



(Knowledge for Development)

KIBABII UNIVERSITY
UNIVERSITY EXAMINATIONS
2021/2022 ACADEMIC YEAR
FIRST YEAR SECOND SEMESTER
SUPPLEMENTARY EXAMINATION
FOR THE DEGREE OF BACHELOR OF INFORMATION
TECHNOLOGY

COURSE CODE: STA 114

COURSE TITLE: PROBABILITY AND STATISTICS 1

DATE: 29/07/2022

TIME: 2:00 PM - 4:00 PM

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

QUESTION ONE (30 MARKS)

- a) Explain briefly:
- The difference between qualitative and quantitative data (4mrks)
 - The difference between Independent and Dependent variables (4mrks)
 - The difference between Discrete and continuous variable (2mrks)
 - Six sources of data (6mrks)

b.) Students were asked how long it took them to travel to college on a particular morning. A cumulative frequency distribution was formed.

Time taken (minutes)	Cumulative frequency
< 5	28
< 10	45
< 15	81
< 20	143
< 25	280
< 30	349
< 35	374
< 40	395
< 45	400

- Construct a Frequency Distribution Table (4marks)
- Compute the following: (3mrks)
 - The Mean (4mrks)
 - Standard deviation (3mrks)
 - The median

QUESTION TWO (20 MARKS)

- Identify three measures of central tendency and list major characteristics of each measure of central tendency. (12marks)
- The marks of 500 candidates in an examination are normally distributed with mean of 45 marks and standard deviation of 20 marks.
 - Given that the pass mark is 43, estimate the number of candidates who passed the examination (3marks)
 - If 5% of the candidates obtained a distinction by scoring x marks or more, estimate the value of x (3marks)
 - Estimate the interquartile range of the distribution (2marks)

QUESTION THREE (20 MARKS)

- a. Thirty random observations are taken from each of the following distributions and samples mean calculated. Find, in each case the probability that the sample mean exceeds 5.
- i. X is the number of telephone call made in an evening to a counseling service, where $X \sim Po(4.5)$ (4marks)
 - ii. X is the number of heads obtained when an unbiased coin is tossed nine times
i. (3marks)
 - iii. X is distributed uniformly throughout the range $2 \leq X \leq 7$ (3marks)
- b. Discuss the role of statistics as used in different fields. (10marks)

QUESTION FOUR (20 MARKS)

- a.) Define regression analysis and state its assumptions (4marks)
- b.) Outline the importance of a scatter plot (3marks)
- c.) An old film is treated with a chemical in order to improve the contrast. Preliminary tests – on nine samples drawn from a segment of the film produced the produced results.

Sample	A	B	C	D	E	F	G	H	I
x	1.0	1.5	2.0	2.5	3.0	3.5	4.0	4.5	5.0
y	49	60	66	62	72	64	89	90	96

The quantity x is a measure of the amount of chemical applied and Y is the contrast index, which takes values between 0 no contrast and 100 (maximum contrast)

- i. Plot a scatter plot diagram to illustrate the data (5marks)
- ii. It is subsequently discovered that one of film was damaged and produced on incorrect result. State which sample you think this was (1mark)
- iii. Ignoring the incorrect sample, calculate the product moment correlation coefficient
 $\sum x^2 = 83.75, \sum y = 584, \sum x = 23.5, \sum y^2 = 44622, \sum xy = 1883, n = 8$ (3marks)
- iv. The line of regression of y on x has equation $y = a + bx$. Calculate the values of a and b , each correct to 3 significant figures (4marks)

QUESTION FIVE (20 MARKS)

The height of X cm, of each man in a random sample of 200 men living in Mombasa was measured. The following results were obtained: $\sum X = 35,050$, $\sum x^2 = 6163109$

- a) Calculate unbiased estimates of the mean and variance of the heights of men living in Mombasa (12marks)
- b) Determine an appropriate 90% confidence interval for the mean height of men living in Mombasa and name the theorem that you have assumed (8marks)