exercises_week10

Lindsey Greenhill

4/7/2021

Question 2

Part a

In the code below, I created both an ordered multinomial logit model (mod_1) and probit model (mod_2). See the stargazer table below for the estimates for both these models. For the dependent variable, partyid3, 1 = Independent, 2 = Democrat, and 3 = Republican. The reference category is Independent.

```
# creating multinomial logit (ordered). Chose to include race so the graphical
# plot below looks more normal

mod_1 <- polr(factor(partyid3) ~ age + educ1 + income +
                ideo7 + race,
              data = nes_clean)

# creating probit model

mod_2 <- polr(factor(partyid3) ~ age + educ1 + income + ideo7 + race,
              data = nes_clean,
              method = "probit")

# summary of both models

 summary(mod_1)
```

```
##
## Re-fitting to get Hessian
```

Call: polr(formula = factor(partyid3) ~ age + educ1 + income + ideo7 + race, data = nes_clean)

Coefficients: Value Std. Error t value age -0.006863 0.006103 -1.12453 educ1 0.342579 0.124638 2.74858 income 0.196140 0.093874 2.08939 ideo7 0.723619 0.075035 9.64380 raceblack -1.325481 0.725417 -1.82720 racehispanic -0.060660 0.711235 -0.08529 racenative_american 0.472878 0.761278 0.62116 racewhite 0.496945 0.632633 0.78552

Intercepts: Value Std. Error t value 1|2 4.8669 0.8907 5.4644 2|3 5.2780 0.8955 5.8937

Residual Deviance: 832.957 AIC: 852.957

```
# summary(mod_2)

stargazer(mod_1, mod_2, type = "latex")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Apr 13, 2021 - 17:25:29

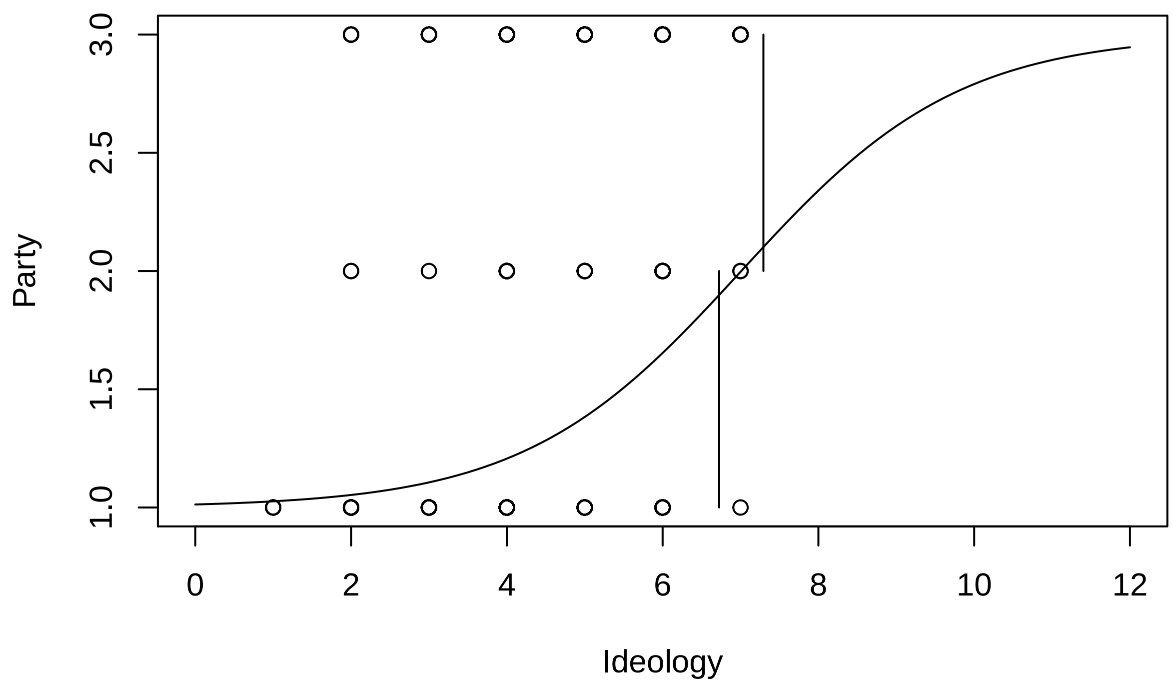
Table 1

	Dependent variable:	
	partyid3	
	ordered logistic	*ordered probit*
	(1)	(2)
age	−0.007	−0.004
	(0.006)	(0.004)
educ1	0.343***	0.195***
	(0.125)	(0.073)
income	0.196**	0.118**
	(0.094)	(0.056)
ideo7	0.724***	0.429***
	(0.075)	(0.043)
raceblack	−1.325*	−0.820*
	(0.725)	(0.431)
racehispanic	−0.061	−0.020
	(0.711)	(0.427)
racenative_american	0.473	0.281
	(0.761)	(0.455)
racewhite	0.497	0.286
	(0.633)	(0.381)
Observations	559	559
Note:	*p<0.1; **p<0.05; ***p<0.01	

In the code below, I create a graphical representation of the multinomial logit model. The plot is based on figure 6.4 from the text book, where the curves show the expected responses for party identification based off of ideology. The data points do not seem to follow the curve very well, suggesting that the model might not be the best fit.

