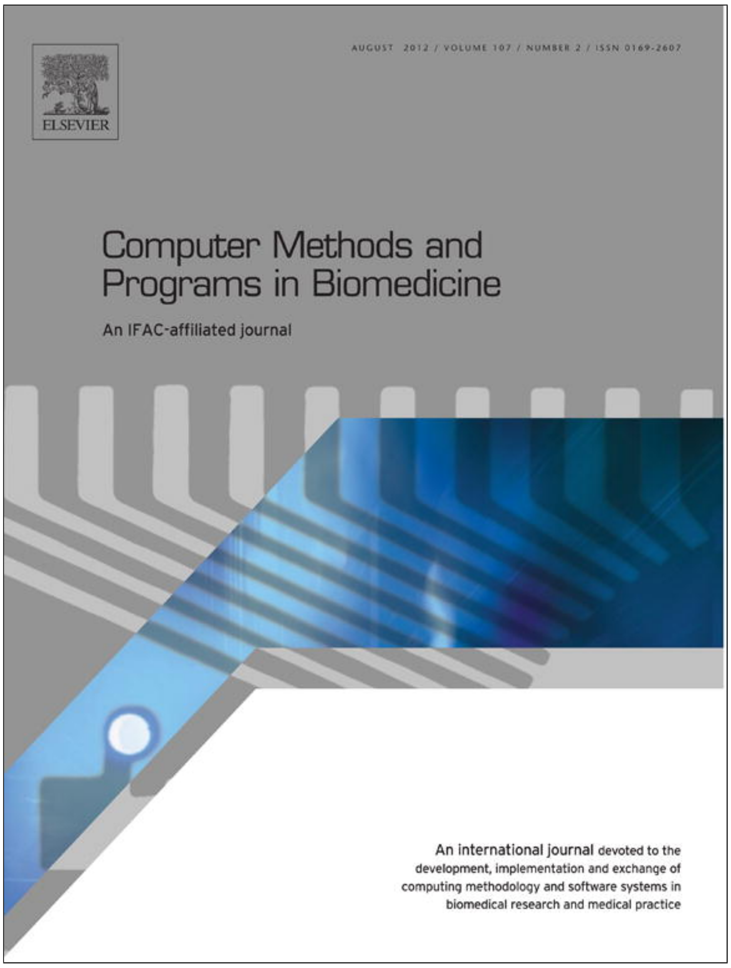
AUGUST 2012 / VOLUME 107 / NUMBER 2 / ISSN 0169-2607
ELSEVIER
Computer Methods and Programs in Biomedicine
An IFAC-affiliated journal
An international journal devoted to the development, implementation and exchange of computing methodology and software systems in biomedical research and medical practice

COMPUTER METHODS AND PROGRAMS IN BIOMEDICINE 107 (2012) 341–346

ELSEVIER

journal homepage: www.intl.elsevierhealth.com/journals/cmpb

A warning concerning the estimation of multinomial logistic models with correlated responses in SAS☆

*Mark de Rooij**, *Hailemichael M. Worku*

Leiden University, Psychological Institute, Methodology and Statistics Unit, Netherlands

ARTICLE INFO

Article history:
Received 4 January 2011
Received in revised form
18 January 2012
Accepted 23 January 2012

Keywords:
Multinomial logistic model
Random effects
marginal models
SAS
GEE

ABSTRACT

Kuss and McLerran [1] in a paper in this journal provide SAS code for the estimation of multinomial logistic models for correlated data. Their motivation derived from two papers that recommended to estimate such models using a Poisson likelihood, which is according to Kuss and McLerran "statistically correct but computationally inefficient". Kuss and McLerran propose several estimating methods. Some of these are based on the fact that the multinomial model is a multivariate binary model. Subsequently a procedure proposed by Wright [5] is exploited to fit the models. In this paper we will show that the new computation methods, based on the approach by Wright, are statistically incorrect because they do not take into account that for multinomial data a multivariate link function is needed. An alternative estimation strategy is proposed using the clustered bootstrap.

