# Estimating the disattenuated correlation

In this example from Lord (1957), two measures $x_1$ and $x_2$ are 15-item vocabulary tests administered under liberal time limits. Two other measures $x_3$ and $x_4$ are highly speeded 75-item vocabulary tests. The covariance matrix based on NO = 649 examinees, is given in the table below.

**Table: Covariance matrix for four vocabulary measures**

| | $x_1$ | $x_2$ | $x_3$ | $x_4$ |
|---|---|---|---|---|
| $x_1$ | 86.3979 | | | |
| $x_2$ | 57.7751 | 86.2632 | | |
| $x_3$ | 56.8651 | 59.3177 | 97.2850 | |
| $x_4$ | 58.8986 | 59.6683 | 73.8201 | 97.8192 |

The disattenuated correlation, $\phi$, between the two latent variables is estimated, and the hypothesis $\phi = 1$ is tested. In addition, the hypothesis that the measures are parallel is tested.

The measurement model is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_2 & 0 \\ 0 & \lambda_3 \\ 0 & \lambda_4 \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix},$$

with covariance matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \lambda_1 & 0 \\ \lambda_2 & 0 \\ 0 & \lambda_3 \\ 0 & \lambda_4 \end{bmatrix} \begin{bmatrix} 1 & \phi \\ \phi & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_2 & 0 & 0 \\ 0 & 0 & \lambda_3 & \lambda_4 \end{bmatrix} + \begin{bmatrix} \theta_1 & 0 & 0 & 0 \\ 0 & \theta_2 & 0 & 0 \\ 0 & 0 & \theta_3 & 0 \\ 0 & 0 & 0 & \theta_4 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1^2 + \theta_1 & & & \\ \lambda_1\lambda_2 & \lambda_2^2 + \theta_2 & & \\ \lambda_1\lambda_3\phi & \lambda_2\lambda_3\phi & \lambda_3^2 + \theta_3 & \\ \lambda_1\lambda_4\phi & \lambda_2\lambda_4\phi & \lambda_3\lambda_4 & \lambda_4^2 + \theta_4 \end{bmatrix}.$$

In this model, $x_1$ and $x_2$ are congeneric measures of $\xi_1$, and $x_3$ and $x_4$ are congeneric measures of $\xi_2$. The disattenuated correlation $\phi$ is the correlation between $\xi_1$ and $\xi_2$. To analyze the data, one can set up four hypotheses:

$$\begin{aligned} H_1 &: \quad \lambda_1 = \lambda_2, \lambda_3 = \lambda_4, \theta_1 = \theta_2, \theta_3 = \theta_4, \phi = 1 \\ H_2 &: \quad \lambda_1 = \lambda_2, \lambda_3 = \lambda_4, \theta_1 = \theta_2, \theta_3 = \theta_4 \\ H_3 &: \quad \phi = 1 \\ H_4 &: \quad \lambda_1, \lambda_2, \lambda_3, \lambda_4, \theta_1, \theta_2, \theta_3, \theta_4, \phi \text{ unconstrained,} \end{aligned}$$

and estimate the model under each of these. Under hypotheses $H_1$, $H_2$, and $H_3$, the model involves *equality constraints*, imposed on the parameters of the base model $H_4$.

