Measuring Data

1. For each of the following, indicate whether the data is measured on a nominal, ordinal, interval, or ratio scale.
    a. Heights of The University of Jordan college students.
    b. Amount of snow that falls in Amman in January.
    c. Languages spoken by JU students
    d. Waist size of men who attend weight watchers every week.
    e. Strength categories of hurricanes (1,2,3,4,5).
    f. Average lifespan of alcoholics.
    g. Amount of alcohol (in ounces) consumed by high school seniors.
    h. How often college students use mouthwash in a typical week.
    i. Scores on the GRE Exam (range is from 200 to 800)
2. Indicate which of the following are discrete measurements and which are continuous measurements: Discrete Continuous
    a. life of a Dell monitor
    b. Waist size of Jordanian football players.
    c. Mileage of Japanese cars
    d. Weight of human brains.
    e. number of left-handed people on basketball teams
    f. Time to complete the task of assembling a computer.
    g. Number of foreign students in each statistics class.
    h. Number of bedrooms in your home.
3. Indicate which of the following is a parameter and which is a statistic:
    a. Population mean Parameter Statistic
    b. Population standard deviation.
    c. Sample standard deviation.
    d. Population variance.
    e. Population median.
    f. Sample mean.
    g. A mean obtained from the U.S. census.
    h. A mean obtained from sampling 2000 American Adults.
4. For each of the following, indicate the appropriate statistical measures that may be used for analysis (proportions, mean, median, quartiles, mode…). List as many as are appropriate.
    a. For data measured on a nominal scale.
    b. For data measured on an ordinal scale.
    c. For data measured on an interval scale.
5. A Dell computer manufactures 1,000 super computers. A researcher wants a sample of 50 of them. Each of the computers is assigned numbers and then random numbers are used so that every computer has an equal chance of being selected (5%).
    a. This is known as a…..
    b. Measurements obtained from this are known as:….
    c. Suppose the researcher decides to sample 1,000 computers. This is called a…..
    d. Any measurement obtained is known as a……

Chi Square

1. Please critically analyze and critique Chi Square research report and make judgment on the accuracy of the statistical techniques employed on those report. It will be evaluated on clarity, relevance, comprehensiveness, completeness, and correctness.
2. Sometimes you will be working with data that have already been put into a table of frequencies, for the second question I'll show you how to get SPSS to do a χ^2 on such data. Say someone hands you some data where the frequencies have already been recorded in the following table and you want to have SPSS perform a test for association.

	Downhill skiing	Cross country skiing	Snow boarding	Total
Older	**14**	**30**	**11**	**55**
Younger	**25**	**10**	**42**	**77**
Total	**39**	**40**	**53**	**132**

Begin with a new data window. You need to create three variables, 'age', 'sport', and 'freq'. To do this go to the data window, and click on the 'Variable View' tab, there create each variable and set 'age' and 'sport' to Type 'String', and 'freq' to Type 'Numeric'.

After creating the variables go back to the 'Data View' and input the following data for your variables (note you will have six lines of data, with three values on each line).

age	sport	freq
Older	Down	14
Older	Cross	30
Older	Board	11
Younger	Down	25
Younger	Cross	10
Younger	Board	42

After putting in the above information, go to Data>>WeightCases and indicate that you want to weight cases by variable 'frequency'. Click Ok. Then analyze the relationship of age and sport as you did before: go to Analyze>>Descriptive Statistics>>CrossTab, ask for chi-square and Cramer's V.

1. χ^2 = _____
2. p = _____
3. State your decision regarding H0:
   - Do not reject H0
   - Reject H0
4. Assuming you have no confounding variables, what can you conclude?
   - Can conclude the variables are associated in the population.
   - Cannot determine whether or not the variables are associated in the population.
5. Cramer's V = _____

Sample Multiple Choice Questions

1. Measures of central tendency are:
    a. Inferential statistics that identify the best single value for representing a set of data.
    b. Descriptive statistics that identify the best single value for representing a set of data.
    c. Inferential statistics that identify the spread of the scores in a data set.
    d. Descriptive statistics that identify the spread of the scores in a data set.
2. Which of the following is not a characteristic of the mean?
    a. It is affected by extreme scores.
    b. It minimizes the sum of squared deviations.
    c. The sum of the deviations about the mean is 0.
    d. It is best used with ordinal data.
3. The sum of squared deviations is a good measure of variability except
    a. it uses the mean in its calculation.
    b. it can only be used with ratio data.
    c. it does not take N into account.
    d. it cannot be used in any other statistical analyses.
4. Which of the following is used to represent a known value for the population variance?
    a. s
    b. σ
    c. σ^2
    d. s^2
5. Which of the following is true regarding factors affecting variability?
    a. Greater variability is found in smaller samples.
    b. Random sampling results in distributions with large amounts of variability.
    c. The independent variable should not affect variability.
    d. An insensitive dependent measure will produce a lot of variability.

6. Name a sampling procedure in which initial respondents are selected by probability methods, and then additional respondents are obtained from information provided by initial respondents.
    a. Semi-random
    b. Initial
    c. Quota
    d. Snowball
7. A researcher divides the population of product users into three groups based on degree of use. If the researcher then draws a random sample from each user group independently, she has created a _____ sample.
    a. Random.
    b. Stratified.
    c. Judgment.
    d. group data
8. Suppose a researcher is concerned with a nominal scale that identifies users versus nonusers of bank credit cards. The measure of central tendency appropriate to this scale is the
    a. Mean.
    b. Median.

c. Mode.
d. Average.

9. The formula μ = 0 ± a sampling error
   a. Expresses the idea of the central limit theorem.
   b. Expresses the idea of the confidence interval.
   c. Expresses the idea that a sample frame is not a perfect representation of the population.
   d. None of the above.
10. Constructing a frequency distribution
   a. Is one of the most common means of summarizing data.
   b. Begins by recording the number of times a particular value occurs.
   c. Is the basis for construction of a percentage distribution.
   d. All of the above
11. The practical result of the central limit theorem is that
   a. Researchers must take a large number of samples before inferences about the population can be made.
   b. The researcher must know the shape of the population distribution before inferences about the population can be made.
   c. The concept of the sampling distribution is unimportant to researchers.
   d. none of the above
12. A frequency distribution (or probability distribution) of all possible values of sample means for a sample size of 200 is the
   a. Population distribution.
   b. Sample distribution.
   c. Sampling distribution.
   d. Standard normal distribution.
13. Hypothesis tests are designed so that the _____ hypothesis will be rejected.
   a. Null
   b. Alternative
   c. Alpha
   d. beta
14. When using the chi-square analysis in bivariate problems, degrees of freedom are calculated as (R=row and C=column)
   a. (R - 1)(C - 1).
   b. R x C + 1.
   c. (R + 1)(C + 1).
   d. R + (C - 1).