Thailand Statistician
January 2012; 10(1) : 87-105
http://statassoc.or.th
Contributed paper

A New Generalized Ordinal Logit Model for Multicategory Response Data

Somsri Jamroenpinyo [a] *, Timothy E. O'Brien [b], and Chinnaphong Bumrungsup [a]

[a] Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathum Thani, 12121, Thailand.

[b] Department of Mathematical Sciences, Loyola University Chicago, USA.

* Author for correspondence; e-mail: huajam@gmail.com

Received: 7 June 2011
Accepted: 6 February 2012

Abstract

This paper introduces and illustrates a new generalized ordinal logit (GOL) model which connects the four commonly-used multicategory logit models by using two hyper-parameters. The commonly used models in multicategory models are the adjacent-categories logit model (AC), the proportional odds (PO) model, and two variants of the continuation-ratio logit (CR) models. The GOL model generalizes these four models in the sense that each is a special case of the larger GOL model, and this GOL model is used for multicategory response data. In this article, we discuss (maximum likelihood) estimation and testing related to the GOL model, providing SAS/IML computer programs for the same, and illustrating the use of the proposed model with two real datasets.

Keywords: adjacent-categories, baseline-category logits, continuation-ratios, multinomial distribution, nominal responses, ordinal responses, proportional odds

1. Introduction

Regression analysis is a statistical method tool that has been used in the fields of many disciplines such as in business, social sciences, medicine, etc. The purpose for this tool is to predict response variables Y by using explanatory variables X. When the response variable Y is qualitative and is classified into k categories. Multicategory Logistic regression models have been handled. These models assume that the response counts for the categories of Y have a multinomial distribution. They include many common models [1] such as the baseline-category logit (BCL), adjacent-categories logits (AC), proportional odds (PO) and continuation-ratio logits models (CR). In Minitab, the default for "ordinal logistic regression" is BCL model. This is the same as the default for "nominal logistic regression" which is AC model. These are the defaults for other software packages. Thus, we will concern only with AC. For some datasets, these models are not sufficiently rich to fit these datasets. In order to solve this problem, the multicategory logit models need to be extended. So the generalized ordinal logit (GOL) model is proposed to connect multicategory logit models.

In addition, we use the model name followed by "s" to denote that the independent variable x has been scaled using the usual Box-Cox transformation formula in Schabenberger and Pierce [2]. Thus, "ACs" denotes the AC (adjacent category) model with scaled x variable.

This article is organized as follows: In Section 2, we describe the methodology for connecting the multicategory logit models. Results from GOL model are given in Section 3. We present two examples to illustrate the GOL model in Section 4, and our discussion is provided in Section 5.

2. Methodology

Considering the common models in multicategory logit models such as AC, PO and CR. The CR model has two forms of logits called CRA and CRB. If PO model has different slopes, we call it the UPO model. The associated logit transformations for these various models are given in Table 1.

Table 1. Multinomial Logits for the AC, UPO, CRA and CRB Models.

k	logits	AC	UPO	CRA	CRB
3	1	$\log\left[\frac{\pi_1}{\pi_2}\right]$	$\log\left[\frac{\pi_1}{\pi_2+\pi_3}\right]$	$\log\left[\frac{\pi_1}{\pi_2}\right]$	$\log\left[\frac{\pi_1}{\pi_2+\pi_3}\right]$
	2	$\log\left[\frac{\pi_2}{\pi_3}\right]$	$\log\left[\frac{\pi_1+\pi_2}{\pi_3}\right]$	$\log\left[\frac{\pi_1+\pi_2}{\pi_3}\right]$	$\log\left[\frac{\pi_2}{\pi_3}\right]$
4	1	$\log\left[\frac{\pi_1}{\pi_2}\right]$	$\log\left[\frac{\pi_1}{\pi_2+\pi_3+\pi_4}\right]$	$\log\left[\frac{\pi_1}{\pi_2}\right]$	$\log\left[\frac{\pi_1}{\pi_2+\pi_3+\pi_4}\right]$
	2	$\log\left[\frac{\pi_2}{\pi_3}\right]$	$\log\left[\frac{\pi_1+\pi_2}{\pi_3+\pi_4}\right]$	$\log\left[\frac{\pi_1+\pi_2}{\pi_3}\right]$	$\log\left[\frac{\pi_2}{\pi_3+\pi_4}\right]$
	3	$\log\left[\frac{\pi_3}{\pi_4}\right]$	$\log\left[\frac{\pi_1+\pi_2+\pi_3}{\pi_4}\right]$	$\log\left[\frac{\pi_1+\pi_2+\pi_3}{\pi_4}\right]$	$\log\left[\frac{\pi_3}{\pi_4}\right]$

We present the GOL model to connect the multicategory logit models by using two extra parameters θ_1 and θ_2, called "GOL parameters", where $0 \le \theta_1 \le 1$ and $0 \le \theta_2 \le 1$. When we fix the value of θ_1 and θ_2 to be 0 or 1, the GOL model collapses to be AC, PO, CRA and CRB models (see Table 3). As such, the GOL model is the generalized of these commonly-used multicategory logit models.

