Robust semi-parametric estimators: missing data and causal inference

E. Cantoni¹ and X. de Luna²*

¹ Research Center for Statistics and Geneva School of Economics and Management, University of Geneva, Switzerland; Eva.Cantoni@unige.ch
² Department of Statistics, USBE, Umeå University, Umeå, Sweden; xavier.deluna@umu.se.
* Presenting author

Keywords. bounded influence function; inverse probability weighting; double robust.

We consider situations where we aim at estimating location and scale parameters from a distribution law of interest, from which a random sample has been drawn. We introduce semi-parametric estimators, which are able to deal simultaneously with two common challenges within this general context: (i) not all observations from the random sample intended are available (incomplete data due to dropout, selection, potential outcomes framework), and (ii) some of the available observations in the sample may be contaminated (generated by a nuisance distribution, outliers). Under an assumption of ignorable missingness, popular semi-parametric estimators of the parameters of interest are augmented inverse probability weighted (AIPW, doubly robust) estimators (e.g., ?). They use two auxiliary models, one for the missingness mechanism, and another for an outcome of interest, both given observed covariates. AIPW estimators are then robust to misspecification of one of these two models (but not both simultaneously - a so called double robustness property). We introduce versions of AIPW, which provide, moreover, robustness to contamination of the distribution of interest. Asymptotic properties are described and finite sample results are presented. We motivate the need for robust AIPW estimators with a follow up study on BMI combining data from an intervention study and population wide record linked data.

References