



MAASAI MARA UNIVERSITY

REGULAR UNIVERSITY EXAMINATIONS

2021/2022 ACADEMIC YEAR

SECOND YEAR FIRST SEMESTER

**SCHOOL OF PURE APPLIED AND HEALTH
SCIENCES**

DIPLOMA IN NUTRITION

COURSE CODE: DND 2107

**COURSE TITLE: INTRODUCTION TO
BIostatISTICS**

DATE: 8TH APRIL, 2022

TIME: 0830-1030

INSTRUCTIONS TO CANDIDATES

- i. This paper consists of three sections*
- ii. Answer **ALL** Question in Section **One and two***
- iii. Answer two questions in section **Three***

1. Which of the following is not a characteristic of an estimator
 - a) Unbiased
 - b) Consistency
 - c) Efficiency
 - d) Normal

2. In determination of an estimator what is the minimum sample size for a t test .
 - a) 30
 - b) 45
 - c) 100
 - d) 10

3. A manager wants an estimate of sales of salesmen in his company. A random sample 100 out of 500 salesmen is selected and average sales are found to be Shs. 75,000. if a sample standard deviation is Shs. 15000 then find out the population mean at 99% level of confidence.
 - a) 13420 to 13500
 - b) 75000 to 78464
 - c) 34640 to 35400
 - d) 71536 to 78464

4. Given two samples A and B of 100 and 400 items respectively, they have the means $\bar{X}_1 = 7$ and $\bar{X}_2 = 10$ and standard deviations of 2 and 3 respectively. Determine the standard error of the samples.
 - a) 0.25
 - b) 0.7
 - c) 1.04
 - d) 0.5

5. In a sample of 800 candidates, 560 were male. Estimate the population proportion at 95% confidence level.
 - a) 0.67 to 0.73
 - b) 0.60 to 0.70
 - c) 0.54 to 0.64
 - d) 0.73 to 0.78

6. Chi square test is used to test independence of attributes
 - a) True
 - b) False

7. What is the value of standard error of one tailed test at a 5% level of significance
 - a) 1.96
 - b) 1.65

- c) 2.33
 - d) 2.50
8. Which of the following is not a probability sampling method
- a) Simple random sampling
 - b) Multistage sampling
 - c) Purposive sampling
 - d) Cluster sampling
9. The values which separate the rejection region from the acceptance region are called critical values
- a) True
 - b) False
10. Acceptance of a false hypothesis is called type II error
- c) True
 - d) False
11. A positive Skewness is where the mean is less than the mode
- a) True
 - b) False
12. In a symmetrical bell shaped distribution, approximately 95% of the distribution will lie between ± 2
- a) True
 - b) False
13. The following regression line $y = 5 + 0.785x$, shows the relationship between two variables, estimate the value of y when x is 100.
- a) 83.5
 - b) 78.5
 - c) 500
 - d) 5.785
14. A p-value: A measure of how much evidence you have against the null hypothesis.
- a) True
 - b) False

15. Parametric tests assumes that the data under consideration fulfill normality condition and so standard statistical tests can be used.

a) True

b) False

16. *Test statistic*: A value, determined from sample information, used to determine whether or not to reject the null hypothesis

a) True

b) False

17. A statistical hypothesis is an assertion about a parameter of a population.

a) True

b) False

18. One way Analysis of Variance (ANOVA) is a test used for Comparing more than 2 two population means with known population variance .

a) True

b) False

19. .The Coefficient of Correlation is a measure of the strength of the relationship between two variables.

a) True

b) False

20. A Dependent Variable is a variable that is being predicted or estimated.

a) True

b)False

Section 2: Answer all questions (40 marks)

1.The table shows whether or not the subjects suffered from heart disease and how their snoring habits were classified by their partners.

	Never	Occasionally Snores	Snores every night
Heart disease	50	90	120
No heart disease	70	90	80

Use a χ^2 test, at the 5% significance level, to investigate whether frequency of snoring is related to heart disease. **(10 marks)**

2. In an opinion poll conducted by means of a simple random sample of 2000 voters ; 45% of those questioned stated that they would vote for party B if there were an immediate general election. Calculate a 99% confidence interval for the percentage voting for the party in the elections . **(10 marks)**

3. Distinguish between the following concepts as used in hypothesis testing

a) Type 1 and type II error **(4 marks)**

b) Nominal and ordinal data measurements **(4 marks)**

c) null and alternative hypothesis **(2marks)**

4. The following is data obtained by a student in a biostatistics experiment

24	13	28	15	25	29	15	46
9	10	17	22	23	17	16	32
11	12	18	20	13	27	18	22
20	14	26	14	19	19	40	31
17	21	23	26	18	24	21	27

a) From the data make a frequency distribution table starting with the following class (5-10) **(5 marks)**

b) Calculate the mean, standard deviation and coefficient of variance

(5 arks)

section C : Answer any three questions 40 marks

1. a) The following data gives the age in months of a child and the corresponding weight in Kg

Age	15	17	19	21	23	25	27	29	31	33	35
Weight	13	21	21	24	23	25	22	24	32	31	35

- i. Plot a scatter diagram for the data and make comments **(4 marks)**
 ii. Calculate the regression line ($y = a + bx$) and use it to estimate the value of Y if X is 120 **(6 marks)**

- b) In a Nutrition competition 2 assessors were asked to rank the 10 contestants using the professional assessment skills on their response to new food supplement. The results obtained were given as shown in the table below

Contestants	A	B	C	D	E	F	G	H
1 st assessor	6	1	3	7	8	2	4	5
2 nd assessor	5	3	4	6	7	1	8	10

Hint. $r = 1 - \frac{6 \sum d^2}{n(n^2 - 1)}$

REQUIRED

Calculate the rank correlation coefficient and hence comment briefly on the value obtained **(10 Marks)**

2. a) A sample of 8 students were given a diagnostic test before studying a particular module and then again after completing the module. The following data gives their scores before and after the training.

Score 1: Before 19 21 17 21 23 18 14

Score 2: After 26 29 23 23 19 29 19

Test at 0.1 and 0.05 levels of significant if the teaching leads to improvements in students. **(10 Marks)**

- b) Brandways company indicate on the label that their loaves weigh 400g. A sample of 40 loaves is selected hourly from their processing line and the contents weighed. Last hour a sample of 40 loaves had a mean weight of 403g

with a standard deviation of 8g. Test at .05 significance level whether their process is out of control? **(10 marks)**

3. a) The following table gives three treatments made to some groups
Treatment 1 (x1): 15 20 19 14
Treatment 2 (x2): 10 15 11
Treatment 3 (x3): 18 19 23
Test at $\alpha = 0$ whether the Treatments have different effects or are the same .
(10 marks)

b)The daily water usage per person in Thika is normally distributed with a mean of 20 gallons and a standard deviation of 5 gallons. What is the probability that a person from Thika selected at random will use;
(a) less than 20 gallons per day?
(b) less than 25 gallons per day?
(c) more than 30 gallons per day?
(d) Uses between 25 and 30 gallons? **(10 marks)**

/////END/////