Chapter 1

Introduction

Definition of Statistics (page 3)

Definition 1.1.

Statistics is the branch of science that deals with the collection, organization, analysis, interpretation and presentation of data.

Definition of Statistical Inquiry (page 11)

Definition 1.7.

A **_statistical inquiry_** is a designed research that provides information needed to solve a research problem.

Population vs. Sample (page 4)

Definition 1.2.

Population is the collection of all elements under consideration in a statistical inquiry.

Sample is a subset of the population.

Figure 1.1 (page 4) shows a diagram depicting the relationship between the population and the sample. The big circle is the population while the small circle within is the sample. This emphasizes the requirement that all elements of the sample must belong in the population.
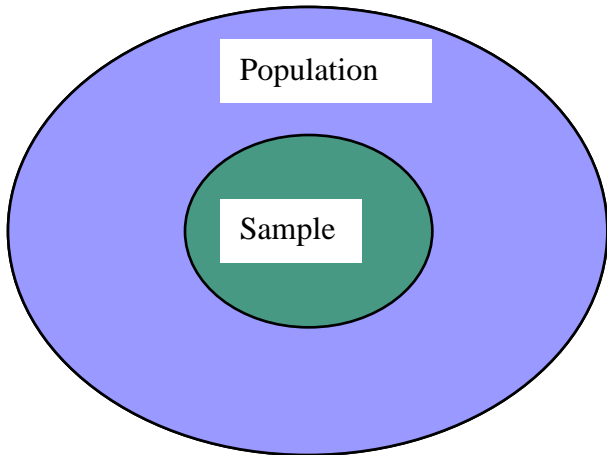
Population

Sample

Figure 1.1 Population and the Sample

Examples (page 4)

Problem: What is the average expenditure of households in Metro Manila?
Population: set of all households in Metro Manila

Problem: What is the average expenditure of households in Quezon City?
Population: set of all households in Quezon City

Example of population with people as elements:
- set of farmers in Central Luzon

Examples of population with animals/objects as elements:
- collection of milkfish cultured in Luzon
- set of fluorescent bulbs manufactured for a month

Examples of population with areas as elements:
- set of barangays in Metro Manila
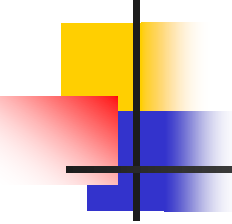- collection of fishponds in Pampanga

Variable, Observation, and Data (page 5)

Definition 1.3.

- The ***variable*** is a characteristic or attribute of the elements in a collection that can assume different values for the different elements.

- An ***observation*** is a realized value of a variable measured from an element.

- ***Data*** is the collection of all the observations.

Examples of variables with their possible values: (Example 1.1, page 5)

Variable Name	Possible Values
I = monthly income of a person in pesos	$i \geq 0$
N = number of children of a teacher	$n = 0, 1, 2, 3, \ldots$
H = height of a basketball player	$h > 0$ cms.

Example of Population and Variable/s of Interest
(Example 1.2b, page 5)

The research division of a certain pharmaceutical company is investigating the effectiveness of a new diet pill in reducing weight on female adults.

Population: set of all female adults who will use the diet pill

Variable/s of interest: weight before taking the diet pill, weight after taking the diet pill

Exercise #2 (page 7)

A doctor claims that three tablespoons of pure virgin coconut oil daily can reduce weight. To test the doctor's claim, a researcher studied two groups of 25 women aged 35 to 40 years with weights between 130 to 140 pounds. He administered the three tablespoons of pure virgin coconut oil daily for a period of three months to one group of women only. After three months, he weighed the two groups of women.

a) Identify the two populations of interest.
b) What are the two samples?
c) What is the variable of interest?

Definition of Summary Measure (page 6)

A ***summary measure*** is a single numeric figure that describes a particular feature of the whole collection.

Parameter and Statistic (page 6)

Definition 1.4.

