MARGIN OF ERROR BETWEEN SIMPLE RANDOM SAMPLING AND STRATIFIED SAMPLING

Permadina Kanah Arieska[1], Novera Herdiani[2]

[1] Public Health Department/ Health Faculty, University of Nahdlatul Ulama Surabaya, Indonesia
permadina.kanah@unusa.ac.id

[2] Public Health Department/ Health Faculty, University of Nahdlatul Ulama Surabaya, Indonesia
novera.herdiani@unusa.ac.id

Abstract

The sample is a small part or quantity intended to show what the whole is like. While the sampling technique is part of the statistical method related to sampling methods. The sampling method that can be used for sampling includes Simple Random sampling and Stratified Sampling. A simple random sample is a subset of a statistical population in which each member of the subset has an equal probability of being chosen. Stratified Sampling is a way of sampling by paying attention to the strata (levels) in the population. Both of these sampling methods were compared to obtain a smaller margin of error in the Body Mass Index (BMI) data of the College Student of Health Faculty, Nahdlatul Ulama University, Surabaya. The confidence level used is 99%, 95% and 90% with each sample number as many as 634 samples, 252 samples and 88 samples. The results show that as the sample size increases, the margin of error decreases. Variance in the Stratified Sampling technique is smaller than the variance in Simple Random sampling. So that it can be stated that descriptively, the Stratified Sampling technique is more effective than the Simple Random sampling technique.

Keywords: ***Simple Random Sampling, Stratified Sampling, Margin of error***

Introduction

Simple Random Sampling or commonly abbreviated as Random Sampling is a method of sampling where each element that forms a population is given the same opportunity to be selected as a sample. This method is very easy if there is a complete list of elements of the population. The procedure that is quite accurate for random sampling is to use a Table of random numbers, besides that it can also be done by raffling. Random sampling carried out according to the procedure is not at all a guarantee that a sample will be a perfect representation of the population because random sampling can actually result in a unique sample but random sampling must be understood in the context of the probable process if the sample randomly taken from a population repeatedly, in general, all of the samples will be able to provide a more accurate estimate of the population as well as variability or error can be estimated and the test of statistical significance also shows the probability of results taking into account Sampling Errors.

Stratified Sampling is a way of taking samples by paying attention to the strata (levels) in the population. In stratified data previously grouped into certain levels, such as levels, high, low, medium / good, samples taken from each level. Populations are grouped into sub-populations based on certain criteria possessed by population size. Each sub-population is homogeneous. From each sub then some members are randomly taken with a proportional / disproportionate composition. The total members taken are determined as the number of members of the study sample.

The margin of error represents the number of errors in sampling in a survey. Margin of error measures how representative data is obtained from samples with data that is in the actual population. The greater the margin of error, the further a sample can be said to represent the actual population. The smaller the margin of error, the more data can be said in the sample representing the actual population data. Margin of error arises because in practice the sample cannot always explain the population perfectly. How well a sample represents a population

can be seen from two things, namely the margin of error and the confidence levels. From these two aspects it can be seen how we choose good samples that truly represent the population.

Body Mass Index (BMI) data from students of the Nahdlatul Ulama University, Health Faculty in Surabaya can be done using Random Sampling and Stratified Sampling. Both sampling will be compared to the margin of error and the confidence level to compare which types of sampling methods provide a smaller margin of error.

Research methods

The study was conducted by collecting data on BMI Student of the Health Faculty, Nahdlatul Ulama University, Surabaya (UNUSA). BMI data is calculated the size of the concentration and distribution. Then the sampling method is applied using Random sampling and Stratified sampling. The sampling method chosen is a sampling method that has a smaller margin of error and a smaller variance. The variables observed in this study were the Body Mass Index (BMI). The Body Mass Index is calculated from body weight data in Kilograms and data Height in meters. Population data is the BMI data of the Student's Health Faculty in Surabaya's Nahdlatul Ulama University which consists of three study programs with a total of 676 students. BMI data will be applied using Random sampling and Stratified Sampling. These two methods will compare the value of the margin of error with different confidence interval.

Research Results and Discussion

In descriptive analysis, the description show about comparison of error margins based on the number of samples and sampling results using two sampling methods. Table 1 shows data on the number of Health Faculty students according to the Study program. There are 676 students who are members of the population.

