MACHAKOS UNIVERSITY

University Examinations 2016/2017

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

FOURTH YEAR FIRST SEMESTER 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: 27/7/2017 **TIME: 11:00 – 1:00 PM**

Instructions to the Candidate:

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 a Scientific Calculator and Statistical Tables for this paper.*

1. (a) (i) Explain the term *non-parametric* test of hypothesis. (2 *marks*)

(ii) Non-parametric tests are said to be less powerful compared to parametric tests. Justify this statement. (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) The following is a set of marks scored in Mathematics in an examination by a class comprising of 32 students from a certain high school.

50	44	25	28	69	58	42	81	16	20	08	84	28	26	46	66
35	24	32	36	60	48	45	86	30	17	34	34	85	40	54	20

Test the set of marks for randomness at the 5% level of significance. Note: median = 38. (10 *marks*)

(c) The following data set represents the yields of maize in thousand bags per farm block recorded in two agricultural zones – Zone A and Zone B.

Zone A	15.2, 8.9, 9.3, 14.4, 15.6, 11.8, 16.3, 17.8
Zone B	11.5, 12.6, 19.4, 21.3, 32.5, 18.6, 17.0, 23.4, 29.6

Using the Mann-Whitney test, assess whether there is a statistically significant difference in the mean yields between the two agricultural zones. Take level of significance $\alpha = 5\%$ and test statistic $U_{0.05} = 18$. (10 *marks*)

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

(b) The marks scored in two papers in an examination by a random sample of 12 students from a certain high school is as shown in the table below.

Student	A	B	C	D	E	F	G	H	J	K	L	M
Paper 1	12	08	34	48	49	43	54	18	32	15	36	52
Paper 2	43	25	40	50	10	24	30	33	53	42	09	20

Using the Wilcoxon signed rank test for paired data, test whether the two sets of marks are drawn from the same subject (population) at the 5% level of significance. (16 *marks*)

3. A researcher on family finance carried out a study on the relationships between house rent and salaries of both men and women among employed couples. A random sample of 10 couples was taken and the monthly rent paid for the house and gross salaries for the man and his wife was recorded with all figures in thousand Kenya shillings. The data is as shown in the table below:

Couple	A	B	C	D	E	F	G	H	J	K
Rent	26	18	30	25	16	18	45	24	18	12
Male	180	230	250	280	150	120	240	150	140	110
Female	180	260	160	250	110	200	160	100	160	120

(i) Compute the Spearman's rank correlation co-efficient between the monthly rent for the house and the salary of male employees, and between the monthly rent for the house and the salary of female employees. (12 *marks*)

(ii) What conclusions can you draw from the results obtained (i) above. (2 *marks*)

(iii) Following public outcry on high house rent, the government has decided to adjust salaries upwards but funds are not enough for all employees. As a statistician, what would you recommend to the government based on the findings in (ii)? (2 *marks*)

(iv) Test for significance each of the two correlation co-efficients computed in (i) above.(4 *marks*)

4. (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.

Tree	A	B	C	D
Sample 1	19	20	60	130
Sample 2	21	61	80	129
Sample 3	40	99	100	149

Using the Kruskal-Wallis test, test 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 two dimensional 2 X 2 contingency table for a chi-square test for independence with observed table cell frequencies *a*, *b*, *c*, and *d*, show that the chi-square test statistic χ^2 is given by:

$$\chi^2 = \sum \frac{N\,(ad - bc)^2}{(a+b)(a+c)(b+d)(c+d)}$$ (12 *marks*)

5. (a) A university has a total student population of 5000 distributed in six schools as follows: Engineering 400, Business 1800, Humanities 1200, Sciences 800, Agriculture 400 and Hospitality 300. Fee bursary is to be given to a total of 120 students and this should be fairly proportional to the number of students in the school. The number of student who were awarded bursaries in each school were as follows: Engineering 16, Business 24, Humanities 25, Sciences 22, Agriculture 18 and Hospitality 15. Using chi-square test for goodness of fit, evaluate whether the bursaries were equitably distributed among the schools. Test this at the 5% level of significance (8 *marks*)

(d) A road safety researcher carried out a study on road accidents and impact on victims with an objective of finding out the relationship between the class of a road and the level of injury of the victims on Kenyan roads. A random sample of 500 accident victims was collected from all the hierarchical classes of roads and the data summarized in the contingency table below.

		Level of Injury		
		Light	Serious	Dead
Road	Class A	120	70	24
	Class B	84	54	16
	Class C	96	26	10

It is required to test whether there is a relationship between the level of injury of the victim and the classification of the road. Test this at the 5% level and 1% level. (12 *marks*)