R Packages

Ivan Jacob Agaloos Pesigan

January 21, 2024

betaDela

- Ivan Jacob Agaloos Pesigan, Rong Wei Sun, and Shu Fai Cheung. "betaDelta and betaSandwich: Confidence intervals for standardized regression coefficients in R". in: *Multivariate Behavioral Research* (Apr. 2023), pp. 1–4. DOI: 10.1080/00273171.2023.2201277

betaSandwich

- Ivan Jacob Agaloos Pesigan, Rong Wei Sun, and Shu Fai Cheung. "betaDelta and betaSandwich: Confidence intervals for standardized regression coefficients in R". in: *Multivariate Behavioral Research* (Apr. 2023), pp. 1–4. DOI: 10.1080/00273171.2023.2201277

dynr

- Lu Ou, Michael D. Hunter, and Sy-Miin Chow. "What's for dynr: A package for linear and nonlinear dynamic modeling in R". in: *The R Journal* 11.1 (2019), p. 91. DOI: 10.32614/rj-2019-012

fungible

- Niels G. Waller. *fungible: Psychometric functions from the Waller Lab*. The R Foundation, 2022. URL: https://CRAN.R-project.org/package=fungible

lavaan

- Yves Rosseel. "lavaan: An R package for structural equation modeling". In: *Journal of Statistical Software* 48.2 (2012). DOI: 10.18637/jss.v048.i02

MASS

- W. N. Venables and B. D. Ripley. *Modern applied statistics with S*. Springer New York, 2002. DOI: 10.1007/978-0-387-21706-2

mice

- Stef van Buuren and Karin Groothuis-Oudshoorn. "mice: Multivariate Imputation by Chained Equations in R". in: *Journal of Statistical Software* 45.3 (2011). DOI: 10.18637/jss.v045.i03

OpenMx

- Michael C. Neale et al. "OpenMx 2.0: Extended Structural Equation and Statistical Modeling". In: *Psychometrika* 81.2 (Jan. 2015), pp. 535–549. DOI: 10.1007/s11336-014-9435-8

Rcpp

- Dirk Eddelbuettel. *Seamless R and C++ integration with Rcpp*. Springer New York, 2013. ISBN: 978-1-4614-6868-4. DOI: 10.1007/978-1-4614-6868-4
- Dirk Eddelbuettel and James Joseph Balamuta. "Extending R with C++: A brief introduction to Rcpp". In: *PeerJ Preprints* 3188v1.3 (Aug. 2017). DOI: 10.7287/peerj.preprints.3188v1
- Dirk Eddelbuettel and Romain François. "Rcpp: Seamless R and C++ integration". In: *Journal of Statistical Software* 40.8 (2011). DOI: 10.18637/jss.v040.i08
- Dirk Eddelbuettel, Romain Francois, et al. *Rcpp: Seamless R and C++ Integration*. 2023. URL: https://CRAN.R-project.org/package=Rcpp

RcppArmadillo

- Dirk Eddelbuettel and Conrad Sanderson. "RcppArmadillo: Accelerating R with high-performance C++ linear algebra". In: *Computational Statistics & Data Analysis* 71 (Mar. 2014), pp. 1054–1063. DOI: 10.1016/j.csda.2013.02.005

semmcci

- Ivan Jacob Agaloos Pesigan and Shu Fai Cheung. "Monte Carlo confidence intervals for the indirect effect with missing data". In: *Behavior Research Methods* (Aug. 2023). DOI: 10.3758/s13428-023-02114-4

semlbci

- Shu Fai Cheung and Ivan Jacob Agaloos Pesigan. "semlbci: An R package for forming likelihood-based confidence intervals for parameter estimates, correlations, indirect effects,

and other derived parameters". In: *Structural Equation Modeling: A Multidisciplinary Journal* (May 2023), pp. 1–15. DOI: 10.1080/10705511.2023.2183860

semTools

- Terrence D. Jorgensen et al. *semTools: Useful tools for structural equation modeling.* 2022. URL: https://CRAN.R-project.org/package=semTools

References

Cheung et al.: semlbci: An R package for forming likelihood-based confidence intervals for parameter estimates, correlations, indirect effects, and other derived parameters

Cheung-Pesigan-2023b

Shu Fai Cheung and Ivan Jacob Agaloos Pesigan. "semlbci: An R package for forming likelihood-based confidence intervals for parameter estimates, correlations, indirect effects, and other derived parameters". In: *Structural Equation Modeling: A Multidisciplinary Journal* (May 2023), pp. 1–15. DOI: 10.1080/10705511.2023.2183860.

