A: 6.0

IX. Polytomous Regression:

A. Introduction

1. Regression Context
   a. One categorical response variable, now allowing for more than two levels.
   b. One or more explanatory variables
      i. Response variables satisfy earlier requirements.
2. Regression Parameterization
   a. Make probability associated with category $j$ for individual $k$ depend on covariates $\boldsymbol{x}_k$ through parameters $\boldsymbol{\beta}_j$
   b. Often force some components of $\boldsymbol{\beta}_j$ not to depend on $j$
      i. Generally first component of $\boldsymbol{x}_k$ is 1
      ii. Hence first component of $\boldsymbol{\beta}_j$ is intercept
         - Generally intercept depends on $j$
         - Generally the other components do not.

A: 6.1

B. Baseline-Category logits

1. Baseline-Category Logit Model
   a. Allow a different set of parameters $\boldsymbol{\beta}_j$ for difference between baseline and each non-baseline value for response.
   b. $\log(\mathsf{P}\left[Y_k = j\right] / \mathsf{P}\left[Y_k = 0\right]) = \boldsymbol{x}_k\boldsymbol{\beta}_j$
      i. $(1 + \sum_{j>0} \exp(\boldsymbol{x}_k\boldsymbol{\beta}_j))\mathsf{P}\left[Y_k = 0\right] = 1$
      ii. $\mathsf{P}\left[Y_k = 0\right] = 1/(1 + \sum_{j>0} \exp(\boldsymbol{x}_k\boldsymbol{\beta}_j))$
   c. $\log(\mathsf{P}\left[Y_k = j\right] / \mathsf{P}\left[Y_k = l\right]) = \boldsymbol{x}_k(\boldsymbol{\beta}_j - \boldsymbol{\beta}_l)$



   d. Note that aside from baseline category, all other categories are treated symmetrically.
      i. Hence no use of category ordering.
2. Model likelihood
   a. Let $W_{kj} = \begin{cases} 1 & \text{if } Y_k = j \\ 0 & \text{otherwise} \end{cases}$
   b. Series of separate logistic regressions conditional on sum of that category and baseline category.
   c. Multinomial likelihood $L(\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \ldots, \boldsymbol{\beta}_{J-1}) = \prod_k \prod_j \mathsf{P}\left[Y_k = j\right]^{W_{kj}}$
   d. $\ell = \sum_k \sum_j W_{kj} \log(\mathsf{P}\left[Y_k = j\right])$ .
      i. $\ell = \sum_k [\sum_{j>0} W_{kj}\boldsymbol{x}_k\boldsymbol{\beta}_j + \sum_j W_{kj} \log(\mathsf{P}\left[Y_k = 0\right])]$
      ii. $\ell = \sum_k [\sum_{j>0} W_{kj}\boldsymbol{x}_k\boldsymbol{\beta}_j - \sum_j W_{kj} \log(1 + \sum_{j>0} \exp(\boldsymbol{x}_k\boldsymbol{\beta}_j))]$
   e. $\frac{d}{d\boldsymbol{\beta}_j}\ell = \sum_k [W_{kj} - \sum_l W_{kl}\pi_{lk}]\boldsymbol{x}_k$
      i. $\pi_{jk} = \exp(\boldsymbol{x}_k\boldsymbol{\beta}_j)/(1 + \sum_{l>0} \exp(\boldsymbol{x}_k\boldsymbol{\beta}_j)))$ R Code SAS Code

A: 6.2

C. Models using category ordering

1. Cumulative Logit Model:
   a. Suppose $\boldsymbol{\beta}_j = (\theta_j, \boldsymbol{\alpha})$ .
   b. Suppose that $T_k - \boldsymbol{x}_k\boldsymbol{\alpha}$ has CDF $\exp(t)/(1 + \exp(t))$
      i. Mean 0, standard deviation 1.8138
   c. Pick an increasing sequence $\theta_j$
   d. Suppose that $Y_k = j$ if $T_k \in [\theta_{j-1}, \theta_j)$ .
   e. Probabilties for outcomes.
      i. $\mathsf{P}\left[Y_k > j\right] = (1 + \exp(\theta_j + \boldsymbol{x}_k\boldsymbol{\alpha}))^{-1}$



      ii. $\mathsf{P}\left[Y_k \leq j\right] = \exp(\theta_j + \boldsymbol{x}_k\boldsymbol{\alpha})(1 + \exp(\theta_j + \boldsymbol{x}_k\boldsymbol{\alpha}))^{-1}$
      iii. $\log(\mathsf{P}\left[Y_k > j\right] / \mathsf{P}\left[Y_k \leq j\right]) = \theta_j + \boldsymbol{x}_k\boldsymbol{\alpha}$ : cumulative logit model R Code SAS Code

A: 6.3

2. Continuation logits:
   a. Model log odds ratio for membership in category vs. all below it
   b. or all above it.
   c. or adjacent categories. R Code