1. Introduction

Recently we were looking for SPSS or SAS code to fit the multinomial model to clustered data and we came across the paper of Kuss and McLerran [1] that provides SAS code to fit both random effects and marginal multinomial logistic models. At first we were quite enthusiastic about the code but later on we noted some strange results, which were also present in the analyses presented in Kuss and McLerran. After a detailed scrutiny we found out that some of the fitting procedures are incorrect. In this note we will first show the Kuss and McLerran (K&M) procedure and then show their mistake.

2. The Kuss and McLerran procedure

Let Y_{ij} denote the jth response ($j = 1, \ldots, n_i$) in cluster i ($i = 1, \ldots, n$) where this response is from one of R distinct categories ($r = 1, \ldots, R$). Further, $\mathbf{x}_{ij}$ denotes a column vector of p covariates for the jth observation in the ith cluster. To specify a multinomial model K&M reorganised the response vector as a $((R-1) \times 1)$-vector Y^*_{ij} of binary indicator variables Y^*_{ijr} such that $Y_{ij} = 2, \ldots, R$ results in $Y^*_{ijr} = 1$ in column r and 0 anywhere else. In the case of $Y_{ij} = 1$ (i.e. the reference category is chosen) $Y^*_{ijr} = 0$ in all $R-1$ columns. This reorganisation of the response vector can be interpreted as transforming the multinomial model into a multivariate binary model. Such a transformation is inspired by the work of [2–4].

Wright [5] proposed a general way for the estimation of multivariate models. Wright shows that multivariate models may be estimated by software for univariate models by reordering the multivariate responses in a vector and reorganizing the matrix with explanatory variables. Suppose we have two response variables Y1 and Y2 with observations $Y1_i$ and $Y2_i$ and a predictor variable X with observations X_i. The multivariate regression model:

☆ This research was conducted while the first author was sponsored by the Netherlands Organisation for Scientific Research (NWO), Innovational Grant, no. 452-06-002.

* *Corresponding author*. Tel.: +31 71 5274102.
E-mail address: rooijm@fsw.leidenuniv.nl (M. de Rooij).

0169-2607/$ – see front matter © 2012 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2012.01.008

$$\begin{bmatrix} Y1_1 & Y2_1 \\ Y1_2 & Y2_2 \\ Y1_3 & Y2_3 \end{bmatrix} = \begin{bmatrix} 1 & X_1 \\ 1 & X_2 \\ 1 & X_3 \end{bmatrix} \begin{bmatrix} \beta_{01} & \beta_{02} \\ \beta_{11} & \beta_{12} \end{bmatrix} + \begin{bmatrix} \epsilon 1_1 & \epsilon 2_1 \\ \epsilon 1_2 & \epsilon 2_2 \\ \epsilon 1_3 & \epsilon 2_3 \end{bmatrix},$$

may be estimated by

$$\begin{bmatrix} Y1_1 \\ Y2_1 \\ Y1_2 \\ Y2_2 \\ Y1_3 \\ Y2_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & X_1 & 0 \\ 0 & 1 & 0 & X_1 \\ 1 & 0 & X_2 & 0 \\ 0 & 1 & 0 & X_2 \\ 1 & 0 & X_3 & 0 \\ 0 & 1 & 0 & X_3 \end{bmatrix} \begin{bmatrix} \beta_{01} \\ \beta_{02} \\ \beta_{11} \\ \beta_{12} \end{bmatrix} + \begin{bmatrix} \epsilon 1_1 \\ \epsilon 2_1 \\ \epsilon 1_2 \\ \epsilon 2_2 \\ \epsilon 1_3 \\ \epsilon 2_3 \end{bmatrix}.$$

This trick can be used in case when the data are truly bivariate (or more generally multivariate). K&M use this idea to estimate multinomial models. The model equation then becomes:

$$\log\left(\frac{\pi^*_{ir}}{1-\pi^*_{ir}}\right) = \theta^*_r + \mathbf{x}_{ij}^{\mathrm{T}}\beta^*_r, \quad r = 2,\ldots,R, \tag{1}$$

where π^*_{ir} denotes the expectation of all elements of Y^*_i belonging to response category r, and $Y^*_i = \left(Y^{*\mathrm{T}}_{i1},\ldots,Y^{*\mathrm{T}}_{in_i}\right)^{\mathrm{T}}$.