Abstract: There are three common types of confidence interval (CI) in structural equation modeling (SEM): Wald-type CI, bootstrapping CI, and likelihood-based CI (LBCI). LBCI has the following advantages: (1) it has better coverage probabilities and Type I error rate compared to Wald-type CI when the sample size is finite; (2) it correctly tests the null hypothesis of a parameter based on likelihood ratio chi-square difference test; (3) it is less computationally intensive than bootstrapping CI; and (4) it is invariant to transformations. However, LBCI is not available in many popular SEM software packages. We developed an R package, semlbci, for forming LBCI for parameters in models fitted by lavaan, a popular open-source SEM package, such that researchers have more options in forming CIs for parameters in SEM. The package supports both unstandardized and standardized

estimates, derived parameters such as indirect effect, multisample models, and the robust LBCI proposed by Falk.

Eddelbuettel: Seamless R and C++ integration with Rcpp **Eddelbuettel-2013**

Dirk Eddelbuettel. *Seamless R and C++ integration with Rcpp*. Springer New York, 2013. ISBN: 978-1-4614-6868-4. DOI: 10.1007/978-1-4614-6868-4.

Abstract: Illustrates a range of statistical computations in R using the Rcpp package. Provides a general introduction to extending R with C++ code. Features an appendix for R users new to the C++ programming language Rcpp packages are presented in the context of useful application case studies.

Eddelbuettel et al.: Extending R with C++: A brief introduction to Rcpp **Eddelbuettel-Balamuta-2017**

Dirk Eddelbuettel and James Joseph Balamuta. "Extending R with C++: A brief introduction to Rcpp". In: *PeerJ Preprints* 3188v1.3 (Aug. 2017). DOI: 10.7287/peerj.preprints.3188v1.

Abstract: R has always provided an application programming interface (API) for extensions. Based on the C language, it uses a number of macros and other low-level constructs to exchange data structures between the R process and any dynamically-loaded component modules authors added to it. With the introduction of the Rcpp package, and its later refinements, this process has become considerably easier yet also more robust. By now, Rcpp has become the most popular extension mechanism for R. This article introduces Rcpp, and illustrates with several examples how the Rcpp Attributes mechanism in particular eases the transition of objects between R and C++ code.

Eddelbuettel et al.: Rcpp: Seamless R and C++ Integration
Eddelbuettel-Francois-Allaire-etal-2023

Dirk Eddelbuettel, Romain Francois, et al. *Rcpp: Seamless R and C++ Integration*. 2023. URL: https://CRAN.R-project.org/package=Rcpp.

Eddelbuettel et al.: Rcpp: Seamless R and C++ integration
Eddelbuettel-Francois-2011

Dirk Eddelbuettel and Romain François. "Rcpp: Seamless R and C++ integration". In: *Journal of Statistical Software* 40.8 (2011). DOI: 10.18637/jss.v040.i08.

Abstract: The Rcpp package simplifies integrating C++ code with R. It provides a consistent C++ class hierarchy that maps various types of R objects (vectors, matrices, functions, environments, ...) to dedicated C++ classes. Object interchange between R and C++ is managed by simple, flexible and extensible concepts which include broad support for C++ Standard Template Library idioms. C++ code can both be compiled, linked and loaded on the fly, or added via packages. Flexible error and exception code handling is provided. Rcpp substantially lowers the barrier for programmers wanting to combine C++ code with R.

Eddelbuettel et al.: RcppArmadillo: Accelerating R with high-performance C++ linear algebra
Eddelbuettel-Sanderson-2014

Dirk Eddelbuettel and Conrad Sanderson. "RcppArmadillo: Accelerating R with high-performance C++ linear algebra". In: *Computational Statistics & Data Analysis* 71 (Mar. 2014), pp. 1054–1063. DOI: 10.1016/j.csda.2013.02.005.

