



East Los Angeles College Math 227 Statistics

Instructor: Gabriel Castro
Winter 2011
1:00 – 4:30pm

1



Chapter 1

Introduction to Statistics

2

STATISTICS???

- A collection of methods for planning studies and experiments, obtaining data, and then analyzing, interpreting, organizing, presenting, and drawing conclusions based on the data that was collected.
- Data:
Observations that are collected such as responses and measurements.

3

Types of Data:

Example of Measurements:

Scores on an exam worth 100 points

78 81 98 56 34 100 66 12 89 5 65 68

*Note: A single value in a data set is called a datum
(or data point)

Examples of Responses:

Yes or No, True or False, Male or Female, and
Ethnicity: Asian, African American, Caucasian,
Latino and Native American

4

Types of Statistics

- Descriptive Statistics:

Methods for organizing and summarizing information. These methods consist of tables, charts, graphs and calculations such as averages and percentages.

We will primarily discuss this type of statistics in Chapters 1 – 6

- Inferential Statistics:

Gathers data from a sample to make inferences about the larger population from which the sample was drawn.

We will primarily discuss this type of statistics in Chapters 1 – 6

5

Let us define and differentiate some terms that play a big role in statistics.

- Population:

The entire group of subjects or elements about which inferences will be made.

- Sample:

A subset or sub collection of subjects or elements selected from a population. These subjects are actually measured or surveyed.

- Census:

A collection of data from every subject or element of the population.

6

A student in Math 227 conducts a research project about how college undergrads use Facebook. She begins with a survey mailed to 100 people she knows. She asks them to mail back a response to this question: “Do you use Facebook for networking, communicating with friends, or finding a romantic relationship?” She receives 60 responses, with 35 replying communicating with friends, 20 replying networking and 5 replying romantic relationship.

Identify the sample, population and determine whether the sample is representative of the population.

Sample: The 40 responses

Population: College Students

The sample is not representative of the population.

7

Some important terms when differentiating samples and populations.

■ Parameter:

A descriptive measure about the population.

■ Statistic:

A descriptive measure about the sample.

Examples:

1. Females make up 60% of the ELAC student body.
 1. Parameter.
2. The average cost of books is for 25 Students is \$96.47.
 2. Statistic.
3. The current Senate of the United States consists of 87 men and 13 women.
 3. Parameter.

8

Classifying Data

- Qualitative Data:

Also known as categorical data, it can be separated into distinct categories with respect to a particular, non-numerical characteristic.

- Quantitative Data:

Numerical data representing measurements or counts.

Examples:

1. Race = *Qualitative*
2. Weight = *Quantitative*
3. Age = *Quantitative*
4. Color = *Qualitative*

9

Types of Quantitative Data

- Discrete Data:

Data with values that is either a finite number or a countable number. "Observed Values"

Examples:

1. Number of cars in the student parking lot.
2. Shoes sold at a retail store.

- Continuous Data:

Data which results from infinitely many possible values between any two specific values without any gaps.

Values that **can not be counted**.

Examples:

1. Amount of money made in a two week pay period.
2. Ounces of soda in a can of Diet Coke.

10

Levels of Measurement

■ Nominal Level:

Data that is made up of categories or names with no ordering system.

Examples:

1. Colors of t-shirts in a store.
2. Zip codes.
3. Response of Yes or No.
4. Political Affiliation.

■ Ordinal Level:

Data that can be ranked or ordered, but the difference between ranks mean nothing.

Examples:

1. Size of t-shirts in a store.
2. Letter Grades.
3. Movie ratings.
4. Outcome of a race.

11

■ Interval Level:

The data is like the ordinal level, but two additional properties must be met:

- Differences between ranks are meaningful.
- No natural starting point. Zero is meaningless.

Examples:

1. Days.
2. Temperature.

■ Ratio Level:

The data is like the interval level; however, additional properties are:

- Natural starting point.
- Ratios, just like differences, are meaningful.

Examples:

1. Weights.
2. Income.

12

■ Difficulty arises when working with the interval and ratio levels of measurement. To simplify that distinction, use a simple “ratio test:” Consider two quantities where one number is twice the other, and ask whether “twice” can be used to correctly describe the quantities.

Example:

A person that earns \$20 per hour earns *twice* as much money as a person who earns \$10 per hour, but water that is 100°F is *not twice* as hot as water that is 50°F. Income is at the ratio level while Fahrenheit temperatures are at the interval level.

13

Design of Experiments

■ Observational Study vs. Experiment

In an observational study, we observe and measure specific characteristics without any modifications. With an experiment, we will first apply some **treatment** then observe its effects.

■ *Confounding:*

Ruins experiments because the experimenter is not able to differentiate between the effects of different factors. Plan the experiment so that confounding does not occur.

14

Types of Studies

- Cross-sectional Study:

Conducted at a single point in time or over a short period of time. No follow-up.

- Retrospective Study:

Data are collected from the past by going back in time through examination of records, interviews, and so on. Usually of subjects having a particular outcome are identified and studied.

- Prospective Study:

A cohort study that follows over time a group of similar individuals (a cohort) who differ with respect to certain factors under study, in order to determine how these factors affect rates of a certain outcome.

15

Sampling Methods

- Random Sample:

A sampling technique where we select a sample for study. Each individual is chosen entirely by chance and each member of the population has a known chance of being included in the sample.

- Simple Random Sample (SRS):

The basic sampling technique where we select a group of subjects for study. Each individual is chosen entirely by chance and each member of the population has an equal chance of being included in the sample. Every possible sample of size n , n is a natural number, has an equal chance of being selected.

16

Example of a Simple Random Sample:

We have seen that when students want to add a class the Instructor conducts a lottery.

Say there are 80 students who want to add the class and there are only 15 seats available. A way to select a *simple random sample* is to take 80 ping-pong balls and number them from 1 to 80. Put them into a large barrel and mix them up, and then select 15 ping-pong balls all at one time.

If we selected the ping-pong balls one at a time, then it would be considered a *random sample*.

*NOTE: In this case we must consider selecting with replacement or selecting without replacement. We will further discuss this in chapter 4.

17

Sampling Methods

Systematic Samples:

Each member in the population is numbered and every k^{th} element is selected.

Stratified Samples:

Subdivide the population into at least 2 different groups according to a particular characteristic, then draw a sample from each group.

18

Sampling Methods

Cluster Samples:

Divide the population into sections (usually areas) called clusters, then randomly select some of those clusters and then select all members from those clusters.

Convenience Samples:

Take samples that are readily available. Use results that are very easy to get.

19

Examples:

1. An instructor wants information on ELAC students transferring to a university. She gathers data by a show of hands from all her students in her Sociology class.

Convenience.

2. At Vons supermarket, the checkers give out surveys to every ninth customer that they assist.

Systematic.

3. The high school principle divides the school into groups, 14, 15, 16, 17, 18 and 19 year olds then draws a sample from each age group.

Stratified.

4. The high school principle divides the school in groups of 14, 15, 16, 17, 18 and 19 year olds then randomly selects 3 of those groups and surveys all of its members.

Cluster.

20