```r
# graphs of both models. Make these like the ones in 6.4 of the textbook. Had to
# create the inverse logit function. Maybe I could have gotten this from the arm
# lib? Not sure

invlogit <- function(Xb) {
  1/(1+exp(-Xb))
}

# taken pretty much directly from the book
```

```r
expected <- function (x, c1.5, c2.5, sigma){
  p1.5 <- invlogit ((x-c1.5)/sigma)
  p2.5 <- invlogit ((x-c2.5)/sigma)
  return ((1*(1-p1.5) + 2*(p1.5-p2.5) + 3*p2.5))
}

# graph of the mod_1. We don't need to do the probit model for this part. It's a
# bit weird that the cut points are outside of the range of the data.

{
plot (nes_clean$ideo7, nes_clean$partyid3, xlim=c(0,12), ylim=c(1,3),
      xlab="Ideology", ylab="Party")
lines (rep (mod_1$zeta[1]/mod_1$coefficients[4], 2), c(1,2))
lines (rep (mod_1$zeta[2]/mod_1$coefficients[4], 2), c(2,3))
curve (expected (
  x, mod_1$zeta[1]/mod_1$coefficients[4],
  mod_1$zeta[2]/mod_1$coefficients[4],
  1/mod_1$coefficients[4]),
  add=TRUE)
}
```



Part b

Explain the results of the fitted model

I thought about the results of the fitted model by looking at the statistical significance of each estimate and the value of that estimate relative to the cutoff points of the ordered model

- Cut points: Democrat:Independent 4.8669, Independent:Republican 5.278
- Statistical significance: the educ1, income, and ideo7 coefficients are statistically significant at the alpha = .05 level. The age and race variables are not statistically significant.
- educ1 coefficient = .343: the education coefficient (.343) is less than the difference between the cut points (~.41). This means that an increase in education by itself will not actually move a voter from one political party to another.
- income coefficient = .196: the income bracket coefficient (.196) is elss than the difference between the cutoff points, similar to the education coefficient. Thus, a one increase in income categories doesn't move a voter on average from one party to another.
- ideo7 coefficient = .724: the ideology coefficient (.724) is greater than the difference between the cut points, meaning that an increase on the ideology scale has a relatively larger effect than the education or income scales and can bring a voter from one party to another party on average.

Question 3

Data

I am using the rodents data from the week 7 pset.

Part a

In the code below I create a logit (mod_3) and probit (mod_4) model.

```
# logit model

mod_3 <- glm(rodent2 ~ race + dilap + regext , data = rodents,
             family = binomial(link = "logit"))

# probit model

mod_4 <- glm(rodent2 ~ race + dilap + regext , data = rodents,
             family = binomial(link = "probit"))

# displaying the two models

stargazer(mod_3, mod_4, type = "latex")
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Tue, Apr 13, 2021 - 17:25:29

The logit and the probit estimates should be more or less the same after scaling by 1.6. We can check this by multiplying the probit coefficients by 1.6. After making this adjustments, we see that the estimates roughly match up with eachother (see table below).

Table 2

	Dependent variable:	
	rodent2	
	logistic	*probit*
	(1)	(2)
racePuerto Rican	1.396***	0.789***
	(0.226)	(0.131)
raceBlack	1.545***	0.880***
	(0.179)	(0.101)
raceOther Hispanic	1.944***	1.124***
	(0.200)	(0.116)
raceAsian/Pacific Islander	0.363	0.196
	(0.302)	(0.165)
raceAmer-Indian/Native Alaskan	2.137***	1.249**
	(0.829)	(0.517)
raceTwo or More Races	0.801	0.432
	(0.803)	(0.458)
dilap	1.257***	0.750***
	(0.294)	(0.177)
regext	−0.196	−0.099
	(0.137)	(0.079)
Constant	−2.072***	−1.216***
	(0.147)	(0.078)
Observations	1,341	1,341
Log Likelihood	−671.830	−672.341
Akaike Inf. Crit.	1,361.660	1,362.682
Note:	*p<0.1; **p<0.05; ***p<0.01	

```r
# scaling the coefficients by 1.6

new_coefs <- mod_4$coefficients * 1.6

# making table of new coefficients. Using gt for better presentation

coefs_tb <- tibble(mod_3_est = mod_3$coefficients,
       mod_4_est = mod_4$coefficients,
       mod_4_est_scaled_1.6 = new_coefs) %>%
  gt() %>%
  fmt_number(columns = everything(),
             decimals = 2)

# displaying table

coefs_tb
```

mod_3_est	mod_4_est	mod_4_est_scaled_1.6
−2.07	−1.22	−1.95
1.40	0.79	1.26
1.55	0.88	1.41
1.94	1.12	1.80
0.36	0.20	0.31
2.14	1.25	2.00
0.80	0.43	0.69
1.26	0.75	1.20
−0.20	−0.10	−0.16