Statistica Sinica: Preprint
doi:10.5705/ss.202016.0320

In observational studies, propensity score methods are popular for estimating causal effects. With completely observed data, this approach is valid under several assumptions; however, in practice data are often missing which can have a substantial impact on the estimation. Current remedies to deal with missing covariates in propensity score methods generally fall into two categories. Some authors propose to account for the missing data patterns in propensity score estimation. Others propose to first impute the missing data, then utilize conventional propensity score adjustment methods. Both approaches assume that the data are missing at random (MAR), and there is little discussion regarding the impact on treatment effect estimation if covariates are missing not at random (MNAR). In this paper, we first examine the implication of the MAR assumption under the potential outcome framework. We then propose a sensitivity analysis method for assessing the impact of a MNAR covariate on treatment effect estimation with a matching estimator, with varying magnitudes of unmeasured confounding effect due to the missing covariate. Our method takes full advantage of the information contained in the partially missing covariate by matching on the observed portion and identifying a bounding distribution for the missing portion. It can be interpreted similarly as Rosenbaum's sensitivity analysis, and the results are robust since we make few parametric assumptions. We illustrate the application of the method using the 2012 Ohio Medicaid Assessment Survey (OMAS) to investigate the effect of health insurance on health outcomes, where an important covariate, household income, is partially missing.