BINDURA UNIVERSITY OF SCIENCE EDUCATION   SFM224

BACHELOR OF STATISTICS AND FINANCIAL MATHEMATICS

APPLIED STATISTICS

Time : 3 hours

~ MAR 2024

Candidates may attempt ALL questions in Section A and at most TWO questions in Section B. Each question should start on a fresh page.

SECTION A (40 marks)

Candidates may attempt ALL questions being careful to number them A1 to A4.

A1. State and explain any four measurement scales in non parametric statistics   [8]

A2. Distinguish between the following terms

    (a) Significance level and critical level   [3]

    (b) Parametric and non parametric statistics   [3]

    (c) Skewness and Kurtosis   [3]

    (d) Descriptive statistics and Inferential statistics   [3]

A3. The table below shows 50 values generated from a process. We would like to know whether we may conclude that the pattern of departures above and below the average is the result of a non random process   [10]

55	48	57	47	42
52	57	48	47	55
51	55	46	54	49
52	42	42	60	59
52	55	49	51	42
52	55	42	54	44
58	47	49	52	50
51	47	45	46	54
46	49	49	56	59
46	39	49	52	51

A4. Consider the samples from two normal populations

Sample 1	8.2	5.3	6.5	5.1	9.7	10.8		
Sample 2	9.5	8.3	7.5	10.9	11.3	9.3	8.8	8.0

Test at 10% significance level whether the two populations variances are the same. [8]

SECTION B (60 marks)

Candidates may attempt TWO questions being careful to number them B5 to B7.

B5. (a) The following data shows the age at diagnosis of type II diabetes in young adults

Male	19	22	16	29	24	20	11	23
Female	20	11	17	12	13	14	19	25

Use the Mann Whitney U test to determine if age at diagnosis is different for males and females [13]

(b) Outline the main steps of hypotheses testing. [6]

(c) Two quantities of wool of standard lengths are stretched with a given force before and after washing. The Lengths of each quantity of wool before (X) and after (Y) are given in the following table.

X	1.2	0.5	2.1	1.5	1.6	2.0	0.75	2.5	1.5	1.2
Y	2.0	0.75	2.8	1.5	1.9	2.5	1.0	3.1	2.0	1.8

(i) Is this a matched pair or not? Explain. [3]

(ii) Is there significant difference between X and Y? [8]

B6. (a) Suppose that we want to test the hypotheses

$$H_0 : p = 0.6 \text{ against } H_1 : p = 0.7$$

on the basis of a random sample of 10. If $T{:}Y = \sum_{i=1}^{10} X > 6$ is the critical region. Calculate

(i) the probability of type I error. [4]

(ii) the probability of type II error. [4]

(iii) power of the test. [1]

(b) 17 patients were taken at random to asses the levels of CD4 count before and after anti retroviral therapy. The results are given in the table below.

Patient no	Before	After
1	443	57
2	421	352
3	436	587
4	376	415
5	458	458
6	408	424
7	422	463
8	431	583
9	459	432
10	369	379
11	360	370
12	431	584
13	403	422
14	436	587
15	376	415
16	370	419
17	443	57

The following hypotheses are to be tested

H_0: No difference in median of the signed differences
H_1: Median of the signed differences is less than 0

[13]

(c) Let $30, 22, 32, 26, 24, 40, 34, 36, 32, 33, 28, 30$ be a realisation from a certain population using the data-set. Calculate the coefficient of skewness and comment on the result. [4]

(d) Econometric data was generated to asses the effects of dollarization (x_1) and inflation (x_2)on economic growth (y). The analysis generated the linear model below.

$$y = 100 + 0.5x_1 - 7.6x_2$$

Comment on the model. [4]

B7. (a) An analysis of accidents was made to determine the distribution of numbers of fatal accidents for commuter omnibus of different sizes. The results for 346 accidents are as follows

Size of omnibus	Small	Medium	Large
Fatal	67	26	16
Not fatal	128	63	46

Test at 1% level of significance whether the fatality of accidents depends on the size of the commuter omnibus. [13]

(b) A study on a group pf 8 people was conducted to ascertain whether or not physical exercise alleviate depression. Each person was randomly allocated to one of the three groups: 'no exercise', 20 minutes of jogging per day and 60 minutes of jogging per day. At the end of a month, each participant was asked to rate how depressed they now feel, on a Likert scale that runs from 1 − 100. Below is a table of ratings

no exercise	20 minutes jogging	60 minutes jogging
23	22	59
26	27	66
51	39	38
49	29	49
58	46	56
37	48	60
29	49	56
44	65	62

Use the Kruskal Wallis test to test if physical exercise alleviate depression. [11]

(c) Define the following terms

    (i) Run as used in runs test. [2]

    (ii) P- value . [2]

END OF QUESTION PAPER