MACHAKOS UNIVERSITY

University Examinations 2017/2018

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS AND STATISTICS

FOURTH YEAR SECOND SEMESTER EXAMINATION FOR

BACHELOR OF EDUCATION (SCIENCE)

BACHELOR OF EDUCATION (ARTS)

SMA 467: TEST OF HYPOTHESIS II

DATE: 6/12/2017 **TIME: 8:30 – 10:30AM**

***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 30 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	17	34	84	85	50	34

Test the set of marks for randomness at the 5% level of significance. Note: median = 43. (10 *marks*)

(d) 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 $\alpha = 5\%$.

(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	16	34	48	49	43	54	18	32	15	36	52
Paper 2	43	25	40	50	10	24	30	33	53	42	28	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 chain of hotels analysed data on their profits with an objective of forecasting the profit for each month of a given year. They used two time series forecasting techniques – the moving average and the least squares. Monthly profit in million Kenya shillings was recorded for actual figures and also for the two forecasting techniques as shown in the table below:

Month	Jan	Feb	Mar	Apr	May	June	July	Aug	Sept	Oct	Nov	Dec
Actual	15	19	30	18	62	30	78	68	30	55	23	17
Moving Average	24	26	64	48	40	54	45	72	48	32	48	25
Least Squares	20	25	23	28	48	64	82	90	45	65	28	10

(i) Compute the Spearman's rank correlation co-efficient between the actual monthly profit for the hotel and the moving average forecasts, and between the actual profit for the hotel and the least squares forecasts. (12 *marks*)

(ii) Recommend the most suitable forecasting technique for the hotel, justifying your choice based on the results obtained in (i) above. (3 *marks*)

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

4. (a) A public university has a total student population of 5000 distributed in six schools as follows: Engineering 400, Business 1800, Humanities 1200, Sciences 800, Agriculture 500 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 students who were awarded bursaries per 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 at the 5% level of significance. (8 *marks*)

(b) A crime researcher carried out a study on the relationship between the type of crime committed and the age group of the convict. A random sample of 400 convicted criminals was collected, and the data pertaining to age and type of crime summarized in the contingency table below.

		Youth	Middle Age	Old Age
Crime	Robbery	78	96	41
	Fraud	30	54	21
	Murder	32	25	23

Test whether there is a relationship between the type of crime and the age group. Use the 5% level of significance. (12 *marks*)

5. (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	2.9	3.0	2.5	2.6	3.2
Sample 2	3.8	2.7	4.0	2.4	
Sample 3	2.8	3.4	3.7	2.2	2.0

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*)