Natasha ANGELOSKA-GALEVSKA

UDK: 303.725.3
Review article

STRATEGIES AND METHODOLOGICAL CONSIDERATIONS IN CHOOSING A RESEARCH SAMPLE

Abstract:

The text defines the fundamental terms relevant to the process of sampling, elaborates the characteristics of individual samples, selection strategies, and the key methodological issues and dilemmas of the researchers during the procedures of sampling.

The selection of a sample is an essential aspect when planning the methodology of empirical research and has a key role in ensuring the reliability and validity of research results. Researchers should carefully consider the characteristics of the population and the specific research objectives and accordingly choose the sampling strategy.

Representative samples minimize bias and provide a higher level of confidence when generalizing results to a larger population. Non-representative samples are valuable when the selection of a representative sample is impractical or impossible. They involve different forms of bias but offer flexibility, cost-effectiveness and practicality. Non-representative samples are applicable in qualitative studies and in the research with populations that are hard to reach. However, it is important to be aware of their limitations and the justification for generalization findings outside the specific sample.

Regardless of the used strategy, researchers must strive for transparency, clearly documenting the unit selection procedure and stress the potential limitations as key elements to ensure the reliability and validity of the research results.

Keywords: *methodology, representative samples, non-representative samples*

Introduction

One of the main methodological issues when planning scientific research is choosing of a sample. At this stage, the key questions that the researcher should think about and decide on are the following: what sample should be selected for the research, what strategy should be applied to select the subjects in the separate samples and what should be the size of the sample in order to perform valid and reliable conclusions.

In the following text, the key terms relevant to working with samples are explained, the characteristics of individual samples and their selection strategies are considered, and the key methodological issues and dilemmas of the researchers in the procedure of working with samples are elaborated.

1. Defining key terms related to the process of sampling

The key terms encountered in the sampling phase are population, sample, sampling frame, sampling rate or fraction, sampling interval and sampling error.

A population is a basic set of units relevant to the subject of research that possess at least one common feature. The population refers to the entire group of individuals, cases, or elements that we want to study and draw conclusions about. The population can be finite or infinite depends on the possibility units to be counted and census to be made of all units. Finite population has a known and limited number of members or cases so it can be fully enumerated and observed, meaning that every member of the population is identifiable and accessible for study. In practical research, many populations are finite. For example, a finite population could be all students enrolled in a particular school, all employees in a company, or all households in a specific neighbourhood. These populations have a definite size that can be determined, and it is possible to study every individual or case within the population if resources allow.

With the infinite or unlimited population, it is not possible to enumerate all potential units or subjects and in such a case a selection frame is made from the existing units. *The sampling frame*actually represents a list of all the units of the population based on which the sample is selected.

A sample is the studied part of the population on the basis of which we draw conclusions about the phenomenon. If we understand the population as a set that includes all the relevant subjects for the research, the sample is its part or subset. In the Macedonian professional terminology, the term sample is used synonymously with the term *mostra*.

The calculations obtained from the population are called parameters and are denoted by letters of the Greek alphabet, and the calculations obtained from the sample are called statistics and are usually denoted by Latin letters. For example, the Greek sigma is used for the standard deviation calculated from the population, while the standard deviation calculated from a sample is denoted by the Latin s.

There are several reasons why in some cases we cannot work with the entire population of the research. Most often, the population is large and inaccessible as a whole, and working with all units of the population would cost a lot of time and resources. That is why the research is carried out on a specific sample, and with the help of statistics we estimate the parameters of the population. In doing so, calculations from descriptive statistics (arithmetic mean, standard deviation and other measures of central tendency and variability) are used, while estimates and conclusions regarding the hypotheses are performed with procedures from inferential statistics (testing the significance of differences, correlation, non-parametric tests).

The sample should be large enough to generalize the results with a high degree of confidence to the entire population and to reduce the error of inference due to sampling. **Sampling error** actually refers to the discrepancy or difference between the characteristics or statistics of a sample and the corresponding values of the population from which the sample is drawn. Sampling error can be influenced by various factors, including the size of the sample, the sampling method used, and the variability or diversity within the population.

