

Ivan Jacob Agaloos Pesigan

February 21, 2024

References

Andrews: Heteroskedasticity and autocorrelation consistent covariance matrix estimation **Andrews-1991**

Donald W. K. Andrews. “Heteroskedasticity and autocorrelation consistent covariance matrix estimation”. In: *Econometrica* 59.3 (May 1991), p. 817. DOI: [10.2307/2938229](https://doi.org/10.2307/2938229).

Abstract: This paper is concerned with the estimation of covariance matrices in the presence of heteroskedasticity and autocorrelation of unknown forms. Currently available estimators that are designed for this context depend upon the choice of a lag truncation parameter and a weighting scheme. Results in the literature provide a condition on the growth rate of the lag truncation parameter as $T \rightarrow \infty$ that is sufficient for consistency. No results are available, however, regarding the choice of lag truncation parameter for a fixed sample size, regarding data-dependent automatic lag truncation parameters, or regarding the choice of weighting scheme. In consequence, available estimators are not entirely operational and the relative merits of the estimators are unknown. This paper addresses these problems. The asymptotic truncated mean squared errors of estimators in a given class are determined and compared. Asymptotically optimal kernel/weighting scheme and bandwidth/lag truncation parameters are obtained using an asymptotic truncated mean squared error criterion. Using these results, data-dependent automatic bandwidth/lag truncation parameters are introduced. The finite sample properties of the estimators are analyzed via Monte Carlo simulation.

Andrews: Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space **Andrews-2000**

Donald W. K. Andrews. “Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space”. In: *Econometrica* 68.2 (Mar. 2000), pp. 399–405. DOI: [10.1111/1468-0262.00114](https://doi.org/10.1111/1468-0262.00114).

Andrews et al.: An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator **Andrews-Monahan-1992**

Donald W. K. Andrews and J. Christopher Monahan. “An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator”. In: *Econometrica* 60.4 (July 1992), p. 953. DOI: [10.2307/2951574](https://doi.org/10.2307/2951574).

Arbuckle: Full information estimation in the presence of incomplete data **Arbuckle-1996**

James L. Arbuckle. “Full information estimation in the presence of incomplete data”. In: *Advanced structural equation modeling*. Ed. by George A. Marcoulides and Randall E. Schumacker. 1996. DOI: [10.4324/9781315827414](https://doi.org/10.4324/9781315827414).

Arbuckle: Amos 27.0 user’s guide **Arbuckle-2020**

James L. Arbuckle. *Amos 27.0 user’s guide*. Chicago: IBM SPSS, 2020.

Arbuckle: Amos 28.0 user’s guide **Arbuckle-2021**

James L. Arbuckle. *Amos 28.0 user’s guide*. Chicago: IBM SPSS, 2021.

Aroian: The probability function of the product of two normally distributed variables

Aroian-1947

Leo A. Aroian. “The probability function of the product of two normally distributed variables”. In: *The Annals of Mathematical Statistics* 18.2 (June 1947), pp. 265–271. DOI: [10.1214/aoms/1177730442](https://doi.org/10.1214/aoms/1177730442).

Abstract: Let x and y follow a normal bivariate probability function with means \bar{X}, \bar{Y} , standard deviations σ_1, σ_2 , respectively, r the coefficient of correlation, and $\rho_1 = \bar{X}/\sigma_1, \rho_2 = \bar{Y}/\sigma_2$. Professor C. C. Craig [1] has found the probability function of $z = xy/\sigma_1\sigma_2$ in closed form as the difference of two integrals. For purposes of numerical computation he has expanded this result in an infinite series involving powers of z, ρ_1, ρ_2 , and Bessel functions of a certain type; in addition, he has determined the moments, semin-variants, and the moment generating function of z . However, for ρ_1 and ρ_2 large, as Craig points out, the series expansion converges very slowly. Even for ρ_1 and ρ_2 as small as 2, the expansion is unwieldy. We shall show that as ρ_1 and $\rho_2 \rightarrow \infty$, the probability function of z approaches a normal curve and in case $r = 0$ the Type III function and the Gram-Charlier Type A series are excellent approximations to the z distribution in the proper region. Numerical integration provides a substitute for the infinite series wherever the exact values of the probability function of z are needed. Some extensions of the main theorem are given in section 5 and a practical problem involving the probability function of z is solved.

Asparouhov et al.: Dynamic structural equation models

Asparouhov-Hamaker-Muthen-2017

Tihomir Asparouhov, Ellen L. Hamaker, and Bengt Muthén. “Dynamic structural equation models”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.3 (Dec. 2017), pp. 359–388. DOI: [10.1080/10705511.2017.1406803](https://doi.org/10.1080/10705511.2017.1406803).

Abstract: This article presents dynamic structural equation modeling (DSEM), which can be used to study the evolution of observed and latent variables as well as the structural equation models over time. DSEM is suitable for analyzing intensive longitudinal data where observations from multiple individuals are collected at many points in time. The modeling framework encompasses previously published DSEM models and is a comprehensive attempt to combine time-series modeling with structural equation modeling. DSEM is estimated with Bayesian methods using the Markov chain Monte Carlo Gibbs sampler and the Metropolis-Hastings sampler. We provide a detailed description of the estimation algorithm as implemented in the Mplus software package. DSEM can be used for longitudinal analysis of any duration and with any number of observations across time. Simulation studies are used to illustrate the framework and study the performance of the estimation method. Methods for evaluating model fit are also discussed.

Asparouhov et al.: Multiple imputation with Mplus

Asparouhov-Muthen-2022

Tihomir Asparouhov and Bengt O. Muthén. *Multiple imputation with Mplus*. Tech. rep. [http: www.statmodel.com](http://www.statmodel.com), 2022. URL: <http://www.statmodel.com/download/Imputations7.pdf>.

Baltes et al.: History and rationale of longitudinal research

Baltes-Nesselroade-1979

Paul B. Baltes and John R. Nesselroade. “History and rationale of longitudinal research”. In: *Longitudinal research in the study of behavior and development*. Ed. by John R. Nesselroade and Paul B. Baltes. New York, NY: Academic Press, 1979. ISBN: 012515660X.

Abstract: Within the context of developmental psychology, longitudinal research is defined and reviewed from a historical perspective. Longitudinal research is shown always to include repeated-measurement methodology as the defining attribute, with individuals being the entity under study in developmental psychology. Additional characterizations vary, depending on historical and theoretical contexts. The need for longitudinal research was recognized at least as early as the nineteenth century. Terminology and specification of rationale, however, did not appear until the second or

third decade of the twentieth century. The term longitudinal was initially identified in the context of age-based definitions of development. Recent decades, however, have seen an expansion of developmental theory beyond monolithic views to include age-irrelevant and multidirectional conceptions of the nature of development, particularly if a life-span perspective is taken. Such a pluralistic conception of behavioral development implies a more generic definition of longitudinal methodology than is associated with the traditional age-developmental view. Finally, it is important to recognize that the objective of longitudinal methodology is not only the descriptive identification of change. The objective includes explanatory goals also. Only recently has the unique strength of longitudinal research for explanatory efforts been recognized. In the second section of this chapter, a series of rationales for longitudinal research are outlined. These rationales are developed within the context of developmental psychology. They deal with (1) the direct identification of intraindividual change; (2) the identification of interindividual differences in intraindividual change; (3) the analysis of interrelationships in behavioral change; (4) the analysis of causes (determinants) of intraindividual change; and (5) the analysis of causes (determinants) of interindividual differences in intraindividual change. In a third section, selected issues in longitudinal designs and analysis are briefly reviewed. The need for complex longitudinal designs and control groups is emphasized to help counteract the rather widespread assumption that simple longitudinal studies are invariably sufficient for answering developmental questions. Furthermore, general limitations on aspects of developmental research associated with the study of assigned variables such as age, sex, or cohort are outlined. These limitations place constraints on design purity and mandate the use of and familiarity with alternative quasi-experimental designs. As an example, some of the problems associated with causal analysis involving distal (delayed, mediated) influences and the use of lagged paradigms and causal modeling are discussed.

Barnard et al.: Nonparametric standard errors and confidence intervals: Discussion

Barnard-Collins-Farewell-et-al-1981

George A. Barnard et al. “Nonparametric standard errors and confidence intervals: Discussion”. In: *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* 9.2 (1981), pp. 158–170. DOI: [10.2307/3314609](https://doi.org/10.2307/3314609).

Baron et al.: The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations

Baron-Kenny-1986

Reuben M. Baron and David A. Kenny. “The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations”. In: *Journal of Personality and Social Psychology* 51.6 (1986), pp. 1173–1182. DOI: [10.1037/0022-3514.51.6.1173](https://doi.org/10.1037/0022-3514.51.6.1173).

Abstract: In this article, we attempt to distinguish between the properties of moderator and mediator variables at a number of levels. First, we seek to make theorists and researchers aware of the importance of not using the terms moderator and mediator interchangeably by carefully elaborating, both conceptually and strategically, the many ways in which moderators and mediators differ. We then go beyond this largely pedagogical function and delineate the conceptual and strategic implications of making use of such distinctions with regard to a wide range of phenomena, including control and stress, attitudes, and personality traits. We also provide a specific compendium of analytic procedures appropriate for making the most effective use of the moderator and mediator distinction, both separately and in terms of a broader causal system that includes both moderators and mediators.

D. J. Bauer et al.: Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations

Bauer-Preacher-Gil-2006

Daniel J. Bauer, Kristopher J. Preacher, and Karen M. Gil. “Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations”. In: *Psychological Methods* 11.2 (2006), pp. 142–163. DOI: [10.1037/1082-989x.11.2.142](https://doi.org/10.1037/1082-989x.11.2.142).

Beran: The impact of the bootstrap on statistical algorithms and theory **Beran-2003**

Rudolf Beran. “The impact of the bootstrap on statistical algorithms and theory”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994972](https://doi.org/10.1214/ss/1063994972).

Abstract: Bootstrap ideas yield remarkably effective algorithms for realizing certain programs in statistics. These include the construction of (possibly simultaneous) confidence sets and tests in classical models for which exact or asymptotic distribution theory is intractable. Success of the bootstrap, in the sense of doing what is expected under a probability model for data, is not universal. Modifications to Efron’s definition of the bootstrap are needed to make the idea work for modern procedures that are not classically regular.

Biesanz et al.: Assessing mediational models: Testing and interval estimation for indirect effects

Biesanz-Falk-Savalei-2010

Jeremy C. Biesanz, Carl F. Falk, and Victoria Savalei. “Assessing mediational models: Testing and interval estimation for indirect effects”. In: *Multivariate Behavioral Research* 45.4 (Aug. 2010), pp. 661–701. DOI: [10.1080/00273171.2010.498292](https://doi.org/10.1080/00273171.2010.498292).

Abstract: Theoretical models specifying indirect or mediated effects are common in the social sciences. An indirect effect exists when an independent variable’s influence on the dependent variable is mediated through an intervening variable. Classic approaches to assessing such mediational hy-

potheses (Baron & Kenny, 1986; Sobel, 1982) have in recent years been supplemented by computationally intensive methods such as bootstrapping, the distribution of the product methods, and hierarchical Bayesian Markov chain Monte Carlo (MCMC) methods. These different approaches for assessing mediation are illustrated using data from Dunn, Biesanz, Human, and Finn (2007). However, little is known about how these methods perform relative to each other, particularly in more challenging situations, such as with data that are incomplete and/or nonnormal. This article presents an extensive Monte Carlo simulation evaluating a host of approaches for assessing mediation. We examine Type I error rates, power, and coverage. We study normal and nonnormal data as well as complete and incomplete data. In addition, we adapt a method, recently proposed in statistical literature, that does not rely on confidence intervals (CIs) to test the null hypothesis of no indirect effect. The results suggest that the new inferential method—the partial posterior p value—slightly outperforms existing ones in terms of maintaining Type I error rates while maximizing power, especially with incomplete data. Among confidence interval approaches, the bias-corrected accelerated (BCa) bootstrapping approach often has inflated Type I error rates and inconsistent coverage and is not recommended. In contrast, the bootstrapped percentile confidence interval and the hierarchical Bayesian MCMC method perform best overall, maintaining Type I error rates, exhibiting reasonable power, and producing stable and accurate coverage rates.

Blanca et al.: Skewness and kurtosis in real data samples

Blanca-Arnau-LopezMontiel-et-al-2013

Maria J. Blanca et al. “Skewness and kurtosis in real data samples”. In: *Methodology* 9.2 (May 2013), pp. 78–84. DOI: [10.1027/1614-2241/a000057](https://doi.org/10.1027/1614-2241/a000057).

Abstract: Parametric statistics are based on the assumption of normality. Recent findings suggest that Type I error and power can be adversely affected when data are non-normal. This paper aims to assess the distributional shape of real data by examining the values of the third and fourth central moments as a measurement of skewness and kurtosis in small samples. The analysis concerned 693 distributions with a sample size ranging from 10 to 30. Measures of cognitive ability and of other

psychological variables were included. The results showed that skewness ranged between -2.49 and 2.33. The values of kurtosis ranged between -1.92 and 7.41. Considering skewness and kurtosis together the results indicated that only 5.5% of distributions were close to expected values under normality. Although extreme contamination does not seem to be very frequent, the findings are consistent with previous research suggesting that normality is not the rule with real data.

Boettiger et al.: An introduction to Rocker: Docker containers for R

Boettiger-Eddelbuettel-2017

Carl Boettiger and Dirk Eddelbuettel. “An introduction to Rocker: Docker containers for R”. In: *The R Journal* 9.2 (2017), p. 527. DOI: [10.32614/rj-2017-065](https://doi.org/10.32614/rj-2017-065).

Abstract: We describe the Rocker project, which provides a widely-used suite of Docker images with customized R environments for particular tasks. We discuss how this suite is organized, and how these tools can increase portability, scaling, reproducibility, and convenience of R users and developers.

Bolger et al.: Diary methods: Capturing life as it is lived **Bolger-Davis-Rafaeli-2003**

Niall Bolger, Angelina Davis, and Eshkol Rafaeli. “Diary methods: Capturing life as it is lived”. In: *Annual Review of Psychology* 54.1 (Feb. 2003), pp. 579–616. ISSN: 1545-2085. DOI: [10.1146/annurev.psych.54.101601.145030](https://doi.org/10.1146/annurev.psych.54.101601.145030).

Abstract: In diary studies, people provide frequent reports on the events and experiences of their daily lives. These reports capture the particulars of experience in a way that is not possible using traditional designs. We review the types of research questions that diary methods are best equipped to answer, the main designs that can be used, current technology for obtaining diary reports, and appropriate data analysis strategies. Major recent developments include the use of electronic forms of data collection and multilevel models in data analysis. We identify several areas of research opportunities: 1. in technology, combining electronic diary reports with collateral measures such

as ambulatory heart rate; 2. in measurement, switching from measures based on between-person differences to those based on within-person changes; and 3. in research questions, using diaries to (a) explain why people differ in variability rather than mean level, (b) study change processes during major events and transitions, and (c) study interpersonal processes using dyadic and group diary methods.

Bollen et al.: Direct and indirect effects: Classical and bootstrap estimates of variability

Bollen-Stine-1990

Kenneth A. Bollen and Robert Stine. “Direct and indirect effects: Classical and bootstrap estimates of variability”. In: *Sociological Methodology* 20 (1990), p. 115. DOI: [10.2307/271084](https://doi.org/10.2307/271084).

Abstract: The decomposition of effects in structural equation models has been of considerable interest to social scientists. Finite-sample or asymptotic results for the sampling distribution of estimators of direct effects are widely available. Statistical inferences about indirect effects have relied exclusively on asymptotic methods which assume that the limiting distribution of the estimator is normal, with a standard error derived from the delta method. We examine bootstrap procedures as another way to generate standard errors and confidence intervals and to estimate the sampling distributions of estimators of direct and indirect effects. We illustrate the classical and the bootstrap methods with three empirical examples. We find that in a moderately large sample, the bootstrap distribution of an estimator is close to that assumed with the classical and delta methods but that in small samples, there are some differences. Bootstrap methods provide a check on the classical and delta methods when the latter are applied under less than ideal conditions.

Boos: Introduction to the bootstrap world

Boos-2003

Dennis D. Boos. “Introduction to the bootstrap world”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994971](https://doi.org/10.1214/ss/1063994971).

Abstract: The bootstrap has made a fundamental impact on how we carry out statistical inference in problems without analytic solutions. This fact is illustrated with examples and comments that emphasize the parametric bootstrap and hypothesis testing.

Bradley: Robustness?**Bradley-1978**

James V. Bradley. “Robustness?” In: *British Journal of Mathematical and Statistical Psychology* 31.2 (Nov. 1978), pp. 144–152. DOI: [10.1111/j.2044-8317.1978.tb00581.x](https://doi.org/10.1111/j.2044-8317.1978.tb00581.x).

Abstract: The actual behaviour of the probability of a Type I error under assumption violation is quite complex, depending upon a wide variety of interacting factors. Yet allegations of robustness tend to ignore its highly particularistic nature and neglect to mention important qualifying conditions. The result is often a vast overgeneralization which nevertheless is difficult to refute since a standard quantitative definition of what constitutes robustness does not exist. Yet under any halfway reasonable quantitative definition, many of the most prevalent claims of robustness would be demonstrably false. Therefore robustness is a highly questionable concept.

Brockwell et al.: Time series: Theory and methods**Brockwell-Davis-1991**

Peter J. Brockwell and Richard A. Davis. *Time series: Theory and methods*. Springer New York, 1991. ISBN: 9781441903204. DOI: [10.1007/978-1-4419-0320-4](https://doi.org/10.1007/978-1-4419-0320-4).

Abstract: This edition contains a large number of additions and corrections scattered throughout the text, including the incorporation of a new chapter on state-space models. The companion diskette for the IBM PC has expanded into the software package ITSM: An Interactive Time Series Modelling Package for the PC, which includes a manual and can be ordered from Springer-Verlag. * We are indebted to many readers who have used the book and programs and made suggestions for improvements. Unfortunately there is not enough space to acknowledge all who have contributed in this way; however, special mention must be made of our prize-winning fault-finders, Sid Resnick and F. Pukelsheim. Special mention should also be made of Anthony Brockwell, whose advice and

support on computing matters was invaluable in the preparation of the new diskettes. We have been fortunate to work on the new edition in the excellent environments provided by the University of Melbourne and Colorado State University. We thank Duane Boes particularly for his support and encouragement throughout, and the Australian Research Council and National Science Foundation for their support of research related to the new material. We are also indebted to Springer-Verlag for their constant support and assistance in preparing the second edition.

Browne: Asymptotically distribution-free methods for the analysis of covariance structures **Browne-1984**

Michael W. Browne. "Asymptotically distribution-free methods for the analysis of covariance structures". In: *British Journal of Mathematical and Statistical Psychology* 37.1 (May 1984), pp. 62–83. DOI: [10.1111/j.2044-8317.1984.tb00789.x](https://doi.org/10.1111/j.2044-8317.1984.tb00789.x).

Abstract: Methods for obtaining tests of fit of structural models for covariance matrices and estimator standard error which are asymptotically distribution free are derived. Modifications to standard normal theory tests and standard errors which make them applicable to the wider class of elliptical distributions are provided. A random sampling experiment to investigate some of the proposed methods is described.

Casella: Introduction to the silver anniversary of the bootstrap **Casella-2003**

George Casella. "Introduction to the silver anniversary of the bootstrap". In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994967](https://doi.org/10.1214/ss/1063994967).

Chen et al.: Vector autoregression, structural equation modeling, and their synthesis in neuroimaging data analysis **Chen-Daniel-Ziad-et-al-2011**

Gang Chen et al. "Vector autoregression, structural equation modeling, and their synthesis in

neuroimaging data analysis”. In: *Computers in Biology and Medicine* 41.12 (Dec. 2011), pp. 1142–1155. DOI: [10.1016/j.compbiomed.2011.09.004](https://doi.org/10.1016/j.compbiomed.2011.09.004).

Abstract: Vector autoregression (VAR) and structural equation modeling (SEM) are two popular brain-network modeling tools. VAR, which is a data-driven approach, assumes that connected regions exert time-lagged influences on one another. In contrast, the hypothesis-driven SEM is used to validate an existing connectivity model where connected regions have contemporaneous interactions among them. We present the two models in detail and discuss their applicability to fMRI data, and their interpretational limits. We also propose a unified approach that models both lagged and contemporaneous effects. The unifying model, structural vector autoregression (SVAR), may improve statistical and explanatory power, and avoid some prevalent pitfalls that can occur when VAR and SEM are utilized separately.

Cheong et al.: Investigation of mediational processes using parallel process latent growth curve modeling **Cheong-MacKinnon-Khoo-2003**

JeeWon Cheong, David P. MacKinnon, and Siek Toon Khoo. “Investigation of mediational processes using parallel process latent growth curve modeling”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 10.2 (Apr. 2003), pp. 238–262. DOI: [10.1207/s15328007sem1002_5](https://doi.org/10.1207/s15328007sem1002_5).

Abstract: This study investigated a method to evaluate mediational processes using latent growth curve modeling. The mediator and the outcome measured across multiple time points were viewed as 2 separate parallel processes. The mediational process was defined as the independent variable influencing the growth of the mediator, which, in turn, affected the growth of the outcome. To illustrate modeling procedures, empirical data from a longitudinal drug prevention program, Adolescents Training and Learning to Avoid Steroids, were used. The program effects on the growth of the mediator and the growth of the outcome were examined first in a 2-group structural equation model. The mediational process was then modeled and tested in a parallel process latent growth

curve model by relating the prevention program condition, the growth rate factor of the mediator, and the growth rate factor of the outcome.

Chesher et al.: The bias of a heteroskedasticity consistent covariance matrix estimator

Chesher-Jewitt-1987

Andrew Chesher and Ian Jewitt. “The bias of a heteroskedasticity consistent covariance matrix estimator”. In: *Econometrica* 55.5 (Sept. 1987), p. 1217. DOI: [10.2307/1911269](https://doi.org/10.2307/1911269).

G. W. Cheung et al.: Testing mediation and suppression effects of latent variables

Cheung-Lau-2007

Gordon W. Cheung and Rebecca S. Lau. “Testing mediation and suppression effects of latent variables”. In: *Organizational Research Methods* 11.2 (July 2007), pp. 296–325. DOI: [10.1177/1094428107300343](https://doi.org/10.1177/1094428107300343).

Abstract: Because of the importance of mediation studies, researchers have been continuously searching for the best statistical test for mediation effect. The approaches that have been most commonly employed include those that use zero-order and partial correlation, hierarchical regression models, and structural equation modeling (SEM). This study extends MacKinnon and colleagues (MacKinnon, Lockwood, Hoffmann, West, & Sheets, 2002; MacKinnon, Lockwood, & Williams, 2004, MacKinnon, Warsi, & Dwyer, 1995) works by conducting a simulation that examines the distribution of mediation and suppression effects of latent variables with SEM, and the properties of confidence intervals developed from eight different methods. Results show that SEM provides unbiased estimates of mediation and suppression effects, and that the bias-corrected bootstrap confidence intervals perform best in testing for mediation and suppression effects. Steps to implement the recommended procedures with Amos are presented.

M. W.-L. Cheung: Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models
Cheung-2007

Mike W.-L. Cheung. "Comparison of approaches to constructing confidence intervals for mediating effects using structural equation models". In: *Structural Equation Modeling: A Multidisciplinary Journal* 14.2 (May 2007), pp. 227–246. DOI: [10.1080/10705510709336745](https://doi.org/10.1080/10705510709336745).

Abstract: Mediators are variables that explain the association between an independent variable and a dependent variable. Structural equation modeling (SEM) is widely used to test models with mediating effects. This article illustrates how to construct confidence intervals (CIs) of the mediating effects for a variety of models in SEM. Specifically, mediating models with 1 mediator, 2 intermediate mediators, 2 specific mediators, and 1 mediator in 2 independent groups are illustrated. By using phantom variables (Rindskopf, 1984), a Wald CI, percentile bootstrap CI, bias-corrected bootstrap CI, and a likelihood-based CI on the mediating effect are easily constructed with some existing SEM packages, such as LISREL, Mplus, and Mx. Monte Carlo simulation studies are used to compare the coverage probabilities of these CIs. The results show that the coverage probabilities of these CIs are comparable when the mediating effect is large or when the sample size is large. However, when the mediating effect and the sample size are both small, the bootstrap CI and likelihood-based CI are preferred over the Wald CI. Extensions of this SEM approach for future research are discussed.

M. W.-L. Cheung: Comparison of methods for constructing confidence intervals of standardized indirect effects
Cheung-2009a

Mike W.-L. Cheung. "Comparison of methods for constructing confidence intervals of standardized indirect effects". In: *Behavior Research Methods* 41.2 (May 2009), pp. 425–438. DOI: [10.3758/brm.41.2.425](https://doi.org/10.3758/brm.41.2.425).

Abstract: Mediation models are often used as a means to explain the psychological mechanisms between an independent and a dependent variable in the behavioral and social sciences. A major limitation of the unstandardized indirect effect calculated from raw scores is that it cannot be

interpreted as an effect-size measure. In contrast, the standardized indirect effect calculated from standardized scores can be a good candidate as a measure of effect size because it is scale invariant. In the present article, 11 methods for constructing the confidence intervals (CIs) of the standardized indirect effects were evaluated via a computer simulation. These included six Wald CIs, three bootstrap CIs, one likelihood-based CI, and the PRODCLIN CI. The results consistently showed that the percentile bootstrap, the bias-corrected bootstrap, and the likelihood-based approaches had the best coverage probability. Mplus, LISREL, and Mx syntax were included to facilitate the use of these preferred methods in applied settings. Future issues on the use of the standardized indirect effects are discussed.

M. W.-L. Cheung: Constructing approximate confidence intervals for parameters with structural equation models **Cheung-2009b**

Mike W.-L. Cheung. "Constructing approximate confidence intervals for parameters with structural equation models". In: *Structural Equation Modeling: A Multidisciplinary Journal* 16.2 (Apr. 2009), pp. 267–294. DOI: [10.1080/10705510902751291](https://doi.org/10.1080/10705510902751291).

Abstract: Confidence intervals (CIs) for parameters are usually constructed based on the estimated standard errors. These are known as Wald CIs. This article argues that likelihood-based CIs (CIs based on likelihood ratio statistics) are often preferred to Wald CIs. It shows how the likelihood-based CIs and the Wald CIs for many statistics and psychometric indexes can be constructed with the use of phantom variables (Rindskopf, 1984) in some of the current structural equation modeling (SEM) packages. The procedures to form CIs for the differences in correlation coefficients, squared multiple correlations, indirect effects, coefficient alphas, and reliability estimates are illustrated. A simulation study on the Pearson correlation is used to demonstrate the advantages of the likelihood-based CI over the Wald CI. Issues arising from this SEM approach and extensions of this approach are discussed.

Mike W.-L. Cheung. “Synthesizing indirect effects in mediation models with meta-analytic methods”. In: *Alcohol and Alcoholism* 57.1 (June 2021), pp. 5–15. DOI: [10.1093/alcalc/agab044](https://doi.org/10.1093/alcalc/agab044).

Abstract: Aims A mediator is a variable that explains the underlying mechanism between an independent variable and a dependent variable. The indirect effect indicates the effect from the predictor to the outcome variable via the mediator. In contrast, the direct effect represents the predictor’s effort on the outcome variable after controlling for the mediator. **Methods** A single study rarely provides enough evidence to answer research questions in a particular domain. Replications are generally recommended as the gold standard to conduct scientific research. When a sufficient number of studies have been conducted addressing similar research questions, a meta-analysis can be used to synthesize those studies’ findings. **Results** The main objective of this paper is to introduce two frameworks to integrating studies using mediation analysis. The first framework involves calculating standardized indirect effects and direct effects and conducting a multivariate meta-analysis on those effect sizes. The second one uses meta-analytic structural equation modeling to synthesize correlation matrices and fit mediation models on the average correlation matrix. We illustrate these procedures on a real dataset using the R statistical platform. **Conclusion** This paper closes with some further directions for future studies.

Shu Fai Cheung and Ivan Jacob Agaloos Pesigan. “FINDOUT: Using either SPSS commands or graphical user interface to identify influential cases in structural equation modeling in AMOS”. In: *Multivariate Behavioral Research* (Jan. 2023), pp. 1–5. DOI: [10.1080/00273171.2022.2148089](https://doi.org/10.1080/00273171.2022.2148089).

Abstract: The results in a structural equation modeling (SEM) analysis can be influenced by just a few observations, called influential cases. Tools have been developed for users of R to identify them. However, similar tools are not available for AMOS, which is also a popular SEM software package. We introduce the FINDOUT toolset, a group of SPSS extension commands, and an AMOS plugin, to identify influential cases and examine how these cases influence the results. The SPSS commands can be used either as syntax commands or as custom dialogs from pull-down menus, and the AMOS plugin can be run from AMOS pull-down menu. We believe these tools can help researchers to examine the robustness of their findings to influential cases.

S. F. 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](https://doi.org/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.

