MACHAKOS UNIVERSITY

University Examinations 2018/2019
SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

FIRST YEAR SPECIAL/SUPPLEMENTARY EXAMINATION FOR
BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE
BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING
BACHELOR OF SCIENCE IN MATHEMATICS
BACHELOR OF ECONOMICS AND STATISTICS
BACHELOR OF EDUCATION
BACHELOR OF ARTS

SMA 160 : INTRODUCTION TO PROBABILITY AND STATISTICS

DATE: 23/7/2019 **TIME: 8:30 – 10:30 AM**

Instructions to the Candidate:

1. *Answer Question* 1 *and any other* ***two*** *questions;*
2. *You must have a Scientific calculator for this paper.*

1. (a) Differentiate between the terms *primary data* and *secondary data* as used in Statistics. (4 *marks*)

(b) Explain in words each of the following terms as used in Statistics:

(i) mean;
(ii) median. (4 *marks*)

(c) The data given below represents the frequency distribution of grades scored in Statistics by 200 diploma students the national examination last year.

Grade scored	1	2	3	4	5	6	7	8
No. of students	5	15	30	42	46	36	18	8

Determine each of the following measures about the distribution the grades scored:

(i) mean;
(ii) median;
(iii) standard deviation. (6 *marks*)

(d) An agricultural scientist carried out a study on the relationship between maize crop yield and amount of fertilizer per unit area of land. The statistical analysis for the two variables yielded the linear least squares regression model $y = 10.254 + 2.582x$: where y is the number of sacks of maize produced and x in the number of bags of fertilizer used per acre of land.

Interpret the regression model above. (4 *marks*)

(e) A committee of 6 members is to be formed from a group of 8 men and their wives. Determine the number of ways in which the committee can be constituted if a man and his wife cannot both serve in the same committee. (4 *marks*)

(f) Two identical boxes contain apples. The first box contains 8 red apples and 4 green apples while the second box contains 6 red apples and 10 green apples. The apples are identical in all aspects except colour. A customer picks an apple at random one from one of the boxes selected at random.

(i) Present this information using a probability tree diagram. (3 *marks*)

(ii) Determine the probability that a red apple is picked. (2 *marks*)

(iii) If a red apple is picked, determine the probability that it is from the first box. (3 *marks*)

2. (a) The table below shows the age distribution of 800 employees working at the headquarters of a county government.

Age (years)	20 - 24	25 - 29	30 - 34	35 - 39	40 - 44	45 - 49	50 - 54	55 - 59	60 - 64
No. of employees	10	45	120	180	160	125	95	50	15

(i) Estimate the standard deviation of the age of the employees. (4 *marks*)

(ii) Due to continuous burden of the huge wage bill, it is desired to reduce the present number of employees according to the following scheme:

- To retrain the first 15% from the lower age group.
- To retain the next 45% in their current workstations.
- To transfer the next 30% to the national government.
- To retrench the 10% from the highest age group.

Determine the age limits of the persons retained and those to be transferred to the national government. (6 *marks*)

(iii) Suppose the employees whose age is between 32 and 42 years are to be retained at their current workstations, determine the proportion of the employees who will be retained. (4 *marks*)

(b) There is an outbreak of a certain mysterious disease in a certain area. This disease has a known symptom. However, medical experts realised that not all people who have the disease display the symptom, and not all those who display the symptom have the disease. After conducting a medical survey among a random sample of 540 residents in the area, the findings were as follows:

- 200 patients who displayed the symptom had the disease;
- 120 patients who displayed the symptom did not have the disease;
- 150 patients who had the disease did not display the symptom;
- 70 patients did not have the disease and did not display the symptom;

Determine the following probabilities:

(i) A patient who has the disease is selected at random from the area, determine the probability that he displays the symptom. (3 *marks*)

(ii) A patient who displays the symptom is selected at random from the area, determine the probability that he has the disease. (3 *marks*)

3. An economist carried out a study in the relationship between the monthly salary of an employee and the amount of sacco saving among Machakos County employees. A random sample of 12 typical employees of Machakos County was taken. The data relating the monthly salary in thousand Kenya shillings and the corresponding monthly saving in thousand Kenya shillings is shown in the table below.

Employee	A	B	C	D	E	F	G	H	J	K	L	M
Salary	48	12	48	12	48	12	48	12	48	12	48	12
Saving	78	16	78	16	78	16	78	16	78	16	78	16

(a) (i) Compute the Pearson's product moment co-efficient of correlation between the monthly salary and the monthly saving of the employees. (9 *marks*)

(ii) Interpret the result obtained in (a) (i) above. (2 *marks*)

(b) (i) Determine the least squares regression line of the monthly saving on the monthly salary for the employees. (4 *marks*)

(ii) Interpret the regression line obtained in (b) (i) above. (3 *marks*)

(iii) Using the regression line obtained in (b)(i), estimate by calculation the monthly saving for a typical employee whose monthly salary is Ksh 50,000. (2 *marks*)

4. (a) By taking the simple linear regression model $y_i = a + bx_i + e_i$ derive each of the following using the least squares technique:

(i) The two normal equations;

(ii) The regression intercept a and the regression co-efficient b in the regression model $\hat{y}_i = a + bx_i$ (10 *marks*)

where $b = \dfrac{\sum xy - n\,\bar{x}\,\bar{y}}{\sum x^2 - n\bar{x}^2}$ and $a = \bar{y} - b\,\bar{x}$

(b) A property ownership survey conducted by the Ministry of Planning among 700 residents randomly selected from a group of estates in Nairobi showed the following information about the ownership of a personal car, a rental house and a matatu:

- 295 residents own a personal car.
- 225 residents own a rental house.
- 235 residents own a matatu.
- 100 residents own a personal car and a rental house.
- 105 residents own a personal car and a matatu.

- 75 residents own a rental house and a matatu.
- 30 residents own a personal car, a rental house and a matatu.

(i) Present this information in a Venn diagram. (6 *marks*)

(ii) Determine the number of residents in the study who own:

I. exactly one type of asset.

II. own a car or house but not a matatu.

III. own a car and house but not a matatu. (4 *marks*)

5. (a) Differentiate between *categorical data* and *numerical data*, illustrating with relevant examples from a real life situation. (4 *marks*)

(b) A landline phone number system in Nairobi uses 10 digits to assign phone numbers to its subscribers. However, the first three digits must be a 0 (zero), 2 and 0 in that order, while the, the fourth digit must not be a 0 (zero). Determine the maximum number of subscribers which can be accommodated by this coding system. (4 *marks*)

(c) In a certain machine, the following actions are possible; the top can blow or not; the spiggots can go up or go down or remain still; and the cum-lever can wriggle or thump. The top blows with a probability of $\frac{1}{5}$. The spiggots go up with a probability of $\frac{1}{2}$ if the top blows; otherwise they go up or go down or remain still with equal probabilities. The cum-lever wriggles if the spiggots go up; otherwise it wriggles with a probability of $\frac{1}{6}$. The machine is robust and works provided the cum-lever is thumping.

(i) Present this information in a probability tree diagram;

(ii) Determine the probability that the machine works;

(iii) Determine the probability that the top blows and the cum lever wriggles;

(iv) If the machine breaks down, determine the probability that the top is blowing. (12 *marks*)