts ..... 6363
Summary of Analyses ..... 6364
Specifying Value Lists in Analysis Statements ..... 6366
Keyword-Lists ..... 6366
Number-Lists ..... 6366
Grouped-Number-Lists ..... 6367
Name-Lists ..... 6368
Grouped-Name-Lists ..... 6368
Sample Size Adjustment Options ..... 6369
Error and Information Output ..... 6369
Displayed Output ..... 6371
ODS Table Names ..... 6372
Computational Resources ..... 6372
Memory ..... 6372
CPU Time ..... 6372
Computational Methods and Formulas ..... 6373
Common Notation ..... 6373
Analyses in the LOGISTIC Statement ..... 6374
Analyses in the MULTREG Statement ..... 6377
Analyses in the ONECORR Statement ..... 6379
Analyses in the ONESAMPLEFREQ Statement ..... 6381
Analyses in the ONESAMPLEMEANS Statement ..... 6399
Analyses in the ONEWAYANOVA Statement ..... 6403
Analyses in the PAIREDFREQ Statement ..... 6404
Analyses in the PAIREDMEANS Statement ..... 6408
Analyses in the TWOSAMPLEFREQ Statement ..... 6413
Analyses in the TWOSAMPLEMEANS Statement ..... 6416
Analyses in the TWOSAMPLESURVIVAL Statement ..... 6422
Analyses in the TWOSAMPLEWILCOXON Statement ..... 6426
ODS Graphics ..... 6428
Examples: POWER Procedure ..... 6429
Example 77.1: One-Way ANOVA ..... 6429
Example 77.2: The Sawtooth Power Function in Proportion Analyses ..... 6434
Example 77.3: Simple AB/BA Crossover Designs ..... 6443
Example 77.4: Noninferiority Test with Lognormal Data ..... 6446
Example 77.5: Multiple Regression and Correlation ..... 6450
Example 77.6: Comparing Two Survival Curves ..... 6454
Example 77.7: Confidence Interval Precision ..... 6457
Example 77.8: Customizing Plots ..... 6459
Assigning Analysis Parameters to Axes ..... 6461
Fine-Tuning a Sample Size Axis ..... 6466
Adding Reference Lines ..... 6470
Linking Plot Features to Analysis Parameters ..... 6473
Choosing Key (Legend) Styles ..... 6478
Modifying Symbol Locations ..... 6482
Example 77.9: Binary Logistic Regression with Independent Predictors ..... 6484
Example 77.10: Wilcoxon-Mann-Whitney Test ..... 6486
References ..... 6489

## Overview: POWER Procedure

Power and sample size analysis optimizes the resource usage and design of a study, improving chances of conclusive results with maximum efficiency. The POWER procedure performs prospective power and sample size analyses for a variety of goals, such as the following:

- determining the sample size required to get a significant result with adequate probability (power)
- characterizing the power of a study to detect a meaningful effect
- conducting what-if analyses to assess sensitivity of the power or required sample size to other factors

Here prospective indicates that the analysis pertains to planning for a future study. This is in contrast to retrospective power analysis for a past study, which is not supported by the procedure.

A variety of statistical analyses are covered:

- $t$ tests, equivalence tests, and confidence intervals for means
- tests, equivalence tests, and confidence intervals for binomial proportions
- multiple regression
- tests of correlation and partial correlation
- one-way analysis of variance
- rank tests for comparing two survival curves
- logistic regression with binary response
- Wilcoxon-Mann-Whitney (rank-sum) test

For more complex linear models, see Chapter 47, "The GLMPOWER Procedure."
Input for PROC POWER includes the components considered in study planning:

- design
- statistical model and test
- significance level (alpha)
- surmised effects and variability
- power
- sample size

You designate one of these components by a missing value in the input, in order to identify it as the result parameter. The procedure calculates this result value over one or more scenarios of input values for all other components. Power and sample size are the most common result values, but for some analyses the result can be something else. For example, you can solve for the sample size of a single group for a two-sample $t$ test.

In addition to tabular results, PROC POWER produces graphs. You can produce the most common types of plots easily with default settings and use a variety of options for more customized graphics. For example, you can control the choice of axis variables, axis ranges, number of plotted points, mapping of graphical features (such as color, line style, symbol and panel) to analysis parameters, and legend appearance.

If ODS Graphics is enabled, then PROC POWER uses ODS Graphics to create graphs; otherwise, traditional graphs are produced.

For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 606 in Chapter 21, "Statistical Graphics Using ODS."

For specific information about the statistical graphics and options available with the POWER procedure, see the PLOT statement and the section "ODS Graphics" on page 6428.

The POWER procedure is one of several tools available in SAS/STAT software for power and sample size analysis. PROC GLMPOWER supports more complex linear models. The Power and Sample Size application provides a user interface and implements many of the analyses supported in the procedures. See Chapter 47, "The GLMPOWER Procedure," and Chapter 78, "The Power and Sample Size Application," for details.

The following sections of this chapter describe how to use PROC POWER and discuss the underlying statistical methodology. The section "Getting Started: POWER Procedure" on page 6268 introduces PROC POWER with simple examples of power computation for a one-sample $t$ test and sample size determination for a two-sample $t$ test. The section "Syntax: POWER Procedure" on page 6274 describes the syntax of the procedure. The section "Details: POWER Procedure" on page 6363 summarizes the methods employed by PROC POWER and provides details on several special topics. The section "Examples: POWER Procedure" on page 6429 illustrates the use of the POWER procedure with several applications.

For an overview of methodology and SAS tools for power and sample size analysis, see Chapter 18, "Introduction to Power and Sample Size Analysis." For more discussion and examples, see O'Brien and Castelloe (2007); Castelloe (2000); Castelloe and O'Brien (2001); Muller and Benignus (1992); O'Brien and Muller (1993); Lenth (2001).

## Getting Started: POWER Procedure

## Computing Power for a One-Sample $t$ Test

Suppose you want to improve the accuracy of a machine used to print logos on sports jerseys. The logo placement has an inherently high variability, but the horizontal alignment of the machine can be adjusted. The operator agrees to pay for a costly adjustment if you can establish a nonzero mean horizontal displacement in either direction with high confidence. You have 150 jerseys at your disposal to measure, and you want to determine your chances of a significant result (power) by using a one-sample $t$ test with a two-sided $\alpha=0.05$.

You decide that 8 mm is the smallest displacement worth addressing. Hence, you will assume a true mean of 8 in the power computation. Experience indicates that the standard deviation is about 40.

Use the ONESAMPLEMEANS statement in the POWER procedure to compute the power. Indicate power as the result parameter by specifying the POWER= option with a missing value (.). Specify your conjectures for the mean and standard deviation by using the MEAN= and STDDEV= options and for the sample size by using the NTOTAL= option. The statements required to perform this analysis are as follows:

```
proc power;
    onesamplemeans
        mean = 8
        ntotal = 150
        stddev = 40
        power = .;
run;
```

Default values for the TEST=, DIST=, ALPHA=, NULLMEAN=, and SIDES= options specify a two-sided $t$ test for a mean of 0 , assuming a normal distribution with a significance level of $\alpha=0.05$.

Figure 77.1 shows the output.
Figure 77.1 Sample Size Analysis for One-Sample $t$ Test
The POWER Procedure One-Sample t Test for Mean

| Fixed Scenario Elements |  |
| :---: | :---: |
| Distribution | Normal |
| Method | Exact |
| Mean | 8 |
| Standard Deviation | 40 |
| Total Sample Size | 150 |
| Number of Sides | 2 |
| Null Mean | 0 |
| Alpha | 0.05 |
| Computed Power |  |
| Power |  |
| 0.682 |  |

The power is about 0.68 . In other words, there is about a $2 / 3$ chance that the $t$ test will produce a significant result demonstrating the machine's average off-center displacement. This probability depends on the assumptions for the mean and standard deviation.

Now, suppose you want to account for some of your uncertainty in conjecturing the true mean and standard deviation by evaluating the power for four scenarios, using reasonable low and high values, 5 and 10 for the mean, and 30 and 50 for the standard deviation. Also, you might be able to measure more than 150 jerseys, and you would like to know under what circumstances you could get by with fewer. You want to plot power for sample sizes between 100 and 200 to visualize how sensitive the power is to changes in sample size for these four scenarios of means and standard deviations. The following statements perform this analysis:

```
ods graphics on;
proc power;
    onesamplemeans
            mean = 5 10
            ntotal = 150
            stddev = 30 50
            power = .;
    plot x=n min=100 max=200;
run;
ods graphics off;
```

The new mean and standard deviation values are specified by using the MEAN= and STDDEV= options in the ONESAMPLEMEANS statement. The PLOT statement with $\mathrm{X}=\mathrm{N}$ produces a plot with sample size on the X axis. (The result parameter, in this case the power, is always plotted on the other axis.) The $\mathrm{MIN}=$ and MAX = options in the PLOT statement determine the sample size range. The ODS GRAPHICS ON statement enables ODS Graphics.

Figure 77.2 shows the output, and Figure 77.3 shows the plot.

Figure 77.2 Sample Size Analysis for One-Sample $t$ Test with Input Ranges

> The POWER Procedure One-Sample $t$ Test for Mean

| Fixed Scenario Elements |  |  |  |
| :---: | :---: | :---: | :---: |
| Distribution <br> Method <br> Total Sample Size <br> Number of Sides <br> Null Mean <br> Alpha |  |  | Normal |
|  |  |  | Exact |
|  |  |  | 150 |
|  |  |  | 2 |
|  |  |  | 0 |
|  |  |  | 0.05 |
|  | pu | d Pow |  |
| Index |  | Std Dev | Power |
| 1 | 5 | 30 | 0.527 |
| 2 | 5 | 50 | 0.229 |
|  | 10 | 30 | 0.982 |
|  | 10 | 50 | 0.682 |

Figure 77.3 Plot of Power versus Sample Size for One-Sample t Test with Input Ranges


The power ranges from about 0.23 to 0.98 for a sample size of 150 depending on the mean and standard deviation. In Figure 77.3, the line style identifies the mean, and the plotting symbol identifies the standard deviation. The locations of plotting symbols indicate computed powers; the curves are linear interpolations of these points. The plot suggests sufficient power for a mean of 10 and standard deviation of 30 (for any of the sample sizes) but insufficient power for the other three scenarios.

## Determining Required Sample Size for a Two-Sample $t$ Test

In this example you want to compare two physical therapy treatments designed to increase muscle flexibility. You need to determine the number of patients required to achieve a power of at least 0.9 to detect a group mean difference in a two-sample $t$ test. You will use $\alpha=0.05$ (two-tailed).

The mean flexibility with the standard treatment (as measured on a scale of 1 to 20) is well known to be about 13 and is thought to be between 14 and 15 with the new treatment. You conjecture three alternative scenarios for the means:

1. $\mu_{1}=13, \mu_{2}=14$
2. $\mu_{1}=13, \mu_{2}=14.5$
3. $\mu_{1}=13, \mu_{2}=15$

You conjecture two scenarios for the common group standard deviation:

1. $\sigma=1.2$
2. $\sigma=1.7$

You also want to try three weighting schemes:

1. equal group sizes (balanced, or $1: 1$ )
2. twice as many patients with the new treatment (1:2)
3. three times as many patients with the new treatment (1:3)

This makes $3 \times 2 \times 3=18$ scenarios in all.
Use the TWOSAMPLEMEANS statement in the POWER procedure to determine the sample sizes required to give $90 \%$ power for each of these 18 scenarios. Indicate total sample size as the result parameter by specifying the NTOTAL= option with a missing value (.). Specify your conjectures for the means by using the GROUPMEANS= option. Using the "matched" notation (discussed in the section "Specifying Value Lists in Analysis Statements" on page 6366), enclose the two group means for each scenario in parentheses. Use the STDDEV= option to specify scenarios for the common standard deviation. Specify the weighting schemes by using the GROUPWEIGHTS= option. You could again use the matched notation. But for illustrative purposes, specify the scenarios for each group weight separately by using the "crossed" notation, with scenarios for each group weight separated by a vertical bar (I). The statements that perform the analysis are as follows:

```
proc power;
    twosamplemeans
    groupmeans = (13 14) (13 14.5) (13 15)
    stddev = 1.2 1.7
    groupweights = 1 | 1 2 3
    power = 0.9
    ntotal = .;
run;
```

Default values for the TEST=, DIST=, NULLDIFF $=$, ALPHA $=$, and SIDES $=$ options specify a two-sided $t$ test of group mean difference equal to 0 , assuming a normal distribution with a significance level of $\alpha=0.05$. The results are shown in Figure 77.4.

Figure 77.4 Sample Size Analysis for Two-Sample $t$ Test Using Group Means
The POWER Procedure Two-Sample t Test for Mean Difference

| Fixed Scenario <br> Elements |  |
| :--- | ---: |
| Distribution | Normal |
| Method | Exact |
| Group 1 Weight | 1 |
| Nominal Power | 0.9 |
| Number of Sides | 2 |
| Null Difference | 0 |
| Alpha | 0.05 |


| Computed N Total |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Mean1 | Mean2 | Std Dev | Weight2 | Actual Power | N Total |
| 1 | 13 | 14.0 | 1.2 | 1 | 0.907 | 64 |
| 2 | 13 | 14.0 | 1.2 | 2 | 0.908 | 72 |
| 3 | 13 | 14.0 | 1.2 | 3 | 0.905 | 84 |
| 4 | 13 | 14.0 | 1.7 | 1 | 0.901 | 124 |
| 5 | 13 | 14.0 | 1.7 | 2 | 0.905 | 141 |
| 6 | 13 | 14.0 | 1.7 | 3 | 0.900 | 164 |
| 7 | 13 | 14.5 | 1.2 | 1 | 0.910 | 30 |
| 8 | 13 | 14.5 | 1.2 | 2 | 0.906 | 33 |
| 9 | 13 | 14.5 | 1.2 | 3 | 0.916 | 40 |
| 10 | 13 | 14.5 | 1.7 | 1 | 0.900 | 56 |
| 11 | 13 | 14.5 | 1.7 | 2 | 0.901 | 63 |
| 12 | 13 | 14.5 | 1.7 | 3 | 0.908 | 76 |
| 13 | 13 | 15.0 | 1.2 | 1 | 0.913 | 18 |
| 14 | 13 | 15.0 | 1.2 | 2 | 0.927 | 21 |
| 15 | 13 | 15.0 | 1.2 | 3 | 0.922 | 24 |
| 16 | 13 | 15.0 | 1.7 | 1 | 0.914 | 34 |
| 17 | 13 | 15.0 | 1.7 | 2 | 0.921 | 39 |
| 18 | 13 | 15.0 | 1.7 | 3 | 0.910 | 44 |

The interpretation is that in the best-case scenario (large mean difference of 2, small standard deviation of 1.2 , and balanced design), a sample size of $N=18\left(n_{1}=n_{2}=9\right)$ patients is sufficient to achieve a power of at least 0.9 . In the worst-case scenario (small mean difference of 1 , large standard deviation of 1.7, and a 1:3 unbalanced design), a sample size of $N=164\left(n_{1}=41, n_{2}=123\right)$ patients is necessary. The Nominal Power of 0.9 in the "Fixed Scenario Elements" table represents the input target power, and the Actual Power column in the "Computed N Total" table is the power at the sample size ( N Total) adjusted to achieve the specified sample weighting exactly.

Note the following characteristics of the analysis, and ways you can modify them if you want:

- The total sample sizes are rounded up to multiples of the weight sums ( 2 for the $1: 1$ design, 3 for the $1: 2$ design, and 4 for the $1: 3$ design) to ensure that each group size is an integer. To request raw fractional sample size solutions, use the NFRACTIONAL option.
- Only the group weight that varies (the one for group 2) is displayed as an output column, while the weight for group 1 appears in the "Fixed Scenario Elements" table. To display the group weights together in output columns, use the matched version of the value list rather than the crossed version.
- If you can specify only differences between group means (instead of their individual values), or if you want to display the mean differences instead of the individual means, use the MEANDIFF= option instead of the GROUPMEANS $=$ option.

The following statements implement all of these modifications:

```
proc power;
    twosamplemeans
            nfractional
            meandiff = 1 to 2 by 0.5
            stddev = 1.2 1.7
            groupweights = (1 1) (1 2) (1 3)
            power = 0.9
            ntotal = .;
run;
```

Figure 77.5 shows the new results.
Figure 77.5 Sample Size Analysis for Two-Sample $t$ Test Using Mean Differences
The POWER Procedure Two-Sample t Test for Mean Difference

| Fixed Scenario <br> Elements |  |
| :--- | ---: |
| Distribution | Normal |
| Method | Exact |
| Nominal Power | 0.9 |
| Number of Sides | 2 |
| Null Difference | 0 |
| Alpha | 0.05 |

Figure 77.5 continued

| Computed Ceiling N Total |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Mean Diff | Std <br> Dev | Weight1 | Weight2 | Fractional N Total | Actual Power | Ceiling Total |
| 1 | 1.0 | 1.2 | 1 | 1 | 62.507429 | 0.902 | 63 |
| 2 | 1.0 | 1.2 | 1 | 2 | 70.065711 | 0.904 | 71 |
| 3 | 1.0 | 1.2 | 1 | 3 | 82.665772 | 0.901 | 83 |
| 4 | 1.0 | 1.7 | 1 | 1 | 123.418482 | 0.901 | 124 |
| 5 | 1.0 | 1.7 | 1 | 2 | 138.598159 | 0.901 | 139 |
| 6 | 1.0 | 1.7 | 1 | 3 | 163.899094 | 0.900 | 164 |
| 7 | 1.5 | 1.2 | 1 | 1 | 28.961958 | 0.900 | 29 |
| 8 | 1.5 | 1.2 | 1 | 2 | 32.308867 | 0.906 | 33 |
| 9 | 1.5 | 1.2 | 1 | 3 | 37.893351 | 0.901 | 38 |
| 10 | 1.5 | 1.7 | 1 | 1 | 55.977156 | 0.900 | 56 |
| 11 | 1.5 | 1.7 | 1 | 2 | 62.717357 | 0.901 | 63 |
| 12 | 1.5 | 1.7 | 1 | 3 | 73.954291 | 0.900 | 74 |
| 13 | 2.0 | 1.2 | 1 | 1 | 17.298518 | 0.913 | 18 |
| 14 | 2.0 | 1.2 | 1 | 2 | 19.163836 | 0.913 | 20 |
| 15 | 2.0 | 1.2 | 1 | 3 | 22.282926 | 0.910 | 23 |
| 16 | 2.0 | 1.7 | 1 | 1 | 32.413512 | 0.905 | 33 |
| 17 | 2.0 | 1.7 | 1 | 2 | 36.195531 | 0.907 | 37 |
| 18 | 2.0 | 1.7 | 1 | 3 | 42.504535 | 0.903 | 43 |

Note that the Nominal Power of 0.9 applies to the raw computed sample size (Fractional N Total), and the Actual Power column applies to the rounded sample size (Ceiling N Total). Some of the adjusted sample sizes in Figure 77.5 are lower than those in Figure 77.4 because underlying group sample sizes are allowed to be fractional (for example, the first Ceiling N Total of 63 corresponding to equal group sizes of 31.5).

## Syntax: POWER Procedure

The following statements are available in the POWER procedure:

```
PROC POWER < options> ;
    LOGISTIC <options> ;
    MULTREG <options> ;
    ONECORR < options > ;
    ONESAMPLEFREQ < options> ;
    ONESAMPLEMEANS < options> ;
    ONEWAYANOVA < options> ;
    PAIREDFREQ < options > ;
    PAIREDMEANS < options> ;
    PLOT < plot-options> </ graph-options> ;
    TWOSAMPLEFREQ < options>;
    TWOSAMPLEMEANS < options> ;
    TWOSAMPLESURVIVAL < options> ;
    TWOSAMPLEWILCOXON < options> ;
```

The statements in the POWER procedure consist of the PROC POWER statement, a set of analysis statements (for requesting specific power and sample size analyses), and the PLOT statement (for producing graphs). The PROC POWER statement and at least one of the analysis statements are required. The analysis statements are LOGISTIC, MULTREG, ONECORR, ONESAMPLEFREQ, ONESAMPLEMEANS, ONEWAYANOVA, PAIREDFREQ, PAIREDMEANS, TWOSAMPLEFREQ, TWOSAMPLEMEANS, TWOSAMPLESURVIVAL, and TWOSAMPLEWILCOXON.

You can use multiple analysis statements and multiple PLOT statements. Each analysis statement produces a separate sample size analysis. Each PLOT statement refers to the previous analysis statement and generates a separate graph (or set of graphs).

The name of an analysis statement describes the framework of the statistical analysis for which sample size calculations are desired. You use options in the analysis statements to identify the result parameter to compute, to specify the statistical test and computational options, and to provide one or more scenarios for the values of relevant analysis parameters.

Table 77.1 summarizes the basic functions of each statement in PROC POWER. The syntax of each statement in Table 77.1 is described in the following pages.

Table 77.1 Statements in the POWER Procedure

| Statement | Description |
| :--- | :--- |
| PROC POWER | Invokes the procedure |
| LOGISTIC | Likelihood ratio chi-square test of a single predictor in logistic <br> regression with binary response |
| MULTREG | Tests of one or more coefficients in multiple linear regression <br> ONECORR |
| Fisher's $z$ test and $t$ test of (partial) correlation |  |
| Tests, confidence interval precision, and equivalence tests of a |  |
| single binomial proportion |  |

See the section "Summary of Analyses" on page 6364 for a summary of the analyses available and the syntax required for them.

## PROC POWER Statement

PROC POWER < options > ;
The PROC POWER statement invokes the POWER procedure. You can specify the following option.

## PLOTONLY

specifies that only graphical results from the PLOT statement should be produced.

## LOGISTIC Statement

LOGISTIC < options > ;
The LOGISTIC statement performs power and sample size analyses for the likelihood ratio chi-square test of a single predictor in binary logistic regression, possibly in the presence of one or more covariates that might be correlated with the tested predictor.

## Summary of Options

Table 77.2 summarizes the options available in the LOGISTIC statement.

## Table 77.2 LOGISTIC Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| COVARIATES $=$ | Specifies the distributions of predictor variables |
| TESTPREDICTOR= | Specifies the distribution of the predictor variable being tested |
| VARDIST= | Defines a distribution for a predictor variable |
| Specify effects |  |
| CORR= | Specifies the multiple correlation between the predictor and the covariates |
| COVODDSRATIOS= | Specifies the odds ratios for the covariates |
| COVREGCOEFFS= | Specifies the regression coefficients for the covariates |
| DEFAULTUNIT= | Specifies the default change in the predictor variables |
| INTERCEPT= | Specifies the intercept |
| RESPONSEPROB= | Specifies the response probability |
| TESTODDSRATIO= | Specifies the odds ratio being tested |
| TESTREGCOEFF= | Specifies the regression coefficient for the predictor variable |
| UNITS= | Specifies the changes in the predictor variables |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NTOTAL= | Specifies the sample size |
| Specify power |  |
| POWER= | Specifies the desired power of the test |

Table 77.2 continued

| Option | Description |
| :--- | :--- |
| Specify computational | method |
| DEFAULTNBINS $=$ | Specifies the default number of categories for each predictor variable |
| NBINS $=$ | Specifies the number of categories for predictor variables |
| Control ordering in output |  |
| OUTPUTORDER $=$ | Controls the output order of parameters |

Table 77.3 summarizes the valid result parameters in the LOGISTIC statement.

Table 77.3 Summary of Result Parameters in the LOGISTIC Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=LRCHI | Power | POWER=. |
|  | Sample size | NTOTAL=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## CORR=number-list

specifies the multiple correlation $(\rho)$ between the tested predictor and the covariates. If you also specify the COVARIATES = option, then the sample size is either multiplied (if you are computing power) or divided (if you are computing sample size) by a factor of ( $1-\rho^{2}$ ). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## COVARIATES=grouped-name-list

specifies the distributions of any predictor variables in the model but not being tested, using labels specified with the VARDIST= option. The distributions are assumed to be independent of each other and of the tested predictor. If this option is omitted, then the tested predictor specified by the TESTEDPREDICTOR= option is assumed to be the only predictor in the model. For information about specifying the grouped-name-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## COVODDSRATIOS=grouped-number-list

specifies the odds ratios for the covariates in the full model (including variables in the TESTPREDICTOR= and COVARIATES $=$ options). The ordering of the values corresponds to the ordering in the COVARIATES $=$ option. If the response variable is coded as $Y=1$ for success and $Y=0$ for failure, then the odds ratio for each covariate $X$ is the odds of $Y=1$ when $X=a$ divided by the odds of $Y=1$ when $X=b$, where $a$ and $b$ are determined from the DEFAULTUNIT $=$ and UNITS $=$ options. Values must be greater than zero. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## COVREGCOEFFS=grouped-number-list

specifies the regression coefficients for the covariates in the full model including the test predictor (as specified by the TESTPREDICTOR= option). The ordering of the values corresponds to the ordering in the COVARIATES= option. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DEFAULTNBINS=number

specifies the default number of categories (or "bins") into which the distribution for each predictor variable is divided in internal calculations. Higher values increase computational time and memory requirements but generally lead to more accurate results. However, if the value is too high, then numerical instability can occur. Lower values are less likely to produce "No solution computed" errors. Each test predictor or covariate that is absent from the NBINS= option derives its bin number from the DEFAULTNBINS $=$ option. The default value is DEFAULTNBINS $=10$.

There are two variable distributions for which the number of bins can be overridden internally:

- For an ordinal distribution, the number of ordinal values is always used as the number of bins.
- For a binomial distribution, if the requested number of bins is larger than $n+1$, where $n$ is the sample size parameter of the binomial distribution, then exactly $n+1$ bins are used.


## DEFAULTUNIT=change-spec

specifies the default change in the predictor variables assumed for odds ratios specified with the COVODDSRATIOS $=$ and TESTODDSRATIO $=$ options. Each test predictor or covariate that is absent from the UNITS = option derives its change value from the DEFAULTUNIT $=$ option. The value must be nonzero. The default value is DEFAULTUNIT $=1$. This option can be used only if at least one of the COVODDSRATIOS $=$ and TESTODDSRATIO $=$ options is used.

Valid specifications for change-spec are as follows:
number defines the odds ratio as the ratio of the response variable odds when $X=a$ to the odds when $X=a-$ number for any constant $a$.
$<+\mid->$ SD defines the odds ratio as the ratio of the odds when $X=a$ to the odds when $X=a-\sigma$ (or $X=a+\sigma$, if SD is preceded by a minus sign $(-)$ ) for any constant $a$, where $\sigma$ is the standard deviation of $X$ (as determined from the VARDIST= option).
multiple*SD defines the odds ratio as the ratio of the odds when $X=a$ to the odds when $X=a-$ multiple $* \sigma$ for any constant $a$, where $\sigma$ is the standard deviation of $X$ (as determined from the VARDIST= option).

PERCENTILES $(p 1, p 2)$ defines the odds ratio as the ratio of the odds when $X$ is equal to its $p 2 \times$ 100th percentile to the odds when $X$ is equal to its $p 1 \times 100$ th percentile (where the percentiles are determined from the distribution specified in the VARDIST= option). Values for $p 1$ and $p 2$ must be strictly between 0 and 1 .

## INTERCEPT=number-list

specifies the intercept in the full model (including variables in the TESTPREDICTOR= and COVARIATES $=$ options). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

```
NBINS=("name" = number < . . . "name" = number >)
```

specifies the number of categories (or "bins") into which the distribution for each predictor variable (identified by its name from the VARDIST= option) is divided in internal calculations. Higher values increase computational time and memory requirements but generally lead to more accurate results. However, if the value is too high, then numerical instability can occur. Lower values are less likely to produce "No solution computed" errors. Each predictor variable that is absent from the NBINS= option derives its bin number from the DEFAULTNBINS= option.

There are two variable distributions for which the NBINS = value can be overridden internally:

- For an ordinal distribution, the number of ordinal values is always used as the number of bins.
- For a binomial distribution, if the requested number of bins is larger than $n+1$, where $n$ is the sample size parameter of the binomial distribution, then exactly $n+1$ bins are used.


## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). Values must be at least one. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- DEFAULTNBINS=
- NBINS=
- $\mathrm{ALPHA}=$
- RESPONSEPROB=
- INTERCEPT=
- TESTPREDICTOR=
- TESTODDSRATIO=
- TESTREGCOEFF=
- COVARIATES=
- COVODDSRATIOS=
- COVREGCOEFFS=
- CORR=
- NTOTAL=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the LOGISTIC statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the LOGISTIC statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## RESPONSEPROB=number-list

specifies the response probability in the full model when all predictor variables (including variables in the TESTPREDICTOR= and COVARIATES $=$ options) are equal to their means. The $\log$ odds of this probability are equal to the intercept in the full model where all predictor are centered at their means. If the response variable is coded as $Y=1$ for success and $Y=0$ for failure, then this probability is equal to the mean of $Y$ in the full model when all Xs are equal to their means. Values must be strictly between zero and one. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TEST=LRCHI

specifies the likelihood ratio chi-square test of a single model parameter in binary logistic regression. This is the default test option.

## TESTODDSRATIO=number-list

specifies the odds ratio for the predictor variable being tested in the full model (including variables in the TESTPREDICTOR $=$ and COVARIATES $=$ options). If the response variable is coded as $Y=1$ for success and $Y=0$ for failure, then the odds ratio for the $X$ being tested is the odds of $Y=1$ when $X=a$ divided by the odds of $Y=1$ when $X=b$, where $a$ and $b$ are determined from the DEFAULTUNIT= and UNITS $=$ options. Values must be greater than zero. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TESTPREDICTOR=name-list

specifies the distribution of the predictor variable being tested, using labels specified with the VARDIST= option. This distribution is assumed to be independent of the distributions of the covariates as defined in the COVARIATES= option. For information about specifying the name-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TESTREGCOEFF=number-list

specifies the regression coefficient for the predictor variable being tested in the full model including the covariates specified by the COVARIATES= option. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## UNITS=("name" = change-spec < . . . "name" = change-spec >)

specifies the changes in the predictor variables assumed for odds ratios specified with the COVODDSRATIOS $=$ and TESTODDSRATIO $=$ options. Each predictor variable whose name (from the VARDIST= option) is absent from the UNITS option derives its change value from the DEFAULTUNIT $=$ option. This option can be used only if at least one of the COVODDSRATIOS= and TESTODDSRATIO $=$ options is used.

Valid specifications for change-spec are as follows:
number defines the odds ratio as the ratio of the response variable odds when $X=a$ to the odds when $X=a-$ number for any constant $a$.
<+ $\mid->$ SD defines the odds ratio as the ratio of the odds when $X=a$ to the odds when $X=a-\sigma$ (or $X=a+\sigma$, if SD is preceded by a minus sign (-)) for any constant $a$, where $\sigma$ is the standard deviation of $X$ (as determined from the VARDIST= option).
multiple*SD defines the odds ratio as the ratio of the odds when $X=a$ to the odds when $X=a-$ multiple $\times \sigma$ for any constant $a$, where $\sigma$ is the standard deviation of $X$ (as determined from the VARDIST= option).
PERCENTILES $(p 1, p 2)$ defines the odds ratio as the ratio of the odds when $X$ is equal to its $p 2 \times$ 100th percentile to the odds when $X$ is equal to its $p 1 \times 100$ th percentile (where the percentiles are determined from the distribution specified in the VARDIST= option). Values for $p 1$ and $p 2$ must be strictly between 0 and 1 .

Each unit value must be nonzero.

## VARDIST("/abe/")=distribution (parameters)

defines a distribution for a predictor variable.
For the VARDIST= option,
label identifies the variable distribution in the output and with the COVARIATES $=$ and TESTPREDICTOR= options.
distribution specifies the distributional form of the variable.
parameters specifies one or more parameters associated with the distribution.
The distributions and parameters are named and defined in the same way as the distributions and arguments in the CDF SAS function; for more information, see SAS Language Reference: Dictionary. Choices for distributional forms and their parameters are as follows:

ORDINAL ((values) : (probabilities)) is an ordered categorical distribution. The values are any numbers separated by spaces. The probabilities are numbers between 0 and 1 (inclusive) separated by spaces. Their sum must be exactly 1 . The number of probabilities must match the number of values.
BETA $(a, b<, I, r>)$ is a beta distribution with shape parameters $a$ and $b$ and optional location parameters $/$ and $r$. The values of $a$ and $b$ must be greater than 0 , and $/$ must be less than $r$. The default values for $l$ and $r$ are 0 and 1 , respectively.
BINOMIAL ( $p, n$ ) is a binomial distribution with probability of success $p$ and number of independent Bernoulli trials $n$. The value of $p$ must be greater than 0 and less than 1 , and $n$ must be an integer greater than 0 . If $n=1$, then the distribution is binary.
EXPONENTIAL ( $\lambda$ ) is an exponential distribution with scale $\lambda$, which must be greater than 0 .
GAMMA $(a, \lambda)$ is a gamma distribution with shape $a$ and scale $\lambda$. The values of $a$ and $\lambda$ must be greater than 0 .
$\operatorname{LAPLACE}(\theta, \lambda) \quad$ is a Laplace distribution with location $\theta$ and scale $\lambda$. The value of $\lambda$ must be greater than 0 .
LOGISTIC $(\theta, \lambda) \quad$ is a logistic distribution with location $\theta$ and scale $\lambda$. The value of $\lambda$ must be greater than 0 .

LOGNORMAL $(\theta, \lambda)$ is a lognormal distribution with location $\theta$ and scale $\lambda$. The value of $\lambda$ must be greater than 0 .
$\operatorname{NORMAL}(\theta, \lambda) \quad$ is a normal distribution with mean $\theta$ and standard deviation $\lambda$. The value of $\lambda$ must be greater than 0 .
POISSON $(m) \quad$ is a Poisson distribution with mean $m$. The value of $m$ must be greater than 0 .
UNIFORM $(I, r)$ is a uniform distribution on the interval $[I, r]$, where $I<r$.

## Restrictions on Option Combinations

To specify the intercept in the full model, choose one of the following two parameterizations:

- intercept (using the INTERCEPT= options)
- $\operatorname{Prob}(Y=1)$ when all predictors are equal to their means (using the RESPONSEPROB=option)

To specify the effect associated with the predictor variable being tested, choose one of the following two parameterizations:

- odds ratio (using the TESTODDSRATIO= options)
- regression coefficient (using the TESTREGCOEFFS= option)

To describe the effects of the covariates in the full model, choose one of the following two parameterizations:

- odds ratios (using the COVODDSRATIOS= options)
- regression coefficients (using the COVREGCOEFFS= options)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the LOGISTIC statement.

## Likelihood Ratio Chi-Square Test for One Predictor

You can express effects in terms of response probability and odds ratios, as in the following statements:

```
proc power;
    logistic
        vardist("x1a") = normal(0, 2)
        vardist("x1b") = normal(0, 3)
        vardist("x2") = poisson(7)
        vardist("x3a") = ordinal((-5 0 5) : (.3 .4 .3))
        vardist("x3b") = ordinal((-5 0 5) : (.4 .3 .3))
        testpredictor = "x1a" "x1b"
        covariates = "x2" | "x3a" "x3b"
        responseprob = 0.15
        testoddsratio = 1.75
        covoddsratios = (2.1 1.4)
        ntotal = 100
        power = .;
run;
```

The VARDIST= options define the distributions of the predictor variables. The TESTPREDICTOR= option specifies two scenarios for the test predictor distribution, $\operatorname{Normal}(10,2)$ and $\operatorname{Normal}(10,3)$. The COVARIATES $=$ option specifies two covariates, the first with a Poisson(7) distribution. The second covariate has an ordinal distribution on the values $-5,0$, and 5 with two scenarios for the associated probabilities: (.3, $.4, .3)$ and $(.4, .3, .3)$. The response probability in the full model with all variables equal to zero is specified by the RESPONSEPROB $=$ option as 0.15 . The odds ratio for a unit decrease in the tested predictor is specified by the TESTODDSRATIO= option to be 1.75. Corresponding odds ratios for the two covariates in the full model are specified by the COVODDSRATIOS $=$ option to be 2.1 and 1.4. The POWER=. option requests a solution for the power at a sample size of 100 as specified by the NTOTAL= option.

Default values of the TEST = and ALPHA= options specify a likelihood ratio test of the first predictor with a significance level of 0.05 . The default of DEFAULTUNIT $=1$ specifies that all odds ratios are defined in terms of unit changes in predictors. The default of DEFAULTNBINS $=10$ specifies that each of the three predictor variables is discretized into a distribution with 10 categories in internal calculations.

You can also express effects in terms of regression coefficients, as in the following statements:

```
proc power;
    logistic
        vardist("x1a") = normal(0, 2)
        vardist("x1b") = normal(0, 3)
        vardist("x2") = poisson(7)
        vardist("x3a") = ordinal((-5 0 5) : (.3 .4 .3))
        vardist("x3b") = ordinal((-5 0 5) : (.4 .3 .3))
        testpredictor = "x1a" "x1b"
        covariates = "x2" | "x3a" "x3b"
        intercept = -6.928162
        testregcoeff = 0.5596158
        covregcoeffs = (0.7419373 0.3364722)
        ntotal = 100
        power = .;
run;
```

The regression coefficients for the tested predictor (TESTREGCOEFF $=0.5596158$ ) and covariates (COVREGCOEFFS $=(0.74193730 .3364722)$ ) are determined by taking the logarithm of the corresponding odds ratios. The intercept in the full model is specified as -6.928162 by the INTERCEPT= option. This number is calculated according to the formula at the end of "Analyses in the LOGISTIC Statement" on page 6374, which expresses the intercept in terms of the response probability, regression coefficients, and predictor means:

$$
\text { Intercept }=\log \left(\frac{0.15}{1-0.15}\right)-(0.5596158(0)+0.7419373(7)+0.3364722(0))
$$

## MULTREG Statement

## MULTREG <options> ;

The MULTREG statement performs power and sample size analyses for Type III $F$ tests of sets of predictors in multiple linear regression, assuming either fixed or normally distributed predictors.

## Summary of Options

Table 77.4 summarizes the options available in the MULTREG statement.

Table 77.4 MULTREG Statement Options

| Option | Description |
| :--- | :--- |
| Define analysis |  |
| TEST $=$ | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA $=$ | Specifies the significance level |
| MODEL $=$ | Specifies the assumed distribution of the predictors |
| NFULLPREDICTORS= | Specifies the number of predictors in the full model |
| NOINT | Specifies a no-intercept model |
| NREDUCEDPREDICTORS= $=$ | Specifies the number of predictors in the reduced model |
| NTESTPREDICTORS= | Specifies the number of predictors being tested |
| Specify effects |  |
| PARTIALCORR= | Specifies the partial correlation |
| RSQUAREDIFF= | Specifies the difference in $R^{2}$ |
| RSQUAREFULL= | Specifies the $R^{2}$ of the full model |
| RSQUAREREDUCED= | Specifies the $R^{2}$ of the reduced model |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NTOTAL= | Specifies the sample size |
| Specify power |  |
| POWER= | Specifies the desired power |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the order of parameters |

Table 77.5 summarizes the valid result parameters in the MULTREG statement.

Table 77.5 Summary of Result Parameters in the MULTREG Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=TYPE3 | Power | POWER=. |
|  | Sample size | NTOTAL=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## MODEL=keyword-list

specifies the assumed distribution of the tested predictors. MODEL=FIXED indicates a fixed predictor distribution. MODEL=RANDOM (the default) indicates a joint multivariate normal distribution for the response and tested predictors. You can use the aliases CONDITIONAL for FIXED and UNCONDITIONAL for RANDOM. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

FIXED fixed predictors
RANDOM random (multivariate normal) predictors

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NFULLPREDICTORS=number-list

## NFULLPRED=number-list

specifies the number of predictors in the full model, not counting the intercept. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366 .

## NOINT

specifies a no-intercept model (for both full and reduced models). By default, the intercept is included in the model. If you want to test the intercept, you can specify the NOINT option and simply consider the intercept to be one of the predictors being tested. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NREDUCEDPREDICTORS=number-list

## NREDUCEDPRED=number-list

## NREDPRED=number-list

specifies the number of predictors in the reduced model, not counting the intercept. This is the same as the difference between values of the NFULLPREDICTORS $=$ and NTESTPREDICTORS $=$ options. Note that supplying a value of 0 is the same as specifying an $F$ test of a Pearson correlation. This option cannot be used at the same time as the NTESTPREDICTORS= option. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTESTPREDICTORS=number-list

## NTESTPRED=number-list

specifies the number of predictors being tested. This is the same as the difference between values of the NFULLPREDICTORS= and NREDUCEDPREDICTORS= options. Note that supplying identical values for the NTESTPREDICTORS $=$ and NFULLPREDICTORS $=$ options is the same as specifying an $F$ test of a Pearson correlation. This option cannot be used at the same time as the NREDUCEDPREDICTORS = option. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). The minimum acceptable value for the sample size depends on the MODEL=, NOINT, NFULLPREDICTORS $=$, NTESTPREDICTORS $=$, and NREDUCEDPREDICTORS $=$ options. It ranges from $p+$ 1 to $p+3$, where $p$ is the value of the NFULLPREDICTORS option. For further information about minimum NTOTAL values, see Table 77.30. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- MODEL=
- NFULLPREDICTORS=
- NTESTPREDICTORS=
- NREDUCEDPREDICTORS=
- ALPHA=
- PARTIALCORR=
- RSQUAREFULL=
- RSQUAREREDUCED=
- RSQUAREDIFF=
- NTOTAL=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the MULTREG statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the MULTREG statement.

## PARTIALCORR=number-list

PCORR=number-list
specifies the partial correlation between the tested predictors and the response, adjusting for any other predictors in the model. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## RSQUAREDIFF=number-list

RSQDIFF=number-list
specifies the difference in $R^{2}$ between the full and reduced models. This is equivalent to the proportion of variation explained by the predictors you are testing. It is also equivalent to the squared semipartial correlation of the tested predictors with the response. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## RSQUAREFULL=number-list

RSQFULL=number-list
specifies the $R^{2}$ of the full model, where $R^{2}$ is the proportion of variation explained by the model. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## RSQUAREREDUCED=number-list

RSQREDUCED=number-list
RSQRED=number-list
specifies the $R^{2}$ of the reduced model, where $R^{2}$ is the proportion of variation explained by the model. If the reduced model is an empty or intercept-only model (in other words, if NREDUCEDPREDICTORS $=0$ or NTESTPREDICTORS=NFULLPREDICTORS), then RSQUAREREDUCED=0 is assumed. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TEST=TYPE3

specifies a Type III $F$ test of a set of predictors adjusting for any other predictors in the model. This is the default test option.

## Restrictions on Option Combinations

To specify the number of predictors, use any two of these three options:

- the number of predictors in the full model (NFULLPREDICTORS=)
- the number of predictors in the reduced model (NREDUCEDPREDICTORS=)
- the number of predictors being tested (NTESTPREDICTORS=)

To specify the effect, choose one of the following parameterizations:

- partial correlation (by using the PARTIALCORR=option)
- $R^{2}$ for the full and reduced models (by using any two of RSQUAREDIFF=, RSQUAREFULL=, and RSQUAREREDUCED=)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the MULTREG statement.

## Type III F Test of a Set of Predictors

You can express effects in terms of partial correlation, as in the following statements. Default values of the TEST $=$, MODEL $=$, and ALPHA $=$ options specify a Type III $F$ test with a significance level of 0.05 , assuming normally distributed predictors.

```
proc power;
    multreg
        model = random
        nfullpredictors = 7
        ntestpredictors = 3
        partialcorr = 0.35
        ntotal = 100
        power = .;
run;
```

You can also express effects in terms of $R^{2}$ :

```
proc power;
    multreg
        model = fixed
        nfullpredictors = 7
        ntestpredictors = 3
        rsquarefull = 0.9
        rsquarediff = 0.1
        ntotal = .
        power = 0.9;
run;
```


## ONECORR Statement

## ONECORR < options > ;

The ONECORR statement performs power and sample size analyses for tests of simple and partial Pearson correlation between two variables. Both Fisher's $z$ test and the $t$ test are supported.

## Summary of Options

Table 77.6 summarizes the options available in the ONECORR statement.

Table 77.6 ONECORR Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| DIST= | Specifies the underlying distribution assumed for the test statistic |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| MODEL= | Specifies the assumed distribution of the variables |
| NPARTIALVARS= | Specifies the number of variables adjusted for in the correlation |
| NULLCORR= | Specifies the null value of the correlation |
| SIDES= | Specifies the number of sides and the direction of the statistical test |
| Specify effects |  |
| CORR= | Specifies the correlation |

Table 77.6 continued

| Option | Description |
| :---: | :---: |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NTOTAL= | Specifies the sample size |
| Specify power |  |
| POWER= | Specifies the desired power of the test |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.7 summarizes the valid result parameters in the ONECORR statement.

Table 77.7 Summary of Result Parameters in the ONECORR Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=PEARSON | Power | POWER $=$. |
|  | Sample size | NTOTAL=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## CORR=number-list

specifies the correlation between two variables, possibly adjusting for other variables as determined by the NPARTIALVARS = option. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DIST=FISHERZ | T

specifies the underlying distribution assumed for the test statistic. FISHERZ corresponds to Fisher's $z$ normalizing transformation of the correlation coefficient. T corresponds to the $t$ transformation of the correlation coefficient. Note that DIST=T is equivalent to analyses in the MULTREG statement with NTESTPREDICTORS=1. The default value is FISHERZ.

## MODEL=keyword-list

specifies the assumed distribution of the first variable when DIST=T. The second variable is assumed to have a normal distribution. MODEL=FIXED indicates a fixed distribution. MODEL=RANDOM (the default) indicates a joint bivariate normal distribution with the second variable. You can use the aliases CONDITIONAL for FIXED and UNCONDITIONAL for RANDOM. This option can be used only for DIST=T. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

FIXED fixed variables
RANDOM random (bivariate normal) variables

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NPARTIALVARS=number-list

## NPVARS=number-list

specifies the number of variables adjusted for in the correlation between the two primary variables. The default value is 0 , corresponding to a simple correlation. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). Values for the sample size must be at least $p+3$ when DIST=T and MODEL=CONDITIONAL, and at least $p+4$ when either DIST=FISHER or when DIST=T and MODEL=UNCONDITIONAL, where $p$ is the value of the NPARTIALVARS option. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLCORR=number-list

## NULLC=number-list

specifies the null value of the correlation. The default value is 0 . This option can be used only with the DIST=FISHERZ analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- MODEL=
- SIDES=
- NULL=
- $\mathrm{ALPHA}=$
- NPARTIALVARS=
- CORR=
- NTOTAL=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the ONECORR statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the ONECORR statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test. Valid keywords are
1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value
The default value is 2 .

## TEST=PEARSON

specifies a test of the Pearson correlation coefficient between two variables, possibly adjusting for other variables. This is the default test option.

## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the ONECORR statement.

## Fisher's z Test for Pearson Correlation

The following statements demonstrate a power computation for Fisher's z test for correlation. Default values of TEST=PEARSON, ALPHA= 0.05 , SIDES=2, and NPARTIALVARS $=0$ are assumed.

```
proc power;
    onecorr dist=fisherz
        nullcorr = 0.15
        corr = 0.35
        ntotal = 180
        power = .;
run;
```

$t$ Test for Pearson Correlation
The following statements demonstrate a sample size computation for the $t$ test for correlation. Default values of TEST=PEARSON, MODEL=RANDOM, ALPHA=0.05, and SIDES=2 are assumed.

```
proc power;
    onecorr dist=t
        npartialvars = 4
        corr = 0.45
        ntotal = .
        power = 0.85;
run;
```


## ONESAMPLEFREQ Statement

ONESAMPLEFREQ < options > ;
The ONESAMPLEFREQ statement performs power and sample size analyses for exact and approximate tests (including equivalence, noninferiority, and superiority) and confidence interval precision for a single binomial proportion.

## Summary of Options

Table 77.8 summarizes the options available in the ONESAMPLEFREQ statement.
Table 77.8 ONESAMPLEFREQ Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| $\mathrm{CI}=$ | Specifies an analysis of precision of a confidence interval |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| EQUIVBOUNDS= | Specifies the lower and upper equivalence bounds |
| LOWER= | Specifies the lower equivalence bound |
| MARGIN= | Specifies the equivalence or noninferiority or superiority margin |
| NULLPROPORTION= | Specifies the null proportion |
| SIDES= | Specifies the number of sides and the direction of the statistical test |
| UPPER= | Specifies the upper equivalence bound |
| Specify effect |  |
| HALFWIDTH= | Specifies the desired confidence interval half-width |
| PROPORTION= | Specifies the binomial proportion |
| Specify variance estimation |  |
| VAREST= | Specifies how the variance is computed |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NTOTAL= | Specifies the sample size |
| Specify power and related probabilities |  |
| POWER= | Specifies the desired power of the test |
| PROBWIDTH= | Specifies the probability of obtaining a confidence interval half-width less than or equal to the value specified by HALFWIDTH= |
| Choose computational method |  |
| METHOD= | Specifies the computational method |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.9 summarizes the valid result parameters for different analyses in the ONESAMPLEFREQ statement.

Table 77.9 Summary of Result Parameters in the ONESAMPLEFREQ Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| CI=WILSON | Prob(width) | PROBWIDTH=. |
| CI=AGRESTICOULL | Prob(width) | PROBWIDTH=. |
| CI=JEFFREYS | Prob(width) | PROBWIDTH=. |
| CI=EXACT | Prob(width) | PROBWIDTH=. |
| CI=WALD | Prob(width) | PROBWIDTH=. |
| CI=WALD_CORRECT | Prob(width) | PROBWIDTH=. |
| TEST=ADJZ METHOD=EXACT | Power | POWER=. |
| TEST=ADJZ METHOD=NORMAL | Power | POWER=. |
|  | Sample size | NTOTAL=. |
| TEST=EQUIV_ADJZ METHOD=EXACT | Power | POWER=. |
| TEST=EQUIV_ADJZ METHOD=NORMAL | Power | POWER=. |
|  | Sample size | NTOTAL=. |
| TEST=EQUIV_EXACT | Power | POWER=. |
| TEST=EQUIV_Z METHOD=EXACT | Power | POWER=. |
| TEST=EQUIV_Z METHOD=NORMAL | Power | POWER=. |
|  | Sample size | NTOTAL=. |
| TEST=EXACT | Power | POWER=. |
| TEST=Z METHOD=EXACT | Power | POWER=. |
| TEST=Z METHOD=NORMAL | Power | POWER=. |
|  | Sample size | NTOTAL=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. If the CI= and SIDES=1 options are used, then the value must be less than 0.5 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

CI
CI=AGRESTICOULL | AC
CI=JEFFREYS
Cl=EXACT | CLOPPERPEARSON | CP
CI=WALD
Cl=WALD_CORRECT
Cl=WILSON | SCORE
specifies an analysis of precision of a confidence interval for the sample binomial proportion.

The value of the CI= option specifies the type of confidence interval. The CI=AGRESTICOULL option is a generalization of the "Adjusted Wald / add 2 successes and 2 failures" interval of Agresti and Coull (1998) and is presented in Brown, Cai, and DasGupta (2001). It corresponds to the TABLES / BINOMIAL (AGRESTICOULL) option in PROC FREQ. The CI=JEFFREYS option specifies the equal-tailed Jeffreys prior Bayesian interval, corresponding to the TABLES / BINOMIAL (JEFFREYS) option in PROC FREQ. The CI=EXACT option specifies the exact Clopper-Pearson confidence interval based on enumeration, corresponding to the TABLES / BINOMIAL (EXACT) option in PROC FREQ. The CI=WALD option specifies the confidence interval based on the Wald test (also commonly called the $z$ test or normal-approximation test), corresponding to the TABLES / BINOMIAL (WALD) option in PROC FREQ. The CI=WALD_CORRECT option specifies the confidence interval based on the Wald test with continuity correction, corresponding to the TABLES / BINOMIAL (CORRECT WALD) option in PROC FREQ. The CI=WILSON option (the default) specifies the confidence interval based on the score statistic, corresponding to the TABLES / BINOMIAL (WILSON) option in PROC FREQ.

Instead of power, the relevant probability for this analysis is the probability of achieving a desired precision. Specifically, it is the probability that the half-width of the confidence interval will be at most the value specified by the HALFWIDTH= option.

## EQUIVBOUNDS=grouped-number-list

specifies the lower and upper equivalence bounds, representing the same information as the combination of the LOWER = and UPPER= options but grouping them together. The EQUIVBOUNDS= option can be used only with equivalence analyses (TEST=EQUIV_ADJZ I EQUIV_EXACT I EQUIV_Z). Values must be strictly between 0 and 1 . For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## HALFWIDTH=number-list

specifies the desired confidence interval half-width. The half-width for a two-sided interval is the length of the confidence interval divided by two. This option can be used only with the CI= analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## LOWER=number-list

specifies the lower equivalence bound for the binomial proportion. The LOWER= option can be used only with equivalence analyses (TEST=EQUIV_ADJZ I EQUIV_EXACT I EQUIV_Z). Values must be strictly between 0 and 1 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## MARGIN=number-list

specifies the equivalence or noninferiority or superiority margin, depending on the analysis.
The MARGIN= option can be used with one-sided analyses (SIDES $=1|\mathrm{U}| \mathrm{L}$ ), in which case it specifies the margin added to the null proportion value in the hypothesis test, resulting in a noninferiority or superiority test (depending on the agreement between the effect and hypothesis directions and the sign of the margin). A test with a null proportion $p_{0}$ and a margin $m$ is the same as a test with null proportion $p_{0}+m$ and no margin.

The MARGIN= option can also be used with equivalence analyses (TEST=EQUIV_ADJZ I EQUIV_EXACT I EQUIV_Z) when the NULLPROPORTION= option is used, in which case it specifies the lower and upper equivalence bounds as $p_{0}-m$ and $p_{0}+m$, where $p_{0}$ is the value of the NULLPROPORTION $=$ option and $m$ is the value of the MARGIN=option.

The MARGIN= option cannot be used in conjunction with the SIDES=2 option. (Instead, specify an equivalence analysis by using TEST=EQUIV_ADJZ or TEST=EQUIV_EXACT or TEST=EQUIV_Z). Also, the MARGIN= option cannot be used with the $\mathrm{CI}=$ option.

Values must be strictly between -1 and 1 . In addition, the sum of NULLPROPORTION and MARGIN must be strictly between 0 and 1 for one-sided analyses, and the derived lower equivalence bound ( 2 * NULLPROPORTION - MARGIN) must be strictly between 0 and 1 for equivalence analyses.

For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## METHOD=EXACT | NORMAL

specifies the computational method. METHOD=EXACT (the default) computes exact results by using the binomial distribution. METHOD=NORMAL computes approximate results by using the normal approximation to the binomial distribution.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. This option is invalid when the METHOD=EXACT option is specified. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLPROPORTION=number-list

## NULLP=number-list

specifies the null proportion. A value of 0.5 corresponds to the sign test. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES=
- NULLPROPORTION=
- ALPHA=
- PROPORTION=
- NTOTAL=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the ONESAMPLEFREQ statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the ONESAMPLEFREQ statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PROBWIDTH=number-list

specifies the desired probability of obtaining a confidence interval half-width less than or equal to the value specified by the HALFWIDTH= option. A missing value (PROBWIDTH=.) requests a solution for this probability. Values are expressed as probabilities (for example, 0.9 ) rather than percentages. This option can be used only with the $\mathrm{CI}=$ analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PROPORTION=number-list

## $\mathrm{P}=$ number-list

specifies the binomial proportion-that is, the expected proportion of successes in the hypothetical binomial trial. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value
If the effect size is zero, then SIDES=1 is not permitted; instead, specify the direction of the test explicitly in this case with either SIDES $=\mathrm{L}$ or $\operatorname{SIDES}=\mathrm{U}$. The default value is 2 .

## TEST= ADJZ | EQUIV_ADJZ |EQUIV_EXACT |EQUIV_Z|EXACT | Z <br> TEST

specifies the statistical analysis. TEST=ADJZ specifies a normal-approximate $z$ test with continuity adjustment. TEST=EQUIV_ADJZ specifies a normal-approximate two-sided equivalence test based on the $z$ statistic with continuity adjustment and a TOST (two one-sided tests) procedure. TEST=EQUIV_EXACT specifies the exact binomial two-sided equivalence test based on a TOST (two one-sided tests) procedure. TEST=EQUIV_Z specifies a normal-approximate two-sided equivalence test based on the $z$ statistic without any continuity adjustment, which is the same as the chi-square statistic, and a TOST (two one-sided tests) procedure. TEST or TEST=EXACT (the default) specifies the exact binomial test. TEST $=\mathrm{Z}$ specifies a normal-approximate $z$ test without any continuity adjustment, which is the same as the chi-square test when SIDES=2.

## UPPER=number-list

specifies the upper equivalence bound for the binomial proportion. The UPPER= option can be used only with equivalence analyses (TEST=EQUIV_ADJZ I EQUIV_EXACT I EQUIV_Z). Values must be strictly between 0 and 1 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## VAREST=keyword-list

specifies how the variance is computed in the test statistic for the TEST=Z, TEST=ADJZ, TEST=EQUIV_Z, and TEST=EQUIV_ADJZ analyses. For information about specifying the keywordlist, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords are as follows:

NULL (the default) estimates the variance by using the null proportion(s) (specified by some combination of the NULLPROPORTION=, MARGIN=, LOWER=, and UPPER= options). For TEST=Z and TEST=ADJZ, the null proportion is the value of the NULLPROPORTION= option plus the value of the MARGIN= option (if it is used). For TEST=EQUIV_Z and TEST=EQUIV_ADJZ, there are two null proportions, corresponding to the lower and upper equivalence bounds, one for each test in the TOST (two one-sided tests) procedure.
SAMPLE estimates the variance by using the observed sample proportion.
This option is ignored if the analysis is one other than TEST=Z, TEST=ADJZ, TEST=EQUIV_Z, or TEST=EQUIV_ADJZ.

## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the ONESAMPLEFREQ statement.

## Exact Test of a Binomial Proportion

The following statements demonstrate a power computation for the exact test of a binomial proportion. Defaults for the SIDES = and ALPHA = options specify a two-sided test with a 0.05 significance level.

```
proc power;
    onesamplefreq test=exact
        nullproportion = 0.2
        proportion = 0.3
        ntotal = 100
        power = .;
run;
```


## z Test

The following statements demonstrate a sample size computation for the $z$ test of a binomial proportion. Defaults for the SIDES=, ALPHA=, and VAREST= options specify a two-sided test with a 0.05 significance level that uses the null variance estimate.

```
proc power;
    onesamplefreq test=z method=normal
        nullproportion = 0.8
        proportion = 0.85
        sides = u
        ntotal = .
        power = .9;
run;
```


## z Test with Continuity Adjustment

The following statements demonstrate a sample size computation for the $z$ test of a binomial proportion with a continuity adjustment. Defaults for the SIDES=, ALPHA=, and VAREST= options specify a two-sided test with a 0.05 significance level that uses the null variance estimate.

```
proc power;
    onesamplefreq test=adjz method=normal
        nullproportion = 0.15
        proportion = 0.1
        sides = l
        ntotal = .
        power = .9;
run;
```


## Exact Equivalence Test for a Binomial Proportion

You can specify equivalence bounds by using the EQUIVBOUNDS= option, as in the following statements:

```
proc power;
    onesamplefreq test=equiv_exact
        proportion = 0.35
        equivbounds = (0.2 0.4)
        ntotal = 50
        power = .;
run;
```

You can also specify the combination of NULLPROPORTION= and MARGIN=options:

```
proc power;
    onesamplefreq test=equiv_exact
        proportion = 0.35
        nullproportion = 0.3
        margin = 0.1
        ntotal = 50
        power = .;
run;
```

Finally, you can specify the combination of LOWER= and UPPER= options:

```
proc power;
    onesamplefreq test=equiv_exact
        proportion = 0.35
        lower = 0.2
        upper = 0.4
        ntotal = 50
        power = .;
run;
```

Note that the three preceding analyses are identical.

## Exact Noninferiority Test for a Binomial Proportion

A noninferiority test corresponds to an upper one-sided test with a negative-valued margin, as demonstrated in the following statements:

```
proc power;
    onesamplefreq test=exact
    sides = U
    proportion = 0.15
    nullproportion = 0.1
    margin = -0.02
    ntotal = 130
    power = .;
run;
```


## Exact Superiority Test for a Binomial Proportion

A superiority test corresponds to an upper one-sided test with a positive-valued margin, as demonstrated in the following statements:

```
proc power;
    onesamplefreq test=exact
        sides = U
        proportion = 0.15
        nullproportion = 0.1
        margin = 0.02
        ntotal = 130
        power = .;
run;
```


## Confidence Interval Precision

The following statements performs a confidence interval precision analysis for the Wilson score-based confidence interval for a binomial proportion. The default value of the ALPHA= option specifies a confidence level of 0.95 .

```
proc power;
    onesamplefreq ci=wilson
        halfwidth = 0.1
        proportion = 0.3
        ntotal = 70
        probwidth = .;
run;
```


## Restrictions on Option Combinations

To specify the equivalence bounds for TEST=EQUIV_ADJZ, TEST=EQUIV_EXACT, and TEST=EQUIV_Z, use any of these three option sets:

- lower and upper equivalence bounds, using the EQUIVBOUNDS= option
- lower and upper equivalence bounds, using the LOWER= and UPPER= options
- null proportion (NULLPROPORTION=) and margin (MARGIN=)


## ONESAMPLEMEANS Statement

ONESAMPLEMEANS < options> ;
The ONESAMPLEMEANS statement performs power and sample size analyses for $t$ tests, equivalence tests, and confidence interval precision involving one sample.

## Summary of Options

Table 77.10 summarizes the options available in the ONESAMPLEMEANS statement.

## Table 77.10 ONESAMPLEMEANS Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| $\mathrm{CI}=$ | Specifies an analysis of precision of the confidence interval for the mean |
| DIST= | Specifies the underlying distribution assumed for the test statistic |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| LOWER= | Specifies the lower equivalence bound for the mean |
| NULLMEAN= | Specifies the null mean |
| SIDES= | Specifies the number of sides and the direction of the statistical test |
| UPPER= | Specifies the upper equivalence bound for the mean |
| Specify effect |  |
| HALFWIDTH= | Specifies the desired confidence interval half-width |
| MEAN= | Specifies the mean |
| Specify variability |  |
| CV= | Specifies the coefficient of variation |
| STDDEV= | Specifies the standard deviation |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NTOTAL= | Specifies the sample size |
| Specify power and related probabilities |  |
| POWER= | Specifies the desired power of the test |
| PROBTYPE= | Specifies the type of probability for the PROBWIDTH= option |
| PROBWIDTH= | Specifies the probability of obtaining a confidence interval half-width less than or equal to the value specified by HALFWIDTH= |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.11 summarizes the valid result parameters for different analyses in the ONESAMPLEMEANS statement.

## Table 77.11 Summary of Result Parameters in the

 ONESAMPLEMEANS Statement| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=T DIST=NORMAL | Power | POWER=. |
|  | Sample size | NTOTAL=. |
|  | Alpha | ALPHA=. |
|  | Mean | MEAN=. |
|  | Standard Deviation | STDDEV=. |
| TEST=T DIST=LOGNORMAL | Power | POWER=. |
|  | Sample size | NTOTAL=. |
|  | Power | POWER=. |
|  | Sample size | NTOTAL=. |
| CI=T | Prob(width) | PROBWIDTH=. |
|  | Sample size | NTOTAL=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test or requests a solution for alpha with a missing value ( $\mathrm{ALPHA}=$.). The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. If the $\mathrm{CI}=$ and $\mathrm{SIDES}=1$ options are used, then the value must be less than 0.5 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Cl

## $\mathrm{Cl}=\mathrm{T}$

specifies an analysis of precision of the confidence interval for the mean. Instead of power, the relevant probability for this analysis is the probability of achieving a desired precision. Specifically, it is the probability that the half-width of the confidence interval will be at most the value specified by the HALFWIDTH = option. If neither the $\mathrm{CI}=$ option nor the TEST= option is used, the default is TEST=T.

## CV=number-list

specifies the coefficient of variation, defined as the ratio of the standard deviation to the mean on the original data scale. You can use this option only with DIST=LOGNORMAL. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DIST=LOGNORMAL | NORMAL

specifies the underlying distribution assumed for the test statistic. NORMAL corresponds the normal distribution, and LOGNORMAL corresponds to the lognormal distribution. The default value is NORMAL.

## HALFWIDTH=number-list

specifies the desired confidence interval half-width. The half-width is defined as the distance between the point estimate and a finite endpoint. This option can be used only with the CI=T analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## LOWER=number-list

specifies the lower equivalence bound for the mean. This option can be used only with the TEST=EQUIV analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## MEAN=number-list

specifies the mean, in the original scale, or requests a solution for the mean with a missing value (MEAN=.). The mean is arithmetic if DIST=NORMAL and geometric if DIST=LOGNORMAL. This option can be used only with the TEST=T and TEST=EQUIV analyses. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLMEAN=number-list

NULLM =number-list
specifies the null mean, in the original scale (whether DIST=NORMAL or DIST=LOGNORMAL). The default value is 0 when DIST=NORMAL and 1 when DIST=LOGNORMAL. This option can be used only with the TEST=T analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES=
- NULLMEAN=
- LOWER=
- UPPER=
- ALPHA=
- MEAN=
- HALFWIDTH=
- STDDEV=
- CV=
- NTOTAL=
- POWER=
- PROBTYPE=
- PROBWIDTH=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the ONESAMPLEMEANS statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the ONESAMPLEMEANS statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. This option can be used only with the TEST=T and TEST=EQUIV analyses. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PROBTYPE=keyword-list

specifies the type of probability for the PROBWIDTH= option. A value of CONDITIONAL (the default) indicates the conditional probability that the confidence interval half-width is at most the value specified by the HALFWIDTH= option, given that the true mean is captured by the confidence interval. A value of UNCONDITIONAL indicates the unconditional probability that the confidence interval half-width is at most the value specified by the HALFWIDTH= option. You can use the alias GIVENVALIDITY for CONDITIONAL. The PROBTYPE= option can be used only with the CI=T analysis. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

CONDITIONAL width probability conditional on interval containing the mean
UNCONDITIONAL unconditional width probability

## PROBWIDTH=number-list

specifies the desired probability of obtaining a confidence interval half-width less than or equal to the value specified by the HALFWIDTH= option. A missing value (PROBWIDTH=.) requests a solution for this probability. The type of probability is controlled with the PROBTYPE= option. Values are expressed as probabilities (for example, 0.9) rather than percentages. This option can be used only with the $\mathrm{CI}=\mathrm{T}$ analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test or confidence interval. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords and their interpretation for the TEST= analyses are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value

For confidence intervals, SIDES=U refers to an interval between the lower confidence limit and infinity, and SIDES=L refers to an interval between minus infinity and the upper confidence limit. For both of these cases and SIDES $=1$, the confidence interval computations are equivalent. The SIDES $=$ option can be used only with the $\mathrm{TEST}=\mathrm{T}$ and $\mathrm{CI}=\mathrm{T}$ analyses. The default value is 2 .

