



**JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND
TECHNOLOGY**

**UNIVERSITY EXAMINATIONS
RESITS FOR 2019/2020 ACADEMIC YEAR**

**THIRD YEAR, FIRST SEMESTER EXAMINATION FOR THE DEGREE OF
BACHELOR OF SCIENCE (BIOLOGICAL SCIENCES)**

COURSE CODE: SBI 3336 / SZL 308

COURSE TITLE: Biostatistics

DATE

TIME

DURATION: 2 HOURS

INSTRUCTIONS:

- 1. This paper contains two sections (A and B)**
- 2. Answer ALL questions in Section A and any Two (2) questions in Section B**
- 3. Write ALL answers in the booklet provided**
- 4. You may use illustrations in your answers as you deem necessary**

SECTION A:**(30 MARKS)**

1. Describe the three measures of central tendency (3 marks)
2. Calculate the variance, standard deviation and mean of the data set below (3 marks)

29	34	39	21
40	35	35	23
20	33	29	39
32	31	36	27
23	26	27	24

3. Suppose a random sample of 100 12-year-old boys were chosen and the heights of these 100 boys recorded. The sample mean height is 64 inches, and the sample standard deviation is 5 inches. Calculate the 95% confidence interval of the mean. You may assume heights of 12-year-old boys are normally distributed. (3 marks)
4. List the statistical tests you would use to test the following:
 - a. Determine the variation in wing length of butterflies from five different locations
 - b. To validate the statement “people having high cholesterol suffer more from hypertension”
 - c. Determine the difference in the weights in two populations.
5. Explain why scientists do not conduct a population census when they set out to carry scientific research. (3 marks)
6. Suppose you intend to compare two populations, state the null and alternate hypothesis for the following scenarios
 - a. A two-tailed test (1 mark)
 - b. A left tailed test (1 mark)
 - c. A right-tailed test. (1 mark)
7. State three assumptions of analysis of variance (3 marks).
8. Explain the following terminologies:
 - a. Simple linear regression (1 mark)
 - b. Cause and effect relationship (1 mark)
 - c. Spurious correlations (1 mark)
9. List three advantages of non-parametric tests.
10. Describe the Kruskal-Wallis test indicating scenarios when it is used. (3 marks).

SECTION B**(40 MARKS)**

11.
 - a. Discuss the Chi-square (χ^2) test procedure describing its use, its characteristics and how to make inferences from its results. (8 marks)
 - b. In an experiment with peas, the results included 360 round and yellow seeds, 130 round and green seeds, 118 wrinkled and yellow seeds and 32 wrinkled and green seeds. According to the Mendelian theory of heredity, the numbers should be in the ratio 9:3:3:1. Is there any evidence of difference between the expected and observed frequencies for each seed category? (12 marks)
12. You have been asked to conduct a study in Siaya County to find out if a new invasive species of a lepidopteran pest that lays its eggs in the stems of maize is present in all districts.
 - a. Discuss the sampling technique that you would use. (5 marks)
 - b. Describe the type of data that you would collect. (4 marks)
 - c. State the hypothesis that you would use in your study. (3 marks)
 - d. Based on your hypothesis, describe a statistical test that you would use. (8 marks)
- 13.

- a. Discuss the paired t-test procedure describing its use, test statistics and how to make inferences from its results. (10 marks)
- b. The data set below indicates the yield of wheat during the years 2011 and 2012 from the same fields. At 95% confidence interval, determine whether there is a difference in yield between the two years.

2011	2012	2011	2012	2011	2012
81	80.7	82.3	103.1	98.9	66.4
105.4	82.3	77.3	105.1	89	49.9
119.7	80.4	78.4	116.5	69.1	96.7
109.7	87.2	131.3	139.9	89.3	61.9
98.3	84.2	89.6	129.6	104.1	80.3
146.6	100.4	119.8	98.9	86.9	67.7
142	115.5	121.4	61.9	77.1	66.7
150.7	112.2	124	96.2	78.9	67.4
191.5	147.7	140.8	125.5	101.8	91.8
145.7	108.1	124.8	75.7	96	94.1

14. Consider the data set below which shows the results of a study that investigated the number of animal species in three different habitats. At 95% confidence interval determine whether there is a difference in the number of animal species between the three habitats. Summarize your results in a table.

Number of animal species		
Island	Mainland	Peninsula
5	5	7
2	7	5
1	6	6
5	3	6
2	9	5
4	8	6
3	10	5
4	9	6
2	9	6
3	7	6