BRAND. Broad Research in Accounting, Negotiation, and Distribution
e-ISSN: 2067-8177 | p-ISSN: 2068-0473
Covered in: EBSCO; RePEc (Ideas, EconPapers); Google Scholar; Zenodo; Academia.edu;
2020, Volume 11, Issue 1, pages: 32-36.
Submitted: February 20th, 2020 | *Accepted for publication:* May 5th, 2020

Theoretical Aspects Related to Statistical Research Using a Questionnaire

Monica Muntean (Chitescu)
Ioan Slavici University of Timișoara
Professor Doctor Aurel Păunescu
Podeanu Street, Timișoara, 307200,
Romania.

Abstract: *Statistical research relies on various data collection methods, among which surveys play a crucial role in inferential statistics. Unlike total observation, which includes the entire population, survey-based research uses a sample to estimate population parameters and test hypotheses. This paper explores the methodological foundations of survey research, its relationship with statistical populations and samples, and the principles of parameter estimation, including point estimation and confidence intervals. Special attention is given to the properties of estimators, such as unbiasedness, efficiency, and consistency, which influence the accuracy of statistical inference.*

Keywords: *survey research; statistical inference; population; sample; estimation; confidence interval; statistical parameters.*

How to cite: Muntean, M. (2020). Theoretical aspects related to statistical research using a questionnaire. *BRAND. Broad Research in Accounting, Negotiation, and Distribution. 11*(1), 32-36. https://doi.org/10.70594/brand/11.1/5

1. Introduction

In classical statistics, data collection can be performed through total or partial observation. While total observation ensures complete coverage, it is often impractical due to time, cost, and logistical constraints. Instead, partial observation, or survey research, relies on extracting a sample from the population to infer characteristics about the whole. This approach falls within inferential statistics and allows researchers to make predictions and validate hypotheses efficiently.

The fundamental goal of survey research is to estimate the characteristics of a statistical population using a subset of its units. By applying probability theory, researchers can infer unknown population parameters based on sample data. This paper discusses the conceptual framework of survey-based research, highlighting key statistical properties that ensure the reliability of estimation procedures.

2. Specific Concepts of Research Using Surveys

Classical statistics include total observation, which is a method of obtaining data. However, to acquire more concise information, it is necessary to apply the method of partial observation, which is part of inferential statistics (survey research) and consists of using a sample from the statistical population.

Survey research aims to understand the statistical population based on observations of multiple samples or just one. It is characterized by estimating the characteristics of statistical distributions to verify the desired hypothesis and make potential statistical predictions.

The purpose of partial research is to estimate the results of data processing obtained from a sample using probability theory on the unknown parameters of the studied population.

In survey research, two characteristics are studied: the statistical population to be examined and the sample used to obtain the population parameters. These must have the same methodological content.

The statistical population consists of all units subjected to the research phenomenon, where the volume is denoted by N, and the alternative variable is denoted by M.

The sample (partial community, selection, specimen, etc.) is the unit extracted from the statistical population to collect data and generalize the results obtained from processing them. The volume of this sample is denoted by n.

3. Relationships between Partial and Total Research

- Partial Research: n<Nn
- Total Research: n=Nn

In other words, from a statistical population (a specific population, not just any), multiple samples of different volumes and structures can be drawn.

Mean and variance are specific values for the sample and statistical population, and as such, they are denoted and calculated differently, depending on the reference level. Their determination is illustrated in the table below:

Parametru statistic	*La nivel de populaţie statistică*	*La nivel de eşantion (selecţie)*
Media aritmetică	$m = \frac{\sum_{i=1}^{N} x_i}{N}$	$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$
Varianța	$\sigma^2 = \frac{\sum_{i=1}^{N} (x_i - m)^2}{N}$	$s^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}$

Figure 1. Mean and Variance Formula (Mnerie, et. al., 2012)

4. Estimation as a Statistical Parameter

Estimation is the process through which a parameter is estimated with a certain probability, either through point estimation (a single value) or confidence interval estimation, where the unknown significance of the population parameters is evaluated based on the recorded data from a sample extracted from it. (UMFCV.RO, 2020)

- Point estimation involves calculating a single, admissible value for the sought parameter based on data from the sample units. (Spătaru, 2004)
- Confidence interval estimation defines the limits of a possible range in which the true value of the population parameter is expected to fall.

The estimator represents a statistical measure used to evaluate an unknown population parameter. This results from the process of statistical inference or induction and has an associated probability that determines its accuracy or confidence level.

Estimation illustrates the effectiveness of an estimator for the examined parameter, based on a sample. It is defined so that, for various drawn samples, the estimated values should closely approximate the unknown population parameter as much as possible.

The specific value of the estimator and the estimation is determined by three properties:

1. Unbiasedness
2. Convergence
3. Efficiency

An estimator is considered unbiased if its mean is equal to the true value of the corresponding parameter. The estimator's value is closer to the true value of the population parameter as the sample size increases ($n \rightarrow N$).

An estimator is biased if its mean is not consistently equal to the true value of the corresponding parameter but has some deviation or difference from the true value.

