# MACHAKOS UNIVERSITY

**University Examinations 2021/2022 Academic Year**

**SCHOOL OF PURE AND APPLIED SCIENCES**

**DEPARTMENT OF MATHEMATICS AND STATISTICS**

**SECOND YEAR FIRST SEMESTER EXAMINATION FOR**

**BACHELOR OF SCIENCE COMPUTER SCIENCE**

**SCO 212 PROBABILITY & STATISTICS FOR COMPUTER SCIENCE**

**DATE: 14/12/2021** **TIME: 2.00-4.00 PM**

**INSTRUCTION:**

***Answer Question One and Any Other Two Questions***

**QUESTION ONE (30 MARKS) COMPULSORY**

a) Explain the meaning of the following terms as applied in Statistics

i. Sample

ii. Variable (4 marks)

b) Differentiate between EACH of the following terms:

i. Descriptive and inferential statistics (2 marks)

ii. Point and interval estimation (2 marks)

c) The *pmf* of a discrete random variable X is given $p(X = x) = kx \ for\ x = 1,2,3,4,5,6$

Determine

i. the value of k. (3 marks)

ii. the mean of the distribution. (2 marks)

iii. the $p(X < 4)$ (2 marks)

d) Given that x is a continuous random variable with a density function defined as

$$p(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0, x < 0 \end{cases}$$

Show that p(x) is a probability density function (5 marks)

e) The pdf of a random variable is given by

$$f(x) = \begin{cases} \frac{1}{4}, x = 0,2 \\ \frac{1}{2}, x = 1 \\ 0, elsewhere \end{cases}$$

Determine the moment generating function for x. (4 marks)

f) A study of 49 randomly chosen 8-year-olds shows that they watch television an average of 38 hours per week with a standard deviation of 6.4 hours. Construct a 99% confidence interval for the average time per week that all such children watch television (6 marks)

**QUESTION TWO (20 MARKS)**

a) A random variable x has a probability function given by

$$f(x) = \begin{cases} kxe^{-3x}, for\ x > 0 \\ 0, elsewhere \end{cases}$$ Where k is aconstant,

Determine

i. The value of k (5 marks)

ii. The expected value of x (5 marks)

b) The mean number of strikes in a particular industry was found to be 1.2 per week. Determine the probability that during a given week there will be:

i. No strikes (3 marks)

ii. More than 2 strikes (4 marks)

iii. Exactly 4 strikes (3 marks)

**QUESTION THREE (20 MARKS)**

a) Explain the meaning of the following terms as used in probability theory:

i. An event

ii. Random experiment

iii. Mutually exclusive events

iv. Independent events (8 marks)

b) The data below shows the number of hours worked in one week by employees in certain IT organization:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 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 |

i. Tabulate a frequency distribution table with class intervals by 38.9 – 40.4, … etc (6 marks)

ii. Use the table in 2(a) above to calculate the quartile deviation (6 marks)

**QUESTION FOUR (20 MARKS)**

a) A continuous random variable x has a probability density function f(x) given by

$$f(x) = \begin{cases} 6(x - x^2), for\ 0 \leq x \leq 1 \\ 0, elsewhere \end{cases}$$

Calculate the first and the second moments hence find the variance of x. (8 marks)

b) The data below shows gross weekly earnings of employees by age of an IT company in the year 2008.

| Age (years) | 18 | 20 | 22 | 27 | 35 | 45 | 55 |
|---|---|---|---|---|---|---|---|
| Weekly earnings ('000) | 15.50 | 23.20 | 34.0 | 44.90 | 53.10 | 55.0 | 57.20 |

i. Calculate the least squares regression line of gross weekly earnings on age. (10 marks)

ii. Use the equation in (i) to estimate the weekly earnings of an employee aged 50 years (2 marks)

**QUESTION FIVE (20 MARKS)**

a) A TV network is concerned about the high cost of producing many of its programs. A study is conducted to relate the production costs for 30 minutes of programming (in hundreds of thousands of dollars) to the ratings that the program gets in the national ratings survey. The results are shown below:

| Production cost | 1.2 | 1.6 | 1.8 | 2.5 | 2.7 | 3.0 | 3.5 | 4.4 |
|---|---|---|---|---|---|---|---|---|
| Ratings | 3.3 | 3.9 | 5.7 | 4.2 | 4.5 | 8.2 | 6.1 | 4.6 |

Calculate the correlation coefficient for the ratings and the production cost. (8 marks)

b) A sample of 50 items is drawn from a population of manufactured products and weight x of each item is recorded. Prior experience has shown that the weight has a normal probability distribution with mean $\mu$=6kg and standard deviation $\sigma$ =2.5kg

i. Calculate the mean and variance of the sample (2 marks)

ii. Determine the probability that the sample has a mean weight of more than 6.25kg (4 marks)

iii. Determine the probability that the sample has a mean weight of less than 5.5kg (4 marks)

iv. How would the sampling distribution of $\bar{x}$ change if the sample size was increased to 100 (2 marks)