Ordinal response regression models

Categorical Data Analysis, Summer 2015

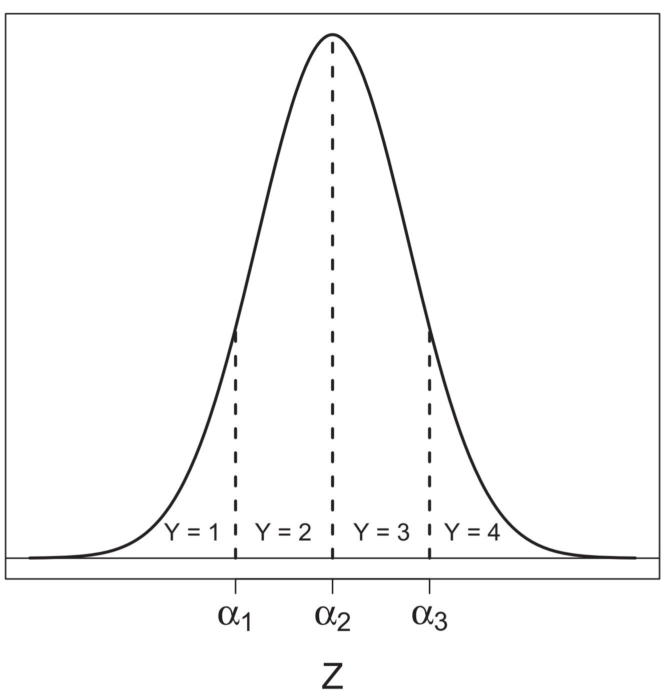
Overview

- In the previous lecture, we discussed the baseline-category logit model for multivariable regression modeling of a multinomial response variable Y with J categories.
- Recall that this model estimates a complete set of p parameters (intercept and slopes) for each of $J - 1$ baseline-referenced logits.
- When the response variable is ordinal, we can use this structure to our advantage to develop more parsimonious and powerful models.

Ordinal variables

- Ordinal outcomes are common in medical research
  - Likert scale ('Strongly disagree', 'Disagree', 'Neither agree nor disagree', 'Agree', 'Strongly agree')
  - Disease severity ('Normal', 'Mild', 'Moderate', 'Severe')
- It is often possible to consider an ordinal variable to be a discretized version of a continuous latent variable.
- In this context, the ordinal variable is a discrete version of an unmeasured (and unobserved) continuous variable.

Ordinal variables (cont.)



Cumulative logits

- A natural approach for regression modeling with an ordinal response variable is to construct logit models for dichotomized versions of Y while moving in sequence 'up' the ordinal scale.
- Such logits are called cumulative logits and are constructed from cumulative probabilities

$$P(Y \leq j|\mathbf{x}) = \pi_1(\mathbf{x}) + \ldots + \pi_j(\mathbf{x}),\ j = 1, \ldots, J$$

- The cumulative logits are defined as

$$\begin{aligned}\text{logit}[P(Y \leq j|\mathbf{x})] &= \log \frac{P(Y \leq j|\mathbf{x})}{P(Y > j|\mathbf{x})} \\ &= \log \frac{P(Y \leq j|\mathbf{x})}{1 - P(Y \leq j|\mathbf{x})}\end{aligned}$$

Cumulative logits (cont.)

In contrast to the baseline-category logits, each cumulative logit uses <u>all</u> response categories.

$$\begin{aligned}L_1 = \text{logit}[P(Y \leq 1|\mathbf{x})] &= \log \frac{\pi_1(\mathbf{x})}{\pi_2(\mathbf{x}) + \pi_3(\mathbf{x}) \ldots + \pi_J(\mathbf{x})} \\ L_2 = \text{logit}[P(Y \leq 2|\mathbf{x})] &= \log \frac{\pi_1(\mathbf{x}) + \pi_2(\mathbf{x})}{\pi_3(\mathbf{x}) + \ldots + \pi_{J-1}(\mathbf{x}) + \pi_J(\mathbf{x})} \\ &\vdots \\ L_{J-1} = \text{logit}[P(Y \leq J-1|\mathbf{x})] &= \log \frac{\pi_1(\mathbf{x}) + \pi_2(\mathbf{x}) + \ldots + \pi_{J-1}(\mathbf{x})}{\pi_J(\mathbf{x})}\end{aligned}$$