- The ***parameter*** is a summary measure describing a specific characteristic of the population. (It is computed using population data.)

- The ***statistic*** is a summary measure describing a specific characteristic of the sample. (It is computed using sample data.)

Example of Parameter and Statistic (Example 1.4, page 6-7)

Consider the case where our population consists of 35 students in an Elementary Statistics class. The parameter of interest is P=proportion of students in this class with cellular phones.

$$P = \frac{\text{number of elements possessing a certain characteristic}}{\text{number of elements in the collection}} = \frac{no.\ of\ students\ with\ cellular\ phones}{no.\ of\ students\ in\ the\ population}$$

The variable under study, X, is whether or not the student owns a cellular phone. This variable has only two possible values:

    0 - student does not own a cellular phone   1 - student owns a cellular phone.

Suppose that among the 35 students, 28 own a cellular phone. Thus,

$$P = \frac{no.\ of\ students\ with\ cellular\ phones}{no.\ of\ students\ in\ the\ population} = \frac{28}{35} = 0.8$$

Suppose we were not able to collect data from all the 35 students. Instead, we only took a sample of 10 students from this class. Among the 10 students in the sample, 7 own cellular phones. Can you compute for the parameter, P?

Example of Parameter and Statistic
(Example 1.4 cont'd, page 6-7)

We cannot compute for the parameter, P=proportion of students in the population with cellular phones but we can compute for $\hat{P}$ (read as "P hat"), where $\hat{P}$ =proportion of students in the sample with cellular phones as follows:

$$\hat{P} = \frac{no.\ of\ students\ with\ cellular\ phones\ in\ the\ sample}{no.\ of\ students\ in\ the\ sample} = \frac{7}{10} = 0.7$$

The proportion of students in our sample with cellular phones is an example of a statistic because it is a summary measure describing a characteristic of the sample.

Suppose we redefine the population as the collection of all students enrolled in all sections of Elementary Statistics so that the class consisting of 35 students earlier is now just a sample taken from this new population. Is the earlier computed proportion of 0.8 a parameter or a statistic?

Exercise #3 (page 7)

Mr. Donaldo Chan, a candidate for vice-mayor in Orion, Bataan, wants to find out if there is a need to intensify his campaign efforts against his opponents. He requested the services of a group of students to interview 1,000 of the 3,000 registered voters of Orion, Bataan. The survey results showed that 75% of the 1,000 voters in the sample will vote for him as vice-mayor.

a) Identify the population and the sample.
b) Identify the variable of interest.
c) Identify the parameter and the statistic.

Exercise #4: (page 8)

The average weekly allowance of students last year at a private high school was P600.00 per week, based on an enrollment of 1,080 students. The third year students who did not have this information interviewed 50 students and found their average weekly allowance last year to be P550.00.

a) Identify the population and sample.
b) Identify the variable of interest.
c) Identify the parameter and the statistic.

Exercise #5: (page 8)

Define suitable populations, variable/s, and the parameter of interest for the following studies:

a) The manufacturer of badminton shuttlecocks wishes to determine how many games their brand of shuttlecock will last on the average.

b) A market researcher wishes to determine the market share of the four leading brands of cereal among children in Metro Manila who are 12 years old or younger.

Two Major Fields of Statistics
(page 8)

1. **Applied Statistics**
   - Definition: ***Applied statistics*** is concerned with the procedures and techniques used in the collection, presentation, organization, analysis, and interpretation of data.
   - The study of applied statistics allows us to select and properly implement the most appropriate statistical methods that will provide solutions to the research problem

2. **Theoretical or Mathematical Statistics**
   - Definition: ***Mathematical Statistics*** is concerned with the development of the mathematical foundations of the methods used in applied statistics.
   - The study of mathematical statistics permits us to understand the rationale behind the methods we use in analysis and to establish new theories that will validate the use of new statistical methods or modifications of existing statistical methods in solving research problems that are more complex.