2. Types of samples and sampling procedures

In the research project, and accordingly in the research report, the population of the research should be defined and the characteristics and size of the sample on which the research is carried out should be determined. This means that it should be stated what kind ofsample is used in research (random, systematic, stratified, cluster, convenience sample, judgement sample, chain or snow ball sample) and to describe the way in which the units will be selected. It is also recommended to include tables and/or graphics to clarify the structure and demographic characteristics of the sample.

The specified characteristics should be explained for each separate category of respondents that will be covered by the research, for example, sample of teachers, students, parents, sample of documents for content analysis, time sample for observation, etc.

In quantitative research that includes hypothesis testing, representative samples should be used to provide a reliable empirical basis for generalization of the obtained results. The selection of a sample is not a simple process and there are numerous methodological issues related to it, which refer to the characteristics of the certain types of samples and the validity and reliability of the derived conclusions. In the following text, we will talk about the different types of samples, their advantages and weaknesses, their suitability in relation to the research objectives and the reliability of conclusions.

2.1. Representative samples

2.1.1. Random sample

The random sample is of great importance in statistics because it usually provides a normal distribution of cases, which is the basis for applying a large number of statistical procedures. The random sample is the basis for testing hypotheses and making generalizations of the obtained conclusions.

The simple random sample has the following advantages:

- It is unbiased
- It facilitates the application of various statistical tests and techniques
- Each unit of the population has an equal probability of being selected in the sample
- It is chosen in a simple way without prior study of the structure
- It saves time and money
- Theoretically it is the most elaborate
- It has the best features
- It is best suited for statistical inference
- Enables comparisons across different studies or populations.

2.1.1.1.Sampling method

When forming the sample, if it is a finite set, the sampling fraction and the sampling interval of the selection are determined.

The sampling fraction, also known as the selection rate or sampling rate, is a term used in sampling to describe the proportion of elements or individuals selected from a population to be included in a sample.

The selection rate is typically calculated by dividing the sample size (the number of elements or individuals in the sample - n) by the population size (the total number of elements or individuals in the population - N). It is expressed as a decimal or percentage or expressed as a proportion $\frac{n}{N}$. For example, if from a population consisting of 800 students, we decide to examine 200 cases, then the proportion will be $\frac{200}{800} = 0,25$.

Expressed as a percentage, this means that 25 % cases from the population are included in a sample or, in other words, the probability of each unit being selected in the sample is 0.25 or 25 %.

The sampling fraction depends on the precise inference we strive to achieve, but in general it can be said that the higher selection rate provides more precise estimates and reduce the sampling errors.

The sampling interval is the reciprocal value of the sampling fraction and is calculated by dividing the population size (N) by the desired sample size (n) or $\frac{N}{n}$. Using the data from the previous example, this would be $\frac{800}{200} = 4$, which means that every fourth unit in the population would be selected to form the sample.

Sampling interval can only be applied to a finite population. If the population is infinite, then we use the sampling frameconsists of those units that currently exist as a basis of selection.
The units are assigned with serial numbers and afterwards selected by lottery. One possible way is the serial numbers of the units to be written on slips of paper which will then be randomly drawn. The second and more practical way is to use a table of random numbers (Fisher, Kendal or other). Numbers in random number tables can be read in any direction, left, right, up, down or diagonally.

Figure 1. Example of random numbers for selecting the units in the sample

73735	45963	78134	63873
02965	58303	90708	20025
98859	23851	27965	62394
33666	62570	64775	78428
81666	26440	20422	05720
15838	47174	76866	14330
89793	34378	08730	56522
78155	22466	81978	57323
16381	66207	11698	99314
75002	80827	53867	37797
99982	27601	62686	44711
84543	87442	50033	14021
77757	54043	46176	42391
80871	32792	87989	72248
30500	28220	12444	71840

The starting place for reading is determined arbitrarily or a first digit is randomly selected with a pencil to begin the counting. If the number of cases in the population is two-digit, two-digit numbers are read, if it is three-digit, three digits are read, etc. From the series of numbers, those that are greater than the number of units in the population and those that occurred more than once are then eliminated.

