**Topic 8**

**Measures of Center**

The feature of a distribution we most often care about is the notion of its center.

First, lets calculate some measure of center:

For a variable with n data values {x1,x2,x3, …, xn}

**Sample Mean** (or Average) = $\frac{(x\_{1}+x\_{2}+x\_{3}+…+x\_{n})}{n}$

 **Median** (physical center of the data)

 First, order the data from smallest to largest

If **n odd:** located in the position (n+1)/2

 If **n even:** mean of two values located at n/2 and (n+1)/2

 **Mode** the most frequent value(s) (also the highest point on the graph)

The last two are **physical** descriptions of the data. The first is a **mathematical** description. The mean is the most common of the three. This is remarkable because symmetric distributions, for which they should be restricted, are not so common.

There are other means:

Distribution is *symmetric*,

* Mean
* Median
* Mode

Distribution is *asymmetric*,

1. The asymmetry is triangular
	* Median
	* Mean
2. Asymmetry has a long tail.
	* Median

Distribution is *symmetric with outliers*

* Median
* Mode

Distribution is *asymmetric with outliers*

* Median

Distribution is *Multimodal*

* Mean *does not apply* in this situation

Mean is an **Unbiased Estimator,**

* its expected value is the same as the parameter it is estimating.
* As n->N, the estimator->parameter

Median and Mode are a **Resistant** **Estimators**

* it is not sensitive to extreme values or outliers

 Population **Mean** = $\frac{(x\_{1}+x\_{2}+x\_{3}+…+x\_{N})}{N}$