(1)

Name : Ruchit Bhardwaj

USCID : 1111417799

Ans (1) $n = 64$

$$\hat{y} = 2.178 + 0.469x_1 + 3.369x_2 + 3.054x_3$$
$$\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3$$

a) $\hat{\beta}_1$ interprets the studen't's actual score in the exam based on the expected score in the exam for the same student.

b) $CI(\beta_2) = \hat{\beta_2} \pm 2 \cdot SE(\hat{\beta_2})$
$= 3.369 \pm 2(0.456)$
$= 3.369 \pm 0.912$
$= [3.369 - 0.912, 3.369 + 0.912]$
$= [2.457, 4.281]$ = confidence interval.

c) $H_0 \Rightarrow \beta_3 = 0$, $\alpha = 0.05$
$H_A \Rightarrow \beta_3 \neq 0$

calculating t-statistic for $\beta_3 = \frac{\hat{\beta_3} - 0}{SE(\hat{\beta_3})} = \frac{3.054}{1.457} = 2.096$.

$t_{n-p-1, \alpha/2} \Rightarrow t_{64-3-1, 0.05/2} \Rightarrow t_{60, 0.025}$

for t-table, $t_{60, 0.025} = 2.000$

Since $2.096 > 2.000$, we can reject Null Hyp.
It means that β_3 is corresponding to x_3 is significant.
Interpret = studen't score on actual exam, depends on his/her grade point average.

②

d) $t\text{-statistic}(\beta_1) = \frac{\hat{\beta}_1 - 0}{SE(\hat{\beta}_1)} = \frac{0.469}{0.090} = 5.211$

$\text{——"——}(\beta_2) = \frac{\hat{\beta}_2 - 0}{SE(\hat{\beta}_2)} = \frac{3.369}{0.456} = 7.388$

$\text{——"——}(\beta_3) = \frac{\hat{\beta}_3 - 0}{SE(\hat{\beta}_3)} = \frac{3.054}{1.457} = 2.096.$

& $t_{n-p-1,\alpha/2} = t_{60,0.01/2} = t_{60,0.005} = 2.660$

$H_0 = \beta_1 = \beta_2 = \beta_3 = 0.$

$H_A = \beta_1 = \beta_2 = \beta_3 \neq 0.$

For β_1, $5.211 > 2.660$, so, we reject null hypo.
This means $\beta_1 \neq 0$.

For β_2, $7.388 > 2.660$, reject NH, means $\beta_2 \neq 0$.

For β_3, $2.096 < 2.660$, we cannot reject Null Hyp.
so, $\beta_3 = 0$.

Overall, we cannot reject Null hypo.
Interpret : for $\alpha = 0.01$, x_1 & x_2 are significant predictors
i.e. it means students pred. marks depend on
actual score & hours/week working on course but
does not depend on GPA.

3

e) $x_1 = 80$
$x_2 = 8$
$x_3 = 3.0$

$$\hat{y} = 2.178 + 0.469(80) + 3.369(8) + 3.054(3)$$
$$= \boxed{75.812}$$

Q2) According to the given dataset

Index	x_1	x_2	Y	$\hat{Y}$	Errors
1	0	-1	+	+	0
2	0	0	−	+	1
3	-2	3	−	+	1
4	12	1	★	+	1
5	-5	7	−	+	1
6	1	-9	+	+	0
7	19	-10	+	+	0
8	0	15	★	+	1
9	12	-4	+	+	0
					$\Sigma = 5$

If we use LOOCV, we will leave out every index once. After that classification will happen based on the remaining instances (8 instances).
eg: leaving index = 1, Y are (3+, 2★, 3−).
Based on rule given, + will be preferred.
so, $\hat{Y}$ for $(0,-1) = +$
We will do this for all. Summarized in table

$\Sigma_{errors} = 5$

Total count $= 9$.