Some computer programs have the ability to provide a series of random numbers that are generated by a certain algorithm and these can be used to select units from the underlying population in the sample.

2.1.2. A systematic sample

When a large population and a large sample are available, units may be selected by systematic enumeration rather than random selection. Depending on the sample size required for the survey, a sampling interval is determined and a counting is made starting from some unit below that interval. For example, if we should select 250 cases out of 1000, the sampling interval is 4. This means that every fourth unit will be selected, starting with the first, second, third, or fourth. Which will be the starting unit is determined by random selection and then systematically counted down with the appropriate interval. For example, if, by random selection, we choose the ordinal number two as the starting unit, then in the sample with sampling interval four, the following units would be selected: two, six, ten, fourteen, eighteen, twenty-two, etc. Below is an example of a table of numbers that can be used to select units in a systematic sample.

Figure 2. Table of numbers for selecting units in the systematic sample

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

However, it's important to stress that systematic sampling can be bias if there is a hidden pattern or periodicity in the population that aligns with the sampling interval. For example, if a population list is sorted in a particular order, and the sampling interval coincides with that order, the sample may not be highly representative. To mitigate this, it's crucial to ensure the population is randomly ordered or that the starting point is randomly selected. Otherwise, the application of the random sample will be more suitable. More information about the systematic sample can be found in the book by Devore & Berk (2012).

2.1.3. Stratified sample

The stratified sample got its name from the Latin word *stratum*, which means layer. Stratified sampling can be used when the structure of the underlying population is known and a conclusion needs to be made about the characteristics of the certain categories. The stratumrepresents a part of the population which is more homogenous than the homogeneity of the whole population. The homogeneity of the stratum is concluded through measures of dispersion, usually through the standard deviation. Strata in the sample can be formed according to place of residence, according to sex, age, social status, level of intelligence, level of development of a trait or level of hearing impairment or any characteristic according to which one category of the population stands out as more homogeneous in relation to another.

There are two ways of selecting the units in the stratified sample, and therefore there are two subtypes of this sample: proportional stratified and optimal stratified sample.

In ***proportional stratified sampling***, units from individual strata are selected in the same proportion as they are in the population. For example, if the population consists of 500 subjects, where 300 are from the urban and 200 from the rural area, then that ratio is expressed as a proportion and the same ratio is applied when selecting the units in the sample. For example, if in the population 60 % subjects live in a city and 40% in a village and a sample of 100 units should be selected, then 60 subjects should be from a city and 40 from a village. Once the number of units in the strata is determined, the units themselves can then be selected randomly or systematically.

The optimal stratified sample has the best possible characteristics, but the statistical procedures for sampling are more complex. In this sample, the units are selected in an optimal ratio, which means that we try to achieve greater precision on a smaller sample. The principle of selection implies that a smaller number of units is selected from the more homogeneous stratum, and a larger number of units is selected from the more heterogeneous stratum. We conclude about the homogeneity of the strata through the standard deviations and a proportion is calculated, using the values of the standard deviations and the number of units in the separate strata or with symbols:

$$\frac{n_1 s_1}{n_2 s_2}$$

Whereby:

- n_1 are the number of units in one stratum of the population
- n_2 are the number of units in the second stratum of the population
- s_1 is the standard deviation in the first stratum
- s_2 is the standard deviation in the second stratum.

For example, if in one stratum there are 2500 respondents and a standard deviation of 20, and in the second stratum 1500 respondents with a standard deviation of 10, then calculated according to the formula, we will get:

$$\frac{2500x20}{1500x10} = \frac{10}{3}$$

That means that every tenth unit from the large sample and every third unit from the smaller sample should be selected.