All four models can be estimated in one run using stacked input. The command file analyzes the models in the order $H_4$, $H_3$, $H_2$, and $H_1$. The covariance matrix of the four variables is in the file **EX33.COV** (see the **LISREL Examples** folder) which is rewound after each problem. Because $\boldsymbol{\Phi}$ is singular by definition in $H_1$ and $H_3$, AD must be set OFF in these cases. Here is the contents of **EX33.LIS**:

```
ESTIMATING THE DISSATTENUATED CORRELATION    HYPOTHESIS 4
DA NI=4 NO=649
CM FI=EX33.COV RE
MO NX=4 NK=2 PH=FI
FR LX 1 1 LX 2 1 LX 3 2 LX 4 2
FR PH 2 1
VA 1 PH 1 1 PH 2 2
OU SE
ESTIMATING THE DISSATTENUATED CORRELATION    HYPOTHESIS 3
DA NI=4 NO=649
CM FI=EX33.COV RE
MO NX=4 NK=2 PH=FI
FR LX 1 1 LX 2 1 LX 3 2 LX 4 2
VA 1 PH 1 1 PH 2 1 PH 2 2
OU SE AD=OFF
ESTIMATING THE DISSATTENUATED CORRELATION    HYPOTHESIS 2
```

```
DA NI=4 NO=649
CM FI=EX33.COV RE
MO NX=4 NK=2 PH=FI
FR LX 1 1 LX 2 1 LX 3 2 LX 4 2
FR PH 2 1
VA 1 PH 1 1 PH 2 2
EQ LX 1 1 LX 2 1
EQ TD 1 TD 2
EQ LX 3 2 LX 4 2
EQ TD 3 TD 4
OU SE
ESTIMATING THE DISSATTENUATED CORRELATION    HYPOTHESIS 1
DA NI=4 NO=649
CM FI=EX33.COV
MO NX=4 NK=2 PH=FI
FR LX 1 1 LX 2 1 LX 3 2 LX 4 2
VA 1 PH 1 1 PH 2 1 PH 2 2
EQ LX 1 1 LX 2 1
EQ TD 1 TD 2
EQ LX 3 2 LX 4 2
EQ TD 3 TD 4
OU SE AD=OFF
```

Results for the four analyses are reported in the tables below.

**Table: Four vocabulary measures: summary of analyses**

| Hypothesis | No. of par. | $\chi^2$ | *Df* | *P* |
|---|---|---|---|---|
| $H_1$ | 4 | 37.33 | 6 | 0.00 |
| $H_2$ | 5 | 1.93 | 5 | 0.86 |
| $H_3$ | 8 | 36.21 | 2 | 0.00 |
| $H_4$ | 9 | 0.70 | 1 | 0.40 |

**Table: Four vocabulary measures: test of hypotheses**

| | Parallel | Congeneric | |
|---|---|---|---|
| $\varphi = 1$ | $\chi_6^2 = 37.33$ | $\chi_2^2 = 36.21$ | $\chi_4^2 = 1.12$ |
| $\varphi \neq 1$ | $\chi_5^2 = 1.93$ | $\chi_1^2 = 0.70$ | $\chi_4^2 = 1.23$ |
| | $\chi_1^2 = 35.40$ | $\chi_1^2 = 35.51$ | |

Each hypothesis is tested against the general alternative that $\mathbf{\Sigma}$ is unconstrained. To consider various hypotheses that can be tested, the four $\chi^2$ values from the first table are also recorded in the second. Test of $H_1$ against $H_2$ gives $\chi_1^2 = 35.40$ with one degree of freedom. An alternative test is $H_3$ against $H_4$, which gives $\chi_1^2 = 35.51$ with one degree of freedom. Thus, regardless of whether we treat the two pairs of measures as parallel or congeneric, the hypothesis $\phi = 1$ is rejected. There is strong evidence that the unspeeded and speeded measures do not measure the same trait. The hypothesis that the two pairs of measures are parallel can also be tested by means of the table above. This gives $\chi_4^2 = 1.12$ or $\chi_4^2 = 1.23$ with four degrees of freedom, depending on whether we assume $\phi = 1$ or $\phi \neq 1$. Thus we cannot reject the hypothesis that the two pairs of measures are parallel. It appears that $H_2$ is the most reasonable of the four hypotheses. The maximum likelihood estimate of $\phi$ under $H_2$ is $\hat{\phi}$ = 0.899 with a standard error of 0.019. An approximate 95% confidence interval for $\phi$ is 0.86 < $\phi$ <0.94.