Scientific Software International

Transposing data for use in HMLM/HMLM2

In the case of repeated measures data, it is often necessary to transpose the data.

Consider the example of data collected on children at the end of grade 3 and followed up annually thereafter until grade 5. The variables in the data set are the identification number of the child, the grade the child was in at the time of measurement and the mathematics score obtained by the child at the end of each year. Data for the first two children are shown below. Note that the third measurement for child 0002 was not made.

Data set 1:

ID	Grade 3	Math 3	Grade 4	Math 4	Grade 5	Math 5
0001	3	1.15	4	1.13	5	2.30
0002	3	2.43	4	3.87	.	.

In order to analyze the changes in mathematics scores over time, the data have to be rewritten in the format shown below.

Data set 2:

ID	Grade	Math
0001	3	1.15
0001	4	1.13
0001	5	2.30
0002	3	2.43
0002	4	3.87

If, moreover, the data are to be used as input for the creation of a MDM file to be used in a multivariate analysis with the HMLM and HMLM2 modules, where indicator variables matching the order of the measurements have to be added to the data set, the transposed data set should also contain the following variables:

Data set 3:

ID	Grade	Math	Indicator 1	Indicator 2	Indicator 3
0001	3	1.15	1.00	0.00	0.00
0001	4	1.13	0.00	1.00	0.00
0001	5	2.30	0.00	0.00	1.00

0002	3	2.43	1.00	0.00	0.00
0002	4	3.87	0.00	1.00	0.00

The following lines of SAS code may be used to rewrite data set 1 in the form of data set 3. Note that the code should be changed to reflect the number and names of the variables from a different data set.

```
data d1;

infile 'school.dat';

input id grade3 math3 grade4 math4 grade5 math5;

*change counter to reflect number of variables and list variable names; array aitems[3] math3 math4 math5;

*change counter to reflect number of variables and list indicator variable names followed by a 0 for each indicator variable to be created;

array avec[3] adum3-adum5 (0 0 0);

do i=1 to 3;*change counter to reflect number of variables; outcome = aitems[i];

do j=1 to 3;*change counter to reflect number of variables; avec[j] = (j=i);

end;

if outcome ne . then output; end;

do i=1 to 3; *change counter to reflect number of variables; avec[i] = 0;

end;
```