

WEIGHTING A NON-PROBABILITY WEB SAMPLE SURVEY CARRIED OUT DURING THE PANDEMIC

Marcelo Trindade Pitta
Brazilian Network Information Center (NIC.br)
Pedro Luis do Nascimento Silva
National School of Statistical Sciences

Abstract

Research on the use of non-probability samples for producing reliable statistics is on the rise. Several approaches have been proposed to overcome the difficulties arising from such samples, such as response or selection bias. At the same time, official and public statistics producers face increasing challenges arising from the demand for more timely and more disaggregated data, while coping with decreasing resources and increasing non-response rates. The COVID-19 pandemic has added urgency to the need for alternatives that enable such producers to meet the demands for timely, relevant and accurate information, while facing the additional constraints on survey-taking imposed by social-distancing and other health protection protocols.

The Brazilian Network Information Center (NIC.br) is a non-profit organization responsible for the planning, evaluation, and monitoring the use of information and communication technologies (ICT) in Brazil. NIC.br has been taking yearly probabilistic face-to-face national surveys on the use of ICT by households (ICTHS). In 2020, after social distancing restrictions took effect, NIC.br moved to experiment alternative modes for collecting information during the pandemic.

One project was a study targeting Internet users collecting data via a self-administered web-based questionnaire and telephone interviewing. Recruitment of the sample for this survey was carried out by a private marketing research company. This sample was approached and invited to participate in the survey, with limited success (14% response rate). Both the recruitment strategy and subsequent volunteer style response pattern implied that the sample available for analysis displayed non-negligible representation bias compared to the target population for the study.

This paper presents the approach used to weight the sample, that attempted to mitigate selection and response biases. The approach relied on combining data from a traditional sample survey recently carried out (ICTHS 2019) to develop weights for units observed in the panel survey.

Keywords: survey, non-probability samples, Web panel, weighting.