In their example K&M denote their reorganized matrix with explanatory variables `tbidouble`. K&M provide code using the `GENMOD` and `GLIMMIX` to fit both marginal model as well as random effect models using the idea of Wright [5].

To demonstrate the models and the respective SAS codes K&M use a data set on physicians recommendations and preferences in traumatic brain injury (TBI) rehabilitation. In this study, 36 physicians were asked to decide on the optimal rehabilitation setting (in-patient, day-clinic, out-patient) for each of 10 typical TBI disease histories. Four covariates, all binary, were included in the model, two of them referring to physicians characteristics (1: Is the physician a neurologist [NEURO]; and 2: Is the physician a specialist [SPECIAL]) and two describing the disease history (3: Is the time since the event longer than 3 months [TIME]; and 4: Is the patient severely handicapped after the TBI [SEVERITY]). As the reference category of the response the stationary in-patient setting was chosen, and compare day-clinic (DC) and out-patient (OP) to this.

The results presented in K&M are reproduced in Table 1. In the first column the estimates of the standard multinomial logistic regression model (ignoring the clustering) are presented, in the second of the multinomial logistic with random effects using a `NLMIXED` procedure, in the third, fourth, and fifth column K&M give estimates obtained by the above explained multivariate procedure using `GENMOD` (4th column) and `GLIMMIX` (3rd and 5th columns).

3. What went wrong?

The multinomial logistic model cannot be fitted with standard generalized linear model software, nor generalized estimating equation software. Although the multinomial logistic model is a kind of multivariate binary logistic model (as shown by [2–4]) it is different in the sense that it needs a *multivariate link function*. Therefore the procedure of Wright [5] cannot be applied in a straightforward manner, since in this procedure the link function does not change compared to the univariate case (binary logistic regression). The effect can be seen in formula (1) which is not the equation of the multinomial logistic regression model; the correct formula should be:

$$\log\left(\frac{\pi^*_{ir}}{\pi^*_{i1}}\right) = \theta^*_r + \mathbf{x}_{ij}^{\mathrm{T}}\beta^*_r, \quad r = 2,\ldots,R, \tag{2}$$

where thus $(1-\pi^*_{ir})$ is changed into π^*_{i1}. In a multinomial logistic model π^*_{i1} is not equal to $(1-\pi^*_{ir})$ but to $(1-\sum_{r\neq 1}\pi^*_{ir})$. Moreover, Eq. (1) leads to estimated 'probabilities':

$$\pi^*_{ir} = \frac{\exp(\theta^*_r + \mathbf{x}_{ij}^{\mathrm{T}}\beta^*_r)}{1+\exp(\theta^*_r + \mathbf{x}_{ij}^{\mathrm{T}}\beta^*_r)}, \quad r = 2,\ldots,R,$$

which do not necessarily sum to one. Under the multinomial logistic model this should be:

$$\pi^*_{ir} = \frac{\exp(\theta^*_r + \mathbf{x}_{ij}^{\mathrm{T}}\beta^*_r)}{1+\sum_{s=2}^{R}\exp(\theta^*_s + \mathbf{x}_{ij}^{\mathrm{T}}\beta^*_s)},$$

with $\theta^*_1 = 0$ and $\beta^*_1 = 0$.

The results using the Wright [5] approach are presented in columns 3–5 of Table 1. Looking at the table again in more detail we see some strange results. For example, in the upper part of the table (DC) the negative regression weights for [NEURO] in the PQL, GEE, and MQL methods. The numbers in the column under 'Multinomial' give consistent estimates for the marginal model (only the standard errors are wrong), and it can be expected that the estimates of the marginal model are very close to these estimates. In Table 1 it can be seen that the estimates in columns 3–5 are quite different from the estimates in column 1.

