$\frac{\partial S6}{\partial c} = -2ab t_i e^{ct_i}\left[b^{...} + e^{ct_i}(H_i - a)\right] / (e^{ct_i}+a)^3$

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

CONTINUOUS ASSESSMENT TEST 2018/2019

STA 2408: REGRESSION MODELLING II

TIME: 1 HOUR

INSTRUCTIONS TO CANDIDATES:

* *Answer question all the questions*

* *Be neat and show all your workings*

1. The height of a child at different ages Estimate the height of the child as an adult of 30

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	69.2	70

years of age using the growth model, $H = \frac{a}{1+be^{-ct}}$. [5 marks]

2. There is a functional relationship between the mass density p of air and the altitude h above the sea level.

Altitude above the sea level, h (km)	0.32	0.64	1.28	1.60
Mass Density, $\rho(kg/m^3)$	1.15	1.10	1.05	0.95

In the regression model $\rho = k_1 e^{-k_2 h}$, the constant k_2 is found as $k_2 = 0.1315$. assuming that the mass density of air at the top of the atmosphere is $\frac{1}{1000^{th}}$ of the mass density of air at the sea level.
Calculate the altitude in kilometers of the top of the atmosphere. [8 marks]

3. For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l$ and $T_{nl}(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

4. Given the data

X	20	30	15	25	28
Y	45 6	35.3	40.3	20.0	43.2

where X and Y are predictor and response variables,respectively.Using a rectangular Kernel function and 3-nearest neighbour(K-NN)find the estimate of

i) Density function of X at point x=29.5 and at x=34 $p(x)$ [3 marks]

ii) regression function at point x=29.5 and at x=34 $m(x)$ [3 marks]

5. Let $X_1, X_2 \in [0,1]$ and

$$\begin{cases} Z = 1, & \text{if } X_1 = 1 \text{ or } X_2 = 1 \\ Z = 0, & \text{if } X_1 = 0 \text{ or } X_2 = 0 \end{cases}$$

be the logical XOR classification to be learned by a perception. The training set then consists of input vectors including the first co-ordinate $x_0 = 1$;

$$X^1 = \begin{pmatrix}1\\1\\0\end{pmatrix}, \quad X^2 = \begin{pmatrix}1\\0\\1\end{pmatrix}, \quad X^3 = \begin{pmatrix}1\\1\\1\end{pmatrix}, \quad X^4 = \begin{pmatrix}1\\0\\0\end{pmatrix}$$

with correction classification $Z^{(1)} = Z^{(2)} = Z^{(3)} = 1, Z^{(4)} = 0$.

The perception with weights w_0, w_1, w_2 classifies an object as 1 if and only if $w_0x_0 + w_1x_1 + w_2x_2 > 0$ and as 0 elsewhere. Taking $w = (0,0,0)^t$ as initial weight and $\eta = 1$ as relaxation factor, train the perception and find the weight that achieves correct classification for all input vectors in the training set. [12marks]

Capsule = -6.3896 + 0.0266 (psa) - 0.0208 (age) + 1.0790 (gleason)

P = e^{capsule} = $\frac{P}{1-P}$ = 0.1149

P = 10.31 %

Odds Ratio of Psa

Exp(0.0266) = 1.0267

C.I = [(0.0266 - 1.96 × 0.00894), (0.0266 + 1.96 × 0.00894)]

$$\frac{\partial SSE}{\partial a} = \sum_{i=1}^{n} \frac{2e^{ct_i}\left[ae^{ct_i} - H_i(e^{ct_i}+b)\right]}{(e^{ct_i}+b)^2}$$

$$\frac{\partial SSE}{\partial b} = \sum_{i=1}^{n} \frac{2ae^{ct_i}\left[bH_i + e^{ct_i}(H_i - a)\right]}{(e^{ct_i}+b)^3}$$

$$\frac{\partial SSE}{\partial b} = \sum_{i=1}^{n} \frac{-2abt_i e^{ct_i}\left[bH_i + e^{ct_i}(H_i - a)\right]}{(e^{ct_i}+b)^3}$$

W-1-2-60-1-6

JOMO KENYATTA UNIVERSITY

110

OF

AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2020/2021

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

THE DEGREE OF BACHELOR OF SCIENCE IN STATISTICS/BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING/ BACHELOR OF SCIENCE IN BIOSTATISTICS/ BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE.

STA 2408: REGRESSION MODELLING II

DATE: AUGUST 2021 TIME: 2 HOURS

INST RUCTIONS TO CANDIDATES:

1. *Answer question ONE and any other two questions*
2. *Be neat and show all your workings*
3. *All questions except question one carry equal marks*

This paper consists of 4 printed pages
STACS Examination board 2020/2021.

A kernel function is a bounded continuous function on all reals, $\mathbb{R}$ satisfying $K_1: \int K(u)du = 1$, $K_2: K(u) \geq 0 \ \forall u$, $K_3: \int uK(u)du > 0$, $K_4: \int u^2 K(u)du < \infty$

Splines are functions which are piecewise polynomial.

smoothing → A non-parametric technique for ~~[illegible]~~ fitting curves where the form of curve is not pre-determined but estimated through data.

QUESTION ONE (30 MARKS)

(a) (i) Differentiate between parametric and non parametric regression. [2 marks]

relies on statistical distributions on data → do not depend on any distribution

(ii) Define the term smoothing with respect to non parametric regression. [2 marks]

(b) Define the following terms giving examples in each case.

non-parametric technique to estimate the conditional expectation of a random variable.

(i) Kernel regression. → e.g. Gauss kernel: $K(u) = \frac{1}{\sqrt{2\pi}} e^{-u^2/2}$ Quartic: $K(u) = \frac{5}{6}(1-m^2)^2$ [3 mks]

(ii) Spline regression. → [3 mks]

$y = ae^{bx}$

$S_r = \sum (y_i - ae^{bx_i})^2$

(c) For an exponential model $y = ae^{bx}$ that is best fit to the data $(x_1, y_1), \ldots, (x_n, y_n)$ derive a non-linear equation that can be used to estimate the value b. [5 mks]

$a = \frac{\sum_{i=1}^{n} y_i e^{bx_i}}{\sum_{i=1}^{n} e^{2bx_i}}$

$\sum y_i x_i e^{bx_i} - \frac{\sum y_i e^{bx_i}}{\sum e^{2bx_i}} \left(\sum_{i=1}^{n} x_i e^{2bx_i}\right) = 0$

(d) Discuss Local polynomial regression also called Locally Weighted Regression (LOWESS) stating its 5-step procedure. [5 mks]

(e) A logistic regression model was fitted to predict Tumor penetration of prostatic Capsule (0 = no penetration, 1 = penetration) using Prostatic Specific (PSA), age, and total gleason score. SAS output was generated as in Figure 1.

predictor

predic

PC =

PSA + Age + Gleas

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	514.289	411.208
SC	518.229	426.969
-2 Log L	512.289	403.208

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	-6.3896	1.4976	18.2045	<.0001
PSA	1	0.0266	0.00894	8.8442	0.0029
AGE	1	-0.0208	0.0188	1.2351	0.2664
GLEASON	1	1.0790	0.1611	44.8373	<.0001

Figure 1: Output

(i) Write the resulting logistic regression equation for the above model. [2mks]

(ii) What is the predicted probability of having capsule=1 for a 69-year old man with a psa level of 10 mg/ml and a gleason score of 5, according to model 1?. [2mks]

(iii) What does the intercept from the model tell you? [1mk]

(f) There exists a functional relationship between the mass density ρ of air and the altitude h above the sea level. A sample data of the two variables is given below

Altitude above sea level, h(km)	0.32	0.64	1.28	1.60
Mass Density, ρ (kg/m^3)	1.15	1.10	1.05	0.95

$= 3.84$

$= 0.2326$

$\ln \rho = \ln K_1 - K_2 h$

$\ln \rho = \ln K_1 - 0.1315h$

$\ln \rho = y \quad \ln K_1 = a_0$

$y = a_0 - 0.1315h$

therefore

$a_0 = \bar{y} + 0.1315\bar{h}$

$a_0 = \frac{0.2326}{4} + 0.1315\left(\frac{3.84}{4}\right)$

$\ln \rho$

$a_0 = 0.18439$

$K_1 = e^{0.18439} = 1.2025$

At sea level, $h = 0$ $\rho = 1.2025$

at the top $P = \frac{1}{1000}(1.2025)$

model $\rho = 1.2025 e^{-0.1315h}$

$1.2025 \times 10^{-3} = 1.2025 e^{-0.1315h}$

$h = \frac{\ln 1.0 \times 10^{-3}}{-0.1315} = 52.53 \text{km}$

The functional relationship can be expressed using the regression model $\overset{y}{\rho} = k_1e^{-k_2h}$, where the constant k_2 is found as $k_2 = 0.1315$. Assuming that the mass density of air at the top of the atmosphere is $1/1000^{\text{th}}$ of the mass density of air at sea level find the altitude in kilometers of the top of the atmosphere. [5 mks]

QUESTION TWO (20 MARKS)

→ relates the variables linearly e.g $y = mx + b$

(a) Differentiate between linear and non-linear models. [2mks]

↳ relates variables in a Non-linear (curved) relationship. e.g $y = x_1 x_2^2$

(b) Many patients get concerned when a test involves injection of a radioactive material. For example for scanning a gallbladder, a few drops of Technetium-99m isotope is used. Half of the technetium-99m would be gone in about 6 hours. It, however, takes about 24 hours for the radiation levels to reach what we are exposed to in day-to-day activities. Below is given the relative intensity of radiation as a function of time.

using bisection method. $\lambda = -0.120$ & $\lambda = -0$

$A = \frac{\sum_{i=1} \rho_i e^{\lambda t_i}}{\sum_{i=1} e^{2\lambda t_i}} = 0$

$\rho = 0.99983\, e^{-0.11508t}$

$f(-0.120) = 6.2501 - \left(\frac{2.9062}{2.3763}\right)(6.0954) = 0.091357$

$f(-0.110) = -0.10099$

$f(-0.120) * f(-0.110) < 0$ → then $\lambda = -0.115$

t (hrs)	0	1	3	5	7	9
ρ	1.000	0.891	0.708	0.562	0.447	0.355

λ can be found $\sum_{i=1} \rho_i t_i e^{\lambda t_i} - \left(\frac{\sum \rho_i e^{\lambda t_i}}{\sum e^{2\lambda t_i}}\right) \sum t_i e^{2\lambda t_i} = 0$ → Draw a table

The correct value of

The level of the relative intensity of radiation is related to time via an exponential model $\rho = Ae^{\lambda t}$, where λ is found as $\lambda = -0.11505$. Find the value of the constant term A, the half-life of Technium-99m, and the radiation intensity after 24 hours. [10 marks]

$f(-0.11508) = 0$

$\lambda = -0.11508$

(ii) $\rho = \frac{1}{2}\rho|_{t=0}$

$t = 6.0232$ years

$A = \frac{\sum \rho_i e^{\lambda t_i}}{\sum e^{2\lambda t_i}} = \frac{2.9373}{2.9378} = 0.99983$

(c) The height of a child is measured at different ages as follows

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	69.2	70

iii) $= 0.99983\, e^{-0.11508(24)}$

$0.99983(e^{-0.11508(6)})$

$= 6.37\%$

Estimate the height of the child as an adult of 30 years of age using the growth model, $H = \frac{a}{1+be^{-ct}}$. [8 marks]

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. Below is given the flow rate, in gallons per minute as a function of pressure.