S. F. Cheung et al.: DIY bootstrapping: Getting the nonparametric bootstrap confidence interval in SPSS for any statistics or function of statistics (when this bootstrapping is appropriate)
Cheung-Pesigan-Vong-2022

Shu Fai Cheung, Ivan Jacob Agaloos Pesigan, and Weng Ngai Vong. “DIY bootstrapping: Getting the nonparametric bootstrap confidence interval in SPSS for any statistics or function of statistics (when this bootstrapping is appropriate)”. In: *Behavior Research Methods* 55.2 (Mar. 2022), pp. 474–490. DOI: [10.3758/s13428-022-01808-5](https://doi.org/10.3758/s13428-022-01808-5).

Abstract: Researchers can generate bootstrap confidence intervals for some statistics in SPSS using the BOOTSTRAP command. However, this command can only be applied to selected procedures, and only to selected statistics in these procedures. We developed an extension command and prepared some sample syntax files based on existing approaches from the Internet to illustrate how researchers can (a) generate a large number of nonparametric bootstrap samples, (b) do desired analysis on all these samples, and (c) form the bootstrap confidence intervals for selected statistics using the OMS commands. We developed these tools to help researchers apply nonparametric bootstrapping to any statistics for which this method is appropriate, including statistics derived from other statistics, such as standardized effect size measures computed from the t test results. We also discussed how researchers can extend the tools for other statistics and scenarios they encounter.

Chow et al.: Age differences in dynamical emotion-cognition linkages

Chow-Hamagani-Nesselroade-2007

Sy-Miin Chow, Fumiaki Hamagani, and John R. Nesselroade. “Age differences in dynamical emotion-cognition linkages”. In: *Psychology and Aging* 22.4 (Dec. 2007), pp. 765–780. ISSN: 0882-7974. DOI: [10.1037/0882-7974.22.4.765](https://doi.org/10.1037/0882-7974.22.4.765).

Abstract: The ability to maintain the separation between positive emotion and negative emotion in times of stress has been construed as a resilience mechanism. Emotional resiliency is particularly relevant in old age given concomitant declines in cognitive performance. In the present study, the

authors examined the dynamical linkages among positive emotion, negative emotion, and cognition as individuals performed a complex cognitive task. Comparisons were made between younger ($n = 63$) and older ($n = 52$) age groups. Older adults manifested significant unidirectional coupling from negative emotion to cognitive performance; younger adults manifested significant unidirectional coupling from negative emotion to positive emotion and from cognitive performance to both positive and negative emotions. Implications for age differences in emotion regulatory strategies are discussed.

Chow et al.: Equivalence and differences between structural equation modeling and state-space modeling techniques **Chow-Ho-Hamaker-et-al-2010**

Sy-Miin Chow, Moon-ho R. Ho, et al. “Equivalence and differences between structural equation modeling and state-space modeling techniques”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 17.2 (Apr. 2010), pp. 303–332. DOI: [10.1080/10705511003661553](https://doi.org/10.1080/10705511003661553).

Abstract: State-space modeling techniques have been compared to structural equation modeling (SEM) techniques in various contexts but their unique strengths have often been overshadowed by their similarities to SEM. In this article, we provide a comprehensive discussion of these 2 approaches’ similarities and differences through analytic comparisons and numerical simulations, with a focus on their use in representing intraindividual dynamics and interindividual differences. To demonstrate the respective strengths and weaknesses of the 2 approaches in representing these 2 aspects, we simulated data under (a) a cross-sectional common factor model, (b) a latent difference score model with random effects in intercept and slope, and (c) a bivariate dynamic factor analysis model with auto- and cross-regression parameters. Possible ways in which SEM and state-space modeling can be utilized as complementary tools in representing human developmental and other related processes are discussed.

Chow et al.: Continuous-time dynamic models: Connections to structural equation models and other discrete-time models **Chow-Losardo-Park-et-al-2023**

Sy-Miin Chow, Diane Losardo, et al. “Continuous-time dynamic models: Connections to structural equation models and other discrete-time models”. In: *Handbook of structural equation modeling*. Ed. by Rick H. Hoyle. 2nd ed. New York: The Guilford Press, 2023. ISBN: 9781462550722.

Cochran: The χ^2 test of goodness of fit **Cochran-1952**

William G. Cochran. “The χ^2 test of goodness of fit”. In: *The Annals of Mathematical Statistics* 23.3 (Sept. 1952), pp. 315–345. DOI: [10.1214/aoms/1177729380](https://doi.org/10.1214/aoms/1177729380).

Abstract: This paper contains an expository discussion of the chi square test of goodness of fit, intended for the student and user of statistical theory rather than for the expert. Part I describes the historical development of the distribution theory on which the test rests. Research bearing on the practical application of the test—in particular on the minimum expected number per class and the construction of classes—is discussed in Part II. Some varied opinions about the extent to which the test actually is useful to the scientist are presented in Part III. Part IV outlines a number of tests that have been proposed as substitutes for the chi square test (the ω^2 test, the smooth test, the likelihood ratio test) and Part V a number of supplementary tests (the run test, tests based on low moments, subdivision of chi square into components).

Cohen: Statistical power analysis for the behavioral sciences **Cohen-1988**

Jacob Cohen. *Statistical power analysis for the behavioral sciences*. 2nd ed. Routledge, 1988. ISBN: 9780203771587. DOI: [10.4324/9780203771587](https://doi.org/10.4324/9780203771587).

Abstract: Statistical Power Analysis is a nontechnical guide to power analysis in research planning that provides users of applied statistics with the tools they need for more effective analysis. The Second Edition includes:

- a chapter covering power analysis in set correlation and multivariate methods;
- a chapter considering effect size, psychometric reliability, and the efficacy of “qualifying” dependent variables and;
- expanded power and sample size tables for multiple regression/correlation.

Library: HA29 .C66 1988.

Cole et al.: Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling. **Cole-Maxwell-2003**

David A. Cole and Scott E. Maxwell. “Testing mediational models with longitudinal data: Questions and tips in the use of structural equation modeling.” In: *Journal of Abnormal Psychology* 112.4 (Nov. 2003), pp. 558–577. DOI: [10.1037/0021-843x.112.4.558](https://doi.org/10.1037/0021-843x.112.4.558).

Abstract: R. M. Baron and D. A. Kenny (1986; see record 1987-13085-001) provided clarion conceptual and methodological guidelines for testing mediational models with cross-sectional data. Graduating from cross-sectional to longitudinal designs enables researchers to make more rigorous inferences about the causal relations implied by such models. In this transition, misconceptions and erroneous assumptions are the norm. First, we describe some of the questions that arise (and misconceptions that sometimes emerge) in longitudinal tests of mediational models. We also provide a collection of tips for structural equation modeling (SEM) of mediational processes. Finally, we suggest a series of 5 steps when using SEM to test mediational processes in longitudinal designs: testing the measurement model, testing for added components, testing for omitted paths, testing the stationarity assumption, and estimating the mediational effects.

Collins et al.: Best methods for the analysis of change: Recent advances, unanswered questions, future directions **Collins-Horn-1991**

Linda M. Collins and John L. Horn, eds. *Best methods for the analysis of change: Recent advances, unanswered questions, future directions*. Washington, DC: American Psychological Association, 1991. ISBN: 978-1-55798-113-4. DOI: [10.1037/10099-000](https://doi.org/10.1037/10099-000).

Abstract: The chapters making up this book represent a rich offering of current research on the analysis of change.

Library: BF637.C4 B48 1991.

Collins et al.: New methods for the analysis of change **Collins-Sayer-2002**

Linda M. Collins and Aline Sayer, eds. *New methods for the analysis of change. [based on a conference held in 1998 at The Pennsylvania State University, a follow-up to the Los Angeles conference Best Methods for the Analysis of Change]*. 2. print. Decade of behavior. Washington, DC: American Psychological Association, 2002. 442 pp. ISBN: 1557987548.

Craig: On the frequency function of xy **Craig-1936**

Cecil C. Craig. "On the frequency function of xy ". In: *The Annals of Mathematical Statistics* 7.1 (Mar. 1936), pp. 1–15. DOI: [10.1214/aoms/1177732541](https://doi.org/10.1214/aoms/1177732541).

Cribari-Neto: Asymptotic inference under heteroskedasticity of unknown form **CribariNeto-2004**

Francisco Cribari-Neto. "Asymptotic inference under heteroskedasticity of unknown form". In: *Computational Statistics & Data Analysis* 45.2 (Mar. 2004), pp. 215–233. DOI: [10.1016/S0167-9473\(02\)00366-3](https://doi.org/10.1016/S0167-9473(02)00366-3).

Abstract: We focus on the finite-sample behavior of heteroskedasticity-consistent covariance matrix estimators and associated quasi- t tests. The estimator most commonly used is that proposed by Halbert White. Its finite-sample behavior under both homoskedasticity and heteroskedasticity is analyzed using Monte Carlo methods. We also consider two other consistent estimators, namely: the HC3 estimator, which is an approximation to the jackknife estimator, and the weighted bootstrap estimator. Additionally, we evaluate the finite-sample behavior of two bootstrap quasi- t tests: the test based on a single bootstrapping scheme and the test based on a double, nested bootstrapping scheme. The latter is very computer-intensive, but proves to work well in small samples. Finally, we propose a new estimator, which we call HC4; it is tailored to take into account the effect of leverage points in the design matrix on associated quasi- t tests.

Cribari-Neto et al.: A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model **CribariNeto-daSilva-2010**

Francisco Cribari-Neto and Wilton Bernardino da Silva. “A new heteroskedasticity-consistent covariance matrix estimator for the linear regression model”. In: *AStA Advances in Statistical Analysis* 95.2 (Nov. 2010), pp. 129–146. DOI: [10.1007/s10182-010-0141-2](https://doi.org/10.1007/s10182-010-0141-2).

Abstract: The assumption that all random errors in the linear regression model share the same variance (homoskedasticity) is often violated in practice. The ordinary least squares estimator of the vector of regression parameters remains unbiased, consistent and asymptotically normal under unequal error variances. Many practitioners then choose to base their inferences on such an estimator. The usual practice is to couple it with an asymptotically valid estimation of its covariance matrix, and then carry out hypothesis tests that are valid under heteroskedasticity of unknown form. We use numerical integration methods to compute the exact null distributions of some quasi- t test statistics, and propose a new covariance matrix estimator. The numerical results favor testing inference based on the estimator we propose.

Cribari-Neto et al.: Inference under heteroskedasticity and leveraged data

CribariNeto-Souza-Vasconcellos-2007

Francisco Cribari-Neto, Tatiene C. Souza, and Klaus L. P. Vasconcellos. “Inference under heteroskedasticity and leveraged data”. In: *Communications in Statistics - Theory and Methods* 36.10 (Aug. 2007), pp. 1877–1888. DOI: [10.1080/03610920601126589](https://doi.org/10.1080/03610920601126589).

Abstract: We evaluate the finite-sample behavior of different heteroskedasticity-consistent covariance matrix estimators, under both constant and unequal error variances. We consider the estimator proposed by Halbert White (HC0), and also its variants known as HC2, HC3, and HC4; the latter was recently proposed by Cribari-Neto (2004). We propose a new covariance matrix estimator: HC5. It is the first consistent estimator to explicitly take into account the effect that the maximal leverage has on the associated inference. Our numerical results show that quasi- t inference based on HC5 is typically more reliable than inference based on other covariance matrix estimators.

Cribari-Neto et al.: Errata: Inference under heteroskedasticity and leveraged data, Communications in Statistics, Theory and Methods, 36, 1877–1888, 2007

CribariNeto-Souza-Vasconcellos-2008

Francisco Cribari-Neto, Tatiene C. Souza, and Klaus L. P. Vasconcellos. “Errata: Inference under heteroskedasticity and leveraged data, Communications in Statistics, Theory and Methods, 36, 1877–1888, 2007”. In: *Communications in Statistics - Theory and Methods* 37.20 (Sept. 2008), pp. 3329–3330. DOI: [10.1080/03610920802109210](https://doi.org/10.1080/03610920802109210).

Cronbach et al.: How we should measure “change”: Or should we?

Cronbach-Furby-1970

Lee J. Cronbach and Lita Furby. “How we should measure “change”: Or should we?” In: *Psychological Bulletin* 74.1 (July 1970), pp. 68–80. DOI: [10.1037/h0029382](https://doi.org/10.1037/h0029382).

Abstract: Examines procedures previously recommended by various authors for the estimation of “change” scores, “residual,” or “basefree” measures of change, and other kinds of difference scores. A procedure proposed by F. M. Lord is extended to obtain more precise estimates, and an alternative to the L. R. Tucker, F. Damarin, and S. A. Messick (see 41:3) procedure is offered. A consideration of the purposes for which change measures have been sought in the past leads to a series of recommended procedures which solve research and personnel-decision problems without estimation of change scores for individuals.

Curran et al.: The disaggregation of within-person and between-person effects in longitudinal models of change **Curran-Bauer-2011**

Patrick J. Curran and Daniel J. Bauer. “The disaggregation of within-person and between-person effects in longitudinal models of change”. In: *Annual Review of Psychology* 62.1 (Jan. 2011), pp. 583–619. DOI: [10.1146/annurev.psych.093008.100356](https://doi.org/10.1146/annurev.psych.093008.100356).

Abstract: Longitudinal models are becoming increasingly prevalent in the behavioral sciences, with key advantages including increased power, more comprehensive measurement, and establishment of temporal precedence. One particularly salient strength offered by longitudinal data is the ability to disaggregate between-person and within-person effects in the regression of an outcome on a time-varying covariate. However, the ability to disaggregate these effects has not been fully capitalized upon in many social science research applications. Two likely reasons for this omission are the general lack of discussion of disaggregating effects in the substantive literature and the need to overcome several remaining analytic challenges that limit existing quantitative methods used to isolate these effects in practice. This review explores both substantive and quantitative issues related to the disaggregation of effects over time, with a particular emphasis placed on the multilevel model. Existing analytic methods are reviewed, a general approach to the problem is proposed, and both the existing and proposed methods are demonstrated using several artificial data sets. Potential limitations and directions for future research are discussed, and recommendations for the disaggregation of effects in practice are offered.

Davidson et al.: Estimation and inference in econometrics Davidson-MacKinnon-1993

Russell Davidson and James G. MacKinnon. *Estimation and inference in econometrics*. New York, NY: Oxford University Press, 1993. ISBN: 9780195060119.

Library: HB139 .D368 1993.

Davison et al.: Bootstrap methods and their application Davison-Hinkley-1997

Anthony Christopher Davison and David Victor Hinkley. *Bootstrap methods and their application*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge and New York, NY, USA: Cambridge University Press, 1997. ISBN: 9780521573917. DOI: [10.1017/CB09780511802843](https://doi.org/10.1017/CB09780511802843).

Abstract: Bootstrap methods are computer-intensive methods of statistical analysis, which use simulation to calculate standard errors, confidence intervals, and significance tests. The methods apply for any level of modelling, and so can be used for fully parametric, semiparametric, and completely nonparametric analysis. This 1997 book gives a broad and up-to-date coverage of bootstrap methods, with numerous applied examples, developed in a coherent way with the necessary theoretical basis. Applications include stratified data; finite populations; censored and missing data; linear, nonlinear, and smooth regression models; classification; time series and spatial problems. Special features of the book include: extensive discussion of significance tests and confidence intervals; material on various diagnostic methods; and methods for efficient computation, including improved Monte Carlo simulation. Each chapter includes both practical and theoretical exercises. S-Plus programs for implementing the methods described in the text are available from the supporting website.

Library: QA276.8 .D38 1997.

Anthony Christopher Davison, David Victor Hinkley, and George Alastair Young. “Recent developments in bootstrap methodology”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994969](https://doi.org/10.1214/ss/1063994969).

Abstract: Ever since its introduction, the bootstrap has provided both a powerful set of solutions for practical statisticians, and a rich source of theoretical and methodological problems for statistics. In this article, some recent developments in bootstrap methodology are reviewed and discussed. After a brief introduction to the bootstrap, we consider the following topics at varying levels of detail: the use of bootstrapping for highly accurate parametric inference; theoretical properties of nonparametric bootstrapping with unequal probabilities; subsampling and the m out of n bootstrap; bootstrap failures and remedies for superefficient estimators; recent topics in significance testing; bootstrap improvements of unstable classifiers and resampling for dependent data. The treatment is telegraphic rather than exhaustive.

de Haan-Rietdijk et al.: Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data **HaanRietdijk-Voelkle-Keijsers-Hamaker-2017**

Silvia de Haan-Rietdijk et al. “Discrete- vs. continuous-time modeling of unequally spaced experience sampling method data”. In: *Frontiers in Psychology* 8 (Oct. 2017). ISSN: 1664-1078. DOI: [10.3389/fpsyg.2017.01849](https://doi.org/10.3389/fpsyg.2017.01849).

Abstract: The Experience Sampling Method is a common approach in psychological research for collecting intensive longitudinal data with high ecological validity. One characteristic of ESM data is that it is often unequally spaced, because the measurement intervals within a day are deliberately varied, and measurement continues over several days. This poses a problem for discrete-time (DT) modeling approaches, which are based on the assumption that all measurements are equally

spaced. Nevertheless, DT approaches such as (vector) autoregressive modeling are often used to analyze ESM data, for instance in the context of affective dynamics research. There are equivalent continuous-time (CT) models, but they are more difficult to implement. In this paper we take a pragmatic approach and evaluate the practical relevance of the violated model assumption in DT AR(1) and VAR(1) models, for the $N = 1$ case. We use simulated data under an ESM measurement design to investigate the bias in the parameters of interest under four different model implementations, ranging from the true CT model that accounts for all the exact measurement times, to the crudest possible DT model implementation, where even the nighttime is treated as a regular interval. An analysis of empirical affect data illustrates how the differences between DT and CT modeling can play out in practice. We find that the size and the direction of the bias in DT (V)AR models for unequally spaced ESM data depend quite strongly on the true parameter in addition to data characteristics. Our recommendation is to use CT modeling whenever possible, especially now that new software implementations have become available.

Deboeck et al.: No need to be discrete: A method for continuous time mediation analysis **Deboeck-Preacher-2015**

Pascal R. Deboeck and Kristopher J. Preacher. “No need to be discrete: A method for continuous time mediation analysis”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 23.1 (June 2015), pp. 61–75. DOI: [10.1080/10705511.2014.973960](https://doi.org/10.1080/10705511.2014.973960).

Abstract: Mediation is one concept that has shaped numerous theories. The list of problems associated with mediation models, however, has been growing. Mediation models based on cross-sectional data can produce unexpected estimates, so much so that making longitudinal or causal inferences is inadvisable. Even longitudinal mediation models have faults, as parameter estimates produced by these models are specific to the lag between observations, leading to much debate over appropriate lag selection. Using continuous time models (CTMs) rather than commonly employed discrete time models, one can estimate lag-independent parameters. We demonstrate methodology that allows for continuous time mediation analyses, with attention to concepts such as indirect and direct effects,

partial mediation, the effect of lag, and the lags at which relations become maximal. A simulation compares common longitudinal mediation methods with CTMs. Reanalysis of a published covariance matrix demonstrates that CTMs can be fit to data used in longitudinal mediation studies.

Demeshko et al.: A novel continuous and structural VAR modeling approach and its application to reactor noise analysis **Demeshko-Washio-Kawahara-et-al-2015**

Marina Demeshko et al. “A novel continuous and structural VAR modeling approach and its application to reactor noise analysis”. In: *ACM Transactions on Intelligent Systems and Technology* 7.2 (Nov. 2015), pp. 1–22. DOI: [10.1145/2710025](https://doi.org/10.1145/2710025).

Abstract: A vector autoregressive model in discrete time domain (DVAR) is often used to analyze continuous time, multivariate, linear Markov systems through their observed time series data sampled at discrete timesteps. Based on previous studies, the DVAR model is supposed to be a non-canonical representation of the system, that is, it does not correspond to a unique system bijectively. However, in this article, we characterize the relations of the DVAR model with its corresponding Structural Vector AR (SVAR) and Continuous Time Vector AR (CTVAR) models through a finite difference method across continuous and discrete time domain. We further clarify that the DVAR model of a continuous time, multivariate, linear Markov system is canonical under a highly generic condition. Our analysis shows that we can uniquely reproduce its SVAR and CTVAR models from the DVAR model. Based on these results, we propose a novel Continuous and Structural Vector Autoregressive (CSVAR) modeling approach to derive the SVAR and the CTVAR models from their DVAR model empirically derived from the observed time series of continuous time linear Markov systems. We demonstrate its superior performance through some numerical experiments on both artificial and real-world data.

Didier et al.: Signal processing and machine learning with transdermal alcohol concentration to predict natural environment alcohol consumption.

Didier-King-Polley-et-al-2023

Nathan A. Didier et al. "Signal processing and machine learning with transdermal alcohol concentration to predict natural environment alcohol consumption." In: *Experimental and Clinical Psychopharmacology* (Oct. 2023). ISSN: 1064-1297. DOI: [10.1037/pha0000683](https://doi.org/10.1037/pha0000683).

Abstract: Wrist-worn alcohol biosensors continuously and discreetly record transdermal alcohol concentration (TAC) and may allow alcohol researchers to monitor alcohol consumption in participants' natural environments. However, the field lacks established methods for signal processing and detecting alcohol events using these devices. We developed software that streamlines analysis of raw data (TAC, temperature, and motion) from a wrist-worn alcohol biosensor (BACtrack Skyn) through a signal processing and machine learning pipeline: biologically implausible skin surface temperature readings ($> 28^{\circ}\text{C}$) were screened for potential device removal and TAC artifacts were corrected, features that describe TAC (e.g., rise duration) were calculated and used to train models (random forest and logistic regression) that predict self-reported alcohol consumption, and model performances were measured and summarized in autogenerated reports. The software was tested using 60 Skyn data sets recorded during 30 alcohol drinking episodes and 30 nonalcohol drinking episodes. Participants ($N = 36$; 13 with alcohol use disorder) wore the Skyn during one alcohol drinking episode and one nonalcohol drinking episode in their natural environment. In terms of distinguishing alcohol from nonalcohol drinking, correcting artifacts in the data resulted in 10% improvement in model accuracy relative to using raw data. Random forest and logistic regression models were both accurate, correctly predicting 97% (58/60; AUC-ROCs = 0.98, 0.96) of episodes. Area under TAC curve, rise duration of TAC curve, and peak TAC were the most important features for predictive accuracy. With promising model performance, this protocol will enhance the efficiency and reliability of TAC sensors for future alcohol monitoring research.

Dudgeon: Some improvements in confidence intervals for standardized regression coefficients **Dudgeon-2017**

Paul Dudgeon. “Some improvements in confidence intervals for standardized regression coefficients”. In: *Psychometrika* 82.4 (Mar. 2017), pp. 928–951. DOI: [10.1007/s11336-017-9563-z](https://doi.org/10.1007/s11336-017-9563-z).

Abstract: Yuan and Chan (Psychometrika 76:670-690, 2011. doi:10.1007/S11336-011-9224-6) derived consistent confidence intervals for standardized regression coefficients under fixed and random score assumptions. Jones and Waller (Psychometrika 80:365-378, 2015. doi:10.1007/S11336-013-9380-Y) extended these developments to circumstances where data are non-normal by examining confidence intervals based on Browne’s (Br J Math Stat Psychol 37:62-83, 1984. doi:10.1111/j.2044-8317.1984.tb00789.x) asymptotic distribution-free (ADF) theory. Seven different heteroscedastic-consistent (HC) estimators were investigated in the current study as potentially better solutions for constructing confidence intervals on standardized regression coefficients under non-normality. Normal theory, ADF, and HC estimators were evaluated in a Monte Carlo simulation. Findings confirmed the superiority of the HC3 (MacKinnon and White, J Econ 35:305-325, 1985. doi:10.1016/0304-4076(85)90158-7) and HC5 (Cribari-Neto and Da Silva, Adv Stat Anal 95:129-146, 2011. doi:10.1007/s10182-010-0141-2) interval estimators over Jones and Waller’s ADF estimator under all conditions investigated, as well as over the normal theory method. The HC5 estimator was more robust in a restricted set of conditions over the HC3 estimator. Some possible extensions of HC estimators to other effect size measures are considered for future developments.

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](https://doi.org/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](https://doi.org/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-et-al-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](https://doi.org/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](https://doi.org/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.

Efron: Bootstrap methods: Another look at the jackknife**Efron-1979a**

Bradley Efron. “Bootstrap methods: Another look at the jackknife”. In: *The Annals of Statistics* 7.1 (Jan. 1979). DOI: [10.1214/aos/1176344552](https://doi.org/10.1214/aos/1176344552).

Abstract: We discuss the following problem: given a random sample $\mathbf{X} = (X_1, X_2, \dots, X_n)$ from an unknown probability distribution F , estimate the sampling distribution of some prespecified random variable $R(\mathbf{X}, F)$, on the basis of the observed data \mathbf{x} . (Standard jackknife theory gives an approximate mean and variance in the case $R(\mathbf{X}, F) = \theta(\hat{F}) - \theta(F)$, θ some parameter of interest.) A general method, called the “bootstrap” is introduced, and shown to work satisfactorily on a variety of estimation problems. The jackknife is shown to be a linear approximation method for the bootstrap. The exposition proceeds by a series of examples: variance of the sample median, error rates in a linear discriminant analysis, ratio estimation, estimating regression parameters, etc.

Efron: Computers and the theory of statistics: Thinking the unthinkable**Efron-1979b**

Bradley Efron. “Computers and the theory of statistics: Thinking the unthinkable”. In: *SIAM Review* 21.4 (Oct. 1979), pp. 460–480. DOI: [10.1137/1021092](https://doi.org/10.1137/1021092).

Abstract: This is a survey article concerning recent advances in certain areas of statistical theory, written for a mathematical audience with no background in statistics. The topics are chosen to illustrate a special point: how the advent of the high-speed computer has affected the development of statistical theory. The topics discussed include nonparametric methods, the jackknife, the bootstrap, cross-validation, error-rate estimation in discriminant analysis, robust estimation, the influence function, censored data, the EM algorithm, and Cox’s likelihood function. The exposition is mainly by example, with only a little offered in the way of theoretical development.

Bradley Efron. “Nonparametric standard errors and confidence intervals”. In: *Canadian Journal of Statistics / La Revue Canadienne de Statistique* 9.2 (1981), pp. 139–158. DOI: [10.2307/3314608](https://doi.org/10.2307/3314608).

Abstract: We investigate several nonparametric methods; the bootstrap, the jackknife, the delta method, and other related techniques. The first and simplest goal is the assignment of nonparametric standard errors to a real-valued statistic. More ambitiously, we consider setting nonparametric confidence intervals for a real-valued parameter. Building on the well understood case of confidence intervals for the median, some hopeful evidence is presented that such a theory may be possible.

Efron: Nonparametric standard errors and confidence intervals: Rejoinder**Efron-1981b**

Bradley Efron. “Nonparametric standard errors and confidence intervals: Rejoinder”. In: *The Canadian Journal of Statistics / La Revue Canadienne de Statistique* 9.2 (1981), pp. 170–172. DOI: [10.2307/3314610](https://doi.org/10.2307/3314610).

Efron: Better bootstrap confidence intervals**Efron-1987**

Bradley Efron. “Better bootstrap confidence intervals”. In: *Journal of the American Statistical Association* 82.397 (Mar. 1987), pp. 171–185. DOI: [10.1080/01621459.1987.10478410](https://doi.org/10.1080/01621459.1987.10478410).

Abstract: We consider the problem of setting approximate confidence intervals for a single parameter θ in a multiparameter family. The standard approximate intervals based on maximum likelihood theory, $\hat{\theta} \pm \hat{\sigma}z^{(\alpha)}$, can be quite misleading. In practice, tricks based on transformations, bias corrections, and so forth, are often used to improve their accuracy. The bootstrap confidence intervals discussed in this article automatically incorporate such tricks without requiring the statistician to think them through for each new application, at the price of a considerable increase in computational effort. The new intervals incorporate an improvement over previously suggested methods,

which results in second-order correctness in a wide variety of problems. In addition to parametric families, bootstrap intervals are also developed for nonparametric situations.

Efron: Bootstrap confidence intervals: Good or bad?**Efron-1988**

Bradley Efron. “Bootstrap confidence intervals: Good or bad?” In: *Psychological Bulletin* 104.2 (1988), pp. 293–296. DOI: [10.1037/0033-2909.104.2.293](https://doi.org/10.1037/0033-2909.104.2.293).

Abstract: The bootstrap is a nonparametric technique for estimating standard errors and approximate confidence intervals. Rasmussen has used a simulation experiment to suggest that bootstrap confidence intervals perform very poorly in the estimation of a correlation coefficient. Part of Rasmussen’s simulation is repeated. A careful look at the results shows the bootstrap intervals performing quite well. Some remarks are made concerning the virtues and defects of bootstrap intervals in general.

Efron: Second thoughts on the bootstrap**Efron-2003**

Bradley Efron. “Second thoughts on the bootstrap”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994968](https://doi.org/10.1214/ss/1063994968).

Abstract: This brief review article is appearing in the issue of Statistical Science that marks the 25th anniversary of the bootstrap. It concerns some of the theoretical and methodological aspects of the bootstrap and how they might influence future work in statistics.

Efron: Bayesian inference and the parametric bootstrap**Efron-2012**

Bradley Efron. “Bayesian inference and the parametric bootstrap”. In: *The Annals of Applied Statistics* 6.4 (Dec. 2012). DOI: [10.1214/12-aos571](https://doi.org/10.1214/12-aos571).