Misclassification estimate $= 5/9 = 0.\bar{5} = \boxed{0.5555}$

Q3) a) $k = \{1, 2\}$ & single feature binary classification.

$$f_1(x) = \frac{1}{2} \exp\left(-\frac{x}{2}\right) \quad \& \quad f_2(x) = \frac{x}{4} \exp\left(-\frac{x}{2}\right).$$

We know,

$$\gamma_1(x) = \cancel{\frac{\pi_1 f_1(x)}{\sum_{l=1}^{2} \pi_l f_l(x)}} = \cancel{\frac{\pi_1 f_1(x)}{\pi_1 f_1(x) + \pi_2 f_2(x)}} = \frac{\pi_1 p_{x_1}(x)}{\sum_{l=1}^{2} \pi_l p_{x_l}(x)}$$

Calculating, $p_{x_1}(x)$, ~~assuming gaussian dis~~

$$\gamma_1(x) = \frac{\pi_1 f_1(x)}{\sum_{l=1}^{2} \pi_l f_l(x)} = \frac{\pi_1 f_1(x)}{\pi_1 f_1(x) + \pi_2 f_2(x)}$$

$$= \frac{\pi_1 \left[\frac{1}{2} e^{-x/2}\right]}{\pi_1 \left[\frac{1}{2} e^{-x/2}\right] + \pi_2 \left[\frac{x}{4} e^{-x/2}\right]} =$$

(5)

given $\pi_1 = \pi_2 = \pi$

$$\gamma_1(x) = \frac{\pi/2 \left(e^{-x/2}\right)}{\frac{\pi}{2} e^{-x/2} + \frac{\pi x}{4} e^{-x/2}} = \frac{\pi/2}{\frac{\pi}{2} + \frac{\pi x}{4}} = \boxed{\frac{1}{1 + x/2}}$$

similarly for $\gamma_2(x) = \dfrac{\pi_2 f_2(x)}{\pi_1 f_1(x) + \pi_2 f_2(x)}$

$$= \frac{\frac{\pi x}{4} e^{-x/2}}{\frac{\pi}{2} e^{-x/2} + \frac{\pi x}{4} e^{-x/2}} = \frac{\pi x/4}{\pi/2 + \pi x/4} = \boxed{\frac{x/2}{1 + x/2}}$$

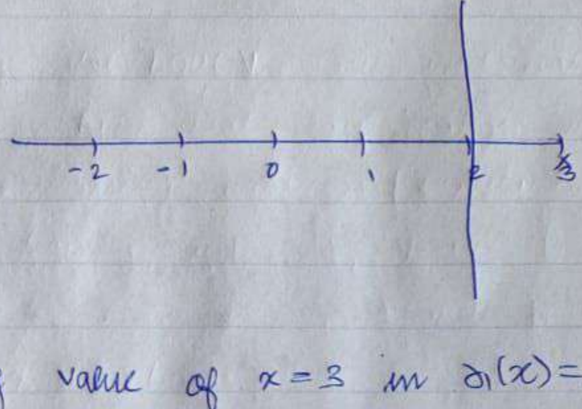
b) To find decision boundary, $\gamma_1(x) = \gamma_2(x) = 1/2$

$$\frac{1}{1 + x/2} = \frac{x/2}{1 + x/2}$$

$$1 = x/2$$

$$x = 2$$

Drawing a graph.



so, $x = 3$ is classified in class $\boxed{k = 2}$

setting value of $x = 3$ in $\gamma_1(x) = \dfrac{1}{1 + 3/2} = 0.4$.

& $\gamma_2(x) = \dfrac{3/2}{1 + 3/2} = \dfrac{1.5}{2.5} = 3/5 = 0.6$

⑥

Q3) $f_1(x) = \frac{1}{2} e^{-x/2}$

$$P_1(x) = \frac{\pi_1 f_1(x)}{\sum_{l=1}^{2} \pi_l f_l(x)} = \frac{\pi_1 \cdot \frac{1}{2} e^{-x/2}}{\pi_1 \cdot \frac{1}{2} e^{-x/2} + \pi_2 \cdot \frac{x}{4} e^{-x/2}}$$

= Taking log, ignoring denominater since will be same for $\delta_1(x)$ & $\delta_2(x)$.

$$\boxed{\delta_1(x) = \log(P_1(x)) = \log \pi_1 + \log \frac{1}{2} - \frac{x}{2}}$$

Doing same for $P_2(x) = \frac{\pi_2 f_2(x)}{\sum_{l=1}^{2} \pi_l f_l(x)}$.

