



DEDAN KIMATHI UNIVERSITY OF TECHNOLOGY

University Examinations 2022

SECOND YEAR SECOND SEMESTER EXAMINATION FOR THE DEGREE IN BACHELOR OF SCIENCE IN CIVIL ENGINEERING AND BACHELOR OF EDUCATION IN CIVIL ENGINEERING

Date: 5th April 2022

SAS 2130: Probability and Statistics 1

Time: 8:30-10:30 am

Answer question **ONE (COMPULSORY)** and any other **TWO** questions. Show clearly your work and the number of questions attempted.

QUESTION ONE

[30 MARKS]

- (a) Define the term statistics and briefly explain why you need to be knowledgeable in the subject in your area of specialization. [3 marks]
- (b) Giving an example in each case differentiate between the following terms as used in statistics.
- i. Qualitative and quantitative data [2 marks]
 - ii. Nominal and ordinal data [2 marks]
 - iii. Census and sample survey [2 marks]
- (c) An assembly consists of three mechanical components. Suppose that the probabilities that the first, second, and third components meet specifications are 0.95, 0.98, and 0.99. Assume that the components are independent. Determine the probability mass function of the number of components in the assembly that meet specifications [4 marks]
- (d) In *Applied Life Data Analysis* (Wiley, 1982), Wayne Nelson presents the breakdown time of an insulating fluid between electrodes at 34 kV. The times, in minutes, are as follows: 0.19, 0.78, 0.96, 1.31, 2.78, 3.16, 4.15, 4.67, 4.85, 6.50, 7.35, 8.01, 8.27, 12.06, 31.75, 32.52, 33.91, 36.71 and 72.89. Calculate the sample mean and sample standard deviation [4 marks]
- (e) In a certain type of metal test specimen, the normal stress on a specimen is known to be functionally related to the shear resistance. The Following is a set of coded experimental data on the two variables:

<i>Normalstress(x)</i>	26.8	25.4	28.9	23.6	27.7	23.9	24.7	28.1	26.9
<i>Shearstress(y)</i>	26.5	27.3	24.2	27.1	23.6	25.9	26.3	22.5	21.7

- i. Estimate the regression line $y = \alpha + \beta x$ [5 marks]
- ii. Estimate the shear resistance for a normal stress of 24.5 kilograms per square centimeter [2 marks]

- (f) For a given data set the regression lines y on x and x on y are $y + 0.29x = 20.8$ and $0.785y + x = 16.2$ respectively. Find
- The product moment correlation coefficient [2 marks]
 - \bar{X} and \bar{Y} [4 marks]

QUESTION TWO

[20 MARKS]

- (a) The mode for the following incomplete distribution of weights of 160 students is 56.

<i>Weights(kg)</i>	30-40	40-50	50-60	60-70	70-80	80-90
<i>No.of students</i>	20	36	x	y	15	5

Assuming the weights are linearly distributed in each group, estimate the;

- Values of x and y [5 marks]
 - Median weight [3 marks]
 - Mean and standard deviation of the distribution, using the coding method and an assumed mean of 55kgs [6 marks]
 - Are values obtained in (ii) and (iii) above the actual values of those quantities? Give a reason for your answer. [2 marks]
- (b) The marks for a sample of 25 students from a large class in a recent test have a sample mean of 57.2 and a standard deviation 7.3. The marks are subsequently adjusted; each mark is multiplied by 1.1 and the result then increased by 8. Calculate the sample mean and standard deviation of the adjusted marks.. [4 marks]

QUESTION THREE

[20 MARKS]

- (a) Two events A and B are such that $P(A) = \frac{8}{15}$, $P(B) = \frac{1}{3}$ and $P(A | B) = \frac{1}{5}$. calculate the probability that :
- Both events occur [2 marks]
 - Only one of the events occur [3 marks]
 - None of the events occur [2 marks]
- (b) The table below shows the frequency distribution of the masses of 52 students in a college. The measurements were recorded to the nearest kg.

<i>Marks</i>	40-44	45-49	50-54	55-59	60-64	65-69	70-74
<i>Numberof students</i>	3	2	7	18	18	3	1

Estimate

- Median [2 marks]
- The value of x if 40% of the students were heavier than x [3 marks]
- The mean absolute deviation [4 marks]
- The standard deviation of the distribution [4 marks]

QUESTION FOUR

[20 MARKS]

- (a) The marks obtained by eight students in Maths and Programming are as shown below.

<i>Maths</i>	67	24	85	51	39	97	81	70
<i>Programming</i>	70	59	71	38	55	62	80	76

Calculate to 4d.p, the Spearmans Rank correlation coefficient and comment on your results.(6 marks)

- (b) Customers who purchase a certain make of car can order an engine in any of three sizes. Of all cars sold, 45% have the smallest engine, 35% have the medium-sized one, and 20% have the largest. Of cars with the smallest engine, 10% fail an emissions test within two years of purchase, while 12% of those with the medium size and 15% of those with the largest engine fail.
- What is the probability that a randomly chosen car will fail an emissions test within two years? . [3 marks]
 - A record for a failed emissions test is chosen at random. What is the probability that it is for a car with a small engine? [2 marks]
- (c) The following data shows two groups of casual workers, their number and average wages they are paid.

	Group A	Group B	Combined Groups
Number of workers	X	100	250
Mean Age	600	Y	6800
Standard Deviation	8	9	Z

- Calculate the missing entries X,Y and Z. [5 marks]
- Calculate the coefficient of variation for groups A and B [3 marks]
- From (ii) above, which has a greater variability in income? [1 marks]

QUESTION FIVE

[20 MARKS]

- (a) The heights and corresponding weights of a group of 9 randomly selected students were measured and the following results obtained.

<i>Height(m)</i> X	1.6	1.7	1.6	1.4	1.6	1.7	1.4	1.3	1.2
<i>Weight(kg)</i> Y	70	75	65	55	60	76	55	50	49

Obtain

- The spearman's rank correlation coefficient and interpret it. [5 marks]
- The product moment correlation coefficient and comment on its value. [4 marks]
- The coefficient of determination and interpret it [3 marks]

- (b) A computer program generates random questions in arithmetic that have to be answered within a fixed time. The probability of answering the first question correctly is 0.8. Whenever a question is answered correctly, the next question generated is more difficult and the probability of a correct answer being given is reduced by 0.1. Whenever a question is answered wrongly, the next question is of the same standard and the probability of answering it correctly is not changed.
- i. Draw a tree diagram to show this information for the first two generated questions [2 marks]
 - ii. Find the probability that the second question was answered wrongly [2 marks]
 - iii. By extending the tree diagram find the probability that the second question is answered correctly given that the third question is answered correctly. [4 marks]

END OF EXAM