



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

LOCATION SWANSEA

LLEOLIAD ABERTAWE

LEVEL 5

LEFEL 5

EXAMINATION PERIOD
APRIL 2021

CYFNOD ARHOLIADAU
2021

MODULE CODE: M2X9624
RESEARCH METHODOLOGY

(Statistical methods examination)

PART 2

TIME: 1 HOUR

Equipment or materials required – formula booklets and calculators allowed

Adnoddau neu ddefnyddiau sydd eu hangen -

ALL answers and workings should be included in answer book, scan and submit to Turnitin

Answer ALL questions

STUDENT NAME: FRASER REID

STUDENT NUMBER: 1802658

PART 2
(50%)

- Put the answers to the questions in the BOX or LINES provided.
- Include all workings out in the space provided

1. The following table presents mean diameter of ash (*Fraxinus ornus*) in Rome, Italy and an area 20 miles from the city in a suburban nature reserve with no industrial development or major roads.

Ash sample sites	Ash Nature Reserve Diameter cm (x2)	Ash Rome diameter cm (x2)
1	7.7 (59.29)	6.1 (37.21)
2	8.8 (77.44)	6.3 (39.69)
3	9.0 (81)	4.9 (24.01)
4	7.0 (49)	5.4 (29.16)
5	9.9 (98.01)	5.2 (27.04)
6	8.2 (67.24)	4.8 (23.04)
7	10.2 (104.04)	5.7 (32.49)
8	6.1 (37.21)	7.9 (62.41)

Determine using the T Test within 95% & 99% confidence limits, whether there is a difference in ash trees.

- (a) Construct a Null hypothesis based on diameter and pollution
(2%)

there is no statistical difference between ash tree diameter based on pollution levels

- (b) Calculate the following descriptive statistics

Mean diameter NR. (2%)

8.36

Mean diameter Rome (2%)

5.79

Standard Deviation NR (1%)

1.40

Standard Deviation Rome (1%)

1.005

- (c) Calculate the value of T (5%)

4.21

- (d) Interpretive statement (2%)

the calculated number is higher than the tabulated values in both 95% and 99% confidence limits, meaning we reject the null hypothesis, meaning pollution levels do affect the growth of Ash tree diameter.

TOTAL 15%
Q.1

Workings out

$$Q1 \quad NR - SD = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\frac{573.23 - \frac{4475.61}{8}}{7}}$$

$$\approx 1.97$$

$$R - SD = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}} = \sqrt{\frac{275.05 - \frac{2143.69}{8}}{7}}$$

$$\approx 1.01$$

$$\approx 1.005$$

$$t = \frac{|\bar{x}_{NR} - \bar{x}_R|}{\sqrt{\frac{(n_n-1)s_n^2 + (n_R-1)s_R^2}{n_n+n_R-2} \times \frac{n_n+n_R}{n_n n_R}}}$$

$$= \frac{8.36 - 5.79}{\sqrt{\frac{(8-1)5^2 + (7-1)1^2}{8+7-2} \times \frac{8+7}{8 \times 7}}}$$

$$= \frac{2.57}{\sqrt{\frac{27.125 + 4.2}{14} \times \frac{15}{56}}}$$

$$= \frac{2.57}{\sqrt{1.49 \times 0.25}}$$

$$= \frac{2.57}{\sqrt{0.3725}}$$

$$= \frac{2.57}{0.6103}$$

$$= 4.21$$

$$\approx 4.21$$

TOTAL
(12%) Q. 2

2. The following data set is from a survey of mean number of moths found at different types of street lighting in Bristol during February of the Pandemic Year (2021):

- LED street lighting
- Sodium bulbs

Sites	Mean Moth numbers at Sodium street lighting	Mean Moth numbers at LED street lighting	Differences	Rank	Signed rank
1.	9.5	13.6	-4.1	10	-10
2.	10.4	13.5	-3.1	9	~~-9~~ -9
3.	9.9	12.1	-2.2	7	~~-7~~ -7
4.	11.8	11.9	-0.1	1	-1
5.	10.8	12.2	-1.4	5	-5
6.	10.9	12.9	-2	6	-6
7.	11.2	10.9	0.3	2	+2
8.	11.3	12.2	-0.9	3	-3
9.	10.7	13.2	-2.5	8	-8
10.	11.8	13.1	-1.3	4	-4
TOTAL	108.3	125.6			

(a) Using the Wilcoxon test for matched pairs, determine any significant difference between the results.

