



Strathmore
UNIVERSITY

School of Computing and Engineering and Sciences
Computer Network & Cyber Security
End of Semester Examination
CNS 2206: Probability and Statistics I

DATE: 5th December 2022

Time: 2 Hours

Instructions

1. This examination consists of **FIVE** questions.
2. Answer **Question ONE (COMPULSORY)** and any other **TWO** questions.
3. Do not write on the question paper.

Question One (30 Marks)

(a) Differentiate between:

- | | |
|--|------------------|
| (i) Descriptive and inferential statistics | (2 marks) |
| (ii) Quantitative and qualitative data | (2 Marks) |

(b) The following data represents monthly wages of employees (in US dollars) in Company B.

43, 82, 42, 20, 53, 22, 50, 32, 28, 25, 20, 12, 25, 14

Using the data given obtain:

- | | |
|-------------------------|------------------|
| (i) Arithmetic Mean | (2 Marks) |
| (ii) Modal wage | (1 Mark) |
| (iii) Variance | (3 Marks) |
| (iv) Quartile Deviation | (2 Marks) |

(c) The number of clients a wholesale attendant handles within one minute, X , during a given day at a downtown premise in a city is a random variable with the following probability distribution

X	0	1	2	3	4	5
$P(X = x)$	0.18	0.39	0.24	k	0.04	0.01

- (i) Find the value of k , and determine the expected number of customers the attendant handles within 1 minute. **(2 Marks)**
- (ii) Compute the variance of the random variable, X . **(3 Marks)**
- (d) P and Q are independent events such that $P(X) = 0.35$ and $P(Y) = 0.65$. Find:
- (i) $P(X \cup Y)$ **(1 Mark)**
- (ii) $P(X \cap Y)$ **(1 Mark)**
- (e) An incomplete distribution of data obtained from a certain company is given below.

Variable	0-10	10-20	20-30	30-40	40-50	50-60	60-70
Frequency	10	20	F1	40	F2	25	15

- (i) If median is 35. Find out the missing frequencies, F1 and F2 given that the total frequency is 170. **(4 marks)**
- (ii) Compute the arithmetic mean of the completed table using 30 as an assumed mean. **(3 marks)**
- (f) Given the data below
- | | | |
|--------------------|----|----|
| | X | Y |
| Arithmetic mean | 70 | 50 |
| Standard deviation | 5 | 8 |
- Correlation coefficient between X and Y is 0.6
- (i) Obtain the regression equation of Y on X **(3 Marks)**
- (ii) Estimate the value of Y when $X = 56$. **(1 Mark)**

Question Two (20 Marks)

A survey was conducted to determine the salaries of employees in US dollars in a certain organization. The results of the survey are as follows:

Salary (US Dollars)	30-40	40-50	50-60	60-70	70-80	80-90	90-100	100-110
Number of workers	20	30	60	75	115	100	60	40

- (a) Calculate the first quartile (Q1) and third quartile (Q3) hence find the semi-interquartile range of the salaries. **(5 Marks)**
- (b) Compute the:
 - (i) Median **(4 Marks)**
 - (ii) Standard deviation **(5 Marks)**
- (c) Draw a cumulative frequency curve using the 'less than' method hence find the number of employees earning between 55 and 85 Canadian dollars. **(6 Marks)**

Question Three (20 Marks)

The following data shows the age in years of 200 automobiles.

Age of automobile	0-4	4-8	8-12	12-16	16-20	20-24	24-28
Number of automobiles	32	16	30	25	50	27	20

From the data given, find the:

- (a) Mean age **(4 Marks)**
- (b) Mode of the distribution **(4 Marks)**
- (c) Median age **(4 Marks)**
- (d) Standard deviation of the ages **(4 Marks)**
- (e) Obtain the Karl Pearson's coefficient of skewness and comment on the results. **(4 Marks)**

Question Four (20 Marks)

- (a) State and explain **three** types of correlation. **(3 Marks)**

The data below were extracted from the data base of a certain company regarding the age of employees and the day they were reported sick in a month.

	Employees									
	1	2	3	4	5	6	7	8	9	10
Age (X)	30	32	35	40	48	50	52	55	57	61
Sick Days (Y)	2	1	0	2	5	4	5	6	8	7

- (b) From the data given compute Karl Pearson's coefficient of correlation and interpret your results. **(6 marks)**

Hint:
$$r = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{N \sum X^2 - (\sum X)^2} \sqrt{N \sum Y^2 - (\sum Y)^2}}$$

- (c) Find the coefficient of determination and comment on the results. **(3 Marks)**
(d) Obtain the regression equation of age (X) on sick days (Y). **(5 Marks)**
(e) Predict the age of an employee whose sick days in a month is 13. **(3 Marks)**

Question Five (20 Marks)

- (a) Explain the difference between the following terms as used in probability:
(i) Dependent and independent events **(2 Marks)**
(ii) Mutually exclusive and compound events **(2 Marks)**
(b) The records of 500 interviewees are given below

Score	CNS	BBIT	ICS	Total
Below 60	90	40	90	220
Between 70 and 80	40	75	85	200
Above 80	10	35	35	80
Total	140	150	210	500

If an interviewee is selected from this group of interviewees, find:

- (i) The probability that his score is below 60, given that he is a Bachelor of Informatics and Computer Science (ICS) graduate. **(1 Mark)**
(ii) The probability that he is a CNS graduate. **(1 Mark)**
(iii) The probability that he is a Bachelor of Business and Information Technology (BBIT) graduate given that his score is above 80. **(2 Marks)**
(c) Given that $P(A) = 0.30$, $P(B^c) = 0.22$ and $P(A \cap B) = 0.16$, find $P(A \cup B)$. **(2 Marks)**

- (d) A trader has 65% chance of making a sale to each client. The behavior of successive clients is independent. If two clients Ben and Liz enter the trader's premises, what is the probability that the trader will make a sale? **(2 Marks)**
- (e) The number of mobile handsets sold, X , during a given day at a certain town is a random variable with the following probability mass function

X	0	1	2	3
$P(X = x)$	0.4	0.3	0.1	0.2

- (i) Show that the above distribution is a probability mass function. **(2 Marks)**
- (ii) Compute the $E(3X + 2)$. **(3 Marks)**
- (iii) Compute the $Var(2X - 1)$. **(3 Marks)**