## STDDEV=number-list

## STD=number-list

specifies the standard deviation, or requests a solution for the standard deviation with a missing value (STDDEV=.). You can use this option only with DIST=NORMAL. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TEST=EQUIV | T

TEST
specifies the statistical analysis. TEST=EQUIV specifies an equivalence test of the mean by using a two one-sided tests (TOST) analysis (Schuirmann 1987). TEST or TEST=T (the default) specifies a $t$ test on the mean. If neither the $\mathrm{TEST}=$ option nor the $\mathrm{CI}=$ option is used, the default is TEST=T.

## UPPER=number-list

specifies the upper equivalence bound for the mean, in the original scale (whether DIST=NORMAL or DIST=LOGNORMAL). This option can be used only with the TEST=EQUIV analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Restrictions on Option Combinations

To define the analysis, choose one of the following parameterizations:

- a statistical test (by using the TEST= option)
- confidence interval precision (by using the $\mathrm{CI}=$ option)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the ONESAMPLEMEANS statement.

## One-Sample t Test

The following statements demonstrate a power computation for the one-sample $t$ test. Default values for the DIST=, SIDES=, NULLMEAN=, and ALPHA= options specify a two-sided test for zero mean with a normal distribution and a significance level of 0.05 .

```
proc power;
    onesamplemeans test=t
        mean = 7
        stddev = 3
        ntotal = 50
        power = .;
run;
```


## One-Sample t Test with Lognormal Data

The following statements demonstrate a sample size computation for the one-sample $t$ test for lognormal data. Default values for the SIDES=, NULLMEAN=, and ALPHA= options specify a two-sided test for unit mean with a significance level of 0.05 .

```
proc power;
    onesamplemeans test=t dist=lognormal
        mean = 7
        cv = 0.8
        ntotal = .
        power = 0.9;
run;
```


## Equivalence Test for Mean of Normal Data

The following statements demonstrate a power computation for the TOST equivalence test for a normal mean. Default values for the DIST= and ALPHA= options specify a normal distribution and a significance level of 0.05 .

```
proc power;
    onesamplemeans test=equiv
        lower = 2
        upper = 7
        mean = 4
        stddev = 3
        ntotal = 100
        power = .;
run;
```


## Equivalence Test for Mean of Lognormal Data

The following statements demonstrate a sample size computation for the TOST equivalence test for a lognormal mean. The default of ALPHA $=0.05$ specifies a significance level of 0.05 .

```
proc power;
    onesamplemeans test=equiv dist=lognormal
        lower = 1
        upper = 5
        mean = 3
        cv = 0.6
        ntotal = .
        power = 0.85;
run;
```


## Confidence Interval for Mean

By default $\mathrm{CI}=\mathrm{T}$ analyzes the conditional probability of obtaining the desired precision, given that the interval contains the true mean, as in the following statements. The defaults of SIDES $=2$ and ALPHA $=0.05$ specify a two-sided interval with a confidence level of 0.95 .

```
proc power;
    onesamplemeans ci = t
        halfwidth = 14
        stddev = 8
        ntotal = 50
        probwidth = .;
run;
```


## ONEWAYANOVA Statement

ONEWAYANOVA < options > ;
The ONEWAYANOVA statement performs power and sample size analyses for one-degree-of-freedom contrasts and the overall $F$ test in one-way analysis of variance.

## Summary of Options

Table 77.12 summarizes the options available in the ONEWAYANOVA statement.

Table 77.12 ONEWAYANOVA Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| CONTRAST= | Specifies coefficients for single-degree-of-freedom hypothesis tests |
| NULLCONTRAST= | Specifies the null value of the contrast |
| SIDES= | Specifies the number of sides and the direction of the statistical test |
| Specify effect |  |
| GROUPMEANS= | Specifies the group means |
| Specify variability |  |
| STDDEV= | Specifies the error standard deviation |
| Specify sample size and allocation |  |
| GROUPNS= | Specifies the group sample sizes |
| GROUPWEIGHTS= | Specifies the sample size allocation weights for the groups |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NPERGROUP= | Specifies the common sample size per group |
| NTOTAL= | Specifies the sample size |
| Specify power |  |
| POWER= | Specifies the desired power of the test |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.13 summarizes the valid result parameters for different analyses in the ONEWAYANOVA statement.

## Table 77.13 Summary of Result Parameters in the ONEWAYANOVA Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=CONTRAST | Power | POWER=. |
|  | Sample size | NTOTAL=. |
|  |  | NPERGROUP==. |
| TEST=OVERALL | Power <br> Sample size | POWER=. |
|  |  | NTOTAL=. |
|  | NPERGROUP==. |  |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

CONTRAST $=$ ( values ) < (. . values ) > specifies coefficients for single-degree-of-freedom hypothesis tests. You must provide a coefficient for every mean appearing in the GROUPMEANS= option. Specify multiple contrasts either with additional sets of coefficients or with additional CONTRAST= options. For example, you can specify two different contrasts of five means by using the following:

```
CONTRAST =((1 -1 0 0 0) (1 0 - (1 0 0)
```


## GROUPMEANS=grouped-number-list

## GMEANS=grouped-number-list

specifies the group means. This option is used to implicitly set the number of groups. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPNS=grouped-number-list

GNS=grouped-number-list
specifies the group sample sizes. The number of groups represented must be the same as with the GROUPMEANS = option. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPWEIGHTS=grouped-number-list

## GWEIGHTS=grouped-number-list

specifies the sample size allocation weights for the groups. This option controls how the total sample size is divided between the groups. Each set of values across all groups represents relative allocation weights. Additionally, if the NFRACTIONAL option is not used, the total sample size is restricted to be equal to a multiple of the sum of the group weights (so that the resulting design has an integer sample size for each group while adhering exactly to the group allocation weights). The number of groups represented must be the same as with the GROUPMEANS= option. Values must be integers unless
the NFRACTIONAL option is used. The default value is 1 for each group, amounting to a balanced design. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NPERGROUP=number-list

## NPERG=number-list

specifies the common sample size per group or requests a solution for the common sample size per group with a missing value (NPERGROUP==.). Use of this option implicitly specifies a balanced design. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLCONTRAST=number-list

NULLC=number-list
specifies the null value of the contrast. The default value is 0 . This option can be used only with the TEST=CONTRAST analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- CONTRAST=
- SIDES=
- NULLCONTRAST=
- ALPHA=
- GROUPMEANS=
- STDDEV=
- GROUPWEIGHTS=
- NTOTAL=
- NPERGROUP==
- GROUPNS=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the ONEWAYANOVA statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the ONEWAYANOVA statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value
This option can be used only with the TEST=CONTRAST analysis. The default value is 2 .

## STDDEV=number-list

STD=number-list
specifies the error standard deviation. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TEST=CONTRAST | OVERALL

specifies the statistical analysis. TEST=CONTRAST specifies a one-degree-of-freedom test of a contrast of means. The test is the usual $F$ test for the two-sided case and the usual $t$ test for the one-sided case. TEST=OVERALL specifies the overall $F$ test of equality of all means. The default is TEST=CONTRAST if the CONTRAST= option is used, and TEST=OVERALL otherwise.

## Restrictions on Option Combinations

To specify the sample size and allocation, choose one of the following parameterizations:

- sample size per group in a balanced design (by using the NPERGROUP== option)
- total sample size and allocation weights (by using the NTOTAL= and GROUPWEIGHTS= options)
- individual group sample sizes (by using the GROUPNS= option)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the ONEWAYANOVA statement.

## One-Degree-of-Freedom Contrast

You can use the NPERGROUP== option in a balanced design, as in the following statements. Default values for the SIDES=, NULLCONTRAST=, and ALPHA= options specify a two-sided test for a contrast value of 0 with a significance level of 0.05 .

```
proc power;
    onewayanova test=contrast
        contrast = (1 0 -1)
        groupmeans = 3 | 7 | 8
        stddev = 4
        npergroup = 50
        power = .;
run;
```

You can also specify an unbalanced design with the NTOTAL= and GROUPWEIGHTS= options:

```
proc power;
    onewayanova test=contrast
            contrast = (1 0 -1)
            groupmeans = 3 | 7 | 8
            stddev = 4
            groupweights = (1 2 2)
            ntotal = .
            power = 0.9;
run;
```

Another way to specify the sample sizes is with the GROUPNS= option:

```
proc power;
    onewayanova test=contrast
        contrast = (1 0 -1)
        groupmeans = 3 | 7 | 8
        stddev = 4
        groupns = (20 40 40)
        power = .;
run;
```


## Overall F Test

The following statements demonstrate a power computation for the overall $F$ test in a one-way ANOVA. The default of ALPHA $=0.05$ specifies a significance level of 0.05 .

```
proc power;
    onewayanova test=overall
        groupmeans = 3 | 7 | 8
        stddev = 4
        npergroup = 50
        power = .;
run;
```


## PAIREDFREQ Statement

PAIREDFREQ < options > ;
The PAIREDFREQ statement performs power and sample size analyses for McNemar's test for paired proportions.

## Summary of Options

Table 77.14 summarizes the options available in the PAIREDFREQ statement.

Table 77.14 PAIREDFREQ Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| DIST= | Specifies the underlying distribution assumed for the test statistic |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| NULLDISCPROPRATIO= | Specifies the null value of the ratio of discordant proportions |
| SIDES= | Specifies the number of sides and the direction of the statistical test or confidence interval |
| Specify effects |  |
| CORR= | Specifies the correlation $\phi$ between members of a pair |
| DISCPROPDIFF= | Specifies the discordant proportion difference $p_{01}-p_{10}$ |
| DISCPROPORTIONS= | Specifies the two discordant proportions, $p_{10}$ and $p_{01}$ |
| DISCPROPRATIO= | Specifies the ratio $p_{01} / p_{10}$ |
| ODDSRATIO= | Specifies the odds ratio $[p .1 /(1-p .1)] /\left[p_{1} \cdot /\left(1-p_{1} \cdot\right)\right]$ |
| PAIREDPROPORTIONS= | Specifies the two paired proportions, $p_{1}$. and $p_{\cdot 1}$ |
| PROPORTIONDIFF= | Specifies the proportion difference $p \cdot 1-p_{1}$. |
| REFPROPORTION= | Specifies either the reference first proportion $p_{1}$. or the reference discordant proportion $p_{10}$ |
| RELATIVERISK= | Specifies the relative risk p.1/ $p_{1}$. |
| TOTALPROPDISC= | Specifies the discordant proportion sum, $p_{10}+p_{01}$ |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NPAIRS= | Specifies the total number of proportion pairs |
| Specify power |  |
| POWER= | Specifies the desired power of the test |
| Choose computational method |  |
| METHOD= | Specifies the computational method |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.15 summarizes the valid result parameters in the PAIREDFREQ statement.

Table 77.15 Summary of Result Parameters in the PAIREDFREQ Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=MCNEMAR METHOD=CONNOR | Power | POWER=. |
|  | Sample size | NPAIRS=. |

Table 77.15 continued

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=MCNEMAR METHOD=EXACT | Power | POWER=. |
| TEST=MCNEMAR METHOD=MIETTINEN | Power | POWER=. |
|  | Sample size | NPAIRS=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## CORR=number-list

specifies the correlation $\phi$ between members of a pair. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DISCPROPORTIONS=grouped-number-list

## DISCPS=grouped-number-list

specifies the two discordant proportions, $p_{10}$ and $p_{01}$. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DISCPROPDIFF=number-list

DISCPDIFF=number-list
specifies the difference $p_{01}-p_{10}$ between discordant proportions. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DISCPROPRATIO=number-list

## DISCPRATIO=number-list

specifies the ratio $p_{01} / p_{10}$ of discordant proportions. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DIST=EXACT_COND | NORMAL

specifies the underlying distribution assumed for the test statistic. EXACT_COND corresponds to the exact conditional test, based on the exact binomial distribution of the two types of discordant pairs given the total number of discordant pairs. NORMAL corresponds to the conditional test based on the normal approximation to the binomial distribution of the two types of discordant pairs given the total number of discordant pairs. The default value is EXACT_COND.

## METHOD=CONNOR | EXACT | MIETTINEN

specifies the computational method. METHOD=EXACT (the default) uses the exact binomial distributions of the total number of discordant pairs and the two types of discordant pairs. METHOD=CONNOR uses an approximation from Connor (1987), and METHOD=MIETTINEN uses an approximation from Miettinen (1968). The CONNOR and MIETTINEN methods are valid only for DIST $=$ NORMAL.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option. This option cannot be used with METHOD=EXACT.

## NPAIRS=number-list

specifies the total number of proportion pairs (concordant and discordant) or requests a solution for the number of pairs with a missing value (NPAIRS=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLDISCPROPRATIO=number-list

## NULLDISCPRATIO=number-list

## NULLRATIO=number-list

## NULLR=number-list

specifies the null value of the ratio of discordant proportions. The default value is 1 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## ODDSRATIO=number-list

OR=number-list
specifies the odds ratio $\left[p \cdot 1 /\left(1-p_{\cdot 1}\right)\right] /\left[p_{1} \cdot /\left(1-p_{1}.\right)\right]$. For information about specifying the numberlist, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES=
- NULLDISCPROPRATIO=
- ALPHA=
- PAIREDPROPORTIONS=
- PROPORTIONDIFF=
- ODDSRATIO=
- RELATIVERISK=
- CORR=
- DISCPROPORTIONS=
- DISCPROPDIFF=
- TOTALPROPDISC=
- REFPROPORTION=
- DISCPROPRATIO=
- NPAIRS=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the PAIREDFREQ statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the PAIREDFREQ statement.

## PAIREDPROPORTIONS=grouped-number-list

PPROPORTIONS=grouped-number-list
PAIREDPS=grouped-number-list

## PPS=grouped-number-list

specifies the two paired proportions, $p_{1}$. and $p_{\cdot 1}$. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PROPORTIONDIFF=number-list

## PDIFF=number-list

specifies the proportion difference $p .1-p_{1}$.. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## REFPROPORTION=number-list

## REFP=number-list

specifies either the reference first proportion $p_{1}$. (when used in conjunction with the PROPORTIONDIFF $=$, ODDSRATIO $=$, or RELATIVERISK $=$ option) or the reference discordant proportion $p_{10}$ (when used in conjunction with the DISCPROPDIFF= or DISCPROPRATIO= option). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## RELATIVERISK=number-list

$\mathbf{R R}=$ number-list
specifies the relative risk $p \cdot 1 / p_{1}$.. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test or confidence interval. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords and their interpretation are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value
If the effect size is zero, then SIDES=1 is not permitted; instead, specify the direction of the test explicitly in this case with either SIDES=L or SIDES=U. The default value is 2 .

## TEST=MCNEMAR

specifies the McNemar test of paired proportions. This is the default test option.

## TOTALPROPDISC=number-list

TOTALPDISC=number-list
PDISC=number-list
specifies the sum of the two discordant proportions, $p_{10}+p_{01}$. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Restrictions on Option Combinations

To specify the proportions, choose one of the following parameterizations:

- discordant proportions (using the DISCPROPORTIONS= option)
- difference and sum of discordant proportions (using the DISCPROPDIFF= and TOTALPROPDISC=options)
- difference of discordant proportions and reference discordant proportion (using the DISCPROPDIFF= and REFPROPORTION= options)
- ratio of discordant proportions and reference discordant proportion (using the DISCPROPRATIO= and REFPROPORTION= options)
- ratio and sum of discordant proportions (using the DISCPROPRATIO= and TOTALPROPDISC=options)
- paired proportions and correlation (using the PAIREDPROPORTIONS $=$ and CORR=options)
- proportion difference, reference proportion, and correlation (using the PROPORTIONDIFF=, REFPROPORTION=, and CORR= options)
- odds ratio, reference proportion, and correlation (using the ODDSRATIO=, REFPROPORTION=, and CORR= options)
- relative risk, reference proportion, and correlation (using the RELATIVERISK=, REFPROPORTION=, and CORR= options)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the PAIREDFREQ statement.

## McNemar Exact Conditional Test

You can express effects in terms of the individual discordant proportions, as in the following statements. Default values for the TEST=, SIDES=, ALPHA=, and NULLDISCPROPRATIO= options specify a twosided McNemar test for no effect with a significance level of 0.05.

```
proc power;
    pairedfreq dist=exact_cond
        discproportions = 0.15 | 0.45
        npairs = 80
        power = .;
run;
```

You can also express effects in terms of the difference and sum of discordant proportions:

```
proc power;
    pairedfreq dist=exact_cond
        discpropdiff = 0.3
        totalpropdisc = 0.6
        npairs = 80
        power = .;
run;
```

You can also express effects in terms of the difference of discordant proportions and the reference discordant proportion:

```
proc power;
    pairedfreq dist=exact_cond
        discpropdiff = 0.3
        refproportion = 0.15
        npairs = 80
        power = .;
run;
```

You can also express effects in terms of the ratio of discordant proportions and the denominator of the ratio:

```
proc power;
    pairedfreq dist=exact_cond
        discpropratio = 3
        refproportion = 0.15
        npairs = 80
        power = .;
run;
```

You can also express effects in terms of the ratio and sum of discordant proportions:

```
proc power;
    pairedfreq dist=exact_cond
            discpropratio = 3
            totalpropdisc = 0.6
            npairs = 80
            power = .;
run;
```

You can also express effects in terms of the paired proportions and correlation:

```
proc power;
    pairedfreq dist=exact_cond
            pairedproportions = 0.6 | 0.8
            corr = 0.4
            npairs = 45
            power = .;
run;
```

You can also express effects in terms of the proportion difference, reference proportion, and correlation:

```
proc power;
    pairedfreq dist=exact_cond
        proportiondiff = 0.2
        refproportion = 0.6
```

```
    corr = 0.4
    npairs = 45
    power = .;
run;
```

You can also express effects in terms of the odds ratio, reference proportion, and correlation:

```
proc power;
    pairedfreq dist=exact_cond
        oddsratio = 2.66667
        refproportion = 0.6
        corr = 0.4
        npairs = 45
        power = .;
run;
```

You can also express effects in terms of the relative risk, reference proportion, and correlation:

```
proc power;
    pairedfreq dist=exact_cond
        relativerisk = 1.33333
        refproportion = 0.6
        corr = 0.4
        npairs = 45
        power = .;
run;
```


## McNemar Normal Approximation Test

The following statements demonstrate a sample size computation for the normal-approximate McNemar test. The default value for the METHOD= option specifies an exact sample size computation. Default values for the TEST=, SIDES=, ALPHA=, and NULLDISCPROPRATIO= options specify a two-sided McNemar test for no effect with a significance level of 0.05 .

```
proc power;
    pairedfreq dist=normal method=connor
        discproportions = 0.15 | 0.45
        npairs = .
        power = .9;
run;
```


## PAIREDMEANS Statement

PAIREDMEANS < options > ;
The PAIREDMEANS statement performs power and sample size analyses for $t$ tests, equivalence tests, and confidence interval precision involving paired samples.

## Summary of Options

Table 77.14 summarizes the options available in the PAIREDMEANS statement.

Table 77.16 PAIREDMEANS Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| $\mathrm{CI}=$ | Specifies an analysis of precision of the confidence interval for the mean difference |
| DIST= | Specifies the underlying distribution assumed for the test statistic |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| LOWER= | Specifies the lower equivalence bound |
| NULLDIFF= | Specifies the null mean difference |
| NULLRATIO= | Specifies the null mean ratio |
| SIDES= | Specifies the number of sides and the direction of the statistical test or confidence interval |
| UPPER= | Specifies the upper equivalence bound |
| Specify effects |  |
| HALFWIDTH= | Specifies the desired confidence interval half-width |
| MEANDIFF= | Specifies the mean difference |
| MEANRATIO= | Specifies the geometric mean ratio, $\gamma_{2} / \gamma_{1}$ |
| PAIREDMEANS= | Specifies the two paired means |
| Specify variability |  |
| CORR= | Specifies the correlation between members of a pair |
| $\mathrm{CV}=$ | Specifies the common coefficient of variation |
| PAIREDCVS= | Specifies the coefficient of variation for each member of a pair |
| PAIREDSTDDEVS= | Specifies the standard deviation of each member of a pair |
| STDDEV= | Specifies the common standard deviation |
| Specify sample size |  |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NPAIRS= | Specifies the number of pairs |
| Specify power and related probabilities |  |
| POWER= | Specifies the desired power of the test |
| PROBTYPE= | Specifies the type of probability for the PROBWIDTH= option |
| PROBWIDTH= | Specifies the probability of obtaining a confidence interval half-width less than or equal to the value specified by the HALFWIDTH= |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.17 summarizes the valid result parameters for different analyses in the PAIREDMEANS statement.

Table 77.17 Summary of Result Parameters in the PAIREDMEANS Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=DIFF | Power <br> Sample size | POWER=. |
|  |  |  |


| Table 77.17 continued |  |  |
| :--- | :--- | :--- |
| Analyses | Solve For | Syntax |
| TEST=RATIO | Power | POWER $=$. |
|  | Sample size | NPAIRS $=$. |
| TEST=EQUIV_DIFF | Power | POWER $=$. |
|  | Sample size | NPAIRS=. |
| TEST=EQUIV_RATIO | Power | POWER=. |
|  | Sample size | NPAIRS $=$. |
|  | Prob(width) | PROBWIDTH=. |
|  | Sample size | NPAIRS $=$. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. If the CI= and SIDES $=1$ options are used, then the value must be less than 0.5 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Cl

## CI=DIFF

specifies an analysis of precision of the confidence interval for the mean difference. Instead of power, the relevant probability for this analysis is the probability of achieving a desired precision. Specifically, it is the probability that the half-width of the observed confidence interval will be at most the value specified by the HALFWIDTH= option. If neither the CI= option nor the TEST= option is used, the default is TEST=DIFF.

## CORR=number-list

specifies the correlation between members of a pair. For tests that assume lognormal data (DIST=LOGNORMAL, or TEST=RATIO or TEST=EQUIV_RATIO), values of the CORR= option are restricted to the range ( $\rho_{L}, \rho_{U}$ ), where

$$
\begin{aligned}
& \rho_{L}=\frac{\exp \left(-\left[\log \left(\mathrm{CV}_{1}^{2}+1\right) \log \left(\mathrm{CV}_{2}^{2}+1\right)\right]^{\frac{1}{2}}\right)-1}{\mathrm{CV}_{1} \mathrm{CV}_{2}} \\
& \rho_{U}=\frac{\exp \left(\left[\log \left(\mathrm{CV}_{1}^{2}+1\right) \log \left(\mathrm{CV}_{2}^{2}+1\right)\right]^{\frac{1}{2}}\right)-1}{\mathrm{CV}_{1} \mathrm{CV}_{2}}
\end{aligned}
$$

and $\mathrm{CV}_{1}$ are the $\mathrm{CV}_{2}$ coefficient of variation values specified by the $\mathrm{CV}=$ or PAIREDCVS= option. See "Paired $t$ Test for Mean Ratio with Lognormal Data (TEST=RATIO)" on page 6408 for more information about this restriction on correlation values. For information about specifying the numberlist, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## CV=number-list

specifies the coefficient of variation that is assumed to be common to both members of a pair. The coefficient of variation is defined as the ratio of the standard deviation to the mean on the original data scale. You can use this option only with DIST=LOGNORMAL. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DIST=LOGNORMAL | NORMAL

specifies the underlying distribution assumed for the test statistic. NORMAL corresponds the normal distribution, and LOGNORMAL corresponds to the lognormal distribution. The default value (also the only acceptable value in each case) is NORMAL for TEST=DIFF, TEST=EQUIV_DIFF, and CI=DIFF; and LOGNORMAL for TEST=RATIO and TEST=EQUIV_RATIO.

## HALFWIDTH=number-list

specifies the desired confidence interval half-width. The half-width is defined as the distance between the point estimate and a finite endpoint. This option can be used only with the CI=DIFF analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## LOWER=number-list

specifies the lower equivalence bound for the mean difference or mean ratio, in the original scale (whether DIST=NORMAL or DIST=LOGNORMAL). This option can be used only with the TEST=EQUIV_DIFF and TEST=EQUIV_RATIO analyses. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## MEANDIFF=number-list

specifies the mean difference, defined as the mean of the difference between the second and first members of a pair, $\mu_{2}-\mu_{1}$. This option can be used only with the TEST=DIFF and TEST=EQUIV_DIFF analyses. When TEST=EQUIV_DIFF, the mean difference is interpreted as the treatment mean minus the reference mean. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## MEANRATIO=number-list

specifies the geometric mean ratio, defined as $\gamma_{2} / \gamma_{1}$. This option can be used only with the TEST=RATIO and TEST=EQUIV_RATIO analyses. When TEST=EQUIV_RATIO, the mean ratio is interpreted as the treatment mean divided by the reference mean. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NPAIRS=number-list

specifies the number of pairs or requests a solution for the number of pairs with a missing value (NPAIRS=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLDIFF=number-list

NULLD=number-list
specifies the null mean difference. The default value is 0 . This option can be used only with the TEST=DIFF analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLRATIO=number-list

NULLR=number-list
specifies the null mean ratio. The default value is 1 . This option can be used only with the TEST=RATIO analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES=
- NULLDIFF=
- NULLRATIO=
- LOWER=
- UPPER=
- ALPHA=
- PAIREDMEANS=
- MEANDIFF=
- MEANRATIO=
- HALFWIDTH=
- STDDEV=
- PAIREDSTDDEVS=
- CV=
- PAIREDCVS=
- CORR=
- NPAIRS=
- POWER=
- PROBTYPE=
- PROBWIDTH=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the PAIREDMEANS statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the PAIREDMEANS statement.

## PAIREDCVS=grouped-number-list

specifies the coefficient of variation for each member of a pair. Unlike the CV= option, the PAIREDCVS $=$ option supports different values for each member of a pair. The coefficient of variation is defined as the ratio of the standard deviation to the mean on the original data scale. Values must be nonnegative (unless both are equal to zero, which is permitted). This option can be used only with DIST=LOGNORMAL. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PAIREDMEANS=grouped-number-list

PMEANS=grouped-number-list
specifies the two paired means, in the original scale. The means are arithmetic if DIST=NORMAL and geometric if DIST=LOGNORMAL. This option cannot be used with the CI=DIFF analysis. When TEST=EQUIV_DIFF, the means are interpreted as the reference mean (first) and the treatment mean (second). For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PAIREDSTDDEVS=grouped-number-list

PAIREDSTDS=grouped-number-list
PSTDDEVS=grouped-number-list
PSTDS=grouped-number-list
specifies the standard deviation of each member of a pair. Unlike the STDDEV= option, the PAIREDSTDDEVS = option supports different values for each member of a pair. This option can be used only with DIST=NORMAL. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. This option cannot be used with the CI=DIFF analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PROBTYPE=keyword-list

specifies the type of probability for the PROBWIDTH= option. A value of CONDITIONAL (the default) indicates the conditional probability that the confidence interval half-width is at most the value specified by the HALFWIDTH= option, given that the true mean difference is captured by the confidence interval. A value of UNCONDITIONAL indicates the unconditional probability that the confidence interval half-width is at most the value specified by the HALFWIDTH= option. you can use the alias GIVENVALIDITY for CONDITIONAL. The PROBTYPE= option can be used only with the CI=DIFF analysis. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

CONDITIONAL width probability conditional on interval containing the mean UNCONDITIONAL unconditional width probability

## PROBWIDTH=number-list

specifies the desired probability of obtaining a confidence interval half-width less than or equal to the value specified by the HALFWIDTH= option. A missing value (PROBWIDTH=.) requests a solution for this probability. The type of probability is controlled with the PROBTYPE= option. Values are expressed as probabilities (for example, 0.9 ) rather than percentages. This option can be used only with the CI=DIFF analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test or confidence interval. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords and their interpretation for the TEST= analyses are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value

For confidence intervals, SIDES=U refers to an interval between the lower confidence limit and infinity, and SIDES $=$ L refers to an interval between minus infinity and the upper confidence limit. For both of these cases and SIDES=1, the confidence interval computations are equivalent. The SIDES= option cannot be used with the TEST=EQUIV_DIFF and TEST=EQUIV_RATIO analyses. The default value is 2 .

## STDDEV=number-list

## STD=number-list

specifies the standard deviation assumed to be common to both members of a pair. This option can be used only with DIST=NORMAL. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TEST=DIFF | EQUIV_DIFF | EQUIV_RATIO | RATIO

## TEST

specifies the statistical analysis. TEST or TEST=DIFF (the default) specifies a paired $t$ test on the mean difference. TEST=EQUIV_DIFF specifies an additive equivalence test of the mean difference by using a two one-sided tests (TOST) analysis (Schuirmann 1987). TEST=EQUIV_RATIO specifies a multiplicative equivalence test of the mean ratio by using a TOST analysis. TEST=RATIO specifies a paired $t$ test on the geometric mean ratio. If neither the TEST= option nor the $\mathrm{CI}=$ option is used, the default is TEST=DIFF.

## UPPER=number-list

specifies the upper equivalence bound for the mean difference or mean ratio, in the original scale (whether DIST=NORMAL or DIST=LOGNORMAL). This option can be used only with the TEST=EQUIV_DIFF and TEST=EQUIV_RATIO analyses. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Restrictions on Option Combinations

To define the analysis, choose one of the following parameterizations:

- a statistical test (by using the TEST= option)
- confidence interval precision (by using the CI= option)

To specify the means, choose one of the following parameterizations:

- individual means (by using the PAIREDMEANS= option)
- mean difference (by using the MEANDIFF= option)
- mean ratio (by using the MEANRATIO= option)

To specify the coefficient of variation, choose one of the following parameterizations:

- common coefficient of variation (by using the $\mathrm{CV}=$ option)
- individual coefficients of variation (by using the PAIREDCVS= option)

To specify the standard deviation, choose one of the following parameterizations:

- common standard deviation (by using the STDDEV= option)
- individual standard deviations (by using the PAIREDSTDDEVS= option)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the PAIREDMEANS statement.

## Paired $t$ Test

You can express effects in terms of the mean difference and variability in terms of a correlation and common standard deviation, as in the following statements. Default values for the DIST=, SIDES=, NULLDIFF=, and ALPHA= options specify a two-sided test for no difference with a normal distribution and a significance level of 0.05 .

```
proc power;
    pairedmeans test=diff
        meandiff = 7
        corr = 0.4
        stddev = 12
        npairs = 50
        power = .;
run;
```

You can also express effects in terms of individual means and variability in terms of correlation and individual standard deviations:

```
proc power;
    pairedmeans test=diff
        pairedmeans = 8 | 15
        corr = 0.4
        pairedstddevs = (7 12)
        npairs = .
        power = 0.9;
run;
```


## Paired t Test of Mean Ratio with Lognormal Data

You can express variability in terms of correlation and a common coefficient of variation, as in the following statements. Defaults for the DIST=, SIDES=, NULLRATIO= and ALPHA= options specify a two-sided test of mean ratio $=1$ assuming a lognormal distribution and a significance level of 0.05 .

```
proc power;
    pairedmeans test=ratio
        meanratio = 7
        corr = 0.3
        cv = 1.2
        npairs = 30
        power = .;
run;
```

You can also express variability in terms of correlation and individual coefficients of variation:

```
proc power;
    pairedmeans test=ratio
            meanratio = 7
            corr = 0.3
            pairedcvs = 0.8 | 0.9
            npairs = 30
            power = .;
run;
```


## Additive Equivalence Test for Mean Difference with Normal Data

The following statements demonstrate a sample size computation for a TOST equivalence test for a normal mean difference. Default values for the DIST= and ALPHA= options specify a normal distribution and a significance level of 0.05 .

```
proc power;
    pairedmeans test=equiv_diff
        lower = 2
        upper = 5
        meandiff = 4
        corr = 0.2
        stddev = 8
        npairs = .
        power = 0.9;
run;
```


## Multiplicative Equivalence Test for Mean Ratio with Lognormal Data

The following statements demonstrate a power computation for a TOST equivalence test for a lognormal mean ratio. Default values for the DIST= and ALPHA= options specify a lognormal distribution and a significance level of 0.05 .

```
proc power;
    pairedmeans test=equiv_ratio
        lower = 3
        upper = 7
        meanratio = 5
        corr = 0.2
        cv = 1.1
        npairs = 50
        power = .;
run;
```


## Confidence Interval for Mean Difference

By default CI=DIFF analyzes the conditional probability of obtaining the desired precision, given that the interval contains the true mean difference, as in the following statements. The defaults of SIDES=2 and ALPHA $=0.05$ specify a two-sided interval with a confidence level of 0.95 .

```
proc power;
    pairedmeans ci = diff
        halfwidth = 4
        corr = 0.35
        stddev = 8
        npairs = 30
        probwidth = .;
run;
```


## PLOT Statement

```
PLOT < plot-options> < / graph-options> ;
```

The PLOT statement produces a graph or set of graphs for the sample size analysis defined by the previous analysis statement. The plot-options define the plot characteristics, and the graph-options are SAS/GRAPHstyle options. If ODS Graphics is enabled, then the PLOT statement uses ODS Graphics to create graphs. For example:

```
ods graphics on;
proc power;
    onesamplemeans
        mean = 5 10
        ntotal = 150
        stddev = 30 50
        power = .;
    plot x=n min=100 max=200;
run;
ods graphics off;
```

Otherwise, traditional graphics are produced. For example:

```
ods graphics off;
proc power;
    onesamplemeans
        mean = 5 10
        ntotal = 150
        stddev = 30 50
        power = .;
    plot x=n min=100 max=200;
run;
```

For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 606 in Chapter 21, "Statistical Graphics Using ODS."

## Options

You can specify the following plot-options in the PLOT statement.

## INTERPOL=JOIN | NONE

specifies the type of curve to draw through the computed points. The INTERPOL=JOIN option connects computed points by straight lines. The INTERPOL=NONE option leaves computed points unconnected.

## KEY=BYCURVE < ( bycurve-options ) >

KEY=BYFEATURE < ( byfeature-options ) >
KEY=ONCURVES
specifies the style of key (or "legend") for the plot. The default is KEY=BYFEATURE, which specifies a key with a column of entries for each plot feature (line style, color, and/or symbol). Each entry shows the mapping between a value of the feature and the value(s) of the analysis parameter(s) linked to that feature. The KEY=BYCURVE option specifies a key with each row identifying a distinct curve in the plot. The KEY=ONCURVES option places a curve-specific label adjacent to each curve.
You can specify the following byfeature-options in parentheses after the KEY=BYCURVE option.

## NUMBERS=OFF | ON

specifies how the key should identify curves. If NUMBERS=OFF, then the key includes symbol, color, and line style samples to identify the curves. If NUMBERS $=O N$, then the key includes numbers matching numeric labels placed adjacent to the curves. The default is NUMBERS=ON.

## POS=BOTTOM | INSET

specifies the position of the key. The POS=BOTTOM option places the key below the X axis. The POS=INSET option places the key inside the plotting region and attempts to choose the least crowded corner. The default is POS=BOTTOM.

You can specify the following byfeature-options in parentheses after KEY=BYFEATURE option.

## POS=BOTTOM | INSET

specifies the position of the key. The POS=BOTTOM option places the key below the X axis. The POS=INSET option places the key inside the plotting region and attempts to choose the least crowded corner. The default is POS=BOTTOM.

## MARKERS=ANALYSIS | COMPUTED | NICE | NONE

specifies the locations for plotting symbols.
The MARKERS=ANALYSIS option places plotting symbols at locations corresponding to the values of the relevant input parameter from the analysis statement preceding the PLOT statement.
The MARKERS=COMPUTED option (the default) places plotting symbols at the locations of actual computed points from the sample size analysis.

The MARKERS=NICE option places plotting symbols at tick mark locations (corresponding to the argument axis).

The MARKERS=NONE option disables plotting symbols.

## MAX=number | DATAMAX

specifies the maximum of the range of values for the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). The default is DATAMAX, which specifies the maximum value that occurs for this parameter in the analysis statement that precedes the PLOT statement.

## MIN=number | DATAMIN

specifies the minimum of the range of values for the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). The default is DATAMIN, which specifies the minimum value that occurs for this parameter in the analysis statement that precedes the PLOT statement.

## NPOINTS=number

## NPTS=number

specifies the number of values for the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). You cannot use the NPOINTS= and STEP= options simultaneously. The default value for typical situations is 20.

## $\mathrm{STEP}=$ number

specifies the increment between values of the parameter associated with the "argument" axis (the axis that is not representing the parameter being solved for). You cannot use the STEP= and NPOINTS= options simultaneously. By default, the NPOINTS= option is used instead of the STEP= option.

VARY ( feature < BY parameter-list> < , ..., feature < BY parameter-list>> )
specifies how plot features should be linked to varying analysis parameters. Available plot features are COLOR, LINESTYLE, PANEL, and SYMBOL. A "panel" refers to a separate plot with a heading identifying the subset of values represented in the plot.

The parameter-list is a list of one or more names separated by spaces. Each name must match the name of an analysis option used in the analysis statement preceding the PLOT statement. Also, the name must be the primary name for the analysis option-that is, the one listed first in the syntax description.

If you omit the < BY parameter-list > portion for a feature, then one or more multivalued parameters from the analysis will be automatically selected for you.

## X=EFFECT | N | POWER

specifies a plot with the requested type of parameter on the X axis and the parameter being solved for on the Y axis. When $\mathrm{X}=$ EFFECT, the parameter assigned to the X axis is the one most representative of "effect size." When $\mathrm{X}=\mathrm{N}$, the parameter assigned to the X axis is the sample size. When $\mathrm{X}=\mathrm{POWER}$, the parameter assigned to the X axis is the one most representative of "power" (either power itself or a similar probability, such as Prob(Width) for confidence interval analyses). You cannot use the $\mathrm{X}=$ and $\mathrm{Y}=$ options simultaneously. The default is $\mathrm{X}=\mathrm{POWER}$, unless the result parameter is power or $\operatorname{Prob}($ Width), in which case the default is $\mathrm{X}=\mathrm{N}$.

You can use the $\mathrm{X}=\mathrm{N}$ option only when a scalar sample size parameter is used as input in the analysis. For example, $\mathrm{X}=\mathrm{N}$ can be used with total sample size or sample size per group, or with two group sample sizes when one is being solved for.
Table 77.18 summarizes the parameters representing effect size in different analyses.

Table 77.18 Effect Size Parameters for Different Analyses

| Analysis Statement and Options | Effect Size Parameters |
| :--- | :--- |
| LOGISTIC | None |
| MULTREG | Partial correlation or $R^{2}$ |
| difference |  |
| ONECORR | Correlation |
| ONESAMPLEFREQ TEST | Proportion |
| ONESAMPLEFREQ CI | CI half-width |
| ONESAMPLEMEANS TEST=T, | Mean |
| ONESAMPLEMEANS TEST=EQUIV | CI half-width |
| ONESAMPLEMEANS CI=T | None |
| ONEWAYANOVA | Discordant proportion difference |
| PAIREDFREQ | or ratio |
| PAIREDMEANS TEST=DIFF, | Mean difference |
| PAIREDMEANS TEST=EQUIV_DIFF | Mene |
| PAIREDMEANS TEST=RATIO, | Mean ratio |
| PAIREDMEANS TEST=EQUIV_RATIO | CI half-width |
| PAIREDMEANS CI=DIFF | Proportion difference, odds ratio, |
| TWOSAMPLEFREQ | or relative risk |
| TWOSAMPLEMEANS TEST=DIFF, |  |
| TWOSAMPLEMEANS TEST=DIFF_SATT, | Mean difference ratio |
| TWOSAMPLEMEANS TEST=EQUIV_DIFF |  |
| TWOSAMPLEMEANS TEST=RATIO, | NPAM used, else none |
| TWAMPLEMEANS TEST=EQUIV_RATIO | MPLEMEANS CI=DIFF |

## XOPTS=( $x$-options )

specifies plot characteristics pertaining to the X axis.
You can specify the following $x$-options in parentheses.

## CROSSREF=NO | YES

specifies whether the reference lines defined by the REF= $x$-option should be crossed with a reference line on the Y axis that indicates the solution point on the curve.

## REF=number-list

specifies locations for reference lines extending from the X axis across the entire plotting region. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Y=EFFECT | N | POWER

specifies a plot with the requested type of parameter on the Y axis and the parameter being solved for on the X axis. When $\mathrm{Y}=$ EFFECT, the parameter assigned to the Y axis is the one most representative of "effect size." When $\mathrm{Y}=\mathrm{N}$, the parameter assigned to the Y axis is the sample size. When $\mathrm{Y}=\mathrm{POWER}$, the parameter assigned to the Y axis is the one most representative of "power" (either power itself or a similar probability, such as Prob(Width) for confidence interval analyses). You cannot use the $\mathrm{Y}=$ and $\mathrm{X}=$ options simultaneously. By default, the $\mathrm{X}=$ option is used instead of the $\mathrm{Y}=$ option.

## YOPTS=( $y$-options )

specifies plot characteristics pertaining to the Y axis.
You can specify the following $y$-options in parentheses.

## CROSSREF=NO | YES

specifies whether the reference lines defined by the REF=y-option should be crossed with a reference line on the X axis that indicates the solution point on the curve.

## REF=number-list

specifies locations for reference lines extending from the Y axis across the entire plotting region. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

You can specify the following graph-options in the PLOT statement after a slash (/).

## DESCRIPTION='string '

specifies a descriptive string of up to 40 characters that appears in the "Description" field of the graphics catalog. The description does not appear on the plots. By default, PROC POWER assigns a description either of the form " $Y$ versus $X$ " (for a single-panel plot) or of the form " $Y$ versus $X(S)$," where $Y$ is the parameter on the Y axis, $X$ is the parameter on the X axis, and $S$ is a description of the subset represented on the current panel of a multipanel plot.

## NAME='string '

specifies a name of up to eight characters for the catalog entry for the plot. The default name is PLOTn, where $n$ is the number of the plot statement within the current invocation of PROC POWER. If the name duplicates the name of an existing entry, SAS/GRAPH software adds a number to the duplicate name to create a unique entry-for example, PLOT11 and PLOT12 for the second and third panels of a multipanel plot generated in the first PLOT statement in an invocation of PROC POWER.

## TWOSAMPLEFREQ Statement

## TWOSAMPLEFREQ < options > ;

The TWOSAMPLEFREQ statement performs power and sample size analyses for tests of two independent proportions. The Farrington-Manning score, Pearson's chi-square, Fisher's exact, and likelihood ratio chi-square tests are supported.

## Summary of Options

Table 77.19 summarizes the options available in the TWOSAMPLEFREQ statement.

Table 77.19 TWOSAMPLEFREQ Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| NULLODDSRATIO= | Specifies the null odds ratio |
| NULLPROPORTIONDIFF= | Specifies the null proportion difference |
| NULLRELATIVERISK= | Specifies the null relative risk |
| SIDES= | Specifies the number of sides and the direction of the statistical test or confidence interval |
| Specify effects |  |
| GROUPPROPORTIONS= | Specifies the two independent proportions, $p_{1}$ and $p_{2}$ |
| ODDSRATIO= | Specifies the odds ratio [ $\left.p_{2} /\left(1-p_{2}\right)\right] /\left[p_{1} /\left(1-p_{1}\right)\right]$ |
| PROPORTIONDIFF= | Specifies the proportion difference $p_{2}-p_{1}$ |
| REFPROPORTION= | Specifies the reference proportion $p_{1}$ |
| RELATIVERISK= | Specifies the relative risk $p_{2} / p_{1}$ |
| Specify sample size and allocation |  |
| GROUPNS= | Specifies the two group sample sizes |
| GROUPWEIGHTS= | Specifies the sample size allocation weights for the two groups |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NPERGROUP= | Specifies the common sample size per group |
| NTOTAL= | Specifies the sample size |
| Specify power |  |
| POWER= | Specifies the desired power of the test |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.20 summarizes the valid result parameters for different analyses in the TWOSAMPLEFREQ statement.

| Table 77.20 | $\begin{array}{l}\text { Summary of Result Parameters in the } \\ \text { TWOSAMPLEFREQ Statement }\end{array}$ |
| :---: | :--- |


| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=FISHER | Power | POWER $=$. |
|  | Sample size | NTOTAL=. |
| TEST=FM | Power | NPERGROUP=. |
|  | Sample size $=$. |  |
|  |  | NTOTAL=. |
|  | NPERGROUP=. |  |

Table 77.20 continued

| Analyses | Solve For | Syntax |
| :---: | :---: | :---: |
| TEST=LRCHI | Power | POWER=. |
|  | Sample size | NTOTAL=. |
| TEST $=$ PCHI | Power | POWER=. |
|  | Sample size | NTOTAL=. <br> NPERGROUP= |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPPROPORTIONS=grouped-number-list

GPROPORTIONS=grouped-number-list
GROUPPS=grouped-number-list
GPS=grouped-number-list
specifies the two independent proportions, $p_{1}$ and $p_{2}$. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPNS=grouped-number-list

GNS=grouped-number-list
specifies the two group sample sizes. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPWEIGHTS=grouped-number-list

## GWEIGHTS=grouped-number-list

specifies the sample size allocation weights for the two groups. This option controls how the total sample size is divided between the two groups. Each pair of values for the two groups represents relative allocation weights. Additionally, if the NFRACTIONAL option is not used, the total sample size is restricted to be equal to a multiple of the sum of the two group weights (so that the resulting design has an integer sample size for each group while adhering exactly to the group allocation weights). Values must be integers unless the NFRACTIONAL option is used. The default value is (1 1), a balanced design with a weight of 1 for each group. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NPERGROUP=number-list

## NPERG=number-list

specifies the common sample size per group or requests a solution for the common sample size per group with a missing value (NPERGROUP=.). Use of this option implicitly specifies a balanced design. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLODDSRATIO=number-list

## NULLOR=number-list

specifies the null odds ratio. You can specify this option only if you also specify the ODDSRATIO= and TEST=PCHI options. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. By default, NULLOR=1.

## NULLPROPORTIONDIFF=number-list

NULLPDIFF=number-list
specifies the null proportion difference. You can specify this option only if you also specify the GROUPPROPORTIONS $=$ or PROPORTIONDIFF= option and the TEST=FM or TEST=PCHI option. If you are using a non-default null value, then TEST=FM is recommended. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. By default, NULLPDIFF=0.

## NULLRELATIVERISK=number-list

## NULLRR=number-list

 specifies the null relative risk. You can specify this option only if you also specify the RELATIVERISK = and TEST=PCHI options. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. By default, NULLRR=1.
## ODDSRATIO=number-list

OR=number-list
specifies the odds ratio $\left[p_{2} /\left(1-p_{2}\right)\right] /\left[p_{1} /\left(1-p_{1}\right)\right]$. For information about specifying the numberlist, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES=
- NULLPROPORTIONDIFF=
- NULLODDSRATIO=
- NULLRELATIVERISK=
- ALPHA=
- GROUPPROPORTIONS=
- REFPROPORTION=
- PROPORTIONDIFF=
- ODDSRATIO=
- RELATIVERISK=
- GROUPWEIGHTS=
- NTOTAL=
- NPERGROUP=
- GROUPNS=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the TWOSAMPLEFREQ statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the TWOSAMPLEFREQ statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PROPORTIONDIFF=number-list

## PDIFF=number-list

specifies the proportion difference $p_{2}-p_{1}$. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## REFPROPORTION=number-list

## REFP=number-list

specifies the reference proportion $p_{1}$. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## RELATIVERISK=number-list

RR=number-list
specifies the relative risk $p_{2} / p_{1}$. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test or confidence interval. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords and their interpretation are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value
If the effect size is zero, then SIDES=1 is not permitted; instead, specify the direction of the test explicitly in this case with either SIDES=L or SIDES=U. The default value is 2 .

## TEST=FISHER | FM | LRCHI | PCHI

specifies the statistical analysis. You can specify the following values:
FISHER specifies Fisher's exact test.
FM specifies the score test of Farrington and Manning (1990).
LRCHI specifies the likelihood ratio chi-square test.
PCHI specifies Pearson's chi-square test.
If you are using a non-default null value for a noninferiority or superiority test, then TEST=FM is the most appropriate choice. For information about the power and sample size computational methods and formulas, see the section "Analyses in the TWOSAMPLEFREQ Statement" on page 6413. By default, TEST=PCHI.

## Restrictions on Option Combinations

To specify the proportions, choose one of the following parameterizations:

- individual proportions (by using the GROUPPROPORTIONS= option)
- difference between proportions and reference proportion (by using the PROPORTIONDIFF= and REFPROPORTION= options)
- odds ratio and reference proportion (by using the ODDSRATIO= and REFPROPORTION=options)
- relative risk and reference proportion (by using the RELATIVERISK= and REFPROPORTION= options)

To specify the sample size and allocation, choose one of the following parameterizations:

- sample size per group in a balanced design (by using the NPERGROUP=option)
- total sample size and allocation weights (by using the NTOTAL= and GROUPWEIGHTS= options)
- individual group sample sizes (by using the GROUPNS= option)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the TWOSAMPLEFREQ statement.

## Pearson Chi-Square Test for Two Proportions

You can use the NPERGROUP= option in a balanced design and express effects in terms of the individual proportions, as in the following statements. Default values for the SIDES $=$ and ALPHA= options specify a two-sided test with a significance level of 0.05 .

```
proc power;
    twosamplefreq test=pchi
        groupproportions = (.15 .25)
        nullproportiondiff = . 03
        npergroup = 50
        power = .;
run;
```

You can also specify an unbalanced design by using the NTOTAL= and GROUPWEIGHTS= options and express effects in terms of the odds ratio. The default value of the NULLODDSRATIO= option specifies a test of no effect.

```
proc power;
    twosamplefreq test=pchi
            oddsratio = 2.5
            refproportion = 0.3
            groupweights = (1 2)
            ntotal = .
            power = 0.8;
run;
```

You can also specify sample sizes with the GROUPNS $=$ option and express effects in terms of relative risks. The default value of the NULLRELATIVERISK= option specifies a test of no effect.

```
proc power;
    twosamplefreq test=pchi
            relativerisk = 1.5
            refproportion = 0.2
            groupns = 40 | 60
            power = .;
run;
```

You can also express effects in terms of the proportion difference. The default value of the NULLPROPORTIONDIFF = option specifies a test of no effect, and the default value of the GROUPWEIGHTS= option specifies a balanced design.

```
proc power;
    twosamplefreq test=pchi
            proportiondiff = 0.15
            refproportion = 0.4
            ntotal = 100
            power = .;
run;
```


## Farrington-Manning Score Test for Two Proportions

The following statements demonstrate a sample size computation for the Farrington-Manning score test for two proportions:

```
proc power;
    twosamplefreq test=fm
        proportiondiff = 0.06
        refproportion = 0.32
        nullproportiondiff = -0.02
```

```
    sides = u
    ntotal = .
    power = 0.85;
run;
```


## Fisher's Exact Conditional Test for Two Proportions

The following statements demonstrate a power computation for Fisher's exact conditional test for two proportions. Default values for the SIDES = and ALPHA= options specify a two-sided test with a significance level of 0.05 .

```
proc power;
    twosamplefreq test=fisher
        groupproportions = (.35 . 15)
        npergroup = 50
        power = .;
run;
```


## Likelihood Ratio Chi-Square Test for Two Proportions

The following statements demonstrate a sample size computation for the likelihood ratio chi-square test for two proportions. Default values for the SIDES= and ALPHA= options specify a two-sided test with a significance level of 0.05 .

```
proc power;
    twosamplefreq test=lrchi
        oddsratio = 2
        refproportion = 0.4
        npergroup = .
        power = 0.9;
run;
```


## TWOSAMPLEMEANS Statement

## TWOSAMPLEMEANS < options > ;

The TWOSAMPLEMEANS statement performs power and sample size analyses for pooled and unpooled $t$ tests, equivalence tests, and confidence interval precision involving two independent samples.

## Summary of Options

Table 77.21 summarizes the options available in the TWOSAMPLEMEANS statement.

Table 77.21 TWOSAMPLEMEANS Statement Options

| Option | Description |
| :--- | :--- |
| Define analysis |  |
| CI $=$ | Specifies an analysis of precision of the confidence interval |
| DIST $=$ | Specifies the underlying distribution assumed for the test statistic |
| TEST $=$ | Specifies the statistical analysis |
| Specify analysis information  <br> ALPHA $=$ Specifies the significance level |  |

Table 77.21 continued

| Option | Description |
| :---: | :---: |
| LOWER= | Specifies the lower equivalence bound |
| NULLDIFF= | Specifies the null mean difference |
| NULLRATIO= | Specifies the null mean ratio |
| SIDES= | Specifies the number of sides and the direction of the statistical test or confidence interval |
| UPPER= | Specifies the upper equivalence bound |
| Specify effects |  |
| HALFWIDTH= | Specifies the desired confidence interval half-width |
| GROUPMEANS= | Specifies the two group means |
| MEANDIFF= | Specifies the mean difference |
| MEANRATIO= | Specifies the geometric mean ratio, $\gamma_{2} / \gamma_{1}$ |
| Specify variability |  |
| CV= | Specifies the common coefficient of variation |
| GROUPSTDDEVS= | Specifies the standard deviation of each group |
| STDDEV= | Specifies the common standard deviation |
| Specify sample size and allocation |  |
| GROUPNS= | Specifies the two group sample sizes |
| GROUPWEIGHTS= | Specifies the sample size allocation weights for the two groups |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NPERGROUP= | Specifies the common sample size per group |
| NTOTAL= | Specifies the sample size |
| Specify power and related probabilities |  |
| POWER= | Specifies the desired power of the test |
| PROBTYPE= | Specifies the type of probability for the PROBWIDTH= option |
| PROBWIDTH= | Specifies the desired probability of obtaining a confidence interval halfwidth less than or equal to the value specified |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.22 summarizes the valid result parameters for different analyses in the TWOSAMPLEMEANS statement.

Table 77.22 Summary of Result Parameters in the TWOSAMPLEMEANS Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=DIFF | Power | POWER $=$. |
|  | Sample size | NTOTAL $=$. |
|  | Group sample size | NPERGROUP $=$. |
|  |  | GROUPNS $=n 11$. |
|  |  | GROUPNS $=. \mid n 2$ |
|  |  | GROUPNS $=(n 1)$. |
|  |  | GROUPNS $=(. n 2)$ |

Table 77.22 continued

| Analyses | Solve For | Syntax |
| :---: | :---: | :---: |
|  | Group weight | GROUPWEIGHTS $=w 11$. |
|  |  | GROUPWEIGHTS=. $\mid ~ w 2$ |
|  |  | GROUPWEIGHTS $=(w 1$. |
|  |  | GROUPWEIGHTS $=(. w 2)$ |
|  | Alpha | ALPHA=. |
|  | Group mean | GROUPMEANS $=$ mean 11 . |
|  |  | GROUPMEANS $=1$ I mean2 |
|  |  | GROUPMEANS $=($ mean 1.$)$ |
|  |  | GROUPMEANS $=($. mean 2 ) |
|  | Mean difference | MEANDIFF=. |
|  | Standard deviation | STDDEV=. |
| TEST=DIFF_SATT | Power | POWER=. |
|  | Sample size | NTOTAL=. |
|  |  | NPERGROUP= |
| TEST=RATIO | Power | POWER=. |
|  | Sample size | NTOTAL=. |
|  |  | NPERGROUP=. |
| TEST=EQUIV_DIFF | Power | POWER=. |
|  | Sample size | NTOTAL=. |
|  |  | NPERGROUP=. |
| TEST=EQUIV_RATIO |  | POWER=. |
|  | Sample size | NTOTAL=. |
|  |  | NPERGROUP=. |
| $\mathrm{CI}=$ DIFF | Prob(width) | PROBWIDTH=. |
|  | Sample size | NTOTAL=. |
|  |  | NPERGROUP=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test or requests a solution for alpha with a missing value (ALPHA=.). The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. If the $\mathrm{CI}=$ and $\mathrm{SIDES}=1$ options are used, then the value must be less than 0.5 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Cl

## Cl=DIFF

specifies an analysis of precision of the confidence interval for the mean difference, assuming equal variances. Instead of power, the relevant probability for this analysis is the probability that the interval
half-width is at most the value specified by the HALFWIDTH= option. If neither the TEST= option nor the $\mathrm{CI}=$ option is used, the default is TEST=DIFF.

## CV=number-list

specifies the coefficient of variation assumed to be common to both groups. The coefficient of variation is defined as the ratio of the standard deviation to the mean on the original data scale. You can use this option only with DIST=LOGNORMAL. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## DIST=LOGNORMAL | NORMAL

specifies the underlying distribution assumed for the test statistic. NORMAL corresponds the normal distribution, and LOGNORMAL corresponds to the lognormal distribution. The default value (also the only acceptable value in each case) is NORMAL for TEST=DIFF, TEST=DIFF_SATT, TEST=EQUIV_DIFF, and CI=DIFF; and LOGNORMAL for TEST=RATIO and TEST=EQUIV_RATIO.

## GROUPMEANS=grouped-number-list

GMEANS=grouped-number-list
specifies the two group means or requests a solution for one group mean given the other. Means are in the original scale. They are arithmetic if DIST=NORMAL and geometric if DIST=LOGNORMAL. This option cannot be used with the CI=DIFF analysis. When TEST=EQUIV_DIFF, the means are interpreted as the reference mean (first) and the treatment mean (second). For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPNS=grouped-number-list

## GNS=grouped-number-list

specifies the two group sample sizes or requests a solution for one group sample size given the other. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPSTDDEVS=grouped-number-list

GSTDDEVS=grouped-number-list

## GROUPSTDS=grouped-number-list

## GSTDS=grouped-number-list

specifies the standard deviation of each group. Unlike the STDDEV= option, the GROUPSTDDEVS $==$ option supports different values for each group. It is valid only for the Satterthwaite $t$ test (TEST=DIFF_SATT DIST=NORMAL). For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPWEIGHTS=grouped-number-list

## GWEIGHTS=grouped-number-list

specifies the sample size allocation weights for the two groups, or requests a solution for one group weight given the other. This option controls how the total sample size is divided between the two groups. Each pair of values for the two groups represents relative allocation weights. Additionally, if the NFRACTIONAL option is not used, the total sample size is restricted to be equal to a multiple of the sum of the two group weights (so that the resulting design has an integer sample size for each group while adhering exactly to the group allocation weights). Values must be integers unless the NFRACTIONAL option is used. The default value is (11), a balanced design with a weight of 1 for each group. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## HALFWIDTH=number-list

specifies the desired confidence interval half-width. The half-width is defined as the distance between the point estimate and a finite endpoint. This option can be used only with the CI=DIFF analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## LOWER=number-list

specifies the lower equivalence bound for the mean difference or mean ratio, in the original scale (whether DIST=NORMAL or DIST=LOGNORMAL). Values must be greater than 0 when DIST=LOGNORMAL. This option can be used only with the TEST=EQUIV_DIFF and TEST=EQUIV_RATIO analyses. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## MEANDIFF=number-list

specifies the mean difference, defined as $\mu_{2}-\mu_{1}$, or requests a solution for the mean difference with a missing value (MEANDIFF=.). This option can be used only with the TEST=DIFF, TEST=DIFF_SATT, and TEST=EQUIV_DIFF analyses. When TEST=EQUIV_DIFF, the mean difference is interpreted as the treatment mean minus the reference mean. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366

## MEANRATIO=number-list

specifies the geometric mean ratio, defined as $\gamma_{2} / \gamma_{1}$. This option can be used only with the TEST=RATIO and TEST=EQUIV_RATIO analyses. When TEST=EQUIV_RATIO, the mean ratio is interpreted as the treatment mean divided by the reference mean. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NPERGROUP=number-list

## NPERG=number-list

specifies the common sample size per group or requests a solution for the common sample size per group with a missing value (NPERGROUP=.). Use of this option implicitly specifies a balanced design. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLDIFF=number-list

## NULLD=number-list

specifies the null mean difference. The default value is 0 . This option can be used only with the TEST=DIFF and TEST=DIFF_SATT analyses. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NULLRATIO=number-list

NULLR=number-list
specifies the null mean ratio. The default value is 1 . This option can be used only with the TEST=RATIO analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES $=$
- NULLDIFF=
- NULLRATIO=
- LOWER=
- UPPER=
- $\mathrm{ALPHA}=$
- GROUPMEANS=
- MEANDIFF=
- MEANRATIO=
- HALFWIDTH=
- STDDEV=
- GROUPSTDDEVS==
- $\mathrm{CV}=$
- GROUPWEIGHTS=
- NTOTAL=
- NPERGROUP=
- GROUPNS=
- POWER=
- PROBTYPE=
- PROBWIDTH=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the TWOSAMPLEMEANS statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the TWOSAMPLEMEANS statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. This option cannot be used with the CI=DIFF analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## PROBTYPE=keyword-list

specifies the type of probability for the PROBWIDTH= option. A value of CONDITIONAL (the default) indicates the conditional probability that the confidence interval half-width is at most the value specified by the HALFWIDTH= option, given that the true mean difference is captured by the confidence interval. A value of UNCONDITIONAL indicates the unconditional probability that the confidence interval half-width is at most the value specified by the HALFWIDTH= option. you can
use the alias GIVENVALIDITY for CONDITIONAL. The PROBTYPE= option can be used only with the CI=DIFF analysis. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

CONDITIONAL width probability conditional on interval containing the mean
UNCONDITIONAL unconditional width probability

## PROBWIDTH=number-list

specifies the desired probability of obtaining a confidence interval half-width less than or equal to the value specified by the HALFWIDTH= option. A missing value (PROBWIDTH=.) requests a solution for this probability. The type of probability is controlled with the PROBTYPE= option. Values are expressed as probabilities (for example, 0.9 ) rather than percentages. This option can be used only with the CI=DIFF analysis. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test or confidence interval. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords and their interpretation for the TEST= analyses are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value
For confidence intervals, SIDES=U refers to an interval between the lower confidence limit and infinity, and SIDES $=$ L refers to an interval between minus infinity and the upper confidence limit. For both of these cases and SIDES=1, the confidence interval computations are equivalent. The SIDES= option cannot be used with the TEST=EQUIV_DIFF and TEST=EQUIV_RATIO analyses. The default value is 2 .

## STDDEV=number-list

## STD=number-list

specifies the standard deviation assumed to be common to both groups, or requests a solution for the common standard deviation with a missing value (STDDEV=.). This option can be used only with DIST=NORMAL. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## TEST=DIFF | DIFF_SATT | EQUIV_DIFF | EQUIV_RATIO | RATIO

## TEST

specifies the statistical analysis. TEST or TEST=DIFF (the default) specifies a pooled $t$ test on the mean difference, assuming equal variances. TEST=DIFF_SATT specifies a Satterthwaite unpooled $t$ test on the mean difference, assuming unequal variances. TEST=EQUIV_DIFF specifies an additive equivalence test of the mean difference by using a two one-sided tests (TOST) analysis (Schuirmann 1987). TEST=EQUIV_RATIO specifies a multiplicative equivalence test of the mean ratio by using a TOST analysis. TEST=RATIO specifies a pooled $t$ test on the mean ratio, assuming equal coefficients of variation. If neither the TEST= option nor the CI= option is used, the default is TEST=DIFF.

## UPPER=number-list

specifies the upper equivalence bound for the mean difference or mean ratio, in the original scale (whether DIST=NORMAL or DIST=LOGNORMAL). This option can be used only with the TEST=EQUIV_DIFF and TEST=EQUIV_RATIO analyses. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Restrictions on Option Combinations

To define the analysis, choose one of the following parameterizations:

- a statistical test (by using the TEST= option)
- confidence interval precision (by using the $\mathrm{CI}=$ option)

To specify the means, choose one of the following parameterizations:

- individual group means (by using the GROUPMEANS= option)
- mean difference (by using the MEANDIFF= option)
- mean ratio (by using the MEANRATIO= option)

To specify standard deviations in the Satterthwaite $t$ test (TEST=DIFF_SATT), choose one of the following parameterizations:

- common standard deviation (by using the STDDEV= option)
- individual group standard deviations (by using the GROUPSTDDEVS== option)

To specify the sample sizes and allocation, choose one of the following parameterizations:

- sample size per group in a balanced design (by using the NPERGROUP=option)
- total sample size and allocation weights (by using the NTOTAL= and GROUPWEIGHTS= options)
- individual group sample sizes (by using the GROUPNS= option)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the TWOSAMPLEMEANS statement.

