(15)

(Knowledge for Development)

KIBABII UNIVERSITY

UNIVERSITY EXAMINATIONS

2019/2020 ACADEMIC YEAR

SECOND YEAR SECOND SEMESTER

SPECIAL/SUPPLEMENTARY EXAMINATION

FOR THE DEGREE OF BACHELOR OF SCIENCE

COURSE CODE: STA 205

COURSE TITLE: INTRODUCTION TO STATISTICS AND PROBABILITY

DATE: 04/02/2021 **TIME:** 11 AM -1 PM

INSTRUCTIONS TO CANDIDATES

Answer Question One and Any other TWO Questions

TIME: 2 Hours

This Paper Consists of 5 Printed Pages. Please Turn Over.

QUESTION ONE (30 MARKS)

a) Define what is meant by

i) Critical region

ii) Hypothesis testing

iii) Estimation

iv) Power of the test (4 marks)

b) A machine produced 20 defective articles in the batch of 400. After overhauling it produced 10 defectives in a batch of 300. Has the machine improved? (4 marks)

c) Two independent samples of 7 and 6 gave the following values;

Sample A	9	11	13	15	9	12	15
Sample B	11	12	10	9	8	10	-

Examine whether the difference between the means of the two samples is significant at 5% level of significance. (8 marks)

d) In a hospital 480 female and 520 male babies we\re born in a week. Do these figures confirm the hypothesis that males and females are born in equal number at 5% significance level (4 marks)

e) Differentiate between axiomatic approach to probability and classical approach to probability (2marks)

f) In an investigation to estimate the mean weight in Kg of 15 year old children in a particular region, a random sample of 100 children is selected. Previous study indicate that the variance of weights of such children is 3.0 Kg. Suppose the sample mean weight is 38.4 Kg, estimate the population mean weight of all 15 years old children in the region assuming that these weights are normally distributed. Use $\alpha = 5\%$. (5 marks)

g) The random variable X is distributed B (7, 0.2). Find correct to three decimal places

i. P(X=3)

ii. $P(1 < X \leq 4)$

iii. $P(X > 1)$ (3marks)

QUESTION TWO (20 MARKS)

a) Define the following terms as used in statistical inference

i. Confidence interval

ii. Estimator

iii. Parameter

iv. Statistic

v. Random variable (5marks)

b) In a one hour period, the number of gallons of a certain toxic chemical that is produced at a local plant say Y has the following pmf

Y	0	1	2	3
$P_y(Y)$	0.2	0.3	0.3	0.2

(i) Compute the expected number of gallons produced during a one-hour period (2 marks).

(ii) The cost (in hundreds of dollars) to produce Y gallons is given by the cost function $C(Y) = 3 + 12Y + 2Y^2$. What is the expected cost in a one hour period? (6 marks).

c) A random sample of 11 bags were selected from a machine packaging wheat flour in bags marked 1 kg. The actual weight of each flour in kgs were 1.017, 1.051, 1.078, 0.997, 1.033, 0.996, 1.059, 1.082, 1.014, 1.072 and 0.998. . Construct a 95% C.I for the mean weight of flour in bags marked 1 kg assuming the weights are normally distributed. (7 marks)

QUESTION THREE (20 MARKS)

a) State three assumptions made in the determination of F-test. (3 Marks)

b) A manufacturer of car batteries guarantees that his batteries will last, on the average of 3 years with a standard deviation of 1 year. If 5 of these batteries have lifetimes of 1.9, 2.4, 3.0, 3.5 and 4.2 years. Is the manufacturer still convinced that his batteries have a standard deviation of 1 year at $\alpha = 0.05$? (7 marks)

c) The mean score on a widely given freshman mathematics examination is 75. A mathematics teacher at a very large university wants to determine whether there is statistical evidence for

claiming that this year's class is not average. Test for this at 5% level of significance using the following scores. (10 marks)

94	69	89	49	88	89	85
95	55	93	86	62	83	96
48	51	69	74	83	71	89
58	89	81	79	52	73	
75	91	68	100	63	81	

QUESTION FOUR (20 MARKS)

a) State the reasons for the increase in the use of non-parametric tests in research. (3 marks)

b) Two random sample were drawn from two normal populations and their values were

A	66	67	75	76	82	84	88	90	92		
B	64	66	74	78	82	85	87	92	93	95	97

At 5% level of significance test whether the two populations have the same variance (10 marks)

c) A batch of parts contains 100 parts from a local supplier of tubing and 200 parts from a supplier of tubing in the next state. If four parts are selected randomly and without replacement. What is the probability that;

(i) Two or more parts in the sample are from the local suppliers (4 marks)

(ii) Atleast one part in the sample is from the local supplier (3 marks)

QUESTION FIVE (20 MARKS)

a) A machine produced 20 defective articles in a batch of 400. After over hauling it produced 10 defective in a batch of 300. Has the machine improved? (7 marks)

b) The table below gives observed frequencies in the nine different length-of-stay and type-of-insurance categories into which the sample has been divided. Prof. Simwa wishes to test the hypothesis at $\alpha = 0.01$

H_O: length of stay and type of insurance are independent

H_1: length of stay depends on type of insurance

Fraction of cost covered by insurance	Days in hospital			Total
	< 5	5-10	>10	
< 25%	40	75	65	180
25-50%	30	45	75	150
>50%	40	100	190	330
Total	110	220	330	660

Find whether to reject or accept the null hypothesis (13 marks)