GESIS TRAINING

gesis Leibniz Institute for the Social Sciences

48th GESIS Spring Seminar: Bayesian Modelling in the Social Sciences

Cologne, 11–29 March 2019

Syllabus for week 2: "Bayesian Structural Equation Modelling"

Lecturers:	Prof. Dr. Rens van de Schoot	Dr. Milica Miočević
Email:	A.G.J.vandeSchoot@uu.nl	m.miocevic@uu.nl
Homepage:	https://www.uu.nl/staff/agjvandeschoot	https://www.uu.nl/staff/MMiocevic

Date: 18-22 March 2019
Time: Tuesday to Friday 09:00-17:00 (Monday 10:00-18:00)

About the Lecturers:

Prof. Dr. Rens van de Schoot works as a professor at Utrecht University in the Netherlands and as extra-ordinary professor North-West University in South-Africa. He is a member of the Young Academy of The Royal Netherlands Academy of Arts and Sciences (KNAW), a group of fifty top-researchers (age below 45) selected based on the quality of their research as well as on the ability to communicate science to a broad audience, and a member of the Society of Multivariate Experimental Psychology, a small elective society of researchers who develop and apply multivariate quantitative methods. Rens obtained his PhD cum laude on the topic of applying Bayesian statistics to real life data at the Methodology and Statistics department at Utrecht University, The Netherlands. His research and teaching broadly focus on Bayesian structural equation models in the social sciences.

Dr. Milica Miočević is an Assistant Professor in the department of Methods and Statistics, faculty of Social and Behavioral Sciences, at Utrecht University. Her research focuses on the pros and cons of Bayesian methods for mediation analysis. She has written papers on the statistical properties of the mediated effect with diffuse and informative prior distributions, on statistical properties of effect size measures for the mediated effect computed in the Bayesian framework, and several tutorial papers on how to perform mediation analysis in the Bayesian framework. Milica teaches courses and workshops on Bayesian statistics, mediation analysis, multivariate statistics, and structural equation modeling.

Course Description:

During this course students will be introduced to philosophical underpinnings of Bayesian statistics and will learn how to fit regression, mediation, CFA, and longitudinal growth models in the Bayesian framework. Students will learn the steps in conducting Bayesian analyses and will be able to understand articles that examine and apply Bayesian SEM. The course is highly interactive, and the afternoons will be dedicated to implementing and practicing the material using the participant's software of choice (Mplus or R in tandem with JAGS or stan). We highly recommend bringing your own data for Day 5 of the course; however, the instructors will have example data sets for participants who do not have their own data.

Keywords:

Bayesian statistics, SEM, priors, data analysis, factor analysis, regression analysis, mediation analysis, longitudinal growth models

Course Prerequisites:

Participants should have knowledge of regression analysis and basic SEM. No previous knowledge of Bayesian analysis is assumed. Participants should have a good grasp of the software package they plan to use (R or Mplus).

Target Group:

Participants will find the course useful if they

- Are interested in using Bayesian statistics in their own work.
- Encounter convergence issues using classical methods.
- Have small samples and/or access to prior information from previous studies.
- Wish to understand new methodological developments that make use of Bayesian statistics and/or Markov Chain Monte Carlo (MCMC).

Course and Learning Objectives:

By the end of the course participants will:

- Know the differences between 'classical' and Bayesian statistics, and when to use to Bayesian analyses instead of classical statistics.
- Know how to apply Bayesian SEM to analyze their own data.
- Know how to apply the WAMBS-checklist (When to worry and how to Avoid the Misuse of Bayesian Statistics).
- Critically evaluate applications of Bayesian methods in scientific studies.

Organisational Structure of the Course:

The course is structured around lectures in the morning and hands-on lab sessions in the afternoon with a lunch break in between. The first day provides an overview of the Bayesian framework and how it differs from classical statistics. The second day covers prior distributions and steps in a Bayesian analysis. Bayesian CFA and path models are covered on day 3, and day 4 focuses on two types of Bayesian SEM (mediation models with latent variables and longitudinal growth models). Day 5 consists of hands-on exercises in prior specification and applying the skills learned during days 1-4 using the participants' own data (or example data sets provided by the instructors).

Software and hardware requirements:

Participants will have the choice between using Mplus or R for the course exercises. Participants can bring their own laptops with the software of their choice installed. GESIS also provides participants with workshop computers and all relevant software.

Long Course Description:

The popularity of Bayesian statistics has increased over the years, however, as of now Bayesian methods are not a part of the statistics curricula in most graduate programs internationally. The Bayesian estimation framework can handle some commonly encountered problems in classical statistics, such as the lack of power in small sample research and convergence issues in complex models. Furthermore, some researchers prefer the Bayesian framework because it provides a way of sequentially updating knowledge with new data instead of requiring that each new study tests the null hypothesis that there is no effect in the population.

