UNIVERSITY OF SWAZILND

FACULTY OF HEALTH SCIENCES

FINAL EXAMINATION PAPER, MAY 2010

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 INVIGILATOR HAS GRANTED PERMISSION.

QUESTION 1

SITUATION: In your study among health care professionals, you obtained the following data set.

Cadre	Frequency
Nurses	70
Medical doctors	15
Surgeons	8
Laboratory Technicians	27
Radiologists	10
Total	**130**

A. Present the data in a graph format. (5)

B. Describe the characteristics of a normally distribution. (3)

C. Describe any six (6) features that contribute to the quality of table that presents data

(6)

SITUATION: At your health facility there were 5.000 thousand deliveries in 2008. Only 455 of the women had caesarean sections.

D. What is the probability of having a caesarean section (per 100 women) at your health facility? (5)

E. **Calculate:** Out of ten women, how many are likely to give normal deliveries (normal vaginal deliveries) in your health facility? (6)

TOTAL 25 MARKS

QUESTION 2

Below is a linear regression model between the independent variable, household income, and the dependent variable, self reported health

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.301(a)	.091	.086	.641

a Predictors: (Constant), Household Income

ANOVA(b)

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	8.364	1	8.364	20.334	.000(a)
	Residual	83.913	204	.411		
	Total	92.277	205			

a Predictors: (Constant), Household Income
b Dependent Variable: Self reported Health Status

A. Compute the Coefficient of Determination, and describe its meaning? (2)

B. Present and Interpret the ANOVA output. (4)

C. Describe four (4) assumptions of Linear Regression and state how these assumptions could be violated by a researcher. (8)

Situation: The following are CD4 cells (mg/dL) of six (6) adult clients who are receiving antiretroviral therapy: 23, 72, 50, 98, 110, and 103.

Compute the following:

D. The standard deviation. (4)

E. The skewness of the data set, and interpret its meaning. (4)

F. The Coefficient of Variation. (3)

TOTAL 25 MARKS

QUESTION 3

CLEARLY write the letter that corresponds with the most appropriate answer **e.g. 1. B**

1. A positive correlation is present when _____.

    A. several variables never change

    B. Two variables move in the same direction.

    C. Two variables go in the opposite direction.

    D. One variable goes up and one goes down.

2. Which correlation is the strongest?

    A. +.10

    B. -.95

    C. +.90

    D. -1.00

3. The correlation between intelligence quotient (IQ) test scores and grades is:

    A. positive

    B. negative

    C. perfect

    D. they are not correlated

4. To determine whether noise affects the ability to solve statistics problems, a researcher has one group solve statistics problems in a quiet class-room and another group solve statistics problems in a noisy class-room. The group solving problems in the noisy class-room completes 15 problems in one-hour and the group solving problems in the quiet class-room completes 22 problems in one-hour. In this experiment, the independent variable is _____ and the dependent variable is _____.

    A. the number of problems solved; the difficulty of the problems

    B. the number of problems solved; the noise level in the room

    C. the noise level in the room; the number of problems solved

    D. the noise level in the room; the difficulty of the problems

5. Which of the following statements sounds like a null hypothesis?

    A. The coin is not fair

    B. There is a correlation in the population

    C. There is no difference between male and female incomes in the population

    D. The defendant is guilty

6. Which of the following is the researcher usually interested in supporting when he or she is engaging in hypothesis testing.