The GOL model with *k* ≥ 3 categories has *k* –1 logits, and these are of the form:

The first equation:

$$\log\left[\frac{\pi_1}{\pi_2+\theta_1(\pi_3+\ldots+\pi_k)}\right] = \alpha_1+\beta_1 x$$

The middle (*k*-3) logits:

$$\log\left[\frac{\theta_2(\pi_1+\ldots+\pi_{(k-m)-1})+\pi_{k-m}}{\pi_{(k-m)+1}+\theta_1(\pi_{(k-m)+2}+\ldots+\pi_k)}\right] = \alpha_{k-m}+\beta_{k-m} x \qquad , \text{m = 2, 3 … (k-2);}$$

The last equation:

$$\log\left[\frac{\theta_2(\pi_1+\ldots+\pi_{k-2})+\pi_{k-1}}{\pi_k}\right] = \alpha_{k-1}+\beta_{k-1} x$$

In this paper we focus on the GOL model for the cases *k*= 3 and *k* = 4, but generalizations to larger value of *k* are straightforward. The GOL logits are shown in Table 2.

Table 2. GOL logits.

k	logits	GOL
3	1	$\log\left[\dfrac{\pi_1}{\pi_2+\theta_1\pi_3}\right]$
	2	$\log\left[\dfrac{\theta_2\pi_1+\pi_2}{\pi_3}\right]$
4	1	$\log\left[\dfrac{\pi_1}{\pi_2+\theta_1(\pi_3+\pi_4)}\right]$
	2	$\log\left[\dfrac{\theta_2\pi_1+\pi_2}{\pi_3+\theta_1\pi_4}\right]$
	3	$\log\left[\dfrac{\theta_2(\pi_1+\pi_2)+\pi_3}{\pi_4}\right]$

If we assign the values of θ_1 and θ_2, we obtain the sub-models as shown in Table 3.

Table 3. Corresponding values of θ_1 and θ_2 for Sub-models.

θ_1	θ_2	Sub-models
0	0	AC
1	1	UPO
0	1	CRA
1	0	CRB

Then we derive formulas to find parameter estimates for the GOL model in this manner.

Let $ex_1 = \exp(\alpha_1 + \beta_1 x)$, $ex_2 = \exp(\alpha_2 + \beta_2 x)$, …respectively.

We derive probabilities and find parameter estimates for the GOL model as follows:

From the GOL logit equations that have $k-1$ equations with the constraint $\sum_{i=1}^{k}\pi_i = 1$.

Taking the anti-logarithm, we solve the system of linear equations to get the estimated probabilities.

When $k=3$.The GOL logit equations are

$$\log\left[\frac{\pi_1}{\pi_2+\theta_1\pi_3}\right] = \alpha_1+\beta_1 x$$

$$\log\left[\frac{\theta_2\pi_1+\pi_2}{\pi_3}\right] = \alpha_2+\beta_2 x$$

with constraint $\sum_{i=1}^{3}\pi_i = 1$.Taking the anti-logarithm, we get equations

$$\frac{\pi_1}{\pi_2+\theta_1\pi_3} = ex_1$$

$$\frac{\theta_2\pi_1+\pi_2}{\pi_3} = ex_2$$

or

$$\pi_1 = ex_1(\pi_2+\theta_1\pi_3)$$

$$\theta_2\pi_1+\pi_2 = ex_2\,\pi_3$$

$$\pi_1+\pi_2+\pi_3 = 1$$

When $k=4$.The GOL logit equations are

$$\log\left[\frac{\pi_1}{\pi_2+\theta_1(\pi_3+\pi_4)}\right] = \alpha_1+\beta_1 x$$

$$\log\left[\frac{\theta_2\pi_1+\pi_2}{\pi_3+\theta_1\pi_4}\right] = \alpha_2+\beta_2 x$$

$$\log\left[\frac{\theta_2(\pi_1+\pi_2)+\pi_3}{\pi_4}\right] = \alpha_3+\beta_3 x$$

with constraint $\sum_{i=1}^{4}\pi_i = 1$.Taking the anti-logarithm, we get equations

$$\frac{\pi_1}{\pi_2+\theta_1(\pi_3+\pi_4)} = ex_1$$

$$\frac{\theta_2\pi_1+\pi_2}{\pi_3+\theta_1\pi_4} = ex_2$$

$$\frac{\theta_2(\pi_1+\pi_2)+\pi_3}{\pi_4} = ex_3$$

or

$$\begin{aligned}
\pi_1 &= ex_1(\pi_2+\theta_1(\pi_3+\pi_4)) \\
\theta_2\pi_1+\pi_2 &= ex_2(\pi_3+\theta_1\pi_4) \\
\theta_2(\pi_1+\pi_2)+\pi_3 &= ex_3\pi_4 \\
\pi_1+\pi_2+\pi_3+\pi_4 &= 1
\end{aligned}$$

