EAA PhD Forum 2018: Bayesian Statistics in Accounting Research

Harm H. Schuett – LMU Munich School of Management

E-mail: schuett@bwl.lmu.de Web: github.com/hschuett/EAA2018Bayes

---

"If I'm doing an experiment to save the world, I better use my prior." – Andrew Gelman

Learning Objectives

- Review hypothesis testing and uncertainty
- Understand why arguing "if it is a noisy measure it will work against me finding a result" is dangerous.
- Discuss how to use priors to combat the noise problem
- Discuss how to use priors/updating to identify latent variables
- Start thinking about where you can apply this in your own work

Description

This session is aimed at young accounting researchers who are interested in learning new, robust methods to identify the many latent constructs we deal with in accounting. Because of advances in computing power and the public debate about issues in the application of classical hypothesis testing (e.g., Simmons, Nelson, and Simonsohn 2011; Dyckman and Zeff 2014; Gelman and Carlin 2014; Harvey 2017), Bayesian statistics is gaining more and more traction in various areas of the social sciences. This session is intended to give a short, accessible introduction to Bayesian Statistics and its possible use case for accounting research.

Bayesian statistics excels at two things. First it helps incorporating external knowledge into the model, thereby regularizeing estimates (i.e., reducing the chance of noise fitting). Second it provides a flexible approach to model latent variables. Both use cases hold significant potential for accounting research questions that involve hard to measure constructs from noisy data. Such constructs are for example: disclosure characteristics (e.g., readability), accrual quality, undetected fraud (Hahn, Murray, and Manolopoulou 2016), latent topics and their distributions in a corpus of documents (e.g., Dyer, Lang, and Stice-Lawrence 2017), latent financial news audiences (Schütt 2017), or uncovering incrementally useful variables (Cremers 2002). In 75 minutes we will discuss:

1. What is the difference to frequentist statistics?
2. What are the practical (rather than philosophical) advantages of Bayesian statistics?
3. In which areas and settings is Bayesian statistics useful to accounting researchers?
4. Bayesian statistics is computationally more complex. What are the practical hurdles?

Bayesian and Frequentist statistics are two tools. Each has advantages and disadvantages. Talking about the differences between Bayesian and Frequentist approaches to statistics is instructive not only for understanding how Bayesian data analysis works, but also for better understanding how to apply frequentist methods. Thus, we will spend some time on that. The remainder of the session will use a few examples to illustrate how Bayesian analysis works and when it is most useful for us: Settings where we want to model heterogeneity and settings with noisy data or hard to measure constructs. In such situations, the chance of accidental noise fitting and false positives is high. Here we would like to use every bit of uncontroversial prior knowledge we have to improve the precision of our inferences. Bayesian methods offer a very flexible and intuitive approach to do just that (Gelman et al. 2013). Many people find, as Nobel laureate Christopher Sims remarked: "Once one becomes used to thinking about inference from a Bayesian perspective, it becomes difficult to understand why many econometricians are uncomfortable with that way of thinking" (Sims 2010, 1; Sims 2007). However, the downside is that these methods are more complex to code and computationally intensive. Thus, we will end with a discussion of the practical hurdles of Bayesian approaches.

Intended Audience

Anyone with a basic understanding of classical regression analysis and an interest in new methods for measuring constructs or modelling heterogeneity.

Materials

The examples used will be coded using R and Stan. Slides, R and Stan code will be available at the github repo and ARC platform shortly before the session.

References

Cremers, K.J. Martijn. 2002. "Stock Return Predictability: A Bayesian Model Selection Perspective." *The Review of Financial Studies* 15 (4): 1223–49.

Dyckman, Thomas R, and Stephen A Zeff. 2014. "Some Methodological Deficiencies in Empirical Research Articles in Accounting." *Accounting Horizons* 28 (3): 695–712.

Dyer, Travis, Mark Lang, and Lorien Stice-Lawrence. 2017. "The Evolution of 10-K Textual Disclosure: Evidence from Latent Dirichlet Allocation." *Journal of Accounting and Economics* 64 (2-3): 221–45.

Gelman, Andrew, and John Carlin. 2014. "Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors." *Perspectives on Psychological Science* 9 (6): 641–51.

Gelman, Andrew, John B Carlin, Hal S Stern, David B Dunson, Aki Vehtari, and Donald B Rubin. 2013. *Bayesian Data Analysis*. 3rd ed. Chapman & Hall/CRC.

Hahn, P Richard, Jared S Murray, and Ioanna Manolopoulou. 2016. "A Bayesian Partial Identification Approach to Inferring the Prevalence of Accounting Misconduct." *Journal of the American Statistical Association* 111 (513): 14–26.

Harvey, Campbell R. 2017. "Presidential Address: The Scientific Outlook in Financial Economics." *The Journal of Finance*.

Schütt, Harm H. 2017. "Competition in Financial News Markets and Trading Activity." Available at SSRN: https://ssrn.com/abstract=3051732.

Simmons, Joseph P, Leif D Nelson, and Uri Simonsohn. 2011. "False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22 (11): 1359–66.

Sims, Christopher. 2007. "Bayesian Methods in Applied Econometrics, or Why Econometrics Should Always and Everywhere Be Bayesian." Teaching note, Department of Economics, Princeton University. http://sims.princeton.edu/yftp/EmetSoc607/AppliedBayes.pdf.

———. 2010. "Understanding Non-Bayesians." Unpublished chapter, Department of Economics, Princeton University. http://sims.princeton.edu/yftp/UndrstndgNnBsns/GewekeBookChpter.pdf.