## Two-Sample $t$ Test Assuming Equal Variances

You can use the NPERGROUP= option in a balanced design and express effects in terms of the mean difference, as in the following statements. Default values for the DIST=, SIDES=, NULLDIFF=, and ALPHA= options specify a two-sided test for no difference with a normal distribution and a significance level of 0.05 .

```
proc power;
    twosamplemeans test=diff
            meandiff = 7
            stddev = 12
            npergroup = 50
            power = .;
run;
```

You can also specify an unbalanced design by using the NTOTAL= and GROUPWEIGHTS= options and express effects in terms of individual group means:

```
proc power;
    twosamplemeans test=diff
        groupmeans = 8 | 15
        stddev = 4
        groupweights = (2 3)
        ntotal = .
        power = 0.9;
run;
```

Another way to specify the sample sizes is with the GROUPNS= option:

```
proc power;
    twosamplemeans test=diff
            groupmeans = 8 | 15
            stddev = 4
            groupns = (25 40)
            power = .;
run;
```


## Two-Sample Satterthwaite $t$ Test Assuming Unequal Variances

The following statements demonstrate a power computation for the two-sample Satterthwaite $t$ test allowing unequal variances. Default values for the $\mathrm{DIST}=, \mathrm{SIDES}=$, $\mathrm{NULLDIFF}=$, and $\mathrm{ALPHA}=$ options specify a two-sided test for no difference with a normal distribution and a significance level of 0.05.

```
proc power;
    twosamplemeans test=diff_satt
        meandiff = 3
        groupstddevs = 5 | 8
        groupweights = (1 2)
        ntotal = 60
        power = .;
run;
```


## Two-Sample Pooled $t$ Test of Mean Ratio with Lognormal Data

The following statements demonstrate a power computation for the pooled $t$ test of a lognormal mean ratio. Default values for the DIST $=$, SIDES $=$, NULLRATIO $=$, and ALPHA= options specify a two-sided test of mean ratio $=1$ assuming a lognormal distribution and a significance level of 0.05.

```
proc power;
    twosamplemeans test=ratio
        meanratio = 7
```

```
    cv = 0.8
    groupns = 50 | 70
    power = .;
run;
```


## Additive Equivalence Test for Mean Difference with Normal Data

The following statements demonstrate a sample size computation for the TOST equivalence test for a normal mean difference. A default value of GROUPWEIGHTS=(11) specifies a balanced design. Default values for the DIST= and ALPHA= options specify a significance level of 0.05 and an assumption of normally distributed data.

```
proc power;
    twosamplemeans test=equiv_diff
        lower = 2
        upper = 5
        meandiff = 4
        stddev = 8
        ntotal = .
        power = 0.9;
run;
```


## Multiplicative Equivalence Test for Mean Ratio with Lognormal Data

The following statements demonstrate a power computation for the TOST equivalence test for a lognormal mean ratio. Default values for the DIST $=$ and $\mathrm{ALPHA}=$ options specify a significance level of 0.05 and an assumption of lognormally distributed data.

```
proc power;
    twosamplemeans test=equiv_ratio
        lower = 3
        upper = 7
        meanratio = 5
        cv = 0.75
        npergroup = 50
        power = .;
run;
```


## Confidence Interval for Mean Difference

By default CI=DIFF analyzes the conditional probability of obtaining the desired precision, given that the interval contains the true mean difference, as in the following statements. The defaults of SIDES=2 and ALPHA $=0.05$ specify a two-sided interval with a confidence level of 0.95 .

```
proc power;
    twosamplemeans ci = diff
        halfwidth = 4
        stddev = 8
        groupns = (30 35)
        probwidth = .;
run;
```


## TWOSAMPLESURVIVAL Statement

TWOSAMPLESURVIVAL < options > ;
The TWOSAMPLESURVIVAL statement performs power and sample size analyses for comparing two survival curves. The log-rank, Gehan, and Tarone-Ware rank tests are supported.

## Summary of Options

Table 77.23 summarizes the options available in the TWOSAMPLESURVIVAL statement.

Table 77.23 TWOSAMPLESURVIVAL Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ACCRUALTIME= | Specifies the accrual time |
| ALPHA= | Specifies the significance level |
| FOLLOWUPTIME= | Specifies the follow-up time |
| SIDES= | Specifies the number of sides and the direction of the statistical test or confidence interval |
| TOTALTIME= | Specifies the total time |
| Specify effects |  |
| CURVE= | Defines a survival curve |
| GROUPMEDSURVTIMES= | Specifies the median survival times in each group |
| GROUPSURVEXPHAZARDS= | Specifies exponential hazard rates of the survival curve for each group |
| GROUPSURVIVAL= | Specifies the survival curve for each group |
| HAZARDRATIO= | Specifies the hazard ratio |
| REFSURVEXPHAZARD= | Specifies the exponential hazard rate of the survival curve for the reference group |
| REFSURVIVAL= | Specifies the survival curve for the reference group |
| Specify loss information |  |
| GROUPLOSS= | Specifies the exponential loss survival curve for each group |
| GROUPLOSSEXPHAZARDS= | Specifies the exponential hazards of the loss in each group |
| GROUPMEDLOSSTIMES= | Specifies the median times of the loss in each group |
| Specify sample size and allocation |  |
| ACCRUALRATEPERGROUP= | Specifies the common accrual rate per group |
| ACCRUALRATETOTAL= | Specifies the total accrual rate |
| EVENTSTOTAL= | Specifies the expected total number of events |
| GROUPACCRUALRATES= | Specifies the accrual rate for each group |
| GROUPNS= | Specifies the two group sample sizes |
| GROUPWEIGHTS= | Specifies the sample size allocation weights for the two groups |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NPERGROUP= | Specifies the common sample size per group |
| NTOTAL= | Specifies the sample size |


| Table 77.23 |  |
| :--- | :--- | continued

Table 77.24 summarizes the valid result parameters for different analyses in the TWOSAMPLESURVIVAL statement.

| Analyses | Solve For | Syntax |
| :---: | :---: | :---: |
| TEST=GEHAN | Power <br> Sample size | POWER=. <br> NTOTAL=. <br> NPERGROUP=. <br> EVENTSTOTAL=. <br> ACCRUALRATETOTAL=. <br> ACCRUALRATEPERGROUP=. |
| TEST=LOGRANK | Power <br> Sample size | POWER=. <br> NTOTAL=. <br> NPERGROUP=. <br> EVENTSTOTAL=. <br> ACCRUALRATETOTAL=. <br> ACCRUALRATEPERGROUP=. |
| TEST=TARONEWARE | Power <br> Sample size | POWER=. <br> NTOTAL=. <br> NPERGROUP=. <br> EVENTSTOTAL=. <br> ACCRUALRATETOTAL=. <br> ACCRUALRATEPERGROUP=. |

## Dictionary of Options

ACCRUALRATEPERGROUP=number-list
ACCRUALRATEPERG=number-list
ARPERGROUP=number-list
ARPERG=number-list
specifies the common accrual rate per group or requests a solution for the common accrual rate per group with a missing value (ACCRUALRATEPERGROUP=.). The accrual rate per group is the
number of subjects in each group that enters the study per time unit during the accrual period. Use of this option implicitly specifies a balanced design. The NFRACTIONAL option is automatically enabled when the ACCRUALRATEPERGROUP= option is used. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## ACCRUALRATETOTAL=number-list

## ARTOTAL=number-list

specifies the total accrual rate or requests a solution for the accrual rate with a missing value (ACCRUALRATETOTAL=.). The total accrual rate is the total number of subjects that enter the study per time unit during the accrual period. The NFRACTIONAL option is automatically enabled when the ACCRUALRATETOTAL= option is used. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## ACCRUALTIME=number-list | MAX

## ACCTIME=number-list | MAX

ACCRUALT=number-list | MAX

## ACCT=number-list | MAX

specifies the accrual time. Accrual is assumed to occur uniformly from time 0 to the time specified by the ACCRUALTIME= option. If the GROUPSURVIVAL= or REFSURVIVAL= option is used, then the value of the total time (the sum of accrual and follow-up times) must be less than or equal to the largest time in each multipoint (piecewise linear) survival curve. If the ACCRUALRATEPERGROUP=, ACCRUALRATETOTAL=, or GROUPACCRUALRATES = option is used, then the accrual time must be greater than 0 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

ACCRUALTIME=MAX can be used when each scenario in the analysis contains at least one piecewise linear survival curve (in the GROUPSURVIVAL= or REFSURVIVAL= option). It causes the accrual time to be automatically set, separately for each scenario, to the maximum possible time supported by the piecewise linear survival curve(s) in that scenario. It is not compatible with the FOLLOWUPTIME=MAX option or the TOTALTIME= option.

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## CURVE("label")=points

defines a survival curve.
For the CURVE= option,
label identifies the curve in the output and with the GROUPLOSS=, GROUPSURVIVAL=, and REFSURVIVAL= options.
points specifies one or more (time, survival) pairs on the curve, where the survival value denotes the probability of surviving until at least the specified time.

A single-point curve is interpreted as exponential, and a multipoint curve is interpreted as piecewise linear. Points can be expressed in either of two forms:

- a series of time:survival pairs separated by spaces. For example:

$$
1: 0.92: 0.73: 0.6
$$

- a DOLIST of times enclosed in parentheses, followed by a colon (:), followed by a DOLIST of survival values enclosed in parentheses. For example:

```
(1 to 3 by 1):(0.9 0.7 0.6)
```

The DOLIST format is the same as in the DATA step.
Points can also be expressed as combinations of the two forms. For example:

```
1:0.9 2:0.8 (3 to 6 by 1):(0.7 0.65 0.6 0.55)
```

The points have the following restrictions:

- The time values must be nonnegative and strictly increasing.
- The survival values must be strictly decreasing.
- The survival value at a time of 0 must be equal to 1 .
- If there is only one point, then the time must be greater than 0 , and the survival value cannot be 0 or 1 .


## EVENTSTOTAL=number-list

EVENTTOTAL=number-list

## EETOTAL=number-list

specifies the expected total number of events-that is, deaths, whether observed or censored-during the entire study period, or requests a solution for this parameter with a missing value (EVENTSTOTAL=.). The NFRACTIONAL option is automatically enabled when the EVENTSTOTAL= option is used. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## FOLLOWUPTIME=number-list | MAX

FUTIME=number-list | MAX
FOLLOWUPT=number-list | MAX

## FUT=number-list | MAX

specifies the follow-up time, the amount of time in the study past the accrual time. If the GROUPSURVIVAL $=$ or REFSURVIVAL $=$ option is used, then the value of the total time (the sum of accrual and follow-up times) must be less than or equal to the largest time in each multipoint (piecewise linear) survival curve. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

FOLLOWUPTIME=MAX can be used when each scenario in the analysis contains at least one piecewise linear survival curve (in the GROUPSURVIVAL= or REFSURVIVAL= option). It causes the follow-up time to be automatically set, separately for each scenario, to the maximum possible time supported by the piecewise linear survival curve(s) in that scenario. It is not compatible with the ACCRUALTIME=MAX option or the TOTALTIME= option.

## GROUPACCRUALRATES=grouped-number-list

GACCRUALRATES=grouped-number-list
GROUPARS=grouped-number-list
GARS=grouped-number-list
specifies the accrual rate for each group. The groupwise accrual rates are the numbers of subjects in each group that enters the study per time unit during the accrual period. The NFRACTIONAL option is automatically enabled when the GROUPACCRUALRATES $=$ option is used. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366 .

## GROUPLOSS=grouped-name-list

GLOSS=grouped-name-list
specifies the exponential loss survival curve for each group, by using labels specified with the CURVE= option. Loss is assumed to follow an exponential curve, indicating the expected rate of censoring (in other words, loss to follow-up) over time. For information about specifying the grouped-name-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPLOSSEXPHAZARDS=grouped-number-list

## GLOSSEXPHAZARDS=grouped-number-list

GROUPLOSSEXPHS=grouped-number-list
GLOSSEXPHS=grouped-number-list
specifies the exponential hazards of the loss in each group. Loss is assumed to follow an exponential curve, indicating the expected rate of censoring (in other words, loss to follow-up) over time. If none of the GROUPLOSSEXPHAZARDS $=$, GROUPLOSS $=$, and GROUPMEDLOSSTIMES $=$ options are used, the default of GROUPLOSSEXPHAZARDS $=\left(\begin{array}{ll}0 & 0\end{array}\right)$ indicates no loss to follow-up. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPMEDLOSSTIMES=grouped-number-list

GMEDLOSSTIMES=grouped-number-list
GROUPMEDLOSSTS=grouped-number-list
GMEDLOSSTS=grouped-number-list
specifies the median times of the loss in each group. Loss is assumed to follow an exponential curve, indicating the expected rate of censoring (in other words, loss to follow-up) over time. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPMEDSURVTIMES=grouped-number-list

GMEDSURVTIMES=grouped-number-list
GROUPMEDSURVTS=grouped-number-list

## GMEDSURVTS=grouped-number-list

specifies the median survival times in each group. When the GROUPMEDSURVTIMES= option is used, the survival curve in each group is assumed to be exponential. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPNS=grouped-number-list

GNS=grouped-number-list
specifies the two group sample sizes. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPSURVEXPHAZARDS=grouped-number-list

GSURVEXPHAZARDS=grouped-number-list
GROUPSURVEXPHS=grouped-number-list

## GEXPHS=grouped-number-list

specifies exponential hazard rates of the survival curve for each group. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPSURVIVAL=grouped-name-list

## GSURVIVAL=grouped-name-list

GROUPSURV=grouped-name-list
GSURV=grouped-name-list
specifies the survival curve for each group, by using labels specified with the CURVE= option. For information about specifying the grouped-name-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPWEIGHTS=grouped-number-list

## GWEIGHTS=grouped-number-list

specifies the sample size allocation weights for the two groups. This option controls how the total sample size is divided between the two groups. Each pair of values for the two groups represents relative allocation weights. Additionally, if the NFRACTIONAL option is not used, the total sample size is restricted to be equal to a multiple of the sum of the two group weights (so that the resulting design has an integer sample size for each group while adhering exactly to the group allocation weights). Values must be integers unless the NFRACTIONAL option is used. The default value is (11), a balanced design with a weight of 1 for each group. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## HAZARDRATIO=number-list

## HR=number-list

specifies the hazard ratio of the second group's survival curve to the first group's survival curve. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. This option is automatically enabled when any of the following options are used: ACCRUALRATEPERGROUP=, ACCRUALRATETOTAL=, EVENTSTOTAL=, and GROUPACCRUALRATES=. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NPERGROUP=number-list

## NPERG=number-list

specifies the common sample size per group or requests a solution for the common sample size per group with a missing value (NPERGROUP=.). Use of this option implicitly specifies a balanced design. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NSUBINTERVAL=number-list

## NSUBINTERVALS=number-list

## NSUB=number-list

NSUBS=number-list
specifies the number of subintervals per unit time to use in internal calculations. Higher values increase computational time and memory requirements but generally lead to more accurate results. The default value is 12 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES=
- ACCRUALTIME=
- FOLLOWUPTIME=
- TOTALTIME=
- NSUBINTERVAL=
- $\mathrm{ALPHA}=$
- REFSURVIVAL=
- GROUPSURVIVAL=
- REFSURVEXPHAZARD=
- HAZARDRATIO=
- GROUPSURVEXPHAZARDS=
- GROUPMEDSURVTIMES=
- GROUPLOSSEXPHAZARDS=
- GROUPLOSS=
- GROUPMEDLOSSTIMES=
- GROUPWEIGHTS=
- NTOTAL=
- ACCRUALRATETOTAL=
- EVENTSTOTAL=
- NPERGROUP=
- ACCRUALRATEPERGROUP=
- GROUPNS=
- GROUPACCRUALRATES=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the TWOSAMPLESURVIVAL statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the TWOSAMPLESURVIVAL statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## REFSURVEXPHAZARD=number-list

## REFSURVEXPH=number-list

specifies the exponential hazard rate of the survival curve for the first (reference) group. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## REFSURVIVAL=name-list

REFSURV=name-list
specifies the survival curve for the first (reference) group, by using labels specified with the CURVE= option. For information about specifying the name-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test or confidence interval. For information about specifying the keyword-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366. Valid keywords and their interpretation are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with the alternative hypothesis favoring better survival in the second group
L lower one-sided with the alternative hypothesis favoring better survival in the first (reference) group

The default value is 2 .

## TEST=GEHAN | LOGRANK | TARONEWARE

specifies the statistical analysis. TEST=GEHAN specifies the Gehan rank test. TEST=LOGRANK (the default) specifies the log-rank test. TEST=TARONEWARE specifies the Tarone-Ware rank test.

## TOTALTIME=number-list | MAX

TOTALT=number-list | MAX
specifies the total time, which is equal to the sum of accrual and follow-up times. If the GROUPSURVIVAL = or REFSURVIVAL= option is used, then the value of the total time must be less than or equal to the largest time in each multipoint (piecewise linear) survival curve. For information
about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

TOTALTIME=MAX can be used when each scenario in the analysis contains at least one piecewise linear survival curve (in the GROUPSURVIVAL= or REFSURVIVAL= option). It causes the total time to be automatically set, separately for each scenario, to the maximum possible time supported by the piecewise linear survival curve(s) in that scenario. It is not compatible with the ACCRUALTIME=MAX option or the FOLLOWUPTIME=MAX option.

## Restrictions on Option Combinations

To specify the survival curves, choose one of the following parameterizations:

- arbitrary piecewise linear or exponential curves (by using the CURVE= and GROUPSURVIVAL= options)
- curves with proportional hazards (by using the CURVE=, REFSURVIVAL=, and HAZARDRATIO= options)
- exponential curves, by using one of the following parameterizations:
- median survival times (by using the GROUPMEDSURVTIMES= option)
- the hazard ratio and the hazard of the reference curve (by using the HAZARDRATIO= and REFSURVEXPHAZARD= options)
- the individual hazards (by using the GROUPSURVEXPHAZARDS= option)

To specify the study time, use any two of the following three options:

- accrual time (by using the ACCRUALTIME= option)
- follow-up time (by using the FOLLOWUPTIME= option)
- total time, the sum of accrual and follow-up times (by using the TOTALTIME= option)

To specify the sample size and allocation, choose one of the following parameterizations:

- sample size per group in a balanced design (by using the NPERGROUP=option)
- accrual rate per group in a balanced design (by using the ACCRUALRATEPERGROUP= option)
- total sample size and allocation weights (by using the NTOTAL= and GROUPWEIGHTS= options)
- total accrual rate and allocation weights (by using the ACCRUALRATETOTAL= and GROUPWEIGHTS = options)
- expected total number of events and allocation weights (by using the EVENTSTOTAL= and GROUPWEIGHTS= options)
- individual group sample sizes (by using the GROUPNS= option)
- individual group accrual rates (by using the GROUPACCRUALRATES= option)

The values of parameters that involve expected number of events or accrual rate are converted internally to the analogous sample size parameterization (that is, the NPERGROUP=, NTOTAL=, or GROUPNS= option) for the purpose of sample size adjustments according to the presence or absence of the NFRACTIONAL option.

To specify the exponential loss curves, choose one of the following parameterizations:

- a point on the loss curve of each group (by using the CURVE= and GROUPLOSS= options)
- median loss times (by using the GROUPMEDLOSSTIMES= option)
- the individual loss hazards (by using the GROUPLOSSEXPHAZARDS= option)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the TWOSAMPLESURVIVAL statement.

## Log-Rank Test for Two Survival Curves

You can use the NPERGROUP= option in a balanced design and specify piecewise linear or exponential survival curves by using the CURVE= and GROUPSURVIVAL= options, as in the following statements. Default values for the SIDES=, ALPHA=, NSUBINTERVAL=, and GROUPLOSSEXPHAZARDS= options specify a two-sided test with a significance level of 0.05 , an assumption of no loss to follow-up, and the use of 12 subintervals per unit time in computations.

```
proc power;
    twosamplesurvival test=logrank
        curve("Control") = (1 2 3):(0.8 0.7 0.6)
        curve("Treatment") = (5):(.6)
        groupsurvival = "Control" | "Treatment"
        accrualtime = 2
        followuptime = 1
        npergroup = 50
        power = .;
run;
```

In the preceding example, the "Control" curve is piecewise linear (since it has more than one point), and the "Treatment" curve is exponential (since it has only one point).

You can also specify an unbalanced design by using the NTOTAL= and GROUPWEIGHTS= options and specify piecewise linear or exponential survival curves with proportional hazards by using the CURVE=, REFSURVIVAL=, and HAZARDRATIO= options:

```
proc power;
    twosamplesurvival test=logrank
        curve("Control") = (1 2 3):(0.8 0.7 0.6)
        refsurvival = "Control"
        hazardratio = 1.5
        accrualtime = 2
        followuptime = 1
        groupweights = (1 2)
        ntotal = .
```

```
    power = 0.8;
run;
```

Instead of computing sample size, you can compute the accrual rate by using the ACCRUALRATETOTAL= option:

```
proc power;
    twosamplesurvival test=logrank
        curve("Control") = (1 2 3):(0.8 0.7 0.6)
        refsurvival = "Control"
        hazardratio = 1.5
        accrualtime = 2
        followuptime = 1
        groupweights = (1 2)
        accrualratetotal = .
        power = 0.8;
run;
```

or the expected number of events by using the EVENTSTOTAL= option:

```
proc power;
    twosamplesurvival test=logrank
        curve("Control") = (1 2 3):(0.8 0.7 0.6)
        refsurvival = "Control"
        hazardratio = 1.5
        accrualtime = 2
        followuptime = 1
        groupweights = (1 2)
        eventstotal = .
        power = 0.8;
run;
```

You can also specify sample sizes with the GROUPNS= option and specify exponential survival curves in terms of median survival times:

```
proc power;
    twosamplesurvival test=logrank
        groupmedsurvtimes = (16 22)
        accrualtime = 6
        totaltime = 18
        groupns = 40 | 60
        power = .;
run;
```

You can also specify exponential survival curves in terms of the hazard ratio and reference hazard. The default value of the GROUPWEIGHTS = option specifies a balanced design.

```
proc power;
    twosamplesurvival test=logrank
        hazardratio = 1.2
        refsurvexphazard = 0.7
        accrualtime = 2
        totaltime = 4
        ntotal = 100
        power = .;
run;
```

You can also specify exponential survival curves in terms of the individual hazards, as in the following statements:

```
proc power;
    twosamplesurvival test=logrank
        groupsurvexphazards = 0.7 | 0.84
        accrualtime = 2
        totaltime = 4
        ntotal = .
        power = 0.9;
run;
```


## Gehan Rank Test for Two Survival Curves

In addition to the log-rank test, you can also specify the Gehan tank test, as in the following statements. Default values for the SIDES=, ALPHA=, NSUBINTERVAL=, and GROUPLOSSEXPHAZARDS= options specify a two-sided test with a significance level of 0.05 , an assumption of no loss to follow-up, and the use of 12 subintervals per unit time in computations.

```
proc power;
    twosamplesurvival test=gehan
        groupmedsurvtimes = 5 | 7
        accrualtime = 3
        totaltime = 6
        npergroup = .
        power = 0.8;
run;
```


## Tarone-Ware Rank Test for Two Survival Curves

You can also specify the Tarone-Ware tank test, as in the following statements. Default values for the SIDES=, ALPHA=, NSUBINTERVAL=, and GROUPLOSSEXPHAZARDS = options specify a two-sided test with a significance level of 0.05 , an assumption of no loss to follow-up, and the use of 12 subintervals per unit time in computations.

```
proc power;
    twosamplesurvival test=taroneware
        groupmedsurvtimes = 5 | 7
        accrualtime = 3
        totaltime = 6
        npergroup = 100
        power = .;
run;
```


## TWOSAMPLEWILCOXON Statement

TWOSAMPLEWILCOXON < options > ;
The TWOSAMPLEWILCOXON statement performs power and sample size analyses for the Wilcoxon-Mann-Whitney test (also called the Wilcoxon rank-sum test, Mann-Whitney-Wilcoxon test, or Mann-Whitney U test) for two independent groups.

Note that the O'Brien-Castelloe approach to computing power for the Wilcoxon test is approximate, based on asymptotic behavior as the total sample size gets large. The quality of the power approximation degrades for small sample sizes; conversely, the quality of the sample size approximation degrades if the two distributions are far apart, so that only a small sample is needed to detect a significant difference. But this degradation is rarely a problem in practical situations, in which experiments are usually performed for relatively close distributions.

## Summary of Options

Table 77.25 summarizes the options available in the TWOSAMPLEWILCOXON statement.

Table 77.25 TWOSAMPLEWILCOXON Statement Options

| Option | Description |
| :---: | :---: |
| Define analysis |  |
| TEST= | Specifies the statistical analysis |
| Specify analysis information |  |
| ALPHA= | Specifies the significance level |
| SIDES= | Specifies the number of sides and the direction of the statistical test |
| Specify distributions |  |
| VARDIST= | Defines a distribution for a variable |
| VARIABLES= | Specifies the distributions of two or more variables |
| Specify sample size and allocation |  |
| GROUPNS= | Specifies the two group sample sizes |
| GROUPWEIGHTS= | Specifies the sample size allocation weights for the two groups |
| NFRACTIONAL | Enables fractional input and output for sample sizes |
| NPERGROUP= | Specifies the common sample size per group |
| NTOTAL= | Specifies the sample size |
| Specify power |  |
| POWER= | Specifies the desired power of the test |
| Specify computational options |  |
| NBINS= | Specifies the number of categories for each variable |
| Control ordering in output |  |
| OUTPUTORDER= | Controls the output order of parameters |

Table 77.26 summarizes the valid result parameters in the TWOSAMPLEWILCOXON statement.

Table 77.26 Summary of Result Parameters in the TWOSAMPLEWILCOXON Statement

| Analyses | Solve For | Syntax |
| :--- | :--- | :--- |
| TEST=WMW | Power | POWER=. |
|  | Sample size | NTOTAL=. |
|  |  | NPERGROUP=. |

## Dictionary of Options

## ALPHA=number-list

specifies the level of significance of the statistical test. The default is 0.05 , corresponding to the usual $0.05 \times 100 \%=5 \%$ level of significance. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPNS=grouped-number-list

## GNS=grouped-number-list

specifies the two group sample sizes. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## GROUPWEIGHTS=grouped-number-list

## GWEIGHTS=grouped-number-list

specifies the sample size allocation weights for the two groups. This option controls how the total sample size is divided between the two groups. Each pair of values for the two groups represents relative allocation weights. Additionally, if the NFRACTIONAL option is not used, the total sample size is restricted to be equal to a multiple of the sum of the two group weights (so that the resulting design has an integer sample size for each group while adhering exactly to the group allocation weights). Values must be integers unless the NFRACTIONAL option is used. The default value is (11), a balanced design with a weight of 1 for each group. For information about specifying the grouped-number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NBINS=number-list

specifies the number of categories (or "bins") each variable's distribution is divided into (unless it is ordinal, in which case the categories remain intact) in internal calculations. Higher values increase computational time and memory requirements but generally lead to more accurate results. However, if the value is too high, then numerical instability can occur. Lower values are less likely to produce "No solution computed" errors. The default value is 1000 . For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NFRACTIONAL

## NFRAC

enables fractional input and output for sample sizes. See the section "Sample Size Adjustment Options" on page 6369 for information about the ramifications of the presence (and absence) of the NFRACTIONAL option.

## NPERGROUP=number-list

## NPERG=number-list

specifies the common sample size per group or requests a solution for the common sample size per group with a missing value (NPERGROUP=.). Use of this option implicitly specifies a balanced design. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## NTOTAL=number-list

specifies the sample size or requests a solution for the sample size with a missing value (NTOTAL=.). For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## OUTPUTORDER=INTERNAL | REVERSE | SYNTAX

controls how the input and default analysis parameters are ordered in the output. OUTPUTORDER=INTERNAL (the default) arranges the parameters in the output according to the following order of their corresponding options:

- SIDES
- NBINS=
- $\mathrm{ALPHA}=$
- VARIABLES=
- GROUPWEIGHTS=
- NTOTAL=
- NPERGROUP=
- GROUPNS=
- POWER=

The OUTPUTORDER=SYNTAX option arranges the parameters in the output in the same order in which their corresponding options are specified in the TWOSAMPLEWILCOXON statement. The OUTPUTORDER=REVERSE option arranges the parameters in the output in the reverse of the order in which their corresponding options are specified in the TWOSAMPLEWILCOXON statement.

## POWER=number-list

specifies the desired power of the test or requests a solution for the power with a missing value (POWER=.). The power is expressed as a probability, a number between 0 and 1 , rather than as a percentage. For information about specifying the number-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## SIDES=keyword-list

specifies the number of sides (or tails) and the direction of the statistical test. Valid keywords are as follows:

1 one-sided with alternative hypothesis in same direction as effect
2 two-sided
U upper one-sided with alternative greater than null value
L lower one-sided with alternative less than null value

The default value is 2 .

## TEST=WMW

specifies the Wilcoxon-Mann-Whitney test for two independent groups This is the default test option.

## VARDIST("label")=distribution (parameters)

defines a distribution for a variable.
For the VARDIST= option,
label identifies the variable distribution in the output and with the VARIABLES= option.
distribution specifies the distributional form of the variable.
parameters specifies one or more parameters associated with the distribution.

The distributions and parameters are named and defined in the same way as the distributions and arguments in the CDF SAS function; for more information, see SAS Language Reference: Dictionary. Choices for distributional forms and their parameters are as follows:

ORDINAL ((values) : (probabilities)) is an ordered categorical distribution. The values are any numbers separated by spaces. The probabilities are numbers between 0 and 1 (inclusive) separated by spaces. Their sum must be exactly 1 . The number of probabilities must match the number of values.

BETA $(a, b<, I, r>)$ is a beta distribution with shape parameters $a$ and $b$ and optional location parameters $I$ and $r$. The values of $a$ and $b$ must be greater than 0 , and $/$ must be less than $r$. The default values for $l$ and $r$ are 0 and 1 , respectively.
BINOMIAL ( $p, n$ ) is a binomial distribution with probability of success $p$ and number of independent Bernoulli trials $n$. The value of $p$ must be greater than 0 and less than 1 , and $n$ must be an integer greater than 0 . If $n=1$, then the distribution is binary.
$\operatorname{EXPONENTIAL}(\lambda)$ is an exponential distribution with scale $\lambda$, which must be greater than 0 .
$\operatorname{GAMMA}(a, \lambda)$ is a gamma distribution with shape $a$ and scale $\lambda$. The values of $a$ and $\lambda$ must be greater than 0 .
$\operatorname{LAPLACE}(\theta, \lambda) \quad$ is a Laplace distribution with location $\theta$ and scale $\lambda$. The value of $\lambda$ must be greater than 0 .
LOGISTIC $(\theta, \lambda)$ is a logistic distribution with location $\theta$ and scale $\lambda$. The value of $\lambda$ must be greater than 0 .

LOGNORMAL $\theta, \lambda)$ is a lognormal distribution with location $\theta$ and scale $\lambda$. The value of $\lambda$ must be greater than 0 .
$\operatorname{NORMAL}(\theta, \lambda) \quad$ is a normal distribution with mean $\theta$ and standard deviation $\lambda$. The value of $\lambda$ must be greater than 0 .

POISSON $(m) \quad$ is a Poisson distribution with mean $m$. The value of $m$ must be greater than 0 .
UNIFORM $(I, r)$ is a uniform distribution on the interval $[I, r]$, where $I<r$.

## VARIABLES=grouped-name-list

VARS=grouped-name-list
specifies the distributions of two or more variables, using labels specified with the VARDIST= option. For information about specifying the grouped-name-list, see the section "Specifying Value Lists in Analysis Statements" on page 6366.

## Restrictions on Option Combinations

To specify the sample size and allocation, choose one of the following parameterizations:

- sample size per group in a balanced design (using the NPERGROUP=option)
- total sample size and allocation weights (using the NTOTAL= and GROUPWEIGHTS= options)
- individual group sample sizes (using the GROUPNS= option)


## Option Groups for Common Analyses

This section summarizes the syntax for the common analyses supported in the TWOSAMPLEWILCOXON statement.

## Wilcoxon-Mann-Whitney Test for Comparing Two Distributions

The following statements performs a power analysis for Wilcoxon-Mann-Whitney tests comparing an ordinal variable with each other type of distribution. Default values for the ALPHA=, NBINS=, SIDES=, and TEST= options specify a two-sided test with a significance level of 0.05 and the use of 1000 categories per distribution when discretization is needed.

```
proc power;
    twosamplewilcoxon
        vardist("myordinal") = ordinal ((0 1 2) : (.2 . 3 . 5) )
        vardist("mybeta1") = beta (1, 2)
        vardist("mybeta2") = beta (1, 2, 0, 2)
        vardist("mybinomial") = binomial (.3, 3)
        vardist("myexponential") = exponential (2)
        vardist("mygamma") = gamma (1.5, 2)
        vardist("mylaplace") = laplace (1, 2)
        vardist("mylogistic") = logistic (1, 2)
        vardist("mylognormal") = lognormal (1, 2)
        vardist("mynormal") = normal (3, 2)
        vardist("mypoisson") = poisson (2)
        vardist("myuniform") = uniform (0, 2)
        variables = "myordinal" | "mybeta1" "mybeta2" "mybinomial"
                                    "myexponential" "mygamma" "mylaplace"
                                    "mylogistic" "mylognormal" "mynormal"
                                "mypoisson" "myuniform"
        ntotal = 40
        power = .;
run;
```


## Details: POWER Procedure

## Overview of Power Concepts

In statistical hypothesis testing, you typically express the belief that some effect exists in a population by specifying an alternative hypothesis $H_{1}$. You state a null hypothesis $H_{0}$ as the assertion that the effect does not exist and attempt to gather evidence to reject $H_{0}$ in favor of $H_{1}$. Evidence is gathered in the form of sample data, and a statistical test is used to assess $H_{0}$. If $H_{0}$ is rejected but there really is no effect, this is called a Type I error. The probability of a Type I error is usually designated "alpha" or $\alpha$, and statistical tests are designed to ensure that $\alpha$ is suitably small (for example, less than 0.05 ).

If there really is an effect in the population but $H_{0}$ is not rejected in the statistical test, then a Type II error has been made. The probability of a Type II error is usually designated "beta" or $\beta$. The probability $1-\beta$ of avoiding a Type II error-that is, correctly rejecting $H_{0}$ and achieving statistical significance-is called the power. (NOTE: Another more general definition of power is the probability of rejecting $H_{0}$ for any given set
of circumstances, even those corresponding to $H_{0}$ being true. The POWER procedure uses this more general definition.)

An important goal in study planning is to ensure an acceptably high level of power. Sample size plays a prominent role in power computations because the focus is often on determining a sufficient sample size to achieve a certain power, or assessing the power for a range of different sample sizes.

Some of the analyses in the POWER procedure focus on precision rather than power. An analysis of confidence interval precision is analogous to a traditional power analysis, with "CI Half-Width" taking the place of effect size and "Prob(Width)" taking the place of power. The CI Half-Width is the margin of error associated with the confidence interval, the distance between the point estimate and an endpoint. The $\operatorname{Prob}($ Width ) is the probability of obtaining a confidence interval with at most a target half-width.

## Summary of Analyses

Table 77.27 gives a summary of the analyses supported in the POWER procedure. The name of the analysis statement reflects the type of data and design. The TEST $=, \mathrm{CI}=$, and DIST= options specify the focus of the statistical hypothesis (in other words, the criterion on which the research question is based) and the test statistic to be used in data analysis.

Table 77.27 Summary of Analyses

| Analysis | Statement | Options |
| :--- | :--- | :--- |
| Logistic regression: likelihood ratio <br> chi-square test | LOGISTIC |  |
| Multiple linear regression: Type III $F$ test | MULTREG |  |
| Correlation: Fisher's $z$ test | ONECORR | DIST=FISHERZ |
| Correlation: $t$ test | ONECORR | DIST=T |
| Binomial proportion: exact test | ONESAMPLEFREQ | TEST=EXACT |
| Binomial proportion: $z$ test | ONESAMPLEFREQ | TEST=Z |
| Binomial proportion: $z$ test with continuity <br> adjustment <br> Binomial proportion: exact equivalence test | ONESAMPLEFREQ | ONESAMPLEFREQ |

Table 77.27 continued

| Analysis | Statement | Options |
| :--- | :--- | :--- |
| One-sample $t$ test with lognormal data | ONESAMPLEMEANS | TEST=T |
| One-sample equivalence test for mean of <br> normal data <br> One-sample equivalence test for mean of <br> lognormal data | ONESAMPLEMEANS | DIST=LOGNORMAL |
| Confidence interval for a mean | ONESAMPLEMEANS | TEST=EQUIV |
| One-way ANOVA: one-degree-of-freedom <br> contrast | ONEWAV |  |
| One-way ANOVA: overall $F$ test | ONEWT=LOGNORMAL |  |
| McNemar exact conditional test <br> McNemar normal approximation test | PAIREDFREQ | CI=T |
| Paired $t$ test | PAIREDMEANA | TEST=CONTRAST |
| Paired $t$ test of mean ratio with lognormal data <br> Paired additive equivalence of mean <br> difference with normal data | PAIREDMEANS | PAIREDMEANS |

Table 77.27 continued

| Analysis | Statement | Options |
| :--- | :--- | :--- |
| Gehan rank test for comparing two survival <br> curves | TWOSAMPLESURVIVAL | TEST=GEHAN |
| Tarone-Ware rank test for comparing two <br> survival curves | TWOSAMPLESURVIVAL | TEST=TARONEWARE |
| Wilcoxon-Mann-Whitney (rank-sum) test | TWOSAMPLEWILCOXON |  |

## Specifying Value Lists in Analysis Statements

To specify one or more scenarios for an analysis parameter (or set of parameters), you provide a list of values for the statement option that corresponds to the parameter(s). To identify the parameter you want to solve for, you place missing values in the appropriate list.
There are five basic types of such lists: keyword-lists, number-lists, grouped-number-lists, name-lists, and grouped-name-lists. Some parameters, such as the direction of a test, have values represented by one or more keywords in a keyword-list. Scenarios for scalar-valued parameters, such as power, are represented by a number-list. Scenarios for groups of scalar-valued parameters, such as group sample sizes in a multigroup design, are represented by a grouped-number-list. Scenarios for named parameters, such as reference survival curves, are represented by a name-list. Scenarios for groups of named parameters, such as group survival curves, are represented by a grouped-name-list.
The following subsections explain these five basic types of lists.

## Keyword-Lists

A keyword-list is a list of one or more keywords, separated by spaces. For example, you can specify both two-sided and upper-tailed versions of a one-sample $t$ test as follows:

```
SIDES = 2 U
```


## Number-Lists

A number-list can be one of two things: a series of one or more numbers expressed in the form of one or more DOLISTs, or a missing value indicator (.).
The DOLIST format is the same as in the DATA step language. For example, for the one-sample $t$ test you can specify four scenarios ( $30,50,70$, and 100) for a total sample size in any of the following ways.

```
NTOTAL = 30 50 70 100
NTOTAL = 30 to 70 by 20 100
```

A missing value identifies a parameter as the result parameter; it is valid only with options representing parameters you can solve for in a given analysis. For example, you can request a solution for NTOTAL as follows:

## Grouped-Number-Lists

A grouped-number-list specifies multiple scenarios for numeric values in two or more groups, possibly including missing value indicators to solve for a specific group. The list can assume one of two general forms, a "crossed" version and a "matched" version.

## Crossed Grouped-Number-Lists

The crossed version of a grouped number list consists of a series of number-lists (see the section "NumberLists" on page 6366), one representing each group, with groups separated by a vertical bar (I). The values for each group represent multiple scenarios for that group, and the scenarios for each individual group are crossed to produce the set of all scenarios for the analysis option. For example, you can specify the following six scenarios for the sizes ( $n_{1}, n_{2}$ ) of two groups

$$
(20,30)(20,40)(20,50)
$$

$$
(25,30)(25,40)(25,50)
$$

as follows:

```
GROUPNS = 20 25 | 30 40 50
```

If the analysis can solve for a value in one group given the other groups, then one of the number-lists in a crossed grouped-number-list can be a missing value indicator (.). For example, in a two-sample $t$ test you can posit three scenarios for the group 2 sample size while solving for the group 1 sample size:

```
GROUPNS = . | 30 40 50
```

Some analyses can involve more than two groups. For example, you can specify $2 \times 3 \times 1=6$ scenarios for the means of three groups in a one-way ANOVA as follows:

```
GROUPMEANS = 10 12 | 10 to 20 by 5 | 24
```


## Matched Grouped-Number-Lists

The matched version of a grouped number list consists of a series of numeric lists, each enclosed in parentheses. Each list consists of a value for each group and represents a single scenario for the analysis option. Multiple scenarios for the analysis option are represented by multiple lists. For example, you can express the crossed grouped-number-list

```
GROUPNS = 20 25 | 30 40 50
```

alternatively in a matched format:

```
GROUPNS = (20 30) (20 40) (20 50) (25 30) (25 40) (25 50)
```

The matched version is particularly useful when you want to include only a subset of all combinations of individual group values. For example, you might want to pair 20 only with 50 , and 25 only with 30 and 40 :

```
GROUPNS = (20 50) (25 30) (25 40)
```

If the analysis can solve for a value in one group given the other groups, then you can replace the value for that group with a missing value indicator (.). If used, the missing value indicator must occur in the same group in every scenario. For example, you can solve for the group 1 sample size (as in the section "Crossed Grouped-Number-Lists" on page 6367) by using a matched format:

```
GROUPNS = (. 30) (. 40) (. 50)
```

Some analyses can involve more than two groups. For example, you can specify two scenarios for the means of three groups in a one-way ANOVA:

```
GROUPMEANS =((15 24 32) (12 25 36)
```


## Name-Lists

A name-list is a list of one or more names that are enclosed in single or double quotation marks and separated by spaces. For example, you can specify two scenarios for the reference survival curve in a log-rank test as follows:

```
REFSURVIVAL = "Curve A" "Curve B"
```


## Grouped-Name-Lists

A grouped-name-list specifies multiple scenarios for names in two or more groups. The list can assume one of two general forms, a "crossed" version and a "matched" version.

## Crossed Grouped-Name-Lists

The crossed version of a grouped name list consists of a series of name-lists (see the section "Name-Lists" on page 6368), one representing each group, with groups separated by a vertical bar (I). The values for each group represent multiple scenarios for that group, and the scenarios for each individual group are crossed to produce the set of all scenarios for the analysis option. For example, you can specify the following six scenarios for the survival curves ( $c_{1}, c_{2}$ ) of two groups
("Curve A", "Curve C")("Curve A", "Curve D")("Curve A", "Curve E")
("Curve B", "Curve C")("Curve B", "Curve D")("Curve B", "Curve E")
as follows:

```
GROUPSURVIVAL = "Curve A" "Curve B" | "Curve C" "Curve D"
    "Curve E"
```


## Matched Grouped-Name-Lists

The matched version of a grouped name list consists of a series of name lists, each enclosed in parentheses. Each list consists of a name for each group and represents a single scenario for the analysis option. Multiple scenarios for the analysis option are represented by multiple lists. For example, you can express the crossed grouped-name-list

```
GROUPSURVIVAL = "Curve A" "Curve B" | "Curve C" "Curve D"
    "Curve E"
```

alternatively in a matched format:

```
GROUPSURVIVAL = ("Curve A" "Curve C")
    ("Curve A" "Curve D")
    ("Curve A" "Curve E")
    ("Curve B" "Curve C")
    ("Curve B" "Curve D")
    ("Curve B" "Curve E")
```

The matched version is particularly useful when you want to include only a subset of all combinations of individual group values. For example, you might want to pair "Curve A" only with "Curve C", and "Curve B" only with "Curve D" and "Curve E":

```
GROUPSURVIVAL = ("Curve A" "Curve C")
    ("Curve B" "Curve D")
    ("Curve B" "Curve E")
```


## Sample Size Adjustment Options

By default, PROC POWER rounds sample sizes conservatively (down in the input, up in the output) so that all total sizes (and individual group sample sizes, if a multigroup design) are integers. This is generally considered conservative because it selects the closest realistic design providing at most the power of the (possibly fractional) input or mathematically optimized design. In addition, in a multigroup design, all group sizes are adjusted to be multiples of the corresponding group weights. For example, if GROUPWEIGHTS = (26), then all group 1 sample sizes become multiples of 2 , and all group 2 sample sizes become multiples of 6 (and all total sample sizes become multiples of 8 ).

With the NFRACTIONAL option, sample size input is not rounded, and sample size output (whether total or groupwise) are reported in two versions, a raw "fractional" version and a "ceiling" version rounded up to the nearest integer.

Whenever an input sample size is adjusted, both the original ("nominal") and adjusted ("actual") sample sizes are reported. Whenever computed output sample sizes are adjusted, both the original input ("nominal") power and the achieved ("actual") power at the adjusted sample size are reported.

## Error and Information Output

The Error column in the main output table provides reasons for missing results and flags numerical results that are bounds rather than exact answers. For example, consider the sample size analysis implemented by the following statements:

```
proc power;
    twosamplefreq test=pchi
        method=normal
        oddsratio= 1.0001
        refproportion=.4
        nulloddsratio=1
        power=. }
        ntotal=.;
run;
```

Figure 77.6 Error Column
The POWER Procedure Pearson Chi-square Test for Two Proportions

| Fixed Scenario Elements |  |
| :--- | ---: |
| Distribution | Asymptotic normal |
| Method | Normal approximation |
| Null Odds Ratio | 1 |
| Reference (Group 1) Proportion | 0.4 |
| Odds Ratio | 1.0001 |
| Nominal Power | 0.9 |
| Number of Sides | 2 |
| Alpha | 0.05 |
| Group 1 Weight | 1 |
| Group 2 Weight | 1 |

Computed N Total
Actual
Power N Total Error
0.206 2.15E+09 Solution is a lower bound

The output in Figure 77.6 reveals that the sample size to achieve a power of 0.9 could not be computed, but that the sample size 2.15E+09 achieves a power of 0.206.

The Info column provides further details about Error column entries, warnings about any boundary conditions detected, and notes about any adjustments to input. Note that the Info column is hidden by default in the main output. You can view it by using the ODS OUTPUT statement to save the output as a data set and the PRINT procedure. For example, the following SAS statements print both the Error and Info columns for a power computation in a two-sample $t$ test:

```
proc power;
    twosamplemeans
        meandiff= 07
        stddev=2
        ntotal=2 5
        power=.;
    ods output output=Power;
run;
```

```
proc print noobs data=Power;
    var MeanDiff NominalNTotal NTotal Power Error Info;
run;
```

The output is shown in Figure 77.7.
Figure 77.7 Error and Info Columns

| MeanDiff NominalNTotal | NTotal | Power | Error | Info |
| :---: | :---: | :---: | :---: | :---: |
| 0 | 2 | 2 | . | Invalid input $N$ too small / No effect |
| 0 | 5 | 4 | 0.050 | Input N adjusted / No effect |
| 7 | 2 | 2 | . | Invalid input N too small |
| 7 | 5 | 4 | 0.477 | Input N adjusted |

The mean difference of 0 specified with the MEANDIFF= option leads to a "No effect" message to appear in the Info column. The sample size of 2 specified with the NTOTAL= option leads to an "Invalid input" message in the Error column and an "NTotal too small" message in the Info column. The sample size of 5 leads to an "Input N adjusted" message in the Info column because it is rounded down to 4 to produce integer group sizes of 2 per group.

## Displayed Output

If you use the PLOTONLY option in the PROC POWER statement, the procedure displays only graphical output. Otherwise, the displayed output of the POWER procedure includes the following:

- the "Fixed Scenario Elements" table, which shows all applicable single-valued analysis parameters, in the following order: distribution, method, parameters that are input explicitly, and parameters that are supplied with defaults
- an output table that shows the following when applicable (in order): the index of the scenario, all multivalued input, ancillary results, the primary computed result, and error descriptions
- plots (if requested)

For each input parameter, the order of the input values is preserved in the output.
Ancillary results include the following:

- Actual Power, the achieved power, if it differs from the input (Nominal) power value
- Actual Prob(Width), the achieved precision probability, if it differs from the input (Nominal) probability value
- Actual Alpha, the achieved significance level, if it differs from the input (Nominal) alpha value
- fractional sample size, if the NFRACTIONAL option is used in the analysis statement

If sample size is the result parameter and the NFRACTIONAL option is used in the analysis statement, then both "Fractional" and "Ceiling" sample size results are displayed. Fractional sample sizes correspond to the "Nominal" values of power or precision probability. Ceiling sample sizes are simply the fractional sample sizes rounded up to the nearest integer; they correspond to "Actual" values of power or precision probability.

## ODS Table Names

PROC POWER assigns a name to each table that it creates. You can use these names to reference the table when using the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in Table 77.28. For more information about ODS, see Chapter 20, "Using the Output Delivery System."

Table 77.28 ODS Tables Produced by PROC POWER

| ODS Table Name | Description | Statement |
| :--- | :--- | :--- |
| FixedElements | Factoid with single-valued analysis parameters | Default* |
| Output | All input and computed analysis parameters, error messages, and <br> information messages for each scenario | Default |
| PlotContent | Data contained in plots, including analysis parameters and indices <br> identifying plot features. (NOTE: This table is saved as a data set <br> and not displayed in PROC POWER output.) | PLOT |

*Depends on input.

## Computational Resources

## Memory

In the TWOSAMPLESURVIVAL statement, the amount of required memory is roughly proportional to the product of the number of subintervals (specified by the NSUBINTERVAL= option) and the total time of the study (specified by the ACCRUALTIME=, FOLLOWUPTIME=, and TOTALTIME= options). If you run out of memory, then you can try either specifying a smaller number of subintervals, changing the time scale to a use a longer time unit (for example, years instead of months), or both.

## CPU Time

In the Satterthwaite $t$ test analysis (TWOSAMPLEMEANS TEST=DIFF_SATT), the required CPU time grows as the mean difference decreases relative to the standard deviations. In the PAIREDFREQ statement, the required CPU time for the exact power computation (METHOD=EXACT) grows with the sample size.

## Computational Methods and Formulas

This section describes the approaches that PROC POWER uses to compute power for each analysis. The first subsection defines some common notation. The following subsections describe the various power analyses, including discussions of the data, statistical test, and power formula for each analysis. Unless otherwise indicated, computed values for parameters besides power (for example, sample size) are obtained by solving power formulas for the desired parameters.

## Common Notation

Table 77.29 displays notation for some of the more common parameters across analyses. The Associated Syntax column shows examples of relevant analysis statement options, where applicable.

Table 77.29 Common Notation

| Symbol | Description | Associated Syntax |
| :---: | :---: | :---: |
| $\alpha$ | Significance level | ALPHA= |
| $N$ | Total sample size | NTOTAL $=$, NPAIRS $=$ |
| $n_{i}$ | Sample size in ith group | NPERGROUP=, GROUPNS= |
| $w_{i}$ | Allocation weight for $i$ th group (standardized to sum to 1) | GROUPWEIGHTS= |
| $\mu$ | (Arithmetic) mean | MEAN= |
| $\mu_{i}$ | (Arithmetic) mean in $i$ th group | GROUPMEANS=, PAIREDMEANS= |
| $\mu_{\text {diff }}$ | (Arithmetic) mean difference, $\mu_{2}-\mu_{1}$ or $\mu_{T}-\mu_{R}$ | MEANDIFF= |
| $\mu_{0}$ | Null mean or mean difference (arithmetic) | NULL=, NULLDIFF= |
| $\gamma$ | Geometric mean | MEAN= |
| $\gamma_{i}$ | Geometric mean in ith group | GROUPMEANS=, PAIREDMEANS= |
| $\gamma_{0}$ | Null mean or mean ratio (geometric) | NULL=, NULLRATIO= |
| $\sigma$ | Standard deviation (or common standard deviation per group) | STDDEV= |
| $\sigma_{i}$ | Standard deviation in ith group | GROUPSTDDEVS=, PAIREDSTDDEVS= |
| $\sigma_{\text {diff }}$ | Standard deviation of differences |  |
| CV | Coefficient of variation, defined as the ratio of the standard deviation to the (arithmetic) mean on the original data scale | $\mathrm{CV}=$, PAIREDCVS $=$ |
| $\rho$ | Correlation | CORR= |

Table 77.29 continued

| Symbol | Description | Associated Syntax |
| :--- | :--- | :--- |
| $\mu_{T}, \mu_{R}$ | Treatment and reference (arithmetic) means for | GROUPMEANS $=$, |
|  | equivalence test | PAIREDMEANS $=$ |
| $\gamma_{T}, \gamma_{R}$ | Treatment and reference geometric means for | GROUPMEANS $=$, |
| $\theta_{L}$ | equivalence test | PAIREDMEANS $=$ |
| $\theta_{U}$ | Lower equivalence bound | LOWER= |
| $t(\nu, \delta)$ | Upper equivalence bound | UPPER $=$ |
| $F\left(\nu_{1}, \nu_{2}, \lambda\right)$ | $F$ distribution with $d f v$ and noncentrality $\delta$ |  |
| $t_{p ; \nu}$ | $\nu_{2}$, and noncentrality $\lambda$ <br> $p$ th percentile of $t$ distribution with $d f v$ <br> $F_{p ; \nu_{1}, \nu_{2}}$ <br> $p$ th percentile of $F$ distribution with numerator $d f \nu_{1}$ <br> and denominator $d f \nu_{2}$ |  |
| $\operatorname{Bin}(N, p)$ | Binomial distribution with sample size $N$ and <br> proportion $p$ |  |

A "lower one-sided" test is associated with SIDES=L (or SIDES=1 with the effect smaller than the null value), and an "upper one-sided" test is associated with SIDES=U (or SIDES=1 with the effect larger than the null value).

Owen (1965) defines a function, known as Owen's $Q$, that is convenient for representing terms in power formulas for confidence intervals and equivalence tests:

$$
Q_{\nu}(t, \delta ; a, b)=\frac{\sqrt{2 \pi}}{\Gamma\left(\frac{\nu}{2}\right) 2^{\frac{v-2}{2}}} \int_{a}^{b} \Phi\left(\frac{t x}{\sqrt{v}}-\delta\right) x^{\nu-1} \phi(x) \mathrm{d} x
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and cumulative distribution function of the standard normal distribution, respectively.

## Analyses in the LOGISTIC Statement

## Likelihood Ratio Chi-Square Test for One Predictor (TEST=LRCHI)

The power computing formula is based on Shieh and O'Brien (1998); Shieh (2000); Self, Mauritsen, and Ohara (1992), and Hsieh (1989).

Define the following notation for a logistic regression analysis:

$$
\begin{aligned}
& N=\# \text { subjects (NTOTAL) } \\
& K=\# \text { predictors (not counting intercept) } \\
& \mathrm{x}=\left(x_{1}, \ldots, x_{K}\right)^{\prime}=\text { random variables for predictor vector } \\
& \mathbf{x}_{-1}=\left(x_{2}, \ldots, x_{K}\right)^{\prime} \\
& \boldsymbol{\mu}=\left(\mu_{1}, \ldots, \mu_{K}\right)^{\prime}=\mathrm{Ex}=\text { mean predictor vector } \\
& \mathrm{x}_{i}=\left(x_{i 1}, \ldots, x_{i K}\right)^{\prime}=\text { predictor vector for subject } i \quad(i \in 1, \ldots, N) \\
& Y=\text { random variable for response ( } 0 \text { or 1) } \\
& Y_{i}=\text { response for subject } i \quad(i \in 1, \ldots, N) \\
& p_{i}=\operatorname{Prob}\left(Y_{i}=1 \mid \mathbf{x}_{i}\right) \quad(i \in 1, \ldots, N) \\
& \phi=\operatorname{Prob}\left(Y_{i}=1 \mid \mathbf{x}_{i}=\mu\right) \quad \text { (RESPONSEPROB) } \\
& U_{j}=\text { unit change for } j \text { th predictor (UNITS) } \\
& \mathrm{OR}_{j}=\operatorname{Odds}\left(Y_{i}=1 \mid x_{i j}=c\right) / \operatorname{Odds}\left(Y_{i}=1 \mid x_{i j}=c-U_{j}\right) \quad(c \text { arbitrary }, i \in 1, \ldots, N \text {, } \\
& j \in 1, \ldots, K \text { ) (TESTODDSRATIO if } j=1 \text {, COVODDSRATIOS if } j>1 \text { ) } \\
& \Psi_{0}=\text { intercept in full model (INTERCEPT) } \\
& \boldsymbol{\Psi}=\left(\Psi_{1}, \ldots, \Psi_{K}\right)^{\prime}=\text { regression coefficients in full model } \\
& \text { ( } \left.\Psi_{1}=\text { TESTREGCOEFF, others }=\text { COVREGCOEFFS }\right) \\
& \rho=\operatorname{Corr}\left(\mathrm{x}_{-1}, x_{1}\right) \quad \text { (CORR) } \\
& c_{j}=\# \text { distinct possible values of } x_{i j} \quad(j \in 1, \ldots, K)(f o r ~ a n y ~ i) \quad \text { (NBINS) } \\
& x_{g j}^{\star}=g \text { th possible value of } x_{i j} \quad\left(g \in 1, \ldots, c_{j}\right)(j \in 1, \ldots, K) \\
& \text { (for any } i \text { ) (VARDIST) } \\
& \pi_{g j}=\operatorname{Prob}\left(x_{i j}=x_{g j}^{\star}\right) \quad\left(g \in 1, \ldots, c_{j}\right)(j \in 1, \ldots, K) \\
& \text { (for any } i \text { ) (VARDIST) } \\
& C=\prod_{j=1}^{K} c_{j}=\# \text { possible values of } \mathbf{x}_{i} \quad(\text { for any } i) \\
& \mathbf{x}_{m}^{\star}=m \text { th possible value of } \mathbf{x}_{i} \quad(m \in 1, \ldots, C) \\
& \pi_{m}=\operatorname{Prob}\left(\mathrm{x}_{i}=\mathrm{x}_{m}^{\star}\right) \quad(m \in 1, \ldots, C)
\end{aligned}
$$

The logistic regression model is

$$
\log \left(\frac{p_{i}}{1-p_{i}}\right)=\Psi_{0}+\boldsymbol{\Psi}^{\prime} \mathbf{x}_{i}
$$

The hypothesis test of the first predictor variable is

$$
\begin{aligned}
& H_{0}: \Psi_{1}=0 \\
& H_{1}: \Psi_{1} \neq 0
\end{aligned}
$$

Assuming independence among all predictor variables, $\pi_{m}$ is defined as follows:

$$
\pi_{m}=\prod_{j=1}^{K} \pi_{h(m, j) j} \quad(m \in 1, \ldots, C)
$$

where $h(m, j)$ is calculated according to the following algorithm:

$$
\begin{aligned}
& z=m \\
& \text { do } \quad j=K \text { to } 1 \\
& \quad h(m, j)=\bmod \left(z-1, c_{j}\right)+1 \\
& \quad z=\text { floor }\left((z-1) / c_{j}\right)+1
\end{aligned}
$$

end;

This algorithm causes the elements of the transposed vector $\{h(m, 1), \ldots, h(m, K)\}$ to vary fastest to slowest from right to left as $m$ increases, as shown in the following table of $h(m, j)$ values:

|  |  | $j$ |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $h(m, j)$ | 1 | 2 | $\cdots$ | $K-1$ | $K$ |  |
| $m$ | 1 | 1 | 1 | $\cdots$ | 1 | 1 |
|  | 1 | 1 | 1 | $\cdots$ | 1 | 2 |
|  | $\vdots$ |  |  | $\vdots$ |  |  |
|  | $\vdots$ | 1 | 1 | $\cdots$ | 1 | $c_{K}$ |
|  | $\vdots$ | 1 | 1 | $\cdots$ | 2 | 1 |
|  | $\vdots$ | 1 | 1 | $\cdots$ | 2 | 2 |
|  | $\vdots$ |  |  | $\vdots$ |  |  |
|  | $\vdots$ | 1 | 1 | $\cdots$ | 2 | $c_{K}$ |
|  | $\vdots$ |  |  | $\vdots$ |  |  |
|  | $\vdots$ | $c_{1}$ | $c_{2}$ | $\cdots$ | $c_{K-1}$ | 1 |
|  | $\vdots$ | $c_{1}$ | $c_{2}$ | $\cdots$ | $c_{K-1}$ | 2 |
|  | $\vdots$ |  |  | $\vdots$ |  |  |
|  | $C$ | $c_{1}$ | $c_{2}$ | $\cdots$ | $c_{K-1}$ | $c_{K}$ |

The $\mathrm{x}_{m}^{\star}$ values are determined in a completely analogous manner.
The discretization is handled as follows (unless the distribution is ordinal, or binomial with sample size parameter at least as large as requested number of bins): for $x_{j}$, generate $c_{j}$ quantiles at evenly spaced probability values such that each such quantile is at the midpoint of a bin with probability $\frac{1}{c_{j}}$. In other words,

$$
\begin{aligned}
x_{g j}^{\star}= & \left(\frac{g-0.5}{c_{j}}\right) \text { th quantile of relevant distribution } \\
& \left(g \in 1, \ldots, c_{j}\right)(j \in 1, \ldots, K) \\
\pi_{g j}= & \frac{1}{c_{j}} \quad(\text { same for all } g)
\end{aligned}
$$

The primary noncentrality for the power computation is

$$
\Delta^{\star}=2 \sum_{m=1}^{C} \pi_{m}\left[b^{\prime}\left(\theta_{m}\right)\left(\theta_{m}-\theta_{m}^{\star}\right)-\left(b\left(\theta_{m}\right)-b\left(\theta_{m}^{\star}\right)\right)\right]
$$

where

$$
\begin{aligned}
b^{\prime}(\theta) & =\frac{\exp (\theta)}{1+\exp (\theta)} \\
b(\theta) & =\log (1+\exp (\theta)) \\
\theta_{m} & =\Psi_{0}+\boldsymbol{\Psi}^{\prime} \mathbf{x}_{m}^{\star} \\
\theta_{m}^{\star} & =\Psi_{0}^{\star}+\boldsymbol{\Psi}^{\star} \mathbf{x}_{m}^{\star}
\end{aligned}
$$

where

$$
\begin{aligned}
& \Psi_{0}^{\star}=\Psi_{0}+\Psi_{1} \mu_{1}=\text { intercept in reduced model, absorbing the tested predictor } \\
& \Psi^{\star}=\left(0, \Psi_{2}, \ldots, \Psi_{K}\right)^{\prime}=\text { coefficients in reduced model }
\end{aligned}
$$

The power is

$$
\text { power }=P\left(\chi^{2}\left(1, \Delta^{\star} N\left(1-\rho^{2}\right)\right) \geq \chi_{1-\alpha}^{2}(1)\right)
$$

The factor $\left(1-\rho^{2}\right)$ is the adjustment for correlation between the predictor that is being tested and other predictors, from Hsieh (1989).

Alternative input parameterizations are handled by the following transformations:

$$
\begin{aligned}
& \Psi_{0}=\log \left(\frac{\phi}{1-\phi}\right)-\boldsymbol{\Psi}^{\prime} \boldsymbol{\mu} \\
& \Psi_{j}=\frac{\log \left(\mathrm{OR}_{j}\right)}{U_{j}} \quad(j \in 1, \ldots, K)
\end{aligned}
$$

## Analyses in the MULTREG Statement

## Type III F Test in Multiple Regression (TEST=TYPE3)

Maxwell (2000) discusses a number of different ways to represent effect sizes (and to compute exact power based on them) in multiple regression. PROC POWER supports two of these, multiple partial correlation and $R^{2}$ in full and reduced models.

Let $p$ denote the total number of predictors in the full model (excluding the intercept), and let $Y$ denote the response variable. You are testing that the coefficients of $p_{1} \geq 1$ predictors in a set $X_{1}$ are 0 , controlling for all of the other predictors $X_{-1}$, which consists of $p-p_{1} \geq 0$ variables.
The hypotheses can be expressed in two different ways. The first is in terms of $\rho_{Y X_{1} \mid X_{-1}}$, the multiple partial correlation between the predictors in $X_{1}$ and the response $Y$ adjusting for the predictors in $X_{-1}$ :

$$
\begin{aligned}
& H_{0}: \rho_{Y X_{1} \mid X_{-1}}^{2}=0 \\
& H_{1}: \rho_{Y X_{1} \mid X_{-1}}^{2}>0
\end{aligned}
$$

The second is in terms of the multiple correlations in full $\left(\rho_{Y \mid\left(X_{1}, X_{-1}\right)}\right)$ and reduced $\left(\rho_{Y \mid X_{-1}}\right)$ nested models:

$$
\begin{aligned}
& H_{0}: \rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}-\rho_{Y \mid X_{-1}}^{2}=0 \\
& H_{1}: \rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}-\rho_{Y \mid X_{-1}}^{2}>0
\end{aligned}
$$

Note that the squared values of $\rho_{Y \mid\left(X_{1}, X_{-1}\right)}$ and $\rho_{Y \mid X_{-1}}$ are the population $R^{2}$ values for full and reduced models.

The test statistic can be written in terms of the sample multiple partial correlation $R_{Y X_{1} \mid X_{-1}}$,

$$
F= \begin{cases}(N-1-p) \frac{R_{Y X_{1} \mid X_{-1}}^{2}}{1-R_{Y X_{1} \mid X_{-1}}^{2}}, & \text { intercept } \\ (N-p) \frac{R_{Y X_{1} \mid X_{-1}}^{2}}{1-R_{Y X_{1} \mid X_{-1}}^{2}}, & \text { no intercept }\end{cases}
$$

or the sample multiple correlations in full $\left(R_{Y \mid\left(X_{1}, X_{-1}\right)}\right)$ and reduced $\left(R_{Y \mid X_{-1}}\right)$ models,

$$
F= \begin{cases}(N-1-p) \frac{R_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}-R_{Y \mid X_{-1}}^{2}}{1-R_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}}, & \text { intercept } \\ (N-p) \frac{R_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}-R_{Y \mid X_{-1}}^{2}}{1-R_{Y \mid\left(X_{1}, X_{-1}\right)}^{2},} & \text { no intercept }\end{cases}
$$

The test is the usual Type III $F$ test in multiple regression:

$$
\text { Reject } \quad H_{0} \quad \text { if } \begin{cases}F \geq F_{1-\alpha}\left(p_{1}, N-1-p\right), & \text { intercept } \\ F \geq F_{1-\alpha}\left(p_{1}, N-p\right), & \text { no intercept }\end{cases}
$$

Although the test is invariant to whether the predictors are assumed to be random or fixed, the power is affected by this assumption. If the response and predictors are assumed to have a joint multivariate normal distribution, then the exact power is given by the following formula:

$$
\begin{aligned}
\text { power } & =\left\{\begin{array}{ll}
P\left[\left(\frac{N-1-p}{p_{1}}\right)\left(\frac{R_{Y X_{1} \mid X_{-1}}^{2}}{1-R_{Y X_{1} \mid X_{-1}}^{2}}\right) \geq F_{1-\alpha}\left(p_{1}, N-1-p\right)\right], & \text { intercept } \\
P\left[\left(\frac{N-p}{p_{1}}\right)\left(\frac{R_{Y X_{1} \mid X_{-1}}^{2}}{1-R_{Y X_{1} \mid X_{-1}}^{2}}\right) \geq F_{1-\alpha}\left(p_{1}, N-p\right)\right], & \text { no intercept } \\
& = \begin{cases}P\left[R_{Y X_{1} \mid X_{-1}}^{2} \geq \frac{F_{1-\alpha}\left(p_{1}, N-1-p\right)}{F_{1-\alpha}\left(p_{1}, N-1-p\right)+\frac{N-1-p}{p_{1}}}\right], & \text { intercept } \\
P\left[R_{Y X_{1} \mid X_{-1}}^{2} \geq \frac{F_{1-\alpha}\left(p_{1}, N-p\right)}{F_{1-\alpha}\left(p_{1}, N-p\right)+\frac{N-p}{p_{1}}}\right], & \text { no intercept }\end{cases}
\end{array}> \begin{cases}\end{cases} \right.
\end{aligned}
$$

The distribution of $R_{Y X_{1} \mid X_{-1}}^{2}$ (for any $\rho_{Y X_{1} \mid X_{-1}}^{2}$ ) is given in Chapter 32 of Johnson, Kotz, and Balakrishnan (1995). Sample size tables are presented in Gatsonis and Sampson (1989).

