Quick Start Manual

PASS

Power Analysis and Sample Size System

Published by
NCSS
Dr. Jerry L. Hintze
Kaysville, Utah

PASS Quick Start Manual

Direct inquiries to:

NCSS
329 North 1000 East
Kaysville, Utah 84037
Phone (801) 546-0445
Fax (801) 546-3907
Email: support@ncss.com

About This Manual

Congratulations on your purchase of the ***PASS*** package! ***PASS*** offers:

- Easy parameter entry.
- A comprehensive list of power analysis routines that are accurate and verified, yet are quick and easy to learn and use.
- Straightforward procedures for creating paper printouts and file copies of both the numerical and graphical reports.

Our goal is that with the help of this quick start manual, you will be up and running on ***PASS*** in less than one hour. **You should read the first nine chapters immediately**. After that, you will only need to refer to the chapters corresponding to the procedures you want to use. The discussion of each procedure includes one or more tutorials that will take you step-by-step through the tasks necessary to run the procedure.

I believe you will find that these user's guides provides a quick, easy, efficient, and effective way for first-time ***PASS*** users to get up and running.

I look forward to any suggestions you have to improve the usefulness of this manual and/or the ***PASS*** system. Meanwhile, good computing!

Jerry Hintze, Author

PASS License Agreement

Important: The enclosed Power Analysis and Sample Size software program (PASS) is licensed by NCSS to customers for their use only on the terms set forth below. Your purchase and use of the PASS system indicates your acceptance of these terms.

1. **LICENSE.** NCSS hereby agrees to grant you a non-exclusive license to use the accompanying PASS program subject to the terms and restrictions set forth in this License Agreement.

2. **COPYRIGHT.** PASS and its documentation are copyrighted. You may not copy or otherwise reproduce any part of PASS or its documentation, except that you may load PASS into a computer as an essential step in executing it on the computer and make backup copies for your use on the same computer.

3. **BACKUP POLICY.** PASS may be backed up by you for your use on the same machine for which PASS was purchased.

4. **RESTRICTIONS ON USE AND TRANSFER.** The original and any backup copies of PASS and its documentation are to be used only in connection with a single user. This user may load PASS onto several machines for his/her convenience (such as a desktop and laptop computer), but only for use by the licensee. You may physically transfer PASS from one computer to another, provided that PASS is used in connection with only one user. You may not distribute copies of PASS or its documentation to others. You may transfer this license together with the original and all backup copies of PASS and its documentation, provided that the transferee agrees to be bound by the terms of this License Agreement. PASS licenses may not be transferred more frequently than once in twelve months. Neither PASS nor its documentation may be modified or translated without written permission from NCSS.

*You may not use, copy, modify, or transfer **PASS**, or any copy, modification, or merged portion, in whole or in part, except as expressly provided for in this license.*

5. **NO WARRANTY OF PERFORMANCE.** NCSS does not and cannot warrant the performance or results that may be obtained by using PASS. Accordingly, PASS and its documentation are licensed "as is" without warranty as to their performance, merchantability, or fitness for any particular purpose. The entire risk as to the results and performance of PASS is assumed by you. Should PASS prove defective, you (and not NCSS nor its dealer) assume the entire cost of all necessary servicing, repair, or correction.

6. **LIMITED WARRANTY ON CD.** To the original licensee only, NCSS warrants the medium on which PASS is recorded to be free from defects in materials and faulty workmanship under normal use and service for a period of ninety days from the date PASS is delivered. If, during this ninety-day period, a defect in a CD should occur, the CD may be returned to NCSS at its address, or to the dealer from which PASS was purchased, and NCSS will replace the CD without charge to you, provided that you have sent a copy of your receipt for PASS. Your sole and exclusive remedy in the event of a defect is expressly limited to the replacement of the CD as provided above.

Any implied warranties of merchantability and fitness for a particular purpose are limited in duration to a period of ninety (90) days from the date of delivery. If the failure of a CD has resulted from accident, abuse, or misapplication of the CD, NCSS shall have no responsibility to replace the CD under the terms of this limited warranty. This limited warranty gives you specific legal rights, and you may also have other rights which vary from state to state.

7. **LIMITATION OF LIABILITY.** Neither NCSS nor anyone else who has been involved in the creation, production, or delivery of PASS shall be liable for any direct, incidental, or consequential damages, such as, but not limited to, loss of anticipated profits or benefits, resulting from the use of PASS or arising out of any breach of any warranty. Some states do not allow the exclusion or limitation of direct, incidental, or consequential damages, so the above limitation may not apply to you.

8. **TERM.** The license is effective until terminated. You may terminate it at any time by destroying PASS and documentation together with all copies, modifications, and merged portions in any form. It will also terminate if you fail to comply with any term or condition of this License Agreement. You agree upon such termination to destroy PASS and documentation together with all copies, modifications, and merged portions in any form.

9. **YOUR USE OF PASS ACKNOWLEDGES** that you have read this customer license agreement and agree to its terms. You further agree that the license agreement is the complete and exclusive statement of the agreement between us and supersedes any proposal or prior agreement, oral or written, and any other communications between us relating to the subject matter of this agreement.

Dr. Jerry L. Hintze & **NCSS**, Kaysville, Utah

Preface

PASS (**P**ower **A**nalysis and **S**ample **S**ize) is an advanced, easy-to-use statistical analysis software package. The system was designed and written by Dr. Jerry L. Hintze over the last Seventeen years. Dr. Hintze drew upon his experience both in teaching statistics at the university level and in various types of statistical consulting.

The present version, written for 32-bit versions of Microsoft Windows (Vista, XP, NT, ME, 2000, 98, etc.) computer systems, is the result of several iterations. Experience over the years with several different types of users has helped the program evolve into its present form.

NCSS maintains a website at www.ncss.com where we make the latest edition of ***PASS*** available for free downloading. The software is password protected, so only users with valid serial numbers may use this downloaded edition. We hope that you will download the latest edition routinely and thus avoid any bugs that have been corrected since you purchased your copy.

We believe ***PASS*** to be an accurate, exciting, easy-to-use program. If you find any portion which you feel needs to be changed, please let us know. Also, we openly welcome suggestions for additions and enhancements.

Verification

All calculations used in this program have been extensively tested and verified. First, they have been verified against the original journal article or textbook that contained the formulas. Second, they have been verified against second and third sources when these exist.

Table of Contents

Installation

Running PASS

The PASS Home Window

The Procedure Window

The Output Window

The Map (Quick Launch) Window

Introduction to Power Analysis

Proportions

Means

Chapter 1

Installation

Before You Install

1. Check System Requirements

PASS runs on 32-bit Windows systems. These include Windows Vista, Windows XP, Windows 2000, Windows NT 4.0, Windows ME, and Windows 98. The recommended minimum system is a Pentium PC with at least 64 MB of memory.

PASS takes up about 80 MB of disk space. Once installed, ***PASS*** also requires about 20 MB of temporary disk space while it is running.

2. Find a Home for PASS

Before you start installing, decide on a folder where you want to install ***PASS***. By default, the setup program will install ***PASS*** application files in *C:\Program Files\NCSS\PASS 2008*. You may change this during the installation, but not after. The example data, template, and macro files will be placed in your personal documents folder (usually *C:\...\[My] Documents\NCSS\PASS 2008*) in appropriate subdirectories. The program will save all procedure templates and macros to these folders while the program is running.

What Install Does

The installation procedure creates the necessary folders and copies the ***PASS*** program from the installation file, called *PASS2008SETUP.EXE*, to those folders. The files in *PASS2008SETUP.EXE* are compressed, so the installation program decompresses these files as it copies them to your hard disk.

The following folders are created during installation (assuming defaults are chosen during installation):

Folder	Contents
C:\Program Files\NCSS\PASS 2008	Contains most of the program files including the ***PASS*** executable file, *PASS2008.exe*, and the ***PASS*** Help System file, *PASS Help.exe*.
C:\Program Files\NCSS\PASS 2008\Icons	Contains some program icons.
C:\Program Files\NCSS\PASS 2008\Pdf	Contains printable copies of the documentation in PDF format.
C:\Program Files\NCSS\PASS 2008\Sts	Contains all labels, text, and online messages.

C:\…\[My] Documents\NCSS\PASS 2008\Data	Contains the database files used by some of the tutorials. An empty subfolder called "My Data" is created within this folder for easy storage of your personal data files. You can save the data to any folder you wish.
C:\…\[My] Documents\NCSS\PASS 2008\Junk	Contains temporary files used by the program while it is running. Under normal operation, ***PASS*** will automatically delete temporary files. After closing ***PASS***, you can delete any files left in this folder (but do not delete the folder itself).
C:\…\[My] Documents\NCSS\PASS 2008\Macros	Contains saved macros.
C:\…\[My] Documents\NCSS\PASS 2008\Report	The default folder in which to save your output. You can save the reports to any folder you wish.
C:\…\[My] Documents\NCSS\PASS 2008\Settings	Contains the files used to store your procedure templates. These files are used by the ***PASS*** template system, which is described in a later chapter.

Installing PASS

This section gives instructions for installing ***PASS*** on your computer system. You must use the ***PASS*** setup program to install ***PASS***. The files are compressed, so you cannot simply copy the files to your hard drive.

Follow these basic steps to install ***PASS*** on your computer system:

1. Make sure that you are using a 32-bit or 64-bit version of Windows such as Windows Vista, Windows XP, Windows 2000, Windows NT 4.0, Windows ME, and Windows 98.
2. If you are installing from a CD, insert the CD in the CD drive. The installation program should start automatically. If it does not, on the Start menu, select the Run command. Enter *D:\NCSS\PASS2008Setup*. You may have to substitute the appropriate letter for your CD drive if it is not *D*. If you are installing from a download, simply run the downloaded file (*PASS2008SETUP.exe*).
3. Once the setup starts, follow the instructions on the screen. ***PASS*** will be installed to the drive and folder you designate.

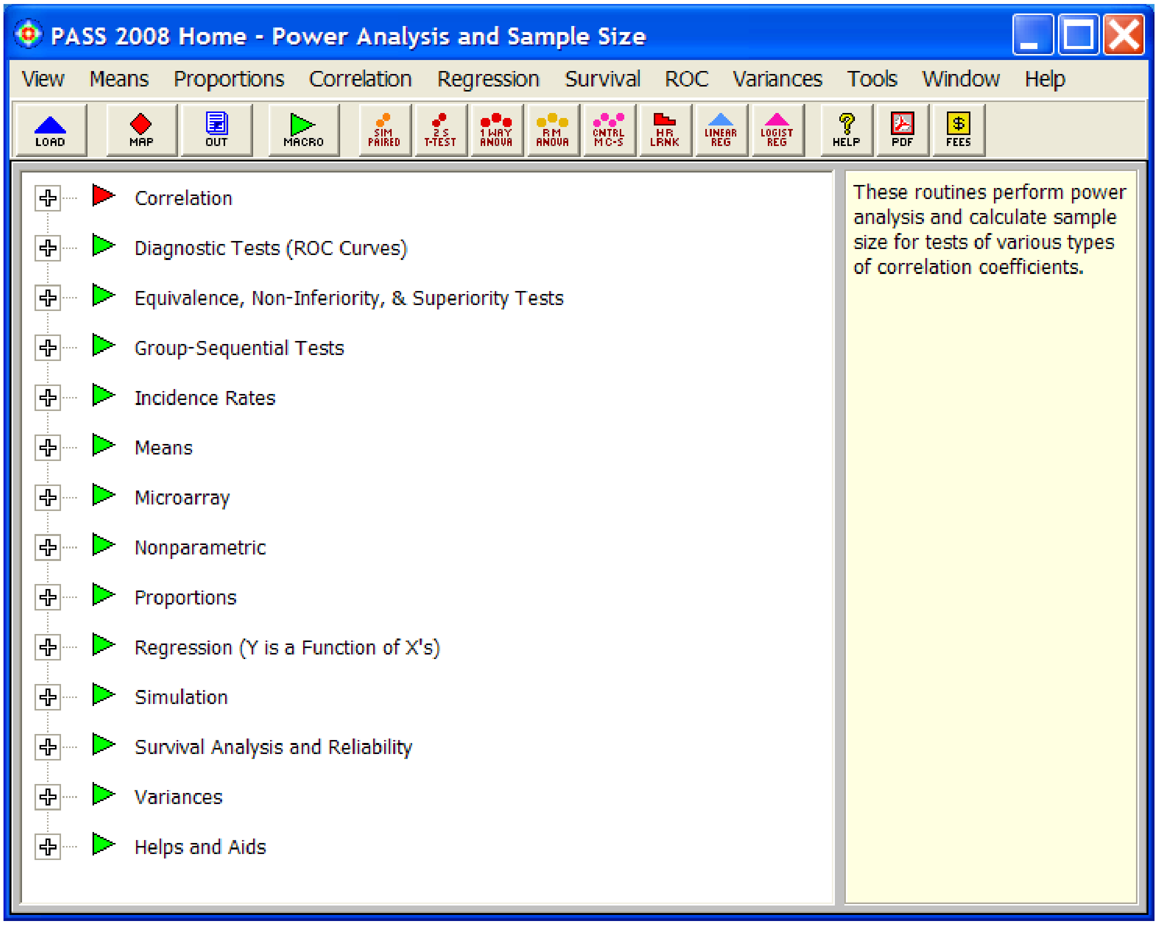
If Something Goes Wrong during Installation

The installation procedure is automatic. If something goes wrong during installation, delete the *C:\Program Files\NCSS\PASS 2008* directory and start the installation process at the beginning. If trouble persists, contact our technical support staff as indicated below.

Starting PASS

PASS may be started using your keyboard or your mouse using the same techniques that you use to start any other Windows application. You can start ***PASS*** by selecting **PASS 2008** from your Start menu using standard mouse or keyboard operations.

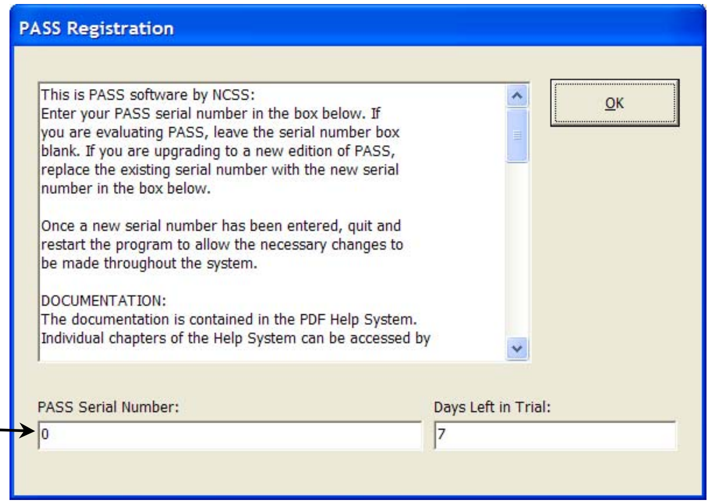
The first time you run ***PASS***, enter your serial number in the pop-up window that appears when the program begins. After entering a serial number, the **PASS Home** window will appear.



Entering Your Serial Number

The first time you run ***PASS***, enter your serial number in the pop-up window that appears. If you do not enter a serial number, ***PASS*** will enter trial mode and you will have 7 days to evaluate ***PASS***. When in trial mode, ***PASS*** is fully-functional but the spreadsheet is limited to 100 rows of data.

Enter your ***PASS*** serial number here.

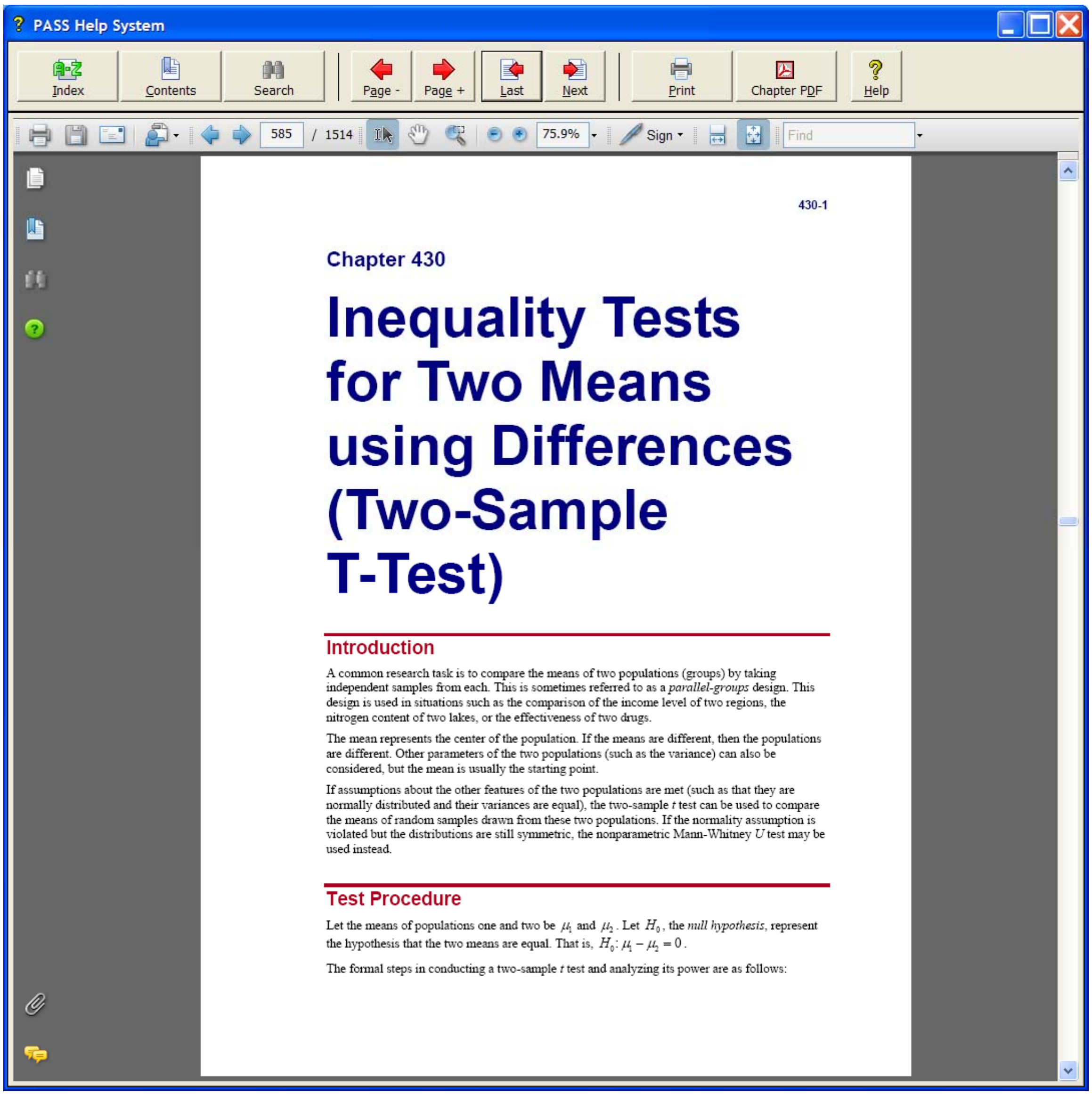


When you click **OK**, the **PASS Home** window will appear.

Obtaining Help

The PASS Help System

To help you learn and use ***PASS*** efficiently, the material in this manual is included in the ***PASS*** Help System. The Help System is started from the Help menu or by clicking on the yellow "?" icon on the right side of the toolbar. ***PASS*** updates, available for download at www.ncss.com, may contain adjustments or improvements of the ***PASS*** Help System. Adobe Acrobat or Adobe Reader version 7 or 8 is required to view the help system. You can download Adobe Reader 8 for free by going to www.adobe.com. Adobe Reader 8 can also be installed from the *Utilities* folder on your ***PASS*** installation CD.

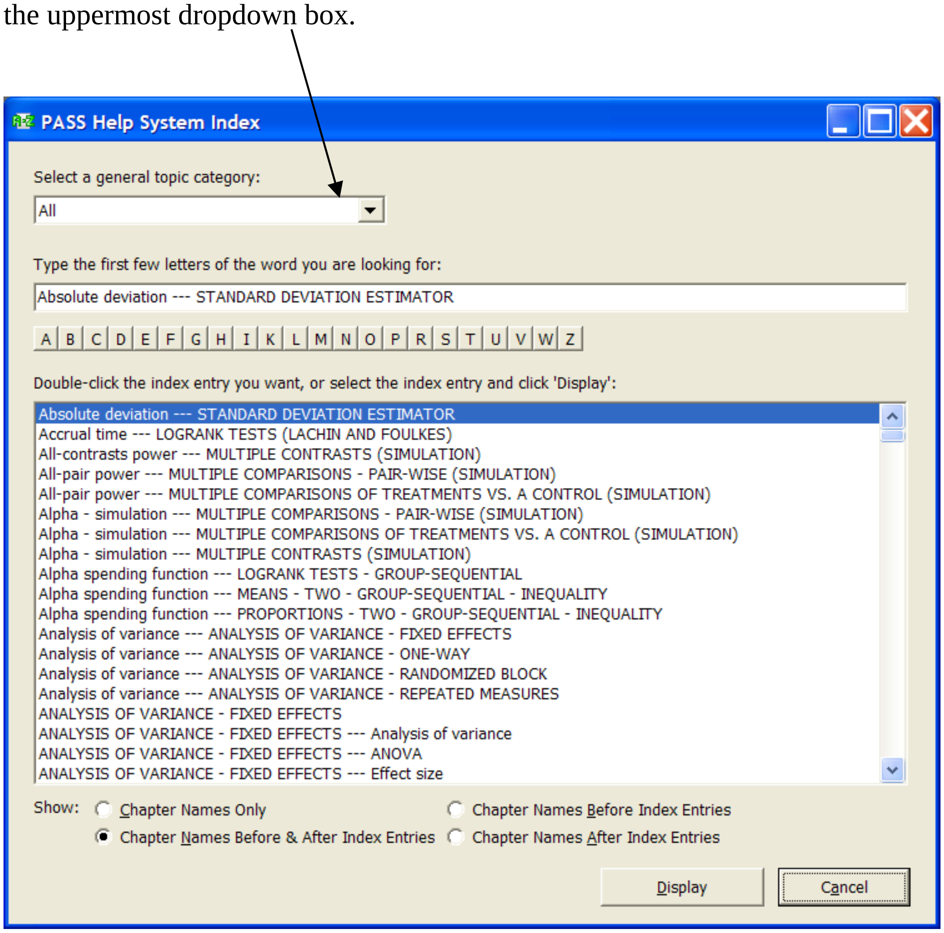


Navigating the Help System

There are a few key features of our help system that will let you use the help system more efficiently. We will now explain each of these features.

Index Window

The Index Window can be launched at any time by clicking on the Index button on the ***PASS*** Help System display window. The index allows you to quickly locate keywords and/or statistical topics. You can narrow the list of index entries displayed by selecting a specific topic category in the uppermost dropdown box.