The Proportional Odds Model

- Modeling a single cumulative logit as a function of covariates is equivalent to logistic regression where categories $1, \ldots, j$ and $j+1, \ldots, J$ form the outcomes of success and failure, respectively.
- The proportional odds model simultaneously uses all $J-1$ cumulative logits in a single model, and is given by

$$\text{logit}[P(Y \leq j|\mathbf{x})] = \alpha_j + \boldsymbol{\beta}'\mathbf{x}, \;\; j = 1, \ldots, J-1$$

- The natural ordering of the cumulative probabilities results in model constraints:

$$\begin{aligned} P(Y \leq 1|\mathbf{x}) &\leq \ldots \leq & P(Y \leq J-1|\mathbf{x}) \\ \text{logit}[P(Y \leq 1|\mathbf{x})] &\leq \ldots \leq & \text{logit}[P(Y \leq J-1|\mathbf{x})] \\ \alpha_1 + \boldsymbol{\beta}'\mathbf{x} &\leq \ldots \leq & \alpha_{J-1} + \boldsymbol{\beta}'\mathbf{x} \end{aligned}$$

- We therefore constrain the $\{\alpha_j\}$ so that $\alpha_1 \leq \ldots \leq \alpha_{J-1}$.

The Proportional Odds Model (cont.)

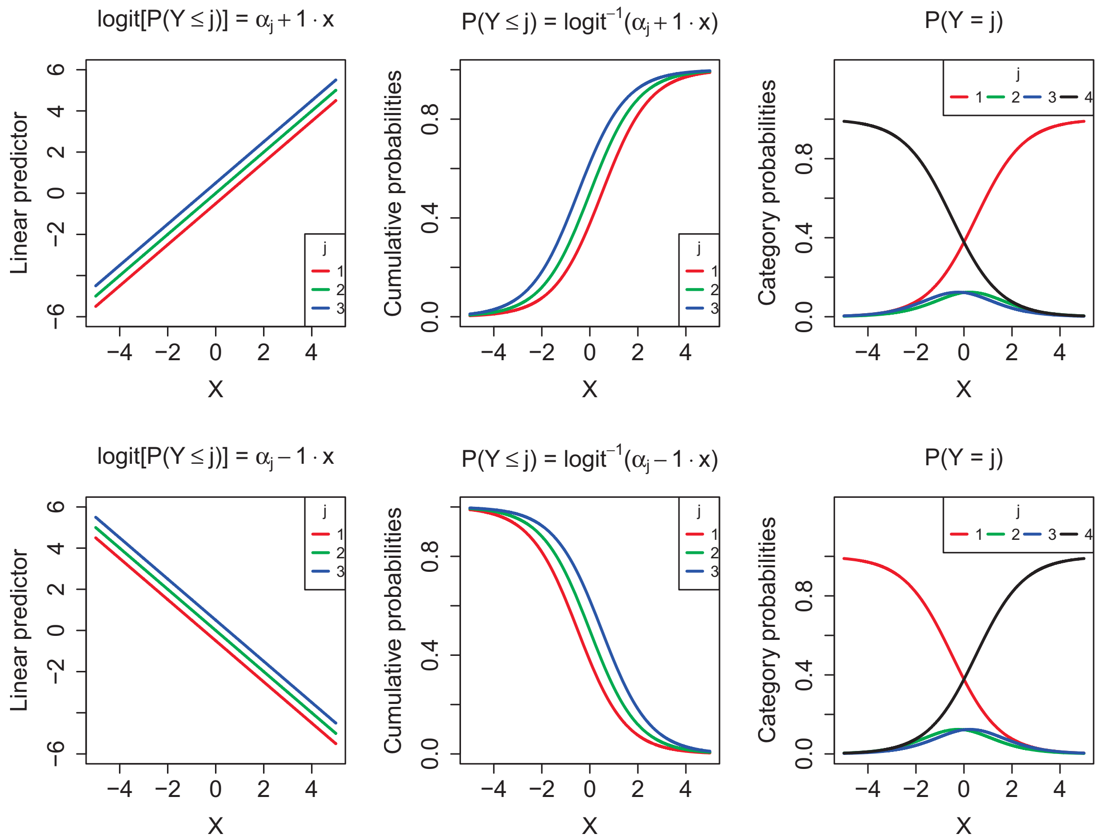
- Additionally, the proportional odds model assumes the <u>same effects</u> $\boldsymbol{\beta}$ across all cumulative logits.
- The common $\boldsymbol{\beta}$ and the ordering constraint on the intercepts are important differences between the proportional odds model and the baseline-category logit model.
- Cumulative probabilities are constructed using the inverse logit function