For the parameter estimates of models, we use the maximum likelihood estimation method. Let L be the log-likelihood expressed at a single value of x and corresponding to the multinomial distribution. θ is vector of parameters in the models. The likelihood function (denoted "l") and log likelihood function (denoted "L") for $k = 3,4$ are shown in Table 4. Maximum likelihood estimates for the various models are obtained by maximizing L (or equivalently l) with respect to the model parameters θ. Thus, the probability mass function of Y, represented by $f(y)$ is as follows;

For $k = 3$, $f(y) = \frac{n!}{y_1!y_2!y_3!}\pi_1^{y_1}\pi_2^{y_2}\pi_3^{y_3}$ Where $\sum_{i=1}^{3}\pi_i = 1$ and $\sum_{i=1}^{3}y_i = n$.

$k = 4$, $f(y) = \frac{n!}{y_1!y_2!y_3!y_4!}\pi_1^{y_1}\pi_2^{y_2}\pi_3^{y_3}\pi_4^{y_4}$ Where $\sum_{i=1}^{4}\pi_i = 1$ and $\sum_{i=1}^{4}y_i = n$.

Table 4. Likelihood and Log-Likelihood Expressions for k = 3 and 4 Cases.

k	Likelihood function	Log likelihood function
3	$l \propto \pi_1^{y_1}\pi_2^{y_2}\pi_3^{y_3}$	$L = c + y_1 \log \pi_1 + y_2 \log \pi_2 + y_3 \log \pi_3$
4	$l \propto \pi_1^{y_1}\pi_2^{y_2}\pi_3^{y_3}\pi_4^{y_4}$	$L = c + y_1 \log \pi_1 + y_2 \log \pi_2 + y_3 \log \pi_3 + y_4 \log \pi_4$

The maximum likelihood estimate (MLE) of θ [3] is given by

$$\hat{\theta} = \arg\max_{\theta} \prod_{i=1}^{n}\prod_{j=1}^{k} \pi_j^{y_{ij}} = \arg\max_{\theta} \sum_{i=1}^{n}\sum_{j=1}^{k} y_{ij} \log \pi_j$$

The variance -covariance matrix of $\hat{\theta}$ for large samples is approximated by

$$Var(\hat{\theta}) \approx \frac{1}{N} M^{-1}(\theta)$$

where N is the total number of observations; in this expression M(θ) is the Fisher information matrix for the given design with jkth (ij)element equal to the negative of the expected value of the second derivative of L with respect to the jth (i)and kth (j)parameters. We assume that $M(\theta)$ is nonsingular; in the event that it is singular, a generalized inverse is used to obtain the corresponding variance-covariance matrix.

Since the models are nonlinear, the maximum likelihood estimate may be found by using iterative methods such as the Gauss-Newton iterative method. Bates & Watts [4] indicated that The Gauss-Newton iterative method is continued until the value of parameter on successive iterations stabilize. This can be measured by the size of each parameter increment relative to the previous parameter value, which is the basis for one of the common criteria used to declare convergence. In SAS, when the Newton-Raphson iterative method is used, criteria for convergence is that at least one element of the (projected) gradient is greater than 0.001.

For each given dataset, we write a SAS/IML program to find parameter estimates in only AC, UPO, CRA, CRB and GOL models, using the Newton-Raphson iterative method. Then we choose the best model to describe each dataset by using the minimum negative log-likelihood.

After the model has been selected, we use parameter estimates in the model to obtain estimated probabilities and fitted values for the model.

3. Results

After solving systems of equations in GOL logit model in section 2, we get probabilities as shown in Table 5. From this table, If we specify the values of θ_1 and θ_2 as shown in Table 3, we will obtain the probabilities in multinomial logits model as shown in Table 6.