If the predictors are assumed to have fixed values, then the exact power is given by the noncentral $F$ distribution. The noncentrality parameter is

$$
\lambda=N \frac{\rho_{Y X_{1} \mid X_{-1}}^{2}}{1-\rho_{Y X_{1} \mid X_{-1}}^{2}}
$$

or equivalently,

$$
\lambda=N \frac{\rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}-\rho_{Y \mid X_{-1}}^{2}}{\left.1-\rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}\right)}
$$

The power is

$$
\text { power }= \begin{cases}P\left(F\left(p_{1}, N-1-p, \lambda\right) \geq F_{1-\alpha}\left(p_{1}, N-1-p\right)\right), & \text { intercept } \\ P\left(F\left(p_{1}, N-p, \lambda\right) \geq F_{1-\alpha}\left(p_{1}, N-p\right)\right), & \text { no intercept }\end{cases}
$$

The minimum acceptable input value of $N$ depends on several factors, as shown in Table 77.30.
Table 77.30 Minimum Acceptable Sample Size Values in the MULTREG Statement

| Predictor Type | Intercept in Model? | $p_{1}=1 \boldsymbol{?}$ | Minimum $\boldsymbol{N}$ |
| :--- | :---: | :--- | :---: |
| Random | Yes | Yes | $p+3$ |
| Random | Yes | No | $p+2$ |
| Random | No | Yes | $p+2$ |
| Random | No | No | $p+1$ |
| Fixed | Yes | Yes or No | $p+2$ |
| Fixed | No | Yes or No | $p+1$ |

## Analyses in the ONECORR Statement

## Fisher's z Test for Pearson Correlation (TEST=PEARSON DIST=FISHERZ)

Fisher's $z$ transformation (Fisher 1921) of the sample correlation $R_{Y \mid\left(X_{1}, X_{-1}\right)}$ is defined as

$$
z=\frac{1}{2} \log \left(\frac{1+R_{Y \mid\left(X_{1}, X_{-1}\right)}}{1-R_{Y \mid\left(X_{1}, X_{-1}\right)}}\right)
$$

Fisher's $z$ test assumes the approximate normal distribution $N\left(\mu, \sigma^{2}\right)$ for $z$, where

$$
\mu=\frac{1}{2} \log \left(\frac{1+\rho_{Y \mid\left(X_{1}, X_{-1}\right)}}{1-\rho_{Y \mid\left(X_{1}, X_{-1}\right)}}\right)+\frac{\rho_{Y \mid\left(X_{1}, X_{-1}\right)}}{2\left(N-1-p^{\star}\right)}
$$

and

$$
\sigma^{2}=\frac{1}{N-3-p^{\star}}
$$

where $p^{\star}$ is the number of variables partialed out (Anderson 1984, pp. 132-133) and $\rho_{Y \mid\left(X_{1}, X_{-1}\right)}$ is the partial correlation between $Y$ and $X_{1}$ adjusting for the set of zero or more variables $X_{-1}$.

The test statistic

$$
z^{\star}=\left(N-3-p^{\star}\right)^{\frac{1}{2}}\left[z-\frac{1}{2} \log \left(\frac{1+\rho_{0}}{1-\rho_{0}}\right)-\frac{\rho_{0}}{2\left(N-1-p^{\star}\right)}\right]
$$

is assumed to have a normal distribution $N(\delta, \nu)$, where $\rho_{0}$ is the null partial correlation and $\delta$ and $v$ are derived from Section 16.33 of Stuart and Ord (1994):

$$
\begin{aligned}
\delta= & \left(N-3-p^{\star}\right)^{\frac{1}{2}}\left[\frac{1}{2} \log \left(\frac{1+\rho_{Y \mid\left(X_{1}, X_{-1}\right)}}{1-\rho_{Y \mid\left(X_{1}, X_{-1}\right)}}\right)+\frac{\rho_{Y \mid\left(X_{1}, X_{-1}\right)}}{2\left(N-1-p^{\star}\right)}\left(1+\frac{5+\rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}}{4\left(N-1-p^{\star}\right)}+\right.\right. \\
& \left.\left.\frac{11+2 \rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}+3 \rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{4}}{8\left(N-1-p^{\star}\right)^{2}}\right)-\frac{1}{2} \log \left(\frac{1+\rho_{0}}{1-\rho_{0}}\right)-\frac{\rho_{0}}{2\left(N-1-p^{\star}\right)}\right] \\
v= & \frac{N-3-p^{\star}}{N-1-p^{\star}}\left[1+\frac{4-\rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}}{2\left(N-1-p^{\star}\right)}+\frac{22-6 \rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{2}-3 \rho_{Y \mid\left(X_{1}, X_{-1}\right)}^{4}}{6\left(N-1-p^{\star}\right)^{2}}\right]
\end{aligned}
$$

The approximate power is computed as

$$
\text { power }= \begin{cases}\Phi\left(\frac{\delta-z_{1-\alpha}}{\nu^{\frac{1}{2}}}\right), & \text { upper one-sided } \\ \Phi\left(\frac{-\delta-z_{1}-\alpha}{\nu^{\frac{1}{2}}}\right), & \text { lower one-sided } \\ \Phi\left(\frac{\delta-z_{1-\frac{\alpha}{2}}}{\nu^{\frac{1}{2}}}\right)+\Phi\left(\frac{-\delta-z_{1-\frac{\alpha}{2}}}{\nu^{\frac{1}{2}}}\right), & \text { two-sided }\end{cases}
$$

Because the test is biased, the achieved significance level might differ from the nominal significance level. The actual alpha is computed in the same way as the power, except that the correlation $\rho_{Y \mid\left(X_{1}, X_{-1}\right)}$ is replaced by the null correlation $\rho_{0}$.

## $t$ Test for Pearson Correlation (TEST=PEARSON DIST=T)

The two-sided case is identical to multiple regression with an intercept and $p_{1}=1$, which is discussed in the section "Analyses in the MULTREG Statement" on page 6377.

Let $p^{\star}$ denote the number of variables partialed out. For the one-sided cases, the test statistic is

$$
t=\left(N-2-p^{\star}\right)^{\frac{1}{2}} \frac{R_{Y X_{1} \mid X_{-1}}}{\left(1-R_{Y X_{1} \mid X_{-1}}^{2}\right)^{\frac{1}{2}}}
$$

which is assumed to have a null distribution of $t\left(N-2-p^{\star}\right)$.
If the $X$ and $Y$ variables are assumed to have a joint multivariate normal distribution, then the exact power is given by the following formula:

$$
\begin{aligned}
& \text { power }= \begin{cases}P\left[\begin{array}{ll}
\left.\left(N-2-p^{\star}\right)^{\frac{1}{2}} \frac{R_{Y X_{1} \mid X_{-1}}}{\left(1-R_{Y X_{1} \mid X_{-1}}^{2}\right)^{\frac{1}{2}}} \geq t_{1-\alpha}\left(N-2-p^{\star}\right)\right], & \text { upper one-sided } \\
P\left[\left(N-2-p^{\star}\right)^{\frac{1}{2}} \frac{R_{Y X_{1} \mid X_{-1}}}{\left(1-R_{Y X_{1} \mid X_{-1}}^{2}\right)^{\frac{1}{2}}} \leq t_{\alpha}\left(N-2-p^{\star}\right)\right], & \text { lower one-sided }
\end{array}\right.\end{cases} \\
& =\left\{\begin{array}{lll}
P\left[R_{Y \mid\left(X_{1}, X_{-1}\right)} \geq \frac{t_{1-\alpha}\left(N-2-p^{\star}\right)}{\left(t_{1-\alpha}^{2}\left(N-2-p^{\star}\right)+N-2-p^{\star}\right)^{\frac{1}{2}}}\right], & \text { upper one-sided } \\
P\left[R_{Y \mid\left(X_{1}, X_{-1}\right)} \leq \frac{t_{\alpha}\left(N-2-p^{\star}\right)}{\left(t_{\alpha}^{2}\left(N-2-p^{\star}\right)+N-2-p^{\star}\right)^{\frac{1}{2}}}\right], & \text { lower one-sided }
\end{array}\right.
\end{aligned}
$$

The distribution of $R_{Y \mid\left(X_{1}, X_{-1}\right)}$ (given the underlying true correlation $\rho_{Y \mid\left(X_{1}, X_{-1}\right)}$ ) is given in Chapter 32 of Johnson, Kotz, and Balakrishnan (1995).

If the $X$ variables are assumed to have fixed values, then the exact power is given by the noncentral $t$ distribution $t\left(N-2-p^{\star}, \delta\right)$, where the noncentrality is

$$
\delta=N^{\frac{1}{2}} \frac{\rho_{Y X_{1} \mid X_{-1}}}{\left(1-\rho_{Y X_{1} \mid X_{-1}}^{2}\right)^{\frac{1}{2}}}
$$

The power is

$$
\text { power }= \begin{cases}P\left(t\left(N-2-p^{\star}, \delta\right) \geq t_{1-\alpha}\left(N-2-p^{\star}\right)\right), & \text { upper one-sided } \\ P\left(t\left(N-2-p^{\star}, \delta\right) \leq t_{\alpha}\left(N-2-p^{\star}\right)\right), & \text { lower one-sided }\end{cases}
$$

## Analyses in the ONESAMPLEFREQ Statement

## Exact Test of a Binomial Proportion (TEST=EXACT)

Let $X$ be distributed as $\operatorname{Bin}(N, p)$. The hypotheses for the test of the proportion $p$ are as follows:

$$
\begin{aligned}
& H_{0}: p=p_{0} \\
& H_{1}: \begin{cases}p \neq p_{0}, & \text { two-sided } \\
p>p_{0}, & \text { upper one-sided } \\
p<p_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

The exact test assumes binomially distributed data and requires $N \geq 1$ and $0<p_{0}<1$. The test statistic is

$$
X=\text { number of successes } \sim \operatorname{Bin}(N, p)
$$

The significance probability $\alpha$ is split symmetrically for two-sided tests, in the sense that each tail is filled with as much as possible up to $\alpha / 2$.

Exact power computations are based on the binomial distribution and computing formulas such as the following from Johnson, Kotz, and Kemp (1992, equation 3.20):

$$
P(X \geq C \mid N, p)=P\left(F_{v_{1}, \nu_{2}} \leq \frac{\nu_{2} p}{v_{1}(1-p)}\right) \quad \text { where } \nu_{1}=2 C \text { and } \nu_{2}=2(N-C+1)
$$

Let $C_{L}$ and $C_{U}$ denote lower and upper critical values, respectively. Let $\alpha_{a}$ denote the achieved (actual) significance level, which for two-sided tests is the sum of the favorable major tail ( $\alpha_{M}$ ) and the opposite minor tail ( $\alpha_{m}$ ).

For the upper one-sided case,

$$
C_{U}=\min \left\{C: P\left(X \geq C \mid p_{0}\right) \leq \alpha\right\}
$$

Reject $H_{0}$ if $X \geq C_{U}$

$$
\begin{aligned}
\alpha_{a} & =P\left(X \geq C_{U} \mid p_{0}\right) \\
\text { power } & =P\left(X \geq C_{U} \mid p\right)
\end{aligned}
$$

For the lower one-sided case,

$$
C_{L}=\max \left\{C: P\left(X \leq C \mid p_{0}\right) \leq \alpha\right\}
$$

Reject $H_{0}$ if $X \leq C_{L}$

$$
\begin{aligned}
\alpha_{a} & =P\left(X \leq C_{L} \mid p_{0}\right) \\
\text { power } & =P\left(X \leq C_{L} \mid p\right)
\end{aligned}
$$

For the two-sided case,

$$
\begin{aligned}
& C_{L}=\max \left\{C: P\left(X \leq C \mid p_{0}\right) \leq \frac{\alpha}{2}\right\} \\
& C_{U}=\min \left\{C: P\left(X \geq C \mid p_{0}\right) \leq \frac{\alpha}{2}\right\}
\end{aligned}
$$

Reject $H_{0}$ if $X \leq C_{L}$ or $X \geq C_{U}$

$$
\begin{aligned}
\alpha_{a} & =P\left(X \leq C_{L} \text { or } X \geq C_{U} \mid p_{0}\right) \\
\text { power } & =P\left(X \leq C_{L} \text { or } X \geq C_{U} \mid p\right)
\end{aligned}
$$

## z Test for Binomial Proportion Using Null Variance (TEST=Z VAREST=NULL)

For the normal approximation test, the test statistic is

$$
Z(X)=\frac{X-N p_{0}}{\left[N p_{0}\left(1-p_{0}\right)\right]^{\frac{1}{2}}}
$$

For the METHOD=EXACT option, the computations are the same as described in the section "Exact Test of a Binomial Proportion (TEST=EXACT)" on page 6381 except for the definitions of the critical values.

For the upper one-sided case,

$$
C_{U}=\min \left\{C: Z(C) \geq z_{1-\alpha}\right\}
$$

For the lower one-sided case,

$$
C_{L}=\max \left\{C: Z(C) \leq z_{\alpha}\right\}
$$

For the two-sided case,

$$
\begin{aligned}
& C_{L}=\max \left\{C: Z(C) \leq z \frac{\alpha}{2}\right\} \\
& C_{U}=\min \left\{C: Z(C) \geq z_{1-\frac{\alpha}{2}}\right\}
\end{aligned}
$$

For the METHOD=NORMAL option, the test statistic $Z(X)$ is assumed to have the normal distribution

$$
\mathrm{N}\left(\frac{N^{\frac{1}{2}}\left(p-p_{0}\right)}{\left[p_{0}\left(1-p_{0}\right)\right]^{\frac{1}{2}}}, \frac{p(1-p)}{p_{0}\left(1-p_{0}\right)}\right)
$$

The approximate power is computed as

$$
\text { power }= \begin{cases}\Phi\left(\frac{z_{\alpha}+\sqrt{N} \frac{p-p_{0}}{\sqrt{p_{0}\left(1-p_{0}\right)}}}{\left.\sqrt{\frac{p(1-p)}{p_{0}\left(1-p_{0}\right.}}\right),}\right. & \text { upper one-sided } \\ \Phi\left(\frac{z_{\alpha}-\sqrt{N} \frac{p-p_{0}}{\sqrt{p_{0}\left(1-p_{0}\right)}}}{\left.\sqrt{\frac{p(1-p)}{p_{0}\left(1-p_{0}\right.}}\right),}\right. & \text { lower one-sided } \\ \Phi\left(\frac{z_{\frac{\alpha}{2}}^{N}+\sqrt{N} \frac{p-p_{0}}{\sqrt{p_{0}\left(1-p_{0}\right)}}}{\sqrt{\frac{p(1-p)}{p_{0}\left(1-p_{0}\right)}}}\right)+\Phi\left(\frac{\left.z_{\frac{\alpha}{2}-\sqrt{N} \frac{p-p_{0}}{\sqrt{p_{p}\left(1-p_{0}\right)}}}^{\sqrt{\frac{p(1-p)}{p_{0}\left(1-p_{0}\right)}}}\right),}{}\right. \text { two-sided }\end{cases}
$$

The approximate sample size is computed in closed form for the one-sided cases by inverting the power equation,

$$
N=\left(\frac{z_{\text {power }} \sqrt{p(1-p)}+z_{1-\alpha} \sqrt{p_{0}\left(1-p_{0}\right)}}{p-p_{0}}\right)^{2}
$$

and by numerical inversion for the two-sided case.

## z Test for Binomial Proportion Using Sample Variance (TEST=Z VAREST=SAMPLE)

For the normal approximation test using the sample variance, the test statistic is

$$
Z_{s}(X)=\frac{X-N p_{0}}{[N \hat{p}(1-\hat{p})]^{\frac{1}{2}}}
$$

where $\hat{p}=X / N$.
For the METHOD=EXACT option, the computations are the same as described in the section "Exact Test of a Binomial Proportion (TEST=EXACT)" on page 6381 except for the definitions of the critical values.

For the upper one-sided case,

$$
C_{U}=\min \left\{C: Z_{s}(C) \geq z_{1-\alpha}\right\}
$$

For the lower one-sided case,

$$
C_{L}=\max \left\{C: Z_{s}(C) \leq z_{\alpha}\right\}
$$

For the two-sided case,

$$
\begin{aligned}
& C_{L}=\max \left\{C: Z_{s}(C) \leq z_{\frac{\alpha}{2}}\right\} \\
& C_{U}=\min \left\{C: Z_{s}(C) \geq z_{1-\frac{\alpha}{2}}\right\}
\end{aligned}
$$

For the METHOD=NORMAL option, the test statistic $Z_{s}(X)$ is assumed to have the normal distribution

$$
\mathrm{N}\left(\frac{N^{\frac{1}{2}}\left(p-p_{0}\right)}{[p(1-p)]^{\frac{1}{2}}}, 1\right)
$$

(see Chow, Shao, and Wang (2003, p. 82)).
The approximate power is computed as

$$
\text { power }= \begin{cases}\Phi\left(z_{\alpha}+\sqrt{N} \frac{p-p_{0}}{\sqrt{p(1-p)}}\right), & \text { upper one-sided } \\ \Phi\left(z_{\alpha}-\sqrt{N} \frac{p-p_{0}}{\sqrt{p(1-p)}}\right), & \text { lower one-sided } \\ \Phi\left(z_{\frac{\alpha}{2}}+\sqrt{N} \frac{p-p_{0}}{\sqrt{p(1-p)}}\right)+\Phi\left(z_{\frac{\alpha}{2}}-\sqrt{N} \frac{p-p_{0}}{\sqrt{p(1-p)}}\right), & \text { two-sided }\end{cases}
$$

The approximate sample size is computed in closed form for the one-sided cases by inverting the power equation,

$$
N=p(1-p)\left(\frac{z_{\mathrm{power}}+z_{1-\alpha}}{p-p_{0}}\right)^{2}
$$

and by numerical inversion for the two-sided case.

## z Test for Binomial Proportion with Continuity Adjustment Using Null Variance (TEST=ADJZ VAREST=NULL)

For the normal approximation test with continuity adjustment, the test statistic is (Pagano and Gauvreau 1993, p. 295):

$$
Z_{c}(X)=\frac{X-N p_{0}+0.5\left(1_{\left\{X<N p_{0}\right\}}\right)-0.5\left(1_{\left\{X>N p_{0}\right\}}\right)}{\left[N p_{0}\left(1-p_{0}\right)\right]^{\frac{1}{2}}}
$$

For the METHOD=EXACT option, the computations are the same as described in the section "Exact Test of a Binomial Proportion (TEST=EXACT)" on page 6381 except for the definitions of the critical values.

For the upper one-sided case,

$$
C_{U}=\min \left\{C: Z_{c}(C) \geq z_{1-\alpha}\right\}
$$

For the lower one-sided case,

$$
C_{L}=\max \left\{C: Z_{c}(C) \leq z_{\alpha}\right\}
$$

For the two-sided case,

$$
\begin{aligned}
& C_{L}=\max \left\{C: Z_{c}(C) \leq z_{\frac{\alpha}{2}}\right\} \\
& C_{U}=\min \left\{C: Z_{c}(C) \geq z_{1-\frac{\alpha}{2}}\right\}
\end{aligned}
$$

For the METHOD=NORMAL option, the test statistic $Z_{c}(X)$ is assumed to have the normal distribution $N\left(\mu, \sigma^{2}\right)$, where $\mu$ and $\sigma^{2}$ are derived as follows.

For convenience of notation, define

$$
k=\frac{1}{2 \sqrt{N p_{0}\left(1-p_{0}\right)}}
$$

Then

$$
E\left[Z_{c}(X)\right]=2 k N p-2 k N p_{0}+k P\left(X<N p_{0}\right)-k P\left(X>N p_{0}\right)
$$

and

$$
\begin{aligned}
\operatorname{Var}\left[Z_{c}(X)\right]= & 4 k^{2} N p(1-p)+k^{2}\left[1-P\left(X=N p_{0}\right)\right]-k^{2}\left[P\left(X<N p_{0}\right)-P\left(X>N p_{0}\right)\right]^{2} \\
& +4 k^{2}\left[E\left(X 1_{\left\{X<N p_{0}\right\}}\right)-E\left(X 1_{\left\{X>N p_{0}\right\}}\right)\right]-4 k^{2} N p\left[P\left(X<N p_{0}\right)-P\left(X>N p_{0}\right)\right]
\end{aligned}
$$

The probabilities $P\left(X=N p_{0}\right), P\left(X<N p_{0}\right)$, and $P\left(X>N p_{0}\right)$ and the truncated expectations $E\left(X 1_{\left\{X<N p_{0}\right\}}\right)$ and $E\left(X 1_{\left\{X>N p_{0}\right\}}\right)$ are approximated by assuming the normal-approximate distribution of $X, N(N p, N p(1-p))$. Letting $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal PDF and CDF, respectively, and defining $d$ as

$$
d=\frac{N p_{0}-N p}{[N p(1-p)]^{\frac{1}{2}}}
$$

the terms are computed as follows:

$$
\begin{aligned}
P\left(X=N p_{0}\right) & =0 \\
P\left(X<N p_{0}\right) & =\Phi(d) \\
P\left(X>N p_{0}\right) & =1-\Phi(d) \\
E\left(X 1_{\left\{X<N p_{0}\right\}}\right) & =N p \Phi(d)-[N p(1-p)]^{\frac{1}{2}} \phi(d) \\
E\left(X 1_{\left\{X>N p_{0}\right\}}\right) & =N p[1-\Phi(d)]+[N p(1-p)]^{\frac{1}{2}} \phi(d)
\end{aligned}
$$

The mean and variance of $Z_{c}(X)$ are thus approximated by

$$
\mu=k\left[2 N p-2 N p_{0}+2 \Phi(d)-1\right]
$$

and

$$
\sigma^{2}=4 k^{2}\left[N p(1-p)+\Phi(d)(1-\Phi(d))-2(N p(1-p))^{\frac{1}{2}} \phi(d)\right]
$$

The approximate power is computed as

$$
\text { power }= \begin{cases}\Phi\left(\frac{z_{\alpha}+\mu}{\sigma}\right), & \text { upper one-sided } \\ \Phi\left(\frac{z_{\alpha}-\mu}{\sigma}\right), & \text { lower one-sided } \\ \Phi\left(\frac{z_{\frac{\alpha}{2}}+\mu}{\sigma}\right)+\Phi\left(\frac{z_{\frac{\alpha}{2}}-\mu}{\sigma}\right), & \text { two-sided }\end{cases}
$$

The approximate sample size is computed by numerical inversion.

## z Test for Binomial Proportion with Continuity Adjustment Using Sample Variance (TEST=ADJZ VAREST=SAMPLE)

For the normal approximation test with continuity adjustment using the sample variance, the test statistic is

$$
Z_{c s}(X)=\frac{X-N p_{0}+0.5\left(1_{\left\{X<N p_{0}\right\}}\right)-0.5\left(1_{\left\{X>N p_{0}\right\}}\right)}{[N \hat{p}(1-\hat{p})]^{\frac{1}{2}}}
$$

where $\hat{p}=X / N$.
For the METHOD=EXACT option, the computations are the same as described in the section "Exact Test of a Binomial Proportion (TEST=EXACT)" on page 6381 except for the definitions of the critical values.

For the upper one-sided case,

$$
C_{U}=\min \left\{C: Z_{c s}(C) \geq z_{1-\alpha}\right\}
$$

For the lower one-sided case,

$$
C_{L}=\max \left\{C: Z_{c s}(C) \leq z_{\alpha}\right\}
$$

For the two-sided case,

$$
\begin{aligned}
& C_{L}=\max \left\{C: Z_{c s}(C) \leq z \frac{\alpha}{2}\right\} \\
& C_{U}=\min \left\{C: Z_{c s}(C) \geq z_{1-\frac{\alpha}{2}}\right\}
\end{aligned}
$$

For the METHOD=NORMAL option, the test statistic $Z_{c s}(X)$ is assumed to have the normal distribution $N\left(\mu, \sigma^{2}\right)$, where $\mu$ and $\sigma^{2}$ are derived as follows.

For convenience of notation, define

$$
k=\frac{1}{2 \sqrt{N p(1-p)}}
$$

Then

$$
E\left[Z_{c s}(X)\right] \approx 2 k N p-2 k N p_{0}+k P\left(X<N p_{0}\right)-k P\left(X>N p_{0}\right)
$$

and

$$
\begin{aligned}
\operatorname{Var}\left[Z_{c s}(X)\right] \approx & 4 k^{2} N p(1-p)+k^{2}\left[1-P\left(X=N p_{0}\right)\right]-k^{2}\left[P\left(X<N p_{0}\right)-P\left(X>N p_{0}\right)\right]^{2} \\
& +4 k^{2}\left[E\left(X 1_{\left\{X<N p_{0}\right\}}\right)-E\left(X 1_{\left\{X>N p_{0}\right\}}\right)\right]-4 k^{2} N p\left[P\left(X<N p_{0}\right)-P\left(X>N p_{0}\right)\right]
\end{aligned}
$$

The probabilities $P\left(X=N p_{0}\right), P\left(X<N p_{0}\right)$, and $P\left(X>N p_{0}\right)$ and the truncated expectations $E\left(X 1_{\left\{X<N p_{0}\right\}}\right)$ and $E\left(X 1_{\left\{X>N p_{0}\right\}}\right)$ are approximated by assuming the normal-approximate distribution of $X, N(N p, N p(1-p))$. Letting $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal PDF and CDF, respectively, and defining $d$ as

$$
d=\frac{N p_{0}-N p}{[N p(1-p)]^{\frac{1}{2}}}
$$

the terms are computed as follows:

$$
\begin{aligned}
P\left(X=N p_{0}\right) & =0 \\
P\left(X<N p_{0}\right) & =\Phi(d) \\
P\left(X>N p_{0}\right) & =1-\Phi(d) \\
E\left(X 1_{\left\{X<N p_{0}\right\}}\right) & =N p \Phi(d)-[N p(1-p)]^{\frac{1}{2}} \phi(d) \\
E\left(X 1_{\left\{X>N p_{0}\right\}}\right) & =N p[1-\Phi(d)]+[N p(1-p)]^{\frac{1}{2}} \phi(d)
\end{aligned}
$$

The mean and variance of $Z_{c s}(X)$ are thus approximated by

$$
\mu=k\left[2 N p-2 N p_{0}+2 \Phi(d)-1\right]
$$

and

$$
\sigma^{2}=4 k^{2}\left[N p(1-p)+\Phi(d)(1-\Phi(d))-2(N p(1-p))^{\frac{1}{2}} \phi(d)\right]
$$

The approximate power is computed as

$$
\text { power }= \begin{cases}\Phi\left(\frac{z_{\alpha}+\mu}{\sigma}\right), & \text { upper one-sided } \\ \Phi\left(\frac{z_{\alpha}-\mu}{\sigma}\right), & \text { lower one-sided } \\ \Phi\left(\frac{z_{2}+\mu}{\sigma}\right)+\Phi\left(\frac{z_{\frac{\alpha}{2}}-\mu}{\sigma}\right), & \text { two-sided }\end{cases}
$$

The approximate sample size is computed by numerical inversion.

## Exact Equivalence Test of a Binomial Proportion (TEST=EQUIV_EXACT)

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: p<\theta_{L} \quad \text { or } \quad p>\theta_{U} \\
& H_{1}: \theta_{L} \leq p \leq \theta_{U}
\end{aligned}
$$

where $\theta_{L}$ and $\theta_{U}$ are the lower and upper equivalence bounds, respectively.
The analysis is the two one-sided tests (TOST) procedure as described in Chow, Shao, and Wang (2003) on p. 84, but using exact critical values as on p. 116 instead of normal-based critical values.

Two different hypothesis tests are carried out:

$$
\begin{aligned}
& H_{a 0}: p<\theta_{L} \\
& H_{a 1}: p \geq \theta_{L}
\end{aligned}
$$

and

$$
\begin{aligned}
& H_{b 0}: p>\theta_{U} \\
& H_{b 1}: p \leq \theta_{U}
\end{aligned}
$$

If $H_{a 0}$ is rejected in favor of $H_{a 1}$ and $H_{b 0}$ is rejected in favor of $H_{b 1}$, then $H_{0}$ is rejected in favor of $H_{1}$.
The test statistic for each of the two tests ( $H_{a 0}$ versus $H_{a 1}$ and $H_{b 0}$ versus $H_{b 1}$ ) is

$$
X=\text { number of successes } \sim \operatorname{Bin}(N, p)
$$

Let $C_{U}$ denote the critical value of the exact upper one-sided test of $H_{a 0}$ versus $H_{a 1}$, and let $C_{L}$ denote the critical value of the exact lower one-sided test of $H_{b 0}$ versus $H_{b 1}$. These critical values are computed in the section "Exact Test of a Binomial Proportion (TEST=EXACT)" on page 6381. Both of these tests are rejected if and only if $C_{U} \leq X \leq C_{L}$. Thus, the exact power of the equivalence test is

$$
\begin{aligned}
\text { power } & =P\left(C_{U} \leq X \leq C_{L}\right) \\
& =P\left(X \geq C_{U}\right)-P\left(X \geq C_{L}+1\right)
\end{aligned}
$$

The probabilities are computed using Johnson and Kotz (1970, equation 3.20).

## z Equivalence Test for Binomial Proportion Using Null Variance (TEST=EQUIV_Z VAREST=NULL)

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: p<\theta_{L} \text { or } p>\theta_{U} \\
& H_{1}: \theta_{L} \leq p \leq \theta_{U}
\end{aligned}
$$

where $\theta_{L}$ and $\theta_{U}$ are the lower and upper equivalence bounds, respectively.
The analysis is the two one-sided tests (TOST) procedure as described in Chow, Shao, and Wang (2003) on p. 84, but using the null variance instead of the sample variance.

Two different hypothesis tests are carried out:

$$
\begin{aligned}
& H_{a 0}: p<\theta_{L} \\
& H_{a 1}: p \geq \theta_{L}
\end{aligned}
$$

and

$$
\begin{aligned}
& H_{b 0}: p>\theta_{U} \\
& H_{b 1}: p \leq \theta_{U}
\end{aligned}
$$

If $H_{a 0}$ is rejected in favor of $H_{a 1}$ and $H_{b 0}$ is rejected in favor of $H_{b 1}$, then $H_{0}$ is rejected in favor of $H_{1}$.
The test statistic for the test of $H_{a 0}$ versus $H_{a 1}$ is

$$
Z_{L}(X)=\frac{X-N \theta_{L}}{\left[N \theta_{L}\left(1-\theta_{L}\right)\right]^{\frac{1}{2}}}
$$

The test statistic for the test of $H_{b 0}$ versus $H_{b 1}$ is

$$
Z_{U}(X)=\frac{X-N \theta_{U}}{\left[N \theta_{U}\left(1-\theta_{U}\right)\right]^{\frac{1}{2}}}
$$

For the METHOD=EXACT option, let $C_{U}$ denote the critical value of the exact upper one-sided test of $H_{a 0}$ versus $H_{a 1}$ using $Z_{L}(X)$. This critical value is computed in the section "z Test for Binomial Proportion Using Null Variance (TEST=Z VAREST=NULL)" on page 6382. Similarly, let $C_{L}$ denote the critical value of the exact lower one-sided test of $H_{b 0}$ versus $H_{b 1}$ using $Z_{U}(X)$. Both of these tests are rejected if and only if $C_{U} \leq X \leq C_{L}$. Thus, the exact power of the equivalence test is

$$
\begin{aligned}
\text { power } & =P\left(C_{U} \leq X \leq C_{L}\right) \\
& =P\left(X \geq C_{U}\right)-P\left(X \geq C_{L}+1\right)
\end{aligned}
$$

The probabilities are computed using Johnson and Kotz (1970, equation 3.20).
For the METHOD=NORMAL option, the test statistic $Z_{L}(X)$ is assumed to have the normal distribution

$$
\mathrm{N}\left(\frac{N^{\frac{1}{2}}\left(p-\theta_{L}\right)}{\left[\theta_{L}\left(1-\theta_{L}\right)\right]^{\frac{1}{2}}}, \frac{p(1-p)}{\theta_{L}\left(1-\theta_{L}\right)}\right)
$$

and the test statistic $Z_{U}(X)$ is assumed to have the normal distribution

$$
\mathrm{N}\left(\frac{N^{\frac{1}{2}}\left(p-\theta_{U}\right)}{\left[\theta_{U}\left(1-\theta_{U}\right)\right]^{\frac{1}{2}}}, \frac{p(1-p)}{\theta_{U}\left(1-\theta_{U}\right)}\right)
$$

(see Chow, Shao, and Wang (2003, p. 84)). The approximate power is computed as

$$
\text { power }=\Phi\left(\frac{z_{\alpha}-\sqrt{N} \frac{p-\theta_{U}}{\sqrt{\theta_{U}\left(1-\theta_{U}\right)}}}{\sqrt{\frac{p(1-p)}{\theta_{U}\left(1-\theta_{U}\right)}}}\right)+\Phi\left(\frac{z_{\alpha}+\sqrt{N} \frac{p-\theta_{L}}{\sqrt{\theta_{L}\left(1-\theta_{L}\right)}}}{\sqrt{\frac{p(1-p)}{\theta_{L}\left(1-\theta_{L}\right)}}}\right)-1
$$

The approximate sample size is computed by numerically inverting the power formula, using the sample size estimate $N_{0}$ of Chow, Shao, and Wang (2003, p. 85) as an initial guess:

$$
N_{0}=p(1-p)\left(\frac{z_{1-\alpha}+z_{(1+\text { power }) / 2}}{0.5\left(\theta_{U}-\theta_{L}\right)-\left|p-0.5\left(\theta_{L}+\theta_{U}\right)\right|}\right)^{2}
$$

## z Equivalence Test for Binomial Proportion Using Sample Variance (TEST=EQUIV_Z VAREST=SAMPLE)

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: p<\theta_{L} \text { or } p>\theta_{U} \\
& H_{1}: \theta_{L} \leq p \leq \theta_{U}
\end{aligned}
$$

where $\theta_{L}$ and $\theta_{U}$ are the lower and upper equivalence bounds, respectively.
The analysis is the two one-sided tests (TOST) procedure as described in Chow, Shao, and Wang (2003) on p. 84.

Two different hypothesis tests are carried out:

$$
\begin{aligned}
& H_{a 0}: p<\theta_{L} \\
& H_{a 1}: p \geq \theta_{L}
\end{aligned}
$$

and

$$
\begin{aligned}
H_{b 0}: p>\theta_{U} \\
H_{b 1}: p \leq \theta_{U}
\end{aligned}
$$

If $H_{a 0}$ is rejected in favor of $H_{a 1}$ and $H_{b 0}$ is rejected in favor of $H_{b 1}$, then $H_{0}$ is rejected in favor of $H_{1}$.
The test statistic for the test of $H_{a 0}$ versus $H_{a 1}$ is

$$
Z_{s L}(X)=\frac{X-N \theta_{L}}{[N \hat{p}(1-\hat{p})]^{\frac{1}{2}}}
$$

where $\hat{p}=X / N$.
The test statistic for the test of $H_{b 0}$ versus $H_{b 1}$ is

$$
Z_{s U}(X)=\frac{X-N \theta_{U}}{[N \hat{p}(1-\hat{p})]^{\frac{1}{2}}}
$$

For the METHOD=EXACT option, let $C_{U}$ denote the critical value of the exact upper one-sided test of $H_{a 0}$ versus $H_{a 1}$ using $Z_{s L}(X)$. This critical value is computed in the section "z Test for Binomial Proportion Using Sample Variance (TEST=Z VAREST=SAMPLE)" on page 6383. Similarly, let $C_{L}$ denote the critical value of the exact lower one-sided test of $H_{b 0}$ versus $H_{b 1}$ using $Z_{s U}(X)$. Both of these tests are rejected if and only if $C_{U} \leq X \leq C_{L}$. Thus, the exact power of the equivalence test is

$$
\begin{aligned}
\text { power } & =P\left(C_{U} \leq X \leq C_{L}\right) \\
& =P\left(X \geq C_{U}\right)-P\left(X \geq C_{L}+1\right)
\end{aligned}
$$

The probabilities are computed using Johnson and Kotz (1970, equation 3.20).
For the METHOD=NORMAL option, the test statistic $Z_{s L}(X)$ is assumed to have the normal distribution

$$
\mathrm{N}\left(\frac{N^{\frac{1}{2}}\left(p-\theta_{L}\right)}{[p(1-p)]^{\frac{1}{2}}}, 1\right)
$$

and the test statistic $Z_{S U}(X)$ is assumed to have the normal distribution

$$
\mathrm{N}\left(\frac{N^{\frac{1}{2}}\left(p-\theta_{U}\right)}{[p(1-p)]^{\frac{1}{2}}}, 1\right)
$$

(see Chow, Shao, and Wang (2003), p. 84).
The approximate power is computed as

$$
\text { power }=\Phi\left(z_{\alpha}-\sqrt{N} \frac{p-\theta_{U}}{\sqrt{p(1-p)}}\right)+\Phi\left(z_{\alpha}+\sqrt{N} \frac{p-\theta_{L}}{\sqrt{p(1-p)}}\right)-1
$$

The approximate sample size is computed by numerically inverting the power formula, using the sample size estimate $N_{0}$ of Chow, Shao, and Wang (2003, p. 85) as an initial guess:

$$
N_{0}=p(1-p)\left(\frac{z_{1-\alpha}+z_{(1+\text { power }) / 2}}{0.5\left(\theta_{U}-\theta_{L}\right)-\left|p-0.5\left(\theta_{L}+\theta_{U}\right)\right|}\right)^{2}
$$

## z Equivalence Test for Binomial Proportion with Continuity Adjustment Using Null Variance (TEST=EQUIV_ADJZ VAREST=NULL)

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: p<\theta_{L} \text { or } p>\theta_{U} \\
& H_{1}: \theta_{L} \leq p \leq \theta_{U}
\end{aligned}
$$

where $\theta_{L}$ and $\theta_{U}$ are the lower and upper equivalence bounds, respectively.
The analysis is the two one-sided tests (TOST) procedure as described in Chow, Shao, and Wang (2003) on p. 84, but using the null variance instead of the sample variance.

Two different hypothesis tests are carried out:

$$
\begin{aligned}
& H_{a 0}: p<\theta_{L} \\
& H_{a 1}: p \geq \theta_{L}
\end{aligned}
$$

and

$$
\begin{aligned}
& H_{b 0}: p>\theta_{U} \\
& H_{b 1}: p \leq \theta_{U}
\end{aligned}
$$

If $H_{a 0}$ is rejected in favor of $H_{a 1}$ and $H_{b 0}$ is rejected in favor of $H_{b 1}$, then $H_{0}$ is rejected in favor of $H_{1}$. The test statistic for the test of $H_{a 0}$ versus $H_{a 1}$ is

$$
Z_{c L}(X)=\frac{X-N \theta_{L}+0.5\left(1_{\left\{X<N \theta_{L}\right\}}\right)-0.5\left(1_{\left\{X>N \theta_{L}\right\}}\right)}{\left[N \hat{\theta_{L}}\left(1-\hat{\theta_{L}}\right)\right]^{\frac{1}{2}}}
$$

where $\hat{p}=X / N$.
The test statistic for the test of $H_{b 0}$ versus $H_{b 1}$ is

$$
Z_{c U}(X)=\frac{X-N \theta_{U}+0.5\left(1_{\left\{X<N \theta_{U}\right\}}\right)-0.5\left(1_{\left\{X>N \theta_{U}\right\}}\right)}{\left[N \hat{\theta_{U}}\left(1-\hat{\theta_{U}}\right)\right]^{\frac{1}{2}}}
$$

For the METHOD=EXACT option, let $C_{U}$ denote the critical value of the exact upper one-sided test of $H_{a 0}$ versus $H_{a 1}$ using $Z_{c L}(X)$. This critical value is computed in the section " z Test for Binomial Proportion with Continuity Adjustment Using Null Variance (TEST=ADJZ VAREST=NULL)" on page 6384. Similarly, let $C_{L}$ denote the critical value of the exact lower one-sided test of $H_{b 0}$ versus $H_{b 1}$ using $Z_{c U}(X)$. Both of these tests are rejected if and only if $C_{U} \leq X \leq C_{L}$. Thus, the exact power of the equivalence test is

$$
\begin{aligned}
\text { power } & =P\left(C_{U} \leq X \leq C_{L}\right) \\
& =P\left(X \geq C_{U}\right)-P\left(X \geq C_{L}+1\right)
\end{aligned}
$$

The probabilities are computed using Johnson and Kotz (1970, equation 3.20).
For the METHOD $=$ NORMAL option, the test statistic $Z_{c L}(X)$ is assumed to have the normal distribution $N\left(\mu_{L}, \sigma_{L}^{2}\right)$, and $Z_{c U}(X)$ is assumed to have the normal distribution $N\left(\mu_{U}, \sigma_{U}^{2}\right)$, where $\mu_{L}, \mu_{U}, \sigma_{L}^{2}$, and $\sigma_{U}^{2}$ are derived as follows.

For convenience of notation, define

$$
\begin{aligned}
k_{L} & =\frac{1}{2 \sqrt{N \theta_{L}\left(1-\theta_{L}\right)}} \\
k_{U} & =\frac{1}{2 \sqrt{N \theta_{U}\left(1-\theta_{U}\right)}}
\end{aligned}
$$

Then

$$
\begin{aligned}
& E\left[Z_{c L}(X)\right] \approx 2 k_{L} N p-2 k_{L} N \theta_{L}+k_{L} P\left(X<N \theta_{L}\right)-k_{L} P\left(X>N \theta_{L}\right) \\
& E\left[Z_{c U}(X)\right] \approx 2 k_{U} N p-2 k_{U} N \theta_{U}+k_{U} P\left(X<N \theta_{U}\right)-k_{U} P\left(X>N \theta_{U}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{Var}\left[Z_{c L}(X)\right] \approx & 4 k_{L}^{2} N p(1-p)+k_{L}^{2}\left[1-P\left(X=N \theta_{L}\right)\right]-k_{L}^{2}\left[P\left(X<N \theta_{L}\right)-P\left(X>N \theta_{L}\right)\right]^{2} \\
& +4 k_{L}^{2}\left[E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right)-E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right)\right]-4 k_{L}^{2} N p\left[P\left(X<N \theta_{L}\right)-P\left(X>N \theta_{L}\right)\right] \\
\operatorname{Var}\left[Z_{c U}(X)\right] \approx & 4 k_{U}^{2} N p(1-p)+k_{U}^{2}\left[1-P\left(X=N \theta_{U}\right)\right]-k_{U}^{2}\left[P\left(X<N \theta_{U}\right)-P\left(X>N \theta_{U}\right)\right]^{2} \\
& +4 k_{U}^{2}\left[E\left(X 1_{\left\{X<N \theta_{U}\right\}}\right)-E\left(X 1_{\left\{X>N \theta_{U}\right\}}\right)\right]-4 k_{U}^{2} N p\left[P\left(X<N \theta_{U}\right)-P\left(X>N \theta_{U}\right)\right]
\end{aligned}
$$

The probabilities $P\left(X=N \theta_{L}\right), P\left(X<N \theta_{L}\right), P\left(X>N \theta_{L}\right), P\left(X=N \theta_{U}\right), P\left(X<N \theta_{U}\right)$, and $P\left(X>N \theta_{U}\right)$ and the truncated expectations $E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right), E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right), E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right)$, and $E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right)$ are approximated by assuming the normal-approximate distribution of $X, N(N p, N p(1-p))$. Letting $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal PDF and CDF, respectively, and defining $d_{L}$ and $d_{U}$ as

$$
\begin{aligned}
d_{L} & =\frac{N \theta_{L}-N p}{[N p(1-p)]^{\frac{1}{2}}} \\
d_{U} & =\frac{N \theta_{U}-N p}{[N p(1-p)]^{\frac{1}{2}}}
\end{aligned}
$$

the terms are computed as follows:

$$
\begin{aligned}
P\left(X=N \theta_{L}\right) & =0 \\
P\left(X=N \theta_{U}\right) & =0 \\
P\left(X<N \theta_{L}\right) & =\Phi\left(d_{L}\right) \\
P\left(X<N \theta_{U}\right) & =\Phi\left(d_{U}\right) \\
P\left(X>N \theta_{L}\right) & =1-\Phi\left(d_{L}\right) \\
P\left(X>N \theta_{U}\right) & =1-\Phi\left(d_{U}\right) \\
E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right) & =N p \Phi\left(d_{L}\right)-[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{L}\right) \\
E\left(X 1_{\left\{X<N \theta_{U}\right\}}\right) & =N p \Phi\left(d_{U}\right)-[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{U}\right) \\
E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right) & =N p\left[1-\Phi\left(d_{L}\right)\right]+[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{L}\right) \\
E\left(X 1_{\left\{X>N \theta_{U}\right\}}\right) & =N p\left[1-\Phi\left(d_{U}\right)\right]+[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{U}\right)
\end{aligned}
$$

The mean and variance of $Z_{c L}(X)$ and $Z_{c U}(X)$ are thus approximated by

$$
\begin{aligned}
\mu_{L} & =k_{L}\left[2 N p-2 N \theta_{L}+2 \Phi\left(d_{L}\right)-1\right] \\
\mu_{U} & =k_{U}\left[2 N p-2 N \theta_{U}+2 \Phi\left(d_{U}\right)-1\right]
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma_{L}^{2} & =4 k_{L}^{2}\left[N p(1-p)+\Phi\left(d_{L}\right)\left(1-\Phi\left(d_{L}\right)\right)-2(N p(1-p))^{\frac{1}{2}} \phi\left(d_{L}\right)\right] \\
\sigma_{U}^{2} & =4 k_{U}^{2}\left[N p(1-p)+\Phi\left(d_{U}\right)\left(1-\Phi\left(d_{U}\right)\right)-2(N p(1-p))^{\frac{1}{2}} \phi\left(d_{U}\right)\right]
\end{aligned}
$$

The approximate power is computed as

$$
\text { power }=\Phi\left(\frac{z_{\alpha}-\mu_{U}}{\sigma_{U}}\right)+\Phi\left(\frac{z_{\alpha}+\mu_{L}}{\sigma_{L}}\right)-1
$$

The approximate sample size is computed by numerically inverting the power formula.

## z Equivalence Test for Binomial Proportion with Continuity Adjustment Using Sample Variance (TEST=EQUIV_ADJZ VAREST=SAMPLE)

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: p<\theta_{L} \text { or } p>\theta_{U} \\
& H_{1}: \theta_{L} \leq p \leq \theta_{U}
\end{aligned}
$$

where $\theta_{L}$ and $\theta_{U}$ are the lower and upper equivalence bounds, respectively.
The analysis is the two one-sided tests (TOST) procedure as described in Chow, Shao, and Wang (2003) on p. 84.

Two different hypothesis tests are carried out:

$$
\begin{aligned}
& H_{a 0}: p<\theta_{L} \\
& H_{a 1}: p \geq \theta_{L}
\end{aligned}
$$

and

$$
\begin{aligned}
& H_{b 0}: p>\theta_{U} \\
& H_{b 1}: p \leq \theta_{U}
\end{aligned}
$$

If $H_{a 0}$ is rejected in favor of $H_{a 1}$ and $H_{b 0}$ is rejected in favor of $H_{b 1}$, then $H_{0}$ is rejected in favor of $H_{1}$. The test statistic for the test of $H_{a 0}$ versus $H_{a 1}$ is

$$
Z_{c s L}(X)=\frac{X-N \theta_{L}+0.5\left(1_{\left\{X<N \theta_{L}\right\}}\right)-0.5\left(1_{\left\{X>N \theta_{L}\right\}}\right)}{[N \hat{p}(1-\hat{p})]^{\frac{1}{2}}}
$$

where $\hat{p}=X / N$.
The test statistic for the test of $H_{b 0}$ versus $H_{b 1}$ is

$$
Z_{c s U}(X)=\frac{X-N \theta_{U}+0.5\left(1_{\left\{X<N \theta_{U}\right\}}\right)-0.5\left(1_{\left\{X>N \theta_{U}\right\}}\right)}{[N \hat{p}(1-\hat{p})]^{\frac{1}{2}}}
$$

For the METHOD=EXACT option, let $C_{U}$ denote the critical value of the exact upper one-sided test of $H_{a 0}$ versus $H_{a 1}$ using $Z_{c s L}(X)$. This critical value is computed in the section " $z$ Test for Binomial Proportion with Continuity Adjustment Using Sample Variance (TEST=ADJZ VAREST=SAMPLE)" on page 6386. Similarly, let $C_{L}$ denote the critical value of the exact lower one-sided test of $H_{b 0}$ versus $H_{b 1}$ using $Z_{c s U}(X)$. Both of these tests are rejected if and only if $C_{U} \leq X \leq C_{L}$. Thus, the exact power of the equivalence test is

$$
\begin{aligned}
\text { power } & =P\left(C_{U} \leq X \leq C_{L}\right) \\
& =P\left(X \geq C_{U}\right)-P\left(X \geq C_{L}+1\right)
\end{aligned}
$$

The probabilities are computed using Johnson and Kotz (1970, equation 3.20).
For the METHOD=NORMAL option, the test statistic $Z_{c s L}(X)$ is assumed to have the normal distribution $N\left(\mu_{L}, \sigma_{L}^{2}\right)$, and $Z_{c s U}(X)$ is assumed to have the normal distribution $N\left(\mu_{U}, \sigma_{U}^{2}\right)$, where $\mu_{L}, \mu_{U}, \sigma_{L}^{2}$ and $\sigma_{U}^{2}$ are derived as follows.
For convenience of notation, define

$$
k=\frac{1}{2 \sqrt{N p(1-p)}}
$$

Then

$$
\begin{aligned}
& E\left[Z_{c s L}(X)\right] \approx 2 k N p-2 k N \theta_{L}+k P\left(X<N \theta_{L}\right)-k P\left(X>N \theta_{L}\right) \\
& E\left[Z_{c s U}(X)\right] \approx 2 k N p-2 k N \theta_{U}+k P\left(X<N \theta_{U}\right)-k P\left(X>N \theta_{U}\right)
\end{aligned}
$$

and

$$
\begin{aligned}
\operatorname{Var}\left[Z_{c s L}(X)\right] \approx & 4 k^{2} N p(1-p)+k^{2}\left[1-P\left(X=N \theta_{L}\right)\right]-k^{2}\left[P\left(X<N \theta_{L}\right)-P\left(X>N \theta_{L}\right)\right]^{2} \\
& +4 k^{2}\left[E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right)-E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right)\right]-4 k^{2} N p\left[P\left(X<N \theta_{L}\right)-P\left(X>N \theta_{L}\right)\right] \\
\operatorname{Var}\left[Z_{c s U}(X)\right] \approx & 4 k^{2} N p(1-p)+k^{2}\left[1-P\left(X=N \theta_{U}\right)\right]-k^{2}\left[P\left(X<N \theta_{U}\right)-P\left(X>N \theta_{U}\right)\right]^{2} \\
& +4 k^{2}\left[E\left(X 1_{\left\{X<N \theta_{U}\right\}}\right)-E\left(X 1_{\left\{X>N \theta_{U}\right\}}\right)\right]-4 k^{2} N p\left[P\left(X<N \theta_{U}\right)-P\left(X>N \theta_{U}\right)\right]
\end{aligned}
$$

The probabilities $P\left(X=N \theta_{L}\right), P\left(X<N \theta_{L}\right), P\left(X>N \theta_{L}\right), P\left(X=N \theta_{U}\right), P\left(X<N \theta_{U}\right)$, and $P\left(X>N \theta_{U}\right)$ and the truncated expectations $E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right), E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right), E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right)$, and $E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right)$ are approximated by assuming the normal-approximate distribution of $X, N(N p, N p(1-p))$. Letting $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal PDF and CDF, respectively, and defining $d_{L}$ and $d_{U}$ as

$$
\begin{aligned}
d_{L} & =\frac{N \theta_{L}-N p}{[N p(1-p)]^{\frac{1}{2}}} \\
d_{U} & =\frac{N \theta_{U}-N p}{[N p(1-p)]^{\frac{1}{2}}}
\end{aligned}
$$

the terms are computed as follows:

$$
\begin{aligned}
P\left(X=N \theta_{L}\right) & =0 \\
P\left(X=N \theta_{U}\right) & =0 \\
P\left(X<N \theta_{L}\right) & =\Phi\left(d_{L}\right) \\
P\left(X<N \theta_{U}\right) & =\Phi\left(d_{U}\right) \\
P\left(X>N \theta_{L}\right) & =1-\Phi\left(d_{L}\right) \\
P\left(X>N \theta_{U}\right) & =1-\Phi\left(d_{U}\right) \\
E\left(X 1_{\left\{X<N \theta_{L}\right\}}\right) & =N p \Phi\left(d_{L}\right)-[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{L}\right) \\
E\left(X 1_{\left\{X<N \theta_{U}\right\}}\right) & =N p \Phi\left(d_{U}\right)-[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{U}\right) \\
E\left(X 1_{\left\{X>N \theta_{L}\right\}}\right) & =N p\left[1-\Phi\left(d_{L}\right)\right]+[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{L}\right) \\
E\left(X 1_{\left\{X>N \theta_{U}\right\}}\right) & =N p\left[1-\Phi\left(d_{U}\right)\right]+[N p(1-p)]^{\frac{1}{2}} \phi\left(d_{U}\right)
\end{aligned}
$$

The mean and variance of $Z_{c s L}(X)$ and $Z_{c s U}(X)$ are thus approximated by

$$
\begin{aligned}
& \mu_{L}=k\left[2 N p-2 N \theta_{L}+2 \Phi\left(d_{L}\right)-1\right] \\
& \mu_{U}=k\left[2 N p-2 N \theta_{U}+2 \Phi\left(d_{U}\right)-1\right]
\end{aligned}
$$

and

$$
\begin{aligned}
\sigma_{L}^{2} & =4 k^{2}\left[N p(1-p)+\Phi\left(d_{L}\right)\left(1-\Phi\left(d_{L}\right)\right)-2(N p(1-p))^{\frac{1}{2}} \phi\left(d_{L}\right)\right] \\
\sigma_{U}^{2} & =4 k^{2}\left[N p(1-p)+\Phi\left(d_{U}\right)\left(1-\Phi\left(d_{U}\right)\right)-2(N p(1-p))^{\frac{1}{2}} \phi\left(d_{U}\right)\right]
\end{aligned}
$$

The approximate power is computed as

$$
\text { power }=\Phi\left(\frac{z_{\alpha}-\mu_{U}}{\sigma_{U}}\right)+\Phi\left(\frac{z_{\alpha}+\mu_{L}}{\sigma_{L}}\right)-1
$$

The approximate sample size is computed by numerically inverting the power formula.

## Wilson Score Confidence Interval for Binomial Proportion (CI=WILSON)

The two-sided $100(1-\alpha) \%$ confidence interval for $p$ is

$$
\frac{X+\frac{z_{1-\alpha / 2}^{2}}{2}}{N+z_{1-\alpha / 2}^{2}} \pm \frac{z_{1-\alpha / 2} N^{\frac{1}{2}}}{N+z_{1-\alpha / 2}^{2}}\left(\hat{p}(1-\hat{p})+\frac{z_{1-\alpha / 2}^{2}}{4 N}\right)^{\frac{1}{2}}
$$

So the half-width for the two-sided $100(1-\alpha) \%$ confidence interval is

$$
\text { half-width }=\frac{z_{1-\alpha / 2} N^{\frac{1}{2}}}{N+z_{1-\alpha / 2}^{2}}\left(\hat{p}(1-\hat{p})+\frac{z_{1-\alpha / 2}^{2}}{4 N}\right)^{\frac{1}{2}}
$$

$\operatorname{Prob}$ (Width) is calculated exactly by adding up the probabilities of observing each $X \in\{1, \ldots, N\}$ that produces a confidence interval whose half-width is at most a target value $h$ :

$$
\operatorname{Prob}(\text { Width })=\sum_{i=0}^{N} P(X=i) 1_{\text {half-width }}<h
$$

For references and more details about this and all other confidence intervals associated with the $\mathrm{CI}=$ option, see "Binomial Proportion" on page 2663 in Chapter 40, "The FREQ Procedure."

## Agresti-Coull "Add $k$ Successes and Failures" Confidence Interval for Binomial Proportion

 (CI=AGRESTICOULL)The two-sided $100(1-\alpha) \%$ confidence interval for $p$ is

$$
\frac{X+\frac{z_{1-\alpha / 2}^{2}}{2}}{N+z_{1-\alpha / 2}^{2}} \pm z_{1-\alpha / 2}\left(\frac{\frac{X+\frac{z_{1-\alpha / 2}^{2}}{N+z_{1-\alpha / 2}^{2}}}{2}\left(1-\frac{X+\frac{z_{1-\alpha / 2}^{2}}{N+z_{1-\alpha / 2}^{2}}}{2}\right.}{N+z_{1-\alpha / 2}^{2}}\right)^{\frac{1}{2}}
$$

So the half-width for the two-sided $100(1-\alpha) \%$ confidence interval is
$\operatorname{Prob}($ Width ) is calculated exactly by adding up the probabilities of observing each $X \in\{1, \ldots, N\}$ that produces a confidence interval whose half-width is at most a target value $h$ :

$$
\operatorname{Prob}(\text { Width })=\sum_{i=0}^{N} P(X=i) 1_{\text {half-width }}<h
$$

## Jeffreys Confidence Interval for Binomial Proportion (CI=JEFFREYS)

The two-sided $100(1-\alpha) \%$ confidence interval for $p$ is

$$
\left[L_{J}(X), U_{J}(X)\right]
$$

where

$$
L_{J}(X)= \begin{cases}0, & X=0 \\ \operatorname{Beta}_{\alpha / 2 ; X+1 / 2, N-X+1 / 2}, & X>0\end{cases}
$$

and

$$
U_{J}(X)= \begin{cases}\operatorname{Beta}_{1-\alpha / 2 ; X+1 / 2, N-X+1 / 2}, & X<N \\ 1, & X=N\end{cases}
$$

The half-width of this two-sided $100(1-\alpha) \%$ confidence interval is defined as half the width of the full interval:

$$
\text { half-width }=\frac{1}{2}\left(U_{J}(X)-L_{J}(X)\right)
$$

$\operatorname{Prob}($ Width ) is calculated exactly by adding up the probabilities of observing each $X \in\{1, \ldots, N\}$ that produces a confidence interval whose half-width is at most a target value $h$ :

$$
\operatorname{Prob}(\text { Width })=\sum_{i=0}^{N} P(X=i) 1_{\text {half-width }<h}
$$

## Exact Clopper-Pearson Confidence Interval for Binomial Proportion (CI=EXACT)

The two-sided $100(1-\alpha) \%$ confidence interval for $p$ is

$$
\left[L_{E}(X), U_{E}(X)\right]
$$

where

$$
L_{E}(X)= \begin{cases}0, & X=0 \\ \operatorname{Beta}_{\alpha / 2 ; X, N-X+1}, & X>0\end{cases}
$$

and

$$
U_{E}(X)= \begin{cases}\operatorname{Beta}_{1-\alpha / 2 ; X+1, N-X}, & X<N \\ 1, & X=N\end{cases}
$$

The half-width of this two-sided $100(1-\alpha) \%$ confidence interval is defined as half the width of the full interval:

$$
\text { half-width }=\frac{1}{2}\left(U_{E}(X)-L_{E}(X)\right)
$$

$\operatorname{Prob}($ Width ) is calculated exactly by adding up the probabilities of observing each $X \in\{1, \ldots, N\}$ that produces a confidence interval whose half-width is at most a target value $h$ :

$$
\operatorname{Prob}(\text { Width })=\sum_{i=0}^{N} P(X=i) 1_{\text {half-width }<h}
$$

## Wald Confidence Interval for Binomial Proportion (CI=WALD)

The two-sided $100(1-\alpha) \%$ confidence interval for $p$ is

$$
\hat{p} \quad \pm \quad z_{1-\alpha / 2}\left(\frac{\hat{p}(1-\hat{p})}{N}\right)^{\frac{1}{2}}
$$

So the half-width for the two-sided $100(1-\alpha) \%$ confidence interval is

$$
\text { half-width }=z_{1-\alpha / 2}\left(\frac{\hat{p}(1-\hat{p})}{N}\right)^{\frac{1}{2}}
$$

Prob(Width) is calculated exactly by adding up the probabilities of observing each $X \in\{1, \ldots, N\}$ that produces a confidence interval whose half-width is at most a target value $h$ :

$$
\operatorname{Prob}(\text { Width })=\sum_{i=0}^{N} P(X=i) 1_{\text {half-width }<h}
$$

Continuity-Corrected Wald Confidence Interval for Binomial Proportion (CI=WALD_CORRECT)
The two-sided $100(1-\alpha) \%$ confidence interval for $p$ is

$$
\hat{p} \pm\left[z_{1-\alpha / 2}\left(\frac{\hat{p}(1-\hat{p})}{N}\right)^{\frac{1}{2}}+\frac{1}{2 N}\right]
$$

So the half-width for the two-sided $100(1-\alpha) \%$ confidence interval is

$$
\text { half-width }=z_{1-\alpha / 2}\left(\frac{\hat{p}(1-\hat{p})}{N}\right)^{\frac{1}{2}}+\frac{1}{2 N}
$$

Prob(Width) is calculated exactly by adding up the probabilities of observing each $X \in\{1, \ldots, N\}$ that produces a confidence interval whose half-width is at most a target value $h$ :

$$
\operatorname{Prob}(\text { Width })=\sum_{i=0}^{N} P(X=i) 1_{\text {half-width }<h}
$$

## Analyses in the ONESAMPLEMEANS Statement

## One-Sample t Test (TEST=T)

The hypotheses for the one-sample $t$ test are

$$
\begin{aligned}
& H_{0}: \mu=\mu_{0} \\
& H_{1}: \begin{cases}\mu \neq \mu_{0}, & \text { two-sided } \\
\mu>\mu_{0} & \text { upper one-sided } \\
\mu<\mu_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

The test assumes normally distributed data and requires $N \geq 2$. The test statistics are

$$
\begin{aligned}
t & =N^{\frac{1}{2}}\left(\frac{\bar{x}-\mu_{0}}{s}\right) \sim t(N-1, \delta) \\
t^{2} & \sim F\left(1, N-1, \delta^{2}\right)
\end{aligned}
$$

where $\bar{x}$ is the sample mean, $s$ is the sample standard deviation, and

$$
\delta=N^{\frac{1}{2}}\left(\frac{\mu-\mu_{0}}{\sigma}\right)
$$

The test is
Reject $\quad H_{0} \quad$ if $\begin{cases}t^{2} \geq F_{1-\alpha}(1, N-1), & \text { two-sided } \\ t \geq t_{1-\alpha}(N-1), & \text { upper one-sided } \\ t \leq t_{\alpha}(N-1), & \text { lower one-sided }\end{cases}$

Exact power computations for $t$ tests are discussed in O'Brien and Muller (1993, Section 8.2), although not specifically for the one-sample case. The power is based on the noncentral $t$ and $F$ distributions:

$$
\text { power }= \begin{cases}P\left(F\left(1, N-1, \delta^{2}\right) \geq F_{1-\alpha}(1, N-1)\right), & \text { two-sided } \\ P\left(t(N-1, \delta) \geq t_{1-\alpha}(N-1)\right), & \text { upper one-sided } \\ P\left(t(N-1, \delta) \leq t_{\alpha}(N-1)\right), & \text { lower one-sided }\end{cases}
$$

Solutions for $N, \alpha$, and $\delta$ are obtained by numerically inverting the power equation. Closed-form solutions for other parameters, in terms of $\delta$, are as follows:

$$
\begin{aligned}
\mu & =\delta \sigma N^{-\frac{1}{2}}+\mu_{0} \\
\sigma & = \begin{cases}\delta^{-1} N^{\frac{1}{2}}\left(\mu-\mu_{0}\right), & |\delta|>0 \\
\text { undefined, }, & \text { otherwise }\end{cases}
\end{aligned}
$$

## One-Sample $t$ Test with Lognormal Data (TEST=T DIST=LOGNORMAL)

The lognormal case is handled by reexpressing the analysis equivalently as a normality-based test on the log-transformed data, by using properties of the lognormal distribution as discussed in Johnson, Kotz, and Balakrishnan (1994, Chapter 14). The approaches in the section "One-Sample $t$ Test (TEST=T)" on page 6399 then apply.

In contrast to the usual $t$ test on normal data, the hypotheses with lognormal data are defined in terms of geometric means rather than arithmetic means. This is because the transformation of a null arithmetic mean of lognormal data to the normal scale depends on the unknown coefficient of variation, resulting in an ill-defined hypothesis on the log-transformed data. Geometric means transform cleanly and are more natural for lognormal data.

The hypotheses for the one-sample $t$ test with lognormal data are

$$
\begin{aligned}
& H_{0}: \frac{\gamma}{\gamma_{0}}=1 \\
& H_{1}: \begin{cases}\frac{\gamma}{\gamma_{0}} \neq 1, & \text { two-sided } \\
\frac{y}{\gamma_{0}}>1, & \text { upper one-sided } \\
\frac{y}{\gamma_{0}}<1, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

Let $\mu^{\star}$ and $\sigma^{\star}$ be the (arithmetic) mean and standard deviation of the normal distribution of the logtransformed data. The hypotheses can be rewritten as follows:

$$
\begin{aligned}
& H_{0}: \mu^{\star}=\log \left(\gamma_{0}\right) \\
& H_{1}: \begin{cases}\mu^{\star} \neq \log \left(\gamma_{0}\right), & \text { two-sided } \\
\mu^{\star}>\log \left(\gamma_{0}\right), & \text { upper one-sided } \\
\mu^{\star}<\log \left(\gamma_{0}\right), & \text { lower one-sided }\end{cases}
\end{aligned}
$$

where $\mu^{\star}=\log (\gamma)$.

