Multiple Response Logit Model

Jee Dong Jun

16.8.2020

From Binary to Multi-response

- For binary response, we only have to consider π
- For multinomial responses, we need to consider all π_j
- With J categories, the log odds for $\binom{J}{2}$ pairs are described by logistic models.
- J-1 of these are enough

Baseline-logit model

- Comparing conditional distributions of response variable for two groups
- When Y has J categories, treat it as multinomial with $\{\pi_1(x), ..., \pi_J(x)\}$
- Logistic model describes log odd of each category with last category.

$$log \frac{\pi_j}{\pi_J} = \alpha_j + \beta_j x \quad j = 1..., J-1$$

Alligator Food example

- Responses are food choice of alligator (five categories)
- Classifies the alligators according to L,S,G
- Tried goodness of fit of baseline-category logit models

TABLE 7.2 Goodness of Fit of Baseline-Category Logit Models for Table 7.1

Model[a]	G^2	X^2	df
( )	116.8	106.5	60
(G)	114.7	101.2	56
(S)	101.6	86.9	56
(L)	73.6	79.6	48
$(L + S)$	52.5	58.0	44
$(G + L + S)$	50.3	52.6	40
Collapsed over G			
( )	81.4	73.1	28
(S)	66.2	54.3	24
(L)	38.2	32.7	16
$(L + S)$	17.1	15.0	12

[a] G, gender; S, size; L, lake of capture. See the text for details.

Estimating Response Probabilities

- We can directly estimate response probability instead of using logit

$$\pi_j(x) = \frac{exp(a_j + \beta_j x)}{1 + \sum_{h=1}^{J-1} exp(a_h + \beta_h x)}$$

- Note we can form likelihood equation like likelihood equation for logistic regression. However, we need to replace binomial distribution with multinomial distribution.

Multivariate GLMs

- For response vectors $y_i = (y_{i1}, ..., y_{iJ-1})$, expected values are $(\pi_1, ..., \pi_J)$
- $\mathbf{g}(\mu_i) = X_i\beta$
- With $g_j(\mu_i) = log\frac{\mu_{ij}}{1-(\mu_{i1}+...+\mu_{iJ-1})}$
- Therefore, multicategory logit model is kind of multivariate GLM
- It comes from utility representation $U_{ij} = a_j + \beta_j x_i + \epsilon_{ij}$

Ordinal Response

- Using ordinality of variable can be beneficial compared to using just nominal response.
- One possible way is to use cumulative logit using the fact that categories are ordered.
- $P(Y \leq j|x) = \pi_1(x) + ... + \pi_j(x)$
- Note each cumulative logit uses all J probabilities.

Proportional Odds Form of Cumulative Logit Model

- $logitP(Y \leq j|x) = \alpha_j + \beta x$
- Each cumulative logit has its own intercept but constant β
- This model assumes same effect of β across all logits
- Effect parameters are invariant to number of categories.

Latent Variable Motivation

- Suppose there exists latent variable Y* with probability distribution $G(y * -\eta(x))$
- $Y = j$ when $a_j < y* < a_{j+1}$
- Appropriate location family distribution gives proportional odds structure.

Checking the Proportional Odds Assumption

- It was assumed that β is constant throughout different cumulative logits
- One can generalize this model replacing β with β_j
- Try score test on whether complex model fits better.
- If proportional odds assumption fails, there are few alternatives.

Alternative models for ordinal responses

- We can use different link function that is the inverse of the continuous cdf G
- Logit link function is special case when G is standard logistic cdf
- Cumulative probit model uses standard normal cdf
- Complementary log-log link can be used also.

Adjacent-Categories Logit Models

- Instead of using cumulative probabilities, we use pair of adjacent response probabilities
- $log \frac{\pi_j(x)}{\pi_{j+1}} = a_j + \beta x$
- This model gives baseline-category logit model with same β but modified x
- Effect of x stacks up (acknowledging order of categories)

Other alternatives

- Continuation-Ratio Logit Models
- Continuation-ratio is $log \frac{\pi_j(x)}{\pi_{j+1}+...+\pi_J}$
- Useful when sequential mechanism determines the response outcome

Stochastic Ordering Location Effect

- Cumulative link models are stochastically ordered on the response.
- For pair of x1 and x2, $P(Y < j|x1) < P(Y < j|x2)$ for all j or vise versa.
- This is violated because the dispersion also varies with x.
- For example, more dispersion might occur at x1 than at x2 even though they concentrate around the same location.

Conditional Independence in IxJxK Tables

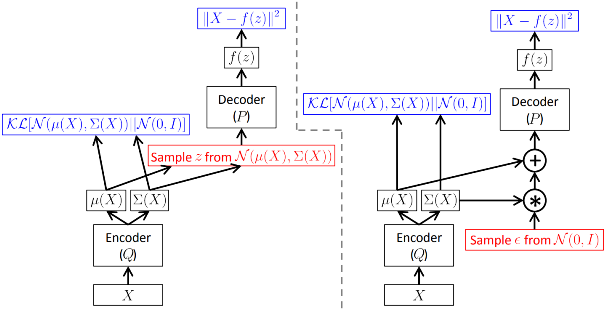

Figure 2: Conditional Test

Generalized CMH test

- It generalized to multiple rows and columns.
- Three possible cases, both nominal, both ordinal or one each
- Conditioning on row and column totlas (I-1)(J-1) nonredundant cell counts
- Generalized CMH test is closely related to score test of conditional indepndence.