BIST P8110: Applied Regression II

17. Ordinal Logistic Regression

Qixuan Chen

Department of Biostatistics
Columbia University

What is Ordinal Logistic Regression?

- A categorical variable is considered ordinal if there is a natural ordering of the possible values. Examples include
  - opinion surveys with responses (ranging from “strongly agree" to “strongly disagree"),
  - medical cost (“low", “median", and “high"),
  - diseases (graded on scales from “least severe" to “most severe").
- Models for this type of data are extensions of the logistic regression model.
- The most well known ordinal logistic regression model is also called the proportional odds model.

Proportional Odds Model

- Considering K ordered categories, we define

$$\Pr(Y \leq k) = p_1 + \cdots + p_k$$

$$p_1 + p_2 + \cdots + p_K = 1$$

$$\text{odds}(Y \leq k) = \frac{P(Y \leq k)}{1 - P(Y \leq k)} = \frac{p_1 + \cdots + p_k}{p_{k+1} + \cdots + p_K}$$

$$\text{logit}(Pr(Y \leq k)) = \log\left[\frac{P(Y \leq k)}{1 - P(Y \leq k)}\right] = \log\left[\text{odds}(Y \leq k)\right]$$

$$k = 1, \ldots, K - 1$$

Proportional Odds Model

- The proportional odds model is given by

$$\text{logit}(Pr(Y \leq k|X)) = \alpha_k + \beta_1 X_1 + \cdots + \beta_m X_m, \; k = 1, \ldots, K-1 \tag{1}$$

  where, the $K - 1$ odds for each cut-off category k differ only in the intercept α_k. In other words, the odds are proportional.
- The probability of belonging to the lowest k categories is

$$Pr(Y \leq k) = \frac{e^{\alpha_k + \beta_1 X_1 + \cdots + \beta_m X_m}}{1 + e^{\alpha_k + \beta_1 X_1 + \cdots + \beta_m X_m}} \tag{2}$$

Model with $K = 4$

- For a four-level categorical outcome variable, three equations will be estimated. The equations are:

	Pooled Categories	vs	Pooled Categories
Equation 1:	1		2 3 4
Equation 2:	1 2		3 4
Equation 3:	1 2 3		4

  - Each equation models the odds of being in the set of categories on the left versus the set of categories on the right.
  - Only one set of coefficients for each X.

Proportional Odds Assumption

- The proportional odds model assumes parallel regression
  - The coefficients for the variables in the equations would not vary significantly if they were estimated separately.
  - The intercepts would be different, but the slopes would be essentially the same.
- The proportional odds assumption needs to be tested

Interpretation of Parameters

- Interpretation of intercepts
  - Intercepts can be interpreted similarly to the intercept in a dichotomous logistic regression.
  - The α_k in Model (1) can be interpreted as the log odds of Y belonging to the lowest k categories when all $X = 0$.
- Interpretation of slopes
  - A positive coefficient indicates an increased chance that a subject with a higher score on X will be observed in a lower category of Y.
  - A negative coefficient indicates an increased chance that a subject with a higher score on X will be observed in a higher category of Y.

Case Study

- Consider a study on taste of various cheese additives. Researchers tested four cheese additives and obtained 52 response ratings for each additive. (McCullagh and Nelder; 1989, p. 175)
  - The variable rating contains the response rating on a scale of nine categories (ranging from strongly dislike (1) to excellent taste (9)).
  - The variable additive specifies the cheese additive (1, 2, 3, or 4).
  - The variable count gives the frequency that each additive received each rating.

Cheese Data

Obs	additive	rating	count
1	1	1	0
2	1	2	0
3	1	3	1
4	1	4	7
5	1	5	8
6	1	6	8
7	1	7	19
8	1	8	8
9	1	9	1
10	2	1	6
11	2	2	9
12	2	3	12
13	2	4	11
14	2	5	7
15	2	6	6
16	2	7	1
17	2	8	0
18	2	9	0

⋮

Proportional Odds Model for Taste Rating

- The ordinal logistic regression model is

$$log\left\{\frac{Pr(Y \leq k)}{1 - Pr(Y \leq k)}\right\} = \alpha_k + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3$$

where

$$X_1 = \begin{cases} 1 & \text{Additive} = 2 \\ 0 & \text{ortherwise} \end{cases}$$

$$X_2 = \begin{cases} 1 & \text{Additive} = 3 \\ 0 & \text{ortherwise} \end{cases}$$

$$X_3 = \begin{cases} 1 & \text{Additive} = 4 \\ 0 & \text{ortherwise} \end{cases}$$

$$k = 1, 2, \ldots, 7, 8$$

SAS Code

The SAS code is:

```
data cheese;
infile "C:\cheese.csv" delimiter = ',' MISSOVER DSD;
input additive rating count;
run;

ods graphics on;
proc logistic data=cheese plots(only)=oddsratio;
   freq count;
   class Additive (ref='1')/param=ref;
   model rating = Additive;
   oddsratio Additive;
   effectplot / polybar;
run;
ods graphics off;
```

The LOGISTIC Procedure Options

- The ODDSRATIO statement computes odds ratios for all combinations of the Additive levels.
- The PLOTS option produces a graphical display of the odds ratios.
  - The ONLY option is used to suppress the default plots.
  - The ODDSRATIO option displays odds ratios and confidence limits.
- The EFFECTPLOT statement displays the predicted probabilities.
  - The POLYBAR option displays polytomous response data as a stacked histogram with bar heights defined by the individual predicted value.

