

DSS Statistics Seminar

June 6, 2023, 12:00

In person Room VI (Building CU002 - Città Universitaria)

Webinar <https://uniroma1.zoom.us/j/86881977368?pwd=SWRFcVFjMDZTa0lXZk05TE1zNm5adz09>
Passcode: 432940

Statistical Inference in the Presence of Imputed Survey Data Through Regression Trees and Random Forests

David Haziza

University of Ottawa

Item nonresponse in surveys is usually handled through some form of imputation. Regression trees and random forests provide flexible tools for obtaining a set of imputed values. In this presentation, we lay out a set of conditions on the imputation model sufficient for the consistency of imputed estimators based on regression trees and random forests. We will introduce a novel variance estimator that accounts for sampling and nonresponse. The choice of hyper-parameters will also be discussed. Finally, we will present the results from a simulation study that investigates the performance of point and variance estimators in terms of bias, efficiency and coverage rate.

Co-authors: Mehdi Dagdoug (University of Ottawa) and Camelia Goga (Université de Bourgogne Franche Comté)



SAPIENZA
UNIVERSITÀ DI ROMA

**Organized in collaboration with
Istat Advisory Committee
on statistical methods**