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

The rate of water flow is related to the nozzle pressure via the regression model $F = \alpha P^{\beta}$. By transforming the above data, find

(a) The value of the regression parameters α and β [5mks]

(b) The rate of water flow if the nozzle pressure is increased to 60 psi [2mks]

(b) State the steps involved in fitting a Local Polynomial Regression Curve [3 mks]

(c) Draw a well labelled diagram of a neural network with two hidden layers and $X_1, X2$ and X_3 as the input variables. [5 mks]

(d) Define what is meant by a spline and illustrate using two examples. Show that splines are linear smoothers. [5 mks]

QUESTION FOUR (20 MARKS)

(a) state and explain the assumptions made by linear mixed models. [6 mks]

(b) Show that a linear regression estimator is a special case of non parametric estimator. [3 mks]

(c) What is a Kernel regression. Illustrate your explanation with examples. [3 mks]

(d) To find contraction of a steel cylinder, one needs to regress the thermal expansion coefficient data to temperature. Consider the following data

Temperature (T)	Thermal expansion Coefficient, α
-320	2.76
-240	3.83
-160	4.72
-80	5.43
0	6.00
80	6.47

If the thermal expansion coefficient of steel behaves as a second order polynomial function of temperature fit the above data to the model $\alpha = a_0 + a_1T + a_2T^2$
[8 Marks]

2c

$$\frac{\partial S_r}{\partial a} = \sum \frac{2e^{ct_i}\left[ae^{ct_i} - H_i\left(e^{ct_i}+b\right)\right]}{\left(e^{ct_i}+b\right)^2}$$

$$\frac{\partial S_r}{\partial b} = \sum \frac{2ae^{ct_i}\left[bH_i + e^{ct_i}\left(H_i - a\right)\right]}{\left(e^{ct_i}+b\right)^3}$$

$$\frac{\partial S_r}{\partial c} = \sum \frac{-2ab\,t_i e^{ct_i}\left[bH_i + e^{ct_i}\left(H_i - a\right)\right]}{\left(e^{ct_i}+b\right)^3}$$

Three guess

(0,20) (12,60) (18,70)

$$20 = \frac{a}{1+be^{-c(0)}}$$

$$60 = \frac{a}{1+be^{-12c}}$$

$$70 = \frac{a}{1+be^{-18c}}$$

By applying newton Raphon

$a = 7.4321 \times 10^1$

$b = 2.8233$ $\quad c = 2.1715 \times 10^{-1}$

$$H = \frac{7.4321 \times 10^1}{1 + 2.8233e^{-2.1715 \times 10^{-1} t}}$$

When $t = 30$

$H_i \simeq 74$

W-1-2-60-1-6

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2018/2019

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

THE DEGREE OF BACHELOR OF SCIENCE IN STATISTICS/BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING/ BACHELOR OF SCIENCE IN BIOSTATISTICS/ BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE.

STA 2408: REGRESSION MODELLING II

DATE: DECEMBER 2018 TIME: 2 HOURS

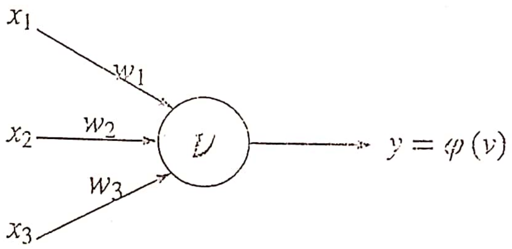
INSTRUCTIONS TO CANDIDATES:

1. *Answer question ONE and any other two questions*
2. *Be neat and show all your workings*
3. *All questions except question one carry equal marks*

This paper consists of 5 printed pages
STACS Examination board 2018/2019.

QUESTION ONE

(a) (i) Below is a diagram of a single artificial neuron (unit)



Flowchart (1).png

The node has three inputs x=(x_1, x_2, x_3) that receive only binary signals (either 0 or 1). How many different input patterns can this code receive? What if the node had four inputs? Five? Can you give a formula that computes the number of binary input patterns for a given number of inputs. [3 marks]

(ii) Suppose that a credit card company decided to deploy a new system for assessing credit worthiness of its customers. The new system is using a feed-forward neural network with a supervised learning algorithm. Suggest in a form of essay what should the bank have before the system can be used? Discuss the problem associated with this requirement. [3 marks]

(b) The height of a child at different ages Estimate the height of the child as an adult of 30

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	69.2	70

years of age using the growth model. $H = \frac{a}{1+be^{-at}}$. [5 marks]

(c) There is a functional relationship between the mass density p of air and the altitude h above the sea level.

Altitude above the sea level, h (km)	0.32	0.64	1.28	1.60
Mass Density. $\rho(kg/m^3)$	1.15	1.10	1.05	0.95

In the regression model $\rho = k_1 e^{-k_2 h}$. the constant k_2 is found as $k_2 = 0.1315$. assuming that the mass density of air at the top of the atmosphere is $\frac{1}{1000^{th}}$ of the mass density of air at the sea level.

Calculate the altitude in kilometers of the top of the atmosphere. [8 marks]

(d) For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)'$ and $T_{nl}(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l = ,0, 1, 2$ as 22.8038, −0.5758, 1.1189 respectively, and $T_{nl}(x_0), l = 0, 1$ as 24.0092, 0.1631 respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

(c) The following result are from a perspective study that were considered in building a logistic regression model for predicting capsule=1 that included psa, age, and gleason in the model(model 1). Part of the resulting SAS output follows:

Figure 1: Model 1

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	514.289	411.208
SC	518.229	426.959
-2 Log L	512.289	403.208

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	-6.3896	1.4976	18.2045	<.0001
psa	1	0.0265	0.00894	8.8442	0.0029
age	1	-0.0203	0.0183	1.2351	0.2664
gleason	1	1.0790	0.1611	44.8373	<.0001

(i) Write the resulting logistic regression equation for model 1 [1 mark]

(ii) What is the predicted probability of having a capsule=1 for a 69-year old man with psa level of 10mg/ml and a gleason score of 5, according to model 1? [1 mark]

(iii) What does the intercept from the model tell you? [1 mark]

(iv) Calculate the odds ratio and 95% confidence interval for psa from the model. Interpret. [1 mark]

QUESTION TWO (20 MARKS)

(a) In Kernel regression estimation,one may choose a deterministic or stochastic design depending on the type of the problem at hand.

(i) Describe a deterministic design model and give its kernel estimate of the mean function. [3 marks]

(ii) Describe a stochastic design model and give its kernel estimate of the mean function [3 marks]

(b) A steel cylinder at 80^0F of length 12" is placed in a commercially available liquid nitrogen bath (-315^0F). If the thermal expansion coefficient of steel behaves as a second order polynomial function of temperature and the polynomial is found by regressing the data below.

Temperature, $T(^0F)$ X	Thermal expansion Coefficient, α Y ($\mu\ in/in/\ ^0F$)
-320	2.76
-240	3.83
-160	4.72
-80	5.43
0	6.00
80	6.47

Y X X^2

(i) Fit the data to $\alpha = a_0 + a_1T + a_2T^2$. [10 marks]

(ii) Calculate the reduction in the length of the cylinder in inches. [4 marks]

QUESTION THREE (20 MARKS)

(a) Given the data

X	20	30	15	25	28
Y	45.6	35.3	40.3	20.0	43.2

where X and Y are predictor and response variables,respectively.Using a rectangular Kernel function and 3-nearest neighbour(K-NN)find the estimate of

i) Density function of X at point x=29 [3 marks]

ii) regression function at point x=29 [3 marks]

(b) We have a sample $(X_i, Y_i), i = 1, 2, \cdots, n$ generated from the NPR model $Y_i = m(X_i) + \epsilon_i$, $i = 1, 2, \cdots, n$ where $m(x)$ is an unknown smooth function and $X_1 < X_2 <, \cdots, < X_n$. Let $Y = [Y_1, \cdots, Y_n]^T$ be the response vector and $\hat{Y} = [\hat{Y}_1, \cdots, \hat{Y}_n]^T$ be the estimated response vector, where $\hat{Y}_i = \hat{m}(X_i)$, $i = 1, 2, \cdots, n$ for some non-parametric estimator $\hat{m}(x)$. When $\hat{m}(x)$ is a linear smoother, we have $\hat{Y} = AY$ where A is known as the associated smoother matrix and df=trace(A) is known as the associated degrees of freedom, measuring how complex the fitting model is.

(i) First assume that $\hat{m}(x)$ is the usual regression spline smoother constructed based on the p-th order truncated power basis $\Phi(x)$ using K distinct knots $\tau_1 < \tau_2 < \cdots \tau_K$. When $n > K + p$ and p fixed, show that df will increase as K increasing. [2 marks]

(ii) Assume now that $\hat{m}(x)$ is the N-W estimator using a bandwidth $h > 0$ and a symmetric kernel $K(.)$ which is a pdf. Show that when K is fixed and n is sufficiently large, df will decrease as h increasing. [3 marks]

(iii) Assume now that $\hat{m}(x)$ is the cubic smoothing spline smoother with a smoothing parameter λ. Show that df will decrease as λ increasing. [3 marks]

(iv) Assume now that $\hat{m}(x)$ is the P-spline smoother with p-th order truncated power basis $\Phi(x)$ using K distinct knots $\tau_1 < \tau_2 < \cdots < \tau_K$ and a smoothing parameter λ. Show that when $n > p + K$, df will decrease as λ increases. [6 marks]

QUESTION FOUR (20 MARKS)

(a) Let $X_1, X_2 \quad \epsilon \ [0, 1]$ and

$$\begin{cases} Z = 1, & \text{if } X_1 = 1 \text{ or } X_2 = 2 \\ Z = 0, & \text{if } X_1 = 0 \text{ or } X_2 = 0 \end{cases}$$

be the logical XOR classification to be learned by a perception. The training set then consists of input vectors including the first co-ordinate $x_0 = 1$;

$$X^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad X^2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad X^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad X^4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

with correction classification$Z^{(1)} = Z^{(2)} = Z^{(3)} = 1, Z^{(4)} = 0$.

