UNIVERSITY OF SWAZILND

FACULTY OF HEALTH SCIENCES

FINAL EXAMINATION PAPER, MAY 2011

TITLE OF PAPER: HEALTH STATISTICS

COURSE CODE: HSC 404

TIME ALLOWED: TWO (2) HOURS

MARKS: 75

INSTRUCTIONS:

1. THERE ARE THREE (3) QUESTIONS IN THIS PAPER
2. ANSWER ALL THREE QUESTIONS
3. EACH QUESTION IS ALLOCATED 25 MARKS
4. WRITE LEGIBLY
5. **All FINAL ANSWERS MUST BE TO THE NEAREST 2/10, SHOW ALL YOUR CALCULATIONS.**

THIS PAPER IS <u>NOT TO BE OPENED</u> UNTIL THE INVESTIGATOR HAS GRANTED PERMISSION.

QUESTION 1

You collected blood pressure (mm Hg) from your 23 study participants. Below are the diastolic pressures that your obtained.

74	56	76	84	62	80	56	82	60	60	72
72	65	72	60	76	72	58	86	80	62	76
86										

A. Develop a table, group the data into class intervals with widths of 5. The apparent lower limit of the first class interval must be 55 mmHg. (5)

For the above data set, compute the:

B. Median (3)

C. Mean (3)

D. Standard deviation. (4)

E. Position of 55 percentile (4)

F. Skewness, and intepret (3)

G. Coefficient of Variation (3)

TOTAL 25 MARKS

QUESTION 2

A study was conducted among participants to determine the likelihood of developing lung cancer. The data of the participants' smoking history (independent variable) was correlated with the participants' overall health (dependent variable), the results are presented below.

Correlations

		Smoking History	overall state of health
Smoking History	Pearson Correlation	1	-.200(**)
	Sig. (2-tailed)	.	.000
	N	442	441
overall state of health	Pearson Correlation	-.200(**)	1
	Sig. (2-tailed)	.000	.
	N	441	444

** Correlation is significant at the 0.01 level (2-tailed).

A. Present the results from the correlation output. (4)

B. What is the meaning of the results? (3)

C. Who (which) are the four (4) sons of Pearsons r, specify when each of the four sons is utilized. (8)

Below is the linear regression model between the independent variable, age, and the dependent variable, to obey.

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.160(a)	.026	.025	1.383

a Predictors: (Constant), Age of Respondent

ANOVA(b)

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	49.301	1	49.301	25.763	.000(a)
	Residual	1873.433	979	1.914		
	Total	1922.734	980			

a Predictors: (Constant), Age of Respondent
b Dependent Variable: To Obey

D. Present (the results) and interpret the ANOVA output. (5)

E. Compute the Coefficient of Determination, and describe its meaning? (5)

TOTAL 25 MARKS

QUESTION 3

Compute the answer as required OR clearly write the letter that corresponds with the most appropriate answer to the statement /question.

1. _____ results if you fail to reject the null hypothesis, when the null hypothesis is actually false.

   A. Type I error

   B. Type II error

   C. Type III error

   D. Type IV error

2. A graph that uses vertical bars to represent data is called a _____.

A. line graph

B. Bar graph

C. Scatter plot

D. Vertical graph

3. The goal of _____ is to focus on summarizing and explaining a specific set of data.

A. inferential statistics

B. descriptive statistics

C. None of the above

D. All of the above

4. As a general rule, the _____ is the best measure of central tendency because it is more precise.

A. Mean

B. Median

C. Mode

D. Range

5. Focusing on describing or explaining data versus going beyond immediate data and making inferences is the difference between _____.

A. Central tendency and common tendency

B. Mutual exclusive and mutual exhaustive properties

C. Descriptive and inferential

D. Positive skew and negative skew

6. Why are variance and standard deviation the most popular measures of variability?

A. They are the most stable and are foundations for more advanced statistical analyses

B. They are the most simple to calculate with large data sets

C. They provide nominally scaled data

D. None of the above

7. _____ is the set of procedures used to explain or predict the values of a dependent variable based on the values of one or more independent variables.

