Math 2300: Statistical Methods

Chapter 1

Section 1.1 How can we investigate using data

The information we gather with experiments and with surveys is collectively called **data**.

What is statistics?

Statistics is the art and science of

- Designing studies
- Analyzing the data that those studies produce
- Translating the data into knowledge and understanding of the world

Three main aspects of statistics

- **Design:** Planning how to obtain data to answer the questions of interest
- **Description:** Summarizing the data that are obtained
- **Inference:** Making decisions and predictions based on the data

Ex 1:

Suppose you want to open a restaurant in Lubbock, but not sure what type of a restaurant it should be. How can you find out what type would bring you more profit?

Design:

You have to plan how to collect data. For instance, by interviewing residents picking telephone numbers randomly from the directory.

Description:

Prepare a table with percentages.

Type of restaurant	Popularity (%)
BBQ	24
Mexican	25
Italian	18
Chinese	13
Sea Food	14
Other	6
Total	**100**

Inference:

Mexican and BBQ restaurants are popular. It should be a good choice to start an Mexican or a BBQ restaurant.

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer true or false.

1) The information we gather with experiments and with surveys is collectively called data. 1) B
A) False B) True

SHORT ANSWER. Write the word or phrase that best completes each statement or answers the question.

Fill in the blank.

2) _____ is the art and science of learning from data. 2) Statistics

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Select the most appropriate answer.

3) The following statement refers to which aspect of a statistical study: "A meteorologist constructs a graph showing the total precipitation in Phoenix, Arizona in each of the months of 2004"? 3) A
A) Description B) Design C) Inference

4) The following statement refers to which aspect of a statistical study: "From past figures, it is predicted that 47% of the registered voters in Virginia will vote in the June primary"? 4) C
A) Description B) Design C) Inference

5) Planning the methods for data collection to study the effects of Vitamin E on athletic strength would be classified as which aspect of statistics: design, inference or description? 5) B
A) Inference B) Design C) Description

Section 1.2 We learn about Populations using Samples

The entities that we measure in a study are called **subjects**. (Usually a person or an object whose characteristics are of interest).

A characteristic being measured in a study is given the name **variable**.

Population and Sample

- **Population:** The total set of subjects in which we are interested.
- **Sample:** The subset of the population for whom we have (or plan to have) data.

Ex 2: We are interested in the average age of the students in Texas Tech.

With regard to this example, the entities we measure are the students. Therefore, the students are the subjects. The characteristic being measured is the age. This is a variable in this study.

The set of all the subjects of interest is the set of students in Texas Tech. Therefore the population is “students in Texas Tech”.

However, it is not feasible to find the age of all the students in Texas Tech. Therefore we select a random set of students from Texas Tech to find the average age of the group. This set of students, which we actually collect data from, is a sample. For instance, this class is could be a sample.

Descriptive Statistics and Inferential Statistics

Descriptive statistics:

Descriptive statistics refers to methods from summarizing the data. The summaries usually consist of graphs and numbers such as averages and percentages.

Inferential statistics:

Inferential statistics refers to methods of making decisions or predictions about a population, based on data obtained from a sample of that population.

Sample statistics and Population parameters

- **Parameter:** A numerical summary of the population
- **Statistic:** A numerical summary of a sample taken from the population

Ex 3: Referring back to ex 2, the numerical summary we are interested in the population is the average age.

Therefore the average age of the students of Texas Tech is the parameter.

The same numerical summary with regard to a sample is called a statistic.

For instance, the average age of this class is a statistic.

Randomness and Variability

If a sample statistic is to be a good reflection of the population parameter, the sample has to be random. In other words, each subject in the population has to have the same chance of being included in the sample. This is the basis of **random sampling**.

Ex 4: We are interested in finding out how concerned people in Texas are about global warming. A survey was done in several colleges in Texas and found out that 40% of the respondents were concerned about global warming. Is this a good reflection of the percentage of people in Texas who are concerned about global warming?

Ex 6: Suppose you need to find the average height of Texas Tech students. You take the basketball team for a sample and find that the average height it 6’ 2”. Is this sample statistic a good reflection of the population parameter?

Variability in this regard refers to how the sample statistics vary from sample to sample.

Ex 6: For instance, suppose that we want to find the average age of the Texas Tech students. There are a lot of possible samples one could choose. Consider two of these samples; for instance, section 6 and section 24 of Math 2300 in spring 2010. Do you expect the sample statistics (average age in each) for the two classes to be the same?

MULTIPLE CHOICE. Choose the one alternative that best completes the statement or answers the question.

Answer true or false.

1) Parameter values are usually known. 1) B

A) True B) False

Fill in the blank.

2) The entities that we measure in a study are called the _____. 2) B

A) parameters B) subjects C) data D) statistics

Determine whether the summary measure is better described as a parameter or a statistic.

3) The average gpa of Harvard's 2007 graduating class. 3) B

A) Statistic B) Parameter

4) The proportion of teenagers in a nationwide survey who stated that they consumed alcohol on a regular basis. 4) A

A) Statistic B) Parameter

Provide an appropriate response.

5) A survey of 1500 American households found that 33% of the households own a computer. Identify the sample. 5) B

A) All American households owning a computer
B) The 1500 American households surveyed
C) 33% of American households
D) The collection of all American households
E) The 33% of the 1500 households sampled that own a computer

6) A survey of 1500 American households found that 33% of the households own a computer. Identify the population. 6) D

A) 33% of American households
B) The 1500 American households surveyed
C) The 33% of the 1500 households sampled that own a computer
D) The collection of all American households
E) All American households owning a computer

Answer true or false.

7) A lobbyist for a major airspace firm wants to get the opinion of state legislators on an issue of importance to his industry. The lobbyist contacts 10 state legislators at the Top of the Mark Restaurant during a lunch break and each one is polled about the issue. This technique produces a random sample. 7) B

A) True B) False

Select the most appropriate answer.

8) The estimation of the population average age of registered voters in the state of Ohio based on the sample average age of 1,000 registered voters in that state and its corresponding margin of error is an example of 8) A

A) inferential statistics.
B) a statistic.
C) a parameter.
D) deductive statistics.
E) descriptive statistics.