A convergent estimator is one whose variance tends to zero as the sample size approaches the total population size ($n \rightarrow N$).

An *estimator is absolutely correct* if it is both unbiased and sufficient.

An *estimator is correct* if it is biased but sufficient.

An *estimator is efficient* if it has the smallest variance compared to any other estimator computed for the same sample of size *n*.

5. Rigorous Estimation of Mean and Variance

Let *m* be the unknown population mean. The sample-based estimator is denoted as $\bar{X}$. In this case, it can be demonstrated that $\bar{X}$ is an unbiased and sufficient estimator of *m*. Thus, the sample mean is considered an absolutely correct estimator of the population mean (*m*).

If the unknown variance of the population is denoted by σ^2, and its estimator, calculated from the sample, is marked as S^2, it can be proven that S^2 is an unbiased and sufficient estimator of σ^2.

Therefore, the sample variance (S^2) is a correct estimator of the population variance (σ^2). Regarding the point estimation of population parameters, it is effective only for very large samples. In economic procedures, parameter evaluation is preferably based on an interval of values (confidence interval), which can also be estimated from smaller samples.

6. Estimation of Variance and Mean Using the Confidence Interval

Confidence interval estimation involves determining a range of values, with a certain probability, within which the unknown population parameter is expected to fall. (Cărbunaru-Băcescu, 2009)

The probability that the population parameter is within the confidence interval is denoted by *p* and is called the confidence level.

Conversely, the probability that the unknown population parameter falls outside the confidence interval is denoted by α, calculated as α=1−p, and is referred to as the *significance level*.

Typically, the confidence interval limits are defined as:

- Lower limit: L_i
- Upper limit: L_s

The probability that an unknown population parameter, denoted by θ, lies within the confidence interval is given by:

$$P(L_i \leq \theta \leq L_s) = 1 - \alpha$$

Figure 2. Probability Calculation Formula

7. Confidence Interval Estimation of Population Mean (m)

In economic applications, this process is commonly applied as follows:

If the sample size is less than 30 (n≤30n), and the population variance (σ^2) is variable, the confidence interval for the mean is estimated using the Student's distribution with v=n−1 degrees of freedom:

$$P\left(\bar{x} - t \cdot \frac{s}{\sqrt{n}} \leq m \leq \bar{x} + t \cdot \frac{s}{\sqrt{n}} \right) = 1 - \alpha$$

Figure 3. Mean Calculation Formula

where:

$\bar{X}$ represents the sample mean.

s is the sample standard deviation estimator.

t is the tabular value from the Student's t-distribution, depending on α/2 and n−1 degrees of freedom.

8. Confidence Interval Estimation of Population Variance (σ²)

The estimation of variance (σ^2) through the confidence interval is determined using the chi-square (X^2) distribution:

$$P\left(\frac{(n-1)\cdot s^2}{\chi^2_{\frac{\alpha}{2}}} \leq \sigma^2 \leq \frac{(n-1)\cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}} \right) = 1 - \alpha$$

Figure 4. Variance Estimation Formula

where:

- *n* is the sample size.
- s^2 is the sample variance estimator, calculated from sample data.
- σ^2 is the unknown population variance.

- X^2 și $X^2_{1-\frac{\alpha}{2}}$
  represents tabular values of the X^2 distribution for n−1 degrees of freedom.

9. Estimation of Standard Deviation

The confidence interval for standard deviation is determined by taking the square root of the variance confidence interval limits:

$$P\left(\sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{\frac{\alpha}{2}}}} \leq \sigma \leq \sqrt{\frac{(n-1)\cdot s^2}{\chi^2_{1-\frac{\alpha}{2}}}}\right) = 1-\alpha$$

Figure 5. Standard Deviation Formula

10. Conclusion

Survey research serves as a fundamental tool in statistical inference, allowing researchers to derive insights about a population from a sample. The accuracy of estimations depends on the properties of the chosen estimators, with unbiasedness, efficiency, and consistency being critical. Confidence intervals enhance reliability by accounting for estimation uncertainty. By applying rigorous statistical methods, researchers can ensure robust conclusions and meaningful predictions based on survey data.

References

Cărbunaru-Băcescu A. (2009). *Statistică-Bazele statisticii*, Editura Universitară, Bucureşti.

Mnerie, G. V., Dorneanu, L., Mnerie, D., Nani, V., Mariş, S., Saizescu, C., & Ularu, N. (2012). *Elemente de matematici aplicate, statistică şi econometrie*, Editura Fundaţiei pentru Cultură şi Învăţământ "Ioan Slavici", Timişoara, Editura Eurostampa, Timişoara.

Petcu, N. (2005). *Statistica in turism.Teorie si aplicatii*, Editura ALBASTRA.

Spătaru L. (2004). *Analiză economico-financiară*, Editura Economică, Bucureşti.

University of Medicine, Pharmacy, Science, and Technology of Târgu Mureș. (2020). *TS_08*. Retrieved from https://www.umfcv.ro/files/t/s/TS_08.pdf