Math 485 HW 6

(6.1)

1)

a) $$\log\left(\frac{\hat{\pi}_D}{\hat{\pi}_I}\right) = 3.3 - 0.2x$$

$$\log\left(\frac{\hat{\pi}_R}{\hat{\pi}_I}\right) = 1.0 + 0.3x$$

$$\log\left(\frac{\hat{\pi}_R}{\hat{\pi}_D}\right) = \log\left(\frac{\hat{\pi}_R/\hat{\pi}_I}{\hat{\pi}_D/\hat{\pi}_I}\right) = \log\left(\frac{\hat{\pi}_R}{\hat{\pi}_I}\right) - \log\left(\frac{\hat{\pi}_D}{\hat{\pi}_I}\right)$$

$$= 1 + .3x - (3.3 - 0.2x)$$

$$= -2.3 + .5x$$

$\Rightarrow \hat{\beta} = .5$

So the estimated odds of preferring Republicans is $e^{.5} = 1.65$ times the estimated odds of that of Democrats.

In other words, the estimated odds of preferring Republicans over Democrats increase by 65% for every unit increase ($10,000) in annual income.

b) If $\hat{\pi}_R = \hat{\pi}_D \Rightarrow \log\left(\frac{\hat{\pi}_R}{\hat{\pi}_D}\right) = \log(1) = 0$

or $-2.3 + .5x = 0 \Rightarrow x = 4.6$

Therefore $\hat{\pi}_R > \hat{\pi}_D$ if $x > 4.6$ or If the annual income > $46,0

(6.1)

1) continue

c) $$\hat{\pi}_I = \frac{1}{1+e^{\alpha_1+\beta_1 x}+e^{\alpha_2+\beta_2 x}}$$

$$= \frac{1}{1+e^{3.3-.2x}+e^{1+.3x}}$$

recall: In multicategory logit model with J category we have

$$\pi_j = \frac{e^{\alpha_j+\beta_j x}}{1+e^{\alpha_1+\beta_1 x}+\cdots+e^{\alpha_{J-1}+\beta_{J-1} x}}, \quad j=1,2,\dots,J-1$$

$$\pi_J = \frac{1}{1+e^{\alpha_1+\beta_1 x}+\cdots+e^{\alpha_{J-1}+\beta_{J-1} x}}$$

In this Problem, we have $j = 1, 2, 3$ ($1 \to D$, $2 \to R$, $3 \to I$), so $J=3$.

(6.3)

2)

a) see the R code.

Let

$$S = \begin{cases} 1 & \text{if size} \leq 2.3 \\ 0 & \text{if siz} > 2.3 \end{cases}$$

$$L_1 = \begin{cases} 1 & \text{lake is "Hancook"} \\ 0 & \text{otherwise} \end{cases}, \quad L_2 = \begin{cases} 1 & \text{lake is "oklawaha"} \\ 0 & \text{otherwise} \end{cases}$$

$$L_3 = \begin{cases} 1 & \text{lake is "Trafford"} \\ 0 & \text{otherwise} \end{cases}$$

When $L_1 = L_2 = L_3 = 0$ then it refers to the lake "George"

We have:

$$Log\left(\frac{\hat{\pi}_j}{\hat{\pi}_J}\right) = \hat{\alpha}_j + \hat{\beta}_{1j} S + \hat{\beta}_{2j} L_1 + \hat{\beta}_{3j} L_2 + \hat{\beta}_{4j} L_3 \quad , j = 1, \dots, J-1$$

From R output we got:

$$\log\left(\frac{\hat{\pi}_I}{\hat{\pi}_F}\right) = -1.55 + 1.46\, S - 1.66\, L_1 + .94\, L_2 + 1.12\, L_3$$

$$\log\left(\frac{\hat{\pi}_R}{\hat{\pi}_F}\right) = -3.31 - .35\, S + 1.24\, L_1 + 2.46\, L_2 + 2.94\, L_3$$

$$\log\left(\frac{\hat{\pi}_B}{\hat{\pi}_F}\right) = -2.09 - .63\, S + .7\, L_1 - .65\, L_2 + 1.09\, L_3$$

$$\log\left(\frac{\hat{\pi}_O}{\hat{\pi}_F}\right) = -1.9 + .33\, S + .83\, L_1 + 0.01\, L_2 + 1.52\, L_3$$

6.3

2) continue

b) We are interested in $\hat{\pi}_F$ when $L_2 = 1$ ( lake oklawaha) and $S = 1$ ( for the short length ) and $S = 0$ ( for long length).

