KISII UNIVERSITY

UNIVERSITY EXAMINATIONS

FIRST YEAR EXAMINATION FOR THE AWARD OF THE DEGREE OF BACHELOR OF SCIENCE IN MATHEMATICS AND APPLIED STATISTICS
SECOND SEMESTER 2022/2023
[JANUARY-APRIL, 2023]

MATH 141: INTRODUCTION TO STATISTICS

STREAM: Y1S2 **TIME: 2 HOURS**

DAY: TUESDAY, 9:00 – 11:00 AM **DATE: 28/03/2023**

INSTRUCTIONS

1. **Do not write anything on this question paper.**
2. ***Answer question ONE and any other TWO questions.***

QUESTION ONE (30 MARKS)

a. Differentiate between the following terms as used in statistics;

i. Primary and Secondary data **(2 marks)**
ii. Discrete and continuous data (**2 marks**)

b. Outline the qualities of a good average (**5 marks**)

c. The annual wages for A sample of 10 employees of an agricultural firm were recorded. The following are the annual salaries (in thousands of Dollars).

41, 49, 50, 82,55, 87,52, 52, 47,44

Using the data,

i. verify that **Arithmetic Mean** $(\bar{X}) \geq$ **Geometric Mean**(**GM**) $\geq$ **Harmonic Mean**(**HM**) (**7 marks**)
ii. find the difference between the median and mode for the data (**3 marks**)

d. (i). Differentiate between a trial and an event as used in probability. (**2 marks**)

(ii). A farmer in Kisii selects 2 cows at random from 4 large, 5 medium and 6 small cows. What is the probability that they are of the same size. (**3 marks**)

e. The coefficient of rank correlation of the marks scored by 10 students in English and Kiswahili was found to be 0.2. it was later discovered that the difference in ranks in the two subjects obtained by one of the students was wrongly taken as 9 instead of 7.

i. Find the correct coefficient of rank correlation (**5 marks**)
ii. Interpret the new coefficient of rank correlation (**1 mark**)

QUESTION TWO (20 MARKS)

a. With the aid of diagrams, differentiate between
   i. positive and negative skewness (4 marks)
   ii. leptokurtic and mesokurtic forms of Kurtosis (4 marks)

b. The monthly profits (in Kshs'000) of 33 shops are distributed as follow *(2 marks)*

82, 96, 99, 102, 103, 103, 106, 107, 108, 108, 108, 108, 109, 110, 110, 111, 113, 113, 113, 113, 115, 115, 118, 118, 119, 121, 122, 122, 127, 132, 136, 140, 146

The data is to be organized into a grouped frequency distribution table.

i. Determine the number of classes hence the class interval (**4 marks**)
ii. From the grouped frequency table, calculate the mode monthly profit. **(4 marks)**

c. Consider the following information on ultimate tensile strength (lb/in) for a sample of 100 hard zirconium copper wire specimens:
$\bar{X} = 76.87$, Standrad deviation $= 18$, mode $= 73.88$.
   i. Find the median tensile strength (2 marks)
   ii. it is discovered that the tensile strength sample values were exaggerated hence two methods are recommended for their re-adjusted
      I. multiply each value 1.25
      II. subtract 7.77 from each value.

Determine how the mean and standard deviation will be affected in each case. *(4 marks)*

QUESTION THREE (30 MARKS)

a. Differentiate between the following statistical terms
   i. Correlation analysis and regression analysis. (**2 marks**)
   ii. Positive correlation and negative correlation (**2 marks**)

b. The marks scored by 5 students in Accounting and Statistics units are recorded in the table below:

Student Number	1	2	3	4	5
Marks in Accounting	45	35	20	23	47
Marks in Statistics	50	25	20	25	45

i. Calculate the Karl Pearson's coefficient of correlation (**7 marks**)
ii. Regress marks scored in accounting against marks scored in statistics. (**7 marks**)

iii. Estimate the score in statistics for a student who scored 55 in accounting

(**2 marks**)

QUESTION FOUR (20 MARKS)

a) (i). Define a probability sample space (2 marks)
(ii). Two unbiased dice are drawn and the sum of the numbers showing up recorded. Draw a table and illustrate the sample space (3 marks)
(iii). Using the results in (ii) above, determine the probability of getting a sum of at least 11 or at most 2. (3 marks)

b) Differentiate between mutually exclusive events, independent events and equally likely outcomes as used in probability (6 marks)

c) In a group of 1,000 persons, there are 650 who can speak Hindi, 400 can speak English and 150 can speak both Hindi and English. If a person is selected at random, what is the probability that he speaks

i. Hindi only, (**2 marks**)
ii. only one of the two languages, (**2 marks**)
iii. at least one of the two llanguages (**2 marks**)

QUESTION FIVE (30 MARKS)

The data below represents the weights of 100 students who attended the Valentine excursion at Uhuru Park

Ages	10-20	20-30	30-40	40-50	50-60	60-70
Number of people	5	15	X	30	15	10

Find

a. the value of x (1 mark)
b. the mean number of people who attended the excursion (4marks)
c. the median number of people who were present in that excursion (4marks)
d. the standard deviation for data above (8marks)
e. the upper quartile for the data above (3marks)