Abstract: The parametric bootstrap can be used for the efficient computation of Bayes posterior distributions. Importance sampling formulas take on an easy form relating to the deviance in ex-

ponential families and are particularly simple starting from Jeffreys invariant prior. Because of the i.i.d. nature of bootstrap sampling, familiar formulas describe the computational accuracy of the Bayes estimates. Besides computational methods, the theory provides a connection between Bayesian and frequentist analysis. Efficient algorithms for the frequentist accuracy of Bayesian inferences are developed and demonstrated in a model selection example.

Efron et al.: An introduction to the bootstrap**Efron-Tibshirani-1993**

Bradley Efron and Robert J. Tibshirani. *An introduction to the bootstrap*. Monographs on statistics and applied probability ; 57. New York: Chapman & Hall, 1993. ISBN: 9780412042317. DOI: [10.1201/9780429246593](https://doi.org/10.1201/9780429246593).

Abstract: Statistics is a subject of many uses and surprisingly few effective practitioners. The traditional road to statistical knowledge is blocked, for most, by a formidable wall of mathematics. The approach in *An Introduction to the Bootstrap* avoids that wall. It arms scientists and engineers, as well as statisticians, with the computational techniques they need to analyze and understand complicated data sets.

Library: QA276.8 .E3745 1993.

Enders: Applied missing data analysis**Enders-2010**

Craig K. Enders. *Applied missing data analysis*. Guilford Publications, May 31, 2010. 377 pp. ISBN: 9781606236390.

Abstract: Walking readers step by step through complex concepts, this book translates missing data techniques into something that applied researchers and graduate students can understand and utilize in their own research. Enders explains the rationale and procedural details for maximum likelihood estimation, Bayesian estimation, multiple imputation, and models for handling missing not at random (MNAR) data. Easy-to-follow examples and small simulated data sets illustrate the tech-

niques and clarify the underlying principles. The companion website (www.appliedmissingdata.com) includes data files and syntax for the examples in the book as well as up-to-date information on software. The book is accessible to substantive researchers while providing a level of detail that will satisfy quantitative specialists.

Library: HA29 .E497 2010.

Epskamp et al.: Estimating psychological networks and their accuracy: A tutorial paper
Epskamp-Borsboom-Fried-2017

Sacha Epskamp, Denny Borsboom, and Eiko I. Fried. “Estimating psychological networks and their accuracy: A tutorial paper”. In: *Behavior Research Methods* 50.1 (Mar. 2017), pp. 195–212. ISSN: 1554-3528. DOI: [10.3758/s13428-017-0862-1](https://doi.org/10.3758/s13428-017-0862-1).

Abstract: The usage of psychological networks that conceptualize behavior as a complex interplay of psychological and other components has gained increasing popularity in various research fields. While prior publications have tackled the topics of estimating and interpreting such networks, little work has been conducted to check how accurate (i.e., prone to sampling variation) networks are estimated, and how stable (i.e., interpretation remains similar with less observations) inferences from the network structure (such as centrality indices) are. In this tutorial paper, we aim to introduce the reader to this field and tackle the problem of accuracy under sampling variation. We first introduce the current state-of-the-art of network estimation. Second, we provide a rationale why researchers should investigate the accuracy of psychological networks. Third, we describe how bootstrap routines can be used to (A) assess the accuracy of estimated network connections, (B) investigate the stability of centrality indices, and (C) test whether network connections and centrality estimates for different variables differ from each other. We introduce two novel statistical methods: for (B) the correlation stability coefficient, and for (C) the bootstrapped difference test for edge-weights and centrality indices. We conducted and present simulation studies to assess the performance of both methods. Finally, we developed the free R-package *bootnet* that allows for estimating psychological networks

in a generalized framework in addition to the proposed bootstrap methods. We showcase bootnet in a tutorial, accompanied by R syntax, in which we analyze a dataset of 359 women with posttraumatic stress disorder available online.

Epskamp et al.: The Gaussian graphical model in cross-sectional and time-series data

Epskamp-Lourens-Mottus-et-al-2018

Sacha Epskamp, Lourens J. Waldorp, et al. “The Gaussian graphical model in cross-sectional and time-series data”. In: *Multivariate Behavioral Research* 53.4 (Apr. 2018), pp. 453–480. DOI: [10.1080/00273171.2018.1454823](https://doi.org/10.1080/00273171.2018.1454823).

Abstract: We discuss the Gaussian graphical model (GGM; an undirected network of partial correlation coefficients) and detail its utility as an exploratory data analysis tool. The GGM shows which variables predict one-another, allows for sparse modeling of covariance structures, and may highlight potential causal relationships between observed variables. We describe the utility in three kinds of psychological data sets: data sets in which consecutive cases are assumed independent (e.g., cross-sectional data), temporally ordered data sets (e.g., $n = 1$ time series), and a mixture of the 2 (e.g., $n \geq 1$ time series). In time-series analysis, the GGM can be used to model the residual structure of a vector-autoregression analysis (VAR), also termed graphical VAR. Two network models can then be obtained: a temporal network and a contemporaneous network. When analyzing data from multiple subjects, a GGM can also be formed on the covariance structure of stationary means-the between-subjects network. We discuss the interpretation of these models and propose estimation methods to obtain these networks, which we implement in the R packages graphicalVAR and mlVAR. The methods are showcased in two empirical examples, and simulation studies on these methods are included in the supplementary materials.

Ernst et al.: Utilizing a quantile function approach to obtain exact bootstrap solutions

Ernst-Hutson-2003

Michael D. Ernst and Alan D. Hutson. “Utilizing a quantile function approach to obtain exact bootstrap solutions”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994978](https://doi.org/10.1214/ss/1063994978).

Abstract: The popularity of the bootstrap is due in part to its wide applicability and the ease of implementing resampling procedures on modern computers. But careful reading of Efron (1979) will show that at its heart, the bootstrap is a “plug-in” procedure that involves calculating a functional $\theta(\hat{F})$ from an estimate of the c.d.f. F . Resampling becomes invaluable when, as is often the case, $\theta(\hat{F})$ cannot be calculated explicitly. We discuss some situations where working with the sample quantile function, \hat{Q} , rather than \hat{F} , can lead to explicit (exact) solutions to $\theta(\hat{F})$.

Fairchild et al.: Using mediation and moderation analyses to enhance prevention research

Fairchild-MacKinnon-2014

Amanda J. Fairchild and David P. MacKinnon. “Using mediation and moderation analyses to enhance prevention research”. In: *Defining Prevention Science*. Springer US, 2014, pp. 537–555. DOI: [10.1007/978-1-4899-7424-2_23](https://doi.org/10.1007/978-1-4899-7424-2_23).

Abstract: Integrating mediating and moderating variables into prevention research can refine interventions and guide program evaluation by demonstrating how and for whom programs work, as well as lending insight into the construct validity of an intervention. In this way, program development and evaluation strategies that incorporate mediation and moderation analyses contribute to our ability to affect behavioral change. This chapter aims to illustrate how mediation and moderation analyses enhance and inform prevention and intervention work. To that end we define and differentiate the models, discuss their application to prevention programming and research, and provide information on their estimation for individuals seeking to implement these analyses.

Emilio Ferrer and John McArdle. “Alternative structural models for multivariate longitudinal data analysis”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 10.4 (Oct. 2003), pp. 493–524. DOI: [10.1207/s15328007sem1004_1](https://doi.org/10.1207/s15328007sem1004_1).

Abstract: Structural equation models are presented as alternative models for examining longitudinal data. The models include (a) a cross-lagged regression model, (b) a factor model based on latent growth curves, and (c) a dynamic model based on latent difference scores. The illustrative data are on motivation and perceived competence of students during their first semester in high school. The 3 models yielded different results and such differences were discussed in terms of the conceptualization of change underlying each model. The last model was defended as the most reasonable for these data because it captured the dynamic interrelations between the examined constructs and, at the same time, identified potential growth in the variables.

Fridberg et al.: Examining features of transdermal alcohol biosensor readings: A promising approach with implications for research and intervention

Fridberg-Wang-Porges-2022

Daniel J. Fridberg, Yan Wang, and Eric Porges. “Examining features of transdermal alcohol biosensor readings: A promising approach with implications for research and intervention”. In: *Alcoholism: Clinical and Experimental Research* 46.4 (Feb. 2022), pp. 514–516. ISSN: 1530-0277. DOI: [10.1111/acer.14794](https://doi.org/10.1111/acer.14794).

Fritz et al.: Required sample size to detect the mediated effect **Fritz-MacKinnon-2007**

Matthew S. Fritz and David P. MacKinnon. “Required sample size to detect the mediated effect”. In: *Psychological Science* 18.3 (Mar. 2007), pp. 233–239. DOI: [10.1111/j.1467-9280.2007.01882.x](https://doi.org/10.1111/j.1467-9280.2007.01882.x).

Abstract: Mediation models are widely used, and there are many tests of the mediated effect. One of the most common questions that researchers have when planning mediation studies is, “How many subjects do I need to achieve adequate power when testing for mediation?” This article presents the necessary sample sizes for six of the most common and the most recommended tests of mediation for various combinations of parameters, to provide a guide for researchers when designing studies or applying for grants.

Fritz et al.: Explanation of two anomalous results in statistical mediation analysis

Fritz-Taylor-MacKinnon-2012

Matthew S. Fritz, Aaron B. Taylor, and David P. MacKinnon. “Explanation of two anomalous results in statistical mediation analysis”. In: *Multivariate Behavioral Research* 47.1 (Feb. 2012), pp. 61–87. DOI: [10.1080/00273171.2012.640596](https://doi.org/10.1080/00273171.2012.640596).

Abstract: Previous studies of different methods of testing mediation models have consistently found two anomalous results. The first result is elevated Type I error rates for the bias-corrected and accelerated bias-corrected bootstrap tests not found in nonresampling tests or in resampling tests that did not include a bias correction. This is of special concern as the bias-corrected bootstrap is often recommended and used due to its higher statistical power compared with other tests. The second result is statistical power reaching an asymptote far below 1.0 and in some conditions even declining slightly as the size of the relationship between X and M, a , increased. Two computer simulations were conducted to examine these findings in greater detail. Results from the first simulation found that the increased Type I error rates for the bias-corrected and accelerated bias-corrected bootstrap are a function of an interaction between the size of the individual paths making up the mediated effect and the sample size, such that elevated Type I error rates occur when the sample size is small and the effect size of the nonzero path is medium or larger. Results from the second simulation found that stagnation and decreases in statistical power as a function of the effect size of the a path occurred primarily when the path between M and Y, b , was small. Two empirical mediation examples are provided using data from a steroid prevention and health promotion program aimed

at high school football players (Athletes Training and Learning to Avoid Steroids; Goldberg et al., 1996), one to illustrate a possible Type I error for the bias-corrected bootstrap test and a second to illustrate a loss in power related to the size of α . Implications of these findings are discussed.

Gates et al.: Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM
Gates-Molenaar-Hillary-et-al-2010

Kathleen M. Gates et al. "Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using formal equivalences among SEM path modeling, VAR, and unified SEM". In: *NeuroImage* 50.3 (Apr. 2010), pp. 1118–1125. DOI: [10.1016/j.neuroimage.2009.12.117](https://doi.org/10.1016/j.neuroimage.2009.12.117).

Abstract: Modeling the relationships among brain regions of interest (ROIs) carries unique potential to explicate how the brain orchestrates information processing. However, hurdles arise when using functional MRI data. Variation in ROI activity contains sequential dependencies and shared influences on synchronized activation. Consequently, both lagged and contemporaneous relationships must be considered for unbiased statistical parameter estimation. Identifying these relationships using a data-driven approach could guide theory-building regarding integrated processing. The present paper demonstrates how the unified SEM attends to both lagged and contemporaneous influences on ROI activity. Additionally, this paper offers an approach akin to Granger causality testing, Lagrange multiplier testing, for statistically identifying directional influence among ROIs and employs this approach using an automatic search procedure to arrive at the optimal model. Rationale for this equivalence is offered by explicating the formal relationships among path modeling, vector autoregression, and unified SEM. When applied to simulated data, biases in estimates which do not consider both lagged and contemporaneous paths become apparent. Finally, the use of unified SEM with the automatic search procedure is applied to an empirical data example.

A. R. Georgeson, Diana Alvarez-Bartolo, and David P. MacKinnon. “A sensitivity analysis for temporal bias in cross-sectional mediation”. In: *Psychological Methods* (Dec. 2023). DOI: [10.1037/met0000628](https://doi.org/10.1037/met0000628).

Abstract: For over three decades, methodologists have cautioned against the use of cross-sectional mediation analyses because they yield biased parameter estimates. Yet, cross-sectional mediation models persist in practice and sometimes represent the only analytic option. We propose a sensitivity analysis procedure to encourage a more principled use of cross-sectional mediation analysis, drawing inspiration from Gollob and Reichardt (1987, 1991). The procedure is based on the two-wave longitudinal mediation model and uses phantom variables for the baseline data. After a researcher provides ranges of possible values for cross-lagged, autoregressive, and baseline Y and M correlations among the phantom and observed variables, they can use the sensitivity analysis to identify longitudinal conditions in which conclusions from a cross-sectional model would differ most from a longitudinal model. To support the procedure, we first show that differences in sign and effect size of the b-path occur most often when the cross-sectional effect size of the b-path is small and the cross-lagged and the autoregressive correlations are equal or similar in magnitude. We then apply the procedure to cross-sectional analyses from real studies and compare the sensitivity analysis results to actual results from a longitudinal mediation analysis. While no statistical procedure can replace longitudinal data, these examples demonstrate that the sensitivity analysis can recover the effect that was actually observed in the longitudinal data if provided with the correct input information. Implications of the routine application of sensitivity analysis to temporal bias are discussed. R code for the procedure is provided in the online supplementary materials.

Harry F. Gollob and Charles S. Reichardt. “Taking account of time lags in causal models”. In: *Child Development* 58.1 (Feb. 1987), p. 80. ISSN: 0009-3920. DOI: [10.2307/1130293](https://doi.org/10.2307/1130293).

Abstract: Although it takes time for a cause to exert an effect, causal models often fail to allow adequately for time lags. In particular, causal models that contain cross-sectional relations (i. e., relations between values of 2 variables at the same time) are unsatisfactory because (a) they omit the values of variables at prior times, (b) they omit effects that variables can have on themselves, and (c) they fail to specify the length of the causal interval that is being studied. These omissions can produce severe biases in estimates of the size of causal effects. Longitudinal models also can fail to take account of time lags properly, and this too can lead to severely biased estimates. The discussion illustrates the biases that can occur in both cross-sectional and longitudinal models, introduces the latent longitudinal approach to causal modeling, and shows how latent longitudinal models can be used to reduce bias by taking account of time lags even when data are available for only 1 point in time.

Leo A. Goodman. “On the exact variance of products”. In: *Journal of the American Statistical Association* 55.292 (Dec. 1960), pp. 708–713. DOI: [10.1080/01621459.1960.10483369](https://doi.org/10.1080/01621459.1960.10483369).

Abstract: A simple exact formula for the variance of the product of two random variables, say, x and y , is given as a function of the means and central product-moments of x and y . The usual approximate variance formula for xy is compared with this exact formula; e.g., we note, in the special case where x and y are independent, that the “variance” computed by the approximate formula is less than the exact variance, and that the accuracy of the approximation depends on the sum of the reciprocals of the squared coefficients of variation of x and y . The case where x and y need not be independent is also studied, and exact variance formulas are presented for several different

“product estimates.” (The usefulness of exact formulas becomes apparent when the variances of these estimates are compared.) When x and y are independent, simple unbiased estimates of these exact variances are suggested; in the more general case, consistent estimates are presented.

J. W. Graham et al.: How many imputations are really needed? Some practical clarifications of multiple imputation theory **Graham-Olchowski-Gilreath-2007**

John W. Graham, Allison E. Olchowski, and Tamika D. Gilreath. “How many imputations are really needed? Some practical clarifications of multiple imputation theory”. In: *Prevention Science* 8.3 (June 2007), pp. 206–213. DOI: [10.1007/s11121-007-0070-9](https://doi.org/10.1007/s11121-007-0070-9).

Abstract: Multiple imputation (MI) and full information maximum likelihood (FIML) are the two most common approaches to missing data analysis. In theory, MI and FIML are equivalent when identical models are tested using the same variables, and when m , the number of imputations performed with MI, approaches infinity. However, it is important to know how many imputations are necessary before MI and FIML are sufficiently equivalent in ways that are important to prevention scientists. MI theory suggests that small values of m , even on the order of three to five imputations, yield excellent results. Previous guidelines for sufficient m are based on relative efficiency, which involves the fraction of missing information (γ) for the parameter being estimated, and m . In the present study, we used a Monte Carlo simulation to test MI models across several scenarios in which γ and m were varied. Standard errors and p -values for the regression coefficient of interest varied as a function of m , but not at the same rate as relative efficiency. Most importantly, statistical power for small effect sizes diminished as m became smaller, and the rate of this power falloff was much greater than predicted by changes in relative efficiency. Based our findings, we recommend that researchers using MI should perform many more imputations than previously considered sufficient. These recommendations are based on γ , and take into consideration one’s tolerance for a preventable power falloff (compared to FIML) due to using too few imputations.

Fei Gu, Kristopher J. Preacher, and Emilio Ferrer. “A state space modeling approach to mediation analysis”. In: *Journal of Educational and Behavioral Statistics* 39.2 (Apr. 2014), pp. 117–143. ISSN: 1935-1054. DOI: [10.3102/1076998614524823](https://doi.org/10.3102/1076998614524823).

Abstract: Mediation is a causal process that evolves over time. Thus, a study of mediation requires data collected throughout the process. However, most applications of mediation analysis use cross-sectional rather than longitudinal data. Another implicit assumption commonly made in longitudinal designs for mediation analysis is that the same mediation process universally applies to all members of the population under investigation. This assumption ignores the important issue of ergodicity before aggregating the data across subjects. We first argue that there exists a discrepancy between the concept of mediation and the research designs that are typically used to investigate it. Second, based on the concept of ergodicity, we argue that a given mediation process probably is not equally valid for all individuals in a population. Therefore, the purpose of this article is to propose a two-faceted solution. The first facet of the solution is that we advocate a single-subject time-series design that aligns data collection with researchers’ conceptual understanding of mediation. The second facet is to introduce a flexible statistical method—the state space model—as an ideal technique to analyze single-subject time series data in mediation studies. We provide an overview of the state space method and illustrative applications using both simulated and real time series data. Finally, we discuss additional issues related to research design and modeling.

Rachel L. Gunn et al. “Characterising patterns of alcohol use among heavy drinkers: A cluster analysis utilising alcohol biosensor data”. In: *Drug and Alcohol Review* 40.7 (May 2021), pp. 1155–1164. ISSN: 1465-3362. DOI: [10.1111/dar.13306](https://doi.org/10.1111/dar.13306).

Abstract: Introduction: Previous research has predominately relied on person-level or single characteristics of drinking episodes to characterise patterns of drinking that may confer risk. This research often relies on self-report measures. Advancements in wearable alcohol biosensors provide a multifaceted objective measure of drinking. The current study aimed to characterise drinking episodes using data derived from a wearable alcohol biosensor. Methods: Participants ($n = 45$) were adult heavy drinkers who wore the Secure Continuous Remote Alcohol Monitoring (SCRAM) bracelet and reported on their drinking behaviours. Cluster analysis was used to evaluate unique combinations of alcohol episode characteristics. Associations between clusters and self-reported person and event-level factors were also examined in univariable and multivariable models. Results: Results suggested three unique clusters: Cluster 1 (most common, slowest rate of rise to and decline from peak), Cluster 2 (highest peak transdermal alcohol concentration and area under the curve) and Cluster 3 (fastest rate of decline from peak). Univariable analyses distinguished Cluster 1 as having fewer self-reported drinks and fewer episodes that occurred on weekends relative to Cluster 2. The effect for number of drinks remained in multivariable analyses. Discussion and Conclusions: This is the first study to characterise drinking patterns at the event-level using objective data. Results suggest that it is possible to distinguish drinking episodes based on several characteristics derived from wearable alcohol biosensors. This examination lays the groundwork for future studies to characterise patterns of drinking and their association with consequences of drinking behaviour.

Hall: A short prehistory of the bootstrap

Hall-2003

Peter Hall. “A short prehistory of the bootstrap”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994970](https://doi.org/10.1214/ss/1063994970).

Abstract: The contemporary development of bootstrap methods, from the time of Efron’s early articles to the present day, is well documented and widely appreciated. Likewise, the relationship of bootstrap techniques to certain early work on permutation testing, the jackknife and cross-validation is well understood. Less known, however, are the connections of the bootstrap to research on survey sampling for spatial data in the first half of the last century or to work from the 1940s to the 1970s

on subsampling and resampling. In a selective way, some of these early linkages will be explored, giving emphasis to developments with which the statistics community tends to be less familiar. Particular attention will be paid to the work of P. C. Mahalanobis, whose development in the 1930s and 1940s of moving-block sampling methods for spatial data has a range of interesting features, and to contributions of other scientists who, during the next 40 years, developed half-sampling, subsampling and resampling methods.

Hamaker et al.: A critique of the cross-lagged panel model

Hamaker-Kuiper-Grasman-2015

Ellen L. Hamaker, Rebecca M. Kuiper, and Raoul P. P. P. Grasman. “A critique of the cross-lagged panel model”. In: *Psychological Methods* 20.1 (2015), pp. 102–116. DOI: [10.1037/a0038889](https://doi.org/10.1037/a0038889).

Abstract: The cross-lagged panel model (CLPM) is believed by many to overcome the problems associated with the use of cross-lagged correlations as a way to study causal influences in longitudinal panel data. The current article, however, shows that if stability of constructs is to some extent of a trait-like, timeinvariant nature, the autoregressive relationships of the CLPM fail to adequately account for this. As a result, the lagged parameters that are obtained with the CLPM do not represent the actual within-person relationships over time, and this may lead to erroneous conclusions regarding the presence, predominance, and sign of causal influences. In this article we present an alternative model that separates the within-person process from stable between-person differences through the inclusion of random intercepts, and we discuss how this model is related to existing structural equation models that include cross-lagged relationships. We derive the analytical relationship between the cross-lagged parameters from the CLPM and the alternative model, and use simulations to demonstrate the spurious results that may arise when using the CLPM to analyze data that include stable, trait-like individual differences. We also present a modeling strategy to avoid this pitfall and illustrate this using an empirical data set. The implications for both existing and future cross-lagged panel research are discussed.

James D. Hamilton. *Time series analysis*. Princeton, NJ: Princeton University Press, 1994. 1799 pp. ISBN: 9780691218632.

Andrew C. Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, Feb. 1990. DOI: [10.1017/cbo9781107049994](https://doi.org/10.1017/cbo9781107049994).

Abstract: In this book, Andrew Harvey sets out to provide a unified and comprehensive theory of structural time series models. Unlike the traditional ARIMA models, structural time series models consist explicitly of unobserved components, such as trends and seasonals, which have a direct interpretation. As a result the model selection methodology associated with structural models is much closer to econometric methodology. The link with econometrics is made even closer by the natural way in which the models can be extended to include explanatory variables and to cope with multivariate time series. From the technical point of view, state space models and the Kalman filter play a key role in the statistical treatment of structural time series models. The book includes a detailed treatment of the Kalman filter. This technique was originally developed in control engineering, but is becoming increasingly important in fields such as economics and operations research. This book is concerned primarily with modelling economic and social time series, and with addressing the special problems which the treatment of such series poses. The properties of the models and the methodological techniques used to select them are illustrated with various applications. These range from the modelling of trends and cycles in US macroeconomic time series to an evaluation of the effects of seat belt legislation in the UK.

Hatemi-J: A new method to choose optimal lag order in stable and unstable VAR models **HatemiJ-2003**

Abdulnasser Hatemi-J. “A new method to choose optimal lag order in stable and unstable VAR models”. In: *Applied Economics Letters* 10.3 (Feb. 2003), pp. 135–137. DOI: [10.1080/1350485022000041050](https://doi.org/10.1080/1350485022000041050).

Abstract: A crucial aspect of empirical research based on the vector autoregressive (VAR) model is the choice of the lag order, since all inference in the VAR model is based on the chosen lag order. Here, a new information criterion is introduced for this purpose. The conducted Monte Carlo simulation experiments show that this new information criterion performs well in picking the true lag order in stable as well as unstable VAR models.

Hatemi-J: Multivariate tests for autocorrelation in the stable and unstable VAR models **HatemiJ-2004**

Abdulnasser Hatemi-J. “Multivariate tests for autocorrelation in the stable and unstable VAR models”. In: *Economic Modelling* 21.4 (July 2004), pp. 661–683. DOI: [10.1016/j.econmod.2003.09.005](https://doi.org/10.1016/j.econmod.2003.09.005).

Abstract: This study investigates the size and power properties of three multivariate tests for autocorrelation, namely portmanteau test, Lagrange multiplier (LM) test and Rao F-test, in the stable and unstable vector autoregressive (VAR) models, with and without autoregressive conditional heteroscedasticity (ARCH) using Monte Carlo experiments. Many combinations of parameters are used in the simulations to cover a wide range of situations in order to make the results more representative. The results of conducted simulations show that all three tests perform relatively well in stable VAR models without ARCH. In unstable VAR models the portmanteau test exhibits serious size distortions. LM and Rao tests perform well in unstable VAR models without ARCH. These results are true, irrespective of sample size or order of autocorrelation. Another clear result that the simulations show is that none of the tests have the correct size when ARCH is present irrespective of VAR models being stable or unstable and regardless of the sample size or order of autocorrelation.

The portmanteau test appears to have slightly better power properties than the LM test in almost all scenarios.

Hayes: Beyond Baron and Kenny: Statistical mediation analysis in the new millennium

Hayes-2009

Andrew F. Hayes. “Beyond Baron and Kenny: Statistical mediation analysis in the new millennium”. In: *Communication Monographs* 76.4 (Dec. 2009), pp. 408–420. DOI: [10.1080/03637750903310360](https://doi.org/10.1080/03637750903310360).

Abstract: Understanding communication processes is the goal of most communication researchers. Rarely are we satisfied merely ascertaining whether messages have an effect on some outcome of focus in a specific context. Instead, we seek to understand how such effects come to be. What kinds of causal sequences does exposure to a message initiate? What are the causal pathways through which a message exerts its effect? And what role does communication play in the transmission of the effects of other variables over time and space? Numerous communication models attempt to describe the mechanism through which messages or other communication-related variables transmit their effects or intervene between two other variables in a causal model. The communication literature is replete with tests of such models. Over the years, methods used to test such process models have grown in sophistication. An example includes the rise of structural equation modeling (SEM), which allows investigators to examine how well a process model that links some focal variable X to some outcome Y through one or more intervening pathways fits the observed data. Yet frequently, the analytical choices communication researchers make when testing intervening variables models are out of step with advances made in the statistical methods literature. My goal here is to update the field on some of these new advances. While at it, I challenge some conventional wisdom and nudge the field toward a more modern way of thinking about the analysis of intervening variable effects.

Hayes: Introduction to mediation, moderation, and conditional process analysis: A regression-based approach **Hayes-2022**

Andrew F. Hayes. *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. 3rd ed. Methodology in the social sciences. Guilford Publications, 2022, p. 732. ISBN: 9781462549030.

Abstract: Lauded for its easy-to-understand, conversational discussion of the fundamentals of mediation, moderation, and conditional process analysis, this book has been fully revised with 50% new content, including sections on working with multicategorical antecedent variables, the use of PROCESS version 3 for SPSS and SAS for model estimation, and annotated PROCESS v3 outputs. Using the principles of ordinary least squares regression, Andrew F. Hayes carefully explains procedures for testing hypotheses about the conditions under and the mechanisms by which causal effects operate, as well as the moderation of such mechanisms. Hayes shows how to estimate and interpret direct, indirect, and conditional effects; probe and visualize interactions; test questions about moderated mediation; and report different types of analyses. Data for all the examples are available on the companion website (www.afhayes.com) along with links to download PROCESS.

Library: HA31.3 .H39 2022.

Hayes et al.: Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation **Hayes-Cai-2007**

Andrew F. Hayes and Li Cai. “Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation”. In: *Behavior Research Methods* 39.4 (Nov. 2007), pp. 709–722. DOI: [10.3758/bf03192961](https://doi.org/10.3758/bf03192961).

Hayes et al.: The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis **Hayes-Scharkow-2013**

Andrew F. Hayes and Michael Scharkow. “The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis”. In: *Psychological Science* 24.10 (Aug. 2013), pp. 1918–1927. DOI: [10.1177/0956797613480187](https://doi.org/10.1177/0956797613480187).

Abstract: A content analysis of 2 years of Psychological Science articles reveals inconsistencies in how researchers make inferences about indirect effects when conducting a statistical mediation analysis. In this study, we examined the frequency with which popularly used tests disagree, whether the method an investigator uses makes a difference in the conclusion he or she will reach, and whether there is a most trustworthy test that can be recommended to balance practical and performance considerations. We found that tests agree much more frequently than they disagree, but disagreements are more common when an indirect effect exists than when it does not. We recommend the bias-corrected bootstrap confidence interval as the most trustworthy test if power is of utmost concern, although it can be slightly liberal in some circumstances. Investigators concerned about Type I errors should choose the Monte Carlo confidence interval or the distribution-of-the-product approach, which rarely disagree. The percentile bootstrap confidence interval is a good compromise test.

