

## Why are *t*-values for Error Variances Equal? A Paradox in Path Models for Observed Variables

Karl G Jöreskog

Let  $\mathbf{y} = (y_1, y_2, \dots, y_p)$  be a set of jointly dependent (endogenous) variables and let  $\mathbf{x} = (x_1, x_2, \dots, x_q)$  be a set of independent (exogenous) variables. Consider any model of the form

$$\mathbf{y} = \alpha + \mathbf{B}\mathbf{y} + \mathbf{\Gamma}\mathbf{x} + \mathbf{z} \quad (1)$$

where  $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_p)$  is a vector of intercept terms and  $\mathbf{z} = (z_1, z_2, \dots, z_p)$  is a vector of error terms assumed to be uncorrelated with  $\mathbf{x}$ . There are no latent variables in the model. The matrix  $\mathbf{B}$  is assumed to have zero elements in and above the diagonal, i.e., (1) is a recursive system, see Jöreskog & Sörbom (1996a, pp 143-145). Any element below the diagonal of  $\mathbf{B}$  may be either a fixed zero or a free parameter. The matrix  $\mathbf{\Gamma}$  may contain any number of fixed zeros and any number of free elements. The error covariance matrix  $\psi = \text{Cov}(\mathbf{z})$  is assumed to be diagonal. Suppose the model is estimated by ML from a sample of size  $N$  under the assumption that  $\mathbf{z}$  is multivariate normal. Nothing needs to be assumed about the distribution of  $\mathbf{x}$ .

In scalar notation, equation (1) is

$$y_i = \alpha_i + \beta_{i1}y_1 + \beta_{i2}y_2 + \dots + \beta_{i,i-1}y_{i-1} + \gamma_{i1}x_1 + \gamma_{i2}x_2 + \dots + \gamma_{iq}x_q + z_i, \quad i = 1, 2, \dots, p, \quad (2)$$

where some of the  $\beta$ 's and  $\gamma$ 's may be zero. If  $\beta_{im} = 0$ ,  $y_i$  does not depend on  $y_m$  and if  $\gamma_{in} = 0$ ,  $y_i$  does not depend on  $x_n$ . The assumptions made are sufficient to show that each equation in (2) is a regression equation in the sense that  $z_i$  is uncorrelated with all variables appearing on the right side of (2). Then the asymptotic variance of the error variance  $\text{Var}(z_i)$  is given by

$$N\text{AVar}(\hat{\psi}_{ii}) = 2\psi_{ii}^2, \quad i=1,2,\dots, p, \quad (3)$$

where  $\psi_{ii}$  is the true error variance. An unbiased estimate of the variance of  $\hat{\psi}_{ii}$  - hat is

$$\text{Var}(\hat{\psi}_{ii}) = 2\hat{\psi}_{ii}^2/(N-1-s), \quad i=1,2,\dots, p, \quad (4)$$

where  $s$  is the number of genuine variables appearing on the right side of (2).

LISREL estimates model (1) by fitting the covariance matrix implied by the model to the sample covariance matrix and the asymptotic variances of parameter estimates are obtained from the information matrix. LISREL does not make use of the fact that each equation can be estimated separately as a regression model. All that LISREL knows is that  $\mathbf{z}$  is uncorrelated with  $\mathbf{x}$ . Therefore, LISREL estimates the variance of the error variance using  $s = q$  in (4), i.e. uses the asymptotically equivalent formula

$$\text{Est}[\text{Var}(\hat{\psi}_{ii})] = 2\hat{\psi}_{ii}^2/(N-1-q), \quad i=1,2,\dots, p, \quad (5)$$

A consequence of (5) is that the t-value for  $(\hat{\psi})_{ii}$  will be

$$t(\hat{\psi}_{ii}) = \sqrt{(N-1-q)/2}, i=1,2,\dots, p, \quad (6)$$

In other words, the t-values of  $(\hat{\psi})_{ii}$  are the same for all  $i=1,2, \dots, p$ . This is a paradoxical result since the t-values are not only the same for all error variances but also for all models containing the same number of x-variables and estimated from the same sample, regardless of what restrictions are imposed on **B** and **G**. Furthermore, these t-values do not depend on the data other than through the number of x-variables.

This paradox can be observed in the examples EX44.LS8, EX45A.LS8, and EX44B.LS8 (see Jöreskog & Sörbom, 1996a) and in the examples EX2A.SPL, EX2B.SPL, and EX3A.SPL (see Jöreskog & Sörbom, 1996b).

