MASINDE MULIRO UNIVERSITY OF SCIENCE AND TECHNOLOGY (MMUST)

UNIVERSITY EXAMINATIONS (MAIN PAPER) 2023/2024 ACADEMIC YEAR

THIRD YEAR FIRST SEMESTER EXAMINATIONS

FOR THE DIPLOMA
IN
MEDICAL BIOTECHNOLOGY

COURSE CODE: BBD 311

COURSE TITLE: FUNDAMENTALS OF MEDICAL BIOSTATISTICS

DATE: 6ᵀᴴ DECEMBER 2023   **TIME: 2.00-4.00PM**

INSTRUCTIONS TO CANDIDATES

This paper is divided into three sections, A B and C, carrying respectively: Multiple Choice Questions (**MCQs**), Short Answer Questions (**SAQs**) and Long Answer Questions (**LAQs**). Answer all questions. **DO NOT WRITE ON THE QUESTION PAPER**

TIME: 2 Hours

MMUST observes ZERO tolerance to examination cheating

This Paper Consists of 5 Printed Pages. Please Turn Over

SECTION A: Multiple Choice Questions (20Marks)

1. Which one of the following variables is **not** categorical?

A. Age of a person.

B. Gender of a person: male or female.

C. Choice on a test item: true or false.

D. Marital status of a person (single, married, divorced, other)

2. Which one of these statistics is **unaffected** by outliers?

A. Mean

B. Interquartile range

C. Standard deviation

D. Range

3. A list of 5 pulse rates is: **70, 64, 80, 74, 92**. What is the **median** for this list?

A. 74

B. 76

C. 77

D. 80

4. Which one of the following would indicate that a dataset is not bell-shaped?

A. The range is equal to 5 standard deviations.

B. The range is larger than the interquartile range.

C. The mean is much smaller than the median.

D. There are no outliers.

5. The value of a correlation is reported by a researcher to be r = −0.5. Which one of the following Statements is **correct**?

A. The x-variable explains 25% of the variability in the y-variable.

B. The x-variable explains −25% of the variability in the y-variable.

C. The x-variable explains 50% of the variability in the y-variable.

D. The x-variable explains −50% of the variability in the y-variable.

6. What is the effect of an outlier on the value of a correlation coefficient?

A. An outlier will always decrease a correlation coefficient.

B. An outlier will always increase a correlation coefficient.

C. An outlier might either decrease or increase a correlation coefficient, depending on

D. Where it is in relation to the other points.

7. A chi-square test involves a set of counts called "expected counts." What are the expected counts?

A. Hypothetical counts that would occur of the alternative hypothesis were true.

B. Hypothetical counts that would occur if the null hypothesis were true.

C. The actual counts that did occur in the observed data.

D. The long-run counts that would be expected if the observed counts are representative

8. Pick the choice that best completes the following sentence. If a relationship between two Variables is called statistically significant, it means the investigators think the variables are _____

A. Not related in the population represented by the sample.

B. Related in the sample due to chance alone.

C. Very important.

9. In which case is the following equation valid: $p(A + B) = p(A) + p(B)$

A. If "A" and "B" are independent events.

B. If "A" and "B" are mutually exclusive events.

C. For any "A" and "B" events.

D. If "A" and "B" are dependent events.

10. which one of the following is not a requirement for a good estimation?

A. A good estimation should be consistent.

B. A good estimation should be efficient.

C. A good estimation should be unbiased.

D. A good estimation should be unpredictable.

11. The definition of frequency in statistics is?

A. The frequency of an event is equal to the count of its occurrence in the population.

B. The frequency of an event is equal to the count of observations.

C. The frequency of an event is equal to the count of its occurrence per unit time.

D. The frequency of an event is equal to the count of its occurrence in a series of observations

12. Choose the correct statement(s).

A. In medical practice, confidence level usually corresponds to 95% probability.

B. The confidence level can be chosen freely.

C. The confidence interval contains 95% of the data.

D. The confidence interval contains 68% of the data.

13. In a one-sample t-test the calculated t value is 1.897 and the t value that belongs to the significance level is 2.013. What should be your decision?