The test assumes lognormally distributed data and requires $N \geq 2$.
The power is

$$
\text { power }= \begin{cases}P\left(F\left(1, N-1, \delta^{2}\right) \geq F_{1-\alpha}(1, N-1)\right), & \text { two-sided } \\ P\left(t(N-1, \delta) \geq t_{1-\alpha}(N-1)\right), & \text { upper one-sided } \\ P\left(t(N-1, \delta) \leq t_{\alpha}(N-1)\right), & \text { lower one-sided }\end{cases}
$$

where

$$
\begin{aligned}
\delta & =N^{\frac{1}{2}}\left(\frac{\mu^{\star}-\log \left(\gamma_{0}\right)}{\sigma^{\star}}\right) \\
\sigma^{\star} & =\left[\log \left(\mathrm{CV}^{2}+1\right)\right]^{\frac{1}{2}}
\end{aligned}
$$

## Equivalence Test for Mean of Normal Data (TEST=EQUIV DIST=NORMAL)

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: \mu<\theta_{L} \text { or } \mu>\theta_{U} \\
& H_{1}: \theta_{L} \leq \mu \leq \theta_{U}
\end{aligned}
$$

The analysis is the two one-sided tests (TOST) procedure of Schuirmann (1987). The test assumes normally distributed data and requires $N \geq 2$. Phillips (1990) derives an expression for the exact power assuming a two-sample balanced design; the results are easily adapted to a one-sample design:

$$
\begin{aligned}
\text { power }= & Q_{N-1}\left(\left(-t_{1-\alpha}(N-1)\right), \frac{\mu-\theta_{U}}{\sigma N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\theta_{U}-\theta_{L}\right)}{2 \sigma N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)- \\
& Q_{N-1}\left(\left(t_{1-\alpha}(N-1)\right), \frac{\mu-\theta_{L}}{\sigma N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\theta_{U}-\theta_{L}\right)}{2 \sigma N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)
\end{aligned}
$$

where $Q .(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.

## Equivalence Test for Mean of Lognormal Data (TEST=EQUIV DIST=LOGNORMAL)

The lognormal case is handled by reexpressing the analysis equivalently as a normality-based test on the log-transformed data, by using properties of the lognormal distribution as discussed in Johnson, Kotz, and Balakrishnan (1994, Chapter 14). The approaches in the section "Equivalence Test for Mean of Normal Data (TEST=EQUIV DIST=NORMAL)" on page 6401 then apply.

In contrast to the additive equivalence test on normal data, the hypotheses with lognormal data are defined in terms of geometric means rather than arithmetic means. This is because the transformation of an arithmetic mean of lognormal data to the normal scale depends on the unknown coefficient of variation, resulting in an ill-defined hypothesis on the log-transformed data. Geometric means transform cleanly and are more natural for lognormal data.

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: \gamma \leq \theta_{L} \quad \text { or } \quad \gamma \geq \theta_{U} \\
& H_{1}: \theta_{L}<\gamma<\theta_{U}
\end{aligned}
$$

$$
\text { where } 0<\theta_{L}<\theta_{U}
$$

The analysis is the two one-sided tests (TOST) procedure of Schuirmann (1987) on the log-transformed data. The test assumes lognormally distributed data and requires $N \geq 2$. Diletti, Hauschke, and Steinijans (1991) derive an expression for the exact power assuming a crossover design; the results are easily adapted to a one-sample design:

$$
\begin{aligned}
\text { power }= & Q_{N-1}\left(\left(-t_{1-\alpha}(N-1)\right), \frac{\log (\gamma)-\log \left(\theta_{U}\right)}{\sigma^{\star} N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\log \left(\theta_{U}\right)-\log \left(\theta_{L}\right)\right)}{2 \sigma^{\star} N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)- \\
& Q_{N-1}\left(\left(t_{1-\alpha}(N-1)\right), \frac{\log (\gamma)-\log \left(\theta_{L}\right)}{\sigma^{\star} N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\log \left(\theta_{U}\right)-\log \left(\theta_{L}\right)\right)}{2 \sigma^{\star} N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)
\end{aligned}
$$

where

$$
\sigma^{\star}=\left[\log \left(\mathrm{CV}^{2}+1\right)\right]^{\frac{1}{2}}
$$

is the standard deviation of the log-transformed data, and $Q .(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.

## Confidence Interval for Mean (CI=T)

This analysis of precision applies to the standard $t$-based confidence interval:

$$
\begin{array}{lll}
{\left[\bar{x}-t_{1-\frac{\alpha}{2}}(N-1) \frac{s}{\sqrt{N}},\right.} & \left.\bar{x}+t_{1-\frac{\alpha}{2}}(N-1) \frac{s}{\sqrt{N}}\right], & \text { two-sided } \\
\left.\bar{x}-t_{1-\alpha}(N-1) \frac{s}{\sqrt{N}}, \quad \infty\right), & \text { upper one-sided } \\
\left(-\infty, \quad \bar{x}+t_{1-\alpha}(N-1) \frac{s}{\sqrt{N}}\right], & \text { lower one-sided }
\end{array}
$$

where $\bar{x}$ is the sample mean and $s$ is the sample standard deviation. The "half-width" is defined as the distance from the point estimate $\bar{x}$ to a finite endpoint,

$$
\text { half-width }= \begin{cases}t_{1-\frac{\alpha}{2}}(N-1) \frac{s}{\sqrt{N}}, & \text { two-sided } \\ t_{1-\alpha}(N-1) \frac{s}{\sqrt{N}}, & \text { one-sided }\end{cases}
$$

A "valid" conference interval captures the true mean. The exact probability of obtaining at most the target confidence interval half-width $h$, unconditional or conditional on validity, is given by Beal (1989):

$$
\begin{aligned}
& \operatorname{Pr}(\text { half-width } \leq h)= \begin{cases}P\left(\chi^{2}(N-1) \leq \frac{h^{2} N(N-1)}{\sigma^{2}\left(t_{1-\frac{\alpha}{2}}^{2}(N-1)\right)}\right), & \text { two-sided } \\
P\left(\chi^{2}(N-1) \leq \frac{h^{2} N(N-1)}{\sigma^{2}\left(t_{1-\alpha}^{2}(N-1)\right)}\right), & \text { one-sided }\end{cases} \\
& \operatorname{Pr}(\text { half-width } \leq h \mid \\
&\text { validity })
\end{aligned}=\left\{\begin{array}{cc}
\left(\frac{1}{1-\alpha}\right) 2\left[Q _ { N - 1 } \left(\left(t_{1-\frac{\alpha}{2}}(N-1)\right), 0 ;\right.\right. & \text { two-sided } \\
\left.\left.0, b_{1}\right)-Q_{N-1}\left(0,0 ; 0, b_{1}\right)\right], & \text { one-sided } \\
\left(\frac{1}{1-\alpha}\right) Q_{N-1}\left(\left(t_{1-\alpha}(N-1)\right), 0 ; 0, b_{1}\right), & \text { one }
\end{array}\right.
$$

where

$$
\begin{aligned}
b_{1} & =\frac{h(N-1)^{\frac{1}{2}}}{\sigma\left(t_{1-\frac{\alpha}{c}}^{c}(N-1)\right) N^{-\frac{1}{2}}} \\
c & =\text { number of sides }
\end{aligned}
$$

and $Q .(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.
A "quality" confidence interval is both sufficiently narrow (half-width $\leq h$ ) and valid:

$$
\begin{aligned}
\operatorname{Pr}(\text { quality }) & =\operatorname{Pr}(\text { half-width } \leq h \text { and validity }) \\
& =\operatorname{Pr}(\text { half-width } \leq h \mid \text { validity })(1-\alpha)
\end{aligned}
$$

## Analyses in the ONEWAYANOVA Statement

## One-Degree-of-Freedom Contrast (TEST=CONTRAST)

The hypotheses are

$$
\begin{aligned}
& H_{0}: c_{1} \mu_{1}+\cdots+c_{G} \mu_{G}=c_{0} \\
& H_{1}: \begin{cases}c_{1} \mu_{1}+\cdots+c_{G} \mu_{G} \neq c_{0}, & \text { two-sided } \\
c_{1} \mu_{1}+\cdots+c_{G} \mu_{G}>c_{0}, & \text { upper one-sided } \\
c_{1} \mu_{1}+\cdots+c_{G} \mu_{G}<c_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

where $G$ is the number of groups, $\left\{c_{1}, \ldots, c_{G}\right\}$ are the contrast coefficients, and $c_{0}$ is the null contrast value. The test is the usual $F$ test for a contrast in one-way ANOVA. It assumes normal data with common group variances and requires $N \geq G+1$ and $n_{i} \geq 1$.

O'Brien and Muller (1993, Section 8.2.3.2) give the exact power as

$$
\text { power }= \begin{cases}P\left(F\left(1, N-G, \delta^{2}\right) \geq F_{1-\alpha}(1, N-G)\right), & \text { two-sided } \\ P\left(t(N-G, \delta) \geq t_{1-\alpha}(N-G)\right), & \text { upper one-sided } \\ P\left(t(N-G, \delta) \leq t_{\alpha}(N-G)\right), & \text { lower one-sided }\end{cases}
$$

where

$$
\delta=N^{\frac{1}{2}}\left(\frac{\sum_{i=1}^{G} c_{i} \mu_{i}-c_{0}}{\sigma\left(\sum_{i=1}^{G} \frac{c_{i}^{2}}{w_{i}}\right)^{\frac{1}{2}}}\right)
$$

## Overall F Test (TEST=OVERALL)

The hypotheses are

$$
\begin{aligned}
& H_{0}: \mu_{1}=\mu_{2}=\cdots=\mu_{G} \\
& H_{1}: \mu_{i} \neq \mu_{j} \text { for some } i, j
\end{aligned}
$$

where $G$ is the number of groups.

The test is the usual overall $F$ test for equality of means in one-way ANOVA. It assumes normal data with common group variances and requires $N \geq G+1$ and $n_{i} \geq 1$.
O'Brien and Muller (1993, Section 8.2.3.1) give the exact power as

$$
\text { power }=P\left(F(G-1, N-G, \lambda) \geq F_{1-\alpha}(G-1, N-G)\right)
$$

where the noncentrality is

$$
\lambda=N\left(\frac{\sum_{i=1}^{G} w_{i}\left(\mu_{i}-\bar{\mu}\right)^{2}}{\sigma^{2}}\right)
$$

and

$$
\bar{\mu}=\sum_{i=1}^{G} w_{i} \mu_{i}
$$

## Analyses in the PAIREDFREQ Statement

Overview of Conditional McNemar Tests
Notation:

|  |  | Case |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Failure | Success |  |
| Control | Failure | $n_{00}$ | $n_{01}$ | $n_{0}$. |
|  | Success | $n_{10}$ | $n_{11}$ | $n_{1}$. |
|  |  | $n \cdot 0$ | $n \cdot 1$ | $N$ |

$$
\begin{aligned}
n_{00} & =\#\{\text { control=failure, case=failure }\} \\
n_{01} & =\#\{\text { control=failure, case=success }\} \\
n_{10} & =\#\{\text { control=success, case=failure }\} \\
n_{11} & =\#\{\text { control=success, case=success }\} \\
N & =n_{00}+n_{01}+n_{10}+n_{11} \\
n_{D} & =n_{01}+n_{10} \equiv \# \text { discordant pairs } \\
\hat{\pi}_{i j} & =\frac{n_{i j}}{N} \\
\pi_{i j} & =\text { theoretical population value of } \hat{\pi}_{i j} \\
\pi_{1 \cdot} & =\pi_{10}+\pi_{11} \\
\pi_{\cdot 1} & =\pi_{01}+\pi_{11} \\
\phi & =\text { Corr(control observation, case observation) } \quad \text { (within a pair) } \\
\text { DPR } & =\text { "discordant proportion ratio" }=\frac{\pi_{01}}{\pi_{10}}
\end{aligned}
$$

$\mathrm{DPR}_{0}=$ null DPR

Power formulas are given here in terms of the discordant proportions $\pi_{10}$ and $\pi_{01}$. If the input is specified in terms of $\left\{\pi_{1} ., \pi_{\cdot 1}, \phi\right\}$, then it can be converted into values for $\left\{\pi_{10}, \pi_{01}\right\}$ as follows:

$$
\begin{aligned}
& \pi_{01}=\pi_{\cdot 1}\left(1-\pi_{1} \cdot\right)-\phi\left(\left(1-\pi_{1} \cdot\right) \pi_{1} \cdot\left(1-\pi_{\cdot 1}\right) \pi_{\cdot 1}\right)^{\frac{1}{2}} \\
& \pi_{10}=\pi_{01}+\pi_{1} \cdot-\pi \cdot 1
\end{aligned}
$$

All McNemar tests covered in PROC POWER are conditional, meaning that $n_{D}$ is assumed fixed at its observed value.

For the usual $\mathrm{DPR}_{0}=1$, the hypotheses are

$$
\begin{aligned}
& H_{0}: \pi \cdot 1=\pi_{1} . \\
& H_{1}: \begin{cases}\pi \cdot 1 \neq \pi_{1}, & \text { two-sided } \\
\pi \cdot 1>\pi_{1}, & \text { upper one-sided } \\
\pi \cdot 1<\pi_{1}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

The test statistic for both tests covered in PROC POWER (DIST=EXACT_COND and DIST=NORMAL) is the McNemar statistic $Q_{M}$, which has the following form when $\mathrm{DPR}_{0}=1$ :

$$
Q_{M_{0}}=\frac{\left(n_{01}-n_{10}\right)^{2}}{n_{01}+n_{10}}
$$

For the conditional McNemar tests, this is equivalent to the square of the $Z(X)$ statistic for the test of a single proportion (normal approximation to binomial), where the proportion is $\frac{\pi_{01}}{\pi_{01}+\pi_{10}}$, the null is 0.5 , and " $N$ " is $n_{D}$ (see, for example, Schork and Williams 1980):

$$
\begin{aligned}
Z(X) & =\frac{n_{01}-n_{D}(0.5)}{\left[n_{D} 0.5(1-0.5)\right]^{\frac{1}{2}}} \dot{\sim}\left(\frac{n_{D}^{\frac{1}{2}}\left(\frac{\pi_{01}}{\pi_{01}+\pi_{10}}-0.5\right)}{[0.5(1-0.5)]^{\frac{1}{2}}}, \frac{\frac{\pi_{01}}{\pi_{01}+\pi_{10}}\left(1-\frac{\pi_{01}}{\pi_{01}+\pi_{10}}\right)}{0.5(1-0.5)}\right) \\
& =\frac{n_{01}-\left(n_{01}+n_{10}\right)(0.5)}{\left[\left(n_{01}+n_{10}\right) 0.5(1-0.5)\right]^{\frac{1}{2}}} \\
& =\frac{n_{01}-n_{10}}{\left[n_{01}+n_{10}\right]^{\frac{1}{2}}} \\
& =\sqrt{Q_{M_{0}}}
\end{aligned}
$$

This can be generalized to a custom null for $\frac{\pi_{01}}{\pi_{01}+\pi_{10}}$, which is equivalent to specifying a custom null DPR:

$$
\left[\frac{\pi_{01}}{\pi_{01}+\pi_{10}}\right]_{0} \equiv\left[\frac{1}{\left.1+\frac{1}{\frac{\pi_{01}}{\pi_{10}}}\right]_{0} \equiv \frac{1}{1+\frac{1}{\mathrm{DPR}_{0}}} \text {. }}\right.
$$

So, a conditional McNemar test (asymptotic or exact) with a custom null is equivalent to the test of a single proportion $p_{1} \equiv \frac{\pi_{01}}{\pi_{01}+\pi_{10}}$ with a null value $p_{0} \equiv \frac{1}{1+\frac{1}{\mathrm{DPR}_{0}}}$, with a sample size of $n_{D}$ :

$$
\begin{aligned}
& H_{0}: p_{1}=p_{0} \\
& H_{1}: \begin{cases}p_{1} \neq p_{0}, & \text { two-sided } \\
p_{1}>p_{0}, & \text { one-sided } \mathrm{U} \\
p_{1}<p_{0}, & \text { one-sided } \mathrm{L}\end{cases}
\end{aligned}
$$

which is equivalent to

$$
\begin{aligned}
& H_{0}: \mathrm{DPR}=\mathrm{DPR}_{0} \\
& H_{1}: \begin{cases}\mathrm{DPR} \neq \mathrm{DPR}_{0}, & \text { two-sided } \\
\mathrm{DPR}>\mathrm{DPR}_{0}, & \text { one-sided } \mathrm{U} \\
\mathrm{DPR}<\mathrm{DPR}_{0}, & \text { one-sided } \mathrm{L}\end{cases}
\end{aligned}
$$

The general form of the test statistic is thus

$$
Q_{M}=\frac{\left(n_{01}-n_{D} p_{0}\right)^{2}}{n_{D} p_{0}\left(1-p_{0}\right)}
$$

The two most common conditional McNemar tests assume either the exact conditional distribution of $Q_{M}$ (covered by the DIST=EXACT_COND analysis) or a standard normal distribution for $Q_{M}$ (covered by the DIST=NORMAL analysis).

## McNemar Exact Conditional Test (TEST=MCNEMAR DIST=EXACT_COND)

For DIST=EXACT_COND, the power is calculated assuming that the test is conducted by using the exact conditional distribution of $Q_{M}$ (conditional on $n_{D}$ ). The power is calculated by first computing the conditional power for each possible $n_{D}$. The unconditional power is computed as a weighted average over all possible outcomes of $n_{D}$ :

$$
\text { power }=\sum_{n_{D}=0}^{N} P\left(n_{D}\right) P\left(\text { Reject } p_{1}=p_{0} \mid n_{D}\right)
$$

where $n_{D} \sim \operatorname{Bin}\left(\pi_{01}+\pi_{10}, N\right)$, and $P\left(\right.$ Reject $\left.p_{1}=p_{0} \mid n_{D}\right)$ is calculated by using the exact method in the section "Exact Test of a Binomial Proportion (TEST=EXACT)" on page 6381.

The achieved significance level, reported as "Actual Alpha" in the analysis, is computed in the same way except by using the actual alpha of the one-sample test in place of its power:

$$
\text { actual alpha }=\sum_{n_{D}=0}^{N} P\left(n_{D}\right) \alpha^{\star}\left(p_{1}, p_{0} \mid n_{D}\right)
$$

where $\alpha^{\star}\left(p_{1}, p_{0} \mid n_{D}\right)$ is the actual alpha calculated by using the exact method in the section "Exact Test of a Binomial Proportion (TEST=EXACT)" on page 6381 with proportion $p_{1}$, null $p_{0}$, and sample size $n_{D}$.

## McNemar Normal Approximation Test (TEST=MCNEMAR DIST=NORMAL)

For DIST=NORMAL, power is calculated assuming the test is conducted by using the normal-approximate distribution of $Q_{M}$ (conditional on $n_{D}$ ).

For the METHOD=EXACT option, the power is calculated in the same way as described in the section "McNemar Exact Conditional Test (TEST=MCNEMAR DIST=EXACT_COND)" on page 6406, except that $P\left(\right.$ Reject $\left.p_{1}=p_{0} \mid n_{D}\right)$ is calculated by using the exact method in the section "z Test for Binomial Proportion Using Null Variance (TEST=Z VAREST=NULL)" on page 6382. The achieved significance level is calculated in the same way as described at the end of the section "McNemar Exact Conditional Test (TEST=MCNEMAR DIST=EXACT_COND)" on page 6406.

For the METHOD=MIETTINEN option, approximate sample size for the one-sided cases is computed according to equation (5.6) in Miettinen (1968):

$$
N=\frac{\left\{z_{1-\alpha}\left(p_{10}+p_{01}\right)+z_{\text {power }}\left[\left(p_{10}+p_{01}\right)^{2}-\frac{1}{4}\left(p_{01}-p_{10}\right)^{2}\left(3+p_{10}+p_{01}\right)\right]^{\frac{1}{2}}\right\}^{2}}{\left(p_{10}+p_{01}\right)\left(p_{01}-p_{10}\right)^{2}}
$$

Approximate power for the one-sided cases is computed by solving the sample size equation for power, and approximate power for the two-sided case follows easily by summing the one-sided powers each at $\alpha / 2$ :

$$
\text { power }= \begin{cases}\Phi\left(\frac{\left(p_{01}-p_{10}\right)\left[N\left(p_{10}+p_{01}\right)\right]^{\frac{1}{2}}-z_{1-\alpha}\left(p_{10}+p_{01}\right)}{\left[\left(p_{10}+p_{01}\right)^{2}-\frac{1}{4}\left(p_{01}-p_{10}\right)^{2}\left(3+p_{10}+p_{01}\right)\right]^{\frac{1}{2}}}\right), & \text { upper one-sided } \\ \Phi\left(\frac{-\left(p_{01}-p_{10}\right)\left[N\left(p_{10}+p_{01}\right)\right]^{\frac{1}{2}}-z_{1-\alpha}\left(p_{10}+p_{01}\right)}{\left[\left(p_{10}+p_{01}\right)^{2}-\frac{1}{4}\left(p_{01}-p_{10}\right)^{2}\left(3+p_{10}+p_{01}\right)\right]^{\frac{1}{2}}}\right), & \text { lower one-sided } \\ \Phi\left(\frac{\left(p_{01}-p_{10}\right)\left[N\left(p_{10}+p_{01}\right)\right]^{\frac{1}{2}}-z_{1-\frac{\alpha}{2}}\left(p_{10}+p_{01}\right)}{\left[\left(p_{10}+p_{01}\right)^{2}-\frac{1}{4}\left(p_{01}-p_{10}\right)^{2}\left(3+p_{10}+p_{01}\right)\right]^{\frac{1}{2}}}\right)+ & \\ \Phi\left(\frac{-\left(p_{01}-p_{10}\right)\left[N\left(p_{10}+p_{01}\right)\right]^{\frac{1}{2}-z_{1-\frac{\alpha}{2}}}\left(p_{10}+p_{01}\right)}{\left[\left(p_{10}+p_{01}\right)^{2}-\frac{1}{4}\left(p_{01}-p_{10}\right)^{2}\left(3+p_{10}+p_{01}\right)\right]^{\frac{1}{2}}}\right), & \text { two-sided }\end{cases}
$$

The two-sided solution for $N$ is obtained by numerically inverting the power equation.
In general, compared to METHOD=CONNOR, the METHOD=MIETTINEN approximation tends to be slightly more accurate but can be slightly anticonservative in the sense of underestimating sample size and overestimating power (Lachin 1992, p. 1250).

For the METHOD=CONNOR option, approximate sample size for the one-sided cases is computed according to equation (3) in Connor (1987):

$$
N=\frac{\left\{z_{1-\alpha}\left(p_{10}+p_{01}\right)^{\frac{1}{2}}+z_{\text {power }}\left[p_{10}+p_{01}-\left(p_{01}-p_{10}\right)^{2}\right]^{\frac{1}{2}}\right\}^{2}}{\left(p_{01}-p_{10}\right)^{2}}
$$

Approximate power for the one-sided cases is computed by solving the sample size equation for power, and approximate power for the two-sided case follows easily by summing the one-sided powers each at $\alpha / 2$ :

$$
\text { power }= \begin{cases}\Phi\left(\frac{\left(p_{01}-p_{10}\right) N^{\frac{1}{2}}-z_{1-\alpha}\left(p_{10}+p_{01}\right)^{\frac{1}{2}}}{\left.\left[p_{10}+p_{01}-\left(p_{01}-p_{10}\right)^{2}\right]^{\frac{1}{2}}\right),}\right. & \text { upper one-sided } \\ \Phi\left(\frac{-\left(p_{01}-p_{10}\right) N^{\frac{1}{2}}-z_{1-\alpha}\left(p_{10}+p_{01}\right)^{\frac{1}{2}}}{\left[p_{10}+p_{01}-\left(p_{01}-p_{10}\right)^{2}\right]^{\frac{1}{2}}}\right), & \text { lower one-sided } \\ \Phi\left(\frac{\left(p_{01}-p_{10}\right) N^{\frac{1}{2}}-z_{1-\frac{\alpha}{2}}\left(p_{10}+p_{01}\right)^{\frac{1}{2}}}{\left[p_{10}+p_{01}-\left(p_{01}-p_{10}\right)^{2}\right]^{\frac{1}{2}}}\right)+ & \\ \Phi\left(\frac{-\left(p_{01}-p_{10}\right) N^{\frac{1}{2}}-z_{1-\frac{\alpha}{2}}\left(p_{10}+p_{01}\right)^{\frac{1}{2}}}{\left[p_{10}+p_{01}-\left(p_{01}-p_{10}\right)^{2}\right]^{\frac{1}{2}}}\right), & \text { two-sided }\end{cases}
$$

The two-sided solution for $N$ is obtained by numerically inverting the power equation.
In general, compared to METHOD=MIETTINEN, the METHOD=CONNOR approximation tends to be slightly less accurate but slightly conservative in the sense of overestimating sample size and underestimating power (Lachin 1992, p. 1250).

## Analyses in the PAIREDMEANS Statement

## Paired $t$ Test (TEST=DIFF)

The hypotheses for the paired $t$ test are

$$
\begin{aligned}
& H_{0}: \mu_{\text {diff }}=\mu_{0} \\
& H_{1}: \begin{cases}\mu_{\text {diff }} \neq \mu_{0}, & \text { two-sided } \\
\mu_{\text {diff }}>\mu_{0}, & \text { upper one-sided } \\
\mu_{\text {diff }}<\mu_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

The test assumes normally distributed data and requires $N \geq 2$. The test statistics are

$$
\begin{aligned}
t & =N^{\frac{1}{2}}\left(\frac{\bar{d}-\mu_{0}}{s_{d}}\right) \sim t(N-1, \delta) \\
t^{2} & \sim F\left(1, N-1, \delta^{2}\right)
\end{aligned}
$$

where $\bar{d}$ and $s_{\boldsymbol{d}}$ are the sample mean and standard deviation of the differences and

$$
\delta=N^{\frac{1}{2}}\left(\frac{\mu_{\text {diff }}-\mu_{0}}{\sigma_{\text {diff }}}\right)
$$

and

$$
\sigma_{\text {diff }}=\left(\sigma_{1}^{2}+\sigma_{2}^{2}-2 \rho \sigma_{1} \sigma_{2}\right)^{\frac{1}{2}}
$$

The test is

$$
\text { Reject } \quad H_{0} \quad \text { if } \begin{cases}t^{2} \geq F_{1-\alpha}(1, N-1), & \text { two-sided } \\ t \geq t_{1-\alpha}(N-1), & \text { upper one-sided } \\ t \leq t_{\alpha}(N-1), & \text { lower one-sided }\end{cases}
$$

Exact power computations for $t$ tests are given in O'Brien and Muller (1993, Section 8.2.2):

$$
\text { power }= \begin{cases}P\left(F\left(1, N-1, \delta^{2}\right) \geq F_{1-\alpha}(1, N-1)\right), & \text { two-sided } \\ P\left(t(N-1, \delta) \geq t_{1-\alpha}(N-1)\right), & \text { upper one-sided } \\ P\left(t(N-1, \delta) \leq t_{\alpha}(N-1)\right), & \text { lower one-sided }\end{cases}
$$

## Paired $t$ Test for Mean Ratio with Lognormal Data (TEST=RATIO)

The lognormal case is handled by reexpressing the analysis equivalently as a normality-based test on the log-transformed data, by using properties of the lognormal distribution as discussed in Johnson, Kotz, and Balakrishnan (1994, Chapter 14). The approaches in the section "Paired $t$ Test (TEST=DIFF)" on page 6408 then apply.

In contrast to the usual $t$ test on normal data, the hypotheses with lognormal data are defined in terms of geometric means rather than arithmetic means.

The hypotheses for the paired $t$ test with lognormal pairs $\left\{Y_{1}, Y_{2}\right\}$ are

$$
\begin{aligned}
& H_{0}: \frac{\gamma_{2}}{\gamma_{1}}=\gamma_{0} \\
& H_{1}: \begin{cases}\frac{\gamma_{2}}{\gamma_{1}} \neq \gamma_{0}, & \text { two-sided } \\
\frac{\gamma_{2}}{\gamma_{1}}>\gamma_{0}, & \text { upper one-sided } \\
\frac{\gamma_{2}}{\gamma_{1}}<\gamma_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

Let $\mu_{1}^{\star}, \mu_{2}^{\star}, \sigma_{1}^{\star}, \sigma_{2}^{\star}$, and $\rho^{\star}$ be the (arithmetic) means, standard deviations, and correlation of the bivariate normal distribution of the $\log$-transformed data $\left\{\log Y_{1}, \log Y_{2}\right\}$. The hypotheses can be rewritten as follows:

$$
\begin{aligned}
& H_{0}: \mu_{2}^{\star}-\mu_{1}^{\star}=\log \left(\gamma_{0}\right) \\
& H_{1}: \begin{cases}\mu_{2}^{\star}-\mu_{1}^{\star} \neq \log \left(\gamma_{0}\right), & \text { two-sided } \\
\mu_{2}^{\star}-\mu_{1}^{\star}>\log \left(\gamma_{0}\right), & \text { upper one-sided } \\
\mu_{2}^{\star}-\mu_{1}^{\star}<\log \left(\gamma_{0}\right), & \text { lower one-sided }\end{cases}
\end{aligned}
$$

where

$$
\begin{aligned}
\mu_{1}^{\star} & =\log \gamma_{1} \\
\mu_{2}^{\star} & =\log \gamma_{2} \\
\sigma_{1}^{\star} & =\left[\log \left(\mathrm{CV}_{1}^{2}+1\right)\right]^{\frac{1}{2}} \\
\sigma_{2}^{\star} & =\left[\log \left(\mathrm{CV}_{2}^{2}+1\right)\right]^{\frac{1}{2}} \\
\rho^{\star} & =\frac{\log \left\{\rho \mathrm{CV}_{1} \mathrm{CV}_{2}+1\right\}}{\sigma_{1}^{\star} \sigma_{2}^{\star}}
\end{aligned}
$$

and $\mathrm{CV}_{1}, \mathrm{CV}_{2}$, and $\rho$ are the coefficients of variation and the correlation of the original untransformed pairs $\left\{Y_{1}, Y_{2}\right\}$. The conversion from $\rho$ to $\rho^{\star}$ is given by equation (44.36) on page 27 of Kotz, Balakrishnan, and Johnson (2000) and due to Jones and Miller (1966).

The valid range of $\rho$ is restricted to ( $\rho_{L}, \rho_{U}$ ), where

$$
\begin{aligned}
& \rho_{L}=\frac{\exp \left(-\left[\log \left(\mathrm{CV}_{1}^{2}+1\right) \log \left(\mathrm{CV}_{2}^{2}+1\right)\right]^{\frac{1}{2}}\right)-1}{\mathrm{CV}_{1} \mathrm{CV}_{2}} \\
& \rho_{U}=\frac{\exp \left(\left[\log \left(\mathrm{CV}_{1}^{2}+1\right) \log \left(\mathrm{CV}_{2}^{2}+1\right)\right]^{\frac{1}{2}}\right)-1}{\mathrm{CV}_{1} \mathrm{CV}_{2}}
\end{aligned}
$$

These bounds are computed from equation (44.36) on page 27 of Kotz, Balakrishnan, and Johnson (2000) by observing that $\rho$ is a monotonically increasing function of $\rho^{\star}$ and plugging in the values $\rho^{\star}=-1$ and $\rho^{\star}=1$. Note that when the coefficients of variation are equal ( $\mathrm{CV}_{1}=\mathrm{CV}_{2}=\mathrm{CV}$ ), the bounds simplify to

$$
\begin{aligned}
& \rho_{L}=\frac{-1}{\mathrm{CV}^{2}+1} \\
& \rho_{U}=1
\end{aligned}
$$

The test assumes lognormally distributed data and requires $N \geq 2$. The power is

$$
\text { power }= \begin{cases}P\left(F\left(1, N-1, \delta^{2}\right) \geq F_{1-\alpha}(1, N-1)\right), & \text { two-sided } \\ P\left(t(N-1, \delta) \geq t_{1-\alpha}(N-1)\right), & \text { upper one-sided } \\ P\left(t(N-1, \delta) \leq t_{\alpha}(N-1)\right), & \text { lower one-sided }\end{cases}
$$

where

$$
\delta=N^{\frac{1}{2}}\left(\frac{\mu_{1}^{\star}-\mu_{2}^{\star}-\log \left(\gamma_{0}\right)}{\sigma^{\star}}\right)
$$

and

$$
\sigma^{\star}=\left(\sigma_{1}^{\star 2}+\sigma_{2}^{\star 2}-2 \rho^{\star} \sigma_{1}^{\star} \sigma_{2}^{\star}\right)^{\frac{1}{2}}
$$

## Additive Equivalence Test for Mean Difference with Normal Data (TEST=EQUIV_DIFF)

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: \mu_{\text {diff }}<\theta_{L} \quad \text { or } \quad \mu_{\text {diff }}>\theta_{U} \\
& H_{1}: \theta_{L} \leq \mu_{\text {diff }} \leq \theta_{U}
\end{aligned}
$$

The analysis is the two one-sided tests (TOST) procedure of Schuirmann (1987). The test assumes normally distributed data and requires $N \geq 2$. Phillips (1990) derives an expression for the exact power assuming a two-sample balanced design; the results are easily adapted to a paired design:

$$
\begin{aligned}
\text { power }= & Q_{N-1}\left(\left(-t_{1-\alpha}(N-1)\right), \frac{\mu_{\text {diff }}-\theta_{U}}{\sigma_{\text {diff }} N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\theta_{U}-\theta_{L}\right)}{2 \sigma_{\text {diff }} N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)- \\
& Q_{N-1}\left(\left(t_{1-\alpha}(N-1)\right), \frac{\mu_{\text {diff }}-\theta_{L}}{\sigma_{\text {diff }} N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\theta_{U}-\theta_{L}\right)}{2 \sigma_{\text {diff }} N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)
\end{aligned}
$$

where

$$
\sigma_{\mathrm{diff}}=\left(\sigma_{1}^{2}+\sigma_{2}^{2}-2 \rho \sigma_{1} \sigma_{2}\right)^{\frac{1}{2}}
$$

and $Q .(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.

## Multiplicative Equivalence Test for Mean Ratio with Lognormal Data (TEST=EQUIV_RATIO)

The lognormal case is handled by reexpressing the analysis equivalently as a normality-based test on the log-transformed data, by using properties of the lognormal distribution as discussed in Johnson, Kotz, and Balakrishnan (1994, Chapter 14). The approaches in the section "Additive Equivalence Test for Mean Difference with Normal Data (TEST=EQUIV_DIFF)" on page 6410 then apply.

In contrast to the additive equivalence test on normal data, the hypotheses with lognormal data are defined in terms of geometric means rather than arithmetic means.

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: \frac{\gamma_{T}}{\gamma_{R}} \leq \theta_{L} \quad \text { or } \quad \frac{\gamma_{T}}{\gamma_{R}} \geq \theta_{U} \\
& H_{1}: \theta_{L}<\frac{\gamma_{T}}{\gamma_{R}}<\theta_{U}
\end{aligned}
$$

where $0<\theta_{L}<\theta_{U}$
The analysis is the two one-sided tests (TOST) procedure of Schuirmann (1987) on the log-transformed data. The test assumes lognormally distributed data and requires $N \geq 2$. Diletti, Hauschke, and Steinijans (1991) derive an expression for the exact power assuming a crossover design; the results are easily adapted to a paired design:

$$
\begin{aligned}
\text { power }= & Q_{N-1}\left(\left(-t_{1-\alpha}(N-1)\right), \frac{\log \left(\frac{\gamma_{T}}{\gamma_{R}}\right)-\log \left(\theta_{U}\right)}{\sigma^{\star} N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\log \left(\theta_{U}\right)-\log \left(\theta_{L}\right)\right)}{2 \sigma^{\star} N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)- \\
& Q_{N-1}\left(\left(t_{1-\alpha}(N-1)\right), \frac{\log \left(\frac{\gamma_{T}}{\gamma_{R}}\right)-\log \left(\theta_{L}\right)}{\sigma^{\star} N^{-\frac{1}{2}}} ; 0, \frac{(N-1)^{\frac{1}{2}}\left(\log \left(\theta_{U}\right)-\log \left(\theta_{L}\right)\right)}{2 \sigma^{\star} N^{-\frac{1}{2}}\left(t_{1-\alpha}(N-1)\right)}\right)
\end{aligned}
$$

where $\sigma^{\star}$ is the standard deviation of the differences between the log-transformed pairs (in other words, the standard deviation of $\log \left(Y_{T}\right)-\log \left(Y_{R}\right)$, where $Y_{T}$ and $Y_{R}$ are observations from the treatment and reference, respectively), computed as

$$
\begin{aligned}
\sigma^{\star} & =\left(\sigma_{R}^{\star 2}+\sigma_{T}^{\star 2}-2 \rho^{\star} \sigma_{R}^{\star} \sigma_{T}^{\star}\right)^{\frac{1}{2}} \\
\sigma_{R}^{\star} & =\left[\log \left(\mathrm{CV}_{R}^{2}+1\right)\right]^{\frac{1}{2}} \\
\sigma_{T}^{\star} & =\left[\log \left(\mathrm{CV}_{T}^{2}+1\right)\right]^{\frac{1}{2}} \\
\rho^{\star} & =\frac{\log \left\{\rho \mathrm{CV}_{R} \mathrm{CV}_{T}+1\right\}}{\sigma_{R}^{\star} \sigma_{T}^{\star}}
\end{aligned}
$$

where $\mathrm{CV}_{R}, \mathrm{CV}_{T}$, and $\rho$ are the coefficients of variation and the correlation of the original untransformed pairs $\left\{Y_{T}, Y_{R}\right\}$, and $Q .(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function. The conversion from $\rho$ to $\rho^{\star}$ is given by equation (44.36) on page 27 of Kotz, Balakrishnan, and Johnson (2000) and due to Jones and Miller (1966), and Owen's Q function is defined in the section "Common Notation" on page 6373.

The valid range of $\rho$ is restricted to ( $\rho_{L}, \rho_{U}$ ), where

$$
\begin{aligned}
\rho_{L} & =\frac{\exp \left(-\left[\log \left(\mathrm{CV}_{R}^{2}+1\right) \log \left(\mathrm{CV}_{T}^{2}+1\right)\right]^{\frac{1}{2}}\right)-1}{\mathrm{CV}_{R} \mathrm{CV}_{T}} \\
\rho_{U} & =\frac{\exp \left(\left[\log \left(\mathrm{CV}_{R}^{2}+1\right) \log \left(\mathrm{CV}_{T}^{2}+1\right)\right]^{\frac{1}{2}}\right)-1}{\mathrm{CV}_{R} \mathrm{CV}_{T}}
\end{aligned}
$$

These bounds are computed from equation (44.36) on page 27 of Kotz, Balakrishnan, and Johnson (2000) by observing that $\rho$ is a monotonically increasing function of $\rho^{\star}$ and plugging in the values $\rho^{\star}=-1$ and $\rho^{\star}=1$. Note that when the coefficients of variation are equal $\left(\mathrm{CV}_{R}=\mathrm{CV}_{T}=\mathrm{CV}\right)$, the bounds simplify to

$$
\begin{aligned}
& \rho_{L}=\frac{-1}{\mathrm{CV}^{2}+1} \\
& \rho_{U}=1
\end{aligned}
$$

## Confidence Interval for Mean Difference (CI=DIFF)

This analysis of precision applies to the standard $t$-based confidence interval:

$$
\begin{array}{ll}
{\left[\bar{d}-t_{1-\frac{\alpha}{2}}(N-1) \frac{s_{d}}{\sqrt{N}}, \quad \bar{d}+t_{1-\frac{\alpha}{2}}(N-1) \frac{s_{d}}{\sqrt{N}}\right],} & \text { two-sided } \\
{\left[\bar{d}-t_{1-\alpha}(N-1) \frac{s_{d}}{\sqrt{N}}, \quad \infty\right),} & \text { upper one-sided } \\
\left(-\infty, \quad \bar{d}+t_{1-\alpha}(N-1) \frac{s_{d}}{\sqrt{N}}\right], & \text { lower one-sided }
\end{array}
$$

where $\bar{d}$ and $s_{d}$ are the sample mean and standard deviation of the differences. The "half-width" is defined as the distance from the point estimate $\bar{d}$ to a finite endpoint,

$$
\text { half-width }= \begin{cases}t_{1-\frac{\alpha}{2}}(N-1) \frac{s_{d}}{\sqrt{N}}, & \text { two-sided } \\ t_{1-\alpha}(N-1) \frac{s_{d}}{\sqrt{N}}, & \text { one-sided }\end{cases}
$$

A "valid" conference interval captures the true mean difference. The exact probability of obtaining at most the target confidence interval half-width $h$, unconditional or conditional on validity, is given by Beal (1989):

$$
\begin{aligned}
& \operatorname{Pr}(\text { half-width } \leq h)= \begin{cases}P\left(\chi^{2}(N-1) \leq \frac{h^{2} N(N-1)}{\sigma_{\text {diff }}^{2}\left(t_{1-\frac{\alpha}{2}}^{2}(N-1)\right)}\right), & \text { two-sided } \\
P\left(\chi^{2}(N-1) \leq \frac{h^{2} N(N-1)}{\sigma_{\text {diff }}^{2}\left(t_{1-\alpha}^{2}(N-1)\right)}\right), & \text { one-sided }\end{cases} \\
& \operatorname{Pr}\left(\text { half-width } \leq h \left\lvert\, \quad=\left\{\begin{array}{cr}
\left(\frac{1}{1-\alpha}\right) 2\left[Q _ { N - 1 } \left(\left(t_{1-\frac{\alpha}{2}}(N-1)\right), 0 ;\right.\right. & \\
\left.\left.0, b_{1}\right)-Q_{N-1}\left(0,0 ; 0, b_{1}\right)\right], & \text { two-sided } \\
\text { validity }) \\
\left(\frac{1}{1-\alpha}\right) Q_{N-1}\left(\left(t_{1-\alpha}(N-1)\right), 0 ; 0, b_{1}\right), & \text { one-sided }
\end{array}\right.\right.\right.
\end{aligned}
$$

where

$$
\begin{aligned}
\sigma_{\text {diff }} & =\left(\sigma_{1}^{2}+\sigma_{2}^{2}-2 \rho \sigma_{1} \sigma_{2}\right)^{\frac{1}{2}} \\
b_{1} & =\frac{h(N-1)^{\frac{1}{2}}}{\sigma_{\text {diff }}\left(t_{1-\frac{\alpha}{c}}(N-1)\right) N^{-\frac{1}{2}}} \\
c & =\text { number of sides }
\end{aligned}
$$

and $Q \cdot(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.
A "quality" confidence interval is both sufficiently narrow (half-width $\leq h$ ) and valid:

$$
\begin{aligned}
\operatorname{Pr}(\text { quality }) & =\operatorname{Pr}(\text { half-width } \leq h \text { and validity }) \\
& =\operatorname{Pr}(\text { half-width } \leq h \mid \text { validity })(1-\alpha)
\end{aligned}
$$

## Analyses in the TWOSAMPLEFREQ Statement

## Overview of the $2 \times 2$ Table

Notation:

|  |  | Outcome |  |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  | Failure | Success |  |
| Group | 1 | $n_{1}-x_{1}$ | $x_{1}$ | $n_{1}$ |
|  | 2 | $n_{2}-x_{2}$ | $x_{2}$ | $n_{2}$ |
|  |  | $N-m$ | $m$ | $N$ |

$$
\begin{aligned}
x_{1} & =\# \text { successes in group } 1 \\
x_{2} & =\# \text { successes in group } 2 \\
m & =x_{1}+x_{2}=\text { total } \# \text { successes } \\
\hat{p_{1}} & =\frac{x_{1}}{n_{1}} \\
\hat{p_{2}} & =\frac{x_{2}}{n_{2}} \\
\hat{p} & =\frac{m}{N}=w_{1} \hat{p_{1}}+w_{2} \hat{p_{2}}
\end{aligned}
$$

The hypotheses are

$$
\begin{aligned}
& H_{0}: p_{2}-p_{1}=p_{0} \\
& H_{1}: \begin{cases}p_{2}-p_{1} \neq p_{0}, & \text { two-sided } \\
p_{2}-p_{1}>p_{0}, & \text { upper one-sided } \\
p_{2}-p_{1}<p_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

where $p_{0}$ is constrained to be 0 for the likelihood ratio and Fisher's exact tests. If $p_{0}<0$ in an upper one-sided test or $p_{0}>0$ in a lower one-sided test, then the test is a noninferiority test. If $p_{0}>0$ in an upper one-sided test or $p_{0}<0$ in a lower one-sided test, then the test is a superiority test. Although $p_{0}$ is unconstrained for the Pearson chi-square test, $p_{0} \neq 0$ is not recommended for that test. The Farrington-Manning score test is a better choice when $p_{0} \neq 0$.

Internal calculations are performed in terms of $p_{1}, p_{2}$, and $p_{0}$. An input set consisting of $\mathrm{OR}, p_{1}$, and $\mathrm{OR}_{0}$ is transformed as follows:

$$
\begin{aligned}
p_{2} & =\frac{(\mathrm{OR}) p_{1}}{1-p_{1}+(\mathrm{OR}) p_{1}} \\
p_{10} & =p_{1} \\
p_{20} & =\frac{\mathrm{OR}_{0} p_{10}}{1-p_{10}+\left(\mathrm{OR}_{0}\right) p_{10}} \\
p_{0} & =p_{20}-p_{10}
\end{aligned}
$$

An input set consisting of $\mathrm{RR}, p_{1}$, and $\mathrm{RR}_{0}$ is transformed as follows:

$$
\begin{aligned}
p_{2} & =(\mathrm{RR}) p_{1} \\
p_{10} & =p_{1} \\
p_{20} & =\left(\mathrm{RR}_{0}\right) p_{10} \\
p_{0} & =p_{20}-p_{10}
\end{aligned}
$$

Note that the transformation of either $\mathrm{OR}_{0}$ or $\mathrm{RR}_{0}$ to $p_{0}$ is not unique. The chosen parameterization fixes the null value $p_{10}$ at the input value of $p_{1}$. Some values of $\mathrm{OR}_{0}$ or $\mathrm{RR}_{0}$ might lead to invalid values of $p_{0}$ ( $p_{0} \leq 0$ or $p_{0} \geq 1$ ), in which case an "Invalid input" error occurs.

## Farrington-Manning Score Test for Two Proportions (TEST=FM)

The Farrington-Manning score test is based on equations (1), (2), and (12) in Farrington and Manning (1990). The test statistic, which is assumed to have a null distribution of $N(0,1)$ under $H_{0}$, is

$$
z_{F M}=\frac{\hat{p_{2}}-\hat{p_{1}}-p_{0}}{\left[\frac{\tilde{p_{1}}\left(1-\tilde{p_{1}}\right)}{n_{1}}+\frac{\tilde{\tilde{2}_{2}}\left(1-\tilde{p_{2}}\right)}{n_{2}}\right]^{\frac{1}{2}}}=\left[N w_{1} w_{2}\right]^{\frac{1}{2}} \frac{\hat{p_{2}}-\hat{p_{1}}-p_{0}}{\left[w_{2} \tilde{p_{1}}\left(1-\tilde{p_{1}}\right)+w_{1} \tilde{p_{2}}\left(1-\tilde{p_{2}}\right)\right]^{\frac{1}{2}}}
$$

where $\tilde{p_{1}}$ and $\tilde{p_{2}}$ are the maximum likelihood estimates of the proportions under the restriction $\tilde{p_{2}}-\tilde{p_{1}}=p_{0}$.
Sample size for the one-sided cases is given by equations (4) and (12) in Farrington and Manning (1990). One-sided power is computed by inverting the sample size formula. Power for the two-sided case is computed by adding the lower-sided and upper-sided powers, each evaluated at $\alpha / 2$. Sample size for the two-sided case is obtained by numerically inverting the power formula,

$$
\text { power }=\left\{\begin{array}{lll}
\Phi\left(\frac{\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\alpha}\left[w_{2} \tilde{p}_{1}\left(1-\tilde{p}_{1}\right)+w_{1} \tilde{p}_{2}\left(1-\tilde{p}_{2}\right)\right]^{\frac{1}{2}}}{\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right]^{\frac{1}{2}}}\right), & \text { upper one-sided } \\
\Phi\left(\frac{-\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\alpha}\left[w_{2} \tilde{p}_{1}\left(1-\tilde{p}_{1}\right)+w_{1} \tilde{p}_{2}\left(1-\tilde{p}_{2}\right)\right]^{\frac{1}{2}}}{\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right]^{\frac{1}{2}}}\right), & \text { lower one-sided } \\
\Phi\left(\frac{\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\frac{\alpha}{2}}\left[w_{2} \tilde{p_{1}}\left(1-\tilde{p}_{1}\right)+w_{1} \tilde{p_{2}}\left(1-\tilde{p}_{2}\right)\right]^{\frac{1}{2}}}{\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right)^{\frac{1}{2}}}\right)+ & \\
\Phi\left(\frac{-\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\frac{\alpha}{2}}\left[w_{2} \tilde{p_{1}}\left(1-\tilde{p}_{1}\right)+w_{1} \tilde{p_{2}}\left(1-\tilde{p_{2}}\right)\right]^{\frac{1}{2}}}{\left.\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right]^{\frac{1}{2}}\right),}\right. & \text { two-sided }
\end{array}\right.
$$

where

$$
\begin{aligned}
\tilde{p}_{2} & =2 u \cos (w)-b /(3 a) \\
\tilde{p}_{1} & =\tilde{p}_{2}-p_{0} \\
w & =\left(\pi+\cos ^{-1}\left(v / u^{3}\right)\right) / 3 \\
v & =b^{3} /(3 a)^{3}-b c /\left(6 a^{2}\right)+d /(2 a) \\
u & =\operatorname{sign}(v) \sqrt{b^{2} /(3 a)^{2}-c /(3 a)} \\
a & =1+w_{1} / w_{2} \\
b & =-\left[1+w_{1} / w_{2}+p_{2}+\left(w_{1} / w_{2}\right) p_{1}+p_{0}\left(w_{1} / w_{2}+2\right)\right] \\
c & =p_{0}^{2}+p_{0}\left(2 p_{2}+w_{1} / w_{2}+1\right)+p_{2}+\left(w_{1} / w_{2}\right) p_{1} \\
d & =-p_{2} p_{0}\left(1+p_{0}\right)
\end{aligned}
$$

For the one-sided cases, a closed-form inversion of the power equation yields an approximate total sample size of

$$
N=\frac{\left[z_{1-\alpha}\left\{w_{2} \tilde{p}_{1}\left(1-\tilde{p}_{1}\right)+w_{1} \tilde{p}_{2}\left(1-\tilde{p}_{2}\right)\right\}^{\frac{1}{2}}+z_{\text {power }}\left\{w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right\}^{\frac{1}{2}}\right]^{2}}{w_{1} w_{2}\left(p_{2}-p_{1}-p_{0}\right)^{2}}
$$

For the two-sided case, the solution for $N$ is obtained by numerically inverting the power equation.

## Pearson Chi-Square Test for Two Proportions (TEST=PCHI)

The usual Pearson chi-square test is unconditional. The test statistic

$$
z_{P}=\frac{\hat{p}_{2}-\hat{p}_{1}-p_{0}}{\left[\hat{p}(1-\hat{p})\left(\frac{1}{n_{1}}+\frac{1}{n_{2}}\right)\right]^{\frac{1}{2}}}=\left[N w_{1} w_{2}\right]^{\frac{1}{2}} \frac{\hat{p_{2}}-\hat{p}_{1}-p_{0}}{[\hat{p}(1-\hat{p})]^{\frac{1}{2}}}
$$

is assumed to have a null distribution of $N(0,1)$.
Sample size for the one-sided cases is given by equation (4) in Fleiss, Tytun, and Ury (1980). One-sided power is computed as suggested by Diegert and Diegert (1981) by inverting the sample size formula. Power for the two-sided case is computed by adding the lower-sided and upper-sided powers each evaluated at $\alpha / 2$. Sample size for the two-sided case is obtained by numerically inverting the power formula. A custom null value $p_{0}$ for the proportion difference $p_{2}-p_{1}$ is also supported, but it is not recommended. If you are using a non-default null value, then the Farrington-Manning score test is a better choice.

$$
\text { power }=\left\{\begin{array}{lll}
\Phi\left(\frac{\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\alpha}\left[\left(w_{1} p_{1}+w_{2} p_{2}\right)\left(1-w_{1} p_{1}-w_{2} p_{2}\right)\right]^{\frac{1}{2}}}{\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right),},\right. & \text { upper one-sided } \\
\Phi\left(\frac{-\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\alpha}\left[\left(w_{1} p_{1}+w_{2} p_{2}\right)\left(1-w_{1} p_{1}-w_{2} p_{2}\right)\right]^{\frac{1}{2}}}{\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right]^{\frac{1}{2}}}\right), & \text { lower one-sided } \\
\Phi\left(\frac{\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\frac{\alpha}{2}}\left[\left(w_{1} p_{1}+w_{2} p_{2}\right)\left(1-w_{1} p_{1}-w_{2} p_{2}\right)\right]^{\frac{1}{2}}}{\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right]^{\frac{1}{2}}}\right)+ & \\
\Phi\left(\frac{-\left(p_{2}-p_{1}-p_{0}\right)\left(N w_{1} w_{2}\right)^{\frac{1}{2}}-z_{1-\frac{\alpha}{2}}\left[\left(w_{1} p_{1}+w_{2} p_{2}\right)\left(1-w_{1} p_{1}-w_{2} p_{2}\right)\right]^{\frac{1}{2}}}{\left[w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right]^{\frac{1}{2}}}\right), & \text { two-sided }
\end{array}\right.
$$

For the one-sided cases, a closed-form inversion of the power equation yields an approximate total sample size

$$
N=\frac{\left[z_{1-\alpha}\left\{\left(w_{1} p_{1}+w_{2} p_{2}\right)\left(1-w_{1} p_{1}-w_{2} p_{2}\right)\right\}^{\frac{1}{2}}+z_{\text {power }}\left\{w_{2} p_{1}\left(1-p_{1}\right)+w_{1} p_{2}\left(1-p_{2}\right)\right\}^{\frac{1}{2}}\right]^{2}}{w_{1} w_{2}\left(p_{2}-p_{1}-p_{0}\right)^{2}}
$$

For the two-sided case, the solution for $N$ is obtained by numerically inverting the power equation.

## Likelihood Ratio Chi-Square Test for Two Proportions (TEST=LRCHI)

The usual likelihood ratio chi-square test is unconditional. The test statistic

$$
z_{\mathrm{LR}}=\left(-1_{\left\{p_{2}<p_{1}\right\}}\right) \sqrt{2 N \sum_{i=1}^{2}\left[w_{i} \hat{p}_{i} \log \left(\frac{\hat{p}_{i}}{\hat{p}}\right)+w_{i}\left(1-\hat{p}_{i}\right) \log \left(\frac{1-\hat{p}_{i}}{1-\hat{p}}\right)\right]}
$$

is assumed to have a null distribution of $N(0,1)$ and an alternative distribution of $N(\delta, 1)$, where

$$
\delta=N^{\frac{1}{2}}\left(-1_{\left\{p_{2}<p_{1}\right\}}\right) \sqrt{2 \sum_{i=1}^{2}\left[w_{i} p_{i} \log \left(\frac{p_{i}}{w_{1} p_{1}+w_{2} p_{2}}\right)+w_{i}\left(1-p_{i}\right) \log \left(\frac{1-p_{i}}{1-\left(w_{1} p_{1}+w_{2} p_{2}\right)}\right)\right]}
$$

The approximate power is

$$
\text { power }= \begin{cases}\Phi\left(\delta-z_{1-\alpha}\right), & \text { upper one-sided } \\ \Phi\left(-\delta-z_{1-\alpha}\right), & \text { lower one-sided } \\ \Phi\left(\delta-z_{1-\frac{\alpha}{2}}\right)+\Phi\left(-\delta-z_{1-\frac{\alpha}{2}}\right), & \text { two-sided }\end{cases}
$$

For the one-sided cases, a closed-form inversion of the power equation yield an approximate total sample size

$$
N=\left(\frac{z_{\text {power }}+z_{1-\alpha}}{\delta}\right)^{2}
$$

For the two-sided case, the solution for $N$ is obtained by numerically inverting the power equation.
Fisher's Exact Conditional Test for Two Proportions (Test=FISHER)
Fisher's exact test is conditional on the observed total number of successes $m$. Power and sample size computations are based on a test with similar power properties, the continuity-adjusted arcsine test. The test statistic

$$
\begin{aligned}
z_{A}= & \left(4 N w_{1} w_{2}\right)^{\frac{1}{2}}\left[\arcsin \left(\left[\hat{p}_{2}+\frac{1}{2 N w_{2}}\left(1_{\left\{\hat{p}_{2}<\hat{p}_{1}\right\}}-1_{\left\{\hat{p}_{2}>\hat{p}_{1}\right\}}\right)\right]^{\frac{1}{2}}\right)\right. \\
& \left.-\arcsin \left(\left[\hat{p}_{1}+\frac{1}{2 N w_{1}}\left(1_{\left\{\hat{p}_{1}<\hat{p}_{2}\right\}}-1_{\left\{\hat{p}_{1}>\hat{p}_{2}\right\}}\right)\right]^{\frac{1}{2}}\right)\right]
\end{aligned}
$$

is assumed to have a null distribution of $N(0,1)$ and an alternative distribution of $N(\delta, 1)$, where

$$
\begin{aligned}
\delta= & \left(4 N w_{1} w_{2}\right)^{\frac{1}{2}}\left[\operatorname { a r c s i n } \left(\left[p_{2}+\frac{1}{2 N w_{2}}\left(1_{\left\{p_{2}<p_{1}\right\}}-1_{\left\{p_{2}>p_{1}\right\}}\right]^{\frac{1}{2}}\right)\right.\right. \\
& \left.-\arcsin \left(\left[p_{1}+\frac{1}{2 N w_{1}}\left(1_{\left\{p_{1}<p_{2}\right\}}-1_{\left\{p_{1}>p_{2}\right\}}\right)\right]^{\frac{1}{2}}\right)\right]
\end{aligned}
$$

The approximate power for the one-sided balanced case is given by Walters (1979) and is easily extended to the unbalanced and two-sided cases:

$$
\text { power }= \begin{cases}\Phi\left(\delta-z_{1-\alpha}\right), & \text { upper one-sided } \\ \Phi\left(-\delta-z_{1-\alpha}\right), & \text { lower one-sided } \\ \Phi\left(\delta-z_{1-\frac{\alpha}{2}}\right)+\Phi\left(-\delta-z_{1-\frac{\alpha}{2}}\right), & \text { two-sided }\end{cases}
$$

The approximation is valid only for $N \geq 1 /\left(2 w_{1} w_{2}\left|p_{1}-p_{2}\right|\right)$.

## Analyses in the TWOSAMPLEMEANS Statement

## Two-Sample $t$ Test Assuming Equal Variances (TEST=DIFF)

The hypotheses for the two-sample $t$ test are

$$
\begin{aligned}
& H_{0}: \mu_{\text {diff }}=\mu_{0} \\
& H_{1}: \begin{cases}\mu_{\text {diff }} \neq \mu_{0}, & \text { two-sided } \\
\mu_{\text {diff }}>\mu_{0}, & \text { upper one-sided } \\
\mu_{\text {diff }}<\mu_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

The test assumes normally distributed data and common standard deviation per group, and it requires $N \geq 3$, $n_{1} \geq 1$, and $n_{2} \geq 1$. The test statistics are

$$
\begin{aligned}
t & =N^{\frac{1}{2}}\left(w_{1} w_{2}\right)^{\frac{1}{2}}\left(\frac{\bar{x}_{2}-\bar{x}_{1}-\mu_{0}}{s_{p}}\right) \sim t(N-2, \delta) \\
t^{2} & \sim F\left(1, N-2, \delta^{2}\right)
\end{aligned}
$$

where $\bar{x}_{1}$ and $\bar{x}_{2}$ are the sample means and $s_{p}$ is the pooled standard deviation, and

$$
\delta=N^{\frac{1}{2}}\left(w_{1} w_{2}\right)^{\frac{1}{2}}\left(\frac{\mu_{\mathrm{diff}}-\mu_{0}}{\sigma}\right)
$$

The test is
Reject $H_{0} \quad$ if $\begin{cases}t^{2} \geq F_{1-\alpha}(1, N-2), & \text { two-sided } \\ t \geq t_{1-\alpha}(N-2), & \text { upper one-sided } \\ t \leq t_{\alpha}(N-2), & \text { lower one-sided }\end{cases}$
Exact power computations for $t$ tests are given in O'Brien and Muller (1993, Section 8.2.1):

$$
\text { power }= \begin{cases}P\left(F\left(1, N-2, \delta^{2}\right) \geq F_{1-\alpha}(1, N-2)\right), & \text { two-sided } \\ P\left(t(N-2, \delta) \geq t_{1-\alpha}(N-2)\right), & \text { upper one-sided } \\ P\left(t(N-2, \delta) \leq t_{\alpha}(N-2)\right), & \text { lower one-sided }\end{cases}
$$

Solutions for $N, n_{1}, n_{2}, \alpha$, and $\delta$ are obtained by numerically inverting the power equation. Closed-form solutions for other parameters, in terms of $\delta$, are as follows:

$$
\begin{aligned}
\mu_{\text {diff }} & =\delta \sigma\left(N w_{1} w_{2}\right)^{-\frac{1}{2}}+\mu_{0} \\
\mu_{1} & =\delta \sigma\left(N w_{1} w_{2}\right)^{-\frac{1}{2}}+\mu_{0}-\mu_{2} \\
\mu_{2} & =\delta \sigma\left(N w_{1} w_{2}\right)^{-\frac{1}{2}}+\mu_{0}-\mu_{1} \\
\sigma & = \begin{cases}\delta^{-1}\left(N w_{1} w_{2}\right)^{\frac{1}{2}}\left(\mu_{\text {diff }}-\mu_{0}\right), & |\delta|>0 \\
\text { undefined, } & \text { otherwise }\end{cases} \\
w_{1} & = \begin{cases}\frac{1}{2} \pm \frac{1}{2}\left[1-\frac{4 \delta^{2} \sigma^{2}}{N\left(\mu_{\text {diff }}-\mu_{0}\right)^{2}}\right]^{\frac{1}{2}}, & 0<|\delta| \leq \frac{1}{2} N^{\frac{1}{2}} \frac{\left|\mu_{\text {diff }}-\mu_{0}\right|}{\sigma} \\
\text { undefined, } & \text { otherwise }\end{cases} \\
w_{2} & = \begin{cases}\frac{1}{2} \pm \frac{1}{2}\left[1-\frac{4 \delta^{2} \sigma^{2}}{N\left(\mu_{\text {diff }}-\mu_{0}\right)^{2}}\right]^{\frac{1}{2}}, & 0<|\delta| \leq \frac{1}{2} N^{\frac{1}{2}} \frac{\left|\mu_{\text {diff }}-\mu_{0}\right|}{\sigma} \\
\text { undefined, } & \text { otherwise }\end{cases}
\end{aligned}
$$

Finally, here is a derivation of the solution for $w_{1}$ :
Solve the $\delta$ equation for $w_{1}$ (which requires the quadratic formula). Then determine the range of $\delta$ given $w_{1}$ :

$$
\begin{aligned}
& \min _{w_{1}}(\delta)= \begin{cases}0, & \text { when } \quad w_{1}=0 \quad \text { or } \quad 1, \quad \text { if } \quad\left(\mu_{\text {diff }}-\mu_{0}\right) \geq 0 \\
\frac{1}{2} N^{\frac{1}{2}} \frac{\left(\mu_{\text {diff }}-\mu_{0}\right)}{\sigma}, & \text { when } \quad w_{1}=\frac{1}{2}, \quad \text { if } \quad\left(\mu_{\text {diff }}-\mu_{0}\right)<0\end{cases} \\
& \max _{w_{1}}(\delta)= \begin{cases}0, & \text { when } \quad w_{1}=0 \quad \text { or } \quad 1, \quad \text { if } \quad\left(\mu_{\text {diff }}-\mu_{0}\right)<0 \\
\frac{1}{2} N^{\frac{1}{2}} \frac{\left(\mu_{\text {diff }}-\mu_{0}\right)}{\sigma}, & \text { when } \quad w_{1}=\frac{1}{2}, \quad \text { if } \quad\left(\mu_{\text {diff }}-\mu_{0}\right) \geq 0\end{cases}
\end{aligned}
$$

This implies

$$
|\delta| \leq \frac{1}{2} N^{\frac{1}{2}} \frac{\left|\mu_{\text {diff }}-\mu_{0}\right|}{\sigma}
$$

Two-Sample Satterthwaite $t$ Test Assuming Unequal Variances (TEST=DIFF_SATT)
The hypotheses for the two-sample Satterthwaite $t$ test are

$$
\begin{aligned}
& H_{0}: \mu_{\text {diff }}=\mu_{0} \\
& H_{1}: \begin{cases}\mu_{\text {diff }} \neq \mu_{0}, & \text { two-sided } \\
\mu_{\text {diff }}>\mu_{0}, & \text { upper one-sided } \\
\mu_{\text {diff }}<\mu_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

The test assumes normally distributed data and requires $N \geq 3, n_{1} \geq 1$, and $n_{2} \geq 1$. The test statistics are

$$
\begin{aligned}
t & =\frac{\bar{x}_{2}-\bar{x}_{1}-\mu_{0}}{\left[\frac{s_{1}^{2}}{n_{1}}+\frac{s_{2}^{2}}{n_{2}}\right]^{\frac{1}{2}}}=N^{\frac{1}{2}} \frac{\bar{x}_{2}-\bar{x}_{1}-\mu_{0}}{\left[\frac{s_{1}^{2}}{w_{1}}+\frac{s_{2}^{2}}{w_{2}}\right]^{\frac{1}{2}}} \\
F & =t^{2}
\end{aligned}
$$

where $\bar{x}_{1}$ and $\bar{x}_{2}$ are the sample means and $s_{1}$ and $s_{2}$ are the sample standard deviations.
DiSantostefano and Muller (1995, p. 585) state, the test is based on assuming that under $H_{0}, F$ is distributed as $F(1, v)$, where $v$ is given by Satterthwaite's approximation (Satterthwaite 1946),

$$
v=\frac{\left[\frac{\sigma_{1}^{2}}{n_{1}}+\frac{\sigma_{2}^{2}}{n_{2}}\right]^{2}}{\left[\frac{\left[\frac{\sigma_{1}^{2}}{n_{1}}\right]^{2}}{n_{1}-1}+\frac{\left[\frac{\sigma_{2}^{2}}{n_{2}}\right]^{2}}{n_{2}-1}\right.}=\frac{\left[\frac{\sigma_{1}^{2}}{w_{1}}+\frac{\sigma_{2}^{2}}{w_{2}}\right]^{2}}{\left[\frac{\sigma_{1}^{2}}{w_{1}}\right]^{2}} \frac{\left[\frac{\sigma_{2}^{2}}{w_{2}}\right]^{2}}{N w_{2}-1}+1 .
$$

Since $v$ is unknown, in practice it must be replaced by an estimate

$$
\hat{v}=\frac{\left[\frac{s_{1}^{2}}{n_{1}}+\frac{s_{2}^{2}}{n_{2}}\right]^{2}}{\left[\frac{\left[\frac{s_{1}^{2}}{n_{1}}\right]^{2}}{n_{1}-1}+\frac{\left[\frac{s_{2}^{2}}{n_{2}}\right]^{2}}{n_{2}-1}\right.}=\frac{\left[\frac{s_{1}^{2}}{w_{1}}+\frac{s_{2}^{2}}{w_{2}}\right]^{2}}{\frac{\left[\frac{s_{1}^{2}}{w_{1}}\right]^{2}}{N w_{1}-1}+\frac{\left[\frac{s_{2}^{2}}{w_{2}}\right]^{2}}{N w_{2}-1}}
$$

So the test is

$$
\text { Reject } \quad H_{0} \quad \text { if } \begin{cases}F \geq F_{1-\alpha}(1, \hat{v}), & \text { two-sided } \\ t \geq t_{1-\alpha}(\hat{v}), & \text { upper one-sided } \\ t \leq t_{\alpha}(\hat{v}), & \text { lower one-sided }\end{cases}
$$

Exact solutions for power for the two-sided and upper one-sided cases are given in Moser, Stevens, and Watts (1989). The lower one-sided case follows easily by using symmetry. The equations are as follows:

$$
\text { power }=\left\{\begin{array}{cl}
\int_{0}^{\infty} P(F(1, N-2, \lambda)> & \\
\left.h(u) F_{1-\alpha}(1, v(u)) \mid u\right) f(u) \mathrm{d} u, & \text { two-sided } \\
\int_{0}^{\infty} P\left(t\left(N-2, \lambda^{\frac{1}{2}}\right)>\right. & \\
\left.\left.[h(u)]^{\frac{1}{2}} t_{1-\alpha}(v(u)) \right\rvert\, u\right) f(u) \mathrm{d} u, & \text { upper one-sided } \\
\int_{0}^{\infty} P\left(t\left(N-2, \lambda \frac{1}{2}\right)<\right. & \\
\left.\left.[h(u)]^{\frac{1}{2}} t_{\alpha}(v(u)) \right\rvert\, u\right) f(u) \mathrm{d} u, & \text { lower one-sided }
\end{array}\right.
$$

where

$$
\begin{aligned}
h(u) & =\frac{\left(\frac{1}{n_{1}}+\frac{u}{n_{2}}\right)\left(n_{1}+n_{2}-2\right)}{\left[\left(n_{1}-1\right)+\left(n_{2}-1\right) \frac{u \sigma_{1}^{2}}{\sigma_{2}^{2}}\right]\left(\frac{1}{n_{1}}+\frac{\sigma_{2}^{2}}{\sigma_{1}^{2} n_{2}}\right)} \\
v(u) & =\frac{\left(\frac{1}{n_{1}}+\frac{u}{n_{2}}\right)^{2}}{\frac{1}{n_{1}^{2}\left(n_{1}-1\right)}+\frac{u^{2}}{n_{2}^{2}\left(n_{2}-1\right)}} \\
\lambda & =\frac{\left(\mu_{\text {diff }}-\mu_{0}\right)^{2}}{\frac{\sigma_{1}^{2}}{n_{1}}+\frac{\sigma_{2}^{2}}{n_{2}}} \\
f(u) & =\frac{\Gamma\left(\frac{n_{1}+n_{2}-2}{2}\right)}{\Gamma\left(\frac{n_{1}-1}{2}\right) \Gamma\left(\frac{n_{2}-1}{2}\right)}\left[\frac{\sigma_{1}^{2}\left(n_{2}-1\right)}{\sigma_{2}^{2}\left(n_{1}-1\right)}\right]^{\frac{n_{2}-1}{2}} u^{\frac{n_{2}-3}{2}}\left[1+\left(\frac{n_{2}-1}{n_{1}-1}\right) \frac{u \sigma_{1}^{2}}{\sigma_{2}^{2}}\right]^{-\left(\frac{n_{1}+n_{2}-2}{2}\right)}
\end{aligned}
$$

The density $f(u)$ is obtained from the fact that

$$
\frac{u \sigma_{1}^{2}}{\sigma_{2}^{2}} \sim F\left(n_{2}-1, n_{1}-1\right)
$$

Because the test is biased, the achieved significance level might differ from the nominal significance level. The actual alpha is computed in the same way as the power, except that the mean difference $\mu_{\text {diff }}$ is replaced by the null mean difference $\mu_{0}$.

