

How to deal with sampling weights in Pandemic Times?

Campos, Pedro^{1,2} and Cabral, Mafalda¹

¹Statistics Portugal

²FEP, Universidade do Porto and INESC TEC

Abstract

In official statistics, probabilistic sampling designs are commonly used, namely in social surveys. For estimation purposes, these methods use sampling weights that can be thought as the number of population units each sample unit is representative of. For the total Y of a population, the corresponding benchmark Horvitz-Thompson estimator is given by

$\hat{Y}_\pi = \sum_{k \in U} \frac{1}{\pi_k} y_k = \sum_{k \in S} d_k y_k$, where π is the selection probability of unit k and d_k is the initial

weight, i.e. the inverse of the selection probability. In several social surveys, final weights are also computed, and reflect different adjustments that occur after the collection of the sample units, such as frame integration, nonresponse and calibration. However, the final weights often differ from the initial weights.

In this paper we address the effect of the pandemics on the difference between initial and final weights, due to non-response motivated by changes in the mode interview. The substitution of face-to-face interviews (CAPI) by telephone interviews (CATI) during the COVID-19 crisis led to new solutions, such as smoothing the initial weights and recalculating the weights based on a predefinition of limits between the initial and final weights. These procedures allow the estimation of quantities in a coherent way, in rotation-scheme surveys such as Labour Force Survey and the Information and knowledge society - household survey.