Sum of the positive ranks (T^+) (3%)

Sum of negative ranks (T^-) (3%)

Critical value at $p=0.05$ (3%)

2

-53

8

(b) Interpretive statement (3%)

The lowest Sum of ranks is less than the tabulated value, therefore we reject the null hypothesis, meaning LED street lighting does attract more moths.

3. A biologist collects samples of midges *Dixella pupae* from vegetation on the edge of a pond in Gower (2020). Adult midges identified, counted and results input into Table below. (Assumption species were randomly distributed.)

	<i>Dixella autumnalis</i>	<i>Dixella aestivalis</i>	<i>Dixella amphibia</i>	<i>Dixella attica</i>	<i>Dixella fatica</i>	Total
Observed	40	35	5	10	10	100
Expected	20	20	20	20	20	100
O-E	20	15	-15	-10	-10	
O-E ²	400	225	225	100	100	
O-E ² /E	20	11.25	11.25	5	5	52.5

Perform a Chi square test to assess if there was any difference between the midge species identified. P 0.05 & P0.01. Is the distribution homogeneous?

- (a) Construct a Null hypothesis based on this research

(2%)

there is no statistical difference between the midge species identified

- (b) Calculate the Chi statistic at p. 0.05

(6%)

$$\chi^2 = \sum \frac{(O-E)^2}{E} = 52.5$$

$$\text{Chi stat } 0.05 = 9.488$$

$$5-1 = 4$$

(c) Interpretative statement

(2%)

the Calculated Value is higher than the tabulated value, therefore the results are significant meaning that the species of midges are distributed is not homogenous

TOTAL 10%
Q.3

Workings out

4. The following data set highlights Brecon Beacon visitor numbers pre and during pandemic lockdown on consecutive weekends in March 2020.

23 March – With the UK death toll hitting 335 deaths and 16 in Wales, Boris Johnson announced a nationwide 'Stay at Home' order would come into effect as of midnight and it would be reviewed every 3 weeks.[6] This would become known as the UK lockdown.

National Park	Visitor numbers (000,s) 16 th March	X ²	Visitor numbers (000,s) 23 rd March	Y ²	XY
1	38	1444	20	400	760
2	58	3364	26	676	1508
3	41	1681	18	324	738
4	21	441	29	841	609
5	48	2304	24	576	1152
6	28	784	14	196	392
7	26	676	19	361	494
8	36	1296	16	256	576
SUMS	Σx	Σx^2	Σy	Σy^2	Σxy
	296	11990	166	3630	6229

From linear regression analysis identify whether a prediction for NP visitor numbers can be calculated (First stages of linear regression) linking visitor numbers and pandemic lockdown. The following calculations need to be undertaken.

(a) Construct a Null hypothesis based for visitors (2%)

There is no statistical difference between visitor numbers pre/post lockdown.

(b) Calculate the value of **b** (gradient) (3%)

0.08

Calculate the value of **a** (Intercept) (3%)

14.23

(c) What would be the predicted visitor numbers during lockdown for a pre-lockdown figure of 46,000. (3%)

(d) Interpretative statement (2%)

Workings out

$$b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}} = \frac{6229 - \frac{296 \times 166}{8}}{11990 - \frac{(296)^2}{8}}$$

$$= \frac{87}{1038} = 0.08$$

$$a = \bar{y} - b\bar{x} = 16.6 - (0.08 \times 29.6) = 14.23$$

TOTAL 13%
Q.4