The perception with weightsw_0, w_1, w_2 classifies an object as 1 if and only if $w_0x_0 + w_1x_1 + w_2x_2 > 0$ and as 0 elsewhere.Taking $w = (0, 0, 0)^t$ as initial weight and $\eta = 1$ as relaxation factor,train the perception and find the weight that achieves correct classification for all input vectors in the training set. [12marks]

1(b) We have a sample (X_i, Y_i), $i = 1, 2, \cdots, n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$ where $E(\epsilon_i|X_1, \cdots, X_n) = 0$ and $E(\epsilon_i^2|X_1, \cdots, X_n) = \sigma^2(X_i)$ and $X_1, X_2, \cdots, X_n$ has a pdf $f(x)$. Moreover, $E(\epsilon_i\epsilon_j|X_1, \cdots, X_n) = 0$ for $i \neq j$.Define

$$C_1(h) = -\frac{2}{n}\sum_{j=1}^{n} \epsilon_j[\hat{m}_h(X_j) - m(X_j)], \quad C_2(h) = -\frac{2}{n}\sum_{j=1}^{n} \epsilon_j[\hat{m}_h^{(-j)}(X_j) - m(X_j)]$$

Where $[\hat{m}_h(X_j)$ is the N-W estimator of $m(x)$ at X_j and $\hat{m}_h^{(-j)}$ is the N-W of $m(x)$ at X_j obtained using all the data except (X_j, Y_j).

Find the asymptotic expression of $E(C_1(h)|X_1, \cdots, X_n)$ [8mks]

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

CONTINUOUS ASSESSMENT TEST 2018/2019

STA 2408: REGRESSION MODELLING II

TAKE AWAY ASSIGNMENT

INSTRUCTIONS TO CANDIDATES:

* *Answer question all the questions*

* *Be neat and show all your workings*

1. For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l$ and $T_{nl}(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

2. For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l$ and $T_{nl}(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

CONTINUOUS ASSESSMENT TEST 2018/2019

STA 2408: REGRESSION MODELLING II

TIME: 1 HOUR

INSTRUCTIONS TO CANDIDATES:

* *Answer question all the questions*

* *Be neat and show all your workings*

1. The height of a child at different ages Estimate the height of the child as an adult of 30

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	69.2	70

years of age using the growth model, $H = \frac{a}{1+be^{-ct}}$. [5 marks]

2. There is a functional relationship between the mass density ρ of air and the altitude h above the sea level.

Altitude above the sea level, h (km)	0.32	0.64	1.28	1.60
Mass Density, $\rho(kg/m^3)$	1.15	1.10	1.05	0.95

In the regression model $\rho = k_1e^{-k_2h}$, the constant k_2 is found as $k_2 = 0.1315$. assuming that the mass density of air at the top of the atmosphere is $\frac{1}{1000^{th}}$ of the mass density of air at the sea level.
Calculate the altitude in kilometers of the top of the atmosphere. [8 marks]

3. For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l$ and $T_{nl}(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

4. Given the data

X	20	30	15	25	28
Y	45.6	35.3	40.3	20.0	43.2

where X and Y are predictor and response variables,respectively.Using a rectangular Kernel function and 3-nearest neighbour(K-NN)find the estimate of

$\hat{p}(x) = \frac{K-1}{2n\, d_x(x)}$

$\hat{m}_2(29) = \frac{1}{K-1}\sum_{i=1}^{n} Y_i \Pi(x - d_i ...)$

i) Density function of X at point x=29.5 and at x=34 [3 marks]

ii) regression function at point x=29.5 and at x=34 [3 marks]

5. Let $X_1, X_2 \in [0,1]$ and

$$\begin{cases} Z = 1, & \text{if } X_1 = 1 \text{ or } X_2 = 1 \\ Z = 0, & \text{if } X_1 = 0 \text{ or } X_2 = 0 \end{cases}$$

be the logical XOR classification to be learned by a perception. The training set then consists of input vectors including the first co-ordinate $x_0 = 1$;

$$X^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}; \quad X^2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad X^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad X^4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

with correction classification $Z^{(1)} = Z^{(2)} = Z^{(3)} = 1, Z^{(4)} = 0$.

The perception with weightsw_0, w_1, w_2 classifies an object as 1 if and only if $w_0x_0 + w_1x_1 + w_2x_2 > 0$ and as 0 elsewhere.Taking $w = (0,0,0)^t$ as initial weight and $\eta = 1$ as relaxation factor,train the perception and find the weight that achieves correct classification for all input vectors in the training set. [12marks]

W-1-2-60-1-6

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2020/2021

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

THE DEGREE OF BACHELOR OF SCIENCE IN STATISTICS/BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING/ BACHELOR OF SCIENCE IN BIOSTATISTICS/ BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE.

STA 2408: REGRESSION MODELLING II

DATE: AUGUST 2021 TIME: 2 HOURS

INST RUCTIONS TO CANDIDATES:

1. *Answer question ONE and any other two questions*
2. *Be neat and show all your workings*
3. *All questions except question one carry equal marks*

This paper consists of 4 printed pages
STACS Examination board 2020/2021.

QUESTION ONE (30 MARKS)

(a) (i) Differentiate between parametric and non parametric regression [2 marks]

(ii) Define the term smoothing with respect to non parametric regression. [2 marks]

(b) Define the following terms giving examples in each case.

(i) Kernel regression. [3 mks]

(ii) Spline regression. [3 mks]

(c) For an exponential model $y = ae^{bx}$ that is best fit to the data $(x_1, y_1), ..., (x_n, y_n)$ derive a non-linear equation that can be used to estimate the value b. [5 mks]

(d) Discuss Local polynomial regression also called Locally Weighted Regression (LOWESS) stating its 5-step procedure. [5 mks]

(e) A logistic regression model was fitted to predict Tumor penetration of prostatic Capsule (0 = no penetration, 1 = penetration) using Prostatic Specific (PSA), age, and total gleason score. SAS output was generated as in Figure 1.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	514.289	411.208
SC	518.229	426.969
-2 Log L	512.289	403.208

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	-6.3896	1.4976	18.2045	<.0001
PSA	1	0.0266	0.00894	8.8442	0.0029
AGE	1	-0.0208	0:0188	1.2351	0.2664
GLEASON	1	1.0790	0.1611	44.8373	<.0001

Figure 1: Output

(i) Write the resulting logistic regression equation for the above model. [2mks]

(ii) What is the predicted probability of having capsule=1 for a 69-year old man with a psa level of 10 mg/ml and a gleason score of 5, according to model 1?. [2mks]

(iii) What does the intercept from the model tell you? [1mk]

(f) There exists a functional relationship between the mass density ρ of air and the altitude h above the sea level. A sample data of the two variables is given below

Altitude above sea level, h(km)	0.32	0.64	1.28	1.60
Mass Density, ρ (kg/m^3)	1.15	1.10	1.05	0.95

he functional relationship can be expressed using the regression model $\rho = k_1 e^{-k_2 h}$, where the constant k_2 is found as $k_2 = 0.1315$. Assuming that the mass density of air at the top of the atmosphere is $1/1000^{th}$ of the mass density of air at sea level find the altitude in kilometers of the top of the atmosphere. [5 mks]

QUESTION TWO (20 MARKS)

(a) Differentiate between linear and non-linear models. [2mks]

(b) Many patients get concerned when a test involves injection of a radioactive material.For example for scanning a gallbladder, a few drops of Technetium-99m isotope is used. Half of the technetium-99m would be gone in about 6 hours. It, however, takes about 24 hours for the radiation levels to reach what we are exposed to in day-to-day activities. Below is given the relative intensity of radiation as a function of time.

t (hrs)	0	1	3	5	7	9
ρ	1.000	0.891	0.708	0.562	0.447	0.355

The level of the relative intensity of radiation is related to time via an exponential model $\rho = Ae^{\lambda t}$, where λ is found as $\lambda = 0.11505$. Find the value of the constant term A, the half-life of Technium-99m, and the radiation intensity after 24 hours. [10 marks]

(c) The height of a child is measured at different ages as follows

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	69.2	70

Estimate the height of the child as an adult of 30 years of age using the growth model, $H = \frac{a}{1+be^{-ct}}$. [8 marks]

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. Below is given the flow rate, in gallons per minute as a function of pressure.

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

The rate of water flow is related to the nozzle pressure via the regression model $F = \alpha P^{\beta}$. By transforming the above data, find

(a) The value of the regression parameters α and β [5mks]

(b) The rate of water flow if the nozzle pressure is increased to 60 psi [2mks]

(b) State the steps involved in fitting a Local Polynomial Regression Curve [3 mks]

(c) Draw a well labelled diagram of a neural network with two hidden layers and $X_1, X2$ and X_3 as the input variables. [5 mks]

(d) Define what is meant by a spline and illustrate using two examples. Show that splines are linear smoothers. [5 mks]

linear
cubic splines.

QUESTION FOUR (20 MARKS)

(a) state and explain the assumptions made by linear mixed models. [6 mks]

(b) Show that a linear regression estimator is a special case of non parametric estimator. [3 mks]

(c) What is a Kernel regression. Illustrate your explanation with examples. [3 mks]

(d) To find contraction of a steel cylinder, one needs to regress the thermal expansion coefficient data to temperature. Consider the following data

Temperature (T)	Thermal expansion Coefficient, α
-320	2.76
-240	3.83
-160	4.72
-80	5.43
0	6.00
80	6.47

If the thermal expansion coefficient of steel behaves as a second order polynomial function of temperature fit the above data to the model $\alpha = a_0 + a_1 T + a_2 T^2$
[8 Marks]

W-1-2-60-1-6

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2018/2019

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

THE DEGREE OF BACHELOR OF SCIENCE IN STATISTICS/BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING/ BACHELOR OF SCIENCE IN BIOSTATISTICS/ BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE.

STA 2408: REGRESSION MODELLING II

DATE: DECEMBER 2018 TIME: 2 HOURS

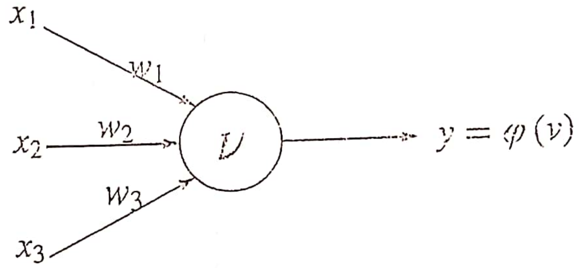
INSTRUCTIONS TO CANDIDATES:

1. *Answer question ONE and any other two questions*
2. *Be neat and show all your workings*
3. *All questions except question one carry equal marks*

This paper consists of 5 printed pages
STACS Examination board 2018/2019.