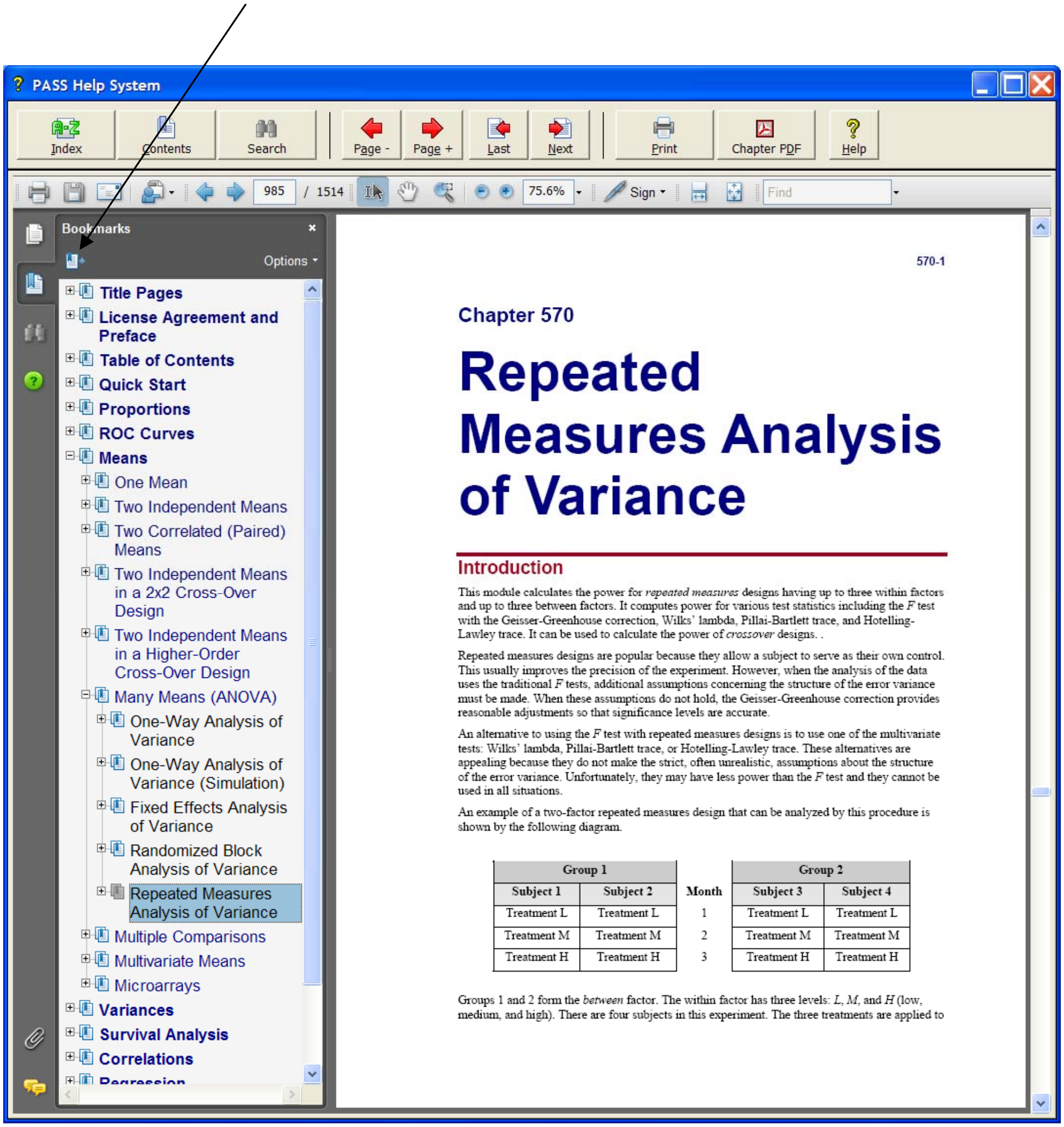
Index entries are displayed in the format

Index Entry --- CHAPTER   or   CHAPTER --- Index Entry.

You can control which entries are displayed by clicking on the radio buttons at the bottom of the window.

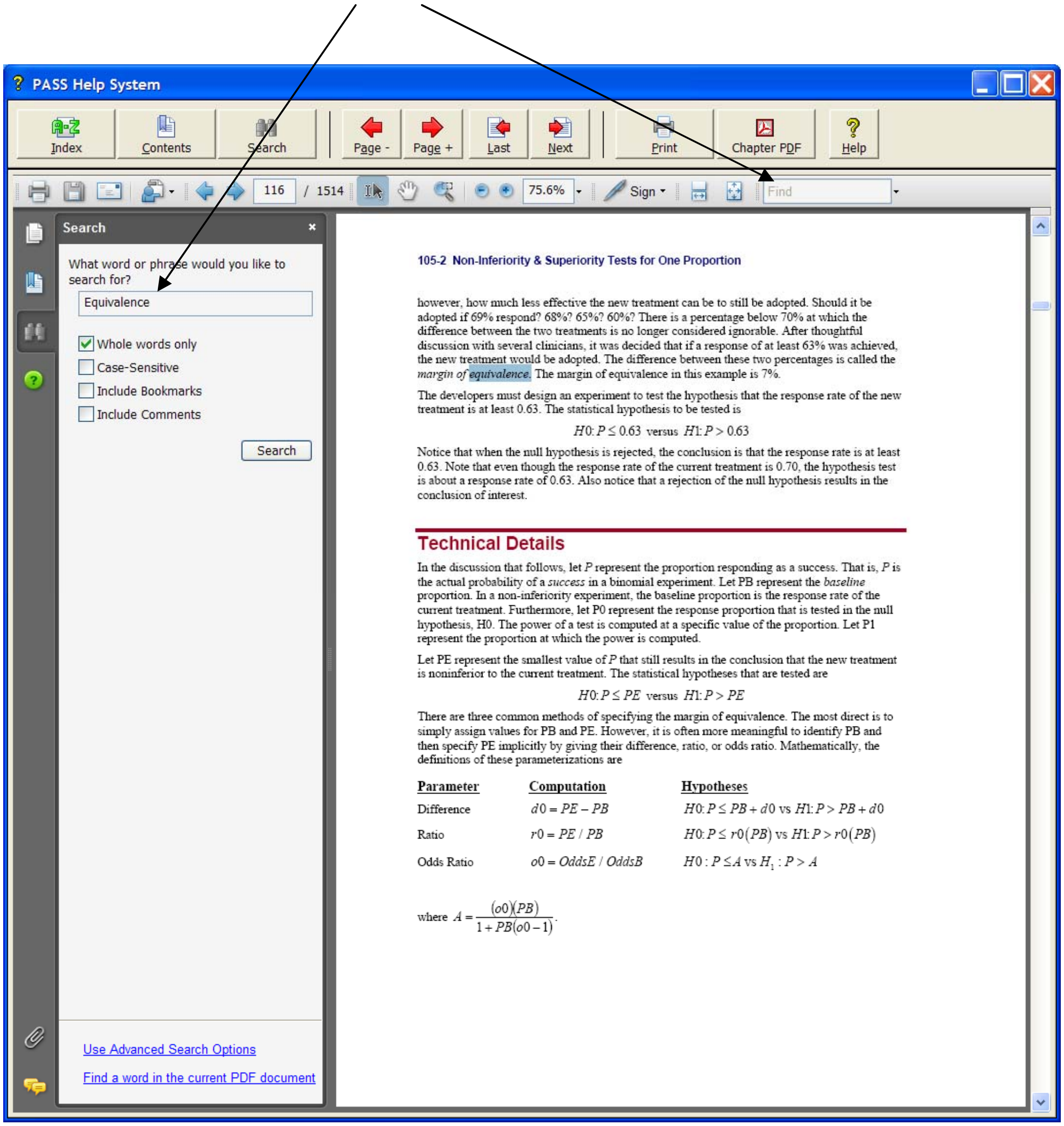
Contents Window

Clicking on the Contents button opens the Contents (Bookmarks) Window of the viewer. From this window you can expand the table of contents to view nested headings. You can click on the "Expand Current Bookmark" icon to instantly find the bookmark location for the currently-displayed page in the help document.



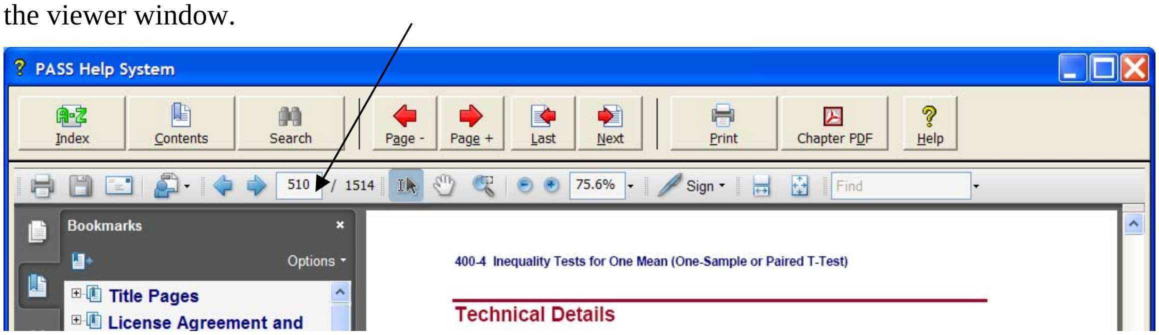
Search Window

Clicking on the Search button opens the Search Window of the viewer. From this window you can search the entire help system for any word or phrase. A search can also be initiated from the Find box in the viewer toolbar.



Printing the Documentation

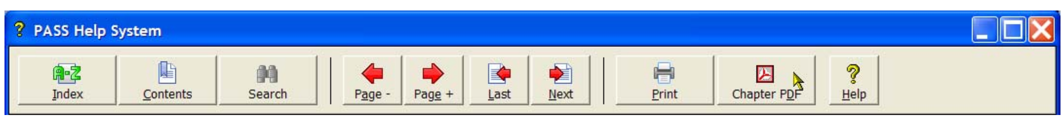
To print pages from the documentation, click on the **Print** button on the ***PASS*** Help System toolbar. This will launch the Adobe Reader print dialogue screen. You can choose to print a single page or a range of pages from the help file. When entering page numbers, remember to use the PDF file page numbers (e.g., 510-514) and not the page numbers found in the document pages (e.g., 400-4 to 400-8 is not a valid page range). The Adobe Reader page numbers can be seen in the viewer window.



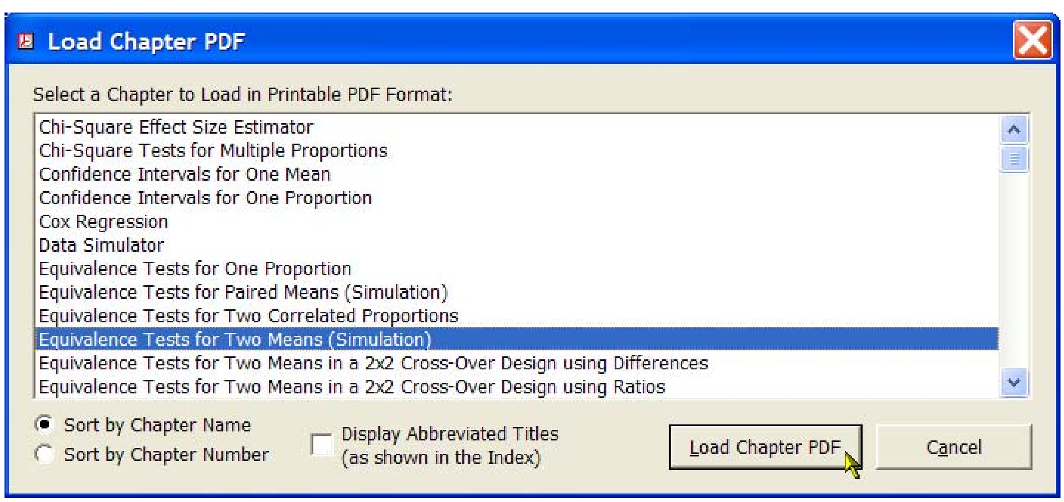
If you are using Adobe Reader 7, then the page numbers are found at the bottom of the viewer window.

One of the benefits of the ***PASS*** Help System is the ease with which you can print any chapter or topic from the electronic help manual. To print a single chapter or topic using your default PDF viewer, take the following steps:

1. Click on the **Chapter PDF** icon in the ***PASS*** Help System toolbar.



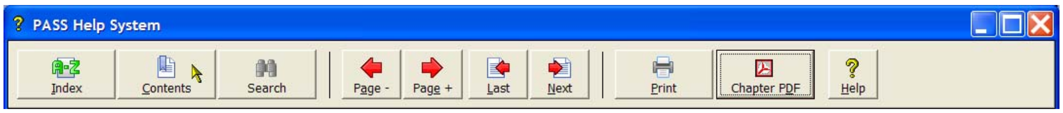
2. Choose the chapter you would like to print from the list and click **Load Chapter PDF**. This will launch the individual chapter PDF in a separate window using your default PDF viewer (e.g., Adobe Reader).



3. Use the **Print** function of your PDF viewer to print the entire chapter or individual pages from the chapter.

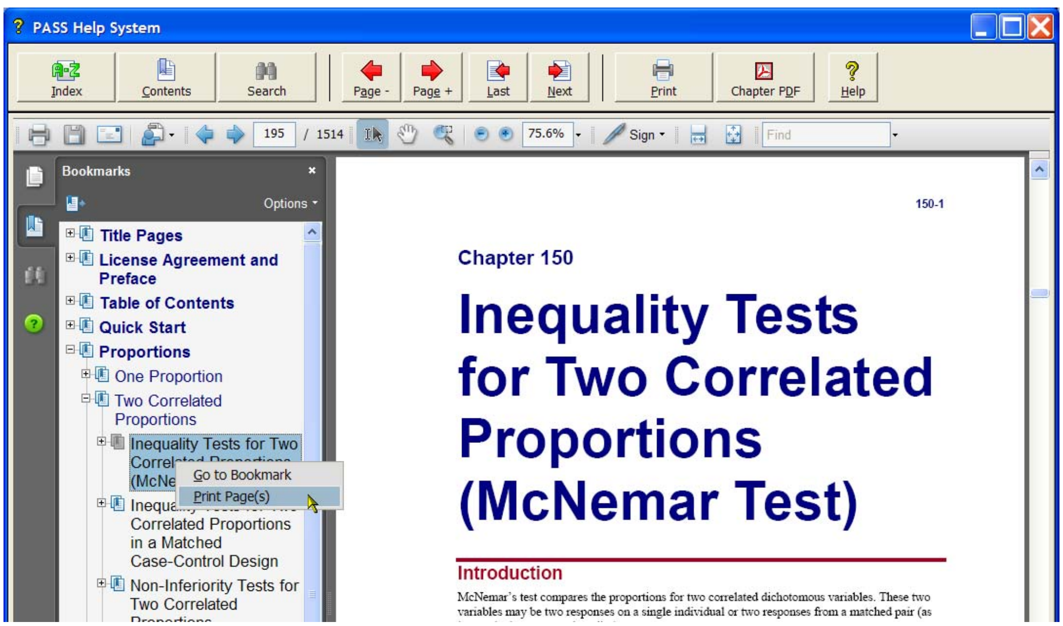
If you have Adobe Reader 8 or later, you can print entire chapters using an alternative method as follows (**This will not work with Adobe Reader 7**):

1. Open the Contents (Bookmarks) Window by clicking on the **Contents button** at the top of the ***PASS*** Help System display window.



2. Expand the bookmarks to display the chapter or topic name you wish to print (e.g., the Two-Sample T-Test Chapter). Then, **highlight** the chapter name, **right-click** on the highlighted selection (or select Options in the panel above), and select **Print Page(s)**. This will automatically print only the pages from the selected chapter.

   CAUTION: When you click Print Page(s), the command is sent to the printer automatically without any intermediate Print Setup window being displayed. Make sure that you have selected only the topic you want before clicking Print Page(s).



If you do not want to print the entire chapter, continue to expand the bookmark tree to the topic you wish to print before completing step 2. The Print Page(s) command prints all pages containing bookmarks that are nested within the highlighted bookmark.

Technical Support

If you have a question about ***PASS***, you should first look to the printed documentation and the included Help system. If you cannot find the answer there, look for help on the web at www.ncss.com/support.html. If you are unable to find the answer to your question by these means, contact ***NCSS*** technical support for assistance by calling (801) 546-0445 between 8 a.m. and 5 p.m. (MST). You can contact us by email at support@ncss.com or by fax at (801) 546-3907. Our technical support staff will help you with your question.

If you encounter problems or errors while using ***PASS***, please view our list of recent corrections before calling by going to www.ncss.com/release_notes.html to find out if you problem or error has been corrected by an update. You can download updates anytime by going to http://www.ncss.com/download.html. If updating your software does not correct the problem, contact us by phone or email.

To help us answer your questions more accurately, we may need to know about your computer system. Please have pertinent information about your computer and operating system available.

Chapter 2

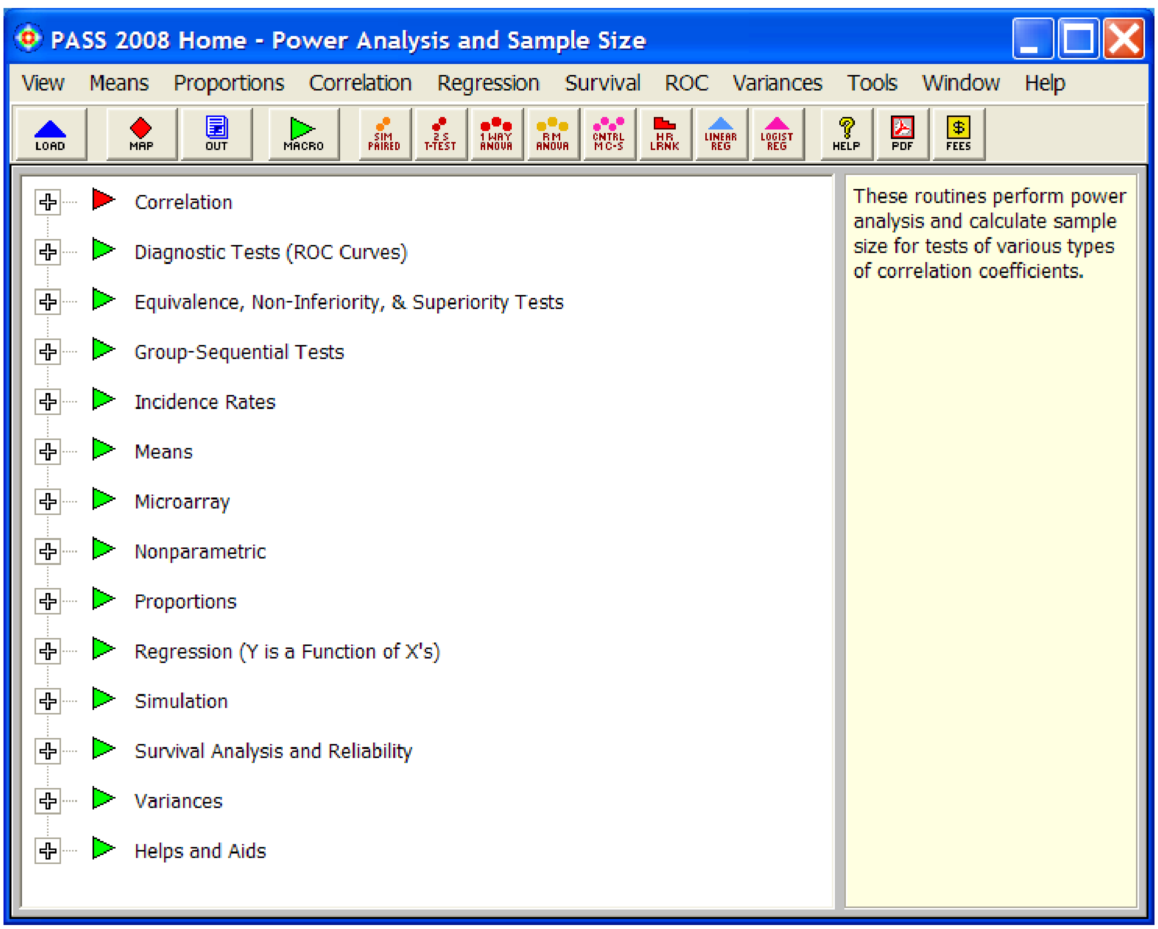
Running PASS

About This Chapter

This chapter will show you how to start up and run a power analysis of the two-sample *t* test. It will give you a brief introduction to the windows used in ***PASS***: the *PASS Home* window, the *procedure* window, and the *output* window.

Starting *PASS*

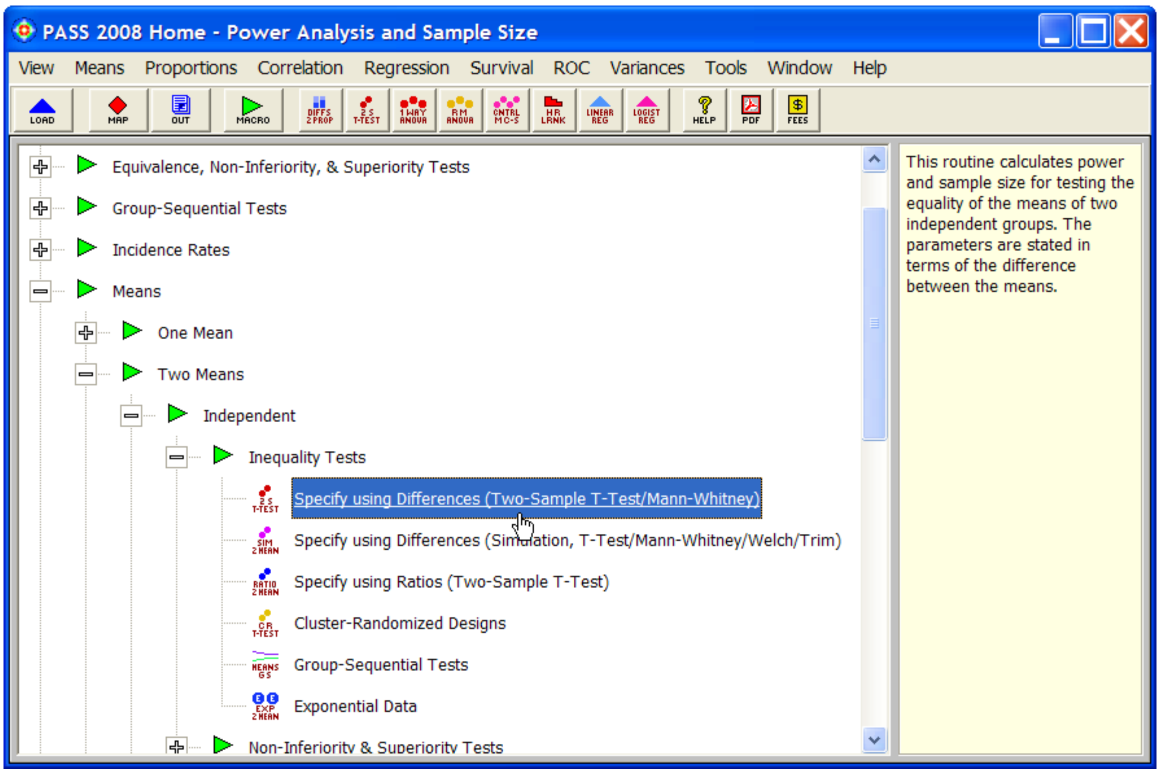
To start ***PASS***, select ***PASS 2008*** from the Windows Start menu or double-click the ***PASS*** icon. If you are licensed for ***PASS***, the following PASS Home window will appear.



This window gives you access to all of the ***PASS*** procedures. Clicking on the plus sign or double-clicking on a phrase will expand the list so that you can see the procedures in that group. To load a specific procedure window, double-click on it or highlight it and click the Load button.

Loading a Procedure

The *Two-Sample T-test* is a procedure to test the inequality of two means from independent samples. Take the following steps to load this procedure. Expand the Means topic by double-clicking on the word **Means**. Drilling down, double-click on **Two Means**, and then on **Independent**, then **Inequality Tests**. The first topic in the list is **Specify using Differences (Two-Sample T-Test/Mann-Whitney)**. This is the Two-Sample T-test. Double-click it.



