Use of R at NIS Romania: Focus on Calibration in Household Surveys

Ana-Maria Ciuhu; Monica Rafu

National Institute of Statistics, Romania

The 11th International Conference
uRos2023
Bucharest, Romania
December 12-14, 2023

Summary

1. Context
   - Rsolution for calibration
   - ReGenesees Pros
2. Case study: Calibration of LFS
   - Calibration steps
   - Calibration model
   - Known population totals

Context

- Migration from SAS to R in Sampling Unit for Household Surveys:
  - sampling
  - calibration
  - variance estimation for quality reports
  - data cleaning
  - production of tables for dissemination
- Being a **challenge** for every user to involve himself and to exchange knowledge

Rsolution for calibration

- ReGenesees package (Zardetto 2015 [1] and Zardetto 2023 [2])
- current version 2.3 just released two weeks ago: functions that estimate the minimum sample size required to satisfy specific precision constraints

ReGenesees Pros

- Aggregate figures derived from estimators, achieved by the calibration, correspond with the known population totals.
- Tools for diagnosing of the calibration process, ie convergence to population totals and Unequal Weighting Effect.
- If required, a consistent trimming can be applied to the coefficients obtained by calibration.

Calibration steps

- Treating non-response and adjusting design weights
- Defining sample design
- Setup population template
- Compute optimal limits for weights
- Calibration and diagnosis
- Trimming (if needed)

Calibration model

$$\text{calmodel} = \text{Houdensity} + \text{Sex} * \text{Residarea} * \text{AgeGroup} - 1,$$

$$\text{partition} = \text{region}$$

where:

- houdensity: known total of households
- region: NUTS2

Known population totals

- Population of private households by sex, residential area, 5-years age groups by NUTS2 *(ref. dates: 1st of January and 1st of July)*
- Number of households by NUTS2 *(ref. date: 1st of January)*

References

D. Zardetto.
ReGenesees: an Advanced R System for Calibration, Estimation and Sampling Error Assessment in Complex Sample Surveys.
Journal of Official Statistics, 31(2), 177 - 203, 2015.
https://sciendo.com:443/article/10.1515/jos-2015-0013

D. Zardetto.
ReGenesees: R Evolved Generalized Software for Sampling Estimates and Errors in Surveys.
R package version 2.3, 2023.

Thank you!