QUESTION ONE

(a) (i) Below is a diagram of a single artificial neuron (unit)



Flowchart (1).png

The node has three inputs x=(x_1, x_2, x_3) that receive only binary signals (either 0 or 1). How many different input patterns can this code receive? What if the node had four inputs? Five? Can you give a formula that computes the number of binary input patterns for a given number of inputs. [3 marks]

(ii) Suppose that a credit card company decided to deploy a new system for assessing credit worthiness of its customers. The new system is using a feed-forward neural network with a supervised learning algorithm. Suggest in a form of essay what should the bank have before the system can be used? Discuss the problem associated with this requirement. [3 marks]

(b) The height of a child at different ages Estimate the height of the child as an adult of 30

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	63.2	70

years of age using the growth model. $H = \frac{a}{1+be^{-ct}}$. [5 marks]

(c) There is a functional relationship between the mass density p of air and the altitude h above the sea level.

Altitude above the sea level, h (km)	0.32	0.64	1.28	1.60
Mass Density, $\rho(kg/m^3)$	1.15	1.10	1.05	0.95

In the regression model $\rho = k_1 e^{-k_2 h}$, the constant k_2 is found as $k_2 = 0.1315$. assuming that the mass density of air at the top of the atmosphere is $\frac{1}{1000^{th}}$ of the mass density of air at the sea level.
Calculate the altitude in kilometers of the top of the atmosphere. [8 marks]

(d) For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^n K_h(X_i - x)(X_i - x)'$ and $T_{nl}(x) = \sum_{i=1}^n K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

(e) The following result are from a perspective study that were considered in building a logistic regression model for predicting capsule=1 that included psa, age, and gleason in the model(model 1). Part of the resulting SAS output follows:

Figure 1: Model 1

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	514.289	411.208
SC	518.229	426.959
-2 Log L	512.289	403.208

$y = a + bn_i$

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	-6.3896	1.4976	18.2045	<.0001
psa	1	0.0266	0.00894	8.8442	0.0029
age	1	-0.0208	0.0188	1.2351	0.2664
gleason	1	1.0790	0.1611	44.8373	<.0001

$y = mx + 2u + \varepsilon$

(i) Write the resulting logistic regression equation for model 1 [1 mark]

(ii) What is the predicted probability of having a capsule=1 for a 69-year old man with psa level of 10mg/ml and a gleason score of 5, according to model 1? [1 mark]

(iii) What does the intercept from the model tell you? [1 mark]

(iv) Calculate the odds ratio and 95% confidence interval for psa from the model. Interpret. [1 mark]

QUESTION TWO (20 MARKS)

(a) In Kernel regression estimation,one may choose a deterministic or stochastic design depending on the type of the problem at hand.

(i) Describe a deterministic design model and give its kernel estimate of the mean function. [3 marks]

(ii) Describe a stochastic design model and give its kernel estimate of the mean function [3 marks]

(b) A steel cylinder at 80^0F of length 12" is placed in a commercially available liquid nitrogen bath (-315^0F). If the thermal expansion coefficient of steel behaves as a second order polynomial function of temperature and the polynomial is found by regressing the data below.

Temperature, $T(^0F)$	Thermal expansion Coefficient, α ($\mu\ in/in/\ ^0F$)
-320	2.76
-240	3.83
-160	4.72
-80	5.43
0	6.00
80	6.47

(i) Fit the data to $\alpha = a_0 + a_1T + a_2T^2$. [10 marks]

(ii) Calculate the reduction in the length of the cylinder in inches. [4 marks]

QUESTION THREE (20 MARKS)

(a) Given the data

X	20	30	15	25	28
Y	45.6	35.3	40.3	20.0	43.2

where X and Y are predictor and response variables,respectively.Using a rectangular Kernel function and 3-nearest neighbour(K-NN)find the estimate of

i) Density function of X at point x=29 [3 marks]

ii) regression function at point x=29 [3 marks]

(b) We have a sample $(X_i, Y_i), i = 1, 2, \cdots, n$ generated from the NPR model $Y_i = m(X_i) + \epsilon_i$, $i = 1, 2, \cdots, n$ where $m(x)$ is an unknown smooth function and $X_1 < X_2 <, \cdots, < X_n$. Let $Y = [Y_1, \cdots, Y_n]^T$ be the response vector and $\hat{Y}=[\hat{Y}_1, \cdots, \hat{Y}_n]^T$ be the estimated response vector, where $\hat{Y}_i = \hat{m}(X_i)$, $i = 1, 2, \cdots, n$ for some non-parametric estimator $\hat{m}(x)$. When $\hat{m}(x)$ is a linear smoother, we have $\hat{Y} = AY$ where A is known as the associated smoother matrix and df=trace(A) is known as the associated degrees of freedom, measuring how complex the fitting model is.

(i) First assume that $\hat{m}(x)$ is the usual regression spline smoother constructed based on the p-th order truncated power basis $\Phi(x)$ using K distinct knots $\tau_1 < \tau_2 < \cdots \tau_K$. When $n > K - p$ and p fixed, show that df will increase as K increasing. [2 marks]

(ii) Assume now that $\hat{m}(x)$ is the N-W estimator using a bandwidth $h > 0$ and a symmetric kernel $K(.)$ which is a pdf. Show that when K is fixed and n is sufficiently large, df will decrease as h increasing. [3 marks]

(iii) Assume now that $\hat{m}(x)$ is the cubic smoothing spline smoother with a smoothing parameter λ. Show that df will decrease as λ increasing. [3 marks]

(iv) Assume now that $\hat{m}(x)$ is the P-spline smoother with p-th order truncated power basis $\Phi(x)$ using K distinct knots $\tau_1 < \tau_2 < \cdots < \tau_K$ and a smoothing parameter λ. Show that when $n > p + K$, df will decrease as λ increases. [6 marks]

2

10
2

QUESTION FOUR (20 MARKS)

(a) Let $X_1, X_2 \quad \epsilon \ [0, 1]$ and

done as example. Kusi's notes

$$\begin{cases} Z = 1, & \text{if } X_1 = 1 \text{ or } X_2 = 2 \\ Z = 0, & \text{if } X_1 = 0 \text{ or } X_2 = 0 \end{cases}$$

be the logical XOR classification to be learned by a perception. The training set then consists of input vectors including the first co-ordinate $x_0 = 1$;

$$X^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad X^2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad X^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad X^4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

with correction classification$Z^{(1)} = Z^{(2)} = Z^{(3)} = 1, Z^{(4)} = 0$.

The perception with weightsw_0, w_1, w_2 classifies an object as 1 if and only if $w_0x_0 + w_1x_1 + w_2x_2 > 0$ and as 0 elsewhere.Taking $w = (0, 0, 0)^t$ as initial weight and $\eta = 1$ as relaxation factor,train the perception and find the weight that achieves correct classification for all input vectors in the training set. [12marks]

(b) We have a sample (X_i, Y_i), $i = 1, 2, \cdots, n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$ where $E(\epsilon_i|X_1, \cdots, X_n) = 0$ and $E(\epsilon_i^2|X_1, \cdots, X_n) = \sigma^2(X_i)$ and $X_1, X_2, \cdots, X_n$ has a pdf $f(x)$. Moreover, $E(\epsilon_i\epsilon_j|X_1, \cdots, X_n) = 0$ for $i \neq j$.Define

$$C_1(h) = -\frac{2}{n}\sum_{j=1}^{n} \epsilon_j[\hat{m}_h(X_j) - m(X_j)], \quad C_2(h) = -\frac{2}{n}\sum_{j=1}^{n} \epsilon_j[\hat{m}_h^{(-j)}(X_j) - m(X_j)]$$

Where $[\hat{m}_h(X_j)$ is the N-W estimator of $m(x)$ at X_j and $\hat{m}_h^{(-j)}$ is the N-W of $m(x)$ at X_j obtained using all the data except (X_j, Y_j).

Find the asymptotic expression of $E(C_1(h)|X_1, \cdots, X_n)$ [8mks]

$[1, 1, 0] + [0 \ -1] \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$

$[1 \ 1 \ 0] + (-1 \ 0 \ 0) = \begin{bmatrix} \cancel{0} \\ \cancel{1} \\ \cancel{0} \end{bmatrix} \quad [0 \ 1 \ 0]^T$

12
20
10
42

42
23
65

$1+1, \ \cancel{1}\times 0, \ 6+0$

0 1

12

0.0655360,000

W-1-2-60-1-6

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2019/2020

FOURTH YEAR FIRST SEMESTER EXAMINATION FOR

THE DEGREE OF BACHELOR OF SCIENCE IN STATISTICS/BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING/ BACHELOR OF SCIENCE IN BIOSTATISTICS/ BACHELOR OF SCIENCE IN ACTUARIAL SCIENCE.

STA 2408: REGRESSION MODELLING II

DATE: DECEMBER 2019 TIME: 2 HOURS

INSTRUCTION:

1. *Answer question ONE and any other two questions*
2. *Be neat and show all your workings*
3. *All questions except question one carry equal marks*

QUESTION ONE (30 MARKS)

(a) (i) Differentiate between parametric and non parametric regression. [2 marks]

(ii) Differentiate between kernel regression and neural network regression. [2 marks]

(b) Discuss Local polynomial regression also called Locally Weighted Regression (LOWESS) stating its 5-step procedure. [5 marks]

(c) A logistic regression model was fitted to predict Tumor penetration of prostatic Capsule (0 = no penetration, 1 = penetration) using Prostatic Specific (PSA), age, and total gleason score. SAS output was generated as shown below.

