Multinomial Logit Models for Variable Response Categories Ordered

Malika CHIKHI[1*], Thierry MOREAU[2] and Michel CHAVANCE[2]

[1] **Mathematics Department, University of Constantine 1, Ain El Bey, 25000, Algeria**

[2] **Biostatistics Department, INSERM U 780, 94807 Villejuif Cedex, France**

[2] **Biostatistics Departments, INSERM U 780, 94807 Villejuif Cedex, France**

Abstract

This paper present three type of logits for ordered response to c categories. We interpreted in term of distribution two logistics models: the cumulative and continuation-ratio logit for ordinal response variables to c categories.

Keywords: *statistics, ordered response, categorical variable, multinomial logit, adjacent-categories logit , continuation-ratio logit, cumulative logit.*

1. Introduction

Although described for several years, the extension of logistic regression in case the ordered categorical variable [1, 2] is not always used when it could be. The objective of this note is to present the different appropriate models for this estimation as described in detail in the reference book of A. Agresti [2, 3].
An aspect of some of these models, which can be useful for their implementation, is particularly emphasized; it is their interpretation in the case where the classes of the dependent variable can be considered from the partition of the variation interval of a random variable underlying continuous. For many models, the interpretation is described and the distribution family is deducted (3, 4).

2. Results and Discussion

2.1. Multinomial logit models for variable response categories

2.1.1. The generalized Logit Model

Let $\left(\pi_j(x)\right)$ denote probabilities of response j, j = 1... c, at the i th setting of values of k explanatory variables $X_i = (x_{i1} \dots x_{ik})'$. The generalized logit model is [4, 5]:

$$L_j(x_i) = log\left(\frac{\pi_j(x_i)}{\pi_c(x_i)}\right)$$

$$or \quad L_j(x_i) = log\left(\frac{pr(\, y=j/x_i)}{pr(\, y=c/x_i)}\right) \qquad j= 1\dots c.$$

In terms of the response probabilities, the model is writing:

$$\pi_j(x_i) = \frac{exp(\beta'_j \, x_i)}{\sum_{h=1} exp(\beta'_h \, x_i)}$$

2.1.2. Multinomial Logit Model

For subject i and response choice j, let $X_{ij} = (x_{ij1} \dots x_{ijk})'$ denote the values of the k explanatory variables. Conditional on the set of response choices C_i for subject i , in terms of the response probabilities, general multinomial logit model is defined as:

$$\pi_j(xij) = \frac{exp(\beta' \, x_{ij})}{\sum_{h \in ci} exp(\beta' \, x_{ih})}$$

All models given below are special cases of the multinomial logit model.

2.1.3. Multinomial logit models for variable response categories ordered.

When variable response categories have a natural ordering, we utilize the ordering directly in the way we construct logits.

We present three types of logit for ordered response [1, 2] to c categories.

Let $X'(x_1, \dots, x_k)$ denote a vector of k explanatory variables and the dependent variable Y for which c classes are defined and ordered actually supposed to be continuous; different logit below expressed.

a) Adjacent-categories Logit Models

Let$\{\pi_1(x), \ldots + \pi_c(x)\}$denote response probabilities at value x for a set of explanatory variables.

The representations of the adjacent - categories logits are:

$$L_j(x) = log\left(\frac{\pi_j(x)}{\pi_{j+1}(x)}\right)$$

or $L_j(x) = log\left(\frac{pr(\,y=j/x)}{pr(\,y=\,j+1/x)}\right)$, j=1,..., c. (1)

b) Continuation- Ratio Logit Models

Continuation- Ratio logits [2, 3] are defined as:

$$L_j(x) = log\left(\frac{\pi_j(x)}{\pi_{j+1}(x) + \cdots + \pi_c(x)}\right)$$

or $L_j(x) = log\left(\frac{pr(\,y=j/x)}{pr(\,y\geq\,j+1/x)}\right)$, j=1... c. (2)

or as:

$$L_j(x) = log\left(\frac{\pi_{j+1}(x)}{\pi_1(x)+\cdots+\pi_j(x)}\right)$$

or $L_j(x) = log\left(\frac{pr(\,y=j+1/x)}{pr(\,y\leq\,j/x)}\right)$, j= 1,....,c. (2')

c) Cumulative Logit Models

Another way to use ordered response categories is by forming logits of cumulative probabilities

$$pr(\,y \leq j/x) = \pi_1(x) + \cdots + \pi_j(x), j = 1\ldots \text{c.}$$

The cumulative logits are defined as:

$$L_j(x) = log\left(\frac{\pi_1(x)+\cdots+\pi_j(x)}{\pi_{j+1}(x)+\cdots+\pi_c(x)}\right) \quad ,$$

$$L_j(x) = log\left(\frac{pr(\,y\leq j/x)}{pr(\,y\geq\,j+1/x)}\right), \text{j= 1,..., c. (3)}$$

or as

$$L_j(x) == log\left(\frac{\pi_{j+1}(x)+\cdots+\pi_c(x)}{\pi_1(x)+\cdots+\pi_j(x)}\right) \quad ,$$

$$L_j(x) = log\left(\frac{pr(\,y\geq j+1/x)}{pr(\,y\leq\,j/x)}\right), \text{j= 1,..., c. (3')}$$

Each cumulative logit uses all c response categories.

Note. It is noticed that these three logits are identical in the case c = 2.The general model is given by:

$$L_j(x) = \alpha_j + \beta' x$$

β' is a vector of unknown parameters (β_1, ..., β_k) and the estimate with α_j ,j = 1,... ,c-1.

We note in the case where one of the explanatory variables (for example x_1) is categorical and ordinal, if score u_i can be assigned to each category i with $i = 1 \ldots k$, we can write:

$$\begin{aligned} &L_j(\,x_1 = u_i, x_2, \ldots, x_k) \\ &\qquad - \; L_j(\,x_1 = u_{i-1}, x_2, \ldots, x_k) \\ &\qquad = \beta'(u_i - u_{i-1}) \end{aligned}$$

Hence, the exp β' (u_i - u_{i-1}) can be interpreted as the odds ratio and the model assumes that the odds ratio is the same for all j, in particular in the case where scores u_i, are consecutive integers .

If $u_i = i$, the odds ratio is exp(β) and $L_j(x)$ is given by (3,3'), it is the odds ratio between the binary variable defined by:

$$F_j(x) = pr(y \leq j/x) \text{ and } 1 - F_j(x)\,;$$

and the binary variable defined by membership in one or other of two adjacent categories x_1.

• In the case where x_1 is the only explanatory variable as ordinal above, the model was called [1, 2]: logistic model association uniform.

• In the case where one of the k variables is qualitative nominal categories, the general model applies in defining example (k-1) binary variables indicative of each category, choosing a reference category for which (k-1) variables are zero. Note that when only explanatory variable of this type is considered, the model is denominated: row effects model [5]

• Using logit models to adjacent categories were mainly presented by Goodman [6]. Among the authors who have studied models in various forms include Snell [7], Williams and Grizzle [8], Mc Cullagh [9] and Bock [10].The authors suggested the continuation- ratio logits models are Thompson [11], Cox [12], Fienberg and Mason[13] and Mc Cullagh and Nelder [5].

2.1.4. Estimates the parameters of the model

The authors suggested that the row effects logistic models are Grizzle and Williams [8]. They adjusted these models using the method of weighted least squares. In 1973, Bock and Yates used the method of maximum likelihood [10].

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 6, No 2, November 2012
ISSN (Online): 1694-0814
www.IJCSI.org

2.1.5. Test of parameters signification

The hypothesis concerning the nullity of one or more parameters of the general logistic model can be tested by a test log likelihood ratio.
In the case of categorical the explanatory variables, the test is similar to the chi-square test defined above. It is interpreted in the same way in terms of conditional independence between the explanatory variables and the response variable.

3. Interpretation models with variable logistics ordinal response

In this section, the assumptions of the logistic model, [formulas (1), (2) and (3)], are explained in terms of the distribution of the response variable Y conditional on the explanatory variables. Two particular models are considered corresponding to cumulative logit and continuation-ratio logit [2, 3].

3.1. Cumulative Logit Models

Definitions and notations

The dependent variable Y, where c ordered classes are defined, is assumed to be continuous in reality, different classes constitute a partition of the interval of variation. This situation corresponds to the majority of cases encountered in practice.
Noting $a_1 < a_2 < \ldots < a_c$, limits (observed or not) of contiguous intervals where c Y takes its values, the probability that Y belongs to the class j is:

$$\pi_j = pr\left(a_{j-1} < Y < a_j\right),$$

and the distribution function of Y in a_j (j = 1… c) is:

$$F(a_j) = pr(Y \le a_j) = \pi_1 +, \ldots, +\pi_j$$

Whether X is an explanatory variable quantitative or qualitative. The calculations are made below with the only explanatory variable X but would be unchanged if other variables were included in the model.
The distribution function of Y conditional X in $(\mathrm{Y} = a_j)$ is:

$$Fx(a_j) = pr\{(Y \le a_j) / \mathrm{X} = \mathrm{x}\}$$

and $Sx(a_j) = 1 - Fx(a_j)$ the corresponding survivals function.