$$= \frac{\pi_2 \cdot \frac{x}{4} e^{-x/2}}{\pi_1 \cdot \frac{1}{2} e^{-x/2} + \pi_2 \frac{x}{4} e^{-x/2}}$$

$$\boxed{\delta_2(x) = \text{Taking log}, \quad \log \pi_2 + \log \frac{1}{4} + \log x - \frac{x}{2}}$$

ⓑ To find decision boundary $\delta_1(x) = \delta_2(x)$

since $\pi_1 = \pi_2$,

$$\delta_1(x) = \log \pi + \log \frac{1}{2} - \frac{x}{2}$$

$$\delta_2(x) = \log \pi + \log \frac{1}{4} + \log x - \frac{x}{2}$$

Equating, $\log \frac{1}{2} = \log \frac{1}{4} + \log x$.

6.5

$\log x = \log \frac{1}{2} - \log \frac{1}{4}$.

$\log x = \log \frac{1/2}{1/4}$

$\boxed{x = 2}$.

so, $x = 3$ classified to class $\boxed{k = 2}$.

⑦

Q4) $n = 498$ & $p_k(x) = \dfrac{e^{\beta_{0k} + \beta_{1k}x_1 + \beta_{2k}x_2 + \beta_{3k}x_3}}{\sum_{k=1}^{3} e^{\beta_{0k} + \beta_{1k}x_1 + \beta_{2k}x_2 + \beta_{3k}x_3}}$

$p = 3$

$\alpha = 0.05$

Calculating $t_{n-p-1, \alpha/2} = t_{498-3-1, 0.05/2} = t_{494, 0.025}$

From table, $\boxed{t_{494, 0.025} = 1.960}$

$t(\hat{\beta}_{11}) = \dfrac{\hat{\beta}_{11} - 0}{SE(\hat{\beta}_{11})} = \dfrac{-2}{s_1}$

For features to be statistically significant, we need to reject null hypothesis.

so, $\dfrac{-2}{s_1} \geq 1.960 \rightarrow s_1 < \dfrac{-2}{1.960} \rightarrow s_1 \leq -1.020$.

$\boxed{(s_1)_{max} = -1.020}$

similarly for all,

$\dfrac{-1}{s_2} \geq 1.960 \rightarrow s_2 \leq \dfrac{-1}{1.960} \rightarrow s_2 \leq -0.510$

$\boxed{(s_2)_{max} = -0.510}$

$\dfrac{1.5}{s_3} \geq 1.960 \rightarrow s_3 \leq \dfrac{1.5}{1.960} \rightarrow s_3 \leq 0.765$

$\boxed{(s_3)_{max} = 0.765}$

$$\frac{-2.5}{S_4} \geq 1.960 \rightarrow S_4 \leq \frac{-2.5}{1.960} \rightarrow S_4 \leq -1.275.$$

$$\boxed{S_{4max} = 1.275}$$

$$\frac{2}{S_5} \geq 1.960 \rightarrow S_5 \leq \frac{2}{1.960} \rightarrow S_5 \leq 1.020$$

$$\boxed{(S_5)_{max} = 1.020}$$

b) $P_1(0,0,-1) = \dfrac{e^{\beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3}}{Den} = e^{1+(-2)(0)+(-1)(0}$

$$= \frac{e^{1+(-2)(0)+(-1)(0)+1.5(-1)}}{Den} = \frac{e^{1-1.5}}{Den} = \frac{e^{-0.5}}{Den}$$

$$P_2(0,0,-1) = \frac{e^{\beta_{02}+\beta_{12}x_1+\beta_{22}x_2+\beta_{32}x_3}}{Den.}$$

$$= \frac{e^{0+0(0)+(-2.5)(0)+0(-1)}}{Den} = \frac{e^0}{Den} =$$

$$P_3(0,0,-1) = \frac{e^{\beta_{03}+\beta_{13}x_1+\beta_{23}x_2+\beta_{33}x_3}}{Den} = \frac{e^{0+0(0)+0(0)+2(-1)}}{Den}$$

$$= \frac{e^{-2}}{Den}$$

Comparing. $\dfrac{e^0}{Den} = 1$, so, $\boxed{x^* = (0,0,-1) \text{ classified in class 2}}$

