BBS 4·1

Regression Analysis II.

1· QUESTION PAPER.

1b.

t	0	5	8	12	16	18
H	20	36.2	52	60	69.2	70.

soln.

$$H = \frac{a}{1+be^{-ct}} \qquad \frac{dS_r}{da} = \sum_{i=1}^{6} \frac{2e^{ct_i}\left(ae^{ct_i} - H_i\left(e^{ct_i}+b\right)\right)}{\left(e^{ct_i}+b\right)^2}.$$

$$\frac{dS_r}{db} = \sum_{i=1}^{6} \frac{2ae^{ct_i}\left[bH_i + e^{ct_i}(H_i - a)\right]}{\left(e^{ct_i}+b\right)^3} = 0.$$

$$\frac{dS_r}{dc} = \sum_{i=1}^{6} \frac{-2abt_ie^{ct_i}\left(bH_i + e^{ct_i}(H_i-a)\right)}{\left(e^{ct_i}+b\right)^3} = 0.$$

To get the intial roots, we can use three data points (0, 20) (12, 60) (18, 70). We have;

$$20 = \frac{a}{1+be^{-c(0)}} \qquad 60 = \frac{a}{1+be^{-12c}} \qquad 70 = \frac{a}{1+be^{-18c}}.$$

One can solve for the three unknowns a, b, c to obtain $a = 75.534$, $b = 2.7767$ $c = 0.19772$

Applying newton raphson method for simultaneous non-linear equation, one can get the roots:

$a = 74.321$ $b = 2.8233$ $c = 0.21715$.

The saturation growth model of the height of the child is then

$$H = \frac{74.321}{1+2.8233e^{-0.21715t}}$$

when $t = 30$ yrs

$H = 74$ inches.

(Q). h(km) 0.32 0.64 1.28 1.60
P (kg/m³) 1.15 1.10 1.05 0.95

$P_i = k_1 e^{-k_2 h}$ $k_2 = 0.1315$
$P_1 = \cancel{0.1315} \, k_1 e^{-k_2 h}$.

soln.
$p = k_1 e^{-0.1315h}$.

$\ln p = z$ $z = e^p$.
$\ln k_1 = a_0$. $k_1 = e^{a_0}$

$\ln p = \ln k_1 - 0.1315h$.
$z = a_0 - 0.1315h$.
$a_0 = \bar{z} - 0.1315\bar{h}$

$\bar{z} = \frac{0.2326}{4} = 0.05815$

$\bar{h} = \frac{3.84}{4} = 0.96$.

z ln p	h
0.1398	0.32
0.0953	0.64
0.0488	1.28
−0.0513	1.60
Σ = 0.2326	3.84.

$a_0 = 0.05815 - (0.1315 \times 0.96) = 0.18439$
$k_1 = e^{a_0} = 1.2025$

$P_1 = 1.2025 \, e^{-0.1315h}$.

$-Bx_?$

$+e^{-Bx_i}$

$g(1+$

d) $S_{n\,l}(x_0)$ $l = 0, 1, 2.$ as $22.8038, -0.5758, 1.1189$
$T_{n\,l}(x_0)$ $l = 0, 1$ as $24.0092, 0.1631$
Find $M(x_0)$ and $M'(x_0)$.

soln.

$$M(x_0) = \frac{S_{n_2}(x_0)\,T_{n_0}(x_0) - S_{n_1}(x_0)\,T_{n_1}(x_0)}{S_{n_2}(x_0)\,S_{n_0}(x_0) - (S_{n_1}(x_0))^2}.$$

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

$$Ans = 1.07044977 \approx 1.0704.$$

$$M'(x_0) = \frac{-S_{n_1}(x_0)\,T_{n_0}(x_0) + S_{n_0}(x_0)\,T_{n_1}(x_0)}{S_{n_2}(x_0)\,S_{n_0}(x_0) - (S_{n_1}(x_0))^2}.$$

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

$$= 0.6966.$$

E). SAS OUTPUT.

(i) $\ln \frac{\pi(x)}{1-\pi(x)} = -6.3896 + 0.0266\underset{(psa)}{X_1} - 0.0208\underset{(age)}{X_2} + 1.0790\underset{(gleason)}{X_3}$

(ii) $$E(Y) = \frac{e^{B_0 + B_1X_1 + B_2X_2 + B_3X_3}}{1 + e^{B_0 + B_1X_1 + B_2X_2 + B_3X_3}}$$

$$= \frac{e^{-6.396 + 0.0266(10) - 0.0208(69) + (1.0790 \times 5)}}{1 + e^{-6.396 + (0.0266 \times 10) - (0.0208 \times 69) + (1.0790 \times 5)}}$$

Ans = 0.1030.

