



Introduction to the vol. 47, no. 1, 2020

Maomi Ueno¹

Published online: 17 February 2020
© The Behaviormetric Society 2020

Welcome to the Vol. 47, No. 1, 2020 of Behaviormetrika. In this issue, we have the following eleven original papers, one note, and two invited papers.

This issue includes a special feature: “Advances in Composite-based Structural Equation Modeling” (Hwang et al. 2020; Cho and Choi 2020; Jung et al. 2020; Ryoo et al. 2020; Schamberger et al. 2020) which was edited by Heungsun Hwang and Marko Sarstedt. This special issue seeks to serve as a platform for advancing and furthering our understanding of composite-based SEM methods. The regular papers of this issue include the following seven original papers, one note, and one invited paper.

The first invited paper is “The practical importance of understanding placebo effects and their role when approving drugs and recommending doses for medical practice” by Donald B. Rubin who is most well-known for the Rubin Causal Model, a set of methods designed for causal inference with observational data, and for his methods for dealing with missing data. I invited him for the invited paper because Rubin Causal Model will play one of the most important roles in data science. In addition, I found that his unique personal history might have greatly affected his great academic works. Therefore, I asked him to add his personal history in this paper. This paper mainly provides insightful reconsideration of “blind treatments for placebo effects observed using the causal model” in Jin and Rubin (2008). Especially, he emphasizes that we should consider conscious humans do not act like laboratory rats to design experiments.

The original paper “Cognitive diagnosis models for estimation of misconceptions analyzing multiple-choice data” by Koken Ozaki, Shingo Sugawara, and Noriko Arai proposes two statistical models that can estimate examinees’ possession of misconceptions by analyzing multiple-choice data, which are unscored data. By converting multiple-choice data to binary response data (correct: 1 or wrong: 0), the Bug-DINO model can estimate examinees’ possession of misconceptions. However, converting multiple-choice data to binary data causes a loss in information, because which incorrect option an examinee chooses is important information for an examinee’s knowledge state. The three models (two developed

✉ Maomi Ueno
ueno@ai.is.uec.ac.jp

¹ University of Electro-Communications, Tokyo, Japan

models and the Bug-DINO model) are compared in a simulation study, and the developed models are applied to the reading skill test data.

The original paper “Using the discontinuation rule to reduce the effect of random guessing on parameter estimation in the item response theory” by Tian-shu Pan and Youngmi Cho explored the feasibility of the four- and six-consecutive-zero discontinuation rules to reduce the effect of random guessing on parameter estimation in the Rasch model. The results showed that random guessing inflated estimation errors and these discontinuation rules can reduce this effect on item-parameter estimation under the joint and marginal maximum likelihood, but can do so only for person-parameter estimation under the marginal maximum likelihood and expected a posteriori methods.

The original paper “Addressing score comparability in diagnostic classification models: an observed-score equating and linking approach” by Ren Liu tries to revamp and examine the observed-score approach for equating scores under diagnostic classification models (DCMs). The observed score approach was adapted to accommodate the categorical latent traits and scores, a unique property of DCMs. Three simulation studies, each corresponding to a data collection design, were conducted to evaluate the amount of equating error and the improvement in score comparability with the observed-score approach. Findings indicate that (a) DCM scores are robust to form differences when high-quality items are used and (b) the observed-score approach shows promise for yielding a small amount of equating error and increasing classification accuracy under small sample size conditions.

The original paper “Bayesian network Fisher kernel for categorical feature spaces” by Janne Leppä-aho, Tomi Silander, and Teemu Roos proposes a new similarity function between vectors of possibly dependent categorical variables using the Fisher kernel for Bayesian networks. This study also experimentally demonstrates how this kernel can be used to find subsets of observations that we see as representative for the underlying Bayesian network model.

The original paper “An estimation of causal structure based on Latent LiNGAM for mixed data” by Mako Yamayoshi, Jun Tsuchida, and Hiroshi Yadohisa proposes the Latent LiNGAM (L-LiNGAM), where each variable corresponds to a continuous latent variable and is observed as data through transformation via a link function. In the numerical study, when mixing discrete variables, the estimation of causal structure using L-LiNGAM is proven useful in terms of sum of squared error and path recovery. Moreover, from real-world data applications, the causal structure estimated by L-LiNGAM is shown to be the best for evaluation under SEM. The model is also superior to that of existing methods.

