MATH 3070 Applied Statistics I

Nan Shen

Lec 1

Disclaimer: *These notes may NOT be distributed outside this class without the permission of the Instructor.*

1 Overview and Descriptive Statistics

What is Statistics?

visual or quantitative

- Methods to describe/organize/summarize "data"

- Methods to draw conclusions from "info" in data

1.1 Populations, Samples, and Processes

An investigation will typically focus on ___well define set of objects, called a population.___

There are two basic methods for studying a population:

- Census: collecting data on the whole pop.

- Sampling: collecting data on a subset of pop.

What is a variable?

A variable is ___a characteristic of obj. in pop.___

1

Note:

- Variable denoted by letters: x, y, ...

- Actual values make up the data set

Data results from making observations either on a single variable or simultaneously on two or more variables.

A **univariate** data set = each obj. in pop. is given one value

- e.g. height of a student (pop. = students at U)
  5-2, 5-5, 6-1, 4-5, 5-6, ... ← actual data

A **bivariate** data set = obj. → two char. val.

- e.g. (height, weight)

A **multivariate** data set = obj. → a number of char. val.

- e.g. (height, weight, age, gender, major)

Types of variables:

- Categorical: char. is from a list of (usually non-numerical) values ⎤
- Quantitative: char. is a number ⎥

  e.g.
  eye color
  major

quantitative

— Discrete: <u>Possible values are finite or **countably infinite**</u>

<u>e.g. # cars in household (0,1,2,3, ···)</u>

— Continuous: <u>Variable can take any value</u>

<u>in same interval.</u>

Summary: <u>E.g. age, weight</u>

Branches of Statistics:

1. **Descriptive Statistics**, <u>describe data</u>

   A. Some of these methods are <u>graphical</u> in nature; e.g. <u>histogram</u>

<u>box plot, pie charts, scatter plot</u>

   B. Other descriptive methods involve calculation of <u>numerical</u> summary measures,

e.g. <u>mean, standard dev median, etc</u>

2. Having obtained a sample from a population, an investigator would frequently like to

<u>draw conclusions about population</u>

Techniques for generalizing from a sample to a population are gathered within the branch of

our discipline called <u>inferential stats</u>

sample → stats → population

probability

3