Hesterberg: What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum **Hesterberg-2014**

Tim C. Hesterberg. *What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum*. 2014. arXiv: [1411.5279](https://arxiv.org/abs/1411.5279) [stat.OT]. URL: <https://arxiv.org/abs/1411.5279>.

Abstract: I have three goals in this article:

1. To show the enormous potential of bootstrapping and permutation tests to help students un-

derstand statistical concepts including sampling distributions, standard errors, bias, confidence intervals, null distributions, and P-values.

2. To dig deeper, understand why these methods work and when they don't, things to watch out for, and how to deal with these issues when teaching.
3. To change statistical practice—by comparing these methods to common t tests and intervals, we see how inaccurate the latter are; we confirm this with asymptotics. $n \geq 30$ isn't enough—think $n \geq 5000$.

Resampling provides diagnostics, and more accurate alternatives. Sadly, the common bootstrap percentile interval badly under-covers in small samples; there are better alternatives. The tone is informal, with a few stories and jokes.

Hesterberg: What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum **Hesterberg-2015**

Tim C. Hesterberg. “What teachers should know about the bootstrap: Resampling in the undergraduate statistics curriculum”. In: *The American Statistician* 69.4 (Oct. 2015), pp. 371–386. DOI: [10.1080/00031305.2015.1089789](https://doi.org/10.1080/00031305.2015.1089789).

Abstract: Bootstrapping has enormous potential in statistics education and practice, but there are subtle issues and ways to go wrong. For example, the common combination of nonparametric bootstrapping and bootstrap percentile confidence intervals is less accurate than using t -intervals for small samples, though more accurate for larger samples. My goals in this article are to provide a deeper understanding of bootstrap methods—how they work, when they work or not, and which methods work better—and to highlight pedagogical issues. Supplementary materials for this article are available online.

Hingson et al.: Magnitude and trends in heavy episodic drinking, alcohol-impaired driving, and alcohol-related mortality and overdose hospitalizations among emerging adults of college ages 18–24 in the United States, 1998–2014

Hingson-Zha-Smyth-2017

Ralph Hingson, Wenxing Zha, and Daniel Smyth. “Magnitude and trends in heavy episodic drinking, alcohol-impaired driving, and alcohol-related mortality and overdose hospitalizations among emerging adults of college ages 18–24 in the United States, 1998–2014”. In: *Journal of Studies on Alcohol and Drugs* 78.4 (July 2017), pp. 540–548. ISSN: 1938-4114. DOI: [10.15288/jsad.2017.78.540](https://doi.org/10.15288/jsad.2017.78.540).

Abstract: Objective: This article estimates percentages of U.S. emerging adults ages 18-24 engaging in past-month heavy episodic drinking and past-year alcohol-impaired driving, and numbers experiencing alcohol-related unintentional injury deaths and overdose hospitalizations between 1998 and 2014. Method: We analyzed national injury mortality data from coroner, census, and college enrollment statistics, the National Survey on Drug Use and Health, and the Nationwide Inpatient Sample. Results: From 1999 to 2005, percentages of emerging adults ages 18-24 reporting past-month heavy episodic drinking rose from 37.1% to 43.1% and then declined to 38.8% in 2014. Alcohol-impaired driving rose from 24% to 25.5% and then declined to 16.0%. Alcohol-related unintentional injury deaths increased from 4,807 in 1998 to 5,531 in 2005 and then declined to 4,105 in 2014, a reduction of 29% per 100,000 since 1998. Alcohol-related traffic deaths increased from 3,783 in 1998 to 4,114 in 2005 and then declined to 2,614 in 2014, down 43% per 100,000 since 1998. Alcohol-related overdose deaths increased from 207 in 1998 to 891 in 2014, a 254% increase per 100,000. Other types of nontraffic unintentional injury deaths declined. Alcohol-overdose hospitalizations rose 26% per 100,000 from 1998 to 2014, especially from increases in alcohol/other drug overdoses, up 61% (alcohol/opioid overdoses up 197%). Conclusions: Among emerging adults, a trend toward increased alcohol-related unintentional injury deaths, heavy episodic drinking, and alcohol-impaired driving between 1998 and 2005 was reversed by 2014. Persistent high levels of heavy episodic drinking and related problems among emerging adults underscore a need to expand individually oriented in-

terventions, college/community collaborative programs, and evidence-supported policies to reduce their drinking and related problems.

David V. Hinkley: Jackknifing in unbalanced situations**Hinkley-1977**

David V. Hinkley. “Jackknifing in unbalanced situations”. In: *Technometrics* 19.3 (Aug. 1977), pp. 285–292. DOI: [10.1080/00401706.1977.10489550](https://doi.org/10.1080/00401706.1977.10489550).

Abstract: Both the standard jackknife and a weighted jackknife are investigated in the general linear model situation. Properties of bias reduction and standard error estimation are derived and the weighted jackknife shown to be superior for unbalanced data. There is a preliminary discussion of robust regression fitting using jackknife pseudo-values.

Holmes: Bootstrapping phylogenetic trees: Theory and methods**Holmes-2003a**

Susan Holmes. “Bootstrapping phylogenetic trees: Theory and methods”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994979](https://doi.org/10.1214/ss/1063994979).

Abstract: This is a survey of the use of the bootstrap in the area of systematic and evolutionary biology. I present the current usage by biologists of the bootstrap as a tool both for making inferences and for evaluating robustness, and propose a framework for thinking about these problems in terms of mathematical statistics.

Holmes: Bradley Efron: A conversation with good friends**Holmes-2003b**

Susan Holmes. “Bradley Efron: A conversation with good friends”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994981](https://doi.org/10.1214/ss/1063994981).

Abstract: Bradley Efron is Professor of Statistics and Biostatistics at Stanford University. He works on a combination of theoretical and applied topics, including empirical Bayes, survival analysis, exponential families, bootstrap and jackknife methods and confidence intervals. Most of his ap-

plied work has originated in biomedical consulting projects at the Stanford Medical School, mixed in with a few papers concerning astronomy and physics. Even his theoretical papers usually begin with specific applied problems. All three of the interviewers here have been close scientific collaborators. Brad was born in St. Paul, Minnesota, May 1938, to Esther and Miles Efron, Jewish-Russian immigrants. A Merit Scholarship, in the program's inaugural year, brought him to Caltech, graduating in Mathematics in 1960. He arrived at Stanford that Fall, eventually gaining his Ph.D., under the direction of Rupert Miller and Herb Solomon, in the Statistics Department, whose faculty also included Charles Stein, Herman Chernoff, Manny Parzen, Lincoln Moses and Ingram Olkin. Brad has lived at Stanford since 1960, with sabbaticals at Harvard, Imperial College and Berkeley. He has held several administrative positions in the university: Chair of Statistics, Associate Dean of Science, Chairman of the University Advisory Board and Chair of the Faculty Senate. He is currently Chair of the Undergraduate Program in Applied Mathematics. Honors include doctorates from Chicago, Madrid and Oslo, a MacArthur Prize Fellowship, membership in the National Academy of Sciences and the American Academy of Arts and Sciences, fellowship in the IMS and ASA, the Wilks Medal, Parzen Prize, the newly inaugurated Rao Prize and the outstanding statistician award from the Chicago ASA chapter. He has been the Rietz, Wald, and Fisher lecturers and holds the Max H. Stein endowed chair as Professor of Humanities and Sciences at Stanford. Professional service includes Theory and Methods Editor of JASA and President of the IMS. Currently he is President-Elect of the American Statistical Association, becoming President in 2004.

S. D. Horn et al.: Estimating heteroscedastic variances in linear models

Horn-Horn-Duncan-1975

Susan D. Horn, Roger A. Horn, and David B. Duncan. "Estimating heteroscedastic variances in linear models". In: *Journal of the American Statistical Association* 70.350 (June 1975), pp. 380–385. DOI: [10.1080/01621459.1975.10479877](https://doi.org/10.1080/01621459.1975.10479877).

Joel L. Horowitz. “The bootstrap in econometrics”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994976](https://doi.org/10.1214/ss/1063994976).

Abstract: This paper presents examples of problems in estimation and hypothesis testing that demonstrate the use and performance of the bootstrap in econometric settings. The examples are illustrated with two empirical applications. The paper concludes with a discussion of topics on which further research is needed.

Hunter: State space modeling in an open source, modular, structural equation modeling environment**Hunter-2017**

Michael D. Hunter. “State space modeling in an open source, modular, structural equation modeling environment”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.2 (Oct. 2017), pp. 307–324. DOI: [10.1080/10705511.2017.1369354](https://doi.org/10.1080/10705511.2017.1369354).

Abstract: State space models (SSMs) are introduced in the context of structural equation modeling (SEM). In particular, the OpenMx implementation of SSMs using the Kalman filter and prediction error decomposition is discussed. In reflection of modularity, the implementation uses the same full information maximum likelihood missing data procedures for SSMs and SEMs. Similarly, generic OpenMx features such as likelihood ratio tests, profile likelihood confidence intervals, Hessian-based standard errors, definition variables, and the matrix algebra interface are all supported. Example scripts for specification of autoregressive models, multiple lag models (VAR(p)), multiple lag moving average models (VARMA(p, q)), multiple subject models, and latent growth models are provided. Additionally, latent variable calculation based on the Kalman filter and raw data generation based on a model are all included. Finally, future work for extending SSMs to allow for random effects and for presenting them in diagrams is discussed.

Stefano M. Iacus. *Simulation and Inference for Stochastic Differential Equations*. Springer New York, 2008. DOI: [10.1007/978-0-387-75839-8](https://doi.org/10.1007/978-0-387-75839-8).

Lawrence R. James and Jeanne M. Brett. “Mediators, moderators, and tests for mediation”. In: *Journal of Applied Psychology* 69.2 (1984), pp. 307–321. DOI: [10.1037/0021-9010.69.2.307](https://doi.org/10.1037/0021-9010.69.2.307).

Abstract: Discusses mediation relations in causal terms. Influences of an antecedent are transmitted to a consequence through an intervening mediator. Mediation relations may assume a number of functional forms, including nonadditive, nonlinear, and nonrecursive forms. Although mediation and moderation are distinguishable processes, with nonadditive forms (moderated mediation) a particular variable may be both a mediator and a moderator within a single set of functional relations. Current models for testing mediation relations in industrial and organizational psychology often involve an interplay between exploratory (correlational) statistical tests and causal inference. It is suggested that no middle ground exists between exploratory and confirmatory (causal) analysis and that attempts to explain how mediation processes occur require specified causal models.

Jeff A. Jones and Niels G. Waller. “Computing confidence intervals for standardized regression coefficients.” In: *Psychological Methods* 18.4 (2013), pp. 435–453. DOI: [10.1037/a0033269](https://doi.org/10.1037/a0033269).

Abstract: With fixed predictors, the standard method (Cohen, Cohen, West, & Aiken, 2003, p. 86; Harris, 2001, p. 80; Hays, 1994, p. 709) for computing confidence intervals (CIs) for standardized regression coefficients fails to account for the sampling variability of the criterion standard deviation. With random predictors, this method also fails to account for the sampling variability of the pre-

dictor standard deviations. Nevertheless, under some conditions the standard method will produce CIs with accurate coverage rates. To delineate these conditions, we used a Monte Carlo simulation to compute empirical CI coverage rates in samples drawn from 36 populations with a wide range of data characteristics. We also computed the empirical CI coverage rates for 4 alternative methods that have been discussed in the literature: noncentrality interval estimation, the delta method, the percentile bootstrap, and the bias-corrected and accelerated bootstrap. Our results showed that for many data-parameter configurations—for example, sample size, predictor correlations, coefficient of determination (R^2), orientation of β with respect to the eigenvectors of the predictor correlation matrix, R_X —the standard method produced coverage rates that were close to their expected values. However, when population R^2 was large and when β approached the last eigenvector of R_X , then the standard method coverage rates were frequently below the nominal rate (sometimes by a considerable amount). In these conditions, the delta method and the 2 bootstrap procedures were consistently accurate. Results using noncentrality interval estimation were inconsistent. In light of these findings, we recommend that researchers use the delta method to evaluate the sampling variability of standardized regression coefficients.

Jones et al.: The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior **Jones-Waller-2013b**

Jeff A. Jones and Niels G. Waller. *The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior*. Tech. rep. University of Minnesota-Twin Cities, May 25, 2013. URL: <http://users.cla.umn.edu/~nwaller/downloads/techreports/TR052913.pdf> (visited on 07/22/2022).

Abstract: Yuan and Chan (2011) recently showed how to compute the covariance matrix of standardized regression coefficients from covariances. In this paper, we describe a new method for computing this covariance matrix from correlations. We then show that Yuan and Chan’s original equations can also be used when only correlational data are available. Next, we describe an

asymptotic distribution-free (ADF; Browne, 1984) method for computing the covariance matrix of standardized regression coefficients. We show that the ADF method works well with non-normal data in moderate-to-large samples using both simulated and real-data examples. Finally, we provide R code (R Development Core Team, 2012) in an Appendix to make these methods accessible to applied researchers.

Jones et al.: The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior **Jones-Waller-2015**

Jeff A. Jones and Niels G. Waller. “The normal-theory and asymptotic distribution-free (ADF) covariance matrix of standardized regression coefficients: Theoretical extensions and finite sample behavior”. In: *Psychometrika* 80.2 (June 2015), pp. 365–378. DOI: [10.1007/s11336-013-9380-y](https://doi.org/10.1007/s11336-013-9380-y).

Abstract: Yuan and Chan (Psychometrika, 76, 670-690, 2011) recently showed how to compute the covariance matrix of standardized regression coefficients from covariances. In this paper, we describe a method for computing this covariance matrix from correlations. Next, we describe an asymptotic distribution-free (ADF; Browne in British Journal of Mathematical and Statistical Psychology, 37, 62-83, 1984) method for computing the covariance matrix of standardized regression coefficients. We show that the ADF method works well with nonnormal data in moderate-to-large samples using both simulated and real-data examples. R code (R Development Core Team, 2012) is available from the authors or through the Psychometrika online repository for supplementary materials.

Jorgensen et al.: semTools: Useful tools for structural equation modeling **Jorgensen-Pornprasertmanit-Schoemann-et-al-2022**

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

Charles M. Judd and David A. Kenny. “Process analysis”. In: *Evaluation Review* 5.5 (Oct. 1981), pp. 602–619. DOI: [10.1177/0193841x8100500502](https://doi.org/10.1177/0193841x8100500502).

Abstract: This article presents the rationale and procedures for conducting a process analysis in evaluation research. Such an analysis attempts to identify the process that mediates the effects of some treatment, by estimating the parameters of a causal chain between the treatment and some outcome variable. Two different procedures for estimating mediation are discussed. In addition we present procedures for examining whether a treatment exerts its effects, in part, by altering the mediating process that produces the outcome. Finally, the benefits of process analysis in evaluation research are underlined.

Kalman: A new approach to linear filtering and prediction problems **Kalman-1960**

R. E. Kalman. “A new approach to linear filtering and prediction problems”. In: *Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45. DOI: [10.1115/1.3662552](https://doi.org/10.1115/1.3662552).

Abstract: The classical filtering and prediction problem is re-examined using the Bode-Shannon representation of random processes and the “state-transition” method of analysis of dynamic systems. New results are: (1) The formulation and methods of solution of the problem apply without modification to stationary and nonstationary statistics and to growing-memory and infinite-memory filters. (2) A nonlinear difference (or differential) equation is derived for the covariance matrix of the optimal estimation error. From the solution of this equation the co-efficients of the difference (or differential) equation of the optimal linear filter are obtained without further calculations. (3) The filtering problem is shown to be the dual of the noise-free regulator problem. The new method developed here is applied to two well-known problems, confirming and extending earlier results. The discussion is largely self-contained and proceeds from first principles; basic concepts of the theory of random processes are reviewed in the Appendix.

Kauermann et al.: A note on the efficiency of sandwich covariance matrix estimation

Kauermann-Carroll-2001

Göran Kauermann and Raymond J. Carroll. “A note on the efficiency of sandwich covariance matrix estimation”. In: *Journal of the American Statistical Association* 96.456 (Dec. 2001), pp. 1387–1396. DOI: [10.1198/016214501753382309](https://doi.org/10.1198/016214501753382309).

Abstract: The sandwich estimator, also known as robust covariance matrix estimator, heteroscedasticity-consistent covariance matrix estimate, or empirical covariance matrix estimator, has achieved increasing use in the econometric literature as well as with the growing popularity of generalized estimating equations. Its virtue is that it provides consistent estimates of the covariance matrix for parameter estimates even when the fitted parametric model fails to hold or is not even specified. Surprisingly though, there has been little discussion of properties of the sandwich method other than consistency. We investigate the sandwich estimator in quasi-likelihood models asymptotically, and in the linear case analytically. We show that under certain circumstances when the quasi-likelihood model is correct, the sandwich estimate is often far more variable than the usual parametric variance estimate. The increased variance is a fixed feature of the method and the price that one pays to obtain consistency even when the parametric model fails or when there is heteroscedasticity. We show that the additional variability directly affects the coverage probability of confidence intervals constructed from sandwich variance estimates. In fact, the use of sandwich variance estimates combined with t -distribution quantiles gives confidence intervals with coverage probability falling below the nominal value. We propose an adjustment to compensate for this fact.

Kim et al.: State-space models with regime switching: Classical and Gibbs-sampling approaches with applications

Kim-Nelson-1999

Chang-Jin Kim and Charles R. Nelson. *State-space models with regime switching: Classical and Gibbs-sampling approaches with applications*. The MIT Press, 1999. ISBN: 9780262277112. DOI: [10.7551/mitpress/6444.001.0001](https://doi.org/10.7551/mitpress/6444.001.0001).

Abstract: Both state-space models and Markov switching models have been highly productive paths for empirical research in macroeconomics and finance. This book presents recent advances in econometric methods that make feasible the estimation of models that have both features. One approach, in the classical framework, approximates the likelihood function; the other, in the Bayesian framework, uses Gibbs-sampling to simulate posterior distributions from data. The authors present numerous applications of these approaches in detail: decomposition of time series into trend and cycle, a new index of coincident economic indicators, approaches to modeling monetary policy uncertainty, Friedman's "plucking" model of recessions, the detection of turning points in the business cycle and the question of whether booms and recessions are duration-dependent, state-space models with heteroskedastic disturbances, fads and crashes in financial markets, long-run real exchange rates, and mean reversion in asset returns.

Library: HB135 .K515 1999.

Kisbu-Sakarya et al.: The distribution of the product explains normal theory mediation confidence interval estimation **KisbuSakarya-MacKinnon-Miocevic-2014**

Yasemin Kisbu-Sakarya, David P. MacKinnon, and Milica Miočević. "The distribution of the product explains normal theory mediation confidence interval estimation". In: *Multivariate Behavioral Research* 49.3 (May 2014), pp. 261–268. DOI: [10.1080/00273171.2014.903162](https://doi.org/10.1080/00273171.2014.903162).

Abstract: The distribution of the product has several useful applications. One of these applications is its use to form confidence intervals for the indirect effect as the product of 2 regression coefficients. The purpose of this article is to investigate how the moments of the distribution of the product explain normal theory mediation confidence interval coverage and imbalance. Values of the critical ratio for each random variable are used to demonstrate how the moments of the distribution of the product change across values of the critical ratio observed in research studies. Results of the simulation study showed that as skewness in absolute value increases, coverage decreases. And as skewness in absolute value and kurtosis increases, imbalance increases. The difference between testing the

significance of the indirect effect using the normal theory versus the asymmetric distribution of the product is further illustrated with a real data example. This article is the first study to show the direct link between the distribution of the product and indirect effect confidence intervals and clarifies the results of previous simulation studies by showing why normal theory confidence intervals for indirect effects are often less accurate than those obtained from the asymmetric distribution of the product or from resampling methods.

Koopman et al.: Pulling the Sobel test up by its bootstraps

Koopman-Howe-Hollenbeck-2014

Joel Koopman, Michael Howe, and John R. Hollenbeck. "Pulling the Sobel test up by its bootstraps". In: Charles E. Lance and Robert J. Vandenberg. *More statistical and methodological myths and urban legends: Doctrine, verity and fable in organizational and social sciences*. Routledge/Taylor & Francis Group, 2014, pp. 224–243. ISBN: 9780203775851. DOI: [10.4324/9780203775851](https://doi.org/10.4324/9780203775851).

Abstract: In the domain of building and testing theory, mediation relationships are among the most important that can be proposed. Mediation helps to explicate our theoretical models (Leavitt, Mitchell, & Peterson, 2010) and addresses the fundamental question of why two constructs are related (Whetten, 1989). One of the better-known methods for testing mediation is commonly referred to as the "Sobel test," named for the researcher who derived a standard error (Sobel, 1982) to test the significance of the indirect effect. Recently, a number of different research teams (e.g., Preacher & Hayes, 2004; Shrout & Bolger, 2002) have criticized the Sobel test because this standard error requires an assumption of normality for the indirect effect sampling distribution. This distribution tends to be positively skewed (i.e., not normal), particularly in small samples, and so this assumption can be problematic (Preacher & Hayes, 2004; Stone & Sobel, 1990). As a result, the statistical power of the Sobel test may be lessened in these contexts (Preacher & Hayes 2004; Shrout & Bolger, 2002). In light of this concern, some scholars have advocated instead for the use of bootstrapping to test the significance of the indirect effect (e.g., Shrout & Bolger 2002). Bootstrapping requires no a priori assumption about the shape of the sampling distribution because this distribution is

empirically estimated using a resampling procedure (Efron & Tibshirani, 1993). As a result, departures from normality are less troublesome when creating a confidence interval for the indirect effect. For this reason, bootstrapping is now widely believed to be inherently superior to the Sobel test when testing the significance of the indirect effect in organizational research. Our position is that this belief constitutes an urban legend. As with all statistical urban legends, there is an underlying kernel of truth to the belief that bootstrapping is superior to the Sobel test. However, as we discuss in this chapter, there are several reasons to be concerned with a broad belief in the superiority of bootstrapping. We begin with a brief overview of mediation testing focusing on the Sobel test and bootstrapping and then explain the underlying kernel of truth that has propelled bootstrapping to the forefront of mediation testing in organizational research. Subsequently, we discuss four areas of concern that cast doubt on the belief of the inherent superiority of bootstrapping. Finally, we conclude with recommendations concerning the future of mediation testing in organizational research.

Koopman et al.: Small sample mediation testing: Misplaced confidence in bootstrapped confidence intervals **Koopman-Howe-Hollenbeck-et-al-2015**

Joel Koopman, Michael Howe, John R. Hollenbeck, and Hock-Peng Sin. “Small sample mediation testing: Misplaced confidence in bootstrapped confidence intervals”. In: *Journal of Applied Psychology* 100.1 (2015), pp. 194–202. DOI: [10.1037/a0036635](https://doi.org/10.1037/a0036635).

Abstract: Bootstrapping is an analytical tool commonly used in psychology to test the statistical significance of the indirect effect in mediation models. Bootstrapping proponents have particularly advocated for its use for samples of 20-80 cases. This advocacy has been heeded, especially in the *Journal of Applied Psychology*, as researchers are increasingly utilizing bootstrapping to test mediation with samples in this range. We discuss reasons to be concerned with this escalation, and in a simulation study focused specifically on this range of sample sizes, we demonstrate not only that bootstrapping has insufficient statistical power to provide a rigorous hypothesis test in most conditions but also that bootstrapping has a tendency to exhibit an inflated Type I error

rate. We then extend our simulations to investigate an alternative empirical resampling method as well as a Bayesian approach and demonstrate that they exhibit comparable statistical power to bootstrapping in small samples without the associated inflated Type I error. Implications for researchers testing mediation hypotheses in small samples are presented. For researchers wishing to use these methods in their own research, we have provided R syntax in the online supplemental materials.

Kuiper et al.: Drawing conclusions from cross-lagged relationships: Re-considering the role of the time-interval **Kuiper-Oisin-2018**

Rebecca M. Kuiper and Oisin Ryan. “Drawing conclusions from cross-lagged relationships: Re-considering the role of the time-interval”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 25.5 (Mar. 2018), pp. 809–823. DOI: [10.1080/10705511.2018.1431046](https://doi.org/10.1080/10705511.2018.1431046).

Abstract: The cross-lagged panel model (CLPM), a discrete-time (DT) SEM model, is frequently used to gather evidence for (reciprocal) Granger-causal relationships when lacking an experimental design. However, it is well known that CLPMs can lead to different parameter estimates depending on the time-interval of observation. Consequently, this can lead to researchers drawing conflicting conclusions regarding the sign and/or dominance of relationships. Multiple authors have suggested the use of continuous-time models to address this issue. In this article, we demonstrate the exact circumstances under which such conflicting conclusions occur. Specifically, we show that such conflicts are only avoided in general in the case of bivariate, stable, nonoscillating, first-order systems, when comparing models with uniform time-intervals between observations. In addition, we provide a range of tools, proofs, and guidelines regarding the comparison of discrete- and continuous-time parameter estimates.

Kurtzer et al.: hpcng/singularity: Singularity 3.7.3 Kurtzer-cclerget-Bauer-et-al-2021

Gregory M. Kurtzer, cclerget, et al. *hpcng/singularity: Singularity 3.7.3*. 2021. DOI: [10.5281/ZENODO.1310023](https://doi.org/10.5281/ZENODO.1310023).

Kurtzer et al.: Singularity: Scientific containers for mobility of compute**Kurtzer-Sochat-Bauer-2017**

Gregory M. Kurtzer, Vanessa Sochat, and Michael W. Bauer. “Singularity: Scientific containers for mobility of compute”. In: *PLOS ONE* 12.5 (May 2017). Ed. by Attila Gursoy, e0177459. DOI: [10.1371/journal.pone.0177459](https://doi.org/10.1371/journal.pone.0177459).

Kwan et al.: Comparing standardized coefficients in structural equation modeling: A model reparameterization approach

Kwan-Chan-2011

Joyce L. Y. Kwan and Wai Chan. “Comparing standardized coefficients in structural equation modeling: A model reparameterization approach”. In: *Behavior Research Methods* 43.3 (Apr. 2011), pp. 730–745. DOI: [10.3758/s13428-011-0088-6](https://doi.org/10.3758/s13428-011-0088-6).

Abstract: We propose a two-stage method for comparing standardized coefficients in structural equation modeling (SEM). At stage 1, we transform the original model of interest into the standardized model by model reparameterization, so that the model parameters appearing in the standardized model are equivalent to the standardized parameters of the original model. At stage 2, we impose appropriate linear equality constraints on the standardized model and use a likelihood ratio test to make statistical inferences about the equality of standardized coefficients. Unlike other existing methods for comparing standardized coefficients, the proposed method does not require specific modeling features (e.g., specification of nonlinear constraints), which are available only in certain SEM software programs. Moreover, this method allows researchers to compare two or more standardized coefficients simultaneously in a standard and convenient way. Three real examples are

given to illustrate the proposed method, using EQS, a popular SEM software program. Results show that the proposed method performs satisfactorily for testing the equality of standardized coefficients.

Kwan et al.: Comparing squared multiple correlation coefficients using structural equation modeling **Kwan-Chan-2014**

Joyce L. Y. Kwan and Wai Chan. “Comparing squared multiple correlation coefficients using structural equation modeling”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 21.2 (Apr. 2014), pp. 225–238. DOI: [10.1080/10705511.2014.882673](https://doi.org/10.1080/10705511.2014.882673).

Abstract: In social science research, a common topic in multiple regression analysis is to compare the squared multiple correlation coefficients in different populations. Existing methods based on asymptotic theories (Olkin & Finn, 1995) and bootstrapping (Chan, 2009) are available but these can only handle a 2-group comparison. Another method based on structural equation modeling (SEM) has been proposed recently. However, this method has three disadvantages. First, it requires the user to explicitly specify the sample R^2 as a function in terms of the basic SEM model parameters, which is sometimes troublesome and error prone. Second, it requires the specification of nonlinear constraints, which is not available in some popular SEM software programs. Third, it is for a 2-group comparison primarily. In this article, a 2-stage SEM method is proposed as an alternative. Unlike all other existing methods, the proposed method is simple to use, and it does not require any specific programming features such as the specification of nonlinear constraints. More important, the method allows a simultaneous comparison of 3 or more groups. A real example is given to illustrate the proposed method using EQS, a popular SEM software program.

Partha Lahiri. “On the impact of bootstrap in survey sampling and small-area estimation”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994975](https://doi.org/10.1214/ss/1063994975).

Abstract: Development of valid bootstrap procedures has been a challenging problem for survey samplers for the last two decades. This is due to the fact that in surveys we constantly face various complex issues such as complex correlation structure induced by the survey design, weighting, imputation, small-area estimation, among others. In this paper, we critically review various bootstrap methods developed to deal with these challenging issues. We discuss two applications where the bootstrap has been found to be effective.

Leffingwell et al.: Continuous Objective Monitoring of Alcohol Use: Twenty-First Century Measurement Using Transdermal Sensors Leffingwell-Cooney-Murphy-et-al-2012

Thad R. Leffingwell et al. “Continuous Objective Monitoring of Alcohol Use: Twenty-First Century Measurement Using Transdermal Sensors”. In: *Alcoholism: Clinical and Experimental Research* 37.1 (July 2012), pp. 16–22. ISSN: 1530-0277. DOI: [10.1111/j.1530-0277.2012.01869.x](https://doi.org/10.1111/j.1530-0277.2012.01869.x).

