

ADJUSTING FOR NON-CONFOUNDING COVARIATES IN CASE-CONTROL ASSOCIATION STUDIES

Siliang Zhang, Jinbo Chen, Zhiliang Ying and Hong Zhang*

*East China Normal University, University of Pennsylvania,
Columbia University and University of Science and Technology of China*

Abstract: There is a substantial literature in case-control logistic regression on whether or not non-confounding covariates should be adjusted for. However, only limited and ad hoc theoretical results are available on this important topic. A constrained maximum likelihood method was recently proposed, which appears to be generally more powerful than logistic regression methods with or without adjusting for non-confounding covariates. This note provides a theoretical clarification for the case-control logistic regression with and without covariate adjustment and the constrained maximum likelihood method on their relative performances in terms of asymptotic relative efficiencies. We show that the benefit of covariate adjustment in the case-control logistic regression depends on the disease prevalence. We also show that the constrained maximum likelihood estimator gives an asymptotically uniformly most powerful test.

Key words and phrases: Asymptotic relative efficiency, case-control design, constrained maximum likelihood, logistic regression, non-confounding covariate.

1. Introduction

It is well known that adjusting for baseline covariates can lead to an improved statistical inference efficiency (Fisher, 1958). Through rigorous derivation, Robinson and Jewell (1991) showed that in logistic regression, adjusting for non-confounding covariates in randomized clinical trials always benefits the testing for treatment effect in terms of Pitman's asymptotic relative efficiency (ARE), albeit with some estimation precision loss. Neuhaus (1998) extended the results of Robinson and Jewell (1991) to generalized linear models.

It is more complicated, however, when data are collected retrospectively but analyzed using prospective logistic regression. In particular, Kuo and Feingold (2010) showed through simulations that adjusting for non-confounding covariates in case-control studies may decrease the statistical inference efficiency for exposure-disease association; see Xing and Xing (2010) for additional comments. Pirinen, Donnelly and Spencer (2012) showed that adjusting for non-confounding covariates in case-control studies can decrease estimation precision, resulting in a loss of power provided that the estimation is approximately unbiased. In general,

*Corresponding author. E-mail: zhangh@ustc.edu.cn