Sine Fish is the baseline, therefore

$$\hat{\pi}_F = \frac{1}{1 + e^{\alpha_1 + \beta_{11} S + \cdots + \beta_{41} L_3} + \cdots + e^{1 + \beta_{1(J-1)} S + \cdots + \beta_{4(J-1)} L_3}} = \frac{1}{A}$$

$S = 1,\ L_2 = 1$
$L_1 = L_3 = 0$

$$A = 1 + e^{-1.55 + 1.46 + .94} + e^{-3.31 - .35 + 2.46} + e^{-2.09 - .63 - .65}$$
$$+ e^{-1.9 + .33 + .01} = 3.8853$$

$$\Rightarrow \boxed{\hat{\pi}_F = \frac{1}{3.8853} = 0.2573} \longrightarrow$$ Aligators in the lake "oklavaha" with the size of less tha 2.3 who their primery food choice is Fish.

$S = 0,\ L_2 = 1$
$L_1 = L_3 = 0$

$$\Rightarrow A = 1 + e^{-1.55 + .94} + e^{-3.31 + 2.46} + e^{-2.09 - .65} + e^{-1.9 + .01} = 2.1864$$

$$\boxed{\hat{\pi}_F = \frac{1}{2.1864} = .4573}$$

So, the probabity that the longer Aligator in this lake who use Fish is higher than of that with smaller size aligators.

(6.5)

3) $Y$ = job satisfaction

$$\text{logit}(\hat{P}(Y \leq j)) = \hat{\alpha}_j - 0.5x_1 + 0.6x_2 + 1.19x_3$$

Since the order of $Y$ is from least satisfied to very satisfied, and the coefficient of $x_1$ is negative, and the coefficients of $x_2$ & $x_3$ are positive, therefor, job satisfaction will decrease in lower levels of $x_1$ and higher levels of $x_2$ & $x_3$. i.e.

$$\hat{P}(Y \leq j) = \frac{e^{\alpha_j - .5x_1 + .6x_2 + 1.19x_3}}{1 + e^{\alpha_j - .5x_1 + .6x_2 + 1.19x_3}}$$

Notice that $\frac{e^x}{1+e^x}$ is a increasing function respect to $x$.

$$\left( \frac{\partial}{\partial x}\left(\frac{e^x}{1+e^x}\right) = \frac{e^x}{(1+e^x)^2} > 0 \right)$$

In other words, job satisfaction will increase in higher levels of $x_1$ and lower levels of $x_2$ & $x_3$.

(6.5)

3) continue

b) From parat (a), we know that if $x_1$ takes its higher level and $x_2$ & $x_3$ take their lower levels, then the highest job satisfaction will be achieved. That is,

when $x_1 = 4$ and $x_2 = x_3 = 1$, then we have

$$\hat{p}(Y \leq j) = \frac{e^{\alpha_j - (.5)(4) + .6(1) + 1.19(1)}}{e^{\alpha_j - (.5)(4) + .6(1) + 1.19(1)}}$$

$$= \frac{e^{\alpha_j - .21}}{1 + e^{-.21}}$$

---

(6.6)

4)

a)

$\pi_1$ = probability of Not happiness
$\pi_2$ = " " pretty "
$\pi_3$ = " " very "

$x = 1, 2, 3$ different levels of income.

$$\log\left(\frac{\hat{\pi}_1}{\hat{\pi}_3}\right) = -2.5551 - 0.2275\, x$$

$$\log\left(\frac{\hat{\pi}_2}{\hat{\pi}_3}\right) = -0.351 - 0.0962\, x$$

( 6.6 )

4) continue

b) $\log\left(\frac{\hat{\pi}_1}{\hat{\pi}_3}\right) = -2.5552 - 0.2275\,x$

$\hat{\beta} = -0.2275 \Rightarrow e^{\hat{\beta}} = e^{-.2275} = .796$

The estimated odds of being no happy versus being very happy multiplied by $e^{-.2275} = .796$ for each one unit increase in incom

That is, the estimated odds of being no happy increases as income increases.

---

c) $H_0: \beta = 0$
$H_A: \beta \neq 0$

wald statistics = .9432

P-value = .624

$\alpha <$ P-value, therefore, fail to reject the null hypothesis.

i.e., it seems that income has no effect of the happiness of the marital.

d) $H_0$: the fitted model is adequate
$H: \sim H_0$

Deviance = 3.19
df = 2, P-value = .2

$\alpha <$ P-value $\rightarrow$ Fail to Reject $H_0 \rightarrow$

$\Rightarrow$ the model fits adequately.