2.1.4. Cluster sample

Cluster sample has weaker characteristics compared to previous samples, but it is easier to use. Specific to this sample is that the selection frame does not include all the units of the population but separate sums of those units. The units are grouped into clusters, and the clusters are chosen randomly, while all units from the cluster are included in the research. For example, instead of all the inhabitants of a city, we choose certain districts, urban blocks or a certain number of schools with all the students in them.

William G. Cochran, in his valuable book on sampling techniques (Cochran, 1977), explains this cluster sampling as a sampling technique in which the population is divided into clusters or groups, and a random sample of clusters is selected for analysis. Each selected cluster represents a smaller version of the entire population, and all individuals within the selected clusters are included in the sample.

Cochran explains that cluster sampling is often used when it is more practical or cost-effective to sample groups rather than individuals. The clusters may be natural groups, such as households, schools, or neighbourhoods, or they can be artificially created for the purpose of sampling. (Cochran, 1977),

Ideally, clusters should be heterogeneous within themselves but similar to each other to ensure that they are representative of the population.

A variant of the cluster sample is the multiphase cluster sample, which is used on large populations in wider geographic regions. Although this sample is not as reliable as random sampling, it is more economical and suitable for use in macro research.

In addition to numerous advantages, representative samples also have limitations related to their impracticality in cases where the selection frame is very extensive, when a list of population units is not available or possible, as well as in cases where less-represented subgroups of the population may be missing from the sample.

2.2. Non-representative samples

Samples that do not have characteristics of representativeness for the entire population are also used in research, that is, samples whose selection is biased. In **non-representative samples,** units from the population do not have an equal chance of being selected in the sample, but they are selected according to some other principle. Non-representative samples of the type of convenience sample, judgement sample, or "snowball" sample are used in qualitative research in which we do not intend to draw statistical conclusions, but rather to gain knowledge related to the investigated phenomena, no matter it is a subject, group or community (Angeloska- Galevska, 1998, 42-44).

2.2.1. Convenience sample

The convenience sample is a non-representative sample that is selected based on its convenience or availability. For example, a researcher may choose to survey people in a city park because it is easy to find respondents there. Convenience sampling is used in a variety of research contexts, particularly when researchers face limitations in time, resources, or access to a representative sample. (Babbie, 2016)

In the following text there are some examples of how this sample can be used:

Online Surveys: Researchers often use convenience sampling to collect data through online surveys. They may distribute the survey through social media platforms, email lists, or online forums, allowing individuals who are easily accessible and willing to participate and answer the survey questions. Although not considered representative, such a sample can provide valuable insights into certain topics or discover initial trends in a phenomenon.

Educational research with pupils and students: In educational research, convenience sampling is often used by researchers who have easy access to students. For example, if we want to examine the effects of a certain teaching method, we may select available students from a particular school or visitors to a particular course. Although the findings cannot be generalized to the entire student population, they can still provide useful insights into the given context.

Researchers sometimes use a convenience sample for research in *clinical or healthcare settings*. For example, a study investigating the prevalence of a particular medical condition may select patients from a particular hospital where access to potential participants can be provided. Although the findings may not represent the wider population, they may provide useful information within that specific phenomenon or health practice.

We can also use convenience sampling when studying *volunteers* who are self-selected to participate in a particular program, intervention, or event. For example, a study evaluating the impact of a volunteer-based mentoring program may recruit participants from the individuals that are already involved in the program. Although the findings may not count beyond the volunteer group, they may lead to program improvements or generate insights about the experiences of program participants. Some authors distinguish the volunteer sample as a particular type of sample.

Commercial *or marketing* researchers interested in consumer behaviour or market research can use convenience sampling by approaching individuals in stores or shopping centres to participate in their study. Such a choice is not considered random because in the interviewer's approach to the people there is still a subjective component present, because those people who seem more favourable and ready to answer the question are usually chosen. Although the findings may not represent the entire consumer population, they can provide insight into specific consumer behaviours or preferences in a particular environment.

