Journal of the American Statistical Association

Book Reviews

Published online: 02 Oct 2014.

PLEASE SCROLL DOWN FOR ARTICLE

Book Reviews

Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach.

Andrew F. HAYES. New York: The Guilford Press, 2013, xvii + 507 pp., $65.00 (H), ISBN: 978-1-60918-230-4.

While the mere existence of causal effect is often of great interest, scientists are also interested in the underlying causal mechanisms: How and when does a certain causal agent affect the outcome of interest? *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach* deals with this question. This book elegantly presents both the basic and advanced issues of mediation and moderation analysis, and I believe that it will be beneficial for graduate students and applied researchers who are interested in causal mechanisms using linear models.

The first part of this book contains three chapters on the fundamentals of statistical association, causality, and linear regression. Chapter 1 is a very high-level introduction of the whole book and scientific research in general. Chapter 1 introduces mediation, moderation, and the more general conditional process analysis, and it also discusses correlation, causality, and statistical modeling in observational studies and experiments. Chapter 2 and Chapter 3 review the basics of simple and multiple linear regression, which play important roles for later discussion of causal mechanisms.

The second part of this book is about mediation analysis. This part starts with Chapter 4, a very detailed discussion of the simple mediation model. In Chapter 4, examples with both binary and continuous causal agents are used to illustrate the main ingredients of mediation analysis: parameter estimation, testing for indirect effect, and software implementation. Chapter 5 moves to multiple mediator models that allow for multiple causal pathways from the causal agent to the outcome of interest. Chapter 5 discusses both the parallel multiple mediator model and the serial multiple mediator model, and the author uses very illustrative figures to help the reader understand the intuition behind the associated statistical methods. These two more technical chapters are followed by Chapter 6, which critiques different classic approaches and discusses potential shortcomings and dangers of mediation analysis. For example, it has a nice critique of the traditional Baron–Kenny causal step approach, as well as the traditional view of complete and partial mediation. It also discusses the dangers of potential confounding between the mediator and outcome, and the possibly unclear causal order in cross-sectional data.

Part III of this book covers moderation analysis, and it has a parallel structure to Part II. Chapter 7 covers the basics, and Chapter 8 further discusses some special cases. Chapter 9 closes by touching more advanced topics in moderation analysis: truths and myths about mean centering regressors, artificial categorization of continuous moderators, and multiple moderators.

The fourth part of this book is about more general conditional process analysis, which unifies moderated mediation, mediated moderation, and other "hybrids" of these two causal mechanisms. Chapter 10 is mainly about moderated mediation analysis, including examples of moderation of the direct effect, the indirect effect, or both. Here, the new concepts of conditional direct and indirect effects generalize the unconditional effects in the simple mediation model, allowing these effects to depend on the value of the moderator. In Chapter 11, the author shows some further examples of conditional process analysis, and concludes that moderated mediation has a more meaningful interpretation than mediated moderation. Chapter 12 ends with a discussion about the strategy for approaching an analysis, the possibility of simultaneous moderation and mediation of the same variable, and the potential pitfalls for a subgroup analysis.

While I enjoyed reading the author's insightful discussions of methods in Chapters 1, 6, 9, and 12, I wish these discussions went beyond their current scope. For example, while the book discusses the potential confounding between the mediator and the outcome in Chapter 6—a very important issue in practice even if the treatment is randomly assigned—I wish the author had gone more in depth, and had discussed more about sensitivity analysis. Second, the book discusses mean centering in Chapter 9, but the scale of the outcome may be more crucial for effect moderation, as a nonlinear transformation of an outcome may completely remove a moderation. This is not discussed. Finally, the book defines all causal mechanisms in terms of regression coefficients of linear models, but modern statistical causal inference (e.g., the potential outcomes model due to J. Neyman and D. B. Rubin or the causal diagrams due to J. Pearl) allows us to define these causal mechanisms more generally and more precisely. I would have liked to see these alternative approaches at least mentioned.

In general, I think this is a very good textbook for applied researchers in social sciences. It covers mediation and moderation analysis using regression techniques quite nicely. The online materials of this book provide the data and software code for SAS and SPSS, which are very helpful supplements. I think this book could be very useful for both preliminary and advanced readers who are interested in mediation and moderation analysis. In particular, Chapters 1, 2, 3, 4, 5, 7, and 8 are nice introductions of the basic statistical concepts and causal mechanisms, and Chapters 6, 9, 10, 11, and 12 give a taste of more advanced topics. However, researchers should not stop here, but also find texts on alternative approaches based on more general frameworks.

Peng DING
Harvard University

Measures of Interobserver Agreement and Reliability (2nd ed.).

Mohamed M. SHOUKRI. Boca Raton, FL: CRC Press, 2010, xxi + 269 pp., $109.95 (H), ISBN: 978-1-43-981080-4.

This book is the second edition of a previously published book entitled *Measures of Interobserver Agreement*. This edition is a major revision of the first with added new materials and some removal of old. The revision almost doubled the length of the previous book. Specifically, the author added two new chapters: one on "Method Comparison Studies" and the other on "Population Coefficient of Variation as a Measure of Precision and Reproducibility." Three chapters in the first edition were either removed or were integrated into other chapters. The chapter on "Assessing Agreement from Dependent Data" was removed in the second edition, and the author refers the reader to von Eye's book entitled "Analyzing Rater's Agreement" (2005) for most of this content. Materials on "Sample Size Requirements for the Design of a Reliability Study" and "Workshops" are integrated in different chapters in the second edition.

The aim of the book is on evaluation of raters' agreement and reliability for both continuous and categorical data. The author greatly expanded the introductory chapter with added basic probability theory in Section 1.2. The readers who have a background in statistics may skip this section. One important conceptual presentation was added in Section 1.1.2 to differentiate the concepts of agreement and reliability. The new conceptual differentiation of agreement and reliability is interesting but can be confusing. Even though agreement and reliability are often used interchangeably, the author states that they are actually two different concepts, defining reliability as the ability to differentiate among subjects and agreement as conformity. However, I find these new definitions difficult to understand and the new concepts are not consistent between continuous data and categorical data. The notion of "the ability to differentiate among subjects" is particularly abstract and difficult to make sense of. The author subsequently states that the reliability is the ratio of the subjects' variance to the total variance, which coincides with the original definition of reliability. The author uses intraclass correlation coefficient (ICC) as reliability coefficient to illustrate its dependency on between-subject variability. This is similar to Kappa's dependency on prevalence because between-subject variability is determined by prevalence for binary data. The author seems to be aware of the correspondence between ICC and Kappa for continuous and categorical data, but continues to call ICC a reliability measure and Kappa an agreement measure. This causes further confusion because ICC and Kappa are essentially the same measure, but treated as different measures in the book. The author also seems to be unaware of the review article on assessing agreement by Barnhart et al. (2007) where reliability is treated as a narrower concept than agreement, in the sense that reliability is a scaled agreement index (a type of agreement measure). There are other unscaled agreement indices that may correspond to the author's narrower definition of agreement. The author did not specify which indices are agreement measures only nor elaborate what conformity means.

Chapter 2 focuses on using different versions of ICCs based on different analysis of variance (ANOVA) models to evaluate reliability for continuous data. The content of ICCs is presented based on the outline by Shrout and Fleiss (1979) for data without replications. The ICCs for data with replication are presented in a later section. It appears that this chapter did not include all of the different versions of ICCs and the corresponding estimates in McGraw and Wong (1996) or Chen and Barnhart (2008). The author includes new ICC materials on testing equality of ICCs, use of normality assumption on inference of ICC, balanced versus unbalance data, and approximation of confidence