Title

stata.com

Example 24 — Reliability

Description Remarks and examples Also see

Description

Below we demonstrate `sem`'s `reliability()` option with the following data:

```
. use https://www.stata-press.com/data/r18/sem_rel
(measurement error with known reliabilities)

. summarize

    Variable |        Obs        Mean    Std. dev.       Min        Max
-------------+---------------------------------------------------------
           y |      1,234     701.081    71.79378        487        943
          x1 |      1,234     100.278     14.1552         51        149
          x2 |      1,234    100.2066    14.50912         55        150

. notes

_dta:
  1.  Fictional data.
  2.  Variables x1 and x2 each contain a test score designed to measure X.  The
      test is scored to have mean 100.
  3.  Variables x1 and x2 are both known to have reliability 0.5.
  4.  Variable y is the outcome, believed to be related to X.
```

See [SEM] **sem and gsem option reliability( )** for background.

Remarks and examples

stata.com

Remarks are presented under the following headings:

Baseline model (reliability ignored)
Model with reliability
Model with two measurement variables and reliability

Baseline model (reliability ignored)

```
. sem (y <- x1)
Endogenous variables
  Observed: y
Exogenous variables
  Observed: x1
Fitting target model:
Iteration 0:  Log likelihood = -11629.745
Iteration 1:  Log likelihood = -11629.745
Structural equation model                          Number of obs = 1,234
Estimation method: ml
Log likelihood = -11629.745
```

	Coefficient	OIM std. err.	z	P>|z|	[95% conf.	interval]
Structural						
y						
x1	3.54976	.1031254	34.42	0.000	3.347637	3.751882
_cons	345.1184	10.44365	33.05	0.000	324.6492	365.5876
var(e.y)	2627.401	105.7752			2428.053	2843.115

```
LR test of model vs. saturated: chi2(0) = 0.00              Prob > chi2 = .
```

Notes:

1. In these data, variable x1 is measured with error.
2. If we ignore that, we obtain a path coefficient for y<-x1 of 3.55.
3. We also ran this model for y<-x2. We obtained a path coefficient of 3.48.

Model with reliability

```
. sem (x1<-X) (y<-X), reliability(x1 .5)
Endogenous variables
  Measurement: x1 y
Exogenous variables
  Latent: X
Fitting target model:
Iteration 0:  Log likelihood = -11745.845
Iteration 1:  Log likelihood = -11661.626
Iteration 2:  Log likelihood = -11631.469
Iteration 3:  Log likelihood = -11629.755
Iteration 4:  Log likelihood = -11629.745
Iteration 5:  Log likelihood = -11629.745
Structural equation model                          Number of obs = 1,234
Estimation method: ml
Log likelihood = -11629.745
 ( 1)  [x1]X = 1
 ( 2)  [/]var(e.x1) = 100.1036
```

	Coefficient	OIM std. err.	z	P>|z|	[95% conf. interval]	
Measurement						
x1						
X	1	(constrained)				
_cons	100.278	.4027933	248.96	0.000	99.4885	101.0674
y						
X	7.09952	.352463	20.14	0.000	6.408705	7.790335
_cons	701.081	2.042929	343.17	0.000	697.077	705.0851
var(e.x1)	100.1036	(constrained)				
var(e.y)	104.631	207.3381			2.152334	5086.411
var(X)	100.1036	8.060038			85.48963	117.2157

```
LR test of model vs. saturated: chi2(0) = 0.00                 Prob > chi2 = .
```

Notes:

1. We wish to estimate the effect of `y<-x1` when `x1` is measured with error (0.50 reliability). To do that, we introduce latent variable `X` and write our model as `(x1<-X) (y<-X)`.
2. When we ignored the measurement error of `x1`, we obtained a path coefficient for `y<-x1` of 3.55. Taking into account the measurement error, we obtain a coefficient of 7.1.

Model with two measurement variables and reliability

```
. sem (x1 x2<-X) (y<-X), reliability(x1 .5  x2 .5)

Endogenous variables
  Measurement: x1 x2 y

Exogenous variables
  Latent: X

Fitting target model:
Iteration 0:  Log likelihood = -16258.636
Iteration 1:  Log likelihood = -16258.401
Iteration 2:  Log likelihood =   -16258.4

Structural equation model                             Number of obs = 1,234
Estimation method: ml

Log likelihood = -16258.4

 ( 1)  [x1]X = 1
 ( 2)  [/]var(e.x1) = 100.1036
 ( 3)  [/]var(e.x2) = 105.1719

------------------------------------------------------------------------------
             |                 OIM
             | Coefficient  std. err.      z    P>|z|     [95% conf. interval]
-------------+----------------------------------------------------------------
Measurement  |
  x1         |
           X |          1  (constrained)
       _cons |    100.278   .4037851   248.34   0.000     99.48655    101.0694
  -----------+----------------------------------------------------------------
  x2         |
           X |   1.030101   .0417346    24.68   0.000     .9483029      1.1119
       _cons |   100.2066   .4149165   241.51   0.000     99.39342    101.0199
  -----------+----------------------------------------------------------------
  y          |
           X |   7.031299   .2484176    28.30   0.000     6.544409    7.518188
       _cons |    701.081   2.042928   343.17   0.000      697.077    705.0851
-------------+----------------------------------------------------------------
   var(e.x1) |   100.1036  (constrained)
   var(e.x2) |   105.1719  (constrained)
    var(e.y) |    152.329     105.26                      39.31868    590.1553
      var(X) |   101.0907   7.343656                      87.67509    116.5591
------------------------------------------------------------------------------
LR test of model vs. saturated: chi2(2) = 0.59            Prob > chi2 = 0.7430
```

Notes:

1. We wish to estimate the effect of `y<-X`. We have two measures of `X`—`x1` and `x2`—both measured with error (0.50 reliability).
2. In the previous section, we used just `x1`. We obtained path coefficient 7.1 with standard error 0.4. Using both `x1` and `x2`, we obtain path coefficient 7.0 and standard error 0.2.
3. We at StataCorp created these fictional data. The true coefficient is 7.

Also see

[SEM] **sem and gsem option reliability( )** — Fraction of variance not due to measurement error

[SEM] **Example 1** — Single-factor measurement model

For suggested citations, see the FAQ on citing Stata documentation.