STK3100 Exercises, Week 9

Vinnie Ko, Jonas Moss

October 25, 2018

Exercise 5.24

i)
Since f is symmetric around 0, $F^{-1}(0.5) = 0$. Thus, if $0.5 = F(\beta_0 + \beta_1 x_i)$, then $\beta_0 + \beta_1 x_i = 0$ and $x_i = -\dfrac{\beta_0}{\beta_1}$

ii)
The rate of change is $\dfrac{\partial \pi_i}{\partial x_i} = \dfrac{\partial \pi_i}{\partial \eta_i}\dfrac{\partial \eta_i}{\partial x_i} = f(\eta_i)\beta_1$. When $\pi_i = 0.5$, we have $\eta_i = 0$ and the rate of change is $\beta_1 f(0)$.

For the probit link, $f(0) = \dfrac{1}{\sqrt{2\pi}}$, the rate of change is $\dfrac{1}{\sqrt{2\pi}}\beta_{1,\text{probit}}$.
For the logit link, $f(0) = 0.25$, the change of rate is $0.25\beta_{1,\text{logit}}$.

Since the rate of change should be approximately the same whether the probit or logit link is used, the estimates for the logistic model should be approximately $\dfrac{4}{\sqrt{2\pi}} \approx 1.6$ times the estimates for the probit model.

Exercise 5.25

i)
With complementary log-log link we have

$$
\begin{aligned}
\log\left(-\log(1-\pi_i)\right) &= \beta_0 + \beta_1 x_i \\
\log\left(\frac{1}{1-\pi_i}\right) &= \exp\left[\beta_0 + \beta_1 x_i\right] \\
\frac{1}{1-\pi_i} &= \exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right] \\
\pi_i &= 1 - \frac{1}{\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]}
\end{aligned}
$$

When $\pi_i = 0.5$,

$$\begin{aligned} 2 &= \exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right] \\ \log\left(\log 2\right) &= \beta_0 + \beta_1 x_i \\ x_i &= \frac{\log\left(\log 2\right) - \beta_0}{\beta_1} = \frac{-0.3665 - \beta_0}{\beta_1}. \end{aligned}$$

The rate of change is

$$\begin{aligned} \frac{\partial \pi_i}{\partial x_i} &= \frac{\partial \pi_i}{\partial \eta_i}\frac{\partial \eta_i}{\partial x_i} \\ &= \frac{\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right] \cdot \exp\left[\beta_0 + \beta_1 x_i\right]}{\left(\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]\right)^2}\beta_1 \\ &= \frac{\exp\left[\beta_0 + \beta_1 x_i\right]}{\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]}\beta_1. \end{aligned}$$

We find x_i that maximizes the rate of change

$$\begin{aligned} \frac{\partial^2 \pi_i}{\partial x_i^2} &= \frac{\partial}{\partial x_i}\left(\frac{\partial \pi_i}{\partial x_i}\right) \\ &= \frac{\partial}{\partial x_i}\left(\frac{\exp\left[\beta_0 + \beta_1 x_i\right]}{\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]}\beta_1\right) \\ &= \frac{\beta_1 \exp\left[\beta_0 + \beta_1 x_i\right]\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right] - \exp\left[\beta_0 + \beta_1 x_i\right]\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]\exp\left[\beta_0 + \beta_1 x_i\right]\beta_1}{\left(\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]\right)^2}\beta_1 \\ &= \frac{\exp\left[\beta_0 + \beta_1 x_i\right] - \exp\left[\beta_0 + \beta_1 x_i\right]\exp\left[\beta_0 + \beta_1 x_i\right]}{\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]}\beta_1^2 \\ &= \frac{\exp\left[\beta_0 + \beta_1 x_i\right]\left(1 - \exp\left[\beta_0 + \beta_1 x_i\right]\right)}{\exp\left[\exp\left[\beta_0 + \beta_1 x_i\right]\right]}\beta_1^2 \\ &= 0. \end{aligned}$$

So, the rate of change reaches its maximum when $\exp\left[\beta_0 + \beta_1 x_i\right] = 1$. We have then $\beta_0 + \beta_1 x_i = 0$ and $x_i = -\frac{\beta_0}{\beta_1}$. When $x_i = -\frac{\beta_0}{\beta_1}$, $\pi_i = 1 - \frac{1}{e} = 0.6321$.

ii)
Repeat i) for log-log link

$$-\log\left(-\log \pi_i\right) = \beta_0 + \beta_1 x_i$$

Exercise 5.26

It's given that $Y_i \sim \text{Pois}\left(\mu_i = \exp\left[\sum_j \beta_j x_{i,j}\right]\right)$ and $Z_i = I(Y_i > 0)$.
So, $Z_i \sim \text{Bin}\left(1, \text{P}(Y_i > 0)\right)$.
Notice that $\pi_i = \text{P}(Y_i > 0) = 1 - \text{P}(Y_i \leq 0) = 1 - \text{P}(Y_i = 0) = 1 - e^{-\mu_i} = 1 - \exp\left[-\exp\left[\eta_i\right]\right] = 1 - \exp\left[-\exp\left[\sum_j \beta_j x_{i,j}\right]\right]$.

