

Coin's Package Permutation Tests Graphic User Interface Extension for R Commander

Daniel-Corneliu LEUCUȚA¹

¹ „Iuliu Hațieganu” University of Medicine and Pharmacy Cluj-Napoca, Department of Medical Informatics and Biostatistics, 6 Louis Pasteur, 400349 Cluj-Napoca, Romania.

E-mail: dleucuta@umfcluj.ro

* Author to whom correspondence should be addressed; Tel.: +4-0264-431697; Fax: +4-0264-593847.

Received: 5 September 2011 / Accepted: 15 September 2011 / Published online: 22 September 2011

Abstract

Background: For small, skewed, sparse, unbalanced data, or data with lots of ties the exact P values computed with permutation tests can help making valid statistical inferences. In the medical field, especially in exploratory studies, like phase I, phase II trials or correlation studies these techniques can prove of a great help. R is a free software environment for statistical computing and graphics that offers this kind of tests through the coin package. R Commander graphical user interface to R is one of the most comprehensive in respect of functions offered.

The *aim* of this work was to create an extension for R Commander that offers an interface for permutation tests, a handful of non parametric tests in a permutation framework, and new options for some of the existing non parametric tests, making access to the coin package functions.

Materials and Methods: The extension was compiled with R version 2.10. and tested on with R version 2.12.1, and R Commander 1.6-2.

Results: The extension has been released successfully, named RcmdrPlugin.coin, and it is available online on CRAN servers for any R Commander user to download and use. The extension creates a new menu, Coin that provides interface to the following tests: Cochran-Mantel-Haenzel test, Marginal Homogeneity test, linear by linear association test, chi square test; median test, two sample Wilcoxon test, normal quartile test, two/k sample permutation test, Kruskal-Wallis test, Ansari-Bradley test, Fligner-Killeen test, Spearman test, Maximally Selected Statistics, two/k sample test for censored data, paired-samples Wilcoxon test, and Friedman rank-sum test.

Keywords: Coin package; Permutation tests; R Commander; Graphic user interface.

Introduction

Permutation tests are statistical significance tests where the distribution of the statistic under the null hypothesis is found by computing all values of the test statistic under rearrangements of the labels of the observations. They are also known as randomization tests, re-randomization tests, conditional inference procedures or distribution free tests. Permutation tests are a subgroup of non-parametric tests.

Modern statistical methods rely a lot on non-parametric techniques to compare two or more groups. Although not using distributional assumptions about the populations compared they rely on asymptotic theory that is valid only if the sample sizes are sufficient and well balanced across the populations. For small, skewed, sparse, unbalanced data, or data with lots of ties the asymptotic

theory might not apply [1]. In these cases the exact P values (and confidence intervals) computed with permutation tests can help making valid statistical inferences. In the medical field, especially in exploratory studies, like phase I, phase II trials or correlation studies these techniques can prove of a great help, due to the small number of subjects and problems with data nature (skewness, unbalanced data).

The idea behind them is not a new one, since they were first devised by R. A. Fisher [2]. Due to the computational needs not present at that time at a reasonable price, they weren't widely used, but now these constraints faded away. Nowadays many statistical computer programs offer exact tests: StatXact [3], SPSS Exact Tests module [4], SAS [5], R coin (Conditional inference procedures in a permutation test framework) package [6], and others.

R is a free software (under the terms of GNU General Public License) environment for statistical computing and graphics [7]. It runs on many platforms, like Windows, MacOS and UNIX including Linux. R offers a huge number of statistical and graphical techniques in a well organized, flexible, extensible, and coherent system. Beside data manipulation and analysis it is used also for research in statistical methodology. R works in a command line interface. For users new to R it might be a difficult to get to work with it. To overcome this, a handful of packages (extensions to R) were written to offer graphical user interfaces (GUI). There are a number of GUIs for R: R Commander (a R package that provides basic-statistics GUI for R with script, output windows, menus, buttons and information fields [8]), Red-R (a visual programming interface for R, that uses concepts of data-flow programming to make data the center of attention while hiding all the programming complexity [9]) SciViews-R (a series of packages providing a GUI application programming interface for R [10]), JGR (Java Gui for R - a universal and unified Graphical User Interface for R [11]), RKWard (a transparent front-end to R providing a convenient user-interface and integration with an office suite [12]).

