Threshold model

Model description

This example is from Gianola and Foulley: Sire evaluation for ordered categorical data with a threshold model (GSE, 15:201-224, 1983).

The model is:

$$\boldsymbol{\lambda}_y = \mathbf{Xb} + \mathbf{Zs} + \mathbf{e},$$

where

$\boldsymbol{\lambda}_y$ is vector of liability variables for calving ease observations,
$\mathbf{b}$ is vector of fixed herd times year interaction, age of dam and sex of calf effects,
$\mathbf{X}$ is design matrix to link observations to appropriate fixed effects,
$\mathbf{s}$ is vector of random additive genetic sire effects,
$\mathbf{Z}$ is design matrix to link observations to appropriate random effects, and
$\mathbf{e}$ is random residual vector.

It is assumed that $\mathrm{var}(\mathbf{s}) = \mathbf{A}\sigma_s^2$, $\mathrm{var}(\mathbf{e}) = \mathbf{I}\sigma_e^2$ and $\mathrm{cov}(\mathbf{s},\mathbf{e}) = \mathrm{cov}(\mathbf{e},\mathbf{s}) = \mathbf{0}$, where **A** is the numerator relationship matrix for sires. Further, it is assumed that thresholds are known (t_1=0.3753 and t_2=1.0113 according to Gianola and Foulley).

The variance components are $\sigma_s^2 = .05263$ and $\sigma_e^2 = 1$.

Input files

Datafile:

The data file contains information on the data to be analyzed together with class and regression variables for the model. The data file can be in free format (columns are separated by at least one space) or in binary format.

Each record, i.e., line in a free format file, has been divided to two parts: 1) Integer number columns and 2) real number columns. Columns of real numbers are always after the integer number columns.

THR.dat

```
1   2   1   1   1
1   2   2   1   1
1   3   1   1   1
1   2   2   2   2
1   3   1   2   1
1   3   1   2   3
1   3   2   2   1
1   3   2   2   1
1   3   2   2   1
1   2   1   3   1
1   2   1   3   2
.
.
.
```

Column 1: Herd x Year (integer)
Column 2: Age of dam (integer)
Column 3: Sex of calf (integer)
Column 4: Sire (integer)
Column 5: Calving ease (real)

Pedigree file:

All the pedigree information must be given in the pedigree file. Each animal in the pedigree must have a record with four integers of which the forth integer is optional.

THR.ped

```
1   0   0
2   0   0
3   0   0
4   0   0
```

Column 1: Animal ID
Column 2: Sire ID
Column 3: Dam ID

Variance component file:

The variance components file has variances and covariances for all the random effects in the model. The variance component file has a line for each (co)variance. Each line has 3 integers followed by a real number (the (co)variance value). The first integer is the random effect number followed by the row-column combination, and, finally, the (co)variance parameter. Order of lines in the file is irrelevant.

THR.var

```
1 1 1  .052631579  # Genetic variance
2 1 1 1.           # Residual variance must be one
```

Command file:

Threshold model can be given to mix99i only using the old directive method; this model can not be given using CLIM instructions. Everything beyond '#' sign is considered as a comment.

THR.dir

```
# TITLE:
Gianola & Foulley. 1983. GSE 15:201-224, EM algorithm, fixed thresholds
# INT-VAR:
        herd_year dam_age calf_sex sire
# REAL-VAR:
        calv_ease
# TRAITS:
        1
# TRGRP:
        1 -
# SORT_R: Block sorting variable, Relationship code sorting variable
        -                       4
# FIXRAN: Number of different fixed and random effect
        3   1
# MODEL:Sgr,trait, W,
        1   T2   1   -   1 2 3 4
# THR_METHOD:
        em ft
# FIXED THRESHOLDS:
        .3753 1.0113
# BLKORD: Order of effects within herd
        - - - 1
# RANDOM:
        1
# PEDIGR: number: sire
        1        1
# REGRES: VL,VLSITE,VLSITdate,VLSP,vltank,anim
        4   cl       cl      cl     cl
# COMBINE:
        n
# RELMET:
        sm
# INPFIL:
        THR.dat
# VAR:
        4 1 f
# MISSVA:
        -9
# SCALE:
        n
# PEDFIL:
        THR.ped
# PARFIL:
        THR.var
# TMPDIR:
        .
# SOLFOR:
        y
# SOLUNF:
        n
# PRECON: WpW, XpX
        b    f
# NPROC:  Number of processors used by the solver program
        1
# COMBLK: Number of blocks in common area when parallel processing used
        0
```

Solver option file:

Everything beyond '#' sign is considered as a comment.

THR.stop

```
# RAM: RAM demand: H=high, M=medium, L=low
    H
 # STOP: Max_no_PCGiters, Stop_crit, Crit (A/R/D), f=force, Max_no_EMcycles
    110                   1.0e-7     R             f        30
 # RESID: Calculate residuals? (Y/N)
    N
 # VALID: N=no, P=prediction, S=sum of effects, Y=YD, D=DYD, I=IDD, G=generate
    N
 # HETVAR: adjust for HV? (N)o, (S)tart, (C)ontinue), (F)inale iterations.
    N
 # TYPSOL: Solution files? (N)o, (Y)es, (A)itken, (H)alf-Chebychev
    Y
```

Solution files

Structure of the formatted solution files depends on the model. Therefore, explanation of the content of those files is given in the printout of the particular run of the program.

Solver option file for threshold model analysis has been given in 'Input files for Example 8'.

"Solfix"-file contains solutions for fixed effects.

```
 Fact. Trt       Level     N-Obs  Solution                 Factor Trait
   1    1          1         13  0.0000                 herd_yea calv_eas
   1    1          2         15 0.29746                 herd_yea calv_eas
   2    1          2         13  0.0000                  dam_age calv_eas
   2    1          3         15 -.12691                  dam_age calv_eas
   3    1          1         14 -.19690E-03             calf_sex calv_eas
   3    1          2         14 -.39079                 calf_sex calv_eas
```

Column 1: Factor number
Column 2: Trait number
Column 3: Level code
Column 4: Number of observations
Column 5: Solution
Column 6: Name of factor
Column 7: Name of trait

"Solani"-file contains solutions for genetic animal effect.

```
  1       0       7 -.81533E-01
  2       0       9 0.65485E-01
  3       0       6 0.12279
  4       0       6 -.10674
```

Column 1: Animal ID
Column 2: Number of descendants

Column 3: Number of observations
Column 4: Solution for trait 1 calv_eas and factor sire