Special restrictions in multinomial logistic regression

John Hendrickx
Management Studies Group
Wageningen UR
Hollandseweg 1
6706 KN Wageningen
The Netherlands

email: John_Hendrickx@yahoo.com

Introduction

This paper describes two Stata programs, `mclgen` and `mclest`, for imposing special restrictions on multinomial logistic models. *MCL* stands for Multinomial Conditional Logit, a term coined by Breen (1994). An MCL model uses a conditional logit program to estimate a multinomial logistic model. This produces the same log likelihood, estimates and standard errors, but allows greater flexibility in imposing constraints. The MCL approach makes it possible to impose different restrictions on the response variable for different independent variables. For example, linear logits could be imposed for certain independent variables and an unordered response for others. One specific application is to include models for the analysis of square tables, e.g. quasi-independence, uniform association, symmetric association, into a multinomial logistic model (Logan 1983, Breen 1994).

`Mclest` can also estimate two types of models with both linear and multiplicative terms. The Stereotyped Ordered Regression model (SOR) estimates a metric for the dependent variable and a single parameter for each independent variable (Anderson 1984, DiPrete 1990). It is more flexible than `ologit` because it does not assume ordered categories, although it does assume that the response categories can be scaled on a single dimension. This makes it useful for “semi-ordered” variables such as occupation, where the rank of categories such as farmers is not altogether clear.

A second special model that can be estimated by `mclest` is the Row and Columns model 2 (Goodman 1979). This model, originally developed for loglinear analysis, estimates a metric for a categorical independent variable as well as the response variable. The effect of the independent variable can therefore be expressed through a single parameter. The SOR and RC2 models are estimated by iteratively running MCL models, taking first one element of the multiplicative terms as given, then the other.

Multinomial Conditional Logit Models

Multinomial logistic models and conditional logit models are very similar. Any model that can be estimated by `mlogit` can also be estimated by `clogit`, but this involves extra steps that are unnecessary for typical multinomial models. In order to estimate the model with `clogit`, the data must first be transformed into a *person/choice* file, the format for McFadden’s choice model. In a person/choice file, each respondent has a separate record for each category of the response variable. A stratifying variable indexes respondents, the response variable indexes response options, and a dichotomous variable indicates which response option is the respondent's actual choice. This dichotomous variable is entered as the dependent variable in `clogit` and the stratifying variable is specified in the `group` option. The response variable, which in a standard multinomial program would be the dependent variable, is now entered as an independent variable. Its main effects, using one of the categories as reference, correspond with the intercept term of a multinomial model. Interactions of the response variable with explanatory variables correspond with the effects of these variables.

The following example shows how the data can be transformed into a person/choice file and how an MNL model estimated using `clogit` (cf. Stata Reference Manual, version 3). The data here are taken from the 1972-78 GSS data used by Logan (1983: 332-333) and contain 838 cases. The response variable is `occ` (occupational class) with 5 categories: 1 "farm occupations", 2 "operatives, service, and laborers", 3 "craftsmen and kindred workers", 4 "sales and clerical", 5 "professional, technical, and managerial". There are two explanatory variables: `educ` (education in years) and `black` (race; 1=black, 0=non-black).

```
* Logistic Regression (Logan 1983: 333)
use logan
gen strata=_n

expand 5
sort strata
gen respfact=mod(_n-1,5)+1
gen didep=(occ==respfact)
quietly replace occ=respfact

xi: clogit didep i.occ i.occ|black i.occ|educ, strata(strata)
```

The first step in creating the person/choice file is to define the stratifying variable `strata` using the current case numbers. Next, `expand` is used to create a copy of each record for each of the response options. The data is then sorted so that each respondent's records are grouped together. The variable `respfact` is constructed with values 1 to 5 within each stratum in order to index response options. The variable `didep` is then created to indicate which record corresponds with the respondent's actual choice. Once this has been done, the response variable `occ` is no longer needed and its contents can be replaced by those of `respfact`. Of course, `respfact` could be used in the following instead, but this procedure using has the advantage that variable and value labels assigned to `occ` will be used in the output.

