



MAASAI MARA UNIVERSITY

REGULAR UNIVERSITY EXAMINATIONS 2022/2023 ACADEMIC YEAR THIRD YEAR FIRST TRIMESTER

SCHOOL OF PURE, APPLIED AND HEALTH SCIENCES BACHELOR OF SCIENCE IN NURSING

COURSE CODE: NUR 3103 COURSE TITLE: BIostatISTICS

DATE: 17/4/2023

TIME: 1430-1730 HRS

INSTRUCTION TO CANDIDATES

Section A: Multiple Choice Questions. Answer ALL Questions

Section B: Short Answer Questions. Answer ALL Questions

Section C: Long Answer Questions. Answer ANY TWO questions.

This paper consists of 8 printed pages. Please turn over.

SECTION A- MULTIPLE CHOICE QUESTIONS (ANSWER ALL) (20 MARKS)

1. Pearson correlation coefficient, denoted by r , measures
 - A. The scattering strength of data for a statistical series
 - B. The strength of the correlation between the mean and median
 - C. The strength of the correlation between two numerical parameters
 - D. The amount of variation in the dependent variable explained by independent variable.
2. In which scale of measurement does pain rating of a patient classified?
 - A. Nominal.
 - B. Ordinal.
 - C. Interval.
 - D. Ratio.
3. Skewness measure _____
 - A. Symmetry of a dataset.
 - B. Lack of symmetry.
 - C. Dispersion of a data set.
 - D. Outliers within the dataset.
4. When a dataset is skewed which is the best measure of central tendency
 - A. Arithmetic mean
 - B. Median
 - C. Mode
 - D. Quantiles
5. The probability that the patient coming to our office has viral infection is 0.51. The probability of the occurrence of flu infection at our office is 0.1. What is the probability that the patient has flu, IF we know that this patient has viral infection?
 - A. 19.61%.
 - B. 5.1%.
 - C. 0.459
 - D. 0.049.
6. The probability of occurrence of vascular stenosis among smokers is 0.26. The probability of being a smoker in a given population is 0.34. What is the probability in this population that someone is a smoker with vascular stenosis?
 - A. 0.765
 - B. 1
 - C. Data is too few to answer.
 - D. 0.088

7. The data that categories patients as males or females are known as:
- A. Random data
 - B. Nominal data
 - C. Ordinal data
 - D. Interval data
8. Type I error refers to _____
- A. Error resulting from wrong calculation
 - B. Error resulting from bias data
 - C. Error resulting from accepting null hypothesis wrongly
 - D. Error resulting from rejecting null hypothesis wrongly
9. The diagnostic results of a patient for malaria can be classified.
- A. Mutually Exclusive
 - B. Mutually Inclusive
 - C. Dependent
 - D. Independent
10. A sample of 5 body weights (in pounds) is as follows: 116, 168, 124, 132, 110. The *sample median* is:
- A. 124.
 - B. 116.
 - C. 132.
 - D. 130.
 - E. None of the above.
11. The following are true about Standard deviation except:
- A. It is the square root of variance
 - B. It is measured using the unit of the variable
 - C. It is measured using the squared unit of the variable
 - D. It has values generally comparable with the average value
12. The following are true about a regression line except:
- A. located as close as possible to all the points of a scatter chart
 - B. Defined by an equation having 2 parameters: the slope and the intercept
 - C. Provides an approximate relationship between the values of two parameters
 - D. Parallel to one of the coordinate axes
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. Which one of the following is an example of statistical inference?
- A. Calculating the sample t-value.
 - B. Rejecting the null hypothesis as a result of hypothesis testing.
 - C. Calculating the second central moment of the sample.
 - D. Calculating the second moment of the sample.
15. We are studying the applicability of a diagnostic test. What is the name of the parameter given by the ratio of true positive tests and all positive tests?
- A. Positive predictive value
 - B. Sensitivity
 - C. Specificity
 - D. Negative predictive value
16. When a hypothesis for statistical test is rejected at 99% confidence level what will happen at 95%. Give a reason.
- A. Rejected, because the calculated test statistic will still fall in the critical region.
 - B. Fail to reject, because the calculated test statistic will still fall in the critical region.
 - C. Don't know, because we cannot know the exact location of calculated test statistic.
 - D. Rejected, because the calculated test statistic will still fall in the acceptance region.
 - E. Fail to reject, because the calculated test statistic will still fall in the acceptance region.
17. Suppose a random sample of 100 12-year-old boys were chosen and the heights of these 100 boys recorded. The sample mean height is 64 inches, and the sample standard deviation is 5 inches. You may assume heights of 12-year-old boys are normally distributed. Which interval below includes approximately 95% of the heights of 12-year-old boys?
- A. 63 to 65 inches.
 - B. 39 to 89 inches.
 - C. 54 to 74 inches.
 - D. 59 to 69 inches.

- E. Cannot be determined from the information given.
- F. Can be determined from the information given, but none of the above choices is correct.
18. Cholesterol levels are measured on a random sample of 1,000 persons, and the sample standard deviation is calculated. Suppose a second survey were repeated in the same population, but the sample size tripled to 3,000. Then which of the following is true?
- A. The new sample standard deviation would tend to be smaller than the first and approximately about one-third the size.
 - B. The new sample standard deviation would tend to be larger than the first and approximately about three times the size.
 - C. The new sample standard deviation would tend to be larger than the first, but we cannot approximate by how much.
 - D. None of the above is true because there is no reason to believe one standard deviation would tend to be larger than the other.
19. A clinical trial is more valuable when:
- A. Sensitivity and Specificity have higher values
 - B. Sensitivity is higher than specificity
 - C. Specificity is higher than Sensitivity.
 - D. The sensitivity and specificity values are close, even equal, regardless of their values
20. The mode of a series of numerical values is:
- A. A value for which half of the values are higher and half of the values are lower
 - B. The value located exactly midway between the minimum and maximum of the series
 - C. The most commonly encountered values among the series.
 - D. A measure of the eccentricity of the series.

