**Module 1: Homework Questions**

Please indicate your answers by highlighting your choice in yellow.

1. The research process is comprised of five stages. Which of the following is *not* one of those stages?

a. Data collection

b. Theory generation

c. Analysis

d. Dispersion

2. Looking at the distribution of data on the bar chart below, how best can it be described?

|  |
| --- |
| FCEs by age and gender for lung cancer (C33-C34) in 2006-07. |

Source: http://www.hesonline.nhs.uk/Ease/servlet/ContentServer?siteID=1937&categoryID=955

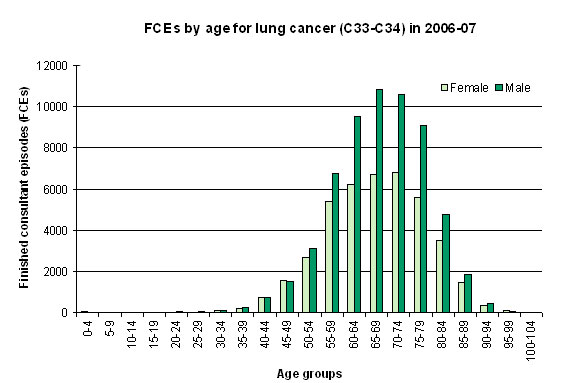
a. Normally distributed

b. Negatively skewed

c. Positively skewed

d. Leptokurtic

3. What does the data in the chart demonstrate?



a. The incidence of cancer is highest among women aged 55–59.

b. The incidence of childhood cancer is highest in girls.

c. Cancer affects everyone.

d. A higher incidence of cancer occurs in men across most age groups.

1. In this scale of measurement the magnitude of the difference between numbers is meaningful, and you can add and subtract but not multiply or divide since there is no true zero.
2. Nominal
3. Ordinal
4. Interval
5. Ratio
6. With ordinal data, a \_\_\_\_\_ is used to depict the frequency distribution.
7. Pie chart
8. Histogram
9. Polygon
10. Bar graph
11. With nominal data, the \_\_\_\_\_ is the measure of central tendency.
12. Mode
13. Median
14. Mean
15. Any of the above are appropriate for nominal data
16. With nominal data, a \_\_\_\_\_ may be used to depict the frequency distribution.
17. Pie chart
18. Histogram
19. Polygon
20. Median
21. The mean is the most common measure of central tendency for which types of measurement scales?
    1. Interval and ratio
    2. Nominal and interval
    3. Ordinal and ratio
    4. Nominal and ordinal
22. A specialized form of categorical variable that has only 2 categories is:
    1. A nominal variable
    2. A continuous variable
    3. An independent variable
    4. A dichotomous or binary variable

Questions 10-13 relate to the Berg Balance Scale that we often use to assess functional balance.

1. The Berg Balance scale rates a patient’s performance on each of 14 items using a scale from 0 to 4 then sums the 14 items for a total score. Each of the 14 items on the Berg Balance scale utilizes what level of measurement?
   1. Nominal
   2. Dichotomous
   3. Ordinal
   4. Ratio
2. The most appropriate measure of central tendency when looking at a frequency distribution portraying the scores of 100 patients on one particular item on the Berg Balance scale is:
   1. Mode
   2. Median
   3. Mean
   4. Range
3. The most appropriate measure of central tendency for a dataset of Berg Balance *total scores* is:
   1. Mode
   2. Median
   3. Mean
   4. Median or mean
4. Which of the following is/are true about the median as a measure of central tendency?
   1. It is relatively unaffected by extreme scores at either end of the distribution of scores
   2. It can be used with ordinal, interval and ratio data
   3. It is relatively unaffected by skewed distributions and therefore useful in nonparametric statistics
   4. All are true
5. Which of the following are measures of dispersion of data in a frequency distribution?
   1. Standard deviation
   2. Interquartile range
   3. Range of scores
   4. All are measure of dispersion
6. Which of the following are measures of dispersion of scores around the mean?
   1. Standard deviation
   2. Interquartile range
   3. Range of scores
   4. All of the above