Once the person/choice file has been created, the multinomial logistic model can be estimated in `clogit`. `Didep` is specified as the dependent variable and `strata` is entered in the `strata` option. The main effects of `occ` using the first category as reference correspond with the intercept term and interactions of `occ` with `educ` and `black` correspond with the effects of these two variables. `Xi` will also include main effects of `educ` and `black` but these will be dropped by `clogit` due to the fact that they are constant within strata. Alternatively, `desmat` (Hendrickx 1999) can be used here to generate the design matrix. `Desmat` provides greater flexibility in specifying interactions and contrasts and the companion program `desrep` can summarize the output using informative labels.

Using mclgen and mclest

The programs `mclgen` and `mclest` automate the above procedure. `Mclgen` transforms the data into a person/choice file and `mclest` enters the dichotomous dependent variable and stratifying variable, then estimates the model. The necessary steps are now reduced to:

```
mclgen occ
xi: mclest i.occ i.occ|black i.occ|educ
```

This provides the following output:

```
. mclgen occ
(3352 observations created)

Your response factor is occ with 5 categories.
Its main effects form the intercept of a multinomial logistic model,
interactions with independent variables form their effects.

.
. xi: mclest i.occ i.occ|black i.occ|educ
i.occ                Iocc_1-5     (naturally coded; Iocc_1 omitted)
i.occ|black          IoXbla_#     (coded as above)
i.occ|educ           IoXedu_#     (coded as above)

Note: educ omitted due to no within-group variance.
Note: black omitted due to no within-group variance.
Iteration 0:   log likelihood = -1223.0058
Iteration 1:   log likelihood = -1025.7296
Iteration 2:   log likelihood = -1009.3479
Iteration 3:   log likelihood = -1007.2919
Iteration 4:   log likelihood = -1007.1621
Iteration 5:   log likelihood = -1007.1614
Iteration 6:   log likelihood = -1007.1614

Conditional (fixed-effects) logistic regression   Number of obs   =       4190
                                                  LR chi2(12)     =     683.10
                                                  Prob > chi2     =     0.0000
Log likelihood = -1007.1614                       Pseudo R2       =     0.2532

------------------------------------------------------------------------------
 __didep |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  Iocc_2 |   2.913547   1.373878      2.121   0.034       .2207963    5.606298
  Iocc_3 |   1.843265   1.381555      1.334   0.182      -.8645331    4.551063
  Iocc_4 |  -3.138894    1.47574     -2.127   0.033      -6.031291   -.2464979
  Iocc_5 |  -6.131355   1.441328     -4.254   0.000      -8.956306   -3.306405
IoXbla_2 |   1.305156   1.043259      1.251   0.211      -.7395935    3.349906
IoXbla_3 |    .628363   1.055375      0.595   0.552      -1.440134     2.69686
IoXbla_4 |   .3326202   1.103486      0.301   0.763      -1.830173    2.495413
IoXbla_5 |  -.2258867   1.093162     -0.207   0.836      -2.368446    1.916672
IoXedu_2 |  -.0505162   .1108293     -0.456   0.649      -.2677376    .1667052
IoXedu_3 |   .0386727   .1111411      0.348   0.728      -.1791599    .2565054
IoXedu_4 |   .3692091   .1163028      3.175   0.002       .1412598    .5971583
IoXedu_5 |   .6439505   .1135536      5.671   0.000       .4213896    .8665115
------------------------------------------------------------------------------
```

The parameters, standard errors, and log likelihood are the same as those of a model estimated using “`mlogit occ educ black, base(1)`”. The number of observations reported is 5 times the sample size, since the data have been transformed into a person/choice file. Note that the `LR chi2` value is the chi-square improvement relative to a model where all response options are equally likely, not an intercept-only multinomial model. Likewise, the `Pseudo R2` value uses an equiprobability model as its baseline, not an intercept model.

