

MAS 403: NONPARAMETRIC METHODS

DATE: _____

Time Allowed: 2 hours

INSTRUCTIONS:

- Attempt **QUESTION ONE** and **ANY OTHER TWO**.
- Start each question on a fresh page.
- Indicate question numbers clearly at the top of each page.
- Observe further instructions on the answer booklet.

Question One (30 marks Compulsory)

- a) The time taken, in days, to diagnose leukemia in en animals infected with leukemia cells is as follows:

239 119 265 278 257 227 286 279 228 145

Is the median time to diagnose the disease in such a population likely to be 200 days? Use sign test with 5% level of significance. [4 marks]

- b) State 4 advantages of non parametric tests over the parametric counterparts. [4 marks]

c) The following data show the length of time in days that control mice, C, and inoculated mice, N, live.

C 20 23 28 30 31 32 44 33

N 20 29 23 48 41 32 36 43 42 46

Is there a difference in survivorship between control and inoculated mice? Use the Mann Whitney test at 5% level of significance. [6 marks]

d) Briefly describe when to use each of the following nonparametric methods:

i). Kruskal-Wallis test. [2 marks]

ii). Kolmogorov-Smirnov test. [2 marks]

e) Arrivals at a telephone both are assumed to be Poisson distributed. The following table gives a distribution of arrivals. Use the chi-square goodness of fit test to test the assumption at 5% level of significance. [6 marks]

ARRIVALS	FREQUENCY
0	14
1	31
2	47
3	41
4	29
5	21
6	10
7	5
8	2
TOTAL	200

f) The following is the distribution of hourly number of trucks arriving at a company's warehouse:

Trucks arrival per hour	0	1	2	3	4	5	6	7	8
Frequency	52	151	130	102	45	12	5	1	2

Find the mean of this distribution and use it (rounded to 1 decimal place) as the parameter λ , and test at 5% significance level if this data can be looked upon as values from a random variable having the Poisson distribution. [6 marks]

Question Two (20 Marks - Optional)

- a) Outline the Wilcoxon signed – rank test procedure stating the null hypothesis to be tested. [6 marks]
- b) Safarinet Corporation has installed speed governors on its salespersons' cars to regulate their speeds. The following table gives the number of contacts made by each of seven randomly selected sales representatives during the week before governors were installed and the number of contacts made during the week after installation.

Salesperson	A	B	C	D	E	F	G
Before	50	63	42	55	44	65	66
After	49	60	47	51	50	60	58

- i. Using the Wilcoxon signed-rank test at the 5% level of significance, can you conclude that the use of governors tends to reduce the number of contacts made per week by Safarinet Corporation's sales representatives? [8 marks]
- ii. Compare your conclusions of part i. with the result of the hypothesis test using the t distribution. [6 marks]

Question Three (20 marks - Optional)

- a) Define a run. [2 marks]
- b) A coin is tossed 30 times and the outcome noted (Head up =H, Tail up=T) as follows:

HHHTTHHTTTTTHHHHTHTHHHHHTTTHTTTTH

- i. State the number of runs in the sequence. [2 marks]
- ii. Use a runs table to determine if the coin is fair. [6 marks]
- c) Tourism is regarded as one of the major foreign exchange earners in our country. The following data was collected between the year 1970 and 1982. Determine whether the frequency of tourists to the country was random (had no trend) within the period. Test at 0.01 level of significance.

<i>Year</i>	<i>Number of tourists (millions)</i>
1970	12,362
1971	12,739
1972	13,057
1973	13,955
1974	14,123
1975	15,698
1976	17,523
1977	18,610
1978	19,842
1979	20,310
1980	22,500
1981	23,080
1982	21,916

[10 marks]

Question Four (20 marks - Optional)

- a) Outline the procedure for Kruskal - Wallis test. [5 marks]
- b) The cattle disease, East Coast Fever, is transmitted by the tick *D. marginatus*. Data were collected to see if the time spent feeding by the female ticks was affected by photoperiodic conditions:

	Simulated hours of day light per day			
	6 hours	12 hours	18 hours	24 hours
Time in	7.1	8.6	12.1	9.1
hours spent	14.3	11.0	13.9	14.5
feeding by	13.4	9.0	14.1	11.5
individual	10.7	12.6	8.6	12.7
ticks		14.8	13.2	11.7
				11.2

Conduct the Kruskal Wallis test to determine whether photoperiodic conditions affected the feeding time of the ticks. Use $\alpha = 0.05$. [15 marks]

Question Five (20 marks - Optional)

- a) Consider a random sample $(x_1, y_1), \dots, (x_n, y_n)$. Suppose there are no ties. Show that the rank correlation coefficient is given by

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)}$$

where d_i is the difference between ranks assigned to x_i and y_i . use the standard results:

$$\sum_{i=1}^n i = \frac{n(n+1)}{2} \quad \text{and} \quad \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6} \quad [7 \text{ marks}]$$