$$Lj(x) = log\left(\frac{pr(y=j/x)}{pr(y\ge j+1/x)}\right) \qquad \text{j= 1,....,c.}$$

The model is written

$$\frac{1 - Fx(a_j)}{Fx(a_j)} = exp(\alpha_{j+}\beta x) \quad ,$$
$$j = 1, \ldots, c \ . \qquad (4)$$

with

$$Fx(a_j) = \frac{1}{1 + exp(\alpha_{j+}\beta x)}$$

x_1 and x_2 are two values taken by X such that $x_1 < x_2$, according to the model:

$$\frac{Sx_{2(a_j)}/Fx_{2(a_j)}}{Sx_{1(a_j)}/Fx_{1(a_j)}} = \exp\beta(x_2 - x_1) \quad ,$$
$$j = 1, \ldots, c \ . \qquad (5)$$

If $\beta \ge 0$, it follows that the above report is greater than 1 for all j which implies:

$$Fx_1(a_j) > Fx_2(a_j) \text{ , j=1… c.}$$

The distribution of Y on each a_j conditional on x_2 appears stochastically larger than Y conditional on x_1.
If $\beta < 0$, the opposite is true.

More precisely, the expression (5) defines a relationship between the distributions of Y conditional on x_1 and x_2.
This relationship is verified in particular for distribution logistics translated from one another. Indeed, if the distribution of Y conditional on x is assumed to logistics, we can always write:

$$Fx(a_j) = \frac{1}{1 + exp(-a_{j+}\beta x)}$$

The model is then tested by asking $\alpha j = (-a_j)$ and the translation parameter between the distributions of Y conditional on x_1 and x_2 is $[exp\,\beta(x_2 - x_1)]$ indeed, it appears that:

$$Fx_2(a_j) = Fx_1\left[a_j + \beta(x_2 - x_1)\right]$$

It may be noted that in the case where x is an ordinal variable taking consecutive integer values, exp (β), is the odds-ratio "local global" variables previously defined between two consecutive x and for any one given day.

3.2. continuation-ratio logit models

In the previous section, the approach has been to seek the distribution of the random variable underlying continuous Y coincides with the distribution function defined by the cumulative logit model, the points a_j are the ends of the classes defined by Y; in what follows the approach will be

different and will continue to find the limit of the discrete model defined for categorical variables [14, 15].
Two different continuation-ratio logit models relationship can be considered.

Noting:

$$P_j = pr\{(\ a_j < Y \leq a_{j+1}) \,/\, Y > a_j\}$$

and $P'_j = pr\{(a_j < Y \leq a_{j+1}) \,/\, Y < a_{j+1}\}$

They express the logit P_j or P'_j conditional on X = x in the form:

$$L_j(x) = (\alpha_j + \beta x)$$

The interpretation presented for the limit model when one a $_{j+1}$ tends to a $_j$ for all j (and c tends to infinity), that is to say, when it is the achievements of the continuous variable Y that are observed.

$$p_j = \frac{pr(\ a_j < Y \leq a_{j-1})}{S\ (a_j)}$$

$$tends\ to\ \frac{f(a_j)}{S\ (a_j)} dy \ = \lambda(a_j)\ dy$$

In this case: Where $\lambda(a_j)$ is by definition the function of «instantanes risk function » Y on a_j and similarly (1-p_j) tends to 1.

The same way:

$$p'_j = \frac{pr(\ a_j < Y \leq a_{j-1})}{F\ (a_{j-1})}$$

$$tends\ to\ \frac{f(a_j)}{F\ (a_j)} dy$$

Where f (a_j) is the probability density function of Y on a_j of more 1-P_j tends to 1.
Let x_1 and x_2 are two values taken by X such that $x_1 < x_2$, conditioning with respect to X, the two models are respectively:

$$\frac{fx(a_j)}{S\ x(a_j)} = exp(\alpha_{j+}\beta x)$$

$$and\ \ \frac{fx(a_j)}{F\ x(a_j)} = exp(\alpha_{j+}\beta x)$$

Writing the first model for x_1 and x_2, we obtain:

$$\frac{f\ x_2(a_j)}{Sx_2(a_j)} dy = \frac{f\ x_1(a_j)}{Sx_1(a_j)} dy \times \exp\beta(\ x_2 - x_1)$$