For a standard multinomial model, the MCL approach is only cumbersome. The advantage of using it lies in the ability to impose different restrictions on the response variable for different independent variables, something that cannot be easily done using `mlogit`. One application of this is to specify models for square tables, such as quasi-independence or uniform association, in a multinomial model with continuous covariates (Logan 1983, Breen 1994). These models were developed as loglinear models with special restrictions on the interaction between the row and column variable. They can be recast as multinomial logistic models where the restrictions on the response (column) variable depend on the category of the row variable.

Specifying a square table design is an MCL model follows the same procedure as for a loglinear model. One creates interactions between dummy variables for the response variable and dummy variables for the categorical independent using non-standard restrictions. This can be illustrated in the following example:

```
use logan
mclgen occ
gen d1=(focc==1)*(occ==1)
gen d2=(focc==2)*(occ==2)
gen d3=(focc==3)*(occ==3)
gen d4=(focc==4)*(occ==4)
gen d5=(focc==5)*(occ==5)
gen u=focc*occ
xi: mclest i.occ d* u i.occ|black i.occ|educ
```

This model specifies a quasi-uniform association model between father's occupation `focc` and the response variable `occ`, with `educ` and `black` as covariates. The dummies `d1-d5` measure the likelihood that father and son have the same occupation. They can be seen as interactions of a dummy for `focc==`*j* and a single dummy for the response variable corresponding with the logit `focc==`*j* versus `focc~=`*j*. The variable `u` for uniform association can be seen as an interaction of `occ` using a linear logits restriction and treating `focc` as a continuous rather than a categorical variable. The linear logits restriction means that for this effect, a unit's change in `focc` will result in a constant change in the logit between any two adjacent categories of `occ`.

In general, any type of restriction can be applied to the response variable and the restriction type can be varied at will for each independent variable. By applying the difference contrast to the response variable (cf. Hendrickx 1999), an adjacent logits model could be estimated. Another application could be to apply linear logits for some independent variables, and standard logits for others.

Stereotyped Ordered Regression

`Mclest` can also estimate two special designs incorporating both linear and multiplicative effects. One is the *Stereotyped Ordered Regression* model (Anderson 1984, DiPrete 1990). The SOR model is an alternative to the proportional odds model used in `ologit`. The SOR model estimates a scaling metric for the response factor based on the effects of independent variables. The model has a standard multinomial intercept with J-1 parameters for a response variable with J categories. It estimates J-2 independent scale values ϕ_j for the response factor and a single scaled beta parameter for each independent variable. This means that the SOR model is less parsimonious than the proportional odds model, since it has an extra J-2 parameters for the scaling metric. On the other hand, the SOR model does not assume that the response categories are ordered, although it does assume that they can be ordered. This makes it particularly useful when the rank of one or more categories is not altogether clear.

The SOR model can be specified as:

$$\log\left(\frac{P(Y=q)}{P(Y=r)}\right)=\text{logit}\left(\frac{q}{r}\right)=\alpha_q-\alpha_r+\left(\phi_q-\phi_r\right)\left(\beta_1 X_1+\beta_2 X_2+\ldots+\beta_K X_K\right)$$

Where Y is the response variable with categories j = 1 to J, q and r are categories of Y, α_j represents the intercept parameters with suitable restrictions, ϕ_j represents the scaling metric with suitable restrictions, X_k represents independent variables with k=1 to K, and β_k represents parameters of the independent variables. Two restrictions must be placed on the scaling metric ϕ_j in order to identify the model. `Mclest` sets the first value to 0 and the last value to 1 while estimating the model. For the final estimates, the scaling metric is also normalized, with a mean of 0 and a sum of squares of 1.

