STAT 151A Lecture 38

Henry Liev

24 November 2025

Remark 0.1 (Nested Dichotomies as an Alternative to Multinomial Input Models)

They are not the same
Pros:

- No fancy tools needed

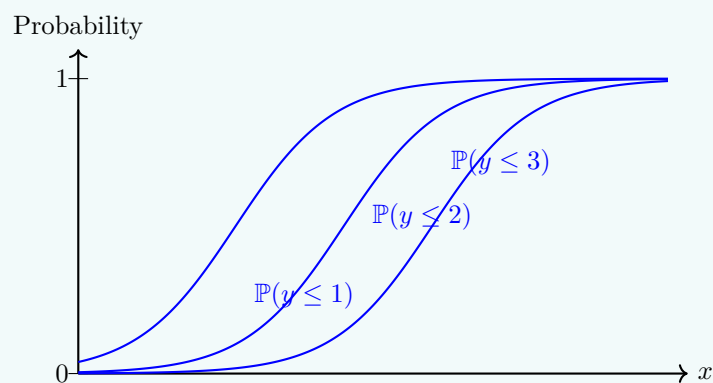- It accounts for order among the decisions

Cons:

- More models are needed

- Specific type you select matters

- Less flexible about comparing arbitrary categories

$$y_i \in \begin{Bmatrix} \text{some hs} \\ \text{hs grad} \\ \text{some college} \\ \text{associates} \\ \text{bachelors} \\ \text{masters} \\ \text{prof/phd} \end{Bmatrix} , \text{natural ordering, ordinal rather than nominal}$$

Toy model: $\mathbb{P}(y_i \leq j) = \sum_{l=1}^{j} \mathbb{P}(y_i = l)$

Remark 0.2 (Ordinal Logistic Regression/Proportional Odds Model)

$\log\left(\frac{\mathbb{P}(y\leq j)}{\mathbb{P}(y>j)}\right) = \alpha_j + \beta_1 x_{1i} + \cdots + \beta_p x_{pi}, j < m$

Probability

$$1$$

$\mathbb{P}(y \leq 3)$

$\mathbb{P}(y \leq 2)$

$\mathbb{P}(y \leq 1)$

$$0$$

x

These grahps all have exactly the same shape: parallel $\rightarrow$ proportional $\rightarrow$ assuming we don't need ? multinomial

Benefit: $(m-1)+p$ parameters vs $(p+1)(m-1)$

Also fit this and do inference using MLE

Interpretation: β_j is still similarly interpreted:

e^{β_j} is factor by which odds that $y \leq$ any category are higher for a unit where x_j is ? ? higher

Intercepts: pick an individual i

$\frac{\text{odd}(y_i|x_i)}{\text{odd}(y_k|x_i)} = \frac{\exp(\alpha_j + x_i^\intercal \beta)}{\exp(\alpha_j + x_i^\intercal \beta)} = \exp(\alpha_j - \alpha_k)$

Difference in intercepts gives the factor by which first (positive) category's odds differs from the second category's odds where all ? are the same