Journal of Biometrics and Biostatistics

Open Access

Powerful Derivation for the Ventured Wedge Plan

Siddiqui Farman*

Department of Entomology, Pennsylvania State University, Pennsylvania, USA

Editorial Note

Stepped wedge designed trials are a type of cluster-randomized study in which the intervention is introduced to each cluster in a random order over time. This design is often used to assess the effect of a new intervention as it is rolled out across a series of clinics or communities. Based on a permutation argument, we derive a closed-form expression for an estimate of the intervention effect, along with its standard error, for a stepped wedge design trial. We show that these estimates are robust to misspecification of both the mean and covariance structure of the underlying data-generating mechanism, thereby providing a robust approach to inference for the intervention effect in stepped wedge designs. We use simulations to evaluate the type 1 error and power of the proposed estimate and to compare the performance of the proposed estimate to the optimal estimate when the correct model specification is known. The limitations, possible extensions, and open problems regarding the method are discussed. Many estimators of the average effect of a treatment on an outcome require estimation of the propensity score, the outcome regression or both. It is often beneficial to utilize flexible techniques, such as semiparametric regression or machine learning, to estimate these quantities. However, optimal estimation of these regressions does not necessarily lead to optimal estimation of the average treatment effect, particularly in settings with strong instrumental variables. A recent proposal addressed these issues via the outcome-adaptive lasso, a penalized regression technique for estimating the propensity score that seeks to minimize the impact of instrumental variables on treatment effect estimators. However, a notable limitation of this approach is that its application is restricted to parametric models. We propose a more flexible alternative that we call the outcome highly adaptive lasso. We discuss the large sample theory for this estimator and propose closed-form confidence intervals based on the proposed estimator. We show via simulation that our method offers benefits over several popular approaches.

Identifiability of statistical models is a fundamental regularity condition that is required for valid statistical inference. Investigation of model identifiability is mathematically challenging for complex models such as latent class models. Jones et al. used Goodman's technique to investigate the identifiability of latent class models with applications to diagnostic tests in the absence of a gold standard test. The tool they used was based on examining the singularity of the Jacobian or the Fisher information matrix, in order to obtain insights into local identifiability (i.e., there exists a neighborhood of a parameter such that no other parameter in the neighborhood leads to the same probability distribution as the parameter). In this paper, we investigate a stronger condition: global identifiability (i.e., no two parameters in the parameter space give rise to the same probability distribution), by introducing a powerful mathematical tool from computational algebra: the Gröbner basis. With several existing well-known examples, we argue that the Gröbner basis method is easy to implement and powerful to study global identifiability of latent class models, and is an attractive alternative to the information matrix analysis by Rothenberg and the Jacobian analysis.

Semiparametric Mixed-Scale Models

In this paper, we propose a unified Bayesian joint modeling framework for studying association between a binary treatment outcome and a baseline matrix-valued predictor. Specifically, a joint modeling approach relating an outcome to a matrix-valued predictor through a probabilistic formulation of multilinear principal component analysis is developed. This framework establishes a theoretical relationship between the outcome and the matrix-valued predictor, although the predictor is not explicitly expressed in the model. Simulation studies are provided showing that the proposed method is superior or competitive to other methods, such as a two-stage approach and a classical principal component regression in terms of both prediction accuracy and estimation of association; its advantage is most notable when the sample size is small and the dimensionality in the imaging covariate is large.

Corresponding Author: Siddiqui Farman, Department of Entomology, Pennsylvania State University, Pennsylvania, USA, Tel: +1 814-865-4700; E-mail:fas5130@psu.edu