**BOTA 302** 



## **BOTA 302: BIOSTATISTICS**

STREAM: Y3 S2

TIME: 2 HOURS

DATE: 10/05/2022

DAY: TUESDAY, 3:00 PM - 5:00 PM

### **INSTRUCTIONS:**

1. Do not write anything on this question paper.

2. Answer all questions in section A and any TWO in section B.

### **SECTION A: ANSWER ALL QUESTIONS (40 MARKS)**

#### (*P value=0.05 for ALL questions*)

1. (a) By giving examples, explain the following types of scales as applied in Biological data

i. Ordinal	(1 mark)
------------	----------

ii. Nominal (1 mark)

(b) Outline reasons why you will be required to use stratified random sampling.

(3 marks)

a) Rank the following numbers from the lowest to the highest. (2 marks)
 2.8, 6.1, 7.4, 2.8, 3.3, 4.7, 4.4, 3.2, 7.5, 4.7, 4.7

b) Calculate the mean and median for the data above in 4 (a). (4 marks)

- 3. Briefly describe three graphical forms of data distributions. (6 marks)
- 4. Explain the advantages and disadvantages of the median. (5 marks)

- 5. Differentiate between the following terms as used in biostatics. (6 marks)
  - a) Null hypothesis and alternate hypothesis
  - b) Type I error and Type II error
  - c) Correlation coefficient and regression analysis
- 6. A sample of seedlings has been found to have the following heights (mm) in a nursery: 121, 125, 128, 134, 136, 138, 139, 141, 144, 145, 149, 151.
  Calculate the standard error of the mean. (5 marks)
- 7. Explain the assumptions for ANOVA (analysis of variance). (4 marks)
- 8. Outline the conditions that must be met for application of parametric statistics. (3 marks)

# SECTION B: ANSWER ANY TWO QUESTIONS (30 MARKS)(P value=0.05)

11. In a study of the nutritive value of plants in a habitat, the crude protein content (mg/g of plant material) was determined as shown in the following data set:

Protein content (mg/g plant material)	Frequency
8.35-8.45	3
8.45-8.55	7
8.55-8.65	9
8.65-8.75	2
8.75-8.85	19
8.85-8.95	19
8.95-9.05	26
9.05 -9.15	20
9.15 -9.25	15
9.25 -9.35	12
9.35-9.45	6

- (a) What is the mode for this data set? (2 marks)
- (b) Draw the histogram and the frequency polygon for these data. (3 marks)
- (c) Calculate the cumulative frequencies for the different classes. (5 marks)
- (d) Calculate the mean and the standard deviation. (5 marks)
- 12. (a) In an experiment on barley-genetics an F2 generation was

	Green	Green	Chlorina	Chlorina	
	Non two row	Two row	Non two row	Two row	
Observed	1.262	301	282	143	

Test the hypotheses that there is a 9:3:3:1 ration dihybrid segregation ( $P=0.05, x^2=7.81$ ).

(8 marks)

(b) An attempt was made to test whether there is a 1:1 sex ratio among white throated bee eaters living in 4 different locations around Naivasha.

Location	Males	Females		
1	64	74		
2	33	44		
3	19	25		
4	15	16		

Perform a heterogeneity chi-square analysis to test whether the four samples may be justifiably pooled.  $(P=0.05,x^2=7.81)$  (7 marks)

13. (a) The following values were obtained during digestibility trials of two feeds by tame buffaloes. (P=0.05, *t critical one tail* =1.7139, *t critical two tail* = 2.0687)

Feed 1 6.25, 5.78, 3.5, 4.7, 3.8, 4.0, 4.7, 5.8, 4.0, 5.7, 6.0, 5.0, 6.5

Feed 2 11.2, 10.9, 10.8, 11.3, 7.4, 8.2, 7.5, 9.2, 10.2, 8.4, 8.6, 10.2

Formulate null hypothesis to test whether the two feeds have similar digestibility and if not which of the two is more digestible. (10 marks)

(b) Describe the requirements or assumptions that have to be met for regression analysis (5 marks)

14. (a) The following are body weights of impalas caught from different parts of the country.Does body weight vary with area of origin? (10 marks) (P=0.05) (*F critical=3.49*)

Nairobi	Amboseli	Nakuru	Tsavo
70.8	78.7	108.6	97.9
67.0	77.7	108.1	94.2
75.0	84.0	106.2	93.1
68.6	76.3	102.5	95.7

(b) Measurements of serum cholesterol (mg/100mls) and arterial calcium deposition (mg/100g) were made on 12 animals. The data was as follows:

Calcium	64	57	47	64	29	29	45	37	68	62	41	29
Cholesterol	298	303	233	287	236	245	265	233	286	290	264	239

Calculate the correlation coefficient.

(5 marks)