BINDURA UNIVERSITY OF SCIENCE EDUCATION SFM224

BACHELOR OF STATISTICS AND FINANCIAL MATHEMATICS SWE103

APPLIED STATISTICS CSH204

Time : 3 hours NOV 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. Define the following terms

(a) Null hypothesis [3]

(b) Significance level [3]

(c) Test statistic [3]

(d) Skewness [3]

A3. The table below shows departures from normal daily temperatures recorded in Harare in 2019. We would like to know whether we may conclude that the pattern of departures above and below the normal is the result of an non random process [10]

Day	Departure from normal	Day	Departure from normal
1	12	16	-9
2	13	17	6
3	12	18	7
4	11	19	10
5	5	20	10
6	2	21	1
7	-1	22	1
8	2	23	3
9	-1	24	7
10	3	25	-2
11	2	26	-6
12	-6	27	-6
13	-7	28	-5
14	-7	29	-2
15	-12	30	-2

A4. Two independent random samples were selected from each of the normal populations. The data is in the following table

Sample 1	13	3	8	5	6	7	4	9	10	9	
Sample 2	14	7	7	9	5	8	9	10	11	12	13

(a) Calculate S_p^2, the pooled estimator of σ^2. [2]

(b) Test at 5% significance level whether the two populations are the same. [7]

SECTION B (60 marks)

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

B5. (a) The effectiveness of advertising for two rival products (Brand X and Brand Y) was compared. Market research at a local shopping center was carried out with participants being shown adverts for two rival brands of coffee which they rated on the overall likelihood of them buying the product (out of 10, with 10 being definitely going to buy the product). Half of the participants gave ratings for one of the products, the other half gave ratings for the other products.

	Brand X		Brand Y
Participant	Rating	Participant	Rating
1	3	1	9
2	4	2	7
3	2	3	5
4	6	4	10
5	2	5	6
6	5	6	8

Is there significant difference between the ratings given to each brand in terms of the likelihood of buying the product? (Use the Mann Whitney test) [13]

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

(c) 10 patients were taken at random to asses the levels of blood pressure before and after oral contraception. The blood pressure of each patient before (X) and after (Y) are given in the following table.

X	4.61	6.42	5.4	4.54	3.98	3.82	5.01	4.34	3.8
Y	3.84	5.57	5.85	4.80	3.68	2.96	4.41	4.34	3.8

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

(ii) Is there significant difference between blood pressure before (X) and after (Y) oral contraception? [8]

B6. (a) A group of students have been asked to rate a sentence on a scale of acceptability from 0 (totally unacceptable) to 5 (totally acceptable) from informal spoken and formal written english. An investigator predicts that the sentence will be judged as more acceptable in informal spoken that formal written english. The scores are given in the table below.

Subject no	Informal spoken	Formal written
1	5	5
2	4	2
3	5	3
4	4	4
5	3	1
6	2	3
7	4	3
8	5	1
9	4	2
10	2	3
11	4	2
12	4	3
13	5	3
14	3	5
15	3	0

Is there significant difference between the two sets of scores? Apply the sign test. [13]

(b) Let X_1 be a random sample from the Poisson distribution with parameter μ. Suppose that we want to test the hypotheses

$$H_0 : \mu = 2 \text{ against } H_1 : \mu = 3$$

using the decision T: reject H_0 if $X > 2$. Calculate

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

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

(iii) power of the test. [1]

(c) Let30, 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) The following table is the result of analysing a random sample invoices submitted by branches of a larger chain of book shops

	Novel	Textbook	General Interest
Hardback	3	16	2
Paperback	4	2	3

Asses at 5% level whether or not there is an association between the type of book sold and it's cover. [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) Test statistics [3]

(ii) Run as used in runs test. [3]

END OF QUESTION PAPER