

Weighting units of observations

Surveyanalysis.org

**Sampling weights, probability weights, sample balancing weights*

Frequency weights

Replication weights

Replicate weights (different from replication weights)

**Expansion weights, expansion factors, inflation factors, grossing-up factors*

Analytic weights

Importance weights

We can add

Design weights

Post-stratification weights

Sample-size weights

Population weights

Influence function weights

Panel weights, quota-sampling weights, accidental-sample weights

General assumptions

Complicated data set

- 1) A national survey (NS) is a collection of information obtained by using a sampling method that aims to represent an adult population of a given country or “national territory.”**
- 2) A national survey is conducted within a given survey project (SP) and within a given project wave (PW).**

Different type of weights

- **Survey weights = weights for respondents (IDs) within a given national survey (NS) = within-survey weights**
- **Weights for national surveys (NS) as units of analysis = between-survey weights**

Within survey weights

- **Survey weights are numbers associated with the respondents that specify the influence the various observations should have in analysis.**
- **Weights are used to adjust for unequal probabilities of selection in the survey sampling.**

Adjustment

- Ideally, a selected sample = miniature of the population it came from → Sample should be representative with respect to all variables measured in the survey.

Not usually the case.

- Weighting adjustment as commonly applied correction technique: assigns an adjustment weight to each survey respondent.

Persons in under-represented groups get a weight > 1 , and those in over-represented groups get a weight < 1 .

In the computation of means, totals and percentages: use values of the variables and the weighted values.

Weights in Analysis of Longitudinal (Panel) Data

- **Panel data sets could have two types of weights:**
 - **Cross-sectional weights for the first wave (if needed);**
 - **Panel weights for second and further waves**
- **Panel weights are computed from two components:**
 - 1. The weights from the first wave, or previous wave**
 - 2. A weight calculated to adjust for attrition between waves.**
- **Calculating the non-response (attrition) weight component:**
 - **Usually use logistic regression with response to the wave as outcome variable (0= no; 1=yes).**
 - **Predict probability of responding**
 - **Inverse of this probability is the attrition weight.**

The impact of weights on the standard errors: standard errors should be based on the actual N and not the weighted N.

- **SPSS treats weights incorrectly in inferential statistics.**
- **SVY procedures in Stata; also *pweight***

Normalization of weights:

- **Always set the weights so the N in the weighted data equals the N in the unweighted data. To calculate, multiply the weight by $(\text{Unweighted N}) / (\text{Weighted N})$**

Weights in regression analysis:

Reasonable choice of not using weights at all under the condition that one includes all the variables for the weights as independent variables.

Doing so you obtain unbiased regression estimates and unbiased standard errors.