## Two-Sample Pooled $t$ Test of Mean Ratio with Lognormal Data (TEST=RATIO)

The lognormal case is handled by reexpressing the analysis equivalently as a normality-based test on the log-transformed data, by using properties of the lognormal distribution as discussed in Johnson, Kotz, and Balakrishnan (1994, Chapter 14). The approaches in the section "Two-Sample $t$ Test Assuming Equal Variances (TEST=DIFF)" on page 6416 then apply.

In contrast to the usual $t$ test on normal data, the hypotheses with lognormal data are defined in terms of geometric means rather than arithmetic means. The test assumes equal coefficients of variation in the two groups.

The hypotheses for the two-sample $t$ test with lognormal data are

$$
\begin{aligned}
& H_{0}: \frac{\gamma_{2}}{\gamma_{1}}=\gamma_{0} \\
& H_{1}: \begin{cases}\frac{\gamma_{2}}{\gamma_{1}} \neq \gamma_{0}, & \text { two-sided } \\
\frac{\gamma_{2}}{\gamma_{1}}>\gamma_{0}, & \text { upper one-sided } \\
\frac{\gamma_{2}}{\gamma_{1}}<\gamma_{0}, & \text { lower one-sided }\end{cases}
\end{aligned}
$$

Let $\mu_{1}^{\star}, \mu_{2}^{\star}$, and $\sigma^{\star}$ be the (arithmetic) means and common standard deviation of the corresponding normal distributions of the log-transformed data. The hypotheses can be rewritten as follows:

$$
\begin{aligned}
& H_{0}: \mu_{2}^{\star}-\mu_{1}^{\star}=\log \left(\gamma_{0}\right) \\
& H_{1}: \begin{cases}\mu_{2}^{\star}-\mu_{1}^{\star} \neq \log \left(\gamma_{0}\right), & \text { two-sided } \\
\mu_{2}^{\star}-\mu_{1}^{\star}>\log \left(\gamma_{0}\right), & \text { upper one-sided } \\
\mu_{2}^{\star}-\mu_{1}^{\star}<\log \left(\gamma_{0}\right), & \text { lower one-sided }\end{cases}
\end{aligned}
$$

where

$$
\begin{aligned}
\mu_{1}^{\star} & =\log \gamma_{1} \\
\mu_{2}^{\star} & =\log \gamma_{2}
\end{aligned}
$$

The test assumes lognormally distributed data and requires $N \geq 3, n_{1} \geq 1$, and $n_{2} \geq 1$.
The power is

$$
\text { power }= \begin{cases}P\left(F\left(1, N-2, \delta^{2}\right) \geq F_{1-\alpha}(1, N-2)\right), & \text { two-sided } \\ P\left(t(N-2, \delta) \geq t_{1-\alpha}(N-2)\right), & \text { upper one-sided } \\ P\left(t(N-2, \delta) \leq t_{\alpha}(N-2)\right), & \text { lower one-sided }\end{cases}
$$

where

$$
\begin{aligned}
\delta & =N^{\frac{1}{2}}\left(w_{1} w_{2}\right)^{\frac{1}{2}}\left(\frac{\mu_{2}^{\star}-\mu_{1}^{\star}-\log \left(\gamma_{0}\right)}{\sigma^{\star}}\right) \\
\sigma^{\star} & =\left[\log \left(\mathrm{CV}^{2}+1\right)\right]^{\frac{1}{2}}
\end{aligned}
$$

Additive Equivalence Test for Mean Difference with Normal Data (TEST=EQUIV_DIFF)
The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: \mu_{\text {diff }}<\theta_{L} \quad \text { or } \quad \mu_{\text {diff }}>\theta_{U} \\
& H_{1}: \theta_{L} \leq \mu_{\text {diff }} \leq \theta_{U}
\end{aligned}
$$

The analysis is the two one-sided tests (TOST) procedure of Schuirmann (1987). The test assumes normally distributed data and requires $N \geq 3, n_{1} \geq 1$, and $n_{2} \geq 1$. Phillips (1990) derives an expression for the exact power assuming a balanced design; the results are easily adapted to an unbalanced design:

$$
\begin{aligned}
\text { power }= & Q_{N-2}\left(\left(-t_{1-\alpha}(N-2)\right), \frac{\mu_{\text {diff }}-\theta_{U}}{\sigma N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}} ; 0, \frac{(N-2)^{\frac{1}{2}}\left(\theta_{U}-\theta_{L}\right)}{2 \sigma N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}\left(t_{1-\alpha}(N-2)\right)}\right)- \\
& Q_{N-2}\left(\left(t_{1-\alpha}(N-2)\right), \frac{\mu_{\text {diff }}-\theta_{L}}{\sigma N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}} ; 0, \frac{(N-2)^{\frac{1}{2}}\left(\theta_{U}-\theta_{L}\right)}{2 \sigma N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}\left(t_{1-\alpha}(N-2)\right)}\right)
\end{aligned}
$$

where $Q .(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.

## Multiplicative Equivalence Test for Mean Ratio with Lognormal Data (TEST=EQUIV_RATIO)

The lognormal case is handled by reexpressing the analysis equivalently as a normality-based test on the log-transformed data, by using properties of the lognormal distribution as discussed in Johnson, Kotz, and Balakrishnan (1994, Chapter 14). The approaches in the section "Additive Equivalence Test for Mean Difference with Normal Data (TEST=EQUIV_DIFF)" on page 6420 then apply.

In contrast to the additive equivalence test on normal data, the hypotheses with lognormal data are defined in terms of geometric means rather than arithmetic means.

The hypotheses for the equivalence test are

$$
\begin{aligned}
& H_{0}: \frac{\gamma_{T}}{\gamma_{R}} \leq \theta_{L} \quad \text { or } \quad \frac{\gamma_{T}}{\gamma_{R}} \geq \theta_{U} \\
& H_{1}: \theta_{L}<\frac{\gamma_{T}}{\gamma_{R}}<\theta_{U} \\
& \text { where } 0<\theta_{L}<\theta_{U}
\end{aligned}
$$

The analysis is the two one-sided tests (TOST) procedure of Schuirmann (1987) on the log-transformed data. The test assumes lognormally distributed data and requires $N \geq 3, n_{1} \geq 1$, and $n_{2} \geq 1$. Diletti, Hauschke, and Steinijans (1991) derive an expression for the exact power assuming a crossover design; the results are easily adapted to an unbalanced two-sample design:

$$
\begin{aligned}
\text { power }= & Q_{N-2}\left(\left(-t_{1-\alpha}(N-2)\right), \frac{\log \left(\frac{\gamma_{T}}{\gamma_{R}}\right)-\log \left(\theta_{U}\right)}{\sigma^{\star} N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}} ; 0, \frac{(N-2)^{\frac{1}{2}}\left(\log \left(\theta_{U}\right)-\log \left(\theta_{L}\right)\right)}{2 \sigma^{\star} N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}\left(t_{1-\alpha}(N-2)\right)}\right)- \\
& Q_{N-2}\left(\left(t_{1-\alpha}(N-2)\right), \frac{\log \left(\frac{\gamma_{T}}{\gamma_{R}}\right)-\log \left(\theta_{L}\right)}{\sigma^{\star} N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}} ; 0, \frac{(N-2)^{\frac{1}{2}}\left(\log \left(\theta_{U}\right)-\log \left(\theta_{L}\right)\right)}{2 \sigma^{\star} N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}\left(t_{1-\alpha}(N-2)\right)}\right)
\end{aligned}
$$

where

$$
\sigma^{\star}=\left[\log \left(\mathrm{CV}^{2}+1\right)\right]^{\frac{1}{2}}
$$

is the (assumed common) standard deviation of the normal distribution of the log-transformed data, and $Q .(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.

## Confidence Interval for Mean Difference (CI=DIFF)

This analysis of precision applies to the standard $t$-based confidence interval:

$$
\begin{array}{ll}
{\left[\left(\bar{x}_{2}-\bar{x}_{1}\right)-t_{1-\frac{\alpha}{2}}(N-2) \frac{s_{p}}{\sqrt{N w_{1} w_{2}}},\right.} & \\
\left.\left(\bar{x}_{2}-\bar{x}_{1}\right)+t_{1-\frac{\alpha}{2}}(N-2) \frac{s_{p}}{\sqrt{N w_{1} w_{2}}}\right], & \text { two-sided } \\
{\left[\left(\bar{x}_{2}-\bar{x}_{1}\right)-t_{1-\alpha}(N-2) \frac{s_{p}}{\sqrt{N w_{1} w_{2}}},\right.} & \infty), \\
\left(-\infty, \quad\left(\bar{x}_{2}-\bar{x}_{1}\right)+t_{1-\alpha}(N-2) \frac{s_{p}}{\sqrt{N w_{1} w_{2}}}\right], & \text { upper one-sided } \\
(-\infty w e r \text { one-sided }
\end{array}
$$

where $\bar{x}_{1}$ and $\bar{x}_{2}$ are the sample means and $s_{p}$ is the pooled standard deviation. The "half-width" is defined as the distance from the point estimate $\bar{x}_{2}-\bar{x}_{1}$ to a finite endpoint,

$$
\text { half-width }= \begin{cases}t_{1-\frac{\alpha}{2}}(N-2) \frac{s_{p}}{\sqrt{N w_{1} w_{2}}}, & \text { two-sided } \\ t_{1-\alpha}(N-2) \frac{s_{p}}{\sqrt{N w_{1} w_{2}}}, & \text { one-sided }\end{cases}
$$

A "valid" conference interval captures the true mean. The exact probability of obtaining at most the target confidence interval half-width $h$, unconditional or conditional on validity, is given by Beal (1989):

$$
\begin{aligned}
\operatorname{Pr}(\text { half-width } \leq h) & =\left\{\begin{array}{cc}
P\left(\chi^{2}(N-2) \leq \frac{h^{2} N(N-2)\left(w_{1} w_{2}\right)}{\sigma^{2}\left(t_{1-\frac{\alpha}{2}}^{2}(N-2)\right)}\right), & \text { two-sided } \\
P\left(\chi^{2}(N-2) \leq \frac{h^{2} N(N-2)\left(w_{1} w_{2}\right)}{\sigma^{2}\left(t_{1-\alpha}^{2}(N-2)\right)}\right), & \text { one-sided } \\
\text { Pr(half-width } \leq h \mid \\
\text { validity })
\end{array}=\left\{\begin{array}{cc}
\left(\frac{1}{1-\alpha}\right) 2\left[Q _ { N - 2 } \left(\left(t_{1-\frac{\alpha}{2}}(N-2)\right), 0 ;\right.\right. & \\
\left.\left.0, b_{2}\right)-Q_{N-2}\left(0,0 ; 0, b_{2}\right)\right], & \text { two-sided } \\
\left(\frac{1}{1-\alpha}\right) Q_{N-2}\left(\left(t_{1-\alpha}(N-2)\right), 0 ; 0, b_{2}\right), & \text { one-sided }
\end{array}\right.\right.
\end{aligned}
$$

where

$$
\begin{aligned}
b_{2} & =\frac{h(N-2)^{\frac{1}{2}}}{\sigma\left(t_{1-\frac{\alpha}{c}}(N-2)\right) N^{-\frac{1}{2}}\left(w_{1} w_{2}\right)^{-\frac{1}{2}}} \\
c & =\text { number of sides }
\end{aligned}
$$

and $Q \cdot(\cdot, \cdot ; \cdot, \cdot)$ is Owen's Q function, defined in the section "Common Notation" on page 6373.
A "quality" confidence interval is both sufficiently narrow (half-width $\leq h$ ) and valid:

$$
\begin{aligned}
\operatorname{Pr}(\text { quality }) & =\operatorname{Pr}(\text { half-width } \leq h \text { and validity }) \\
& =\operatorname{Pr}(\text { half-width } \leq h \mid \text { validity })(1-\alpha)
\end{aligned}
$$

## Analyses in the TWOSAMPLESURVIVAL Statement

Rank Tests for Two Survival Curves (TEST=LOGRANK, TEST=GEHAN, TEST=TARONEWARE)
The method is from Lakatos (1988) and Cantor (1997, pp. 83-92).

Define the following notation:

$$
\begin{aligned}
X_{j}(i) & =i \text { th input time point on survival curve for group } j \\
S_{j}(i) & =\text { input survivor function value corresponding to } X_{j}(i) \\
h_{j}(t) & =\text { hazard rate for group } j \text { at time } t \\
\Psi_{j}(t) & =\text { loss hazard rate for group } j \text { at time } t \\
\lambda_{j} & =\text { exponential hazard rate for group } j \\
R & =\text { hazard ratio of group } 2 \text { to group } 1 \equiv \text { (assumed constant) value of } \frac{h_{2}(t)}{h_{1}(t)} \\
m_{j} & =\text { median survival time for group } j \\
b & =\text { number of subintervals per time unit } \\
T & =\text { accrual time } \\
\tau & =\text { follow-up time after accrual } \\
L_{j} & =\text { exponential loss rate for group } j \\
X L_{j} & =\text { input time point on loss curve for group } j \\
S L_{j} & =\text { input survivor function value corresponding to } X L_{j} \\
m L_{j} & =\text { median survival time for group } j \\
r_{i} & =\text { rank for } i \text { th time point }
\end{aligned}
$$

Each survival curve can be specified in one of several ways.

- For exponential curves:
- a single point ( $\left.X_{j}(1), S_{j}(1)\right)$ on the curve
- median survival time
- hazard rate
- hazard ratio (for curve 2, with respect to curve 1)
- For piecewise linear curves with proportional hazards:
- a set of points $\left\{\left(X_{1}(1), S_{1}(1)\right),\left(X_{1}(2), S_{1}(2)\right), \ldots\right\}$ (for curve 1)
- hazard ratio (for curve 2 , with respect to curve 1)
- For arbitrary piecewise linear curves:
- a set of points $\left\{\left(X_{j}(1), S_{j}(1)\right),\left(X_{j}(2), S_{j}(2)\right), \ldots\right\}$

A total of $M+1$ evenly spaced time points $\left\{t_{0}=0, t_{1}, t_{2}, \ldots, t_{M}=T+\tau\right\}$ are used in calculations, where

$$
M=\text { floor }((T+\tau) b)
$$

The hazard function is calculated for each survival curve at each time point. For an exponential curve, the (constant) hazard is given by one of the following, depending on the input parameterization:

$$
h_{j}(t)=\left\{\begin{array}{l}
\lambda_{j} \\
\lambda_{1} R \\
\frac{-\log \left(\frac{1}{2}\right)}{m_{j}} \\
\frac{-\log \left(S_{j}(1)\right)}{X_{j}(1)} \\
\frac{-\log \left(S_{1}(1)\right)}{X_{1}(1)} R
\end{array}\right.
$$

For a piecewise linear curve, define the following additional notation:

$$
\begin{aligned}
t_{i}^{-} & =\text {largest input time } X \text { such that } X \leq t_{i} \\
t_{i}^{+} & =\text {smallest input time } X \text { such that } X>t_{i}
\end{aligned}
$$

The hazard is computed by using linear interpolation as follows:

$$
h_{j}\left(t_{i}\right)=\frac{S_{j}\left(t_{i}^{-}\right)-S_{j}\left(t_{i}^{+}\right)}{\left[S_{j}\left(t_{i}^{+}\right)-S_{j}\left(t_{i}^{-}\right)\right]\left[t_{i}-t_{i}^{-}\right]+S_{j}\left(t_{i}^{-}\right)\left[t_{i}^{+}-t_{i}^{-}\right]}
$$

With proportional hazards, the hazard rate of group 2's curve in terms of the hazard rate of group 1's curve is

$$
h_{2}(t)=h_{1}(t) R
$$

Hazard function values $\left\{\Psi_{j}\left(t_{i}\right)\right\}$ for the loss curves are computed in an analogous way from $\left\{L_{j}, X L_{j}, S L_{j}, m L_{j}\right\}$.
The expected number at risk $N_{j}(i)$ at time $i$ in group $j$ is calculated for each group and time points 0 through M-1, as follows:

$$
\begin{aligned}
N_{j}(0) & =N w_{j} \\
N_{j}(i+1) & =N_{j}(i)\left[1-h_{j}\left(t_{i}\right)\left(\frac{1}{b}\right)-\Psi_{j}\left(t_{i}\right)\left(\frac{1}{b}\right)-\left(\frac{1}{b\left(T+\tau-t_{i}\right)}\right) 1_{\left\{t_{i}>\tau\right\}}\right]
\end{aligned}
$$

Define $\theta_{i}$ as the ratio of hazards and $\phi_{i}$ as the ratio of expected numbers at risk for time $t_{i}$ :

$$
\begin{aligned}
\theta_{i} & =\frac{h_{2}\left(t_{i}\right)}{h_{1}\left(t_{i}\right)} \\
\phi_{i} & =\frac{N_{2}(i)}{N_{1}(i)}
\end{aligned}
$$

The expected number of deaths in each subinterval is calculated as follows:

$$
D_{i}=\left[h_{1}\left(t_{i}\right) N_{1}(i)+h_{2}\left(t_{i}\right) N_{2}(i)\right]\left(\frac{1}{b}\right)
$$

The rank values are calculated as follows according to which test statistic is used:

$$
r_{i}= \begin{cases}1, & \text { log-rank } \\ N_{1}(i)+N_{2}(i), & \text { Gehan } \\ \sqrt{N_{1}(i)+N_{2}(i)}, & \text { Tarone-Ware }\end{cases}
$$

The distribution of the test statistic is approximated by $N(E, 1)$ where

$$
E=\frac{\sum_{i=0}^{M-1} D_{i} r_{i}\left[\frac{\phi_{i} \theta_{i}}{1+\phi_{i} \theta_{i}}-\frac{\phi_{i}}{1+\phi_{i}}\right]}{\sqrt{\sum_{i=0}^{M-1} D_{i} r_{i}^{2} \frac{\phi_{i}}{\left(1+\phi_{i}\right)^{2}}}}
$$

Note that $N^{\frac{1}{2}}$ can be factored out of the mean $E$, and so it can be expressed equivalently as

$$
E=N^{\frac{1}{2}} E^{\star}=N^{\frac{1}{2}}\left[\frac{\sum_{i=0}^{M-1} D_{i}^{\star} r_{i}^{\star}\left[\frac{\phi_{i} \theta_{i}}{1+\phi_{i} \theta_{i}}-\frac{\phi_{i}}{1+\phi_{i}}\right]}{\sqrt{\sum_{i=0}^{M-1} D_{i}^{\star} r_{i}^{\star 2} \frac{\phi_{i}}{\left(1+\phi_{i}\right)^{2}}}}\right]
$$

where $E^{\star}$ is free of $N$ and

$$
\begin{aligned}
D_{i}^{\star} & =\left[h_{1}\left(t_{i}\right) N_{1}^{\star}(i)+h_{2}\left(t_{i}\right) N_{2}^{\star}(i)\right]\left(\frac{1}{b}\right) \\
r_{i}^{\star} & = \begin{cases}1, & \text { log-rank } \\
N_{1}^{\star}(i)+N_{2}^{\star}(i), & \text { Gehan } \\
\sqrt{N_{1}^{\star}(i)+N_{2}^{\star}(i),} & \text { Tarone-Ware }\end{cases} \\
N_{j}^{\star}(0) & =w_{j} \\
N_{j}^{\star}(i+1) & =N_{j}^{\star}(i)\left[1-h_{j}\left(t_{i}\right)\left(\frac{1}{b}\right)-\Psi_{j}\left(t_{i}\right)\left(\frac{1}{b}\right)-\left(\frac{1}{b\left(T+\tau-t_{i}\right)}\right) 1_{\left\{t_{i}>\tau\right\}}\right]
\end{aligned}
$$

The approximate power is

$$
\text { power }= \begin{cases}\Phi\left(-N^{\frac{1}{2}} E^{\star}-z_{1-\alpha}\right), & \text { upper one-sided } \\ \Phi\left(N^{\frac{1}{2}} E^{\star}-z_{1-\alpha}\right), & \text { lower one-sided } \\ \Phi\left(-N^{\frac{1}{2}} E^{\star}-z_{1-\frac{\alpha}{2}}\right)+\Phi\left(N^{\frac{1}{2}} E^{\star}-z_{1-\frac{\alpha}{2}}\right), & \text { two-sided }\end{cases}
$$

Note that the upper and lower one-sided cases are expressed differently than in other analyses. This is because $E^{\star}>0$ corresponds to a higher survival curve in group 1 and thus, by the convention used in PROC power for two-group analyses, the lower side.
For the one-sided cases, a closed-form inversion of the power equation yield an approximate total sample size

$$
N=\left(\frac{z_{\text {power }}+z_{1-\alpha}}{E^{\star}}\right)^{2}
$$

For the two-sided case, the solution for $N$ is obtained by numerically inverting the power equation.
Accrual rates are converted to and from sample sizes according to the equation $a_{j}=n_{j} / T$, where $a_{j}$ is the accrual rate for group $j$.

Expected numbers of events-that is, deaths, whether observed or censored-are converted to and from sample sizes according to the equation

$$
e_{j}= \begin{cases}n_{j}\left[1-S_{j}(\tau)\right], & T=0 \\ n_{j}\left[1-\frac{1}{T} \int_{0}^{T} S_{j}(T+\tau-t) d t\right], & T>0\end{cases}
$$

where $e_{j}$ is the expected number of events in group $j$. For an exponential curve, the equation simplifies to

$$
e_{j}= \begin{cases}n_{j}\left[1-\exp \left(-\lambda_{j} \tau\right)\right], & T=0 \\ n_{j}\left[1-\frac{1}{\lambda_{j} T}\left(\exp \left(-\lambda_{j} \tau\right)-\exp \left(-\lambda_{j}(T+\tau)\right)\right)\right], & T>0\end{cases}
$$

For a piecewise linear curve, first define $K_{j}$ as the number of time points in the following collection: $\tau$, $T+\tau$, and input time points for group $j$ strictly between $\tau$ and $T+\tau$. Denote the ordered set of these points as $\left\{u_{j 1}, \ldots, u_{j K_{j}}\right\}$. The survival function values $S_{j}(\tau)$ and $S_{j}(T+\tau)$ are calculated by linear interpolation between adjacent input time points if they do not coincide with any input time points. Then the equation for a piecewise linear curve simplifies to

$$
e_{j}= \begin{cases}n_{j}\left[1-S_{j}(\tau)\right], & T=0 \\ n_{j}\left[1-\frac{1}{2 T} \sum_{i=1}^{K_{j}-1}\left(u_{j, i+1}-u_{j i}\right)\left(S_{j}\left(u_{j i}\right)+S_{j}\left(u_{j, i+1}\right)\right)\right], & T>0\end{cases}
$$

## Analyses in the TWOSAMPLEWILCOXON Statement

## Wilcoxon-Mann-Whitney Test for Comparing Two Distributions (TEST=WMW)

The power approximation in this section is applicable to the Wilcoxon-Mann-Whitney (WMW) test as invoked with the WILCOXON option in the PROC NPAR1WAY statement of the NPAR1WAY procedure. The approximation is based on O'Brien and Castelloe (2006) and an estimator called $\widehat{\mathrm{WMW}_{\text {odds }}}$. See O'Brien and Castelloe (2006) for a definition of $\mathrm{WMW}_{\text {odds }}$, which need not be derived in detail here for purposes of explaining the power formula.

Let $Y_{1}$ and $Y_{2}$ be independent observations from any two distributions that you want to compare using the WMW test. For purposes of deriving the asymptotic distribution of $\widehat{\mathrm{WMW}_{\text {odds }}}$ (and consequently the power computation as well), these distributions must be formulated as ordered categorical ("ordinal") distributions.

If a distribution is continuous, it can be discretized using a large number of categories with negligible loss of accuracy. Each nonordinal distribution is divided into $b$ categories, where $b$ is the value of the NBINS parameter, with breakpoints evenly spaced on the probability scale. That is, each bin contains an equal probability $1 / b$ for that distribution. Then the breakpoints across both distributions are pooled to form a collection of $C$ bins (heretofore called "categories"), and the probabilities of bin membership for each distribution are recalculated. The motivation for this method of binning is to avoid degenerate representations of the distributions-that is, small handfuls of large probabilities among mostly empty bins-as can be caused by something like an evenly spaced grid across raw values rather than probabilities.

After the discretization process just mentioned, there are now two ordinal distributions, each with a set of probabilities across a common set of $C$ ordered categories. For simplicity of notation, assume (without loss of generality) the response values to be $1, \ldots, C$. Represent the conditional probabilities as

$$
\tilde{p}_{i j}=\operatorname{Prob}\left(Y_{i}=j \mid \operatorname{group}=i\right), i \in\{1,2\} \quad \text { and } \quad j \in\{1, \ldots, C\}
$$

and the group allocation weights as

$$
w_{i}=\frac{n_{i}}{N}=\operatorname{Prob}(\text { group }=i), \quad i \in\{1,2\}
$$

The joint probabilities can then be calculated simply as

$$
p_{i j}=\operatorname{Prob}\left(\operatorname{group}=i, Y_{i}=j\right)=w_{i} \tilde{p}_{i j}, i \in\{1,2\} \quad \text { and } \quad j \in\{1, \ldots, C\}
$$

The next step in the power computation is to compute the probabilities that a randomly chosen pair of observations from the two groups is concordant, discordant, or tied. It is useful to define these probabilities as functions of the terms $R s_{i j}$ and $R d_{i j}$, defined as follows, where $Y$ is a random observation drawn from the joint distribution across groups and categories:

$$
\begin{aligned}
R s_{i j}= & \operatorname{Prob}(Y \text { is concordant with cell }(i, j))+\frac{1}{2} \operatorname{Prob}(Y \text { is tied with cell }(i, j)) \\
= & \operatorname{Prob}((\text { group }<i \text { and } Y<j) \text { or }(\text { group }>i \text { and } Y>j))+ \\
& \frac{1}{2} \operatorname{Prob}(\text { group } \neq i \text { and } Y=j) \\
= & \sum_{g=1}^{2} \sum_{c=1}^{C} w_{g} \tilde{p}_{g c}\left[\mathrm{I}_{(g-i)(c-j)>0}+\frac{1}{2} \mathrm{I}_{g \neq i, c=j}\right]
\end{aligned}
$$

and

$$
\begin{aligned}
R d_{i j}= & \operatorname{Prob}(Y \text { is discordant with cell }(i, j))+\frac{1}{2} \operatorname{Prob}(Y \text { is tied with cell }(i, j)) \\
= & \operatorname{Prob}((\text { group }<i \text { and } Y>j) \text { or }(\text { group }>i \text { and } Y<j))+ \\
& \frac{1}{2} \operatorname{Prob}(\text { group } \neq i \text { and } Y=j) \\
= & \sum_{g=1}^{2} \sum_{c=1}^{C} w_{g} \tilde{g}_{g c}\left[\mathrm{I}_{(g-i)(c-j)<0}+\frac{1}{2} \mathrm{I}_{g \neq i, c=j}\right]
\end{aligned}
$$

For an independent random draw $Y_{1}, Y_{2}$ from the two distributions,

$$
\begin{aligned}
P_{c} & =\operatorname{Prob}\left(Y_{1}, Y_{2} \text { concordant }\right)+\frac{1}{2} \operatorname{Prob}\left(Y_{1}, Y_{2} \text { tied }\right) \\
& =\sum_{i=1}^{2} \sum_{j=1}^{C} w_{i} \tilde{p}_{i j} R s_{i j}
\end{aligned}
$$

and

$$
\begin{aligned}
P_{d} & =\operatorname{Prob}\left(Y_{1}, Y_{2} \text { discordant }\right)+\frac{1}{2} \operatorname{Prob}\left(Y_{1}, Y_{2} \text { tied }\right) \\
& =\sum_{i=1}^{2} \sum_{j=1}^{C} w_{i} \tilde{p}_{i j} R d_{i j}
\end{aligned}
$$

Then

$$
\mathrm{WMW}_{\text {odds }}=\frac{P_{c}}{P_{d}}
$$

Proceeding to compute the theoretical standard error associated with $\mathrm{WMW}_{\text {odds }}$ (that is, the population analogue to the sample standard error),

$$
\mathrm{SE}\left(\mathrm{WMW}_{\text {odds }}\right)=\frac{2}{P_{d}}\left[\sum_{i=1}^{2} \sum_{j=1}^{C} w_{i} \tilde{p}_{i j}\left(\mathrm{WMW}_{\text {odds }} R d_{i j}-R s_{i j}\right)^{2} / N\right]^{\frac{1}{2}}
$$

Converting to the natural log scale and using the delta method,

$$
\mathrm{SE}\left(\log \left(\mathrm{WMW}_{\text {odds }}\right)\right)=\frac{\mathrm{SE}\left(\mathrm{WMW}_{\text {odds }}\right)}{\mathrm{WMW}_{\text {odds }}}
$$

The next step is to produce a "smoothed" version of the $2 \times C$ cell probabilities that conforms to the null hypothesis of the Wilcoxon-Mann-Whitney test (in other words, independence in the $2 \times C$ contingency table of probabilities). Let $\mathrm{SE}_{H_{0}}\left(\log \left(\mathrm{WMW}_{\text {odds }}\right)\right)$ denote the theoretical standard error of $\log \left(\mathrm{WMW}_{\text {odds }}\right)$ assuming $H_{0}$.

Finally, compute the power using the noncentral chi-square and normal distributions:

$$
\text { power }=\left\{\begin{array}{l}
P\left(Z \geq \frac{\mathrm{SE}_{H_{0}}\left(\log \left(\mathrm{WMW}_{\text {odds }}\right)\right)}{\mathrm{SE}\left(\log \left(\mathrm{WMW} \mathrm{odds}^{2}\right)\right)} z_{1-\alpha}-\delta^{\star} N^{\frac{1}{2}}\right), \quad \text { upper one-sided } \\
P\left(Z \leq \frac{\mathrm{SE}_{H_{0}}\left(\log \left(\mathrm{WMW} \mathrm{~W}_{\text {odds }}\right)\right)}{\mathrm{SE}\left(\log \left(\mathrm{WMW} \mathrm{odds}^{2}\right)\right)} z_{\alpha}-\delta^{\star} N^{\frac{1}{2}}\right), \quad \text { lower one-sided } \\
P\left(\chi^{2}\left(1,\left(\delta^{\star}\right)^{2} N\right) \geq\left[\frac{\mathrm{SE}_{H_{0}}\left(\log \left(\mathrm{WMW}_{\text {odds }}\right)\right)}{\mathrm{SE}\left(\log \left(\mathrm{WM} W_{\text {odds }}\right)\right)}\right]^{2} \chi_{1-\alpha}^{2}(1)\right), \quad \text { two-sided }
\end{array}\right.
$$

where

$$
\delta^{\star}=\frac{\log \left(\mathrm{WMW}_{\text {odds }}\right)}{N^{\frac{1}{2}} \mathrm{SE}\left(\log \left(\mathrm{WMW}_{\text {odds }}\right)\right)}
$$

is the primary noncentrality-that is, the "effect size" that quantifies how much the two conjectured distributions differ. $Z$ is a standard normal random variable, $\chi^{2}(d f, n c)$ is a noncentral $\chi^{2}$ random variable with degrees of freedom $d f$ and noncentrality $n c$, and $N$ is the total sample size.

## 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, by specifying the ODS GRAPHICS ON statement). For more information about enabling and disabling ODS Graphics, see the section "Enabling and Disabling ODS Graphics" on page 606 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 605 in Chapter 21, "Statistical Graphics Using ODS."

If ODS Graphics is not enabled, then PROC POWER creates traditional graphics.
You can reference every graph produced through ODS Graphics with a name. The names of the graphs that PROC POWER generates are listed in Table 77.31, along with the required statements and options.

Table 77.31 Graphs Produced by PROC POWER

| ODS Graph Name | Plot Description | Option |
| :--- | :--- | :--- |
| PowerPlot | Plot with two of the following three parameters on the <br> X and Y axes: power, sample size, and effect size | PLOT |
| PowerAbort | Empty plot that shows an error message when a plot <br> could not be produced | PLOT |

## Examples: POWER Procedure

## Example 77.1: One-Way ANOVA

This example deals with the same situation as in Example 47.1 of Chapter 47, "The GLMPOWER Procedure."
Hocking (1985, p. 109) describes a study of the effectiveness of electrolytes in reducing lactic acid buildup for long-distance runners. You are planning a similar study in which you will allocate five different fluids to runners on a 10 -mile course and measure lactic acid buildup immediately after the run. The fluids consist of water and two commercial electrolyte drinks, EZDure and LactoZap, each prepared at two concentrations, low (EZD1 and LZ1) and high (EZD2 and LZ2).

You conjecture that the standard deviation of lactic acid measurements given any particular fluid is about 3.75, and that the expected lactic acid values will correspond roughly to those in Table 77.32. You are least familiar with the LZ1 drink and hence decide to consider a range of reasonable values for that mean.

Table 77.32 Mean Lactic Acid Buildup by Fluid

| Water | EZD1 | EZD2 | LZ1 | LZ2 |
| :---: | :---: | :---: | :---: | :---: |
| 35.6 | 33.7 | 30.2 | 29 or 28 | 25.9 |

You are interested in four different comparisons, shown in Table 77.33 with appropriate contrast coefficients.
Table 77.33 Planned Comparisons

|  | Contrast Coefficients |  |  |  |  |
| :--- | ---: | ---: | ---: | ---: | ---: |
| Comparison | Water | EZD1 | EZD2 | LZ1 | LZ2 |
| Water versus electrolytes | 4 | -1 | -1 | -1 | -1 |
| EZD versus LZ | 0 | 1 | 1 | -1 | -1 |
| EZD1 versus EZD2 | 0 | 1 | -1 | 0 | 0 |
| LZ1 versus LZ2 | 0 | 0 | 0 | 1 | -1 |

For each of these contrasts you want to determine the sample size required to achieve a power of 0.9 for detecting an effect with magnitude in accord with Table 77.32. You are not yet attempting to choose a single sample size for the study, but rather checking the range of sample sizes needed for individual contrasts. You plan to test each contrast at $\alpha=0.025$. In the interests of reducing costs, you will provide twice as many runners with water as with any of the electrolytes; in other words, you will use a sample size weighting scheme of 2:1:1:1:1. Use the ONEWAYANOVA statement in the POWER procedure to compute the sample sizes.

The statements required to perform this analysis are as follows:

```
proc power;
    onewayanova
        groupmeans = 35.6 | 33.7 | 30.2 | 29 28 | 25.9
        stddev = 3.75
        groupweights =(\begin{array}{llllll}{2}&{1}&{1}&{1}&{1}\end{array})
        alpha = 0.025
        ntotal = .
        power = 0.9
        contrast =( (\begin{array}{llllll}{4}&{-1}&{-1}&{-1}&{-1}\end{array})(\begin{array}{llllll}{0}&{1}&{1}&{-1}&{-1}\end{array})
            (0 1 - -1 0
run;
```

The NTOTAL= option with a missing value (.) indicates total sample size as the result parameter. The GROUPMEANS = option with values from Table 77.32 specifies your conjectures for the means. With only one mean varying (the LZ1 mean), the "crossed" notation is simpler, showing scenarios for each group mean, separated by vertical bars (I). For more information about crossed and matched notations for grouped values, see the section "Specifying Value Lists in Analysis Statements" on page 6366. The contrasts in Table 77.33 are specified with the CONTRAST= option, by using the "matched" notation with each contrast enclosed in parentheses. The STDDEV=, ALPHA=, and POWER= options specify the error standard deviation, significance level, and power. The GROUPWEIGHTS= option specifies the weighting schemes. Default values for the NULLCONTRAST $=$ and SIDES $=$ options specify a two-sided $t$ test of the contrast equal to 0 . See Output 77.1.1 for the results.

Output 77.1.1 Sample Sizes for One-Way ANOVA Contrasts
The POWER Procedure Single DF Contrast in One-Way ANOVA

| Fixed Scenario Elements |  |
| :--- | ---: |
| Method | Exact |
| Alpha | 0.025 |
| Standard Deviation | 3.75 |
| Group Weights | 21111 |
| Nominal Power | 0.9 |
| Number of Sides | 2 |
| Null Contrast Value | 0 |

## Output 77.1.1 continued

| Computed N Total |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Contrast |  |  | Means |  |  |  |  | Actual Power |  |
| 1 | 4 | -1-1 | -1 -1 | 35.6 | 33.7 | 30.2 | 29 | 25.9 | 0.947 | 30 |
| 2 | 4 | -1-1 | -1 -1 | 35.6 | 33.7 | 30.2 | 28 | 25.9 | 0.901 | 24 |
| 3 | 0 | 11 | -1 -1 | 35.6 | 33.7 | 30.2 |  | 25.9 | 0.929 | 60 |
| 4 | 0 | 11 | -1 -1 | 35.6 | 33.7 | 30.2 | 28 | 25.9 | 0.922 | 48 |
| 5 | 0 | 1-1 | 0 | 035.6 | 33.7 | 30.2 | 29 | 25.9 | 0.901 | 174 |
| 6 | 0 | 1-1 | 0 | 035.6 | 33.7 | 30.2 |  | 25.9 | 0.901 | 174 |
| 7 | 0 | 00 | $1-1$ | 135.6 | 33.7 | 30.2 | 29 | 25.9 | 0.902 | 222 |
| 8 | 0 | 00 | 1-1 | 135.6 | 33.7 | 30.2 | 28 | 25.9 | 0.902 | 480 |

The sample sizes in Output 77.1.1 range from 24 for the comparison of water versus electrolytes to 480 for the comparison of LZ1 versus LZ2, both assuming the smaller LZ1 mean. The sample size for the latter comparison is relatively large because the small mean difference of $28-25.9=2.1$ is hard to detect.

The Nominal Power of 0.9 in the "Fixed Scenario Elements" table in Output 77.1.1 represents the input target power, and the Actual Power column in the "Computed N Total" table is the power at the sample size ( N Total) adjusted to achieve the specified sample weighting. Note that all of the sample sizes are rounded up to multiples of 6 to preserve integer group sizes (since the group weights add up to 6). You can use the NFRACTIONAL option in the ONEWAYANOVA statement to compute raw fractional sample sizes.

Suppose you want to plot the required sample size for the range of power values from 0.5 to 0.95 . First, define the analysis by specifying the same statements as before, but add the PLOTONLY option to the PROC POWER statement to disable the nongraphical results. Next, specify the PLOT statement with X=POWER to request a plot with power on the X axis. (The result parameter, here sample size, is always plotted on the other axis.) Use the MIN= and MAX= options in the PLOT statement to specify the power range. The following statements produce the plot shown in Output 77.1.2.

```
ods graphics on;
proc power plotonly;
    onewayanova
        groupmeans = 35.6 | 33.7 | 30.2 | 29 28 | 25.9
        stddev = 3.75
        groupweights = (\begin{array}{lllll}{2}&{1}&{1}&{1}&{1}\end{array})
        alpha = 0.025
        ntotal = .
        power = 0.9
            contrast = (\begin{array}{llllll}{4}&{-1}&{-1}&{-1}&{-1}\end{array})(\begin{array}{lllll}{0}&{1}&{1}&{-1}&{-1}\end{array})
                    (0 1-1 -1 0
    plot x=power min=.5 max=.95;
run;
```

Output 77.1.2 Plot of Sample Size versus Power for One-Way ANOVA Contrasts


In Output 77.1.2, the line style identifies the contrast, and the plotting symbol identifies the group means scenario. The plot shows that the required sample size is highest for the (0001-1) contrast, corresponding to the test of LZ1 versus LZ2 that was previously found to require the most resources, in either cell means scenario.

Note that some of the plotted points in Output 77.1.2 are unevenly spaced. This is because the plotted points are the rounded sample size results at their corresponding actual power levels. The range specified with the MIN= and MAX= values in the PLOT statement corresponds to nominal power levels. In some cases, actual power is substantially higher than nominal power. To obtain plots with evenly spaced points (but with fractional sample sizes at the computed points), you can use the NFRACTIONAL option in the analysis statement preceding the PLOT statement.

Finally, suppose you want to plot the power for the range of sample sizes you will likely consider for the study (the range of 24 to 480 that achieves 0.9 power for different comparisons). In the ONEWAYANOVA statement, identify power as the result (POWER=.), and specify NTOTAL=24. The following statements produce the plot:

```
proc power plotonly;
    onewayanova
        groupmeans = 35.6 | 33.7 | 30.2 | 29 28 | 25.9
        stddev = 3.75
        groupweights = (\begin{array}{lllll}{2}&{1}&{1}&{1}&{1}\end{array})
        alpha = 0.025
        ntotal = 24
        power = .
        contrast = (\begin{array}{rrrrrrlllll}{4}&{-1}&{-1}&{-1}&{-1)}&{(0}&{1}&{1}&{-1}&{-1)}\\{(0}&{1}&{-1}&{0}&{0)}&{(0}&{0}&{0}&{1}&{-1}\end{array})
    plot x=n min=24 max=480;
run;
ods graphics off;
```

The $\mathrm{X}=\mathrm{N}$ option in the PLOT statement requests a plot with sample size on the X axis.
Note that the value specified with the NTOTAL=24 option is not used. It is overridden in the plot by the MIN= and MAX= options in the PLOT statement, and the PLOTONLY option in the PROC POWER statement disables nongraphical results. But the NTOTAL= option (along with a value) is still needed in the ONEWAYANOVA statement as a placeholder, to identify the desired parameterization for sample size.

Output 77.1.3 shows the resulting plot.

Output 77.1.3 Plot of Power versus Sample Size for One-Way ANOVA Contrasts


Although Output 77.1.2 and Output 77.1.3 surface essentially the same computations for practical power ranges, they each provide a different quick visual assessment. Output 77.1.2 reveals the range of required sample sizes for powers of interest, and Output 77.1.3 reveals the range of achieved powers for sample sizes of interest.

## Example 77.2: The Sawtooth Power Function in Proportion Analyses

For many common statistical analyses, the power curve is monotonically increasing: the more samples you take, the more power you achieve. However, in statistical analyses of discrete data, such as tests of proportions, the power curve is often nonmonotonic. A small increase in sample size can result in a decrease in power, a decrease that is sometimes substantial. The explanation is that the actual significance level (in other words, the achieved Type I error rate) for discrete tests strays below the target level and varies with sample size. The power loss from a decrease in the Type I error rate can outweigh the power gain from an increase in sample size. The example discussed here demonstrates this "sawtooth" phenomenon. For additional discussion on the topic, see Chernick and Liu (2002).

Suppose you have a new scheduling system for an airline, and you want to determine how many flights you must observe to have at least an $80 \%$ chance of establishing an improvement in the proportion of late arrivals
on a specific travel route. You will use a one-sided exact binomial proportion test with a null proportion of $30 \%$, the frequency of late arrivals under the previous scheduling system, and a nominal significance level of $\alpha=0.05$. Well-supported predictions estimate the new late arrival rate to be about $20 \%$, and you will base your sample size determination on this assumption.
The POWER procedure does not currently compute exact sample size directly for the exact binomial test. But you can get an initial estimate by computing the approximate sample size required for a $z$ test. Use the ONESAMPLEFREQ statement in the POWER procedure with TEST=Z and METHOD=NORMAL to compute the approximate sample size to achieve a power of 0.8 by using the $z$ test. The following statements perform the analysis:

```
proc power;
    onesamplefreq test=z method=normal
        sides = 1
        alpha = 0.05
        nullproportion = 0.3
        proportion = 0.2
        ntotal = .
        power = 0.8;
run;
```

The NTOTAL= option with a missing value (.) indicates sample size as the result parameter. The SIDES=1 option specifies a one-sided test. The ALPHA=, NULLPROPORTION=, and POWER= options specify the significance level of 0.05 , null value of 0.3 , and target power of 0.8 , respectively. The PROPORTION= option specifies your conjecture of 0.3 for the true proportion.

Output 77.2.1 Approximate Sample Size for $z$ Test of a Proportion
The POWER Procedure Z Test for Binomial Proportion

| Fixed Scenario Elements |  |
| :---: | :---: |
| Method Normal | approximation |
| Number of Sides | 1 |
| Null Proportion | 0.3 |
| Alpha | 0.05 |
| Binomial Proportion | 0.2 |
| Nominal Power | 0.8 |
| Variance Estimate | Null Variance |
| Computed N Total |  |
| Actual $\mathbf{N}$ <br> Power Total |  |
| 0.800119 |  |

The results, shown in Output 77.2.1, indicate that you need to observe about $N=119$ flights to have an $80 \%$ chance of rejecting the hypothesis of a late arrival proportion of $30 \%$ or higher, if the true proportion is $20 \%$, by using the $z$ test. A similar analysis (Output 77.2.2) reveals an approximate sample size of $N=129$ for the $z$ test with continuity correction, which is performed by using TEST=ADJZ:

```
proc power;
    onesamplefreq test=adjz method=normal
    sides = 1
    alpha = 0.05
        nullproportion = 0.3
        proportion = 0.2
        ntotal = .
        power = 0.8;
run;
```

Output 77.2.2 Approximate Sample Size for $z$ Test with Continuity Correction
The POWER Procedure Z Test for Binomial Proportion with Continuity Adjustment

| Fixed Scenario Elements |  |
| :--- | ---: |
| Method | Normal approximation |
| Number of Sides | 1 |
| Null Proportion | 0.3 |
| Alpha | 0.05 |
| Binomial Proportion | 0.2 |
| Nominal Power | 0.8 |
| Variance Estimate | Null Variance |


| Computed $\mathbf{N}$ <br> Total |
| :--- |
| Actual <br> Power Total |
| 0.801 |

Based on the approximate sample size results, you decide to explore the power of the exact binomial test for sample sizes between 110 and 140. The following statements produce the plot:

```
ods graphics on;
proc power plotonly;
    onesamplefreq test=exact
        sides = 1
        alpha = 0.05
        nullproportion = 0.3
        proportion = 0.2
        ntotal = 119
        power = .;
    plot x=n min=110 max=140 step=1
        yopts=(ref=.8) xopts=(ref=119 129);
run;
```

The TEST=EXACT option in the ONESAMPLEFREQ statement specifies the exact binomial test, and the missing value (.) for the POWER= option indicates power as the result parameter. The PLOTONLY option in the PROC POWER statement disables nongraphical output. The PLOT statement with $\mathrm{X}=\mathrm{N}$ requests a plot with sample size on the X axis. The MIN= and MAX= options in the PLOT statement specify the sample size range. The YOPTS $=(\mathrm{REF}=)$ and $\mathrm{XOPTS}=(\mathrm{REF}=)$ options add reference lines to highlight the approximate sample size results. The $\mathrm{STEP}=1$ option produces a point at each integer sample size. The sample size value
specified with the NTOTAL= option in the ONESAMPLEFREQ statement is overridden by the MIN= and MAX= options in the PLOT statement. Output 77.2.3 shows the resulting plot.

Output 77.2.3 Plot of Power versus Sample Size for Exact Binomial Test


Note the sawtooth pattern in Output 77.2.3. Although the power surpasses the target level of 0.8 at $N$ $=119$, it decreases to 0.79 with $N=120$ and further to 0.76 with $N=122$ before rising again to 0.81 with $N=123$. Not until $N=130$ does the power stay above the 0.8 target. Thus, a more conservative sample size recommendation of 130 might be appropriate, depending on the precise goals of the sample size determination.

In addition to considering alternative sample sizes, you might also want to assess the sensitivity of the power to inaccuracies in assumptions about the true proportion. The following statements produce a plot including true proportion values of 0.18 and 0.22 . They are identical to the previous statements except for the additional true proportion values specified with the PROPORTION= option in the ONESAMPLEFREQ statement.

```
proc power plotonly;
    onesamplefreq test=exact
        sides = 1
        alpha = 0.05
        nullproportion = 0.3
        proportion = 0.18 0.2 0.22
        ntotal = 119
        power = .;
    plot x=n min=110 max=140 step=1
        yopts=(ref=.8) xopts=(ref=119 129);
run;
```

Output 77.2.4 shows the resulting plot.
Output 77.2.4 Plot for Assessing Sensitivity to True Proportion Value


The plot reveals a dramatic sensitivity to the true proportion value. For $N=119$, the power is about 0.92 if the true proportion is 0.18 , and as low as 0.62 if the proportion is 0.22 . Note also that the power jumps occur at the same sample sizes in all three curves; the curves are only shifted and stretched vertically. This is because spikes and valleys in power curves are invariant to the true proportion value; they are due to changes in the critical value of the test.

A closer look at some ancillary output from the analysis sheds light on this property of the sawtooth pattern. You can add an ODS OUTPUT statement to save the plot content corresponding to Output 77.2.3 to a data set:

```
proc power plotonly;
    ods output plotcontent=PlotData;
    onesamplefreq test=exact
            sides = 1
            alpha =0.05
            nullproportion = 0.3
            proportion = 0.2
            ntotal = 119
            power = .;
    plot x=n min=110 max=140 step=1
            yopts=(ref=.8) xopts=(ref=119 129);
run;
```

The PlotData data set contains parameter values for each point in the plot. The parameters include underlying characteristics of the putative test. The following statements print the critical value and actual significance level along with sample size and power:

```
proc print data=PlotData;
    var NTotal LowerCritVal Alpha Power;
run;
```

Output 77.2.5 shows the plot data.

Output 77.2.5 Numerical Content of Plot

| Obs | NTotal LowerCritVal | Alpha | Power |  |
| ---: | ---: | ---: | ---: | ---: |
| $\mathbf{1}$ | 110 | 24 | 0.0356 | 0.729 |
| $\mathbf{2}$ | 111 | 24 | 0.0313 | 0.713 |
| $\mathbf{3}$ | 112 | 25 | 0.0446 | 0.771 |
| $\mathbf{4}$ | 113 | 25 | 0.0395 | 0.756 |
| $\mathbf{5}$ | 114 | 25 | 0.0349 | 0.741 |
| $\mathbf{6}$ | 115 | 26 | 0.0490 | 0.795 |
| $\mathbf{7}$ | 116 | 26 | 0.0435 | 0.781 |
| $\mathbf{8}$ | 117 | 26 | 0.0386 | 0.767 |
| $\mathbf{9}$ | 118 | 26 | 0.0341 | 0.752 |
| $\mathbf{1 0}$ | 119 | 27 | 0.0478 | 0.804 |
| $\mathbf{1 1}$ | 120 | 27 | 0.0425 | 0.790 |
| $\mathbf{1 2}$ | 121 | 27 | 0.0377 | 0.776 |
| $\mathbf{1 3}$ | 122 | 27 | 0.0334 | 0.762 |
| $\mathbf{1 4}$ | 123 | 28 | 0.0465 | 0.812 |
| $\mathbf{1 5}$ | 124 | 28 | 0.0414 | 0.799 |
| $\mathbf{1 6}$ | 125 | 28 | 0.0368 | 0.786 |
| $\mathbf{1 7}$ | 126 | 28 | 0.0327 | 0.772 |
| $\mathbf{1 8}$ | 127 | 29 | 0.0453 | 0.820 |
| $\mathbf{1 9}$ | 128 | 29 | 0.0404 | 0.807 |
| $\mathbf{2 0}$ | 129 | 29 | 0.0359 | 0.794 |
| $\mathbf{2 1}$ | 130 | 30 | 0.0493 | 0.838 |
| $\mathbf{2 2}$ | 131 | 30 | 0.0441 | 0.827 |
| $\mathbf{2 3}$ | 132 | 30 | 0.0394 | 0.815 |
| $\mathbf{2 4}$ | 133 | 30 | 0.0351 | 0.803 |
| $\mathbf{2 5}$ | 134 | 31 | 0.0480 | 0.845 |
| $\mathbf{2 6}$ | 135 | 31 | 0.0429 | 0.834 |
| $\mathbf{2 7}$ | 136 | 31 | 0.0384 | 0.823 |
| $\mathbf{2 8}$ | 137 | 31 | 0.0342 | 0.811 |
| $\mathbf{2 9}$ | 138 | 32 | 0.0466 | 0.851 |
| $\mathbf{3 0}$ | 139 | 32 | 0.0418 | 0.841 |
| $\mathbf{3 1}$ | 140 | 32 | 0.0374 | 0.830 |

Note that whenever the critical value changes, the actual $\alpha$ jumps up to a value close to the nominal $\alpha=$ 0.05 , and the power also jumps up. Then while the critical value stays constant, the actual $\alpha$ and power slowly decrease. The critical value is independent of the true proportion value. So you can achieve a locally maximal power by choosing a sample size corresponding to a spike on the sawtooth curve, and this choice is locally optimal regardless of the unknown value of the true proportion. Locally optimal sample sizes in this case include 115, 119, 123, 127, 130, and 134.

As a point of interest, the power does not always jump sharply and decrease gradually. The shape of the sawtooth depends on the direction of the test and the location of the null proportion relative to 0.5 . For example, if the direction of the hypothesis in this example is reversed (by switching true and null proportion values) so that the rejection region is in the upper tail, then the power curve exhibits sharp decreases and gradual increases. The following statements are similar to those producing the plot in Output 77.2.3 but with values of the PROPORTION= and NULLPROPORTION= options switched:

```
proc power plotonly;
    onesamplefreq test=exact
        sides = 1
        alpha = 0.05
        nullproportion = 0.2
        proportion = 0.3
        ntotal = 119
        power = .;
    plot x=n min=110 max=140 step=1;
run;
```

The resulting plot is shown in Output 77.2.6.
Output 77.2.6 Plot of Power versus Sample Size for Another One-sided Test


Finally, two-sided tests can lead to even more irregular power curve shapes, since changes in lower and upper critical values affect the power in different ways. The following statements produce a plot of power versus sample size for the scenario of a two-sided test with high alpha and a true proportion close to the null value:

```
proc power plotonly;
    onesamplefreq test=exact
    sides = 2
    alpha = 0.2
    nullproportion = 0.1
    proportion = 0.09
    ntotal = 10
    power = .;
    plot x=n min=2 max=100 step=1;
run;
ods graphics off;
```

Output 77.2.7 shows the resulting plot.
Output 77.2.7 Plot of Power versus Sample Size for a Two-Sided Test


Due to the irregular shapes of power curves for proportion tests, the question "Which sample size should I use?" is often insufficient. A sample size solution produced directly in PROC POWER reveals the smallest possible sample size to achieve your target power. But as the examples in this section demonstrate, it is helpful to consult graphs for answers to questions such as the following:

- Which sample size will guarantee that all higher sample sizes also achieve my target power?
- Given a candidate sample size, can I increase it slightly to achieve locally maximal power, or perhaps even decrease it and get higher power?


## Example 77.3: Simple AB/BA Crossover Designs

Crossover trials are experiments in which each subject is given a sequence of different treatments. They are especially common in clinical trials for medical studies. The reduction in variability from taking multiple measurements on a subject allows for more precise treatment comparisons. The simplest such design is the $\mathrm{AB} / \mathrm{BA}$ crossover, in which each subject receives each of two treatments in a randomized order.
Under certain simplifying assumptions, you can test the treatment difference in an $A B / B A$ crossover trial by using either a paired or two-sample $t$ test (or equivalence test, depending on the hypothesis). This example will demonstrate when and how you can use the PAIREDMEANS statement in PROC POWER to perform power analyses for $\mathrm{AB} / \mathrm{BA}$ crossover designs.

Senn (1993, Chapter 3) discusses a study comparing the effects of two bronchodilator medications in treatment of asthma, by using an $A B / B A$ crossover design. Suppose you want to plan a similar study comparing two new medications, "Xilodol" and "Brantium." Half of the patients would be assigned to sequence AB, getting a dose of Xilodol in the first treatment period, a wash-out period of one week, and then a dose of Brantium in the second treatment period. The other half would be assigned to sequence BA, following the same schedule but with the drugs reversed. In each treatment period you would administer the drugs in the morning and then measure peak expiratory flow (PEF) at the end of the day, with higher PEF representing better lung function.

You conjecture that the mean and standard deviation of PEF are about $\mu_{A}=330$ and $\sigma_{A}=40$ for Xilodol and $\mu_{B}=310$ and $\sigma_{B}=55$ for Brantium, and that each pair of measurements on the same subject will have a correlation of about 0.3 . You want to compute the power of both one-sided and two-sided tests of mean difference, with a significance level of $\alpha=0.01$, for a sample size of 100 patients and also plot the power for a range of 50 to 200 patients. Note that the allocation ratio of patients to the two sequences is irrelevant in this analysis.

The choice of statistical test depends on which assumptions are reasonable. One possibility is a $t$ test. A paired or two-sample $t$ test is valid when there is no carryover effect and no interactions between patients, treatments, and periods. See Senn (1993, Chapter 3) for more details. The choice between a paired or a two-sample test depends on what you assume about the period effect. If you assume no period effect, then a paired $t$ test is the appropriate analysis for the design, with the first member of each pair being the Xilodol measurement (regardless of which sequence the patient belongs to). Otherwise, the two-sample $t$ test approach is called for, since this analysis adjusts for the period effect by using an extra degree of freedom.
Suppose you assume no period effect. Then you can use the PAIREDMEANS statement in PROC POWER with the TEST=DIFF option to perform a sample size analysis for the paired $t$ test. Indicate power as the result parameter by specifying the POWER= option with a missing value (.). Specify the conjectured means and standard deviations for each drug by using the PAIREDMEANS = and PAIREDSTDDEVS= options and the correlation by using the CORR= option. Specify both one- and two-sided tests by using the SIDES= option, the significance level by using the ALPHA= option, and the sample size (in terms of number of pairs) by using the NPAIRS = option. Generate a plot of power versus sample size by specifying the PLOT statement with $\mathrm{X}=\mathrm{N}$ to request a plot with sample size on the X axis. (The result parameter, here power, is always plotted on the other axis.) Use the MIN= and MAX= options in the PLOT statement to specify the sample size range (as numbers of pairs).

The following statements perform the sample size analysis:

```
ods graphics on;
proc power;
    pairedmeans test=diff
        pairedmeans =( 330 310)
        pairedstddevs = (40 55)
        corr = 0.3
        sides = 1 2
        alpha = 0.01
            npairs = 100
            power = .;
    plot x=n min=50 max=200;
run;
ods graphics off;
```

Default values for the NULLDIFF= and DIST= options specify a null mean difference of 0 and the assumption of normally distributed data. The output is shown in Output 77.3.1 and Output 77.3.2.

Output 77.3.1 Power for Paired $t$ Analysis of Crossover Design
The POWER Procedure
Paired t Test for Mean Difference

| Fixed Scenario Elements |  |
| :--- | ---: |
| Distribution | Normal |
| Method | Exact |
| Alpha | 0.01 |
| Mean 1 | 330 |
| Mean 2 | 310 |
| Standard Deviation 1 | 40 |
| Standard Deviation 2 | 55 |
| Correlation | 0.3 |
| Number of Pairs | 100 |
| Null Difference | 0 |

Computed Power
Index Sides Power
$\begin{array}{lll}1 & 1 & 0.865 \\ 2 & 2 & 0.801\end{array}$

Output 77.3.2 Plot of Power versus Sample Size for Paired $t$ Analysis of Crossover Design


The "Computed Power" table in Output 77.3.1 shows that the power with 100 patients is about 0.8 for the two-sided test and 0.87 for the one-sided test with the alternative of larger Brantium mean. In Output 77.3.2, the line style identifies the number of sides of the test. The plotting symbols identify locations of actual computed powers; the curves are linear interpolations of these points. The plot demonstrates how much higher the power is in the one-sided test than in the two-sided test for the range of sample sizes.

Suppose now that instead of detecting a difference between Xilodol and Brantium, you want to establish that they are similar-in particular, that the absolute mean PEF difference is at most 35 . You might consider this goal if, for example, one of the drugs has fewer side effects and if a difference of no more than 35 is considered clinically small. Instead of a standard $t$ test, you would conduct an equivalence test of the treatment mean difference for the two drugs. You would test the hypothesis that the true difference is less than -35 or more than 35 against the alternative that the mean difference is between -35 and 35 , by using an additive model and a two one-sided tests ("TOST") analysis.

Assuming no period effect, you can use the PAIREDMEANS statement with the TEST=EQUIV_DIFF option to perform a sample size analysis for the paired equivalence test. Indicate power as the result parameter by specifying the POWER= option with a missing value (.). Use the LOWER= and UPPER= options to specify the equivalence bounds of -35 and 35. Use the PAIREDMEANS=, PAIREDSTDDEVS=, CORR=, and ALPHA= options in the same way as in the $t$ test at the beginning of this example to specify the remaining parameters.

The following statements perform the sample size analysis:

```
proc power;
    pairedmeans test=equiv_add
        lower = -35
        upper = 35
        pairedmeans =( 330 310)
        pairedstddevs = (40 55)
        corr =0.3
        alpha =0.01
        npairs = 100
        power = .;
run;
```

The default option DIST=NORMAL specifies an assumption of normally distributed data. The output is shown in Output 77.3.3.

Output 77.3.3 Power for Paired Equivalence Test for Crossover Design
The POWER Procedure
Equivalence Test for Paired Mean Difference

| Fixed Scenario Elements |  |
| :---: | :---: |
| Distribution | Normal |
| Method | Exact |
| Lower Equivalence Bound | -35 |
| Upper Equivalence Bound | 35 |
| Alpha | 0.01 |
| Reference Mean | 330 |
| Treatment Mean | 310 |
| Standard Deviation 1 | 40 |
| Standard Deviation 2 | 55 |
| Correlation | 0.3 |
| Number of Pairs | 100 |
| Computed Power |  |
| Power |  |
| 0.598 |  |

The power for the paired equivalence test with 100 patients is about 0.6 .

