STAT 2132
Orly olbum

HOMEWORK 5

1. $Y_i = \{1, 2, 3\}$, continuous x_i

$P(Y_i = j) = \pi_{ij} > 0$

$\log\left(\frac{\pi_{i1}}{\pi_{i2}+\pi_{i3}}\right) = \alpha_1 - \beta x_i = \log\left(\frac{\pi_{i1}}{1-\pi_{i1}}\right) = \log\left(\frac{Pr(Y\leq 1)}{Pr(Y>1)}\right)$

$\log\left(\frac{\pi_{i1}+\pi_{i2}}{\pi_{i3}}\right) = \alpha_2 - \beta x_i = \log\left(\frac{Pr(Y\leq 2)}{Pr(Y>2)}\right)$

ⓐ Find π_{i2} in terms of $\alpha_1, \alpha_2, \beta, x_i$

$\log\left(\frac{\pi_{i1}}{1-\pi_{i1}}\right) = \text{logit}(\pi_{i1}) = \alpha_1 - \beta x_i$,

so $\pi_{i1} = \dfrac{\exp(\alpha_1 - \beta x_i)}{1+\exp(\alpha_1 - \beta x_i)}$

$\log\left(\frac{\pi_{i1}+\pi_{i2}}{\pi_{i3}}\right) = \log\left(\frac{\pi_{i1}+\pi_{i2}}{1-(\pi_{i1}+\pi_{i2})}\right) = \text{logit}(\pi_{i1}+\pi_{i2}) = \alpha_2 - \beta x_i$,

so $\pi_{i1} + \pi_{i2} = \dfrac{\exp(\alpha_2 - \beta x_i)}{1+\exp(\alpha_2 - \beta x_i)}$

$$\pi_{i2} = (\pi_{i1} + \pi_{i2}) - \pi_{i1}$$

$$= \frac{\exp(\alpha_2 - \beta x_i)}{1+\exp(\alpha_2 - \beta x_i)} - \frac{\exp(\alpha_1 - \beta x_i)}{1+\exp(\alpha_1 - \beta x_i)}$$

$$= \frac{\exp(\alpha_2 - \beta x_i)[1+\exp(\alpha_1 - \beta x_i)] - [1+\exp(\alpha_2 - \beta x_i)]\exp(\alpha_1 - \beta x_i)}{[1+\exp(\alpha_2 - \beta x_i)][1+\exp(\alpha_1 - \beta x_i)]}$$

$$= \frac{\exp(\alpha_2 - \beta x_i) + \cancel{\exp(\alpha_1 + \alpha_2 - 2\beta x)} - \exp(\alpha_1 - \beta x_i) - \cancel{\exp(\alpha_1 + \alpha_2 - 2\beta x)}}{1 + \exp(\alpha_1 - \beta x_i) + \exp(\alpha_2 - \beta x_i) + \exp(\alpha_1 + \alpha_2 - 2\beta x_i)}$$

$$= \frac{\exp(\alpha_2 - \beta x_i) - \exp(\alpha_1 - \beta x_i)}{[1+\exp(\alpha_2 - \beta x_i)][1+\exp(\alpha_1 - \beta x_i)]} = \pi_{i2}$$

ⓑ Show that $\alpha_1 \leq \alpha_2$

If $\alpha_1 \leq \alpha_2$,

$$\log\left(\frac{\pi_{i1}}{\pi_{i2}+\pi_{i3}}\right) - \cancel{\beta x_i} \leq \log\left(\frac{\pi_{i1}+\pi_{i2}}{\pi_{i3}}\right) - \cancel{\beta x_i}$$

$$\frac{\pi_{i1}}{\pi_{i2}+\pi_{i3}} \leq \frac{\pi_{i1}+\pi_{i2}}{\pi_{i3}}$$

$$\cancel{\pi_{i1}\pi_{i3}} \leq \pi_{i1}\pi_{i2} + \cancel{\pi_{i1}\pi_{i3}} + \pi_{i2}^2 + \pi_{i2}\pi_{i3}$$

$$0 \leq \pi_{i1}\pi_{i2} + \pi_{i2}^2 + \pi_{i2}\pi_{i3}$$

$$0 \leq \pi_{i2}(\pi_{i1} + \pi_{i2} + \pi_{i3})$$

$\pi_{i1} + \pi_{i2} + \pi_{i3} = 1$, so $0 \leq \pi_{i2}$ holds because $\pi_{ij} > 0$, and if $Pr(Y_i = 2) = 0$, $\pi_{i2} = 0$

(c) (Mi) $\text{logit}(\pi_{i1}) = 0.593 + 1.963x_i$
(Mii) $\text{logit}(\pi_{i1} + \pi_{i2}) = 3.707 + 2.258x_i$

(see attached plots)

(d) (Mi) $\text{logit}(\pi_{i1}) = 1.584 + 4.684x_i$
(Mii) $\text{logit}(\pi_{i1} + \pi_{i2}) = 2.441 + 1.153x_i$