A. Regression analysis

B. Regression coefficient

C. Regression equation

D. Regression line

8. Which of the following is NOT a common measure of central tendency?

A. Mode

B. Range

C. Median

D. Mean

9. What is the median of the following data set? 18, 14, 6, 10, 12?

A. 10

B. 14

C. 18

D. 12

10. What is the mean of this set of numbers: 4, 6, 7, 9. 200,000.0?

A. 7.5

B. 400,005.2

C. 7

D. 4

11. The measure of central tendency that is most affected by a few large or small numbers is the _____.

A. mean

B. median

C. mode

D. range

12. In a choice reaction time experiment, one of four lights comes on. If the first light comes on, the participant is to press the first of the four keys. If the second light comes on, the participant is to press the second of the four keys, and so on. The time between when the light comes on and when the key is pressed is recorded. What is the measurement for this study?

A. nominal

B. Ordinal

C. Interval

D. Ratio

13. In a study, people are asked to rate how much they agree with a given statement. If they strongly agree with the statement, they circle the number 1, if they agree with the statement, they circle the number 2. If they disagree with the statement, they circle the number 3. If they strongly disagree with the statement, they circle the number 4. What is the level of measurement in this study?

A. nominal

B. Ordinal

C. Interval

D. Ratio

14. The percentile rank is defined as:

A. the percentage of observations that are below a given score

B. the percentile of observations that are at or below a given score

C. the percentage of observations that are above a given score

D. the percentage of observations that are at or above a given score

15. The cutoff the researcher uses to decide whether to reject the null hypothesis is called the:

A. Significance level

B. Alpha level

C. Probability value

D. Both A and B are correct

16. The standard deviation is:

A. The square root of the variance

B. A measure of variability

C. An approximate indicator of how numbers vary from the mean

D. All of the above

17. In a grouped frequency distribution, the intervals should be what?

A. Mutually exclusive

B. Exhaustive

C. Both A and B

D. Neither A nor B

18. Non-overlapping categories or intervals are known as _____.

A. Inclusive

B. Exhaustive

C. Mutually exclusive

D. Mutually exclusive and exhaustive

19. What is the difference between a statistic and a parameter?

A. Statistics are used to describe data while parameters are used to decide if two (or more) groups are likely to be different from each other.

B. Statistics are used with experiments and parameters are used with quasi-experiments.

C. Statistics are based on samples while parameters are based on populations.

D. Statistics are based on populations while parameters are based on samples.

20. What is the difference between s^2 and σ^2?

A. s^2 is a measure of skewness while σ^2 is a measure of dispersion.

B. s^2 is the standard deviation while σ^2 is the variance.

C. s^2 is the variance of a sample while σ^2 is the variance of a population.

D. s^2 is the variance of a population while σ^2 is the variance of a sample.

21. The mean of a distribution is 23, the median is 24, and the mode is 25.5. It is most likely that this distribution is:

A. negatively skewed

B. positively skewed

C. Symmetric

D. Asymptotic

22. If the sample variance for a frequency distribution consisting of hourly wages was computed to be E10, what is the sample standard deviation?

A) E1.96

B) E4.67

C) E3.16

D) E10.00

23. The coefficient of variation for a set of annual incomes is 18%; the coefficient of variation for the length of service with the company is _____.

29%. What does this indicate?

A. More dispersion in the distribution of the incomes compared with the dispersion of their length of service

B. More dispersion in the lengths of service compared with income

C. Dispersion in the two distributions (income and service) cannot be compared using percents

D. Dispersions are equal

24. How many dependent variables are used in multiple regression?

a. One

b. One or more

c. Two or more

d. Two

25. The members of each basketball team wear numbers on the back of their jerseys. What scale of measurement are these numbers considered?

A. Nominal

B. Ordinal

C. Interval

D. Ratio

TOTAL 25 MARKS