The effect of sampling technique training using UNPAD statistical analysis series on sampling technique ability

R Jatnika¹,*, M Haffas² and H Agustiani¹

¹ Faculty of Psychology, Universitas Padjadjaran, Jl Raya Bandung-Sumedang KM 21 Jatinangor 45363, Indonesia
² Faculty of Law, Universitas Padjadjaran, Jl Raya Bandung-Sumedang KM 21 Jatinangor 45363, Indonesia

*ratna@unpad.ac.id

Abstract. UNPAD SAS (Statistical Analysis Series) is a software for analysing Statistical data. UNPAD SAS can be used to analyse descriptive statistics, nonparametric analysis, correlation analysis, and sample size determination. This research is an experimental study to determine the effect of sampling technique training using UNPAD SAS on sampling technique ability. The research design used one group pretest-posttest design. The independent variable in this study was the provision of sampling technique training using UNPAD SAS. The dependent variable was the ability of the sampling technique. The participants in this study were 32 students at the Faculty of Psychology, Universitas Padjadjaran. The ability of sampling techniques was measured using a questionnaire consisting of 28 items. The results of the data analysis showed a significant increase in the ability of the sampling technique after training compared to before taking the training (t = -13.145, p = 0.000).

1. Introduction

Sampling is an integral part of scientific research. One of the most widely used in research is the sampling technique and minimum sample size calculation. Sampling is important in research to be able to produce accurate results. A well-drawn sample more effectively mirrors the population of interest, allowing relatively accurate generalization of relationships from the sample to the population [1]. A poorly drawn sample, on the other hand, may contain systematic biases that distort findings. The method by which the sample is selected from a sampling frame is integral to the external validity of a survey: the sample has to be representative of the larger population to obtain a composite profile of that population [2]. According to Gupta et al., the sample size is a first and important step in planning a clinical trial [3]. Although techniques for sample size calculation are described in various statistical books, performing these calculations can be complicated and it is desirable to consult an experienced statistician in estimation of this vital study parameter. To have meaningful dialogue with the statistician every research worker should be familiar with the basic concepts of sample size calculations. Figure 1 illustrates the stages that are likely to go through when conducting sampling according to Taherdoost [4]:
One problem that is often experienced by researchers in conducting their research is determining the sample size. The sample size is a significant feature of any empirical study in which the goal is to make inferences about a population from a sample. In order to generalize from a random sample and avoid sampling errors or biases, a random sample needs to be of adequate size [5]. One of the most common reasons why researchers seek help from statistician is sample size calculation. However, despite the common belief that it only involves formula and calculation, researchers often ignore other aspects of research design that leads to proper sample size calculation [6]. Although sample size can be determined in many ways, the most commonly used approach has relied on classical hypothesis-based methods of power analysis. Several challenges often arise with traditional power analysis. The American Psychological Association currently accepts and endorses the use of two methods, power analysis, and precision estimation, for sample size determination [7].

There are many software packages for performing sample size/power calculations:

- Minitab
- PS
- GPower
- Epi-info/StatCalc

In summary, PS is probably the best choice of software, as it covers the most commonly used study designs, is relatively easy to use and importantly, is easy and free to download. However, the user may discover his/her own personal preference [8].

Jatnika and Haffas developed software for sample size determination as an alternative sample size calculation software. The resulting software is named UNPAD SAS (Universitas Padjadjaran Statistical Analysis Series) [9]. UNPAD SAS is made using C++ programming language and method (SDLC = Software Development Life Cycle). SDLC is a systematic process developed by the system analyst and programmer for building information systems. This method consists of stages: Planning, Analysis, Design, Implementation, Testing, Maintenance. UNPAD SAS consists of database management, descriptive statistics, non-parametric statistics, correlation analysis and sample size determination [9]. UNPAD SAS for sample size determination consists of:
1.1. Simple random sampling:
- Sample Size Required to Estimate Mean
- Sample Size Requires to Estimate Proportion

1.2. Stratified random sampling:
- Sample Size Required to Estimate Mean
- Sample Size Requires to Estimate Proportion

1.3. Cluster random sampling:
- Sample Size Required to Estimate Mean
- Sample Size Requires to Estimate Proportion

1.4. Systematic sampling:
- Sample Size Required to Estimate Mean
- Sample Size Requires to Estimate Proportion

1.5. Random generator

Figure 2. UNPAD SAS.

