IX. Polytomous Regression:

A. Parameterization in general:
1. Take a probability associated with category \( j \) for individual \( k \)
2. Make it depend on covariates \( x_k \) through parameters \( \beta_j \)
3. Often force some components of \( \beta_j \) not to depend on \( j \)
   a. Generally first component of \( x_k \) is 1
   b. Hence first component of \( \beta_j \) is intercept
      i. Generally intercept depends on \( j \)
      ii. Generally the other components do not.

B. Baseline-Category logits
1. \( \log(P[Y_k = j]/P[Y_k = 0]) = \beta_j x_k \)
   a. \((1 + \sum_{j>0} \exp(\beta_j x_k))P[Y_k = 0] = 1\)
   b. \(P[Y_k = 0] = 1/(1 + \sum_{j>0} \exp(\beta_j x_k))\)
2. \( \log(P[Y_k = j]/P[Y_k = l]) = (\beta_j - \beta_l) x_k \)

C. How do I fit this?
1. Series of separate logistic regressions conditional on sum of that category and baseline category.
   a. Let \( Z_{kj} = \begin{cases} 1 & \text{if } Y_k = j \\ 0 & \text{otherwise} \end{cases} \)
   b. \( L_j(\beta_j) = \prod_k P[Y_k = 0]^{Z_{k0}} P[Y_k = j]^{Z_{kj}} \)
   c. \( \ell_j(\beta_j) = \sum_k[Z_{k0} \log(P[Y_k = 0])] + \sum_k[Z_{kj} \log(P[Y_k = j])] + \sum_k[Z_{kj} \beta_j x_k - (Z_{k0} + Z_{kj}) \log(1 + \exp(\beta_j x_k))] \)
   d. \( \ell_j'(\beta_j) = \sum_k[Z_{kj} - (Z_{k0} + Z_{kj}) \exp(\beta_j x_k)/[1 + \exp(\beta_j x_k)]^{-1} x_k = \sum_k[Z_{kj} - (Z_{k0} + Z_{kj}) \pi_k] x_k \)

2. All at once:
   a. \( L(\beta_1, \beta_2, \ldots, \beta_{J-1}) = \prod_k \prod_j P[Y_k = j]^{Z_{kj}} \)
   b. \( \ell = \sum_k \sum_j Z_{kj} \log(P[Y_k = j]^{Z_{kj}}) \)
   c. \( \frac{d}{d\beta_j} \ell = \sum_k [Z_{kj} \beta_j x_k + \sum_j Z_{kj} \log(P[Y_k = 0]^{Z_{kj}})] \)
   d. \( = \sum_k [Z_{kj} \beta_j x_k - \sum_j Z_{kj} \log(1 + \sum_j \exp(\beta_j x_k))] \)
   e. \( = \sum_k \frac{d}{d\beta_j} l = \sum_k [Z_{kj} - \sum_j Z_{kl} \pi_{lk}] x_k \)
      i. \( \pi_{jk} = \exp(\beta_j x_k)/[1 + \sum_{l>0} \exp(\beta_l x_k)] \)

3. I don’t know how to fit this model in SAS or R.

D. Cumulative logits:
1. Suppose \( \beta_j = (\theta_j, \alpha) \).
2. Suppose that \( W_k - \alpha x_k \) has CDF \( \exp(w)/(1 + \exp(w)) \)
   a. Mean 0, standard deviation 1.8138
3. Pick an increasing sequence \( \theta_j \)
4. Suppose that \( Y_k = j \) if \( W_k \in [\theta_{j-1}, \theta_j) \).
5. \( P[Y_k = j] = (1 + \exp(\theta_j + \alpha x_k))^{-1} \)
6. \( P[Y_k \leq j] = \exp(\theta_j + \alpha x_k)/(1 + \exp(\theta_j + \alpha x_k))^{-1} \)
7. \( \log(P[Y_k > j]/P[Y_k \leq j]) = \theta_j + \alpha x_k \) : cumulative logit model

E. Complementary Log-Log Link:
1. Previous analysis, with CDF \( 1 - \exp(-\exp(x)) \)
   a. Mean -0.577216, standard deviation 1.28255
2. \( \log(-\log(P[Y_k > j])) = \theta_j + \alpha x_k \)