Statistical Science
2007, Vol. 22, No. 4, 544–559
DOI: 10.1214/07-STS227D
Main article DOI: 10.1214/07-STS227

Comment: Performance of Double-Robust Estimators When "Inverse Probability" Weights Are Highly Variable

James Robins, Mariela Sued, Quanhong Lei-Gomez and Andrea Rotnitzky

1. GENERAL CONSIDERATIONS

We thank the editor Ed George for the opportunity to discuss the paper by Kang and Schaeffer.

The authors' paper provides a review of double-robust (equivalently, double-protected) estimators of (i) the mean $\mu = E(Y)$ of a response Y when Y is missing at random (MAR) (but not completely at random) and of (ii) the average treatment effect in an observational study under the assumption of strong ignorability. In our discussion we will depart from the notation in Kang and Schaeffer (throughout, K&S) and use capital letters to denote random variables and lowercase letter to denote their possible values.

In the missing-data setting (i), one observes n i.i.d. copies of $O = (T, X, TY)$, where X is a vector of always observed covariates and T is the indicator that the response Y is observed. An estimator of μ is double-robust (throughout, DR) if it remains consistent and asymptotically normal (throughout, CAN) when either (but not necessarily both) a model for the propensity score $\pi(X) \equiv P(T = 1|X) = P(T = 1|X, Y)$ or a model for the conditional mean $m(X) \equiv E(Y|X) = E(Y|X, T = 1)$ is correctly specified, where the equalities follow from the MAR assumption. The authors demonstrate, via simulation, that when a linear logistic model for the propensity score and a linear model for the mean of Y given X are both moderately misspecified, there exists a joint distribution under which the OLS regression estimator $\hat{\mu}_{OLS}$ of μ outperforms all candidate estimators that depend on a linear logistic maximum likelihood estimate of the propensity score, including all the DR estimators considered by the authors.

Near the end of their Section 1, the authors state that their simulation example "appears to be precisely the type of situation for which the DR estimators of Robins et al. were developed." They then suggest that their simulation results imply that the cited quotation from Bang and Robins (2005) is incorrect or, at the very least, misguided. We disagree with both the authors' statement and suggestion. First, the cited quote neither claims nor implies that when a linear logistic model for the propensity score and a linear model for the mean of Y given X are moderately misspecified, DR estimators always outperform estimators—such as regression, maximum likelihood, or parametric (multiple) imputation estimators—that do not depend on the estimated propensity score. Indeed, Robins and Wang (2000) in their paper "Inference for Imputation Estimators" stated the following:

> If nonresponse is ignorable, a locally semiparametric efficient estimator is doubly protected; i.e., it is consistent if either a model for nonresponse or a parametric model for the complete data can be correctly specified. On the other hand, consistency of a parametric multiple imputation estimator requires correct specification of a parametric model for the complete data. However, in cases in which the variance of the 'inverse probability' weights is very large, the sampling distribution of a locally semiparametric efficient (augmented inverse probability of response weighted) estimator can be markedly skew and highly variable, and a parametric imputation estimator may be preferred.

The just-quoted cautionary message of Robins and Wang (2000) is not far from K&S's take-home mes-

James Robins is Professor, Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts 02115, USA (e-mail: akaris@hsph.harvard.edu). Mariela Sued is Assistant Professor, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and CONICET, Argentina. Quanhong Lei-Gomez is a Graduate Student, Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts 02115, USA. Andrea Rotnitzky is Professor, Department of Economics, Di Tella University, Buenos Aires, Argentina.