

Sample Design for Ukrainian business Survey

Oksana Honchar

Scientific and Technical Complex of Statistical Research,
State Academy of Statistics, Accounting and Auditing of the State Statistics Committee of Ukraine
e-mail: ohonchar@list.ru

Abstract

In this paper we consider construction of sample design for semi-annual Ukrainian business survey in 2007.

1 Introduction

Every year the State Statistics Committee of Ukraine organizes annual and semi-annual enterprise structural surveys. Since 2002 the semi-annual enterprise structural survey has been conducted using a sample. The sample is constructed in June each year and the final results of this survey are obtained in December. In 2008 the Scientific and Technical Complex of Statistical Research has begun to develop methodological background for implementation of sample into the annual enterprise structural survey. This report discusses the design of the semi-annual enterprise sample survey.

2 Population definition

The sample frame is formed with data base of the annual structural survey of small enterprises. A small enterprise is one that has 50 or less permanent employees and a turnover less than 500 thousands Euros. We corrected the frame taking consideration dead and new enterprises according to the Administrative Enterprise Register of Ukraine.

In May 2007 the corrected frame of enterprises consisted of 308'974 units. These small enterprises were separated into three parts:

- active enterprises (204'722 units);
- zero enterprises, that are enterprises with zero turnover and number of employees (85'779 units);
- new enterprises (18'473 units).

3 Sample design

Active enterprises

Array of active enterprises is stratified according to the basic economic activity (51 groups), the region and the size of enterprise (number of employees). We stratified the enterprises by the region into 6 groups:

1. East (Dnipropetrovsk, Donetsk, Zaporizhzhya, Luhansk, Kharkiv regions);
2. North (Kyiv, Zhytomyr, Poltava, Sumy and Chernihiv regions);
3. Center (Vinnitsya, Kirovograd, Khmelnytskyi and Cherkasy regions);
4. West (Volyn, Zakarpattia, Ivano-Frankivsk, Lviv, Rivne, Ternopil and Chernivtsi regions);
5. South (Autonomous Republic of Crimea, Mykolayiv, Odessa, Kherson regions and Sevastopol city);
6. Kyiv city.

The stratification by the number of employees resulted in four groups:

- 1) enterprises with the number of employees less than 4;
- 2) from 5 to 9 employees;
- 3) from 10 to 19 employees;
- 4) more than 20 employees.

As a result of stratification by these three criteria 1'114 strata were obtained.

The next step of sample design construction was to unify some strata in order to decrease a number of small strata. To save the homogeneity we determined criteria for unification of strata. The strata can be unified according to the similar economic activities, the regional groups and groups by number of employees.

For unification similar region groups in each economical activity group we used Independent - Samples t-Test. In addition to the mathematical criterion, we examined the expediency of such unification. For example, can we unify fishery enterprises from the southern and the northern regional groups? Certainly not. In each economical activity and regional group we have also unified strata that had similar mean of the turnover but we also examined them in order to have homogeneity in the stratum.

After unification we obtained 1'003 strata.

Zero enterprises

Group of zero enterprises consisted of two parts:

- 1) enterprises with zero turnover in 2006 (35'827 units – 41,8%);

2) enterprises with zero turnover and number of employees during two years (49'952 units – 58,%).

From the first part we selected 5% (1'792 units). The second one included the lifeless enterprises that are not observed.

New enterprises

The population of established enterprises during January-April 2007 consisted of 18473 units. The group of new enterprises was stratified by the two criteria: the basic economic activity and the region, because we do not know the sizes of this enterprises (number of employees). Thus in a new enterprise population we obtained 270 strata. The sample size in this population is equal 10%.

4 Atypical enterprises detection

For detection of atypical enterprises statistical tools are used (it allow quantitative to detect them) and logical analysis that consider qualitative factors of detection and considering of atypical enterprises. The statistical tools include formulating statistical hypothesis (variance method), expert judgment and diagram method (Box-plot).

Variance method bases on estimate of relative contribution of each enterprise to variance of analyzed characteristic in stratum. First enterprises are ranked in decreasing order of turnover. Then we calculate relative contribution of enterprise to variance

$$I_n(i) = \frac{V(Y) - \left(\frac{n-1}{n}\right)V(Y - y_i)}{V(Y)},$$

where $V(Y - y_i)$ is variance of turnover for all units in stratum except i ;

$V(Y)$ is total variance of turnover in stratum, $i = \overline{1, n}$, n – number of units in stratum.

Inflection $I_n(i)$ is point after which all units consider as atypical.

Expert judgment used after variance method for ground of decision about atypical enterprise. Expert considers atypical enterprises detected by variance method in dynamics, space, and take final decision about each atypical enterprise. Atypical enterprises and small strata (which have less than fifteen units) are 100% included in the sample.

5 Sampling

Thus sample frame is structured:

- main array of enterprises is stratified by basic economic activity, region and size of enterprise;
- array of atypical enterprises and small strata;
- enterprises with zero turnover in 2006 is stratified by basic economic activity, region and size of enterprise;
- enterprises with zero turnover and number of employees during two years is stratified by basic economic activity and region;
- new enterprises is stratified by basic economic activity and region.

Stratified random sample with Neyman allocation is applied for main array of enterprises. Systematic sampling is used within each stratum.

Table 1. Sample size for semi-annual Ukrainian business survey in 2007

Array	Size, units	Sampling, units	Sampling, %
Main	202046	40945	20
Atypical enterprises and small strata	2676	2676	100
Zero1	35827	1792	5
Zero 2 (lifeless enterprises)	49952	0	0
New enterprises	18473	1847	10

For proportional allocation sample in regions we used latent stratification. Before realization of systematic sampling enterprises within each stratum were ranked by regions.

References

1. Cochran, W. G. (1977) *Sampling Techniques*. John Wiley & Sons, New York.
2. Grun-Rehomme M., Vasechko O. (2004) Outliers Detection in Ukrainian Business Statistics/ *European Conference on Quality and Methodology in Official Statistics*. – Mainz, Germany.
3. Särndal, C.-E., B. Swensson and J. Wretman (1992) *Model Assisted Survey Sampling*. New York.