Table 5. Probabilities in GOL models for *k* = 3, 4.

k	Categories	Probabilities
3	1	$\pi_1 = \dfrac{(\theta_1 + ex_2)ex_1}{den}$
	2	$\pi_2 = \dfrac{ex_2 - \theta_1\theta_2 ex_1}{den}$
	3	$\pi_3 = \dfrac{1+\theta_2 ex_1}{den}$
	den	$1+(\theta_1+\theta_2-\theta_1\theta_2)ex_1 + ex_2 + ex_1\,ex_2$
4	1	$\pi_1 = \dfrac{ex_1[\theta_1\{1+(1+\theta_2-\theta_1\theta_2)ex_2+ex_3\}+ex_2ex_3]}{den}$
	2	$\pi_2 = \dfrac{-\theta_1\theta_2 ex_1 + \theta_1 ex_2 - \theta_1\theta_2(1-\theta_1)ex_1ex_2 - \theta_1\theta_2 ex_1ex_3 + ex_2ex_3}{den}$
	3	$\pi_3 = \dfrac{\theta_1\theta_2(\theta_2-1)ex_1 - \theta_1\theta_2 ex_2 + ex_3 - \theta_1\theta_2 ex_1ex_2 + \theta_2 ex_1ex_3}{den}$
	4	$\pi_4 = \dfrac{1+\theta_2(1+\theta_1-\theta_1\theta_2)ex_1 + \theta_2 ex_2 + \theta_2 ex_1ex_2}{den}$
	den	$1+(\theta_1+\theta_2-\theta_1\theta_2)[ex_1+ex_2+ex_1\,ex_2+ex_1\,ex_3]$ $+ex_3+ex_2\,ex_3+ex_1\,ex_2ex_3$

Note: "den" is represented for denominator in probability terms.

Table 6. Probabilities in multinomial logits model.

k	Probabilities	AC	UPO	CRA	CRB
3	π_1	$\frac{ex_1 ex_2}{den}$	$\frac{ex_1}{1+ex_1}$	$\frac{ex_1 ex_2}{den}$	$\frac{ex_1}{1+ex_1}$
	π_2	$\frac{ex_2}{den}$	$\frac{ex_2 - ex_1}{den}$	$\frac{ex_2}{den}$	$\frac{ex_2}{den}$
	π_3	$\frac{1}{den}$	$\frac{1}{1+ex_2}$	$\frac{1}{1+ex_2}$	$\frac{1}{den}$
	AC UPO CRA CRB	den den	$1+ex_2+ex_1 ex_2$ $(1+ex_1)(1+ex_2)$		
4	π_1	$\frac{ex_1 ex_2 ex_3}{den}$	$\frac{ex_1}{1+ex_1}$	$\frac{ex_1 ex_2 ex_3}{den}$	$\frac{ex_1}{1+ex_1}$
	π_2	$\frac{ex_2 ex_3}{den}$	$\frac{ex_2 - ex_1}{dena}$	$\frac{ex_2 ex_3}{den}$	$\frac{ex_2}{dena}$
	π_3	$\frac{ex_3}{den}$	$\frac{ex_3 - ex_2}{denb}$	$\frac{ex_2}{denb}$	$\frac{ex_3}{den}$
	π_4	$\frac{1}{den}$	$\frac{1}{1+ex_3}$	$\frac{1}{1+ex_3}$	$\frac{1}{den}$
	AC UPO CRA CRB	den dena denb den	$1+ex_3+ex_2 ex_3+ex_1 ex_2 ex_3$ $(1+ex_1)(1+ex_2)$ $(1+ex_2)(1+ex_3)$ $(1+ex_1)(1+ex_2)(1+ex_3)$		

4. Application to datasets

In this section, we apply the GOL model described in section 2 to two real datasets from various references. We first consider the pregnant mice dataset given in Agresti [5]. This dataset was the outcome from a developmental toxicity study for 1435 pregnant mice. Each mouse was exposed to one of five concentration levels of diethylene glycol dimethyl ether for ten days early in the pregnancy. The uterine contents of the pregnant mice were examined for defects in two days later. There are three possible outcomes for each fetus, viz, dead, malformation and normal; see Table 7.

The second illustration is the house fly dataset given in Zocchi and Atkinson [6]. This dataset were resulted from an experiment in which seven sets of 500 pupae each were exposed to one of several doses of radiation. The interested number of flies were recorded. The response data are trinomial with variables:

x is dose of radiation.

y_1 is number of flies that died before the opening of the pupae.

y_2 is number of flies that died before complete emerged.

y_3 is number of flies that complete emerged.

In this latter dataset, for reasons discussed below we use Box and Cox transformation for the independent variable x, $z = \dfrac{\left(\frac{x}{10}\right)^{\gamma} - 1}{\gamma}$, where γ is an additional (scale) parameter to be estimated for the dataset/model combination.

Table 7. Diethylene glycol dimethyl ether pregnant mice data and fitted values.