The original paper “The weak-instruments problem in factor models” by Kazuhiko Hayakawa addresses the instrumental variable estimation of factor models. Specifically, this study investigates the weak-instruments problem, which is not well investigated in the literature, in detail. It shows that the signal-to-noise ratios, which are defined by the variance ratios of the common components to the errors of the scaling variable and the instruments mainly determine the strength of instruments in a confirmatory factor model, while the structure of the factor loading is closely related to the strength of instruments in an explanatory factor model.

Experiments' results are consistent with the theoretical implications, and Stock and Yogo test is useful in practice in detecting the weak instruments.

The original paper “Variational Bayesian inference for the multiple-choice DINA model” by Kazuhiro Yamaguchi proposes, a fully Bayesian formulation for the multiple-choice item version of the deterministic input noisy “AND” gate (MC-DINA) model, which represents a cognitive diagnostic model for extracting information from multiple-choice response data. In addition, a variational inference algorithm containing an empirical Bayesian estimation procedure was developed to solve heavy computational burden problems in Bayesian statistics procedure. The proposed method is as fast as the expectation–maximization algorithm because it does not require random sampling integration like the Markov chain Monte Carlo technique. Moreover, this algorithm can automatically extract optimal hyper-parameters from analyzed data by maximizing the lower bound of the logarithm of the marginal likelihood function. The simulation results show that the proposed technique can successfully recover the true item and student parameters.

The note “Classification analysis of Kouji Uno’s novels using topic model” by Xueqin Liu and Mingzhe Jin tries to analyze Japanese littérateur Kouji Uno’s products using a statistical topic model. Specifically, they classify Uno’s creative phase using Latent Dirichlet Allocation to conduct an investigation into the stylistic characteristics of his novels. As revealed by the results, Uno’s novels can be classified into three groups separated approximately by the two non-productive periods and there are different stylistic characteristics displayed by novels in each group. Moreover, one interesting observation is that his stylistic characteristics have changed even prior to the interruptions caused to writing. It is more reasonable that Uno’s writing style started to change before the interruptions with achievements made to some extent after the resumption.

References

- Cho G, Choi JY (2020) An empirical comparison of generalized structured component analysis and partial least squares path modeling under variance-based structural equation models. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00098-0>
- Hayakawa H (2020) The weak-instruments problem in factor models. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00097-1>
- Hwang H, Sarstedt M, Cheah JH, Ringle CM (2020) A concept analysis of methodological research on composite-based structural equation modeling: bridging PLSPM and GSCA. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00085-5>
- Jin H, Rubin DB (2008) Principal stratification for causal inference with extended partial compliance: application to Efron–Feldman data. *J Am Stat Assoc* 103:101–111
- Jung K, Cho SS, Lee J, Kim S, Ryoo JH (2020) An illustrative application of generalized structured component analysis for brain connectivity research. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00080>
- Leppä-aho J, Silander T, Roos T (2020) Bayesian network Fisher kernel for categorical feature spaces. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00103-6>
- Liu R (2020) Addressing score comparability in diagnostic classification models: an observed-score equating and linking approach. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00102-7>

- Liu X, Jin M (2020) Classification analysis of Kouji Uno's novels using topic model. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00099-z>
- Ozaki K, Sugawara S, Arai N (2020) Cognitive diagnosis models for estimation of misconceptions analyzing multiple-choice data. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00100-9>
- Pan T, Cho Y (2020) Using the discontinuation rule to reduce the effect of random guessing on parameter estimation in the item response theory. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00101-8>
- Rubin DB (2020) The practical importance of understanding placebo effects and their role when approving drugs and recommending doses for medical practice. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00091-7>
- Ryoo JH, Park S, Kim S (2020) Categorical latent variable modeling utilizing fuzzy clustering generalized structured component analysis as an alternative to latent class analysis. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00084-6>
- Schamberger T, Schuberth F, Henseler J, Dijkstra TK (2020) Robust partial least squares path modeling. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00088-2>
- Yamaguchi K (2020) Variational Bayesian inference for the multiple-choice DINA model. *Behaviormetrika*. <https://doi.org/10.1007/s41237-020-00104-w>
- Yamayoshi M, Tsuchida J, Yadohisa H (2020) An estimation of causal structure based on Latent LiNGAM for mixed data. *Behaviormetrika*. <https://doi.org/10.1007/s41237-019-00095-3>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.