MACHAKOS UNIVERSITY

University Examinations 2019/2020 Academic Year

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

FOURTH YEAR SPECIAL /SUPPLEMENTARY EXAMINATION FOR

BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE

BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING

BACHELOR OF SCIENCE IN MATHEMATICS

SMA 467: TEST OF HYPOTHESIS II

DATE: 21/1/2021 **TIME: 2.00-4.00 PM**

INSTRUCTIONS

1. *Answer **Question 1** and any other **two** questions.*
2. *Out of the **three** questions answered, each question must start on a new page.*
3. *You need the following items for this paper**:***
   - *Scientific Calculator.*
   - *Statistical Tables.*

QUESTION ONE (30 MARKS)

a) (i) Define the term *test of hypothesis* as used in Statistics. (2 marks)

(ii) Differentiate between a *one-tailed test* and a *two-tailed test* as used in statistical test of hypothesis. (4 marks)

b) The following is a set of marks scored in Mathematics in an examination by a class comprising of 24 students from a certain high school.

50	44	25	28	32	58	42	81	66	20	18	35
28	68	46	72	24	37	32	36	48	45	86	64

Test the set of marks for randomness at the 5% level of significance. Note: median = 43. (6 marks)

c) The following data set represents the marks scored by a random sample of 14 male students and 12 female students who sat for end of semester paper in general statistics.

Male	58	68	32	64	54	44	56	62	66	60	52	55	75	61
Female	45	72	48	55	65	35	70	55	53	74	73	67		

Using the Mann-Whitney U test, assess whether there is a statistically significant difference in the academic performance between the male and female students. Take level of significance $\alpha = 5\%$. (8 marks)

d) A educational researcher carried out a study on the relationship between the student academic performance in Statistics and the degree course programme taken by the students. A random sample of 500 students was selected, and the data pertaining to the academic performance and the course programme summarised in the contingency table below.

	Mathematics	Economics	Agriculture
Distinction	96	26	10
Credit	120	70	24
Pass	84	54	16

Test whether there is a relationship between the student academic performance and the degree programme taken by the students. Use the 5% and the 1% levels of significance. (10 marks)

QUESTION TWO (20 MARKS)

a) State *two* advantages and *two* disadvantages of non-parametric tests of hypothesis. (4 marks)

b) Explain any two circumstances in which the Spearman's rank correlation co-efficient is suitable in the analysis of statistical data. (4 marks)

c) A study was carried out to determine whether there is a difference in salaries between men and women among married couples. A random sample of 15 couples was selected in a certain town and the monthly salaries in thousand Kenya shillings were as shown in the table below.

Couple	A	B	C	D	E	F	G	H	I	K	L	M	N	P	Q
Male	72	42	66	64	25	18	70	65	82	28	45	34	52	41	75
Female	36	40	35	48	34	24	32	22	37	30	52	24	44	43	40

Using the Wilcoxon signed rank test for paired data, evaluate whether the two sets of salaries are the same. Use the 5% level of significance. (12 marks)

QUESTION THREE (20 MARKS)

An engineering company has developed two types of devices for estimating long distances between two points on the earth surface. One device uses terrestrial communication while the other uses satellite communication. In evaluating them, 12 random distances were taken between various points and the actual distance and also the distance estimated by each of the two devices recorded in kilometres as shown in the table below:

Distance	A	B	C	D	E	F	G	H	J	K	L	M
Actual	15	19	18	30	62	30	78	23	68	55	30	17
Terrestrial	24	26	48	64	40	54	45	48	72	32	48	25
Satellite	20	25	28	23	48	64	82	28	90	65	45	10

a) (i) Compute the Spearman's rank correlation co-efficient between the actual distance and the estimated distance by terrestrial device, and between the actual distance and the estimated distance by satellite based device. (10 marks)

(ii) Interpret each of the co-efficients of correlation obtained in (i) above. (3 marks)

b) Recommend the most suitable device for estimating distance, justifying your choice based on the results obtained in (a) above. (2 marks)

c) Test for significance each of the two co-efficients of correlation obtained in (a)(i) above. (5 marks)

QUESTION FOUR (20 MARKS)

a) An educational researcher claims that the pass rates for subjects in KCSE vary with the proportion of students who pass in some of the subjects is as shown below:

English	60%	History	70%
Mathematics	35%	Physics	20%
Kiswahili	50%	Biology	60%
Geography	75%		

A random sample of **500** KCSE candidates in a certain year was taken for students who sat for all the seven listed subjects, and the number of candidates who passed in the seven subjects is as shown below:

English	240	History	420
Mathematics	105	Physics	125
Kiswahili	210	Biology	330
Geography	420		

Using chi-square test for goodness of fit, evaluate whether this claim about KCSE pass rates is true or not at the 5% level of significance. (8 marks)

b) A manufacturer claims that only 1% of the items from the production line are faulty or defective. The items are produced in batches of 300 items per batch.

A quality control official took a random sample of 250 batches and tested all the items in each batch and found out the following on the number of defectives items per batch.

Defective items (x)	0	1	2	3	4	5	6	7	8	9
No of batches (f)	10	28	44	48	45	35	24	10	4	2

Using the chi-square test for goodness of fit, evaluate whether this claim by the manufacturer is true or not at the 5% level of significance. (12 marks)

QUESTION FOUR (20 MARKS)

a) In a certain mango farm, there are different varieties of mangoes planted randomly and grown under similar conditions. The number of fruits in thousands from three random samples of mango trees in the farm is as shown in the table below.

Sample 1	3.8	3.0	4.2	2.5	2.6	3.2
Sample 2	2.9	2.7	3.3	4.0	2.4	
Sample 3	2.8	3.4	3.6	3.7	2.2	3.5

Using the Kruskal-Wallis test, evaluate whether the three samples of mangoes are drawn from the same variety or not at the 5% level of significance. (8 marks)

b) Given a set of n bi-variate observations comprising pairs of variables (x_i, y_i), derive the formula for the Spearman's rank co-efficient of correlation given by:

$$r_{xy} = 1 - \frac{6\sum(R_x - R_y)^2}{n(n^2-1)}$$ (12 marks)