

Information projection approach to smoothed propensity score weighting for handling selection bias under missing at random

Hengfang Wang¹ · Jae Kwang Kim²

Received: 25 May 2023 / Revised: 25 May 2024 / Accepted: 5 August 2024 / Published online: 21 September 2024 © The Institute of Statistical Mathematics, Tokyo 2024

Abstract

Propensity score weighting is widely used to correct the selection bias in the sample with missing data. The propensity score function is often developed using a model for the response probability, which completely ignores the outcome regression model. In this paper, we explore an alternative approach by developing smoothed propensity score weights that provide a more efficient estimation by removing unnecessary auxiliary variables in the propensity score model. The smoothed propensity score function is obtained by applying the information projection of the original propensity score function to the space that satisfies the moment conditions on the balancing scores obtained from the outcome regression model. By including the covariates for the outcome regression models only in the density ratio model, we can achieve an efficiency gain. Penalized regression is used to identify important covariates. Some limited simulation studies are presented to compare with the existing methods.

Keywords Calibration estimation \cdot Density ratio model \cdot Self-efficiency \cdot Missing data

The online version of this article contains supplementary material. The research of the first author was partially supported by the National Natural Science Foundation of China (12401339), the Open Research Fund of Key Laboratory of Analytical Mathematics and Applications (Fujian Normal University), Ministry of Education, P. R. China (JAM2403) and Middle-aged and Young Teachers' Training Program (SDPY2023024). The research of the second author was partially supported by a grant from US National Science Foundation (2242820), a grant from the Iowa Agriculture and Home Economics Experiment Station, Ames, Iowa, and from the U.S. Department of Agriculture's National Resources Inventory.

[☐] Jae Kwang Kim jkim@iastate.edu

¹ School of Mathematics and Statistics & Fujian Provincial Key Laboratory of Statistics and Artificial Intelligence, Fujian Normal University, Fuzhou, Fujian 350117, China

² Department of Statistics, Iowa State University, 2438 Osborn Dr., Ames, IA 50011, USA