

Stat49Pro

By Scott Guth

Email: sguth@mtsac.edu

Copyrights & Disclaimer

All files of the Stat49Pro library are copyrighted (c) by Scott Guth unless otherwise noted.

The Stat49Pro library is distributed freely in the hope that it will be useful, but is provided “as is” and is subject to change without notice. No warranty of any kind is made with regard to the software or documentation. The author shall not be liable for incidental or consequential damages in connection with the software and the documentation.

Permission to copy the whole, unmodified Stat49Pro package is granted provided that the copies are not made or distributed for resale (excepting nominal copying fees).

Acknowledgements

| | |
|-------------------|--|
| Christian Meland: | This package was originally Christian’s idea. For a time we worked on it together, but eventually Christian became too busy. Christian wrote the following Stat49Pro commands: ZALPHA, and FALPHA, along with subroutines used by these commands. TALPHA and CHIALPHA were originally written by Christian, but then replaced by me to correct a few problems. |
| Eric Hubert | Suggestions & beta testing. |

Thanks also to the many users of the original Stat48 and Stat49 programs who have sent comments & suggestions which had some influence on the development of this package as well.

Introduction

Stat49Pro is made up of two main components: the data manager (“Manage”) and the statistical inference package (“Stat49Pro”). The data manager is used for entering sample data, and is used as an interface for feeding statistics computed “on the fly” directly into the inferential statistics package. Stat49Pro contains most inferential statistics applications found in an elementary statistics course. This includes confidence intervals, hypothesis tests, and sample size computations. All applications may be executed from the built-in graphical interface, or directly from the command line.

Installation

To install the library to port 0: (a similar procedure applies for port 1 or 2)

- a) Remove possible older version with :0:1043 PURGE
- b) Download the library to your HP49
- c) Recall the library to the stack and purge the variable created by the download procedure.
- d) Type 0 STO
- e) Press ON-C

Using Stat49Pro

To use Stat49Pro, enter the STAT menu by pressing [RS] [5], then select “7. Stat49Pro” or “8. Manage”. “Stat49Pro” executes the graphical interface, which guides the user to an appropriate application, receives input, then executes the application. “Manage” executes the data manager, which stores data, and later sends statistics into the input forms in Stat49Pro. Learn to use the data manager! It is one of the strengths of the package.

The Data Manager

The Data Manager opens with a choose menu which offers the following choices

- *Add New*
This feature adds a new data set. Use the same convention for naming sample data as for naming variables on the HP49
The data manager supports the following data types:
 - A. 1-variable
 - B. 1-variable with frequencies (or weights)
 - C. 2-variable
 - D. 2-variable with frequencies (or weights)
 - E. Matrix (for 2-way contingency table (test for independence), or 2-way ANOVA with one observation per cell).
 - F. Matrix of Lists (for 2-way ANOVA with more than one observation per cell).
- *Edit Existing*
- *Delete Data*

- *Describe Data*
This option computes whatever statistics are necessary to describe a selected data set, depending on the data type.
- *Export Data*
This option copies a selected data set to either the stack, or to 'ΣDAT' for use in the HP49's own statistics package.
- *Import Data*
This option imports data from the stack (or as a passed parameter when in 'algebraic' mode), or from 'ΣDAT'.

Use the data manager to store "raw" data. An unlimited number of data samples may be stored. Each sample is given a name. This name must be a valid global ID – use the same conventions used to name variables. When ready to analyze data, enter Stat49Pro and select an application. After entering your application's input form, you may enter statistics manually, *or press the [IMPRT] menu key to import statistics from the data manager.* When using the [IMPRT] feature, some applications require two (independent) samples to be selected, while others (like one-way ANOVA), require that more than two independent samples be selected. Whenever this is the case, simply use the menu key [CHK] to select several samples before pressing [OK]. Please see the table at the end of this document that describes which data types may be imported into any given application.

