Problem Set 3

Cianna Devitt (17321885)

Due: March 26, 2023

Instructions

- Please show your work! You may lose points by simply writing in the answer. If the problem requires you to execute commands in `R`, please include the code you used to get your answers. Please also include the `.R` file that contains your code. If you are not sure if work needs to be shown for a particular problem, please ask.
- Your homework should be submitted electronically on GitHub in `.pdf` form.
- This problem set is due before 23:59 on Sunday March 26, 2023. No late assignments will be accepted.

Question 1

We are interested in how governments' management of public resources impacts economic prosperity. Our data come from Alvarez, Cheibub, Limongi, and Przeworski (1996) and is labelled `gdpChange.csv` on GitHub. The dataset covers 135 countries observed between 1950 or the year of independence or the first year forwhich data on economic growth are available ("entry year"), and 1990 or the last year for which data on economic growth are available ("exit year"). The unit of analysis is a particular country during a particular year, for a total $>$ 3,500 observations.

- Response variable:
  - `GDPWdiff`: Difference in GDP between year t and $t-1$. Possible categories include: "positive", "negative", or "no change"
- Explanatory variables:
  - `REG`: 1=Democracy; 0=Non-Democracy
  - `OIL`: 1=if the average ratio of fuel exports to total exports in 1984-86 exceeded 50%; 0= otherwise

Please answer the following questions:

1. Construct and interpret an unordered multinomial logit with `GDPWdiff` as the output and "no change" as the reference category, including the estimated cutoff points and coefficients.
   `Load in data and set factors, releveled with 'no change' as reference category`

```
gdpChange$GDPWdiff <- factor(ifelse(sign(gdpChange$GDPWdiff) == -1, 'negative',
ifelse(sign(gdpChange$GDPWdiff) == 0, 'no change','positive')))


gdpChange$GDPWdiff <- relevel(gdpChange$GDPWdiff, ref = 'no change')

```

   `Running base multinomial logit unordered`

```
multinom_model1 <- multinom(GDPWdiff ~ REG + OIL,
data = gdpChange)

summary(multinom_model1)


Coefficients:
(Intercept)      REG      OIL
negative     3.805370 1.379282 4.783968
positive     4.533759 1.769007 4.576321

Std. Errors:
(Intercept)       REG      OIL
negative    0.2706832 0.7686958 6.885366
positive    0.2692006 0.7670366 6.885097

Residual Deviance: 4678.77
AIC: 4690.77
```

   `In a given country, there is an increase baseline odds of 1.76 that the difference in GDP will be positive.`

2. Construct and interpret an ordered multinomial logit with `GDPWdiff` as the outcome variable, including the estimated cutoff points and coefficients.
   `Running base multinomial logit ordered`

```
multinom_model2 <- polr(GDPWdiff ~ REG + OIL,
data = gdpChange)

summary(multinom_model2)


Coefficients:
Value Std. Error t value
```

```
REG  0.4102     0.07518    5.456
OIL −0.1788     0.11546   −1.549

Intercepts:
Value     Std. Error t value
no change|negative   −5.3199    0.2523    −21.0878
negative|positive    −0.7036    0.0476    −14.7933

Residual Deviance: 4686.606
AIC: 4694.606
```

Finding proportional odds ratios

```
ci <- confint(multinom_model2)
confint.default(multinom_model2)
exp(cbind(OR= coef(multinom_model2), ci))


  OR      2.5 %    97.5 %
REG 1.5070726 1.3012858 1.747374
OIL 0.8362455 0.6680959 1.050857
```

For a unit increase in democratic status (REG), the odds of GDP difference increase by 1.5, holding constant all other variables For a unit increase in ratio of fuel exports (OIL), the odds of being more likely positive GDP difference is 0.836, holding all other variables constant.

Question 2

Consider the data set `MexicoMuniData.csv`, which includes municipal-level information from Mexico. The outcome of interest is the number of times the winning PAN presidential candidate in 2006 (`PAN.visits.06`) visited a district leading up to the 2009 federal elections, which is a count. Our main predictor of interest is whether the district was highly contested, or whether it was not (the PAN or their opponents have electoral security) in the previous federal elections during 2000 (`competitive.district`), which is binary (1=close/swing district, 0="safe seat"). We also include `marginality.06` (a measure of poverty) and `PAN.governor.06` (a dummy for whether the state has a PAN-affiliated governor) as additional control variables.

(a) Run a Poisson regression because the outcome is a count variable. Is there evidence that PAN presidential candidates visit swing districts more? Provide a test statistic and p-value.

Running Poisson Regression Model

```
Mex_poisson <- glm(PAN.visits.06 ~ competitive.district + marginality
.06
```

```
    + PAN.governor.06, data = MexicoMuniData, family = poisson)

    summary(Mex_poisson)


    Coefficients:
    Estimate Std. Error z value
    (Intercept)           -3.81023    0.22209 -17.156
    competitive.district  -0.08135    0.17069  -0.477
    marginality.06        -2.08014    0.11734 -17.728
    PAN.governor.06       -0.31158    0.16673  -1.869
    Pr(>|z|)
    (Intercept)            <2e-16 ***
    competitive.district   0.6336
    marginality.06         <2e-16 ***
    PAN.governor.06        0.0617 .
    ---
    Signif. codes:
    0    ***    0.001    **    0.01    *    0.05    .    0.1         1

    (Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1473.87  on 2406  degrees of freedom
    Residual deviance:  991.25  on 2403  degrees of freedom
    AIC: 1299.2

    Number of Fisher Scoring iterations: 7
```

(b) With a unit increase in 'marginality.06' ie a measure of poverty, had a diminished liklihood by 2.08 of having candidate visitations, holding all other variables constant. With a unit increase in 'PAN.governor.06' ie whetther state has a PAN=affiliated governor, had a diminished liklihood by 0.31 of having candidate visitations, holding all other variables constant.

(c) Provide the estimated mean number of visits from the winning PAN presidential candidate for a hypothetical district that was competitive (`competitive.district`=1), had an average poverty level (`marginality.06` = 0), and a PAN governor (`PAN.governor.06`=1).

Getting Fitted Values

$\lambda_i = e_o^{\beta} + \beta_{1Xi}$

```
    lambda30 <- exp(coeffs[1] + coeffs[2]*1)
    lambda30
```

```
    (Intercept)
    0.02041293

```