Compare the SOR model to a standard multinomial model:

$$\text{logit}\left(\frac{q}{r}\right)=\alpha_q-\alpha_r+\left(\beta_{q1}-\beta_{r1}\right)X_1+\left(\beta_{q2}-\beta_{r2}\right)X_2+\ldots+\left(\beta_{qK}-\beta_{rK}\right)X_K$$

In a multinomial model, the difference between β_{qk} and β_{rk} shows how the logit(q/r) is affected by X_k. In the SOR model, this effect is equal to $(\phi_q-\phi_r)\beta_k X_k$. The effect on the logit of any two outcomes in the SOR model is proportional for all independent variables. Differences between scale values indicate how strongly the logit for two options is affected by the independent variables. The greater the difference between scale values, the more the logit between two outcomes is affected by the independent variables. The β_k parameters show how independent variable X_k affects the logit of higher versus lower scores, where "higher" and "lower" are defined by the ϕ_j metric.

A SOR model can be requested by specifying a *varlist* in the `sor` option. A SOR model with only one X_k variable would be trivial and equivalent to standard multinomial model since it contains the same number of parameters. A simple SOR model with two variables could be specified as:

```
use logan
mclgen occ
xi: mclest i.occ, sor(educ black)
```

This model will contain 9 parameters: 4 intercept parameters, 3 independent ϕ_j parameters, and 2 β_k parameters. This is only slightly 3 less than for an unrestricted multinomial model.

However, the parsimony of a SOR model does increase as the number of X_k variables increases.

The SOR model contains both linear and multiplicative elements. To estimate it, `mclest` iteratively estimates MCL models, first taking the ϕ_j scaling metric as given and estimating the β_k parameters, then taking the β_k parameters as given and estimating the ϕ_j parameters. This continues until the change in log likelihood between successive MCL models is less than the value specified in the `sortol` option (default .0001) or the maximum number of iterations specified in the `soriter` option is exceeded (default 10). The standard errors for effects are conditional, given the scaling metric ϕ_j. A likelihood ratio test is therefore advisable before drawing any definite conclusions on the significance of effects.

Row and Columns model 2

A second special model that can be estimated by `mclest` is Goodman's (1979) Row and Columns model 2 (RC2). Originally developed for frequency tables, the RC2 model estimates scaling metrics for both the dependent variable and one of the independent variables. The association between the two variables can then be expressed through a single parameter μ. The scaling metric for the dependent variable is ϕ_j as in the SOR model and the scaling metric for the independent variable is σ_v. Two restrictions must be imposed on ϕ_j and σ_v in order to identify the model. During estimation, `mclest` sets $\phi_1 = \sigma_1 = 0$ and $\phi_J = \sigma_V = 1$. The final estimates are also given for normalized scale values, with mean(ϕ_j)=mean(σ_v)=0 and SS(ϕ_j)=SS(σ_v)=1.

A model containing an RC2 effect could be specified as:

$$\operatorname{logit}\left(\frac{q}{r}\right) = \alpha_q - \alpha_r + (\phi_q - \phi_r)\cdot \mu \cdot \sigma_v$$

This model can be seen as the SOR effects of the categorical variable scaled by $\mu \cdot \sigma_v$. In fact, entering dummies for the categorical variable in a SOR model results in an equivalent model. Using the RC2 specification has the advantages that it expresses the effect of the categorical variable through a single parameter μ and allows a comparison between the scale for the response variable and that of the categorical independent.

A variation of the RC2 model is the EQual Row and Columns model 2 (EQRC2). As the name suggests, this model uses the same metric for both the response variable and the categorical independent.

$$\operatorname{logit}\left(\frac{q}{r}\right) = \alpha_q - \alpha_r + (\phi_q - \phi_r)\cdot \mu \cdot \phi_v$$

This is basically the same model, except that the effects of the categorical variable are scaled by ϕ_v instead of σ_v, thus saving J-2 degrees of freedom.

Another variation implemented in `mclest` allows the association μ between the dependent and independent variable to vary by one or more other independent variables.

$$\mathrm{logit}\left(\frac{q}{r}\right)=\alpha_q-\alpha_r\ +(\phi_q-\phi_r)\cdot(\mu_0+\mu_t X_t)\mu\cdot\sigma_v$$

An overall association parameter μ_0 is estimated, together with μ_t parameters indicating how the association changes for each independent variable X_t, t=1 to T. This is basically a parsimonious interaction effect of the categorical variable and the X_t variables.