The data manager stores sample data in a matrix, and the all such matrices are bundled together and stored as a single library object in the variable 'ΣPRORAW'. Data that are entered into input forms are bundled together & stored in another library object called 'ΣPRODAT'. Deleting either of these variables from memory will destroy your data, so exercise caution. If you need to remove data stored by the data manager, use the delete function included in the data manager.

Conventions

Stat49Pro uses the following notational conventions:

| | |
|-----------|------------------------------------|
| μ | population mean |
| \bar{x} | sample mean |
| π | population proportion |
| p | sample proportion |
| σ | population standard deviation |
| s | sample standard deviation |
| ρ | population correlation coefficient |
| r | sample correlation coefficient |

Other conventions:

- "TS" stands for "test statistic", while "pval" stands for "p-value".
- Confidence intervals that use the z or t distributions return the interval, as well as a value E. Most texts refer to E as the "maximum error in the estimate" of whatever parameter is being estimated.
- One sided confidence intervals are supported indirectly. To perform one, simply use the application which performs the 2-sided interval, and double your α -- by changing your confidence level to $1-2\alpha$.
- Two applications, namely TT2I μ and TC2I μ require a parameter POOL? which tells the application what to do regarding the pooling of variances. The choices are as follows:

A. "Unpooled" using degrees of freedom given by $df = \frac{(V_1 + V_2)^2}{\frac{V_1^2}{n_1 - 1} + \frac{V_2^2}{n_2 - 1}}$ where $V_1 = \frac{s_1^2}{n_1}$ and $V_2 = \frac{s_2^2}{n_2}$.

B. "Pooled" using $df = n_1 + n_2 - 2$.

C. "Unpooled EZ DF" uses degrees of freedom given by $df = \min\{n_1 - 1, n_2 - 1\}$.

D. "F-test decides" runs an F test to test for equal variances, and then runs the current application according to the results of the F test. Degrees of freedom used are either those given in A or B above.

E. "F-test EZ DF" runs an F test to test for equal variances, then runs the current application according to the results of the F test. Degrees of freedom used are either those given in B or C above.

F. The "Normal Probability Plot" application is used to check for normality in a sample. There are several different formulas for computing the area to the left of the I 'th quantile on the normal distribution. Stat49Pro allows the user to modify the formula for this area. The default is $(I-3/8)/(N+1/4)$. The user may change this in the input form, but the symbolic entered must be a function of I and N . There is no need to purge global variables I or N if they exist.

G. Notations: TS = Test Statistic, CV = Critical Value, pval = p-value.

The following conventions are necessary *only* when running applications directly from a command prompt. You may ignore this section if you plan to use the graphical interface exclusively to launch applications.

- All hypothesis tests require a parameter [α ,TAILS], where α is the level of significance, and TAILS is equal to -1 for a left-tail test, 0 for a 2-tail test, and 1 for a 1-tail test.

I hope that you find this software useful. If you do, please drop me a note at sguth@mtsac.edu. Comments & suggestions are welcome. On the following page is a complete list of all applications offered in Stat49Pro.

Stat49Pro Inferential Statistics Applications

The current applications performed by Stat49Pro are listed below. All are accessible through the graphical interface of Stat49Pro with a few exceptions. In addition, all commands may be executed directly from the command line. Executing any particular command without arguments will give online help for that command.