When using the convenience sample, it is important to consider its limitations, because the findings are not justified to generalize to larger populations, but the results should be interpreted in the specific context of the study.

Non-representativeness is one of the main weaknesses of convenience sampling. Participants are selected based on their availability, which may introduce bias. The sample may not adequately reflect the diversity or characteristics of the target population, potentially compromising the external validity of the research. An additional source of bias arises from the voluntary participation of respondents in research, as those who choose to participate may have different characteristics or attitudes compared to those who refuse to participate. This bias can affect the validity of research findings, especially if certain groups are over- or under-represented in the sample.

Another disadvantage of convenience sampling is that it often results in a smaller number of respondents, as researchers usually rely on participants who are currently available. The small sample may limit the application of statistical procedures, reducing the possibility of detecting significant associations or differences. It may also limit the generalizability of findings and the precision of estimates.

In convenience sampling, participants are selected based on convenience or availability, and therefore randomization is usually absent, that is, the possibility that each unit of the population has an equal chance of being selected in the sample. This lack of randomization may lead to biased results and compromise the internal validity of the study.

2.2.2. Judgement sample

Judgement sampling is also known as purposive or selective sampling. It uses a non-random sampling strategy that is characteristic of qualitative research where the researcher personally decides which specific respondents to include in the sample. (Palinkas et al., 2015). In such purposive unit selection, the researcher usually selects participants who possess specific characteristics or have knowledge or experiences relevant to the research question. The goal is to obtain a sample that can provide the richest and highest quality information about the investigated phenomenon. In selecting this sample, the expertise, knowledge and judgement of the researcher play a significant role in order to identify individuals or cases that can provide valuable insights into the research topic.

Unlike representative samples that aim to ensure generalizability of conclusions, a judgement sample is not intended to be statistically representative of the entire population. Instead, the focus is on selecting cases that can offer in-depth understanding or unique perspectives related to the research objective. (Angeloska- Galevska, 1998)

The judgement sample is typical for qualitative research using interviews, case studies, ethnography and content analysis. Researchers may select participants based on specific criteria such as expertise, diversity, extreme or deviant cases to gain a comprehensive understanding of the research topic.

Potential limitations of this type of sample are bias and subjectivity of the researcher. The selected sample is not representative of the entire population, and the findings cannot be generalized. However, it is valuable because it can provide rich and detailed insight into complex research questions.

When using judgment sampling it is essential to document the selection process and provide a clear rationale for why specific individuals or cases were selected. This transparency helps ensure the reliability and validity of research findings.

2.2.3. Chain or snowball sample

In this sample, initial participants are selected based on specific criteria, usually by recommendation from someone they know, and then they help recruit additional participants from their social network or contacts. This type of sampling is often used when it is difficult to access or locate the target population.

In this strategy, the researcher begins by identifying and selecting a small number of individuals who meet the desired criteria for the study. These people are often called "key informants". After initial participants are included, the researcher asks them to suggest other individuals who may be relevant to the research topic or possess valuable insights, and the researcher may set specific criteria for participant inclusion.

This procedure is repeated several times, with each new participant proposing additional participants. In this way, the sample grows and progresses in size like a rolling snowball until the researcher reaches a point of data saturation, that is, concludes that new participants provide minimal additional information.

The use of chain sampling is particularly valuable when studying populations that are difficult to access or identify, such as marginalized groups, hidden or stigmatized communities. It can also be effective for researching sensitive topics or when participants are interconnected in a particular social network. Bjernacki and Waldorf describe the application of this sample to qualitative research that aimed to discover how some heroin addicts managed to recover from their addiction without any treatment or therapy. In doing so, they used a sample of 100 people who independently stopped taking drugs versus 100 others who stopped taking heroin after receiving certain therapy. (Biernacki, P., & Waldorf, D., 1981).

