Problem Set 3

Lucy Kinnear

Due: March 28, 2022

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.

```r
    # data wrangling
data$GDPWdiff <- replace(data$GDPWdiff, data$GDPWdiff < 0, '-A')
data$GDPWdiff <- replace(data$GDPWdiff, data$GDPWdiff > 0, 'B')

data$GDPWdiff <- factor(data$GDPWdiff,
                        levels = c( "-A" , 0, "B"),
                        labels = c("negative",
                                   "no change",
                                   "positive"))
# unordered multinomial logit
data$GDPWdiff <- relevel(data$GDPWdiff, ref= "no change")
mult.log <- multinom(GDPWdiff ~ REG + OIL, data = data)
summary(mult.log)
stargazer(mult.log)

# p-values
z <- summary(mult.log)$coefficients/summary(mult.log)$standard.errors
(p <- (1 - pnorm(abs(z), 0, 1)) * 2)
stargazer(p)

# coefficients
exp(coef(mult.log))

```

```
24  # confidence
25  confidence <- (ci <- confint(mult.log))
26  stargazer(p)
```

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

```
1   ordinal.log <- polr(GDPWdiff ~ REG + OIL, data = data, Hess = TRUE)
2   summary(ordinal.log)
3   stargazer(ordinal.log)
4
5   # p value
6   ctable <- coef(summary(ordinal.log))
7   p <- pnorm(abs(ctable[, "t value"]), lower.tail = FALSE) * 2
8   (ctable <- cbind(ctable, "p value" = p))
9
10  # confidence
11  (ci <- confint(ordinal.log))
12
13  # converting to an odds ratio
14  exp(cbind(OR = coef(ordinal.log), ci))
```

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.

```
1  mexico <- read.csv('MexicoMuniData.csv')
2  summary(mexico)
3  # outcome: PAN.visits.06
4  # predictors: competitive.district, marginality.06, PAN.governor.06
5
6  # poisson regression
7  # is there evidence that PAN presidential candidates visit swing districts more?
8  # t statistic and p-value
9
10 pos.reg <- glm(PAN.visits.06 ~ competitive.district + marginality.06 + PAN.governor.06,
       data = mexico, family = poisson)
11 summary(pos.reg)
12 stargazer(pos.reg)
```

(b) Interpret the `marginality.06` and `PAN.governor.06` coefficients. The coefficient of marginality is statistically distinguishable from zero, but PAN.governor's coefficient is not. Marginality appears to have a negative relationship with winning candidate visits.

(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`).
0.01494818

```
1   # estimated mean number of visits from the winning PAN presidential candidate
2  # hypothetical competitive district (competitive.district = 1)
3  # average poverty level (marginality.06 = 0)
4  # PAN governor (PAN.governor.06 = 1)
5  coeff <- coef(pos.reg)
6  coeff
7  est.visits <- exp(coeff[1] + coeff[2]*1 + coeff[3]*0 + coeff[4]*1)
8  est.visits
9
10
```