This gives $\eta_i = \log(-\log(1-\pi_i))$.

Exercise 6.3

We use vector notation $\boldsymbol{x}_i = \left[x_{i,1}, \cdots, x_{i,p}\right]$ and $\boldsymbol{\beta}_j = \left[\beta_{j,1}, \cdots, \beta_{j,p}\right]^{\mathrm{T}}$.

$\pi_{i,j}$ is defined as

$$\pi_{i,j} = \begin{cases} \dfrac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]}{1+\sum_{h=1}^{c-1}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]} & \text{for} \quad j = 1, \cdots, c-1 \\ \\ \dfrac{1}{1+\sum_{h=1}^{c-1}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]} & \text{for} \quad j = c \end{cases}.$$

Since the restriction we impose is $\boldsymbol{\beta}_c = \mathbf{0}$, I prefer to merge these two cases into one:

$$\pi_{i,j} = \frac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]} \quad \text{for} \quad j = 1, \cdots, c.$$

Thus, the rate of change is

$$\begin{aligned}
\frac{\partial \pi_{i,j}}{\partial x_{i,k}} &= \frac{\partial}{\partial x_{i,k}}\left(\frac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}\right) \\
&= \frac{\beta_{j,k}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]\cdot\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right] - \exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]\cdot\sum_{h=1}^{c}\beta_{h,k}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}{\left(\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]\right)^2} \\
&= \frac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}\left(\frac{\beta_{j,k}\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right] - \sum_{h=1}^{c}\beta_{h,k}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}\right) \\
&= \frac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}\left(\beta_{j,k} - \frac{\sum_{h=1}^{c}\beta_{h,k}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}\right) \\
&= \frac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}\left(\beta_{j,k} - \sum_{h=1}^{c}\left[\frac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]}\beta_{h,k}\right]\right) \\
&= \pi_{i,j}\left(\beta_{j,k} - \sum_{h=1}^{c}\left[\pi_{i,h}\beta_{h,k}\right]\right) \\
&= \pi_{i,j}\left(\beta_{j,k} - \sum_{h=1}^{c-1}\left[\pi_{i,h}\beta_{h,k}\right]\right)
\end{aligned}$$

In binary case (i.e. $c = 2$), $j = 1$. So, the rate of change simplifies to

$$\frac{\partial \pi_i}{\partial x_{i,k}} = \beta_k\pi_i(1-\pi_i).$$

Exercise 6.4

When there is 1 covariant and $c = 3$ (and assuming that the last group is the reference case),

$$\pi_{i,j} = \frac{\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_j\right]}{\sum_{h=1}^{c}\exp\left[\boldsymbol{x}_i\boldsymbol{\beta}_h\right]} = \frac{\exp\left[\beta_{j,0}+\beta_{j,1}x_i\right]}{1+\exp\left[\beta_{1,0}+\beta_{1,1}x_i\right]+\exp\left[\beta_{2,0}+\beta_{2,1}x_i\right]}.$$

Thus,

$$\pi_{i,3} = \frac{1}{1+\exp\left[\beta_{1,0}+\beta_{1,1}x_i\right]+\exp\left[\beta_{2,0}+\beta_{2,1}x_i\right]}$$

and

$$\frac{\partial \pi_{i,3}}{\partial x_i} = \frac{\partial}{\partial x_i}\left(\frac{1}{1+\exp\left[\beta_{1,0}+\beta_{1,1}x_i\right]+\exp\left[\beta_{2,0}+\beta_{2,1}x_i\right]}\right) = -\frac{\beta_{1,1}\exp\left[\beta_{1,0}+\beta_{1,1}x_i\right]+\beta_{2,1}\exp\left[\beta_{2,0}+\beta_{2,1}x_i\right]}{\left(1+\exp\left[\beta_{1,0}+\beta_{1,1}x_i\right]+\exp\left[\beta_{2,0}+\beta_{2,1}x_i\right]\right)^2}.$$

a)

If $\beta_{1,1} > 0$ and $\beta_{2,1} > 0$, then $\frac{\partial \pi_{i,3}}{\partial x_i} < 0$. So, $\pi_{i,3}$ is decreasing in x_i.

b)

If $\beta_{1,1} < 0$ and $\beta_{2,1} < 0$, then $\frac{\partial \pi_{i,3}}{\partial x_i} > 0$. So, $\pi_{i,3}$ is increasing in x_i.

c)

If $\beta_{1,1}$ and $\beta_{2,1}$ have opposite signs, $\pi_{i,3}$ can be both negative and positive. So, there is no guarantee that $\pi_{i,3}$ is monotone in x_i.