4) continue $\pi_3 = \cdots$ a very happy ~

Average family in income $\Rightarrow \boxed{x = 2}$

e)

Recall: $$\hat{\pi}_1 = \frac{e^{\hat{\alpha}_1 + \hat{\beta}_1 x}}{1 + e^{\hat{\alpha}_1 \hat{\beta}_1 x} + e^{\hat{\alpha}_2 + \hat{\beta}_2 x}}$$

$$\hat{\pi}_2 = \frac{e^{\hat{\alpha}_2 + \hat{\beta}_2 x}}{1 + e^{\hat{\alpha}_1 + \hat{\beta}_1 x} + e^{\hat{\alpha}_2 + \hat{\beta}_2 x}}$$

$$\hat{\pi}_3 = \frac{1}{1 + e^{\hat{\alpha}_1 + \hat{\beta}_1 x + \hat{\alpha}_2 + \hat{\beta}_2 x}}$$

$$= \frac{1}{1 + e^{[-2.5551 - .2275(2)]} + e^{[-.3514 - .0962(2)]}}$$

$$= 0.61.$$

---

(6.7)

5) a) With $\underline{3}$ response categories, there are two cumulative probabilities to model, and therefore two intercept parameters. The propotional odds have the same predictor effect for each cumulative probability. That is, the curve has the same shape. Therefore, only one effect is reported for income.

6.1

5) continue

b) $\hat{\beta} = -.1117$, here again the coefficient is negative.

Therefore, the estimated odds of being less happy decrease as income increases.

---

c) $H_0: \beta = 0$

$H_A: \beta \neq 0$

LR = .8876, df=1, P-value= .3461

cannot reject the null hypothesis, so, it seems that income has no effect on the happiness of marital.

---

d) $H_0$: the fitted model is adequate

$H_A: \sim H_0$

Deviance = 3.2472, df=3, P-value= .355

$\alpha <$ P-value $\Rightarrow$ cannot reject $H_0$ $\Rightarrow$ the fitted model is adequate.

---

Very happy

e) $\hat{P}(Y=3) = 1 - P(\hat{Y} \leq 2) = 1 - \dfrac{e^{(-.2378-(.1117\times 2))}}{1+e^{(-.2378-(.1117\times 2))}}$

$x = 2$
Average income

$= 1 - .386 = .613$

6.8

6) let Y: Response to chemotherapy
( Progressive Disease, No-change, Partial Remission,
and complete Remission )

$x_1$ = gender ( 1 = M , 0 = F )

$X_2$ = Therapy ( 1 = sequential , 0 = Alternating )

$$logit(P(Y \leq j)) = \alpha_j + \beta_1 X_1 + \beta_2 X_2, \quad j = 1, 2, 3$$

a) see the R code. Based on the results from R output, we have

$$logit(\hat{P}(Y \leq j)) = \alpha_j - 0.5414 X_1 - 0.5807 X_2$$

$\beta_1 = -.5414$ , $\beta_2 = -0.580$

controling for gender, estimated odds that the sequential therapy is in negative direction (no remission, $Y \leq j$) rather than positive direction (partial / or complete remission, $Y > j$) are $e^{-.5807} \approx .56$ times estimated odds for Alternating therapy.

6) continue

a) continue

Seems, there is no evidence of gender effect when controlling for therapy.

$H_0: \beta_1 = 0$

$H_A: \beta_1 \neq 0$

Wald $z = -1.83$

P-value $= 0.066$

$\alpha = .05 < .066 \Rightarrow$ Fail to reject the null hypothesis.

However, if we assume that the gender is somewhat significant, then controlling for therapy, the estimated odds that males are in negative direction $(Y \leq j)$ rather than positive direction $(Y > j)$ is $e^{-.5414} = .58$ times estimated odds for Female.

b) $logit(\hat{P}(Y \leq j)) = \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_1 X_2$

$= \alpha_j - 0.864 X_1 - 1.078 X_2 + .59041 X_1 X_2$

$H_0: \beta_3 = 0$

$H_A: \beta_3 \neq 0$

, Wald $z = 0.995$, p-value $= 0.319$

the interaction is not significant.