SAS Output

Model Information

Data Set	WORK.CHEESE
Response Variable	rating
Number of Response Levels	9
Frequency Variable	count
Model	cumulative logit
Optimization Technique	Fisher's scoring

Number of Observations Read	36
Number of Observations Used	28
Sum of Frequencies Read	208
Sum of Frequencies Used	208

Response Profile

Ordered Value	rating	Total Frequency
1	1	7
2	2	10
3	3	19
4	4	27
5	5	41
6	6	28
7	7	39
8	8	25
9	9	12

Probabilities modeled are cumulated over the lower Ordered Values.

SAS Output

Score Test for the Proportional Odds Assumption

Chi-Square	DF	Pr > ChiSq
17.2866	21	0.6936

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	875.802	733.348
SC	902.502	770.061
-2 Log L	859.802	711.348

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	148.4539	3	<.0001
Score	111.2670	3	<.0001
Wald	115.1504	3	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
additive	3	115.1504	<.0001

SAS Output

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept 1	1	-5.4673	0.5202	110.4514	<.0001
Intercept 2	1	-4.4121	0.4247	107.9168	<.0001
Intercept 3	1	-3.3126	0.3697	80.2992	<.0001
Intercept 4	1	-2.2440	0.3262	47.3307	<.0001
Intercept 5	1	-0.9077	0.2748	10.9125	0.0010
Intercept 6	1	0.0443	0.2598	0.0291	0.8646
Intercept 7	1	1.5459	0.3042	25.8287	<.0001
Intercept 8	1	3.1058	0.4044	58.9727	<.0001
additive 2	1	3.3517	0.4235	62.6335	<.0001
additive 3	1	1.7098	0.3731	21.0072	<.0001
additive 4	1	-1.6128	0.3778	18.2265	<.0001

Wald Confidence Interval for Odds Ratios

Label	Estimate	95% Confidence Limits	
additive 1 vs 2	0.035	0.015	0.080
additive 1 vs 3	0.181	0.087	0.376
additive 1 vs 4	5.017	2.393	10.520
additive 2 vs 3	5.165	2.482	10.746
additive 2 vs 4	143.241	56.558	362.777
additive 3 vs 4	27.734	12.055	63.805

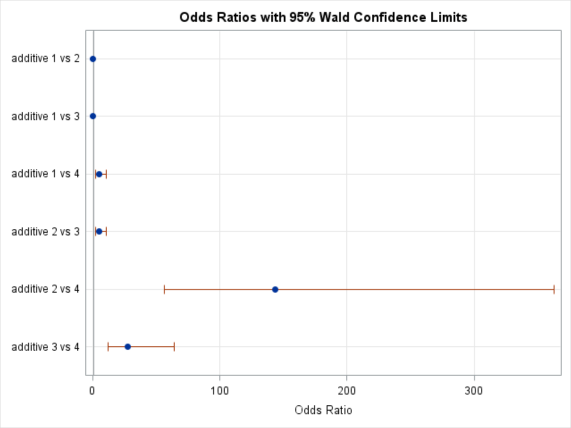
Testing for Proportional Odds Assumption

- The score chi-square test can be used for testing the proportional odds assumption:
  - H_0: the slopes of the X-variables are equal across logit equations; H_α: different slopes are needed.
  - Score test statistics: S=17.287, with df=21.
  - *p*-value: $Pr(\chi^2_{21} \geq 17.287) = 0.6936$
  - Conclusion: We fail to reject H_0 at $\alpha = 0.05$. The proportional odds assumption is reasonable.

What the Model Tells?

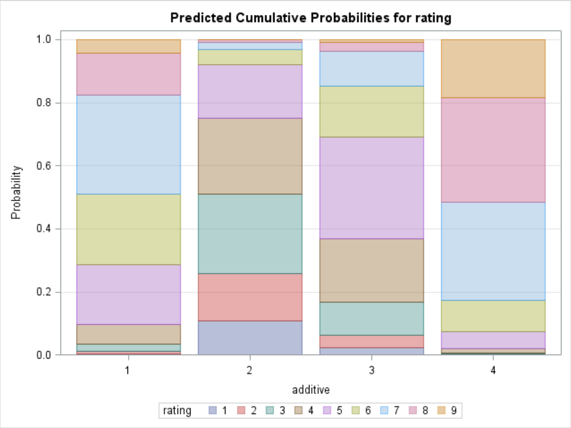
- The positive values ($\hat{\beta}_1 = 3.3517$ and $\hat{\beta}_2 = 1.7098$, $p < .0001$) for the parameter estimates for Additive2 and Additive3 indicate a tendency toward the lower-numbered categories of the second and third cheese additives relative to the first. In other words, The second and third additives are both less favorable than the first additive.
- The negative value ($\hat{\beta}_3 = -1.6128$, $p < .0001$) for the parameter estimate for Additive4 indicates that the fourth additive tastes better than the first.
- The relative magnitudes of these slope estimates imply the preference ordering: fourth, first, third, second. They are all significantly at $\alpha = 0.05$. Why?

OR and Interpretation



- OR Interpretation: the "Additive 1 vs 4" odds ratio says that the first additive has 5.017 times the odds of receiving a lower score than the fourth additive. (See SAS output on page 15)

Predicted Probabilities for Rating



$$\widehat{Pr}(Y \leq k) = \frac{e^{\hat{\alpha}_k + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3}}{1 + e^{\hat{\alpha}_k + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3}}, k = 1, 2, \ldots, 8$$

Summary: Key points

- How to write ordinal logistic regression model?
- How to test proportional odds assumption and determine the degrees of freedom?
- How to interpret intercepts and slopes in ordinal logistic regression?
- Why does the intercept α_k increase as k increases?
- What are the rationales underlying the conclusions on page 17?
- How to calculate the predicted probability of $Y = k$?

Readings

Flom P.L “Multinomial and ordinal logistic regression using PROC LOGISTIC". NESUG 18.