Ordinal Logistic Regression

Dr Wan Nor Arifin

Biostatistics and Research Methodology Unit
Universiti Sains Malaysia
wnarifin@usm.my / wnarifin.github.io

Last update: Jun 24, 2024

Expected outcomes

- Understand the concept of ordinal logistic regression
- Perform ordinal logistic regression
- Perform model assessment
- Present and interpret results

Outlines

- Introduction
- Ordinal logistic regression model
- Model building:
  - Variable selection
  - Variable assessment
  - Interaction term assessment
  - Model fit assessment

Introduction

Introduction

- A regression method to model relationship between:
  - Outcome: ordinal categorical variable
  - Independent variables: numerical, categorical variables
- Ordinal i.e. rank order of categories > than two levels

Introduction

- Ordinal <u>measurement scale</u>
  - Ordinal categorical variable
  - Ordered
  - Examples:
    - Disease severity: Mild, Moderate, Severe
    - Opinion: Disagree, Neutral, Agree
    - Cancer stages etc.

Introduction

- Model the relationship

ordinal outcome = numerical predictors +
categorical predictors

Introduction

- For a $K + 1$ <u>ordered</u> outcome (e.g. $k = \{0, 1, 2, 3\}$ for four categories), THREE main models:

Model	**Proportional Odds**	**Adjacent-category Logit**	**<u>Constrained</u> Continuation-ratio Logit**
Other Name	<u>Constrained</u> Cumulative Logit	<u>Constrained</u> Baseline Logit	-
Specification	$Y \leq k$ vs $Y > k$	$Y = k$ vs $Y = k - 1$	$Y = k$ vs $Y < k$
Example	0 vs > 0 (1,2,3) ≤ 1 (0,1) vs > 1 (2,3) ≤ 2 (0,1,2) vs > 2 (3)	1 vs 0 2 vs 1 3 vs 2	1 vs < 1 (0) 2 vs < 2 (0,1) 3 vs < 3 (0,1,2)

Introduction

- Let's compare the logits:

Model	Proportional Odds	Adjacent-category Logit	Constrained Continuation-ratio Logit
Logits	$c_k(\boldsymbol{x}) = ln\left[\frac{P(Y \leq k \mid \boldsymbol{x})}{P(Y > k \mid \boldsymbol{x})}\right] = \tau_k - \boldsymbol{x}'\boldsymbol{\beta}$	$a_k(\boldsymbol{x}) = ln\left[\frac{P(Y = k \mid \boldsymbol{x})}{P(Y = k-1 \mid \boldsymbol{x})}\right] = \alpha_k + \boldsymbol{x}'\boldsymbol{\beta}$	$r_k(\boldsymbol{x}) = ln\left[\frac{P(Y = k \mid \boldsymbol{x})}{P(Y < k \mid \boldsymbol{x})}\right] = \theta_k + \boldsymbol{x}'\boldsymbol{\beta}$

Minus here to be consistent with most software packages

- All models have different intercepts for each logit (i.e. by *k*)
- … but have same slope coefficients across all logits (i.e. for all *k*) – constrained
- … i.e. a single odds ratio – easier to interpret than multinomial logistic regression
- **Proportion odds** model – frequently used, most intuitive (less than or equal *vs* more) – focus of this lecture.

Proportional Odds Logistic Regression Model for Ordinal Outcome

Logit Functions

Compare 1 to 0

- A logit function for proportional odds is given as:

$$c_k(\boldsymbol{x}) = ln\left[\frac{P(Y \le k \mid \boldsymbol{x})}{P(Y > k \mid \boldsymbol{x})}\right]$$
$$= ln\left[\frac{P(Y=0 \mid \boldsymbol{x}) + P(Y=1 \mid \boldsymbol{x}) \ldots P(Y=k \mid \boldsymbol{x})}{P(Y=k+1 \mid \boldsymbol{x}) + P(Y=k+2 \mid \boldsymbol{x}) \ldots P(Y=K \mid \boldsymbol{x})}\right]$$
$$= \tau_k - \boldsymbol{x}'\boldsymbol{\beta}$$

For vector $\boldsymbol{x}$ comprising of p covariates and k = {0, 1, 2, …, K - 1} for K + 1 categories

Logit Functions

- An example of a logit function for proportional odds when $K = 1$:

$$
\begin{aligned}
c_1(\boldsymbol{x}) &= ln\left[\frac{P(Y \le 1 \mid \boldsymbol{x})}{P(Y > 1 \mid \boldsymbol{x})}\right] \\
&= ln\left[\frac{P(Y = 0 \mid \boldsymbol{x}) + P(Y = 1 \mid \boldsymbol{x})}{P(Y = 2 \mid \boldsymbol{x}) + P(Y = 3 \mid \boldsymbol{x})}\right] \\
&= \tau_1 - \boldsymbol{x}'\boldsymbol{\beta}
\end{aligned}
$$