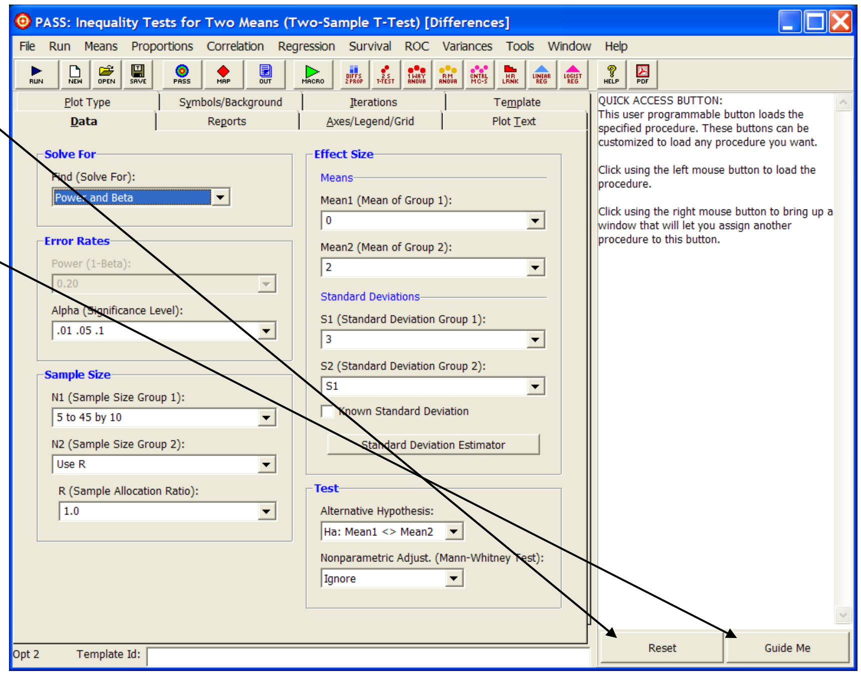
The **PASS: Inequality Tests for Two Means (Two-Sample T-Test) [Differences]** procedure window will appear. Procedure windows let you specify parameters, load and save templates, and run the analyses.

Entering Parameters and Running the Procedure

We will run a power analysis using the default values except that the value of **Mean2** will be **2** and the value of **S1** will be **3**.

1. Click the **Reset** button to set all options to their default values.

2. Click the **Guide Me** button to have PASS prompt you for the necessary options.



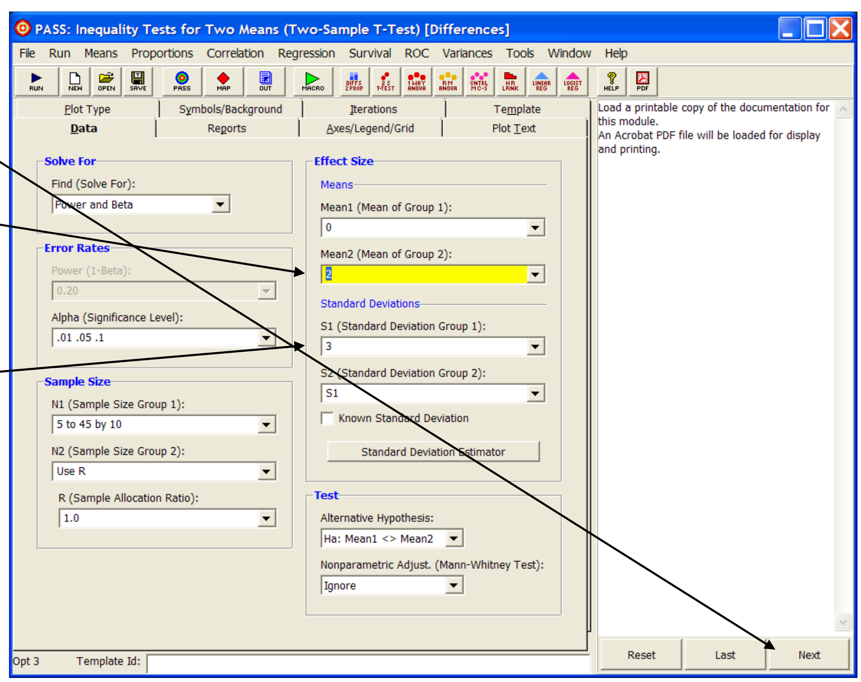
3. Click the **Next** button until you get to the **Mean2** option.

4. Enter **2**.

5. Click the **Next** button until you get to the **S1** option.

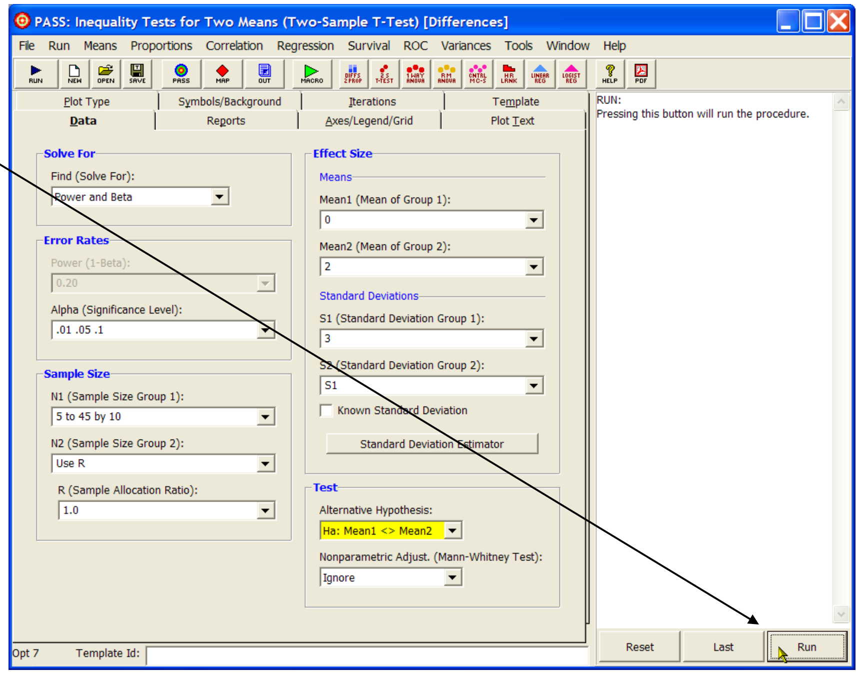
6. Enter **3**.

7. Click the **Next** button until he Next button changes into the Run button.



The completed window will appear as follows.

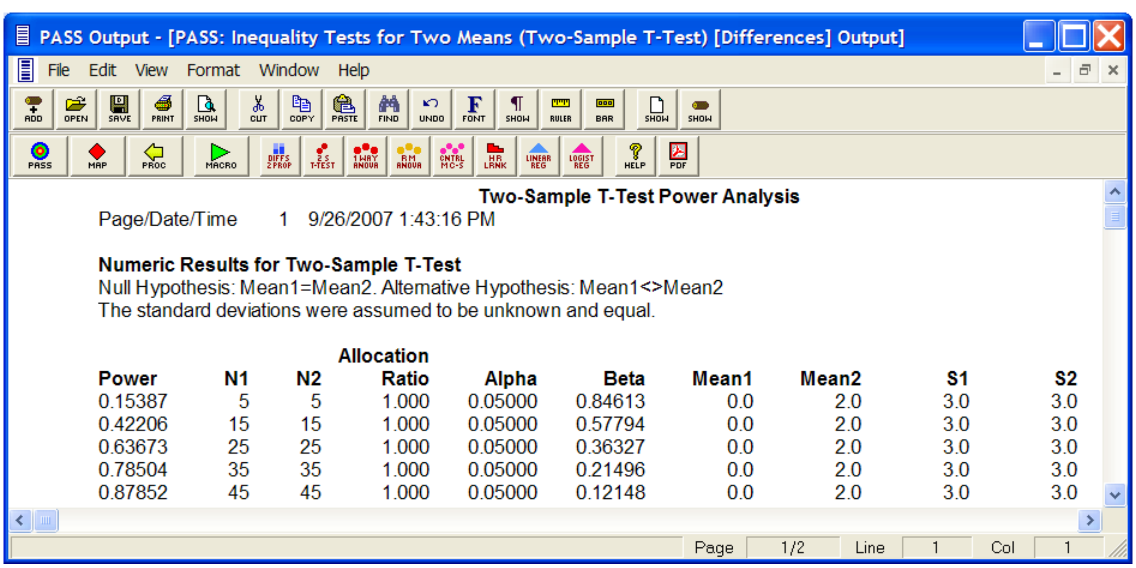
8. Click the **Run** button to perform the power analysis and display the report.



Viewing the Output

The Output window displays the output of the power analysis. It serves as a mini word processor—allowing you to view, edit, save, and print your output. You may want to scroll down to view the graph at the end of the report.

When you are finished, you can quit ***PASS*** by selecting **Exit PASS** from the File menu.



Two-Sample T-Test Power Analysis

Page/Date/Time 1 9/26/2007 1:43:16 PM

Numeric Results for Two-Sample T-Test

Null Hypothesis: Mean1=Mean2. Alternative Hypothesis: Mean1<>Mean2

The standard deviations were assumed to be unknown and equal.

Power	N1	N2	Allocation Ratio	Alpha	Beta	Mean1	Mean2	S1	S2
0.15387	5	5	1.000	0.05000	0.84613	0.0	2.0	3.0	3.0
0.42206	15	15	1.000	0.05000	0.57794	0.0	2.0	3.0	3.0
0.63673	25	25	1.000	0.05000	0.36327	0.0	2.0	3.0	3.0
0.78504	35	35	1.000	0.05000	0.21496	0.0	2.0	3.0	3.0
0.87852	45	45	1.000	0.05000	0.12148	0.0	2.0	3.0	3.0

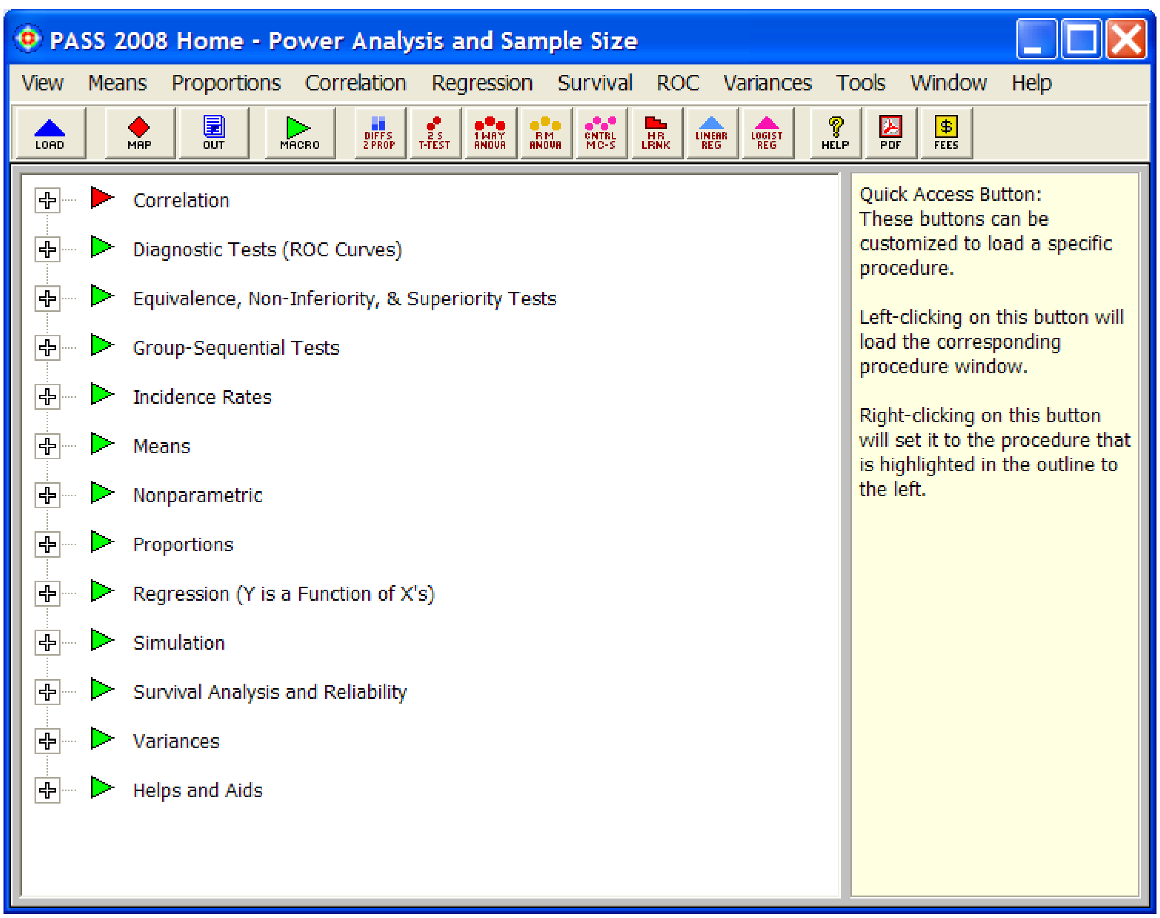
Chapter 3

The PASS Home Window

Introduction

The PASS Home window lets you quickly and easily find the appropriate procedure to be loaded. Using an outline format, it lists every procedure in ***PASS*** along with a brief statement that describes what the procedure is for and when it might be used.

The PASS Home window also lets you configure the eight quick-access buttons that appear on the toolbars of the Map, Procedure, and Output windows. These buttons give you immediate access to your favorite procedures. Right click on any procedure button to change it.



Using the PASS Home Window

The PASS Home window was designed to be easy to use. The window has a set of menus, a toolbar, and a large display area. On the left side of the display area is an outline list of all the procedures in ***PASS***. On the right side of the display area is the immediate help area that displays a brief statement explaining the currently selected item to the left.

Menus

The menus provide a convenient way to transfer from module to module within the ***PASS*** system. Each set of menus will be briefly described here.

Collapse Outline

This option collapses the outline so that only the main headings are displayed.

Expand to First Level

This option expands the outline so that the main headings and first-level subheadings are displayed.

Expand All

This option completely expands the outline so that all entries are displayed.

Bold Text

This option toggles the bolding of the text.

Goto Selected Procedure

This option loads the window of the procedure selected in the outline.

Options

This brings up a window allowing you to personalize your ***PASS*** installation and set various options affecting reports, constants, plots, and views.

Exit PASS

This option closes the PASS Home window and exits ***PASS***.

Procedure Menus

The procedure menus allow you to quickly find and load ***PASS*** power analysis procedures.

Tools Menu

The tools menu contains links to various ***PASS*** utilities.

Window Menu

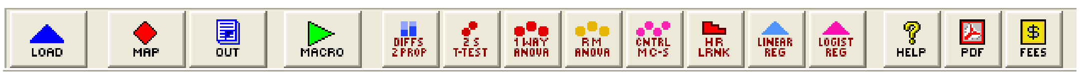
This menu allows you to open other windows in the ***PASS*** system such as the Spreadsheet, the Map (Quick Launch), or the Output window.

Help Menu

This menu allows you to view the ***PASS*** Help System, modify your serial numbers, get information about the program and load various portions of the printable PDF documentation.

Toolbar

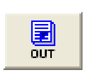
The toolbar gives you one-click access to several of the menu items. The menu item assigned to each button on the toolbar is displayed when the mouse is held over the button for a few seconds.



The action caused by each of these icons is discussed next.

Load Procedure. This button causes the window of the currently selected procedure to be displayed. You can accomplish the same action by double-clicking on the procedure name.

Map. This button causes the PASS Map (Quick Launch) window to be displayed. This window allows you to quickly select any procedure using icon buttons. This window can also be used to change the procedure quick-access buttons in the toolbar.



Output. This button causes the output window to be displayed.

Macro. This button can be used to interface with the macro system. Left-click on this button to run the active macro. Hold your mouse over the button to display the active macro name. Right-click on this button to load the Macro Command Center window.

Quick-Access. These buttons show up on all toolbars throughout the ***PASS*** system. Clicking on them with the left mouse button will display the corresponding procedure. Clicking on any of these buttons with the right mouse button allows you to change the procedure assigned to each button.

Help. This button loads the ***PASS*** Help System at the appropriate topic.

Printable PDF. This button loads the appropriate printable PDF chapter.

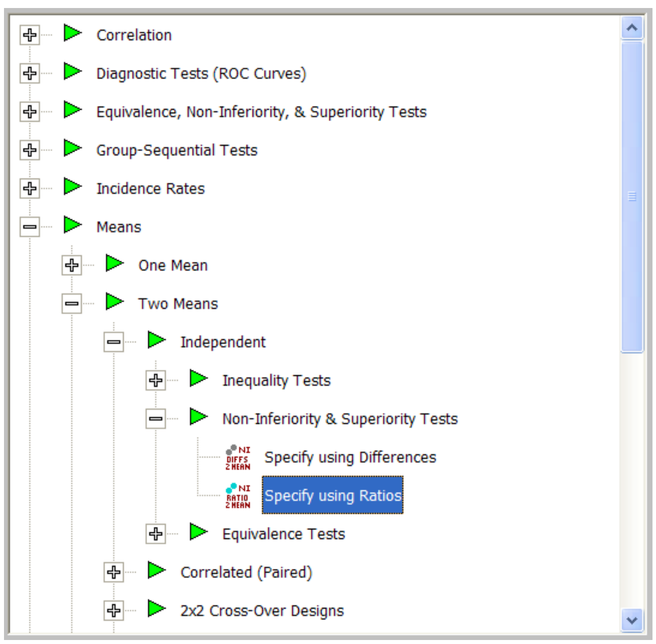
Pricing. This button loads pricing information and product brochures in PDF format.

Customizing the Toolbars

The eight quick-access procedure buttons that show up on all toolbars throughout the program may be changed using the PASS Home Window, the Map (Quick Launch), any procedure window, or the output window. To change the procedures available in the toolbar, right-click on any quick-access button. The buttons can also be changed by dragging and dropping buttons from the Map to the toolbar.

Outline

The outline expands and contracts as you either click on a plus or minus sign, or double-click on a topic. This gives you quick, intuitive access to all of the procedures in ***PASS***.



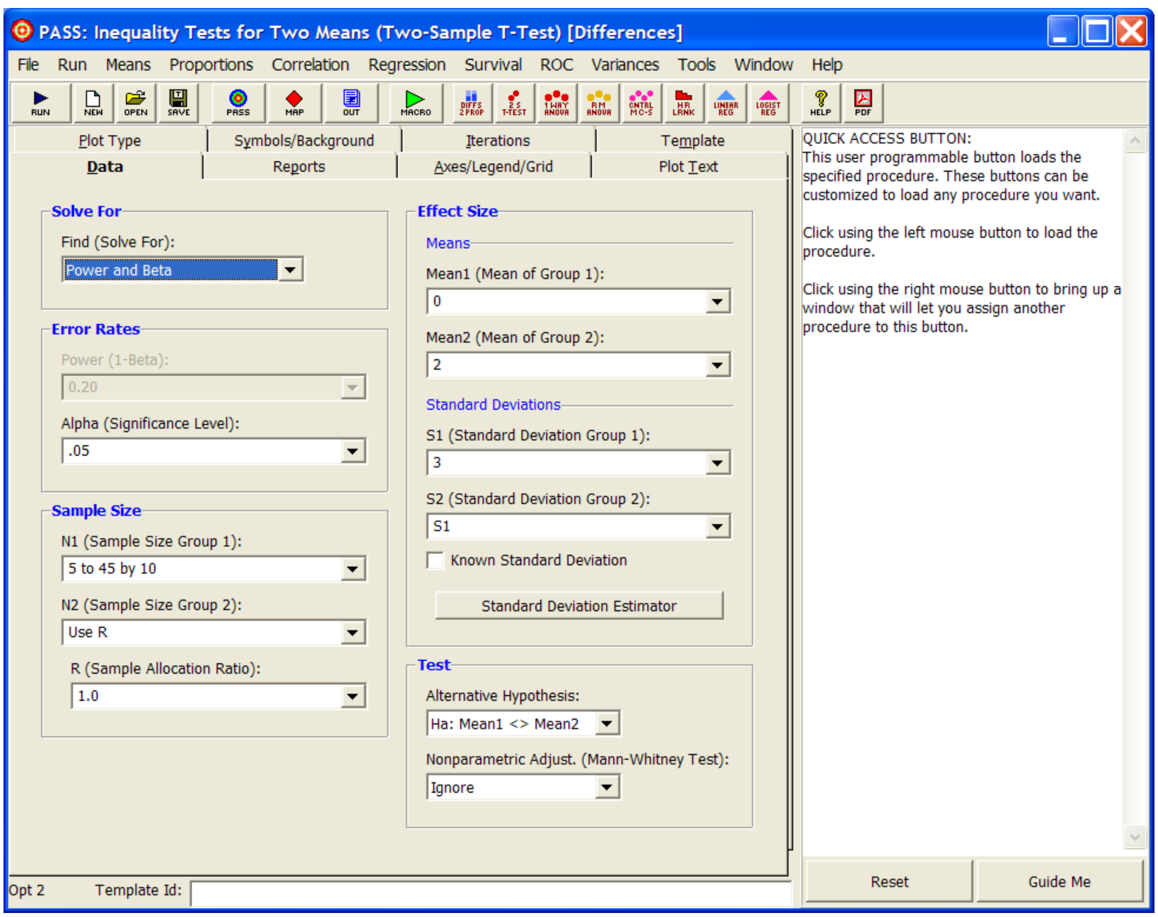
In the example shown here, we clicked on **Means**, then on **Two Means**, then on **Independent**, then on **Non-Inferiority & Superiority Tests**, and finally on **Specify using Ratios** to highlight it. If we double-clicked on **Specify using Ratios**, the Non-Inferiority & Superiority Tests for Two Mans [Ratios] procedure would be displayed.

Chapter 4

The Procedure Window

Introduction

All ***PASS*** procedures are controlled by a *procedure* window that contains all of the settings, options, and parameters that determine the input and output of the program. These options are separated into groups called *panels* or *tabs*. A particular panel is viewed by pressing the corresponding tab that appears near the top of the window. For example, in the window below, the Data panel is active. Other panels include Reports, Axes/Legend/Grid, and Plot Text.



The values of all options available for a procedure are referred to as a *template*. A template may be stored for future use in a *template file*. By creating and saving template files (often referred to as *templates*), you can tailor each procedure to your own specific needs. Each time you use a procedure, you simply load your template and run the analysis you have preset. You do not have to set all the options every time. The specific operations needed to do this are shown later.

At most six procedure windows can be opened at a time. You can widen the window to increase the size of the immediate help window by dragging the corners of the window.

Default Template

Whenever you close a procedure, the current settings are automatically saved in a default template file named *default*. This template file is automatically loaded when the procedure is next opened. This allows you to continue using the template without resetting all of the options.

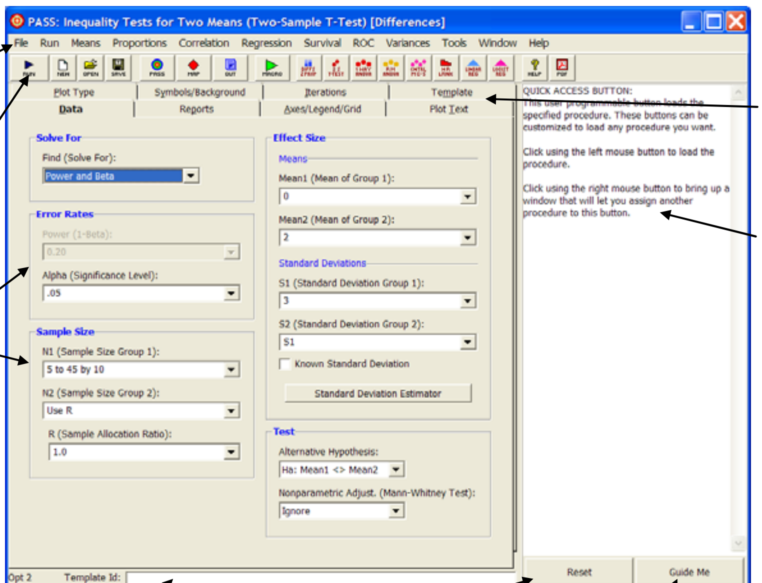
Procedure Window Anatomy

This section explains the various objects found on the procedure window.

