exercises_week6

Lindsey Greenhill

3/3/2021

Exercises 5.1, 5.4

Question 1

Part a

```
# model 1 is a regression of bush vote vs gender, education, party, ideology,
# race, and income

mod <- glm(bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + income, data =  df_clean,
    family = "binomial")

# model 2 is a regression of bush vote vs gender, education, party, ideology,
# race, and an interaction between income and gender

mod_2 <- glm(bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + income + income:gender,
             data =  df_clean,
    family = "binomial")

# model 3 is a regression of bush vote vs gender, education, party, ideology,
# race, and an interaction between education and gender. I took out income
# because it wasn't statistically significant in the past 2 models.

mod_3 <- glm(bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + educ1:gender,
             data =  df_clean,
    family = "binomial")


display(mod)
```

glm(formula = bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + income, family = "binomial", data = df_clean) coef.est coef.se (Intercept) -4.25 0.79
gender 0.41 0.21
educ1 0.17 0.12
partyid3I 1.86 0.31
partyid3R 4.06 0.24
ideo7 0.48 0.08
race1 -0.79 0.61
race2 -2.69 0.75

race3 -0.59 1.02
race5 -0.03 0.71
income 0.01 0.10
— n = 1133, k = 11 residual deviance = 675.5, null deviance = 1534.1 (difference = 858.6)

```
summary(mod_2)
```

Call: glm(formula = bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + income + income:gender, family = "binomial", data = df_clean)

Deviance Residuals: Min 1Q Median 3Q Max
-2.5311 -0.4056 -0.2007 0.4064 3.1769

Coefficients: Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.39386 0.87248 -5.036 4.75e-07 ***gender 0.64167 0.64147 1.000 0.317162***
educ1 0.17189 0.12369 1.390 0.164631
partyid3I 1.86030 0.31064 5.989 2.12e-09 partyid3R 4.05710 0.23544 17.232 < 2e-16 ***ideo7 0.47867***
0.07890 6.067 1.30e-09 race1 -0.77495 0.60700 -1.277 0.201707
race2 -2.69654 0.74642 -3.613 0.000303 *** race3 -0.56925 1.02013 -0.558 0.576832
race5 -0.00981 0.70866 -0.014 0.988955
income 0.05878 0.15783 0.372 0.709588
gender:income -0.07483 0.19385 -0.386 0.699500
— Signif. codes: 0 ' *** 0.001 ** 0.01* " 0.05 ' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 1534.10  on 1132  degrees of freedom
```

Residual deviance: 675.35 on 1121 degrees of freedom AIC: 699.35

Number of Fisher Scoring iterations: 6

```
summary(mod_3)
```

Call: glm(formula = bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + educ1:gender, family = "binomial", data = df_clean)

Deviance Residuals: Min 1Q Median 3Q Max
-2.5166 -0.4109 -0.2069 0.4018 3.1692

Coefficients: Estimate Std. Error z value Pr(>|z|)
(Intercept) -4.18566 0.84864 -4.932 8.13e-07 ***gender 0.32817 0.65566 0.501 0.616704***
educ1 0.16547 0.16223 1.020 0.307745
partyid3I 1.85927 0.31089 5.981 2.22e-09 partyid3R 4.06470 0.23517 17.284 < 2e-16 ***ideo7 0.48209***
0.07877 6.121 9.32e-10 race1 -0.79209 0.60948 -1.300 0.193736
race2 -2.70234 0.74812 -3.612 0.000304 *** race3 -0.59870 1.01671 -0.589 0.555958
race5 -0.04490 0.70880 -0.063 0.949496
gender:educ1 0.02801 0.22271 0.126 0.899904
— Signif. codes: 0 ' *** 0.001 ** 0.01* " 0.05 ' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

```
Null deviance: 1534.1  on 1132  degrees of freedom
```

Residual deviance: 675.5 on 1122 degrees of freedom AIC: 697.5

Number of Fisher Scoring iterations: 6

```r
library(gtsummary)
```

```
##
## Attaching package: 'gtsummary'
```

```
## The following object is masked from 'package:MASS':
##
##     select
```

```r
stargazer(mod, mod_2, mod_3, type = "latex",
          keep.stat = c("null.dev", "res.dev"))
```

% Table created by stargazer v.5.2.2 by Marek Hlavac, Harvard University. E-mail: hlavac at fas.harvard.edu
% Date and time: Thu, Mar 04, 2021 - 09:40:52

```r
# need to add null.dev and res.dev

anova(mod, mod_2, mod_3)
```

Analysis of Deviance Table

Model 1: bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + income Model 2: bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + income + income:gender Model 3: bush_vote ~ gender + educ1 + partyid3 + ideo7 + race + educ1:gender Resid. Df Resid. Dev Df Deviance 1 1122 675.50
2 1121 675.35 1 0.14924 3 1122 675.50 -1 -0.14905

Table 1:

	Dependent variable:		
	bush_vote		
	(1)	(2)	(3)
gender	0.408*	0.642	0.328
	(0.211)	(0.641)	(0.656)
educ1	0.174	0.172	0.165
	(0.123)	(0.124)	(0.162)
partyid3I	1.856***	1.860***	1.859***
	(0.310)	(0.311)	(0.311)
partyid3R	4.060***	4.057***	4.065***
	(0.235)	(0.235)	(0.235)
ideo7	0.481***	0.479***	0.482***
	(0.079)	(0.079)	(0.079)
race1	−0.786	−0.775	−0.792
	(0.606)	(0.607)	(0.609)
race2	−2.692***	−2.697***	−2.702***
	(0.746)	(0.746)	(0.748)
race3	−0.593	−0.569	−0.599
	(1.017)	(1.020)	(1.017)
race5	−0.032	−0.010	−0.045
	(0.706)	(0.709)	(0.709)
income	0.013	0.059	
	(0.103)	(0.158)	
gender:income		−0.075	
		(0.194)	
gender:educ1			0.028
			(0.223)
Constant	−4.253***	−4.394***	−4.186***
	(0.790)	(0.872)	(0.849)
Note:			*p<0.1; **p<0.05; ***p<0.01