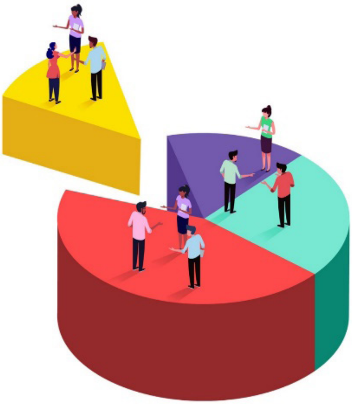
However, it's important to note that snowball sampling has certain limitations. The sample may be prone to biases as referrals tend to come from people with similar characteristics or networks. The results may not be generalizable to the broader population, and there is a risk of creating a homogeneous sample.

Researchers using snowball sampling should be transparent about their sampling method, document the referral process, and acknowledge the limitations of the approach when reporting their findings.

2.2.4. Quota sample

Quota sampling according to characteristics can be considered in between representative and non-representative samples. In quota sampling, the researcher determines in advance a number of respondents from certain categories that are relevant to the research problem (that is why it is called quota) and does not select them according to all the previously mentioned principles of random selection. One of the first authors to write about this type of sample is Kish (1965). Selecting units in this sample involves setting specific quotas based on certain characteristics (such as age, gender, or occupation) and then selecting individuals who meet those quotas.

Figure 3. Graphical illustration of quota sampling



It is most often used when the researcher wants to ensure proportional representation of different groups in the sample. That is why the quota sample is considered by some as a variant of a "non-random" stratified sample

3. Conclusion

Sampling is a critical aspect of research methodology and has a fundamental role in ensuring the reliability and validity of the study findings. The researchers should carefully consider the characteristics of the target population and specific research objectives and choose the sampling strategy accordingly, whether representative or non-representative.

Representative samples such as random sample, systematic, stratified, and cluster provide researchers with a higher level of confidence in generalizing their findings to the larger population. These methods allow random selection and ensure that each member of the population has an equal chance of being included in the sample. By minimizing bias and increasing representativeness, these samples improve the external validity of the research.

On the other hand, non-representative samples such as convenience sample, judgement sample and chain sample are valuable in situations where the selection of a representative sample is impractical or impossible. These strategies involve different forms of bias, but offer flexibility, cost-effectiveness and practicality. They are particularly useful for qualitative studies or when studying hard-to-reach populations. However, it is important to keep in mind their limitations and to be cautious when generalizing the findings beyond the specific sample.

Regardless of the sampling method used, researchers must strive for transparency, clearly documenting the sampling procedure and potential limitations of the chosen samples. Proper sample size, appropriate sampling techniques, and rigorous data collection procedures are the key elements in ensuring the reliability and validity of research results.

We can conclude that using an appropriate sampling strategy is a key decision that affects the quality and generalizability of research findings. Researchers should carefully consider the nature of their research, the characteristics of the target population, available resources, and the desired level of confidence in conclusions. Choosing an adequate strategy will result in better credibility of the research study and will contribute to the advancement of knowledge in the relevant field.

REFERENCES

1. Ангелоска-Галевска, Н. (1998). *Квалитативни истражувања во воспитанието и образованието*. Киро Дандаро. (Angeloska-Galevska, N. (1998). *Qualitative research in upbringing and education*. Kiro Dandaro.)
2. Babbie, E. (2016). The Practice of Social Research.
3. Biernacki, P., & Waldorf, D. (1981). Snowball Sampling: Problems and Techniques of Chain Referral Sampling.
4. Cochran, W. G. (1977). Sampling techniques. John Wiley & Sons.
5. Devore, J. L., & Berk, K. N. (2012). Modern Mathematical Statistics with Applications. Springer.
6. Kish, L. (1965). Survey sampling. John Wiley & Sons.
7. Lohr, S. L. (2019). Sampling: Design and analysis. Cengage Learning.
8. Palinkas, L. A., et al. (2015). Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. Administration and Policy in Mental Health and Mental Health Services Research, 42, 533-544.