Concentration (mg/kg per day)	**Response**			**Total**
x	**Dead** y_1	**Malformation** y_2	**Normal** y_3	
0	15 (11.11)	1 (0.95)	281 (284.94)	297
62.5	17 (13.26)	0 (2.24)	225 (226.50)	242
125	22 (24.83)	7 (8.20)	283 (278.97)	312
250	38 (48.26)	59 (51.54)	202 (199.20)	299
500	144 (139.09)	132 (138.98)	9 (6.93)	285

The fitted values that shown in parentheses in Table 7 are obtained using the CRB model,

$$\log\left[\frac{\pi_1}{\pi_2 + \pi_3}\right] = \alpha_1 + \beta_1 x$$

$$\log\left[\frac{\pi_2}{\pi_3}\right] = \alpha_2 + \beta_2 x$$

We next fit each of the above models to this dataset; the results – both the maximum likelihood estimates and the asymptotic standard errors – are given in Table 8.

Table 8. Parameter estimates and standard errors (in parentheses) for the pregnant mice dataset using the AC, UPO, CRA, CRB and GOL models.

Model	Parameter Estimates (Asymptotic Standard Errors)						- Log-likelihood
	α_1	α_2	β_1	β_2	θ_1	θ_2	
AC	0.9834 (0.2708)	-4.9528 (0.2492)	-0.0021 (0.0006)	0.0140 (0.0008)	----	----	753.2203
UPO	-3.5988 (0.1571)	-3.5733 (0.1559)	0.0072 (0.0004)	0.0122 (0.0006)	----	----	743.5459
CRA	0.9046 (0.2600)	-3.6783 (0.1697)	-0.0019 (0.0006)	0.0127 (0.0007)	----	----	753.7615
CRB	-3.2479 (0.1577)	-5.7019 (0.3323)	0.0064 (0.0004)	0.0174 (0.0012)	----	----	730.3872
GOL	-1.8540 (0.8220)	-5.5382 (0.4117)	0.0037 (0.0016)	0.0169 (0.0013)	0.3004 (0.2208)	0.0470 (0.0550)	728.3356

By using the minimum negative log-likelihood, GOL model is chosen. But the negative log-likelihood is very close to the value in CRB model. Now consider the different between these two values of the negative log-likelihood, which is 730.3872 - 728.3356 = 2.0516. When we use model testing for nested model of this form, we see that the test statistic is χ_2^2 = 2(2.0516) = 4.1032, p-value = 0.1285. Thus, the different of these two values of negative log-likelihood is not significant at $\alpha = 0.05$, and so we choose CRB model for this dataset for reasons of parsimony. That is, the fitted model is:

$$\log\left[\frac{\pi_1}{\pi_2+\pi_3}\right] = -3.2479 + 0.0064\,x$$

$$\log\left[\frac{\pi_2}{\pi_3}\right] = -5.7019 + 0.0174\,x$$

Equivalently,

$$\hat{\pi}_1 = \frac{e^{-3.2478+0.0064\,x}}{1+e^{-3.2478+0.0064\,x}}$$

$$\hat{\pi}_2 = \frac{e^{-5.7019+0.0174x}}{(1+e^{-3.2478+0.0064x})(1+e^{-5.7019+0.0174x})}$$

$$\hat{\pi}_3 = \frac{1}{(1+e^{-3.2478+0.0064x})(1+e^{-5.7019+0.0174x})}$$

We now turn to the second illustrative dataset.

Table 9. House fly data and fitted values.

Dose of radiation	**Response categories**			**Total**
	Unopened Pupae	**Died before finishing emergence**	**Complete emergence**	
x	***y*₁**	***y*₂**	***y*₃**	
80	**62 (58.23)**	**5 (4.63)**	**433 (437.15)**	**500**
100	**94 (99.41)**	**24 (23.45)**	**382 (377.14)**	**500**
120	**179 (184.77)**	**60 (63.93)**	**261 (251.30)**	**500**
140	**335 (319.99)**	**80 (82.98)**	**85 (97.03)**	**500**
160	**432 (438.05)**	**46 (42.53)**	**22 (19.42)**	**500**
180	**487 (487.06)**	**11 (10.63)**	**2 (2.32)**	**500**
200	**498 (498.18)**	**2 (1.65)**	**0 (0.18)**	**500**

The fitted values that shown in parentheses in Table 9 are obtained using the UPOs model

$$\log\left[\frac{\pi_1}{\pi_2+\pi_3}\right] = \alpha_1 + \beta_1 z$$

$$\log\left[\frac{\pi_1+\pi_2}{\pi_3}\right] = \alpha_2 + \beta_2 z$$

Table 10. Parameters estimate and standard errors (in parentheses) for the house fly dataset using the ACs, UPOs, CRAs, CRBs and GOLs models.