Logit Functions

- An example of a logit function for proportional odds when $K = 2$:

$$
\begin{aligned}
c_2(\boldsymbol{x}) &= ln\left[\frac{P(Y \leq 2 \mid \boldsymbol{x})}{P(Y > 2 \mid \boldsymbol{x})}\right] \\
&= ln\left[\frac{P(Y=0 \mid \boldsymbol{x}) + P(Y=1 \mid \boldsymbol{x}) + P(Y=2 \mid \boldsymbol{x})}{P(Y=3 \mid \boldsymbol{x})}\right] \\
&= \tau_2 - \boldsymbol{x}'\boldsymbol{\beta}
\end{aligned}
$$

Odds Ratios

- Since the <u>constraint</u> gives us a single coefficient, the odds ratio is straight forward to calculate – similar to a binary logistic regression
- This is calculated for a covariate x_i as follows:

$$\mathrm{OR}\left(x_i\right)=e^{\beta_i}$$

regardless of the outcome categories to be compared i.e. only concerned with <u>less than or equal</u> vs <u>more</u>

- So it does not matter (0,1) vs (2,3) OR (0) vs (1,2,3) because the odds is proportionate → **proportional odds assumption**

Cumulative Probabilities

- In order to obtain individual outcome probabilities, for proportional odds model, it requires the calculation for cumulative probabilities as follows:

$$\pi_k(\boldsymbol{x}) = \frac{e^{c_k(\boldsymbol{x})}}{1+e^{c_k(\boldsymbol{x})}}$$

Individual Outcome Probabilities

- Following the cumulative probabilities calculation, we may then calculate individual probabilities as follows:

$$P(Y=k \mid \boldsymbol{x}) = \begin{cases} \pi_0(\boldsymbol{x}), & k=0 \\ \pi_k(\boldsymbol{x}) - \pi_{k-1}(\boldsymbol{x}), & k=1,\ldots,K-1 \\ 1-\pi_{K-1}(\boldsymbol{x}), & k=K \end{cases}$$

Individual Outcome Probabilities

- An example when $K = 1$:

$$\pi_1(\boldsymbol{x}) = \frac{e^{c_1(\boldsymbol{x})}}{1 + e^{c_1(\boldsymbol{x})}}$$

$$\pi_0(\boldsymbol{x}) = \frac{e^{c_0(\boldsymbol{x})}}{1 + e^{c_0(\boldsymbol{x})}}$$

$$\begin{aligned} P(Y = 1 \mid \boldsymbol{x}) &= \pi_k(\boldsymbol{x}) - \pi_{k-1}(\boldsymbol{x}) \\ &= \pi_1(\boldsymbol{x}) - \pi_0(\boldsymbol{x}) \end{aligned}$$

Individual Outcome Probabilities

- An example when $K = 2$:

$$\pi_2(\boldsymbol{x}) = \frac{e^{c_2(\boldsymbol{x})}}{1 + e^{c_2(\boldsymbol{x})}}$$

$$\pi_1(\boldsymbol{x}) = \frac{e^{c_1(\boldsymbol{x})}}{1 + e^{c_1(\boldsymbol{x})}}$$

$$\begin{aligned} P(Y = 2 \mid \boldsymbol{x}) &= \pi_k(\boldsymbol{x}) - \pi_{k-1}(\boldsymbol{x}) \\ &= \pi_2(\boldsymbol{x}) - \pi_1(\boldsymbol{x}) \end{aligned}$$

Testing Significance

- Wald test, W
- Likelihood ratio test, G

Testing Significance

- Wald test, W:

$$W = \frac{\hat{\beta}}{\widehat{SE}(\hat{\beta})}$$

then, two-tailed P-value is $P(|z| > W)$, as W follows standard normal distribution.

- More suitable for testing a single variable.

Testing Significance

- Likelihood ratio test, *G*:

Log Likelihood of model withOUT x variable(s) –
Log Likelihood of model with x variable(s)

$$G = -2(L_0 - L_1) \text{ OR}$$

$$G = D_0 - D_1$$

D = Deviance =
-2 Log Likelihood of model

then, *P*-value is $P[\chi^2(\text{df}) > G]$, as *G* follows standard normal distribution, and df = difference in number of parameters between the models.

- Suitable for testing single/many variables.

Model Building

Model-building Steps

1. Variable selection
   – Univariable
   – Multivariable

   → Preliminary main effects model

2. Variable assessment
   – Linearity in logit – numerical variable
   – Other numerical issues
     • Small cell counts
     • Multicollinearity

   → Main effects model

Model-building Steps

3. Interaction term assessment
   – Two-way between selected variables – clinically sensible

   → Preliminary final model
4. Model fit assessment
   – Proportional odds assumption check – Brant Test
   – Goodness-of-fit
     • Lipsitz Test, Ordinal Hosmer-Lemeshow Test
   – Pseudo-R^2
   – Regression diagnostics – from separate binary logistic models

   → Final model