Abstract: Transdermal alcohol sensors continuously collect reliable and valid data on alcohol consumption in vivo over the course of hours to weeks. Transdermal alcohol readings are highly correlated with breath alcohol measurements, but transdermal alcohol levels lag behind breath alcohol levels by one or more hours owing to the longer time required for alcohol to be expelled through perspiration. By providing objective information about alcohol consumption, transdermal alcohol sensors can validate self-report and provide important information not previously available. In this article, we describe the development and evaluation of currently available transdermal alcohol sensors, present the strengths and limitations of the technology, and give examples of recent research using the sensors.

Subhash R. Lele. “Impact of bootstrap on the estimating functions”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994973](https://doi.org/10.1214/ss/1063994973).

Abstract: Estimating functions form an attractive statistical methodology because of their dependence on only a few features of the underlying probabilistic structure. They also put a premium on developing methods that obtain model-robust confidence intervals. Bootstrap and jackknife ideas can be fruitfully used toward this purpose. Another important area in which bootstrap has proved its use is in the context of detecting the problem of multiple roots and searching for the consistent root of an estimating function. In this article, I review, compare and contrast various approaches for bootstrapping estimating functions.

K. H. Li et al.: Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution

Li-Raghunathan-Rubin-1991

K. H. Li, Trivellore Eachambadi Raghunathan, and Donald B. Rubin. “Large-sample significance levels from multiply imputed data using moment-based statistics and an F reference distribution”. In: *Journal of the American Statistical Association* 86.416 (Dec. 1991), pp. 1065–1073. DOI: [10.1080/01621459.1991.10475152](https://doi.org/10.1080/01621459.1991.10475152).

Abstract: We present a procedure for computing significance levels from data sets whose missing values have been multiply imputed data. This procedure uses moment-based statistics, $m \leq 3$ repeated imputations, and an F reference distribution. When $m = \infty$, we show first that our procedure is essentially the same as the ideal procedure in cases of practical importance and, second, that its deviations from the ideal are basically a function of the coefficient of variation of the canonical ratios of complete to observed information. For small m our procedure’s performance is largely governed by this coefficient of variation and the mean of these ratios. Using simulation

techniques with small m , we compare our procedure’s actual and nominal large-sample significance levels and conclude that it is essentially calibrated and thus represents a definite improvement over previously available procedures. Furthermore, we compare the large-sample power of the procedure as a function of m and other factors, such as the dimensionality of the estimand and fraction of missing information, to provide guidance on the choice of the number of imputations; generally, we find the loss of power due to small m to be quite modest in cases likely to occur in practice.

Y. Li et al.: Bayesian forecasting with a regime-switching zero-inflated multilevel poisson regression model: An application to adolescent alcohol use with spatial covariates

Li-Oravec-Zhou-et-al-2022

Yanling Li, Zita Oravec, et al. “Bayesian forecasting with a regime-switching zero-inflated multilevel poisson regression model: An application to adolescent alcohol use with spatial covariates”. In: *Psychometrika* 87.2 (Jan. 2022), pp. 376–402. DOI: [10.1007/s11336-021-09831-9](https://doi.org/10.1007/s11336-021-09831-9).

Abstract: In this paper, we present and evaluate a novel Bayesian regime-switching zero-inflated multilevel Poisson (RS-ZIMLP) regression model for forecasting alcohol use dynamics. The model partitions individuals’ data into two phases, known as regimes, with: (1) a zero-inflation regime that is used to accommodate high instances of zeros (non-drinking) and (2) a multilevel Poisson regression regime in which variations in individuals’ log-transformed average rates of alcohol use are captured by means of an autoregressive process with exogenous predictors and a person-specific intercept. The times at which individuals are in each regime are unknown, but may be estimated from the data. We assume that the regime indicator follows a first-order Markov process as related to exogenous predictors of interest. The forecast performance of the proposed model was evaluated using a Monte Carlo simulation study and further demonstrated using substance use and spatial covariate data from the Colorado Online Twin Study (CoTwins). Results showed that the proposed model yielded better forecast performance compared to a baseline model which predicted all cases as non-drinking and a reduced ZIMLP model without the RS structure, as indicated by higher AUC (the area under the receiver operating characteristic (ROC) curve) scores, and lower mean absolute

errors (MAEs) and root-mean-square errors (RMSEs). The improvements in forecast performance were even more pronounced when we limited the comparisons to participants who showed at least one instance of transition to drinking.

Y. Li et al.: Fitting multilevel vector autoregressive models in Stan, JAGS, and Mplus

Li-Wood-Ji-et-al-2021

Yanling Li, Julie Wood, et al. “Fitting multilevel vector autoregressive models in Stan, JAGS, and Mplus”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 29.3 (Sept. 2021), pp. 452–475. DOI: [10.1080/10705511.2021.1911657](https://doi.org/10.1080/10705511.2021.1911657).

Abstract: The influx of intensive longitudinal data creates a pressing need for complex modeling tools that help enrich our understanding of how individuals change over time. Multilevel vector autoregressive (mlVAR) models allow for simultaneous evaluations of reciprocal linkages between dynamic processes and individual differences, and have gained increased recognition in recent years. High-dimensional and other complex variations of mlVAR models, though often computationally intractable in the frequentist framework, can be readily handled using Markov chain Monte Carlo techniques in a Bayesian framework. However, researchers in social science fields may be unfamiliar with ways to capitalize on recent developments in Bayesian software programs. In this paper, we provide step-by-step illustrations and comparisons of options to fit Bayesian mlVAR models using Stan, JAGS and Mplus, supplemented with a Monte Carlo simulation study. An empirical example is used to demonstrate the utility of mlVAR models in studying intra- and inter-individual variations in affective dynamics.

Little et al.: Statistical analysis with missing data

Little-Rubin-2019

Roderick J. A. Little and Donald B. Rubin. *Statistical analysis with missing data*. 3rd ed. Wiley, Apr. 2019. ISBN: 9781119482260. DOI: [10.1002/9781119482260](https://doi.org/10.1002/9781119482260).

Abstract: An up-to-date, comprehensive treatment of a classic text on missing data in statistics.

The topic of missing data has gained considerable attention in recent decades. This new edition by two acknowledged experts on the subject offers an up-to-date account of practical methodology for handling missing data problems. Blending theory and application, authors Roderick Little and Donald Rubin review historical approaches to the subject and describe simple methods for multivariate analysis with missing values. They then provide a coherent theory for analysis of problems based on likelihoods derived from statistical models for the data and the missing data mechanism, and then they apply the theory to a wide range of important missing data problems. *Statistical Analysis with Missing Data*, Third Edition starts by introducing readers to the subject and approaches toward solving it. It looks at the patterns and mechanisms that create the missing data, as well as a taxonomy of missing data. It then goes on to examine missing data in experiments, before discussing complete-case and available-case analysis, including weighting methods. The new edition expands its coverage to include recent work on topics such as nonresponse in sample surveys, causal inference, diagnostic methods, and sensitivity analysis, among a host of other topics.

- An updated “classic” written by renowned authorities on the subject
- Features over 150 exercises (including many new ones)
- Covers recent work on important methods like multiple imputation, robust alternatives to weighting, and Bayesian methods
- Revises previous topics based on past student feedback and class experience
- Contains an updated and expanded bibliography

The authors were awarded The Karl Pearson Prize in 2017 by the International Statistical Institute, for a research contribution that has had profound influence on statistical theory, methodology or applications. Their work “has been no less than defining and transforming.” (ISI) *Statistical Analysis with Missing Data*, Third Edition is an ideal textbook for upper undergraduate and/or beginning graduate level students of the subject. It is also an excellent source of information for applied statisticians and practitioners in government and industry.

Library: QA276.

Long et al.: Using heteroscedasticity consistent standard errors in the linear regression model **Long-Ervin-2000**

J. Scott Long and Laurie H. Ervin. "Using heteroscedasticity consistent standard errors in the linear regression model". In: *The American Statistician* 54.3 (Aug. 2000), pp. 217–224. DOI: [10.1080/00031305.2000.10474549](https://doi.org/10.1080/00031305.2000.10474549).

Lütkepohl: New introduction to multiple time series analysis **Luetkepohl-2005**

Helmut Lütkepohl. *New introduction to multiple time series analysis*. Berlin: Springer Berlin Heidelberg, 2005. 764 pp. ISBN: 978-3-540-27752-1. DOI: [10.1007/978-3-540-27752-1](https://doi.org/10.1007/978-3-540-27752-1).

Abstract: This reference work and graduate level textbook considers a wide range of models and methods for analyzing and forecasting multiple time series. The models covered include vector autoregressive, cointegrated, vector autoregressive moving average, multivariate ARCH and periodic processes as well as dynamic simultaneous equations and state space models. Least squares, maximum likelihood and Bayesian methods are considered for estimating these models. Different procedures for model selection and model specification are treated and a wide range of tests and criteria for model checking are introduced. Causality analysis, impulse response analysis and innovation accounting are presented as tools for structural analysis. The book is accessible to graduate students in business and economics. In addition, multiple time series courses in other fields such as statistics and engineering may be based on it. Applied researchers involved in analyzing multiple time series may benefit from the book as it provides the background and tools for their tasks. It bridges the gap to the difficult technical literature on the topic.

D. P. MacKinnon: Analysis of mediating variables in prevention and intervention research. **MacKinnon-1994**

David P. MacKinnon. "Analysis of mediating variables in prevention and intervention research." In: *NIDA research monograph* 139 (1994), pp. 127–153.

Abstract: Mediation analysis is one way to test specific hypotheses derived from theory. Although this analysis has been suggested in the prevention literature, mediation analysis rarely is conducted. As the field of prevention matures, more questions about how prevention programs work (or fail to work) will emerge. Studies of mediation can address these questions, thereby reducing the cost and enhancing the impact of prevention programs. The methods outlined here can be applied in the evaluation of primary, secondary, and tertiary prevention programs. Since most prevention studies include measurement of some mediating constructs, mediation effects can be assessed on many existing data sets. Mediation analysis can be used to test ideas about prevention.

D. P. MacKinnon: Introduction to statistical mediation analysis MacKinnon-2008

David P. MacKinnon. *Introduction to statistical mediation analysis*. Multivariate applications. Hoboken: Erlbaum Psych Press, 2008, p. 488. ISBN: 9780805864298. DOI: [10.4324/9780203809556](https://doi.org/10.4324/9780203809556).

Abstract: This volume introduces the statistical, methodological, and conceptual aspects of mediation analysis. Applications from health, social, and developmental psychology, sociology, communication, exercise science, and epidemiology are emphasized throughout. Single-mediator, multilevel, and longitudinal models are reviewed. The author's goal is to help the reader apply mediation analysis to their own data and understand its limitations. Each chapter features an overview, numerous worked examples, a summary, and exercises (with answers to the odd numbered questions). The accompanying downloadable resources contain outputs described in the book from SAS, SPSS, LISREL, EQS, MPLUS, and CALIS, and a program to simulate the model. The notation used is consistent with existing literature on mediation in psychology. The book opens with a review of the types of research questions the mediation model addresses. Part II describes the estimation of mediation effects including assumptions, statistical tests, and the construction of confidence limits. Advanced models including mediation in path analysis, longitudinal models, multilevel data, categorical variables, and mediation in the context of moderation are then described. The book closes with a discussion of the limits of mediation analysis, additional approaches to identifying mediating variables, and future directions. *Introduction to Statistical Mediation Analysis* is intended for re-

searchers and advanced students in health, social, clinical, and developmental psychology as well as communication, public health, nursing, epidemiology, and sociology. Some exposure to a graduate level research methods or statistics course is assumed. The overview of mediation analysis and the guidelines for conducting a mediation analysis will be appreciated by all readers.

Library: QA278.2 .M29 2008.

Mackinnon et al.: Estimating mediated effects in prevention studies

Mackinnon-Dwyer-1993

David P. Mackinnon and James H. Dwyer. "Estimating mediated effects in prevention studies". In: *Evaluation Review* 17.2 (Apr. 1993), pp. 144–158. DOI: [10.1177/0193841x9301700202](https://doi.org/10.1177/0193841x9301700202).

Abstract: The purpose of this article is to describe statistical procedures to assess how prevention and intervention programs achieve their effects. The analyses require the measurement of intervening or mediating variables hypothesized to represent the causal mechanism by which the prevention program achieves its effects. Methods to estimate mediation are illustrated in the evaluation of a health promotion program designed to reduce dietary cholesterol and a school-based drug prevention program. The methods are relatively easy to apply and the information gained from such analyses should add to our understanding of prevention.

D. P. MacKinnon et al.: Distribution of the product confidence limits for the indirect effect: Program PRODCLIN

MacKinnon-Fritz-Williams-et al-2007

David P. MacKinnon, Matthew S. Fritz, et al. "Distribution of the product confidence limits for the indirect effect: Program PRODCLIN". In: *Behavior Research Methods* 39.3 (Aug. 2007), pp. 384–389. DOI: [10.3758/bf03193007](https://doi.org/10.3758/bf03193007).

Abstract: This article describes a program, PRODCLIN (distribution of the PRODuct Confidence Limits for INdirect effects), written for SAS, SPSS, and R, that computes confidence limits for

the product of two normal random variables. The program is important because it can be used to obtain more accurate confidence limits for the indirect effect, as demonstrated in several recent articles (MacKinnon, Lockwood, & Williams, 2004; Pituch, Whittaker, & Stapleton, 2005). Tests of the significance of and confidence limits for indirect effects based on the distribution of the product method have more accurate Type I error rates and more power than other, more commonly used tests. Values for the two paths involved in the indirect effect and their standard errors are entered in the PRODCLIN program, and distribution of the product confidence limits are computed. Several examples are used to illustrate the PRODCLIN program. The PRODCLIN programs in rich text format may be downloaded from www.psychonomic.org/archive.

D. P. MacKinnon et al.: A comparison of methods to test mediation and other intervening variable effects **MacKinnon-Lockwood-Hoffman-et-al-2002**

David P. MacKinnon, Chondra M. Lockwood, Jeanne M. Hoffman, et al. "A comparison of methods to test mediation and other intervening variable effects". In: *Psychological Methods* 7.1 (2002), pp. 83–104. DOI: [10.1037/1082-989x.7.1.83](https://doi.org/10.1037/1082-989x.7.1.83).

Abstract: A Monte Carlo study compared 14 methods to test the statistical significance of the intervening variable effect. An intervening variable (mediator) transmits the effect of an independent variable to a dependent variable. The commonly used R. M. Baron and D. A. Kenny (1986) approach has low statistical power. Two methods based on the distribution of the product and 2 difference-in-coefficients methods have the most accurate Type I error rates and greatest statistical power except in 1 important case in which Type I error rates are too high. The best balance of Type I error and statistical power across all cases is the test of the joint significance of the two effects comprising the intervening variable effect.

D. P. MacKinnon et al.: Confidence limits for the indirect effect: Distribution of the product and resampling methods **MacKinnon-Lockwood-Williams-2004**

David P. MacKinnon, Chondra M. Lockwood, and Jason Williams. “Confidence limits for the indirect effect: Distribution of the product and resampling methods”. In: *Multivariate Behavioral Research* 39.1 (Jan. 2004), pp. 99–128. DOI: [10.1207/s15327906mbr3901_4](https://doi.org/10.1207/s15327906mbr3901_4).

Abstract: The most commonly used method to test an indirect effect is to divide the estimate of the indirect effect by its standard error and compare the resulting z statistic with a critical value from the standard normal distribution. Confidence limits for the indirect effect are also typically based on critical values from the standard normal distribution. This article uses a simulation study to demonstrate that confidence limits are imbalanced because the distribution of the indirect effect is normal only in special cases. Two alternatives for improving the performance of confidence limits for the indirect effect are evaluated: (a) a method based on the distribution of the product of two normal random variables, and (b) resampling methods. In Study 1, confidence limits based on the distribution of the product are more accurate than methods based on an assumed normal distribution but confidence limits are still imbalanced. Study 2 demonstrates that more accurate confidence limits are obtained using resampling methods, with the bias-corrected bootstrap the best method overall.

J. G. MacKinnon et al.: Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties **MacKinnon-White-1985**

James G. MacKinnon and Halbert White. “Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties”. In: *Journal of Econometrics* 29.3 (Sept. 1985), pp. 305–325. DOI: [10.1016/0304-4076\(85\)90158-7](https://doi.org/10.1016/0304-4076(85)90158-7).

Abstract: We examine several modified versions of the heteroskedasticity-consistent covariance matrix estimator of Hinkley (1977) and White (1980). On the basis of sampling experiments which compare the performance of quasi t-statistics, we find that one estimator, based on the jackknife,

performs better in small samples than the rest. We also examine the finite-sample properties of using modified critical values based on Edgeworth approximations, as proposed by Rothenberg (1984). In addition, we compare the power of several tests for heteroskedasticity, and find that it may be wise to employ the jackknife heteroskedasticity-consistent covariance matrix even in the absence of detected heteroskedasticity.

Manthey et al.: What are the economic costs to society attributable to alcohol use? A systematic review and modelling study **Manthey-Hassan-Carr-et-al-2021**

Jakob Manthey et al. “What are the economic costs to society attributable to alcohol use? A systematic review and modelling study”. In: *PharmacoEconomics* 39.7 (May 2021), pp. 809–822. ISSN: 1179-2027. DOI: [10.1007/s40273-021-01031-8](https://doi.org/10.1007/s40273-021-01031-8).

Abstract: Background: Alcohol-attributable costs to society are captured by cost-of-illness studies, however estimates are often not comparable, e.g. due to the omission of relevant cost components. In this contribution we (1) summarize the societal costs attributable to alcohol use, and (2) estimate the total costs under the assumption that all cost components are considered. Methods: A systematic review and meta-analyses were conducted for studies reporting costs from alcohol consumption for the years 2000 and later, using the EMBASE and MEDLINE databases. Cost estimates were converted into 2019 international dollars (Int\$) per adult and into percentage of gross domestic product (GDP). For each study, weights were calculated to correct for the exclusion of cost indicators. Results: Of 1708 studies identified, 29 were included, and the mean costs of alcohol use amounted to 817.6 Intperadult(95%confidenceinterval[CI]601.8–1033.4), equivalentto1.5%oftheGDP(95%CI1.2–1.7%).Adjustingforomissionofcostcomponents,theeconomiccostsofalcoholconsumptionwereestimatedtoamountto130 per adult (95% CI 873-1738), or 2.6% (95% CI 2.0-3.1%) of the GDP. About one-third of costs (38.8%) were incurred through direct costs, while the majority of costs were due to losses in productivity (61.2%). Discussion: The identified cost studies were mainly conducted in high-income settings, with high heterogeneity in the employed methodology. Accounting for some methodologi-

cal variations, our findings demonstrate that alcohol use continues to incur a high level of cost to many societies.

Maxwell et al.: Bias in cross-sectional analyses of longitudinal mediation

Maxwell-Cole-2007

Scott E. Maxwell and David A. Cole. “Bias in cross-sectional analyses of longitudinal mediation”. In: *Psychological Methods* 12.1 (2007), pp. 23–44. DOI: [10.1037/1082-989x.12.1.23](https://doi.org/10.1037/1082-989x.12.1.23).

Abstract: Most empirical tests of mediation utilize cross-sectional data despite the fact that mediation consists of causal processes that unfold over time. The authors considered the possibility that longitudinal mediation might occur under either of two different models of change: (a) an autoregressive model or (b) a random effects model. For both models, the authors demonstrated that cross-sectional approaches to mediation typically generate substantially biased estimates of longitudinal parameters even under the ideal conditions when mediation is complete. In longitudinal models where variable M completely mediates the effect of X on Y, cross-sectional estimates of the direct effect of X on Y, the indirect effect of X on Y through M, and the proportion of the total effect mediated by M are often highly misleading.

Maxwell et al.: Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model

Maxwell-Cole-Mitchell-2011

Scott E. Maxwell, David A. Cole, and Melissa A. Mitchell. “Bias in cross-sectional analyses of longitudinal mediation: Partial and complete mediation under an autoregressive model”. In: *Multivariate Behavioral Research* 46.5 (Sept. 2011), pp. 816–841. DOI: [10.1080/00273171.2011.606716](https://doi.org/10.1080/00273171.2011.606716).

Abstract: Maxwell and Cole (2007) showed that cross-sectional approaches to mediation typically generate substantially biased estimates of longitudinal parameters in the special case of complete mediation. However, their results did not apply to the more typical case of partial mediation. We extend their previous work by showing that substantial bias can also occur with partial mediation. In

particular, cross-sectional analyses can imply the existence of a substantial indirect effect even when the true longitudinal indirect effect is zero. Thus, a variable that is found to be a strong mediator in a cross-sectional analysis may not be a mediator at all in a longitudinal analysis. In addition, we show that very different combinations of longitudinal parameter values can lead to essentially identical cross-sectional correlations, raising serious questions about the interpretability of cross-sectional mediation data. More generally, researchers are encouraged to consider a wide variety of possible mediation models beyond simple cross-sectional models, including but not restricted to autoregressive models of change.

J. J. McArdle: Latent variable modeling of differences and changes with longitudinal data **McArdle-2009**

John J. McArdle. “Latent variable modeling of differences and changes with longitudinal data”. In: *Annual Review of Psychology* 60.1 (Jan. 2009), pp. 577–605. DOI: [10.1146/annurev.psych.60.110707.163612](https://doi.org/10.1146/annurev.psych.60.110707.163612).

Abstract: This review considers a common question in data analysis: What is the most useful way to analyze longitudinal repeated measures data? We discuss some contemporary forms of structural equation models (SEMs) based on the inclusion of latent variables. The specific goals of this review are to clarify basic SEM definitions, consider relations to classical models, focus on testable features of the new models, and provide recent references to more complete presentations. A broader goal is to illustrate why so many researchers are enthusiastic about the SEM approach to data analysis. We first outline some classic problems in longitudinal data analysis, consider definitions of differences and changes, and raise issues about measurement errors. We then present several classic SEMs based on the inclusion of invariant common factors and explain why these are so important. This leads to newer SEMs based on latent change scores, and we explain why these are useful.

Daniel McNeish and Ellen L. Hamaker. “A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus”. In: *Psychological Methods* 25.5 (Oct. 2020), pp. 610–635. DOI: [10.1037/met0000250](https://doi.org/10.1037/met0000250).

Abstract: Technological advances have led to an increase in intensive longitudinal data and the statistical literature on modeling such data is rapidly expanding, as are software capabilities. Common methods in this area are related to time-series analysis, a framework that historically has received little exposure in psychology. There is a scarcity of psychology-based resources introducing the basic ideas of time-series analysis, especially for data sets featuring multiple people. We begin with basics of $N = 1$ time-series analysis and build up to complex dynamic structural equation models available in the newest release of Mplus Version 8. The goal is to provide readers with a basic conceptual understanding of common models, template code, and result interpretation. We provide short descriptions of some advanced issues, but our main priority is to supply readers with a solid knowledge base so that the more advanced literature on the topic is more readily digestible to a larger group of researchers.

Daniel McNeish and David P. MacKinnon. “Intensive longitudinal mediation in Mplus”. In: *Psychological Methods* (Dec. 2022). DOI: [10.1037/met0000536](https://doi.org/10.1037/met0000536).

Abstract: Much of the existing longitudinal mediation literature focuses on panel data where relatively few repeated measures are collected over a relatively broad timespan. However, technological advances in data collection (e.g., smartphones, wearables) have led to a proliferation of short duration, densely collected longitudinal data in behavioral research. These intensive longitudinal data differ in structure and focus relative to traditionally collected panel data. As a result, existing methodological resources do not necessarily extend to nuances present in the recent influx of inten-

sive longitudinal data and designs. In this tutorial, we first cover potential limitations of traditional longitudinal mediation models to accommodate unique characteristics of intensive longitudinal data. Then, we discuss how recently developed dynamic structural equation models (DSEMs) may be well-suited for mediation modeling with intensive longitudinal data and can overcome some of the limitations associated with traditional approaches. We describe four increasingly complex intensive longitudinal mediation models: (a) stationary models where the indirect effect is constant over time and people, (b) person-specific models where the indirect effect varies across people, (c) dynamic models where the indirect effect varies across time, and (d) cross-classified models where the indirect effect varies across both time and people. We apply each model to a running example featuring a mobile health intervention designed to improve health behavior of individuals with binge eating disorder. In each example, we provide annotated Mplus code and interpretation of the output to guide empirical researchers through mediation modeling with this increasingly popular type of longitudinal data.

Merkel: Docker: Lightweight Linux containers for consistent development and deployment
Merkel-2014

Dirk Merkel. “Docker: Lightweight Linux containers for consistent development and deployment”. In: *Linux Journal* 2014.239 (2014), p. 2. URL: <https://www.linuxjournal.com/content/docker-lightweight-linux-containers-consistent-development-and-deployment>.

Micceri: The unicorn, the normal curve, and other improbable creatures
Micceri-1989

Theodore Micceri. “The unicorn, the normal curve, and other improbable creatures”. In: *Psychological Bulletin* 105.1 (1989), pp. 156–166. DOI: [10.1037/0033-2909.105.1.156](https://doi.org/10.1037/0033-2909.105.1.156).

Roger E. Millsap. *Statistical approaches to measurement invariance*. Routledge, 2011. ISBN: 9780203821961.

DOI: [10.4324/9780203821961](https://doi.org/10.4324/9780203821961).

Peter C. M. Molenaar. “Equivalent Dynamic Models”. In: *Multivariate Behavioral Research* 52.2 (Feb. 2017), pp. 242–258. ISSN: 1532-7906. DOI: [10.1080/00273171.2016.1277681](https://doi.org/10.1080/00273171.2016.1277681).

Abstract: Equivalences of two classes of dynamic models for weakly stationary multivariate time series are discussed: dynamic factor models and autoregressive models. It is shown that exploratory dynamic factor models can be rotated, yielding an infinite set of equivalent solutions for any observed series. It also is shown that dynamic factor models with lagged factor loadings are not equivalent to the currently popular state-space models, and that restriction of attention to the latter type of models may yield invalid results. The known equivalent vector autoregressive model types, standard and structural, are given a new interpretation in which they are conceived of as the extremes of an innovating type of hybrid vector autoregressive models. It is shown that consideration of hybrid models solves many problems, in particular with Granger causality testing.

Peter C.M. Molenaar and Cynthia G. Campbell. “The new person-specific paradigm in psychology”. In: *Current Directions in Psychological Science* 18.2 (Apr. 2009), pp. 112–117. ISSN: 1467-8721. DOI: [10.1111/j.1467-8721.2009.01619.x](https://doi.org/10.1111/j.1467-8721.2009.01619.x).

Abstract: Most research methodology in the behavioral sciences employs interindividual analyses, which provide information about the state of affairs of the population. However, as shown by classical mathematical-statistical theorems (the ergodic theorems), such analyses do not provide

information for, and cannot be applied at, the level of the individual, except on rare occasions when the processes of interest meet certain stringent conditions. When psychological processes violate these conditions, the interindividual analyses that are now standardly applied have to be replaced by analysis of intraindividual variation in order to obtain valid results. Two illustrations involving analysis of intraindividual variation of personality and emotional processes are given.

Moneta et al.: Causal search in structural vector autoregressive models

Moneta-Chlas-Entner-et-al-2011

Alessio Moneta et al. “Causal search in structural vector autoregressive models”. In: *Journal of Machine Learning Research - Proceedings Track 12* (Jan. 2011), pp. 95–114.

Abstract: This paper reviews a class of methods to perform causal inference in the framework of a structural vector autoregressive model. We consider three different settings. In the first setting the underlying system is linear with normal disturbances and the structural model is identified by exploiting the information incorporated in the partial correlations of the estimated residuals. Zero partial correlations are used as input of a search algorithm formalized via graphical causal models. In the second, semi-parametric, setting the underlying system is linear with non-Gaussian disturbances. In this case the structural vector autoregressive model is identified through a search procedure based on independent component analysis. Finally, we explore the possibility of causal search in a nonparametric setting by studying the performance of conditional independence tests based on kernel density estimations.

B. O. Muthén et al.: General longitudinal modeling of individual differences in experimental designs: A latent variable framework for analysis and power estimation.

Muthen-Curran-1997

Bengt O. Muthén and Patrick J. Curran. “General longitudinal modeling of individual differences

in experimental designs: A latent variable framework for analysis and power estimation.” In: *Psychological Methods* 2.4 (Dec. 1997), pp. 371–402. DOI: [10.1037/1082-989x.2.4.371](https://doi.org/10.1037/1082-989x.2.4.371).

Abstract: The generality of latent variable modeling of individual differences in development over time is demonstrated with a particular emphasis on randomized intervention studies. First, a brief overview is given of biostatistical and psychometric approaches to repeated measures analysis. Second, the generality of the psychometric approach is indicated by some nonstandard models. Third, a multiple-population analysis approach is proposed for the estimation of treatment effects. The approach clearly describes the treatment effect as development that differs from normative, control-group development. This framework allows for interactions between treatment and initial status in their effects on development. Finally, an approach for the estimation of power to detect treatment effects in this framework is demonstrated. Illustrations of power calculations are carried out with artificial data, varying the sample sizes, number of timepoints, and treatment effect sizes. Real data are used to illustrate analysis strategies and power calculations. Further modeling extensions are discussed.

