

**Methods and formulas for sem** — Methods and formulas for sem[Description](#)[Remarks and examples](#)[References](#)[Also see](#)

## Description

The methods and formulas for the `sem` commands are presented below.

## Remarks and examples

stata.com

Remarks are presented under the following headings:

- Variable notation*
- Model and parameterization*
- Summary data*
- Maximum likelihood*
- Weighted least squares*
- Groups*
- Fitted parameters*
- Satorra–Bentler variance estimation*
- Standardized parameters*
- Reliability*
- Postestimation*
  - Model framework*
  - Goodness of fit*
  - Group goodness of fit*
  - Equation-level goodness of fit*
  - Wald tests*
  - Score tests*
  - Residuals*
  - Testing standardized parameters*
  - Stability of nonrecursive systems*
  - Direct, indirect, and total effects*
  - Predictions*

## Variable notation

We will use the following convention to keep track of the five variable types recognized by the `sem` estimation command:

1. Observed endogenous variables are denoted  $y$ .
2. Observed exogenous variables are denoted  $x$ .
3. Latent endogenous variables are denoted  $\eta$ .
4. Latent exogenous variables are denoted  $\xi$ .
5. Error variables are denoted with prefix  $e$ . on the associated endogenous variable.
  - a. Error variables for observed endogenous are denoted  $e.y$ .
  - b. Error variables for latent endogenous are denoted  $e.\eta$ .

In any given analysis, there are typically several variables of each type. Vectors of the four main variable types are denoted  $\mathbf{y}$ ,  $\mathbf{x}$ ,  $\boldsymbol{\eta}$ , and  $\boldsymbol{\xi}$ . The vector of all endogenous variables is

$$\mathbf{Y} = \begin{pmatrix} \mathbf{y} \\ \boldsymbol{\eta} \end{pmatrix}$$

The vector of all exogenous variables is

$$\mathbf{X} = \begin{pmatrix} \mathbf{x} \\ \boldsymbol{\xi} \end{pmatrix}$$

The vector of all error variables is

$$\boldsymbol{\zeta} = \begin{pmatrix} e.\mathbf{y} \\ e.\boldsymbol{\eta} \end{pmatrix}$$

## Model and parameterization

sem can fit models of the form

$$\mathbf{Y} = \mathbf{B}\mathbf{Y} + \boldsymbol{\Gamma}\mathbf{X} + \boldsymbol{\alpha} + \boldsymbol{\zeta}$$

where  $\mathbf{B} = [\beta_{ij}]$  is the matrix of coefficients on endogenous variables predicting other endogenous variables,  $\boldsymbol{\Gamma} = [\gamma_{ij}]$  is the matrix of coefficients on exogenous variables,  $\boldsymbol{\alpha} = [\alpha_i]$  is the vector of intercepts for the endogenous variables, and  $\boldsymbol{\zeta}$  is assumed to have mean 0 and

$$\text{Cov}(\mathbf{X}, \boldsymbol{\zeta}) = \mathbf{0}$$

Let

$$\boldsymbol{\kappa} = [\kappa_j] = E(\mathbf{X})$$

$$\boldsymbol{\Phi} = [\phi_{ij}] = \text{Var}(\mathbf{X})$$

$$\boldsymbol{\Psi} = [\psi_{ij}] = \text{Var}(\boldsymbol{\zeta})$$

Then the mean vector of the endogenous variables is

$$\boldsymbol{\mu}_Y = E(\mathbf{Y}) = (\mathbf{I} - \mathbf{B})^{-1}(\boldsymbol{\Gamma}\boldsymbol{\kappa} + \boldsymbol{\alpha})$$

the variance matrix of the endogenous variables is

$$\boldsymbol{\Sigma}_{YY} = \text{Var}(\mathbf{Y}) = (\mathbf{I} - \mathbf{B})^{-1}(\boldsymbol{\Gamma}\boldsymbol{\Phi}\boldsymbol{\Gamma}' + \boldsymbol{\Psi})\{(\mathbf{I} - \mathbf{B})^{-1}\}'$$

and the covariance matrix between the endogenous variables and the exogenous variables is

$$\boldsymbol{\Sigma}_{YX} = \text{Cov}(\mathbf{Y}, \mathbf{X}) = (\mathbf{I} - \mathbf{B})^{-1}\boldsymbol{\Gamma}\boldsymbol{\Phi}$$