Menus. The menus let you move to other windows.

Run. Clicking this button runs the program and generates output.

Options. These fields set values that control the analysis.



Tabs. The tabs let you view different groups of options.

Immediate Help. This box displays a brief help message about the field that the over-which the mouse is currently positioned.

Template Id. This box can contain a phrase that identifies this template.

Reset. This button resets all options under all tabs to their default values.

Guide Me. This button instructs the program to step you through the main options that must be set for an analysis.

Menus

The menus provide a convenient way to transfer from module to module within the ***PASS*** system. Each set of menus will be briefly described here.

File Menu

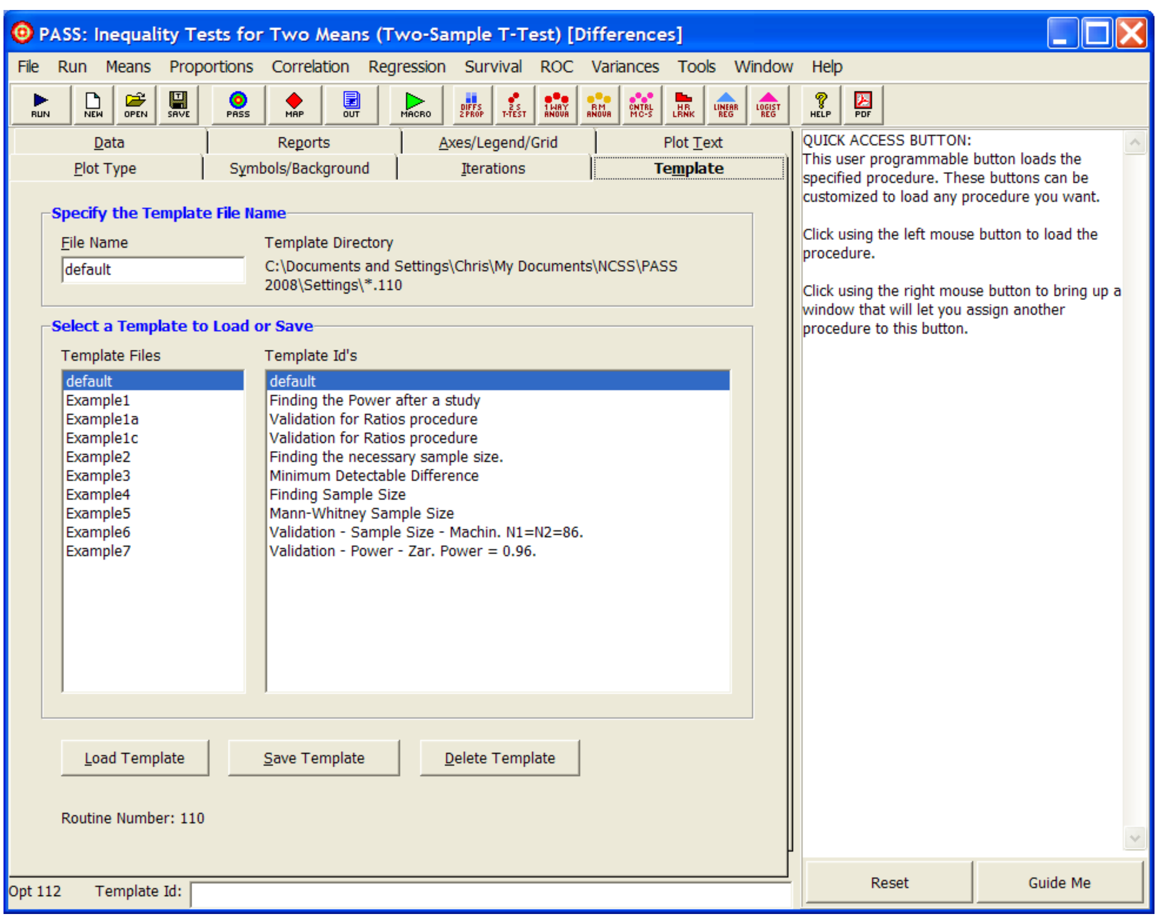
The File Menu is used for initializing, loading, and saving a copy of a template. Each set of options for a procedure, called a *template*, may be saved for future use. In this way, you do not have to set the options every time you use a procedure. Instead, you set the options the first time, save them as a template, and re-use the template whenever you re-use the procedure.

New Template (Reset)

This menu item resets all options to their default values. It performs the same function as the Reset buttons.

Open Template Panel

This option sets the Template panel as the active procedure panel. The Template panel lets you load or save template files. It displays all templates associated with this procedure along with the Template Id (the optional phrase at the bottom of the window).



Load Template (button)

To load a template file, select it from the list of files given in the Template Files box. Once the desired file is selected, press the Load Template button to load the template.

Save Template (button)

To save a template, enter the name you want to give the template file in the File Name box. You may also enter an identifying phrase in the box at the bottom of the window since this will be displayed along side of the file names. Finally, press the Save Template button to save the file. The template files are stored in the folder specified under Template Directory (*C:\...\[My] Documents\NCSS\PASS 2008\Settings*).

Note that there is no automatic connection between the template in memory and the copy on the disk. If you want to save the changes you have made to a template, you must use the Save Template option to save them.

Delete Template (button)

This button deletes the highlighted template file.

Save Template

This option saves the current option settings to the template file that is currently specified in the File Name option of the Template panel. You can be viewing any panel of the procedure when you issue this command—you do not have to be viewing the Template panel.

The template files are stored in the folder specified under Template Directory (*C:\...\[My] Documents\NCSS\PASS 2008\Settings*). You can erase any unwanted template files by deleting them from this folder using the Windows Explorer program.

The template files for each procedure have different file name extensions. Thus, you can use the same name for a template saved from the T-Test procedure as for a template saved from the Multiple Regression procedure. For example, if the Save Template command is issued in the window shown above, the current settings will be saved in a file called *default.110* in the Settings folder.

The **Save** button on the toolbar provides this same operation. It may be more convenient than selecting this menu item.

Options

This brings up a window allowing you to personalize your ***PASS*** installation and set various options affecting reports, constants, plots, and views.

Close Procedure

This option closes this procedure window.

Exit PASS

This option terminates the ***PASS*** system. Before using this option, you should save all spreadsheets, templates, and output documents that you want to keep.

Run Menu

This menu controls the execution of the program.

Run Procedure

The Run Procedure option runs the analysis, displaying the output in the Output document of the word processor. After you have set all options to their appropriate values, select this option to perform the analysis.

Note that the procedure may also be run by pressing the *F9* function key or by pressing the left-most key on the toolbar (the dark-blue-arrow button).

Abort

After starting a procedure, you may find that it is taking longer than you anticipated to finish. You can stop the running of the procedure by pressing this button. The red stop-sign icon that appears on the top right of the screen may be pressed for the same purpose.

Procedure and Tools Menus

These menus allow you to transfer to various ***PASS*** procedures and utilities.

Window Menu

This menu lets you display any of the other windows in the ***PASS*** system that are currently open such as the Output window, the Spreadsheet window, the PASS Home window, or any procedure windows.

Output

Select this option to display the output window.

Spreadsheet

Select this option to display the spreadsheet.

PASS Home

Select this option to display the PASS Home window.

Map – Quick Launch

Select this option to display the procedure Map.

Reset Window Positions

Occasionally, ***PASS*** windows will be loaded, but will not display. This menu item will load the Options window to a tab that will let you reset the position of all program windows.

Help Menu

This menu gives you access to the ***PASS*** Help System and PDF documentation, including references and the Quick Start Manual.

Toolbar

The toolbar is a series of small buttons that appear just below the menus at the top of the procedure window. Each of these buttons provides quick access to a menu item.

The action caused by each of these icons is discussed next.

Run. This button runs the current procedure with the current settingsand generates output.

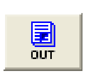
Reset (New Template). This button sets all parameters to their default values.

Open. This button opens the template panel so you can load a saved template.

Save. This button saves the current settings to a template file.

Map. This button causes the PASS Map (Quick Launch) window to be displayed. This window allows you to quickly select any procedure using icon buttons. This window can also be used to change the procedure quick-access buttons in the toolbar.

PASS Home. This button causes the output window to be displayed.



Output. This button causes the output window to be displayed.

Macro. This button can be used to interface with the macro system. Left-click on this button to run the active macro. Hold your mouse over the button to display the active macro name. Right-click on this button to load the Macro Command Center window.

Quick-Access. These buttons show up on all toolbars throughout the ***PASS*** system. Clicking on them with the left mouse button will display the corresponding procedure. Clicking on any of these buttons with the right mouse button allows you to change the procedure assigned to each button.

Help. This button loads the ***PASS*** Help System at the appropriate topic.

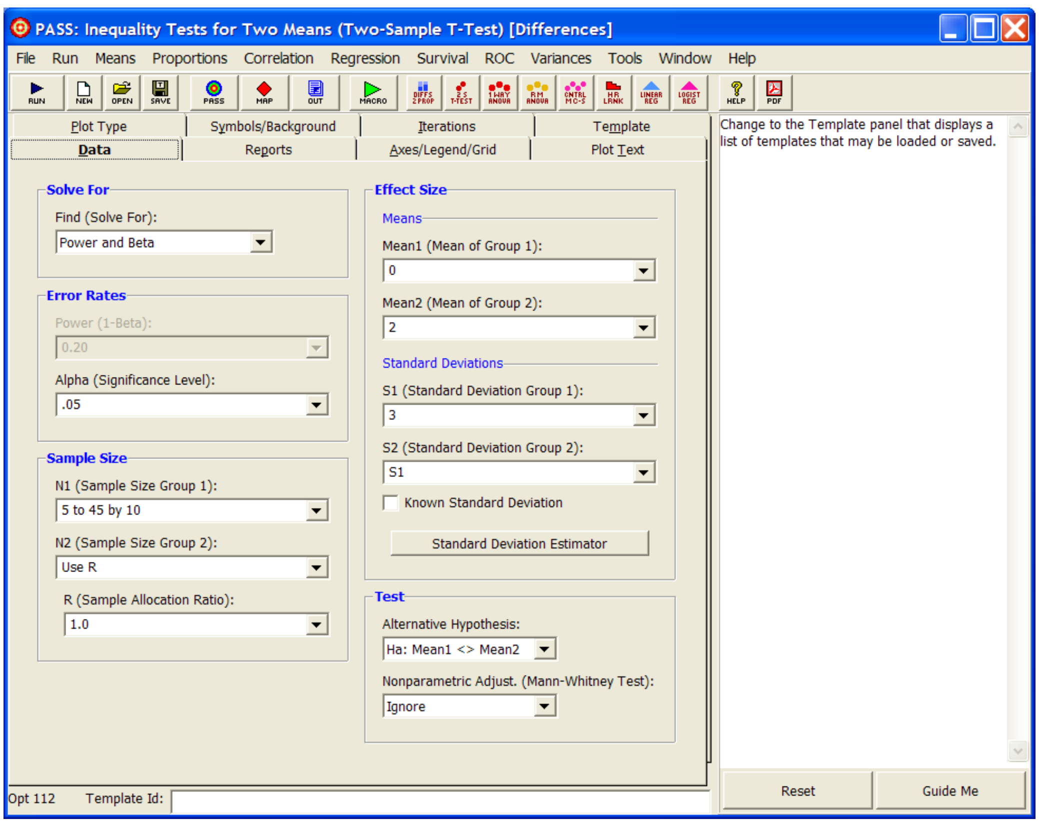
Printable PDF. This button loads the appropriate printable PDF chapter.

The Procedure Window Tabs

The procedure window contains several sets of options on panels (or tabs). Each panel is displayed by clicking on the appropriate tab. We will now describe the purpose and operation of each tab.

Data Tab

The Data tab displays most of the options specific to the procedure. This is where you set the values of power, sample size, alpha, etc. These options are described in detail in the chapters corresponding to each procedure. Once you have set the options, click the **Run** button to generate the output.



Entering Multiple Values

In most cases, boxes that are extra wide allow you to enter multiple values. When this is done, a separate analysis is done for each combination of all multiple values. For example, if you enter four sample sizes and three alpha values, the resulting report will contain 3 × 4 = 12 rows, one for each combination.

You can enter multiple options using list or the *to-by* syntax. The *to-by* syntax is most easily described by an example.

The *to-by* phrase *20 to 100 by 20* is translated to the values: *20,40,60,80,100*.

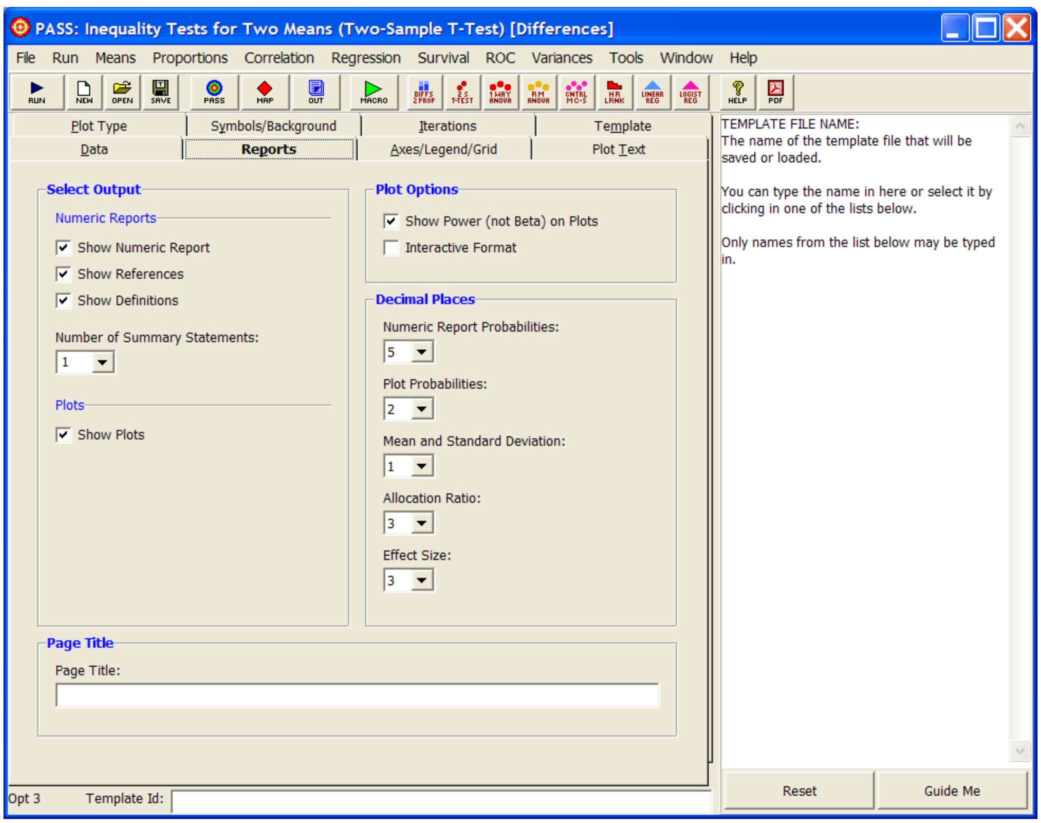
Find (Solve For)

Specify the parameter that is to be solved for in terms of the other parameters. For example, you might want to solve for power or sample size.

In most cases, the algorithm for the calculating the power is programmed within ***PASS***. When other parameters (such as sample size or difference) are selected, a binary search is conducted using the power algorithm. These searches can be time consuming, so the best place to start is with **Find (Solve For)** set to *Power and Beta*.

Reports Tab

The Reports tab displays the options that control the output reports.



Select Output – Numeric Reports

Show Numeric Report

Determines whether the numeric report is displayed in the output.

Show References

Check this box to cause the literature reference(s) to be displayed on the report.

Show Definitions

Check this box to show the definitions at the end of the numeric report. Although these definitions are helpful at first, they tend to clutter the output and this option lets you skip them.

Number of Summary Statements

The program will output a text statement summarizing the results for each scenario. This option specifies the number of scenarios (rows) from the Numerical Report that will have a summary statement displayed. Select *0* to omit the summary statements.

Select Output – Plots

Show Plots

Check this box to display plots in the output.

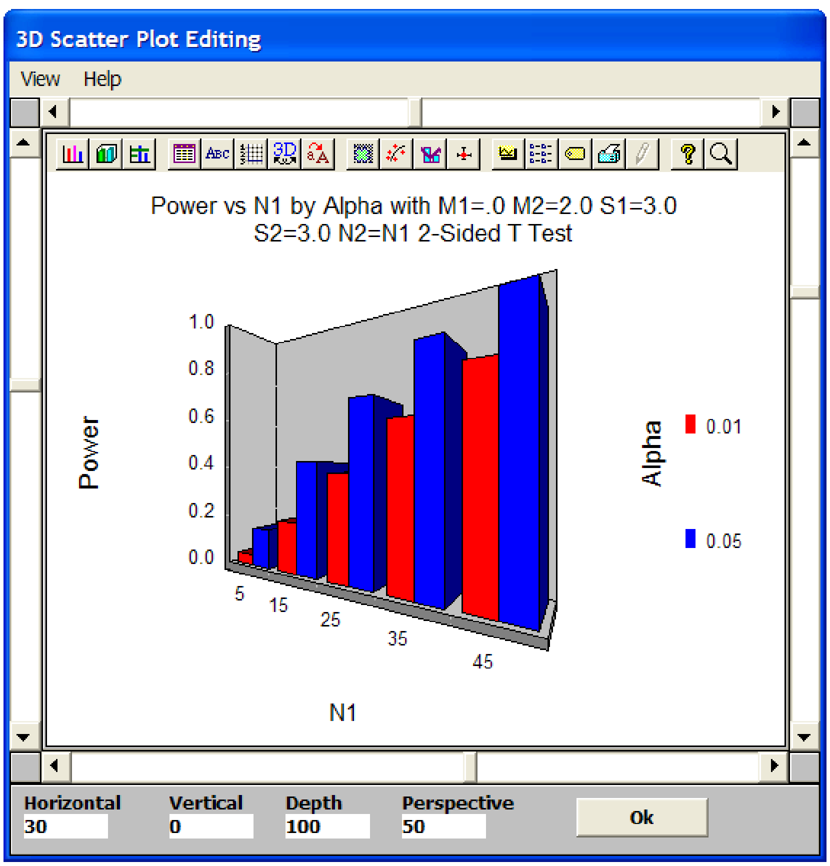
Plot Options

Show Power (not Beta) on Plots

When checked, power is displayed on the plots. When unchecked, power quantities are displayed on the plots as beta using the relationship: **Beta = 1 – Power**.

Interactive Format

This option controls whether the plot may be reformatted interactively after it has been generated. When checked, this option allows charts to be formatted interactively using a plot-editing window.

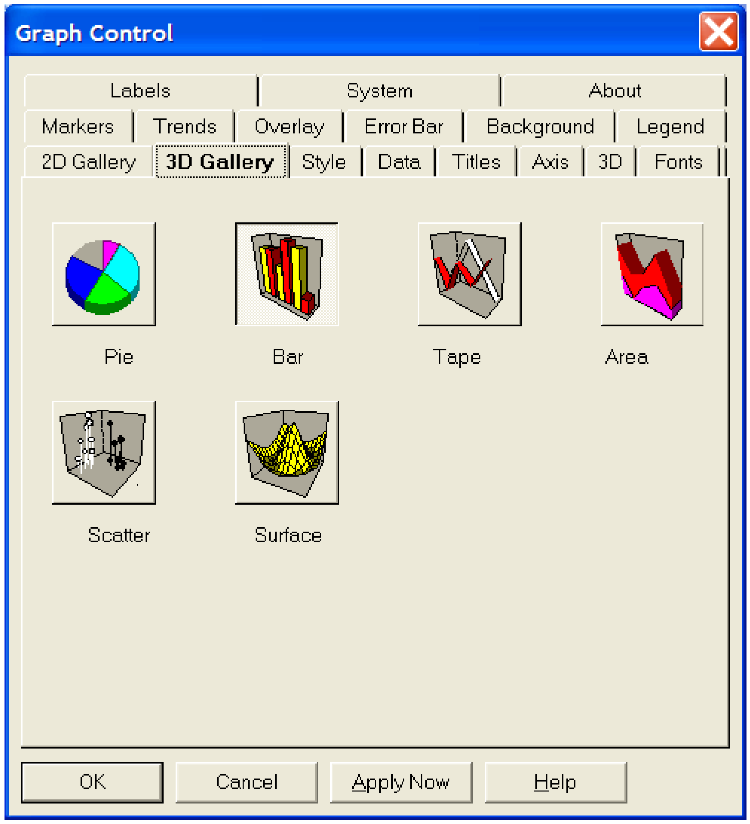


The four scroll bars around the edge of this window control the vertical axis, horizontal axis, depth, and perspective. The current values of these parameters are shown in the boxes at the bottom of the screen.

Once you are finished editing chart, click the **Ok** button to proceed.

Each of the buttons along the top of the Scatter Plot Editing window will display a different tab of the Graph Control window. Each tab provides options which allow detailed modification of the chart.

We will not document these options here since most of them are not necessary to the running of ***PASS***. If you want to explore these options further, choose the Help button at the bottom of the window. This will bring up a special help system that describes all graphics options in detail.



Decimal Places

Decimals

These options set the number of decimal places in corresponding values of the numeric and graphic output.

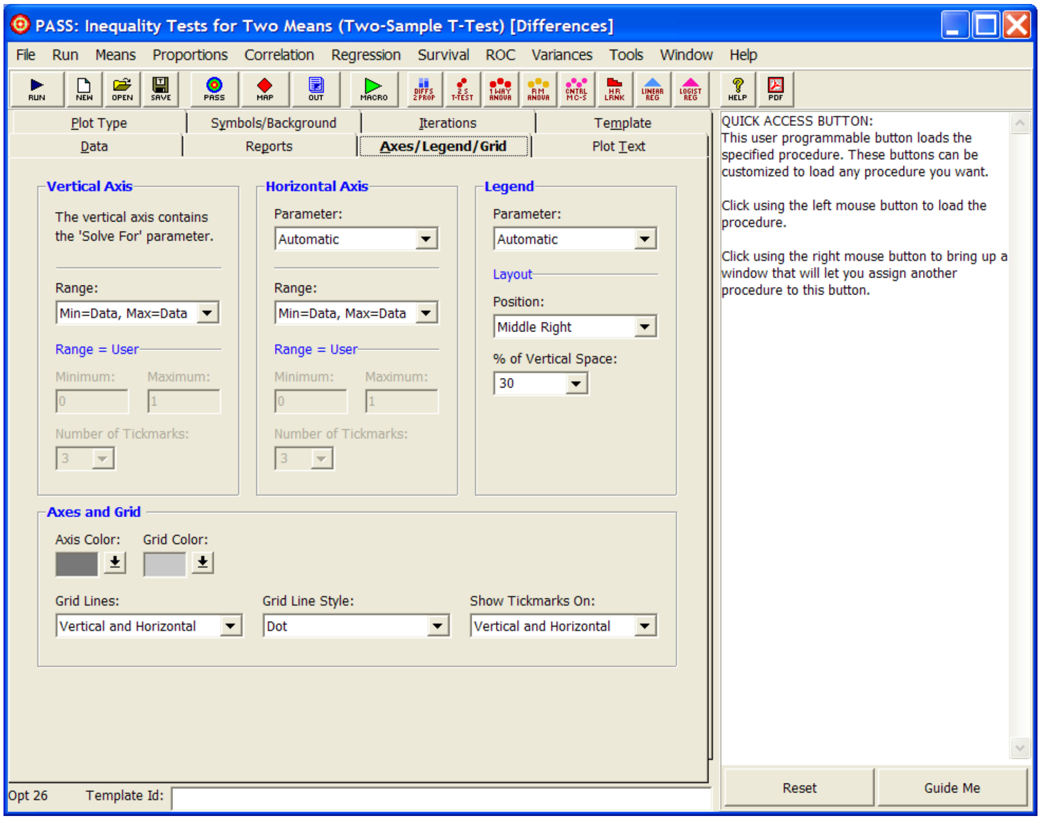
Page Title

Page Title

This option allows you to enter an option title phrase that will appear in the heading of each page of the output.