L. K. Muthén et al.: Mplus user’s guide. Eighth edition **Muthen-Muthen-2017**

Linda K. Muthén and Bengt O. Muthén. *Mplus user’s guide. Eighth edition*. Los Angeles, CA: Muthén & Muthén, 2017.

National Research Council: An assessment of research-doctorate programs in the United States: Social and behavioral sciences **NationalResearchCouncil-1982**

National Research Council. *An assessment of research-doctorate programs in the United States: Social and behavioral sciences*. Washington, D.C.: National Academies Press, Jan. 1982. DOI: [10.17226/9781](https://doi.org/10.17226/9781).

Neale et al.: OpenMx 2.0: Extended structural equation and statistical modeling

Neale-Hunter-Pritikin-et al-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](https://doi.org/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.

Nesselroade et al.: Sometimes, it’s okay to factor difference scores” - The separation of state and trait anxiety

Nesselroade-Cable-1974

John R. Nesselroade and Dana G. Cable. “Sometimes, it’s okay to factor difference scores” - The separation of state and trait anxiety”. In: *Multivariate Behavioral Research* 9.3 (July 1974), pp. 273–284. DOI: [10.1207/s15327906mbr0903_3](https://doi.org/10.1207/s15327906mbr0903_3).

Abstract: Contemporary psychometric policy and practice have tended to make the use of algebraic difference scores in psychological research taboo. Within the more limited domain of factor analytic research on personality, difference scores have been the subject of sporadic debate for more than 30 years. Using the personality trait versus state distinction as a substantive context, the fit of the factor

analytic model to difference score data is investigated and found to be quite good. Methodological issues related to properties of difference scores and their implications for personality research are briefly discussed.

Newey et al.: A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix **Newey-West-1987**

Whitney K. Newey and Kenneth D. West. “A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix”. In: *Econometrica* 55.3 (May 1987), p. 703. DOI: [10.2307/1913610](https://doi.org/10.2307/1913610).

Nüst et al.: The Rockerverse: Packages and applications for containerisation with R **Nust-Eddelbuettel-Bennett-et-al-2020**

Daniel Nüst et al. “The Rockerverse: Packages and applications for containerisation with R”. In: *The R Journal* 12.1 (2020), p. 437. DOI: [10.32614/rj-2020-007](https://doi.org/10.32614/rj-2020-007).

Abstract: The Rocker Project provides widely used Docker images for R across different application scenarios. This article surveys downstream projects that build upon the Rocker Project images and presents the current state of R packages for managing Docker images and controlling containers. These use cases cover diverse topics such as package development, reproducible research, collaborative work, cloud-based data processing, and production deployment of services. The variety of applications demonstrates the power of the Rocker Project specifically and containerisation in general. Across the diverse ways to use containers, we identified common themes: reproducible environments, scalability and efficiency, and portability across clouds. We conclude that the current growth and diversification of use cases is likely to continue its positive impact, but see the need for consolidating the Rockerverse ecosystem of packages, developing common practices for applications, and exploring alternative containerisation software.

O’Laughlin et al.: Cross-sectional analysis of longitudinal mediation processes

OLaughlin-Martin-Ferrer-2018

Kristine D. O’Laughlin, Monica J. Martin, and Emilio Ferrer. “Cross-sectional analysis of longitudinal mediation processes”. In: *Multivariate Behavioral Research* 53.3 (Apr. 2018), pp. 375–402. ISSN: 1532-7906. DOI: [10.1080/00273171.2018.1454822](https://doi.org/10.1080/00273171.2018.1454822).

Abstract: Statistical mediation analysis can help to identify and explain the mechanisms behind psychological processes. Examining a set of variables for mediation effects is a ubiquitous process in the social sciences literature; however, despite evidence suggesting that cross-sectional data can misrepresent the mediation of longitudinal processes, cross-sectional analyses continue to be used in this manner. Alternative longitudinal mediation models, including those rooted in a structural equation modeling framework (cross-lagged panel, latent growth curve, and latent difference score models) are currently available and may provide a better representation of mediation processes for longitudinal data. The purpose of this paper is twofold: first, we provide a comparison of cross-sectional and longitudinal mediation models; second, we advocate using models to evaluate mediation effects that capture the temporal sequence of the process under study. Two separate empirical examples are presented to illustrate differences in the conclusions drawn from cross-sectional and longitudinal mediation analyses. Findings from these examples yielded substantial differences in interpretations between the cross-sectional and longitudinal mediation models considered here. Based on these observations, researchers should use caution when attempting to use cross-sectional data in place of longitudinal data for mediation analyses.

**O’Rourke et al.: Reasons for testing mediation in the absence of an intervention effect:
A research imperative in prevention and intervention research**

ORourke-MacKinnon-2018

Holly P. O’Rourke and David P. MacKinnon. “Reasons for testing mediation in the absence of an

intervention effect: A research imperative in prevention and intervention research”. In: *Journal of Studies on Alcohol and Drugs* 79.2 (Mar. 2018), pp. 171–181. DOI: [10.15288/jsad.2018.79.171](https://doi.org/10.15288/jsad.2018.79.171).

Abstract: **Objective:** Mediation models are used in prevention and intervention research to assess the mechanisms by which interventions influence outcomes. However, researchers may not investigate mediators in the absence of intervention effects on the primary outcome variable. There is emerging evidence that in some situations, tests of mediated effects can be statistically significant when the total intervention effect is not statistically significant. In addition, there are important conceptual and practical reasons for investigating mediation when the intervention effect is non-significant. **Method:** This article discusses the conditions under which mediation may be present when an intervention effect does not have a statistically significant effect and why mediation should always be considered important. **Results:** Mediation may be present in the following conditions: when the total and mediated effects are equal in value, when the mediated and direct effects have opposing signs, when mediated effects are equal across single and multiple-mediator models, and when specific mediated effects have opposing signs. Mediation should be conducted in every study because it provides the opportunity to test known and replicable mediators, to use mediators as an intervention manipulation check, and to address action and conceptual theory in intervention models. **Conclusions:** Mediators are central to intervention programs, and mediators should be investigated for the valuable information they provide about the success or failure of interventions.

O’Rourke et al.: The importance of mediation analysis in substance-use prevention

ORourke-MacKinnon-2019

Holly P. O’Rourke and David P. MacKinnon. “The importance of mediation analysis in substance-use prevention”. In: *Advances in Prevention Science*. Springer International Publishing, 2019, pp. 233–246. DOI: [10.1007/978-3-030-00627-3_15](https://doi.org/10.1007/978-3-030-00627-3_15).

Abstract: This chapter describes the theoretical and practical importance of mediation analysis in substance-use prevention research. The most important reason for including mediators in a re-

search study is to examine the mechanisms by which prevention programs influence substance-use outcomes. Understanding the mechanisms by which prevention programs achieve effects helps reduce the cost and increases effectiveness of prevention programs. This chapter first describes the theoretical foundations of the mediation model in prevention, and reasons for using mediation analysis in substance-use prevention. Next, we provide an overview of statistical mediation analysis for single and multiple mediator models. We summarize mediation analyses in substance-use prevention and discuss future directions for application of mediation analysis in substance-use research.

Oravecz et al.: A hierarchical latent stochastic differential equation model for affective dynamics
Oravecz-Tuerlinckx-Vandekerckhove-2011

Zita Oravecz, Francis Tuerlinckx, and Joachim Vandekerckhove. “A hierarchical latent stochastic differential equation model for affective dynamics”. In: *Psychological Methods* 16.4 (2011), pp. 468–490. DOI: [10.1037/a0024375](https://doi.org/10.1037/a0024375).

Abstract: In this article a continuous-time stochastic model (the Ornstein-Uhlenbeck process) is presented to model the perpetually altering states of the core affect, which is a 2-dimensional concept underlying all our affective experiences. The process model that we propose can account for the temporal changes in core affect on the latent level. The key parameters of the model are the average position (also called home base), the variances and covariances of the process, and the regulatory mechanisms that keep the process in the vicinity of the average position. To account for individual differences, the model is extended hierarchically. A particularly novel contribution is that in principle all parameters of the stochastic process (not only the mean but also its variance and the regulatory parameters) are allowed to differ between individuals. In this way, the aim is to understand the affective dynamics of single individuals and at the same time investigate how these individuals differ from one another. The final model is a continuous-time state-space model for repeated measurement data taken at possibly irregular time points. Both time-invariant and time-varying covariates can be included to investigate sources of individual differences. As

an illustration, the model is applied to a diary study measuring core affect repeatedly for several individuals (thereby generating intensive longitudinal data).

Osborne et al.: A Longitudinal Investigation of the Intellectual Differentiation Hypothesis
Osborne-Suddick-1972

R. T. Osborne and D. E. Suddick. “A Longitudinal Investigation of the Intellectual Differentiation Hypothesis”. In: *The Journal of Genetic Psychology* 121.1 (Sept. 1972), pp. 83–89. ISSN: 1940-0896. DOI: [10.1080/00221325.1972.10533131](https://doi.org/10.1080/00221325.1972.10533131).

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](https://doi.org/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 discrete and 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.

Johan H. Oud, John H. van den Bercken, and Raymond J. Essers. “Longitudinal factor score estimation using the Kalman filter”. In: *Applied Psychological Measurement* 14.4 (Dec. 1990), pp. 395–418. DOI: [10.1177/014662169001400406](https://doi.org/10.1177/014662169001400406).

Abstract: The advantages of the Kalman filter as a factor score estimator in the presence of longitudinal data are described. Because the Kalman filter presupposes the availability of a dynamic state space model, the state space model is reviewed first, and it is shown to be translatable into the LISREL model. Several extensions of the LISREL model specification are discussed in order to enhance the applicability of the Kalman filter for behavioral research data. The Kalman filter and its main properties are summarized. Relationships are shown between the Kalman filter and two well-known cross-sectional factor score estimators: the regression estimator, and the Bartlett estimator. The indeterminacy problem of factor scores is also discussed in the context of Kalman filtering, and the differences are described between Kalman filtering on the basis of a zero-means and a structured-means LISREL model. By using a structured-means LISREL model, the Kalman filter is capable of estimating absolute latent developmental curves. An educational research example is presented. Index terms: factor score estimation, indeterminacy of factor scores, Kalman filter, LISREL longitudinal LISREL modeling, longitudinal factor analysis, state space modeling.

Johan H. L. Oud and Robert A. R. G. Jansen. “Continuous time state space modeling of panel data by means of SEM”. In: *Psychometrika* 65.2 (June 2000), pp. 199–215. DOI: [10.1007/bf02294374](https://doi.org/10.1007/bf02294374).

Abstract: Maximum likelihood parameter estimation of the continuous time linear stochastic state space model is considered on the basis of largeN discrete time data using a structural equation

modeling (SEM) program. Random subject effects are allowed to be part of the model. The exact discrete model (EDM) is employed which links the discrete time model parameters to the underlying continuous time model parameters by means of nonlinear restrictions. The EDM is generalized to cover not only time-invariant parameters but also the cases of stepwise time-varying (piecewise time-invariant) parameters and parameters varying continuously over time according to a general polynomial scheme. The identification of the continuous time parameters is discussed and an educational example is presented.

Park et al.: Subgrouping with chain graphical VAR models

Park-Chow-Epskamp-et-al-2023

Jonathan J. Park, Sy-Miin Chow, et al. “Subgrouping with chain graphical VAR models”. In: *Multivariate Behavioral Research* (2023).

Abstract: Recent years have seen the emergence of an “idio-thetic” class of methods to bridge the gap between nomothetic and idiographic inference. These methods describe nomothetic trends in idiographic processes by pooling intraindividual information across individuals to inform group-level inference or vice versa. The current work introduces a novel “idio-thetic” model: the subgrouped chain graphical vector autoregression (scGVAR). The scGVAR is unique in its ability to identify subgroups of individuals who share common dynamic network structures in both lag(1) and contemporaneous effects. Results from Monte Carlo simulations indicate that the scGVAR shows promise over similar approaches when clusters of individuals differ in their contemporaneous dynamics and in showing increased sensitivity in detecting nuanced group differences while keeping Type-I error rates low. In contrast, a competing approach—the Alternating Least Squares VAR (ALS VAR) performs well when groups were separated by larger distances. Further considerations are provided regarding applications of the ALS VAR and scGVAR on real data and the strengths and limitations of both methods.

Park et al.: On subgrouping continuous processes in discrete time

Park-Fisher-Chow-etal-2023a

Jonathan J. Park, Zachary Fisher, et al. “On subgrouping continuous processes in discrete time”. In: *Multivariate Behavioral Research* 58.1 (Jan. 2023), pp. 154–155. ISSN: 1532-7906. DOI: [10.1080/00273171.2022.2160957](https://doi.org/10.1080/00273171.2022.2160957).

Park et al.: Evaluating discrete time methods for subgrouping continuous processes

Park-Fisher-Chow-etal-2023b

Jonathan J. Park, Zachary F. Fisher, et al. “Evaluating discrete time methods for subgrouping continuous processes”. In: *Multivariate Behavioral Research* (Aug. 2023), pp. 1–13. ISSN: 1532-7906. DOI: [10.1080/00273171.2023.2235685](https://doi.org/10.1080/00273171.2023.2235685).

Abstract: Rapid developments over the last several decades have brought increased focus and attention to the role of time scales and heterogeneity in the modeling of human processes. To address these emerging questions, subgrouping methods developed in the discrete-time framework—such as the vector autoregression (VAR)—have undergone widespread development to identify shared nomothetic trends from idiographic modeling results. Given the dependence of VAR-based parameters on the measurement intervals of the data, we sought to clarify the strengths and limitations of these methods in recovering subgroup dynamics under different measurement intervals. Building on the work of Molenaar and collaborators for subgrouping individual time-series by means of the subgrouped chain graphical VAR (scgVAR) and the subgrouping option in the group iterative multiple model estimation (S-GIMME), we present results from a Monte Carlo study aimed at addressing the implications of identifying subgroups using these discrete-time methods when applied to continuous-time data. Results indicate that discrete-time subgrouping methods perform well at recovering true subgroups when the measurement intervals are large enough to capture the full range of a system’s dynamics, either via lagged or contemporaneous effects. Further implications and limitations are discussed therein.

Patrick et al.: Monitoring the Future Panel Study annual report: National data on substance use among adults ages 19 to 60, 1976-2022

Patrick-Miech-Johnston-et-al-2023

Megan Patrick et al. *Monitoring the Future Panel Study annual report: National data on substance use among adults ages 19 to 60, 1976-2022*. Ann Arbor, MI: Institute for Social Research, The University of Michigan, 2023. DOI: [10.7826/isr-um.06.585140.002.07.0002.2023](https://doi.org/10.7826/isr-um.06.585140.002.07.0002.2023).

Pawitan: In all likelihood: Statistical modelling and inference using likelihood

Pawitan-2013

Yudi Pawitan. *In all likelihood: Statistical modelling and inference using likelihood*. Oxford University Press, Jan. 17, 2013. 544 pp. ISBN: 9780199671229.

Abstract: Based on a course in the theory of statistics this text concentrates on what can be achieved using the likelihood/Fisherian method of taking account of uncertainty when studying a statistical problem. It takes the concept of the likelihood as providing the best methods for unifying the demands of statistical modelling and the theory of inference. Every likelihood concept is illustrated by realistic examples, which are not compromised by computational problems. Examples range from a simple comparison of two accident rates, to complex studies that require generalised linear or semiparametric modelling. The emphasis is that the likelihood is not simply a device to produce an estimate, but an important tool for modelling. The book generally takes an informal approach, where most important results are established using heuristic arguments and motivated with realistic examples. With the currently available computing power, examples are not contrived to allow a closed analytical solution, and the book can concentrate on the statistical aspects of the data modelling. In addition to classical likelihood theory, the book covers many modern topics such as generalized linear models and mixed models, non parametric smoothing, robustness, the EM algorithm and empirical likelihood.

Pesigan: Confidence intervals for standardized coefficients: Applied to regression coefficients in primary studies and indirect effects in meta-analytic structural equation modeling **Pesigan-2022**

Ivan Jacob Agaloos Pesigan. “Confidence intervals for standardized coefficients: Applied to regression coefficients in primary studies and indirect effects in meta-analytic structural equation modeling”. PhD thesis. University of Macau, 2022.

Pesigan et al.: SEM-based methods to form confidence intervals for indirect effect: Still applicable given nonnormality, under certain conditions **Pesigan-Cheung-2020**

Ivan Jacob Agaloos Pesigan and Shu Fai Cheung. “SEM-based methods to form confidence intervals for indirect effect: Still applicable given nonnormality, under certain conditions”. In: *Frontiers in Psychology* 11 (Dec. 2020). DOI: [10.3389/fpsyg.2020.571928](https://doi.org/10.3389/fpsyg.2020.571928).

Abstract: A SEM-based approach using likelihood-based confidence interval (LBCI) has been proposed to form confidence intervals for unstandardized and standardized indirect effect in mediation models. However, when used with the maximum likelihood estimation, this approach requires that the variables are multivariate normally distributed. This can affect the LBCIs of unstandardized and standardized effect differently. In the present study, the robustness of this approach when the predictor is not normally distributed but the error terms are conditionally normal, which does not violate the distributional assumption of ordinary least squares (OLS) estimation, is compared to four other approaches: nonparametric bootstrapping, two variants of LBCI, LBCI assuming the predictor is fixed (LBCI-Fixed-X) and LBCI based on ADF estimation (LBCI-ADF), and Monte Carlo. A simulation study was conducted using a simple mediation model and a serial mediation model, manipulating the distribution of the predictor. The Monte Carlo method performed worst among the methods. LBCI and LBCI-Fixed-X had suboptimal performance when the distributions had high kurtosis and the population indirect effects were medium to large. In some conditions, the problem was severe even when the sample size was large. LBCI-ADF and nonparametric boot-

strapping had coverage probabilities close to the nominal value in nearly all conditions, although the coverage probabilities were still suboptimal for the serial mediation model when the sample size was small with respect to the model. Implications of these findings in the context of this special case of nonnormal data were discussed.

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](https://doi.org/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; Tofghi & 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](https://doi.org/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.

Peugh et al.: Missing data in educational research: A review of reporting practices and suggestions for improvement
Peugh-Enders-2004

James L. Peugh and Craig K. Enders. “Missing data in educational research: A review of reporting practices and suggestions for improvement”. In: *Review of Educational Research* 74.4 (Dec. 2004), pp. 525–556. DOI: [10.3102/00346543074004525](https://doi.org/10.3102/00346543074004525).

Abstract: Missing data analyses have received considerable recent attention in the methodological literature, and two “modern” methods, multiple imputation and maximum likelihood estimation, are recommended. The goals of this article are to (a) provide an overview of missing-data theory, maximum likelihood estimation, and multiple imputation; (b) conduct a methodological review of missing-data reporting practices in 23 applied research journals; and (c) provide a demonstration of multiple imputation and maximum likelihood estimation using the Longitudinal Study of American Youth data. The results indicated that explicit discussions of missing data increased substantially between 1999 and 2003, but the use of maximum likelihood estimation or multiple imputation was rare; the studies relied almost exclusively on listwise and pairwise deletion.

Piasecki: Assessment of alcohol use in the natural environment**Piasecki-2019**

Thomas M. Piasecki. “Assessment of alcohol use in the natural environment”. In: *Alcoholism: Clinical and Experimental Research* 43.4 (Mar. 2019), pp. 564–577. ISSN: 1530-0277. DOI: [10.1111/acer.13975](https://doi.org/10.1111/acer.13975).

Abstract: The current article critically reviews 3 methodological options for assessing drinking episodes in the natural environment. Ecological momentary assessment (EMA) typically involves using mobile devices to collect self-report data from participants in daily life. This technique is now widely used in alcohol research, but investigators have implemented diverse assessment strategies. This article focuses on “high-resolution” EMA protocols that oversample experiences and behaviors within individual drinking episodes. A number of approaches have been used to accomplish this, including using signaled follow-ups tied to drinking initiation, asking participants to log entries before and after individual drinks or drinking episodes, and delivering frequent signaled assessments during periods of the day when alcohol use is most common. Transdermal alcohol sensors (TAS) are devices that are worn continuously and are capable of detecting alcohol eliminated through the skin. These methods are appealing because they do not rely upon drinkers’ self-report. Studies using TAS have been appearing with greater frequency over the past several years. New methods are making the use of TAS more tractable by permitting back-translation of transdermal alcohol

concentration data to more familiar estimates of blood alcohol concentration or breath alcohol concentration. However, the current generation of devices can have problems with missing data and tend to be relatively insensitive to low-level drinking. An emerging area of research investigates the possibility of using mobile device data and machine learning to passively detect the user's drinking, with promising early findings. EMA, TAS, and sensor-based approaches are all valid, and tend to produce convergent information when used in conjunction with one another. Each has a unique profile of advantages, disadvantages, and threats to validity. Therefore, the nature of the underlying research question must dictate the method(s) investigators select.

Politis: The impact of bootstrap methods on time series analysis

Politis-2003

Dimitris N. Politis. "The impact of bootstrap methods on time series analysis". In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994977](https://doi.org/10.1214/ss/1063994977).

Abstract: Sparked by Efron's seminal paper, the decade of the 1980s was a period of active research on bootstrap methods for independent data—mainly i.i.d. or regression set-ups. By contrast, in the 1990s much research was directed towards resampling dependent data, for example, time series and random fields. Consequently, the availability of valid nonparametric inference procedures based on resampling and/or subsampling has freed practitioners from the necessity of resorting to simplifying assumptions such as normality or linearity that may be misleading.

Preacher et al.: SPSS and SAS procedures for estimating indirect effects in simple mediation models

Preacher-Hayes-2004

Kristopher J. Preacher and Andrew F. Hayes. "SPSS and SAS procedures for estimating indirect effects in simple mediation models". In: *Behavior Research Methods, Instruments, & Computers* 36.4 (Nov. 2004), pp. 717–731. DOI: [10.3758/bf03206553](https://doi.org/10.3758/bf03206553).

Abstract: Researchers often conduct mediation analysis in order to indirectly assess the effect of a proposed cause on some outcome through a proposed mediator. The utility of mediation analysis

stems from its ability to go beyond the merely descriptive to a more functional understanding of the relationships among variables. A necessary component of mediation is a statistically and practically significant indirect effect. Although mediation hypotheses are frequently explored in psychological research, formal significance tests of indirect effects are rarely conducted. After a brief overview of mediation, we argue the importance of directly testing the significance of indirect effects and provide SPSS and SAS macros that facilitate estimation of the indirect effect with a normal theory approach and a bootstrap approach to obtaining confidence intervals, as well as the traditional approach advocated by Baron and Kenny (1986). We hope that this discussion and the macros will enhance the frequency of formal mediation tests in the psychology literature. Electronic copies of these macros may be downloaded from the Psychonomic Society's Web archive at www.psychonomic.org/archive/.

Preacher et al.: Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models **Preacher-Hayes-2008**

Kristopher J. Preacher and Andrew F. Hayes. "Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models". In: *Behavior Research Methods* 40.3 (Aug. 2008), pp. 879–891. DOI: [10.3758/brm.40.3.879](https://doi.org/10.3758/brm.40.3.879).

Abstract: Hypotheses involving mediation are common in the behavioral sciences. Mediation exists when a predictor affects a dependent variable indirectly through at least one intervening variable, or mediator. Methods to assess mediation involving multiple simultaneous mediators have received little attention in the methodological literature despite a clear need. We provide an overview of simple and multiple mediation and explore three approaches that can be used to investigate indirect processes, as well as methods for contrasting two or more mediators within a single model. We present an illustrative example, assessing and contrasting potential mediators of the relationship between the helpfulness of socialization agents and job satisfaction. We also provide SAS and SPSS macros, as well as Mplus and LISREL syntax, to facilitate the use of these methods in applications.

Kristopher J. Preacher and James P. Selig. “Advantages of Monte Carlo confidence intervals for indirect effects”. In: *Communication Methods and Measures* 6.2 (Apr. 2012), pp. 77–98. DOI: [10.1080/19312458.2012.679848](https://doi.org/10.1080/19312458.2012.679848).

Abstract: Monte Carlo simulation is a useful but underutilized method of constructing confidence intervals for indirect effects in mediation analysis. The Monte Carlo confidence interval method has several distinct advantages over rival methods. Its performance is comparable to other widely accepted methods of interval construction, it can be used when only summary data are available, it can be used in situations where rival methods (e.g., bootstrapping and distribution of the product methods) are difficult or impossible, and it is not as computer-intensive as some other methods. In this study we discuss Monte Carlo confidence intervals for indirect effects, report the results of a simulation study comparing their performance to that of competing methods, demonstrate the method in applied examples, and discuss several software options for implementation in applied settings.

R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria, 2021. URL: <https://www.R-project.org/>.

R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria, 2022. URL: <https://www.R-project.org/>.

R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria, 2023. URL: <https://www.R-project.org/>.

Trivellore E. Raghunathan et al.: A multivariate technique for multiply imputing missing values using a sequence of regression models

Raghunathan-Lepkowski-Hoewyk-et-al-2001

Trivellore E. Raghunathan et al. “A multivariate technique for multiply imputing missing values using a sequence of regression models”. In: *Survey Methodology* 27.1 (2001), pp. 85–95.

Abstract: This article describes and evaluates a procedure for imputing missing values for a relatively complex data structure when the data are missing at random. The imputations are obtained by fitting a sequence of regression models and drawing values from the corresponding predictive distributions. The types of regression models used are linear, logistic, Poisson, generalized logit or a mixture of these depending on the type of variable being imputed. Two additional common features in the imputation process are incorporated: restriction to a relevant subpopulation for some variables and logical bounds or constraints for the imputed values. The restrictions involve subsetting the sample individuals that satisfy certain criteria while fitting the regression models. The bounds involve drawing values from a truncated predictive distribution. The development of this method was partly motivated by the analysis of two data sets which are used as illustrations. The sequential regression procedure is applied to perform multiple imputation analysis for the two applied problems. The sampling properties of inferences from multiply imputed data sets created using the sequential regression method are evaluated through simulated data sets.

Rasmussen: Estimating correlation coefficients: Bootstrap and parametric approaches

Rasmussen-1987

Jeffrey L. Rasmussen. “Estimating correlation coefficients: Bootstrap and parametric approaches”. In: *Psychological Bulletin* 101.1 (1987), pp. 136–139. DOI: [10.1037/0033-2909.101.1.136](https://doi.org/10.1037/0033-2909.101.1.136).

Abstract: The bootstrap, a computer-intensive approach to statistical data analysis, has been recommended as an alternative to parametric approaches. Advocates claim it is superior because it is not burdened by potentially unwarranted normal theory assumptions and because it retains information about the form of the original sample. Empirical support for its superiority, however, is quite limited. The present article compares the bootstrap and parametric approaches to estimating confidence intervals and Type I error rates of the correlation coefficient. The parametric approach is superior to the bootstrap under both assumption violation and nonviolation. The bootstrap results in overly restricted confidence intervals and overly liberal Type I error rates.

Reichardt: Commentary: Are three waves of data sufficient for assessing mediation?

Reichardt-2011

Charles S. Reichardt. “Commentary: Are three waves of data sufficient for assessing mediation?” In: *Multivariate Behavioral Research* 46.5 (Sept. 2011), pp. 842–851. ISSN: 1532-7906. DOI: [10.1080/00273171.2011.606740](https://doi.org/10.1080/00273171.2011.606740).

Abstract: Maxwell, Cole, and Mitchell (2011) demonstrated that simple structural equation models, when used with cross-sectional data, generally produce biased estimates of mediated effects. I extend those results by showing how simple structural equation models can produce biased estimates of mediated effects when used even with longitudinal data. Even with longitudinal data, simple autoregressive structural equation models can imply the existence of indirect effects when only direct effects exist and the existence of direct effects when only indirect effects exist.

Randall R. Robey and Robert S. Barcikowski. "Type I error and the number of iterations in Monte Carlo studies of robustness". In: *British Journal of Mathematical and Statistical Psychology* 45.2 (Nov. 1992), pp. 283–288. DOI: [10.1111/j.2044-8317.1992.tb00993.x](https://doi.org/10.1111/j.2044-8317.1992.tb00993.x).

Abstract: A recent survey of simulation studies concluded that an overwhelming majority of papers do not report a rationale for the decision regarding the number of Monte Carlo iterations. A surprisingly large number of reports do not contain a justifiable definition of robustness and many studies are conducted with an insufficient number of iterations to achieve satisfactory statistical conclusion validity. The implication is that we do not follow our own advice regarding the management of Type I and Type II errors when conducting Monte Carlo experiments. This paper reports a straightforward application of a well-known procedure for the purpose of objectively determining the exact number of iterations necessary to confidently detect departures from robustness in Monte Carlo results. A table of the number of iterations necessary to detect departures from a series of nominal Type I error rates is included.

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

Rousselet et al.: The percentile bootstrap: A primer with step-by-step instructions in R **Rousselet-Pernet-Wilcox-2021**

Guillaume A. Rousselet, Cyril R. Pernet, and Rand R. Wilcox. “The percentile bootstrap: A primer with step-by-step instructions in R”. In: *Advances in Methods and Practices in Psychological Science* 4.1 (Jan. 2021), pp. 1–10. DOI: [10.1177/2515245920911881](https://doi.org/10.1177/2515245920911881).