Table 1. Number of Health Faculty Students According to the Study Program

Study Program	Semester				Total
	Semester 8	Semester 6	Semester 4	Semester 2	
Public Health	55	46	54	85	240
Laboratory Analyst	54	61	55	94	264
Nutritionist	37	35	49	51	172
TOTAL POPULATION					676

Figure 1 shows the number of students based on gender. Male students have 12% and women have 88%. Student data based on this gender will be used to take samples in the Sampling Stratification Method.
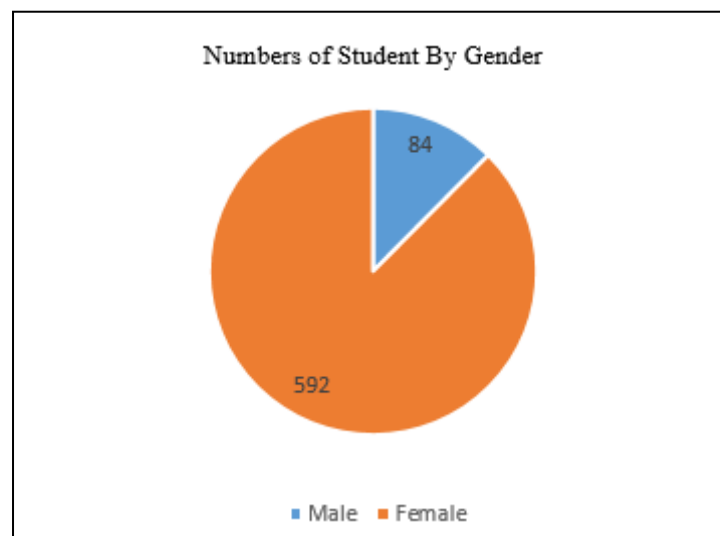
Figure 1. Numbers of Student By Gender

The number of samples used a confidence level of 99%, 95% and 90%. The determination of this confidence interval is an interval of level that is often used in research. Table 2 shows the number of samples taken based on various confidence levels.

Table 2. Number of Samples with Different Confidence Levels

Population	Confidence Level	Sample
676	99%	634
676	95%	252
676	90%	88

Simple Random Sampling technique is one simple and widely used sampling technique. The selection of respondents is based on random numbers using Microsoft Excel. There are a number of respondents selected according to the number of samples obtained. The parameter estimator contain mean and variance with various types of samples are explained in Table 3. Table 3 show an error margin calculation will be performed to describe that the number of samples affects the margin of error. The large number of samples makes the error margin of a sampling will be small and vice versa.

Table 3. Parameter Estimator in Simple Random Sampling

Parameter Estimator	634 Sample	252 Sample	88 Sample
Mean	22.45	22.57	23.75
Variance	0.002	0.054	0.286

The use of Stratified sampling is used by distinguishing BMI data between male and female. Table 4 shows the population based on gender. The use of strata by gender is based on information that BMI between male and female were difference.

Table 4. Number of Health Faculty Students According to Gender

Gender	Population Number
Male	84
Female	592
Total	676

Stratified sampling was also carried out using a confidence level of 90% 95% and 99%. Table 5 shows the number of samples in each strata according to the proportion of the population. The number of samples at each level of confidence will be calculated by estimating the parameters, mean and variance, as in the simple random sampling technique stated in Table 6.

**Table 5.** Total Samples at Each Strata According to Proportion

Strata	Total Population	Total Samples		
		Confidence Level 99%	Confidence Level 95%	Confidence Level 90%
Male	84	79	31	11
Female	592	555	221	77
Total	676	634	252	88

Table 6. Parameter Estimator in Stratified Sampling

Parameter Estimator	634 sample	252 sample	88 sample
Mean	22.54	22.7	22.69
Variance	0.002	0.04	0.218

Figure 2 shows that the smaller number of samples, the greater margin of error. This causes of confidence interval to be even wider than the average estimator value. Margin of error represents the number of errors in sampling survey. Margin of errors measure how representative data is obtained from samples with data that is in the actual population. The greater margin of error, the further a sample could not be said to represent the actual population. The smaller margin of error, the more data can be said that sample represent the actual population.