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	514.289	411.208
SC	518.229	426.969
-2 Log L	512.289	403.208

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	-6.3896	1.4976	18.2045	<.0001
PSA	1	0.0266	0.00894	8.8442	0.0029
AGE	1	-0.0208	0.0188	1.2351	0.2664
GLEASON	1	1.0790	0.1611	44.8373	<.0001

(i) Write the resulting logistic regression equation for the above model. [2 marks]

(ii) What is the predicted probability of having capsule=1 for a 69-year old man with a psa level of 10 mg/ml and a gleason score of 5?. [2 marks]

(iii) What does the intercept from the model tell you? [1 mark]

(d) K-NN is a weighted average in a varying neighborhood(defined through those X variables which are among the K nearest neighbors of X in Euclidean distance (linear real distance)). Let $(X_i, Y_i)_{i=1}^{5}$ be $(1,5), (7,12), (3,1), (2,0), (5,4)$ Compute (K-NN estimate) $\hat{m}_k(x)$ (i.e) for $x = 4$ and $k = 3$. [5 marks]

(e) State and explain the assumptions made by linear mixed models. [6 marks]

(f) Show that a linear regression estimator is a special case of non parametric estimator. [5 marks]

QUESTION TWO (20 MARKS)

(a) Differentiate between linear and non-linear models. [2 marks]

(b) Many patients get concerned when a test involves injection of a radioactive material.For example for scanning a gallbladder, a few drops of Technetium-99m isotope is used. Half of the technetium-99m would be gone in about 6 hours. It, however, takes about 24 hours for the radiation levels to reach what we are exposed to in day-to-day activities. Below is given the relative intensity of radiation as a function of time.

t (hrs)	0	1	3	5	7	9
ρ	1.000	0.891	0.708	0.562	0.447	0.355

In the regression model $\rho = Ae^{\lambda t}$. Find the value of the constant terms A and λ, the half-life of Technium-99m, and the radiation intensity after 24 hours. [10 marks]

(c) The height of a child at different ages is recorded in the table below. Estimate the height

$b = 1.5572 \times 10^{-1}$

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	69.2	70

of the child as an adult of 30 years of age using the growth model, $H = \frac{a}{1+be^{-ct}}$. [8 marks]

QUESTION THREE (20 MARKS)

(a) For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l$ and $T_{nl}(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

(b) Draw a well labelled diagram of a neural network with two hidden layers and $X_1, X2$ and X_3 as the input variables. [5 marks]

(c) Define what is meant by a spline and illustrate using an examples. Show that splines are linear smoothers. [8 marks]

1.704

QUESTION FOUR (20 MARKS)

(a) In Kernel regression estimation , one may choose a deterministic or a stochastic design depending on the type of problem at hand

i) Describe a deterministic design model and give its kernel estimate of the mean functon. [3 marks]

ii) Describe a ~~deterministic~~ Stochastic design model and give its kernel estimate of the mean functon. [3 marks]

(b) Given the data

X	20	30	15	25	28
Y	45.6	35.3	40.3	20.0	43.2

where X and Y are predictor and response variables,respectively.Using a rectangular Kernel function and 3-nearest neighbour(K-NN)find the estimate of

i) Density function of X at point x=29.5 and at x=34 [3 marks]

ii) Regression function at point x=29.5 and at x=34 [3 marks]

(c) We have a sample $(X_i, Y_i),\ i = 1, 2, \cdots, n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$ where $E(\epsilon_i|X_1, \cdots, X_n) = 0$ and $E(\epsilon_i^2|X_1, \cdots, X_n) = \sigma^2(X_i)$ and $X_1, X_2, \cdots, X_n$ has a pdf $f(x)$. Moreover, $E(\epsilon_i\epsilon_j|X_1, \cdots, X_n) = 0$ for $i \neq j$.Define

$$C_1(h) = -\frac{2}{n}\sum_{j=1}^{n} \epsilon_j[\hat{m}_h(X_j) - m(X_j)], \quad C_2(h) = -\frac{2}{n}\sum_{j=1}^{n} \epsilon_j[\hat{m}_h^{(-j)}(X_j) - m(X_j)]$$

Where $[\hat{m}_h(X_j)$ is the N-W estimator of $m(x)$ at X_j and $\hat{m}_h^{(-j)}$ is the N-W of $m(x)$ at X_j obtained using all the data except (X_j, Y_j).
Find the asymptotic expression of $E(C_1(h)|X_1, \cdots, X_n)$ [8mks]

Scanned with CamScanner

W1-2-60-1-6

JOMO KENYATTA UNIVERSITY
OF
AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2017/2018

FOURTH YEAR FIRST SEMESTER EXAMINATIONS FOR THE DEGREE OF BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING

STA 2408: REGRESSION MODELLING II

DATE: JANUARY 2018 TIME: 2 HOURS

INSTRUCTIONS TO CANDIDATES:

1. *Answer question ONE (section A) and any other two questions in section B.*
2. *Be neat and show all your workings*
3. *All questions except question one carry equal marks*

This paper consists of 4 printed pages.

3

SECTION A (30 MARKS)

1. (a) For the regression model

$$Y = X\beta + \epsilon$$

with $\epsilon \sim N(0, \sigma^2 W)$ where W is not an identity matrix, a student argues that the ordinary multiple regression cannot be applied and one should regress $W^{-1}Y$ against $W^{-1}X$ by ordinary multiple regression to estimate β. Do you agree with the student? If so, show how this can be done [6 marks]

(b) An Actuarial Science student is trying to analyze data from a particular stock using an equation of the form $y = f(x) = \frac{a_0}{x} + \frac{a_1}{x^2}$. The measured values of (x, y) are listed in the table below

x	1	2	3	4	5	6	7	8	9	10
y	3	7	0.9	6.8	0.7	0.96	19	37	0.8	0.1

Use nonlinear regression method to determine a_0 and a_1 [7 marks]

(c) Define the term Artificial Neural Network (ANN) [2 marks]

(d) Stock prices (Y, in dollars) are assumed to be affected by the annual rate of dividends of stock (X). A simple linear regression analysis ($Y_i = \beta_o + \beta_1 X_i + \epsilon_i$) was performed on 21 observations where ϵ_i were assumed to follow $\sim iidN(0, \sigma^2)$ and summary statistics were as listed below

$$\bar{X} = 0.4, \bar{Y} = 4, S_X^2 = 0.1, S_Y^2 = 20, S_{XY} = 1.25$$

Perform an ANOVA analysis for the level of significance 5% and give your conclusions[7 marks]

(e) Let $Y = (9, 4, 1, 3)^\top$, $X_1 = (9, 3, 9, 4)^\top$, $X_2 = (5, 3, 5, 4)^\top$

(i) Find the matrix $(X^\top X)^{-1}$ [2 marks]

(ii) Fit a regression model $Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \epsilon$. Give the fitted coefficients and estimate of error variance [4 marks]

(iii) Construct the 95% confidence interval for $E(Y|X_1 = 6, X_2 = 4)$ [3 marks]

2. In a study of the effect of hormones on the productivity of a certain variety of tomato plant, ten plants were treated at different hormone strengths and their yields, in Kgs, noted. A simple linear regression model is fitted in R, with the depended variable (yield in kgs) stored as the vector y and the independent variable (hormone strength) stored as the vector x. Some of the R commands and edited output are shown below

```
> lm1 = lm(y~x)
> summary(lm1)

Coefficients:
              Estimate    Std. Error
(Intercept) -3.72335     0.28072
```

```
x            0.42999     0.01899

> qt(0.975,8)
[1] 2.306004
```

(a) Write down the equation of the model that has been fitted to the data, defining your notations carefully. State the distribution of any error terms in your model [5 marks]

(b) If X is the design matrix for the model fitted in R, then

$$(X^\top X) = \begin{pmatrix} 10.60278 & 154.047 \\ 154.047 & 2320.782 \end{pmatrix}$$

Also,

$$\sum_{i=1}^{10} y_i = 26.76 \qquad \sum_{i=1}^{10} x_i y_i = 424.33$$

where y_i is the ith yield and x_i is the ith hormone strength. Give suitable calculations that show how the parameter estimates have been obtained in the R output [5 marks]

(c) If $x_3 = 11$ and $y_3 = 1.66$, calculate the corresponding fitted value and residual [4 marks]

(d) Test the hypothesis that there is no relationship between hormone strength and yield, stating your conclusion clearly. State the size of your hypothesis test [4 marks]

(e) Calculate the 95% confidence interval for the intercept and gradient in the regression model [2 marks]

3. (a) Consider data from the simple linear model $y_i = \beta_o + \beta_1 x_i + \epsilon_i$, $i = 1, 2, \cdots, n$ where x_i's are fixed constants, β_o, β_1 are the unknown coefficients and ϵ_i's are unobserved i.i.d random variables from $N(0, \sigma^2)$. Let $\widehat{\beta}_0$ and $\widehat{\beta}_1$ be the least squares estimators of β_0 and β_1 respectively. Find the distribution of the vector $(\widehat{\beta}_0, \widehat{\beta}_1)$ [6 marks]

(b) Suppose the model is modified by including more independent variables so that the model is now written as $Y = X\beta + \epsilon$ and assuming same assumptions for the error model distribution.

(i) Show that the vector covariance of residuals $e = y - X\widehat{\beta}$ and y is given by $cov(e, y) = \sigma^2 M$ where M is the idempotent and symmetric matrix [4 marks]

(ii) Suppose the variance of the error term (σ^2) is unknown, demonstrate how such can be estimated from the model given in b(i) above [7 marks]

(c) Suppose we have a data set (x_i, y_i) for $i = 1, 2, \cdots, n$. Consider two different models $y_i = \alpha + \beta x_i^2 + \epsilon_i$ and $y_i = \alpha + \beta x_i^2 + \gamma \exp^{x_i} + \epsilon_i$. Compare the residual sum of squares of the two models. Explain [3 marks]

4. The data given in the following table are the numbers of deaths from AIDS in Kenya for 12 consecutive quarters starting from the second quarter of 1998

Quarter (i)	1	2	3	4	5	6	7	8	9	10	11	12
Number of deaths (n_i)	1	2	3	1	4	9	18	23	31	20	25	37

(a) (i) Draw a scatterplot of the data

(ii) Comment on the nature of the relationship between the number of deaths and the quarter in this early phase of the epidemic [4 marks]

(b) A statistician has suggested that a model of the form

$$E\left[N_i\right] = \gamma i^2$$

might be appropriate for these data, where γ is a parameter to be estimated form the above data. She has proposed two methods for estimating γ given in (i) and (ii) below

(i) Show that the least squares estimate of γ, is obtained by minimizing $q = \sum_{i=1}^{12}\left(n_i - \gamma i^2\right)^2$ is given by

$$\widehat{\gamma} = \frac{\sum_{i=1}^{12} i^2 n_i}{\sum_{i=1}^{12} i^4}$$

(ii) Show that an alternative (weighted) least squares estimate of γ obtained by minimizing $q^* = \sum_{i=1}^{12} \frac{\left(n_i - \gamma i^2\right)^2}{i^2}$ is given by

$$\tilde{\gamma} = \frac{\sum_{i=1}^{12} n_i}{\sum_{i=1}^{12} i^2}$$

(c) Noting that $\sum_{i=1}^{12} i^4 = 60,710$ and $\sum_{i=1}^{12} i^2 = 650$, calculate $\widehat{\gamma}$ and $\tilde{\gamma}$ for the above data[8 marks]

(d) To asses whether the single parameter model which was used in part (b) is appropriate for the data, a two parameter model is now considered. The model is of the form $E\left(N_i\right) = \gamma i^{\theta}$ for $i = 1, 2, \cdots, 12$