Axes/Legend/Grid Tab

The Axes/Legend/Grid tab presents the options that control the plot axes (including the parameters plotted), legend, and gridlines.



Vertical and Horizontal Axes

Parameter

This option selects which of the parameters from the Data tab is displayed on the horizontal axis. The vertical axis always contains the **Solve For** parameter, so you cannot select the parameter that was listed in the Find option. Also, you would normally only select a parameter that has multiple entries.

When this option is set to *Automatic*, the parameter with the most values is selected.

Range

This option designates how the minimum and maximum along this axis are specified. Available options are:

- **Min=0, Max=Data**

  The axis minimum is set to zero. The maximum is selected from the data values. The values of the Minimum, Maximum, and Number of Tickmarks are ignored.

- **Min=Data, Max=Data**

  Both the minimum and the maximum of the axis are determined from the data. The values of the Minimum, Maximum, and Number of Tickmarks are ignored.

- **User (Given Below)**

  This option lets you set the Minimum, Maximum, and Tumber of Tickmarks to scale the axis. These options determine which of the axes have grid lines displayed. This option is particularly useful when you want to make sure that the axis displaying power values displays a grid between zero and one.

Vertical and Horizontal Axes – Range = User

These options are only used when **Range** is set to *User*.

Minimum and Maximum

Specify the axis minimum and/or maximum.

Number of Tickmarks

Specify the number of tickmarks along this axis.

Legend

Parameter

A separate line is drawn for each value of this parameter. The lines are labeled in the legend. When this option is set to *Automatic*, the parameter with the second most values is selected.

Legend – Layout

Position

This option sets the position of the legend.

% of Vertical Space

Specify the size of the legend area as a percentage of the maximum possible. This option lets you shrink a legend that is too large.

Axes and Grid

Axis Color

Specify the color of the axis lines.

Grid Color

Specify the color of the grid lines.

Grid Lines

This option determines which of the axes have grid lines displayed.

Grid Line Style

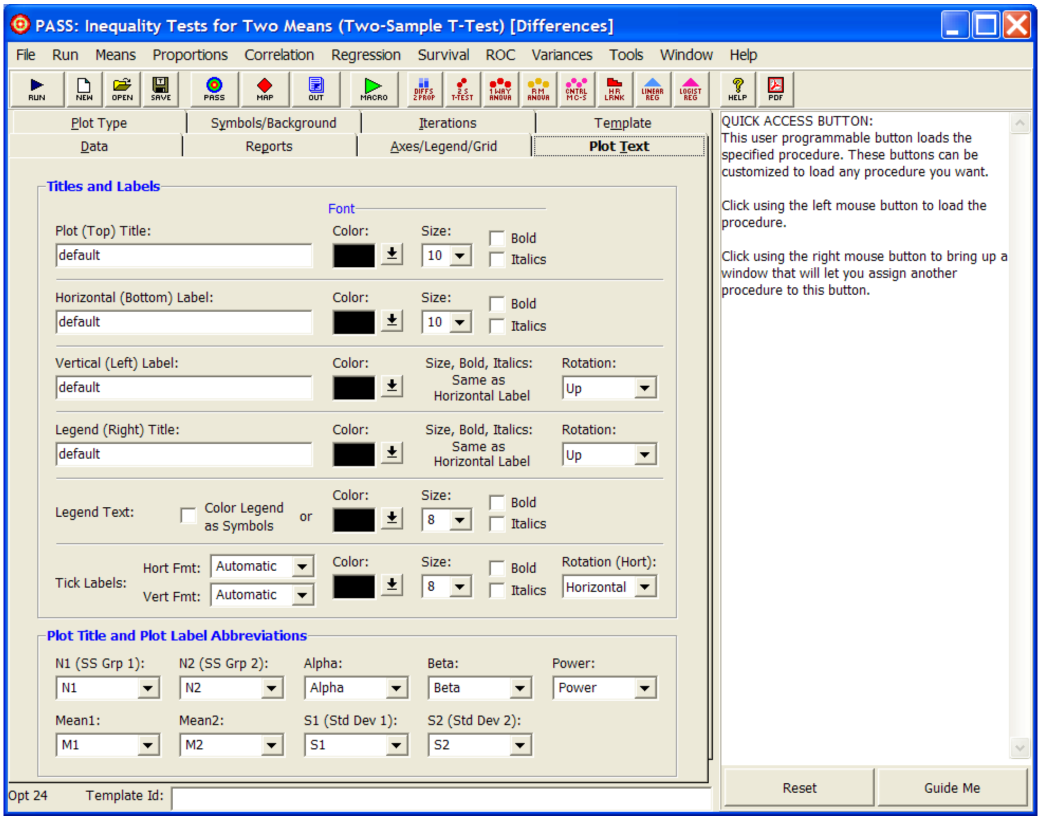
Specify the pattern of the grid lines.

Show Tickmarks On

This option controls which of the axes have tickmarks displayed.

Plot Text Tab

This tab controls the titles, labels, and abbreviations on the plots.



Titles and Labels

Plot Titles and Labels

These options specify the text of the titles and labels displayed on the plots.

Rotation

This option allows you to change the orientation of the corresponding titles or labels.

Titles and Labels – Font

Color, Size, Bold, and Italics

These options specify the font of text displayed on the plot.

Titles and Labels – Legend Text

Color Legend as Symbols

Normally, text in the legend is displayed using the color selected by the Color option. This option indicates that each legend entry is to be displayed in the corresponding group color.

Titles and Labels – Tick Labels

Hort Fmt and Vert Fmt

This option allows you to specify the format of the tick labels on the vertical axis. Select a format from the list or type your own format string according to the syntax rules that follow. Select *Automatic* to use the default format.

Syntax

Character	Function	Description
0	Digit Placeholder	Displays a digit or a zero.
#	Digit Placeholder	Displays a digit or nothing.
.	Decimal Placeholder	Determines how many digits are displayed to the left and right of the decimal separator.
,	Thousand Separator	Separates thousands from hundreds within a number that has four or more places to the left of the decimal separator.
%	Percentage Placeholder	The number is multiplied by 100. The percent character (%) is inserted in the position where it appears in the format string.
E- E+ e- e+	Scientific Format	Displays the number in scientific format.
\	Literal Character	Displays the character immediately following "\" in the format string.

Syntax Examples

Number	Format String	Number Displayed on Plots
1234	0	1234
1234	00000	01234
0.1234	0.00%	12.34%
0.1234	0%	12%
1234	#,##0	1,234
123456	#,##0,\k	123k
12345678	#,##0,,\m\i\l	12mil
12345678	0.0E+00	1.2E+07
0.1234	0.00	0.12
0.1234	0.00000	0.12340
0.1234	0.0E-00	1.2E-01

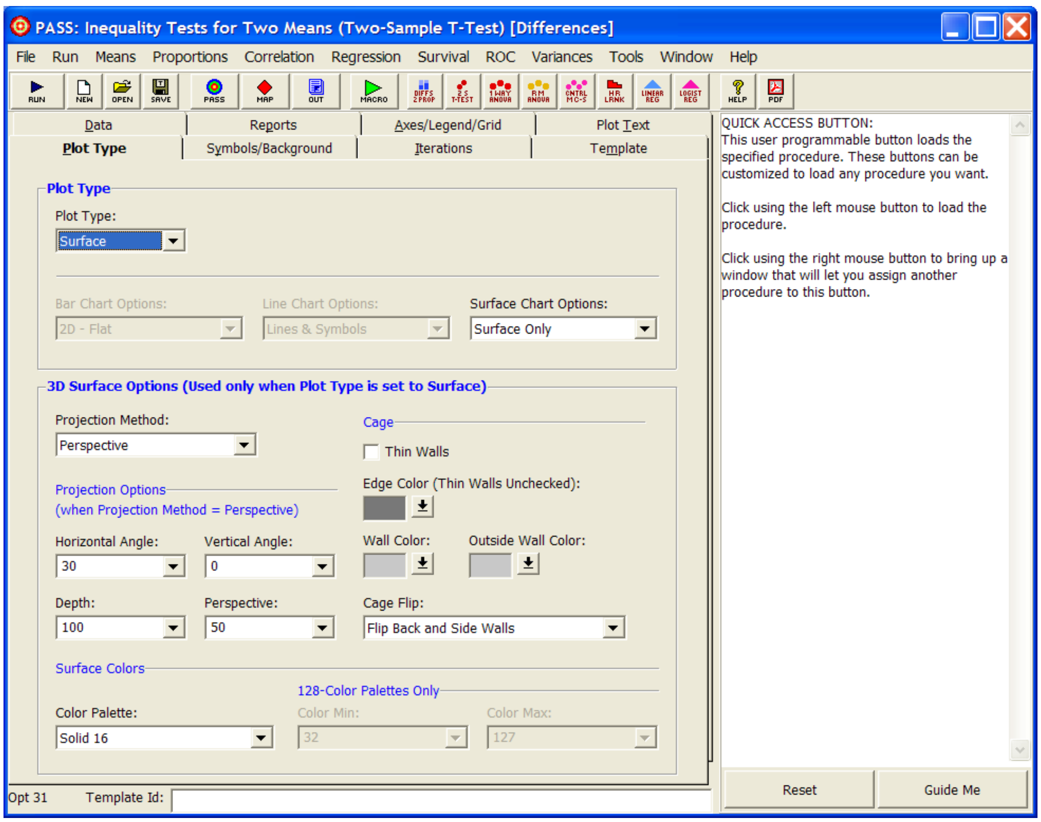
Plot Title and Plot Label Abbreviations

Parameter Abbreviations

These options specify the abbreviations that are used for the parameters in the titles of the plots and the axis labels. It is usually necessary to keep these abbreviations as short as possible since the titles can only contain 80 characters.

Plot Type Tab

These options allow you to specify to type of plot to output.



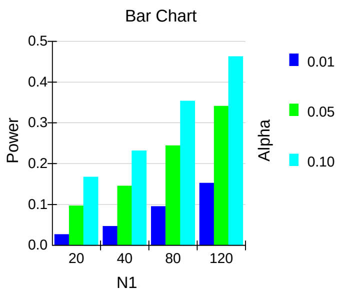
Plot Type

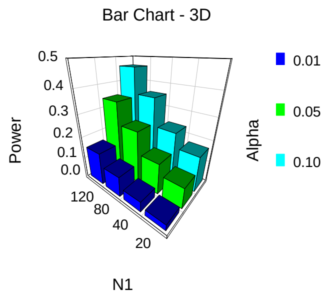
Plot Type

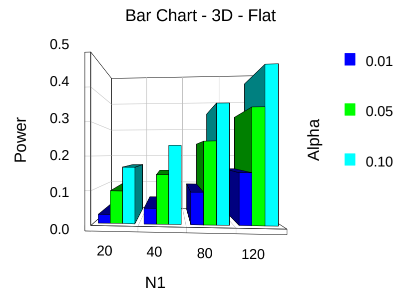
This option controls the type of plot that is displayed. Bar charts, line charts, and surface charts are available.

Bar Chart Options

The following plots are available when **Plot Type** is set to *Bar*.

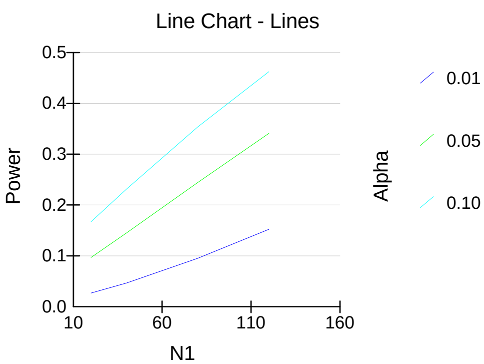


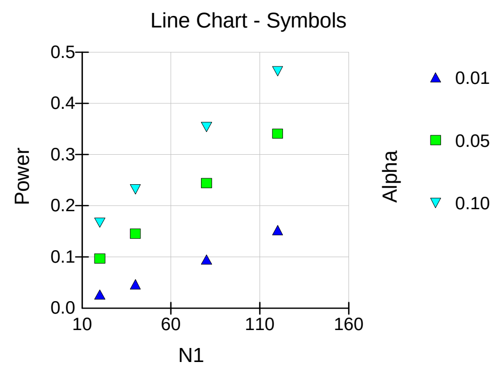


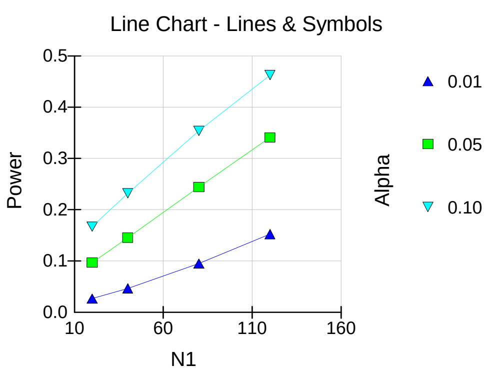


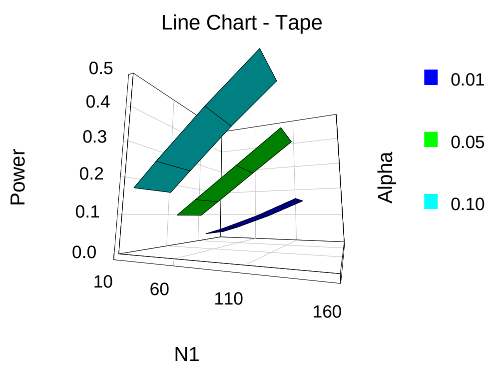
Line Chart Options

The following plots are available when **Plot Type** is set to *Line*.



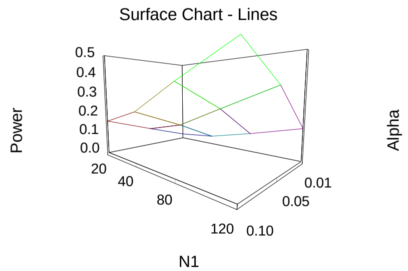


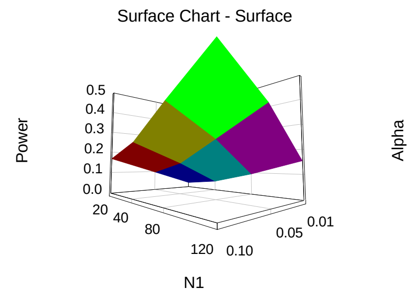


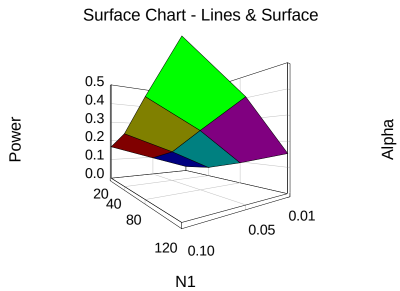


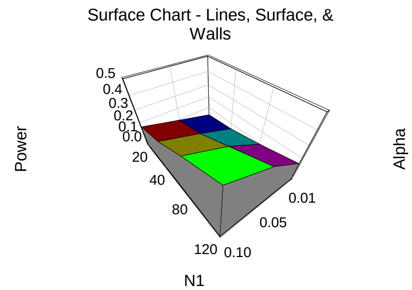
Surface Chart Options

The following plots are available when **Plot Type** is set to *Surface*.









3D Surface Options

These options are only used when **Plot Type** is set to *Surface*.

Projection Method

Sets the projection method of 3D charts.

- **Off**

  No graph is drawn.

- **Perspective**

  The axes are tilted to give a 3D perspective to the plot.

- **Isometric**

  The graph is drawn, but no perspective is attempted.

3D Surface Options – Projection Options

These options are only used when **Projection Method** is set to *Perspective*.

Horizontal Angle

This option sets the horizontal viewing angle (in degrees) for 3D plots. It represents an angle around the base of the plot. The range of values is -180 to 180 degrees. This option may be changed interactively when the Interactive Format option is checked.

Vertical Angle

This option sets the vertical viewing angle (in degrees) for 3D plots. It represents an angle above or below a point halfway up the graph. Values may range from -60 to 90 degrees. This option may be changed interactively when the Interactive Format option is checked.

Depth

This option sets the projected depth of 3D plots. Depth is a percentage of 100, calculated to provide equal increments in the X and Z directions. Values may range from 5 to 400. This option may be changed interactively when the Interactive Format option is checked.

Perspective

This option sets the degree of perspective foreshortening in 3D plots. Perspective is the perceived distance of the viewer from the graph. The range of values is 0 to 100. This option may be changed interactively when the Interactive Format option is checked.

3D Surface Options – Cage

Thin Walls

This option specifies whether the walls of the axis grid that form the background of the chart are thick or thin.

Edge Color

Specify the color of the cage (grid) edge. This option is only used if **Thin Walls** is *unchecked*.

Wall Color and Outside Wall Color

Specify the color of the cage (grid) wall.

Cage Flip

This option controls whether the back and side walls of the graph cage are allowed to switch to the opposite edge for better viewing as the graph is rotated.

3D Surface Options – Surface Colors

Color Palette

Specify a color palette for the surface chart. Using a setting of, say, Black to Red will allow the surface plot to show a continuous array of red hues from lowest to highest.

3D Surface Options – Surface Colors – 128-Color Palettes Only

These options are only used when **Color Palette** is set to a value containing 128 colors.

Color Min

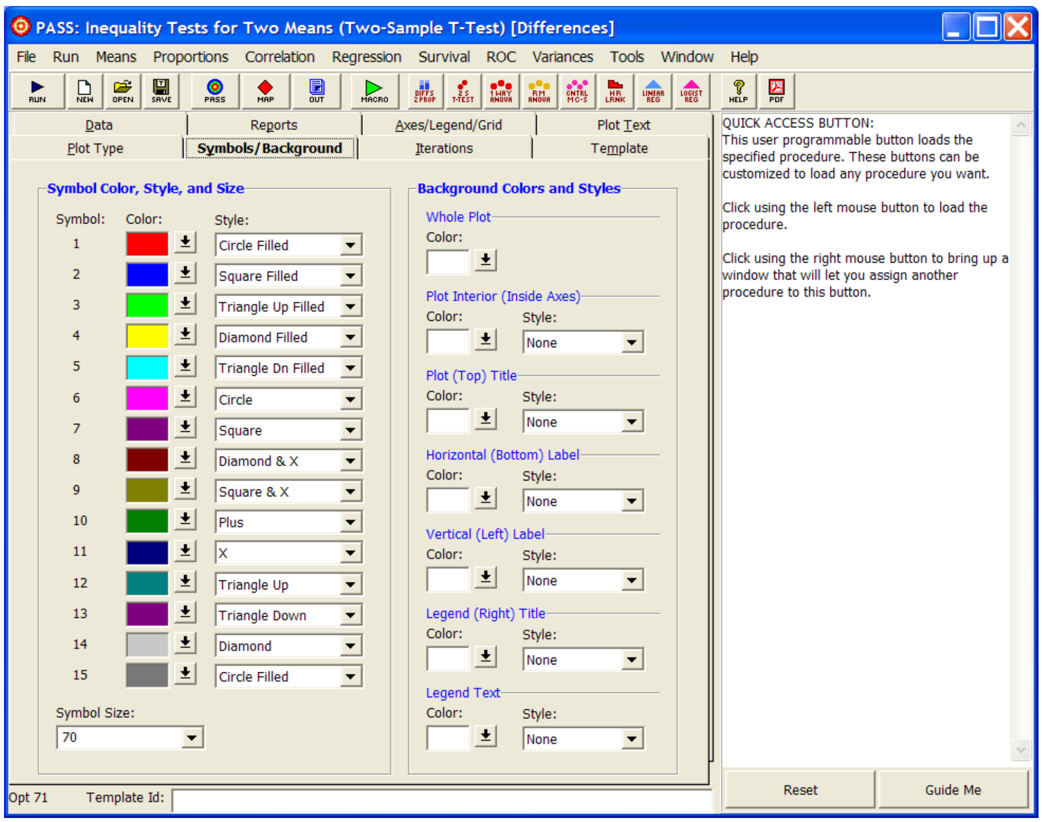
Specifies the number of the color to be associated with the lowest numerical value. Possible values are 32 to 127. A value near 50 usually works well. Note that this option only works with 128-color palettes.

Color Max

Specifies the number of the color to be associated with the largest numerical value. Possible values are 32 to 127. A value near 120 usually works well. Note that this option only works with 128-color palettes.

Symbols/Background Tab

This tab specifies the appearance of up to fifteen symbols. If more than fifteen symbols are needed, the first fifteen are repeated. The plot background is also controlled by this tab.



Symbol Color, Style, and Size

Color and Style

These options specify the color and shape of the plotting symbols.

Symbol Size

This option sets the radius (size) of all plotting symbols.

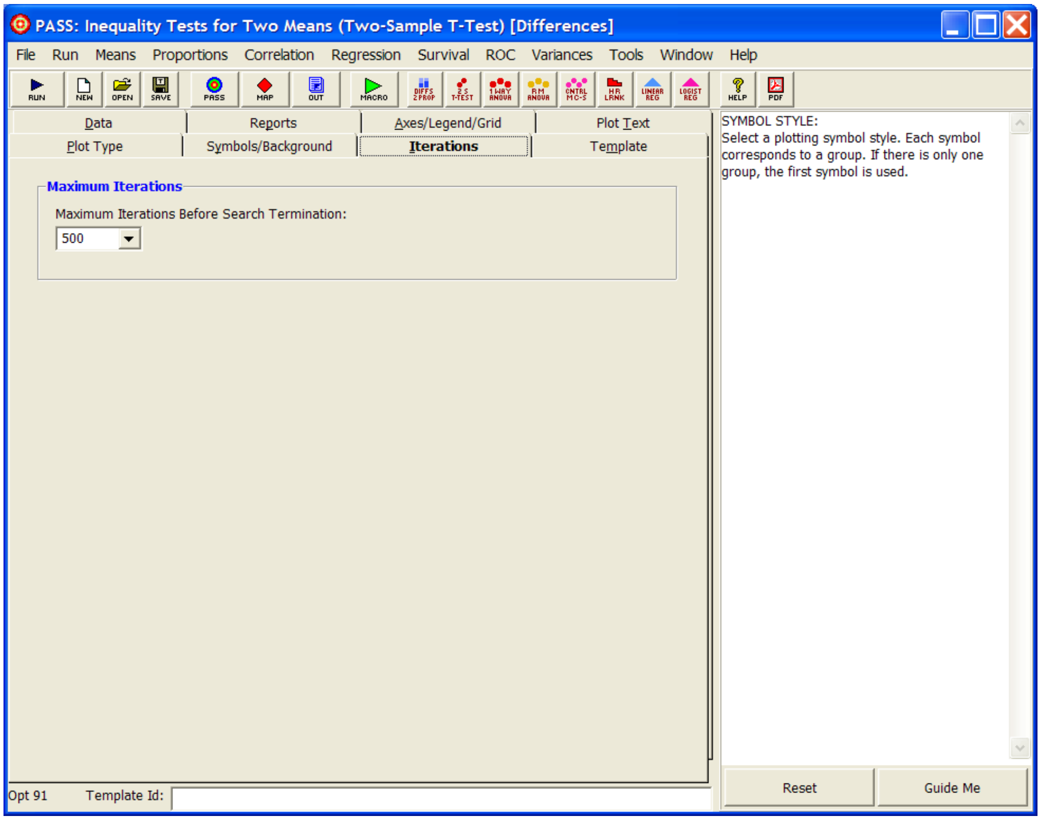
Background Colors and Styles

Color and Style

These options specify the style and color of the plot background.

