MURANG'A UNIVERSITY OF TECHNOLOGY

SCHOOL OF PURE AND APPLIED SCIENCE

DEPARTMENT OF APPLIED SCIENCE

UNIVERSITY ORDINARY EXAMINATION

2017/2018 ACADEMIC YEAR

FOURTH YEAR , **SECOND** SEMESTER EXAMINATION FOR BACHELOR OF SCIENCE IN MATHEMATICS AND COMPUTER SCIENCE AND **THIRD** YEAR, **SECOND** SEMESTER EXAMINATION FOR BACHELOR OF SCIENCE IN APPLIED STATISTICS WITH PROGRAMMING.

AMS 311: REGRESSION MODELLING

DURATION: 2 HOURS

DATE: 18TH APRIL 2018

TIME: 8.00AM – 11.00AM.

Instructions to Candidates:

1. Answer **Section A** and **Any Other Two** questions in **Section B**.
2. Mobile phones are not allowed in the examination room.
3. You are not allowed to write on this examination question paper.

SECTION A – ANSWER ALL QUESTIONS IN THIS SECTION (30 Marks)

QUESTION ONE

a) State and explain any two assumptions used in regression analysis. (2 Marks)

b) Given that $Y_1 = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon_i$ is a multiple linear regression with two independent variables, derive the normal equations using the least square method and write down the matrix notations used in solving the regression coefficiently. (5 Marks)

c) The following data show the number of hours that 10 students studied for a statistics test and there score on the test.

Hours Studied (x)	3	10	15	5	11	7	12	20	1	17
Test Scores (y)	32	60	75	40	65	45	60	90	20	80

i. Fond the regression equation of least squares **???** that approximates the regression of the test scores on the number of hours studied. (5 Marks)

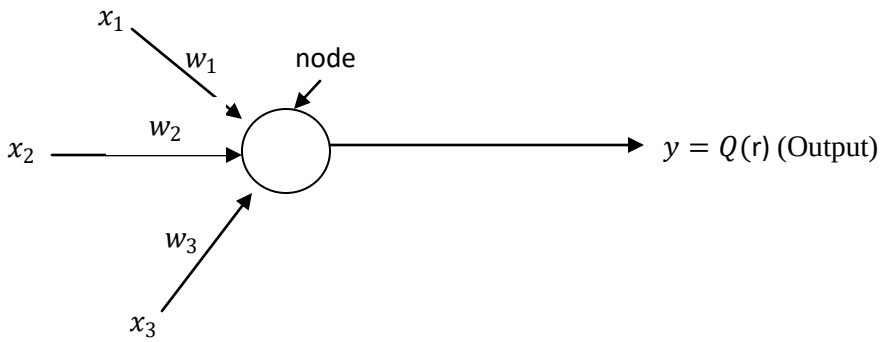
ii. Predict the test score of a student who studied for 16 hours. (2 Marks)

d) Many patients get concerned when a test involves injecting of radioactive material. For example for scanning the gull bladder, a few drops of the Technetium-99mm would gone in about 6 hours. It however takes about 24 hours for the radiation levels to reach what we are exposed to in a day-to day activities. Below is a data od relative intensity of radiation as a function of time.

t (hours)	0	1	3	5	7	9
r	1.00	0.891	0.708	0.562	0.477	0.355

???????????

e) The figure below shows a diagram of a single artificial neuron unit.



The model has 3 inputs $X = (X_1, X_2, X_3)$ that receive pnly binary signal (either0 or 1). How many different input patters can the node receive if it had four inputs? (2 Marks)

f) In generalized linear models the response variable distribution must be a member of the exponential family. Given the probability man function (pmf) of binomial distribution $Bin(n, p)$with parameters n and p is

$$f(\mu, p) = \binom{n}{\mu} p^{\mu}(1-p)^{n-\mu}, \mu = 0,1,\cdots,n$$

i. Express $Bin(n,p)$ as a member of exponential family and show that the natural parameter if given by

$d(\mu) = l_n\left(\frac{p}{1-p}\right)$ (2 Marks)

ii. Show that the mean of $Bin(n,p)$ is np using

$E(a/\mu) = \frac{-c^{'}(\theta)}{b^{'}(\theta)}$ (3 Marks)

g) A researcher studying the characteristics of a population was given a set of data which represented the ages at the time of death for 10 people working in a factory. Individual age: 35, 40, 45, 45, 47, 49, 55, 57, 59, 65. Using a uniform Kernel with a bandwidth of $b = 10$, determine the Kernel density estimate of the probability????

