Michael Leibert

Math 661

Homework 4

1. **Exercise 1 – Agresti # 6.20**

   The following R output shows output from fitting a cumulative logit model to data from the US 2008 General Social Survey. For subject i, let

   - y_i = belief in existence of heaven (1 = yes, 2 = unsure, 3 = no),
   - x_{i1} = gender (1 = female, 0 = male) and
   - x_{i2} = race (1 = black, 0 = white).

```
> cbind(race, gender, y1, y2, y3)
     race gender  y1  y2 y3
[1,]    1      1  88  16  2
[2,]    1      0  54   7  5
[3,]    0      1 397 141 24
[4,]    0      0 235 189 39

> summary(vglm(cbind(y1,y2,y3) ~ gender+race, family=cumulative(parallel=T)))
              Estimate Std. Error z value Pr(>|z|)
(Intercept):1  0.07631    0.08963   0.851    0.395
(Intercept):2  2.32238    0.13522  17.175  < 2e-16 ***
gender         0.76956    0.12253   6.281 3.37e-10 ***
race           1.01645    0.21059   4.827 1.39e-06 ***
---
Residual deviance: 9.2542 on 4 degrees of freedom
Log-likelihood: -23.3814 on 4 degrees of freedom
```

   (a) State the model fitted here and interpret the race and gender effects.

   Belief in heaven is ordinal, with categories (1=yes, 2 = unsure, 3 = no). The survey related y =belief in heaven to two explanatory variables, gender x_{i1} and race x_{i1}. The cumulative logit model of the proportional odds form with main effects has ML fit:

$$\text{logit}\left[\hat{P}\left(y_i \leq j\right)\right] = \hat{\alpha}_j + 0.76956x_{i1} + 1.01645x_{i2}$$

   We interpret the race and gender effects.

   $\hat{\alpha}_1 = 0.07631$: Is the estimated log odds of falling into belief in heaven yes versus all other categories for white male.

   $\hat{\alpha}_2 = 2.32238$: Is the estimated log odds of falling into belief in heaven: yes or belief in heaven: unsure for white male.

$\hat{\beta}_1 = -1.0165$. For a one unit increase in x_{i1}, controlling for x_{i2}, the cumulative probability of falling into group j or lower is higher. That is controlling for gender, blacks are more likely than whites to fall into the categories: belief in heaven: yes or belief in heaven: unsure.

$\hat{\beta}_2 =$. For a one unit increase in x_{i2}, controlling for x_{i1}, the cumulative probability of falling into group j or lower is higher. That is controlling for race, females are more likely than males to fall into the categories: belief in heaven: yes or belief in heaven: unsure.

(b) Test goodness-of-fit and construct confidence intervals for the effects.

1-pchisq(9.254,4)

The p-value, $P\left(\chi_4^2\right) = 0.05505495$, is $> .05$, so there's no evidence that the model does not fit the data well. However, the model has a relatively low goodness-of-fit p-value $= 0.055$, which makes its fit to the data questionable.

A 95% confidence interval given for β_1 is given by:

$$\hat{\beta}_1 \pm z_{0.025}\ SE\left(\hat{\beta}_1\right) = 0.76956 \pm 1.96 \cdot 0.12253 = \big(0.5294012, 1.009719\big).$$

A 95% confidence interval given for β_2 is given by:

$$\hat{\beta}_2 \pm z_{0.025}\ SE\left(\hat{\beta}_2\right) = 1.01645 \pm 1.96 \cdot 0.21059 = \big(0.6036936, 1.429206\big).$$

INTERPRET???

2. **Exercise 2 – Agresti # 6.21**

Refer to the previous exercise. Consider the model

$$\log \frac{\pi_{ij}}{\pi_{i3}} = \alpha_j + \beta_j^G x_{i1} + \beta_j^R x_{i2}, \qquad j = 1, 2.$$

(a) Fit the model and report prediction equations for

$$\log \frac{\pi_{i1}}{\pi_{i3}}, \;\; \log \frac{\pi_{i2}}{\pi_{i3}}, \;\; \log \frac{\pi_{i1}}{\pi_{i2}}.$$

$$\begin{aligned}\log \frac{\pi_{i1}}{\pi_{i3}} &= \log \frac{\dfrac{\exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_1\right)}{\cancel{1 + \sum_{h=1}^{c-1} \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_h\right)}}}{\dfrac{1}{\cancel{1 + \sum_{h=1}^{c-1} \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_h\right)}}} \\ &= \boldsymbol{x}_i^T \boldsymbol{\beta}_1 \\ &= 1.7943 + 1.0339 x_{i1} + 0.6727 x_{i2}\end{aligned}$$

$$\begin{aligned}\log \frac{\pi_{i2}}{\pi_{i3}} &= \log \frac{\dfrac{\exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_2\right)}{\cancel{1 + \sum_{h=1}^{c-1} \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_h\right)}}}{\dfrac{1}{\cancel{1 + \sum_{h=1}^{c-1} \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_h\right)}}} \\ &= \boldsymbol{x}_i^T \boldsymbol{\beta}_2 \\ &= 1.5309 + 0.3087 x_{i1} - 0.4757 x_{i2}\end{aligned}$$

$$\begin{aligned}\log \frac{\pi_{i1}}{\pi_{i2}} &= \log \left(\frac{\exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_1\right)}{\cancel{1 + \sum_{h=1}^{c-1} \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_h\right)}} \frac{\cancel{1 + \sum_{h=1}^{c-1} \exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_h\right)}}{\exp\left(\boldsymbol{x}_i^T \boldsymbol{\beta}_2\right)} \right) \\ &= \log\Big(\exp\left[\boldsymbol{x}_i^T \boldsymbol{\beta}_1\right] \Big) - \log\Big(\exp\left[\boldsymbol{x}_i^T \boldsymbol{\beta}_2\right] \Big) \\ &= \boldsymbol{x}_i^T \boldsymbol{\beta}_1 - \boldsymbol{x}_i^T \boldsymbol{\beta}_2 \\ &= (1.7943 + 1.5309) + (1.0339 + 0.3087)\, x_{i1} + (0.6727 - 0.4757)\, x_{i2} \\ &= 3.3252 + 1.3426 x_{i1} + 0.197 x_{i2}\end{aligned}$$

(b) Using the "yes" and "no" response categories, interpret the conditional gender effect using a 95% confidence interval for the odds ratio.

A 95% confidence interval given for the log odds of β_1^G, the conditional gender effect, is given by:

$$\exp\left(\hat{\beta}_1^G \pm z_{0.025}\; SE\left(\hat{\beta}_1^G\right) \right) = \Big(\exp\left(0.526848\right), \exp\left(1.540952\right) \Big) = \big(1.693586, 4.669033\big).$$

The odds of belief of an afterlife for females is $\exp(1.0339) = 2.8$ times higher than for males, controlling for race. The 95% confidence interval does not contain 1, so there is evidence of the conditional gender effect, controlling for race.

(c) Conduct a likelihood ratio test of the hypothesis that opinion is independent of gender, given race. Interpret.

H_0: opinion is independent of gender, given race