

Methodological Guidelines EVS2017

App A2: Sampling

Related to the EVS 2017 integrated dataset ZA7500

- Current dataset version: v1-0-0, 2018-12-18 (first pre-release)
 - DOI: 10.4232/1.13090 (URL: <http://dx.doi.org/10.4232/1.13090>)
-

The Sampling Design Form (SDF) hereby attached constitutes the template of the form that country teams were required to fill in before the start for the fieldwork. The SDF had to be submitted to the Methodology Group for the review, and approval, of the sampling strategy adopted in the country.

It is divided into different sections:

- Target Population (as described in the Methodological Guidelines)
- Summary of the Sampling Design
- Sampling Design (domains, stages, etc.)
- Planning the Sample Size (to compute the net and gross sample size from the set effective sample size, the estimated design effect, and the expected response rate)
- Analytical Inclusion Probabilities (containing the formulas necessary to compute them).

Country: [Name of Country]
Program Director: [Responsible National Coordinator]
Survey Institute: [Name of the Organization entrusted with the fieldwork]
Sampling Director: [Responsible National Coordinator Sampling Expert]
Reference Survey: [If the current sampling design is recycled from a previous round, then which round is it?]
Date: DRAFT - 2016-11-24 11: 47: 59

Target Population

[The definition of the EVS target population is: All persons 18 years or older at the day of selection who are residents of the country within private households. There is no upper age limit and no exclusion regarding nationality, citizenship or language.

Potential under-coverage of certain groups, say because of language problems or sampling frame deficiencies, or for any other reason, must be stated explicitly. Sampling experts should ask explicitly if there are any groups or areas (e.g. islands or over-sea territories) that are deliberately excluded by the sampling frame, although they are part of the target population.

A number for the size of the target population should be given.]

Summary of the Sampling Design

Sampling Procedure

[A short verbal description of the sampling design should be given here. How many sampling domains, how many sampling stages, how is the selection done at each stage, was a stratification used, and what are the used sampling frames? The text should be self-explanatory so that it can be exported to other EVS documentation files]

Sampling Frame

[A short verbal description of the main sampling frame, i.e. the source of the addresses from which persons are selected]

Sampling Design

1 Sampling Stage

unit: [The definition of the primary sampling units (PSUs).]

fram: [The sampling frame of the PSUs.]

size: [The number of sampled PSUs.]

stra: [The stratification of the PSUs. If possible a table with the number of PSUs in each stratum should be appended.]

- allc: [The allocation method for PSU sample. If possible a table with the number of sampled PSUs in each stratum should be appended.]
- algo: [The selection algorithm of the PSUs. If possible the software used to implement the algorithm should be named.]

2 Sampling Stage

- unit: [The definition of the secondary sampling units (SSUs).]
- fram: [The sampling frame of the SSUs.]
- size: [The number of sampled SSUs.]
- calc: [The allocation method of the SSU sample to the selected PSUs. If possible a table with the number of SSUs in each PSU should be appended.]
- stra: [The stratification of the SSUs within PSUs. If possible a table with the number of SSUs in each stratum should be appended.]
- allc: [The allocation method for SSU sample. If possible a table with the number of sampled SSUs in each stratum should be appended.]
- algo: [The selection algorithm of the SSUs. If possible the software used to implement the algorithm should be named.]

1.X Sampling Stage

[The description of any further sampling stage is analog to the second.]

Planning the Sample Size

Parameters of the Planned Gross Sample Size

\bar{b} : Net Cluster Size	ρ : Intraclass Correlation Coefficient	$Deff_p$: Weighting Effect	rr: Response Rate	ri: Ineligible Rate	n_{eff} : Effective Sample Size
6	0.02	1.2	0.55	0.03	1200

Design Effect

$$Deff = Deff_c * Deff_p$$

$$Deff_c = 1 + (\bar{b} - 1) \cdot \rho$$

Determining Design Effect

Design Effect	
$Deff_c$	$= 1 + (6 - 1) \cdot 0.02$ $= 1.1^*$
$Deff_p$	$= 1.2^*$
$Deff$	$= Deff_p \cdot Deff_c$

$$= 1.32^*$$

*results have been rounded to 3 d.p.s

Gross Sample Size

Gross Sample Size		
n_{net}	=	$Deff \cdot n_{eff}$
	=	$1.32 \cdot 1200$
	=	1584^{**}
n_{gross}	=	$\frac{n_{net}}{rr \cdot (1 - ri)}$
	=	$\frac{1584}{0.55 \cdot (1 - 0.03)}$
	=	2969^{**}
n_{gross}	:=	2969^\dagger

† results have been rounded to nearest integer multiple of 1

** results have been rounded to 0 d.p.s

Analytical Inclusion Probabilities

[A mathematical formula must be given that can be used to calculate the inclusion probabilities of the sampling design] variables that are required for the sampling design]

For example, for a two-stage sampling design with a stratified sample at the first sampling stage and the selection of 9 persons from each sample PSU, this could look the following way:

1. $PROB1_{i|h} = n \frac{N_h N_{hi}}{N N_h} = n \frac{N_{hi}}{N}$
2. $PROB2_{k|ih} = \frac{9}{N_{hi}}$
3. $PROB_{k|jih} = PROB_{i|h} PROB_{j|ih}$

where, $PROB1_{i|h}$ is the inclusion frequency of the i -th PSU in the h -th stratum, $PROB2_{k|ih}$ the inclusion probability of all person within the i -th PSU in the h -th stratum, given its selection at the first sampling stage. Further is N total target population size, N_h target population size in the h -th stratum, n the number of sample PSUs, N_{hi} the target population size in the i -th PSU in the h -th stratum. Finally $PROB_{k|ih}$ is the inclusion probability of the k -th person in i -th PSU in the h -th stratum.

Appendix

[E.g. Tables related to stratification and allocation.]