JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF MATHEMATICS AND ACTUARIAL SCIENCE

UNIVERSITY EXAMINATION FOR DEGREE OF MASTER OF SCIENCE

IN APPLIED STATISTICS

1^{ST} YEAR 2^{nd} SEMESTER 2023/2024 ACADEMIC YEAR

MAIN CAMPUS

COURSE CODE: SAS 812

COURSE TITLE: NON PARAMETRIC METHODS

EXAM VENUE: **STREAM: (MSc. Applied Statistics)**

DATE: **EXAM SESSION:**

TIME: 3.00 HOURS

<u>Instructions:</u>

Answer ANY 3 questions

Candidates are advised not to write on the question paper.

Candidates must hand in their answer booklets to the invigilator while in the examination room.

QUESTION ONE (20 MARKS)

a) Let $X_1, X_2, \ldots, X_n$ be a random sample from a continuous distribution function $F(x)$. Let M be the median of the distribution. Give the Ordinary Sign-test for testing the null hypothesis $H_0 : M = M_0$ against the alternative $H_1 : M > M_0$, where M_0 is a specified number. Show that the test based on the Ordinary-sign test is consistent for H_1 [10 marks]

b) The median age of the onset of diabetes is thought to be 45 years. The ages at onset of diabetes for a random sample of 16 people with diabetes are:

26.2	30.5	35.5	38.0	39.8	40.3	45.0	45.6
45.9	46.8	48.9	51.4	52.4	55.6	60.9	65.4

Perform Wilcoxon signed rank test at 5% to determine if there is any evidence to conclude that the median age of the onset of diabetes differs significantly from 45 years. [10Marks]

QUESTION TWO (20 MARKS)

a) Recent studies on physicians who saw patients suggested that the median length of each patients visit was 22 minutes. It is believed that the median visit length is shorter than 22 minutes for Physicians with a big load of patients. A random sample of 20 visits to physicians assumed to have a big load of patients yielded visit times as follows:

9.4	13.4	15.6	16.2	16.4	16.8	18.1	18.7	18.9	19.1
19.3	20.1	20.4	21.6	21.9	23.4	23.5	24.8	24.9	26.8

Use the large sample approximation of the sign test to determine if there is sufficient evidence to conclude, at 1% level of significance, that the average visit length is shorter than 22 minutes? [10 Marks]

b) Let $X_1, X_2, \ldots, X_n$ be a random sample from unknown continuous distribution function $F(x)$. The aim is to test the null hypothesis is : $H_0 : F(x) = F_0(x)$ for all x against : $H_1 : F(x) \geq F_0(x)$ with strict inequality for some x where $F_0(x)$ is a completely specified continuous distribution function. Give a suitable size α Kolmogorov-Smirnov test for this problem and show that it is consistent. [10 Marks]

QUESTION THREE (20 MARKS)

a) Let $X_1, X_2, \ldots, X_m$ be a random sample from a continuous distribution function $F_1(x)$ and $Y_1, Y_2, \ldots, Y_n$ be a random sample from another continuous distribution function $F_2(y)$. To test for the null hypothesis $H_0: F_1(x) = F_2(y)$ for all x against a suitable alternative, the MannWhitney U-statistic can be used where

$U = \sum\sum D_{ij}$, $D_{ij} = \begin{cases} 1, if\ Y_j < X_i \\ 0, if\ Y_j > X_i \end{cases}$ Derive the null distribution of U. Hence for m=2, n=3 tabulate the values of U up to U=2 and give the rejection region for the Mann-Whitney U-test with nominal significance level 0.10. State the exact significance level. [10 Marks]

b) Five patients went to a diagnostic center to check their fasting blood sugar levels. After the first test they were not satisfied and so they decided to visit three other centers for confirmatory tests. The outcomes were recorded as follows

	Fasting blood sugar levels in 4 centres			
Patient	1	2	3	4
1	150	145	160	155
2	190	190	180	190
3	120	130	130	115
4	140	140	150	140
5	110	110	120	120

Apply Friedman test to check the similarities of the test centre results. [10 Marks]

QUESTION FOUR (20 MARKS)

a) Let Y_n denote the n^{th} order statistic of a random sample of size n from a distribution of the continuous type. Find the smallest value of n for which

$$(Y_1 < \xi_{0.5} < Y_n) \geq 0.95$$ [10 Marks]

b) A survey was carried out on salaries of top management officers of two companies and the amounts earned per month for some officers of the companies noted as follows in thousands of shillings.

Company A	65	58	55	50	59	30	65	90
Company B	76	66	48	49	62	45	55	88

Do the two companies pay their managers the similar salaries? Apply Mann – Whitney U-Test. [10 Marks]

QUESTION FIVE (20 MARKS)

a) Three treatments were replicated six time each in an experiment and the yields per treatment recorded in thousands of kilograms recorded as follows;

Treatment	Observation					
A	2.7	4.6	2.6	3.0	3.2	3.8
B	4.9	4.6	5.0	4.2	3.6	4.2
C	4.6	3.4	2.9	3.5	4.1	5.1

Use the Kruskal-Wallis procedure at 5% to test the differences between treatments.

[10 Marks]

b) Suppose $((X_1, Y_1)(X_2, Y_2), \ldots, (X_n, Y_n)$ is a random sample from a continuous bivariate distribution with the pdf f(x, y) . Let $D_i = Y_i - X_i$, $i = 1, 2, \ldots, n$. Assuming that D_1,$D_2, \ldots, D_n$ is a random sample from a continuous and symmetric distribution, Construct a 100(1-α)% confidence interval for the unknown median of differences based on the Wilcoxon signed-rank test statistic. [10 Marks]