Chapter 1
Overview

Chapter Contents

Chapter 1
Overview

Introduction

When you are planning a study or an experiment, you often need to know how many units to sample to obtain a certain power, or you may want to know the power you would obtain with a specific sample size. The power of a hypothesis test is the probability of rejecting the null hypothesis when the alternative hypothesis is true. With an inadequate sample size, you may not reach valid conclusions with your work; with an excessive sample size, you may waste valuable resources. Thus, performing sample size and power computations is often quite important.

The power and sample size calculations depend on the planned data analysis strategy. That is, if the primary hypothesis test is a two-sample *t* test, then the power calculations must be based on that test. Otherwise, if the sample size calculations and data analyses are not aligned, the results may not be correct.

The SAS/STAT Power and Sample Size application (PSS) is a data analysis tool 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.

Determining sample size requirements ahead of the study is a prospective exercise. You then proceed to select the appropriate number of sampling units and perform data collection and analysis. However, power and sample size calculations are also useful retrospectively. Only prospective power analysis is offered by PSS.

Power and sample size calculations are a function of the specific alternative hypothesis of interest, in addition to other parameters. That is, the power results will vary depending on which value of the alternative hypothesis you specify, so sometimes it is useful to do these analyses for a range of values to determine how sensitive the power analysis is to changes in the alternative

hypothesis value. Often, you produce plots of power versus sample size, called power curves, to see how sample size and power affect each other.

The PSS Application

The PSS application is a Web browser application. It is accessed either from your own machine or your organization's intranet using the Microsoft Internet Explorer browser. It relies on the SAS/STAT procedures POWER and GLMPOWER for its computations.
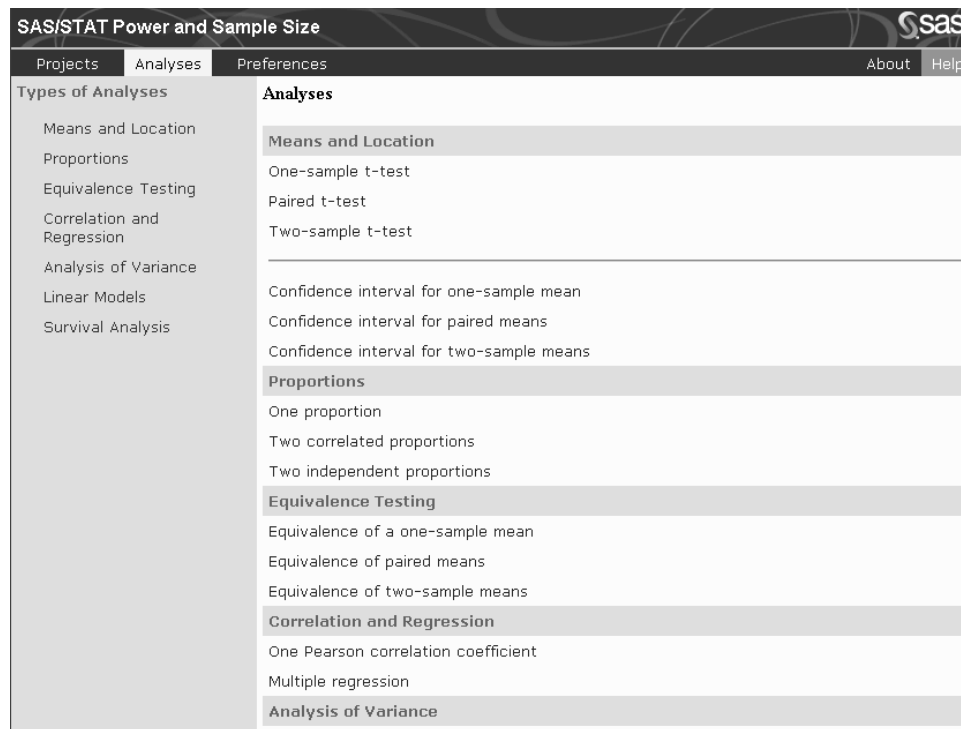

Figure 1.1. PSS Application

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, equivalence tests, and confidence intervals for means; exact binomial, chi-square, Fisher's exact, and McNemar tests for proportions; multiple regression and correlation; one-way analysis of variance; linear models, and rank tests for comparing survival curves.

Table 1.1 lists the analyses that are available.

Table 1.1.   Available Analyses

Category	Analysis
Means and location	One-sample *t* test
	Paired *t* test
	Two-sample *t* test
Confidence intervals	One-sample means
	Paired means
	Two-sample means
Equivalence tests	One-sample means
	Paired means
	Two-sample means
Proportions	One proportion
	Two correlated proportions
	Two independent proportions
Correlation and regression	Pearson correlation coefficient
	Multiple regression
Analysis of variance	One-way ANOVA
Linear models	General linear univariate models
Survival analysis	Two-sample survival rank tests

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 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.