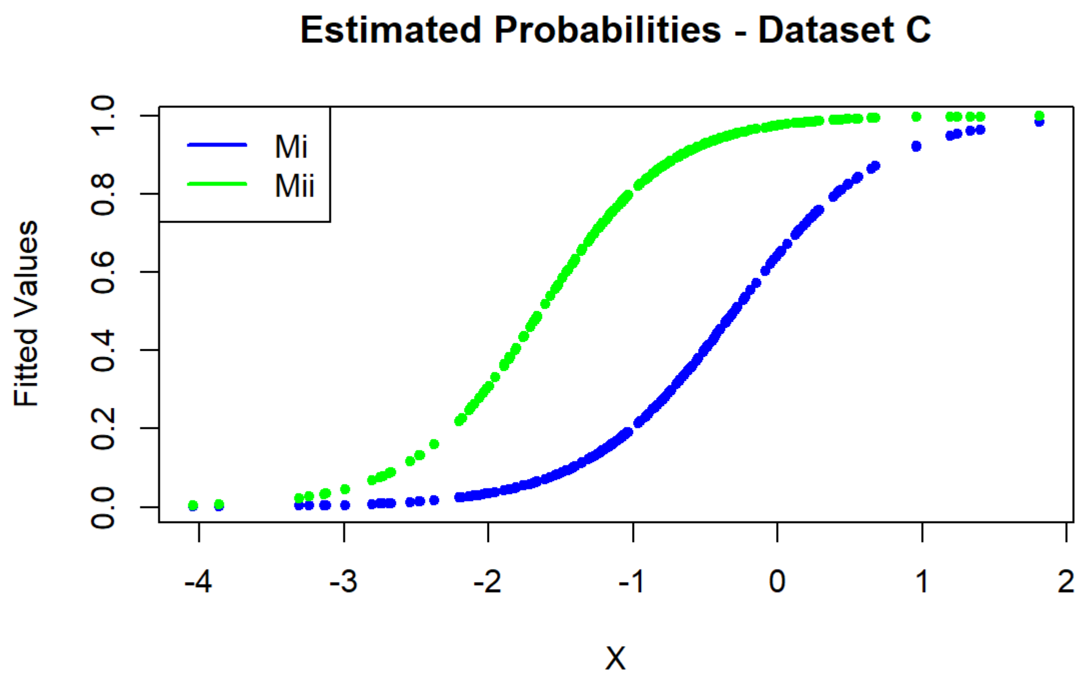
(see attached plots)

(e) Based on the plots, we see that the slopes of c models are very similar, while the slopes of d models are not. Based on the proportional odds model fit, we don't have different slopes for each model, so I wouldn't feel so comfortable fitting a proportional odds model to Dataset D.

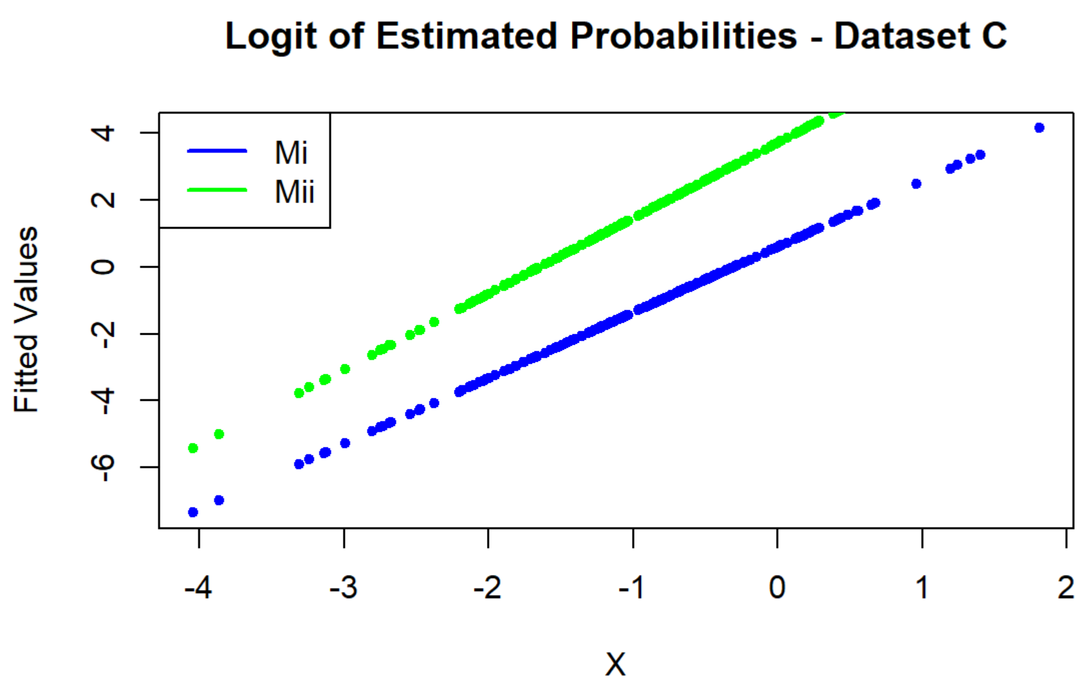
(c)

Using R, make two plots:

(i) On the same axes, plot the estimated probabilities from (Mi) and the estimated probabilities from (Mii) as a function of xi.



(ii) On the same axes, plot the logit of the estimated probabilities from (Mi) and the logit of the estimated probabilities from (Mii) as a function of xi.



(d)

The file Q1d.txt on the Courseweb also contains a sample data set with n = 200. Repeat part (c) using this data set.