

SEMINAR DEPARTMENT OF STATISTICS THE CHINESE UNIVERSITY OF HONG KONG

Sample size and power calculation for propensity score analysis of observational studies

INVITED SPEAKER

Fan Li

Professor

Department of Statistical Science & Department of Biostatistics and Bioinformatics

Duke University

TIME

February 10th, 2025 (Mon) · 2:30 pm - 3:30 pm

VENUE

KKB 101 (1/F) · Leung Kau Kui Building 101 · CUHK

ABSTRACT

Sample size and power calculations in causal inference with observational data are increasingly desired, but related tools were lacking. This paper develops theoretically justified analytical formulas for sample size and power calculation in the propensity score analysis of causal inference. By analyzing the variance of the inverse probability weighting estimator of the average treatment effect (ATE), we clarify the three key components for sample size calculations: propensity score distribution, potential outcome distribution, and their correlation. We devise analytical procedures to identify these components based on commonly available and interpretable summary statistics. We elucidate the critical role of covariate overlap between treatment groups in determining the sample size. In particular, we propose to use the Bhattacharyya coefficient as a measure of covariate overlap, which, together with the treatment proportion, leads to a uniquely identifiable and easily computable propensity score distribution. The proposed method is applicable to both continuous and binary outcomes. We show that the standard two-sample z-test and variance inflation factor methods often lead to, sometimes vastly, inaccurate sample size estimates, especially with limited overlap. We also derive formulas for the average treatment effects for the treated (ATT) and overlapped population (ATO) estimands. We provide simulated and real examples to illustrate the proposed method. We develop an associated R package PSpower. This is a joint work with Bo Liu and Xiaoxiao Zhou.