During this course, students will be gently introduced to Bayesian statistics using class examples. The instructors will clarify the differences between the philosophies and interpretations in classical and Bayesian frameworks, and will illustrate types of research questions that can be answered only using Bayesian methods. This course will also give students experience with running Bayesian analyses and interpreting results, and will instruct participants on the prevailing "best practices" for Bayesian estimation in structural equation models. Participants will emerge from the course with knowledge about how to apply Bayesian methods to answer their research questions, and with the ability to understand articles that examine and apply Bayesian methods for structural equation modeling. We highly recommend bringing your own data as well; however, we have plenty of data available for participants to analyze. Using these examples, we will explore the benefits of Bayesian statistics and discuss what is needed to fit your first Bayesian structural equation model.

This highly interactive 5-day course gently introducing Bayesian estimation for linear regression analysis, factor analysis, mediation analysis with manifest and latent variables, and longitudinal growth models. The first four days are designed to teach participants on how to estimate the above models in the Bayesian framework, and day 5 is dedicated solely to practicing new skills on the participants' data sets (or example data sets provided by instructors).

Participants from a variety of fields including psychology, education, human development, public health, prevention science, sociology, marketing, business, biology, medicine, political science, and communication will benefit from the course.

Preparatory Reading:

- van de Schoot, R., & Depaoli, S. (2014). Bayesian analyses: Where to start and what to report. European Health Psychologist, 16(2), 75-84.
- Kaplan, D., & Depaoli, S. (2012). Bayesian Structural Equation Modeling. In: R. Hoyle (Ed.), Handbook of structural equation modeling. New York: Guilford Press.

Recommended Literature to Look at in Advance:

N/A

Day-to-day Schedule and Literature:

Day	Topic(s)
1	*Morning:* Conceptual introduction + reasons for using Bayesian methods + discussion on interpretability of results when using p-values/95%intervals + model selection+ posterior distributions + credibility intervals *Afternoon:* Computer lab software intro (choose from Mplus and R in combination with JAGS or STAN) and fitting your first Bayesian model.
	Suggested reading: ▪ Gigerenzer, G. (1993). The superego, the ego, and the id in statistical reasoning. *A handbook for data analysis in the behavioral sciences: Methodological issues*, 311-339. ▪ Van de Schoot, R., Kaplan, D., Denissen, J., Asendorpf, J. B., Neyer, F. J., & van Aken, M. A. (2014). A gentle introduction to Bayesian analysis: Applications to developmental research. *Child development*, *85*(3), 842-860.
2	*Morning:* Q&A + WAMBS-checklist (when to worry and how to avoid the misuse of Bayesian Statistics) + Bayesian regression analysis *Afternoon:* Computer lab on sensitivity to prior distributions + Bayesian regression analysis + WAMBS-checklist
	Suggested reading: ▪ Depaoli, S., & van de Schoot, R. (2017). Improving transparency and replication in Bayesian statistics: The WAMBS-Checklist. *Psychological Methods, 22*(2), 240-261.
3	*Morning:* Q&A + Bayesian CFA + Bayesian mediation analysis *Afternoon:* Computer lab on Bayesian CFA and Bayesian mediation analysis
	Suggested reading: ▪ Miočević, M., MacKinnon, D. P., & Levy, R. (2017). Power in Bayesian mediation analysis for small sample research. *Structural equation modeling: a multidisciplinary journal*, *24*(5), 666-683. ▪ Van De Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., & Muthén, B. (2013). Facing off with Scylla and Charybdis: a comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in psychology*, *4*, 770.
4	*Morning:* Q&A + Bayesian mediation analysis with latent variables + Bayesian longitudinal growth models (LGMs) *Afternoon:* Computer lab on Bayesian latent variable mediation analysis and Bayesian LGMs
	Suggested reading: ▪ Chen, J., Choi, J., Weiss, B. A., & Stapleton, L. (2014). An empirical evaluation of mediation effect analysis with manifest and latent variables using Markov Chain Monte Carlo and alternative estimation methods. Structural Equation Modeling: A Multidisciplinary Journal, 21(2), 253-262. ▪ Zhang, Z., Hamagami, F., Lijuan Wang, L., Nesselroade, J. R., & Grimm, K. J. (2007). Bayesian analysis of longitudinal data using growth curve models. International Journal of Behavioral Development, 31(4), 374-383.
5	*Morning:* Q&A + exercise on finding prior information for your Bayesian analysis *Afternoon:* Analyzing your own data (or example data) using Bayesian methods
	Suggested reading: ▪ N/A