D. Complimentary Log-Log Link:

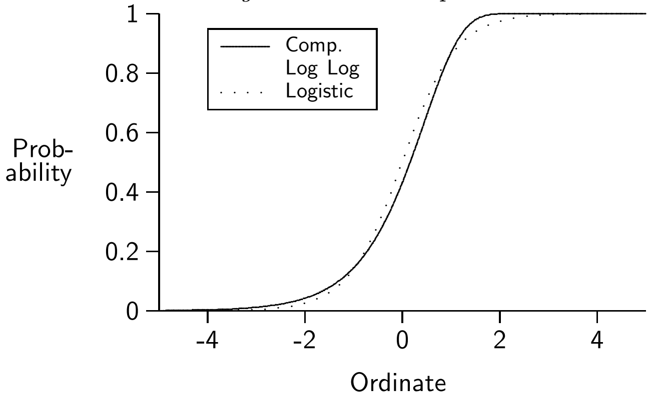
1. Interval censoring
   a. Assume underlying unobserved continuous variables $T_i$
      i. Let $S_i(t) = \mathsf{P}\left[T_i \geq t\right]$ .
   b. $Y_i = j$ iff $T_i \in (t_{j-1}, t_j]$ .
      i. Patients are screened at fixed intervals $t_0, t_1, t_2, \ldots, t_J$ for $t_{J+1} = \infty$ .
   c. $L = \prod_i \prod_{j=1}^{J+1} [S_i(t_{j-1}) - S_i(t_j)]$
   d. Let $\pi_{ij} = \mathsf{P}\left[T_i \leq t_j | T_i > t_{j-1}\right] = 1 - S_i(t_j)/S_i(t_{j-1})$
   e. $\mathsf{P}\left[T_i > t_j\right] = \prod_{l=1}^{j} (1 - \pi_{il})$
   f. $W_{ij}$ indicate which interval subject $i$ has event in.
      i. $1$ if subject $i$ had the event in interval $(t_j, t_{j+1}]$ ,
      ii. $W_{iJ+1} = 1$ if item not observed to fail,
      iii. 0 otherwise.
   g. Likelihood is $\prod_{i=1}^{n} \pi_{ij_i} \prod_{l=1}^{j_i - 1} (1 - \pi_{il}) = \prod_{i=1}^{n} \pi_{ij_i}^{W_{ij_i}} \prod_{l=1}^{j_i - 1} (1 - \pi_{il})^{1 - W_{il}}$ ,
      i. Likelihood for Bernoulli trials $W_{ij}$ with successs probabilities $\pi_{ij}$



      ii. Likelihood contributions multiply through conditioning rather than through independence.
2. Introduced regression parameters via proportional hazards.
   a. For any two subjects, $i$ and $m$ , any time index $j$ , $\mathsf{P}\left[T_i > t_j\right] = \mathsf{P}\left[T_m > t_j\right]^{\exp((z_i - z_m)\beta)}$
      i. $\prod_{l=1}^{j} (1 - \pi_{il}) = \prod_{l=1}^{j} (1 - \pi_{ml})^{\exp((z_i - z_m)\beta)}$
         - $(1 - \pi_{ij}) = (1 - \pi_{mj})^{\exp((z_i - z_m)\beta))}$
         - $(1 - \pi_{ij})^{\exp(-z_i\beta)} = (1 - \pi_{mj})^{\exp(-z_m\beta)}$
         - $\exp(-\boldsymbol{z}_i\boldsymbol{\beta})[-\log(1 - \pi_{ij})] = \exp(-\boldsymbol{z}_m\boldsymbol{\beta})[-\log(1 - \pi_{mj})]$
            - ▷ Note that because $1 - \pi_{ij} \in (0, 1)$ , quantities in brackests are positive.
         - $-\boldsymbol{z}_i\boldsymbol{\beta} + \log(-\log(1 - \pi_{ij})) = -\boldsymbol{z}_m\boldsymbol{\beta} + \log(-\log(1 - \pi_{mj})) \; \forall i, m$ .
   b. $\log(-\log(1 - \pi_{ij})) = \alpha_j + \boldsymbol{z}_i\boldsymbol{\beta}$
      i. for $\alpha_j = -\boldsymbol{z}_i\boldsymbol{\beta} + \log(-\log(1 - \pi_{ij}))$ .
   c. Gives *complimentary log log link* for regression model for $\pi_{ij}$ .
      i. Derived binary variables are same as for baseline logit model. R Code SAS Code
3. Link function
   a. Latent variable formulation with CDF $1 - \exp(-\exp(x))$ .
      i. Mean -0.577216, standard deviation 1.28255.
   b. $\log(-\log(\mathsf{P}\left[Y_k > j\right])) = \theta_j + \boldsymbol{x}_k\boldsymbol{\alpha}$ .
   c. Fig. 16/ compares the link functions

R Code SAS Code

*Fig. 16: Link Comparison*






This page intentionally left blank.



This page intentionally left blank.



This page intentionally left blank.