    A. the alternative hypothesis

    B. the null hypothesis

    C. both the alternative and null hypothesis

    D. neither the alternative nor the null hypothesis

7. A study was conducted to measure the effect of smoking upon the birth weight of a baby. The following variables were measured for each woman in the study:

    (i) smoking status (yes, no),

    (ii) baby birth weight (kg),

    (iii) APGAR score,

    (iv) birth date

The scales (following sequence of above variables) of these variables are:

    A. nominal, ratio, ratio, ratio.

    B. ordinal, ratio, interval, interval

    C. nominal, ratio, ordinal, ratio

    D. nominal, ratio, ordinal, interval

7. Which of the following divides a group of data into four subgroups?

    A. quartiles

    B. percentiles

    C. standard deviation

    D. median

8. If the standard deviation of a population is 9, the population variance is:

    A. 3

    B. 9

    C. 21.35

    D. 81

9. Which of the following is not a measure of central tendency?

    A. percentiles

    B. quartiles

    C. standard deviation

    D. mode

10. Which of the following statements is / are **TRUE** according to the logic of hypothesis testing?

A. when the null hypothesis is true, it should be rejected

B. when the null hypothesis is true, it should not be rejected

C. when the null hypothesis is false, it should be rejected

D. Both B and C are true

E. Both A and C are true

11. The cut-off the researcher use to decide whether to reject the null hypothesis is called the:

A. significance level

B. alpha level

C. probability level

D. Both A and B

12. The use of laws of probability to make inferences and draw statistical conclusions about populations based on sample data is referred to as _____ .

 A. descriptive statistics

B. inferential statistics

C. sample statistics

D. Population statistics

13. Which of the following in **NOT** correct about constructing histograms?

A. all class intervals should be of equal width

B. the bars of the histogram are centred over the class-mark |(midpoint)

C. the first and last classes should be open-ended to account for extreme scores

D. there should be no spaces between bars

14. Which of the following statements in **NOT CORRECT**?

A. the frequency table is useful for summarizing data from a nominal scaled variable

B. means and standard deviations of nominal or ordinal scaled variables are useful summary measures

C. pie charts don't perform well because people have difficulty in accurately quantifying angles

D. the scatterplot is the basic graphical tool for investigating relationships between two interval or ratio scaled variable

SITUATION: the following are birth weights (kg) of 26 male babies born to a group of mothers who smoked during pregnancy.

3, 4, 6, 7, 7, 8, 8, 8, 9, 2, 2, 3, 4, 6, 7, 8, 9, 1, 2, 2, 3, 4, 3, 5, 5, and 6

15. The median birth weight is: **(3 marks)**

A. 13.5 kg

B. 3.2 kg

C. 3.5 kg

C. 5.0 kg

D. average of 13 and 14 kg

16. Refer to the previous situation (grouping the data into class intervals that have widths of 2). The first quartile (25th percentile) of the weight is _____: **(3 marks)**

A. 2.93 kg

B. 2.73 kg

C. 0.25 kg

D. 6.58 kg

E. 2.88 kg

17. In a study that was conducted with antenatal women, the following variables were measured on each woman:

sex, initial weight (kg), temperature (^{0}C), weight gain (kg)

The scales (following sequence of above variables) of these variables are

_____:

    A. nominal, ratio, ratio, ratio

    B. nominal, ratio, interval, interval

    C. ordinal, ratio, interval, ratio

    D. nominal, ratio, interval, ratio

    E. ordinal, interval, ratio, interval

18. If the correlation between body weight and annual income were high and positive, we could conclude that _____:

A. high incomes cause people to eat more

B. low incomes make people to eat less

C. high income people tend to spend a greater proportion of their income on food than low income people, on average

D. high income people tend to be heavier than low income people, on average

E. high incomes cause people to gain weight

19. The sum of the deviations about the mean are always _____:

A. the range

B. the standard deviation total

C. zero

D. positive

20. A random sample of 40 smoking people is classified in the following table:

Ages	Cigarettes smoked
11 - 20	4
21 - 30	6
31 - 40	12
41 - 50	10
51 - 60	8

The mean number of cigarettes smoked in this group is _____:

(2marks)

    A. 4.5 cigarettes

    B. 38.45 cigarettes

    C. 34.45 cigarettes

    D. 38.0 cigarettes

    E. 1537.4 cigarettes

TOTAL = 25 MARKS