Abstract: The percentile bootstrap is the Swiss Army knife of statistics: It is a nonparametric method based on data-driven simulations. It can be applied to many statistical problems, as a substitute to standard parametric approaches, or in situations for which parametric methods do not exist. In this Tutorial, we cover R code to implement the percentile bootstrap to make inferences about central tendency (e.g., means and trimmed means) and spread in a one-sample example and in an example comparing two independent groups. For each example, we explain how to derive a bootstrap distribution and how to get a confidence interval and a p value from that distribution. We also demonstrate how to run a simulation to assess the behavior of the bootstrap. For some purposes, such as making inferences about the mean, the bootstrap performs poorly. But for other purposes, it is the only known method that works well over a broad range of situations. More broadly, combining the percentile bootstrap with robust estimators (i.e., estimators that are not overly sensitive to outliers) can help users gain a deeper understanding of their data than they would using conventional methods.

Rubin: Inference and missing data **Rubin-1976**

Donald B. Rubin. “Inference and missing data”. In: *Biometrika* 63.3 (1976), pp. 581–592. DOI: [10.1093/biomet/63.3.581](https://doi.org/10.1093/biomet/63.3.581).

Abstract: When making sampling distribution inferences about the parameter of the data, θ , it is appropriate to ignore the process that causes missing data if the missing data are ‘missing at random’ and the observed data are ‘observed at random’, but these inferences are generally conditional on the observed pattern of missing data. When making direct-likelihood or Bayesian inferences about θ , it is appropriate to ignore the process that causes missing data if the missing data are missing at random and the parameter of the missing data process is ‘distinct’ from θ . These conditions are the weakest general conditions under which ignoring the process that causes missing data always leads to correct inferences.

Rubin: Multiple imputation for nonresponse in surveys

Rubin-1987

Donald B. Rubin. *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons, Inc., June 1987. ISBN: 9780470316696. DOI: [10.1002/9780470316696](https://doi.org/10.1002/9780470316696).

Abstract: Demonstrates how nonresponse in sample surveys and censuses can be handled by replacing each missing value with two or more multiple imputations. Clearly illustrates the advantages of modern computing to such handle surveys, and demonstrates the benefit of this statistical technique for researchers who must analyze them. Also presents the background for Bayesian and frequentist theory. After establishing that only standard complete-data methods are needed to analyze a multiply-imputed set, the text evaluates procedures in general circumstances, outlining specific procedures for creating imputations in both the ignorable and nonignorable cases. Examples and exercises reinforce ideas, and the interplay of Bayesian and frequentist ideas presents a unified picture of modern statistics.

Library: HA31.2 .R83 1987.

Russell et al.: Baseline protective behavioral strategy use predicts more moderate transdermal alcohol concentration dynamics and fewer negative consequences of drinking in young adults' natural settings. **Russell-Smyth-Turrisi-Rodriguez-2023**

Michael A. Russell, Joshua M. Smyth, et al. "Baseline protective behavioral strategy use predicts more moderate transdermal alcohol concentration dynamics and fewer negative consequences of drinking in young adults' natural settings." In: *Psychology of Addictive Behaviors* (June 2023). ISSN: 0893-164X. DOI: [10.1037/adb0000941](https://doi.org/10.1037/adb0000941).

Abstract: Objective: Test whether frequent protective behavioral strategies (PBS) users report (a) fewer alcohol-related consequences and (b) less risky alcohol intoxication dynamics (measured via transdermal alcohol concentration [TAC] sensor "features") in daily life. Method: Two hundred twenty-two frequently heavy-drinking young adults ($M_{\text{age}} = 22.3$ years) wore TAC sensors for 6 consecutive days. TAC features peak (maximum TAC), rise rate (speed of TAC increase), and area under the curve (AUC) were derived for each day. Negative alcohol-related consequences were measured in the morning after each self-reported drinking day. Past-year PBS use was measured at baseline. Results: Young adults reporting more frequent baseline PBS use showed (a) fewer alcohol-related consequences and (b) lower intoxication dynamics on average (less AUC, lower peaks, and slower rise rates). Limiting/stopping and manner of drinking PBS showed the same pattern of findings as the total score. Serious harm reduction PBS predicted fewer negative alcohol-related consequences, but not TAC features. Multilevel path models showed that TAC features peak and rise rate partially explained associations between PBS (total, limiting/stopping, and manner of drinking) and consequences. Independent contributions of PBS subscales were small and nonsignificant, suggesting that total PBS use was a more important predictor of risk/protection than the specific types of PBS used. Conclusions: Young adults using more total PBS may experience fewer alcohol-related consequences during real-world drinking episodes in part through less risky intoxication dynamics (TAC features). Future research measuring PBS at the daily level is needed to formally test TAC features as day-level mechanisms of protection from acute alcohol-related consequences.

Michael A. Russell, Robert J. Turrisi, and Joshua M. Smyth. "Transdermal sensor features correlate with ecological momentary assessment drinking reports and predict alcohol-related consequences in young adults' natural settings". In: *Alcoholism: Clinical and Experimental Research* 46.1 (Jan. 2022), pp. 100–113. ISSN: 1530-0277. DOI: [10.1111/acer.14739](https://doi.org/10.1111/acer.14739).

Abstract: Background: Wearable transdermal alcohol concentration (TAC) sensors allow passive monitoring of alcohol concentration in natural settings and measurement of multiple features from drinking episodes, including peak intoxication level, speed of intoxication (absorption rate) and elimination, and duration. These passively collected features extend commonly used self-reported drink counts and may facilitate the prediction of alcohol-related consequences in natural settings, aiding risk stratification and prevention efforts. Method: A total of 222 young adults aged 21-29 ($M_{\text{age}} = 22.3$, 64 female, 79% non-Hispanic white, 84% undergraduates) who regularly drink heavily participated in a 5-day study that included the ecological momentary assessment (EMA) of alcohol consumption (daily morning reports and participant-initiated episodic EMA sequences) and the wearing of TAC sensors (SCRAM-CAM anklets). The analytic sample contained 218 participants and 1274 days (including 554 self-reported drinking days). Five features—area under the curve (AUC), peak TAC, rise rate (rate of absorption), fall rate (rate of elimination), and duration—were extracted from TAC-positive trajectories for each drinking day. Day- and person-level associations of TAC features with drink counts (morning and episodic EMA) and alcohol-related consequences were tested using multilevel modeling. Results: TAC features were strongly associated with morning drink reports ($r = 0.60.7$) but only moderately associated with episodic EMA drink counts ($r = 0.30.5$) at both day and person levels. Higher peaks, larger AUCs, faster rise rates, and faster fall rates were significantly predictive of day-level alcohol-related consequences after adjusting for both morning and episodic EMA drink counts in separate models. Person means of TAC features added little above daily scores to the prediction of alcohol-related consequences. Conclusions: These

results support the utility of TAC sensors in studies of alcohol misuse among young adults in natural settings and outline the specific TAC features that contribute to the day-level prediction of alcohol-related consequences. TAC sensors provide a passive option for obtaining valid and unique information predictive of drinking risk in natural settings.

Ryan et al.: Time to intervene: A continuous-time approach to network analysis and centrality
Ryan-Hamaker-2021

Oisin Ryan and Ellen L. Hamaker. “Time to intervene: A continuous-time approach to network analysis and centrality”. In: *Psychometrika* 87.1 (June 2021), pp. 214–252. DOI: [10.1007/s11336-021-09767-0](https://doi.org/10.1007/s11336-021-09767-0).

Abstract: Network analysis of ESM data has become popular in clinical psychology. In this approach, discrete-time (DT) vector auto-regressive (VAR) models define the network structure with centrality measures used to identify intervention targets. However, VAR models suffer from time-interval dependency. Continuous-time (CT) models have been suggested as an alternative but require a conceptual shift, implying that DT-VAR parameters reflect total rather than direct effects. In this paper, we propose and illustrate a CT network approach using CT-VAR models. We define a new network representation and develop centrality measures which inform intervention targeting. This methodology is illustrated with an ESM dataset.

Sacks et al.: 2010 national and state costs of excessive alcohol consumption
Sacks-Gonzales-Bouchery-et al-2015

Jeffrey J. Sacks et al. “2010 national and state costs of excessive alcohol consumption”. In: *American Journal of Preventive Medicine* 49.5 (Nov. 2015), e73–e79. ISSN: 0749-3797. DOI: [10.1016/j.amepre.2015.05.031](https://doi.org/10.1016/j.amepre.2015.05.031).

Abstract: Introduction: Excessive alcohol use cost the U.S. 223.5 billion in 2006. Given economic shifts in the U.S. since 2006, current estimates are needed to help inform the planning of prevention strategies. Methods : From March 2012 to March 201

attributable deaths) and price (e.g., inflation rate in cost of medical care). The total cost, cost to government, and costs for binge drinking cost the U.S. \$249.0 billion in 2010, or about \$2.05 per drink. Government paid for 100.7 billion (40.4%) of these costs. Binge drinking accounted for 191.1 billion (76.7%) of costs; underage drinking 24.3 billion (9.7%) of costs; and drinking while pregnant 5.5 billion (2.2%) of costs. The median cost per state was \$3.5 billion. Binge drinking was responsible for 70% of these costs in all states, and 40% of the binge drinking-related costs were paid by government. Conclusions: Excessive drinking cost the nation almost \$250 billion in 2010. Two of every five of the total cost was paid by government, and three quarters of the costs were due to binge drinking. Several evidence-based strategies can help reduce excessive drinking and related costs, including increasing alcohol excise taxes, limiting alcohol outlet density, and commercial host liability.

SAMHSA: Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health (HHS Publication No. PEP20-07-01-001, NSDUH Series H-55) **SAMHSA-2020**

SAMHSA. *Key substance use and mental health indicators in the United States: Results from the 2019 National Survey on Drug Use and Health (HHS Publication No. PEP20-07-01-001, NSDUH Series H-55)*. Rockville, MD: Center for Behavioral Health Statistics, Quality, Substance Abuse, and Mental Health Services Administration, 2020. URL: <https://www.samhsa.gov/data/>.

SAMHSA: Key substance use and mental health indicators in the United States: Results from the 2022 National Survey on Drug Use and Health (HHS Publication No. PEP23-07-01-006, NSDUH Series H-58) **SAMHSA-2023**

SAMHSA. *Key substance use and mental health indicators in the United States: Results from the 2022 National Survey on Drug Use and Health (HHS Publication No. PEP23-07-01-006, NSDUH Series H-58)*. Rockville, MD: Center for Behavioral Health Statistics, Quality, Substance Abuse, and Mental Health Services Administration, 2023. URL: <https://www.samhsa.gov/data/report/2022-nsduh-annual-national-report>.

Victoria Savalei and Yves Rosseel. “Computational options for standard errors and test statistics with incomplete normal and nonnormal data in SEM”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 29.2 (Oct. 2021), pp. 163–181. DOI: [10.1080/10705511.2021.1877548](https://doi.org/10.1080/10705511.2021.1877548).

Abstract: This article provides an overview of different computational options for inference following normal theory maximum likelihood (ML) estimation in structural equation modeling (SEM) with incomplete normal and nonnormal data. Complete data are covered as a special case. These computational options include whether the information matrix is observed or expected, whether the observed information matrix is estimated numerically or using an analytic asymptotic approximation, and whether the information matrix and the outer product matrix of the score vector are evaluated at the saturated or at the structured estimates. A variety of different standard errors and robust test statistics become possible by varying these options. We review the asymptotic properties of these computational variations, and we show how to obtain them using lavaan in R. We hope that this article will encourage methodologists to study the impact of the available computational options on the performance of standard errors and test statistics in SEM.

Joseph L. Schafer. *Analysis of incomplete multivariate data*. Chapman and Hall/CRC, Aug. 1997. ISBN: 9780367803025. DOI: [10.1201/9780367803025](https://doi.org/10.1201/9780367803025).

Abstract: The last two decades have seen enormous developments in statistical methods for incomplete data. The EM algorithm and its extensions, multiple imputation, and Markov Chain Monte Carlo provide a set of flexible and reliable tools from inference in large classes of missing-data problems. Yet, in practical terms, those developments have had surprisingly little impact on the way most data analysts handle missing values on a routine basis. Analysis of Incomplete Multivariate Data helps bridge the gap between theory and practice, making these missing-data tools accessible

to a broad audience. It presents a unified, Bayesian approach to the analysis of incomplete multi-variate data, covering datasets in which the variables are continuous, categorical, or both. The focus is applied, where necessary, to help readers thoroughly understand the statistical properties of those methods, and the behavior of the accompanying algorithms. All techniques are illustrated with real data examples, with extended discussion and practical advice. All of the algorithms described in this book have been implemented by the author for general use in the statistical languages S and S Plus. The software is available free of charge on the Internet.

Schafer et al.: Missing data: Our view of the state of the art **Schafer-Graham-2002**

Joseph L. Schafer and John W. Graham. “Missing data: Our view of the state of the art”. In: *Psychological Methods* 7.2 (2002), pp. 147–177. DOI: [10.1037/1082-989x.7.2.147](https://doi.org/10.1037/1082-989x.7.2.147).

Abstract: Statistical procedures for missing data have vastly improved, yet misconception and unsound practice still abound. The authors frame the missing-data problem, review methods, offer advice, and raise issues that remain unresolved. They clear up common misunderstandings regarding the missing at random (MAR) concept. They summarize the evidence against older procedures and, with few exceptions, discourage their use. They present, in both technical and practical language, 2 general approaches that come highly recommended: maximum likelihood (ML) and Bayesian multiple imputation (MI). Newer developments are discussed, including some for dealing with missing data that are not MAR. Although not yet in the mainstream, these procedures may eventually extend the ML and MI methods that currently represent the state of the art.

Schenker: Better bootstrap confidence intervals: Comment **Schenker-1987**

Nathaniel Schenker. “Better bootstrap confidence intervals: Comment”. In: *Journal of the American Statistical Association* 82.397 (Mar. 1987), p. 192. DOI: [10.2307/2289150](https://doi.org/10.2307/2289150).

Schermerhorn et al.: Developmental family processes and interparental conflict: Patterns of microlevel influences.
Schermerhorn-Chow-Cummings-2010

Alice C. Schermerhorn, Sy-Miin Chow, and E. Mark Cummings. “Developmental family processes and interparental conflict: Patterns of microlevel influences.” In: *Developmental Psychology* 46.4 (2010), pp. 869–885. ISSN: 0012-1649. DOI: [10.1037/a0019662](https://doi.org/10.1037/a0019662).

Abstract: Although there are frequent calls for the study of effects of children on families and mutual influence processes within families, little empirical progress has been made. We address these questions at the level of microprocesses during marital conflict, including children’s influence on marital conflict and parents’ influence on each other. Participants were 111 cohabiting couples with a child (55 male, 56 female) age 8–16 years. Data were drawn from parents’ diary reports of interparental conflict over 15 days and were analyzed with dynamic systems modeling tools. Child emotions and behavior during conflicts were associated with interparental positivity, negativity, and resolution at the end of the same conflicts. For example, children’s agentic behavior was associated with more marital conflict resolution, whereas child negativity was linked with more marital negativity. Regarding parents’ influence on each other, among the findings, husbands’ and wives’ influence on themselves from one conflict to the next was indicated, and total number of conflicts predicted greater influence of wives’ positivity on husbands’ positivity. Contributions of these findings to the understanding of developmental family processes are discussed, including implications for advanced understanding of interrelations between child and adult functioning and development.

Schouten et al.: Generating missing values for simulation purposes: A multivariate amputation procedure
Schouten-Lugtig-Vink-2018

Rianne Margaretha Schouten, Peter Lugtig, and Gerko Vink. “Generating missing values for simulation purposes: A multivariate amputation procedure”. In: *Journal of Statistical Computation and Simulation* 88.15 (July 2018), pp. 2909–2930. DOI: [10.1080/00949655.2018.1491577](https://doi.org/10.1080/00949655.2018.1491577).

Abstract: Missing data form a ubiquitous problem in scientific research, especially since most statistical analyses require complete data. To evaluate the performance of methods dealing with missing data, researchers perform simulation studies. An important aspect of these studies is the generation of missing values in a simulated, complete data set: the amputation procedure. We investigated the methodological validity and statistical nature of both the current amputation practice and a newly developed and implemented multivariate amputation procedure. We found that the current way of practice may not be appropriate for the generation of intuitive and reliable missing data problems. The multivariate amputation procedure, on the other hand, generates reliable amputations and allows for a proper regulation of missing data problems. The procedure has additional features to generate any missing data scenario precisely as intended. Hence, the multivariate amputation procedure is an efficient method to accurately evaluate missing data methodology.

Schulenberg et al.: Monitoring the Future national survey results on drug use, 1975-2020: Volume II, College students and adults ages 19–60

Schulenberg-Patrick-Johnston-et-al-2021

John E. Schulenberg et al. *Monitoring the Future national survey results on drug use, 1975-2020: Volume II, College students and adults ages 19–60*. Ann Arbor, MI: Institute for Social Research, The University of Michigan, 2021.

Selig et al.: Mediation models for longitudinal data in developmental research

Selig-Preacher-2009

James P. Selig and Kristopher J. Preacher. “Mediation models for longitudinal data in developmental research”. In: *Research in Human Development* 6.2-3 (June 2009), pp. 144–164. DOI: [10.1080/15427600902911247](https://doi.org/10.1080/15427600902911247).

Abstract: Mediation models are used to describe the mechanism(s) by which one variable influences another. These models can be useful in developmental research to explicate the relationship between

variables, developmental processes, or combinations of variables and processes. In this article we describe aspects of mediation effects specific to developmental research. We focus on three central issues in longitudinal mediation models: the theory of change for variables in the model, the role of time in the model, and the types of indirect effects in the model. We use these themes as we describe three different models for examining mediation in longitudinal data.

Serlin: Testing for robustness in Monte Carlo studies**Serlin-2000**

Ronald C. Serlin. "Testing for robustness in Monte Carlo studies". In: *Psychological Methods* 5.2 (2000), pp. 230–240. DOI: [10.1037/1082-989x.5.2.230](https://doi.org/10.1037/1082-989x.5.2.230).

Abstract: Monte Carlo studies provide the information needed to help researchers select appropriate analytical procedures under design conditions in which the underlying assumptions of the procedures are not met. In Monte Carlo studies, the 2 errors that one could commit involve (a) concluding that a statistical procedure is robust when it is not or (b) concluding that it is not robust when it is. In previous attempts to apply standard statistical design principles to Monte Carlo studies, the less severe of these errors has been wrongly designated the Type I error. In this article, a method is presented for controlling the appropriate Type I error rate; the determination of the number of iterations required in a Monte Carlo study to achieve desired power is described; and a confidence interval for a test's true Type I error rate is derived. A robustness criterion is also proposed that is a compromise between W. G. Cochran's (1952) and J. V. Bradley's (1978) criteria.

Serlin et al.: Rationality in psychological research: The good-enough principle**Serlin-Lapsley-1985**

Ronald C. Serlin and Daniel K. Lapsley. "Rationality in psychological research: The good-enough principle". In: *American Psychologist* 40.1 (1985), pp. 73–83. DOI: [10.1037/0003-066x.40.1.73](https://doi.org/10.1037/0003-066x.40.1.73).

Abstract: Reexamines methodological and procedural issues raised by P. Meehl (1967; see also PA, Vol 62:5042) that question the rationality of psychological inquiry. Issues concern the asymmetry

in theory testing between psychology and physics and the slow progress observed in psychological research. A good-enough principle is proposed to resolve Meehl's methodological paradox, and a more powerful reconstruction of science developed by I. Lakatos (1978) is suggested to account for the actual practice of psychological researchers.

Shao: Impact of the bootstrap on sample surveys

Shao-2003

Jun Shao. "Impact of the bootstrap on sample surveys". In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994974](https://doi.org/10.1214/ss/1063994974).

Abstract: This article discusses the impact of the bootstrap on sample surveys and introduces some of the main developments of the bootstrap methodology for sample surveys in the last twenty five years.

Shiffman: Ecological momentary assessment (EMA) in studies of substance use

Shiffman-2009

Saul Shiffman. "Ecological momentary assessment (EMA) in studies of substance use". In: *Psychological Assessment* 21.4 (Dec. 2009), pp. 486–497. DOI: [10.1037/a0017074](https://doi.org/10.1037/a0017074).

Abstract: Ecological momentary assessment (EMA) is particularly suitable for studying substance use, because use is episodic and thought to be related to mood and context. This article reviews EMA methods in substance use research, focusing on tobacco and alcohol use and relapse, where EMA has been most applied. Common EMA designs combine event-based reports of substance use with time-based assessments. Approaches to data organization and analysis have been very diverse, particularly regarding their treatment of time. Compliance with signaled assessments is often high. Compliance with recording of substance use appears good but is harder to validate. Treatment applications of EMA are emerging. EMA captures substance use patterns not measured by questionnaires or retrospective data and holds promise for substance use research.

Saul Shiffman, Arthur A. Stone, and Michael R. Hufford. "Ecological momentary assessment". In: *Annual Review of Clinical Psychology* 4.1 (Apr. 2008), pp. 1–32. DOI: [10.1146/annurev.clinpsy.3.022806.091415](https://doi.org/10.1146/annurev.clinpsy.3.022806.091415).

Abstract: Assessment in clinical psychology typically relies on global retrospective self-reports collected at research or clinic visits, which are limited by recall bias and are not well suited to address how behavior changes over time and across contexts. Ecological momentary assessment (EMA) involves repeated sampling of subjects' current behaviors and experiences in real time, in subjects' natural environments. EMA aims to minimize recall bias, maximize ecological validity, and allow study of microprocesses that influence behavior in real-world contexts. EMA studies assess particular events in subjects' lives or assess subjects at periodic intervals, often by random time sampling, using technologies ranging from written diaries and telephones to electronic diaries and physiological sensors. We discuss the rationale for EMA, EMA designs, methodological and practical issues, and comparisons of EMA and recall data. EMA holds unique promise to advance the science and practice of clinical psychology by shedding light on the dynamics of behavior in real-world settings.

Shrout: Commentary: Mediation analysis, causal process, and cross-sectional data

Shrout-2011

Patrick E. Shrout. "Commentary: Mediation analysis, causal process, and cross-sectional data". In: *Multivariate Behavioral Research* 46.5 (Sept. 2011), pp. 852–860. DOI: [10.1080/00273171.2011.606718](https://doi.org/10.1080/00273171.2011.606718).

Abstract: Maxwell, Cole, and Mitchell (2011) extended the work of Maxwell and Cole (2007), which raised important questions about whether mediation analyses based on cross-sectional data can shed light on longitudinal mediation process. The latest article considers longitudinal processes that can only be partially explained by an intervening variable, and Maxwell et al. showed that

the same general conclusions are obtained, namely that analyses of cross-sectional data will not reveal the longitudinal mediation process. While applauding the advances of the target article, this comment encourages the detailed exploration of alternate causal models in psychology beyond the autoregressive model considered by Maxwell et al. When inferences based on cross-sectional analyses are compared to alternate models, different patterns of bias are likely to be observed. I illustrate how different models of the causal process can be derived using examples from research on psychopathology.

Shrout et al.: Mediation in experimental and nonexperimental studies: New procedures and recommendations
Shrout-Bolger-2002

Patrick E. Shrout and Niall Bolger. “Mediation in experimental and nonexperimental studies: New procedures and recommendations”. In: *Psychological Methods* 7.4 (2002), pp. 422–445. DOI: [10.1037/1082-989X.7.4.422](https://doi.org/10.1037/1082-989X.7.4.422).

Abstract: Mediation is said to occur when a causal effect of some variable X on an outcome Y is explained by some intervening variable M . The authors recommend that with small to moderate samples, bootstrap methods (B. Efron & R. Tibshirani, 1993) be used to assess mediation. Bootstrap tests are powerful because they detect that the sampling distribution of the mediated effect is skewed away from 0. They argue that R. M. Baron and D. A. Kenny’s (1986) recommendation of first testing the $X \rightarrow Y$ association for statistical significance should not be a requirement when there is a priori belief that the effect size is small or suppression is a possibility. Empirical examples and computer setups for bootstrap analyses are provided.

Shumway et al.: Time series analysis and its applications: With R examples
Shumway-Stoffer-2017

Robert H. Shumway and David S. Stoffer. *Time series analysis and its applications: With R exam-*

ples. Springer International Publishing, 2017. ISBN: 978-3-319-52452-8. DOI: [10.1007/978-3-319-52452-8](https://doi.org/10.1007/978-3-319-52452-8).

Abstract: The fourth edition of this popular graduate textbook, like its predecessors, presents a balanced and comprehensive treatment of both time and frequency domain methods with accompanying theory. Numerous examples using nontrivial data illustrate solutions to problems such as discovering natural and anthropogenic climate change, evaluating pain perception experiments using functional magnetic resonance imaging, and monitoring a nuclear test ban treaty. The book is designed as a textbook for graduate level students in the physical, biological, and social sciences and as a graduate level text in statistics. Some parts may also serve as an undergraduate introductory course. Theory and methodology are separated to allow presentations on different levels. In addition to coverage of classical methods of time series regression, ARIMA models, spectral analysis and state-space models, the text includes modern developments including categorical time series analysis, multivariate spectral methods, long memory series, nonlinear models, resampling techniques, GARCH models, ARMAX models, stochastic volatility, wavelets, and Markov chain Monte Carlo integration methods. This edition includes R code for each numerical example in addition to Appendix R, which provides a reference for the data sets and R scripts used in the text in addition to a tutorial on basic R commands and R time series. An additional file is available on the book's website for download, making all the data sets and scripts easy to load into R.

Library: QA280.

Smith et al.: From ecological momentary assessment (EMA) to ecological momentary intervention (EMI): Past and future directions for ambulatory assessment and interventions in eating disorders **Smith-Juarascio-2019**

Kathryn E. Smith and Adrienne Juarascio. "From ecological momentary assessment (EMA) to ecological momentary intervention (EMI): Past and future directions for ambulatory assessment

and interventions in eating disorders”. In: *Current Psychiatry Reports* 21.7 (June 2019). DOI: [10.1007/s11920-019-1046-8](https://doi.org/10.1007/s11920-019-1046-8).

Abstract: Purpose of Review: Ambulatory assessment methods, including ecological momentary assessment (EMA), have often been used in eating disorders (EDs) to assess the type, frequency, and temporal sequencing of ED symptoms occurring in naturalistic environments. Relatedly, growing research in EDs has explored the utility of ecological momentary interventions (EMIs) to target ED symptoms. The aims of the present review were to (1) synthesize recent literature pertaining to ambulatory assessment/EMA and EMI in EDs, and (2) identify relevant limitations and future directions in these domains. Recent Findings: With respect to ambulatory assessment and EMA, there has been substantial growth in the expansion of constructs assessed with EMA, the exploration of state- vs. trait-level processes, integration of objective and passive assessment approaches, and consideration of methodological issues. The EMI literature in EDs also continues to grow, though most of the recent research focuses on mobile health (mHealth) technologies with relatively minimal EMI components that adapt to momentary contextual information. Summary: Despite these encouraging advances, there remain several promising areas of ambulatory assessment research and clinical applications in EDs going forward. These include integration of passive data collection, use of EMA in treatment evaluation and design, evaluation of dynamic system processes, inclusion of diverse samples, and development and evaluation of adaptive, tailored EMIs such as just-in-time adaptive interventions. While much remains to be learned in each of these domains, the continual growth in mobile technology has potential to facilitate and refine our understanding of the nature of ED psychopathology and ultimately improve intervention approaches.

Sobel: Asymptotic confidence intervals for indirect effects in structural equation models
Sobel-1982

Michael E. Sobel. “Asymptotic confidence intervals for indirect effects in structural equation models”. In: *Sociological Methodology* 13 (1982), p. 290. DOI: [10.2307/270723](https://doi.org/10.2307/270723).

Sobel: Some new results on indirect effects and their standard errors in covariance structure models **Sobel-1986**

Michael E. Sobel. “Some new results on indirect effects and their standard errors in covariance structure models”. In: *Sociological Methodology* 16 (1986), p. 159. DOI: [10.2307/270922](https://doi.org/10.2307/270922).

Sobel: Direct and indirect effects in linear structural equation models **Sobel-1987**

Michael E. Sobel. “Direct and indirect effects in linear structural equation models”. In: *Sociological Methods & Research* 16.1 (Aug. 1987), pp. 155–176. DOI: [10.1177/0049124187016001006](https://doi.org/10.1177/0049124187016001006).

Abstract: This article discusses total indirect effects in linear structural equation models. First, I define these effects. Second, I show how the delta method may be used to obtain the standard errors of the sample estimates of these effects and test hypotheses about the magnitudes of the indirect effects. To keep matters simple, I focus throughout on a particularly simple linear structural equation system; for a treatment of the general case, see Sobel (1986). To illustrate the ideas and results, a detailed example is presented.

P. S. Soltis et al.: Applying the Bootstrap in Phylogeny Reconstruction

Soltis-Soltis-2003

Pamela S. Soltis and Douglas E. Soltis. “Applying the Bootstrap in Phylogeny Reconstruction”. In: *Statistical Science* 18.2 (May 2003). DOI: [10.1214/ss/1063994980](https://doi.org/10.1214/ss/1063994980).

