## CHUKA



## UNIVERSITY

## UNIVERSITY EXAMINATIONS

EXAMINATION FOR THE AWARD OF DEGREE OF BACHELOR OF SCIENCE IN ENVIRONMENTA SCIENCE

ENSC 261: STATISTICS FOR ENVIRONMENTAL SCIENCE
STREAMS: BSC (TOUR, WIEM, NARE) Y2S2
TIME: 2 HOURS
DAY/DATE: MONDAY 03/12/2018
11.30 AM - 1.30 PM

INSTRUCTIONS:

- Answer ALL Questions in section A and any Two from Section B
- Do not write anything on the question paper


## SECTION A (30 MARKS)

1. The number of malaria cases and the population size in different towns in January were as given below:

| Category | Nairobi | Eldoret | Meru | Kisumu |
| :--- | :--- | :--- | :--- | :--- |
| Cases | 350 | 150 | 80 | 400 |
| Population | $3,500,000$ | 550,000 | 150,000 | 380,000 |

(a) Determine the proportion of the malaria cases of the total in each of the 4 towns.
[2 marks]
(b) Determine the ratio of malaria cases between the 4 towns
[2 marks]
(c) Determine the rate of malaria cases per 100,000 in each of the four towns. [2 marks]
2. In a given farm, the weights of sheep in a farm taken and recorded as shown below.

| Weight | 25 | 30 | 35 | 40 | 45 | 50 | 60 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Frequenc <br> y | 15 | 12 | 18 | 20 | 14 | 8 | 2 |

Determine the mean absolute deviation.
3. The concentration of lead in soil samples collected from Kiambu dumpsite is given below

| Sample | A | B | C | D | E | F | G |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Concentratio <br> $\mathrm{n}(\mathrm{g} / \mathrm{kg})$ | 10 | 15 | 12 | 20 | 18 | 13 | 22 |

Determine the confidence interval of the mean at $95 \%$ confidence level.
[6 marks]
4. Write short notes on questionnaires and schedules.
5. Differentiate between nominal and ordinal type of data.
6. The correlation coefficient between level of damage by pest in Mango trees and the total weight of fruits in each tree was determined to be 0.90 . briefly explain the nature of the relationship represented by this value.
[4 marks]

## SECTION B (40 MARKS)

7. (a) The height (cm) of trees in five research plots with different treatments was taken every year for a maximum of eight years as shown below.

|  | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | Year 6 | Year 7 | Year 8 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Plot 1 | 30 | 35 | 48 | 52 | 55 | 60 | 66 | 70 |
| Plot 2 | 20 | 22 | 27 | 28 | 29 | - | - | - |
| Plot 3 | 35 | 45 | 53 | 60 | 65 | 68 | - | - |
| Plot 4 | 22 | 25 | 28 | 32 | 35 | 37 | 40 | 42 |
| Plot 5 | 35 | 40 | 48 | 52 | 58 | 62 | 68 | 75 |

Perform an analysis of variance at $95 \%$ confidence level to determine whether there were significant differences in growth between the different research plots.
[12 marks]
(b) In order to test whether a given regime of physical training was important in weight reduction, the weights of participants were taken before and after the training as shown below

| Participan <br> t | A | B | C | D | E | F | G | H |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Before | 60 | 55 | 70 | 65 | 58 | 62 | 80 | 75 |
| After | 58 | 55 | 71 | 62 | 58 | 60 | 81 | 73 |

By use of Wilcoxon Matched-Pairs Signed-Ranks Test determine whether the training regime was significant in affecting the weight of participants.
[8 marks]
8. (a) The weight of tomatoes harvested per plant for different watering regimes was observed and recorded as shown below

| Water <br> $(\mathrm{ml})$ | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 | 65 | 70 | 75 | 80 | 85 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Weight <br> $(\mathrm{g})$ | 100 | 120 | 140 | 160 | 180 | 200 | 220 | 240 | 260 | 280 | 300 | 305 | 305 | 305 |

(i) Determine the correlation coefficient ' $r$ ' that represent the relationship between the two variables. marks]
(ii) Determine whether the relationship is significant at $95 \%$ confidence level.
[4 marks]
(b) Discuss any four biases that can affect the process of sampling. [8 marks]
9. (a) A drug that cures insomnia was tested on 12 participants. They were given the drug and then given a placebo to test the time they would take to fall
asleep. The
The time taken
participants were required to count numbers until they fell asleep. is given below:

| Participant | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| Placebo(min <br> $)$ | 30 | 35 | 47 | 38 | 53 | 45 | 60 | 33 | 34 | 45 | 40 | 48 |
| Drug (min) | 10 | 15 | 20 | 28 | 22 | 26 | 24 | 14 | 20 | 10 | 16 | 27 |

using a paired $t$-test determine whether there is any significant difference between the drug and the placebo.
[12 marks]
(b) Briefly discuss the procedure that should be followed in testing hypothesis.

Given

$$
\frac{\sum f / x-x / i}{n}, C l=\dot{x} \pm\left(\begin{array}{lll}
z & x & \left.s e_{m}\right), t=\frac{D}{\sqrt{\frac{s^{2} D}{N}}}
\end{array}\right.
$$

$$
t=\frac{\dot{x}_{1}-\dot{x}_{2}}{s e_{d}}, M A D=i
$$

$$
r=\frac{N \sum X Y-\sum X \sum Y}{\sqrt{\left[N \sum X^{2}-\left(\sum X\right)^{2}\right]\left[N \sum Y^{2}-\left(\sum Y\right)^{2}\right]}}, t = r \longdiv { \frac { n - 2 } { 1 - r ^ { 2 } } , S _ { 2 } ^ { D } = \frac { \sum D ^ { 2 } - \frac { ( \sum D ) ^ { 2 } } { N } } { N - 1 } }
$$