SECTION B: SHORT ANSWER QUESTIONS (ANSWER ALL) (40 MARKS)

1. Differentiate between the following terms;
- i. Critical value and Critical region. **(2 marks)**
 - ii. Variable and Constant. **(2 marks)**
 - iii. Sensitivity and Specificity. **(2 marks)**
 - iv. Significance level and Confidence level. **(2 marks)**

2. The birth weight for a number of newborns in a certain hospital was recorded as 4kg, 5kg, 5.6kg, 4.5kg, 7kg, 4.2kg, 5.8kg, 6.1kg, 7.5kg, 4.6kg
- Determine the sample mean. **(2 marks)**
 - Determine the sample standard deviation. **(3 marks)**
 - Determine the coefficient of variation. **(2 marks)**
 - Construct a 95% confidence interval for the true average birthweight in the hospital and hence comment on the assumption that the average birthweight was 5kg. **(4 marks)**
3. The I.Q for patients admitted in a hospital is known to be normally distributed with a mean of 105 and a standard deviation of 15. Determine the probability that a randomly selected patient has an I.Q of
- At least 135. **(3 marks)**
 - At most 98. **(3 marks)**
 - Between 125 and 142. **(3 marks)**
4. A case-control study was performed to investigate the association between use of exogenous estrogen during menopause and endometrial cancer. Enrolled were 23 women with newly diagnosed endometrial carcinoma and 17 women free of this disease. Study participants were interviewed regarding their history of use of exogenous estrogen. The data are the following

	Endometrial Cancer		Total	
	Yes	No		
Estrogen	User	8	2	10
	Non - User	15	15	30
	Total	23	17	40

- Determine the probability that a patient has Endometrial Cancer. **(1 mark)**
- Determine the probability that a patient has Endometrial Cancer or uses exogenous estrogen. **(2 marks)**
- Determine the probability that a patient has Endometrial Cancer given that they use exogenous estrogen. **(3 marks)**
- At 95% level of confidence, test whether there is a relationship between use of exogenous estrogen and endometrial cancer infection. **(6 marks)**

SECTION C: LONG ESSAY QUESTIONS (ANSWER ANY TWO) (40 MARKS)

Question 1

The data below shows the weights of patients who visited a certain hospital in a particular week.

Weights	50-55	55-60	60-65	65-70	70-75	75-80
No of Male Patients	8	14	26	22	15	5
No of Female Patients	17	34	17	14	8	2

- i. Determine the average weight for both male and Female patients. **(6 marks)**
- ii. Determine the modal weight for male patients. **(3 marks)**
- iii. Determine the median weight for female patients. **(3 marks)**
- iv. Determine the standard deviation in weight for male patients. **(3 marks)**
- v. Determine the standard deviation in weight for female patients. **(3 marks)**
- vi. Which patients had the highest variation in weight that week? **(2 marks)**

Question 2

- a. In order to reduce the cost of health care in the country NHIF intends to put pressure on hospitals to shorten the average length of stay of their patients. In 2019 the average length of stay in Narok county hospital for men was 6.5 days and the average for women was 5.6 days. A random sample of 20 hospitals selected in 2022 in Narok county had a mean length of stay of 3.6 days for women with a standard deviation of 1.2 days.
 - i. Construct a 95% confidence interval for the average length of stay for women in Narok County hospitals in 2023. **(3 marks)**
 - ii. Interpret the results in (i). **(2 marks)**
 - iii. What will happen to the confidence interval in (i) if we increase the sample size to 80? Justify your reasoning. **(2 marks)**
 - iv. What will happen to the confidence interval in (i) if we increase the confidence level to 99%? Justify your reasoning. **(2 marks)**
 - v. Based on the results in (i) is there sufficient evidence at 95% level of confidence to show that Narok has significantly reduced the length of stay in hospital. Justify your reasoning. **(2 marks)**
- b. Suppose a 95% confidence interval for proportion of patients admitted at a hospital with drinking disorder is (0.61, 0.67).
 - i. Interpret the confidence interval. **(2 marks)**

- ii. Determine the sample proportion of patients with drinking disorder. **(2 marks)**
- iii. Determine the sample size that was used for the study to construct the confidence interval. **(3 marks)**
- iv. It is believed that 55% of patients admitted at the hospital have drinking disorder, based on the confidence interval given, comment on this assumption. **(2 marks)**

Question 3

- a. The data below shows the BMI for a number of male and female patients admitted at a certain hospital. The distributions of BMI for both groups was assumed not to be normally distributed.

Male	Female
23.8	19.6
23.2	23.8
24.6	19.6
26.2	29.1
23.5	25.2
24.5	21.4
21.5	22.0
31.4	27.5
26.4	33.5
22.7	20.6
27.8	29.9
28.1	17.7
25.2	

At 95% level of confidence test whether there was a significant difference in BMI for male and female patients at the hospital. **(15 marks)**

- b. Suppose that, among 25-34-year-old males in the general population, the average daily intake of linoleic acid is 15 g. As part of a dietary-instruction program, ten 25-34-year-old males adopted a vegetarian diet for one month. While on the diet, the average daily intake of linoleic acid in this sample was a sample mean =13 g with a sample standard deviation = 4 g. At 95% level of confidence test whether vegetarian diet reduce the level of linoleic-acid. **(5 marks)**

//END//