ADS research project, ordinal data

F. Meijvis, D. Hessen, C. Spitoni

February 11, 2025

1 Definition of ordinal regression

The dependent random variable Y has ordinal level of measurement. The number of possible values that Y can take is $m+1$. A realization of Y is denoted by y and $y \in \{0, 1, \dots, m\}$. The conditional probability distribution of Y given a set of interval predictors $\mathbf{x} = (x_1, \dots, x_k)'$ is $P(Y = y \,|\, \mathbf{x})$. The regression of Y on $\mathbf{x}$ is ordinal if and only if

$$P(Y \leq y \,|\, \vec{x}) = \sum_{t=0}^{y} P(Y = t \,|\, \vec{x}), \quad \text{for all } y, \tag{1}$$

is a monotonic function of $\mathbf{x}$. If higher values of y are associated with higher values of $\mathbf{x}$, then this function should be non-increasing, for all y. If $P(Y \leq y \,|\, \mathbf{x})$ is a monotonic function of x_i, then its derivative with respect to x_i is either non-negative or non-positive for all x_i.

2 Multinomial logistic regression

Let $z_s = 1$ if $y = s$ and $z_s = 0$ otherwise, for $s = 1, \dots, m$, then the conditional probability distribution of Y given $\mathbf{x}$ can be written as

$$P(Y = y \,|\, \vec{x}) = P(Y = 0 \,|\, \vec{x}) \prod_{s=1}^{m} \left\{ \frac{P(Y = s \,|\, \vec{x})}{P(Y = 0 \,|\, \vec{x})} \right\}^{z_s}, \tag{2}$$

where

$$P(Y = 0 \,|\, \vec{x}) = \left\{ 1 + \sum_{s=1}^{m} \frac{P(Y = s \,|\, \vec{x})}{P(Y = 0 \,|\, \vec{x})} \right\}^{-1}. \tag{3}$$

In the multinomial logistic regression model,

$$\frac{P(Y = s \,|\, \vec{x})}{P(Y = 0 \,|\, \vec{x})} = \exp(\alpha_{0s} + \vec{\alpha}_s' \vec{x}), \text{ for } s = 1, \dots, m, \tag{4}$$

where $\vec{\alpha}_s = (\alpha_{1s}, \dots, \alpha_{ks})'$, so that

$$P(Y = y \,|\, \vec{x}) = \frac{\exp\left\{ \sum_{s=1}^{m} z_s (\alpha_{0s} + \vec{\alpha}_s' \vec{x}) \right\}}{1 + \sum_{s=1}^{m} \exp(\alpha_{0s} + \vec{\alpha}_s' \vec{x})}. \tag{5}$$

It follows that

$$P(Y \leq y \,|\, \vec{x}) = \begin{cases} \frac{1}{1+\sum\limits_{s=1}^{m} \exp(\alpha_{0s}+\vec{\alpha}_s'\vec{x})}, & \text{for } y = 0, \\ \frac{1+\sum\limits_{s=1}^{y} \exp(\alpha_{0s}+\vec{\alpha}_s'\vec{x})}{1+\sum\limits_{s=1}^{m} \exp(\alpha_{0s}+\vec{\alpha}_s'\vec{x})}, & \text{for } y > 0. \end{cases} \tag{6}$$

The multinomial logistic regression model has $(k+1)m$ parameters.

3 conditions for α

In order to have the multinomial regression model satisfy the definition of ordinal regression, we require $P(Y \leq y \mid \vec{x})$ to be monotonic in $\vec{x}$. We will assume that higher values in $\vec{x}$ correspond to higher y. Furthermore, we treat both these statements are component wise.

For monotonicity, we require:

$$\frac{\partial P(Y \leq y|\vec{x})}{\partial x_i} \leq 0$$

For every i and every $\vec{x}$. For class $y = 0$, this ultimately gives:

$$\sum_{s=1}^{m} \alpha_{is} \exp(\alpha_{0s} + \vec{\alpha_s}'\vec{x}) \geq 0$$

This is satisfied if every $\alpha_{is} \geq 0$, but this may not be the strongest condition yet.

Note to self: I still want to test if it is possible to have 1 particular $\alpha \leq 0$

For clas $y > 0$, we find:

$$(1 + \sum_{s=1}^{m} \exp(\alpha_{0s} + \vec{\alpha_s}'\vec{x}))(\sum_{s=1}^{y} \alpha_{is} \exp(\alpha_{0s} + \vec{\alpha_s}'\vec{x}))$$
$$-(1 + \sum_{s=1}^{y} \exp(\alpha_{0s} + \vec{\alpha_s}'\vec{x}))(\sum_{s=1}^{m} \alpha_{is} \exp(\alpha_{0s} + \vec{\alpha_s}'\vec{x})) \leq 0$$

Which reduces back to the previous equation when $y = 0$, and is trivially satisfied when $y = m$.

4 Adjacent categories model

In this section we reparametrize the multinomial logistic regression model. In general we have:

$$P(Y = y \,|\, \vec{x}) = P(Y = 0 \,|\, \vec{x}) \prod_{u=1}^{y} \frac{P(Y = y \,|\, \vec{x})}{P(Y = y-1 \,|\, \vec{x})}$$

For $y = 1, \ldots, m$. In an adjacent categories model we assume

$$\frac{P(Y = y \,|\, \vec{x})}{P(Y = y-1 \,|\, \vec{x})} = \exp(\beta_{0u} + \vec{\beta}_u{}' \vec{x})$$

Which can be seen as the probability of advancing from 1 category to the next (note that this requires an intrinsic ordering of the data i.e. ordinal data). Under these assumptions we find

$$P(Y = y \,|\, \vec{x}) = P(Y = 0 \,|\, \vec{x}) \exp(\sum_{u=1}^{y} [\beta_{0u} + \vec{\beta}_u{}' \vec{x}])$$

By stating that $\vec{\beta}_u = (\beta_{1u}, \ldots, \beta_{ku})'$, so that $\alpha_{0s} = \sum_{u=1}^{s} \beta_{0u}$ and $\vec{\alpha}_s = \sum_{u=1}^{s} \vec{\beta}_u$, we find

$$P(Y = y \,|\, \vec{x}) = \frac{\exp\Big(\alpha_{0y} + y\vec{\beta}' \vec{x}\Big)}{1 + \sum\limits_{s=1}^{m} \exp(\alpha_{0s} + s\vec{\beta}' \vec{x})}$$

Which reduces to

$$P(Y = y \,|\, \vec{x}) = \frac{\exp\big(\alpha_{0y} + \vec{\alpha}_y{}' \vec{x}\big)}{1 + \sum\limits_{s=1}^{m} \exp(\alpha_{0s} + \vec{\alpha}_s{}' \vec{x})}$$

by assuming the coefficients are independent of s such that $\beta_{is} = \beta_i$, for $i = 1, \ldots, k$ and all s, so that $\alpha_{is} = s\beta_i$.