SECTION B – ANSWER ANY TWO QUESTIONS IN THIS SECTION

QUESTION TWO (20 MARKS)

a) Let $y_{1,}y_2, \cdots, y_n$ be independent normal random variables with means $\beta_0 + \beta_1 X_1, \cdots, \beta_0 + \beta_n X_n, i = 1, \cdots, n$ and common variance r^2. Find

i. The point estimators of $\beta_0, \beta_1\ and\ \partial^2$ using the maximum likelihood method (7 Marks)

ii. Show that the estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased estimates of β_0 and β_1 respectively. (6 Marks)

b) Given the non-linear model $Y_u = f(\tilde{t}u, \tilde{\theta}) + \varepsilon u\ where\ E(\varepsilon_u) = 0, var(\varepsilon_u) = 6^2\ and\ \varepsilon \sim N(0, 6^2), cov(\varepsilon_i \varepsilon_i) = 0, u = 1, \cdots, n$. Use Taylor Series to show that the estimate of β is given by

$\hat{\tilde{\beta}} = (Z_0^1 Z_0)^{-1} Z_0^{-1} Y_0$ (7 marks)

QUESTION THREE (20 MARKS)

a) A study was done on the effects of temperature on the yield of a chemical process. The following data was obtained.

Temperature (X)	5	9	3	2	4	0	7	2	3	4	5
Yield (Y)	3	5	4	7	10	8	9	14	13	13	20

i. Assuming a model $Y = \beta_0 + \beta_1 X + \varepsilon_1$ what are the least square estimates of $\beta_0\ a\text{nd}\ \beta_1$? (5 Marks)

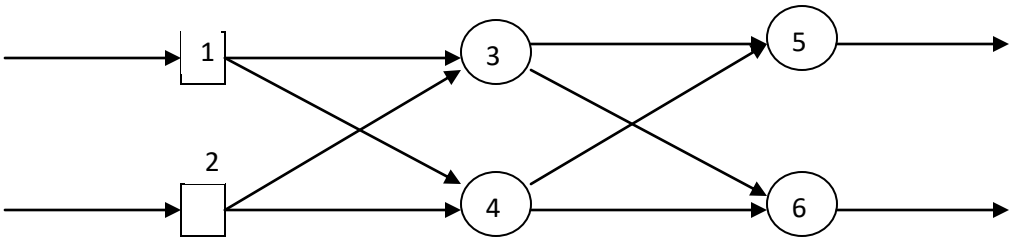
ii. Construct an Anoval table and test the hypothesis $H_0: \beta_1 = 0 \; vs \; H_1: \beta_1 \neq 0$ with $\propto = 0.05$ for β_1. (5 Marks)

b) The data given below shows the weights of 10 mothers at the time they delivered their first born in a given hospital.

Weight 55, 60, 65, 65, 67, 69, 75, 77, 79, 85.

Using triangular Kernel density with bandwidth $h = 10$, find the Kernel smoothed density estimate $\hat{f}(70)$. (4 Marks)

c) The following diagram represents a feed forward neural network with hidden layer.



The table below lists all the weight in the network

$$w_{13} = -2 \qquad w_{35} = 1$$
$$w_{23} = 3 \qquad w_{45} = -1$$
$$w_{14} = 4 \qquad w_{36} = -1$$
$$w_{24} = -1 \qquad w_{46} = 1$$

Each of the nodes 3, 4, 5 and 6 uses the following action function

$$\varphi(r) = \begin{cases} 1 & if \;\; r = 0 \\ 0 & otherwise \end{cases}$$

Where r denotes the weighted sum of a node. Each input nodes (1 and 2) can only receive binary values (0 or 1). Calculate the output of the network for each of the input patterns.

Pattern:	P_1	P_2	P_3
Node 1:	0	1	0
Node 2:	0	0	1

(6 Marks)

QUESTION FOUR (20 MARKS)

a) Show that the least square estimates of the multiple regression coefficients in matrix form are given by

$$\beta = (X'X)^{-1}X'Y$$

Where X' is the transpose of X and $(X'X)^{-1}$ is the inverse of $(X'Y)$. (6 Marks)

b) The following data shows the number of bedrooms (X_1) the number of bathrooms (X_2) and the price at which a random sample of four family houses sold recently. Use matrix notations in (a) above to determine the least squares estimates of multiple regression coefficients for the following data. (14 Marks)

No. of Bedrooms (X_1)	5	3	2	4
No. of Bathrooms (X_2)	2	2	1	3