R Commander GUI is one of the most comprehensive in respect of functions offered. One thing that helped this was the add-ins functionality that allowed many developers to add new functionality to the standard GUI. Till now it didn't provide a specific interface for permutation test.

The **aim** of this work was to create an extension for R Commander that offers an interface for permutation tests, a handful of non parametric tests in a permutation framework, and new options for some of the existing non parametric tests, making access to the coin package functions.

Material and Method

The extension (plug-in) package for the R Commander GUI was created on an Ubuntu 10.04 LTS (Lucid Lynx). It was written in Scite [13] text editor, and compiled in with R version 2.10.1. The code was installed and tested on a Windows 7 workstation, with R version 2.12.1, and R Commander 1.6-2. The extension was named RcmdrPlugin.coin, and was uploaded on CRAN network of servers that stores up-to-date, mirrored, code and documentation for R [14].

The extension code was written in R programming language that provides access to Tcl/Tk programming language [15] functions, for building GUIs. The Tcl/Tk functions are further abstracted by the R Commander GUI functions. The code was based on other pieces of code for other windows/menus in the standard R Commander installation. The code was released as free software under the terms of GNU General Public License).

The extension was created to provide a GUI for the following tests offered by the coin **[Error! Bookmark not defined.]** package in the permutation framework: Cochran-Mantel-Haenzsel test, Marginal Homogeneity test, linear by linear association test, chi square test; median test, two sample Wilcoxon test, normal quartile test, two/k sample permutation test, Kruskal-Wallis test, Ansari-Bradley test, Fligner-Killeen test, Spearman test, Maximally Selected Statistics, two/k sample test for censored data, paired-samples Wilcoxon test, Friedman rank-sum test.

Results

The extension has been released successfully, and it is available online on CRAN servers for any R Commander user to download and use, on the web: <http://CRAN.R-project.org/package=RcmdrPlugin.coin>, or more commonly through the Install packages menu command in R console window.

The extension adds the Coin menu that has three submenus: Contingency tests (tests for contingency tables), Independent Location tests and Symmetry tests. These submenus offer a suite of non parametric tests (see Figure 1).

For each command in the menu a window was built, that allows the selections of the variables involved in the test (usually the group variable, and the response variable, as well as a block variable); the alternative hypothesis (two-sided, difference < 0, difference > 0), the type of test (default, exact, Monte Carlo resampling approximation, asymptotic null distribution), a subset expression (to allow using the test on a subset of the actual data set), and the help for each function (see Figure 2 for an example of such a window). Depending on the test the window might offer options to choose the Ties method (e.g. for the censored data test: logrank, average-scores, Horton and Lausen), the Zero method (e.g. for the two sample Wilcoxon test: Pratt, Wilcoxon), the confidence level (e.g. for the two sample Wilcoxon test:), the number of approximate replications (e.g. for the Friedman rank-sum test), the possibility to compute pairwise comparisons of groups (for Friedman rank-sum test, and for Kruskal-Wallis test), or other specific options.