Furthermore, for the binary logistic case, [6] shows that the estimates of the random effects and the marginal model are related by the following equation:

$$\beta_M \approx \beta_{RE} \times (1-\rho),$$

where β_M are regression weights under the marginal model (columns 4 and 5), β_{RE} regression weights under a similar random effects model (column 2), and ρ represents the intra-cluster correlation. Given that the intra-cluster correlation is in general positive (subjects within clusters tend to answer in a similar manner), the estimates of the marginal model are generally smaller than those of the random effects model. Assuming that a similar relationship is present between estimates of marginal and random effects multinomial models, we see that in Table 1 this relationship is violated. Furthermore, K&M also noted that GQ and PQL differ sometimes substantially, which was unexpected. This can now be understood, because the two procedures fit different models. The GQ approach fits the multinomial logistic regression model, the PQL as proposed by K&M fits Eq. (1).

Thus, only the results in columns 1 and 2 of Table 1 are correct; the estimates in columns 3–5 are incorrect. The results presented in columns 3–5 are obtained with the `GENMOD` and `GLIMMIX` procedures using the multivariate model but without the multivariate link function. These procedures do not estimate the multinomial logistic model.

Table 1 – Results reported in Kuss and McLerran (estimates and respective standard errors in parenthesis) from the standard fixed effects multinomial (PROC LOGISTIC), Gaussian Quadrature (GQ, PROC NLMIXED), PQL (PROC GLIMMIX), GEE (PROC GENMOD) and MQL (PROC GLIMMIX) estimation for the TBI data set.

	Multinomial	GQ	PQL	GEE	MQL
Fixed effects – DC					
$\hat{\theta}$	−0.879 (0.430)	−1.103 (0.606)	−1.468 (0.470)	−1.380 (0.441)	−1.391 (0.446)
$\hat{\beta}_{\text{Neuro}}$	0.088 (0.395)	0.254 (0.731)	−0.220 (0.515)	−0.225 (0.445)	−0.215 (0.471)
$\hat{\beta}_{\text{Special}}$	−0.446 (0.403)	−0.455 (0.714)	−0.572 (0.511)	−0.581 (0.423)	−0.574 (0.469)
$\hat{\beta}_{\text{Time}}$	1.723 (0.324)	2.365 (0.405)	1.096 (0.312)	1.044 (0.265)	1.053 (0.305)
$\hat{\beta}_{\text{Severity}}$	−1.204 (0.414)	−1.602 (0.475)	−0.368 (0.360)	−0.361 (0.352)	−0.355 (0.351)
Fixed effects – OP					
$\hat{\theta}$	−2.429 (0.566)	−3.009 (0.820)	−3.267 (0.643)	−3.034 (0.628)	−3.046 (0.595)
$\hat{\beta}_{\text{Neuro}}$	1.073 (0.481)	1.329 (0.916)	1.085 (0.658)	1.054 (0.416)	1.050 (0.591)
$\hat{\beta}_{\text{Special}}$	0.296 (0.426)	0.281 (0.855)	0.447 (0.607)	0.466 (0.458)	0.435 (0.540)
$\hat{\beta}_{\text{Time}}$	3.149 (0.456)	4.138 (0.588)	2.929 (0.469)	2.654 (0.483)	2.665 (0.441)
$\hat{\beta}_{\text{Severity}}$	−2.022 (0.441)	−2.591 (0.513)	−1.589 (0.395)	−1.485 (0.371)	−1.470 (0.381)
Random effects					
$\hat{\sigma}^2_{\text{DC}}$	–	1.650 (0.804)	0.590 (0.336)	–	0.436 (0.273)
$\hat{\sigma}^2_{\text{OP}}$	–	2.611 (1.200)	1.032 (0.521)	–	0.715 (0.369)
$\hat{\sigma}_{\text{DC-OP}}$	–	1.888 (0.856)	−0.119 (0.324)	–	−0.131 (0.244)
−2LL	492.9	461.2	–	–	–
npar	10	13	–	–	–

4. How to proceed?

How can multinomial logistic models for correlated responses then be fitted? If one is interested in a random effects model the NLMIXED procedure from SAS can be used, as outlined in Kuss and McLerran [1]. This procedure uses Gauss–Hermite quadrature methodology to integrate out the random effects of the likelihood function. This approach is for generalized linear mixed models preferred over penalized and marginal quasi likelihood methods since the latter can be severely biased when applied to binary response variables (Molenberghs and Verbeke, 2005, p. 272 [7]). It cannot be expected that this situation alters for multinomial data.