A. I accept the null hypothesis.

B. I repeat my calculation, because this situation cannot happen in one-sample t-test.

C. I cannot say anything without knowing the probabilities.

D. I reject the null hypothesis.

14. In which case(s) is it appropriate to use Wilcoxon sign test?

A. To test the change of a non-parametric variable in two paired samples.

B. To compare the parametric variable of two samples with different numbers of elements.

C. To test normally distributed numerical variables in one sample.

D. To compare the non-parametric variable of two samples with different numbers of elements.

15. What test can we use if we have 1 numerical, continuous variable in 3 (not paired) groups and the groups are normally distributed?

A. ANOVA

B. Mann-Whitney U

C. Kruskal-Wallis

D. Kolmogorov-Smirnov

16. Null hypothesis is rejected if _____

A. The sample statistical parameter is less than the critical statistical parameter.

B. The significance level (considering a two tailed test) is more than 5%.

C. The significance level is less than 5%.

D. The sample statistical parameter is greater than the critical statistical parameter

17. A test for homogeneity is to be conducted. Which method shall be used?

   A. Student's t-test for two samples.

   B. Chi-square test.

   C. Student's t-test for one sample.

   D. Mann-Whitney U-test.

18. When the data are skewed to the right, measure of skewness will be _____

   A. Negative.

   B. Zero.

   C. Positive.

   D. One.

19. Which one of the following is not a measure of dispersion?

   A. The range.

   B. The 50ᵗʰ percentile.

   C. The standard deviation.

   D. The interquartile range.

20. The most frequently occurring value of data set is called _____

   A. Range

   B. Median

   C. Mode

   D. Mean

SECTION B: Short Answer Questions (40 Marks)

1. The following set of data: **9, 8, 6, 5, 8, 6**. Determine the variance and standard deviation (8 Marks).

2. Describe **any four** non parametric tests used in biostatistics   (8 Marks).

3. Identify the type of data below as **nominal, ordinal, discrete or continuous**   (8 Marks).

   i.   Blood group

   ii.   Temperature (Celsius)

   iii.   Ethnic group

   iv.   Job satisfaction index (1-5)

   v.   Number of heart attacks

   vi.   Calendar year

   vii.   Serum uric acid (mg/100ml)

   viii.   Number of accidents in 3 - year period

4. Given the following data set (age of patients): **18, 59,24,42,21,23,24,32.**

   i.   Determine the **range** of the data set above   (2 Marks).

   ii.   Find the **interquartile range** of the data set above   (6 Marks).

5. Highlight the properties of **mode, median and mean** respectively   (8 Marks).

SECTION C: LONG ANSWER QUESTIONS (60 Marks)

TABLE 1: The table represents total circulating albumin in gm for 30 normal males, age 20-29 years.

CLASS LIMITS (CL)	FREQUENCY (F)
100 – 109	2
110 – 119	6
120 – 129	6
130 – 139	7
140 – 149	8
140 – 159	1
TOTAL	30

1. Using grouped data from **TABLE 1** above, calculate: **the mean, median and standard deviation** (20 Marks).

2. The following table shows the relation between the number of accidents in 1 year and the age of the driver in a random sample of 500 drivers between 18 and 50. Test, at a 01 level of significance, the hypothesis that the number of accidents is independent of the driver's age. **According to Chi – square, what is the predicted outcome?** (20 Marks).

Number of Accidents	18 - 25	26 – 40	Above 40	Total
Zero	75	115	110	300
One	50	65	35	150
Above Two	25	20	5	30
Total	150	200	150	500

3. Three different brands of magnetron tubes were subjected to stress testing. The number of hours each operated before needing repair was recorded.

Brand A	36	48	6	67	53
Brand B	49	33	60	2	55
Brand C	71	31	140	59	224

Use one way Analysis of Variance (**ANOVA**) procedure to test the hypothesis that the mean lifetime under stress is the same for the three brands (20 Marks).