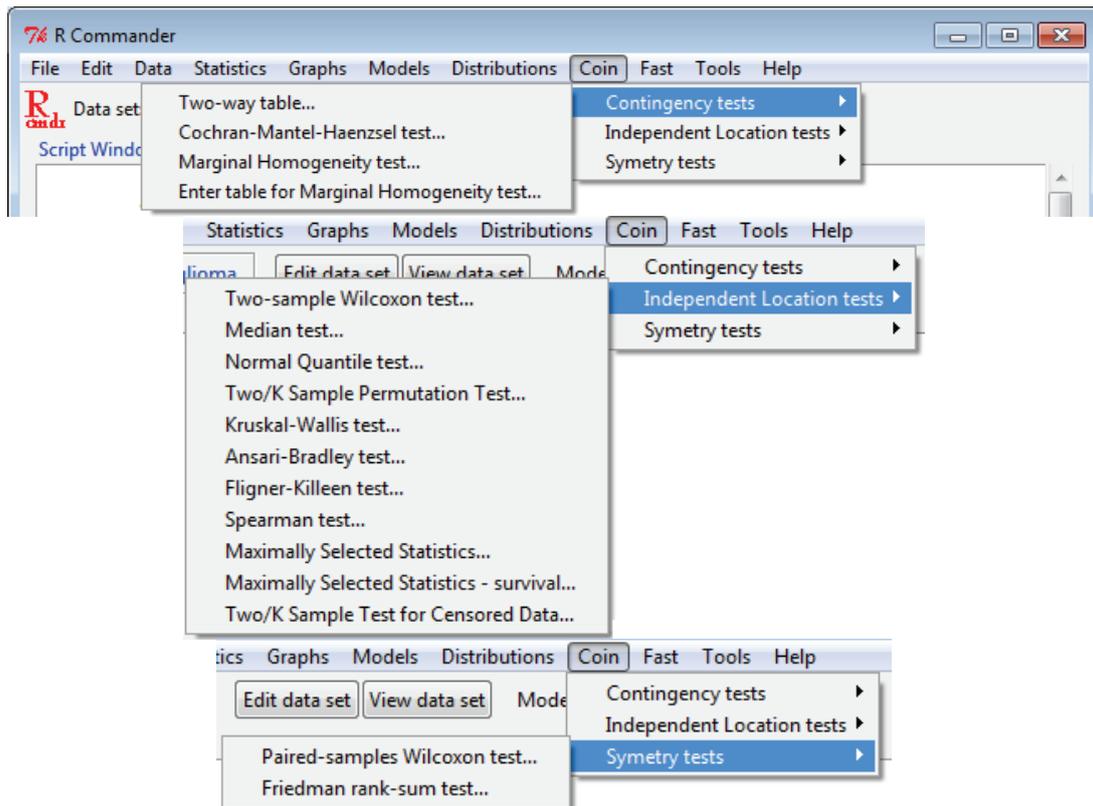


Figure 1. The menus added by the extension to R Commander GUI to access the coin package tests

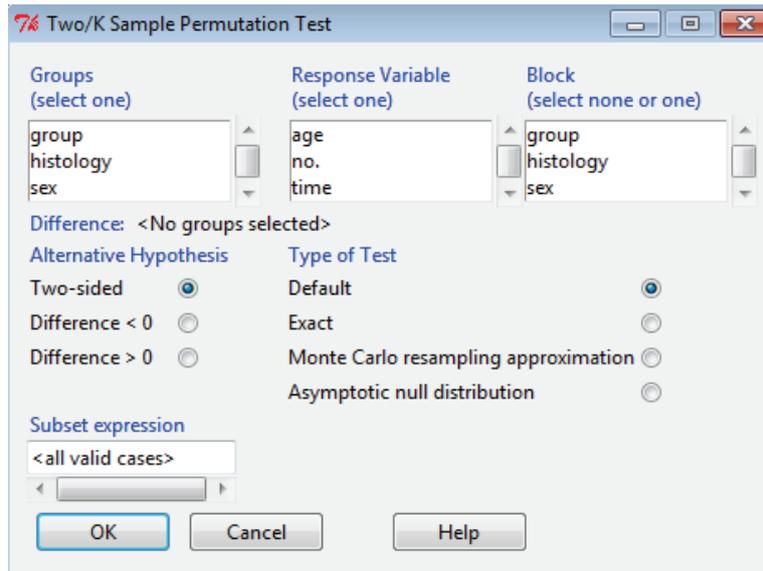


Figure 2. An example of a dialog window for the two/k sample permutation test

For the two/k sample test, and for the maximally selected statistics test for censored data window, if the survival data definition was filled in the R Commander's Data menu, then the variables will be automatically selected in the window.

The output for the tests is as that provided by the functions called in the coin package. For some tests central tendency indicators like median, are shown by groups (see Figure 3).

```
> oneway_test(pain ~ group, alternative='two.sided',
+ | distribution='exact', data=neuropathy)
      Exact 2-Sample Permutation Test
data: pain by group (control, treat)
Z = -1.3191, p-value = 0.1916
alternative hypothesis: true mu is not equal to 0
```

Figure 3. Output for the two/k sample permutation test

Discussion

The R Commander GUI extension for conditional inference tests in a permutation framework (coin [6] package) was successfully built.

This extension adds new functionality to the R Commander GUI helping users with easier access to a wider variety of tests, and options for tests. The extension adds GUI to tests that are not available in the R Commander package, and adds new functionality to some of those that are available in R Commander.

Thus it adds Cochran-Mantel-Haenzsel test, Marginal Homogeneity test, linear by linear association test, median test, normal quartile test, two/k sample permutation test, Ansari-Bradley test, Fligner-Killeen test, Maximally Selected Statistics, two/k sample test for censored data. It also adds post-hoc pairwise comparison tests for Friedman rank-sum test, and for Kruskal-Wallis test.

