Tetrachoric correlations, with asymptotic variances estimated from grouped data

Bock & Lieberman (1970) and Christoffersson (1975) published the data in the table below giving observed frequencies for the 32 response patterns arising from five items (11 through 15) of Section 6 of the Law School Admissions Test (LSAT). All items are dichotomous.

Table: Observed frequencies of response patterns for five items of LSAT6

	Response pattern					
Index	**1**	**2**	**3**	**4**	**5**	**Frequency**
1	0	0	0	0	0	3
2	0	0	0	0	1	6
3	0	0	0	1	0	2
4	0	0	0	1	1	11
5	0	0	1	0	0	1
6	0	0	1	0	1	1
7	0	0	1	1	0	3
8	0	0	1	1	1	4
9	0	1	0	0	0	1
10	0	1	0	0	1	8
11	0	1	0	1	0	0
12	0	1	0	1	1	16
13	0	1	1	0	0	0
14	0	1	1	0	1	3
15	0	1	1	1	0	2
16	0	1	1	1	1	15
17	1	0	0	0	0	10
18	1	0	0	0	1	29
19	1	0	0	1	0	14
20	1	0	0	1	1	81
21	1	0	1	0	0	3
22	1	0	1	0	1	28
23	1	0	1	1	0	15
24	1	0	1	1	1	80
25	1	1	0	0	0	16
26	1	1	0	0	1	56

27	1	1	0	1	0	21
28	1	1	0	1	1	173
29	1	1	1	0	0	11
30	1	1	1	0	1	61
31	1	1	1	1	0	28
32	1	1	1	1	1	298
Total						1000

The sample is a subsample of 1000 from a larger sample of those who took the test. In this example, we use this data to illustrate how quickly PRELIS can compute tetrachoric correlations for these five items and asymptotic variances of these correlations. Data are in the file **LSAT6.DAT** in the **PRELIS Examples** folder.

```
EXAMPLE 5A: LSAT SECTION 6
DA NI=6 NO=1000
RA FI=LSAT6.DAT
WE 6
OR ALL
OU MA=PM SV=AVAR.TST PV
```

Some sections of output follow.

```
Correlation Matrix

              VAR 1      VAR 2      VAR 3      VAR 4      VAR 5
            --------   --------   --------   --------   --------
    VAR 1      1.000
    VAR 2      0.170      1.000
    VAR 3      0.228      0.189      1.000
    VAR 4      0.107      0.111      0.187      1.000
    VAR 5      0.067      0.172      0.105      0.201      1.000

 Total Variance = 5.000 Generalized Variance = 0.803

 Largest Eigenvalue = 1.619 Smallest Eigenvalue = 0.731

 Condition Number = 1.488
```

These correlations essentially agree with those reported by Christoffersson (1975). The largest difference is two units in the third decimal.

```
Asymptotic Variances of Correlations

             R(2,1)     R(3,1)     R(3,2)     R(4,1)     R(4,2)     R(4,3)
            --------   --------   --------   --------   --------   --------
            0.00550    0.00505    0.00263    0.00612    0.00322    0.00284
```

Asymptotic Variances of Correlations

R(5,1)	R(5,2)	R(5,3)	R(5,4)
0.00819	0.00409	0.00392	0.00426

These are large sample estimates of the variances of the estimated tetrachoric correlations. The square roots of these variances are the standard error of the estimated correlations. These can be used to set uyp approximate confidence intervals for the correlations. For example, an approximate 95 percent confidence interval for ρ_{21} is $0.170 \pm 2\sqrt{-/00549} = 0.170 \pm 0.148.$