Censored regressions

The data in the file **DATA.EX3** consist of scores of Swedish school children on reading and spelling tests that relate to metaphonological aspects of the Swedish language. Each schore is the number of correctly answered items. There are 11 tests and 90 cases. In the first run, **EX3A.PRL**, we treat all variables as continuous and estimate the product-moment correlations. In the second run (**EX3B.PRL**), after observing that some of the tests have large "floor" and "ceiling" effects, we declare all variables as censored and re-estimate the product-moment correlations. The data and syntax files are available in the **PRELIS examples** folder.

```
EXAMPLE 3A: TEST SCORE DATA
DA NI=11 NO=90
LA FI=LABELS.EX3
RA FI=DATA.EX3
CO ALL
OU MA=KM
```

In this example, all variables are continuous. The labels are read from a file called **LABELS.EX3**.

Univariate Summary Statistics for Continuous Variables

Variable	Mean	St. Dev.	Skewness	Kurtosis	Minimum	Freq.	Maximum	Freq.
V01	21.789	7.856	-1.117	0.333	0.000	2	30.000	6
V02	14.622	7.048	-0.173	-0.568	0.000	2	28.000	5
V07	11.489	3.069	-0.175	2.673	0.000	1	20.000	2
V08	14.478	3.660	-0.490	2.947	0.000	1	25.000	1
V09	19.122	6.830	-0.635	0.359	0.000	3	32.000	1
V10	21.622	5.560	-0.553	0.163	6.000	1	33.000	2
V21	15.022	2.998	-1.956	3.424	4.000	1	17.000	41
V22	13.122	3.304	-1.123	0.935	2.000	1	17.000	8
V23	12.578	3.402	-1.141	0.735	3.000	3	16.000	21
V24	8.611	3.550	-0.271	-0.551	0.000	1	15.000	2
V25	16.589	4.271	-1.338	2.425	1.000	1	22.000	9

V21 has a high negative skewness; V07, V08, V21, and V25 have high kurtoses; and V21 and V23 have high ceiling effects. For data like these, it is likely that the correlations will be biased due to restrictions of range.

```
Correlation Matrix

             V01        V02        V07        V08        V09        V10
          --------   --------   --------   --------   --------   --------
   V01     1.000
   V02     0.755      1.000
   V07     0.634      0.599      1.000
   V08     0.640      0.618      0.872      1.000
   V09     0.474      0.429      0.345      0.377      1.000
   V10     0.309      0.423      0.339      0.330      0.581      1.000
   V21     0.554      0.568      0.393      0.423      0.296      0.322
   V22     0.509      0.527      0.434      0.399      0.309      0.360
   V23     0.518      0.596      0.484      0.474      0.425      0.489
   V24     0.662      0.672      0.603      0.569      0.503      0.529
   V25     0.726      0.756      0.528      0.516      0.380      0.397

Correlation Matrix

             V21        V22        V23        V24        V25
          --------   --------   --------   --------   --------
   V21     1.000
   V22     0.625      1.000
   V23     0.564      0.519      1.000
   V24     0.538      0.613      0.558      1.000
   V25     0.731      0.634      0.619      0.628      1.000
```

In the second run, Example 3B (**ex3b.prl**), we treat all variables as censored below and above.

```
EXAMPLE 3B: TEST SCORE DATA
DA NI=11 NO=90
LA FI=LABELS.EX3
RA FI=DATA.EX3
CE ALL
OU MA=KM
```

Output from the second run was as follows:

```
Univariate Summary Statistics for Continuous Variables
```

Variable	Mean	St. Dev.	Skewness	Kurtosis	Minimum	Freq.	Maximum	Freq.
V01	21.789	7.856	-1.117	0.333	0.000	2	30.000	6
V02	14.622	7.048	-0.173	-0.568	0.000	2	28.000	5
V07	11.489	3.069	-0.175	2.673	0.000	1	20.000	2
V08	14.478	3.660	-0.490	2.947	0.000	1	25.000	1
V09	19.122	6.830	-0.635	0.359	0.000	3	32.000	1
V10	21.622	5.560	-0.553	0.163	6.000	1	33.000	2

```
V21    15.022     2.998    -1.956     3.424     4.000        1   17.000     41
V22    13.122     3.304    -1.123     0.935     2.000        1   17.000      8
V23    12.578     3.402    -1.141     0.735     3.000        3   16.000     21
V24     8.611     3.550    -0.271    -0.551     0.000        1   15.000      2
V25    16.589     4.271    -1.338     2.425     1.000        1   22.000      9
```

Correlation Matrix

```
              V01        V02        V07        V08        V09        V10

            --------   --------   --------   --------   --------   --------
V01          1.000
V02          0.761      1.000
V07          0.641      0.595      1.000
V08          0.626      0.603      0.870      1.000
V09          0.462      0.390      0.333      0.372      1.000
V10          0.302      0.413      0.346      0.326      0.556      1.000
V21          0.407      0.443      0.355      0.357      0.250      0.249
V22          0.476      0.513      0.411      0.386      0.285      0.336
V23          0.429      0.512      0.442      0.398      0.356      0.406
V24          0.665      0.676      0.612      0.568      0.483      0.531
V25          0.696      0.766      0.525      0.494      0.352      0.395
```

 Correlation Matrix

```
              V21        V22        V23        V24        V25

            --------   --------   --------   --------   --------
V21          1.000
V22          0.462      1.000
V23          0.437      0.405      1.000
V24          0.408      0.606      0.485      1.000
V25          0.583      0.604      0.555      0.629      1.000
```

Both skewness and kurtosis are now smaller than before. Many correlations are smaller than in the previous run, especially those for variables V21 and V23.