The new GUI functionality added to existing tests in R Commander is represented by the option to have blocks in analyzing the data (e.g. when doing studies in multiple centers), for the Wilcoxon rank sum and signed rank tests [16, 17], Kruskal-Wallis rank sum test [16, 18], logrank test [16, 19]. Another useful effect is the possibility to have exact p values for the Fligner-Killeen test of homogeneity of variances [16, 20], logrank test [16, 19], Friedman rank sum test [16, 21] Kruskal-Wallis rank sum test [16, 18].

Conclusion

This work ended up with a GUI extension for R Commander that offers an interface for permutation tests, a handful of non parametric tests in a permutation framework, and new options for some of the existing non parametric tests, making easier access even for not so knowledgeable R users, to the coin package functions.

Conflict of Interest

The author declares that has no conflict of interest.

References

1. Corcoran CD, Senchaudhuri P, Mehta CR, Patel NR. Exact Inference for Categorical Data. In: Armitage P, Colton T, editors. Encyclopedia of Biostatistics: 8-Volume Set. 2nd Edition. John Wiley & Sons; 2005.
2. Fisher RA. Statistical Methods for Research Workers. Edinburgh: Oliver and Boyd; 1925.
3. StatXact Version 6 for Windows. (2004). Software for Exact Nonparametric Inference, featuring Cytel studio. Cytel Software Corporation, Cambridge, MA. URL <http://www.cytel.com/>.
4. SPSS Exact Tests for Windows. (1995). SPSS Inc., Chicago.
5. SAS Version 9 (2004). SAS Institute Inc., Cary, NC.
6. Hothorn T, Hornik K, Wiel MA, Zeileis A. Implementing a Class of Permutation Tests: The coin Package. Journal of Statistical Software 2008;28(8):1-23. Available from: <http://www.jstatsoft.org/v28/i08/>.
7. R Development Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria; 2008. ISBN 3-900051-07-0, Available from: <http://www.R-project.org/>.
8. Fox J. The R Commander: A Basic-Statistics GUI for R [Internet]. [cited 2011 July]. Available from: <http://socserv.mcmaster.ca/jfox/Misc/Rcmdr/>
9. Red-R Visual programming for R [Internet], [cited 2011 July]. Available from: <http://www.red-r.org/>
10. Sciviews-R [Internet], [cited 2011 July]. Available from: <http://www.sciviews.org/SciViews-R/index.html>
11. JGR – Java GUI for R [Internet] [cited 2011 July]. Available from: <http://rforge.net/JGR/index.html>
12. RKWard – An overview [Internet] [cited 2011 July]. Available from: <http://rkwad.sourceforge.net/?content=overview>
13. Scintilla A free source code editing component for Win32 and GTK+ [Internet]. [cited 2011 July] Available from: <http://www.scintilla.org>
14. The Comprehensive R Archive Network [Internet]. [cited 2011 July]. Available from: <http://cran.r-project.org/>
15. About Tcl/Tk [Internet]. [cited 2011 July] Available from: <http://www.tcl.tk/about/>
16. Hothorn T, Hornik K, Wiel MA, Zeileis A. Package “Coin” [Internet]. [cited 2011 July] Available from: <http://cran.r-project.org/web/packages/coin/coin.pdf>
17. Wilcoxon Rank Sum and Signed Rank Tests [Internet]. [cited 2011 July] Available from <http://stat.ethz.ch/R-manual/R-patched/library/stats/html/wilcox.test.html>
18. Kruskal-Wallis Rank Sum Test [Internet]. [cited 2011 July] Available from: <http://stat.ethz.ch/R-manual/R-patched/library/stats/html/kruskal.test.html>
19. Test Survival Curve Differences [Internet]. [cited 2011 July] Available from: <http://stat.ethz.ch/R-manual/R-patched/library/survival/html/survdiff.html>

20. Fligner-Killeen Test of Homogeneity of Variances [Internet]. [cited 2011 July] Available from: <http://stat.ethz.ch/R-manual/R-patched/library/stats/html/fligner.test.html>
21. Friedman Rank Sum Test [Internet]. [cited 2011 July] Available from: <http://stat.ethz.ch/R-manual/R-patched/library/stats/html/friedman.test.html>