An RC2 model is requested by specifying a *varname* in the `rc2` option. At present, only one variable can be used for the RC2 effect. Similarly, an EQRC2 model can be requested by specifying a *varname* in the `eqrc2` option. The `rc2` and `eqrc2` options are mutually exclusive. To let the overall association vary by one or more independent variables, specify a *varlist* in the `muby` option.

Models containing RC2 or EQRC2 effects are estimated by iteratively running MCL models, as is the case for SOR models. Convergence criterion and maximum iterations are determined by the `sortol` and `soriter` options. The standard errors for effects are conditional, given the scaling metrics ϕ_j and σ_v c.q. ϕ_v. A likelihood ratio test is therefore advisable before drawing any definite conclusions on the significance of effects.

The following example estimates a quasi RC2 model for father's occupation, including effects for the likelihood of having the same occupation as father (`diag`) and an rc2 effect. The overall association μ between father's occupation and respondent's occupation is allowed to vary by race. Further more, race and education are included in the model as covariates using a SOR effect.

```
use logan
mclgen occ
gen diag=(focc==occ)*focc
xi: mclest i.occ i.diag, sor(educ black) rc2(focc) muby(black)
```

This produces the following results:

```
. xi: mclest i.occ i.diag, sor(educ black) rc2(focc) muby(black)
i.occ                 Iocc_1-5     (naturally coded; Iocc_1 omitted)
i.diag                Idiag_0-5    (naturally coded; Idiag_0 omitted)

Estimating Stereotype Ordered Regression for educ black

Estimating rc2 effects for focc

mu varies by black


iteration       log likelihood         sub-changes        main changes
----------------------------------------------------------------------
1.1                  -985.5668           -985.5668           -985.5668
1.2                  -971.8872             13.6796           -971.8872
1.3                  -971.4130              0.4742           -971.4130
2.1                  -970.7999              0.6131             14.7669
```

```
2.2                 -970.9284         -0.1285          0.9588
2.3                 -970.7092          0.2192          0.7038
3.1                 -970.7035          0.0057          0.0964
3.2                 -970.6918          0.0117          0.2366
3.3                 -970.6900          0.0018          0.0192
4.1                 -970.6896          0.0004          0.0140
4.2                 -970.6895          0.0001          0.0023
4.3                 -970.6894          0.0002          0.0007
5.1                 -970.6893          0.0000          0.0003
5.2                 -970.6893          0.0000          0.0002
5.3                 -970.6893          0.0000          0.0001
----------------------------------------------------------------------
Convergence criterion .0001 reached in 5 iterations
```

When the `muby` option is used in conjunction with an RC2 model, `mclest` uses three sub-iterations per iteration. First ϕ_j is taken as given and β_k and the product $\sigma_v(\mu_0 + \mu_1 X_1)$ are estimated. Next, β_k and $\sigma_v(\mu_0 + \mu_1 X_1)$ are taken as given and ϕ_j is estimated. In the third sub-iteration, ϕ_j and σ_v are taken as given and β_k, μ_0, and μ_1 are estimated. If the `muby` option is not used or if it is used in conjunction with and EQRC2 model, only two sub-iterations are used. The sub-changes and main changes indicate the change in log likelihood between sub-iterations and main iterations respectively. The change between main iterations is the criterion for determining whether the model has converged.

```
Phi scale for occ

-------------------------------------------------
                                     |      Coef.
-------------------------------------+-----------
phi(1): Farm                         |          0
phi(2): Operatives                   |    -0.4818
phi(3): Craftsmen                    |    -0.2938
phi(4): Sales                        |     0.3653
phi(5): Professional                 |          1
-------------------------------------------------
```