Model	Parameter Estimates (Asymptotic Standard Errors)							- Log-likelihood
	α_1	α_2	β_1	β_2	γ	θ_1	θ_2	
ACs	0.6466 (0.0843)	-3.3646 (0.0650)	0.0005 (0.0005)	0.0013 (0.0012)	3.3651 (0.3596)	----	----	1805.9313
UPOs	-2.8560 (0.1669)	-3.0120 (0.1750)	0.0110 (0.0036)	0.0142 (0.0045)	2.5298 (0.1250)	----	----	1782.7272
CRAs	0.6462 (0.1268)	-2.2944 (0.1244)	0.0004 (0.0002)	0.0015 (0.0005)	3.4375 (0.1451)	----	----	1805.1136
CRBs	-3.4441 (0.3819)	-5.1437 (0.4997)	0.0371 (0.0227)	0.0441 (0.0269)	2.0562 (0.2376)	----	----	1789.9706
GOLs	-2.0271 (0.8162)	-3.2872 (0.5209)	0.0086 (0.0045)	0.0130 (0.0057)	2.5700 (0.1741)	0.5016 (0.3400)	0.7045 (0.3440)	1782.2119

By using the minimum negative log-likelihood, the GOLs model is chosen. But the negative log-likelihood is very close to the value in UPOs model. Now consider the different between these two values of the negative log-likelihood, which is 1782.7272 - 1782.2119 = 0.5153. When we use model testing for nested model of this form, we see that the test statistic is χ_2^2 = 2(0.5153) = 1.0306, p-value = 0.5973. So the difference of these two values of negative log-likelihood is not significant at $\alpha = 0.05$, and so we choose the UPOs model for this dataset. Thus, the fitted model is

$$\log\left[\frac{\pi_1}{\pi_2+\pi_3}\right] = -2.8560 + 0.0110\, z$$

$$\log\left[\frac{\pi_1+\pi_2}{\pi_3}\right] = -3.0120 + 0.0142\, z$$

with

$$\hat{\pi}_1 = \frac{e^{-2.8560+0.0110\, z}}{1+e^{-2.8560+0.0110\, z}}$$

$$\hat{\pi}_2 = \frac{e^{-3.0120+0.0142\, z} - e^{-2.8560+0.0110\, z}}{(1+e^{-2.8560+0.0110\, z})(1+e^{-3.0120+0.0142\, z})}$$

$$\hat{\pi}_3 = \frac{1}{1+e^{-3.0120+0.0142\, z}}$$

5. Conclusion and Discussion

In this study, we have developed the generalized ordinal logit (GOL) model which is generalized of multicategory logit models. This model is used for multicategory response data with k categories, where $k = 3,4$. It seems to see that GOL models can be fitted well with these two real datasets. Since the GOL model has two hyper-parameters and since the models are nested, whenever the change in the -2*log-likelihood values is not too great, we choose the smaller model for the sake of parsimony. We note that there is a potential use of our proposed model for real datasets with k categories response data where $k = 3,4$. Therefore the GOL models are useful for prediction of future values response in disciplines of medicine, social sciences, etc. Another important consequence of our new model is by providing researchers with a way to connect the various multicategory logit models with an eye to obtaining robust optimal designs (good for testing goodness of fit). We take up this latter topic in a forthcoming paper as this is the topic of our ongoing research.

Acknowledgement

The authors gratefully acknowledge Thammasat University for financial support of this research.