(iii)

iv) $OR = e^{B_1} = e^{0.0266} = 1.0269$

$e^{\hat{B_i} \pm Z_{\alpha/2} SE}$

$e^{0.0266 \pm 1.96 \times 0.00894}$

$= e^{[0.0091,\ 0.0441]}$.

QUEST $(1.0091,\ 1.045086)$

QUESTION TWO:

b. Initial $80^{0}F$ length 12 inches placed in $-315^{0}F$.

Temp (T)	Thermal exp	T^2	T^3	T^4	$T\alpha$	$T^2\alpha$
-320	2.76×10^{-6}	102400	-32768000	1.048576×10^{10}	-8.832×10^{-4}	0.282624 ~~3.2768x10~~
-240	3.83×10^{-6}	57600	-13824000	0.33176×10^{10}	-9.192×10^{-4}	0.220608
-160	4.72×10^{-6}	25600	-4096000	0.065536×10^{10}	-7.552×10^{-4}	0.120832
-80	5.43×10^{-6}	6400	-512,000	40960000	-4.344×10^{-4}	0.034752
0	6.00×10^{-6}	0	0	0	0	0
80	6.47×10^{-6}	6400	512,000	40960000	5.176×10^{-4}	0.041408
-720	2.921×10^{-5}	1.984×10^{5}	-5.0688×10^{7}	1.45408×10^{10}	-2.4744×10^{-3}	0.700021

Soln. $\alpha = a_0 + a_1T + a_2T^2$.

$$\begin{bmatrix} n & \Sigma T & \Sigma T^2 \\ \Sigma T & \Sigma T^2 & \Sigma T^3 \\ \Sigma T^2 & \Sigma T^3 & \Sigma T^4 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} \Sigma \alpha \\ \Sigma T\alpha \\ \Sigma T^2\alpha \end{bmatrix}$$

$$\begin{bmatrix} 6 & -7.2\times10^{2} & 1.984\times10^{5} \\ -7.2\times10^{2} & 1.984\times10^{5} & -5.0688\times10^{7} \\ 1.984\times10^{5} & -5.0688\times10^{7} & 1.45408\times10^{10} \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2.921\times10^{-5} \\ -2.4744\times10^{-3} \\ 0.700024 \end{bmatrix}$$

$a_0 = 6.0243\times10^{-6}$

$a_1 = 6.3061\times10^{-9}$

$a_2 = -1.2073\times10^{-11}$

$$\alpha = (6.0243\times10^{-6}) + (6.3061\times10^{-9})T + (-1.2073\times10^{-11})T^2.$$

$T = -315^{0}F$ $\quad \alpha = (6.0243\times10^{-6}) + (6.3061\times10^{-9})(-315) + (-1.2073\times10^{-11})(-315)^2 = 2.839967\times10^{-6}$.

$9 \ (e^{-Rx?}$

QUESTION THREE

x	20	30	15	25	28
y	45.6	35.3	40.3	20.0	43.2

k = 3. find p(x) and m(x).

soln.

Nearest obs. to 29

25, 28, 30.

distance = 4, 1, 1 highest = 4.

$$p(x) = \frac{k-1}{2N\,d_k(x)}$$

$$\frac{3-1}{2 \times 5 \times 4} = \frac{1}{20} = 0.05$$

$$m(x) = \frac{1}{k-1} \sum_{j=1}^{n} y_j \, \mathbb{I}\left(x - d_k(x),\ x + d_k(x)\right) x$$

$$\frac{1}{2} \sum_{j=1}^{n} y_i \, \mathbb{I}\,[25, 33]\, x.$$

x	28	30
y	43.2	35.3

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

$$= \underline{\underline{39.25}}$$

Discuss local polynomial Regression stating the five steps

- It uses the data from a neighborhood around the specific location so that the neighbourhood is defined as a span which is the function of the total points used to form the neighborhood.

- It then uses the points in the neighborhood to generate a weighted least squares estimation of the specific response.

