UNIVERSITY *of* WEST FLORIDA

Ordinal Logistic Regression

Statistics for Data Science II

Introduction

Suppose our response variable has c ordered categories

e.g., classification of student: freshman, sophomore, junior, senior

We again will create $c - 1$ models.

The $\hat{\beta}_i$ will be the same across the models.

The $\hat{\beta}_0$ will change for each category.

Model

We will use the cumulative logit model,

$$
\begin{aligned}
\text{logit}\left(P[Y \le j]\right) &= \hat{\beta}_{0j} + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k \\
\log\left(\frac{P[Y \le j]}{1 - P[Y \le j]}\right) &= \hat{\beta}_{0j} + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k \\
\log\left(\frac{\pi_1 + \ldots + \pi_j}{\pi_{j+1} + \ldots \pi_c}\right) &= \hat{\beta}_{0j} + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k
\end{aligned}
$$

Model

As noted previously, the intercept depends on j.

- This means that curves will have the same shape $\forall\ j$.
- We are just shifting the curve along the x-axis, depending on the response category.

This model assumes *proportional odds*.

- For each predictor included in the model, the slopes across two outcomes response levels are the same, regardless of which two responses we consider.

Model

Example:

Consider data from a General Social Survey, relating political ideology to political party affiliation. Political ideology has a five-point ordinal scale, ranging from very liberal ($Y = 1$) to very conservative ($Y = 5$). Let X be an indicator variable for political party, with $X = 1$ for Democrats and $X = 0$ for Republicans. We will construct an ordinal logistic regression model that models political ideology as a function of political party and sex.

Model

Example:

```
head(data)
```

```
## # A tibble: 6 x 4
##   Sex    Party      Ideology                 Count
##   <chr>  <chr>      <fct>                    <dbl>
## 1 Female Democrat   1 - Very Liberal            44
## 2 Female Democrat   2 - Liberal                 47
## 3 Female Democrat   3 - Moderate               118
## 4 Female Democrat   4 - Conservative            23
## 5 Female Democrat   5 - Very Conservative       32
## 6 Female Republican 1 - Very Liberal            18
```

Model

Example:

```r
data <- dummy_cols(data, select_columns = c("Party", "Sex"))
m1 <- polr(Ideology ~ Party_Republican + Sex_Male,
           data = data, weights = Count, Hess = TRUE)
summary(m1)
```

```
## Call:
## polr(formula = Ideology ~ Party_Republican + Sex_Male, data = data,
##     weights = Count, Hess = TRUE)
##
## Coefficients:
##                   Value Std. Error t value
## Party_Republican 0.9636     0.1297  7.4311
## Sex_Male         0.1169     0.1273  0.9177
##
## Intercepts:
##                                      Value    Std. Error t value
## 1 - Very Liberal|2 - Liberal          -1.4518   0.1226   -11.8373
## 2 - Liberal|3 - Moderate              -0.4583   0.1048    -4.3746
## 3 - Moderate|4 - Conservative          1.2550   0.1142    10.9873
## 4 - Conservative|5 - Very Conservative   2.0890   0.1293    16.1587
##
## Residual Deviance: 2474.142
## AIC: 2486.142
```

Model

Example:

The resulting models are:

$$
\begin{aligned}
\text{logit}\left(P[Y \leq \text{V. Lib.}]\right) &= -1.45 + 0.96\text{repub.} + 0.12\text{male} \\
\text{logit}\left(P[Y \leq \text{Lib.}]\right) &= -0.46 + 0.96\text{repub.} + 0.12\text{male} \\
\text{logit}\left(P[Y \leq \text{Mod.}]\right) &= 1.26 + 0.96\text{repub.} + 0.12\text{male} \\
\text{logit}\left(P[Y \leq \text{Cons.}]\right) &= 2.09 + 0.96\text{repub.} + 0.12\text{male}
\end{aligned}
$$

Note that $P[Y \leq \text{V. Cons.}] = 1$, thus, does not need a model.

Interpretations

Odds ratios are interpreted slightly different due to the model being cumulative.

The change in odds does not depend on the category of the response.

For continuous predictors:

For a one [predictor unit] increase in [predictor], the odds in favor of [the response category j] or lower, as compared to higher than [the response category j], are multiplied by $e^{\hat{\beta}_i}$.

