scikit-posthocs: Pairwise multiple comparison tests in Python

Maksim A. Terpilowski[1]

1 Institute of Evolutionary Physiology and Biochemistry, Saint Petersburg, Russia

Summary

Python currently lacks implementation of many multiple pairwise (post hoc) comparison tests that are routinely performed following a statistically significant result of a parametric or nonparametric analysis of variance (ANOVA) test to assess the differences between group levels.

The `scikit-posthocs` package is aimed at filling this gap by providing a number of nonparametric and parametric pairwise comparisons tests as well as outlier detection algorithms implemented in Python:

1. Conover, Dunn, and Nemenyi tests (for use with Kruskal-Wallis test).
2. Conover, Nemenyi, Siegel, and Miller tests (for use with Friedman test).
3. Quade, van Waerden, Durbin, Student, Mann-Whitney, Wilcoxon, TukeyHSD, Anderson-Darling, Mack-Wolfe, Nashimoto and Wright (NPM), Scheffe, and Tamhane T2 tests.
4. Interquartile range (IQR), Grubbs, Tietjen-Moore, and Generalized Extreme Studentized Deviate (ESD) tests.

It also has plotting functionality to present the results of pairwise comparisons as a heatmap (significance plot).

This package is compatible with Python 2 and 3 versions, relies heavily and extends the functionality of `statsmodels`, `SciPy` and `PMCMRplus` packages (Seabold & Perktold, 2010), (Jones, Oliphant, & Peterson, 2001), (Pohlert, 2018). It is also integrated with `Pandas` (McKinney, 2010) and `Numpy` (T. E. Oliphant, 2006) for efficient computations and data analysis. The package is fully documented and comes with a Jupyter notebook example.

Figures

References

Jones, E., Oliphant, T., & Peterson, P. (2001). SciPy: Open source scientific tools for python. Retrieved from http://www.scipy.org/

McKinney, W. (2010). Data structures for statistical computing in python. In *Proceedings of the 9th python in science conference* (Vol. 445, pp. 51–56). Austin, TX.
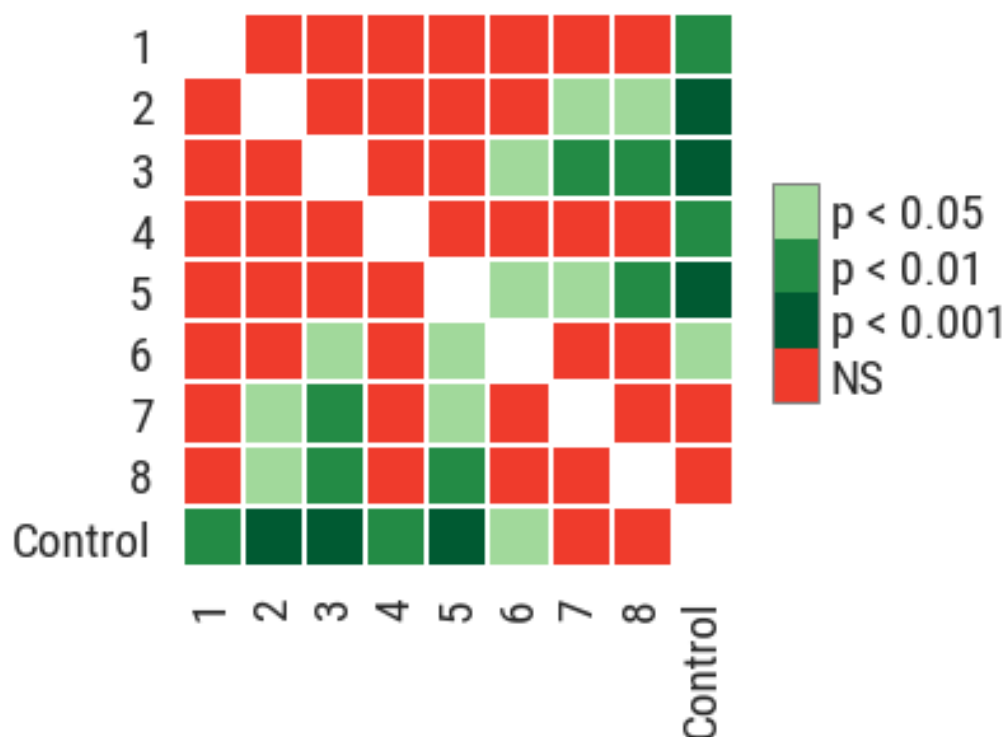
Figure 1: Significance plot

Oliphant, T. E. (2006). *A guide to numpy* (Vol. 1). Trelgol Publishing USA.

Pohlert, T. (2018). PMCMRplus: Calculate pairwise multiple comparisons of mean rank sums extended. Retrieved from https://CRAN.R-project.org/package=PMCMRplus

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In *Proceedings of the 9th python in science conference* (Vol. 57, p. 61). SciPy society Austin.