If one is interested in a marginal model the procedures as outlined in Kuss and McLerran cannot be used. If one is, however, only interested in the regression weights and not in their standard errors one can use the standard multinomial model (column 1). To obtain correct standard errors the bootstrap can be used as shown by Sherman and Le Cessie [8]. It is important to notice that such a resampling procedure should be applied on the level of the cluster such that the correlation between the responses is present in every sample.

To illustrate such a method consider data from the McKinney Homeless Research Project (MHRP) in San Diego as described in chapters 10 and 11 in the book by Hedeker and Gibbons [9]. The aim of this project was to evaluate the effectiveness of using an incentive as a means of providing independent housing to homeless people with severe mental illness. Housing certificates were provided from the Department of Housing and Urban Development to local authorities in San Diego. These housing certificates were designed to make it possible for low income individuals to choose and obtain independent housing in the community. A sample of 361 individuals took part in this longitudinal study and were randomly assigned to the experimental or control condition. Eligibility for the project was restricted to individuals diagnosed with a severe and persistent mental illness who were either homeless or at high risk of becoming homeless at the start of the study. Individuals' housing status was assessed using three categories (living on the street/living in a community center/living independently) at baseline and at 6-, 12-, and 24-month follow up.

In Table 2 proportions for each of the three response categories at the four time points are given for both groups. There it can be seen that some proportions first go up and later down and other proportions first go down and then go up again.

Table 2 – Housing condition across time by group: proportions and sample size.

Group	Status	Time point			
		Baseline	6	12	24
Control	Street	.555	.186	.089	.124
	Community	.339	.578	.582	.455
	Independent	.106	.236	.329	.421
	N	180	161	146	145
Experimental	Street	.442	.093	.121	.120
	Community	.414	.280	.146	.228
	Independent	.144	.627	.732	.652
	N	181	161	157	158

Table 3 – Parameter estimates of the mixed effects multinomial model for the MHRP data. In the C/S column the parameter estimates for the community versus street contrast are given, whereas in the I/S column estimates for the contrast independent versus street are given. The correlation between the two random intercepts equals 0.692.

Effect	C/S	I/S
α	2.436 (0.301)	1.293 (0.363)
β_T	0.160 (0.019)	0.236 (0.024)
β_{T^2}	−0.014 (0.002)	−0.015 (0.003)
β_G	−0.914 (0.420)	2.072 (0.485)
β_{GT}	−0.086 (0.025)	0.022 (0.031)
β_{GT^2}	0.008 (0.003)	−0.008 (0.004)
σ_u	1.519 (0.199)	2.299 (0.225)

Moreover, the proportions for the two groups are rather different. This information lead us to define a model with different quadratic time trends for the two groups and random intercepts. The linear predictor for this model equals:

$$\eta_{itc} = \alpha_c + \beta_{Gc}G_i + \beta_{Tc}T_{it} + \beta_{T^2c}T_{it}^2 + \beta_{GTc}G_iT_{it} + \beta_{GT^2c}G_iT_{it}^2 \; (+u_{ic}),$$

where G_i is an indicator for group membership ($G_i = 1$ for incentive) of participant i, and T_{it} represents the time variable defined as T_{it} = Month − 10. For the marginal model the last term between parenthesis is deleted, for mixed effects models this term is included. For the mixed effects model it is assumed that u_{i1} and u_{i2} come from a bivariate normal distribution.

Results of the mixed effects model are shown in Table 3. These are obtained by using the NLMIXED procedure in SAS as outlined in Kuss and McLerran.