(i) To estimate the parameters γ and θ, a simple linear regression of the form $E\left(Y_i\right) = \alpha + \beta x_i$ is used , where $x_i = \log(i)$ and $Y_i = \log(N_i)$ for $i = 1, 2, \cdots, 12$. Relate the parameters γ and θ to the regression parameters α and β

(ii) The least squares estimates of α and β are -0.6112 and 1.6008 with standard errors 0.4586 and 0.2525 respectively. Using the value for the estimate of β, conduct a formal statistical test to asses whether the form of the model suggested in (b) is adequate [8 marks]

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

CONTINUOUS ASSESSMENT TEST 2018/2019

STA 2408: REGRESSION MODELLING II

TIME: 1 HOUR

INSTRUCTIONS TO CANDIDATES:

* *Answer question all the questions*
* *Be neat and show all your workings*

1. The height of a child at different ages Estimate the height of the child as an adult of 30 years of age using the growth model, $H = \frac{a}{1+be^{-ct}}$. [5 marks]

t(years)	0	5.0	8.0	12	16	18
H(inches)	20	36.2	52	60	69.2	70

2. There is a functional relationship between the mass density p of air and the altitude h above the sea level.

Altitude above the sea level, h (km)	0.32	0.64	1.28	1.60
Mass Density, $\rho(kg/m^3)$	1.15	1.10	1.05	0.95

In the regression model $\rho = k_1 e^{-k_2 h}$, the constant k_2 is found as $k_2 = 0.1315$. assuming that the mass density of air at the top of the atmosphere is $\frac{1}{1000^{th}}$ of the mass density of air at the sea level.
Calculate the altitude in kilometers of the top of the atmosphere. [8 marks]

3. For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^n K_h(X_i - x)(X_i - x)^l$ and $T_{nl}(x) = \sum_{i=1}^n K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively. [7 marks]

4. Given the data

X	20	30	15	25	28
Y	45.6	35.3	40.3	20.0	43.2

20, 28, 25
4 6

$a_0 x e^{kx}$

$\frac{1}{1000^{th}}$

where X and Y are predictor and response variables,respectively.Using a rectangular Kernel function and 3-nearest neighbour(K-NN)find the estimate of

$mx = \frac{1}{K-1} \sum y_i$ $[25, 43]$

$\frac{1}{2}(43.2 + 35.3)$

$\ln k_1$ $a_1 = 0.1315$

$y = a_0 - 0.1315h$

Scanned with CamScanner

i) Density function of X at point x=29.5 and at x=34 [3 ma

ii) regression function at point x=29.5 and at x=34 [3 marks

3. Let $X_1, X_2 \in [0, 1]$ and

$$\begin{cases} Z = 1, & \text{if } X_1 = 1 \text{ or } X_2 = 1 \\ Z = 0, & \text{if } X_1 = 0 \text{ or } X_2 = 0 \end{cases}$$

be the logical XOR classification to be learned by a perception. The training set then consists of input vectors including the first co-ordinate $x_0 = 1$;

$$X^1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad X^2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad X^3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad X^4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

with correction classification$Z^{(1)} = Z^{(2)} = Z^{(3)} = 1, Z^{(4)} = 0$.

The perception with weightsw_0, w_1, w_2 classifies an object as 1 if and only if $w_0x_0 + w_1x_1 + w_2x_2 > 0$ and as 0 elsewhere.Taking $w = (0, 0, 0)^t$ as initial weight and $\eta = 1$ as relaxation factor,train the perception and find the weight that achieves correct classification for all input vectors in the training set. [12marks]

2408

2019
7th 1 DEC 2018

W1-2-60-1-6

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

UNIVERSITY EXAMINATIONS 2017/2018

FOURTH YEAR FIRST SEMESTER EXAMINATIONS FOR THE DEGREE OF BACHELOR OF SCIENCE IN FINANCIAL ENGINEERING

STA 2408: REGRESSION MODELLING II

DATE: JANUARY 2018 TIME: 2 HOURS

INSTRUCTIONS TO CANDIDATES:

1. *Answer question ONE (section A) and any other two questions in section B.*
2. *Be neat and show all your workings*
3. *All questions except question one carry equal marks*

This paper consists of 4 printed pages.

SECTION A (30 MARKS)

1. (a) For the regression model

$$Y = X\beta + \epsilon$$

with $\epsilon \sim N(0, \sigma^2 W)$ where W is not an identity matrix, a student argues that the ordinary multiple regression cannot be applied and one should regress $W^{-1}Y$ against $W^{-1}X$ by ordinary multiple regression to estimate β. Do you agree with the student? If so, show how this can be done [6 marks]

(b) An Actuarial Science student is trying to analyze data from a particular stock using an equation of the form $y = f(x) = \frac{a_0}{x} + \frac{a_1}{x^2}$. The measured values of (x, y) are listed in the table below

x	1	2	3	4	5	6	7	8	9	10
y	3	7	0.9	6.8	0.7	0.96	19	37	0.8	0.1

Use nonlinear regression method to determine a_0 and a_1 [7 marks]

(c) Define the term Artificial Neural Network (ANN) [2 marks]

(d) Stock prices (Y, in dollars) are assumed to be affected by the annual rate of dividends of stock (X). A simple linear regression analysis ($Y_i = \beta_o + \beta_1 X_i + \epsilon_i$) was performed on 21 observations where ϵ_i were assumed to follow $\sim iidN(0, \sigma^2)$ and summary statistics were as listed below

$$\bar{X} = 0.4, \bar{Y} = 4, S_X^2 = 0.1, S_Y^2 = 20, S_{XY} = 1.25$$

Perform an ANOVA analysis for the level of significance 5% and give your conclusions[7 marks]

(e) Let $Y = (9, 4, 1, 3)^T$, $X_1 = (9, 3, 9, 4)^T$, $X_2 = (5, 3, 5, 4)^T$

(i) Find the matrix $(X^T X)^{-1}$ [2 marks]

(ii) Fit a regression model $Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \epsilon$. Give the fitted coefficients and estimate of error variance [4 marks]

(iii) Construct the 95% confidence interval for $E(Y|X_1 = 6, X_2 = 4)$ [3 marks]

2. In a study of the effect of hormones on the productivity of a certain variety of tomato plant, ten plants were treated at different hormone strengths and their yields, in Kgs, noted. A simple linear regression model is fitted in R, with the depended variable (yield in kgs) stored as the vector y and the independent variable (hormone strength) stored as the vector x. Some of the R commands and edited output are shown below

```
> lm1 = lm(y~x)
> summary(lm1)


Coefficients:
             Estimate    Std. Error
(Intercept) -3.72335     0.28072
```

$W_{new} = w + \eta(z^t - y^{(t)}) x^t$

22/32 + 3

JKUAT CONTINUOUS ASSESSEMENT TESTS

NAME: ESTHER MOSE KOMBO. REG NO: SC283-6589/2014.

COURSE: BSc FINANCIAL ENGINEERING YEAR OF STUDY: 4th. SHEET NO: ......

UNIT CODE: ...... TITLE: REGRESSION MODELLING II DATE: 29th Oct 2018

NOTE: This stationery will be used for Continuous Assessment work only.

It will be a breach of examination regulations to use it otherwise.

$AV \neq \underset{i+b/e+gt}{a}$

4) i) $K = 3$. $X = 29.5$ and at $X = 34$

$d_K(X) = 2(K-1)$

$d_K(X) = 2(3-1) = 2 \times 2 = 4$

$\hat{P}(X) = \frac{K-1}{2n\,d_K(X)} = \frac{3-1}{2(5)(4)} = 0.05$

ii) $X = 29.5$

$\hat{m}_3(29) = \frac{1}{K-1} \sum_{j=1}^{n} Y_j \, \Pi(x - d_K(x), x + d_K(x)) X_j$

$= \frac{1}{3-1} \sum_{j=1}^{n} Y_j \, \Pi(29.5 - 4, 29.5 + 4) X_j$

$= \frac{1}{2} \sum_{j=1}^{n} Y_j \, \Pi_{(25.5, 33.5)} X_j$

$= \frac{1}{2}(43.2 + 35.3)$

$= 39.25$

At $X = 34$

$\hat{m}_3 = \frac{1}{2} \sum_{j=1}^{n} Y_j \, \Pi(34-4, 34+4) X_j$

$\frac{1}{2} \sum_{j=1}^{n} Y_j \, \Pi(30, 38)$.

$= \frac{1}{2}(35.3) = 17.65$

5) $Y^{(1)} = W^T X^{(1)}$

$= (0,0,0)\begin{pmatrix}1\\1\\0\end{pmatrix} = 0 \neq z^{(1)}$

so $X^{(1)}$ is not correctly classified.

$W_{new} = (0,0,0) + (1-0)(1\ 1\ 0) = (1\ 1\ 0)$

$Y^{(2)} = (1\ 1\ 0)\begin{pmatrix}1\\0\\1\end{pmatrix} = 1 = z^{(2)}$

$X^{(2)}$ is correctly classified

$Y^{(3)} = (1\ 1\ 0)\begin{pmatrix}1\\1\\1\end{pmatrix} = 2 = z^3$

so X^3 is correctly classified

$Y^{(4)} = (1\ 1\ 0)\begin{pmatrix}1\\0\\0\end{pmatrix} = 1 \neq z^{(4)}$

so X^4 is incorrectly classified

$W_{new} = (1\ 1\ 0) + (0-1)(1\ 0\ 0) =$
$(1\ 1\ 0) + (-1\ 0\ 0) = (0\ 1\ 0)$

$Y^{(1)} = (0\ 1\ 0)\begin{pmatrix}1\\1\\0\end{pmatrix} = 1 = z^{(1)}$ so

$X^{(1)}$ is correctly classified.

$a = 4.3221 \times 10^{1}$

$b = 2.8233$

$c = 2.1715 \times 10^{-1}$

$y^{(2)} = (0\ 1\ 0)\begin{pmatrix}1\\0\\1\end{pmatrix} = 0 \neq z^{(2)}$

so X^2 is incorrectly classified

$W_{new} = (0\ 1\ 0) + (1-0)(1\ 0\ 1)$
$(0\ 1\ 0) + (1\ 0\ 1) = (1\ 1\ 1)$

$y^{(3)} = (1\ 1\ 1)\begin{pmatrix}1\\1\\1\end{pmatrix} = 3 \equiv z^3$

so X^3 is correctly clasified.

