MIXOR2: MIXOR version 2

- $i = 1 \ldots N$ level-2 units
- $j = 1 \ldots n_i$ level-1 units for level-2 unit i
- $c = 1 \ldots C$ categories

The response model for MIXOR2 can be written as:

$$\mathrm{z}_{ijc} = \gamma_c \pm \left[\boldsymbol{u}'_{ij}\boldsymbol{\zeta}_c + \boldsymbol{w}'_{ij}\boldsymbol{\alpha} + \boldsymbol{x}'_{ij}\boldsymbol{\beta}_i \right]$$

with $\gamma_1 = 0$, $\boldsymbol{\zeta}_1 = \mathbf{0}$, and $\boldsymbol{\beta}_i \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, or by standardizing the random effects as

$$\mathrm{z}_{ijc} = \gamma_c \pm \left[\boldsymbol{u}'_{ij}\boldsymbol{\zeta}_c + \boldsymbol{w}'_{ij}\boldsymbol{\alpha} + \boldsymbol{x}'_{ij}\boldsymbol{\mu} + \boldsymbol{x}'_{ij}\boldsymbol{T}\boldsymbol{\theta}_i \right]$$

with $\boldsymbol{\theta}_i \sim N(\mathbf{0}, \boldsymbol{I})$ and $\boldsymbol{\Sigma} = \boldsymbol{T}\boldsymbol{T}'$

- $\boldsymbol{x}$ = R variables for the random effects
- $\boldsymbol{w}$ = P explanatory variables
- $\boldsymbol{u} = h$ variables (a subset of the P explanatory variables) that interact with the $C - 2$ thresholds $(\gamma_2, \ldots, \gamma_{C-1})$

$\Rightarrow$ for each of the h variables in both $\boldsymbol{w}$ and $\boldsymbol{u}$:

- $\boldsymbol{\alpha}$ includes the effect on the first threshold
- $\boldsymbol{\zeta}_c$ includes $C - 2$ deviations from the effect on the first threshold

New Features

- $\pm$
- $\boldsymbol{\zeta}_c$ non-proportional odds (for the logistic response function, and the analogous generalization under the normal, log-log, and complementary log-log response functions) for explanatory variables $\boldsymbol{u}$
- Response functions allowed (for $c = 1 \ldots C - 1$)

$$P(v_{ij} \le c) = \begin{cases} \Phi(\mathrm{z}) & \text{normal} \\ \Psi(\mathrm{z}) & \text{logistic} \quad 1/(1 + \exp(-\mathrm{z})) \\ \Upsilon(\mathrm{z}) & \text{clog-log} \quad 1 - \exp(-\exp(\mathrm{z})) \\ 1 - \Upsilon(\mathrm{z}) & \text{log-log} \quad \exp(-\exp(\mathrm{z})) \end{cases}$$

- Structure of random effect(s)

$$T = \begin{cases} \vdots\ldots & \text{correlated random effects} \\ \ddots & \text{independent random effects} \\ \vdots & \text{varying } \sigma^2_\beta \text{ across } j \text{ or groups of } i \end{cases}$$

  The latter option is useful for:

  - separate σ^2_β for groups of level-2 units (*e.g.*, DZ and MZ twins)
  - separate σ^2_β for distinct level-1 units (*e.g.*, test items within subjects)

- Mean (vector) of random effects

$$\boldsymbol{\mu} = \begin{cases} \text{estimated} \\ \text{set equal to zero} \end{cases}$$

- Allows right censoring
  - only know $v_{ij} > c$
  - $P(v_{ij} > c) = 1 - P(v_{ij} \leq c)$
  - note: right-censoring for last category C is akin to adding category $C + 1$
  - can estimate γ_C and $\boldsymbol{\zeta}_C$
- Provides linear tranformations of estimated parameters and standard errors
  - $\alpha_1 + \alpha_2$
  - $\alpha_1 + \zeta_1$

- Produces additional output files

  MIXOR.EST - a file containing the estimated parameters (with labels).

  MIXOR.VAR - a file containing the large-sample variance covariance matrix of the parameter estimates (the inverse of the information matrix). The full rectangular matrix is printed out, row by row, with the order of the parameters identical to that of MIXOR.EST (*i.e.*, no labels are given in MIXOR.VAR).

  MIXOR.RES - if there is only one random effect (either R=1, or R> 1 and VGRP=1, see below) then a file containing empirical Bayes estimates of the random effect for each level-2 unit is produced. This file lists for each level-2 unit: level-2 ID, the number of level-1 units n_i, the empirical Bayes estimate (posterior mean), and the posterior standard deviation. Additionally, if each level-2 unit has a frequency weight, then this weight is also output to this file immediately following the level-2 ID (and before n_i).

MIXOR.DEF specifications for MIXOR2: lines 1-5 (*same as MIXOR*)

Line 1 - A title of 60 characters

Line 2 - A subtitle of 60 characters

Line 3 - name of input data file. Any legal filename of 80 characters or less can be specified.

Line 4 - name of main output file. Any legal filename of 80 characters or less can be specified.

Line 5 - name of definition file to be saved or retrieved. Any legal filename of 80 characters or less can be specified. Note that a name for this file *must* be specified even in batch processing, although in batch processing nothing is done to this file.

MIXOR.DEF specifications for MIXOR2: line 6 (*new options are* **bolded**)

Line 6 - NPR NF R P CONV MAXJ MISS START WT CATYX NQUAD **FUNC CEN PARTIAL ADD LINFN DIAG NOMU VGRP**

NPR = # of level-2 units whose data is listed on the screen.

NF = # of fields of data to read from input data file.

R = # of random effects.

P = # of fixed effects (not including mean of random effects).

CONV = convergence criterion (*usually .001 or .0001*).

