

Theoretical Approaches Concerning Statistical Survey

Podașcă Raluca
Petroleum-Gas University of Ploiești
raluca.podasca@yahoo.com

Abstract

In the current modern age, socio-statistical surveys are particularly important in researching social life. Thus, individual data obtained through total or partial observations are used. The main issue proposed to solve by selective research is to determine a partial collectivity from the total statistical community with features that characterize the entire statistical population. Partial collectivity is an element of statistical research through which are researched the features of the entire collectivity. Statistical survey represents a method of characterizing the statistical population based on the research of a subset named statistical sample, which is obtained by various methods from general population. The ultimate goal of the survey is not the knowledge and research of the sample, but the knowledge of the whole statistical population.

Key words: statistical survey, statistical investigation, survey method

J.E.L. Classification: C10, J10

1. Introduction

Statistical survey method has a large-scale application, with public and private institutions that organize various research samples. The results of the survey are made public in the press or by other methods of communication. Results of surveys conducted in the economic, social, political or other areas of national interest confirm their importance and usefulness. The survey results will extend to the general population through specific methods, complying the principles of probability, mathematical statistics and the law of large numbers.

2. The survey- method of statistical investigation

With many advantages, statistical survey applies to many areas of social and economic life, so research surveys are given particular importance both in theory and in practice.

Surveys research has two stages:

- Observing or statistical description- consists in extraction of the sample and calculation of statistical values that define it: average, variance, standard deviation, relative indicators, coefficient of variation, coefficients of asymmetry;
- Indicators obtained by processing the sample data extend to the entire population in order to characterize it in terms of statistics.

Survey research use pair concepts such as:

- General collectivity and selection collectivity;
- Average of general collectivity and average of selection collectivity;
- Calculated statistical value and estimated value.

The most important advantages of the survey research are:

- High quality of the results;
- Reduced working time;
- Saving money;
- The research sample comprises a greater number of features than the total observation.

In industry, the survey method is used to:

- Determining productivity in natural units;
- During the technological processes to prevent spoilage;
- Research of the product quality;
- Research of the causes that conduct to employment fluctuations;
- Study the importance of investments.

In trade, the survey is used to:

- Testing the acceptance of new products;
- Studying the demand for goods depending on the influence factors;
- Estimating specific trends for the sales activity;
- Studying the provision of services to the population.

General collectivity or statistical population represents all simple or complex units forming the phenomenon under research. The statistical survey begins by strict definition of the general community in time and space, in terms of content and form of organization. General collectivity consists of statistical units, which can be simple or complex. The volume of the general collectivity, from which are extracted units to create the sample is denoted by "N" for simple units, by "M" for alternative variable, by "R" for complex units.

The selection collectivity (sample) represents a part or a number of elements of the general collectivity from which data is processed in order to extend these results to the entire community. From a statistical population can be extracted more samples, which differ in structure and volume. The indicators that characterize each sample can be considered random variables that can determine the distributions of frequencies. The volume of selection collectivity is denoted by "n" when the sample consists of simple units, by "r" when it refers to complex units, by "m" for the alternative feature (Anghelache et al, 2016).

Table no. 1. Formulas for calculating the average and variance for general collectivity and selection collectivity

Notion	Indicator	Computing relations	
		Non alternative feature	Alternative feature
Selection collectivity	Average	$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$	$w = \frac{m}{n}$
	Variance	$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n}$	$\sigma^2 = w \cdot (1 - w)$
General collectivity	Average	$\bar{x}_0 = \frac{\sum_{i=1}^n x_i}{N}$	$p = \frac{M}{N}$
	Variance	$\sigma_0^2 = \frac{\sum_{i=1}^n (x_i - \bar{x}_0)^2}{N}$	$\sigma^2 = p \cdot (1 - p)$

Source: (adapted from Mendenhall et al, 2012)

Selective research purpose is achieved if the preparation should be carried out according to a plan that will produce the most accurate information. It should include:

1) Objectives- the survey beneficiary should clearly state the objective of the proposed project. The statistical population should be properly defined and the purpose and use of results should be clearly stated. It is very important the delimitation in time and space of the general collectivity by identifying all individual cases in which these phenomena manifest and verify the homogeneity degree. There are organized several successive surveys to verify the representativeness of the average and dispersion characteristics after which the sample is made.

2) The sampling frame- any systematization units allowing random selection of units that will contain the sample are called a random basis and must contain the complete population, without double counting.

3) The sampling units might be human population (cities, households, individuals) in agriculture (farms), industry (companies, batch production).

4) Extract of the sample – to determine its size, the type and selection process, of the indicators.

5) Methods of obtaining data – by direct path, by e-mail, by telephone, by internet.

6) Tracking the survey labor in terms of field operators.