Appendices

a) APPENDIX A. Analysis of the Pregnant Mice Dataset using SAS/IML

```
proc iml;

start NLLUPO(th) global(xx,y1,y2,y3);
b01=th[1]; b02=th[2]; b11=th[3]; b12=th[4];
ex1=exp(b01+b11*xx); ex2=exp(b02+b12*xx);
ee1=ex1/(1+ex1); ee2=ex2/(1+ex2);
pi1=ee1; pi2=ee2-ee1; pi3=1/(1+ex2);
tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
return(-tomax);
finish NLLUPO;
th0upo={-5 -4 0.01 0.01 };
conupo={-9 -9 0.0001 0.0001,-3 1 1 1 };

start NLLAC(th) global(xx,y1,y2,y3);
```

```
b01=th[1]; b02=th[2]; b11=th[3]; b12=th[4];
ex1=exp(b01+b11*xx); ex2=exp(b02+b12*xx);
on=1+0*ex1; den=on+ex2+ex1#ex2;
pi1=ex1#ex2/den; pi2=ex2/den; pi3=1/den;
tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
return(-tomax);
finish NLLAC;
th0ac={ 1 -6 -0.002 0.02};
conac={0.1 -9 -0.1 0.00001,3 -3 -0.0000001 0.1 };

start NLLCRA(th) global(xx,y1,y2,y3);
b01=th[1]; b02=th[2]; b11=th[3]; b12=th[4];
ex1=exp(b01+b11*xx); ex2=exp(b02+b12*xx);
on=1+0*ex1; den=(on+ex1)#(on+ex2);
pi1=ex1#ex2/den; pi2=ex2/den; pi3=1/(on+ex2);
tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
return(-tomax);
finish NLLCRA;
th0cra={-5 -4 0.01 0.01 };
concra={-9 -9 -1 0.0001,3 1 1 1 };

start NLLCRB(th) global(xx,y1,y2,y3);
b01=th[1]; b02=th[2]; b11=th[3]; b12=th[4];
ex1=exp(b01+b11*xx); ex2=exp(b02+b12*xx);
on=1+0*ex1; den=(on+ex1)#(on+ex2);
pi1=ex1/(on+ex1); pi2=ex2/den; pi3=1/den;
tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
return(-tomax);
finish NLLCRB;
th0crb={-5 -4 0.01 0.01 };
concrb={-9 -9 -1 0.0001,3 1 1 1 };

start NLLALL(th) global(xx,y1,y2,y3);
th1=th[1]; th2=th[2]; b01=th[3]; b02=th[4]; b11=th[5]; b12=th[6];
ex1=exp(b01+b11*xx); ex2=exp(b02+b12*xx);
```

```
on=1+0*ex1; den=on+(th1+th2-th1*th2)*ex1+ex2+ex1#ex2;
pi1=(th1*ex1+ex1#ex2)/den; pi2=(ex2-th1*th2*ex1)/den;
pi3=(on+th2*ex1)/den;
tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
return(-tomax);
finish NLLALL;
th0all={0.95 0.05 -3.2 -5.7 0.0064 0.0174};
conall={0.000000001 0.000000001 -5 -7 -0.1 0.0001,
0.999999999 0.999999999 1.5 -2 0.1 0.1 };

xx={0.00000000000000001,62.5,125,250,500};
y1={15,17,22,38,144}; y2={1,0,7,59,132};
y3={281,225,283,202,9};opt={.,0};

call nlptr(rc,thupo,"NLLUPO",th0upo,opt,conupo); NLLUPOmin=NLLUPO(thupo);
print thupo NLLUPOmin;
call nlptr(rc,thac,"NLLAC",th0ac,opt,conac); NLLACmin=NLLAC(thac);
print thac NLLACmin;
call nlptr(rc,thcra,"NLLCRA",th0cra,opt,concra); NLLCRAmin=NLLCRA(thcra);
print thcra NLLCRAmin;
call nlptr(rc,thcrb,"NLLCRB",th0crb,opt,concrb); NLLCRBmin=NLLCRB(thcrb);
print thcrb NLLCRBmin;
call nlptr(rc,thall,"NLLALL",th0all,opt,conall); NLLALLmin=NLLALL(thall);
print thall NLLALLmin;
```

