ISI2F3 – STATISTIKA INDUSTRI

# LECTURE NOTES – CHAPTER 1

## SAMPLING THEORY

RISKA YANU FA'RIFAH, S.Si., M.Si. (RYF)

## Learning Objectives

1. students can understand about the population and sample
2. students can understand the concept of sampling
3. Students can understand the kinds of sampling technique
4. Students can understand how to determine the right sample size.

## Population and Sample

1. **Population** is a group of objects or individuals with certain characteristics in an area at a certain time which is considered for the study. The object in the population can be anything
2. There are two types of Population, **Finite Population** and **Infinite Population**

   **Finite Population** is the population which can be counted. It is also known as a countable population.

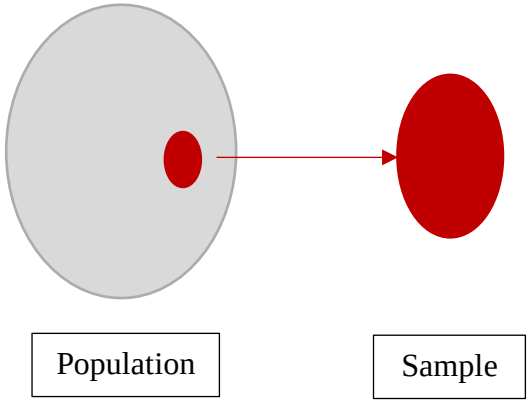
   **Infinite Population** is a collection of objects or individuals that **are no boundaries or can't be counted** about the total number of individuals or objects in the occupied territories

   **For Example:**

Infinite Population


Finite Population

3. **Sample** is the section containing the characteristics of the population.
   Sample is used in statistical testing when population sizes are too large.
4. The relationship between Population and Sample can be described as follows:




If population is N and sample is n, then n <N

5. **Parameter** and **Statistic** are the characteristics of the population and sample
   **Parameter** is a value or characteristic that is measured from all elements of the population
   **Statistic** is a value or characteristic that is measured from the sample elements taken from a population
6. The collection or measurement of data on the population is called a **census**, while the sample is called **sampling**

## Sampling Concept

**Sampling** is a process carried out to select and retrieve sample members appropriately from the population so that the sample taken can represent the population.

1. **Sampling can be done due to several factors**, namely:
   a. The original population has too many (infinite) numbers
   b. The census is not economical and inefficient
   c. The results of the census are not qualified
   d. Destructive experimental research
2. **Good and precise sampling should be:**
   a. The sample size must be sufficient
   b. Representative (describes the state of the original population)

   **To get the two criteria above, a sampling technique is needed**

## Sampling Technique

1. **Sampling techniques** are used to take samples that can represent the population.
2. There are **three important things** related to **sampling**
   a. Why is sampling done?
   b. How to take sample? → **Sampling Technique**
   c. How many sample members should be drawn? → **Sample size**
3. Sampling techniques can be categorized into two types, namely probability and nonprobability sampling
4. **Probability Sampling** is a way of taking a randomly selected sample or using probability theory.
5. **Kinds of Probability sampling**
   a. **Simple Random Sampling**

      It is a random sampling process, in which every individual or element in the population has an equal chance of being selected to be the sample. It can be said that the researcher can freely choose the members of the population to be sampled.

The terms:

- Homogeneous population
- Geographically, the members of the population are located in an accessible place (a certain place or area)

Example:

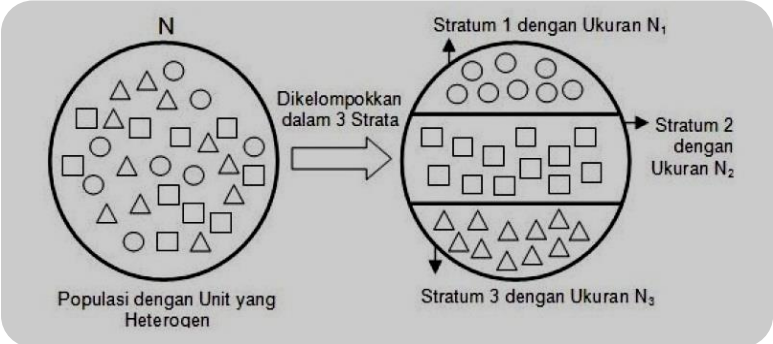
The population is 200 fun bike participants, each of whom has 1 lottery coupon, so everyone has the same chance to win the raffle. There are 40 prizes, 40 out of 200 coupons will be drawn randomly.

b. **Stratified Random Sampling**

It is a sampling method by dividing members of the population into several stratums (according to certain characteristics), then the sample is selected from each formed stratum. The purpose of making this stratum is to get a homogeneous sub-population. There are two types of stratified random sampling, namely:

- If each stratum has a proportional sample size, then it is called Proportionate Stratified Random Sampling
- If some stratum sample size is not proportional, it is called Disproportionate Stratified Random Sampling

Example:




**$N_1$, $N_2$, and $N_3$** are sub-population in stratified random sampling

Sampling is done randomly for each stratum with a certain number (the proportion of the sample is in accordance with the number of members in each stratum)

