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 $\mathrm{E}\epsilon_1^2 = \sigma^2$, 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 $\mathrm{E}\epsilon_1^2 = \sigma^2$, 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]

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**

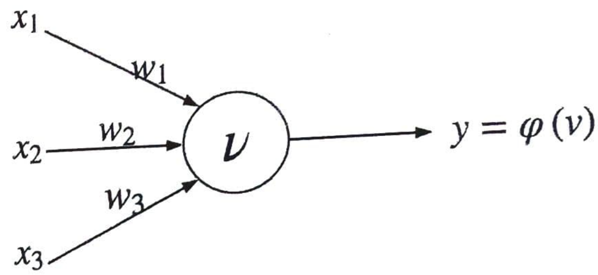
INSTRUCTIONS TO CANDIDATES:

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, $H = \frac{a}{1+be^{-ct}}$. [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 $s_n(x) = \sum_{i=1}^{n} K_h(X_i - x)(X_i - x)^l$ 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]

(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	426.969
-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 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 (μ $in/in/$ ^0F)
-320	2.76
-240	3.83
-160	4.72
-80	5.43
0	6.00
80	6.47

(i) Fit the data to $\alpha = a_0 + a_1 T + a_2 T^2$. [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 $\boldsymbol{Y} = [Y_1, \cdots, Y_n]^T$ be the response vector and $\hat{\boldsymbol{Y}}$=$[\hat{Y}_1, \cdots, \hat{Y}_n]^T$ 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 $\hat{\boldsymbol{Y}} = \boldsymbol{AY}$ where $\boldsymbol{A}$ 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]

(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]