MAXJ = # of ordered outcome categories.

MISS = 0 if no missing values are present in the data, or 1 if missing values are present (codes will later be defined).

START = 0 if automatic starting values are to be used, or 1 for user-defined starting values.

WT = 0 if each 2nd level unit is weighted equally, or 1 for differential weighting.

CATYX = 0 if a crosstab of any variable by the outcome variable is not requested, and 1 if such a crosstab is requested.

NQUAD = # of quadrature points (per random-effect dimension) for numerical integration (*usually set between 10 and 20 for models with 1 random effect, and between 5 and 10 for models with* > 1 *random effects*).

FUNC = 0, 1, 2, 3 for the probit, logistic, complementary log-log, or log-log response function, respectively.

CEN = 0 for no right-censoring or 1 to include right-censoring.

PARTIAL = # of P fixed effects to interact with the threshold parameters (these are assumed to be the first PARTIAL variables of the P fixed effects indicated on line 9)

ADD = 1 to add the model terms to the thresholds or -1 to subtract the model terms from the thresholds. Changing this option from 1 to -1 (or vice versa) reverses the sign (- or +) of all estimates except the thresholds and variance terms.

LINFN = # of linear transforms of the estimated parameters to estimate.

DIAG = 0 for correlated random effects or 1 for independent random effects.

NOMU = 0 to estimate the mean of the random effects or 1 to fix them to zero.

VGRP = 0 (no) or 1 (yes) for random-effects grouping variables. Specify yes only if R> 1 and the R random-effects variables are dummy-coded level-1 or level-2 grouping variables; otherwise specify no (*ordinarily set to 0*). If yes is specified, then R random-effect variance terms are estimated: one for each of the (level-1 or level-2) groups determined by the dummy-codes.

Note, DIAG and VGRP are mutually exclusive options. Thus, if $R > 1$, there are three possibilities: (1) correlated multiple random effects (DIAG=0 and VGRP=0), (2) independent multiple random effects (DIAG=1 and VGRP=0), or (3) multiple random effects indicating level-1 or level-2 groupings (DIAG=0 and VGRP=1).

MIXOR.DEF specifications for MIXOR2: lines 7-9 (*same as in MIXOR*)

Line 7 - two parameters are to be read on this line: the field of the input data file that contains the (level-2) IDs, followed by the field of the input data file that contains the outcome variable.

Line 8 - R parameters are to be read on this line: the field(s) of the input data file that contain(s) the R random effects. The list of fields does not have to be in ascending order, thus 1 3 2 is an acceptable list of variable fields.

Line 9 - P parameters are to be read on this line: the field(s) of the input data file that contain(s) the P fixed effects. As with the listing of the R fields above, the fields do not have to be in ascending order on this list. **Note that if PARTIAL > 0, the ordering of the P fields is important: it is the first PARTIAL variables on this list that are specified to interact with the threshold parameters**.

MIXOR2 specifications: lines after line 9

next line - (*if WT = 1*) - the field of the input data file that contains the weight to be assigned to each level-2 unit.

next line - **(***if CEN = 1***) - the field of the input data file that contains the CENSOR variable (coded 0=censor and 1=event)**.

next line - the MAXJ values of the ordinal outcome variable.

next line - (*if CATYX = 1*) - two parameters and a list of values: the field of the input data file that contains the variable that is to be crosstabulated with the outcome variable, followed by the # of levels of this variable, and a list of the values for all of these levels.

next line *(if MISS = 1)* - missing value code for the outcome variable.

next line *(if MISS = 1)* - R missing value codes for the random-effect variables.

next line *(if MISS = 1)* - P missing value codes for the fixed effects.

next line - an 8 character label for the ordinal outcome variable.

next line - R labels for the random effects in 8 character width fields.

next line *(if START = 1)* - R starting values for the means of the random effects (*unless NOMU=1*).

next line - P labels for the covariates in 8 character width fields (a maximum of 10 labels per line).

next line *(if START = 1)* - P starting values for the covariate effects.

next line *(if START = 1)* - ((R × (R+1)) / 2) starting values for the variance and covariance terms of the random effects given in "packed" form, *e.g.*, for a 2 x 2 covariance matrix, the order of the starting values should be: variance(1), covariance(1,2) and variance(2). *Note: if either DIAG=1 or VGRP=1 then only R starting values are needed.*

next line - (*if START = 1*) - MAXJ-1 starting values for thresholds (if CEN=1) or MAXJ-2 starting values for thresholds (if CEN=0).

final lines - (*if LINFN* > 0) - **LINFN by NPAR coefficients for the linear transforms of the estimates.** NPAR terms in order are

- R random effect mean vector (unless NOMU=1).
- P effects of explanatory variables.
- random-effect variance-covariance terms:
  - if DIAG=0 and VGRP=0: (R × (R+1)/2) unique elements of the random-effect variance-covariance matrix (in packed form).
  - if DIAG=1 or VGRP=1: R variance terms.
- NCUT thresholds, where NCUT = MAXJ-2 (if CEN=0) or MAXJ-1 (if CEN=1).
- PARTIAL × NCUT interactions:
  - NCUT coefficients for the first of PARTIAL variables
  - NCUT coefficients for the second of PARTIAL variables
  - etc.

Each of these LINFN sets of coefficients are multiplied by the "original" estimates according to the order given. Standard errors for these LINFN transforms are also printed out.

The total number of estimated parameters, NPAR, is equal to NA + NV + NG, where

NA = R + P	if NOMU=0
= P	if NOMU=1
NV = R × (R+1)/2	if DIAG=0 and VGRP=0
= R	if DIAG=1 or VGRP=1
NG = (MAXC-2) × (PARTIAL+1)	if ICEN=0
= (MAXC-1) × (PARTIAL+1)	if ICEN=1