Iterations Tab

The Iterations tab presents the parameters that control the searching process.



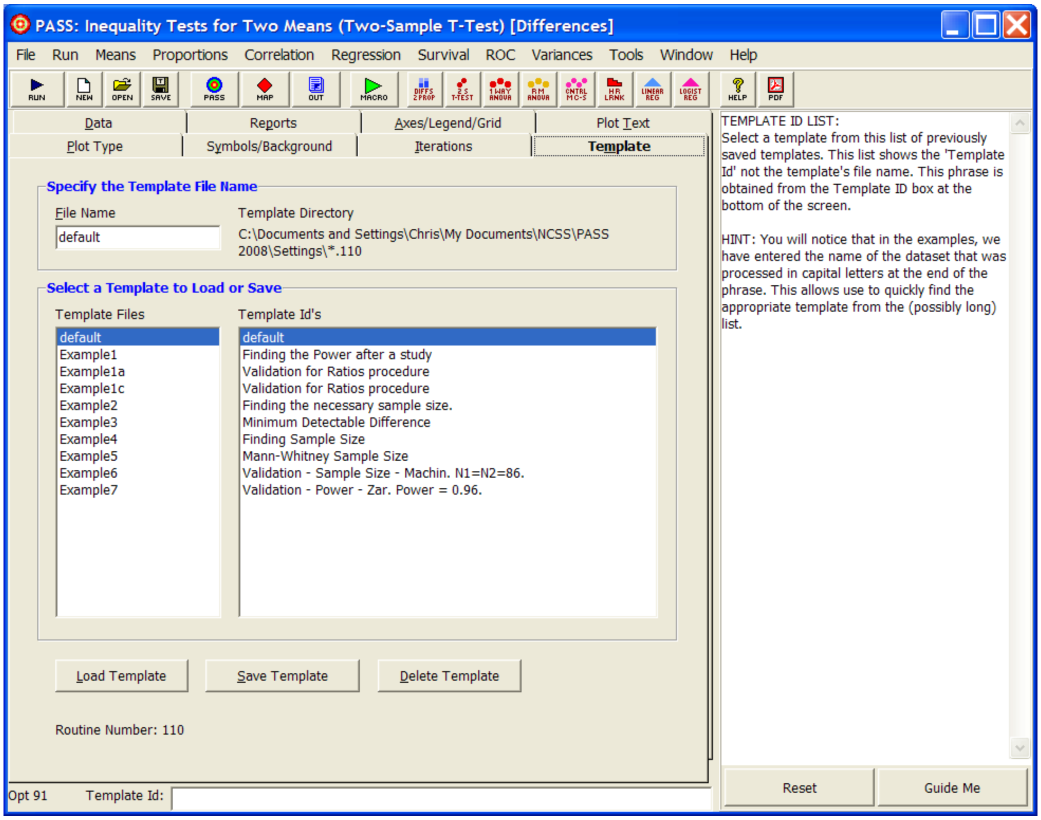
Maximum Iterations

Maximum Iterations Before Search Termination

This option specifies the maximum number of iterations before a search for the parameter of interest is halted. When the maximum number of iterations is reached without convergence, the criterion is left blank. We recommend that at least 500 iterations be specified.

Template Tab

The options on this panel allow various sets of options to be loaded (File menu: Load Template) or stored (File menu: Save Template). A template file contains all the settings for this procedure.



Specify the Template File Name

File Name

Designate the name of the template file either to be loaded or stored.

Select a Template to Load or Save

Template Files

A list of previously stored template files for this procedure.

Template Id's

A list of the Template ID's of the corresponding files. This ID value is loaded in the box at the bottom of the panel.

Template Action Buttons

Load Template

To load a template file, select it from the list of files given in the Template Files box. Once the desired file is selected, press the Load Template button to load the template.

Save Template

This option saves the current option settings to the template file that is currently specified in the File Name box. You may also enter an identifying phrase in the box at the bottom of the window since this will be displayed along side of the file names. The template files are stored in the folder specified under Template Directory (*C:\...\[My] Documents\NCSS\PASS 2008\Settings*). You can erase any unwanted template files by deleting them from this folder using the Windows Explorer program.

The template files for each procedure have different file name extensions. Thus, you can use the same name for a template saved from the T-Test procedure as for a template saved from the Multiple Regression procedure. For example, if the Save Template command is issued in the window shown previously, the current settings will be saved in a file called *default.110* in the

Note that there is no automatic connection between the template in memory and the copy on the disk. If you want to save the changes you have made to a template, you must use the Save Template button to save them.

Delete Template

Move the currently selected template file to the Windows Recycle Bin from which it will be automatically deleted the next time the Windows Recycle Bin is emptied. If you wish to undo the delete operation, move the file back to the ***PASS*** Settings directory (*C:\...\[My] Documents\NCSS\PASS 2008\Settings*) from the Windows Recycle Bin.

Chapter 5

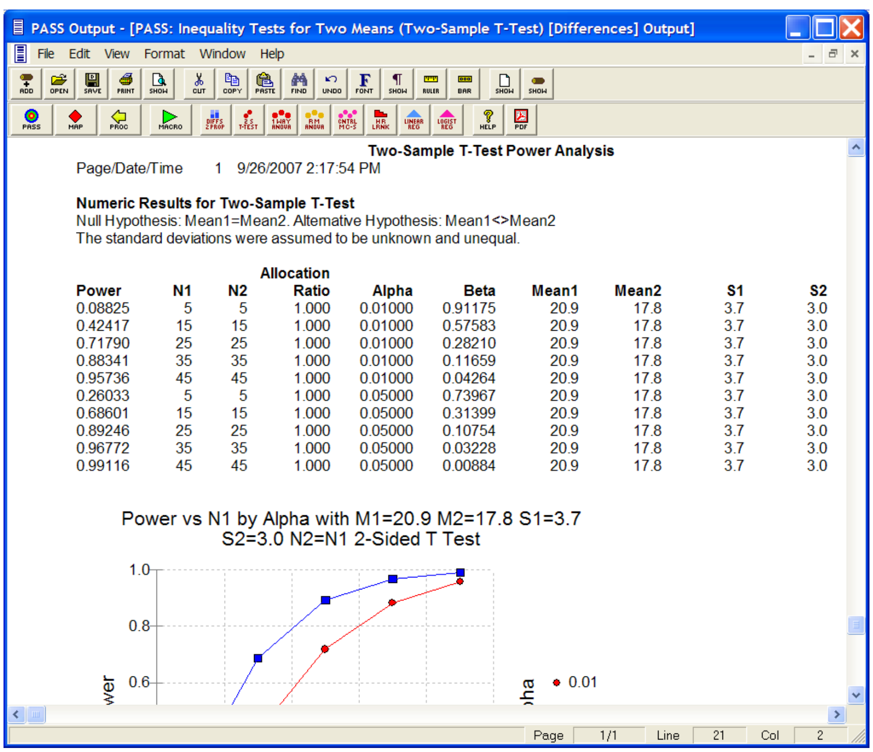
The Output Window

Introduction

PASS sends all statistics and graphics output to its built-in word processor from where they can be viewed, edited, printed, or saved. Reports and graphs are saved in rich text format (RTF). Since RTF is a standard document transfer format, these files may be loaded directly into your word processor for further processing. This chapter covers the basics of the built-in word processor in ***PASS***.

Viewing the Output

The output of the Example 1 template of the Two-Sample T-Test procedure is shown below. The output window is in full-screen mode. The screen will look similar to this. Note that the actual size of your screen depends on the resolution of your monitor, so it may vary.



Power	N1	N2	Allocation Ratio	Alpha	Beta	Mean1	Mean2	S1	S2
0.08825	5	5	1.000	0.01000	0.91175	20.9	17.8	3.7	3.0
0.42417	15	15	1.000	0.01000	0.57583	20.9	17.8	3.7	3.0
0.71790	25	25	1.000	0.01000	0.28210	20.9	17.8	3.7	3.0
0.88341	35	35	1.000	0.01000	0.11659	20.9	17.8	3.7	3.0
0.95736	45	45	1.000	0.01000	0.04264	20.9	17.8	3.7	3.0
0.26033	5	5	1.000	0.05000	0.73967	20.9	17.8	3.7	3.0
0.68601	15	15	1.000	0.05000	0.31399	20.9	17.8	3.7	3.0
0.89246	25	25	1.000	0.05000	0.10754	20.9	17.8	3.7	3.0
0.96772	35	35	1.000	0.05000	0.03228	20.9	17.8	3.7	3.0
0.99116	45	45	1.000	0.05000	0.00884	20.9	17.8	3.7	3.0

Documents

The ***PASS*** word processor maintains two documents: *Output* and *Log*. Although both of these documents allow you to view your data, the *Output* document serves as a viewer while the *Log* document serves as a recorder.

You can load additional documents as well. For example, you might want to view the output from a previous analysis to compare the results with the current analysis. To do this, you open a third document that is actually the log file from a previous analysis.

All ***PASS*** documents are stored in the RTF format. This is a common format that is used by most word processors, including MS-Word and MS-Write. When you save a ***PASS*** report, you will be able to load that report directly into your own word processor. All text, formatting, and graphics will appear in your word processor ready for further editing. You can then save the document in your word processor's native format. In this way, you can easily transfer the output of a ***PASS*** procedure to almost any format you desire.

Output Document

The Output document displays the output report from the current analysis. Whenever you run a ***PASS*** procedure, the resulting reports and graphs are displayed in the Output document. Each new run clears the existing Output document, so if you want to save a report, you must do so before running the next report.

The Output document provides four main functions: display, print, save to the Log document, and save as an RTF file.

Log Document

The Log document provides a place to store a permanent record of your analysis. Since the Output document is erased by each new analysis, you need a place to store your permanent work. The Log document serves this purpose. When you have a report or graph that you want to keep, copy it from the Output document to the Log document.

The Log document provides four main word processing functions: load, display and edit, print, and save. When you load a file into the Log document, you can add new output to it. In this way, you can record your work on a project in a single file, even though your work on that project is spread out over several days.

Menus

The menus provide a convenient way to transfer from module to module within the ***PASS*** system. Each set of menus will be briefly described here.

File Menu

The File Menu is used for opening, saving, and printing ***PASS*** word processor files. All options apply to the currently active document (the document whose title bar is selected). We will now discuss each of the options on this menu.

New

This option opens an empty document. You might use this when you want to make notes about your analysis.

New Log

This option opens an empty log document. You might use this when you want to start a new project.

Open

This option opens an existing file. When this item is selected, the Open Report File dialog box appears. Note that no connection is maintained between a loaded file and its image on the disk. If you make changes to a file, you must save those changes to the disk.

Open Log

This option opens an existing log file. When this item is selected, the Open Report File dialog box appears. The requested file is loaded into the Log document. Note that no connection is maintained between a loaded file and its image on the disk. If you make changes to a file, you must save those changes to the disk.

You might use this option when you want to continue using a certain file as the Log file.

Toggle Auto-Log

When Auto-Log is on, the contents of the Output document are automatically copied to the Log document. The Output document remains unchanged. If you want to keep a copy of all the output that has been placed in the Log document, you will still need to save it manually.

This function allows you to automatically save all output for further use.

Add Output to Log

Selecting this option copies the contents of the Output document to the Log document. The Output document remains unchanged. This allows you to save the current output document for further use.

Save As

This option lets you save the contents of the currently active document to a designated file using the RTF format. Note that only the active document is saved. Also note that all file names should have the “RTF” extension so that other systems can recognize their format.

Printer Setup

This option allows you to set printing options on your printer.

Print Preview

This option displays the output report as it will appear on the printed page. Use it to preview your report before printing it out.

Print

This option lets you print the entire document or a range of pages. When you select this option, a Print Dialog box will appear that lets you control which pages are printed.

Close Output Window

Minimizes the document that is currently being viewed. Note that this option does not clear the document, it simply minimizes it.

Exit PASS

This option exits the ***PASS*** system. All documents, spreadsheets, and procedure windows are closed.

Edit Menu

This menu contains options that let you edit a document.

Undo

This item reverses the last edit action. It is particularly useful for replacing something that was accidentally deleted.

Cut

This item copies the currently selected text to the Windows clipboard and erases it from the document. You can paste the information from the clipboard to a different location in the current document, into another document, into a datasheet in the spreadsheet, or into another application. The selected text is erased.

Copy

This item copies the currently selected text from the document to the Windows clipboard. You can paste this information from the clipboard to a different location in the current document, into another document, into a datasheet in the spreadsheet, or into another application. The selected text is not modified.

Paste

This item copies the contents of the clipboard to the current document at the insertion point. This command is especially useful for moving selected information from the Output document to the Log document.

Select All

This item selects the entire document. Although you can select a portion of the document using the mouse or a shift-arrow key, this is much faster if you want to select the entire document.

Toggle Page Break

Changes the status of the page break on the line at which the insertion point resides. If a page break exists (shown by a horizontal line), it is removed. If a page break does not currently exist at that point, one is added.

Note that ***PASS*** does not repaginate your document for you. Once you make changes, it will be up to you to repaginate your document.

Find

This item opens the Search dialog box. You can specify text that you want to search for. This is especially useful when you are looking for a certain topic or data value in a large report.

Find Next

This item continues finding the text you entered in the Search Dialog box.

Replace

This item opens the Search and Replace dialog box. This allows you to make repetitive changes. For example, you might want to change the name of one of the variables to a more useful name.

Goto Section

This item does not modify the document. Instead, it lets you reposition the insertion point to one of the major topics. When ***PASS*** runs a procedure, it stores the major report topics in this list box. You can quickly position the view to a desired topic using this screen.

View Menu

This menu lets you designate which editing tools you want to use.

Ruler

This option controls whether the ruler and the tabs bar are displayed. The ruler displays the physical dimensions of the document. The tabs bar, found just below the ruler bar, lets you set the margins and tabs of your document. Only the currently selected part of your document is affected by a change in the tabs and margins.

Format Toolbar

This option controls whether the Format Toolbar is displayed. The function of each of the buttons is discussed below.

Status Bar

This option controls whether the Status Bar is displayed at the bottom of the output window.

Show All

Selecting this menu item causes the Ruler, Tabs Bar, Format Toolbar, and Status Bar to be displayed.

Hide All

Selecting this menu item causes the Ruler, Tabs Bar, Format Toolbar, and Status Bar to be hidden. This gives you more screen space to view your output.

Redraw

Occasionally the Output Window becomes cluttered. If this happens, select this option to redisplay the output.

Format Menu

This menu lets you set the format for a selected block of text.

Font

This option displays the Replace Font dialog box, which lets you specify the font and style of the selected text.

Paragraph

This option displays the Paragraph dialog box, which lets you specify the tabs and margins of the selected text.

Format Markers

Indicates whether the (usually hidden) tab arrows and the end-of-paragraph marks are displayed in the document. Note that these characters are never printed.

Window Menu

This menu lets you designate how you want the documents arranged on the screen and which window you want displayed on top of your output desktop.

Cascade

This item arranges the documents in a cascading display from the upper left to the lower right of the screen.

Tile Horizontal

This item arranges the documents horizontally down the word processor window.

Tile Vertical

This item arranges the documents vertically across the word processor window.

Arrange Icons

When a document is minimized, it is represented as an icon at the bottom of the word processor window. This option arranges all document icons. It is usually applied when the word processor window has been resized.

Current Output

This item causes the Output window to be displayed.

Log

This item causes the Log window to be displayed.

PASS Home

This item causes the PASS Home window to be displayed.

Map – Quick Launch

This item causes the Map window to be displayed.

Help Menu

This menu controls the display of the serial numbers and help system.

PASS Help System

This item displays the help system.

About

This item displays the release date and version number of your software.

Serial Numbers

This item displays the ***PASS*** Registration window where your serial numbers can be modified.

Toolbars

A toolbar is a series of small buttons that appear just below the menus at the top of the procedure window. The output window has two toolbars. Each button on the toolbar provides quick access to a menu item.

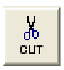
Add Output to Log. This button copies the contents of the Output document to the Log document. The Output document remains unchanged. This allows you to save the current output document for further use.

Open Log. This button opens an existing log file. When this item is selected, the Open Report File dialog box appears. The requested file is loaded into the Log document.

Save As. This button lets you save the contents of the currently active document to a designated file using the RTF format. Note that only the active document is saved. Also note that all file names should have the "RTF" extension so that other systems can recognize their format.

Print. This button lets you print the entire document or a range of pages. When you select this option, a Print Dialog box will appear that lets you control which pages are printed.

Print Preview. This button displays the output report as it will appear on the printed page. Use it to preview your report before printing it out.

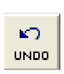


Cut. This button copies the currently selected text to the Windows clipboard and erases it from the document. You can paste the information from the clipboard to a different location in the current document, into another document, into a datasheet in the spreadsheet, or into another application. The selected text is erased.

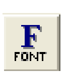
Copy. This button copies the currently selected text from the document to the Windows clipboard. You can paste this information from the clipboard to a different location in the current document, into another document, into a datasheet in the spreadsheet, or into another application. The selected text is not modified.

Paste. This button copies the contents of the clipboard to the current document at the insertion point. This command is especially useful for moving selected information from the Output document to the Log document.

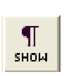
Find. This button opens the Search dialog box. You can specify text that you want to search for. This is especially useful when you are looking for a certain topic or data value in a large report.



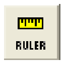
Undo. This button reverses the last edit action. It is particularly useful for replacing something that was accidentally deleted.



Font. This button displays the Replace Font dialog box, which lets you specify the font and style of the selected text.

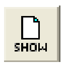


Format Marks. This button is used to toggle the display of the tab arrows and the end-of-paragraph marks in the document. Note that these characters are never printed.



Ruler. This button controls whether the ruler and the tabs bar are displayed. The ruler displays the physical dimensions of the document. The tabs bar, found just below the ruler bar, lets you set the margins and tabs of your document. Only the currently selected part of your document is affected by a change in the tabs and margins.

Format Bar. This button controls whether the Format Toolbar is displayed.



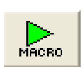
Show Output. This button causes the Output window to be displayed.

Show Log. This button causes the Log window to be displayed.

PASS Home. This button causes the PASS Home window to be displayed.

Map. This button causes the PASS Map (Quick Launch) window to be displayed. This window allows you to quickly select any procedure using icon buttons. This window can also be used to change the procedure quick-access buttons in the toolbar.

Back to Procedure. This button displays the procedure window used to create the current output. This allows you to quickly move between the procedure and the output windows.



Macro. This button can be used to interface with the macro system. Left-click on this button to run the active macro. Hold your mouse over the button to display the active macro name. Right-click on this button to load the Macro Command Center window.

Quick-Access. These buttons show up on all toolbars throughout the ***PASS*** system. Clicking on them with the left mouse button will display the corresponding procedure. Clicking on any of these buttons with the right mouse button allows you to change the procedure assigned to each button.

Help. This button loads the ***PASS*** Help System at the appropriate topic.

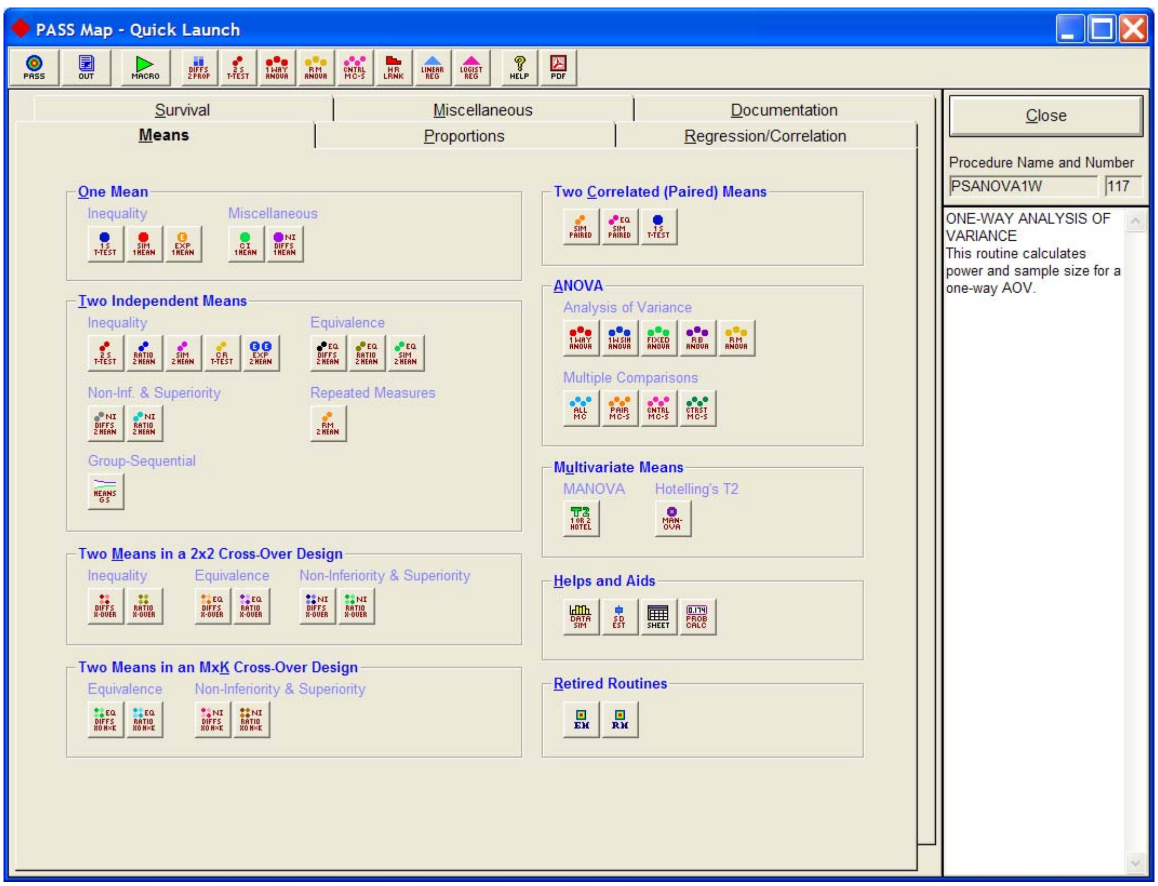
Printable PDF. This button loads the appropriate printable PDF chapter.

Chapter 6

The Map (Quick Launch) Window

Introduction

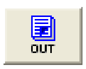
The Map or Quick Launch window allows you to easily navigate to any ***PASS*** procedure using icons. Click on the tabs to view the associated procedure icons. The Map window drag-and-drop feature allows you to easily edit the eight quick-access buttons that appear on all ***PASS*** toolbars.



Toolbar

A toolbar is a series of small buttons that appear just below the menus at the top of the procedure window. Each button on the Map toolbar provides quick access to a ***PASS*** procedure, window, or help file.

PASS Home. This button causes the PASS Home window to be displayed.



Output. This button causes the output window to be displayed.

Macro. This button can be used to interface with the macro system. Left-click on this button to run the active macro. Hold your mouse over the button to display the active macro name. Right-click on this button to load the Macro Command Center window.

Quick-Access. These buttons show up on all toolbars throughout the ***PASS*** system. Clicking on them with the left mouse button will display the corresponding procedure. Clicking on any of these buttons with the right mouse button allows you to change the procedure assigned to each button.