The ϕ_j scale, using the restriction that the first category is fixed to 0 and the last category is fixed to 1. Differences between ϕ_j scale values show how the logit of one occupation versus another is affected by education, race, and scaled father's occupation (controlling for the likelihood of having the same occupation as father). The presence of negative values show that respondents with a higher education, who are non-black, who have a "well-placed" father, are more likely to become a farmer than either an operative or a craftsman. The greatest impact of the independent variables on a logit between adjacent categories is for professionals versus sales, where the difference between the scale values is .635. The smallest impact is on craftsmen versus operatives, a difference of only .188.

```
Sigma scale for focc:

-------------------------------------------------
                                     |      Coef.
-------------------------------------+-----------
sig(1): Farm                         |          0
sig(2): Operatives                   |     0.4401
sig(3): Craftsmen                    |     0.3165
sig(4): Sales                        |     0.6886
sig(5): Professional                 |          1
-------------------------------------------------
```

The σ_v metric defines a "well-placed father", in the context of the model and given the data. As a resource for obtaining a good position and controlling for the likelihood of having the same occupation as one's father, a farm occupation scores the lowest. An operative rather than a craftsman as father increases the likelihood of the respondent getting a better occupation, but the difference is rather small. A father who is a craftsman rather than a farmer, or a professional rather than a salesman, has the greatest impact of any two adjacent scale values on the respondent's occupation.

```
Mu scaled association between occ and focc:

-------------------------------------------------------------------------------
                                     |      Coef.   Std. Err.       z     P>|z|
-------------------------------------+-----------------------------------------
MU: RC2 association between occ and  |     1.0241     0.3551     2.8840  0.0039
MU by black: race                    |     1.1291     1.0633     1.0619  0.2883
-------------------------------------------------------------------------------
```

The μ_0 and μ_1 parameters indicate the magnitude of the effect of father's occupation on respondent's occupation. These standard errors given here are conditional on ϕ_j and σ_v. The estimates indicate that there is a strong association between father's occupation and respondent's occupation but that this is not significantly different for blacks and non-blacks. The impact of having a father who is a professional rather than a salesman on the logit of becoming a professional rather than a salesman is (1-.689)* 1.024* (1-.365) = .202 for non-blacks and (1-.689)* (1.024+1.239)* (1-.365) = .426 (noting that μ_1 is large but also has a large standard error and is not significant).

```
Beta parameters:

-------------------------------------------------------------------------------
                                     |      Coef.   Std. Err.       z     P>|z|
-------------------------------------+-----------------------------------------
SOR effect for educ: education in ye |     0.4406     0.0343    12.8461  0.0000
SOR effect for black: race           |    -1.4045     0.5878    -2.3896  0.0169
-------------------------------------------------------------------------------
```

These are the SOR effects of education and race. The impact of one year of education on the logit of becoming a professional versus a salesman is .441* (1-.365) = .280. The impact of being black rather than non-black is –1.405* (1-.365) = -.891.

```
Full parameter listing:

Conditional (fixed-effects) logistic regression   Number of obs   =       4190
                                                  LR chi2(13)     =     756.04
                                                  Prob > chi2     =     0.0000
Log likelihood =  -970.6893                       Pseudo R2       =     0.2803

------------------------------------------------------------------------------
 __didep |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  Iocc_2 |   6.528278   .5511616     11.845   0.000       5.448021    7.608535
  Iocc_3 |   5.436115   .5270407     10.314   0.000       4.403134    6.469096
  Iocc_4 |   .8334484   .5414884      1.539   0.124      -.2278494    1.894746
  Iocc_5 |  -2.480348   .7093216     -3.497   0.000      -3.870593   -1.090103
```

```
 Idiag_1 |   3.442951   .5782425      5.954   0.000       2.309617    4.576285
 Idiag_2 |   .5049115   .1881059      2.684   0.007       .1362307    .8735924
 Idiag_3 |   .4020211   .1897849      2.118   0.034       .0300496    .7739927
 Idiag_4 |  -.4116733   .3947225     -1.043   0.297      -1.185315    .3619686
 Idiag_5 |  -.5496774   .2975123     -1.848   0.065      -1.132791    .0334361
 __beta1 |   .4406277   .0343005     12.846   0.000       .3734001    .5078554
 __beta2 |  -1.404548   .5877656     -2.390   0.017      -2.556548   -.252549
    __mu |   1.024126   .3551078      2.884   0.004       .3281276    1.720125
 __muby1 |   1.129056    1.06328      1.062   0.288      -.9549348    3.213046
------------------------------------------------------------------------------
```