Abstract: The R statistical environment and language has demonstrated particular strengths for interactive development of statistical algorithms, as well as data modelling and visualisation. Its current implementation has an interpreter at its core which may result in a performance penalty in

comparison to directly executing user algorithms in the native machine code of the host CPU. In contrast, the C++ language has no built-in visualisation capabilities, handling of linear algebra or even basic statistical algorithms; however, user programs are converted to high-performance machine code, ahead of execution. A new method avoids possible speed penalties in R by using the Rcpp extension package in conjunction with the Armadillo C++ matrix library. In addition to the inherent performance advantages of compiled code, Armadillo provides an easy-to-use template-based meta-programming framework, allowing the automatic pooling of several linear algebra operations into one, which in turn can lead to further speedups. With the aid of Rcpp and Armadillo, conversion of linear algebra centred algorithms from R to C++ becomes straightforward. The algorithms retain the overall structure as well as readability, all while maintaining a bidirectional link with the host R environment. Empirical timing comparisons of R and C++ implementations of a Kalman filtering algorithm indicate a speedup of several orders of magnitude.

Jorgensen et al.: semTools: Useful tools for structural equation modeling

Jorgensen-Pornprasertmanit-Schoemann-etal-2022

Terrence D. Jorgensen et al. *semTools: Useful tools for structural equation modeling.* 2022. URL: https://CRAN.R-project.org/package=semTools.

Neale et al.: OpenMx 2.0: Extended Structural Equation and Statistical Modeling

Neale-Hunter-Pritikin-etal-2015

Michael C. Neale et al. "OpenMx 2.0: Extended Structural Equation and Statistical Modeling". In: *Psychometrika* 81.2 (Jan. 2015), pp. 535–549. DOI: 10.1007/s11336-014-9435-8.

Abstract: The new software package OpenMx 2.0 for structural equation and other statistical modeling is introduced and its features are described. OpenMx is evolving in a modular direction and now allows a mix-and-match computational approach that separates model expectations from fit functions and optimizers. Major backend architectural improvements include a move to swappable

open-source optimizers such as the newly written CSOLNP. Entire new methodologies such as item factor analysis and state space modeling have been implemented. New model expectation functions including support for the expression of models in LISREL syntax and a simplified multigroup expectation function are available. Ease-of-use improvements include helper functions to standardize model parameters and compute their Jacobian-based standard errors, access to model components through standard R $ mechanisms, and improved tab completion from within the R Graphical User Interface.

Ou et al.: What's for dynr: A package for linear and nonlinear dynamic modeling in R Ou-Hunter-Chow-2019

Lu Ou, Michael D. Hunter, and Sy-Miin Chow. "What's for dynr: A package for linear and nonlinear dynamic modeling in R". In: *The R Journal* 11.1 (2019), p. 91. DOI: 10.32614/rj-2019-012.

Abstract: Intensive longitudinal data in the behavioral sciences are often noisy, multivariate in nature, and may involve multiple units undergoing regime switches by showing discontinuities interspersed with continuous dynamics. Despite increasing interest in using linear and nonlinear differential/difference equation models with regime switches, there has been a scarcity of software packages that are fast and freely accessible. We have created an R package called dynr that can handle a broad class of linear and nonlinear discreteand continuous-time models, with regime-switching properties and linear Gaussian measurement functions, in C, while maintaining simple and easy-to learn model specification functions in R. We present the mathematical and computational bases used by the dynr R package, and present two illustrative examples to demonstrate the unique features of dynr.

Pesigan et al.: Monte Carlo confidence intervals for the indirect effect with missing data Pesigan-Cheung-2023

Ivan Jacob Agaloos Pesigan and Shu Fai Cheung. "Monte Carlo confidence intervals for the indirect

effect with missing data”. In: *Behavior Research Methods* (Aug. 2023). DOI: 10.3758/s13428-023-02114-4.

Abstract: Missing data is a common occurrence in mediation analysis. As a result, the methods used to construct confidence intervals around the indirect effect should consider missing data. Previous research has demonstrated that, for the indirect effect in data with complete cases, the Monte Carlo method performs as well as nonparametric bootstrap confidence intervals (see MacKinnon et al., Multivariate Behavioral Research, 39(1), 99–128, 2004; Preacher & Selig, Communication Methods and Measures, 6(2), 77–98, 2012; Tofighi & MacKinnon, Structural Equation Modeling: A Multidisciplinary Journal, 23(2), 194–205, 2015). In this manuscript, we propose a simple, fast, and accurate two-step approach for generating confidence intervals for the indirect effect, in the presence of missing data, based on the Monte Carlo method. In the first step, an appropriate method, for example, full-information maximum likelihood or multiple imputation, is used to estimate the parameters and their corresponding sampling variance-covariance matrix in a mediation model. In the second step, the sampling distribution of the indirect effect is simulated using estimates from the first step. A confidence interval is constructed from the resulting sampling distribution. A simulation study with various conditions is presented. Implications of the results for applied research are discussed.

