JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF BIOLOGICAL, PHYSICAL, MATHEMATICS AND ACTURIAL SCIENCES

UNIVERSITY EXAMINATION FOR BACHELOR OF ACTUARIAL SCIENCE

2023/24

MAIN REGULAR

COURSE CODE: WAB 2201

COURSE TITLE: Statistical Modelling

EXAM VENUE **STREAM: Dip. Applied Statistics**

DATE:…… EXAM SESSION: ONE

TIME: $1\frac{1}{2}$ HOURS

Instructions to the Candidate:

1. **Answer ALL in-section A and any other two questions only in Section B.**
2. **Candidates are advised not to write on the question paper.**
3. **Candidates must hand in their answer booklets to the invigilator while in the examination room.**

SECTION A: [30 Marks]

In this section, you are required to provide short responses in the answer sheets provided.

1. State three forms of data analysis: [3 Marks]

2. Based on the data gathered at a particular weather station on the monthly rainfall in mm(r) and the average number of hours of sunshine per day (s), a statistician has determined the following relationship.

$$\boldsymbol{S = 9 - 0.1r}$$

   i. Estimate the average number of hours of sunshine per day if monthly rainfall is 50mm [2 Marks]

3. State three main data sampling techniques. [3 Marks]

4. Transform the relationship $\boldsymbol{Y = ab^x}$ to linear form [4 Marks]

5. Because heat and humidity can cause stress and lower milk portion, a dairy farmer is investigating the relationship between cattle respiration rate (breath per minute) and temperature. The data for a random sample of day and Senepol cattle are given below.

Temperature	85	89	81	86	91	92	87	84	80	82	83	85
Breaths per min.	58	61	42	56	67	65	62	44	46	52	45	53

   i.) Identify the independent and dependent variables [2 Marks]

   ii.) List the ordered pairs in the data set [2 Marks]

6. a.) Define correlation [2 Marks]

   b.) State two main advantages of correlation [2 Marks]

7. State four assumption of analysis of variance $ANOVA$ [4 Marks]

8. Differentiate between Null hypothesis and Alternate hypothesis [2 Marks]

9. State two properties of least square line [2 Marks]

SECTION B: [40 MARKS]

In this section, you are required to select ANY TW0 questions

10. (a) The developer of a residential community in Nairobi in which each all-electric home is equipped with air conditioner believe that monthly electricity bill (in dollars) is linearly related to cooling degrees days (CDD). A random sample months and home was selected. The electric bill and CDD were recorded for each home. The data are in the following table

CDD	18	18	328	211	261	94	261	110	110	261	328
Bill	4.99	49.53	70.31	69.79	65.61	37.03	100.22	56.05	70.33	103.64	134.9982.27
	94	328	328	94	94	94	211				
	82.27	89.06	91.38	112.4	107.82	94.23	25.31				

(a.) find the estimated regression line and explain the meaning of the estimated coefficient $\hat{B}_1$ [10 Marks]

(b) complete the ANOVA table (without replacement) and interpret the coefficient of determination. [10 Marks]

11. The Delmonte company sells a large sterilizer with four extendable shelves for medical tools. Company engineers believe that the time to reach operating temperature from cold start (y, measured in minutes) is linearly related to the thickness of installation (x, in inches) A random sample of size $n = 12$ thickness was selected and time to reach operating temperature recorded for each. The data and summary statistics are as follows:

X	1.3	1.8	0.9	1.6	2.6	1.5	2.1	3.0	0.8	2.4	2.5	2.6
Y	8.0	6.9	8.1	7.0	6.3	6.5	6.4	5.8	8.3	8.3	6.6	6.6

$\sum x_1 = 23.1, \quad \sum y_1 = 84.8, \quad \sum x_1 y_1 = 1585\,, \quad \sum x_1^2 = 50.13\,, \quad \sum y_1^2 = 607.66$

i Complete the ANOVA table and compute F test for a significance regression. Use a significance level of 0.05

12. (a) A schoolteacher is investigating the claim that class size does not affect K.C.S.E results. His observation of nine K.C.S.E classes are as follows:

$Student\ class(c)$	35	32	27	21	34	30	28	24	7
$Average\ K.C.S.E\ score(p$	5.9	4.1	2.4	1.7	6.3	5.3	3.5	2.6	1.6

i. Determine the regression line for p on c [5 Marks]

ii. Assuming X_5 was not included in the result above and contain 15 students, calculate the estimate of the average K.C.S.E points score for this individual class [5 Marks]

(b) A statistician is fitting the following linear regression model through angle;

$$Y_i = \beta x_i + e_i \quad , \quad e \sim N(0, \delta^2)\,; \quad i = 1,2,3, \cdots, n$$

i. Show that least square estimator of β is given by

$\hat{\beta} = \frac{\sum x_i y_i}{\sum x_i^2}$ [5 Marks]

ii. Derive the bias of $\hat{\beta}$ under this model [5 Marks

13. (a) A television execution is trying to decide which sport to bid on in order to obtain the broadcast for next season. A random sample of adults was obtained and each was asked to name their favourite sport to watch on TV. The data and historical proportion are given in the following table:

.

TV sport	*Frequency*	*Historical proportion*
Auto racing	84	0.10
Base ball	119	0.22
Fishing	16	0.05
Football	213	0.33
Golf	36	0.06
Hunting	17	0.03
Soccer	79	0.13
Tennis	43	0.05
Track and field	13	0.03

Is there evidence to suggest that any of the true cell properties differ from historical properties

Use $\propto = 0.01$ [16 Marks]

(b). Explain two types of correlation [4 Marks]