Abstract: With the increasing emphasis in biology on reconstruction of phylogenetic trees, questions have arisen as to how confident one should be in a given phylogenetic tree and how support for phylogenetic trees should be measured. Felsenstein suggested that bootstrapping be applied across characters of a taxon-by-character data matrix to produce replicate “bootstrap data sets,” each of which is then analyzed phylogenetically, with a consensus tree constructed to summarize the results of all replicates. The proportion of trees/replicates in which a grouping is recovered is presented as

a measure of support for that group. Bootstrapping has become a common feature of phylogenetic analysis. However, the interpretation of bootstrap values remains open to discussion, and phylogeneticists have used these values in multiple ways. The usefulness of phylogenetic bootstrapping is potentially limited by a number of features, such as the size of the data matrix and the underlying assumptions of the phylogeny reconstruction program. Recent studies have explored the application of bootstrapping to large data sets and the relative performance of bootstrapping and jackknifing.

Stoffer et al.: Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter **Stoffer-Wall-1991**

David S. Stoffer and Kent D. Wall. “Bootstrapping state-space models: Gaussian maximum likelihood estimation and the Kalman filter”. In: *Journal of the American Statistical Association* 86.416 (Dec. 1991), pp. 1024–1033. DOI: [10.1080/01621459.1991.10475148](https://doi.org/10.1080/01621459.1991.10475148).

Abstract: The bootstrap is proposed as a method for assessing the precision of Gaussian maximum likelihood estimates of the parameters of linear state-space models. Our results also apply to autoregressive moving average models, since they are a special case of state-space models. It is shown that for a time-invariant, stable system, the bootstrap applied to the innovations yields asymptotically consistent standard errors. To investigate the performance of the bootstrap for finite sample lengths, simulation results are presented for a two-state model with 50 and 100 observations; two cases are investigated, one with real characteristic roots and one with complex characteristic roots. The bootstrap is then applied to two real data sets, one used in a test for efficient capital markets and one used to develop an autoregressive integrated moving average model for quarterly earnings data. We find the bootstrap to be of definite value over the conventional asymptotics.

Swift: Transdermal alcohol measurement for estimation of blood alcohol concentration

Swift-2000

Robert Swift. "Transdermal alcohol measurement for estimation of blood alcohol concentration". In: *Alcoholism: Clinical and Experimental Research* 24.4 (Apr. 2000), pp. 422–423. ISSN: 1530-0277. DOI: [10.1111/j.1530-0277.2000.tb02006.x](https://doi.org/10.1111/j.1530-0277.2000.tb02006.x).

Taylor et al.: Four applications of permutation methods to testing a single-mediator model

Taylor-MacKinnon-2012

Aaron B. Taylor and David P. MacKinnon. "Four applications of permutation methods to testing a single-mediator model". In: *Behavior Research Methods* 44.3 (Feb. 2012), pp. 806–844. DOI: [10.3758/s13428-011-0181-x](https://doi.org/10.3758/s13428-011-0181-x).

Abstract: Four applications of permutation tests to the single-mediator model are described and evaluated in this study. Permutation tests work by rearranging data in many possible ways in order to estimate the sampling distribution for the test statistic. The four applications to mediation evaluated here are the permutation test of ab, the permutation joint significance test, and the noniterative and iterative permutation confidence intervals for ab. A Monte Carlo simulation study was used to compare these four tests with the four best available tests for mediation found in previous research: the joint significance test, the distribution of the product test, and the percentile and bias-corrected bootstrap tests. We compared the different methods on Type I error, power, and confidence interval coverage. The noniterative permutation confidence interval for ab was the best performer among the new methods. It successfully controlled Type I error, had power nearly as good as the most powerful existing methods, and had better coverage than any existing method. The iterative permutation confidence interval for ab had lower power than do some existing methods, but it performed better than any other method in terms of coverage. The permutation confidence interval methods are recommended when estimating a confidence interval is a primary concern. SPSS and SAS macros that estimate these confidence intervals are provided.

Aaron B. Taylor, David P. MacKinnon, and Jenn-Yun Tein. “Tests of the three-path mediated effect”. In: *Organizational Research Methods* 11.2 (July 2007), pp. 241–269. DOI: [10.1177/1094428107300344](https://doi.org/10.1177/1094428107300344).

Abstract: In a three-path mediational model, two mediators intervene in a series between an independent and a dependent variable. Methods of testing for mediation in such a model are generalized from the more often used single-mediator model. Six such methods are introduced and compared in a Monte Carlo study in terms of their Type I error, power, and coverage. Based on its results, the joint significance test is preferred when only a hypothesis test is of interest. The percentile bootstrap and bias-corrected bootstrap are preferred when a confidence interval on the mediated effect is desired, with the latter having more power but also slightly inflated Type I error in some conditions.

Davood Tofighi and Ken Kelley. “Indirect effects in sequential mediation models: Evaluating methods for hypothesis testing and confidence interval formation”. In: *Multivariate Behavioral Research* 55.2 (June 2019), pp. 188–210. DOI: [10.1080/00273171.2019.1618545](https://doi.org/10.1080/00273171.2019.1618545).

Abstract: Complex mediation models, such as a two-mediator sequential model, have become more prevalent in the literature. To test an indirect effect in a two-mediator model, we conducted a large-scale Monte Carlo simulation study of the Type I error, statistical power, and confidence interval coverage rates of 10 frequentist and Bayesian confidence/credible intervals (CIs) for normally and nonnormally distributed data. The simulation included never-studied methods and conditions (e.g., Bayesian CI with flat and weakly informative prior methods, two model-based bootstrap methods, and two nonnormality conditions) as well as understudied methods (e.g., profile-likelihood, Monte Carlo with maximum likelihood standard error [MC-ML] and robust standard error [MC-Robust]).

The popular BC bootstrap showed inflated Type I error rates and CI under-coverage. We recommend different methods depending on the purpose of the analysis. For testing the null hypothesis of no mediation, we recommend MC-ML, profile-likelihood, and two Bayesian methods. To report a CI, if data has a multivariate normal distribution, we recommend MC-ML, profile-likelihood, and the two Bayesian methods; otherwise, for multivariate nonnormal data we recommend the percentile bootstrap. We argue that the best method for testing hypotheses is not necessarily the best method for CI construction, which is consistent with the findings we present.

Tofighi et al.: Improved inference in mediation analysis: Introducing the model-based constrained optimization procedure **Tofighi-Kelley-2020**

Davood Tofighi and Ken Kelley. “Improved inference in mediation analysis: Introducing the model-based constrained optimization procedure”. In: *Psychological Methods* 25 (2020), pp. 496–515. DOI: [10.1037/met0000259](https://doi.org/10.1037/met0000259).

Abstract: Mediation analysis is an important approach for investigating causal pathways. One approach used in mediation analysis is the test of an indirect effect, which seeks to measure how the effect of an independent variable impacts an outcome variable through one or more mediators. However, in many situations the proposed tests of indirect effects, including popular confidence interval-based methods, tend to produce poor Type I error rates when mediation does not occur and, more generally, only allow dichotomous decisions of “not significant” or “significant” with regards to the statistical conclusion. To remedy these issues, we propose a new method, a likelihood ratio test (LRT), that uses non-linear constraints in what we term the model-based constrained optimization (MBCO) procedure. The MBCO procedure (a) offers a more robust Type I error rate than existing methods; (b) provides a p-value, which serves as a continuous measure of compatibility of data with the hypothesized null model (not just a dichotomous reject or fail-to-reject decision rule); (c) allows simple and complex hypotheses about mediation (i.e., one or more mediators; different mediational pathways), and (d) allows the mediation model to use observed or latent variables. The MBCO procedure is based on a structural equation modeling framework (even if

latent variables are not specified) with specialized fitting routines, namely with the use of non-linear constraints. We advocate using the MBCO procedure to test hypotheses about an indirect effect in addition to reporting a confidence interval to capture uncertainty about the indirect effect because this combination transcends existing methods.

Tofighi et al.: Monte Carlo confidence intervals for complex functions of indirect effects

Tofighi-MacKinnon-2015

Davood Tofighi and David P. MacKinnon. “Monte Carlo confidence intervals for complex functions of indirect effects”. In: *Structural Equation Modeling: A Multidisciplinary Journal* 23.2 (Aug. 2015), pp. 194–205. DOI: [10.1080/10705511.2015.1057284](https://doi.org/10.1080/10705511.2015.1057284).

Abstract: One challenge in mediation analysis is to generate a confidence interval (CI) with high coverage and power that maintains a nominal significance level for any well-defined function of indirect and direct effects in the general context of structural equation modeling (SEM). This study discusses a proposed Monte Carlo extension that finds the CIs for any well-defined function of the coefficients of SEM such as the product of k coefficients and the ratio of the contrasts of indirect effects, using the Monte Carlo method. Finally, we conduct a small-scale simulation study to compare CIs produced by the Monte Carlo, nonparametric bootstrap, and asymptotic-delta methods. Based on our simulation study, we recommend researchers use the Monte Carlo method to test a complex function of indirect effects.

Uhlenbeck et al.: On the Theory of the Brownian Motion Uhlenbeck-Ornstein-1930

G. E. Uhlenbeck and L. S. Ornstein. “On the Theory of the Brownian Motion”. In: *Physical Review* 36.5 (Sept. 1930), pp. 823–841. DOI: [10.1103/physrev.36.823](https://doi.org/10.1103/physrev.36.823).

Abstract: With a method first indicated by Ornstein the mean values of all the powers of the velocity u and the displacement s of a free particle in Brownian motion are calculated. It is shown that $u - u_0 \exp(-\beta t)$ and $s - u_0 \beta [1 - \exp(-\beta t)]$ where u_0 is the initial velocity and β the friction

coefficient divided by the mass of the particle, follow the normal Gaussian distribution law. For s this gives the exact frequency distribution corresponding to the exact formula for s^2 of Ornstein and Fürth. Discussion is given of the connection with the Fokker-Planck partial differential equation. By the same method exact expressions are obtained for the square of the deviation of a harmonically bound particle in Brownian motion as a function of the time and the initial deviation. Here the periodic, aperiodic and overdamped cases have to be treated separately. In the last case, when β is much larger than the frequency and for values of $t \gg \beta^{-1}$, the formula takes the form of that previously given by Smoluchowski.

van Buuren: Flexible imputation of missing data

vanBuuren-2018

Stef van Buuren. *Flexible imputation of missing data*. 2nd ed. Chapman and Hall/CRC, July 2018. ISBN: 9780429492259. DOI: [10.1201/9780429492259](https://doi.org/10.1201/9780429492259).

Abstract: Missing data pose challenges to real-life data analysis. Simple ad-hoc fixes, like deletion or mean imputation, only work under highly restrictive conditions, which are often not met in practice. Multiple imputation replaces each missing value by multiple plausible values. The variability between these replacements reflects our ignorance of the true (but missing) value. Each of the completed data set is then analyzed by standard methods, and the results are pooled to obtain unbiased estimates with correct confidence intervals. Multiple imputation is a general approach that also inspires novel solutions to old problems by reformulating the task at hand as a missing-data problem. This is the second edition of a popular book on multiple imputation, focused on explaining the application of methods through detailed worked examples using the MICE package as developed by the author. This new edition incorporates the recent developments in this fast-moving field. This class-tested book avoids mathematical and technical details as much as possible: formulas are accompanied by verbal statements that explain the formula in accessible terms. The book sharpens the reader's intuition on how to think about missing data, and provides all the tools needed to execute a well-grounded quantitative analysis in the presence of missing data.

Library: QA278.

van Buuren et al.: Fully conditional specification in multivariate imputation

vanBuuren-Brand-GroothuisOudshoorn-et al-2006

Stef van Buuren, J. P. L. Brand, et al. “Fully conditional specification in multivariate imputation”. In: *Journal of Statistical Computation and Simulation* 76.12 (Dec. 2006), pp. 1049–1064. DOI: [10.1080/10629360600810434](https://doi.org/10.1080/10629360600810434).

Abstract: The use of the Gibbs sampler with fully conditionally specified models, where the distribution of each variable given the other variables is the starting point, has become a popular method to create imputations in incomplete multivariate data. The theoretical weakness of this approach is that the specified conditional densities can be incompatible, and therefore the stationary distribution to which the Gibbs sampler attempts to converge may not exist. This study investigates practical consequences of this problem by means of simulation. Missing data are created under four different missing data mechanisms. Attention is given to the statistical behavior under compatible and incompatible models. The results indicate that multiple imputation produces essentially unbiased estimates with appropriate coverage in the simple cases investigated, even for the incompatible models. Of particular interest is that these results were produced using only five Gibbs iterations starting from a simple draw from observed marginal distributions. It thus appears that, despite the theoretical weaknesses, the actual performance of conditional model specification for multivariate imputation can be quite good, and therefore deserves further study.

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](https://doi.org/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.

van Montfort et al.: Continuous time modeling in the behavioral and related sciences

Montfort-Oud-Voelkle-2018

Kees van Montfort, Johan H. L. Oud, and Manuel C. Voelkle, eds. *Continuous time modeling in the behavioral and related sciences*. Springer International Publishing, 2018. DOI: [10.1007/978-3-319-77219-6](https://doi.org/10.1007/978-3-319-77219-6).

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](https://doi.org/10.1007/978-0-387-21706-2).

D. J. Venzon and S. H. Moolgavkar. “A method for computing profile-likelihood-based confidence intervals”. In: *Applied Statistics* 37.1 (1988), p. 87. DOI: [10.2307/2347496](https://doi.org/10.2307/2347496).

Abstract: The method of constructing confidence regions based on the generalised likelihood ratio statistic is well known for parameter vectors. A similar construction of a confidence interval for a single entry of a vector can be implemented by repeatedly maximising over the other parameters. We present an algorithm for finding these confidence interval endpoints that requires less computation. It employs a modified Newton-Raphson iteration to solve a system of equations that defines the endpoints.

Manuel C. Voelkle and Johan H. L. Oud. “Continuous time modelling with individually varying time intervals for oscillating and non-oscillating processes”. In: *British Journal of Mathematical and Statistical Psychology* 66.1 (Mar. 2012), pp. 103–126. DOI: [10.1111/j.2044-8317.2012.02043.x](https://doi.org/10.1111/j.2044-8317.2012.02043.x).

Abstract: When designing longitudinal studies, researchers often aim at equal intervals. In practice, however, this goal is hardly ever met, with different time intervals between assessment waves and different time intervals between individuals being more the rule than the exception. One of the reasons for the introduction of continuous time models by means of structural equation modelling has been to deal with irregularly spaced assessment waves (e.g., Oud & Delsing, 2010). In the present paper we extend the approach to individually varying time intervals for oscillating and non-oscillating processes. In addition, we show not only that equal intervals are unnecessary but also that it can be advantageous to use unequal sampling intervals, in particular when the sampling rate is low. Two examples are provided to support our arguments. In the first example we compare a continuous

time model of a bivariate coupled process with varying time intervals to a standard discrete time model to illustrate the importance of accounting for the exact time intervals. In the second example the effect of different sampling intervals on estimating a damped linear oscillator is investigated by means of a Monte Carlo simulation. We conclude that it is important to account for individually varying time intervals, and encourage researchers to conceive of longitudinal studies with different time intervals within and between individuals as an opportunity rather than a problem.

Voelkle et al.: An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia **Voelkle-Oud-Davidov-et-al-2012**

Manuel C. Voelkle, Johan H. L. Oud, et al. “An SEM approach to continuous time modeling of panel data: Relating authoritarianism and anomia”. In: *Psychological Methods* 17.2 (2012), pp. 176–192. DOI: [10.1037/a0027543](https://doi.org/10.1037/a0027543).

Abstract: Panel studies, in which the same subjects are repeatedly observed at multiple time points, are among the most popular longitudinal designs in psychology. Meanwhile, there exists a wide range of different methods to analyze such data, with autoregressive and cross-lagged models being 2 of the most well known representatives. Unfortunately, in these models time is only considered implicitly, making it difficult to account for unequally spaced measurement occasions or to compare parameter estimates across studies that are based on different time intervals. Stochastic differential equations offer a solution to this problem by relating the discrete time model to its underlying model in continuous time. It is the goal of the present article to introduce this approach to a broader psychological audience. A step-by-step review of the relationship between discrete and continuous time modeling is provided, and we demonstrate how continuous time parameters can be obtained via structural equation modeling. An empirical example on the relationship between authoritarianism and anomia is used to illustrate the approach.

Niels G. Waller. *fungible: Psychometric functions from the Waller Lab*. The R Foundation, 2022.

URL: <https://CRAN.R-project.org/package=fungible>.

L. Wang et al.: Investigating the impact of the time interval selection on autoregressive mediation modeling: Result interpretations, effect reporting, and temporal designs

Wang-Zhang-2020

Lijuan Wang and Qian Zhang. “Investigating the impact of the time interval selection on autoregressive mediation modeling: Result interpretations, effect reporting, and temporal designs”. In: *Psychological Methods* 25.3 (June 2020), pp. 271–291. DOI: [10.1037/met0000235](https://doi.org/10.1037/met0000235).

Abstract: This study investigates the impact of the time interval (the time passed between 2 consecutive measurements) selection on autoregressive mediation modeling (AMM). For a widely used autoregressive mediation model, via analytical derivations, we explained why and how the conventionally reported time-specific coefficient estimates (e.g., $\hat{a}\hat{b}$ and \hat{c}') and inference results in AMM provide limited information and can arrive in even misleading conclusions about direct and indirect effects over time. Furthermore, under the stationarity assumption, we proposed an approach to calculate the overall direct and indirect effect estimates over time and the time lag lengths at which they reach maxima, using AMM results. The derivation results revealed that the overall direct and indirect effect curves are asymptotically invariant to the time interval selection, under stationarity. With finite samples and thus sampling errors and potential computing problems, however, our simulation results revealed that the overall indirect effect curves were better recovered when the time interval is selected to be closer to half of the time lag length at which the overall indirect effect reaches its maximum. An R function and an R Shiny app were developed to obtain the overall direct and indirect effect curves over time and facilitate the time interval selection using AMM results. Our findings provide another look at the connections between AMM and continuous time mediation modeling and the connections are discussed.

White: A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity **White-1980**

Halbert White. “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity”. In: *Econometrica* 48.4 (May 1980), pp. 817–838. DOI: [10.2307/1912934](https://doi.org/10.2307/1912934).

Abstract: This paper presents a parameter covariance matrix estimator which is consistent even when the disturbances of a linear regression model are heteroskedastic. This estimator does not depend on a formal model of the structure of the heteroskedasticity. By comparing the elements of the new estimator to those of the usual covariance estimator, one obtains a direct test for heteroskedasticity, since in the absence of heteroskedasticity, the two estimators will be approximately equal, but will generally diverge otherwise. The test has an appealing least squares interpretation.

Wright: On the nature of size factors **Wright-1918**

Sewall Wright. “On the nature of size factors”. In: *Genetics* 3.4 (July 1918), pp. 367–374. DOI: [10.1093/genetics/3.4.367](https://doi.org/10.1093/genetics/3.4.367).

Wright: The method of path coefficients **Wright-1934**

Sewall Wright. “The method of path coefficients”. In: *The Annals of Mathematical Statistics* 5.3 (Sept. 1934), pp. 161–215. DOI: [10.1214/aoms/1177732676](https://doi.org/10.1214/aoms/1177732676).

Wu et al.: A new procedure to test mediation with missing data through nonparametric bootstrapping and multiple imputation **Wu-Jia-2013**

Wei Wu and Fan Jia. “A new procedure to test mediation with missing data through nonparametric bootstrapping and multiple imputation”. In: *Multivariate Behavioral Research* 48.5 (Sept. 2013), pp. 663–691. DOI: [10.1080/00273171.2013.816235](https://doi.org/10.1080/00273171.2013.816235).

Abstract: This article proposes a new procedure to test mediation with the presence of missing data by combining nonparametric bootstrapping with multiple imputation (MI). This procedure performs MI first and then bootstrapping for each imputed data set. The proposed procedure is more computationally efficient than the procedure that performs bootstrapping first and then MI for each bootstrap sample. The validity of the procedure is evaluated using a simulation study under different sample size, missing data mechanism, missing data proportion, and shape of distribution conditions. The result suggests that the proposed procedure performs comparably to the procedure that combines bootstrapping with full information maximum likelihood under most conditions. However, caution needs to be taken when using this procedure to handle missing not-at-random or nonnormal data.

Yuan et al.: Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data **Yuan-Bentler-2000**

Ke-Hai Yuan and Peter M. Bentler. "Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data". In: *Sociological Methodology* 30.1 (Aug. 2000), pp. 165–200. DOI: [10.1111/0081-1750.00078](https://doi.org/10.1111/0081-1750.00078).

Abstract: Survey and longitudinal studies in the social and behavioral sciences generally contain missing data. Mean and covariance structure models play an important role in analyzing such data. Two promising methods for dealing with missing data are a direct maximum-likelihood and a two-stage approach based on the unstructured mean and covariance estimates obtained by the EM-algorithm. Typical assumptions under these two methods are ignorable nonresponse and normality of data. However, data sets in social and behavioral sciences are seldom normal, and experience with these procedures indicates that normal theory based methods for nonnormal data very often lead to incorrect model evaluations. By dropping the normal distribution assumption, we develop more accurate procedures for model inference. Based on the theory of generalized estimating equations, a way to obtain consistent standard errors of the two-stage estimates is given. The asymptotic efficiencies of different estimators are compared under various assumptions. We also propose a min-

imum chi-square approach and show that the estimator obtained by this approach is asymptotically at least as efficient as the two likelihood-based estimators for either normal or nonnormal data. The major contribution of this paper is that for each estimator, we give a test statistic whose asymptotic distribution is chisquare as long as the underlying sampling distribution enjoys finite fourth-order moments. We also give a characterization for each of the two likelihood ratio test statistics when the underlying distribution is nonnormal. Modifications to the likelihood ratio statistics are also given. Our working assumption is that the missing data mechanism is missing completely at random. Examples and Monte Carlo studies indicate that, for commonly encountered nonnormal distributions, the procedures developed in this paper are quite reliable even for samples with missing data that are missing at random.

Yuan et al.: Biases and standard errors of standardized regression coefficients

Yuan-Chan-2011

Ke-Hai Yuan and Wai Chan. “Biases and standard errors of standardized regression coefficients”. In: *Psychometrika* 76.4 (Aug. 2011), pp. 670–690. DOI: [10.1007/s11336-011-9224-6](https://doi.org/10.1007/s11336-011-9224-6).

Abstract: The paper obtains consistent standard errors (SE) and biases of order $O(1/n)$ for the sample standardized regression coefficients with both random and given predictors. Analytical results indicate that the formulas for SEs given in popular text books are consistent only when the population value of the regression coefficient is zero. The sample standardized regression coefficients are also biased in general, although it should not be a concern in practice when the sample size is not too small. Monte Carlo results imply that, for both standardized and unstandardized sample regression coefficients, SE estimates based on asymptotics tend to under-predict the empirical ones at smaller sample sizes.

**Yzerbyt et al.: New recommendations for testing indirect effects in mediational models:
The need to report and test component paths** **Yzerbyt-Muller-Batailler-et-al-2018**

Vincent Yzerbyt et al. “New recommendations for testing indirect effects in mediational models: The need to report and test component paths”. In: *Journal of Personality and Social Psychology* 115.6 (Dec. 2018), pp. 929–943. DOI: [10.1037/pspa0000132](https://doi.org/10.1037/pspa0000132).

Abstract: In light of current concerns with replicability and reporting false-positive effects in psychology, we examine Type I errors and power associated with 2 distinct approaches for the assessment of mediation, namely the component approach (testing individual parameter estimates in the model) and the index approach (testing a single mediational index). We conduct simulations that examine both approaches and show that the most commonly used tests under the index approach risk inflated Type I errors compared with the joint-significance test inspired by the component approach. We argue that the tendency to report only a single mediational index is worrisome for this reason and also because it is often accompanied by a failure to critically examine the individual causal paths underlying the mediational model. We recommend testing individual components of the indirect effect to argue for the presence of an indirect effect and then using other recommended procedures to calculate the size of that effect. Beyond simple mediation, we show that our conclusions also apply in cases of within-participant mediation and moderated mediation. We also provide a new R-package that allows for an easy implementation of our recommendations.

Zeileis: Econometric computing with HC and HAC covariance matrix estimators

Zeileis-2004

Achim Zeileis. “Econometric computing with HC and HAC covariance matrix estimators”. In: *Journal of Statistical Software* 11.10 (2004). DOI: [10.18637/jss.v011.i10](https://doi.org/10.18637/jss.v011.i10).

Abstract: Data described by econometric models typically contains autocorrelation and/or heteroskedasticity of unknown form and for inference in such models it is essential to use covariance matrix estimators that can consistently estimate the covariance of the model parameters. Hence,

suitable heteroskedasticity consistent (HC) and heteroskedasticity and autocorrelation consistent (HAC) estimators have been receiving attention in the econometric literature over the last 20 years. To apply these estimators in practice, an implementation is needed that preferably translates the conceptual properties of the underlying theoretical frameworks into computational tools. In this paper, such an implementation in the package `sandwich` in the R system for statistical computing is described and it is shown how the suggested functions provide reusable components that build on readily existing functionality and how they can be integrated easily into new inferential procedures or applications. The toolbox contained in `sandwich` is extremely flexible and comprehensive, including specific functions for the most important HC and HAC estimators from the econometric literature. Several real-world data sets are used to illustrate how the functionality can be integrated into applications.

Zeileis: Object-oriented computation of sandwich estimators

Zeileis-2006

Achim Zeileis. “Object-oriented computation of sandwich estimators”. In: *Journal of Statistical Software* 16.9 (Aug. 2006). DOI: [10.18637/jss.v016.i09](https://doi.org/10.18637/jss.v016.i09).

Abstract: Sandwich covariance matrix estimators are a popular tool in applied regression modeling for performing inference that is robust to certain types of model misspecification. Suitable implementations are available in the R system for statistical computing for certain model fitting functions only (in particular `lm()`), but not for other standard regression functions, such as `glm()`, `nls()`, or `survreg()`. Therefore, conceptual tools and their translation to computational tools in the package `sandwich` are discussed, enabling the computation of sandwich estimators in general parametric models. Object orientation can be achieved by providing a few extractor functions’ most importantly for the empirical estimating functions’ from which various types of sandwich estimators can be computed.

Achim Zeileis, Susanne Köll, and Nathaniel Graham. “Various versatile variances: An object-oriented implementation of clustered covariances in R”. In: *Journal of Statistical Software* 95.1 (Oct. 2020). DOI: [10.18637/jss.v095.i01](https://doi.org/10.18637/jss.v095.i01).

Abstract: Clustered covariances or clustered standard errors are very widely used to account for correlated or clustered data, especially in economics, political sciences, and other social sciences. They are employed to adjust the inference following estimation of a standard least-squares regression or generalized linear model estimated by maximum likelihood. Although many publications just refer to “the” clustered standard errors, there is a surprisingly wide variety of clustered covariances, particularly due to different flavors of bias corrections. Furthermore, while the linear regression model is certainly the most important application case, the same strategies can be employed in more general models (e.g., for zero-inflated, censored, or limited responses). In R, functions for covariances in clustered or panel models have been somewhat scattered or available only for certain modeling functions, notably the (generalized) linear regression model. In contrast, an object-oriented approach to “robust” covariance matrix estimation - applicable beyond `lm()` and `glm()` - is available in the `sandwich` package but has been limited to the case of cross-section or time series data. Starting with `sandwich` 2.4.0, this shortcoming has been corrected: Based on methods for two generic functions (`estfun()` and `bread()`), clustered and panel covariances are provided in `vcovCL()`, `vcovPL()`, and `vcovPC()`. Moreover, clustered bootstrap covariances are provided in `vcovBS()`, using `modelupdate()` on bootstrap samples. These are directly applicable to models from packages including `MASS`, `pscl`, `countreg`, and `betareg`, among many others. Some empirical illustrations are provided as well as an assessment of the methods’ performance in a simulation study.

Zhiyong Zhang and Lijuan Wang. “Methods for mediation analysis with missing data”. In: *Psychometrika* 78.1 (Dec. 2012), pp. 154–184. DOI: [10.1007/s11336-012-9301-5](https://doi.org/10.1007/s11336-012-9301-5).

Abstract: Despite wide applications of both mediation models and missing data techniques, formal discussion of mediation analysis with missing data is still rare. We introduce and compare four approaches to dealing with missing data in mediation analysis including listwise deletion, pairwise deletion, multiple imputation (MI), and a two-stage maximum likelihood (TS-ML) method. An R package *bmem* is developed to implement the four methods for mediation analysis with missing data in the structural equation modeling framework, and two real examples are used to illustrate the application of the four methods. The four methods are evaluated and compared under MCAR, MAR, and MNAR missing data mechanisms through simulation studies. Both MI and TS-ML perform well for MCAR and MAR data regardless of the inclusion of auxiliary variables and for AV-MNAR data with auxiliary variables. Although listwise deletion and pairwise deletion have low power and large parameter estimation bias in many studied conditions, they may provide useful information for exploring missing mechanisms.

Zhiyong Zhang, Lijuan Wang, and Xin Tong. “Mediation analysis with missing data through multiple imputation and bootstrap”. In: *Quantitative Psychology Research*. Springer International Publishing, 2015, pp. 341–355. DOI: [10.1007/978-3-319-19977-1_24](https://doi.org/10.1007/978-3-319-19977-1_24).