UNIVERSITY OF EMBU

2017/2018 ACADEMIC YEAR

SECOND SEMESTER EXAMINATIONS

FIRST YEAR EXAMINATION FOR THE DEGREE OF BACHELOR OF SCIENCE, BACHELOR OF SCIENCE (COMPUTER SCIENCE)

CSC/SIT124: PROBABILITY AND STATISTICS

DATE: APRIL 9, 2018   TIME: 8:30 AM – 10:30 AM

INSTRUCTIONS:

Answer Question ONE and ANY other two Questions

QUESTION ONE (30 MARKS)

a) Differentiate between

   i)   Discrete and continuous data   (2 marks)

   ii) Type I and Type II errors   (2 marks)

b) The table below shows the changes in water temperature measured every two minutes after putting on the heater.

Time (mm)	0	2	4	6	8	10	12
Temperature (^{o}c)	21.5	32.5	37.7	40.1	41.2	43.6	45.6

   i)   Illustrate this type of data using the appropriate presentation method

                                                                        (3 marks)

   ii) Give reasons for selecting this type of data presentation method   (2 marks)

---

c) It is assumed that the number of vehicle accidents (X) on Nakuru – Eldoret road per week assumes a poisson distribution with mean number of accidents per week equal to ($\lambda = 4$). Calculate the probability of more than five accidents occurring in any one week

(4 marks)

d) A survey conducted on 20 households in Njukiri sub location by a student on research showed the number of family members per household as indicated in the table:

Family size	1 - 3	3 – 5	5 - 7	7 - 9	9 - 11
Number of families	7	8	2	2	1

    i)   Find the mode of distribution   (3 marks)

    ii)   Hence comment on household family size of Njukiri sub location   (1 mark)

e) Explain the following three form of statistics

    i)   Descriptive statistics   (2 marks)

    ii)   Inferential statistics   (2 marks)

    iii) Socio-economics statistics   (2 marks)

f) Outline the principles of data tabulation and graphical representation in statistics

(5 marks)

g) If x is a continuous random variable then show that its distribution function:

$$f(x) = \begin{cases} 3x^2 , & o \leq x \leq 1 \\ 0, & otherwise \end{cases}$$

is a probability density function   (3 marks)

QUESTION TWO (20 MARKS)

a) Distinguish between

    i)   Absolute deviation and standard deviation: (2 marks)

    ii)   Percentile and Decile   (2 marks)

b) A survey was carried out one day at University of Embu to find out how many employee repot to work late after 8.00am and with how many minutes. The table below shows the data that was collected.

Time late to work (min)	1 - 10	11 - 20	21 - 30	31 - 40	41 - 50	51 - 60
Number of employees	8	14	12	9	7	4

i) Show that the 50th percentile minute of lateness is the same as the 5th decile.

(6 marks)

ii) Determine the number of university employees that reported later than 80th percentile minute on that day. (3 marks)

iii) Estimate the interquartile range of lateness minutes. (3 marks)

QUESTION THREE (20 MARKS)

a)

i) Differentiate between Permutation and Combination (2 marks)

ii) A company Human Resource Manager has been asked to prepare codes for each employee. Each code should have 4 digits in a specific order, the digits are between 0 to 9. How many different codes are there if one digit may only be used once? (3 marks)

b)

i) Define conditional probability (2 marks)

ii) It has been reported that 0.5% of a population has a particular disease. A test is developed for the disease. The test gives a false positive of 3% of the time and a false negative of 2% of the time. What is the probability that person randomly picked from the area tests positive? (4 marks)

iii) If the person just got the bad news that the test came back positive; what is the probability that Joe has the disease? (5 marks)

c) Explain your understanding of Markov inequalities (4 marks)

QUESTION FOUR (20 MARKS)

a) Explain the difference between Median and Mode of data distribution (2 marks)

b) Data collected from field with 100 observations was grouped into 10 classes and the median measurement was established as 525. The data was summarised in the table as;

Measurement	0-100	100-200	200-300	300-400	400-500	500-600	600-700	700-800	800-900	900-1000
Frequency	2	5	x	12	17	20	y	9	7	4

    i) Determine the value of x and y   (7 marks)

    ii) Calculate the mean and mode of the data distribution   (6 marks)

c) Embu post-office handles on average 10,000 letters per day with a variance of 2,000 letters. Using chebychev's inequalities theorem calculate probability that this post office will handle between 8,000 and 12,000 letters tomorrow?   (5 marks)

QUESTION FIVE (20 MARKS)

a) Outline the process of Hypothesis testing in scientific research   (6 marks)

b) The data in the table gives the distribution of total monthly household expenditure of 200 families of Nemburi Division, Embu county.

Expenditure (Kshs)	1000 - 1500	1500-2000	2000-2500	2500-3000	3000-3500	3500-4000	4000-4500	4000-5000
Number of families	24	40	33	28	30	22	16	7

    i) Illustrate this information using a histogram   (4 marks)

    ii) Explain the message that can be deduced from the histogram   (4 marks)

c)

    i) Define exponential distribution function for variable x.   (2 marks)

    ii) On the average, a certain computer part lasts 10 years. The length of time the computer part lasts is exponentially distributed. Determine probability that a computer part picked randomly will last more than 7 years   (4 marks)

--END--