The conditions for this result can be slightly extended by allowing for some of the error terms to be correlated. Then (3) - (5) will hold only for those equations which are regression equations in the sense that the error term  $z_i$  is uncorrelated with all variables appearing on the right side of the equation.

Here is an example with  $p = 4$  and  $q = 3$ . This example is based on a fictitious sample covariance matrix assumed to be estimated from 767 cases. The covariance matrix is in the file EQTVAL.COV which is

1.525						
1.974	7.172					
1.377	4.066	5.438				
1.404	2.602	1.552	3.948			
-0.214	-0.943	-0.590	-0.329	2.736		
0.702	2.310	1.367	1.064	-0.486	3.740	
0.608	2.181	1.533	0.987	-0.355	2.347	3.944

I present both a LISREL and SIMPLIS input file for this example.

### SIMPLIS

The input is

```

Illustrating Equal T-Values
Observed Variables: Y1 Y2 Y3 Y4 X1 X2 X3
Covariance Matrix from file EQTVAL.COV
Sample size: 767
Relationships:
Y1 = X2 X3
Y2 = Y1 X1 X2 X3
Y3 = Y1 Y2 X3
Y4 = Y1 Y2
Number of Decimals = 3
End of Problem

```

This input gives the following results in the output file

```

Y1 = 0.145*X2 + 0.0678*X3, Errorvar.= 1.382
      (0.0278)      (0.0271)      (0.0707)
      5.222      2.503      19.532

```

$$\begin{aligned}
Y2 &= 1.063*Y1 - 0.188*X1 + 0.256*X2 + 0.220*X3, \text{ Errorvar.} = 3.827 \\
&\quad (0.0602) \quad (0.0433) \quad (0.0474) \quad (0.0452) \quad (0.196) \\
&\quad 17.639 \quad -4.329 \quad 5.399 \quad 4.866 \quad 19.532 \\
\\
Y3 &= 0.296*Y1 + 0.475*Y2 - 0.0482*Y4 + 0.0926*X3, \text{ Errorvar.} = 3.033 \\
&\quad (0.0692) \quad (0.0319) \quad (0.0397) \quad (0.0348) \quad (0.155) \\
&\quad 4.274 \quad 14.896 \quad -1.213 \quad 2.657 \quad 19.532 \\
\\
Y4 &= 0.701*Y1 + 0.170*Y2, \text{ Errorvar.} = 2.522 \\
&\quad (0.0578) \quad (0.0267) \quad (0.129) \\
&\quad 12.128 \quad 6.358 \quad 19.532
\end{aligned}$$

Note that the t-values for the error variances are all equal to 19.532.

## LISREL

The input is

```

Illustrating Equal T-Values
DA NI=7 NO=767
LA=EQTVAL.DAT
CM=EQTVAL.DAT
MO NY=4 NX=3 BE=SD
FI BE(4,3) GA(1,1) GA(3,1) GA(3,2) GA(4,1)-GA(4,3)
OU ND=3

```

This gives the following estimates, standard errors, and t-values for the  $\psi$ 's.

PSI	Y1	Y2	Y3	Y4
	1.382	3.826	3.040	2.523
	(0.071)	(0.196)	(0.156)	(0.129)
	19.532	19.532	19.532	19.532

Now consider the case where  $z_1$  and  $z_3$  are correlated. In the input this may be specified by adding the line

Let the errors of Y1 and Y3 correlate

In the input, the correlated error may be specified by adding PS=SY on the MO line and adding the line

```
FR PS(3,1)
```

This gives the following result

PSI	Y1	Y2	Y3	Y4
Y1	1.382			
	(0.071)			
	19.532			
Y2	- -	3.826		
		(0.196)		
		19.532		

Y3	0.474	- -	3.197
	(0.419)		(0.326)
	1.133		9.802
Y4	- -	- -	- -
			2.523
			(0.129)
			19.532

The  $t$ -value for  $(\psi)_{33}$  - hat is different from the other  $t$ -values because the equation for  $y_3$  is no longer a regression equation.

One can alter the data in any way as long as the sample covariance matrix is positive definite. The  $t$ -values will still be 19.532 for all error variances that correspond to regression equations.

### References

Jöreskog & Sörbom (1996). *LISREL 8 User's Reference Guide*. Chicago: Scientific Software International. (378 pages)

Jöreskog & Sörbom (1996). *LISREL 8: Structural Equation Modeling with the SIMPLIS Command Language*. Chicago: Scientific Software International. (226 pages)