

MACHAKOS UNIVERSITY

DEPARTMENT OF MATHEMATICS AND STATISTICS

BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE

BACHELOR OF SCIENCE IN STATISTICS AND PROGRAMMING

BACHELOR OF SCIENCE IN MATHEMATICS

BACHELOR OF EDUCATION

BACHELOR OF ARTS

UNIVERSITY SUPPLEMENTARY EXAMINATIONS

SMA 467 : TEST OF HYPOTHESIS II

Instructions to the Candidate:

1. Answer **Question 1** and any other **two** questions.
2. 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 *test of hypothesis* as used in statistical inference. (3 marks)
- (ii) Non-parametric tests are said to less powerful compared to parametric tests. Account for this statement. (3 marks)

- (b) The win base record of a certain basketball team for their last 50 consecutive games was as follows:

W, W, W, W, W, W, L, W, W, W, W, W, W, L, W, L, W, W, W, L, L, W, W, W, W, L,
W, W, W, L, L, W, W, W, W, W, L, L, W, W, L, L, L, W, W, L, W, W, W, W

Using the run test, evaluate whether the sequence of wins and losses is random or not at the 5% level of significance. (6 marks)

- (c) A finance researcher in employee salary rates carried out a study to assess whether there is a difference in salary earned among employees in two banks in Kenya. A random sample of 15 employees from Bank A and 13 employees from Bank B was selected and the salary earned in thousand Kenya shillings was recorded as shown in the table below.

Bank A	43	23	42	64	38	52	30	72	28	40	58	12	45	65	60
Bank B	62	44	57	56	54	68	70	85	48	40	82	46	59		

Using the Mann-Whitney U test, evaluate whether there is a statistically significant difference in the salary earned between the two banks. Test at 5% level and 1% level of significance.(8 marks)

- (d) A crime researcher conducted a survey on the relationship between crime committed and the age. A random sample of 400 convicts was collected and the findings from the survey were as shown in the table below:

Age in years	Crime		
	Fraud	Robbery	Murder
16 - 25	60	48	57
26 - 40	40	58	42
41 - 60	20	46	29

Test whether there is a relationship between the type of crime committed and age among the convicted criminals. Test this at the 5% level and 1% level. (10 marks)

2. (a) State *two* advantages and *two* disadvantages of non-parametric tests of hypothesis. (4 marks)
- (b) (i) Explain any *two* circumstances in which the Spearman's rank correlation is a suitable technique in describing data in statistical analysis. (4 marks)
- (ii) Explain *two* circumstances in which the Kolmogorov-Smirnov test can be used as technique in statistical test of hypothesis. (4 marks)
- (c) A researcher in family finance carried out a study to assess whether there is a difference in salary earned among couples in which both partners are employed. A random sample of 12 couples was selected and the salary earned in thousand Kenya shillings was recorded as shown in the table below.

Couple	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
Male	44	23	45	74	38	54	66	92	28	48	58	12
Female	28	11	32	56	35	56	70	85	38	40	48	18

Using the Wilcoxon signed rank test for paired data, evaluate whether there is a significant difference between the male and female salary. Test at the 5% level of significance. (8 marks)

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

Distance	D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
Actual	16	23	25	30	62	25	78	22	68	25	32	20
Terrestrial	36	48	26	64	40	65	48	45	72	32	42	25
Satellite	20	25	30	28	48	30	82	22	70	24	30	18

- (i) Compute the Spearman's rank co-efficient correlation between the actual distance and the distance for the terrestrial device, and between the actual distance and the satellite based device. (10 marks)
- (ii) Recommend the most suitable device for estimating distance, justifying your choice based on the results obtained in (a) above. (2 marks)
- (iii) Test for significance each of the two co-efficients of correlation obtained in (a) above. Use the 1 % level of significance. (6 marks)
- (iv) Determine the 99% confidence interval for the population co-efficient of correlation. (2 marks)

4. (a) Differentiate between each of the following terms as used in statistical test of hypothesis:
- (i) *two tailed test* and *one tailed test*
 - (ii) *right tailed test* and *left tailed test*. (6 marks)
- (b) State *four* conditions which must be satisfied for the chi-square test for goodness of fit to be valid in testing a hypothesis in statistics. (4 marks)
- (c) A researcher in human resource developed a statistical model to be used in estimating the number of job applicants for various positions based on the amount of salary offered. He derived the simple linear regression model $y_i = 2.25 + 4.78x_i$, where y is the number of job applicants and x is the salary offered in thousand Kenya shillings. In order to evaluate the model, data was collected from a sample of 12 job positions with their salary amounts and the corresponding number of job applicants as shown in the table.

Salary	1	2	3	4	5	6	7	8	9	10	11	12
Applicants	3	5	9	12	20	26	35	42	52	56	65	75

Using the chi-square test for goodness of fit, evaluate the model on its suitability to estimate the number of job applicants based on the amount of salary offered. Test at the 5% level and 1% level of significance. (10 marks)

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

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

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:

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