Two Major Areas of Applied Statistics

1. **Descriptive Statistics** (page 9)
    - Definition 1.5. ***Descriptive statistics*** includes all the techniques used in organizing, summarizing and presenting the data on hand.
    - It cannot be used to make generalizations about the population if the data on hand is simply sample data.

2. **Inferential Statistics** (page 10)
    - Definition 1.6. ***Inferential statistics*** includes all the techniques used in analyzing the sample data that will lead to generalizations about a population from which the sample came from.
    - We arrive at our conclusions under conditions of uncertainty because we use partial information only. Conclusions will be subject to some error. Probability theory will help us understand the possible errors that can be committed.

Example of Descriptive Statistics: (Example 1.5a, page 9)

Given the daily sales performance for a product for the previous year, we can draw a line chart or a column chart to emphasize the upward/downward movement of the series. Likewise, we can use descriptive statistics to calculate a quantity index per quarter to compare the sales by quarter for the previous year.

Example of Inferential Statistics: (page 10)

Election polls make use of inferential statistics to predict the winners for the coming election based on data collected from a sample of registered voters.

Example: Krystal Surveys on the 1998 Presidential Elections

Sample size: 500 Metro Manila
                     <u>2387</u> Other provinces
                     2887

	Actual (based on quick count)	Estimate (based on survey)
Estrada	40%	34%
De Venecia	16%	14%

Chapter 1. Introduction

Exercises (page 11)

For each of the following statements below, decide whether it belongs to field of descriptive statistics or inferential statistics.

#2. Janine wants to determine the variability of her six exam scores in Algebra.

#6. A car manufacturer wishes to estimate the average lifetime of batteries by testing a sample of 50 batteries.

#8. a) A marketing research group wishes to determine the number of families not eating three times a day in the sample used for their survey.

   b) A marketing research group wishes to determine the number of families in the Philippines not eating three times a day based on the sample used for their survey.

#9. A politician wants to determine the total number of votes his rival obtained in the past election based on his copies of the tally sheet of electoral returns.

#10. A politician wants to determine the total number of votes his rival obtained in the election based on the sample used in the exit poll.

Definition of Statistical Inquiry (page 11)

Definition 1.7.

A **statistical inquiry** is a designed research that provides information needed to solve a research problem.

Possible Objectives of a Statistical Inquiry (page 12)

1. describe the characteristic of the elements in the population under study through the computation or estimation of a parameter such as the proportion, total, and average;

2. compare the characteristics of the elements in the different subgroups in the population through contrasts of their respective summary measures;

3. justify an assertion made by the researcher about a particular characteristic of the population or subgroups in the population;

4. determine the nature and strength of relationships among the different variables of interest;

5. identify the different groups of inter-related variables under study;

6. reveal the natural groupings of the elements in the population based on the values of a set of variables;

7. determine the effects of one or more variables on a response variable;

8. clarify patterns and trends in the values of a variable over time or space;

9. predict the value of a variable based upon its relationship with another variable; and,

10. forecast future values of a variable using a sequence of observations on the same variable taken over time.

Assignment 1:

Look for 5 different statistical inquiries in the field assigned to your group.

a)   What is the title of the study?
b)   State at least 2 specific objectives of this study that the researchers will achieve by using statistics.
c)   Explain how the achievement of the stated objectives in (b) will be useful in decision-making. In other words, discuss the importance of the achievement of the stated objectives.

Fields:
1.   Public administration and governance
2.   Economics
3.   Marketing
4.   Banking and Finance
5.   Medicine and Epidemiology
6.   Manufacturing and Production
7.   Education
8.   Food science and nutrition
9.   Tourism
10.   Sports

Chapter 1. Introduction

Reading Assignment: Stages in Doing a Statistical Inquiry (pages 13-19)

1. Identify the problem.
2. Plan the study.
3. Collect the data.
4. Explore the data.
5. Analyze data and interpret results.
6. Present the results.

READING ASSIGNMENT:

Brief History of Statistics
(hand-out)