| Application Name | Application Purpose | Allowed Data Types for Import |
|------------------|---|-------------------------------|
| ZT μ | z test for pop. mean | 1Var; 1Var/Freq |
| ZC μ | z conf. int. for pop. mean | 1Var; 1Var/Freq |
| N μ | sample size for estimating pop. mean | 1Var; 1Var/Freq |
| TT μ | t test for pop. mean | 1Var; 1Var/Freq |
| TC μ | t conf. int. for pop. mean | 1Var; 1Var/Freq |
| ZT π | z test for pop. proportion | No import allowed. |
| ZC π | z conf. int. for pop. proportion | No import allowed. |
| N π p | sample size for pop. proportion -- estimate p is known | No import allowed. |
| N π | sample size for pop. proportion -- estimate p is not known | No import allowed. |
| ChiT σ | Chi ² test for pop std. dev. or variance. | 1Var; 1Var/Freq |
| ChiC σ | Chi ² conf. int. for pop. std. dev. or variance. | 1Var; 1Var/Freq |
| ZT2D μ | z test for two means, dependent samples | 2Var; 2Var/Freq |
| ZC2D μ | z conf. int. for two means, dependent samples | 2Var; 2Var/Freq |
| TT2D μ | t test for two means, dependent samples | 2Var; 2Var/Freq |
| TC2D μ | t conf. int. for two means, dependent samples | 2Var; 2Var/Freq |
| ZT2I μ | z test for 2 means, independent samples | 1Var; 1Var/Freq Multi-choose |
| ZC2I μ | z conf. int. for 2 means, independent samples | 1Var; 1Var/Freq Multi-choose |
| TT2I μ | t test for 2 means, independent samples. This test requires a parameter [α ,tails,pool?], where "tails" is equal to -1 for a left-tail test, 0 for a 2-tail test, and 1 for a 1-tail test, and "pool?" is a real number equal to one of the following: 0=unpooled, 1=pooled, 2=unpooled/EZ DF, 3=F-test decides, 4=F-test/EZ DF. For the meaning of these options, see items A through E on the preceding page. | 1Var; 1Var/Freq Multi-choose |
| TC2I μ | t conf. int. for 2 means, independent samples. This command requires a parameter [conf, pool?] where "conf" is the level of confidence, and "pool?" is as described for TT2I μ above. | 1Var; 1Var/Freq Multi-choose |
| ZT2 π | z test for 2 pop. proportions | No import allowed. |
| ZC2 π | z conf. int. for 2 pop. proportions | No import allowed. |
| FT2 σ | F test for two variances | 1Var; 1Var/Freq Multi-choose |
| FC2 σ | F conf. int. for two variances | 1Var; 1Var/Freq Multi-choose |
| CONTING | Contingency tables | Matrix |
| GOODFIT | Chi-square test for fit. | 2Var; 2Var/Freq |
| ANOVA1 | One-way ANOVA | 1Var; 1Var/Freq Multi-choose |
| ANOVA21 | 2-way ANOVA, one observation per cell | Matrix |
| ANOVA21SS | Sums of squares for ANOVA21 | Not available in GUI |
| ANOVA2 | 2-way ANOVA, more than one observation per cell | List of Matrices |
| ANOVA2SS | Sums of squares for ANOVA2 | Not available in GUI |
| ZTp | z test for linear correlation | 2Var; 2Var/Freq |
| TTp | t test for linear correlation | 2Var; 2Var/Freq |
| ZTpFisher | z test for linear correlation (claiming that $\rho = \rho_0 \neq 0$) using Fisher transform. | 2Var; 2Var/Freq |
| ZALPHA | z critical values (inverse z distribution). | No import allowed. |
| TALPHA | t critical values (inverse t distribution). | No import allowed. |
| CHIALPHA | Chi ² critical values (inverse Chi ² distribution). | No import allowed. |
| FALPHA | F critical values (inverse F distribution). | No import allowed. |
| RankSum | Input: { list of real vectors } Output: 3: {list of sample sizes} 2: {list of original vectors - but paired with overall rankings} 1: {list of Rank-Sums} So the input requires each sample stored in vector, and each vector stored in a single list. Execute the command and you get the sum of the ranks for each sample. These ranks may be used in several non-parametric tests (and getting the ranks is really the hardest part of these tests), namely: 1. The Wilcoxon Rank-Sum test; 2. The Mann-Whitney U test; 3. The Kruskal-Wallis test. | 1Var Multi-choose |
| NORMPLOT | Plots a normal probability plot, gives normal scores, and computes r, the correlation coefficient. Note: this function will alter the contents of Σ DAT. | 1Var |
| ZSCORES | Computes normal scores for a given sample size. | Not available in GUI |