Q4)c) To calculate decision boundaries, we take log & calculate

$\delta_1(x_1, x_2, x_3) = \delta_2(x_1, x_2, x_3) = \delta_3(x_1, x_2, x_3) = 0$

denominator will be common, so ignoring that

→ $k = 1$.

$$P_1(x) = \frac{e^{\beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3}}{D}$$

$$\log P_1(x) = \beta_{01} + \beta_{11}x_1 + \beta_{21}x_2 + \beta_{31}x_3$$

$$= 1 + (-2x_1) - x_2 + 1.5x_3$$

$$\boxed{\delta_1(X) = 1 - 2x_1 - x_2 + 1.5x_3} \quad \text{——— (1)}$$

→ $k = 2$.

$$P_2(x) = e^{\beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3}$$

$$\log P_2(x) = \beta_{02} + \beta_{12}x_1 + \beta_{22}x_2 + \beta_{32}x_3$$

$$= 0 + 0 - 2.5x_2 + 0$$

$$\boxed{\delta_2(x) = -2.5x_2} \quad \text{——— (2)}$$

→ $k = 3$

$$P_3(x) = e^{\beta_{03} + \beta_{13}x_1 + \beta_{23}x_2 + \beta_{33}x_3}$$

$$\log(P_3(x)) = \beta_{03} + \beta_{13}x_1 + \beta_{23}x_2 + \beta_{33}x_3$$

$$= 0 + 0 + 0 + 2x_3$$

$$\boxed{\delta_3(x) = 2x_3} \quad \text{——— (3)}$$

9 10

for Decision Boundary between

→ 1 & 2 $\quad g_1(x) = g_2(x)$

$$1 - 2x_1 - x_2 + 1.5x_3 = -2.5x_2$$

$$\boxed{\cancel{1 - 2x_1 + 2x_2 + 1.5x_3 = 0}} \qquad \boxed{1 - 2x_1 + 1.5x_2 + 1.5x_3 = 0}$$

→ 1 & 3 $\quad g_1(x) = g_3(x)$

$$1 - 2x_1 - x_2 + 1.5x_3 = 2x_3$$

$$1 - 2x_1 - x_2 \cancel{++} - 0.5x_3 = 0$$

$$\boxed{1 - 2x_1 + x_2 - 0.5x_3 = 0}$$

→ 2 & 3 $\quad g_2(x) = g_3(x)$

$$-2.5x_2 = 2x_3$$

$$\boxed{2.5x_2 + 2x_3 = 0}$$

(11)

Q5) shrinkage penalty = $\lambda \sum_j \beta_j^2$ is small if coefficient ~ 0

$j = \{1, 4)$

(a) $\beta_2^2 = \lambda(\beta_1^2 + \beta_2^2 + \beta_3^2 + \beta_4^2)$

$$\lambda = \frac{\beta^2}{\sum_{j=1}^{4} \beta_j^2}$$

Not valid as when $\beta \to 0$, $\lambda \uparrow$ shrinkage $\times$ valid

(b) same logic as above.

$$\lambda = \frac{\beta_1^5 + \beta_2^5 + \beta_3^5 + \beta_4^5}{\sum_{j=1}^{4} \beta_j^2}$$

$\beta \to 0$ (β tends to 0), $\lambda \downarrow$ so, valid shrinkage.

(c) $\lambda = \dfrac{|\beta_1| + \beta_2^2 + |\beta_3| + \beta_4^6}{\sum_{j=1}^{4} \beta_j^2}$

valid shrinkage, $\beta \to 0$, $\lambda \downarrow$

(d) $\lambda = \dfrac{\beta_1^2 + |\beta_2| + \beta_3^6 + |\beta_4|}{\sum_{j=1}^{4} \beta_j^2}$ $\beta \to 0$, $\lambda \downarrow$

valid shrinkage

12

(e) $\lambda = \frac{\sqrt{|\beta_1|} + \beta_3^2}{\sum_{j=1}^{4} \beta_j^2}$

$\beta \to 0$, $\lambda \uparrow$ $\rightarrow$ NOT valid shrinkage