These are the estimates from the final sub-iteration. Standard errors are conditional, given ϕ_j and σ_v. The first four parameters are standard multinomial logistic intercepts. The `Idiag_*` parameters measure the likelihood of the respondent having the same occupation as his father. Having a father who is a farmer has a tremendous impact on the logit for becoming a farmer versus any other occupation. The `Idiag_4` and `Idiag_5` parameters for sales and professional occupations are negative, reducing the impact of the RC2 effect if father and son both have either of these occupations. The remaining parameters have been treated above.

```
Normalized Solution:

Normalized phi scale for occ:

-------------------------------------------------
                                     |      Coef.
-------------------------------------+-----------
phi(1): Farm                         |    -0.1003
phi(2): Operatives                   |    -0.5101
phi(3): Craftsmen                    |    -0.3502
phi(4): Sales                        |     0.2104
phi(5): Professional                 |     0.7502
-------------------------------------------------

Normalized Sigma scale for focc:

-------------------------------------------------
                                     |      Coef.
-------------------------------------+-----------
sig(1): Farm                         |    -0.6465
sig(2): Operatives                   |    -0.0647
sig(3): Craftsmen                    |    -0.2281
sig(4): Sales                        |     0.2638
sig(5): Professional                 |     0.6755
-------------------------------------------------
```

Unless the `nonorm` option is used, `mclgen` also produces a normalized solution. The scaling metrics now have a mean of zero and a sum of squares of 1.

```
Iteration 0:   log likelihood = -1218.2183
Iteration 1:   log likelihood = -1000.3561
Iteration 2:   log likelihood = -976.65487
Iteration 3:   log likelihood =  -971.5582
Iteration 4:   log likelihood = -970.73319
Iteration 5:   log likelihood =  -970.6895
Iteration 6:   log likelihood =  -970.6893

Conditional (fixed-effects) logistic regression   Number of obs   =       4190
                                                  LR chi2(13)     =     756.04
                                                  Prob > chi2     =     0.0000
```

```
Log likelihood =  -970.6893                              Pseudo R2       =     0.2803

------------------------------------------------------------------------------
 __didep |      Coef.   Std. Err.       z     P>|z|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  Iocc_2 |   6.286966   .5487816     11.456   0.000       5.211373    7.362558
  Iocc_3 |   5.288965   .5245917     10.082   0.000       4.260784    6.317145
  Iocc_4 |   1.016421   .5376269      1.891   0.059      -.0373089     2.07015
  Iocc_5 |  -1.979513   .7071723     -2.799   0.005      -3.365545   -.5934809
 Idiag_1 |   3.442951   .5782421      5.954   0.000       2.309617    4.576285
 Idiag_2 |   .5049115   .1881059      2.684   0.007       .1362307    .8735924
 Idiag_3 |   .4020211   .1897849      2.118   0.034       .0300496    .7739927
 Idiag_4 |  -.4116732   .3947225     -1.043   0.297      -1.185315    .3619686
 Idiag_5 |  -.5496775   .2975123     -1.848   0.065      -1.132791     .033436
 __beta1 |   .5180692   .0403289     12.846   0.000       .4390261    .5971123
 __beta2 |  -1.002211   .3355206     -2.987   0.003      -1.659819   -.3446024
    __mu |   .9108703   .3158372      2.884   0.004       .2918408      1.5299
 __muby1 |   1.004196   .9456943      1.062   0.288      -.8493306    2.857723
------------------------------------------------------------------------------
```

These are the parameter estimates, given the normalized ϕ_j and σ_v scaling metrics. Which set of parameters is used is basically a matter of personal preference.