$$P(Y \leq j|\mathbf{x}) = \frac{\exp\{\alpha_j + \boldsymbol{\beta}'\mathbf{x}\}}{1 + \exp\{\alpha_j + \boldsymbol{\beta}'\mathbf{x}\}}, \;\; j = 1, \ldots, J-1$$

- Individual category probabilities are constructed by taking differences of cumulative probabilities

$$\begin{aligned} P(Y = 1|\mathbf{x}) &= P(Y \leq 1|\mathbf{x}) \\ P(Y = j|\mathbf{x}) &= P(Y \leq j|\mathbf{x}) - P(Y \leq j-1|\mathbf{x}), \;\; j = 2, \ldots, J-1 \\ P(Y = J|\mathbf{x}) &= 1 - P(Y \leq J-1|\mathbf{x}) \end{aligned}$$

Four-category proportional odds model



Observations

- The proportional odds assumption implies parallel linear predictors for the $J - 1$ cumulative logits (parallelism is not a property of baseline-category logits).
- Changing the sign of β results in a reverse ordering of the response categories.
- This means that the model 'preserves' the ordinal scale of Y, and is therefore an attractive model for regression models of ordinal response variables.

Parameter interpretation

Consider a single continuous predictor x. The difference in cumulative logits for an increment Δ of x is given by

$$\begin{aligned}\text{logit}([P(Y \leq j|x = x_0 + \Delta)]) - \text{logit}([P(Y \leq j|x = x_0)]) &= \\ (\alpha_j + \beta(x_0 + \Delta)) - (\alpha_j + \beta x_0) &= \Delta\beta\end{aligned}$$

- $e^{\Delta\beta}$ is a cumulative odds ratio
- The cumulative odds ratio does not depend on j, and is therefore the same across all $J - 1$ cumulative logits.
- This means that *no matter how you dichotomize* Y, the effect of x on the odds that $Y \leq j$ is constant.
- *This property of a common effect for all cumulative probabilities is referred to as proportional odds*.

Parameter interpretation (cont.)

- Since cumulative odds ratios in proportional odds models are the same for all categories, subsequent inference focuses on the *direction* of response rather than on specific response categories.
- Suppose $e^{\Delta\beta} = 2$ for $\Delta = 1$.
  - This cumulative odds ratio is interpreted as a two-fold increase in the odds of a response as small or smaller for a 1-unit increase in x.
  - Alternatively, since $1/e^{\Delta\beta} = 0.5$, we could also conclude that there is a 50% reduction in the odds of a higher response for a 1-unit increase in x.
- Suppose $e^{\Delta\beta} = 0.6$ for $\Delta = 1$.
  - There is a 40% reduction in the odds of a response as small or smaller for a 1-unit increase in x.
  - Alternatively, since $1/e^{\Delta\beta} \doteq 1.67$, we could also conclude that there is a 67% increase in the odds of a higher response for a 1-unit increase in x.