STA3093 - Non-Parametric & Categorical Data Analysis
Tutorial 1

Submit the tutorial <u>on or before</u> 4.00 pm on October 18th 2024 to the department.

1)

The senior class in a particular high school had 48 boys. Twelve boys lived on farms and the other 36 lived in town. A test was devised to see if farm boys in general were more physically fit than town boys. Each boy in the class was given a physical fitness test in which a low score indicates poor physical condition. The scores of the farm boys (X_i) and the town boys (Y_j) are as follows.

<u>X_i: Farm Boys</u>

14.8	10.6	7.3	12.5	5.6	12.9	6.3	16.1	9.0	11.4	4.2	2.7

<u>Y_j: Town Boys</u>

12.7	16.9	7.6	2.4	6.2	9.9	14.2	7.9	11.3	6.4	6.1	10.6
12.6	16.0	8.3	9.1	15.3	14.8	2.1	10.6	6.7	6.7	10.6	5.0
17.7	5.6	3.6	18.6	1.8	2.6	11.8	5.6	1.0	3.2	5.9	4.0

Test the claim that the farm boys were more physically fit than town boys at a 5% significance level using a suitable nonparametric test procedure.

2)

Twelve volunteers were assigned to each of three weight-reducing plans. The assignment of the volunteers to the plans was at random, and it was assumed that the 36 volunteers in all would resemble a random sample of people who might try a weight-reducing program. Test the null hypothesis that there is no difference in the probability distributions of the amount of weight lost under the three programs against the alternative that there is a difference. The results are given as the number of pounds lost by each person.

Plan A		Plan B		Plan C	
2	17	17	5	29	5
12	4	15	6	3	25
5	25	3	19	25	32
4	6	19	4	28	24
26	21	5	9	11	36
8	6	14	7	7	20

3)

A survey was taken of all seven hospitals in a particular city to obtain the number of babies born over a 12-month period. This time period was divided into the four seasons to test the hypothesis that the birth rate is constant over all four seasons.

The results of the survey are as follows:

Number of Births				
Hospital	**Winter**	**Spring**	**Summer**	**Fall**
A	92	112	94	77
B	9	11	10	12
C	98	109	92	81
D	19	26	19	18
E	21	22	23	24
F	58	71	51	62
G	42	49	44	41

Analyze these data using the Friedman test.

4)

An Olympic diver is rated on ten practice dives, with the following measurements: 1.7, 5.3, 7.6, 8.9, 9.0, 9.1, 9.3, 9.6, 9.9, and 9.9. Test the hypothesis that the distribution function of her scores is given by $F(x)$, where

$$F(x) = 0 \qquad \text{if } x < 0$$

$$F(x) = x^2/100 \qquad \text{if } 0 \le x \le 10$$

$$F(x) = 1 \qquad \text{if } 10 < x$$

5)

In an industrial production line, items are inspected periodically for defectives. The following is a sequence of defective items, D, and non-defective items, N, produced by this production line:

D D N N N D N N D D N N N N

N D D D N N D N N N N D N D

Use the Wald-Wolfowitz runs test with a significance level of 0.05, to determine whether the defectives are occurring at random.

6)

Stitt, Hardy, and Nadel (1971) studied the relationship between the surface area and body weight of squirrel monkeys. The data in the following table represent the total surface areas (cm^3) and body weights (g) for nine squirrel monkeys. Treating body weight as the independent variable, fit a nonparametric linear regression model to the data.

Monkey	Body weight, g	Total Surface Area , cm^3
1	660	780.6
2	705	887.6
3	994	1122.8
4	1129	1125.2
5	1005	1070.4
6	923	1039.2
7	953	1040.0
8	1018	1133.4
9	1181	1148.0

Source: J. T. Stitt, J. D. Hardy, and E. R. Nadel (1971)