Results for the marginal model are shown in Table 4 where the 95% confidence interval is estimated using a clustered bootstrap procedure. SAS code for doing this is shown in Appendix A. Comparing the results in Table 3 and 4 it can be seen that the estimates of the random effect and the marginal model are all in the same direction and that the estimates in Table 4 of the marginal model are overall somewhat smaller (the only exception is the estimate for incentive in the C/S contrast). Furthermore, the same effects reach statistical significance in both analysis. That is, all effects are significant except for the incentive by time interactions in the I/S contrast.

Table 4 – Parameter estimates of the marginal multinomial model for the MHRP data. In the C/S column the parameter estimates for the community versus street contrast are given, whereas in the I/S column estimates for the contrast independent versus street are given. Between parenthesis is a 95% confidence percentile interval based on the clustered bootstrap.

Effect	C/S	I/S
α	1.734	1.054
	(1.340, 2.245)	(0.572, 1.590)
β_T	0.116	0.164
	(0.089, 0.146)	(0.132, 0.202)
β_{T^2}	−0.011	−0.011
	(−0.014, −0.0076)	(−0.0149, −0.0074)
β_G	−1.001	1.098
	(−1.704, −0.267)	(0.374, 1.852)
β_{GT}	−0.078	0.0008
	(−0.122, −0.033)	(−0.050, 0.052)
β_{GT^2}	0.0070	−0.0037
	(0.0016, 0.0120)	(−0.0094, 0.0017)

5. Conclusion

Kuss and McLerran [1] proposed several ways to estimate multinomial logistic regression models for correlated data. The procedures using GENMOD and GLIMMIX are wrong. They do not estimate multinomial logistic models, but something else.

For the random effects model the multinomial regression model can be estimated exploiting the relation between the multinomial and the Poisson distribution (Chen and Kuo [10]). As noted by Kuss and McLerran [1] this is an inefficient procedure. The inefficiency only amounts to computing time. Compared to the time needed to collect the data this is usually only a very small fraction. Therefore, the random effects model can best be estimated using the NLMIXED procedure. For marginal models we proposed to use the clustered bootstrap.

An example of these two methods was presented using SAS. SAS code for the analysis of multinomial data with the clustered bootstrap was provided in Appendix A.

Conflicts of interest

There are no conflicts of interest.

Appendix A. SAS code for Clustered Bootstrap.

In this appendix SAS code is shown for the clustered bootstrap as performed in Section 4. This code is based on a code found in the SAS- L-archives `listserv.uga.edu`.[1] The total code consists of a number of steps, each to be executed. First a standard multinomial logistic regression is performed which is shown in lines 001–004. Note that `time` represents the time variable, while `time2` represents time squared; `treatmentTime` represents the variable obtained by multiplication of `treatment` and `time`, similarly for `treatmentTime2`. To perform a bootstrap we first need to define the number of clusters (in our case subjects) which is 362 (line 005). On lines 006 and 007 an index variable is created, which is needed to draw bootstrap samples. From lines 008–020 the bootstrap samples are drawn. In this case we draw 1000 bootstrap samples as defined in line 009. Subsequently we fit a multinomial logistic regression to each of these bootstrap samples (lines 021–026). This is simply performed using the `by` statement. Note that sometimes an error message is printed, which is due to the combination of the `by` statement and the ods output statement. This error message can be ignored.

To obtain confidence intervals for our regression parameters a number of data management steps need to be undertaken. These are shown in lines 027–054. Finally, using the proc univariate (lines 055–059) we can obtain the 2.5% and 97.5% percentiles of the 1000 estimates of each regression

[1] www.listserv.uga.edu/cgi-bin/wa?A2=ind0801d&L=sas-l&P=44273, last accessed december 15, 2011.