1. The two most commonly used measures of variability with normally distributed interval and ratio data are:
2. Standard deviation and variance
3. Range and variance
4. Interquartile range and variance
5. In a set of scores, the sum of the deviations from the mean will equal \_\_\_\_\_.
6. 1
7. Depends upon the number of scores
8. 0
9. Cannot be determined
10. A non-symmetrical distribution that has more than one mode is called \_\_\_\_\_.
11. Leptokurtic
12. Negatively skewed
13. Multimodal
14. Positively skewed
15. What is the median and range for temperatures of 91, 92, 93, and 94?
    1. 92.5; 3
    2. 92.5; 4
    3. 92; 3
    4. 93; 4
16. Which of the following statements is accurate regarding research design methods?
    1. Repeated measures designs minimize systematic variation
    2. Repeated measures designs minimize unsystematic variation
    3. Independent or between group designs minimize systematic variation
    4. Independent or between group designs minimize unsystematic variation
17. The more varied the scores in a distribution, \_\_\_\_\_\_\_\_\_\_\_\_\_.
    1. The larger the standard deviation will be
    2. The smaller the standard deviation will be
    3. The variance in a distribution have no effect on standard deviation
18. Which of the following represents variation due to random effects such individual ability, motivation, etc.)?
    1. Systematic variation
    2. Unsystematic variation
    3. Variance
    4. Any variation represents measurement error
19. When looking at quartiles in a frequency distribution, which of the following is true?
    1. The second quartile value is the median score
    2. The second quartile value is the mean score
    3. The third quartile value is the median score
    4. The third quartile value is the mean score

Now let’s try an application. This course will not focus so much on calculation but I think it is important that you understand the concept of variance, standard deviation and sum of squares because it is so vital to inferential statistics.

You have a sample of 5 patients seen by their physician (Dr. Timely). Note that it is a “sample” because these 5 patients represent the larger population of patients seen by this particular physician in one year (more on this later!).

The following scores represent the amount of time spent by Dr. Timely with each of the 5 patients in the dataset.

|  |  |  |
| --- | --- | --- |
| **Time in**  **minutes (xi)** | **Deviance (xi-)** | **Deviance squared**  **(xi-)2** |
| **15** | **-17** | **289** |
| **15** | **-17** | **289** |
| **45** | **13** | **169** |
| **40** | **8** | **64** |
| **45** | **13** | **169** |
| **Sum=160** |  | **Sum of squared deviances= ?**  **980** |

Fill out the columns above for Deviance, deviance squared, and sum of squared deviances. To avoid confusion, please do not record your answers in red since red is what will be used to grade your answers.

1. What type of variable is “time in minutes” in the above table?
   1. Nominal/discrete
   2. Ordinal/discrete
   3. Interval/continuous
   4. Ratio/Continuous
2. What is the mean of the patient scores? \_\_\_32\_\_\_\_
3. What is the sum of all deviance scores? \_\_\_0\_\_\_\_
4. What is the sum of squared deviance scores? \_\_\_980\_\_\_\_
5. What is the variance (s2)? 245
6. What is the standard deviation (s)? 15.65
7. What is the value for sum of squares (SS)? 980
8. A standard normal curve has a mean of \_\_\_\_\_\_\_\_ and a standard deviation of \_\_\_\_\_\_\_\_\_.
9. 1; 0
10. 0; 1
11. 10; 100
12. None of the above. It can have any mean and standard deviation.
13. A z-score of +3 means that the data point is:
    1. 1/3 of the mean
    2. 3 times the mean
    3. 3 standard deviations above the mean
    4. 3 standard deviations below the mean
14. On an IQ test, scores are normally distributed with a mean of 100 and standard deviation of 15. Knowing this information, calculate the percentage of people that will score below 90. Use Appendix A.1 in Field. Please show your work.

(90-100)/15= -0.667 or -0.67

z= 0.25143

(0.25143)\*100= 25.143%

1. Using the information on IQ scores in #33, what percentage will score between 90 and 130? Again, use Appendix A.1. and please show your work.

(130-100)/15=2.00

Z= 0.97725

(0.97725)\*100= 97.725%

97.725%-25.143%= 72.582%

1. In a population of 10,000 people who take the IQ test, how many would you expect to score above 130? Please show your work.

100%-97.725%= 2.275

10000\*0.02275= 227.5 people