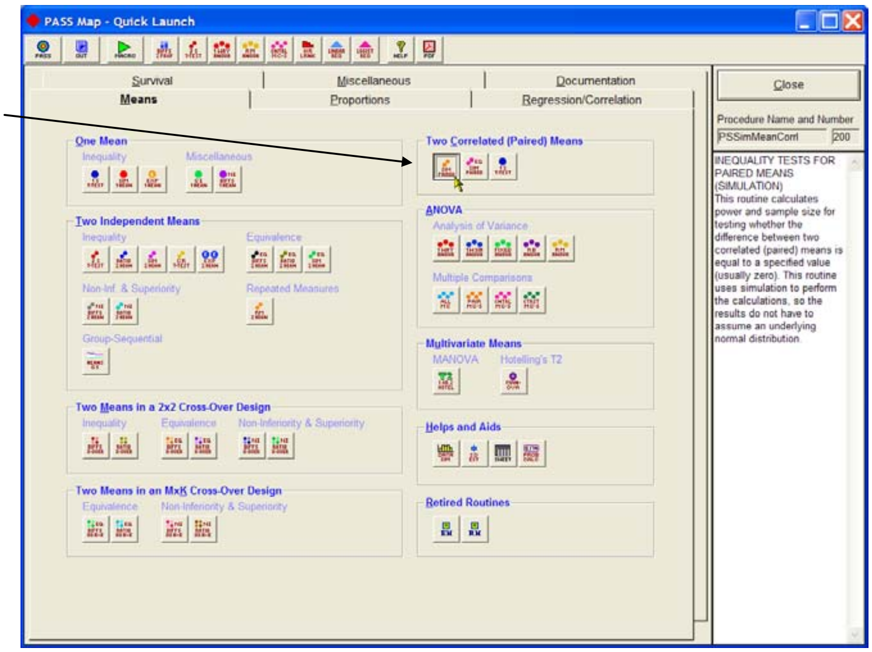
Help. This button loads the ***PASS*** Help System at the appropriate topic.

Printable PDF. This button loads the appropriate printable PDF chapter.

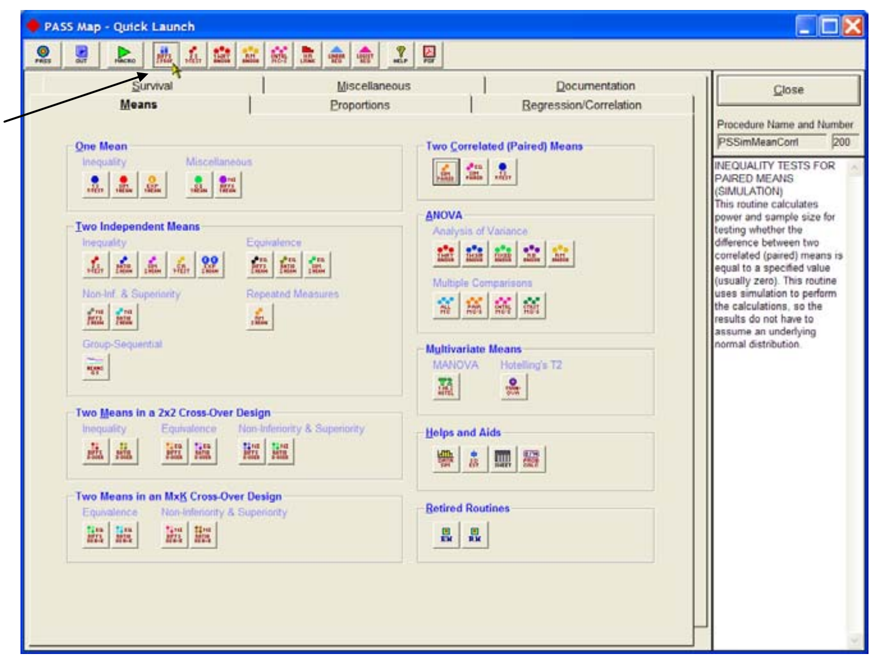
Customizing the Toolbar using Drag-and-Drop

The eight quick-access procedure buttons that show up on all toolbars throughout the program may be changed using the PASS Home Window, the Map (Quick Launch), any procedure window, or the output window. To change the procedures available in the toolbar using the Map window, drag and drop any procedure icon to replace any quick-access button in the toolbar at the top of the window.

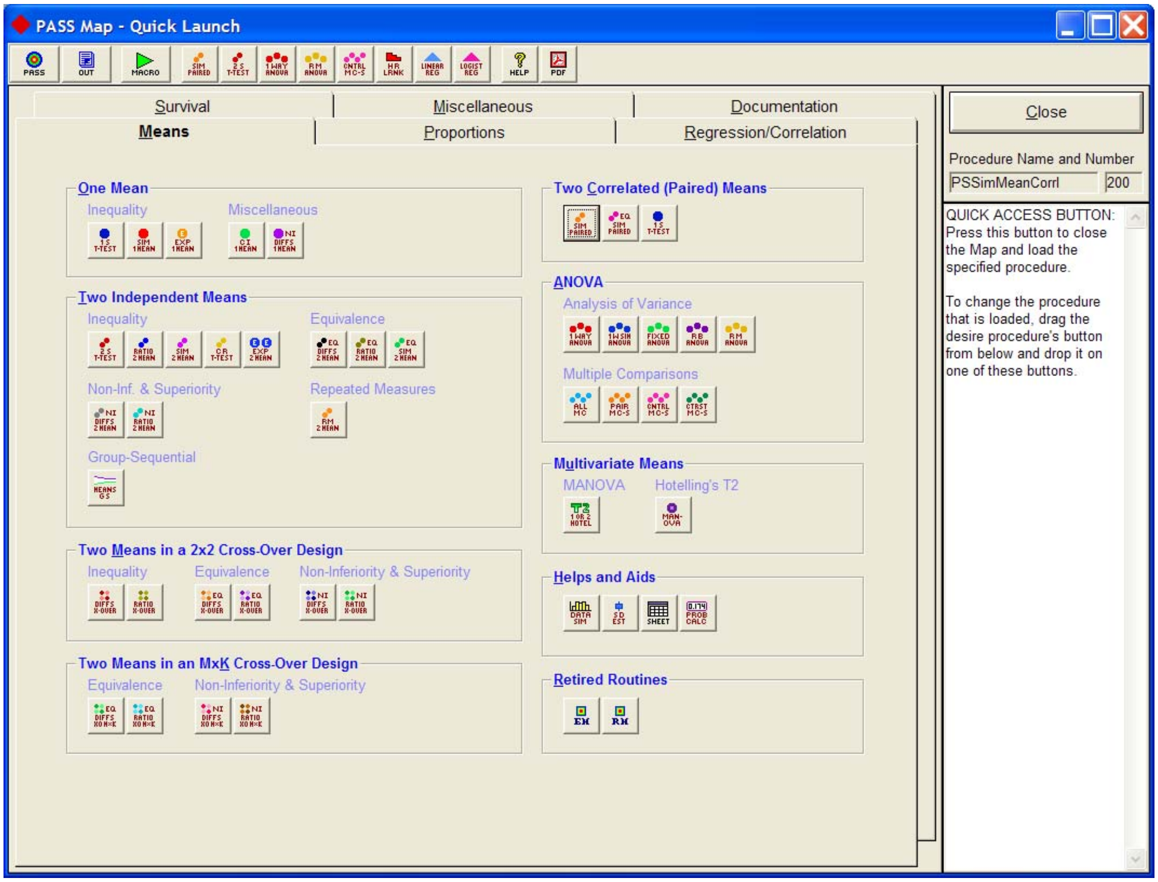
1. **Left-click** on any procedure icon on the Map window and hold down the mouse button.



2. While still holding down the left mouse button, drag the icon to the toolbar to replace any of the eight quick-access button icons.



The toolbar will now display the new icon. This new icon will appear in all toolbars throughout the ***PASS*** system.



Chapter 7

Introduction to Power Analysis

Overview

A statistical test's *power* is the probability that it will result in statistical significance. Since statistical significance is the desired outcome of a study, planning to achieve high power is of prime importance to the researcher. Because of its complexity, however, an analysis of power is often omitted.

PASS calculates statistical power and determines sample sizes. It does so for a broad range of statistical techniques, including the study of means, variances, proportions, survival curves, correlations, bioequivalence, analysis of variance, log rank tests, multiple regression, and contingency tables.

PASS was developed to meet several goals, including ease of learning, ease of use, accuracy, completeness, interpretability, and appropriateness. It lets you study the influence of sample size, effect size, variability, significance level, and power on your statistical analysis.

Brief Introduction to Power Analysis

Statistical power analysis must be discussed in the context of *statistical hypothesis testing*. Hence, this discussion starts with a brief introduction to statistical hypothesis testing, paying particular attention to topics that relate to power analysis and sample size determination. Although the theory behind hypothesis testing is general, its concepts can be reviewed by discussing simple case: testing whether a proportion is greater than a known standard.

Following the usual terminology of statistical hypothesis testing, define two complementary hypotheses

$$H_0: P \le P0 \text{ vs. } H_1: P > P0$$

where P is the response proportion in the population of interest and $P0$ is the known standard value.

H_0 is called the *null hypothesis* because it specifies that the difference between the two proportions is zero (null).

H_1 is called the *alternative hypothesis*. This is the hypothesis of interest to us. Our motivation for conducting the study is to provide evidence that the alternative (or research) hypothesis is true. We do this by showing that the null hypothesis is unlikely—thus establishing that the alternative hypothesis (the only possibility left) is likely.

Outcomes from a statistical test may be categorized as follows:

1. Reject H_0 when H_0 is true. That is, conclude that H_0 is unlikely when it is true. This constitutes a decision error known as the *Type-I error*. The probability of this error is alpha (α) and is often referred to as the *significance level* of the hypothesis test.
2. Do not reject H_0 when H_0 is false. That is, conclude that H_0 is likely when it is false. This constitutes a decision error known as the *Type-II error*. The probability of this error is beta (β). *Power* is $1-\beta$. It is the probability of rejecting H_0 when it is false.
3. Reject H_0 when H_0 is false. This is a correct decision.
4. Do not reject H_0 when H_0 is true. This is also a correct decision.

The basic steps in conducting a study that is analyzed with a hypothesis test are:

1. Specify the statistical hypotheses, H_0 and H_1.
2. Run the experiment on a given number of subjects.
3. Calculate the value of a test statistic, such as the sample proportion.
4. Determine whether the sample values favor H_0 or H_1.

Binomial Probability Table

In the current example, suppose that a random sample of ten individuals is selected, i.e. *N* = 10. The number of individuals, *R*, with the characteristic of interest is counted. Hence, *R* is the test statistic. A table of binomial probabilities gives the probability that *R* takes on each of its eleven possible values for various values for *P*.

	P								
R	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0	0.349	0.107	0.028	0.006	0.001	0.000	0.000	0.000	0.000
1	0.387	0.376	0.121	0.040	0.010	0.002	0.000	0.000	0.000
2	0.194	0.302	0.233	0.121	0.044	0.011	0.001	0.000	0.000
3	0.057	0.201	0.267	0.215	0.117	0.042	0.009	0.001	0.000
4	0.011	0.088	0.200	0.251	0.205	0.111	0.037	0.006	0.000
5	0.001	0.026	0.103	0.201	0.246	0.201	0.103	0.026	0.001
6	0.000	0.006	0.037	0.111	0.205	0.251	0.200	0.088	0.011
7	0.000	0.001	0.009	0.042	0.117	0.215	0.267	0.201	0.057
8	0.000	0.000	0.001	0.011	0.044	0.121	0.233	0.302	0.194
9	0.000	0.000	0.000	0.002	0.010	0.040	0.121	0.376	0.387
10	0.000	0.000	0.000	0.000	0.001	0.006	0.028	0.107	0.349

Let us discuss in detail the interpretation of the values in this table for the simple case in which a coin is flipped ten times and the number of heads is recorded. The column parameter *P* is the probability of obtaining a head on any one toss of the coin. When dealing with coin tossing, one would usually set *P* = 0.5, but this does not have to be the case. The row parameter *R* is the number of heads obtained in ten tosses of a coin.

The body of the table gives the probability of obtaining a particular value of *R*. One way to interpret this probability value is as follows: conduct a simulation in which this experiment is repeated a million times for each value of *P*. Using the results of this simulation, calculate the proportion of experiments that result in each value of *R*. This proportion is recorded in this table. For example, when the probability of obtaining a head on a single toss of a coin is 0.5, ten flips of a coin would result in five heads 24.6% of the time. That is, as the procedure is repeated (flipping a coin ten times) over and over, 24.6% of the outcomes would be five heads.

Calculating the Significance Level, Alpha

We will now explain how the above table is used to set the significance level (the probability of a type-I error) to a pre-specified value. Recall that a type-I error occurs when an experiment results in the rejection of the null hypothesis when, in fact, the null hypothesis is true. By studying the table, the impact of using different *rejection regions* can be determined. A rejection region is a simple rule that states which values of the test statistic will result in the null hypothesis being rejected.

For example, suppose we want to test *P*0 = 0.5. That is, the null hypothesis is that *P* = 0.5 and the alternative hypothesis is that *P* > 0.5. Suppose the rejection region is *R* equal to 8, 9, or 10. That is, H_0 is rejected if *R* = 8, 9, or 10. From the above table, the probability of obtaining 8, 9, or 10 heads in 10 tosses when *P* = 0.5 is calculated as follows:

$$\begin{aligned}\Pr(R=8,9,10 \mid P=0.5) &= \Pr(R=8 \mid P=0.5)+\Pr(R=9 \mid P=0.5)+\Pr(R=10 \mid P=0.5)\\ &= 0.044+0.010+0.001\\ &= 0.055\end{aligned}$$

That is, 5.5% of these coin tossing experiments using this decision rule result in a type-I error. By setting the rejection criterion to *R* = 8, 9, or 10, alpha has been set to 0.055.

It is extremely important to understand what alpha means, so we will go over its interpretation again. If the probability of obtaining a head on a single toss of a coin is 0.5, then 5.5% of the experiments that use the rejection criterion of *R* = 8, 9, or 10 will result in the false conclusion that *P* > 0.5.

The key features of this definition that are often overlooked by researchers are:

1. **The value of alpha is based on a particular value of *P*.** Note that we used the assumption "if the probability of obtaining a head is 0.5" in our calculation of alpha. Hence, if the actual value of *P* is 0.4, our calculations based on the assumption that *P* is 0.5 are wrong. Mathematicians call this a conditional probability since it is based on the condition that *P* is 0.5. Alpha is 0.055 if *P* is 0.5.

   Often, researchers think that setting alpha to 0.05 means that the probability of rejecting the null hypothesis is 0.05. Can you see what is wrong with this statement? They have forgotten to mention the key fact that this statement is based on the assumption that *P* is 0.5!

2. **Alpha is a statement about a proportion in multiple experiments.** Alpha tells us what percentage of a large number of experiments will result in 8, 9, or 10 heads. Alpha is a statement about what to expect from future experiments. It is not a statement about *P*. Occasionally, researchers conclude that the alpha level is the probability that *P* = 0.5. This is not what is meant. Alpha is not a statement about *P*. It is a statement about future experiments, given a particular value of *P*.

Interpreting *P* Values

The term *alpha value* is often used interchangeably with the term *p value*. Although these two terms are closely related, there is an important distinction between them. A *p* value is the largest value of alpha that would result in the rejection of the null hypothesis for a particular set of data. Hence, while the value of alpha is set during the planning of an experiment, the *p* value is calculated from the data after experiment has been run.

Calculating Power and Beta

We will now explain how to calculate the power. Recall that power is the probability of rejecting a false null hypothesis. A false H_0 means that *P* is some value other than *P*0. In order to compute power, we must know the actual value of *P*.

Returning to our coin tossing example, suppose the actual value of *P* is 0.7. What is the power and beta value of this testing procedure? The decision rule is to reject the null hypothesis when *R* is 8, 9, or 10. From the above probability table, the probability of obtaining 8, 9, or 10 heads in 10 tosses of a coin when probability of a head is actually 0.7 is

$$\begin{aligned}\Pr(R = 8,9,10 \mid P = 0.7) &= \Pr(R = 8 \mid P = 0.7) + \Pr(R = 9 \mid P = 0.7) + \Pr(R = 10 \mid P = 0.7) \\ &= 0.233 + 0.121 + 0.028 \\ &= 0.382\end{aligned}$$

This is the power. The value of a type-II error is 1.000 - 0.382, which is 0.618. That is, if *P* is 0.7, then 38.2% of these coin tossing experiments will reject H_0, while 61.8% of them will result in a type-II error.

It is extremely important to understand what beta means, so we will go over its interpretation again. If the probability of obtaining a head on the toss of a coin is 0.7, then 61.8% of the experiments that use the rejection criterion of *R* = 8, 9, or 10 will result in the false conclusion that *P* = 0.5.

The key features of this definition that are often overlooked by researchers are:

1. **The value of beta is based on a particular value of *P*.** Note that we used the assumption "if the probability of obtaining a head is 0.7" in our calculation of beta. Hence, if the actual value of *P* is 0.6, our calculation based on the assumption that *P* was 0.7 is wrong.
2. **Beta is a statement about the proportion of experiments.** Beta tells us what percentage of a large number of experiments will result in 8, 9, or 10 heads. Beta is a statement about what we can expect from future experiments. It is not a statement about *P*.
3. **Beta depends on the value of alpha.** Since the rejection region (8, 9, or 10 heads) depends on the value of alpha, beta depends on alpha.
4. **You cannot make both errors at the same time.** A type-II error can only occur when a type-I error did not occur, and vice versa.

Specifying Alternative Values of the Parameters

We have noted a great deal of confusion about specifying the values of the parameters under the alternative hypothesis. The alternative hypothesis is usually that the value of one parameter is different from another. The hypothesis does not usually specify how different. It simply gives the direction of the difference. The power is calculated at specified alternative values. These values should be considered as values at which the power is calculated, not as the true value.

Effect Size

The *effect size* is the size of the change in the parameter of interest that can be detected by an experiment. For example, in the coin tossing example, the parameter of interest is *P*, the probability of a head. In calculating the sample size, we would need to state what the baseline probability is (probably 0.5) and how large of a deviation from *P* that we want to detect with our experiment. We would expect that it would take a much larger sample size to detect a deviation of 0.01 than it would to detect a deviation of 0.40.

Selecting an appropriate effect size is difficult because it is subjective. The question that must be answered is: what size change in the parameter would be of interest? Note that, in power analysis, the effect size is not the actual difference. Instead, *the effect size is the change in the parameter that is of interest* or is to be detected. This is a fundamental concept that is often forgotten after the experiment is run.

After an experiment is run that leads to non-significance, researchers often ask, “What was the experiment’s power?” and “How large of a sample size would have been needed to detect significance?” To compute the power or sample size, they set the effect size equal to the amount that was seen in their experiment. This is incorrect. *When performing a power analysis after an experiment has completed, the effect size is still the change in the parameter that would be of interest to other scientists*. It is not the change that was actually observed!

Often, the effect size is stated as a percentage change rather than an absolute change. If this is the case, you must convert the percentage change to an absolute change. For example, suppose that you are designing an experiment to determine if tossing a particular coin has exactly a 50% chance of yielding a head. That is, *P*0 is 0.50. Suppose your gambling friends are interested in whether a certain coin has a 10% greater chance. That is, they are concerned with the case where *P* is 0.55 or greater. The effect size is |0.50 - 0.55| or 0.05.

Types of Power Analyses

There are several types of power analyses. Often, power analysis is performed during the design phase of a study to determine the sample size. This type of study would determine the value of *N* for set values of alpha and beta. Another type of power analysis is a post hoc analysis, which is done after the study is concluded. A post hoc analysis studies such questions as:

1. What sample size would have been needed to detect a specific effect size?
2. What is the smallest effect size that could be detected with this sample size?
3. What was the power of the test procedure?

These and similar questions may be answered using power analysis. By considering these kinds of questions after a study is concluded, you can gain important insights into how to make your research more efficient and effective.

Nuisance Parameters

Statistical hypotheses usually make statements about one or more parameters from a set of one or more probability distributions. Often, the hypotheses leave other parameters of the probability distribution unspecified. These unspecified parameters are called ‘nuisance’ parameters.

For example, a common clinical hypothesis is that the response proportions of two drugs are equal. The null hypothesis is that the difference between these two drugs is zero. The alternative is that the difference is non-zero. Note that the actual values of the two proportions are not stated in the hypothesis—just their difference. The actual values of the proportions will be needed to compute the power. That is, different powers will result for the case when $P1 = 0.05$ and $P2 = 0.25$ and for the case $P1 = 0.50$ and $P2 = 0.70$. In this example, the proportion difference ($D = P1 – P2$) is the parameter of interest. The baseline proportion, $P1$, is a nuisance parameter.

Another example of a nuisance parameter occurs when using the t-test to test whether the mean is equal to a particular value. When computing the power or sample size for this test, the hypothesis specifies the value of the mean. However, the value of the standard deviation is also required. In this case, the standard deviation is a nuisance parameter.

When performing a power analysis, you should state all your assumptions, including the values of any nuisance parameters that were used. When you do not have any idea as to reasonable values for nuisance parameters, you should use a range of possible values so that you can analyze how sensitive the results are to the values of the nuisance parameters. Also, do not be tempted to use the nuisance parameter’s value from a previous (or pilot) study. Instead, a reasonable strategy is to compute a confidence interval and use the confidence limit that results in the largest sample size.

Choice of Test Statistics

Many hypothesis tests can be tested with a variety of test statistics. For example, statisticians often have to decide between the t-test and the Wilcoxon test when testing means. Similarly, when testing whether two proportions are equal, they have to decide whether to use a z-test or an exact test. If they choose a z-test, they have to decide whether to apply a continuity correction.

In most cases, each test statistic will have a different power. Thus, it should be obvious that *you must compute the power of the test statistic that will be used in the analysis*. A sample size based on the t-test will not be accurate for a nonparametric test.

The next question is usually “Which test statistic should I use?” You might say “They one that requires the smallest sample size.” However, other issues besides power must be considered. For example, consideration must be given to whether the assumptions of the test statistic will be met be the data. If your data is binary, it is probably unreasonable to assume that they are continuous.

These are simple principles, but they are often overlooked.

Types of Hypotheses

Hypothesis tests work this way. If the null hypothesis if rejected, the alternative hypothesis is concluded to be true. However, if null hypothesis is not rejected, no conclusion is reached--the null hypothesis *is not* concluded to be true. The only way that a conclusion is reach is if the null hypothesis is rejected.

Because of this, it is very important that the null and alternative hypotheses be constructed so that the conclusion of interest is associated with the alternative hypothesis. That way, if the null hypothesis is rejected, the study reaches the desired conclusion.

There are several types of hypotheses. These include inequality, equivalence, non-inferiority, and superiority hypotheses. In the statistical literature, these terms are used with completely different meanings, so it is important to define what is meant by each. We have tried to adopted names that are associated with the alternative hypothesis, since this is the hypothesis of interest.

It is important to note that even though two sets of hypotheses may be similar, they often will have different power and sample size requirements. For example, an equivalence test (see below) appears to be the simple reverse of a two-sided test of inequality, yet the equivalence test requires a much larger sample size to achieve the same power as the inequality test. Hence, you cannot select the sample size for an inequality test and then later decide to run an equivalence test.

Each of the sections to follow will give a brief definition along with an example based on the difference between two proportions.

Inequality Hypothesis

The term 'inequality' is represents the classical one-sided and two-sided hypotheses in which the alternative hypothesis is simply that the two values are unequal. These hypotheses are called tests of superiority by Julious (2004), emphasizing the one-sided versions.

Two-Sided

When the null hypothesis is rejected, the conclusion is simply that the two parameters are unequal. No statement is made about how different. For example, 0.501 and 0.500 are unequal, as are 0.500 and 0.800. Obviously, even though the former are different, the difference is not large enough to be of practical importance in most situations.

$$H_0: p_1 - p_2 = 0 \text{ vs. } H_1: p_1 - p_2 \neq 0 \text{ or } \mathrm{H}_1: p_1 - p_2 < 0 \text{ or } p_1 - p_2 > 0$$

One-Sided

These tests offer a little more information than the two sided tests since the direction of the difference is given. Again, no indication is made about how much higher (or lower) the superior value is to the inferior.

$$H_0: p_1 - p_2 \leq 0 \text{ vs. } H_1: p_1 - p_2 > 0 \text{ or } H_0: p_1 - p_2 \geq 0 \text{ vs. } H_1: p_1 - p_2 < 0$$

Non-Inferiority Hypothesis

These tests are a special case of the one-sided inequality tests. The term 'non-inferiority' is used to indicate that one treatment is not worse than another treatment. That is, one proportion is not less than another proportion by more than a trivial amount called the 'margin of equivalence'.

