

## Lecture 6: Variability

### Variability

Central tendency locates middle of distribution

How are scores distributed around that point?

Low variability vs. high variability

Ways to measure variability

Range

Interquartile range

Sum of squares

Variance

Standard deviation

### Why Variability is Important

Inference

Reliability of estimators

For its own sake

Consistency (manufacturing, sports, etc.)

Diversity (attitudes, strategies)

### Range

Distance from minimum to maximum

$$range = \max(X) - \min(X) + \epsilon$$

$\epsilon$  is measurement unit or precision

Sample range depends on  $n$

More useful as population parameter

Theoretical property of measurement variable

E.g. memory test: min and max *possible*

Rough guidelines, e.g. height

### Interquartile range

Quartiles

Values of  $X$  based on dividing data into quarters

1st quartile: greater than 1/4 of data

3rd quartile: greater than 3/4 of data

2nd quartile = median

Interquartile range

Difference between 1st and 3rd quartiles

Like range, but for middle half of distribution

Not sensitive to  $n \rightarrow$  more stable

## Sum of Squares

Based on deviation of each datum from the mean:  $(X - \mu)$   
Square each deviation and add them up

$$SS = \sum_{pop} (X - \mu)^2$$

## Variance

Most sophisticated statistic for variability  
Sum of squares divided by  $N$   
Mean Square: average squared deviation

$$\sigma^2 = \frac{\sum_{pop} (X - \mu)^2}{N}$$

## Standard deviation

Typical difference between  $X$  and  $\mu$   
Again, based on  $(X - \mu)^2$

$$\sigma = \sqrt{\frac{\sum_{pop} (X - \mu)^2}{N}}$$

Variance is average *squared* deviation, so sqrt(variance) is standard deviation

## Why squared difference?

Could use absolute distances,  $|X - \mu|$   
Would be more intuitive: average distance from the mean  
Squares have special mathematical properties  
Can be broken into different parts

$$\sum_{pop} X^2 = \sum_{pop} (X - \mu)^2 + N\mu^2$$

On right side, first part represents differences among scores; second part represents central tendency of all scores