Measuring the margin of error (MoE) can only be done if the sampling method uses probability sampling. Margin of error is influenced by two things, namely the sample size and sampling method. Figure 2 shows that descriptively, the sampling method with stratified sampling has a smaller MoE compared to simple random sampling, even though the difference is small. In 634 samples (i.e. with a confidence level of 99%), the MoE obtained by the same amount in both sampling techniques. The results show that as the sample size increases, the margin of error decreases. In the case of a small number of samples (with a smaller level of confidence), stratified sampling has smaller margin of error than simple random sampling.
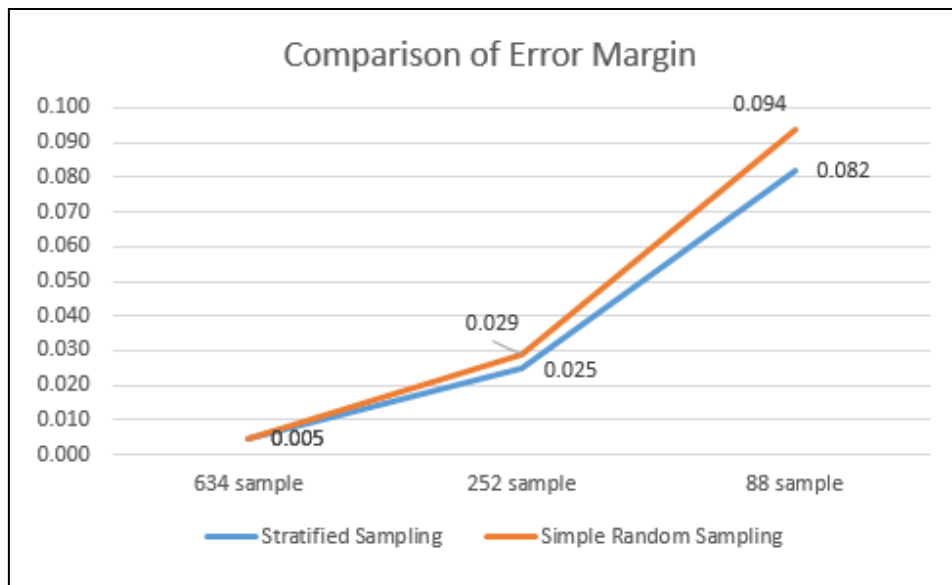
Figure 2. Comparison of Error margin

Conclusions and recommendations

The results show that as the sample size increases, the margin of error decreases. Variance in the Stratified Sampling technique is smaller than the variance in Simple Random sampling. Based on the margin of error, it shows that in the case of the Body Mass Index (BMI) the sample calculation using the stratified sampling method produces a smaller error margin compared to simple random sampling. So descriptively, the stratified sampling method is more efficient than the simple random sampling method.

References

Abuzar Asra dan Achmad Prasetyo. 2015. Pengambilan Sampel dalam Penelitian Survei. Jakarta : RajaGrafindo Persada.

Charles Teddlie and Fen Yu. 2007. *Mixed Methods Sampling : A Typology With Examples.* Journal of Mixed Methods Research.

Habib Ahmed Elsayir. 2004. *Comparison of precision of systematic sampling with some other probability samplings*. American Journal of Theoretical and Applied Statistics.

J. Supranto.2007. Teknik Sampling untuk survei dan Eksperimen. Jakarta : Penerbit Rineka Cipta.

Leslie Kish, Martin Richard Frankel. 2004. *Inference from complex Samples.* JSTOR

Paula Lagares Barreiro. 2001. *Population and sample. Sampling techniques*. Justo Puerto Albandoz. Management Mathematics for European Schools.

Raihan Budiwaskito. 2010. Margin of Error. Program Studi Sistem dan Teknologi Informasi. Sekolah Teknik Elektro dan Informatika. Institut Teknologi Bandung.

Saifuddin Azwar. 2015. Reliabilitas dan Validitas.. Edisi 4. Yogyakarta : Pustaka Pelajar.

Veer B. 2014. Dasar Dasar Biostatistik. Tangerang : Karisma Publishing Group.

Xiangrui Meng. Scalable Simple Random Sampling and Stratifed Sampling. LinkedIn Corporation, 2029 Stierlin Court, Mountain View, CA 94043, USA