Syntax for mclgen

`mclgen` *depvar*

The argument *depvar*, a categorical response factor with a maximum of 12 levels, must be specified for use in the MCL model. Note that the `mclgen` program will modify the data and that the data should therefore be saved before running it. Besides transforming the data into a person/choice file, `mclgen` adds `__didep` and `__strata`, the dichotomous dependent variable and stratifying variable used by `clogit`. In addition, it defines the global macros `$ncat` containing the number of categories of the response variable and `$respfact`, the name of the response factor.

Syntax for mclest:

`mclest` *varlist* [*weight*] [`if` *exp*] [`in` *range*]
      [, `sor(`*varlist*`) soriter(`#`) sortol(`#`)`
      `rc2(`*varname*`) eqrc2(`*varname*`) muby(`*varlist*`) nonorm debug`]

`Mclest` is used to estimate a model after transforming the data to a "person/choice" file using `mclgen`. The *varlist* should contain dummies based on the response factor specified in `mclgen` and interactions of these dummies with independent variables. This design matrix can be specified using `xi` or `desmat` (Hendrickx 1999).

Options

The `mclest` program passes the following arguments on to `clogit` unaltered:

    weight, if, in

See the Stata documentation on `clogit` for further details on these options.

The following options are used to request the special nonlinear models Sterotyped Ordered Regression (SOR) and/or the Row and Columns model 2 (RC2).

`sor(`*varlist*`)` specifies a list of variables for which the SOR constraint should be used. Note that at least two variables should be specified, unless either the `rc2` or `eqrc2` option is being used.

`soriter(`*#*`)` specifies the maximum number of iterations for estimating a SOR or RC2 model. The default value is 10.

`sortol(`*#*`)` specifies the convergence criterion for estimating a SOR or RC2 model. The default value is .0001.

`rc2(`*varname*`)` specifies a categorical independent variable for which the RC2 model is to be used. The eqrc2 option will be ignored if the rc2 option is specified.

`eqrc2(`*varname*`)` specifies a categorical independent variable for which the EQRC2 model is to be used. The rc2 option may not be used together with the eqrc2 option.

`muby(`*varlist*`)` specifies one or more variables that affect the association between the rc2 or eqrc2 variable and the dependent variable. Ignored if not used in conjunction with the rc2 or eqrc2 option.

`nonorm` prevents the `mclest` program from estimating a normalized solution if a SOR and/or RC2 model has been requested.

`debug` prints intermediate results of `clogit`. This can be used to determine the source of error if something goes wrong.

Recovering data

`Mclgen` and `mclest` create a number of variables for internal use when estimating models. To transform the person/choice file back to its original form, specify

```
keep if __didep==1
drop __*
```

References

Anderson, J.A. (1984). "Regression and Ordered Categorical Variables." Journal of the Royal Statistical Society, Series B 46: 1-30.

Breen, Richard. (1994). "Individual Level Models for Mobility Tables and Other Cross-Classifications." Sociological Methods & Research 33: 147-173.

DiPrete, Thomas A. (1990). "Adding Covariates to Loglinear Models for the Study of Social Mobility." American Sociological Review 55: 757-773.

Goodman, Leo A. (1979). "Multiplicative models for the analysis of occupational mobility tables and other kinds of cross-classification tables." American Journal of Sociology 84: 804-819.

Hendrickx, John. (1995). "Multinomial Conditional Logit Models for the Analysis of Status Attainment and Mobility." ICS Working Papers nr. 1.

Hendrickx, John, Ganzeboom, Harry B.G. (1998). "Occupational Status Attainment in the Netherlands, 1920-1990. A Multinomial Logistic Analysis". European Sociological Review 14: 387-403.

Hendrickx, John. (1999). "Using categorical variables in Stata". Stata Technical Bulletin STB-52.

Logan, John A. (1983). "A Multivariate Model for Mobility Tables." American Journal of Sociology 89: 324-349.