Annual survey on constructions

TYPE

Sampling survey. The sampling method used is the single stratified random sampling. The enterprises included in the survey are stratified as following:

- By region - NUTS II
- By Class of NACE Rev.1.1 (4digit level of economic activity), within each administrative region
- By size class of the enterprise. In each of the major strata (major stratum = Geography x Economic Activity), the enterprises were stratified into L = 5 size classes, according to their size, determined by their annual turnover in the business register, as follows.

Size Class	**Turnover (in €)**
Class 1	1 – 89.999
Class 2	90.000 – 249.999
Class 3	250.000 – 1.499.999
Class 4	1.500.000 – 9.999.999
Class 5	10.000.000+

Let h be one of the final strata (Final stratum = Geography X Economic Activity X Size Class). The final strata that contain size classes with L = 4,5, are census strata (take-all).

Survey characteristics estimation

a. Symbols

Defining with index i the selection order of an enterprise from the sampling frame in the stratum h and symbolizing with the y one of the survey characteristics, we can define the following:

y_{hi} : The value of the survey characteristic y of the enterprise of order i in the stratum h

Y_h : The sum of the values of the characteristic y for all enterprises falling into the survey and belonging to the stratum h

Y : The sum of the values of the characteristic y for all enterprises under the survey of the stratum h. That is: $Y = \sum_h Y_{hi}$

N_h : The number of all enterprises falling into the survey and belonging to the stratum h

n_h : The sample size in the stratum h

m_h : The number of respondent units in the stratum h

r_h: Response rate in the stratum h ($r_h = \frac{m_h}{n_h}$)

w_{hi}: The extrapolation factor of the enterprise of order i belonging to the stratum h, ($w_{hi} = 1/(\text{Probability of selected unit } i \text{ in stratum } h) \cdot r^{-1} = \frac{N_h}{n_h} \cdot \frac{n_h}{m_h} = \frac{N_h}{m_h}$)

b. Estimation process

The estimation of Y_h and Y is given by the following formulas:

$$\widehat{Y}_h = \frac{N_h}{m_h} \sum_{i=1}^{m_h} y_{hi}$$

$$\widehat{Y} = \sum_h \widehat{Y}_h$$

c. Variance estimation

The variance estimation of $\widehat{Y}_h$ and $\widehat{Y}$ is given by:

$$V\left(\widehat{Y}_h\right) = \frac{N_h\left(N_h - m_h\right)}{m_h} S_h^2,$$

Where:

$$S_h^2 = \frac{1}{m_h - 1}\left[\sum_{i=1}^{m_h} y_{hi}^2 - \frac{\left(\sum_{i=1}^{m_h} y_{hi}\right)^2}{m_h}\right],$$

$$V\left(\widehat{Y}\right) = \sum_h V\left(\widehat{Y}_h\right)$$

The coefficient of variation (%) of total estimation $\widehat{Y}$ is given by:

$$CV\left(\widehat{Y}\right) = \frac{\sqrt{V\left(\widehat{Y}\right)}}{\widehat{Y}} * 100$$