faoutlier: An R Package for Detecting Influential Cases in Exploratory and Confirmatory Factor Analysis

Applied Psychological Measurement
2015, Vol. 39(7) 573–574
© The Author(s) 2015
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0146621615597894
apm.sagepub.com

R. Philip Chalmers¹ and David B. Flora¹

Like any statistical modeling procedure, results from both exploratory and confirmatory factor analysis are vulnerable to disproportionate influence from unusual or outlying cases. Because the common factor model is a type of multiple regression model, concepts from regression case diagnostics can be applied to factor analysis. For instance, Pek and MacCallum (2011) and Flora, LaBrish, and Chalmers (2012) present several useful techniques for detecting influential cases and demonstrate their utility in applied settings. However, extant factor analysis software packages generally do not implement these important case diagnostic methods; therefore, specialized software is required.

Program Description

The R (R Core Team, 2014) package *faoutlier* implements a number of case diagnostic measures, including robust Mahalanobis distances, model-based outliers obtained using two-step approximations, goodness of fit distances, likelihood distances, generalized Cook's distances, as well as a forward search procedure to iteratively locate clusters of influential cases. The forward search procedure was described by Mavridis and Moustaki (2008), whereas descriptions of the other case diagnostic measures are in Flora et al. (2012). A selection of these procedures is useful in more general structural equation models over and above standard confirmatory factor analysis. The *faoutlier* package is compatible with fitted model objects generated from the *sem* (Fox, 2006), *lavaan* (Rosseel, 2012), and *mirt* (Chalmers, 2012) packages and provides parallel computational support with R's built-in parallel package to help reduce computation times in larger data sets. Depending on the function requested, users may print case diagnostic information to the console to obtain the numerical results or may plot the diagnostics with various generic functions. These numerical and graphical utilities help to identify influential cases that may be distorting the results of exploratory and confirmatory factor analysis procedures.

Availability, Documentation, and Distribution

The *faoutlier* package, associated documentation (Help Pages), and vignette files can be downloaded and installed from the Comprehensive R Archive Network (CRAN) at <http://cran.r-project.org/web/packages/faoutlier/index.html>.

¹York University, Toronto, Ontario, Canada

Corresponding Author:

R. Philip Chalmers, Department of Psychology, York University, 4700 Keele St., Toronto, Ontario, Canada M3J 1P3.
Email: rphilip.chalmers@gmail.com

project.org, or from within R by typing `install.packages('faoutlier')` into the command window. R software is compatible with Windows, Mac OS, and Linux platforms.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

- Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, 48(6), 1-29. Retrieved from <http://www.jstatsoft.org/v48/i06>
- Flora, D. B., LaBrish, C., & Chalmers, R. P. (2012). Old and new ideas for data screening and assumption testing for exploratory and confirmatory factor analysis. *Frontiers in Psychology*, 3, 1-21. doi: 10.3389/fpsyg.2012.00055
- Fox, J. (2006). Structural equation modeling with the sem package in R. *Structural Equation Modeling*, 13, 465-486.
- Mavridis, D., & Moustaki, I. (2008). Detecting outliers in factor analysis using the forward search algorithm. *Multivariate Behavioral Research*, 43, 453-475. doi:10.1080/00273170802285909
- Pek, J., & MacCallum, R. C. (2011). Sensitivity analysis in structural equation models: Cases and their influence. *Multivariate Behavioral Research*, 46, 202-228.
- R Core Team. (2014). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Available from <http://www.R-project.org/>
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36. Retrieved from <http://www.jstatsoft.org/v48/i02>