parameter. Of course, if one would like to assume a normal distribution for the estimates the standard deviation can be computed and using that estimate a confidence interval can be defined. The boundaries of the confidence intervals are printed with the statement in line 061.

```
001    proc logistic data = housing;
002    title 'MBCL model on Original Housing Data';
003    model outcome = time time2 treatment treatmentTime treatmentTime2/ link= glogit aggregate
       scale = none;
004    run;

005    %let Nclusters=362;

006    data housing (index=(Ppn) sortedby=Ppn);
007    set housing; run;

008    data boothousing / view=boothousing;
009    do sample = 1 to 1000;
010    do _i = 1 to Nclusters;
011    PPn = ceil(ranuni(123245)*&nclusters);
012    do until (_iorc_ ne 0);
013    set housing key = Ppn;
014    if _iorc_ eq 0 then output;
015    end;
016    end;
017    end;
018    _error_ = 0;
019    stop;
020    run;

021    proc logistic data = boothousing; class sample
022    title 'MBCL models fit on Bootstrap samples';
023    model outcome = time time2 treatment treatmentTime treatmentTime2/ link = glogit aggregate
       scale = none;
024    by sample;
025    ods output ParamaterEstimates = MBCLparms;
026    run;

027    * Data Management 1;
028    data MBCLparms;
029    set MBCLparms;
030    Resp=0;
031    if Response=1 then Resp=1;
032    if Response=2 then Resp=2;
033    drop Response;
034    run;

035    * Data Management 2;
036    data MBCLparms;
037    set MBCLparms;
038    if Resp=1 then Response="R1";
039    if Resp=2 then Response="R2";
040    drop Resp;
041    run;

042*   Data Management 3;
043    PROC SQL;
044    title 'Sub-Summarized data';
045    CREATE TABLE MBCLparms2 AS
046    SELECT sample, Response||Variable as EstResp, Estimate
047    FROM MBCLparms;
048    QUIT;

049    * Data Management 4;
```

```
PROC TRANSPOSE data=MBCLparms2 out=MBCLparms3(drop=_name_);
by sample;
var Estimate;
id EstResp;
run;

* gives std and percentiles;
proc univariate data=MBCLparms3;
var R1Intercept R2Intercept R1time R2time R1time2 R2time2 R1treatment R2treatment
R1treatmentTime R2treatmentTime R1treatmentTime2 R2treatmentTime2;
output out=percentiles pctlpts=2.5 to 97.5 by 95 pctlpre=R1Intercept_ R2Intercept_ R1time_
R2time_ R1time2_ R2time2_ R1treatment_ R2treatment_ R1treatmentTime_ R2treatmentTime_
R1treatmentTime2_ R2treatmentTime2_;
run;

* Print out the final results per sample;
proc print data=percentiles; run;
```

REFERENCES

[1] O. Kuss, D. McLerran, A note on the estimation of the multinomial logistic model with correlated responses in SAS, Computer Methods and Programs in Biomedicine 87 (2007) 262–269.

[2] L. Fahrmeier, G. Tutz, Multivariate Statistical Modelling based on Generalized Linear Models, 2nd ed., Springer, New York, 2001.

[3] J. Hartzel, A. Agresti, B. Caffo, Multinomial logit random effects models, Statistical Modeling 1 (2001) 81–102.

[4] S.R. Lipsitz, K. Kim, L. Zhao, Analysis of repeated categorical data using generalized estimating equations, Statistics in Medicine 13 (1994) 1149–1163.

[5] S.P. Wright, Multivariate analysis using the MIXED procedure (Paper 229–23), in: Proceedings of the 23rd Annual SAS Users Group (SUGI) International Conference, 1998.

[6] J.M. Neuhaus, Estimation efficiency and tests of covariate effects with clustered binary data, Biometrics 49 (1993) 989–996.

[7] G. Molenberghs, G. Verbeke, Models for Discrete Longitudinal Data, Springer, New York, 2005.

[8] M. Sherman, S. Le Cessie, A comparison between bootstrap methods and generalized estimating equations for correlated outcomes in generalized linear models, Communications in Statistics—Simulation and Computation 26 (1997) 901–925.

[9] D. Hedeker, R.D. Gibbons, Longitudinal Data Analysis, John Wiley and Sons, Inc., Hoboken, New Jersey, 2006.

[10] Z. Chen, L. Kuo, A note on the estimation of the multinomial logit model with random effects, American Statistician 55 (2000) 89–95.