Additional Exercise 21

a)

$y \in \{0,1\}$: death by SIDS
$x_1 \in \{1,2,3,4,5\}$: `kohort`
$x_2 \in \{1,2\}$: `kjønn`
$x_3 \in \mathbb{R}^+$: `vekt`

	Additional parameters	Df	Deviance	Resid. Df	Resid. Dev	P(|Chi|)
NULL	β_0			$570 (= n-1)$	1101.92	
vekt	β_3	1	259.59	$569 (= n-2)$	842.33	< 0.001
factor(kohort)	$\beta_{1,1}, \cdots, \beta_{1,4}$	4	314.59	$565 (= n-6)$	527.74	< 0.001
kjonn	β_2	1	92.81	$564 (= n-7)$	434.93	< 0.001
vekt:factor(kohort)	$\beta_{3:1,1}, \cdots, \beta_{3:1,4}$	4	6.37	$560 (= n-11)$	428.56	0.1732
vekt:kjonn	$\beta_{3:2}$	1	0.19	$559 (= n-12)$	428.37	0.6630
factor(kohort):kjonn	$\beta_{2:1,1}, \cdots, \beta_{2:1,4}$	4	15.32	$555 (= n-16)$	413.05	0.0041
vekt:factor(kohort):kjonn	$\beta_{3:2:1,1}, \cdots, \beta_{3:2:1,4}$	4	5.25	$549 (= n-20)$	407.80	0.2626

The deviance table can be used for Likelihood ratio test of model parameters. For example, if we want to test the significance of parameter β_2 (for the variable `kjønn`). We can read off from the deviance table:

$$\begin{aligned}
-2\log\left(\frac{\max_{H_0}\ell\left(\beta_0,\beta_{1,1},\cdots,\beta_{1,4},\beta_2,\beta_3\right)}{\max_{\text{full}}\ell\left(\beta_0,\beta_{1,1},\cdots,\beta_{1,4},\beta_2,\beta_3\right)}\right) &= -2\log\left(\frac{\max\ \ell\left(\beta_0,\beta_{1,1},\cdots,\beta_{1,4},\beta_3\right)}{\max\ \ell\left(\beta_0,\beta_{1,1},\cdots,\beta_{1,4},\beta_2,\beta_3\right)}\right)\\
&= 527.74 - 424.93\\
&= 92.81\\
&> \chi^2_{1,0.95}\\
&= 3.84
\end{aligned}$$

b)

i)
This is similar to what we have done in Problem 1, c) of mandatory assignment 1.
Interpretation of β_j: log of odds ratio when variable j has increased by 1 unit.

ii)
95% confidence interval for odds ratio can be obtained by:
$\exp\left[\widehat{\beta}_j\right] \cdot \exp\left[\pm z_{0.975} \cdot SE(\widehat{\beta}_j)\right]$

For example, for β_3 (`vekt`):

$$\exp\left[\widehat{\beta}_3\right] \cdot \exp\left[\pm z_{0.975} \cdot SE(\widehat{\beta}_3)\right] = \exp\left[-0.6711\right] \cdot \exp\left[\pm 1.96 \cdot 0.03758\right] = [0.4749, 0.5502]$$

c)

i)
Consider a child with covariate vector $\boldsymbol{x}_i$, and let $y_{i,j} = 1$ if the child dies of cause j ($j = 1, \cdots, J$), and $y_{i,j} = 0$ otherwise. Let $y_{i,0} = 1$ if the child survives, and $y_{i,0} = 0$ if she/he dies. Thus, $\pi_{i,j} = P(y_{i,j} = 1)$.

Now, we let $j = 1$ correspond to SIDS. To only consider those who dies of SIDS and survive, we ignore the irrelevant part of the model and condition only on $\{y_{i,0} = 1 \;\text{ or }\; y_{i,1} = 1\}$. This gives:

$$P(y_{i,1} = 1 | y_{i,0} = 1 \text{ or } y_{i,1} = 1) = \frac{P(y_{i,1} = 1)}{P(y_{i,0} = 1) + P(y_{i,1} = 1)} = \frac{\pi_{i,1}}{\pi_{i,0} + \pi_{i,1}} = \frac{\exp\left[\boldsymbol{x}_i \boldsymbol{\beta}_1\right]}{1 + \exp\left[\boldsymbol{x}_i \boldsymbol{\beta}_1\right]}$$

which is a logistic regression model with the same $\boldsymbol{\beta}_1$ as in the multinomial logit model.
(See p.203 of the book.)

ii)
The advantage of using separate logistic regression for each cause, is simplicity.

The disadvantage is that there is no guarantee that $\sum_{j=0}^{J} \widehat{\pi}_{i,j} = 1$.