JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF MATHEMATICS AND ACTUARIAL SCIENCE

UNIVERSITY EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE ACTUARIAL

SPECIAL RESIT 2 2020/2021 ACADEMIC YEAR

REGULAR (MAIN)

COURSE CODE: SAS 306

COURSE TITLE: Statistical Modeling

EXAM VENUE: **STREAM: (BSc. Actuarial)**

DATE: **EXAM SESSION:**

TIME: 2.00 HOURS

Instructions:

1. **Answer question 1 (Compulsory) and ANY other 2 questions**
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.**

QUESTION ONE

1.a) Explain the linearity of regression. (2mks)

b) Give the appropriate linear form of the following non-linear functions when subjected to transformation

i) $Y = \beta_0 e^{\beta_1 x}$

ii) $Y = \beta_0 x^{\beta_1}$

iii) $Y = \beta_0 10^{\beta_1 x}$ (6mks)

c) Explain the significance of residual analysis in a simple linear regression model. (2mks)

d) Explain ways by which test of hypothesis and analysis of variance is similar and different as far as test of significance of regression coefficient of linear model is concerned. (3mks)

e) Given the regression estimates as; $\hat{\beta}_0 = \overline{y}$ and $\hat{\beta}_1 = \beta_1 + \frac{\sum_{i=1}^{n}(x_i - \overline{x})\ \varepsilon_i}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$, find the variances of the estimates $\hat{\beta}_0$ and $\hat{\beta}_1$. (10mks)

f) Given the data below on the number of hours (x) which 9 students studied for statistics test and their scores (y) on the test

X	4	9	10	14	4	7	12	22	1
Y	31	58	65	73	37	44	60	91	21

.

Find the normal regression line that approximates the regression of the test scores on the number of hours studied. (7mks)

QUESTION TWO

2. The following data shows Maths achievement test scores and the SAC 306 grades for 10 university students.

Students	1	2	3	4	5	6	7	8	9	10
Maths achievement test scores(X)	39	43	21	64	57	47	28	75	34	52
Final SAC 306 Grades(Y)	65	78	52	82	92	89	73	98	56	75

a) Calculate $\overline{X}$ and $\overline{Y}$

b) Calculate least square estimates of $\hat{\beta}_0$ and $\hat{\beta}_1$.

c) Determine the fitted line

d) Determine whether the value of $\hat{\beta}_1$ provides sufficient evidence to indicate that $\hat{\beta}_1$ differs from zero(linear relationship exists between students Maths achievement score x and his final SAC 306 grades y) at $\alpha = 0.05$. (20mks)

QUESTION THREE

3. a) Explain the meaning of standardized residuals as far as testing the adequacy of linear regression model is concerned. (2mks)

b) An article in the *Journal of Sound and Vibration* (Vol. 151, 1991, pp. 383–394) described a study investigating the relationship between noise exposure and hypertension.
The following data are representative of those reported in the article.

Y	1	0	1	2	5	1	4	6	2	3	5	4	6	8	4	5	7	9	7	6
X	60	63	65	70	70	70	80	90	80	80	85	89	90	90	90	90	94	100	100	100

(a) Fit the simple linear regression model using least squares.

(b) Obtain $\widehat{\sigma^2}$

c) Test the adequacy of the regression model by residual analysis (18mks)

QUESTION FOUR

4. In a simple linear regression given as $Y_i = \beta_0 + \beta_1(x_i - \bar{x}) + \varepsilon_i$, β_0 and β_1 are often estimated with $\hat{\beta}_0$ and $\hat{\beta}_1$ respectively.

a) Show that $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased estimates of β_0 and β_1 respectively. (10mks)

b) Obtain the regression line of Y on X for the following data.

X	1	2	3	4	5	6	7	8	9
Y	9	8	10	12	11	13	14	16	15

(10mks).

QUESTION FIVE

5 a) Explain the meaning of non-linear regression. (2mks)

b) A shop sells home computers. The numbers of computers sold in each of five successive years are given in the table below;

Years(X)	1	2	3	4	5
Sales(Y)	10	30	70	140	210

i) Assuming that the sales (y) and the year (x) are related by the equation $y = ab^x$ find the least square regression line, hence estimate the constants a and b.

ii) The shop manager uses this relationship to predict the sales in the sixth year. Find the predicted sales.

iii) Obtain a 95% confidence interval for the slope of the regression line. (18mks)