## Example 77.4: Noninferiority Test with Lognormal Data

The typical goal in noninferiority testing is to conclude that a new treatment or process or product is not appreciably worse than some standard. This is accomplished by convincingly rejecting a one-sided null hypothesis that the new treatment is appreciably worse than the standard. When designing such studies, investigators must define precisely what constitutes "appreciably worse."
You can use the POWER procedure for sample size analyses for a variety of noninferiority tests, by specifying custom, one-sided null hypotheses for common tests. This example illustrates the strategy (often called

Blackwelder's scheme; Blackwelder 1982) by comparing the means of two independent lognormal samples. The logic applies to one-sample, two-sample, and paired-sample problems involving normally distributed measures and proportions.

Suppose you are designing a study hoping to show that a new (less expensive) manufacturing process does not produce appreciably more pollution than the current process. Quantifying "appreciably worse" as $10 \%$, you seek to show that the mean pollutant level from the new process is less than $110 \%$ of that from the current process. In standard hypothesis testing notation, you seek to reject

$$
H_{0}: \frac{\mu_{\text {new }}}{\mu_{\text {current }}} \geq 1.10
$$

in favor of

$$
H_{A}: \frac{\mu_{\text {new }}}{\mu_{\text {current }}}<1.10
$$

This is described graphically in Figure 77.8. Mean ratios below $100 \%$ are better levels for the new process; a ratio of $100 \%$ indicates absolute equivalence; ratios of $100-110 \%$ are "tolerably" worse; and ratios exceeding $110 \%$ are appreciably worse.

Figure 77.8 Hypotheses for the Pollutant Study


An appropriate test for this situation is the common two-group $t$ test on log-transformed data. The hypotheses become

$$
\begin{aligned}
H_{0}: \log \left(\mu_{\text {new }}\right)-\log \left(\mu_{\text {current }}\right) & \geq \log (1.10) \\
H_{A}: \log \left(\mu_{\text {new }}\right)-\log \left(\mu_{\text {current }}\right) & <\log (1.10)
\end{aligned}
$$

Measurements of the pollutant level will be taken by using laboratory models of the two processes and will be treated as independent lognormal observations with a coefficient of variation $(\sigma / \mu)$ between 0.5 and 0.6 for both processes. You will end up with 300 measurements for the current process and 180 for the new one. It is important to avoid a Type I error here, so you set the Type I error rate to 0.01 . Your theoretical work suggests that the new process will actually reduce the pollutant by about $10 \%$ (to $90 \%$ of current), but you need to compute and graph the power of the study if the new levels are actually between $70 \%$ and $120 \%$ of current levels.

Implement the sample size analysis by using the TWOSAMPLEMEANS statement in PROC POWER with the TEST=RATIO option. Indicate power as the result parameter by specifying the POWER=option with a missing value (.). Specify a series of scenarios for the mean ratio between 0.7 and 1.2 by using the MEANRATIO $=$ option. Use the NULLRATIO= option to specify the null mean ratio of 1.10 . Specify SIDES=L to indicate a one-sided test with the alternative hypothesis stating that the mean ratio is lower than the null value. Specify the significance level, scenarios for the coefficient of variation, and the group sample sizes by using the ALPHA=, CV=, and GROUPNS= options. Generate a plot of power versus mean ratio by specifying the PLOT statement with the $\mathrm{X}=E F F E C T$ option to request a plot with mean ratio on the X axis.
(The result parameter, here power, is always plotted on the other axis.) Use the $\mathrm{STEP}=$ option in the PLOT statement to specify an interval of 0.05 between computed points in the plot.

The following statements perform the desired analysis:

```
ods graphics on;
proc power;
    twosamplemeans test=ratio
            meanratio = 0.7 to 1.2 by 0.1
            nullratio = 1.10
            sides = L
            alpha = 0.01
            cv = 0.5 0.6
            groupns = (300 180)
            power = .;
    plot x=effect step=0.05;
run;
ods graphics off;
```

Note the use of $\mathrm{SIDES}=\mathrm{L}$, which forces computations for cases that need a rejection region that is opposite to the one providing the most one-tailed power; in this case, it is the lower tail. Such cases will show power that is less than the prescribed Type I error rate. The default option DIST=LOGNORMAL specifies the assumption of lognormally distributed data. The default MIN $=$ and MAX= options in the plot statement specify an X axis range identical to the effect size range in the TWOSAMPLEMEANS statement (mean ratios between 0.7 and 1.2).

Output 77.4.1 and Output 77.4.2 show the results.
Output 77.4.1 Power for Noninferiority Test of Ratio
The POWER Procedure Two-Sample t Test for Mean Ratio

| Fixed Scenario Elements |  |
| :--- | ---: |
| Distribution | Lognormal |
| Method | Exact |
| Number of Sides | L |
| Null Geometric Mean Ratio | 1.1 |
| Alpha | 0.01 |
| Group 1 Sample Size | 300 |
| Group 2 Sample Size | 180 |

Output 77.4.1 continued

| Computed Power <br> Geo <br> Mean |  |  |  |
| ---: | ---: | ---: | ---: |
| Index |  |  |  |
| $\mathbf{1}$ | 0.7 | 0.5 | $>.999$ |
| $\mathbf{2}$ | 0.7 | 0.6 | $>.999$ |
| $\mathbf{3}$ | 0.8 | 0.5 | $>.999$ |
| $\mathbf{4}$ | 0.8 | 0.6 | $>.999$ |
| $\mathbf{5}$ | 0.9 | 0.5 | 0.985 |
| $\mathbf{6}$ | 0.9 | 0.6 | 0.933 |
| $\mathbf{7}$ | 1.0 | 0.5 | 0.424 |
| $\mathbf{8}$ | 1.0 | 0.6 | 0.306 |
| $\mathbf{9}$ | 1.1 | 0.5 | 0.010 |
| $\mathbf{1 0}$ | 1.1 | 0.6 | 0.010 |
| $\mathbf{1 1}$ | 1.2 | 0.5 | $<.001$ |
| $\mathbf{1 2}$ | 1.2 | 0.6 | $<.001$ |

Output 77.4.2 Plot of Power versus Mean Ratio for Noninferiority Test


The "Computed Power" table in Output 77.4.1 shows that power exceeds 0.90 if the true mean ratio is $90 \%$ or less, as surmised. But power is unacceptably low (0.31-0.42) if the processes happen to be truly equivalent. Note that the power is identical to the alpha level ( 0.01 ) if the true mean ratio is 1.10 and below 0.01 if the true mean ratio is appreciably worse ( $>110 \%$ ). In Output 77.4.2, the line style identifies the coefficient of variation. The plotting symbols identify locations of actual computed powers; the curves are linear interpolations of these points.

## Example 77.5: Multiple Regression and Correlation

You are working with a team of preventive cardiologists investigating whether elevated serum homocysteine levels are linked to atherosclerosis (plaque buildup) in coronary arteries. The planned analysis is an ordinary least squares regression to assess the relationship between total homocysteine level (tHcy) and a plaque burden index (PBI), adjusting for six other variables: age, gender, plasma levels of folate, vitamin $\mathrm{B}_{6}$, vitamin $\mathrm{B}_{12}$, and a serum cholesterol index. You will regress PBI on tHcy and the six other predictors (plus the intercept) and use a Type III $F$ test to assess whether tHcy is a significant predictor after adjusting for the others. You wonder whether 100 subjects will provide adequate statistical power.

This is a correlational study at a single time. Subjects will be screened so that about half will have had a heart problem. All eight variables will be measured during one visit. Most clinicians are familiar with simple correlations between two variables, so you decide to pose the statistical problem in terms of estimating and testing the partial correlation between $X_{1}=\mathrm{tHcy}$ and $Y=$ PBI, controlling for the six other predictor variables ( $R_{Y X_{1} \mid X_{-1}}$ ). This greatly simplifies matters, especially the elicitation of the conjectured effect.
You use partial regression plots like that shown in Figure 77.9 to teach the team that the partial correlation between PBI and tHcy is the correlation of two sets of residuals obtained from ordinary regression models, one from regressing PBI on the six covariates and the other from regressing tHcy on the same covariates. Thus each subject has "expected" tHcy and PBI values based on the six covariates. The cardiologists believe that subjects whose tHcy is relatively higher than expected will also have a PBI that is relatively higher than expected. The partial correlation quantifies that adjusted association just as a standard simple correlation does with the unadjusted linear association between two variables.

Figure 77.9 Partial Regression Plot


Based on previously published studies of various coronary risk factors and after viewing a set of scatterplots showing various correlations, the team surmises that the true partial correlation is likely to be at least 0.35 .

You want to compute the statistical power for a sample size of $N=100$ by using $\alpha=0.05$. You also want to plot power for sample sizes between 50 and 150 . Use the MULTREG statement to compute the power and the PLOT statement to produce the graph. Since the predictors are observed rather than fixed in advanced, and a joint multivariate normal assumption seems tenable, use MODEL=RANDOM. The following statements perform the power analysis:

```
ods graphics on;
proc power;
    multreg
        model = random
        nfullpredictors = 7
        ntestpredictors = 1
        partialcorr = 0.35
        ntotal = 100
        power = .;
    plot x=n min=50 max=150;
run;
ods graphics off;
```

The POWER=. option identifies power as the parameter to compute. The NFULLPREDICTORS= option specifies seven total predictors (not including the intercept), and the NTESTPREDICTORS= option indicates that one of those predictors is being tested. The PARTIALCORR= and NTOTAL= options specify the partial correlation and sample size, respectively. The default value for the ALPHA= option sets the significance level to 0.05 . The $\mathrm{X}=\mathrm{N}$ option in the plot statement requests a plot of sample size on the X axis, and the MIN $=$ and MAX $=$ options specify the sample size range.

Output 77.5.1 shows the output, and Output 77.5.2 shows the plot.

## Output 77.5.1 Power Analysis for Multiple Regression

The POWER Procedure Type III F Test in Multiple Regression

| Fixed Scenario Elements |  |
| :--- | ---: |
| Method | Exact |
| Model | Random X |
| Number of Predictors in Full Model | 7 |
| Number of Test Predictors | 1 |
| Partial Correlation | 0.35 |
| Total Sample Size | 100 |
| Alpha | 0.05 |


| Computed Power |
| :---: |
| Power |
| 0.939 |

Output 77.5.2 Plot of Power versus Sample Size for Multiple Regression


For the sample size $N=100$, the study is almost balanced with respect to Type I and Type II error rates, with $\alpha=0.05$ and $\beta=1-0.937=0.063$. The study thus seems well designed at this sample size.

Now suppose that in a follow-up meeting with the cardiologists, you discover that their specific intent is to demonstrate that the (partial) correlation between PBI and tHcy is greater than 0.2 . You suggest changing the planned data analysis to a one-sided Fisher's $z$ test with a null correlation of 0.2 . The following statements perform a power analysis for this test:

```
proc power;
    onecorr dist=fisherz
        npvars = 6
        corr = 0.35
        nullcorr = 0.2
        sides = 1
        ntotal = 100
        power = .;
run;
```

The DIST=FISHERZ option in the ONECORR statement specifies Fisher's $z$ test. The NPARTIALVARS= option specifies that six additional variables are adjusted for in the partial correlation. The CORR= option specifies the conjectured correlation of 0.35 , and the NULLCORR $=$ option indicates the null value of 0.2 . The SIDES = option specifies a one-sided test.

Output 77.5.3 shows the output.
Output 77.5.3 Power Analysis for Fisher's $z$ Test
The POWER Procedure Fisher's z Test for Pearson Correlation

| Fixed Scenario Elements |  |
| :--- | ---: |
| Distribution | Fisher's z transformation of $r$ |
| Method | Normal approximation |
| Number of Sides | 1 |
| Null Correlation | 0.2 |
| Number of Variables Partialled Out | 6 |
| Correlation | 0.35 |
| Total Sample Size | 100 |
| Nominal Alpha | 0.05 |


| Computed |
| :--- |
| Power |
| Actual |
| Alpha Power |
| $0.05 \quad 0.466$ |

The power for Fisher's $z$ test is less than $50 \%$, the decrease being mostly due to the smaller effect size (relative to the null value). When asked for a recommendation for a new sample size goal, you compute the required sample size to achieve a power of 0.95 (to balance Type I and Type II errors) and 0.85 (a threshold deemed to be minimally acceptable to the team). The following statements perform the sample size determination:

```
proc power;
    onecorr dist=fisherz
        npvars = 6
        corr = 0.35
        nullcorr = 0.2
        sides = 1
        ntotal = .
        power = 0.85 0.95;
run;
```

The NTOTAL=. option identifies sample size as the parameter to compute, and the POWER= option specifies the target powers.

Output 77.5.4 Sample Size Determination for Fisher's $z$ Test
The POWER Procedure
Fisher's z Test for Pearson Correlation

| Fixed Scenario Elements |  |
| :--- | ---: |
| Distribution | Fisher's z transformation of $r$ |
| Method | Normal approximation |
| Number of Sides | 1 |
| Null Correlation | 0.2 |
| Number of Variables Partialled Out | 6 |
| Correlation | 0.35 |
| Nominal Alpha | 0.05 |


| Computed N Total |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Index | Nominal Power | Actual Alpha | Actual Power | Total |
| 1 | 0.85 | 0.05 | 0.850 | 280 |
| 2 | 0.95 | 0.05 | 0.950 | 417 |

The results in Output 77.5.4 reveal a required sample size of 417 to achieve a power of 0.95 and a required sample size of 280 to achieve a power of 0.85 .

## Example 77.6: Comparing Two Survival Curves

You are consulting for a clinical research group planning a trial to compare survival rates for proposed and standard cancer treatments. The planned data analysis is a log-rank test to nonparametrically compare the overall survival curves for the two treatments. Your goal is to determine an appropriate sample size to achieve a power of 0.8 for a two-sided test with $\alpha=0.05$ by using a balanced design.

The survival curve for patients on the standard treatment is well known to be approximately exponential with a median survival time of five years. The research group conjectures that the new proposed treatment will yield a (nonexponential) survival curve similar to the dashed line in Figure 77.6.1.

Patients will be accrued uniformly over two years and then followed for an additional three years past the accrual period. Some loss to follow-up is expected, with roughly exponential rates that would result in about $50 \%$ loss with the standard treatment within 10 years. The loss to follow-up with the proposed treatment is more difficult to predict, but $50 \%$ loss would be expected to occur sometime between years 5 and 20 .

Output 77.6.1 Survival Curves


Use the TWOSAMPLESURVIVAL statement with the TEST=LOGRANK option to compute the required sample size for the log-rank test. The following statements perform the analysis:

```
proc power;
    twosamplesurvival test=logrank
        curve("Standard") = 5 : 0.5
        curve("Proposed") = (1 to 5 by 1):(0.95 0.9 0.75 0.7 0.6)
        groupsurvival = "Standard" | "Proposed"
        accrualtime = 2
        followuptime = 3
        groupmedlosstimes = 10 | 20 5
        power = 0.8
        npergroup = .;
run;
```

The CURVE= option defines the two survival curves. The "Standard" curve has only one point, specifying an exponential form with a survival probability of 0.5 at year 5 . The "Proposed" curve is a piecewise linear curve defined by the five points shown in Figure 77.6.1. The GROUPSURVIVAL= option assigns the survival curves to the two groups, and the ACCRUALTIME= and FOLLOWUPTIME= options specify the accrual and follow-up times. The GROUPMEDLOSSTIMES= option specifies the years at which $50 \%$ loss is expected to occur. The POWER= option specifies the target power, and the NPERGROUP=. option identifies sample size per group as the parameter to compute. Default values for the SIDES $=$ and ALPHA $=$ options specify a two-sided test with $\alpha=0.05$.
Output 77.6.2 shows the results.
Output 77.6.2 Sample Size Determination for Log-Rank Test
The POWER Procedure Log-Rank Test for Two Survival Curves

| Fixed Scenario Elements |  |
| :--- | ---: |
| Method | Lakatos normal approximation |
| Accrual Time | 2 |
| Follow-up Time | 3 |
| Group 1 Survival Curve | Standard |
| Form of Survival Curve 1 | Exponential |
| Group 2 Survival Curve | Proposed |
| Form of Survival Curve 2 | Piecewise Linear |
| Group 1 Median Loss Time | 10 |
| Nominal Power | 0.8 |
| Number of Sides | 2 |
| Number of Time Sub-Intervals | 12 |
| Alpha | 0.05 |


| Computed <br> Median <br> Mer Group |  |  |  |
| ---: | ---: | ---: | ---: |
| Loss | Actual | $\mathbf{N}$ per |  |
| Index | Time $\mathbf{2}$ | Power | Group |
| $\mathbf{1}$ | 20 | 0.800 | 228 |
| $\mathbf{2}$ | 5 | 0.801 | 234 |

The required sample size per group to achieve a power of 0.8 is 228 if the median loss time is 20 years for the proposed treatment. Only six more patients are required in each group if the median loss time is as short as five years.

## Example 77.7: Confidence Interval Precision

An investment firm has hired you to help plan a study to estimate the success of a new investment strategy called IntuiVest. The study involves complex simulations of market conditions over time, and it tracks the balance of a hypothetical brokerage account starting with $\$ 50,000$. Each simulation is very expensive in terms of computing time. You are asked to determine an appropriate number of simulations to estimate the average change in the account balance at the end of three years. The goal is to have a $95 \%$ chance of obtaining a $90 \%$ confidence interval whose half-width is at most $\$ 1,000$. That is, the firm wants to have a $95 \%$ chance of being able to correctly claim at the end of the study that "Our research shows with $90 \%$ confidence that IntuiVest yields a profit of $\$ \mathrm{X}+/-\$ 1,000$ at the end of three years on an initial investment of $\$ 50,000$ (under simulated market conditions)."

The probability of achieving the desired precision (that is, a small interval width) can be calculated either unconditionally or conditionally given that the true mean is captured by the interval. You decide to use the conditional form, considering two of its advantages:

- The conditional probability is usually lower than the unconditional probability for the same sample size, meaning that the conditional form is generally conservative.
- The overall probability of achieving the desired precision and capturing the true mean is easily computed as the product of the half-width probability and the confidence level. In this case, the overall probability is $0.95 \times 0.9=0.855$.

Based on some initial simulations, you expect a standard deviation between $\$ 25,000$ and $\$ 45,000$ for the ending account balance. You will consider both of these values in the sample size analysis.

As mentioned in the section "Overview of Power Concepts" on page 6363, an analysis of confidence interval precision is analogous to a traditional power analysis, with "CI Half-Width" taking the place of effect size and "Prob(Width)" taking the place of power. In this example, the target CI Half-Width is 1000, and the desired $\operatorname{Prob}$ (Width) is 0.95 .

In addition to computing sample sizes for a half-width of $\$ 1,000$, you are asked to plot the required number of simulations for a range of half-widths between $\$ 500$ and $\$ 2,000$. Use the ONESAMPLEMEANS statement with the $\mathrm{CI}=\mathrm{T}$ option to implement the sample size determination. The following statements perform the analysis:

```
ods graphics on;
proc power;
    onesamplemeans ci=t
        alpha = 0.1
        halfwidth = 1000
        stddev = 25000 45000
        probwidth = 0.95
        ntotal = .;
    plot x=effect min=500 max=2000;
run;
ods graphics off;
```

The NTOTAL=. option identifies sample size as the parameter to compute. The ALPHA=0.1 option specifies a confidence level of $1-\alpha=0.9$. The HALFWIDTH= option specifies the target half-width, and the STDDEV= option specifies the conjectured standard deviation values. The PROBWIDTH=option specifies the desired probability of achieving the target precision. The default value PROBTYPE=CONDITIONAL specifies that this probability is conditional on the true mean being captured by the interval. The default of SIDES=2 indicates a two-sided interval.

Output 77.7.1 shows the output, and Output 77.7.2 shows the plot.
Output 77.7.1 Sample Size Determination for Confidence Interval Precision

## The POWER Procedure Confidence Interval for Mean

| Fixed Scenario Elements |  |  |  |  |
| :--- | ---: | ---: | :---: | :---: |
| Distribution | Normal |  |  |  |
| Method | Exact |  |  |  |
| Alpha | 0.1 |  |  |  |
| Cl Half-Width | 1000 |  |  |  |
| Nominal Prob(Width) | 0.95 |  |  |  |
| Number of Sides | 2 |  |  |  |
| Prob Type | Conditional |  |  |  |
| Computed |  |  |  |  |
| N Total |  |  |  |  |

Output 77.7.2 Plot of Sample Size versus Confidence Interval Half-Width


The number of simulations required in order to have a $95 \%$ chance of obtaining a half-width of at most 1000 is between 1788 and 5652, depending on the standard deviation. The plot reveals that more than 20,000 simulations would be required for a half-width of 500 , assuming the higher standard deviation.

## Example 77.8: Customizing Plots

This example demonstrates various ways you can modify and enhance plots:

- assigning analysis parameters to axes
- fine-tuning a sample size axis
- adding reference lines
- linking plot features to analysis parameters
- choosing key (legend) styles
- modifying symbol locations

The example plots are all based on a sample size analysis for a two-sample $t$ test of group mean difference. You start by computing the sample size required to achieve a power of 0.9 by using a two-sided test with $\alpha=$ 0.05 , assuming the first mean is 12 , the second mean is either 15 or 18 , and the standard deviation is either 7 or 9 .

Use the TWOSAMPLEMEANS statement with the TEST=DIFF option to compute the required sample sizes. Indicate total sample size as the result parameter by supplying a missing value (.) with the NTOTAL= option. Use the GROUPMEANS=, STDDEV=, and POWER= options to specify values of the other parameters. The following statements perform the sample size computations:

```
proc power;
    twosamplemeans test=diff
            groupmeans = 12 | 15 18
            stddev = 7 9
            power = 0.9
            ntotal = .;
run;
```

Default values for the NULLDIFF=, SIDES=, GROUPWEIGHTS=, and DIST= options specify a null mean difference of 0 , two-sided test, balanced design, and assumption of normally distributed data, respectively.

Output 77.8.1 shows that the required sample size ranges from 60 to 382 , depending on the unknown standard deviation and second mean.

Output 77.8.1 Computed Sample Sizes
The POWER Procedure Two-Sample t Test for Mean Difference

| Fixed Scenario Elements |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Distribution |  |  | Norm |  |
| Method |  |  |  | xact |
| Group 1 Mean |  |  |  | 12 |
| Nominal Power |  |  |  | 0.9 |
| Number of Side |  |  | des | 2 |
| Null Difference |  |  |  | 0 |
| Alpha |  |  |  | . 05 |
| Group 1 Weight |  |  |  | 1 |
| Group 2 Weight |  |  |  | 1 |
| Computed N Total |  |  |  |  |
| Index | $x$ Mean2 | Std Dev | Actual <br> Power | $\underset{\text { Total }}{\mathrm{N}}$ |
| 1 | 15 | 7 | 0.902 | 232 |
| 2 | 215 | 9 | 0.901 | 382 |
| 3 | 318 | 7 | 0.904 | 60 |
| 4 | 418 | 9 | 9.904 | 98 |

## Assigning Analysis Parameters to Axes

Use the PLOT statement to produce plots for all power and sample size analyses in PROC POWER. For the sample size analysis described at the beginning of this example, suppose you want to plot the required sample size on the Y axis against a range of powers between 0.5 and 0.95 on the X axis. The $\mathrm{X}=$ and $\mathrm{Y}=$ options specify which parameter to plot against the result and which axis to assign to this parameter. You can use either the $\mathrm{X}=$ or the $\mathrm{Y}=$ option, but not both. Use the $\mathrm{X}=$ POWER option in the PLOT statement to request a plot with power on the X axis. The result parameter, here total sample size, is always plotted on the other axis. Use the $\mathrm{MIN}=$ and MAX= options to specify the range of the axis indicated with either the $\mathrm{X}=$ or the $\mathrm{Y}=$ option. Here, specify $\mathrm{MIN}=0.5$ and $\mathrm{MAX}=0.95$ to specify the power range. The following statements produce the plot:

```
ods graphics on;
proc power plotonly;
    twosamplemeans test=diff
        groupmeans = 12 | 15 18
        stddev = 79
        power = 0.9
        ntotal = .;
    plot x=power min=0.5 max=0.95;
run;
```

Note that the value (0.9) of the POWER= option in the TWOSAMPLEMEANS statement is only a placeholder when the PLOTONLY option is used and both the MIN= and MAX= options are used, because the values of the MIN= and MAX= options override the value of 0.9. But the POWER= option itself is still required in the TWOSAMPLEMEANS statement, to provide a complete specification of the sample size analysis.

The resulting plot is shown in Output 77.8.2.

Output 77.8.2 Plot of Sample Size versus Power


The line style identifies the group means scenario, and the plotting symbol identifies the standard deviation scenario. The locations of plotting symbols indicate computed sample sizes; the curves are linear interpolations of these points. By default, each curve consists of approximately 20 computed points (sometimes slightly more or less, depending on the analysis).

If you would rather plot power on the Y axis versus sample size on the X axis, you have two general strategies to choose from. One strategy is to use the $\mathrm{Y}=$ option instead of the $\mathrm{X}=$ option in the PLOT statement:

```
plot y=power min=0.5 max=0.95;
```

Output 77.8.3 Plot of Power versus Sample Size using First Strategy


Note that the resulting plot (Output 77.8.3) is essentially a mirror image of Output 77.8.2. The axis ranges are set such that each curve in Output 77.8.3 contains similar values of Y instead of X. Each plotted point represents the computed value of the X axis at the input value of the Y axis.

A second strategy for plotting power versus sample size (when originally solving for sample size) is to invert the analysis and base the plot on computed power for a given range of sample sizes. This strategy works well for monotonic power curves (as is the case for the $t$ test and most other continuous analyses). It is advantageous in the sense of preserving the traditional role of the Y axis as the computed parameter. A common way to implement this strategy is as follows:

- Determine the range of sample sizes sufficient to cover at the desired power range for all curves (where each "curve" represents a scenario for standard deviation and second group mean).
- Use this range for the X axis of a plot.

To determine the required sample sizes for target powers of 0.5 and 0.95 , change the values in the POWER= option as follows to reflect this range:

```
proc power;
    twosamplemeans test=diff
    groupmeans = 12 | 15 18
    stddev = 7 9
    power = 0.5 0.95
    ntotal = .;
run;
```

Output 77.8.4 reveals that a sample size range of 24 to 470 is approximately sufficient to cover the desired power range of 0.5 to 0.95 for all curves ("approximately" because the actual power at the rounded sample size of 24 is slightly higher than the nominal power of 0.5 ).

Output 77.8.4 Computed Sample Sizes
The POWER Procedure
Two-Sample t Test for Mean Difference

| Fixed Scenario |  |
| :--- | ---: |
| Elements |  |
| Distribution | Normal |
| Method | Exact |
| Group 1 Mean | 12 |
| Number of Sides | 2 |
| Null Difference | 0 |
| Alpha | 0.05 |
| Group 1 Weight | 1 |
| Group 2 Weight | 1 |


|  |  | ompu | ted N Tota |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Mean2 | Std Dev | Nominal Power | Actual Power | N Total |
| 1 | 15 | 7 | 0.50 | 0.502 | 86 |
| 2 | 15 | 7 | 0.95 | 0.951 | 286 |
| 3 | 15 | 9 | 0.50 | 0.505 | 142 |
| 4 | 15 | 9 | 0.95 | 0.950 | 470 |
| 5 | 18 | 7 | 0.50 | 0.519 | 24 |
| 6 | 18 | 7 | 0.95 | 0.953 | 74 |
| 7 | 18 | 9 | 0.50 | 0.516 | 38 |
| 8 | 18 | 9 | 0.95 | 0.952 | 120 |

To plot power on the Y axis for sample sizes between 20 and 500 , use the $\mathrm{X}=\mathrm{N}$ option in the PLOT statement with $\mathrm{MIN}=20$ and $\mathrm{MAX}=500$ :

```
proc power plotonly;
    twosamplemeans test=diff
        groupmeans = 12 | 15 18
        stddev = 7 9
        power = .
        ntotal = 200;
    plot x=n min=20 max=500;
run;
```

Each curve in the resulting plot in Output 77.8.5 covers at least a power range of 0.5 to 0.95 .

## Output 77.8.5 Plot of Power versus Sample Size Using Second Strategy



Finally, suppose you want to produce a plot of sample size versus effect size for a power of 0.9. In this case, the "effect size" is defined to be the mean difference. You need to reparameterize the analysis by using the MEANDIFF= option instead of the GROUPMEANS= option to produce a plot, since each plot axis must be represented by a scalar parameter. Use the $\mathrm{X}=$ EFFECT option in the PLOT statement to assign the mean difference to the X axis. The following statements produce a plot of required sample size to detect mean differences between 3 and 6:

```
proc power plotonly;
    twosamplemeans test=diff
        meandiff = 3 6
        stddev = 7 9
        power = 0.9
        ntotal = .;
    plot x=effect min=3 max=6;
run;
```

The resulting plot Output 77.8 .6 shows how the required sample size decreases with increasing mean difference.

Output 77.8.6 Plot of Sample Size versus Mean Difference


## Fine-Tuning a Sample Size Axis

Consider the following plot request for a sample size analysis similar to the one in Output 77.8.1 but with only a single scenario, and with unbalanced sample size allocation of 2:1:

```
proc power plotonly;
    ods output plotcontent=PlotData;
    twosamplemeans test=diff
        groupmeans = 12 | 18
        stddev = 7
        groupweights = 2 | 1
        power = .
        ntotal = 20;
    plot x=n min=20 max=50 npoints=20;
run;
```

The MIN=, MAX=, and NPOINTS= options in the PLOT statement request a plot with 20 points between 20 and 50 . But the resulting plot (Output 77.8.7) appears to have only 11 points, and they range from 18 to 48 .

Output 77.8.7 Plot with Overlapping Points


The reason that this plot has fewer points than usual is due to the rounding of sample sizes. If you do not use the NFRACTIONAL option in the analysis statement (here, the TWOSAMPLEMEANS statement), then the set of sample size points determined by the $\mathrm{MIN}=$, $\mathrm{MAX}=$, NPOINTS $=$, and STEP= options in the PLOT statement can be rounded to satisfy the allocation weights. In this case, they are rounded down to the nearest multiples of 3 (the sum of the weights), and many of the points overlap. To see the overlap, you can print the NominaINTotal (unadjusted) and NTotal (rounded) variables in the PlotContent ODS object (here saved to a data set called PlotData):

```
proc print data=PlotData;
    var NominalNTotal NTotal;
run;
```

The output is shown in Output 77.8.8.

Output 77.8.8 Sample Sizes

| Obs | NominalNTotal | NTotal |
| ---: | ---: | ---: |
| $\mathbf{1}$ | 18.0 | 18 |
| $\mathbf{2}$ | 19.6 | 18 |
| $\mathbf{3}$ | 21.2 | 21 |
| $\mathbf{4}$ | 22.7 | 21 |
| $\mathbf{5}$ | 24.3 | 24 |
| $\mathbf{6}$ | 25.9 | 24 |
| $\mathbf{7}$ | 27.5 | 27 |
| $\mathbf{8}$ | 29.1 | 27 |
| $\mathbf{9}$ | 30.6 | 30 |
| $\mathbf{1 0}$ | 32.2 | 30 |
| $\mathbf{1 1}$ | 33.8 | 33 |
| $\mathbf{1 2}$ | 35.4 | 33 |
| $\mathbf{1 3}$ | 36.9 | 36 |
| $\mathbf{1 4}$ | 38.5 | 36 |
| $\mathbf{1 4}$ | 40.1 | 39 |
| $\mathbf{1 6}$ | 41.7 | 39 |
| $\mathbf{1 7}$ | 43.3 | 42 |
| $\mathbf{1 8}$ | 44.8 | 42 |
| $\mathbf{1 9}$ | 46.4 | 45 |
| $\mathbf{2 0}$ | 48.0 | 48 |

Besides overlapping of sample size points, another peculiarity that might occur without the NFRACTIONAL option is unequal spacing-for example, in the plot in Output 77.8.9, created with the following statements:

```
proc power plotonly;
    twosamplemeans test=diff
        groupmeans = 12 | 18
        stddev = 7
        groupweights = 2 | 1
        power = .
        ntotal = 20;
    plot x=n min=20 max=50 npoints=5;
run;
```

Output 77.8.9 Plot with Unequally Spaced Points


If you want to guarantee evenly spaced, nonoverlapping sample size points in your plots, you can either (1) use the NFRACTIONAL option in the analysis statement preceding the PLOT statement or (2) use the STEP= option and provide values for the $\mathrm{MIN}=, \mathrm{MAX}=$, and $\mathrm{STEP}=$ options in the PLOT statement that are multiples of the sum of the allocation weights. Note that this sum is simply 1 for one-sample and paired designs and 2 for balanced two-sample designs. So any integer step value works well for one-sample and paired designs, and any even step value works well for balanced two-sample designs. Both of these strategies will avoid rounding adjustments.

The following statements implement the first strategy to create the plot in Output 77.8.10, by using the NFRACTIONAL option in the TWOSAMPLEMEANS statement:

```
proc power plotonly;
    twosamplemeans test=diff
        nfractional
        groupmeans = 12 | 18
        stddev = }
        groupweights = 2 | 1
        power = .
        ntotal = 20;
    plot x=n min=20 max=50 npoints=20;
run;
```

Output 77.8.10 Plot with Fractional Sample Sizes


To implement the second strategy, use multiples of 3 for the $\operatorname{STEP}=$, MIN=, and MAX= options in the PLOT statement (because the sum of the allocation weights is $2+1=3$ ). The following statements use STEP=3, MIN=18, and MAX=48 to create a plot that looks identical to the plot in Output 77.8.7 but suffers no overlapping of points:

```
proc power plotonly;
    twosamplemeans test=diff
        groupmeans = 12 | 18
        stddev = 7
        groupweights = 2 | 1
        power = .
        ntotal = 20;
    plot x=n min=18 max=48 step=3;
run;
```


## Adding Reference Lines

Suppose you want to add reference lines to highlight power $=0.8$ and power $=0.9$ on the plot in Output 77.8.5. You can add simple reference lines by using the YOPTS $=$ option and $\mathrm{REF}=$ suboption in the PLOT statement to produce Output 77.8.11, with the following statements:

```
proc power plotonly;
            twosamplemeans test=diff
            groupmeans = 12 | 15 18
            stddev = 7 9
            power = .
            ntotal = 100;
        plot x=n min=20 max=500
            yopts=(ref=0.8 0.9);
run;
```

Output 77.8.11 Plot with Simple Reference Lines on Y Axis


Or you can specify CROSSREF=YES to add reference lines that intersect each curve and cross over to the other axis:

```
plot x=n min=20 max=500
    yopts=(ref=0.8 0.9 crossref=yes);
```

The resulting plot is shown in Output 77.8.12.

Output 77.8.12 Plot with CROSSREF=YES Style Reference Lines from Y Axis


You can also add reference lines for the X axis by using the XOPTS $=$ option instead of the YOPTS $=$ option. For example, the following PLOT statement produces Output 77.8.13, which has crossing reference lines highlighting the sample size of 100 :

```
plot x=n min=20 max=500
    xopts=(ref=100 crossref=yes);
```

Note that the values that label the reference lines at the X axis in Output 77.8.12 and at the Y axis in Output 77.8.13 are linearly interpolated from two neighboring points on the curves. Thus they might not exactly match corresponding values that are computed directly from the methods in the section "Computational Methods and Formulas" on page 6373-that is, computed by PROC POWER in the absence of a PLOT statement. The two ways of computing these values generally differ by a negligible amount.

Output 77.8.13 Plot with CROSSREF=YES Style Reference Lines from X Axis


## Linking Plot Features to Analysis Parameters

You can use the VARY option in the PLOT statement to specify which of the following features you want to associate with analysis parameters.

- line style
- plotting symbol
- color
- panel

You can specify mappings between each of these features and one or more analysis parameters, or you can simply choose a subset of these features to use (and rely on default settings to associate these features with multiple-valued analysis parameters).

Suppose you supplement the sample size analysis in Output 77.8.5 to include three values of alpha, by using the following statements:

```
proc power plotonly;
    twosamplemeans test=diff
        groupmeans = 12 | 15 18
        stddev = 7 9
        alpha = 0.01 0.025 0.1
        power = .
        ntotal = 100;
    plot x=n min=20 max=500;
run;
```

The defaults for the VARY option in the PLOT statement specify line style varying by the ALPHA= parameter, plotting symbol varying by the GROUPMEANS = parameter, panel varying by the STDDEV= parameter, and color remaining constant. The resulting plot, consisting of two panels, is shown in Output 77.8.14.

## Output 77.8.14 Plot with Default VARY Settings



Output 77.8.14 continued


Suppose you want to produce a plot with only one panel that varies color in addition to line style and plotting symbol. Include the LINESTYLE, SYMBOL, and COLOR keywords in the VARY option in the PLOT statement, as follows, to produce the plot in Output 77.8.15:

```
plot x=n min=20 max=500
    vary (linestyle, symbol, color);
```

Output 77.8.15 Plot with Varying Color Instead of Panel


Finally, suppose you want to specify which features are used and which analysis parameters they are linked to. The following PLOT statement produces a two-panel plot (shown in Output 77.8.16) in which line style varies by standard deviation, plotting symbol varies by both alpha and sides, and panel varies by means:

```
plot x=n min=20 max=500
    vary (linestyle by stddev,
        symbol by alpha sides,
        panel by groupmeans);
```

Output 77.8.16 Plot with Features Explicitly Linked to Parameters


Output 77.8.16 continued


## Choosing Key (Legend) Styles

The default style for the key (or "legend") is one that displays the association between levels of features and levels of analysis parameters, located below the X axis. For example, Output 77.8.5 demonstrates this style of key.

You can reproduce Output 77.8 .5 with the same key but a different location, inside the plotting region, by using the POS=INSET option within the KEY=BYFEATURE option in the PLOT statement. The following statements product the plot in Output 77.8.17:

```
proc power plotonly;
    twosamplemeans test=diff
        groupmeans = 12 | 15 18
        stddev = 7 9
        power = .
        ntotal = 200;
    plot x=n min=20 max=500
        key = byfeature(pos=inset);
run;
```

Output 77.8.17 Plot with a By-Feature Key inside the Plotting Region


Alternatively, you can specify a key that identifies each individual curve separately by number by using the KEY=BYCURVE option in the PLOT statement:

```
plot x=n min=20 max=500
    key = bycurve;
```

The resulting plot is shown in Output 77.8.18.

Output 77.8.18 Plot with a Numbered By-Curve Key


Use the NUMBERS=OFF option within the KEY=BYCURVE option to specify a nonnumbered key that identifies curves with samples of line styles, symbols, and colors:

```
plot x=n min=20 max=500
    key = bycurve(numbers=off pos=inset);
```

The POS=INSET suboption places the key within the plotting region. The resulting plot is shown in Output 77.8.19.

Output 77.8.19 Plot with a Nonnumbered By-Curve Key


Finally, you can attach labels directly to curves with the KEY=ONCURVES option. The following PLOT statement produces Output 77.8.20:

```
plot x=n min=20 max=500
    key = oncurves;
```

Output 77.8.20 Plot with Directly Labeled Curves


## Modifying Symbol Locations

The default locations for plotting symbols are the points computed directly from the power and sample size algorithms. For example, Output 77.8 .5 shows plotting symbols corresponding to computed points. The curves connecting these points are interpolated (as indicated by the INTERPOL=option in the PLOT statement).

You can modify the locations of plotting symbols by using the MARKERS= option in the PLOT statement. The MARKERS=ANALYSIS option places plotting symbols at locations corresponding to the input specified in the analysis statement preceding the PLOT statement. You might prefer this as an alternative to using reference lines to highlight specific points. For example, you can reproduce Output 77.8.5, but with the plotting symbols located at the sample sizes shown in Output 77.8.1, by using the following statements:

```
proc power plotonly;
    twosamplemeans test=diff
        groupmeans = 12 | 15 18
        stddev = 79
        power = .
        ntotal = 232 382 60 98;
```

```
    plot x=n min=20 max=500
    markers=analysis;
run;
```

The analysis statement here is the TWOSAMPLEMEANS statement. The MARKERS=ANALYSIS option in the PLOT statement causes the plotting symbols to occur at sample sizes specified by the NTOTAL= option in the TWOSAMPLEMEANS statement: $232,382,60$, and 98 . The resulting plot is shown in Output 77.8.21.

Output 77.8.21 Plot with MARKERS=ANALYSIS


You can also use the MARKERS=NICE option to align symbols with the tick marks on one of the axes (the X axis when the $\mathrm{X}=$ option is used, or the Y axis when the $\mathrm{Y}=$ option is used):

```
plot x=n min=20 max=500
    markers=nice;
```

The plot created by this PLOT statement is shown in Output 77.8.22.

Output 77.8.22 Plot with MARKERS=NICE


Note that the plotting symbols are aligned with the tick marks on the X axis because the $\mathrm{X}=\mathrm{option}$ is specified.

## Example 77.9: Binary Logistic Regression with Independent Predictors

Suppose you are planning an industrial experiment similar to the analysis in "Getting Started: LOGISTIC Procedure" on page 4492 in Chapter 60, "The LOGISTIC Procedure," but for a different type of ingot. The primary test of interest is the likelihood ratio chi-square test of the effect of heating time on the readiness of the ingots for rolling. Ingots will be randomized independently into one of four different heating times (5, 10, 15 , and 20 minutes) with allocation ratios 2:3:3:2 and three different soaking times ( 2,4 , and 6 minutes) with allocation ratios $2: 2: 1$. The mass of each ingot will be measured as a covariate.

You want to know how many ingots you must sample to have a $90 \%$ chance of detecting an odds ratio as small as 1.2 for a five-minute heating time increase. The odds ratio is defined here as the odds of the ingot not being ready given a heating time of $h$ minutes divided by the odds given a heating time of $h-5$ minutes, for any time $h$. You will use a significance level of $\alpha=0.1$ to balance Type I and Type II errors since you consider their importance to be roughly equal.

The distributions of heating time and soaking time are determined by the design, but you must conjecture the distribution of ingot mass. Suppose you expect its distribution to be approximately normal with mean 4 kg and standard deviation between 1 kg and 2 kg .

You are powering the study for an odds ratio of 1.2 for the heating time, but you must also conjecture odds ratios for soaking time and mass. You suspect that the odds ratio for a unit increase in soaking time is about 1.4, and the odds ratio for a unit increase in mass is between 1 and 1.3.

Finally, you must provide a guess for the average probability of an ingot not being ready for rolling, averaged across all possible design profiles. Existing data suggest that this probability lies between 0.15 and 0.25 .

You decide to evaluate sample size at the two extremes of each parameter for which you conjectured a range. Use the following statements to perform the sample size determination:

```
proc power;
    logistic
        vardist("Heat") = ordinal((5 10 15 20) : (0.2 0.3 0.3 0.2))
        vardist("Soak") = ordinal((2 4 6) : (0.4 0.4 0.2))
        vardist("Mass1") = normal(4, 1)
        vardist("Mass2") = normal(4, 2)
        testpredictor = "Heat"
        covariates = "Soak" | "Mass1" "Mass2"
        responseprob = 0.15 0.25
        testoddsratio = 1.2
        units= ("Heat" = 5)
        covoddsratios = 1.4 | 1 1.3
        alpha = 0.1
        power = 0.9
        ntotal = .;
run;
```

The VARDIST= option is used to define the distributions of the predictor variables. The distributions of heating and soaking times are defined by the experimental design, with ordinal probabilities derived from the allocation ratios. The two conjectured standard deviations for the ingot mass are represented in the Mass1 and Mass2 distributions. The TESTPREDICTOR= option identifies the predictor being tested, and the COVARIATES = option specifies the scenarios for the remaining predictors in the model (soaking time and mass). The RESPONSEPROB= option specifies the overall response probability, and the TESTODDSRATIO $=$ and UNITS= options indicate the odds ratio and increment for heating time. The COVODDSRATIOS $=$ option specifies the scenarios for the odds ratios of soaking time and mass. The default DEFAULTUNIT $=1$ option specifies a unit change for both of these odds ratios. The ALPHA= option sets the significance level, and the POWER= option defines the target power. Finally, the NTOTAL= option with a missing value (.) identifies the parameter to solve for.

Output 77.9.1 shows the results.

Output 77.9.1 Sample Sizes for Test of Heating Time in Logistic Regression
The POWER Procedure
Likelihood Ratio Chi-Square Test for One Predictor

| Fixed Scenario Elements |  |
| :--- | ---: |
| Method | Shieh-O'Brien approximation |
| Alpha | 0.1 |
| Test Predictor | Heat |
| Odds Ratio for Test Predictor | 1.2 |
| Unit for Test Pred Odds Ratio | 5 |
| Nominal Power | 0.9 |


| Computed N Total |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Index | Response Prob | Cova | ariates |  |  | $\begin{aligned} & \mathrm{Co} \\ & \mathrm{Un} \end{aligned}$ |  | Total <br> N <br> Bins | Actual Power | $\underset{\text { Total }}{\mathrm{N}}$ |
| 1 | 0.15 | Soak | Mass1 | 1.4 | 1.0 | 1 | 1 | 120 | 0.900 | 1878 |
| 2 | 0.15 | Soak | Mass1 | 1.4 | 1.3 | 1 | 1 | 120 | 0.900 | 1872 |
| 3 | 0.15 | Soak | Mass2 | 1.4 | 1.0 | 1 | 1 | 120 | 0.900 | 1878 |
| 4 | 0.15 | Soak | Mass2 | 1.4 | 1.3 | 1 | 1 | 120 | 0.900 | 1857 |
| 5 | 0.25 | Soak | Mass1 | 1.4 | 1.0 | 1 | 1 | 120 | 0.900 | 1342 |
| 6 | 0.25 | Soak | Mass1 | 1.4 | 1.3 | 1 | 1 | 120 | 0.900 | 1348 |
| 7 | 0.25 | Soak | Mass2 | 1.4 | 1.0 | 1 | 1 | 120 | 0.900 | 1342 |
| 8 | 0.25 | Soak | Mass2 | 1.4 | 1.3 | 1 | 1 | 120 | 0.900 | 1369 |

The required sample size ranges from 1342 to 1878 , depending on the unknown true values of the overall response probability, mass standard deviation, and soaking time odds ratio. The overall response probability clearly has the largest influence among these parameters, with a sample size increase of almost $40 \%$ going from 0.25 to 0.15 .

## Example 77.10: Wilcoxon-Mann-Whitney Test

Consider a hypothetical clinical trial to treat interstitial cystitis (IC), a painful, chronic inflammatory condition of the bladder with no known cause that most commonly affects women. Two treatments will be compared: lidocaine alone ("lidocaine") versus lidocaine plus a fictitious experimental drug called Mironel ("Mir+lido"). The design is balanced, randomized, double-blind, and female-only. The primary outcome is a measure of overall improvement at week 4 of the study, measured on a seven-point Likert scale as shown in Table 77.34.

Table 77.34 Self-Report Improvement Scale

| Compared to when I started <br> this study, my condition is: |  |
| :--- | ---: |
| Much worse | -3 |
| Worse | -2 |
| Slightly worse | -1 |
| The same | 0 |
| Slightly better | +1 |
| Better | +2 |
| Much better | +3 |

The planned data analysis is a one-sided Wilcoxon-Mann-Whitney test with $\alpha=0.05$ where the alternative hypothesis represents greater improvement for "Mir+lido."

You are asked to graphically assess the power of the planned trial for sample sizes between 100 and 250, assuming that the conditional outcome probabilities given treatment are equal to the values in Table 77.35 .

Table 77.35 Conjectured Conditional Probabilities

|  | Response |  |  |  |  |  |  |  |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Treatment | -3 | -2 | -1 | 0 | +1 | +2 | +3 |  |
| Lidocaine | 0.01 | 0.04 | 0.20 | 0.50 | 0.20 | 0.04 | 0.01 |  |
| Mir+lido | 0.01 | 0.03 | 0.15 | 0.35 | 0.30 | 0.10 | 0.06 |  |

Use the following statements to compute the power at sample sizes of 100 and 250 and generate a power curve:

```
ods graphics on;
proc power;
    twosamplewilcoxon
        vardist("lidocaine") = ordinal ((-3 -2 -1 0 1 2 3) :
                        (.01 . 04 . 20 . 50 . 20 . 04 .01))
        vardist("Mir+lido") = ordinal ((-3 -2 -1 0 1 2 3) :
                            (.01 . 03 . 15 . 35 . 30 . 10 .06))
        variables = "lidocaine" | "Mir+lido"
        sides = u
        ntotal = 100 250
        power = .;
    plot step=10;
run;
ods graphics off;
```

The VARDIST= option is used to define the distribution for each treatment, and the VARIABLES= option specifies the distributions to compare. The SIDES $=\mathrm{U}$ option corresponds to the alternative hypothesis that the second distribution ("Mir+lido") is more favorable. The NTOTAL= option specifies the total sample sizes of interest, and the POWER= option with a missing value (.) identifies the parameter to solve for. The default GROUPWEIGHTS $=$ and ALPHA $=$ options specify a balanced design and significance level $\alpha=0.05$.

The $\mathrm{STEP}=10$ option in the PLOT statement requests a point for each sample size increment of 10 . The default values for the $\mathrm{X}=, \mathrm{MIN}=$, and $\mathrm{MAX}=$ plot options specify a sample size range of 100 to 250 (the same as in the analysis) for the X axis.

The tabular and graphical results are shown in Output 77.10.1 and Output 77.10.2, respectively.
Output 77.10.1 Power Values for Wilcoxon-Mann-Whitney Test
The POWER Procedure Wilcoxon-Mann-Whitney Test

| Fixed Scenario Elements |  |
| :--- | ---: |
| Method | O'Brien-Castelloe approximation |
| Number of Sides | U |
| Group 1 Variable | lidocaine |
| Group 2 Variable | Mir+lido |
| Pooled Number of Bins | 7 |
| Alpha | 0.05 |
| Group 1 Weight | 1 |
| Group 2 Weight | 1 |
| NBins per Group | 1000 |


| Computed Power |  |  |
| ---: | ---: | ---: |
| Index | Total | Power |
| $\mathbf{1}$ | 100 | 0.651 |
| $\mathbf{2}$ | 250 | 0.939 |

Output 77.10.2 Plot of Power versus Sample Size for Wilcoxon Power Analysis


The achieved power ranges from 0.651 to 0.939 , increasing with sample size.

## References

Agresti, A. (1980), "Generalized Odds Ratios for Ordinal Data," Biometrics, 36, 59-67.
Agresti, A. and Coull, B. A. (1998), "Approximate Is Better Than 'Exact' for Interval Estimation of Binomial Proportions," American Statistician, 52, 119-126.

Anderson, T. W. (1984), An Introduction to Multivariate Statistical Analysis, 2nd Edition, New York: John Wiley \& Sons.

Beal, S. L. (1989), "Sample Size Determination for Confidence Intervals on the Population Means and on the Difference between Two Population Means," Biometrics, 45, 969-977.

Blackwelder, W. C. (1982), "Proving the Null Hypothesis' in Clinical Trials," Controlled Clinical Trials, 3, 345-353.

Brown, L. D., Cai, T. T., and DasGupta, A. (2001), "Interval Estimation for a Binomial Proportion," Statistical Science, 16, 101-133.

Cantor, A. B. (1997), Extending SAS Survival Analysis Techniques for Medical Research, Cary, NC: SAS Institute Inc.

Castelloe, J. M. (2000), "Sample Size Computations and Power Analysis with the SAS System," in Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc.

Castelloe, J. M. and O'Brien, R. G. (2001), "Power and Sample Size Determination for Linear Models," in Proceedings of the Twenty-Sixth Annual SAS Users Group International Conference, Cary, NC: SAS Institute Inc.

Chernick, M. R. and Liu, C. Y. (2002), "The Saw-Toothed Behavior of Power Versus Sample Size and Software Solutions: Single Binomial Proportion Using Exact Methods," American Statistician, 56, 149155.

Chow, S.-C., Shao, J., and Wang, H. (2003), Sample Size Calculations in Clinical Research, Boca Raton, FL: CRC Press.

Connor, R. J. (1987), "Sample Size for Testing Differences in Proportions for the Paired-Sample Design," Biometrics, 43, 207-211.

Diegert, C. and Diegert, K. V. (1981), "Note on Inversion of Casagrande-Pike-Smith Approximate SampleSize Formula for Fisher-Irwin Test on $2 \times 2$ Tables," Biometrics, 37, 595.

Diletti, D., Hauschke, D., and Steinijans, V. W. (1991), "Sample Size Determination for Bioequivalence Assessment by Means of Confidence Intervals," International Journal of Clinical Pharmacology, Therapy, and Toxicology, 29, 1-8.

DiSantostefano, R. L. and Muller, K. E. (1995), "A Comparison of Power Approximations for Satterthwaite's Test," Communications in Statistics-Simulation and Computation, 24, 583-593.

Farrington, C. P. and Manning, G. (1990), "Test Statistics and Sample Size Formulae for Comparative Binomial Trials with Null Hypothesis of Non-zero Risk Difference or Non-unity Relative Risk," Statistics in Medicine, 9, 1447-1454.

Fisher, R. A. (1921), "On the 'Probable Error' of a Coefficient of Correlation Deduced from a Small Sample," Metron, 1, 3-32.

Fleiss, J. L., Tytun, A., and Ury, H. K. (1980), "A Simple Approximation for Calculating Sample Sizes for Comparing Independent Proportions," Biometrics, 36, 343-346.

Gatsonis, C. and Sampson, A. R. (1989), "Multiple Correlation: Exact Power and Sample Size Calculations," Psychological Bulletin, 106, 516-524.

Hocking, R. R. (1985), The Analysis of Linear Models, Monterey, CA: Brooks/Cole.
Hsieh, F. Y. (1989), "Sample Size Tables for Logistic Regression," Statistics in Medicine, 8, 795-802.
Johnson, N. L. and Kotz, S. (1970), Distributions in Statistics: Continuous Univariate Distributions, New York: John Wiley \& Sons.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1994), Continuous Univariate Distributions, volume 1, 2nd Edition, New York: John Wiley \& Sons.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995), Continuous Univariate Distributions, volume 2, 2nd Edition, New York: John Wiley \& Sons.

Johnson, N. L., Kotz, S., and Kemp, A. W. (1992), Univariate Discrete Distributions, 2nd Edition, New York: John Wiley \& Sons.

Jones, R. M. and Miller, K. S. (1966), "On the Multivariate Lognormal Distribution," Journal of Industrial Mathematics, 16, 63-76.

Kolassa, J. E. (1995), "A Comparison of Size and Power Calculations for the Wilcoxon Statistic for Ordered Categorical Data," Statistics in Medicine, 14, 1577-1581.

Kotz, S., Balakrishnan, N., and Johnson, N. L. (2000), Continuous Multivariate Distributions, 2nd Edition, New York: Wiley-Interscience.

Lachin, J. M. (1992), "Power and Sample Size Evaluation for the McNemar Test with Application to Matched Case-Control Studies," Statistics in Medicine, 11, 1239-1251.

Lakatos, E. (1988), "Sample Sizes Based on the Log-Rank Statistic in Complex Clinical Trials," Biometrics, 44, 229-241.

Lenth, R. V. (2001), "Some Practical Guidelines for Effective Sample Size Determination," American Statistician, 55, 187-193.

Maxwell, S. E. (2000), "Sample Size and Multiple Regression Analysis," Psychological Methods, 5, 434-458.
Miettinen, O. S. (1968), "The Matched Pairs Design in the Case of All-or-None Responses," Biometrics, 24, 339-352.

Moser, B. K., Stevens, G. R., and Watts, C. L. (1989), "The Two-Sample T Test versus Satterthwaite's Approximate F Test," Communications in Statistics—Theory and Methods, 18, 3963-3975.

Muller, K. E. and Benignus, V. A. (1992), "Increasing Scientific Power with Statistical Power," Neurotoxicology and Teratology, 14, 211-219.

O’Brien, R. G. and Castelloe, J. M. (2006), "Exploiting the Link between the Wilcoxon-Mann-Whitney Test and a Simple Odds Statistic," in Proceedings of the Thirty-First Annual SAS Users Group International Conference, Paper 209-31, Cary, NC: SAS Institute Inc.

O’Brien, R. G. and Castelloe, J. M. (2007), "Sample-Size Analysis for Traditional Hypothesis Testing: Concepts and Issues," in A. Dmitrienko, C. Chuang-Stein, and R. D’Agostino, eds., Pharmaceutical Statistics Using SAS: A Practical Guide, 237-271, Cary, NC: SAS Institute Inc.

O’Brien, R. G. and Muller, K. E. (1993), "Unified Power Analysis for $t$-Tests through Multivariate Hypotheses," in L. K. Edwards, ed., Applied Analysis of Variance in Behavioral Science, 297-344, New York: Marcel Dekker.

Owen, D. B. (1965), "A Special Case of a Bivariate Non-central $t$-Distribution," Biometrika, 52, 437-446.
Pagano, M. and Gauvreau, K. (1993), Principles of Biostatistics, Belmont, CA: Wadsworth.

Phillips, K. F. (1990), "Power of the Two One-Sided Tests Procedure in Bioequivalence," Journal of Pharmacokinetics and Biopharmaceutics, 18, 137-144.

Satterthwaite, F. E. (1946), "An Approximate Distribution of Estimates of Variance Components," Biometrics Bulletin, 2, 110-114.

Schork, M. A. and Williams, G. W. (1980), "Number of Observations Required for the Comparison of Two Correlated Proportions," Communications in Statistics-Simulation and Computation, 9, 349-357.

Schuirmann, D. J. (1987), "A Comparison of the Two One-Sided Tests Procedure and the Power Approach for Assessing the Equivalence of Average Bioavailability," Journal of Pharmacokinetics and Biopharmaceutics, 15, 657-680.

Self, S. G., Mauritsen, R. H., and Ohara, J. (1992), "Power Calculations for Likelihood Ratio Tests in Generalized Linear Models," Biometrics, 48, 31-39.

Senn, S. (1993), Cross-over Trials in Clinical Research, New York: John Wiley \& Sons.
Shieh, G. (2000), "A Comparison of Two Approaches for Power and Sample Size Calculations in Logistic Regression Models," Communications in Statistics-Simulation and Computation, 29, 763-791.

Shieh, G. and O'Brien, R. G. (1998), "A Simpler Method to Compute Power for Likelihood Ratio Tests in Generalized Linear Models," in Annual Joint Statistical Meetings of the American Statistical Association, Alexandria, VA: American Statistical Association.

Stuart, A. and Ord, J. K. (1994), Distribution Theory, volume 1 of Kendall's Advanced Theory of Statistics, 6th Edition, Baltimore: Edward Arnold Publishers.

Walters, D. E. (1979), "In Defence of the Arc Sine Approximation," The Statistician, 28, 219-232.
Wellek, S. (2003), Testing Statistical Hypotheses of Equivalence, Boca Raton, FL: Chapman \& Hall/CRC.

## Chapter 78 <br> The Power and Sample Size Application

## Contents

Overview: PSS Application . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6494
SAS Power and Sample Size . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6494
Getting Started: PSS Application . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6496
Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6496
The Basic Steps . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6496
A Simple Example . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6497
How to Use: PSS Application . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6512
Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6512
SAS Connections . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6512
Setting Preferences . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6515
Creating and Editing PSS Projects . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6519
Importing and Exporting Projects . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6526
Details: PSS Application . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6528
Software Requirements . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6528
Installation . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6528
Configuration . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6528
Example: Two-Sample $t$ Test . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6529
Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6529
Test of Two Independent Means for Equal Variances . . . . . . . . . . . . . . . . . . 6529
Test of Two Independent Means for Unequal Variances . . . . . . . . . . . . . . . . . 6540
Test of Mean Ratios . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6543
Additional Topics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6550
Example: Analysis of Variance . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6553
Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6553
The Example . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6553
Additional Topics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6562
Example: Two-Sample Survival Rank Tests . . . . . . . . . . . . . . . . . . . . . . . . . . 6569
Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6569
The Example . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6569
Additional Topics . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 6582

## Overview: PSS Application

## SAS Power and Sample Size

The SAS Power and Sample Size application (PSS) is a desktop application that provides easy access to power analysis and sample size determination techniques. The application is intended for students and researchers as well as experienced SAS users and statisticians.

Figure 78.1 shows the graphical user interface. PSS relies on the SAS/STAT procedures POWER and GLMPOWER for its computations.

Figure 78.1 PSS Application


This section describes the statistical tasks that are available with the application as well as its principal features.

## Analyses

PSS provides power and sample size computations for a variety of statistical analyses. Included are $t$ tests for means; equivalence tests and confidence intervals for means and proportions; exact binomial, chi-square, Fisher's exact, and McNemar tests for proportions; correlation and regression (multiple and logistic); one-way analysis of variance; linear models; tests of distribution; and rank tests for comparing survival curves.

Table 78.1 lists the analyses that are available.

Table 78.1 Available Analyses

| Category | Analysis |
| :--- | :--- |
| Means | One-sample $t$ test |
|  | Paired $t$ test |
| Confidence intervals | Two-sample $t$ test |
|  | One proportion |
|  | One-sample means |
|  | Paired means |
|  | Two-sample means |
| Equivalence tests | One proportion |
|  | One-sample means |
|  | Paired means |
|  | Two-sample means |
|  | One proportion |
| Proportions | Two correlated proportions |
|  | Two independent proportions |
| Correlation and regression | Pearson correlation coefficient |
|  | Logistic regression with a binary response |
| Multiple regression |  |
| Analysis of variance and linear models | One-way ANOVA |
|  | General linear univariate models |
| Survival analysis | Two-sample survival rank tests |
| Distribution tests | Wilcoxon Mann-Whitney test for two distributions |

## Features

PSS provides multiple input parameter options, stores the results in a project format, displays power curves, and produces narratives for the results. Narratives are descriptions of the input parameters and include a statement about the computed power or sample size. The SAS log and SAS code are also available.

All analyses offer computation of power or sample size. Some analyses offer computation of sample size per group as well as total sample size.

Where appropriate, several alternate ways of entering values for certain parameters are offered. For example, in the two-sample $t$ test analysis, means can be entered for individual groups or as a difference. The null mean difference can be specified as a default of zero or can be explicitly entered.

Information about existing analyses is stored in a project format. You can access each project to review the results or to edit your input parameters and produce another analysis.

## Getting Started: PSS Application

## Overview

This section is intended to get you off to a quick start with PSS. More detailed information about using the application is found in "How to Use: PSS Application" on page 6512 and in the example sections.

To start the application on a PC using the Windows operating system, select Start $\triangleright$ Programs $>$ SAS $\_$SAS Power and Sample Size 3.1 (or the latest release).

When you first use the application for a release, you are asked some configuration questions. For more information see the section "Configuration" on page 6528.
As an initial step, you also must define a SAS connection. If you have Foundation SAS software installed on the PC that you are using for PSS, this step can be done for you automatically. To define a connection or to determine whether one has already been defined, see the section "SAS Connections" on page 6512.

## The Basic Steps

Here are the basic steps that you follow to use PSS.

1. Start a new project by selecting File New on the menu bar or clicking the New icon on the toolbar.
2. In the New window, select the desired analysis type and click OK.

A project window for the analysis type appears with the Edit Properties page displayed. (The tabs on the Edit Properties page and their content vary according to the analysis type.)
3. Click each tab to enter the relevant data for the analysis. (For more information about the types of data to enter, see the example sections.)
4. After you have entered all the data, click the Calculate button.
5. After PSS calculates the results, the project window displays the View Results page with the Summary Table tab displayed by default.
6. To view other results or to review the SAS code or the SAS log, click any of the tabs on the left side of the View Results page.
7. To print any results page, select File $>$ Print on the menu bar.

The remainder of this section takes you through a simple example.

## A Simple Example

Suppose you want to determine the power for a new marketing study. You want to compare car sales in the southeastern region to the national average of 1.0 car per salesperson per day. You believe that the actual average for the region is 1.6 cars per salesperson per day. You want to test if the mean for a single group is larger than a specific value, so the one-sample $t$ test is the appropriate analysis. The conjectured mean is 1.6 and the null mean is 1.0 . You intend to use a significance level of 0.05 for the one-sided test. You want to calculate power for two standard deviations, 0.5 and 0.75 , and two sample sizes, 10 and 20 dealerships.

First, open a new project by selecting File New on the menu bar or clicking the New icon on the toolbar. The New window appears. Then, select the appropriate analysis.

Figure 78.2 New Window


For this example, the selected analysis is the One-sample $\mathbf{t}$ test in the Means section, as shown at the top of Figure 78.2. Select the analysis from the list and click OK. The One-sample $\boldsymbol{t}$ test project window appears with the Edit Properties page displayed, as shown in Figure 78.3.

Figure 78.3 Edit Properties Page


Enter a descriptive label of the project in the Project: field. For the example, change the description to Regional car sales versus the national average. The description is used to identify the project when you reopen it from the Open window.

Select File Save to save the description change. Note in Figure 78.3 that the title bar of the window contains your project description after you have saved the change.

Properties of the project are displayed on several tabs. You can change from tab to tab by clicking a tab or by clicking the Next tab or Previous tab buttons. To display help about the properties for a tab, click the Help button at the bottom of the Edit Properties page.

## Entering Parameter Values

First, click the Solve For tab and choose to calculate power or sample size. For this example, select the Power option, as shown in Figure 78.3.

Next, you must provide values for two analysis options and four parameters. These parameters are set in separate tabs on the Edit Properties page and are labeled Distribution, Hypothesis, Alpha, Mean, Standard Deviation, and Sample Size.

## Distribution

Click the Distribution tab to select a Normal or Lognormal distribution. For the example, you are using means rather than mean ratios, so select Normal, as shown in Figure 78.4.

Figure 78.4 Distribution Tab


## Hypothesis

Click the Hypothesis tab to select a one- or two-sided test. Because you are interested only in whether the southeastern region produces higher daily car sales than the national average, select One-sided test, as shown in Figure 78.5.

Figure 78.5 Hypothesis Tab


There are three one-sided test options: One-sided test, Upper one-sided test, and Lower one-sided test. The Upper one-sided test option would also be appropriate for this example.

## Alpha

Click the Alpha tab to specify one or more significance levels. Enter 0.05, as shown in Figure 78.6.
Figure 78.6 Alpha Tab


This value will be the default unless the default has been changed in the Preferences window. To set preferences, select Tools Preferences on the menu bar. For more information about setting preferences, see the section "Setting Preferences" on page 6515.