(6.8)

6) continue

b) continue

let us to have the intraction term in the model to see how the estimated treatment effect varies by gender:

Estimated odds ratio for treatmen effect $(x_2)$ is:

$$e^{-1.078} = .34 \quad \text{when } x_1 = 0 \text{ (Female)}$$

$$e^{-1.078 + .59041} = .61 \quad \text{when } x_2 = 1 \text{ (Male)}$$

---

c) $H_0: \beta_3 = 0$ $\Leftrightarrow$ $H_0$: model without intraction is adequate

$H_A: \beta_3 \neq 0$ $\quad$ $H_A$: Model with intraction is adequate

LR: $\text{deviance}_0 - \text{deviance}_A = 5.56 - 4.52 = 1.046$

df = 1 $\quad$ P-value = .3062

No, intraction model does not give a better fit.

7) 6.9

a) Since there are five category in the respons variable, therefore there are four cumulative probabilities that are strictly greater than zero, and strictly less than one.

$( 0 < P(Y \leq j) < 1 \ , \text{if } j = 1,2,3,4 )$

For $j=5$, $P(Y \leq 5) = 1$.

When all predictor values are equal zero, the cumulative probability increase across categories.

Sine logit$(x)$ increasing function of $x$, when cumulative probability increese, so logit will increase, as do parameters $\alpha_j$ that specify logits.

b) (i) Since all of the coefficients are negative, the most liberal would happen if all predictors $X_1 = X_2 = X_3 = 0$.

and this is equivalent to "None".

$X_1 = \begin{cases} 1 & \text{if Protestant} \\ 0 & \text{otherwise} \end{cases}$ $X_2 = \begin{cases} 1 & \text{if catholic} \\ 0 & \text{otherwis} \end{cases}$

$X_3 = \begin{cases} 1 & \text{if Jewish} \\ 0 & \text{o.w} \end{cases}$ None if $X_1 = X_2 = X_3 = 0$

6.9

7) continue

b) ii) the most conservative group would be the one with smallest coefficient, and that is $\beta_1 = -1.27$, that is when $X_1 = 1$, i.e. the group of Protestant.

c) $$\text{logit}(\hat{P}(Y \leq j)) = \alpha_j + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

the most libral $\equiv j = 1$, $x_1 = 1$

$$\text{logit}(\hat{P}(Y \leq 1)) = \alpha_1 + \beta_1 X_1$$

$$\Rightarrow \hat{P}(Y \leq 1) = \hat{P}(Y = 1) = \frac{e^{-1.03 - 1.27}}{1 + e^{-1.03 - 1.27}} = .091$$

and for "None" group, we have $X_1 = X_2 = X_3 = 0$

$$\hat{P}(Y = 1) = \hat{P}(Y \leq 1) = \frac{e^{-1.03}}{1 + e^{-1.03}} = .2631$$

7) continue

d) (i) The estimated odds that protestant falls in liberal direction ($Y \leq j$) rather than conservative categories ($Y > j$) is $e^{-1.27} = .28$ times estimated odds for someone with no religious preference.

(ii) Notice that at $j=1$

$$\frac{\text{odds of "protestant"}}{\text{odds of "None"}} = \frac{e^{\alpha_1 + \beta_1}}{e^{\alpha_1}} = e^{\beta_1}$$

$x_1 = 1$

Also at $j=1$

$$\frac{\text{odds of catholic}}{\text{odds of "None"}} = \frac{e^{\alpha_1 + \beta_2}}{e^{\alpha_1}} = e^{\beta_2}$$

Now,

$$\frac{\text{odds of protestant}}{\text{odds of catholic}} = \frac{e^{\alpha_1 + \beta_1}}{e^{\alpha_1 + \beta_2}} = e^{\beta_1 - \beta_2}$$

Therefore, we have the estimated odds ratio comparing protestant to catholic is $e^{\beta_1 - \beta_2} = e^{-1.27-(-1.22)} = e^{-.05} = .951$

6.12

8)

a) The fitted mode is given as

$$\text{logit}(P(Y \leqslant j)) = \alpha_j + \beta x$$
$$= \alpha_j - 0.3843\, x$$

where $x$ = Religion taking 1, 2 and 3.

$y$ = response is the relating happiness

$\beta = -.3843$, so, the estimated probability of being relatively unhappy decreases as religious attendence increases.

Wald $z = -8.45$ for testing $H_0: \beta = 0$, P-value $\simeq 0$

So, the effect is strongly significant.

b) $H_0$: The fitted model is adequate

$H_A$: $\sim H_0$

Deviance $= 0.6178$, df $= 3$, P-value $= .89$

$\Rightarrow$ Fail to reject, $\Rightarrow$ the fit is adequate.

6.22

9) a) True b) True c) False d) True.