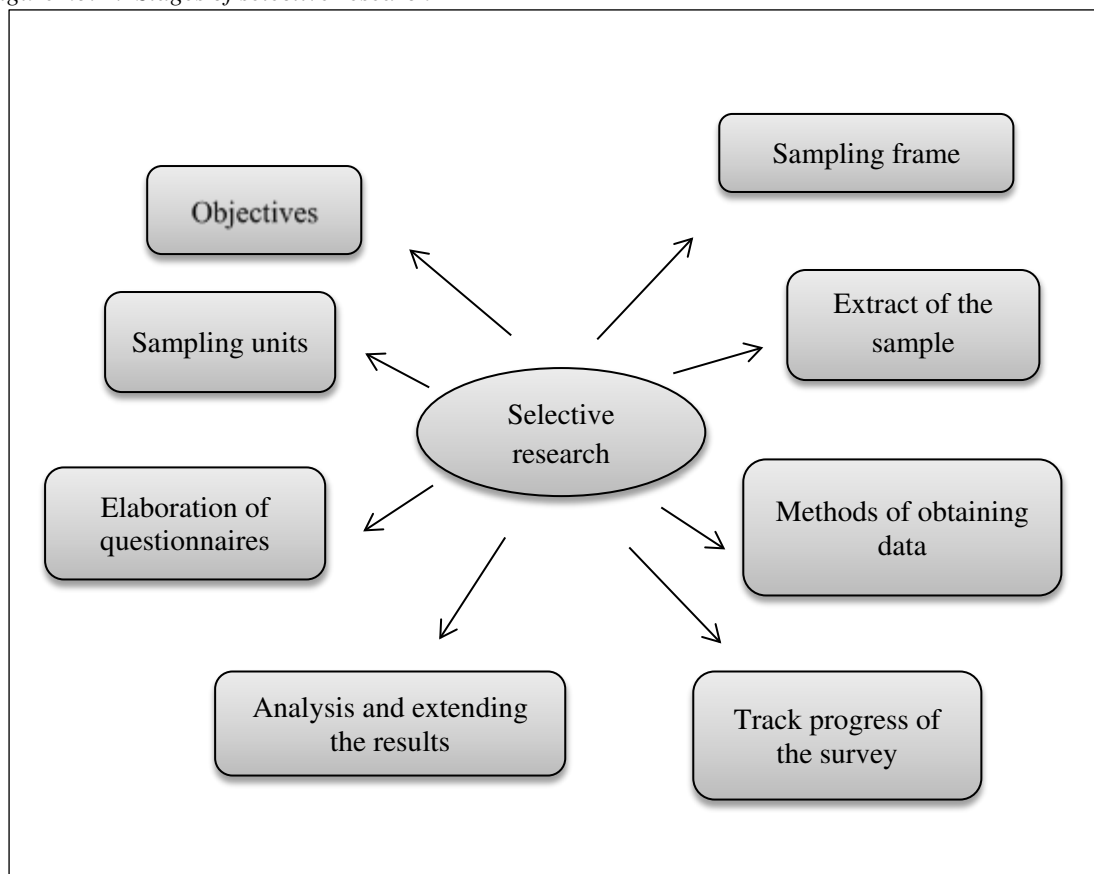
7) The pilot survey – preliminary data is recorded to assist the formation of the new sample, to clarify various issues raised.

8) Elaboration of questionnaires.

9) Establishment the methodology and organization of data processing.

10) Analysis, interpretation and presentation of data: suitable processes are selected to verify the significance of the selected indicators and extrapolation of the results to the general collectivity.

Figure no. 1. Stages of selective research



Source : (adapted from Walpole et al, 2011)

The fundamental principle that should be considered in a selective survey is structured sampling. It is required that the selection collectivity structure should be as close as the general community structure. Theoretically the selection structure and the general collectivity structure can be identically. However, in practice, there will always remain a difference between the two structures. The closer they are together the results of selective research will be expanded with more precision on the general community.

Example:

Suppose 1000 employees perform the same piece to a machine. Distribution of employees by daily production is presented as follows:

Table no.2 Distribution of employees by daily production

Groups after the production of parts (pieces)	General collectivity – employees		Selection collectivity – employees	
	Absolute amount	Share (%)	Absolute amount	Share (%)
60-70	140	14	7	14
70-80	660	66	33	66
80-90	180	18	9	18
90-100	20	2	10	2
Sum	1000	100	50	100

Source: (Podasca, 2017)

Sample average (\bar{x}) is representative for general average (\bar{x}_0) if is satisfied the following relation:

$$\frac{\bar{x} - \bar{x}_0}{\bar{x}_0} \cdot 100 \leq 5\%$$

The difference $\bar{x} - \bar{x}_0$ is called the margin of error or sampling error and occurs due to random character of extraction units. These errors may be:

- Representativeness systematic errors that are the result of non-compliance with the principles for achieving survey, respectively equal opportunity for each element in the sample; these errors can be excluded by proper composition of samples;
- Representative random errors that cannot be avoided.

3. Conclusions

To achieve the establishment of optimal proportions between sample groups and general collectivity, it is necessary to know the distribution of total population. The essential condition to guarantee a selection community is representativeness. This represents the characteristic of the sample to faithfully reproduce the essential characteristics of the total community, to be faithful to its elements. In case of random sampling, the sampling units are randomly selected, so each unit has the same chance of being selected. It is based on the draw principle. The utility of random sampling is based on the fact that its results can be appreciable in probabilistic terms. This method is indicated when the population units are small and there is no significant differences between the various units of the population size.

4. References

- Anghelache. C., Anghel, M. G., 2016. *Fundamental of Economic Statistics. Theoretical Concepts and Case Studies*. Bucharest: Economical Publishing House.
- Mendenhall, W., Beaver, R., Beaver, B., 2012. *Introduction to Probability and Statistics*. Boston: Duxbury Press.
- Podasca, R., 2017. *Economic Statistics: Theory and Applications*. Bucharest: Printech Publishing House.
- Walpole, R., Myers, R., Myers, S., Ye, K., 2011. *Probability and Statistics for Engineers and Scientists*. New Jersey: Pearson Education