JARAMOGI OGINGA ODINGA UNIVERSITY OF SCIENCE AND TECHNOLOGY

SCHOOL OF AGRICULTURAL AND FOOD SCIENCES

UNIVERSITY EXAMINATION FOR DEGREE OF BACHELOR OF SCIENCE ACTUARIAL SCIENCE

3RD YEAR 2ND SEMESTER 2023/2024 ACADEMIC YEAR

MAIN REGULAR

COURSE CODE: WAB 2312

COURSE TITLE: STATISTICAL MODELLING

EXAM VENUE: **STREAM:**

DATE: EXAM SESSION:

TIME: 2.00 HOURS

Instructions:

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

QUESTION ONE COMPULSORY (30 MARKS)

a) Briefly state four assumptions of multiple linear regression model (8marks)

b) Briefly state three statistical properties of Linear Smoothers (6marks)

c) Under what circumstances would you use Poisson Regression? (2marks)

d) What are the assumptions of Poisson Regression? (6marks)

e) The decreasing value of an item that was purchased new 2008 is listed below.

Year	2008	2009	2010	2011	2012	2013	2014
Value of item in $	40	35.5	29.61	21.20	15.73	13.24	10.99

i) Write an equation relating the value of the item and the year it was purchased (4mks)

ii) Predict when the item will be worth $ 1.92 (1mark)

f) Use the data below to regress the data to a second order polynomial and find the value of $\propto$ when Temperature is 700F (3marks)

Temperature (^{0}F)	80	40	-40	-120	-200	-280
$\propto$	6.47	6.24	5.72	5.09	4.30	3.33

QUESTION TWO (20 MARKS)

Evaluate the following dataset to fit a multiple linear regression model (10marks)

Y	x_1	x_2
140	60	22
155	62	25
159	67	24
179	70	20
192	71	15
200	72	14
212	75	14
215	78	11

a) Consider the simple linear regression model: $Y_i=\beta_0+\beta_1 X_i+\varepsilon_i$ with $\varepsilon_i \sim$ i.i.d. $\mathcal{N}(0,\sigma^2)$. Suppose you estimated of the parameters of this model using least squares with a dataset containing 1000 observations. Some calculations using the X matrix, Y vector, and vector of residuals (e) are provided below. Use that information to test the Null Hypothesis that $\beta_1 = 5$ at a 95% confidence level. What do you conclude?

$$[X'X]^{-1} = \begin{pmatrix} 0.5 & 0.1 \\ 0.1 & 3 \end{pmatrix} \qquad X'Y = \begin{pmatrix} -4 \\ 2 \end{pmatrix} \qquad e'e = 212.91$$ (10marks)

QUESTION THREE (20 MARKS)

a) It is suspected from theoretical considerations that the rate of water flow from a firehouse is proportional to some power of the nozzle pressure. Assume pressure data is more accurate. You are transforming the data

Flow rate, F(gallons/min)	96	129	135	145	168	235
Pressure, p(psi)	11	17	20	25	40	55

What is the exponent of the nozzle pressure in the regression model $F = ap^b$ (10marks)

b) When using the transformed data model to find the constants of the regression model $y = ae^{bx}$ To best fit $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, what is the sum of the square of the residuals that is minimized (5marks)

c) Find the transformed data model for the stress-strain curve $\sigma = k_1 \varepsilon e^{-k_1 \varepsilon}$ for concrete in compression, where σ is the stress and ε is the strain, (1mark)

d) Fill in the missing entries of the partially completed one-way ANOVA table. (4marks)

Source	df	SS	MS = SS/df	F-statistic
Treatments		2.124	0.708	0.75
Error	20			
Total				

QUESTION FOUR (20 MARKS)

a) Consider the following training data:

X	y
1	3
2	1
3	0.5

Suppose the data comes from a model $y = cx^{\beta} + noise$, for unknown constants c and β. Use least squares linear regression to find an estimate of c and β (10marks)

b) The sales of a company (in million dollars) for each year are shown in the table below.

x (year)	c) 2005	d) 2006	e) 2007	f) 2008	g) 2009
y (sales)	h) 12	i) 19	j) 29	k) 37	l) 45

i) Find the least square regression line y = a x + b.

ii) Use the least squares regression line as a model to estimate the sales of the company in 2012. (10marks)

QUESTION FIVE (20 MARKS)

a) Consider the linear model $y = \beta_1 X_1 + \beta_2 X_2 + \varepsilon, E(\varepsilon) = 0, V(\varepsilon) = 1$ where the study variable y and the explanatory variables X_1 *and* X_2 are scaled to length unity and the correlation coefficient between X_1 *and* X_2 is 0.5. Let b_1 *and* b_2 be the ordinary least squares estimators of β_1 *and* β_2 respectively. Find the covariance between b_1 *and* b_2 (10marks)

b) A company manufactures an electronic device to be used in a very wide temperature range. The company knows that increased temperature shortens the life time of the device, and a study is therefore performed in which the life time is determined as a function of temperature. The following data is found:
(10marks)

Temperature in Celcius (t)	10	20	30	40	50	60	70	80	90
Life time in hours (y)	420	365	285	220	176	117	69	34	5

I. Calculate the 95% confidence interval for the slope in the usual linear regression model, which expresses the life time as a linear function of the temperature.

II. Can a relation between temperature and life time be documented on level 5%