Let  $\mathbf{Z}$  be the vector of all variables:

$$\mathbf{Z} = \begin{pmatrix} \mathbf{Y} \\ \mathbf{X} \end{pmatrix}$$

Then its mean vector is

$$\boldsymbol{\mu} = E(\mathbf{Z}) = \begin{pmatrix} \boldsymbol{\mu}_Y \\ \boldsymbol{\kappa} \end{pmatrix}$$

and its variance matrix is

$$\boldsymbol{\Sigma} = \text{Var}(\mathbf{Z}) = \begin{pmatrix} \boldsymbol{\Sigma}_{YY} & \boldsymbol{\Sigma}_{YX} \\ \boldsymbol{\Sigma}'_{YX} & \boldsymbol{\Phi} \end{pmatrix}$$

## Summary data

Let  $\mathbf{z}_t$  be the vector of all observed variables for the  $t$ th observation,

$$\mathbf{z}_t = \begin{pmatrix} \mathbf{y}_t \\ \mathbf{x}_t \end{pmatrix}$$

and let  $w_t$  be the corresponding weight value, where  $t = 1, \dots, N$ . If no weights were specified, then  $w_t = 1$ . Let  $w_\cdot$  be the sum of the weights; then the sample mean vector is

$$\bar{\mathbf{z}} = \frac{1}{w_\cdot} \sum_{t=1}^N w_t \mathbf{z}_t$$

and the sample variance matrix is

$$\mathbf{S} = \frac{1}{w_\cdot - 1} \sum_{t=1}^N w_t (\mathbf{z}_t - \bar{\mathbf{z}})(\mathbf{z}_t - \bar{\mathbf{z}})'$$

## Maximum likelihood

Let  $\boldsymbol{\theta}$  be the vector of unique model parameters, such as

$$\boldsymbol{\theta} = \begin{pmatrix} \text{vec}(\mathbf{B}) \\ \text{vec}(\boldsymbol{\Gamma}) \\ \text{vech}(\boldsymbol{\Psi}) \\ \text{vech}(\boldsymbol{\Phi}) \\ \boldsymbol{\alpha} \\ \boldsymbol{\kappa} \end{pmatrix}$$

Then under the assumption of the multivariate normal distribution, the overall log likelihood for  $\boldsymbol{\theta}$  is

$$\log L(\boldsymbol{\theta}) = -\frac{w_\cdot}{2} [k \log(2\pi) + \log \{\det(\boldsymbol{\Sigma}_o)\} + \text{tr}(\mathbf{D}\boldsymbol{\Sigma}_o^{-1})]$$

where  $k$  is the number of observed variables,  $\boldsymbol{\Sigma}_o$  is the submatrix of  $\boldsymbol{\Sigma}$  corresponding to the observed variables, and

$$\mathbf{D} = f\mathbf{S} + (\bar{\mathbf{z}} - \boldsymbol{\mu}_o)(\bar{\mathbf{z}} - \boldsymbol{\mu}_o)'$$

where

$$f = \begin{cases} 1, & \text{if nm1 is specified} \\ \frac{w_\cdot - 1}{w_\cdot}, & \text{otherwise} \end{cases}$$

and  $\boldsymbol{\mu}_o$  is the subvector of  $\boldsymbol{\mu}$  corresponding to the observed variables.

For the BHHH optimization technique and when computing observation-level scores, the log likelihood for  $\boldsymbol{\theta}$  is computed as

$$\log L(\boldsymbol{\theta}) = -\sum_{t=1}^N \frac{w_t}{2} [k \log(2\pi) + \log \{\det(\boldsymbol{\Sigma}_o)\} + (\mathbf{z}_t - \boldsymbol{\mu}_o)' \boldsymbol{\Sigma}_o^{-1} (\mathbf{z}_t - \boldsymbol{\mu}_o)]$$

and the nm1 option is ignored.

When `method(mlmv)` is specified, `sem` groups the data according to missing-value patterns. Each missing-value pattern will have its own summary data:  $k$ ,  $\bar{\mathbf{z}}$ , and  $\mathbf{S}$ . The log likelihood for a missing-value pattern is computed using this summary data and the corresponding elements of  $\boldsymbol{\mu}_o$  and  $\boldsymbol{\Sigma}_o$ . The overall log likelihood is computed by summing the log-likelihood values from each missing-value pattern.