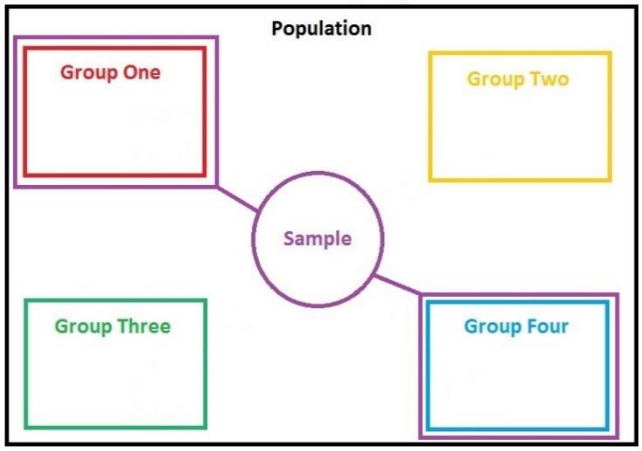
c. **Cluster Random Sampling**

It is a way of taking samples from certain clusters or groups. A cluster can be said to be a sub population of the initial population

Stages in cluster random sampling:

- The population is divided into several clusters or groups
- The first stage sample is taken from some of the clusters that are part of the population (called the primary unit)
- The second stage sample is taken from the primary unit randomly using simple random sampling method or stratified random sampling.

Example:




A lecturer teaches industrial statistics in 4 classes (Class 1 to class 4). If a team is to be formed with 10 members randomly selected from the four classes, then:

- In the first stage, a random cluster / class is taken to determine the 2 selected classes (Class 1 and Class 4).
- The second stage takes samples from each selected cluster with a total sample size of 10 students.

  The arrangement may be one student from grade 1 and 9 students from grade 4 (1 - 9), or (2 - 8), or (3, 7), or (4 - 6), or (5 - 5) depending on the concept used, using simple random sampling or stratified random sampling.

**d. Systematic Random Sampling**

It is a sampling method, where only the first unit of observation from the sample is taken randomly, while each subsequent unit is systematically selected according to the interval.

If the sample size is too large in Simple Random Sampling, it will cause cost and time burdens. Therefore, Systematic Sampling is simpler than Simple Random Sampling

The stages of this Sampling are:

- Determining the sample size you want to take $(n)$ from the population $(N)$ so that the interval is known $(i = N/n)$.
- Taking the first sample unit $(AR_1)$ randomly using conditions $AR_1 \leq i$
- Selecting next sample unit with formula:

  $AR_n = AR_{n-1} + i$

Example:

If we will take 20 people out of 100 people using systematic random sampling, then the stages are:

- Determining an interval

  $i = \frac{N}{n} = \frac{100}{20} = 5$

- Determining the first random sample unit with conditions $AR_1 \leq i$. Suppose a fourth person is taken.
- Determining the second to 20 sample units by means of:

  $AR_n = AR_{n-1} + i$

  $AR_2 = AR_1 + i = 4 + 5 = 9$, so the ninth person who was drawn next.

  $AR_3 = AR_2 + i = 9 + 5 = 14$

  ⋮

  $AR_{20} = AR_{19} + i = 94 + 5 = 99$

6. **Nonprobability sampling** is sampling without paying attention to the opportunities for each unit of the population to become a sample unit (members of the population do not have the same chance of being used as sample members)
7. **Kinds of Nonprobability Sampling**
   a. **Purposive Sampling**

      Is a way of taking samples with a specific purpose or purpose. For example, you want to research about satisfaction with the use of gadgets with brand A. In this case, there will definitely be difficulties when using the probability sampling method, because it is not known exactly who uses the branded gadget A. The easiest way is to come to a place, if someone uses a gadget brand A can be used as a respondent.

   b. **Snowbell Sampling**

      This sampling technique is widely used when the researcher does not know much about the study population, which is known to only one or two members of the sample. Because the researcher wanted more, the researcher asked the first sample to show another sample that could be used as a sample.

For example, the researcher wants to know the economist's view of the implementation of the PSBB. In this case, the researcher first takes data from an economist who he knows. Then the researchers asked the list of names of economists known by the first respondent to be the next respondent.

c. **Accidental Sampling**

It is a way of taking samples only considering convenience without paying attention to other aspects.

For example, the research is to find out the types of social media used by students, then the sample taken is the students in the researcher's class so that it is easier, faster or closer. Things like this turn out to be less objective, thus providing unfavorable general conclusions.

**In general, this nonprobability sampling can be used for social, economic, or marketing research, where research only wants to find out opinions or views without having to carry out further analysis**.

## Sample Size

*How much sample should be taken?*

The higher the level of precision that is desired, the larger the sample size that must be taken. The sample size if the population is known can use the **SLOVIN** formula with the formula:

$$n = \frac{N}{1 + Ne^2}$$

with:

$n$ : number of sample members

$N$ : Number of population members

$e$ : error tolerance (level)

The most commonly used error tolerance is 5% or 10%.

In determining the sample size there are two deviations, namely:

1. ***Sampling Error***

   Represents a deviation due to sample usage. In the sample measurement results obtained statistical values in probabilistic sampling. This statistical value will not exactly match the parameter value. This difference is called a sampling error.

2. ***Non Sampling Error***

   Represents deviations not by use of the sample. Deviation in non-probabilistic sampling, the value of the sample against the population is impossible to measure. Here the deviation that arises due to sampling technique planning errors, sample replacement, enumerator misinterpretation, respondent misinterpretation, respondent intentionally wrong answer or errors due to data processing.

**The first error can be overcome by using an appropriate sampling method. And the second can be overcome with thoroughness in planning, implementing and processing data until publication.**