

Math 122
Introduction to Statistics
Terminology and Introduction

WHAT IS STATISTICS?

Statistics is a collection of tools for

- Collecting data,
- Analyzing the data, and
- Making decisions based on the data.

Rare Event Rule: If a given assumption makes an observed event seem extremely unlikely then that assumption is probably false.

Example: If a coin is assumed to be fair, then it would be very unlikely to see 90 heads out of 100 flips of the coin. If I do flip the coin 100 times and see 90 heads, the coin is probably not fair.

DATA AND STUDIES.

Data are collections of observations such as measurements, survey responses, etc. Data are collected in observational studies and experiments.

Observational Study: Characteristics are observed or measured without trying to affect the subjects being studied.

Examples: Counting squirrels in a park. Measuring heights of students. Counting how many cars stop at an intersection. Asking an opinion without discussing an issue.

Experiment: A treatment is applied to the subjects being studied, and its effects are observed or measured.

Examples: Administering a medication and observing the results. Lecturing on an issue and then asking an opinion.

Quantitative Data consists of numbers representing actual counts or measurements.

Examples: Height, weight, number of siblings, hours of sleep, salary, pulse.

Categorical Data consists of names or labels that do not come from actual counts or measurements. Categorical data often separates subjects into groups.

Examples: Gender, name, home state, room number, eye color.

Continuous Data is quantitative data which comes from a set of values that covers a range without gaps or jumps.

Examples: Height, weight, hours of sleep.

Discrete Data is quantitative data which comes from a set of values which are isolated from each other by gaps or jumps.

Examples: Number of siblings, salary, pulse.

POPULATIONS AND SAMPLES

Population: The population is the collection of all individuals being studied.

Sample: A sample is a set of some members of the population.

Census: A census is a collection of data from every member of a population.

Parameter: A parameter is a measurement based on a census of the entire population.

Statistic: A statistic is a measurement based on a sample of the population.

Statistics Approximate Parameters: Often, a statistic from a sample may be a good approximation of the corresponding population parameter.

THINGS TO CONSIDER IN STATISTICAL ANALYSIS.

Context/Type of data affects what type of statistical analysis can be used.

Sampling Method: Bad sampling methods can render statistics useless. The worst mistake that can be made in sampling is to use a self-selected sample. Small samples can also give meaningless data.

Source: Bias or corruption of data can render statistics useless.

Conclusions: Statistical tests can only render certain types of conclusions. Some types of claims can only be supported (or not). Others can only be rejected (or not).

Practical Implications: Most statistical conclusions need practical interpretations. Not all statistically significant results are practically significant.