

MATH 155 Modern Statistics with Computer Analysis

I. What are our preconceptions?

When I think of statistics

Statistics is

II. What kinds of situations might we wish to consider?

Situation 1: We have found data listing all male and female best actor/actress Oscar winners from 1928 to 2007 inclusive.

Situation 2: Of a random sample of 22 learners, 10 were taught by the new method and 12 were taught by the standard method. We have the results of a reading test given to the 22 learners at the end of a learning period.

III. What are some important statistical terms?

Statistics is the science of data. This involves collecting, classifying, summarizing, organizing, and interpreting numerical information.

Descriptive statistics utilizes numerical and graphical methods to look for patterns in a data set, to summarize the information revealed in a data set, and to present that information in a convenient form.

Inferential statistics utilizes sample data to make estimates, decisions, predictions, or other generalizations about a larger set of data.

Statistical methods are useful for studying *populations* of *experimental units*.

An *experimental unit* is an object (e.g. person, thing, event) upon which we collect data.

A *population* is a set of experimental units that we are interested in studying.

A *variable* is a characteristic or property of an individual population unit.

***Measurement* is the process we use to assign numbers to variables.**

When we measure a variable for each unit in a population, it is called a *census* of the population.

A *sample* is a subset of the units of a population.

A *statistical inference* is an estimate, prediction, or some other generalization about a population based on information contained in a sample.

A *measure of reliability* is a statement about the degree of uncertainty associated with a statistical inference. (How good is the inference?)

***Quantitative data* are measurements that are recorded as naturally occurring numerical values.**

***Qualitative data* are measurements that cannot be measures on a natural numerical scale; they can only be classified into one of a group of categories.**

IV. What are the elements of descriptive statistical problems?

(See p. 11 in text.)

V. What are the elements of inferential statistical problems?

(See p. 11 in text.)

VI. How do we obtain our data?

We obtain data from
a published source
a designed experiment
a survey
an observational study

Usually the data we collect is from a sample from some larger population. If we wish to make meaningful inferences, we must have a representative sample.

A representative sample exhibits characteristics typical of those possessed by the target (larger) population.

The most common way to obtain a representative sample is to select a random sample.

A random sample of n experimental units is a sample selected from the population in such a way that every possible different sample of size n has an equal chance of selection.

What can go wrong in a statistical study?

Selection bias (Some experimental units excluded in selection process)

Nonresponse bias (Data missing on some units in the sample)

Measurement error (Inaccuracies in data recorded, ambiguous or leading questions, interviewer effects)