## Mean

Click the Means tab to enter one or more means and null means. For the example, enter 1.6 in the Mean table and 1.0 in the Null Mean table. Figure 78.7 shows the entered values.

Figure 78.7 Means Tab


Note that additional input rows are available if you want to enter additional sets of parameters. You can also append and delete rows using the $\square$ and $-\bar{l}$ buttons beneath the table. In addition, by selecting a row and right-clicking, you can choose to insert and delete rows in the body of the table from a pop-up menu.

## Standard Deviation

Click the Standard Deviation tab to enter standard deviations. You are interested in two standard deviations, 0.5 and 0.75 . Enter them in the table, as shown in Figure 78.8.

Figure 78.8 Standard Deviation Tab


## Sample Size

You want to be able to sample between 10 and 20 dealerships. Click the Sample Size tab and enter these two values, as shown in Figure 78.9.

Figure 78.9 Sample Size Tab


## Scenarios

The input values are combined into one or more scenarios. In this case, each of the two standard deviations is combined with each of the two sample sizes for a total of four scenarios. Then power is computed for each scenario. In this example, only a single value or setting is present for the mean, null mean, and alpha level, so they are common to all scenarios.

## Results Options

Click the Results tab to select results options including a Summary Table and a Power by Sample Size graph.
Figure 78.10 Results Tab


For this example, select both results check boxes: Create summary table and Create power by sample size graph, as shown in Figure 78.10. These selections can also be set as preferences; see the section "Setting Preferences" on page 6515.

## Customizing the Power by Sample Size Graph

Click the Customize button beside the Create power by sample size graph check box to customize the graph. The Customize Graph window contains two tabs: Axis Orientation and Value Ranges, as shown in Figure 78.11.

Figure 78.11 Customize Graph Window with Axis Orientation Tab


Click the Axis Orientation tab to select which quantity you would like to plot on the vertical axis. You can choose to display the quantity solved for (either power or sample size) on the vertical axis or you can choose to display power or sample size on the vertical axis with the other quantity appearing on the horizontal axis. The default is Quantity solved for (or power) on the vertical axis, which is appropriate for this graph.

The summary table is created using the two sample sizes specified in the Sample Size table, 10 and 20. If you want to create a graph that contains more than these two sample sizes, you can do so by customizing the value ranges for the graph. Click the Value Ranges tab to set the axis range for sample sizes, as shown in Figure 78.12.

Figure 78.12 Customize Graph Window with Value Ranges Tab


Enter 5 for the minimum and 30 for the maximum. Also, select Interval between points in the drop-down list and enter a value of 1 . These values set the sample size axis to range from 5 to 30 in increments of 1 . The completed Value Ranges section of the window is shown in Figure 78.12.

When you solve for power, you can set a range for sample size values, but not for the powers; and vice versa when you solve for sample size. That is, you cannot set the range of axis values for the quantity that you are solving for.

Click OK to save the values that you have entered and return to the Edit Properties page.

## Performing the Analysis

You have now specified all of the necessary input values. Click Calculate to perform the analysis, as shown in Figure 78.13.

Figure 78.13 Calculate Button on the Edit Properties Page


Alternatively, you could choose to save the information that you have entered by selecting File $\rightarrow$ Save from the menu bar or clicking the Save toolbar icon, and perform the analysis at another time. No error checking is done when you save the project.

You can close the project by selecting File Close on the menu bar or clicking the window close $\mathbf{X}$ in the upper right corner of the project window. You can reopen a project by selecting File Open on the menu bar or clicking the Open toolbar icon.

For this example, click Calculate.

## Viewing the Results

Results appear on the View Results page and are viewable in separate tabs. The tabs include Summary
Table, Graph, Narratives, SAS Log, and SAS Code (located on the left side of the View Results page).
The Summary Table and Graph tabs appear if you selected those options on the Results tab of the Edit Properties page. The other tabs always appear.

## Summary Table

Click the Summary Table tab to view the summary table.
Figure 78.14 Summary Table Tab with Fixed Scenario Elements and Computed Power Tables


The Summary table consists of two subtables, as shown in Figure 78.14. The Fixed Scenario Elements table includes the parameters or options that have a single value for the analysis. The Computed Power table contains the input parameters that have been given more than one value, and it shows the computed quantity, power.

Thus, the Computed Power table contains four rows for the four combinations of standard deviation and sample size. From the table you can see that all four powers are high. The smallest value of power, 0.754 , is associated with the largest standard deviation and the smallest sample size. In other words, the probability of rejecting the null hypothesis is greater than $75 \%$ in all four scenarios.

## Power by Sample Size Graph

Click the Graph tab to view the power by sample size graph.
The power by sample size graph in Figure 78.15 contains one curve for each standard deviation. For a standard deviation of 0.5 (the upper curve), increasing sample size above 10 does not lead to much increase in power. If you are satisfied with a power of 0.75 or greater, 10 samples would be adequate for standard deviations between 0.5 and 0.75 .

Figure 78.15 Graph Tab with Power by Sample Size Graph


## Narratives

Click the Narratives tab to display a facility for creating narratives.
Narratives are descriptions of the values that compose each scenario and include a statement about the computed power or sample size.

To create narratives, choose one or more scenarios in the table at the bottom of the tab. A narrative for each selected scenario is displayed in the top portion of the tab. See Figure 78.16.

Figure 78.16 Narrative Tab


For the example, select the first row in the table. The following narrative is displayed for the scenario with a standard deviation of 0.5 and a sample size of 10 :

For a one-sample $t$ test of a normal mean with a one-sided significance level of 0.05 and null mean 1 , assuming a standard deviation of 0.5 , a sample size of 10 has a power of 0.967 to detect a mean of 1.6.

You can select several rows in the table. As you select each one, a corresponding narrative is created and displayed in the top portion of the table. Selecting a second scenario (the third row) produces the following output, where the narrative for the first row is followed by the narrative for the third row:

For a one-sample $t$ test of a normal mean with a one-sided significance level of 0.05 and null mean 1 , assuming a standard deviation of 0.5 , a sample size of 10 has a power of 0.967 to detect a mean of 1.6.

For a one-sample $t$ test of a normal mean with a one-sided significance level of 0.05 and null mean 1 , assuming a standard deviation of 0.75 , a sample size of 10 has a power of 0.754 to detect a mean of 1.6.

## Other Results

Other results include the SAS $\log$ and the SAS code.
The SAS $\log$ that was produced when the Calculate button was last clicked appears on the SAS Log tab.
The SAS statements that produced the results appear on the SAS Code tab.

## Printing Results

To print one or more results, select File Print from the menu bar or click the Print toolbar icon. A window is displayed that lists all available results, as shown in Figure 78.17. Select the results that you want to print and click OK.

Figure 78.17 Print Selection Window


## Changing Properties

If you want to change some values of the properties and rerun the analysis, change to the Edit Properties page and continue. The icons for selecting the Edit Properties and View Results pages are in the command bar just below the project window title.

## Closing the Project

When you are finished working with a project, close it by clicking the $\mathbf{X}$ in the upper right corner of the project window or selecting File Close on the menu bar. If you have not saved the project, you will be asked if you want to save it before closing.

## Opening a Project

You can reopen existing projects using the Open window. Select File» Open on the menu bar or click the Open toolbar icon.

Figure 78.18 Open Window Containing the Analysis Created in the Example


As shown in Figure 78.18, the analysis that you just completed is listed in the table. The label that you assigned to it, Regional car sales versus the national average, appears in the Project column of the table. The table also contains the date that the analysis was last modified. If you do not see the project that you are looking for, change the value of the Display projects by date box to All by selecting All from the drop-down list, and click the Change display button.

You can sort the projects in the table by clicking the header of the desired column. The sort direction is indicated by arrows displayed in the column header.

Select the project that you want to open and click OK. You can also double-click the project entry to open it.

## Changing Values and Rerunning the Analysis

After viewing the graph, you might want to re-create the graph with a different range for sample sizes. On the Results tab of the Edit Properties page, click the Customize button for the power by sample size graph. The Customize Graph window is displayed.

On the Value Ranges tab of the window, change the Maximum value in the Sample Size table from 30 to 20. Click OK.

Rerun the analysis by clicking Calculate. The View Results page is displayed again and the graph now has the new maximum value for the sample size axis.

## How to Use: PSS Application

## Overview

The PSS application is an application that resides on your desktop. It requires a connection to SAS software either on your desktop machine or a remote machine. You can set default values for several parameters and options as preferences. More detail on creating and editing projects is provided. Projects can be imported and exported.

## SAS Connections

Connections to SAS servers are defined in the Preferences window. To access the Preferences window, select Tools $»$ Preferences on the menu bar.

Click the SAS Connection tab to select or define a connection to a SAS server. A connection to a SAS server is required in order to calculate results. The server can be on your local (desktop) machine or on a remote machine.

You can define several SAS connections and choose the one you want to use. To select a previously defined connection, choose it from the Connection list on the SAS Connection tab; see Figure 78.19.

Figure 78.19 SAS Connection Tab


To define a SAS connection, click the Define connection button. The Connection List window appears, as shown in Figure 78.20. To create a new connection, click Add. To edit an existing connection, select it in the Connection List and click Edit.

Figure 78.20 Connection List

| I. Connection List |  |  |
| :--- | :--- | :--- |
| Connection Name | Address (Host) | Port |
| Local SAS Connection | Local Host |  |
| Remote UNIX server | malina.unx.sas.com | 9120 |
| Remote Windows server | d1234.na.sas.com | 5555 |
|  |  |  |

## Defining a SAS Connection

After you click the Add or Edit button, the Define SAS Connection window appears, as shown in Figure 78.21. If you clicked Edit, the previously defined information is available for editing.

Figure 78.21 Define Connection Window


Enter a descriptive label for the connection. The label is used to distinguish among the connections in the connections list.

Then, select Yes or No to specify whether the SAS connection is to the local machine (that is, the one on which PSS is running) or to a remote machine, respectively.

## Defining a Local Connection

To define a connection to the local machine, enter the full path name of (or browse for) the SAS executable file (sas.exe on Microsoft Windows).

Test the SAS connection by clicking the Test SAS Connection button.

## Defining a Remote Connection

To define a connection to a remote machine, select either the UNIX or Windows option to indicate that the remote SAS server is on a machine running the UNIX or Microsoft Windows operating systems, respectively. Then, specify the machine name and port number that the SAS/Connect spawner is using on the remote machine. Contact the SAS server administrator for this information.

If the remote machine is running Microsoft Windows, select the User id and password are required if authentication is required to access the SAS server (that is, if the SAS -security option is used). By default, authentication is required for SAS servers running on UNIX operating systems.

Test the SAS connection by clicking the Test SAS Connection button.

## Additional Settings

Click the Settings button on the Define SAS Connection window to access some additional settings for a remote connection to a SAS server. For the most part these settings are prompts that PSS expects to receive from the SAS/CONNECT spawner on the remote machine, as shown in Figure 78.22.

If the remote SAS server is on a UNIX machine, you must specify the full pathname of the SAS command. Contact the SAS server administrator for this information.

Figure 78.22 Connection Settings Window


## Setting Preferences

In the Preferences window you can set default values for options that are used by all analyses.
To access the Preferences window, select Tools $\downarrow$ Preferences on the menu bar. Figure 78.23 shows the Preferences window.

Figure 78.23 Preferences Window


Preference values are used as the defaults for each newly opened project (that is, those that are opened from the New window). For a specific project, each of these default values can be overridden on the Edit Properties page.

Changes in preferences do not change the state of an existing analysis (that is, one that is accessed from the Open window).

## Selecting the Quantity to Solve For

Click the Solve For tab to select Power or Sample Size as the default value to be solved for; see Figure 78.23. For confidence interval analyses, selecting Power is equivalent to selecting Prob(Width).

For analyses that offer both Sample size per group and Total sample size, the Sample size option on this page corresponds to total sample size.

## Setting Alphas

Click the Alpha tab to enter one or more values for alpha. Alpha is the significance level (false positive probability). For confidence interval analyses, alpha values are transformed into confidence levels by ( $1-$ alpha). For example, an alpha of 0.05 would represent a confidence level of 0.95 .

To set default values of alpha, enter one or more values in the Alpha data entry table. See Figure 78.24. It is not necessary to have any default values for alpha. Add more rows to the table as needed using the ${ }^{-7}$ button at the bottom of the table.

Figure 78.24 Alpha Preference Tab


## Setting Powers

Click the Power tab to enter one or more values for power. It is not necessary to have any default values for power. For confidence interval analyses, power values are treated as prob(width) values.

To set default values of power, enter one or more values in the Power data entry table; see Figure 78.25.

Figure 78.25 Power Preference Tab


## Setting Results Options

Click the Results tab to make default selections for the summary table and the power by sample size graph options; see Figure 78.26.

Figure 78.26 Results Options Preferences Tab


The summary table consists of the input parameter values and the calculated quantity (power or sample size). Select the Create summary table check box to create the table by default.

To request that an analysis create a power by sample size graph by default, select the Create power by sample size graph check box.

## Creating and Editing PSS Projects

A PSS project is an instance of an analysis. The first decision in using PSS is to choose the appropriate test or design. Select the File New on the menu bar or click the New icon on the toolbar. The New window appears with a list of the available analyses. Select the type of analysis that you want from the list and click OK.

When the project is first opened, the Edit Properties page is displayed. It is described in the section "Editing Properties" on page 6520.

After the properties have been specified and the analysis is performed, the View Results page is displayed. See the section "Viewing the Results" on page 6524.

A project that has been saved and closed can be reopened from the Open window. Select File»Open on the menu bar or click the Open icon on the toolbar.

## Editing Properties

The Edit Properties page consists of several analysis options and input parameters that are relevant to the particular analysis. These options and parameters are organized on several tabs, as shown in Figure 78.27.

Figure 78.27 Edit Properties Page


The Edit Properties page contains various controls by which you can enter values or select choices. In addition to the usual data entry controls such as text fields and check boxes, several specialized controls are present: data entry tables and the Alternate Forms control. More detailed descriptions follow.

## Using Data Tables

Data entry tables are composed of data entry fields for one or more rows and columns. Figure 78.28 shows a two-row, two-column table.

Figure 78.28 Two-Column Data Entry Table with Controls


Type an appropriate value in each field. It is not necessary to type data in all rows or to delete empty rows. However, if a table has more than one column, the cells of a row must be completely filled or completely blank. Rows with values in some but not all cells are not allowed.

To append more rows, click the $\square^{-}$button beneath the table. To delete the last row of the table, click the ${ }^{-2}$ button.

Also, you can display a pop-up menu to perform additional actions such as inserting and deleting rows. First, select the row to insert before or delete, then right-click to display the pop-menu and select the desired action.

## Using Alternate Forms

For some input parameters, there are several ways in which data may be entered. For example, in the two-sample $t$ test analysis, group means can be entered as either individual means or a difference between means.

The alternate forms are displayed in a drop-down list with an adjacent button, as shown in Figure 78.29. The button enables you to cycle through the alternatives, displaying each one in turn. To see what forms are available, you can open the drop-down list and select the one you want or you can click the button until the form that you want is displayed.

Figure 78.29 Select a form Drop-Down List and Button


The alternate form last used for an analysis is saved and displayed as the default when a new instance of the analysis is opened.

## Customizing Graphs

The Edit Properties page for all analyses contains a Results tab. You can choose to create a graph, and you can optionally choose to customize the graph by clicking the Customize button that is beside the Create power and sample size graph choice.

As shown in Figure 78.30, the Customize Graph window consists of an Axis Orientation tab and a Value Ranges tab. Use the Axis Orientation options to specify which axes you want used for power and for sample size. Use the Value Ranges settings to specify the axis range for the non-target quantity (that is, the power axis if you are solving for sample size or the sample size axis if you are solving for power).

Figure 78.30 Customize Graph Window


When specifying a value range, you can specify a minimum value and a maximum value. Also, you can select either the Number of points or the Interval between points choice for the axis and specify a value. All of these values are optional; specify only the ones you want.

## Scenarios

A scenario is one instance of a complete set of values for an analysis. For example, if two alpha values and two total sample size values are specified with all other input parameters taking only a single value, there would be four scenarios-the four combinations of two alphas and two sample sizes.

## Performing the Analysis

To perform the analysis, click Calculate at the lower right of the Edit Properties page. The input parameters are checked for validity, and the analysis is performed. The View Results page is then displayed.

## Viewing the Results

The results appear in separate tabs on the View Results Page. These tabs include Summary Table, Graph, Narratives, SAS Log, and SAS Code.

## Viewing the Summary Table

Click the Summary Table tab to view the summary table. It consists of two subtables, as shown in Figure 78.31. The Fixed Scenario Elements table includes the options and parameter values that are constant for the analysis. The computed Power table includes the calculated power or sample size values and the values for input parameters that have multiple values specified for the analysis.

Figure 78.31 View Results Page with Summary Table


## Creating Narratives

Click the Narratives tab to display a facility to create narratives. Narratives are descriptions of the input parameter values and calculated quantities in sentence or paragraph form. Each narrative corresponds to one calculated quantity value.

The Narratives tab is divided into a narrative selector panel and a narrative display panel. To create a narrative, select the row in the narrative selector panel that corresponds to it. You can select as many rows as you want. See Figure 78.32.

Figure 78.32 Narrative Selector and Display


The narrative selector table often contains columns whose values do not vary. For example, in Figure 78.32, the Sides, NullMean, Alpha, and Mean columns contain values that do not vary. You can hide these columns by selecting the Hide columns with constant input values check box.

## Viewing the SAS Log and Code

Click the SAS Code tab to view the SAS statements that are used to generate the analysis results. Click the SAS Log tab to view the SAS log that corresponds to the analysis.

The SAS code differs slightly from the statements in the SAS log. Statements that are used to place the results in the location maintained by the application are not included. This is done to prevent you from overwriting the results stored by the application if you run the SAS code outside of the application.

## Printing Results

To print one or more results, click the Print icon on the toolbar or select File Print on the menu bar. The Select Results to Print window is displayed. You can choose to print one or more of the results by selecting the corresponding options here.

## Saving the Project

To save a project, click the Save toolbar icon or select File Save from the menu bar. Projects can be saved even if some of the information is invalid. Error checking is performed when the Calculate button is clicked.

## Closing the Project

To close a project, click the $\mathbf{X}$ in the upper right corner of the project window or select File Close from the menu bar.

## Importing and Exporting Projects

PSS projects can be imported from the same machine or a different machine. Also, the active project (the project that is open and on top of any other open projects) can be exported.

## Importing Projects

A PSS project that was created on another machine or by another user can be imported and used. Also, importing projects is the recommended way of moving existing PSS projects that were created with PSS release 2.0 (a Web application) to PSS release 3.1 (a desktop application).

PSS files are stored in a folder entitled pss. The pss folder contains a project.xm/ file and individual folders for each project. See Figure 78.33.

Figure 78.33 PSS Directory Structure


If PSS files are on another machine, they must first be copied to a temporary location on the desktop machine that is running PSS. The entire pss folder should be copied.

To import projects, select File Import from the menu bar. Then, specify the full pathname of the pss folder.
Figure 78.34 Import Projects Window


To import PSS 2.0 files, you need to find the pss folder. The easiest way to do this is to search for the project.xm/ file. If you find several files with this name, you need to decide which one or more to import.

## Exporting the Active Project

If you want to send a PSS project to someone, you can export the active project. The active project is the one that is open and that has focus (is displayed on top of any other open projects). Select File Export active project and specify a temporary directory to hold the exported project.

The recipient must import the project using PSS.

## Details: PSS Application

## Software Requirements

PSS is available in SAS/STAT 13.2 or later for the following platforms: Microsoft Windows 7 and 8.
Two configurations are available for SAS connections: local and remote. With the local configuration, PSS and SAS must reside on the same machine. With the remote configuration, PSS and SAS can reside on different machines. SAS connections are defined and selected on the SAS Connection tab on the Preferences window. More information about SAS connections is found in the section "SAS Connections" on page 6512.

For both configurations, Base SAS and SAS/STAT software must be installed and SAS/GRAPH software is recommended.

For the remote configuration, SAS/CONNECT and SAS/IntrNet software must also be installed. For more information about configuring the remote SAS server, click Help $\downarrow$ Contents on the menu bar and then click Configuring a Remote SAS Server under Special Topics in the table of contents.

## Installation

SAS Power and Sample Size is installed separately from the SAS/STAT product. Contact your SAS site representative to have the application installed.

SAS Power and Sample Size is installed using the SAS Software Deployment Wizard. It is listed as an available product with, but separate from, Foundation SAS which contains the SAS/STAT and SAS/GRAPH products that are required for using the application.

## Configuration

When you first run SAS Power and Sample Size 3.1 (PSS), you are asked to provide configuration information.
First, you are asked for the name of a directory to contain the your power and sample size projects. A folder named $p s s$ is created in the specified directory, and projects are stored in the $p s s$ folder. This directory cannot be the same as the one used by PSS 2.0. If it is, PSS requires that another folder name be provided.

Then, if the appropriate release of the SAS System is available on the desktop machine, you are asked whether a connection should be automatically created to it. If you respond no, then PSS informs you that a connection to the SAS server is necessary and asks if you want to select one now or later. A connection to a SAS server is not necessary to use the application until the Calculate button on the Edit Properties page of a project is clicked. More information about connections is available in the section "Setting Preferences" on page 6515.

Then, PSS displays a wizard to help you import existing PSS projects from either a previous release (PSS 2.0) or the current one (PSS 3.1). More information is available in the section "Importing Projects" on page 6526.

## Example: Two-Sample $t$ Test

## Overview

The one-sample $t$ test compares the mean of a sample to a given value. The two-sample $t$ test compares the means of two samples. The paired $t$ test compares the mean of the differences in the observations to a given number. PSS provides power and sample size computations for all of these types of $t$ tests. For more information about power and sample size analysis for $t$ tests, see Chapter 77, "The POWER Procedure."

The two-sample $t$ test tests for differences or ratios between means for two groups. The groups are assumed to be independent. This example describes three examples using the two-sample $t$ test: for equal variances, for unequal variances, and for mean ratios.

## Test of Two Independent Means for Equal Variances

Suppose you are interested in testing whether an experimental drug produces a lower systolic blood pressure than a placebo does. Will 25 subjects per treatment group yield a satisfactory power for this test? From previous work, you expect that the blood pressure is 132 for the control group and 120 for the drug treatment group and that the standard deviation is 15 for both groups. You want to use a one-sided test with a significance level of 0.05 . Because there are two independent groups and you are assuming that blood pressure is normally distributed, the two-sample $t$ test is an appropriate analysis.

Start by creating a new project. Select File New. In the New window, select Two-sample $\boldsymbol{t}$ test from the list. The Two-Sample $t$ test project window appears, with the Edit Properties page displayed.

## Editing Properties

On this page enter a name to describe the project and enter project properties. Click each tab on the Edit Properties page to enter the desired properties. You can also change tabs by clicking the Next tab or Previous tab buttons. See Figure 78.3.

Figure 78.35 Two-Sample $t$ Test


## Project Description

The description is used to identify this particular project in the Open and Delete windows. Type a description for your project in the Project: text box.

For this example, change the description to Experimental blood pressure drug with two groups, as shown in Figure 78.35.

## Solve For

For the two-sample $t$ test analysis, you can choose to solve for power, sample size per group, or total sample size. Specify the desired quantity type on the Solve For tab.

Click the Solve For tab and select the Power option as shown in Figure 78.35. For information about solving for sample size, see the section "Solving for Sample Size" on page 6550.

## Distribution

Click the Distribution tab to select a distribution option that specifies the underlying distribution for the test statistic, as shown in Figure 78.36.

Figure 78.36 Distribution Tab

## Analysis: Two-sample $\boldsymbol{t}$ test

| Project: Experimental blood pressure drug with two groups |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: |
| Properties |  |  |  |  |
| Alpha | Means | Standard Deviation | Sample Size | Results |
| Solve For |  | Distribution | Hypothesis | Test |
| Select the distribution of the testLognormalNormal |  |  |  |  |

For this example, you are interested in means rather than mean ratios, so select the Normal option.

## Hypothesis

Click the Hypothesis tab to select the type of test; see Figure 78.37.
Figure 78.37 Hypothesis Tab


You can choose either a one- or two-sided test. If you do not know the direction of the effect (that is, whether it is positive or negative), the two-sided test is appropriate. If you know the effect's direction, the one-sided test is appropriate. For the one-sided test, the alternative hypothesis is assumed to be in the same direction as the effect. If you specify a one-sided test and the effect is in the unexpected direction, the results of the analysis are invalid.

The One-sided test option assumes that you know the correct direction of the test. Select the Lower one-sided test and Upper one-sided test options to explicitly indicate the direction of the one-sided test.
Because you are interested only in whether the experimental drug lowers blood pressure, select the One-sided test option on the Hypothesis tab.

## Test

Click the Test tab to select either the pooled $t$ test or the Satterthwaite $t$ test.
Figure 78.38 Test Tab


With the independent variances that the example uses, select Pooled $\mathbf{t}$ test option. The Satterthwaite $t$ test is used with unequal variances; it is available only with the normal distribution.

## Alpha

Click the Alpha tab to specify one or more significance levels, as shown in Figure 78.39.

Figure 78.39 Alpha Tab


Alpha is the significance level (that is, the probability of falsely rejecting the null hypothesis). If you frequently use the same values for alpha, set them as defaults in the Preferences window. See the section "Setting Preferences" on page 6515 for more information about setting preferences.

Type the desired significance level of 0.05 in the first cell of the Alpha table (if it is not already the default value).

## Means

Click the Means tab to select one of four possible ways to enter the means and the null mean difference, as shown in Figure 78.40.

Figure 78.40 Means Tab


Select one of the following forms from the Select A Form list. The four available forms are:

## Difference between means

Enter the difference between the group means. The null mean difference is assumed to be 0.

## Group means

Enter the means for each group. The null mean difference is assumed to be 0 . The difference is formed by subtracting the mean for group 1 from the mean for group 2 .

## Difference between means, Null difference

Enter the difference between the group means and a null mean difference.

## Group means, Null difference

Enter the means for each group and a null mean difference. The difference is formed by subtracting the mean for group 1 from the mean for group 2.

For this analysis, you can enter the means for the two groups either individually or as a difference. If your null mean difference is not zero, enter that value in the Null Mean table. (The Null Mean table is displayed only for the Group means, Null Difference and Difference between means, Null difference forms.)

For this example, a null mean difference of 0 is reasonable, so select the Group means form from the list, as shown in Figure 78.40. Enter the control mean of 132 in the first row of the first column and the experimental mean of 120 in the first row of the second column.

## Standard Deviation

Click the Standard Deviation tab to enter the standard deviation for the two groups. It is assumed to be equal for both groups.

For the example, enter a single value of 15, as shown in Figure 78.41.
Figure 78.41 Standard Deviation Tab


## Sample Size

Click the Sample Size tab to select one of three possible ways to enter the sample sizes, as shown in Figure 78.40.

Select one of the following forms from the Select A Form list:

## Sample size per group

Enter the sample size for one of the two groups. The group sizes are assumed to be equal.

## Group sample sizes

Enter the sample size for each of the two groups. The group sizes can be equal or unequal.

## Total N, Group weights

Enter the total sample size for the two groups and the relative sample sizes for each group. For more information about using relative sample sizes, see the section "Using Unequal Group Sizes" on page 6551.

Examine the alternatives by clicking the Select a form down arrow. For this example, select the Sample size per group form. You want to examine a curve of powers in the power by sample size graph, so enter the values 20, 25, and 30 in the Sample Size table, as shown in Figure 78.42. If you need to add more rows to the table, add them by clicking the ${ }^{-7}$ button beneath the table.

Figure 78.42 Sample Size Tab


## Summary of Properties

Table 78.2 contains the values of the input parameters for the example.

Table 78.2 Summary of Input Properties

| Parameter | Value |
| :--- | :--- |
| Solve for | Power |
| Distribution | Normal |
| Hypothesis | One-sided test |
| Test | Pooled t test |
| Alpha | 0.05 |
| Means form | Group means |
| Means | 132,120 |
| Standard deviation | 15 |
| Sample size form | Sample size per group |
| Sample size | $20,25,30$ |

Results
Click the Results tab to request desired results. Summary table and power by sample size graph options are available.

For the example, select the Create summary table and Create power by sample size graph check boxes.
Click Calculate to perform the analysis. If there are no errors in the input values, the View Results page appears. If there are errors in the input parameter values, you are prompted to correct them.

## Viewing Results

The results are listed on separate tabs on the View Results page. Click the tab of each result that you want to view.

## Summary Table

Click the Summary Table tab to view a table that includes the values of the input parameters and the computed quantity (in this example, power). See Figure 78.43.

Figure 78.43 Results Page with Summary Table


The table consists of two subtables: the Fixed Scenario Elements table that contains the input parameters that have only one value for the analysis, and the computed Power table that contains the input parameters that have more than one value for the analysis and the corresponding power. Only the N per group parameter appears in the Computed Power table; all of the other input parameters have a single value. The computed power for a sample size per group of 25 is 0.874 . Thus, you have a probability of 0.87 that the study will find the expected result if the assumptions and conjectured values are correct.

## Power by Sample Size Graph

Click the Graph tab to view a power by sample size graph that displays power on the vertical axis and sample size per group on the horizontal axis. See Figure 78.44.

Figure 78.44 Power by Sample Size Graph


The range of values for the horizontal axis is 20 to 30 , which were the smallest and largest values, respectively, that you entered on the Sample Size tab. You can customize the graph by specifying the values for the sample size axis (see the section "Customizing Graphs" on page 6522).

## Narratives

Click the Narratives tab to create and display a sentence- or paragraph-length text summary of the input parameter values and the computed quantity for combinations of the input parameter values; see Figure 78.45.

Figure 78.45 Narrative Selector and Display

| For a two-sample pooled $t$ test of a normal mean difference with a one-sided significance level of 0.05 , assuming a cormon standard deviation of 15 , a sample size of 20 per group has a power of 0.799 to detect a difference between the means 132 and 120. |  |  |  |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Create Narratives Select one or more scenarios |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |
| Select | Index | Sides | Alpha | Mean 1 | Mean2 | StdDev | NPerGroup | Nulloff | Power | Error | Info |
| $\square$ | 1 | 1 | 0.05 | 132 | 120 | 15 | 20 | 0 | 0.79940818 |  |  |
| $\square$ | 2 | 1 | 0.05 | 132 | 120 | 15 | 25 | 0 | 0.87355245 |  |  |
| $\square$ | 3 | 1 | 0.05 | 132 | 120 | 15 | 30 | 0 | 0.92176938 |  |  |
| $\leqslant$ |  |  |  |  |  |  |  |  |  |  |  |

To create a narrative, selected the desired scenario (row) in the narrative selector table at the bottom of the Narratives tab.

In this example, select the narrative for the sample size per group of 20 , which yields a power of 0.799 . The following text summary is displayed:

```
For a two-sample pooled t test of a normal mean difference with a one-sided
significance level of 0.05, assuming a common standard deviation of
15, a sample size of 20 per group has a power of 0.799 to detect a
difference between the means 132 and 120.
```

To create other narratives, select the desired rows in the narrative selector table. If you also select the second row for the sample size of 25 , another text summary is displayed below the first one:

```
For a two-sample pooled t test of a normal mean difference with a one-sided
significance level of 0.05, assuming a common standard deviation of
15, a sample size of 20 per group has a power of 0.799 to detect a
difference between the means 132 and 120.
```

```
For a two-sample pooled t test of a normal mean difference with a one-sided
significance level of 0.05, assuming a common standard deviation of
15, a sample size of 25 per group has a power of 0.874 to detect a
difference between the means 132 and 120.
```

To change some values of the analysis and rerun it, select the Edit Properties page, change the desired properties, and click the Calculate button again.

## Test of Two Independent Means for Unequal Variances

In the preceding example, you assumed that the population standard deviations were equal. If you believe that the population standard deviations are not equal, use the same two-sample $t$ test analysis as with the preceding example, but change the test option and enter group standard deviations.

You can use the previous example to demonstrate this test. If the project is not already open, open it by selecting File Open on the menu bar, and then selecting the project that you have been using.

Make a copy of the project by selecting File $>$ Save As. Enter a different project description, Experimental blood pressure drug with two groups for unequal variances. Click OK.

The copy of the project is opened, and the current project is closed.

## Editing Properties

## Test

On the Test tab of the copied project, change the test to Satterthwaite $\mathbf{t}$ test, as shown in Figure 78.46.
Figure 78.46 Satterthwaite $t$ Test Option


## Specifying Group Standard Deviations

Click the Standard Deviation tab and enter the group standard deviations of 12 and 15 on a single row, as shown in Figure 78.47.

Figure 78.47 Group Standard Deviations


## Summary of Input Parameters

Table 78.3 contains the values of the input parameters for the example.

Table 78.3 Summary of Input Parameters

| Parameter | Value |
| :--- | :--- |
| Distribution | Normal |
| Hypothesis | One-sided test |
| Test | Satterthwaite $t$ test |
| Alpha | 0.05 |
| Means form | Group means |
| Means | 132,120 |
| Standard deviation | 12,15 |
| Sample size form | Sample size per group |
| Sample size | $20,25,30$ |

Click Calculate to run the analysis.

## Viewing Results

The power for a sample size per group of 25 is 0.924 , as shown in Figure 78.48. Notice that the actual alpha is 0.0499 . This is because the Satterthwaite $t$ test is (slightly) biased.

Figure 78.48 Satterthwaite Test Results
Experimental blood pressure drug with two groups for unequal variances


If you modified the previous example, when you select the Narratives tab, the following message is displayed:

```
Previously selected narratives have been cleared because one or more input
parameter values have changed.
```

In the previous analysis, you created narratives for two scenarios. Because this analysis uses group standard deviations, those selected narratives were cleared. The message would also have appeared if you had changed the number of scenarios.

Use the narrative selector table to create other narratives.

## Test of Mean Ratios

Instead of comparing means for a control and drug treatment group, you might want to investigate whether the blood pressure of the treatment group is lowered by a given percentage of the control group, say 10 percent. That is, you expect the ratio of the treatment group to the control group to be $90 \%$ or less.

PSS provides a two-sample $t$ test of a mean ratio when the data are lognormally distributed.
For mean ratios, the coefficient of variation (CV) is used instead of standard deviation. In this example, you can expect the CV to be between 0.5 and 0.6 . You also want to compare an equally weighted sampling of groups with an overweighted sampling in which the control group contains twice as many subjects as the treatment group: 50 and 25 , respectively.

Make a copy of the project by selecting File Save As. Enter a different project description, Percent improvement with blood pressure drug.

The copy of the project is opened.

## Editing Properties

Several of the input parameters for the test of mean ratios differ from the ones described in the section "Test of Two Independent Means for Equal Variances" on page 6529. Mean ratios and coefficients of variation are used instead of mean differences and standard deviations. These two parameters are discussed in detail in this section. For the input parameters and options that have been discussed previously in this example, only the values for this example are given.

## Solve For Tab

Click the Solve For tab to select the Power option as the quantity to be solved for, as shown in Figure 78.49.

Figure 78.49 Project Description, Solve for Tab


## Distribution

You are interested in mean ratios rather than means, so select the Lognormal option on the Distribution tab, as shown in Figure 78.50.

Figure 78.50 Distribution Tab with Lognormal Option


## Hypothesis and Alpha

Click the Hypothesis tab and select the One-sided test option.
Click the Alpha tab and type 0.05 as the significance level in the first cell of the table, if it is not already there.

## Means

Click the Means tab to select the input form for entering mean ratios. There are four alternate forms for entering means or mean ratios:

## Mean ratio

Enter the ratio of the two group means-that is, the treatment mean divided by the reference mean. The null ratio is assumed to be 1 .

## Group means

Enter the means for each group. The ratio of the means is formed by dividing the mean for group 2 by the mean for group 1. The null ratio is assumed to be 1 .

## Mean ratio, Null ratio

Enter the ratio of the two group means-that is, the treatment mean divided by the reference mean. Enter the null ratio.

## Group means, Null ratio

Enter the means for each group. The ratio of the means is formed by dividing the mean for group 2 by the mean for group 1 . Enter the null ratio.

As shown in Figure 78.51, select the Mean ratio form which uses a default null ratio of 1. Enter a single mean ratio value of 0.9 .

Figure 78.51 Means Tab with Mean Ration Form and Values


## Coefficient of Variation

On the Coefficient of Variation tab, enter the coefficients of variation. They are assumed to be equal for the two groups.

For this example, enter 0.5 and 0.6 , as shown in Figure 78.52.
Figure 78.52 Coefficient of Variation Tab


## Sample Size

On the Sample Size tab, select the Group sample sizes form and enter two sets of values: 25 and 25 in the first row and 25 and 50 in the second row, as shown in Figure 78.53.

Figure 78.53 Sample Sizes


## Summary of Input Parameters

Table 78.4 contains the values of the input parameters for the example.

| Table 78.4 | Summary of Input Parameters |
| :--- | :--- |
| Parameter | Value |
| Hypothesis | One-sided test |
| Distribution | Lognormal |
| Alpha | 0.05 |
| Means form | Mean ratio |
| Mean ratio | 0.9 |
| Coefficients of variation | $0.5,0.6$ |
| Sample size form | Group sample sizes |
| Sample Size | $(25,25),(25,50)$ |

## Results

On the Results tab, select the Create summary table and Create power by sample size graph check boxes.

Click Calculate to perform the analysis.
In this case, the following message is displayed:

```
The power by sample size graph is not available when specifying sample
sizes for two groups.
```

If you want a power by sample size graph, you can choose to plot total sample size instead by using the Total N, Group weights sample size form on the Sample Size tab. For more information about using this input form, see the section "Using Unequal Group Sizes" on page 6551.

## Viewing Results

The first thing that you notice from the summary table in Figure 78.54 is that the calculated powers are quite low-they range from 0.163 to 0.229 . You have less than a $25 \%$ probability of detecting the difference that you are looking for. Clearly, this set of parameter values leads to insufficient power. To increase power, you might choose a larger sample size or a larger alpha.

Figure 78.54 Summary Table


You can also see that oversampling the control group improves power slightly, 0.229 versus 0.193 for the coefficient of variation of 0.5 . However, this is a marginal increase that is probably not worth the added expense.

For the example, use larger sample sizes with equal cell sizes. Return to the Edit Properties page by clicking the Edit Properties icon near the top of the window.

Then, on the Sample size tab, change to the Sample size per group form. Specify sample sizes of 50, 100, 150, and 200, as shown in Figure 78.55.

Figure 78.55 Modified Sample Size Values


Table 78.5 contains the modified values of the input parameters for the example.

Table 78.5 Modified Summary of Input Parameters

| Parameter | Value |
| :--- | :--- |
| Sample size form | Sample size per group |
| Sample size | $50,100,150,200$ |

Rerun the analysis by clicking Calculate.
Figure 78.56 displays the summary table. The largest sample size of 200 (per group) yields a power of 0.72 for a coefficient of variation of 0.5 , and 0.599 for one of 0.6 . With a total of 400 subjects, you still have a $30 \%$ to $40 \%$ probability of not detecting the effect even if it exists.

Figure 78.56 Summary Table for Modified Sample Sizes


## Additional Topics

## Solving for Sample Size

Several types of analysis enable you to solve for either total sample size or sample size per group. The sample size per group choice assumes equal group sizes. When solving for total sample size, the group sizes can be equal or unequal. Select the desired quantity on the Solve For tab. An example of these options is shown in Figure 78.57.

Figure 78.57 Solve For Tab with Sample Size Selected


For either of the two sample size options, you must specify one or more values for power on the Power tab. If you frequently use the same values for power, set them as the default in the Preferences window, which is accessed by Tools Preferences. Changing preferences affects only projects that you create after the change; existing projects are not affected.

If you select total sample size, you must specify whether the group sizes are equal or unequal. Select the appropriate option on the Sample Size tab. For unequal group sizes, you must specify the relative sample sizes for the two groups. For information about providing relative sample sizes, see the section "Using Unequal Group Sizes" on page 6551.

## Using Unequal Group Sizes

When solving for either power or total sample size, you might have unequal group sizes. If so, you must provide relative sample sizes for the groups. Weights must be greater than 0 but do not have to sum to 1 .

Select the Total N, Group weights form on the Sample Size tab. Enter total sample sizes of 30 and 60 in the Total $\mathbf{N}$ table. Select the Unequal group sizes option and click Enter Relative Sample Sizes, as seen in Figure 78.58.

Figure 78.58 Sample Size Tab with Group Weights Form


Figure 78.59 displays the window in which you can enter relative sample sizes. As an example, enter 2 for the first group and 1 for the second. In this case, you are sampling the drug treatment group twice as often as the control group.

Figure 78.59 Relative Sample Sizes Window


The weights control how the total sample size is divided between the two groups. In the example, the sample size for groups 1 and 2 is 20 and 10 , respectively, for a total sample size of 30 .

Click OK to save the values and return to the Edit Properties page.

## Example: Analysis of Variance

## Overview

PSS offers power and sample size calculations for analysis of variance in two tasks: one-way ANOVA and general linear univariate models. Optional contrasts are available in both tasks.

In the one-way ANOVA task, you can solve for sample size per group as well as total sample size. The contrast facility for the one-way ANOVA task enables you to select orthogonal polynomials as well as to specify contrast coefficients. For more information about power and sample size analysis for one-way ANOVA, see Chapter 77, "The POWER Procedure."

In the general linear univariate models task, you specify linear models for a single dependent variable. Type III tests and contrasts of fixed effects are included, and the model can include covariates. For more information about power and sample size analysis for linear univariate models, see Chapter 47, "The GLMPOWER Procedure."

## The Example

Suppose you are interested in testing how two experimental drugs affect systolic blood pressure relative to a standard drug. You want to include both men and women in the study. You have a two-factor design: a drug factor with three levels and a gender factor with two levels. You choose a main-effects-only model because you do not expect a drug by gender interaction. You want to calculate the sample size that will produce a power of 0.9 using a significance level of 0.05 . You believe that the error standard deviation is between 5 and 7 mm pressure. This is a two-way analysis of variance, so the general linear univariate models task is the appropriate one.

## Editing Properties

Start by opening the New window (File - New). In the Analysis of Variance and Linear models section of the New window, select General linear univariate models. The General univariate linear models project appears, with the Edit Properties page displayed.

## Project Description

For the example, change the project description to Three blood pressure drugs and gender.

## Solve For

Click the Solve For tab and select the Sample size option.

## Variables

Click the Variables tab to enter the names of the factors in the design. Click the Add button. The Factor Definition window appears, as shown in Figure 78.60.

Figure 78.60 Factor Definition Window


Enter the name for the first factor, Drug, and enter the number of factor levels in the Number of levels: list box. There are three levels for this factor. Optionally, you can provide a label for each factor level. This label is used to identify factor levels on other tabs of the Edit Properties page. For this example enter the labels Experimental 1, Experimental 2, and Standard for the three levels of the Drug factor. Click OK when you are finished.

Click the Add button again and repeat the process for the second factor, Gender with two levels and labels Female and Male.

Factors can contain blanks and other special characters. Do not use an asterisk (*) because a factor name with an asterisk might be confused with an interaction effect. Factor names can be any length, but they must be distinct from one another in the first 32 characters.

On the Variables tab, you can also specify the name of the dependent variable; in this example, Blood pressure is used.

The completed Variables tab is shown in Figure 78.61.
Figure 78.61 Variables Tab with Factors and Number of Levels


## Model

Click the Model tab, then choose from three model options:

## Main effects

Only the main effects are included in the model.

## Main effects and all interactions

The main effects and all possible interactions are included in the model.

## Custom model

Selected effects are included in the model. The effects are selected in a model builder that is displayed when this model is selected. For more information about specifying a custom model, see the section "Specifying a Custom Model" on page 6566.

For this example, choose the default Main effects model, as shown in Figure 78.62.

Figure 78.62 Model Tab with Main Effects Selected


## Alpha

Click the Alpha tab to specify one or more significance levels. For the example, specify a single significance level of 0.05 .

Alpha is the significance level (that is, the probability of falsely rejecting the null hypothesis). If you frequently use the same values for alpha, set them as the defaults in the Preferences window (Tools Preferences).

## Means

Click the Means tab to enter projected cell means for each cell of the design. The completed means for the example are shown in Figure 78.63.

Figure 78.63 Means Tab with Cell Means


## Standard Deviation

Click the Standard Deviation tab to specify one or more conjectured error standard deviations. The standard deviation is the same as the root mean squared error. For this example, enter two standard deviations, 5 and 7, as shown in Figure 78.64.

Figure 78.64 Standard Deviations Tab


## Relative Sample Size

Click the Sample Size tab to select whether cell sample sizes are equal or unequal.
Figure 78.65 Sample Size Tab with Equal Cell Sample Sizes


For the example, select the Equal cell sizes option, as shown in Figure 78.65.

When solving for sample size, it is necessary to specify whether the cell sample sizes are equal or unequal. If cell sizes are unequal, relative sample size weights must also be specified. For more information about providing sample size weights, see the section "Using Unequal Cell Sizes" on page 6563.

## Power

Click the Power tab to specify one or more powers. For this example, enter a single power of 0.9 , as shown in Figure 78.66.

Figure 78.66 Power Tab


## Summary of Input Parameters

Table 78.6 contains the values of the input parameters for the example.

Table 78.6 Summary of Input Parameters

| Parameter | Value |
| :--- | :--- |
| Model | Main effects |
| Alpha | 0.05 |
| Means | See Table 78.7 |
| Standard deviation | 5,7 |
| Relative sample sizes | Equal cell sizes |
| Power | 0.9 |

Table 78.7 Cell Means

| Gender | Experimental 1 | Drug <br> Experimental 2 | Standard |
| :--- | :---: | :---: | :---: |
| Female | 125 | 121 | 118 |
| Male | 130 | 128 | 125 |

## Results Options

Click the Results tab to select desired results. For the example, select both the Create summary table and Create power by sample size graph check boxes.

The graph consists of four points, one for each of the four scenarios that were created by combining the two factor main effects with the two standard deviations. This graph is not very informative, so specify a range of powers for the horizontal power axis. To change the power axis of the graph, click the Customize button beside the Create power by sample size graph check box to open the Customize Graph window.

Figure 78.67 Value Ranges on Customize Graph Window


Click the Value Ranges tab and enter a minimum power of 0.75 and a maximum power of 0.95 , as shown in Figure 78.67. Click OK to close the window.

Now, click Calculate to perform the analysis.

## Viewing Results

The results are displayed in separate tabs on the View Results page.
Click the Summary Table tab to view the summary table. In the Computed N Total table, sample sizes are listed for each combination of factor and standard deviation (Figure 78.68). You need a total sample size between 60 and 108 to yield a power of 0.9 for the Drug effect if the standard deviation is between 5 and 7 . You need a sample size of half that for the Gender effect.

Figure 78.68 Summary Table

| Three blood pressure drugs and gender |  |  |  |  |  |  |  | $\square$ | x |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 铟 Edit Properties | 渔 Wiew Results |  |  |  |  |  |  |  |  |
| Summary Table |  |  |  |  |  |  |  |  |  |
| Graph |  |  |  |  |  |  |  |  |  |
| Narratives |  |  |  |  |  |  |  |  |  |
| SASLog Fixed Scenario Elements |  |  |  |  |  |  |  |  |  |
| SAS Code |  |  |  |  |  |  |  |  |  |
| Dependent Variable Blood pressure |  |  |  |  |  |  |  |  |  |
| Alpha$0.05$ |  |  |  |  |  |  |  |  |  |
| Mominal Power 0. |  |  |  |  |  |  |  |  |  |
|  |  |  | Com | ted N | tal |  |  |  |  |
|  |  |  | std | Test | Error | Actual | N |  |  |
|  | Index | Source | Dev | DF | DF | Power | Total |  |  |
|  | 1 | Drug | 5 | 2 | 56 | 0.921 | 60 |  |  |
|  | 2 | Drug | 7 | 2 | 104 | 0.905 | 108 |  |  |
|  | 3 | Gender | 5 | 1 | 26 | 0.916 | 30 |  |  |
|  | 4 | Gender | 7 | 1 | 50 | 0.903 | 54 |  |  |

Click the Graph tab to view the power by sample size graph, as shown in Figure 78.69. One approximately linear curve is displayed for each standard deviation and factor combination.

Figure 78.69 Power by Sample Size Graph


Click the Narratives tab to create narratives of one or more scenarios. Select the first scenario, the Drug effect with the standard deviation of 5, in the narrative selector table. Note that the cell means are not included in the following narrative description:

```
For the usual F test of the Drug effect in the general linear univariate
model with fixed class effects [Blood pressure = Drug Gender] using a
significance level of 0.05, assuming the specified cell means and an
error standard deviation of 5, a total sample size of 60 assuming a
balanced design is required to obtain a power of at least 0.9. The
actual power is 0.921.
```

For more information about using the narrative facility, see the section "Creating Narratives" on page 6525 .

## Additional Topics

## Adding Contrasts

Click the Contrasts tab to define one or more contrasts. Contrasts are optional. PSS allows contrasts to be added when using either a main effects model or a main effects and interactions model. At least two factors must have been specified in order to be able to enter contrasts. The contrast tab appears in Figure 78.70.

Figure 78.70 Contrast Tab with Coefficients


To create a contrast, click the New button. Then, select the newly created contrast (Contrast 1) from the list.

Specify a label for the contrast in the Label field. The label should be different from all of the factor names and all interactions in the model, as well as other contrast labels.

Then, for each term you want to include in the contrast, select the term in the Effects list and enter at least two coefficients per term. It is not necessary to enter zeros; blanks are considered to be zeros.

To clear all of the contrast coefficients for a term, click the Clear button. To remove a previously defined contrast, select it from the Contrasts list and click the Remove button.

In this example, you are interested in comparing the two experimental drugs to the standard drug. As shown in Figure 78.70, the contrast coefficients are $0.5,0.5$, and -1 for the three levels of the Drug effect.

Figure 78.71 shows the two scenarios for the contrast at the bottom of the Computed N Total table. The two scenarios also appear in the graph but the graph is not shown here.

Figure 78.71 Computed N Total Table for the Contrast

| 國 Edk Prope | 國Vew Resuls |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Summary Table |  |  |  |  |  |  |  |  |  |
| Graph |  |  |  |  |  |  |  |  | 1 |
| Naratives |  |  |  |  |  |  |  |  |  |
| SASLog |  |  | Fixed Scenario Element |  |  |  |  |  |  |
| SAS Code |  |  |  |  |  |  |  |  |  |
|  |  |  | Dependent Variable Blood | ressu |  |  |  |  |  |
|  |  |  | Alpha |  |  |  |  |  |  |
|  |  |  | Nominal Power |  |  |  |  |  |  |
|  |  |  | Computed N Total |  |  |  |  |  |  |
|  |  |  |  | std | Test | Error | Actual | N |  |
|  | Index | Type | Source | Dev | DF | DF | Power | Total |  |
|  | 1 | Effect | Drug | 5 | 2 | 56 | 0.921 | 60 |  |
|  | 2 | Effect | Drug | 7 | 2 | 104 | 0.905 | 108 |  |
|  | 3 | Effect | Gender | 5 | 1 | 26 | 0.916 | 30 |  |
|  | 4 | Effect | Gender | 7 | 1 | 50 | 0.903 | 54 |  |
|  | 5 | Contrast | Experimental drugs versus standard | 5 | 1 | 62 | 0.924 | 66 |  |
|  | 6 | Contrast | Experimental drugs versus standard | 7 | 1 | 116 | 0.909 | 120 |  |

## Using Unequal Cell Sizes

Click the Sample Size tab to select the equal or unequal cell sizes option.
Figure 78.72 Sample Size Tab


For the example, select the Unequal cell sizes option, as seen in Figure 78.72, and then click the Enter Relative Sample Sizes button.

Figure 78.73 shows the window in which you can enter relative sample sizes. As an example, enter the sample size weights from Table 78.8.

Table 78.8 Sample Size Weights

| Gender | Experimental 1 | Drug <br> Experimental 2 | Standard |
| :--- | :---: | :---: | :---: |
| Males | 1 | 1 | 2 |
| Females | 1 | 1 | 2 |

If you have unequal cell sizes, you must enter relative sample size weights for the cells. Weights do not have to sum to 1 across the cells. Some weights can be zero, but enough weights must be greater than zero so that the effects and contrasts are estimable.

In this case, you want the sample size of the standard group to be twice that of each of the two experimental groups. Click OK to save the values and return to the Edit Properties page.

Figure 78.73 Relative Sample Sizes Window


Figure 78.74 shows the summary table for the Drug by Gender example.
Figure 78.74 Summary Table for Unbalanced Design Example


## Solving for Power

In addition to solving for sample size, you can also solve for power. Figure 78.75 shows the two options. Click the Solve For tab to select the Power option.

Figure 78.75 Solve For Tab with Power Option Selected


When solving for power, you must provide sample size information. For the general linear univariate model analysis, you provide this information by using one of two alternate forms. To choose the desired alternate form, select the desired form from the Select a form list box on the Sample Size tab. The alternate forms are:

## Sample size per cell

Enter the sample size for a cell. Cell sizes are assumed to be equal. Sample size is reported in the summary table as total sample size.
Total N, Cell weights
Enter the total sample size and specify whether cell sizes are to be equal or unequal. Select the Equal cell sizes or Unequal cell sizes option. For unequal cell sizes, you also enter cell weights. Click the Enter Relative Sample Sizes button to display a window that is used to enter the data. For more information about using unequal cell sizes, see the section "Using Unequal Cell Sizes" on page 6563.

## Specifying a Custom Model

Click the Model tab to select from three types of models: a Main effects model, a Main effects and all interactions model, and a Custom model.

To specify a custom model, select the Custom model option; then a model building facility is displayed.
The facility displays a list of the factors on the left. Construct the desired model using the Add, Cross, and Factorial buttons. The example shown in Figure 78.76 has the three main effects and one of the four possible interactions.

Figure 78.76 Model Tab with Custom Model Builder Displayed


Add the three main effects (A, B, C) by selecting them in the Terms list and clicking the Add button. Add the $A * B$ interaction by selecting the $A$ and $B$ factors in the Terms list and clicking the Cross button.

To create the complete factorial design of several factors, select the factors in the Terms list, then click the Factorial button. All possible main effects and interactions are added to the Model Effects list.

To remove effects, select them in the Model Effects list and click the Remove button. Clicking the Remove All button removes all effects in the model.

## Including Covariates

Click the Covariates tab to enter covariate information.
Figure 78.77 Covariates Tab with Proportional Reduction in Variance Form


Figure 78.77 illustrates four covariates and a proportional reduction in variation of 0.3 . The results for the analysis are not shown.

Covariates are optional. If you have covariates, include the total number of degrees of freedom for all covariates. To do this, add the number of continuous covariates and the sum of the degrees of freedom of the classification covariates, and enter this total in the Number of Covariates list box. For example, with two continuous covariates and a single classification covariate factor with three levels, the total would be $2+(3-1)=4$.

Also, you must enter the correlation between the dependent variable and the set of covariates. Two alternate forms are available: Multiple correlation and Proportional reduction in variance. Select the desired form and enter one or more values.

The multiple correlation is between the set of covariates and the dependent variable. Proportional reduction in variation is how much the variance of the dependent variable is reduced by the inclusion of the covariates, expressed as a proportion between 0 and 1 .

## Example: Two-Sample Survival Rank Tests

## Overview

Survival analysis often involves the comparison of survival curves. PSS provides sample size and power calculations for two-sample survival rank analyses. Several rank tests are available: Gehan, log-rank, and Tarone-Ware. There are also several ways to specify the survival functions. For more information about power and sample size analysis for survival rank tests, see Chapter 77, "The POWER Procedure."

## The Example

Suppose you want to compare survival rates for an existing cancer treatment and a new treatment. You intend to use a log-rank test to compare the overall survival curves for the two treatments. You want to determine a sample size to achieve a power of 0.8 for a two-sided test using a balanced design, with a significance level of 0.05 .

The survival curve of patients for the existing treatment is known to be approximately exponential with a median survival time of five years. You think that the proposed treatment will yield a survival curve described by the times and probabilities listed in Table 78.9. Patients are to be accrued uniformly over two years and followed for three years.

Table 78.9 Survival Probabilities for Proposed Treatment

| Time | Probability |
| :---: | :---: |
| 1 | 0.95 |
| 2 | 0.90 |
| 3 | 0.75 |
| 4 | 0.70 |
| 5 | 0.60 |

To create a new survival analysis project, select File New, Then, under the Survival Analysis section, select Two-sample survival rank tests and click OK. The Two-sample survival rank tests project appears with the Edit Properties page displayed.

## Editing Properties

## Project Description

For the example, change the project description to Comparing cancer treatments using two-sample survival rank test.

Figure 78.78 Project Description and Solve For Tab


## Solve For

Click the Solve For tab to select the quantity to solve for. For this example, select the Sample size per group option, as shown in Figure 78.78. For information about calculating total sample size, see the section "Solving for Sample Size" on page 6550.

In this analysis you can solve for power, sample size per group, or total sample size.

## Test

Click the Test tab to select a rank test. For this example, select the Log-rank option, as shown in Figure 78.79.

Figure 78.79 Test Tab


Several rank tests are available: Gehan, log-rank, and Tarone-Ware. The Gehan test is most sensitive to survival differences near the beginning of the study period, the log-rank test is uniformly sensitive throughout the study period, and the Tarone-Ware test is somewhere in between.

## Hypothesis

Click the Hypothesis tab to select a one- or two-sided test. For the example, select the Two-sided test option, as shown in Figure 78.80.

Figure 78.80 Hypothesis Tab


You can choose either a one- or two-sided test. For the one-sided test, the alternative hypothesis is assumed to be in the same direction as the effect. If you do not know the direction of the effect (that is, whether it is positive or negative), the two-sided test is appropriate. If you know the effect's direction, the one-sided test is appropriate. If you specify a one-sided test and the effect is in the unexpected direction, the results of the analysis are invalid.

## Alpha

Click the Alpha tab to enter one or more values for the significance level. For the example, enter the desired significance level of 0.05 in the first cell of the Alpha table, as shown in Figure 78.81, if it is not already the default value.

Figure 78.81 Alpha Tab


The significance level is the probability of falsely rejecting the null hypothesis. If you frequently use the same values for alpha, set them as the defaults in the Preferences window.

## Survival Functions

Click the Survival Functions tab to select the input form for the survival functions.

Figure 78.82 Survival Functions Tab with Number of Curves


Examine the input alternatives available in the Select a form list. There are four alternate forms for entering survival functions. The first three apply only to exponential curves; the fourth applies to both piecewise linear and exponential curves.

## Group median survival times

Enter median survival times for the two groups.

## Group hazards

Enter hazards for the two groups.

## Hazards, Hazard ratios

Enter hazards for the reference group and hazard ratios.

## Survival curves

Enter survival probabilities and their associated times for each of several curves. Select or enter the number of curves from the drop-down list; at least two curves are required. Then, for each curve, select it in the left-hand list, select the Group 1 or Group 2 option, and then define the survival curve by entering pairs of times and probabilities. Enter a time and probability pair only if the probability is less than that of the previous pair.

For information about using the other forms, see the section "Using the Other Survival Curve Forms" on page 6582.

For each survival curve, select the curve in the left-hand list. Then, enter a descriptive label and select which group it is for. The labels should be unique. Finally, enter pairs of survival times and probabilities.

When you enter probabilities, enter a time and probability pair only when the probability for a survival curve changes. For example, if the probability for curve 1 at time 1 and 2 is 0.9 and at time 3 is 0.8 , enter 0.9 for time 1 and 0.8 for time 3.

To specify an exponential survival curve, enter a single time and probability pair. In the example, the exponential curve for the existing treatment is defined by a probability of 0.5 at time 5 .

The units of time for the survival curves must correspond to the units for the accrual, follow-up, and total times, which are described in the section "Accrual Times" on page 6575.

You can also compare several survival curves. For example, if you have two scenarios, A and B, for group 1's curve and two scenarios, C and D, for group 2's curve, then specify probabilities for the four curves and assign A and B to group 1 and C and D to group 2.

For the example, select the Survival curves form, as shown in Figure 78.82. Enter the value, 2, in the Number of survival curves list box.

For the example enter the following values:

- For the first survival curve, enter a label of Existing treatment and select the Group 1 option. For the first curve, enter a time of 5 and a probability of 0.5 . Figure 78.83 shows the resulting values.

Figure 78.83 Survival Times and Probabilities for Curve 1


- For the second curve select Function 2 in the selection list on the left side of the tab. Enter a label of Proposed treatment and select the Group 2 option. Then, enter time values of 1 through 5 and the corresponding probabilities of $0.95,0.9,0.75,0.7$, and 0.6 . To add rows to the table, click the ${ }^{-}$ button beneath the table.

Figure 78.84 shows these values; the last row of the time and probability table is not displayed.
Figure 78.84 Survival Times and Probabilities for Curve 2


## Accrual Times

Click the Accrual times tab to select an input form for accrual times and to enter the times.

Figure 78.85 Accrual Times Tab


Examine the alternatives available in the Select a form list.
Accrual time is the period during which subjects are brought into the study. Follow-up time is the period during which subjects are observed after all subjects have been included in the study. Total time is the sum of accrual and follow-up time. The units of time for the accrual, follow-up, and total times must correspond to the units you used specified for the survival curves.

When you enter survival curves, the sum of the accrual and follow-up times must be less than the largest time for each survival curve. This does not apply to survival curves represented by a single time, which represent exponential curves.

On the Accrual Times tab, there are three alternate forms for entering accrual and follow-up times:

## Accrual times, Follow-up times

Enter accrual and follow-up times.

## Accrual times, Total times

Enter accrual and total times.
Follow-up times, Total times
Enter follow-up and total times.
For the example, select the Accrual times, Follow-up times form. Then enter a single value of 2 in the Accrual table and a value of 3 in the Follow-up table, as shown in Figure 78.85.

## Power

Click the Power tab to enter one or more power values. For the example, enter a single value of 0.8.
When you calculate sample size, it is necessary to specify one or more powers.

## Summary of Input Parameters

Table 78.10 contains the values of the input parameters for the example.

Table 78.10 Summary of Input Parameters

| Parameter | Value |
| :--- | :--- |
| Solve for | Sample size per group |
| Test | Log-rank |
| Hypothesis | Two-sided test |
| Alpha | 0.05 |
| Survival function form | Survival curves |
| Survival curves | See Table 78.11 and Table 78.12 |
| Accrual and follow-up times form | Accrual time, Follow-up times |
| Accrual times | 2 |
| Follow-up times | 3 |
| Power | 0.8 |

Table 78.11 and Table 78.12 contain times and probabilities for the two survival curves, respectively.

Table 78.11 Survival Times and Probabilities for Existing Treatment (Survival Curve 1)

| Time | Probability |
| :---: | :---: |
| 5 | 0.5 |

Table 78.12 Survival Times and Probabilities for Proposed Treatment (Survival Curve 2)

| Time | Probability |
| :---: | :---: |
| 1 | 0.95 |
| 2 | 0.90 |
| 3 | 0.75 |
| 4 | 0.70 |
| 5 | 0.60 |

## Result Options

Click the Results tab to specify the desired result options. For the example, request both results by selecting both the Create summary table and Create power by sample size graph check boxes.

Specifying only one power (as in this example) produces a graph with a single point. You might be interested in a plot of sample sizes for a range of powers-say, between 0.75 and 0.85 . You can customize the graph by specifying the values for the power axis. Also, you might want to change the appearance of the graph to have sample size (per group) on the vertical axis and power on the horizontal axis.

Click the Customize button beside the Create power by sample size graph check box to customize the graph. The Customize Graph window is displayed, as shown in Figure 78.86.

Figure 78.86 Customize Graph Window with Axis Orientation Tab


Click the Axis Orientation tab to select which variable to plot on the vertical axis. For the example, select the Quantity solved for option, as shown in Figure 78.86. This option plots sample size on the vertical axis and power on the horizontal axis. You could also have chosen the Sample size option.

Click the Value Ranges tab to enter minimum and maximum values for a plot axis. For the example, enter a minimum of 0.75 and a maximum of 0.85 in the Powers text boxes. This sets the range of values on the axis for powers. The completed Value Ranges tab of the window is displayed in Figure 78.87. You can set the axis values only for the quantity that is not being solved for.

Figure 78.87 Customize Graph Window with Value Ranges Tab


Click OK to save the values that you have entered and return to the Edit Properties page.
Then, click Calculate to perform the analysis. If there are no errors in the input parameter values, the View Results page appears. If there are errors in the input parameter values, you are prompted to correct them.

## Viewing Results

The results appear in separate tabs on the View Results page of the project. Select the tab of each result that you want to view.

## Summary Table

Click the Summary Table tab to view the summary table. It is composed of two subtables. As shown in Figure 78.88, the Fixed Scenario Elements and Computed N Per Group tables include the values of the input parameters and the computed quantity (in this case, sample size per group, N per group). The sample size per group for the single requested scenario is 226 .

Figure 78.88 Summary Table


## Power by Sample Size Graph

Click the Graph tab to view the power by sample size graph.

Figure 78.89 Power by Sample Size Graph


As you can see in Figure 78.89, the graph is curved slightly upward with larger powers associated with larger sample sizes. Sample size is plotted on the vertical axis as requested in the Customize Graph window.

## Narratives

Click the Narratives tab to create one or more narratives. To generate a narrative, select the single scenario in the narrative selector table at the bottom of the tab. The narrative for this task does not include the survival times and probabilities for the survival curves:

For a log-rank test comparing two survival curves with a two-sided significance level of 0.05, assuming uniform accrual with an accrual time of 2 and a follow-up time of 3 , a sample size of 226 per group is required to obtain a power of at least 0.8 for the exponential curve, "Existing treatment," and the piecewise linear curve, "Proposed treatment." The actual power is 0.800 .

For information about selecting additional narratives when multiple scenarios are present, see the section "Creating Narratives" on page 6525.

## Additional Topics

## Using the Other Survival Curve Forms

Survival functions can be specified as median survival times, hazards, or a combination of hazards for one group and hazard ratios. These all assume exponential curves.
Suppose you are interested in comparing the proposed and existing treatments using their median survival times. The survival times are five years and four years for the two groups, respectively.

Figure 78.90 Median Survival Times and List of Alternate Forms


Click the Survival Functions tab and examine the list of alternate forms available in the Select a form: list. For this example, select the Group median survival times option.
For the example, enter 5 and 4 in the first row of the table. The completed table is shown in Figure 78.90. You can enter one or more sets of two median survival times. The results of the analysis are not shown.