Pesigan et al.: betaDelta and betaSandwich: Confidence intervals for standardized regression coefficients in R **Pesigan-Sun-Cheung-2023**

Ivan Jacob Agaloos Pesigan, Rong Wei Sun, and Shu Fai Cheung. “betaDelta and betaSandwich: Confidence intervals for standardized regression coefficients in R”. In: *Multivariate Behavioral Research* (Apr. 2023), pp. 1–4. DOI: 10.1080/00273171.2023.2201277.

Abstract: The multivariate delta method was used by Yuan and Chan to estimate standard errors and confidence intervals for standardized regression coefficients. Jones and Waller extended the earlier work to situations where data are nonnormal by utilizing Browne’s asymptotic distribution-free

(ADF) theory. Furthermore, Dudgeon developed standard errors and confidence intervals, employing heteroskedasticity-consistent (HC) estimators, that are robust to nonnormality with better performance in smaller sample sizes compared to Jones and Waller's ADF technique. Despite these advancements, empirical research has been slow to adopt these methodologies. This can be a result of the dearth of user-friendly software programs to put these techniques to use. We present the betaDelta and the betaSandwich packages in the R statistical software environment in this manuscript. Both the normal-theory approach and the ADF approach put forth by Yuan and Chan and Jones and Waller are implemented by the betaDelta package. The HC approach proposed by Dudgeon is implemented by the betaSandwich package. The use of the packages is demonstrated with an empirical example. We think the packages will enable applied researchers to accurately assess the sampling variability of standardized regression coefficients.

Rosseel: lavaan: An R package for structural equation modeling **Rosseel-2012**

Yves Rosseel. "lavaan: An R package for structural equation modeling". In: *Journal of Statistical Software* 48.2 (2012). DOI: 10.18637/jss.v048.i02.

Abstract: Structural equation modeling (SEM) is a vast field and widely used by many applied researchers in the social and behavioral sciences. Over the years, many software packages for structural equation modeling have been developed, both free and commercial. However, perhaps the best state-of-the-art software packages in this field are still closed-source and/or commercial. The R package lavaan has been developed to provide applied researchers, teachers, and statisticians, a free, fully open-source, but commercial-quality package for latent variable modeling. This paper explains the aims behind the development of the package, gives an overview of its most important features, and provides some examples to illustrate how lavaan works in practice.

van Buuren et al.: mice: Multivariate Imputation by Chained Equations in R vanBuuren-GroothuisOudshoorn-2011

Stef van Buuren and Karin Groothuis-Oudshoorn. "mice: Multivariate Imputation by Chained Equations in R". In: *Journal of Statistical Software* 45.3 (2011). DOI: 10.18637/jss.v045.i03.

Abstract: The R package mice imputes incomplete multivariate data by chained equations. The software mice 1.0 appeared in the year 2000 as an S-PLUS library, and in 2001 as an R package. mice 1.0 introduced predictor selection, passive imputation and automatic pooling. This article documents mice, which extends the functionality of mice 1.0 in several ways. In mice, the analysis of imputed data is made completely general, whereas the range of models under which pooling works is substantially extended. mice adds new functionality for imputing multilevel data, automatic predictor selection, data handling, post-processing imputed values, specialized pooling routines, model selection tools, and diagnostic graphs. Imputation of categorical data is improved in order to bypass problems caused by perfect prediction. Special attention is paid to transformations, sum scores, indices and interactions using passive imputation, and to the proper setup of the predictor matrix. mice can be downloaded from the Comprehensive R Archive Network. This article provides a hands-on, stepwise approach to solve applied incomplete data problems.

Venables et al.: Modern applied statistics with S Venables-Ripley-2002

W. N. Venables and B. D. Ripley. *Modern applied statistics with S*. Springer New York, 2002. DOI: 10.1007/978-0-387-21706-2.

Waller: fungible: Psychometric functions from the Waller Lab Waller-2022

Niels G. Waller. *fungible: Psychometric functions from the Waller Lab*. The R Foundation, 2022. URL: https://CRAN.R-project.org/package=fungible.