**Module I: Basics of Statistics**

Statistics for the Behavioral Sciences

**•Central Tendency**

•Mean, Median, Mode

•One value that best represents a variable

•Central to all the other values

**Normal Distribution**

**Positive Skew Distribution**

**Negative Skew Distribution**

**•Central Tendency**

•Mean, Median, Mode

•**Mean = arithmetic average**

•Balancing point of a distribution

•Symbolic notation

**•Mean = arithmetic average**

•Ex. Social planning: # of activities

2, 3, 1, 2, 4, 2, 1, 0, 1, 4

•What number best represents social activities?

•Maybe it is the mean (average)

•Symbolic notation for the mean

2, 3, 1, 2, 4, 2, 1, 0, 1, 4

**•Symbolic notation for the mean**

2, 3, 1, 2, 4, 2, 1, 0, 1, 4

**•Balancing point of the distribution**

2, 3, 1, 2, 4, 2, 1, 0, 1, 4

**•Central Tendency**

•Mean, **Median**, Mode

•Median = the middle value when in order

•Reorder numbers from smallest to largest

•Ex. Social planning: # of activities

2, 3, 1, 2, 4, 2, 1, 0, 1, 4

0, 1, 1, 1, 2, 2, 2, 3, 4, 4

•What is the middle value?

**•Median = the middle value when in order**

•Reorder numbers from smallest to largest

•Ex. Social planning: # of activities

2, 3, 1, 1, 4, 2, 1, 0, 1, 4

0, 1, 1, 1, 1, 2, 2, 3, 4, 4

•What is the middle value?

**•Median = the middle value when in order**

•Reorder numbers from smallest to largest

•Ex. Social planning: # of activities

2, 3, 1, 1, 4, 2, 1, 0, 1

0, 1, 1, 1, 1, 2, 2, 3, 4

•What is the middle value?

**•Central Tendency**

•Mean, Median, **Mode**

•Mode = the most frequent value

•Ex. Social planning: # of activities

0, 1, 1, 1, 2, 2, 2, 3, 4, 4

•What is the most frequent value?

**•Ex. Salaries 2005 communication majors**

•Mean = $60,000

•Median = $42,000

•Mode = $41,000

•Which value best represents the data set?

•Typically mean unless skewed

•Use mode when variable is discrete

**•Ex. Salaries 2005 communication majors**

•Mean = $42,500

•Median = $42,000

•Mode = $41,000

•Which value best represents the data set?

**•Central Tendency**

•One value that best represents a variable

•Best value is central to all others

•Mean = arithmetic average

•Median = the middle value when in order

•Mode = the most frequent value

**•Variability**

•Range, Variance, Standard Deviation

•Range

•Range vs. full range

•Difference between extremes of exact values

•Highest value – lowest value

**•Variability**

•Range, Variance, Standard Deviation

•Variance

•Indicator of the spread of the values

•Accounts for differing values from set point

•Variance = average squared deviations from the mean

**•Symbolic notation for variance**

**•Variance = avg. squ. dev. from the mean**

•Ex. Family appreciation scores

•What is the mean?

•What is N?

•What is the variance?

**Variance = avg. squ. dev. from the mean**

Step 1: Subtract the mean from each value

What happens if we sum the dev. scores?

Step 2: Square the deviations

Step 3: Sum the squared (SS) deviations

Step 4: Divide by the number of values

26/4 = 6.5

•Variance and reliability

•Larger variance = lower reliability

•Smaller variance = higher reliability

**•Problem with variance**

•No longer on the same scale as raw data

•Need to square root it

**•Variability**

•Range, Variance, **Standard Deviation**

•Symbolic notation for standard deviation

•**Standard deviation = the typical deviation of the raw scores from the mean**

σ = 2.55

(Note on rounding: How many decimal places?)

**•If we have the mean & stand. dev. then**

•Est. the normal distribution of the data in the population (Note: not a correction)

•Ex. Communications salary

•Mean = $42,500

•Standard deviation = $6,700

**Normal Distribution**

•Normal distribution can be modeled by the mean and standard deviation

•To calculate skewness

**Normal Distribution**

•Check normality by skewness stat and table

•Assume normal if between lower and upper limit

**Correcting Skew**

•Correction for a positive skew

•Take the square root of all the values

•Correction for a negative skew

•Square all of the values

•Once corrected, then calculate measures of central tendency and variability again