ALUPE UNIVERSITY COLLEGE
...Bastion of Knowledge...
P. O.Box 845-50400 Busia(K)
principal@auc.ac.ke
Tel: +254 741 217 185
+254 736 044 469
off Busia-Malaba road

OFFICE OF THE DEPUTY PRINCIPAL
ACADEMICS, STUDENT AFFAIRS AND RESEARCH

UNIVERSITY EXAMINATIONS

2021/2022 ACADEMIC YEAR

SECOND YEAR SECOND SEMESTER REGULAR EXAMINATION

FOR THE DEGREE OF BACHELOR OF SCIENCE (COMPUTER SCIENCE)

COURSE CODE: **STA 205**

COURSE TITLE: **PROBABILITY AND STATISTICS**

DATE: 9TH JUNE, 2022 **TIME: 1400 – 1700 HRS**

INSTRUCTION TO CANDIDATES

- **SEE INSIDE**

THIS PAPER CONSISTS OF 4 PRINTED PAGES **PLEASE TURN OVER**

REGULAR – MAIN EXAM

STA 205: PROBABILITY AND STATISTICS

STREAM: CS **DURATION: 3 Hours**

INSTRUCTION TO CANDIDATES

Answer **ALL** questions from section A and **ANY THREE** Questions in section B.
All questions in section B carry Equal Marks
No sharing of scientific calculators.
Do not write on this question paper

==========

SECTION A (31 MARKS): ANSWER ALL QUESTIONS

QUESTION ONE (16 MARKS)

a) Define the following terms as used in tests of hypothesis
   i) Type I error (1 mark)
   ii) Degrees of freedom (1 mark)

b) The contents of three urns 1, 2 and 3 are as follows

Balls	White	Black	Red
Urn I	1	2	3
Urn II	2	3	1
Urn III	3	1	2

An urn is chosen at random and from it two balls are drawn at random. The two balls are found to be one white and one black. What is the probability that they are from the second urn? (6 marks)

c) Suppose in a given experiment each sample air has a 10% chance of containing particular rare molecules. Assume the samples are independent with regard to the presence of the rare molecule. Find the probability that in the next 18 samples; exactly 2 contain the rare molecules. (3 marks)

d) A sample of 900 items is taken from a population with S.D.15. The mean of the sample is 25. Test whether the sample has come from a population with mean 26.8. (5 marks)

QUESTION TWO (15 MARKS)

a) Given below are the gains in speed in Mbps on two servers, X and Y

X	15	22	20	22	18	14	22			
Y	14	24	12	20	32	21	30	20	22	25

Test at 5% level, whether the two servers differ significantly with regard to increase in speed. (10 marks)

b) From the following data, compute Pearson's correlation coefficient (5 marks)

Price	10	12	14	15	19
Demand	40	41	48	60	50

SECTION B (39 MARKS): ANSWER ANY THREE QUESTIONS

QUESTION THREE (13 MARKS)

a) The mean height obtained from a random sample of 36 children is 30 inches. The standard deviation of the distribution of height of the population is known to be 1.5 inches. Test the statement that the mean height of the population is 33 inches at 5% level of significance. Also, set up 99% confidence limits of the mean height of the population. (8 marks)

b) An aptitude test was conducted for selecting computer scientists in an organization from 1000 students. The average score is 42 and the standard deviation is 24. Assuming normal distribution for the scores, find:

i) The number of candidates whose scores exceeded 58 (2 marks)

ii) The number of students whose scores lied between 30 and 66 (3 marks)

QUESTION FOUR (13 MARKS)

In a capability study of a lathe used in turning a shaft to a diameter of 23.75±0.1 mm a sample of 6 consecutive pieces was taken each day for 8 days. The diameters of these shafts are as given below:

Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8
23.77	23.8	23.77	23.79	23.75	23.78	23.76	23.76
23.80	23.78	23.78	23.76	23.78	23.76	23.78	23.79
23.78	23.76	23.77	23.79	23.78	23.73	23.75	23.77
23.73	23.70	23.77	23.74	23.77	23.76	23.76	23.72
23.76	23.81	23.80	23.82	23.76	23.74	23.81	23.78
23.75	23.77	23.74	23.76	23.79	23.78	23.8	23.78

Construct the $\bar{X}$ and R chart and find out the process capability for the machine. (13 marks)

QUESTION FIVE (13 MARKS)

A public opinion poll surveyed a simple random sample of voters to establish whether there is a relationship between gender and voting preference. Respondents were classified by gender (male or female) and by voting preference (Republican, Democrat, or Independent). Results are shown below.

Voting Preferences			
Gender	**Republican**	**Democrat**	**Independent**
Male	200	150	50
Female	250	300	50

i) Giving reasons, which is the best statistical test for analyzing the data? (2 marks)
ii) State and explain any three assumptions of the test statistic identified in a above (3 marks)
iii) State the null hypothesis of the study. (1 mark)
iv) At 0.05 significance level, does the data support the null hypothesis? (6 marks)
v) Report your findings. (1 mark)

QUESTION SIX (13 MARKS)

a) Consider families with 4 children each. What percentage of families would you expect to have two boys and two girls? (3 marks)

b) The following data relate to the yield of 4 varieties of rice each shown on 5 plots. Find whether there is significant difference between the mean yield of these varieties. (10 marks)

Plot Name/Treatment	1	2	3	4
P	99	103	109	104
Q	101	102	103	100
R	103	100	107	103
S	99	105	97	107
T	98	95	99	106

QUESTION SEVEN (13 MARKS)

a) It is known from past experience that in a certain plant, there are on an average four industrial accidents per year. Find the probability that in a given year there will be less than four accidents. Assume Poisson distribution. (4 marks)

b) From the following bivariate data, you are required to fit the regression line Y on X and predict Y if x = 20. (9 marks)

X	4	12	8	6	4	4	16	8
Y	14	4	2	2	4	6	4	12