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

FIRST YEAR, SECOND SEMESTER EXAMINATION FOR THE DEGREE OF BACHELOR OF ACTUARIAL SCIENCE, BACHELOR OF MATHEMATICS AND COMPUTER SCIENCE, BACHELOR OF SCIENCE IN STATISTICS, BACHELOR OF SCIENCE IN MATHEMATICS, BACHELOR OF SCIENCE IN EDUCATION SCIENCE, BACHELOR OF EDUCATION ARTS, BACHELOR OF SCIENCE IN MATHEMATICS/PHYSICS, BACHELOR OF EDUCATION TECHNOLOGY IN CIVIL ENGINEERING, BACHELOR OF EDUCATION TECHNOLOGY IN ELECTRIAL AND ELECTRONIC ENGINEERING AND BACHELOR OF EDUCATION TECHNOLOGY IN MECHANICAL ENGINEERING

SMS 3151: PROBABILITY AND STATISTICS I

DATE: APRIL 2024 **TIME: 2 HOURS**

INSTRUCTIONS: Answer Question ONE and any other TWO questions.

QUESTION ONE (30 MARKS)

a) Distinguish between quantitative and qualitative data. [2 Marks]

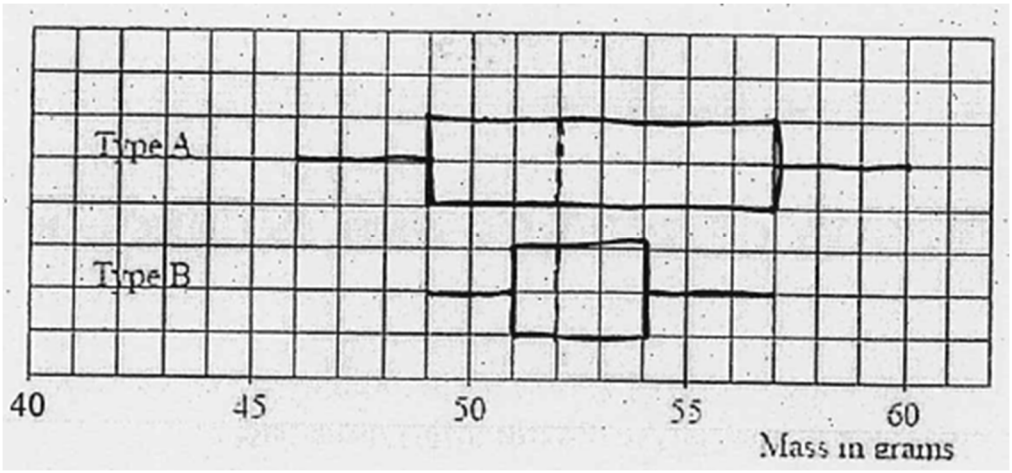
b) For each of the following variables, state an appropriate data collection method.

[4 Marks]

Variable	Proposed data collection method
Blood group	
Gender	
Weight at birth	

Number of children ever born to a woman	

c) A gardener collected data on two types of tomato planted in his farm. The box and whisker plot below shows data for the masses in grams of the tomatoes in the two samples.



i. Give the five number summary for sample A [5 Marks]

ii. Compute the interquartile range for each sample [2 Marks]

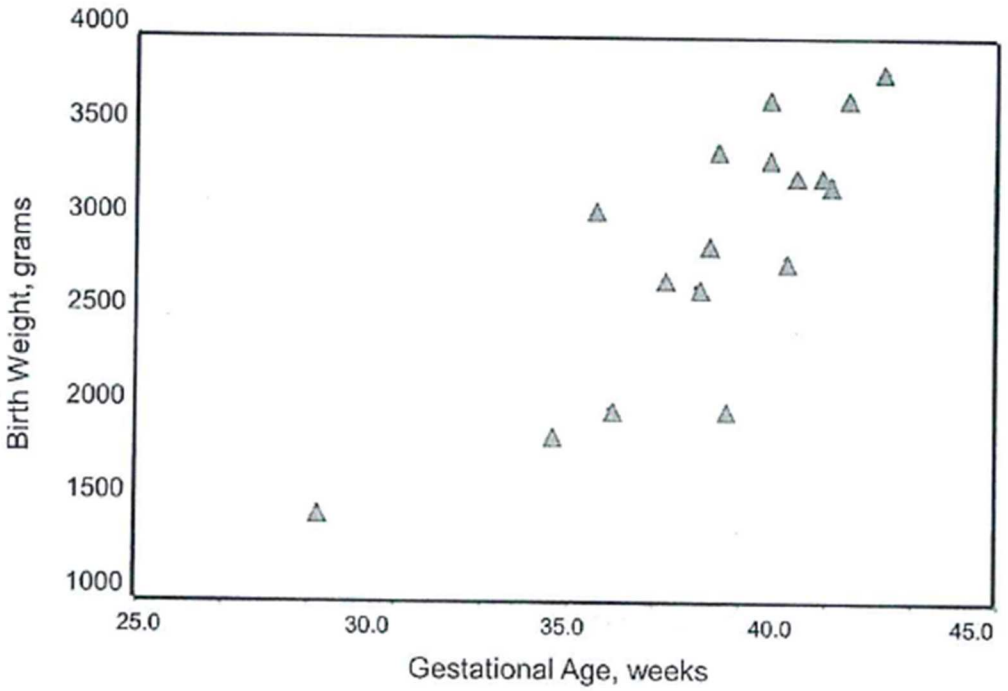
iii. Compute the Bowley's (Quartile) coefficient of skewness for sample B data. Comment on your answer. [2 Marks]

iv. Compare and contrast the two types and advise the gardener which type of tomato he should grow in future. [2 Marks]

d) A small-scale study was conducted involving 17 infants to investigate the association between gestational age at birth (X), measured in weeks, and birth weight (Y), measured in grams. The data is presented in the form of a scatter diagram and the following statistics were computed from the data: Variance of gestational age = 10

