Unit 1: Introduction

Statistics, the science of collecting, analyzing, presenting, and interpreting data. Governmental needs for census data as well as information about a variety of economic activities provided much of the early impetus for the field of statistics. Currently the need to turn the large amounts of data available in many applied fields into useful information has stimulated both theoretical and practical developments in statistics.

Data are the facts and figures that are collected, analyzed, and summarized for presentation and interpretation. Data may be classified as either quantitative or qualitative. Quantitative data measure either how much or how many of something, and qualitative data provide labels, or names, for categories of like items. For example, suppose that a particular study is interested in characteristics such as age, gender, marital status, and annual income for a sample of 100 individuals. These characteristics would be called the variables of the study, and data values for each of the variables would be associated with each individual. Thus, the data values of 28, male, single, and $30,000 would be recorded for a 28-year-old single male with an annual income of $30,000. With 100 individuals and 4 variables, the data set would have $100 \times 4 = 400$ items. In this example, age and annual income are quantitative variables; the corresponding data values indicate how many years and how much money for each individual. Gender and marital status are qualitative variables. The labels male and female provide the qualitative data for gender, and the labels single, married, divorced, and widowed indicate marital status.

Sample survey methods are used to collect data from observational studies, and experimental design methods are used to collect data from experimental studies. The area of descriptive statistics is concerned primarily with methods of presenting and

interpreting data using graphs, tables, and numerical summaries. Whenever statisticians use data from a sample—i.e., a subset of the population—to make statements about a population, they are performing statistical inference. Estimation and hypothesis testing are procedures used to make statistical inferences. Fields such as health care, biology, chemistry, physics, education, engineering, business, and economics make extensive use of statistical inference.

Methods of probability were developed initially for the analysis of gambling games. Probability plays a key role in statistical inference; it is used to provide measures of the quality and precision of the inferences. Many of the methods of statistical inference are described in this article. Some of these methods are used primarily for single-variable studies, while others, such as regression and correlation analysis, are used to make inferences about relationships among two or more variables.

Comprehension Exercises

Choose a, b, c or d which best completes each item.

1) Collecting _____ data is one of the driving forces of the field of statistics.

a) sample   b) census   c) quantitative   d) population

2) Theoretical and practical developments in statistics are _____ by extracting useful information from large amounts of data.

a) defined   b) indicated   c) stimulated   d) provided

3) The annual income of an individual is a _____ variable.

a) quantitative   b) qualitative   c) known   d) categorical

4) In statistical inference, a sample is a _____ of the population.

a) partition   b) fraction   c) division   d) subset

5) The quality and precision of statistical inferences are measured using _____ methods.

a) qualitative   b) probability   c) descriptive   d) economical

Words to Learn

Find the Persian equivalents of the following terms and expressions.

science	analyze	census	data
statistics	collect	theoretical	practical
economics	summarize	quantitative	qualitative
physics	characteristic	variable	value
biology	presentation	suppose	thus
chemistry	interpretation	measure	individual
engineering	survey	label	information
business	procedure	status	inference
education	population	observational	precision
health care	sample	experimental	statistician