b) APPENDIX B. Analysis of the House Fly Dataset using SAS/IML

```
proc iml;

  start NLLUPOs(th) global(xx,y1,y2,y3);
     aa1=th[1]; aa2=th[2]; bb1=th[3]; bb2=th[4]; gam=th[5];
     zz=(1/gam)*(xx##gam-1);
     ex1=exp(aa1+bb1*zz); ex2=exp(aa2+bb2*zz);
     ee1=ex1/(1+ex1); ee2=ex2/(1+ex2);
     pi1=ee1; pi2=(ee2-ee1); pi3=1/(1+ex2);
     tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
     return(-tomax);
```

```
finish NLLUPOs;
th0upos={-5 -4    0.06    0.06     2};
conupos={-9 -9    0.0001  0.0001  -1,
         -2  1    1       1        4};

start NLLACs(th) global(xx,y1,y2,y3);
  aa1=th[1]; aa2=th[2]; bb1=th[3]; bb2=th[4]; gam=th[5];
  zz=(1/gam)*(xx##gam-1);
  ex1=exp(aa1+bb1*zz); ex2=exp(aa2+bb2*zz);
  on=1+0*ex1; den=on+ex2+ex1#ex2;
         eps=0.00001+0*on;
  pi1=ex1#ex2/den+eps; pi2=ex2/den; pi3=1/den;
  tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
  return(-tomax);
finish NLLACs;
th0acs={ -1  -6  0.0002  0.0005    1};
conacs={ -3 -19 -0.1     0.00001  -1,
          3  -3  0.1     0.1       4};

start NLLCRAs(th) global(xx,y1,y2,y3);
  aa1=th[1]; aa2=th[2]; bb1=th[3]; bb2=th[4]; gam=th[5];
  zz=(1/gam)*(xx##gam-1);
  ex1=exp(aa1+bb1*zz); ex2=exp(aa2+bb2*zz);
  on=1+0*ex1; den=(on+ex1)#(on+ex2);
  pi1=ex1#ex2/den; pi2=ex2/den; pi3=1/(on+ex2);
  tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
  return(-tomax);
finish NLLCRAs;
th0cras={-1 -6  0.0001  0.0007  1 };
concras={-9 -9  -1      0.0001 -1,
          3  1  1       1       4};

start NLLCRBs(th) global(xx,y1,y2,y3);
  aa1=th[1]; aa2=th[2]; bb1=th[3]; bb2=th[4]; gam=th[5];
  zz=(1/gam)*(xx##gam-1);
```

```
    ex1=exp(aa1+bb1*zz); ex2=exp(aa2+bb2*zz);
    on=1+0*ex1; den=(on+ex1)#(on+ex2);
    pi1=ex1/(on+ex1); pi2=ex2/den; pi3=1/den;
    tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
    return(-tomax);
  finish NLLCRBs;
  th0crbs={-5  -9  0.0001  0.0007  1};
  concrbs={-9 -19  -1      0.0001 -1,
           3   1  1       1       4};


  th0alla={0.8 0.2 -2 -3 0.001 0.002 3};
  start NLLALLs(th) global(xx,y1,y2,y3);
    th1=th[1]; th2=th[2]; aa1=th[3]; aa2=th[4]; bb1=th[5]; bb2=th[6]; gam=th[7];
    zz=(1/gam)*(xx##gam-1);
    ex1=exp(aa1+bb1*zz); ex2=exp(aa2+bb2*zz);
    on=1+0*ex1; den=on+(th1+th2-th1*th2)*ex1+ex2+ex1#ex2;
    pi1=(th1*ex1+ex1#ex2)/den; pi2=(ex2-th1*th2*ex1)/den; pi3=(on+th2*ex1)/den;
    tomax=t(y1)*log(pi1)+t(y2)*log(pi2)+t(y3)*log(pi3);
    return(-tomax);
  finish NLLALLs;
  conalls={0.001 0.001 -4   -4   0.0001  0.0001    0.9,
          0.999 0.999 -.1  -.1  0.9     0.9       4};

  xx={8,10,12,14,16,18,20};
  y1={62,94,179,335,432,487,498}; y2={5,24,60,80,46,11,2};
y3={433,382,261,85,22,2,0};
  opt={.,0};
  call nlptr(rc,thpos,"nllpos",th0pos,opt,conpos);      nllminpos=nllpos(thpos);     print thpos
nllminpos;
  call nlptr(rc,thupos,"nllupos",th0upos,opt,conupos);  nllminupos=nllupos(thupos);  print
thupos nllminupos;
  call nlptr(rc,thacs,"nllacs",th0acs,opt,conacs);      nllminacs=nllacs(thacs);     print thacs
nllminacs;
```

```
call nlptr(rc,thcras,"nllcras",th0cras,opt,concras); nllmincras=nllcras(thcras); print thcras nllmincras;
call nlptr(rc,thcrbs,"nllcrbs",th0crbs,opt,concrbs); nllmincrbs=nllcrbs(thcrbs); print thcrbs nllmincrbs;
call nlptr(rc,thalls,"nllalls",thallb,opt,conalls); nllminalls=nllalls(thalls); print thalls nllminalls;
```

References

[1] Agresti, A. Categorical Data Analysis: 2nd Edition, John Wiley & Sons, Inc, 2002.

[2] Schabenberger, O. and Pierce, F.J. Contemporary Statistical Models for the Plant and Soil Sciences, CRC Press LLC, 2002.

[3] Federov, V., Gagnon, R., Leonov, S. and Wu, Y. Optimal Design of Experiments in Pharmaceutical Applications, in Dmitrienko, A, Chuang-Stein, C. and D-Agostino, R., Pharmaceutical Statistics Using SAS: A Practical Guide, SAS Publishing, Cary, NC. , 2007: 151-195.

[4] Bates, D. and Watts, D. G. Nonlinear Regression Analysis and its Applications, John Wiley & Sons, Inc, 2007.

[5] Agresti, A. An Introduction to Categorical Data Analysis, John Wiley & Sons, Inc, 2007.

[6] Zocchi, S.S. and Atkinson, A.C. Optimum Experimental Design for Multinomial Logistic Models. Biometrics, 1999: 55: 437-444.