$y^4 = (1\ 1\ 1)\begin{pmatrix}1\\0\\0\end{pmatrix} = 1 \neq z^4$

$W_{new} = (1\ 1\ 1) + (0-1)(1\ 0\ 0)$
$= (1\ 1\ 1) + (-1\ 0\ 0)$
$= (0\ 1\ 1)$

$y^{(1)} = (0\ 1\ 1)\begin{pmatrix}1\\1\\0\end{pmatrix} = 1 = z^{(1)}$

so $X^{(1)}$ is correctly classified

$y^{(2)} = (0\ 1\ 1)\begin{pmatrix}1\\0\\1\end{pmatrix} = 1 = z^{(2)}$

so $X^{(2)}$ is correctly classified

$y^{(3)} = (0\ 1\ 1)\begin{pmatrix}1\\1\\1\end{pmatrix} = 2 = z^{(3)}$

so $X^{(3)}$ is correctly classified.

$y^{(4)} = (0\ 1\ 1)\begin{pmatrix}1\\0\\0\end{pmatrix} = 0 = z^{(4)}$

so X^4 is correctly classified

The weight $(0\ 1\ 1)$ has achieved correct classification for all the input vectors in the training set.

QUE 1

t (years)	0	5.0	8.0	12	16	18
H (inches)	20	36.2	52	60	69.2	70

$$H = \frac{a}{1+be^{-ct}}$$

$$S_r = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n}\left(H_i - \frac{a}{1+be^{-ct_i}}\right)^2$$

$$\frac{\partial S_r}{\partial a} = \sum \frac{2e^{ct_i}[ae^{ct_i} - y_i(e^{ct_i}+b)]}{(b+e^{ct_i})^2} = 0$$

$$\frac{\partial S_r}{\partial b} = \sum \frac{2ae^{ct_i}[bH_i + e^{ct_i}(H_i - a)]}{(b+e^{ct_i})^2} = 0$$

$$\frac{\partial S_r}{\partial c} = \sum \frac{2abe^{ct_i}[bH_i + e^{ct_i}(H_i - a)]}{(b+e^{ct_i})^2} = 0$$

$$\frac{\partial}{\partial a}\sum\left(\frac{2e^{cx_i}[ae^{cx_i} - y(e^{cx_i}-b)]}{(e^{cx_i}+b)^2}\right)$$

$$\frac{\partial S_r}{\partial b} = \sum\left(\frac{2ae^{cx_i}[y_ib + e^{cx_i}(y_i - a)]}{(e^{cx_i}+b)^3}\right)$$

$$\frac{\partial S_r}{\partial c} = \sum \frac{2abx_ie^{cx_i}[by_i + e^{cx_i}(y_i-a)]}{(e^{cx_i}+b)^3}$$

2)

h(km)	0.32	0.64	1.28	1.60
p (kg/m³)	1.15	1.10	1.05	0.95

$p = K_1 e^{-K_2 h}$. $K_2 = 0.1315$.

$\ln p = \ln K_1 - K_2 h$.

let $z = \ln p$. , $a_0 = \ln K_1$ and $a_1 = -K_2$

$= z = a_0 + a_1 h$.

$$a_1 = \frac{n\sum z_i h_i - \sum z_i \sum h_i}{n\sum h_i^2 - (\sum h_i)^2}, \quad n = 4$$

h(km)	p (kg/m³)	$z = \ln p$	$z_i h_i$
0.32	1.15	0.1398.	0.044736.
0.64	1.10	0.0953.	0.060992
1.28	1.05	0.0488.	0.062464
1.60.	0.95	-0.0513.	-0.082208
3.84		0.2326	0.086112.

h_i^2
0.1024
0.4096
1.6384
2.56.
4.7104.

$$a_1 = \frac{4(0.086112) - (0.2326)(3.84)}{4(4.7104) - (3.84)^2}.$$

$$= \frac{-0.548736}{4.096} = -0.13397.$$

$a_0 = \bar{z} - a_1 \bar{h}$

$$= \frac{0.2326}{4} - \left(-0.13397 \times \frac{3.84}{4}\right)$$

$= 0.05815 + 0.1286112$

$a_0 = 0.1867612.$

$K_1 = e^{a_0} = e^{0.1867612} = 1.2053$

$p = 1.21 e^{-0.1315(h)}$.

$\frac{1}{1000} = 0.001$

$0.001 = 1.21 e^{-0.1315(h)}$.

$0.000826 = e^{-0.1315(h)}$.

$-7.0984 = -0.1315(h)$

$h = 53.98.$

Scanned with CamScanner

70

2

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

CONTINUOUS ASSESSMENT TEST 2019/2020

STA 2408: REGRESSION MODELLING
CAT

$j=1,2,\ldots N$

INSTRUCTIONS TO CANDIDATES:
* *Answer question all the questions*
* *Be neat and show all your workings*

$y = mx + c$ $\quad Y = \beta_0 + \beta_1 X$ $\quad Y_j = m(x_j) + e_j$

---

1. Differentiate between parametric and non parametric regression [2 marks]

2. Define what is meant by a spline and illustrate using two examples. Show that splines are linear smoothers.

3. Draw a well labelled diagram of a neural network with two hidden layers and X1; X2 and X3 as the input variables.

4. Define stochastic and deterministic designs under Kernel regression estimation stating their mean functions.

5. K-NN is a weighted average in a varying neighborhood(defined through those X variables which are among the K nearest neighbors of X in Euclidean distance (linear real distance)). Let $(X_i, Y_i)_{i=1}^{5}$ be $(1,5), (7,12), (3,1), (2,0), (5,4)$ Compute (K-NN estimate) $\hat{m}_k(x)$ (i.e) for $x = 4andk = 3$.

6. Show that a linear regression estimator is a special case of non parametric estimator.

7. Given the data

X	20	30	15	25	28
Y	45.6	35.3	40.3	20.0	43.2

where X and Y are predictor and response variables,respectively.Using a rectangular Kernel function and 3-nearest neighbour(K-NN)find the estimate of.

i) Density function of X at point x=29 [3 marks]

ii) regression function at point x=29 [3 marks]

8. There is a functional relationship between the mass density p of air and the altitude h above the sea level.

Altitude above the sea level, h (km)	0.32	0.64	1.28	1.60
Mass Density, $\rho (kg/m^3)$	1.15	1.10	1.05	0.95

In the regression model $\rho = k_1 e^{-k_2 h}$, the constant k_2 is found as $k_2 = 0.1315$. assuming that the mass density of air at the top of the atmosphere is $\frac{1}{1000^{th}}$ of the mass density of air at the sea level.

Calculate the altitude in kilometers of the top of the atmosphere. [8 marks]

$$\rho = \frac{\sum_{i=1} k_i e^{-k_2 h_i}}{\sum_{i=1} e^{-k_2 h_i}}$$

0.000208 = 0.2080

JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY

CONTINUOUS ASSESSMENT TEST 2019/2020

STA 2408: REGRESSION MODELLING

CAT

INSTRUCTIONS TO CANDIDATES:

* *Answer question all the questions*

* *Be neat and show all your workings*

---

1. Differentiate between parametric and non parametric regression [2 marks]
2. Define the term smoothing with respect to non parametric regression. [2 marks]
3. What is Kernel regression. Illustrate your explanation with examples. [4mks]
4. State the steps involved in fitting a Local Polynomial Regression Curve [3mks]
5. K-NN is a weighted average in a varying neighborhood(defined through those X variables which are among the K nearest neighbors of X in Euclidean distance (linear real distance)). Let $(X_i, Y_i)_{i=1}^{5}$ be $(1,5), (7,12), (3,1), (2,0), (5,4)$ Compute (K-NN estimate) $\hat{m}_k(x)$ (i.e) for $x = 4 and k = 3$.
6. Many patients get concerned when a test involves injection of a radioactive material.For example for scanning a gallbladder, a few drops of Technetium-99m isotope is used. Half of the technetium-99m would be gone in about 6 hours. It, however, takes about 24 hours for the radiation levels to reach what we are exposed to in day-to-day activities. Below is given the relative intensity of radiation as a function of time.

t (hrs)	0	1	3	5	7	9
ρ	1.000	0.891	0.708	0.562	0.447	0.355

In the regression model $\rho = Ae^{\lambda t}$, the constant λ is found as $k_2 = 0.11505$. Find the value of the constant term A, the half-life of Technetium-99m, and the radiation intensity after 24 hours.

$A = 0.99983$ $t = 6.0232$

7. For a given sample $(X_i, Y_i), i = 1, 2, \cdots n$ from an NPR model $Y_i = m(X_i) + \epsilon_i$, where $n = 140$, $X_i \in [0.1894, 2.6855]$ and ϵ_i are i.i.d with $E\epsilon_1^2 = \sigma^2$, we define $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l$ and $T_{nl}(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l Y_i$. We used the standard Gaussian kernel and $h = 0.2420$. Let $x_0 = 1$. We obtained $s_{nl}(x_0), l =, 0, 1, 2$ as $22.8038, -0.5758, 1.1189$ respectively, and $T_{nl}(x_0), l = 0, 1$ as $24.0092, 0.1631$ respectively. Compute the local linear estimators of $m(x_0)$ and $m'(x_0)$ respectively.

STA 2408: Regression Modelling II Cat 1

1. Differentiate between the following models

   (a) Linear and non-linear models [2 marks]

   (b) Parametric and non-parametric models [2 marks]

2. For an exponential model $y = \gamma e^{\zeta x}$ that is best fit to the data $(x_1, y_1), ..., (x_n, y_n)$ derive a non-linear equation that can be used to estimate the value ζ. [5 marks]

3. Empirical results have shown that the rate of gas flow from a container is proportional to some power of the nozzle pressure. Below is given the flow rate, in cm^3 per second as a function of pressure.

Flow rate, F (cm^3/sec)	88	134	135	148	172	240
Pressure, P (psi)	15	22	26	28	52	60

The rate of gas flow is related to the nozzle pressure via the regression model $F = \alpha e^{P\beta}$. By transforming the above data, find

   (a) The value of the regression parameters α and β [6 marks]

   (b) The rate of gas flow if the nozzle pressure is increased to 80 psi [2 marks]

4. The following data contains measurements of yield from an experiment done at six different temperature levels.

Temperature (T)	Yield (Y)
5	6.26
10	4.24
15	3.88
20	2.26
25	1.44
30	0.60

If the Yield behaves as a second order polynomial function of temperature fit the above data to the model $Y = a_0 + a_1 T + a_2 T^2$

[7 marks]

5. There exists a functional relationship between the mass density ρ of air and the altitude h above the sea level. A sample data of the two variables is given below

Altitude above sea level, h(km)	0.32	0.64	1.28	1.60
Mass Density, ρ (kg/m^3)	1.15	1.10	1.05	0.95

The functional relationship can be expressed using the regression model $\rho = k_1 e^{-k_2 h}$, where the constant k_2 is found as $k_2 = 0.1315$. Assuming that the mass density of air at the top of the atmosphere is $1/1000^{th}$ of the mass density of air at sea level find the altitude in kilometers of the top of the atmosphere. [6 marks]

9.05 9.50 52.

