**Colorado Technical University**

 **Course:** MATH366 – Probability and Statistics

#### Unit 1 Part 01 Readings: Introduction, Sampling, Types of Data

**Introduction to Statistics**

Why do I have to take *STATISTICS* – terror of all college

students, bane of late-night comics???

Although maligned as "bean counters", statisticians do

more than just count, and rarely have anything to do

with beans (except eating them…)

Statistics is the basis for all the scientific claims you

hear about from Darwin's theory of Evolution to

commercials about the superiority of a certain brand

of aspirin. The government stands (nervously) on

Thanks to Jonathan Bellamy

for the stupendous

Stat Demon drawing !

statistical projections.

While few people will actually complete their own

statistical analysis, everyone is affected by statistics,

for good or ill.

Because it is true that statistics can be used to support anything (true or false),

everyone in the twenty-first century should become skeptical consumers of statistical information.

In this course we will learn how to do statistics, yes, but we will also learn how pervasive statistical lying is, and how to recognize a potential lie when we see one.

numbers DO lie !

statistics is the mathematics of uncertainty

it is the mathematics of our uncertain world

it is the mathematics of making decisions

**Sampling**

**statistics** - collecting, organizing, analyzing, and interpreting data in order to make

decisions

**descriptive statistics**: The branch of statistics that involves organization,

summarization, and display of data

**inferential statistics**: The branch of statistics that involves the use of a sample to

draw conclusions about a population

**population** - a collection of **all** outcomes, responses, measurements, and counts

that are of interest - the UNIVERSE

**parameter** - the numerical description of a characteristic of a population

**census** - a count or measure of an entire population

**sample** - a subset of a population

you want a sample that is **representative** of the population and

**unbiased** so your statistic is a good close estimate of your parameter of interest

**statistic** - the numerical description of a characteristic of a sample

To help ensure a representative sample:

 Make sure every member of the population has an equal chance of being included -

a **random sample**

Take as large a sample as you can - **the Law of Large Numbers**

**Types of data**

**qualitative -** attributes, labels, or nonnumeric entries, ex: gender—male or

female

**quantitative -** numerical measurements or counts, ex: age—1, 2, 10, or 20

**Other types of data:**

Qualitative/quantitative – Some data can be described using a mathematical

formula/Others are so complicated we do not yet have a formula that works

Observation/experiment

Deterministic/nondeterministic

Periodic/nonperiodic – Some data cycle around/Others do not

Continuous/discrete – Some data can be ANY decimal value/Some data can only be

whole numbers

Stationary/nonstationary – Some data have a constant “average” value/Some data

have trend or displacement

**Best descriptor statistics:**

 for qualitative data, use the “mode” (most frequently occurring value)

 for quantitative data, use the traditional average (called an “arithmetic mean”)

**Levels of measurement:**

**nominal -** name only

**ordinal -** an ordering makes sense

**interval -** a meaningful spacing

**ratio -** you can compare one to the other as a percent or ratio

**A weighted average can be used as a best descriptor for ordinal or interval data**

**Interval data with 5 or more categories can be treated as ration**

**Three criteria for good measurement:**

 **reliability** - an indicator of internal consistency; a measure is reliable when

different attempts at measuring something converge on the same result

 **validity** - the accuracy of a measure or the extent to which a score truthfully

represents a concept

 **sensitivity** - the ability to accurately measure the subtle variability in a concept