1

Chris Hanson
DSCI 552
Mid term
10/13/20

Student ID: 2537-9103-07

a) b_0 is the amount of the response y that can not be attributed to the predictors X.

b_5 indicates that if the founders had previous failures, they are likely to obtain an additional \$17,100 in their fundraising

b) β_3 : -200.2 $SE_3 = 101.8$ $n = 26$ $K = 5$ $\alpha = 0.05$; $\alpha/2 = 0.025$

$t_{n-k-1, \alpha/2} = t_{26-5-1, 0.025} = t_{20, 0.025} = 2.086$ (from table)

$t = \frac{\beta_i - \beta_0}{SE(\beta_1)} = \frac{-200.2 - 0}{101.8} = -1.966$; $|t| = 1.966 < 2.086$

Thus we accept the null hypothesis.

c) 80% C.I. for β_1 : $\beta_1 = 700.2$ $SE_1 = 12$

$CI = \beta \pm (t_{, n-k-1, \alpha/2}) \cdot SE$; $700.2 \pm (1.352) \cdot 12 =$
$t_{, 20, 0.1}$
$700.2 \pm 16.2 = \boxed{684 < \beta_1 < 716.4}$ 80% CI

d) SSR = 18147.5 RSS = 17136.5 $\alpha = 0.01$

$F_{p, n-p-1, \alpha} = \frac{(TSS - RSS)/p}{RSS/(n-p-1)} = \frac{18147.5/5}{17136.5/(26-5-1)} = \frac{3,629.5}{856.8} = 4.24$

$F_{5, 20, 0.01}$

From table:

$F_{5,20,0.01} = 4.103$; $\boxed{4.24 > 4.103, \text{ thus we reject the null}}$

e) $R^2 = \frac{TSS - RSS}{TSS}$; $TSS = SSR + RSS = 35,284$; $R^2 = \frac{35,284 - 17,136.5}{35,284} = 0.514$

② $f_{K=1}(x) = \frac{x}{\sigma_1^2} \exp\left(\frac{-x^2}{2\sigma_1^2}\right), \quad x \geq 0$

Chris Hanson
DSCI 552
Midterm 1
10/13/20

$f_{K=2}(x) = \frac{1}{x\sqrt{2\pi}\sigma^2} \exp\left(\frac{-(\ln x - \mu_2)^2}{2\sigma_2^2}\right), \quad x \geq 0$

a) The discriminant function will be a linear function of x when $\sigma_1 = \sigma_2$.

b)

1) $f_1(x) = \frac{10}{1^2} \exp\left(\frac{-10^2}{2 \cdot 1^2}\right) = 10 \cdot \exp\left(\frac{-100}{2}\right) = \underline{1.92 \times 10^{-21}}$

2) $f_2(x) = \frac{1}{10\sqrt{2\pi} \cdot 1} \cdot \exp\left(-\frac{(\ln(10) - 10)^2}{2 \cdot 1^2}\right) = \frac{1}{25} \cdot \exp\left(-\frac{(2.3-10)^2}{2}\right)$

$\cdot \exp\left(\frac{-59.3}{2}\right)$

$= \underline{5.31 \times 10^{-15}}$

$\frac{1}{25} \cdot (1.32 \times 10^{-13})$

$\star \; \Pr(Y=1 \mid X=10) = \frac{\pi_k f_k(x)}{\sum_{l=1}^{K} \pi_l f_l(x)}$ $\quad \pi_1 = \pi_2 = 0.5$

i) $\frac{(0.5)(1.92 \times 10^{-21})}{(0.5)(1.92 \times 10^{-21}) + (0.5)(5.31 \times 10^{-15})}$

$= \frac{9.6 \times 10^{-22}}{(9.6 \times 10^{-22}) + (2.65 \times 10^{-15})}$

$= 3.62 \times 10^{-7}$

ii) $\frac{(0.5)(5.31 \times 10^{-15})}{(0.5)(1.92 \times 10^{-21}) + (0.5)(5.31 \times 10^{-15})}$

$= \frac{2.65 \times 10^{-15}}{(9.6 \times 10^{-22}) + (2.65 \times 10^{-15})}$

$= 0.999$

X=10 will be classified to

K=2

③

Chris Hanson
DSCI 552
Midterm
10/13/20

a) Best subset selection will have the smallest training RSS, (or have the same training RSS as the other models).

b) This would need to be determined with cross validation.

c) i) True

ii) True

iii) False

iv) False

v) False

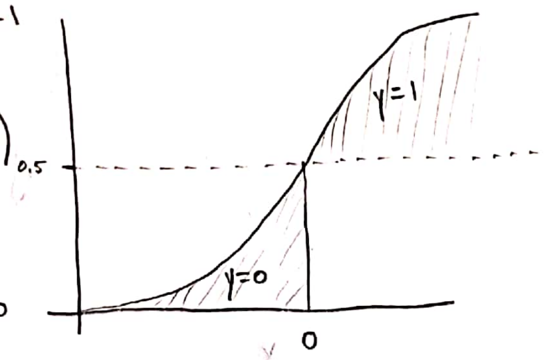
Chris Hanson
DSCI 552
Midterm
10/13/20

4

a) $0.5 = \dfrac{\exp(1+2X_1+X_1^2+X_2)}{1+\exp(1+2X_1+X_1^2+X_2)}$

b) Y=1

P(Y)



Y=0

$\underline{X}$ $(1+2X_1+X_1^2+X_2)$

c) $\underline{\alpha = 0.05}$ $\underline{\text{Insignificant}}$

$\beta_2 = 1$; $SE = s$,

$t = \dfrac{\beta}{s}$

$t_{n-k-1,\alpha/2} = t_{200-3-1,\,0.025}$

$= t_{196,\,0.025} = \sim \underline{1.97}$

$t = \frac{1}{s} < 1.97$; $\boxed{s \geq 0.51}$

5

Chris Hanson
DSCI 552
Midterm
10/13/20