



MACHAKOS UNIVERSITY

University Examinations 2016/2017

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

FIRST YEAR SECOND SEMESTER EXAMINATION FOR DEGREE IN

BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING

BACHELOR OF ECONOMICS AND STATISTICS

BACHELOR OF SCIENCE IN MATHEMATICS

BACHELOR OF EDUCATION (SCIENCE)

BACHELOR OF EDUCATION (ARTS)

BACHELOR OF ARTS

SMA 160: INTRODUCTION TO PROBABILITY AND STATISTICS

Date: 30/5/2017

Time: 2:00 – 4:00 PM

INSTRUCTIONS

Answer Question One and any other two questions

1. (a) Explain each of the following terms as used in Statistics:
 - (i) population;
 - (ii) variable. (4 marks)
- (b) Differentiate between the terms primary data and secondary data as used in Statistics.(4 marks)
- (c) The data given below represents the frequency distribution of marks scored in Mathematics by 250 diploma students in a KCSE National Examination last year.

Grade scored	1	2	3	4	5	6	7	8
No. of students	8	20	36	64	48	40	24	10

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

- (i) Median;
- (ii) Mean;
- (iii) Standard deviation. (7 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 coefficient of correlation between crop yield and amount of fertilizer used was computed and found to be $r = 0.8254$

Compute the co-efficient of determination between the crop yield and amount of fertilizer used, and interpret the result obtained. (5 marks)

- (e) A box contains 4 green apples and 8 red apples which are identical in all aspects except colour. A customer picks 2 apples at random from the box one at a time without replacement.

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

(ii) Hence determine the probability that all the apples picked are of different colours. (2 marks)

- (f) A college collected the following set of bivariate data on the number of Credit grade and the number of hours of study taken during the week. A random sample of 6 students was taken and the data relating the hours of study and the number of credits is as shown in the table.

Student	A	B	C	D	E	F
Hours	20	50	30	50	25	23
Credits	13	16	12	16	15	16

Determine the Spearman's rank correlation co-efficient between the number of hours and credits: (5 marks)

2. The data given below shows the number of hours worked in one week by employees in a certain company:

46.3	39.2	44.2	41.3	45.1	42.3	43.5	40.0
45.6	40.6	42.0	42.6	45.6	39.5	43.1	39.7
46.1	38.9	42.4	42.1	45.0	44.4	42.4	40.8

- (a) Tabulate a frequency distribution table with class interval of 0.5 starting at 38.9 – 40.4, etc. (6 marks)

- (b) Using the table in (a) above, estimate by calculation each of the following statistical measures:

(i) Median; (2 marks)

(ii) Mode; (3 marks)

(iii) Mean; (3 marks)

(iv) Quartile deviation. (6 marks)

3. The Ministry of Lands and Housing is interested in investigating the relationship between the monthly rent of a house and the distance from the nearest tarmac road in Machakos County. A survey was carried out in which a random sample of 10 typical two bed-roomed residential houses in Machakos County was taken. The data relating the distance from the nearest tarmac road in hundred metres and the monthly rent in thousand Kenya shillings is as shown below.

House	A	B	C	D	E	F	G	H	J	K
Distance	24.5	18.8	15.2	35.5	18.6	32.0	26.2	45.4	21.5	12.2
Rent	15	28	20	16	32	10	12	08	20	35

- (a) (i) Compute the Pearson's product moment co-efficient of correlation between the distance from the tarmac road and the monthly rent paid for a house. (9 marks)
- (ii) Interpret the result obtained in (a) (i) above. (2 marks)
- (b) (i) Determine the least squares regression line of the monthly rent paid on the distance from the nearest tarmac road. (4 marks)
- (ii) Interpret the regression line obtained in (b) (i) above. (3 marks)
- (iii) Using the regression line obtained in (b)(i) above, estimate by calculation the monthly rent for a typical two bed-roomed house which is 2500 metres away from the nearest tarmac road. (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 normal equations;
- (ii) The regression intercept a and the regression co-efficient b in the regression model given by $\hat{y}_i = a + bx_i$ (10 marks)

$$\text{where } b = \frac{\sum xy - n \bar{x} \bar{y}}{\sum x^2 - n \bar{x}^2} \quad \text{and} \quad a = \bar{y} - b \bar{x}$$

- (b) A study was conducted by the Ministry of Education on the academic performance of students in compulsory subjects in KCSE. A random sample of 460 students who sat for KCSE in a certain year found out the following:
- 214 students passed in Mathematics
 - 230 students passed in English
 - 194 students passed in Kiswahili
 - 92 students passed in Mathematics and English
 - 80 students passed in Mathematics and Kiswahili
 - 90 students passed in English and Kiswahili
 - 48 students did not pass in any of the three subjects
- (i) Present this information in a Venn diagram. (6 marks)
- (ii) Determine the number of students in the study who passed in:
- exactly two of the subjects;
 - English or Kiswahili but failed in Mathematics. (4 marks)

5. (a) Explain the **four** measurement scales for statistical data as outlined by Stevens (1946) giving two examples in each case. (8 marks)
- (b) 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)