



MERU UNIVERSITY OF SCIENCE AND TECHNOLOGY

P.O. Box 972-60200 – Meru-Kenya.

Tel: +254(0) 799 529 958, +254(0) 799 529 959, +254 (0)712 524 293

Website: www.must.ac.ke Email: info@mucst.ac.ke

UNIVERSITY EXAMINATIONS 2023/2024

SECOND YEAR SECOND SEMESTER EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE IN COMPUTER SCIENCE AND FIRST YEAR SECOND SEMESTER DATA SCIENCE

SSA 3120: PROBABILITY AND STATISTICS

DATE: APRIL 2023

TIME: 2 HOURS

INSTRUCTIONS: Answer Question ONE and any other TWO questions.

QUESTION ONE (30 MARKS)

a) Distinguish between

- i. Primary data and Secondary data (2marks)
- ii. Qualitative data and Quantitative data (2marks)
- iii. Discrete variables and Continuous variables (2marks)

b) Determine whether the function given by $f(x) = \frac{x+2}{25}$ for $x = 1,2,3,4,5$ can serve as a probability distribution of a discrete random variable. (5marks)

c) Roll a fair die twice and define A to be the event that the sum of the scores showing up is greater than 7, B be the event that the sum of the scores showing up is a multiple of 3 and C be the event that the sum of the scores showing up is a prime number. Which of the events A, B and C are independent? Are the three events jointly independent? (5marks)

d) For the following data given below

x	F
10	12
20	23

30	35
40	47
50	38
60	29
70	16

Compute;

- i. Arithmetic mean (3marks)
- ii. Lower quartile (2marks)
- iii. Middle quartile (2marks)
- iv. Upper quartile (2marks)
- v. Quartile deviation (1marks)

e) Find the variance and standard deviation for the data 3, 6, 9, 10, 7, 12, 13, 15, 6, 5, 13
(4marks)

QUESTION TWO(20MARKS)

- a) Explain the difference between correlation and regression (2marks)
- b) The table below shows verbal reasoning test score x and an English test score y for each of random sample of children who took both tests

child	A	B	c	D	E	F	G	H
x	112	113	110	113	112	114	109	113
y	69	65	75	70	70	75	68	76

- i. Calculate the product moment correlation coefficient between the scores (6marks)
- ii. Comment on the results obtained in (i) above (2marks)



- c) A horse was subject to the test of how many minutes it takes to reach a point from the starting point. The horse was made to carry luggage of various weights on 10 trials. The data collected are presented below in the table

Trial	1	2	3	4	5	6	7	8	9	10
Load (Kgs)	11	23	16	32	12	28	29	19	25	20
Time taken (in mins)	13	22	16	47	13	39	43	21	32	22

- i. Find the regression equation between the load and time taken to reach the goal (6marks)
- ii. Estimate the time taken to reach the goal for the loads 35kgs and 23kgs.(4marks)

QUESTION THREE(20MARKS)

- a) The numbers of public libraries in operation in selected states are listed below. Organize the data in a stem and leaf plot.

102,176,210,142,189,176,108,113,205,209,184,144,108,192,176 (4marks)

- b) For the following frequency distribution given below

Class	1-4	5 — 8	9- 12	13 — 16	17 -20	21-24
Freq	10	14	20	16	12	8

Estimate;

- i. Mean (3marks)
- ii. Median (3marks)
- iii. Mode (3marks)
- iv. Construct an ogive for the data (4marks)

- c) A tycoon has 3 house girls who he pays Ksh 4000 each per month, 2 watchmen who he pays Ksh 5000 each and some garden men who receives Ksh 7000 each. If he pays out an verage of Ksh 5700 per month to these people, find the number of garden men. (4marks)

QUESTION FOUR(20MARKS)

- a) Distinguish between kurtosis and skewness (2marks)
- b) Consider the following data which relate to the age distribution of 1000 workers in an industrial establishment

Age	Below 25	25-30	30-35	35-40	40-45	45-50	50-55	55 & above
Freq	120	125	180	160	150	140	100	25

Compute Bowley's coefficient of skewness and interpret the results (10marks)

- c) Using a clearly labeled diagram, show the position of mean, median and mode of a positively Skewed distribution (3marks)
- d) Explain what you understand by the terms Leptokurtic, Mesokurtic and Pletykurtic and illustrate it using a well labelled diagram (5marks)

QUESTION FIVE(20MARKS)

- a) Distinguish between probability mass function and probability density function (2marks)
- b) A random variable X has the following probability distribution

x	1	2	3	4	5	6
P(X=x)	K	3k	5k	7k	9k	11k

Determine k hence compute $p(1 \leq x \leq 5)$ and $p(x > 3)$ (7marks)



- c) Following a flood, 120 tins were recovered from Dharmesh's corner shop. Unfortunately, the water had washed off all the labels. Of the tins 50 contained pet food, 20 contained peas, 35 contained beans and the rest contained soup.
- a) Dharmesh selects a tin at random. Find the probability that it;
- i. Contains soup (1mark)
 - ii. Does not contain pet food (1mark)
- b) Dharmesh selects two tins at random (without replacement). Find the probability the;
- i. Both contains peas (2marks)
 - ii. One contains pet food and the other contains peas (2marks)
- c) Dharmesh selects three tins at random (without replacement). Find the probability that one contains pet food, one contains peas and one contains beans. (4marks)