

STATISTICS – Prof. Richard B. Goldstein

Statistics is the study of how to collect, organize, analyze, and interpret numerical information from data

Data (unorganized) → Information (organized)

collect	experimental results, simulations, census, surveys, samples
organize	tabulate and tally, order, classify, graphics
analyze	<i>descriptive statistics</i> that summarize information, distributions
interpret	<i>inferential statistics</i> which draw conclusions about the population based upon the information from the sample

Variables

Qualitative	Nominal	names, colors, shapes, descriptive data
Quantitative	Ordinal	ordered best to worst, rank order by quality, age groupings (toddler, etc.)
	Interval	ordered with meaningful differences such as the calendar year, time of day, centigrade or Fahrenheit temperature
	Ratio	has a meaningful ratio and zero such as lengths, weights, heights, blood values, exam scores, Kelvin temperature

Samples

Population	every	- refers to <u>all</u> measurements or observations of interest – described by parameters
Sample	some	- subset of an existing population – described by statistics
Random		every sample of size n measurements is <u>equally likely</u> and every member of the population is <u>equally likely</u> to be included
Systematic		the population is arranged in some natural sequential order and starting somewhere one takes every k^{th} value
Cluster		break down the population into sections and take all or some of these sections
Stratified		the population is broken down by characteristics such as age, ethnic groups, etc. and samples are taken in proportion by each group
Convenience		data is collected because it is readily available such as a convenient location
Multistage		uses a variety of methods to create smaller groups which become clusters