Variance of birth weight 485578.8 and

Covariance between gestational age and birth weight = 1798



i. From the scatter diagram, comment on the nature of association, if any, between birth weight and gestational age. [1 Mark]

ii. Determine the product moment correlation coefficient between birth weight and gestational age. Interpret your answer. [2 Marks]

iii. Compute the coefficient of determination between birth weight and gestational age. Interpret your answer. [2 Marks]

a) State the Baye's theorem of conditional probability for some two random events A and B [1 Mark]

b) Suppose a box has 3 red marbles and 2 black ones. We select 2 marbles one after the other, without replacement and check its colour. What is the probability that second marble is red given that the first one is red? [3 Marks]

c) A six-sided die has four green and two red faces and is balanced so that each face is equally likely to come up. The die will be rolled several times. Suppose that the score is 4 if the die is rolled and comes up green, and 1 if it comes up red. Define the random variable X to be the score obtained.

i. Write down the distribution of probability for X. [2 Marks]

ii. Calculate the expectation of X. [2 Marks]

QUESTION TWO (20 MARKS)

a) A group of twelve children participated in a psychological study designed to assess the relationship, if any, between ages, x years, and average total sleep time (ATST) ,y hours. To obtain a measure for ATST, recordings were take on each child on five consecutive nights and then averaged. The results are shown in the table.

Child												
Age (x years)	8.6	6.5	9.3	9.8	7.8	9.4	10.2	8.8	12.1	10.6	7.9	8.2
ATST(y hours)	8.7	6.7	8.7	8.8	8.1	9.0	9.5	9.1	9.7	9.3	8.9	8.5

$\sum x = 109.20 \quad \sum y = 105.00 \sum x^2 = 1{,}017.84 \quad \sum y^2 = 925.42 \quad \sum xy = 966.24$

i. Calculate the mean of x and y. [12 Marks]

ii. Estimate the values of $S_{x,} S_y$ and S_{xy} [13 Marks]

iii. Calculate the value of the product moment correlation coefficient between x and y (round your answer to 4 d.p) [2 Marks]

iv. Highlight two limitations of using correlation in measuring relationship between variables. [2 Marks]

b) The probability and statistics class from Meru University of Science & Technology (MUST) scored the following marks in their continuous assessment test (CAT) x, and final exams, y.

CAT, x	13	18	19	9	11	14	15	8	21	16
Final, y	39	42	48	32	40	46	38	33	51	39

i. Sketch a scatter plot for the above data. Comment on your graph [4 Marks]

ii. Fit a least square regression equation to the data. [6 Marks]

iii. Using the regression equation, calculate the expected final exam marks for a student that scores 17 is the continuous assessment test. [1 Mark]

QUESTION THREE (20 MARKS)

The table below represents the number of visits made to the University library by 100 Year 1 students in a given semester.

Marks	5	10	15	20	25	30	35
Frequency	4	10	20	36	16	12	2

i. Calculate first four moments about mean (central moments). [10 Marks]

ii. Calculate the coefficient of variation for these data. [2 Marks]

iii. Plot a frequency graph for these data; hence discuss the tail behavior and peakedness of this distribution. [4 Marks]

iv. Determine the moment's coefficient of skewness and kurtosis for these data and comment on your answer. [4 Marks]

QUESTION FOUR (20 MARKS)

a) The data below shows the weights (in Kilograms) of 20 gym members.

84, 86, 87, 88, 90, 92, 92, 94, 95, 98, 99, 99, 100, 104, 104, 118, 124,124, 146

i. Determine the range of these data. [1 Mark]

ii. Compute the Kelly's percentile coefficient of skewness and comment on your answer. [5 Marks]

iii. Represent these data in a grouped frequency distribution table with equal class sizes, beginning with 84-88, 89-93, [2 Marks]

iv. State one disadvantage of representing data in the form of grouped frequency tables as opposed to arrays. [1 Mark]

v. Using the frequency table above, estimate the weight of the 14^{th} gym member. [3 Marks]

vi. Construction a well-labelled histogram of weights

vii. On the same chart as the histogram in part (g) above, construct a frequency polygon for these data. [2 Marks]

viii. Taking an assumed mean of 95.5, use the coding method to find the arithmetic mean of weight of gym members. [3 Marks]

QUESTION FIVE (20 MARKS)

a) Differentiate between classical and axiomatic approach to probability. [12 Marks]

b) Suppose we have 60 transistors of which 12 are defective. If 4 of these transistors are selected at random, calculate the probability of selecting at least 3 good transistors. [2 Marks]

c) A small shop has a "pick-n-mix" counter where customers may choose wine gums, jelly beans or cola bottles. The probability that a customer purchases cola bottles is 0.45 ,jelly beans and wine gums 0.19, cola bottles and jelly beans 0.15, cola bottles and wine gums 0.25, cola bottles or jelly beans 0.6, cola bottles or wine gums 0.84, and at least one of them 0.9. Calculate the probability that a customer purchases:

i. jelly beans [2 Marks]

ii. wine gums [2 Marks]

iii. all three [3 Marks]

iv. none of them [1 Mark]

d) A discrete random variable Y has a pmf given by the table below

x	0	1	2	3	4
P(X=x)	k	2k	5k	10k	17k

Find the value of the constant k hence compute $P(1 \leq X < 3)$. [4 Marks)

e) If a random variable X has a mean of 10 and standard deviation of 4, calculate:

i. $E\left[\frac{5X-3}{4}\right]$ [2 Marks]

ii. $var\left[\frac{3X+5}{7}\right]$ [2 Marks)