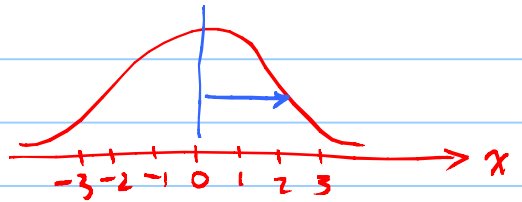


# Lecture 1 (CR.1)

Statistics (at this level) is NOT Math!  
It is extremely ambiguous!

E.g. How wide is this curve?  
About 2 (!) ← not factorial.



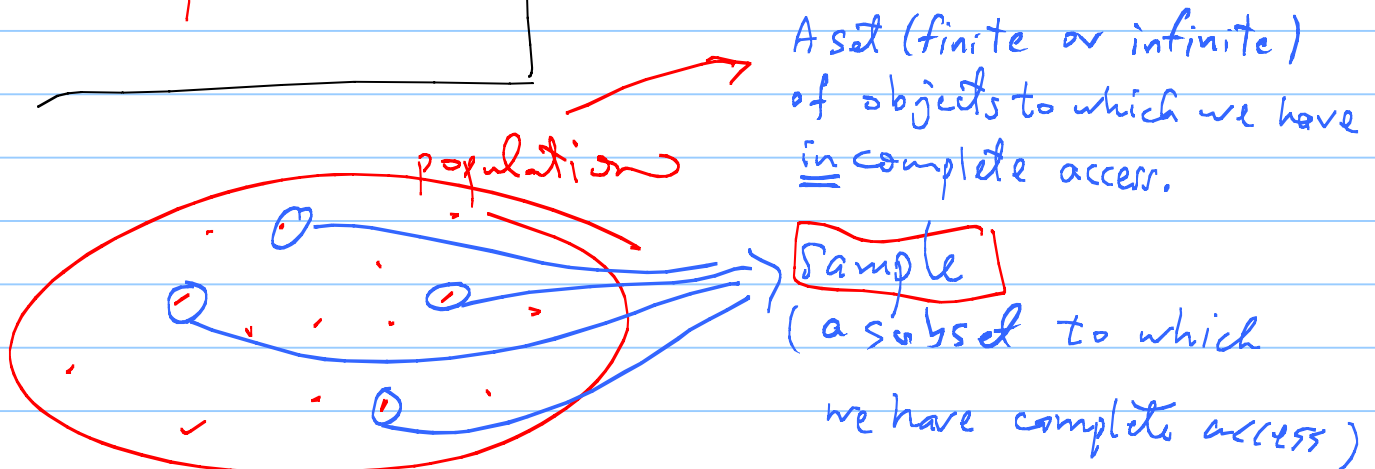
## Two Types of Statistics:

### Descriptive

mean  
median  
mode  
range  
histogram  
scatter plot

### Inferential

To infer something  
about a population  
from a single sample.



E.g. From a sample, we can compute sample mean, sample range, sample ...

What do these say about the population mean?

The "coordinates" of the elements in both pop. and sample are generally called variables. But there are 2 types of variables that we (in 390) worry about

1) continuous : e.g.  $x$  = time it takes to complete a comp. code.  
 $x \in \mathbb{R}$

2) discrete / Categorical :

a) quantitative : e.g.  $x$  = # of defective elements in a <sup>random</sup> computer.  
 $x \in \text{Integers } (0, 1, 2, \dots)$

e.g.  $x$  = # of Macs in a class of 100 students  
 $x \in \{0, 1, \dots, 100\}$  random

b) qualitative : e.g.  $x$  = computer type. called "levels".  
 $x \in \{\text{Mac, Dell, HP}\}$

e.g.  $x$  = state of a coin  
 $x \in \{\text{Heads, Tails}\}$

e.g.  $x$  = Letter grades in a class of 120 students  
 $x \in \{A, B, C, D, F\}$

r.v.

random variable. This is a very important concept in statistics. All we need to know about it is that it is a variable (e.g. length, time, type, ...) that changes values everytime we observe/measure it from a sample.

## hw\_lect1

Construct a data set with The following specifications.  
Any source is allowed: web, books, papers, your own work, etc. Specifications:

- 1) number of cases: 30 or more
- 2) 2 categorical/discrete variables. One of them can have between 2 and 6 levels, and The other must have between 3 and 6 levels.

See part b) for a requirement on the histograms.  
You may want to wait until after the lect on histograms.

- 3) 2 Continuous variables.

- 4) The 4 variables must relate to a common problem; not 4 unrelated variables.

a) print the data in The following format, and turn it in:

	Variable 1	Variable 2	Variable 3	Var. 4
30 or more cases {	:	:	:	:
	:	:	:	:
	:	:	:	:
	:	:	:	:

Keep a copy of The data set because you will need it for other hw problems.

- b) plot histograms for each of The 4 variables. (By hand)  
For The Continuous vars. pick an appropriate # of bins.  
For The discrete vars. it is important for The hist to have at least 2 bars with more than 1 count.

This document was created with Win2PDF available at <http://www.win2pdf.com>.  
The unregistered version of Win2PDF is for evaluation or non-commercial use only.  
This page will not be added after purchasing Win2PDF.