Procedure:

- Take a point say x_0: Find the KNN of x_0 which constitutes of a neighborhood $N(x_0)$. the number of neighbors k is specified as a percentage of the total number of points to be spanned.
- Calculate the largest distance between x_0 and another point in the neighborhood as follows:

$$\Delta x_0 = \max_{N(x_0)} |x_0 - x_1|$$

- Assign weights to each point in $N(x_0)$ using the tri-cube weight function.

$$w = \left| \frac{|x_0 - x_1|}{\Delta x_0} \right| \quad \text{where } w(u) = \begin{cases} (1-u^3)^3 & 0 \le u \le 1 \\ 0 & \text{otherwise} \end{cases}$$

- Calculate the weighted least square fit of y in the neighbourhood $N(x_0)$.

Take the fitted value,

$$\hat{y}_0 = f(x_0) = a$$

$$\hat{y}_i = \underset{w_i}{\operatorname{argmin}} \left\{ \sum_{i=1}^{n} y_i - a - b_i(x - x_0), \ldots - b_p(x_1 - x_0)^p \right\}^2$$

for the i^{th} predictor value.

- Repeat for each predictor value.

Discuss spline regression.

Spline regression is a non-parametric technique that divides the dataset into bins at interval or points known as knots and each bin has its separate fit.

Assumptions of the linear Mixed model.

1. The explanatory variables have a linear relationship with the response variable.
   ✓ It can be tested by plotting model residuals vs predictor.
2. The error term have a constant variance
   ✓ Use ANOVA to test for homogenity in variance of the error term.
3. There is independence of the error term.
   ✓ Use chi-square to test for independence.
4. Error terms are normally distributed.
   ✓ Use Q-Q plots.
5. The dependent variable is qualitative.

Scanned with CamScanner

$x_1, x_2 \in (0,1)$

$z = 1$ if $x_1 = 1$ or $x_2 = 3$ or $x_1 = 3$.

$z = 0$ if $x_1 = 0$, $x_2 = 0$.

$x_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ $x_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ $x_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $x_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$z^{(1)} = z^{(2)} = z^{(3)} = 1$

$w^T = (0, 0, 0)$. $z^{(4)} = 0$.

<u>soln.</u>

$y(1) = w^T x^{(1)}$.

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

$x(1)$ is classified incorrectly. We should modify the weights.

$w_{new} = (0, 0, 0) + (1 - 0) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

$y(1) = w^T (x^{(1)}) = (1 \; 1 \; 0) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 2 = 1 = z^{(1)}$.

$x(1)$ is correctly classified.

$y^{(2)} = w^T x^{(2)} = (1 \; 1 \; 0) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 = z^{(2)}$

x_2 is correctly classified.

$y^{(2)} = w^T x^{(3)} = (1 \; 1 \; 0) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2 = 1 = z^{(3)} =$.

x_3 is correctly classified.

$y(4) = w^T x_4.$

$(1, 1, 0) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1 \neq z^4.$

x_4 is classified incorrectly. We should modify the weights.

$w_{new} = (1, 1, 0) + (0 - 1) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$

$y^{(1)} = w^T x^{(1)}$

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

x_1 is correctly classified.

$y^{(2)} = w^T x^{(2)}$

$(0\ 1\ 0) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 0 \neq z^{(2)}.$

x_2 is classified incorrectly. We should modify the weights.

$w_{new} = (0\ 1\ 0) + 1 - 0 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$

$y(1) = w^T x^{(1)}.$

$[1\ 1\ 1] \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = 2 = 1 = z^{(1)}$

x_1 is correctly classified.

$y^{(2)} = w^T x^{(2)}.$

$(1\ 1\ 1) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 2 = 1 = z^{(2)}.$

x_2 is correctly classified.

$y^{(3)} = w^T x^{(3)}$

$[1\ 1\ 1) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 3 = z^{(3)}$

x_3 is correctly classified.

$y^{(4)} = w^T x^{(4)}$

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

x_4 is classified incorrectly. We should modify the weights.

$w_{new} = 1\ 1\ 1 + (0 - 1) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$.

$y^{(1)} = w^T x^{(1)}$.

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

x_1 is correctly classified.

$y^{(2)} = w^T x^{(2)}$

$(0\ 1\ 1) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = 1 = z^{(2)}$.

x_2 is correctly classified.

$y^{(3)} = w^T x^{(3)}$.

$(0\ 1\ 1) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = 2 = z^{(3)}$.

x_3 is correctly classified.

$y^{(4)} = w^T x^{(4)}$

$(0\ 1\ 1) \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 0 = z^{(4)}$.

x_4 is correctly classified.

Scanned with CamScanner