For categorical predictors:

As compared to [the reference category], the odds in favor of [the response category j] or lower, as compared to higher than [the response category j], for [the predictor category of interest] are multiplied by $e^{\hat{\beta}_i}$.

Interpretations

We can also interpret in terms of a percent increase or decrease

For continuous predictors:

> For a one [predictor unit] increase in [predictor], the odds in favor of [the response category *j*] or lower, as compared to higher than [the response category *j*], are [increased or decreased] by [100($e^{\hat{\beta}_i}$-1)% or 100(1-$e^{\hat{\beta}_i}$)%].

For categorical predictors:

> As compared to [the reference category], the odds in favor of [the response category *j*] or lower, as compared to higher than [the response category *j*], for [the predictor category of interest] are [increased or decreased] by [100($e^{\hat{\beta}_i}$-1)% or 100(1-$e^{\hat{\beta}_i}$)%].

Interpretations

Example:

For the political ideology data,

```
round(exp(coefficients(m1)[1]),2)
```

```
## Party_Republican
##             2.62
```

For any fixed j, the estimated odds that a Republican's response is in the conservative direction rather than the liberal direction are $e^{0.9636} = 2.62$ times the estimated odds for Democrats.

This is a 62% decrease in odds as compared to Democrats.

Interpretations

Example:

For the political ideology data,

```
round(exp(coefficients(m1)[2]),2)
```

```
## Sex_Male
##     1.12
```

For any fixed j, the estimated odds that a male's response is in the conservative direction rather than the liberal direction are $e^{0.1169} = 1.12$ times the estimated odds for females.

This is a 12% increase in odds as compared to females.

Inference

Like in binary and nominal logistic regressions, we can construct Wald Z statistic using the `coeftest()` function.

Example:

```
coeftest(m1)
```

```
##
## t test of coefficients:
##
##                    Estimate Std. Error t value  Pr(>|t|)
## Party_Republican    0.96363    0.12968  7.4311 2.683e-13 ***
## Sex_Male            0.11685    0.12734  0.9177     0.359
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Party affiliation is a significant predictor of political ideology ($p < 0.001$), but biological sex is not ($p = 0.359$).

Inference

Also like in binary and nominal logistic regressions, we can construct confidence intervals by running the model results through the `confint()` function.

Example:

```
round(exp(confint(m1)),2)
```

```
##                   2.5 % 97.5 %
## Party_Republican  2.04   3.38
## Sex_Male          0.88   1.44
```

Thus, the 95% CI for the OR for party affiliation (republican vs. democrat) is (0.30, 0.49) and for biological sex (male vs. female) is (0.88, 1.44).

Predictions

We can use the resulting models to construct specific predictions.

This may require some calculations due to the nature of the cumulative models.

In our example, the model order is V. Lib $\rightarrow$ Lib. $\rightarrow$ Mod. $\rightarrow$ Cons. $\rightarrow$ V. Cons.

If we want the individual probabilities,

$$
\begin{aligned}
P[Y = \text{V. Lib}] &= P[Y \leq \text{V. Lib}] \\
P[Y = \text{Lib.}] &= P[Y \leq \text{Lib.}] - P[Y \leq \text{V. Lib}] \\
P[Y = \text{Mod.}] &= P[Y \leq \text{Mod.}] - P[Y \leq \text{Lib.}] \\
P[Y = \text{Cons.}] &= P[Y \leq \text{Cons.}] - P[Y \leq \text{Mod.}] \\
P[Y = \text{V. Cons.}] &= 1 - P[Y \leq \text{Cons.}]
\end{aligned}
$$

Predictions

Using algebra, we can solve the cumulative logit model for $P[Y \leq j]$:

$$\log\left(\frac{P[Y \leq j]}{1 - P[Y \leq j]}\right) = \hat{\beta}_{0j} + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k$$

$$\frac{P[Y \leq j]}{1 - P[Y \leq j]} = \exp\left\{\hat{\beta}_{0j} + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k\right\}$$

$$\vdots$$

$$P[Y \leq j] = \frac{\exp\left\{\hat{\beta}_{0j} + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k\right\}}{1 + \exp\left\{\hat{\beta}_{0j} + \hat{\beta}_1 X_1 + \ldots + \hat{\beta}_k X_k\right\}}$$

Predictions

We can use the `fitted()` function to find the predicted probabilities of the outcomes.

Example:

```
head(unique(as.tibble(fitted(m1))))
```

```
## # A tibble: 4 x 5
##   `1 - Very Liberal` `2 - Liberal` `3 - Moderate` `4 - Conservative`
##                <dbl>         <dbl>          <dbl>              <dbl>
## 1             0.190          0.198          0.391              0.112
## 2             0.0820         0.112          0.378              0.183
## 3             0.172          0.188          0.397              0.121
## 4             0.0736         0.103          0.367              0.189
## # ... with 1 more variable: 5 - Very Conservative <dbl>
```

Visualization

Ordinal logistic regression models can be visualized like we saw previously with binary and nominal logistic regressions.

> Because we do not have a categorical predictor in the example model, we cannot construct the curve as we did before.

Let us now consider visualizing the expected probabilities for each group to see if we can see a pattern.

> y-axis: predicted probability
>
> x-axis: possible responses
>
> dots: colored by sex and/or political affiliation

Visualization

Example:

```
graph <- data %>%
  mutate(`Sex and Affiliation` = paste(Sex, Party),
         Ideology = str_remove_all(Ideology, "[12345 -]")) %>%
  mutate(Ideology = ifelse(Ideology == "VeryLiberal", "Very Liberal",
                    ifelse(Ideology == "VeryConservative", "Very Conservative",
                    Ideology))) %>%
  mutate(`Very Liberal` = as.tibble(fitted(m1))$`1 - Very Liberal`,
         Liberal = as.tibble(fitted(m1))$`2 - Liberal`,
         Moderate = as.tibble(fitted(m1))$`3 - Moderate`,
         Conservative = as.tibble(fitted(m1))$`4 - Conservative`,
         `Very Conservative` = as.tibble(fitted(m1))$`5 - Very Conservative`) %>%
  dplyr::select(`Sex and Affiliation`, `Very Liberal`, Liberal, Moderate,
                Conservative, `Very Conservative`) %>%
  melt(id.vars="Sex and Affiliation",
       variable.name="Political Ideology",
       value.name="Predicted Probability") %>%
  unique()
```

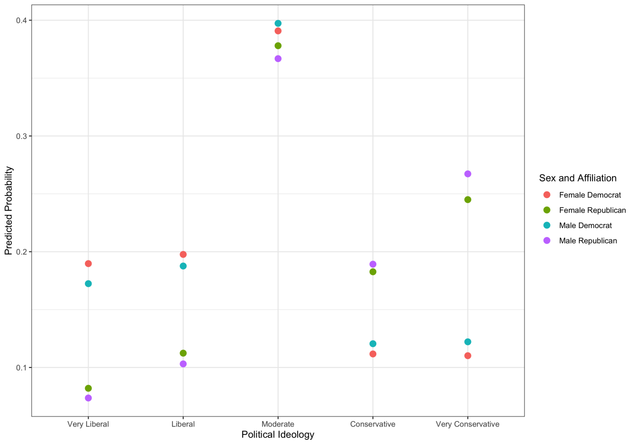
Visualization

Example:

```
plot <- graph %>% ggplot(aes(x = `Political Ideology`,
                             y = `Predicted Probability`,
                             color = `Sex and Affiliation`)) +
  geom_point(size = 3) +
  theme_bw()
```

Visualization

Example:



Proportional Odds Assumption

As mentioned previously, we are assuming proportional odds.

> This means that the slope is the same, regardless of what response category we're looking at.

We will check this assumption with Brant's test (article here).

> Briefly, this will construct a χ^2 test for every predictor in the model.
>
> If $p < \alpha$, the assumption is broken.

If the assumption is broken, we should step back down to nominal logistic regression.

Proportional Odds Assumption

Example:

```
brant(m1)
```

```
## --------------------------------------------------
## Test for        X2   df   probability
## --------------------------------------------------
## Omnibus              0    6    1
## Party_Republican 0    3    1
## Sex_Male        0    3    1
## --------------------------------------------------
##
## H0: Parallel Regression Assumption holds
```

All $p > \alpha$, thus, we meet the proportional odds assumption.