After integration:

$$Sx_2(a_j) = \{\ Sx_1(a_j)\}^{\exp\beta(\ x_2 - x_1)}$$

The second model would lead to the same:

$$Fx_2(a_j) = \{\ Fx_1(a_j)\}^{\exp\beta(\ x_2 - x_1)}$$

Both of these expressions define a special relationship between the distributions of Y conditional on x_1 and x_2.
The first model is known as the "proportional hazards model" or "Cox model" [12,15,16] and is commonly used in survival studies, a special case of this model is one where:

$$S_x(Y) = \exp\ \{(-\lambda y\ \exp\ (\beta x)\}$$

corresponding to an exponential distribution for the distribution of Y conditional on X.
We have been able find a family of distribution satisfying the second model: this family includes the Pareto distribution with the density and distribution functions are:

$$f(y) = \left(\frac{t}{k^t}\right) y^{\ t-1}\ \ ;\ F(y) = \left(\frac{y^t}{k^t}\right) = \left(\frac{y}{k}\right)^t$$

Indeed, writing:

$$Fx_1(a_j) = \left(\frac{a_j}{k}\right)^{t\exp\beta(\ x_1)} ;$$

$$Fx_2(a_j) = \left(\frac{a_j}{k}\right)^{t\exp\beta(\ x_2)}$$

The model is verified.

4. Conclusion

To choose between these different models, we can consider these results as a priori information that is available on the distribution of the variable Y and how it varies according to the categories of the explanatory variable.
According to information from the study that we have we can choose the first model (cumulative logit) if, between the different categories of the variable X, variable Y is translated, the second or third model (continuation-ratio) in the event of a change of scale.

5. Acknowledgements

I express my gratitude and appreciation to the team of the INSERM U780, Villejuif France for giving me the opportunity to realize this paper.

References

[1] A. Agresti, « Categorial Data Analysis », John Wiley and Sons Inc., 1991.

[2] A. Agresti, « Analysis of Ordinal Categorial Data », John Wiley and Sons Inc., 2010.

[3] A. Agresti, « An Introduction to Categorical Data Analysis», John Wiley & Sons, 2007.

[4] R.H. Myers, D. C. Montgomery, G. Geoffrey Vining, Timothy J. Robinson, «Generalized Linear Models», Amazon France, 2012.

[5] M.C. Cullagh and J. Nelder, « Generalized Linear Models », London Chapman and Hall, 1983.

[6] L.A. Goodman, « The Analysis of Dependence In Cross-classification having Ordered Categories using log-linear models for frequencies and log-linear models for odds », *Biometrics*, 39, 1983, pp.149-160.

[7] D. Mc Fadden, « Regression Based Specification Tests for the Multinomial Logit Models », *Journal of Econometrics*, 34, 1987, pp.63-82.

[8] O.D. Williams and J.E. Grizzle, «Analysis of Contingency Tables having Ordered Responses Categories », J. *Amer. Statist. Assoc*., 67, 1972, pp.55-63.

[9] Scott Menard, « Logistic Regression: From Introductory to Advanced Concepts and Applications», Amazon France, 2009.

[10] R.D .Bock and G. Yates « Multiqual log-linear Analysis of Nominal or Ordinal Qualitative by the Method of Maximum Likelihood », Chicago, International Educational Services, 1973.

[11] W.A. Thompson, « On the Treatment of Grouped Observations in Life Studies », *Biometrics*, 33, 1977, pp.463-470.

[12] D.R. Cox, « Regression Models and Life Tables (With discussion) », *J. Roy. Statist. Soc*., B.34, 1972, pp.187-220.

[13] S.E. Fienberg and W.M. Mason, « Identification and Estimation of Age Period Cohort Models in the Analysis of Discrete Archival Data », *Sociological Methodology*, San Francisco, Jossey Bass, 1979, pp. 1-67.

[14] Scott Menard, « Logistic Regression: From Introductory to Advanced Concepts and Applications», Amazon France, 2009.

[15] Raymond H. Myers, Douglas C. Montgomery, G. Geoffrey Vining, Timothy J. Robinson, «Generalized Linear Models», Amazon France, 2012.

[16] D.G. Clayton, « Some Odds-ratio Statistics for the Analysis of Ordered Categorical Data », *Biometrical*, 61, 1974, pp.525-531.