



Compute matrix of polyserial and polychoric correlations

Data for Example 1A consist of the first hundred cases of the artificial data stored in **Data.100**. We examine the data and use the M option to compute the matrix of polychoric and polyserial correlations.

There are six variables. Two of them (variables 1 and 5) are continuous, labeled CONTIN1 and CONTIN2. Four others (Variables 2, 3, 4, and 6) are ordinal, labeled ORDINAL1, ORDINAL2, ORDINAL3, and ORDINAL4. The number “-9” is used to represent missing values.

```

EXAMPLE 1A: PRELIS TEST ON DATA.100
DA NI = 6   NOBS = 100   MISSING = -9   TREATMENT = PAIRWISE
RAW-DATA = DATA.100
LABELS
CONTIN1 ORDINAL1 ORDINAL2 ORDINAL3 CONTIN2 ORDINAL4
ORDINAL ORDINAL1-ORDINAL3 ORDINAL4
OUTPUT MATRIX = PMATRIX
  
```

PRELIS lists all the commands in the output file, so that their correct interpretation can be confirmed. If a syntax error is found, an error message will appear just below the line where the error is located. The program will stop running this problem and will continue with the next problem, if any.

Number of Missing Values per Variable

CONTIN1	ORDINAL1	ORDINAL2	ORDINAL3	CONTIN2	ORDINAL4
21	16	22	25	24	20

Distribution of Missing Values

Total Sample Size(N) = 100

Number of Missing Values	0	1	2	3	4
Number of Cases	23	40	25	10	2

There are 23 cases with no missing values, 40 with one missing value, etc. If listwise deletion had been used, the effective sample size would have been 23. As the data matrix consists of 100 cases, this may seem like a very large reduction in the sample size. But this is actually typical of what will happen if 20 percent (approximately 120 out of 600) of the observations are missing at random.

Effective Sample Sizes

Univariate (in Diagonal) and Pairwise Bivariate (off Diagonal)

	CONTIN1	ORDINAL1	ORDINAL2	ORDINAL3	CONTIN2	ORDINAL4
CONTIN1	79					
ORDINAL1	64	84				
ORDINAL2	60	63	78			
ORDINAL3	58	65	62	75		
CONTIN2	59	63	58	57	76	
ORDINAL4	65	67	64	60	62	80

There are 79 cases with no missing values on CONTIN1; 64 cases with no missing values on both CONTIN1 and ORDINAL1; etc. These are the effective sample sizes that will be used under pairwise deletion.

Percentage of Missing Values

Univariate (in Diagonal) and Pairwise Bivariate (off Diagonal)

	CONTIN1	ORDINAL1	ORDINAL2	ORDINAL3	CONTIN2	ORDINAL4
CONTIN1	21.00					
ORDINAL1	36.00	16.00				
ORDINAL2	40.00	37.00	22.00			
ORDINAL3	42.00	35.00	38.00	25.00		
CONTIN2	41.00	37.00	42.00	43.00	24.00	
ORDINAL4	35.00	33.00	36.00	40.00	38.00	20.00

This table provides basically the same information as the previous table; but in terms of percentage of missing observation. Thus, 21 percent of the 100 cases are missing on CONTIN1; 36 percent are missing on both CONTIN1 and ORDINAL1; etc.

Univariate Marginal Parameters

Variable	Mean	St. Dev.	Thresholds					
ORDINAL1	0.000	1.000	-1.668	-0.751	-0.531	-0.334	0.000	0.531
ORDINAL2	0.000	1.000	-0.361	0.097	0.293	0.615		
ORDINAL3	0.000	1.000	-0.890	0.750				
ORDINAL4	0.000	1.000	-0.524					

CONTIN2 -0.803 0.422 0.314 0.754 0.744 0.689

Correlations and Test Statistics

(PE=Pearson Product Moment, PC=Polychoric, PS=Polyserial)

Variable vs. Variable	Correlation	Test of Model			Test of Close Fit	
		Chi-Squ.	D.F.	P-Value	RMSEA	P-Value
ORDINAL1 vs. CONTIN1	0.706 (PS)	21.456	11	0.029	0.122	0.283
ORDINAL2 vs. CONTIN1	0.382 (PS)	3.464	7	0.839	0.000	0.960
ORDINAL2 vs. ORDINAL1	0.389 (PC)	23.831	23	0.413	0.024	0.925
ORDINAL3 vs. CONTIN1	0.349 (PS)	0.333	3	0.954	0.000	0.979
ORDINAL3 vs. ORDINAL1	0.438 (PC)	8.745	11	0.645	0.000	0.933
ORDINAL3 vs. ORDINAL2	0.562 (PC)	6.187	7	0.518	0.000	0.824
CONTIN2 vs. CONTIN1	0.272 (PE)					
ORDINAL1 vs. CONTIN2	0.243 (PS)	4.403	11	0.957	0.000	0.996
ORDINAL2 vs. CONTIN2	0.304 (PS)	10.668	7	0.154	0.095	0.466
ORDINAL3 vs. CONTIN2	0.275 (PS)	12.588	3	0.006	0.237	0.037
ORDINAL4 vs. CONTIN1	0.258 (PS)	1.587	1	0.208	0.095	0.344
ORDINAL4 vs. ORDINAL1	0.473 (PC)	6.369	5	0.272	0.064	0.601
ORDINAL4 vs. ORDINAL2	0.310 (PC)	8.365	3	0.039	0.167	0.158
ORDINAL4 vs. ORDINAL3	0.160 (PC)	0.061	1	0.804	0.000	0.854
ORDINAL4 vs. CONTIN2	0.044 (PS)	2.941	1	0.086	0.177	0.183

Percentage of Tests Exceeding 0.5% Significance Level: 0.0%

Percentage of Tests Exceeding 1.0% Significance Level: 0.0%

Percentage of Tests Exceeding 5.0% Significance Level: 7.1%

For each pair of variables for which a polychoric or polyserial correlation has been estimated, this table provides a goodness-of-fit test if the model of an underlying bivariate normal distribution. Such a test is not provided when the estimated correlation is a product moment correlation. The table also gives statistics for testing the hypothesis that the correlation in the bivariate distribution is zero.

Correlation Matrix

	CONTIN1	ORDINAL1	ORDINAL2	ORDINAL3	CONTIN2	ORDINAL4
CONTIN1	1.000					
ORDINAL1	0.706	1.000				
ORDINAL2	0.382	0.389	1.000			
ORDINAL3	0.349	0.438	0.562	1.000		
CONTIN2	0.272	0.243	0.304	0.275	1.000	
ORDINAL4	0.258	0.473	0.310	0.160	0.044	1.000

This correlation matrix may be saved in a file by specifying SM = filename on the OU command.