Students often made mistakes in determining sampling technique to be used for their research. These mistakes would result in less accurate research conclusions. Therefore, sampling technique ability needs to be trained to students especially on aspects described by Taherdoost in figure 1. Problems encountered in sample size calculation will be assisted by using UNPAD SAS developed by Jatnika and Haffas [9]. Therefore, this study wants to find out how is the effect of sampling technique training using UNPAD SAS on the sampling technique ability of the students.

2. Methods
This research used experimental design. The design used was the one-group pretest-posttest design. A one-group pretest-posttest design is a type of research design that is most often utilized by behavioral researchers to determine the effect of a treatment or intervention on a given sample. This research design is characterized by two features. The first feature is the use of a single group of participants (i.e., a one-group design). This feature denotes that all participants are part of a single condition—meaning all
participants are given the same treatments and assessments. The second feature is a linear ordering that requires the assessment of a dependent variable before and after treatment is implemented (i.e., a pretest-posttest design). Within pretest-posttest research designs, the effect of a treatment is determined by calculating the difference between the first assessment of the dependent variable [10].

Research variables are:

- Independent Variable is the provision of sampling technique training using UNPAD SAS
- Dependent Variable is student ability in sampling technique

This research design can be described as following

\[ O1 \rightarrow X \rightarrow O2 \]

O1: Student ability in sampling technique
X: Provision of sampling technique training using UNPAD SAS
O2: Student ability in sampling technique

The ability of the sampling technique is measured using a questionnaire consisting of 28 items. This questionnaire is made using multiple choices with 4 answer choices. The questionnaire is developed using aspects:

- Basic concepts of sampling technique
- Sample Size Determination
- UNPAD SAS for Sample Size Determination

The provision of sampling technique training is delivered using the curriculum described in table 1.

\textbf{Table 1. The curriculum of sampling technique training.}

| Aspects                                | Sub Aspects                        |
|----------------------------------------|------------------------------------|
| 1. Basic Concepts of Sampling Technique| a. Non-Probability Sampling |
|                                        | b. Probability Sampling            |
| 2. Sample Size Determination           | a. Simple Random Sampling          |
|                                        |   - Sample Size Required to Estimate Mean |
|                                        |   - Sample Size Required to Estimate Proportion |
|                                        | b. Stratified Random Sampling      |
|                                        |   - Sample Size Required to Estimate Mean |
|                                        |   - Sample Size Required to Estimate Proportion |
|                                        | c. Cluster Random Sampling         |
|                                        |   - Sample Size Required to Estimate Mean |
|                                        |   - Sample Size Required to Estimate Proportion |
|                                        | d. Systematic Sampling             |
|                                        |   - Sample Size Required to Estimate Mean |
|                                        |   - Sample Size Required to Estimate Proportion |
|                                        | Random Generator                   |
| 3. UNPAD SAS installation and procedures|                                    |
| 4. UNPAD SAS for sample size determination|                                |

The participants in this study were 32 students at the Faculty of Psychology, Universitas Padjadjaran. Data obtained from this study is analyzed using the t-test to know the increase of ability in sampling technique after training compared to before taking the training.
3. Results and discussion
The result of data processing can be seen in figure 3 and table 2 as follows:

![Figure 3](image)

**Figure 3.** Mean of sampling technique ability before and after training.

**Table 2.** Technique sampling ability before and after training (n=32).

|       | Mean  | Standard Deviation | T     | p     |
|-------|-------|--------------------|-------|-------|
| Pretest | 14.5625 | 4.0835            | -13.145 | 0.000* |
| Posttest | 24.7812 | 2.1211            |

*: P < 0.01

From Figure 3 and Table 2 we can see that there is a significant increase in sampling technique ability of the students after taking the training than before taking one. It means the given training can increase sampling technique ability significantly.

