MACHAKOS UNIVERSITY

University Examinations 2018/2019

SCHOOL OF PURE AND APPLIED SCIENCES

DEPARTMENT OF MATHEMATICS, STATISTICS AND ACTUARIAL SCIENCE

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

BACHELOR OF EDUCATION

SMA 467: TEST OF HYPOTHESIS II

DATE: 2/5/2019 **TIME: 2:00 – 4:00 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) Differentiate between a *right tailed test* and a *left tailed test* as used in statistical test of hypothesis. (4 *marks*)

b) The following data set represents the weight in kilograms of a group of patients who visited a certain clinic in a general hospital.

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 weights for randomness at the 5% level of significance. Note: median = 43. (8 *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 a mathematics test in Masaku High School.

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, evaluate 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) An educational researcher carried out a study on the relationship between the student academic performance in KCSE Mathematics and the school environment in which the student did KCPE. A random sample of 500 students was selected, and the data pertaining to the academic performance and the KCPE school environment summarised in the contingency table below.

	Urban	Rural
Distinction	92	37
Credit	164	46
Pass	61	28
Fail	48	24

Evaluate whether there is a relationship between the student academic performance and the KCPE school environment. Use the 5% and the 1% levels of significance. (10 *marks*)

2. 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 coefficient is a suitable technique in describing data in statistical analysis. (4 *marks*)

c) A study was carried out to determine whether there is a difference in academic performance between Paper 1 and Paper 2 among the subjects tested in KCSE exams. A random sample of 12 candidates was selected in a certain subject and the marks scored were as shown in the table below.

Student	A	B	C	D	E	F	G	H	J	K	L	M
Paper 1	72	42	66	64	25	18	70	65	82	28	41	75
Paper 2	36	40	35	48	34	24	32	22	37	30	43	40

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

3. In national examination marking centre, a coordination session was conducted in which a random sample of papers from different schools is selected and a copy of each paper is marked by all the ordinary examiners and the chief examiner. In order to evaluate the examiners, the mark awarded by an ordinary examiner is compared by the mark awarded by the chief examiner to check the extent of deviation. Two examiners John and Peter and the chief examiner each marked papers from a random sample of 12 candidates and the marks awarded by the three is as shown in the table below:

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

a) Compute the Spearman's rank correlation co-efficient between the marks awarded by the chief examiner and John, and between the marks awarded by the chief examiner and Peter. (10 *marks*)

b) Assuming that the marks awarded by the chief examiner are considered as the best:

i) Evaluate the *two* examiners based on the results obtained in (a) above. *(3 marks)*

ii) Which of the two examiners between John and Peter would you recommend based on the results obtained in (a) above. Justify your recommendation. (2 *marks*)

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

4. a) State any *three* possible sources of statistical hypothesis. (3 *marks*)

b) State *four* conditions which must be satisfied for the chi-square test for goodness of fit to be valid. (4 *marks*)

c) A researcher in personal finance formulated a statistical model to be used in estimating the number of loan applicants in various commercial banks in Kenya by relating the loan amount applied for and the number of loan applicants. He derived the linear regression model $y = 46.8986 - 2.523434\,x$, where y is the number of loan applicants and x is the loan amount in million Kenya shillings. In order to evaluate the model, data was collected from a sample of 15 loan amounts and their corresponding number of applicants as shown.

Loan amount	1	2	3	4	5	6	7	8	9	10	11	12
No of Applicants	62	54	45	38	28	26	24	21	18	16	14	10

Using the chi-square test for goodness of fit, evaluate the model on its suitability to estimate the number of loan applicants based on amount of loan. Test at the 5% and 1% levels of significance. (13 *marks*)

5. a) In a certain bank, three different types of customers were classified randomly under three samples. The deposit amounts in million Kenya shillings from three random samples of customers are as shown in the table below.

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

Using the Kruskal-Wallis test, evaluate whether the three samples of customers belong to the same income bracket or not at 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 and hence show that:

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