$\frac{23}{32}$

$a_0 + a_1x_1 + a_2x_2^2$

$2\Sigma x_i$ $2\times2=4$

$\Sigma x_i y$

JKUAT CONTINUOUS ASSESSEMENT TESTS

NAME: VERONICAH WAMBUI MACHARIA

REG NO: SCM 214-1595/2015

COURSE: BSC. FINANCIAL ENGINEERING YEAR OF STUDY: 4.1

SHEET NO:

UNIT CODE: STA 2408 TITLE: REGRESSION MODELLING II

DATE:

NOTE: This stationery will be used for Continuous Assessment work only.

It will be a breach of examination regulations to use it otherwise.

3. $n = 140$

$$\hat{m}(x_0) = \frac{S_{n2}(x_0)T_{n0}(x_0) - S_{n1}(x_0)T_{n1}(x_0)}{S_{n2}(x_0)S_{n0}(x_0) - \{S_{n1}(x_0)\}^2} = \frac{(1.1189 * 24.0092) - (-0.5758 * 0.1631)}{(1.1189 * 22.8038) - \{-0.5758\}^2}$$

$$= \frac{26.8639 + 0.09391298}{25.5152 - 0.331546} = \frac{26.9578}{25.18365}$$

$$= 1.0704483$$

and $\hat{m}'(x_0)$ is given by

$$\hat{m}'(x_0) = \frac{-S_{n1}(x_0)T_{n0}(x_0) + S_{n0}(x_0)T_{n1}(x_0)}{S_{n2}(x_0)S_{n0}(x_0) - \{S_{n1}(x_0)\}^2}$$

$$= \frac{(+0.5758 * 24.0092) + (22.8038 * 0.1631)}{1.1189 * 22.8038 - \{-0.5758\}^2}$$

$$= \frac{13.82449736 + 3.719299978}{25.18365} = \frac{17.54379714}{25.18365}$$

$$= 0.6966$$

4. $K = 3$ $K-1 = 2$ $d_K(x) = 4$ $N = 5$

i) $X = 29.5$

$$\hat{P}(x) = \frac{K-1}{2Nd_K(x)} = \frac{2}{2(5)(4)} = 0.05$$

$X = 24$

$$\hat{P}(x) = \frac{K-1}{2Nd_K(x)} = \frac{2}{2(5)(4)} = 0.05$$

4.ii) $X = 29.5$

$$\hat{m}(x) = \frac{1}{k-1} \sum_{i=1}^{N} Y_j \, \Pi_{(x-d_k(x), x+d_k(x))} \cdot X_j$$

$$= \frac{1}{2} \sum_{i=1}^{N} Y_j \, \Pi_{(29.5-4, 29.5+4)} X_j$$

$$= \frac{1}{2} \sum_{i=1}^{N} Y_j \, \Pi_{(25.5, 33.5)} X_j$$

$$= \frac{1}{2} \{ 35.3 + 43.2 \}$$

$$= \frac{1}{2} \{ 78.5 \} = 39.25$$

$$= \underline{39.25}$$

$X = 34$

$$\hat{m}(x) = \frac{1}{k-1} \sum_{i=1}^{N} Y_j \, \Pi_{(x-d_k(x), x+d_k(x))} \cdot X_j$$

$$= \frac{1}{2} \sum_{i=1}^{N} Y_j \, \Pi_{(34-4, 34+4)} X_j$$

(30, 38)

$$= \frac{1}{2} \{ 35.3 + 43.2 \} = \frac{1}{2} \{ 78.5 \}$$

$$= \underline{39.25}$$

0

$= 0$

5. 1. $y^{(1)} = W^T x^{(1)} = (0,0,0)^T \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 0 \neq z^{(1)} = 1$

$\Rightarrow X^{(1)}$ is classified incorrectly. Modifying the weights:

$$W_{new}^T = W^T + \eta (z^{(1)} - y^{(1)}) (x^{(1)})^T$$

$$= [0,0,0]^T + (1-0)[1\ 1\ 0]^T = [1\ 1\ 0]^T$$

2. $y^{(2)} = W_{new}^T x^{(2)} = [1\ 1\ 0]^T \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 = z^{(2)}$

$\Rightarrow X^{(2)}$ is correctly classified.

3. $y^{(3)} = W^T x^{(3)} = [1\ 1\ 0]^T \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2 \equiv 1 = z^{(3)}$

$\Rightarrow x^{(3)}$ is correctly classified.

4. $y^{(4)} = W^T x^{(4)} = [1\ 1\ 0]^T \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 \neq z^{(4)} = 0$

$\Rightarrow x^{(4)}$ is incorrectly classified. Modifying the weights;

$$W_{new}^T = W^T + \eta (z^{(4)} - y^{(4)}) (x^{(4)})^T$$

$$= [1\ 1\ 0]^T + (0-1)[1\ 0\ 0]^T$$

$W^T_{new} = [0\ 1\ 0]^T$

5. $y^{(1)} = W^T x^{(1)} = [0\ 1\ 0]^T \begin{bmatrix}1\\1\\0\end{bmatrix} = 1 = z^{(1)}$

$\Rightarrow x^{(1)}$ is classified correctly.

6. $y^{(2)} = W^T x^{(2)} = [0\ 1\ 0]^T \begin{bmatrix}1\\0\\1\end{bmatrix} = 0 \neq z^{(2)} = 1$

$\Rightarrow x^{(2)}$ is classified incorrectly; modifying the weights;

$W^T_{new} = [0\ 1\ 0]^T + (1-0)[1\ 0\ 1]^T = [1\ 1\ 1]^T$

7. $y^{(3)} = W^T x^{(3)} = [1\ 1\ 1]^T \begin{bmatrix}1\\1\\1\end{bmatrix} = 3 \equiv z^{(3)} = 1$

$\Rightarrow x^{(3)}$ is correctly classified.

8. $y^{(4)} = W^T x^{(4)} = [1\ 1\ 1]^T \begin{bmatrix}1\\0\\0\end{bmatrix} = 1 \neq z^{(4)} = 0$

$\Rightarrow x^{(4)}$ is wrongly classified. Modifying the weights;

$W^T_{new} = [1\ 1\ 1]^T + (0-1)[1\ 0\ 0]^T = [0\ 1\ 1]^T$

which classifies all inputs correctly as follows:

9. $y^{(1)} = W^T x^{(1)} = [0\ 1\ 1]^T \begin{bmatrix}1\\1\\0\end{bmatrix} = 1 = z^{(1)} \Rightarrow x^{(1)}$ is correctly classified.

$y^{(2)} = W^T x^{(2)} = [0\ 1\ 1]^T \begin{bmatrix}1\\0\\1\end{bmatrix} = 1 = z^{(2)} \Rightarrow x^{(2)}$ is correctly classified.

$y^{(3)} = W^T x^{(3)} = [0\ 1\ 1]^T \begin{bmatrix}1\\1\\1\end{bmatrix} = 2 \equiv z^{(3)} \Rightarrow x^{(3)}$ is correctly classified.

$y^{(4)} = W^T x^{(4)} = [0\ 1\ 1]^T \begin{bmatrix}1\\0\\0\end{bmatrix} = 0 = z^{(4)} \Rightarrow x^{(4)}$ is correctly classified.

$W = [0\ 1\ 1]^T$ is the weight that achieves correct classification for all input vectors.

Q. $P = K_1 e^{-K_2 h}$ which when linearized becomes;

$\ln P = \ln K_1 - K_2 h$

But $K_2 = 0.1315$

$\ln P = \ln K_1 - 0.1315h$

Let $A = \ln P$; $K = \ln K_1$

$\Rightarrow P = e^A$ $\Rightarrow K_1 = e^K$

The regression model becomes;

$A = K - 0.1315h$

To obtain the parameters:

$K = \bar{A} + 0.1315\bar{h}$ where $\bar{A} = \frac{\sum_{i=1}^{n} A_i}{n}$

Table:

Summations of the data to obtain parameters:

Table:

i	h_i	P_i	$A_i = \ln P_i$
1	0.32	1.15	0.1398
2	0.64	1.10	0.0953
3	1.28	1.05	0.04879
4	1.60	0.95	-0.05129
	$\Sigma h_i = 3.84$	$\Sigma P_i = 4.25$	$\Sigma A_i = 0.2326$

$K = \bar{A} + 0.1315\bar{h}$

$\frac{\sum_{i=1}^{4} A_i}{4} = \frac{0.2326}{4} = 0.05815$

$\frac{\sum_{i=1}^{4} h_i}{4} = \frac{3.84}{4} = 0.96$

$K = 0.05815 + 0.1315(0.96)$

$K = 0.05815 + 0.12624$

$K = 0.18439$

The regression model $A = K - 0.1315h$ becomes

$A = \ln P$; $K = \ln K_1$

$K_1 = e^{0.18439} = 1.20248$

$A = 0.18439 - 0.1315h$

$\ln P = \ln K_1 - 0.1315h$

$e^A = P = 0.18439\, e^{-0.1315h}$

$P = \frac{1}{100} P \Big|_{h=0}$

$= \frac{1}{100}\{0.18439\, e^{-0.1315(0)}\}$

$= \frac{1}{100}\{0.18439\}$

$0.18439\, e^{-0.1315h} = \frac{1}{100} \cdot 0.18439$

$e^{-0.1315h} = \frac{1}{100}$

$-0.1315h = \ln 0.01 = -4.60517$

$h = 35.02$ km

$K_1 = 1.20248$

$e^A = P = 1.20248\, e^{-0.1315h}$

$\Rightarrow$ The regression model is

$P = 1.20248\, e^{-0.1315h}$

$P = \frac{1}{100} P \Big|_{h=0}$

$= \frac{1}{100}\{1.20248\, e^{-0.1315(0)}\}$

$= \frac{1}{100}\{1.20248\, e^{0}\}$

$1.20248\, e^{-0.1315h} = \frac{1}{100}\{1.20248\, e^{0}\}$

$e^{-0.1315h} = \frac{1}{100}$

$-0.1315h = \ln 0.01$

$h = \frac{-4.60517}{-0.1315}$

$= 35.02$ km

1. $H = \frac{a}{1+be^{-ct}}$

at $t = 0$

$20 = \frac{a}{1+b}$

at $t = 8$

$52 = \frac{a}{1+be^{-8c}}$

at $t = 18$

$70 = \frac{a}{1+be^{-18c}}$

Using Newton-Ralphson iterative method:

$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$