This research also measured the participant's reaction to the given training. Training reaction was measured using 8-items questionnaire with scoring scale: 1 = strongly disagree, 2 = disagree, 3 = agree, 4 = strongly agree. The result of the participant's reaction to the given training can be seen in the following table:

**Table 3.** Training reaction mean.

| No | Statement                                                                 | Mean   |
|----|---------------------------------------------------------------------------|--------|
| 1  | Installing My SQL Server is easy to do                                   | 2.8438 |
| 2  | Installing My SQL Connector is easy to do                                | 2.8125 |
| 3  | Installing UNPAD SAS is easy to do                                       | 2.5312 |
| 4  | Preparing Database is easy to do                                         | 2.6875 |
| 5  | Configuring ODBC is easy to do                                           | 2.5000 |
| 6  | The process of analyzing minimum sample size calculation data is easy to do | 3.1250 |
| 7  | Result presentation is easy to understand                                 | 3.2812 |
| 8  | Result presentation is attractive                                        | 2.6875 |

Based on the result as presented in Table 3, we can see that the process of analyzing minimum sample size calculation data using UNPAD SAS is easy to do with results that are easy to understand. Although the presentation of the result and the process of UNPAD SAS installation still need to be perfected.

The results of this study indicate that sample size calculation is an easy thing to do. Therefore, by using UNPAD SAS, researchers will no longer ignore the sample size in their research as stated by
Arifin [6]. By using UNPAD SAS researchers can also easily calculate sample size, so what Arifin [6] and Gupta et al. [3], said that researchers often ask Statisticians about sample size, can be overcome. But even though sample size calculations can be easily done using UNPAD SAS, researchers still need to understand the sampling procedure as well as told by Taherdoost [4].

4. Conclusion
Based on the research result, it is concluded that the training on sampling technique using UNPAD SAS can increase sampling technique ability significantly. It can conclude that UNPAD SAS is a good software that can be used to increase sampling technique ability especially for determining sample size. The process of analyzing the minimum sample size using UNPAD SAS is easy to do with results that are easy to understand. Therefore, UNPAD SAS can be used as alternative software to calculate the sample size.

References
[1] Scheaffer, Mendenhall and Ott 2012 Elementary Survey Sampling (USA: Brooks/Cole Cengage Learning)
[2] American Statistical Association 2002 USA [Onlien] Retrieved from: Http://www.amstat.org. Accessed 9 December 2002
[3] Gupta K K, Attri J P, Singh A, Kaur H and Kaur G 2016 Basic concepts for sample size calculation: Critical step for any clinical trials! Saudi journal of anaesthesia 10(3) 328
[4] Taherdoost 2016 Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research Int. J. of Academic Research in Management 5 18
[5] Taherdoost 2017 Determining Sample Size; How to Calculate Survey Sample Size Int. J. of Economic and Management Systems 2 (237)
[6] Arifin 2013 Introduction to Sample Size Calculation Education in Medicine Journal 2(89)
[7] Hayat 2013 Understanding Sample Size Determination in Nursing Research Western J. of Nursing Research 35(943)
[8] Gardner 2010 Sample Size and Power Calculation Made Simple International Journal of Therapy and Rehabilitation 17(10)
[9] Jatnika and Haffias 2018 Developing the UNPAD SAS (Universitas Padjajaran Statistical Analysis Series) Software Sisforma J. of Information Systems 5(8)
[10] Allen 2010 The SAGE Encyclopedia of Communication Research (Canada: Sage Publications Inc)