REWEIGHTING WITH MACHINE LEARNING TECHNIQUES IN PANEL SURVEYS. APPLICATION TO THE HEALTH CARE AND SOCIAL SURVEY

L. Castro-Martín[1] and M. M. Rueda[2] and C. Sánchez-Cantalejo[3] and R. Ferri-García[4] and J. Hidalgo Calderón[5] and A. Cabrera[6]

[1] Andalusian School of Public Health, Spain
e-mail: luiscastro193@gmail.com

[2] University of Granada, Spain
e-mail: mrueda@ugr.es

[3] Andalusian School of Public Health, Spain
e-mail: carmen.sanchezcantalejo.easp@juntadeandalucia.es

[4] University of Granada, Spain
e-mail: rferri@ugr.es

[5] University of Granada, Spain
e-mail: jorgehcal@ugr.es

[6] Andalusian School of Public Health, Spain
e-mail: andres.cabrera.easp@juntadeandalucia.es

Abstract

The rapid evolution of COVID-19 required tools to perform a fast and efficient evaluation of the situation. Healthcare services around the world relied on surveys to fill those information gaps regarding the social, economic and health impacts of the disease. Those surveys enabled decision makers to take preventive and protective measures, especially among the most vulnerable population.

The Health Care and Social Survey (ESSOC, by its initials in Spanish), whose protocol was developed in Sánchez-Cantalejo et al. (2021), is a research project that arose from the necessity of information previously described. Its objective is to provide reliable and specific information about the impact of COVID-19 in Andalusia over time in several variables that may be useful in decision-making for controlling the consequences of the pandemic. The survey covers the evolution of health, socioeconomic, psychosocial, behavioral, labor, environmental and clinical variables, both in the general population and among the most vulnerable ones. The study integrates data from various sources, based on surveys, population statistics and official registers (clinical, epidemiological and environmental). Regarding the surveys, they have followed an overlapping panel design, using different adjustment methods to account for non-response and attrition. These methods, including the role of Machine Learning algorithms in their application, will be developed in this session.

Keywords: COVID-19, panel surveys, overlapping design, non-response, Machine Learning

References

Sánchez-Cantalejo, C., Rueda, M. M., Saez, M., Enrique, I., Ferri-García, R., De La Fuente, M., Castro-Martín, L., … & Cabrera-León, A. (2021) Impact of COVID-19 on the health of the general and more vulnerable population and its determinants: Health care and social survey-ESSOC, study protocol. *International Journal of Environmental Research and Public Health*, **18**, 8120.