For example, suppose that a new drug is being developed that is less expensive and has fewer side effects than the standard drug. Producers must show that its effectiveness is no worse than the drug it is to replace.

When testing two proportions in which a higher proportion is better, the non-inferiority of treatment 1 as compared to treatment 2 is expressed as

$$\mathrm{H}_0\!: p_1 - p_2 \le -\delta \text{ vs. } \mathrm{H}_1\!: p_1 - p_2 > -\delta \text{ or } \mathrm{H}_0\!: p_1 \le p_2 - \delta \text{ vs. } \mathrm{H}_1\!: p_1 > p_2 - \delta,$$

where $\delta > 0$ is called the margin of equivalence. Note that when H_0 is rejected, the conclusion is that the first proportion is not less than the second proportion by more than δ.

Perhaps an example will help introduce this type of test. Suppose that the current treatment for a disease works 70% of the time. Unfortunately, this treatment is expensive and occasionally exhibits serious side-effects. A promising new treatment has been developed to the point where it can be tested. One of the first questions that must be answered is whether the new treatment is as good as the current treatment. In other words, do at least 70% of subjects respond to the new treatment?

Because of the many benefits of the new treatment, clinicians are willing to adopt the new treatment even if it is slightly less effective than the current treatment. They must determine, however, how much less effective the new treatment can be and still be adopted. Should it be adopted if 69% respond? 68%? 65%? 60%? There is a percentage below 70% at which the difference between the two treatments is no longer considered ignorable. After thoughtful discussion with several clinicians, it is decided that if a response of at least 63% is achieved, the new treatment will be adopted. The difference between these two percentages is called the *margin of equivalence*. The margin of equivalence in this example is 7% (which is ten percent of the original 70%).

The developers must design an experiment to test the hypothesis that the response rate of the new treatment is at least 0.63. The statistical hypothesis to be tested is

$$H_0\!: p_1 - p_2 \le -0.07 \text{ versus } H_1\!: p_1 - p_2 > -0.07.$$

Notice that when the null hypothesis is rejected, the conclusion is that the response rate is at least 0.63. Note that even though the response rate of the current treatment is 0.70, the hypothesis test is about a response rate of 0.63. Also, notice that a rejection of the null hypothesis results in the conclusion of interest.

Superiority Hypothesis

These tests are a special case of the one-sided inequality tests. The term ‘superiority’ is used to indicate that one treatment is better than another by more than a trivial amount called the ‘margin of equivalence’. For example, suppose that a new drug is being developed that is thought to have superior performance to the existing drug. Producers must show that its effectiveness is better than the drug it is to replace.

When testing two proportions in which a higher proportion is better, the superiority of treatment 1 over treatment 2 is expressed as

$$\mathrm{H}_0\!: p_1 - p_2 \le \delta \text{ vs. } \mathrm{H}_1\!: p_1 - p_2 > \delta \text{ or } \mathrm{H}_0\!: p_1 \le p_2 + \delta \text{ vs. } \mathrm{H}_1\!: p_1 > p_2 + \delta,$$

where $\delta > 0$ is called the margin of equivalence. Note that when H_0 is rejected, the conclusion is that the first proportion is higher than the second proportion by more than δ.

Equivalence Hypothesis

The term ‘equivalence’ is used here to represent tests designed to show that response rates of two treatments do not differ by more than a trivial amount called the ‘margin of equivalence’. These tests are the reverse of the two-sided inequality test.

The typical set of hypotheses are

$$\mathrm{H}_0: p_1 - p_2 \leq \delta_L \text{ or } p_1 - p_2 \geq \delta_U \text{ vs. } \mathrm{H}_1: \delta_L \leq p_1 - p_2 \leq \delta_U ,$$

where $\delta_L < 0$ and $\delta_U > 0$ are called the *equivalence limits*.

Suppose 70% of subjects with a certain disease respond to a certain drug. The company that produces the drug has decided to open a new facility in another city. They must show that the drug produced in the new facility is equivalent (all most the same) as that produced in existing facilities. After thoughtful discussion with several clinicians and regulatory agencies, it is decided that if the response rate of the drug produced at the new facility is between 65% and 75%, the new facility will go into production. In this case, the *margin of equivalence* is 5%.

The statistical hypothesis to be tested is

$$H_0: |p_1 - p_2| \geq 0.05 \text{ vs. } H_1: |p_1 - p_2| < 0.05 .$$

Chapter 8

Proportions

Introduction

This chapter introduces power analysis and sample size calculation for proportions. When the response is binary, the results for each group may be summarized as proportions. Usually, hypothesis tests are conducted to compare two proportions.

There are many issues that must be considered when designing experiments that use proportions. For example, will the proportions be analyzed directly, or will they be converted into differences, ratios, or odds ratios? Which test statistic will be used? Will the design use independent groups or will subjects be measured twice? Will the study have an active control?

The various answers to these and other questions result in a large number of situations. This chapter will introduce you to the issues that are common to all the tests of proportions.

Designs

There are several experimental designs for comparing two proportions. You can you a one-sample design to compare a sample proportion to a specific value. You can compare proportions from two independent samples using independent, stratified, cluster-randomized, or group-sequential designs. You can compare two correlated proportions. And finally, you can compare several categories in a contingency table.

Comparing Proportions

Two proportions may be compared by considering their difference, ratio, or odds ratio. Each of these cases may lead to different test statistics with different powers and sample size requirements.

Assume that p_1 is the response proportion of the experimental group and p_2 is the response proportion of the control (standard or reference) group. Mathematically, these alternative parameterizations are

Parameter	Computation
Difference	$\delta = p_1 - p_2$
Ratio	$\phi = p_1 / p_2$
Odds Ratio	$\psi = \dfrac{p_1 / (1 - p_1)}{p_2 / (1 - p_2)}$

Once you know p_1 and p_2, you can calculate any of these values, and you can easily change from once parameterization to another. Thus, the choice of which parameter you use may seem arbitrary. However, since different parameterizations lead to different test statistics, the choice can lead to a different power and sample size. These parameterizations will be discussed next.

Difference

The difference $\delta = p_1 - p_2$ is perhaps the most common method of comparing two proportions. This parameter is easy to interpret and communicate. It gives the absolute impact of the treatment. However, there are subtle difficulties that can arise with its use.

One difficulty occurs when the event of interest is rare. If a difference of 0.001 is reported for an event with a baseline probability of 0.40, we would dismiss this as being trivial. That is, there is usually little interest in a treatment that decreases the probability from 0.400 to 0.399. However, if the baseline probability of a disease is 0.002, a 0.001 decrease in the disease probability would represent a 50% reduction. Thus, the interpretation of the difference depends on the baseline probability of the event.

When planning studies, the value of p_2 is usually known and the power is calculated at various values of δ. The value of p_1 is then calculated using $p_1 = p_2 + \delta$. Because of the requirement that $0 < p_1 < 1$, the values of δ are restricted to the interval $-p_2 < \delta < 1 - p_2$, not $-1 < \delta < 1$ as you might expect. Likewise, the values of p_2 are restricted to $0 < p_2 < 1 - \delta$ if $\delta > 0$ and $-\delta < p_2 < 1$ if $\delta < 0$.

Because typical values of δ are usually between -0.2 and 0.2, the allowable values of p_2 are restricted to be between 0.2 and 0.8. When the values of p_2 are outside this range, it may be more convenient to use the odds ratio.

Ratio

The (risk) ratio, $\phi = p_1 / p_2$, gives the relative change in the probability of the outcome under each of the hypothesized values. This parameter is direct and easy to interpret. To compare the ratio with the difference, examine the case where p_1 = 0.1437 and p_2 = 0.0793. One should consider which number is more enlightening, δ = -0.0644, or ϕ = 55.18%. In many cases, the relative change (the ratio) is easier to interpret that the absolute change (the difference).

When planning studies, the value of p_2 is usually known and the power is calculated at various values of ϕ. The value of p_1 is then calculated using $p_1 = p_2 \times \delta$. Because of the requirement that $0 < p_1 < 1$, the values of ϕ are restricted to the interval $0 < \phi < 1/p_2$, not $0 < \phi < \infty$ as you might expect. Likewise, the values of p_2 are restricted to $0 < p_2 < 1/\phi$ if $\phi > 1$ and $0 < p_2 < 1$ if $\phi < 1$.

Because typical values of ϕ are usually between 0.5 and 1.5, the values of p_2 are restricted to be between 0 and 0.667. When the values of p_2 are outside this range, it may be more convenient to use the odds ratio.

Odds Ratio

Chances are usually communicated as long-term proportions or probabilities. In betting, chances are often given as odds. For example, the odds of a horse winning a race might be set at 10-to-1 or 3-to-2. Odds can easily be translated into probabilities, and vice versa. An odds of 3-to-2 means that the event is expected to occur three out of five times. That is, an odds of 3-to-2 (1.5) translates to a probability of winning of 0.60.

The odds of an event are calculated by dividing the event risk by the non-event risk. Thus the odds are

$$Odds_1 = \frac{p_1}{1-p_1} \text{ and } Odds_2 = \frac{p_2}{1-p_2}$$

For example, if p is 0.6, the odds are 0.6/0.4 = 1.5. Rather than represent the odds as a decimal amount, it is re-scaled into whole numbers. Thus, instead of presenting the odds as 1.5-to-1, they present as 3-to-2.

Two odds could be compared by considering their difference, but it is more convenient in many situations to form their ratio. Thus, another way to compare proportions is to compute the ratio of their odds. The odds ratio is

$$\psi = \frac{Odds_1}{Odds_2} = \frac{\frac{p_1}{1-p_1}}{\frac{p_2}{1-p_2}}$$

Unlike the difference and the ratio, the odds ratio is not restricted by the value of p_2. The range of possible values of the odds ratio is $-\infty < \psi < \infty$. Because of the freedom in specifying the parameters, the odds ratio is a popular parameterization, even though it is not as easy to interpret as the difference and the ratio.

Specifying the Proportions – Very Important!

It is important to understand the interpretation of p_1 and p_2 within ***PASS***. Suppose p_1 represents the proportion in the treatment group and p_2 represents the proportion in the control group. In most hypothesis tests, these values are equal under the null hypothesis and different under the alternative hypothesis. Thus, under the null hypothesis, all that is needed is the value of p_1 or p_2, but not both. Under the alternative hypothesis, both values are necessary. So, when the input screen asks for p_2 and the difference, these values should be interpreted as follows. The value of p_2 is actually the value of both p_1 and p_2 under H_0. Under H_1, the value of p_1 is calculated from p_2 and δ using the formula $p_1 = p_2 + \delta$.

Also, it is important to understand what we mean by ‘under H_1’ in the above discussion. Notice that H_1 does not specify the exact value of p_1. Instead, it specifies a range of values. For example, H_1 might be $p_1 > p_2$ or $p_1 - p_2 > \delta$. However, even though H_1 gives a range of values for p_1, the power is computed at a specific value of p_1.

Selecting an appropriate value for p_1 must be done very carefully. We recommend the following approach. Select a value of p_1 that represents the change from p_2 that you want the experiment to detect. When you calculate a sample size, it is interpreted as the sample size necessary to detect a difference of at least $p_1 - p_2$ when the significance level is α and the power is $1-\beta$.

The important point is that p_1 is not the value you anticipate obtaining from an experiment. Instead, it is that value of p_1 at which you want to compute the power. This is a very important distinction! This is why, when investigating the power after an experiment is run, we recommend that you do not simply plug in the values of p_1 and p_2 from that experiment. Rather, you use values that represent the size of the difference that you want to detect.

Chapter 9

Means

Introduction

This chapter introduces power analysis and sample size calculation for tests that compare means. In many situations, the results for each treatment group are summarized as means. There are many issues that must be considered when designing experiments for comparing means. For example, are the means independent or correlated? Which test statistic to use? Will a parametric or nonparametric test be used? Are the data normally distributed? Are there more than two treatment groups? The answers to these and other questions result in a large number of situations.

Specifying the Means

Assume that μ_1 is the mean of an experimental group and μ_2 is the mean of a control (standard or reference) group. Suppose δ represents their difference. That is, $\delta = \mu_1 - \mu_2$. In most hypothesis tests, the null hypothesis (H_0) is $\delta = 0$ and the alternative hypothesis (H_1) is $\delta \neq 0$. Since H_0 assumes that $\delta = 0$, all that is really needed to compute the power is the value of δ under H_1. So, when the input screen asks for μ_1 and μ_2, these values should be interpreted as follows. The value of μ_1 is actually the value of both μ_1 and μ_2 under H_0. Under H_1, the values of μ_1 and μ_2 provide the value of δ at which the power is calculated.

The above discussion is summarized in the following chart:

Input Parameter	**Under H_0**	**Under H_1**
Mean1	μ_1, μ_2	μ_1
Mean2	ignored	μ_2

Also, it is important to understand what we mean by ‘under H_1’ in the above discussion. H_1 defines a range of values for δ at which the power can be computed. To compute the power, the specific values of δ must be determined. Thus, there is not a single power value. Instead, there are an infinite number of power values possible, depending on the value of δ.

Selecting an appropriate value for μ_1 must be done very carefully. We recommend the following approach. Select a value of μ_1 that represents the change from μ_2 that you want the experiment to detect. When you calculate a sample size, it is interpreted as the sample size necessary to detect a difference of at least δ when the significance level is α and the power is $1-\beta$.

It is important to realize that δ is not the value you anticipate obtaining from the experiment. Instead, it is that value of δ at which you want to compute the power. This is a very important distinction! This is why, when investigating the power after an experiment is run, we recommend that you do not simply plug in the values of μ_1 and μ_2 from that experiment. Rather, you use values that represent the size of the difference that you want to detect.

Specifying the Standard Deviation

Usually, statistical hypotheses about the means make no direct statement about the standard deviation. However, the standard deviation is a parameter in the normal distribution, so its value must be specified. For this reason, it is called a *nuisance parameter*.

Even though it is not of primary interest, an estimate of the standard deviation is necessary to perform a power analysis. Finding such an estimate is difficult not only because it is required before the data are available, but also because the physical interpretation of the standard deviation is vague. How do you estimate a quantity without data and without a clear understanding of what it is? This section will try to help.

Understanding the Standard Deviation

The standard deviation has two general interpretations. First, it is similar to the average absolute difference between each observation and the mean. Second, it is the average absolute difference between every pair of observations.

The standard deviation of a population of values is calculated using the formula

$$\sigma_X = \sqrt{\frac{\sum_{i=1}^{N}\left(X_i - \mu_X\right)^2}{N}}$$

where *N* is the number of items in the population, *X* is the variable being measured, and μ_X is the mean of *X*. This formula indicates that the standard deviation is the square root of an average of the squared differences between each value and the mean. The differences are squared to remove the sign so that negative values will not cancel out positive values. After summing up these squared differences and dividing by *N*, the square root is taken to put the result back in the original scale. Bottom line—the standard deviation can be thought of as the average absolute difference between the data values and their mean.

Estimating the Standard Deviation

Our task is to find a rough estimate of the standard deviation to use in a power analysis. Several possible methods could be used. These include using the results of a previous study or a pilot study, using the range, using the coefficient of variation, etc. ***PASS*** includes a Standard Deviation Estimator procedure that will help you obtain a standard deviation estimate based on these methods. It is loaded from the Tools menu. Remember that as the standard deviation increases, the power decreases. Hence, an increase in the standard deviation will cause an increase in the sample size. To be conservative in sample size calculation, you should use a large value for the standard deviation.

Simulations

Most of the formulas used in ***PASS*** were derived by analytic methods. That is, based on a series of assumptions, a formula for the power and sample size is derived mathematically. This formula is then programmed and made available in ***PASS***. Unfortunately, the formula is only as realistic as the assumptions upon which it is based. If the assumptions are inaccurate in a certain situation, the power calculations may also be inaccurate. An alternative to using analytic methods is to use *simulation* (or *Monte Carlo*) techniques. Because of the speed of modern computers, simulations can now be run in minutes that would have taken days or weeks only a few years ago.

In power analysis, *simulation* refers to the process of generating several thousand random samples that follow a particular distribution, calculating the test statistic from each sample, and tabulating the distribution of these test statistics so that the significance level and power of the procedure may be estimated.

The steps to a simulation study are

1. Specify how the study is carried out. This includes specifying the randomization procedure, the test statistic that is used, and the significance level that will be used.
2. Generate random samples from the distributions specified by the null hypothesis. Calculate each test statistic from the simulated data and determine if the null hypothesis is accepted or rejected. Tabulate the number of rejections and use this to calculate the test's significance level.
3. Generate random samples from the distributions specified by the alternative hypothesis. Calculate the test statistics from the simulated data and determine if the null hypothesis is accepted or rejected. Tabulate the number of rejections and use this to calculate the test's power.
4. Repeat steps 2 and 3 several thousand times, tabulating the number of times the simulated data leads to a rejection of the null hypothesis. The significance level is the proportion of simulated samples in step 2 that lead to rejection. The power is the proportion of simulated samples in step 3 that lead to rejection.

How Large Should the Simulation Be?

As the number of simulations is increased, the precision and running time of the simulation will be increased also. This section provides a method for estimating of the number simulations needed to achieve a given precision.

Each simulation iteration (or simulation) generates a binary outcome: either the null hypothesis is rejected or not. Thus, the significance level and power estimates each follow the binomial distribution. This knowledge makes it a simple matter to compute confidence intervals for the significance level and power values.

The following table gives one-half the width of a 95% confidence interval for the power when the estimated value is either 0.50 or 0.95.

Simulation Size M	Half-Width when Power = 0.50	Half-Width when Power = 0.95
100	0.100	0.044
500	0.045	0.019
1000	0.032	0.014
2000	0.022	0.010
5000	0.014	0.006
10000	0.010	0.004
50000	0.004	0.002
100000	0.003	0.001

Notice that a simulation size of 1000 gives a precision of plus or minus 0.014 when the true power is 0.95. Also, as the simulation size is increased beyond 5000, there is only a small amount of additional accuracy achieved. Since most sample-size studies require an accuracy of within one or two percentage points, simulation sizes from 2000 to 10000 should be ample.

You are Running Two Simulations

It is important to realize that when you run a simulation in ***PASS***, you are actually running two separate simulations: one to estimate the significance level and the other to estimate the power. The significance-level simulation is defined by the input parameters labeled "|H0". The power simulation is defined by the input parameters labeled "|H1". Even though you have complete flexibility as to what distributions you use in each of these simulations, it usually makes sense to use the same distributions for both simulations—only changing the values of the means.

Unequal Standard Deviations

One of the subtle problems that can make the results of a simulation study misleading is to specify unequal standard deviations unknowingly when you are investigating another feature, such as the amount of skewness. It is well known that if the standard deviations differ (a situation called heteroskedasticity), the accuracy of the significance level and power is doubtful. When investigating the power of the t or F tests in non-normal situations, care must be taken to insure that the standard deviations of the groups remain about the same. Otherwise, the effects of skewness and heteroskedasticity cannot be separated.

Finding the Hypothesized Means

It is important to set the mean difference of each simulation carefully. In the case of analytic formulas, the mean difference is specified easily and directly. Usually, the mean difference is set to zero under the null hypothesis and to a non-zero value under the alternative hypothesis. You must make certain that you follow this pattern when setting up a simulation.

For most distributions, the means are set explicitedly (the exception is the multinomial distribution, where this is impossible). Hence, for both the null and alternative simulations, it is relatively simple to calculate the mean difference. You must do this! We will now present two examples showing how this is done.

For the first example, consider the case of a simulation being run to compare two independent group means using the two-sample t-test. Suppose the ***PASS*** setup is as follows. Note that N(40 2) stands for a normal distribution with a mean of 40 and a standard deviation of 2.

Distribution	*PASS* Input	Mean Value of Simulated Data
Group 1 Distribution | H0	N(40 2)	40.0
Group 2 Distribution | H0	N(40 2)	40.0
Group 1 Distribution | H1	N(42 2)	42.0
Group 2 Distribution | H1	N(40 2)	40.0

The mean difference under H0 is 40 – 40 = 0, which is as it should be. The mean difference under H1 is 42 – 40 = 2. Hence, the power is being estimated for a mean difference of 2.

Next we will consider a more complicated example. Suppose the ***PASS*** setup is as follows. Note that N(40 2)[70];K(0)[30] specifies a mixture distribution made up of 70% from a normal distribution with a mean of 40 and a standard deviation of 2 and 30% from a constant distribution with a value of 30.

Distribution	*PASS* Input	Mean Value of Simulated Data
Group 1 Distribution | H0	N(40 2) [70];K(0)[30]	40(0.7) + 30(0.3) = 37.0
Group 2 Distribution | H0	N(40 2) [70];K(0)[30]	40(0.7) + 30(0.3) = 37.0
Group 1 Distribution | H1	N(42 2) [70];K(0)[30]	42(0.7) + 30(0.3) = 38.4
Group 2 Distribution | H1	N(40 2) [70];K(0)[30]	40(0.7) + 30(0.3) = 37.0

The mean difference under H0 is 37.0 – 37.0 = 0, which is as it should be for the null hypothesis. The mean difference under H1 is 38.4 – 37.0 = 1.4. Hence, the power is being estimated by simulation for a mean difference of 1.4.

You must always be aware of what the mean differences are under both the null and alternative hypotheses.

Adjusting the Significance Level

When faced with the task of designing an experiment that will have a specific significance level for a situation that does not meet the usual assumptions, there are several possibilities.

1. A statistician could be hired to find an appropriate testing procedure.
2. A nonparametric test could be run that (hopefully) corrects for the implausible assumptions.
3. The regular parametric test could be run, relying on the test's 'robustness' to correct for the implausible assumptions.
4. A simulation study could be conducted to determine an appropriate adjustment to the significance level so that the actual significance level is at the required value.

We will now present an example of how to do the simulation adjustment alluded to in item 4, above.

The two-sample t-test is known to be robust to the violation of some assumptions, but it is susceptible to inaccuracies when the data contain outliers. A mixture of two normal distributions will be used to generate data with outliers. The mixture will draw 95% of the data from a normal distribution with a mean of 0 and a standard deviation of 1. The other 5% of the data will come from a normal distribution with a mean of zero and a standard deviation of 10. A simulation study using 10,000 iterations and a sample size of 100 per group produced the following results when the nominal significance level was set to 0.05.

Nominal Alpha	Actual Alpha	Lower 95% Confidence Limit	Upper 95% Confidence Limit	Power
0.050	0.045	0.041	0.049	0.816
0.055	0.051	0.047	0.055	0.843
0.060	0.057	0.053	0.062	0.835

The actual alpha level of the t-test is 0.045, which is below the target value of 0.50. When the nominal alpha level is increased to 0.055, the actual alpha is 0.051—close to the desired level of 0.05. Hence, an adjustment could be applied as follows. Analyze the data with the two-sample t-test even though they contain outliers. However, instead of using an alpha of 0.050, use an alpha of 0.055. When this is done, the simulation shows that the actual alpha will be at the desired 0.05 level.

There is one limitation to this method: the resulting test procedure is not necessarily efficient. That is, it may be possible to derive a testing procedure that is more efficient (requires a smaller sample size to achieve the same power). For example, in this example, a test based on the trimmed mean may be more efficient in the presence of outliers. However, if you do not have the time or ability to derive an alternative test, this adjustment allows you to obtain reasonable testing procedure that achieves a desired significance level and power.

Quick Start Index

Index entries are of the form "chapter-page". A list of chapters is given in the Table of Contents.

V

W