THE CATHOLIC UNIVERSITY OF EASTERN AFRICA

A. M. E. C. E. A

MAIN EXAMINATION

MAY – JULY 2018 TRIMESTER

FACULTY OF SCIENCE

DEPARTMENT OF MATHEMATICS AND ACTUARIAL SCIENCE

REGULAR PROGRAMME

MAT 160: PROBABILITY AND STATISTICS I

Date: JULY 2018	Duration: 2 Hours
INSTRUCTIONS: Answer Question ONE and any other TWO Questions	

Q1. a) Define the following terms   **(6marks)**
- i)   Primary data
- ii)   Positive correlation
- iii)   Skewness

b) A certain data set was collected an a random variable X and the following summary was made
$\sum_{i=1}^{40} x_i = 240$ and $\sum_{i=1}^{40} x_i^2 = 2520$. Calculate the mean and variance of the data   **(4 marks)**

c) If the coefficient of correlation $r = 0$, show that the two regression lines $y\ on\ x$ and $x\ on\ y$ are at right angles   **(4 marks)**

d) For a given set of data, it is known that $\bar{x} = 10$ and $\bar{y} = 4$. The gradient of the regression line $y\ on\ x$ is 0.6. Find the equation of this regression line and estimate y when $x = 12$ **(4 marks)**

e) These are examination marks for a group of 120 students. Represent the data on a histogram and comment on the shape of the distribution   **(5 marks)**

Mark	0-9	10-19	20-29	30-49	50-79
Frequency	8	21	53	28	10

f) Show that $s^2 = \frac{\sum_{i=1}^{n} x_i}{n} - \bar{X}^2$ **(5 marks)**

ISO 9001:2008 Certified by the Kenya Bureau of Standards

g) Distinguish between linear and nonlinear correlation **(2 marks)**

Q2. a) The number of errors X on each of 200 pages of a typescript was monitored. The results showed that $\sum_{i=1}^{n} x_i = 920, \sum_{i=1}^{n} x_i^2 = 5032$.

     i) Calculate the mean and standard deviation of the number of errors per page **(4 marks)**

     ii) A further 50 pages were monitored and it was found that the mean was 4.4 errors and the standard deviation was 2.2 errors. Find the mean and standard deviation of the number of errors per page for the 250 pages **(8 marks)**

b) Consider the following information $\sum_{i=1}^{n} x_i = 165, \sum_{i=1}^{n} y_i = 155, \sum_{i=1}^{n} x_i^2 = 3007, \sum_{i=1}^{n} y_i^2 = 2637, \sum_{i=1}^{n} x_i y_i = 2665 \ and \ n = 11$. Compute

     i) The regression line $X \ on \ Y$ **(4 marks)**
     ii) The regression line $Y \ on \ X$ **(4 marks)**

Q3. a) State the main characteristics of statistics **(10 marks)**

b) Fill in the missing values in the table below given that s is the standard deviation **(10 marks)**

n	$\sum_{i=1}^{n} x_i$	$\sum_{i=1}^{n} x_i^2$	$\bar{X}$	s
63	7, 623	924, 800		
	152.6		10.39	1.7
52		57, 300	33	
18			57	4

Q4. a) Salt is packed in bags which the manufacturer claims to contain 25 kg each. Eighty bags are examined and the mass, $x \ kg$, of each is found. The results are $\sum(x - 25) = 27.2$ and $\sum(x - 25)^2 = 85.1$. Find the mean and standard deviation of the masses **(6 marks)**

b) Two judges rank eight photographs taken in a competition as follows. Calculate Spearman's rank coefficient of correlation for the data **(4 marks)**

Photograph	A	B	C	D	E	F	G	H
Judge 1	2	5	3	6	1	4	7	8
Judge 2	4	3	2	6	1	8	5	7

c) Consider the following data where X represents the female life expectancy and Y represents income.

Country	Afghanistan	Sri Lanka	Bhutan	India	Pakistan	Bangladesh
X	42	50	47	58	57	73
Y	143	179	197	335	384	423

     i)   Using method of regression coefficients, find the equation suitable for this data set **(8 marks)**

    ii)   Estimate the value X for Nepal where the value Y=160 **(1 mark)**

    iii)   Estimate the value Y for Japan where the value Y=858 **(1 mark)**

Q5. a) In a recent test taken by 106 students, the results were distributed as follows.

Marks	1-5	6-10	11-15	16-20	21-25	26-30
Frequency	8	9	20	33	22	14

Use the data to determine the following

    i)   The mean mark **(3 marks)**

    ii)   The standard deviation of the marks **(4 marks)**

    iii)   Pearson's measure of skewness and comment on your answer **(4 marks)**

    iv)   The moment coefficient of kurtosis and comment on your answer **(4 marks)**

    v)   The mode and median of the data **(3 marks)**

  b)   Define the term correlation as used in statistics **(2 marks)**

END