Problem Set 3

Aryan Goyal

March 2024

First, I load in the data and inspect it:

```
gdp_data <- read.csv("https://raw.githubusercontent.com/ASDS-
TCD/StatsII_Spring2024/main/datasets/gdpChange.csv", stringsAsFactors = F)

head(gdp_data)
```

1 Question 1

1.1 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.

```
   # First, I convert the GDP to a factor variable
gdp_data[gdp_data$GDPWdiff==0, "factor_GDP"] <- "no change"
gdp_data[gdp_data$GDPWdiff>0, "factor_GDP"] <- "positive"
gdp_data[gdp_data$GDPWdiff<0, "factor_GDP"] <- "negative"
gdp_data$factor_GDP <- relevel(as.factor(gdp_data$factor_GDP), ref="no change")
```

I run an unordered multinomial logit with releveled GDP variable as the output variable and "no change" as the reference category

```
unorderedLogit <- multinom(factor_GDP ~ REG + OIL, data=gdp_data)
summary(unorderedLogit)
```

These are the results:

	Model 1
negative: (Intercept)	3.81***
	(0.27)
negative: REG	1.38
	(0.77)
negative: OIL	4.78
	(6.89)
positive: (Intercept)	4.53***
	(0.27)
positive: REG	1.77*
	(0.77)
positive: OIL	4.58
	(6.89)
AIC	4690.77
BIC	4728.10
Log Likelihood	−2339.39
Deviance	4678.77
Num. obs.	3721
K	3

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 1: Unordered Multinomial Logit Model

Based on the results, I can make the following interpretations.
The intercepts:
Keeping in mind that the reference category is "no change", the "negative" intercept tells us that when REG and OIL are both = 0, the estimated log odds of going from no change to negative is 3.8.
Similarly, the "positive" intercept tells us that when REG and OIL are both = 0, the estimated log odds of going from no change to positive is 4.53.
For REG:
A 1 unit increase in REG (going from a non-democracy to a democracy) corresponds with a change in the log odds of going from no change to negative by 1.38, while holding OIL constant.
A 1 unit increase in REG (going from a non-democracy to a democracy) corresponds with a change in the log odds of going from no change to positive by 1.77, while holding OIL constant.
For OIL:
A 1 unit increase in OIL corresponds with a change in the log odds of going from no change to negative by 4.78, while holding REG constant.
A 1 unit increase in OIL corresponds with a change in the log odds of going from no change to positive by 4.58, while holding REG constant.

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

First, I relevel the variable to create an ordinal structure

```
gdp_data$factor_GDP1 <- relevel(gdp_data$factor_GDP, ref="negative")
```

Next, I use the polr() function to conduct an ordered multinomial logit regression

```
OrderedModel <- polr(factor_GDP1 ~ REG + OIL, data=gdp_data)
summary(OrderedModel)
```

Here are the results:

	Model 1
REG	0.40^{***}
	(0.08)
OIL	-0.20
	(0.12)
negative—no change	-0.73^{***}
	(0.05)
no change—positive	-0.71^{***}
	(0.05)
AIC	4695.69
BIC	4720.58
Log Likelihood	-2343.84
Deviance	4687.69
Num. obs.	3721

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 2: Ordered Multinomial Logit Model

Next, I interpret the ordered multinomial logit model.
For REG:
A 1 unit increase in REG (going from non-democracy to democracy) corresponds to a change in the log odds of going from negative to no change and from no change to positive by 0.40, while holding OIL constant.
For OIL:
A 1 unit increase in OIL corresponds to a change in the log odds of going from negative to no change and from no change to positive by -0.20, while holding REG constant.

2 Question 2

2.1 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)

I run a Poisson regression using the glm() function:

```
    poissonModel <- glm(PAN.visits.06 ~ competitive.district + marginality.06 +
    PAN.governor.06, data = mexico_elections, family=poisson)
summary(poissonModel)
```

Here are the results:

	Model 1
(Intercept)	-3.81^{***}
	(0.22)
competitive.district	-0.08
	(0.17)
marginality.06	-2.08^{***}
	(0.12)
PAN.governor.06	-0.31
	(0.17)
AIC	1299.21
BIC	1322.36
Log Likelihood	-645.61
Deviance	991.25
Num. obs.	2407

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table 3: Poisson model

To answer this question, I interpret the competitive.district coefficient. A 1 unit increase in competitive.district (going from "safe seat" to "swing district") has a multiplicative effect on the mean of Poisson by $e^{-0.08}$, while holding all other variables constant.
exp(-0.08) = 1.083287.
As $B_1 > 0$, the expected count increases as number of visits increases. However, the test statistic (-0.477) is not large enough and this coefficient is not significant at the critical threshold($p - value = 0.6336 > 0.05$) and hence, we do not have enough evidence to suggest that PAN presidential candidates visit swing districts more.

2.2 Interpret the marginality.06 and PAN.governor.06 coefficients)

For marginality.06:
$exp(-2.08) = 0.13$
A 1 unit increase in poverty corresponds with a decrease in the expected number of visits by a multiplicative factor of 0.13, while holding PAN.governor.06 and competitive.district constant.

For PAN.governor.06:
$exp(-0.31) = 0.73$
In comparison to not having a PAN-affiliated governor, having a PAN-affiliated governor corresponds to a decrease in the expected number of visits by a multiplicative factor of 0.73, while holding marginality.06 and competitive.district constant.

2.3 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).

I make use of the predict() function:

```
predict(poissonModel, newdata=data.frame(competitive.district=1,
                                marginality.06 = 0,
                                PAN.governor.06=1), type="response")
```

Based on this, I get the value: 0.01494818
When a district is competitive, has an average poverty level and has a PAN-affiliated governor, the expected number of visits based on our model is 0.01.