## Weighted least squares

Let  $\mathbf{v}$  be the vector of unique sample moments of the observed variables and  $\boldsymbol{\tau}$  be the corresponding vector of population moments. Then

$$\mathbf{v} = \begin{bmatrix} \bar{\mathbf{z}} \\ \text{vech}(f\mathbf{S}) \end{bmatrix}$$

and

$$\boldsymbol{\tau} = \begin{bmatrix} \boldsymbol{\mu}_o \\ \text{vech}(\boldsymbol{\Sigma}_o) \end{bmatrix}$$

The weighted least-squares (WLS) criterion function to minimize is the quadratic form

$$F_{\text{WLS}}(\boldsymbol{\theta}) = (\mathbf{v} - \boldsymbol{\tau})' \mathbf{W}^{-1} (\mathbf{v} - \boldsymbol{\tau})$$

where  $\mathbf{W}$  is the least-squares weight matrix. For unweighted least squares (ULS), the weight matrix is the identity matrix  $\mathbf{W} = \mathbf{I}$ . Other weight matrices are mentioned in [Bollen \(1989\)](#).

The weight matrix implemented in `sem` is an estimate of the asymptotic covariance matrix of  $\mathbf{v}$ . This weight matrix is derived without any distributional assumptions and is often referred to as derived from an arbitrary distribution function or is asymptotic distribution free (ADF), thus the option `method(adf)`.

## Groups

When the `group()` option is specified, each group has its own summary data and model parameters. The entire collection of model parameters is

$$\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \\ \vdots \\ \boldsymbol{\theta}_G \end{pmatrix}$$

where  $G$  is the number of groups. The group-level criterion values are combined to produce an overall criterion value.

For `method(ml)` and `method(mlmv)`, the overall log likelihood is

$$\log L(\boldsymbol{\theta}) = \sum_{g=1}^G \log L(\boldsymbol{\theta}_g)$$

For `method(adf)`, the overall criterion function to minimize is

$$F_{\text{WLS}}(\boldsymbol{\theta}) = \sum_{g=1}^G F_{\text{WLS}}(\boldsymbol{\theta}_g)$$

## Fitted parameters

sem fits the specified model by maximizing the log likelihood or minimizing the WLS criterion. If  $\theta$  is the vector of model parameters, then the fitted parameter vector is denoted by  $\hat{\theta}$ , and similarly for  $\hat{B}$ ,  $\hat{\Gamma}$ ,  $\hat{\Psi}$ ,  $\hat{\Phi}$ ,  $\hat{\alpha}$ ,  $\hat{\kappa}$ ,  $\hat{\Sigma}$ ,  $\hat{\mu}$ , and their individual elements.

## Satorra–Bentler variance estimation

When `vce(sbentler)` is specified, the variance of  $\hat{\theta}$  is estimated as

$$\widehat{V}_{sb}(\hat{\theta}) = \frac{1}{w} (\Delta' H \Delta)^{-1} \Delta' H W H \Delta (\Delta' H \Delta)^{-1}$$

where  $W$  is the asymptotic covariance matrix of  $v$  used in the WLS criterion function, and  $\Delta$  is the partial derivative matrix of  $\tau$  with respect to  $\theta$  and uses the maximum likelihood fitted values  $\hat{\theta}$ .  $H$  is computed as

$$H = \begin{pmatrix} S^{-1} & \mathbf{o} \\ \mathbf{o} & A \end{pmatrix}$$

where

$$A = .5D'(S^{-1} \otimes S^{-1})D$$

and  $D$  is a duplication matrix. See [Satorra and Bentler \(1994, eq. 16.10\)](#).

## Standardized parameters

Let  $\hat{\sigma}_{ii}$  be the  $i$ th diagonal element of  $\hat{\Sigma}_{YY}$ . Then the standardized parameter estimates are

$$\begin{aligned} \tilde{\beta}_{ij} &= \hat{\beta}_{ij} \sqrt{\frac{\hat{\sigma}_{ii}}{\hat{\sigma}_{jj}}} \\ \tilde{\gamma}_{ij} &= \hat{\gamma}_{ij} \sqrt{\frac{\hat{\phi}_{ii}}{\hat{\sigma}_{jj}}} \\ \tilde{\psi}_{ij} &= \begin{cases} \hat{\psi}_{ij} / \sqrt{\hat{\psi}_{ii} \hat{\psi}_{jj}}, & \text{if } i \neq j \\ \hat{\psi}_{ii} / \hat{\sigma}_{ii}, & \text{if } i = j \end{cases} \\ \tilde{\phi}_{ij} &= \frac{\hat{\phi}_{ij}}{\sqrt{\hat{\phi}_{ii} \hat{\phi}_{jj}}} \\ \tilde{\alpha}_i &= \frac{\hat{\alpha}_i}{\sqrt{\hat{\sigma}_{ii}}} \\ \tilde{\kappa}_j &= \frac{\hat{\kappa}_j}{\sqrt{\hat{\phi}_{jj}}} \end{aligned}$$

The variance matrix of the standardized parameters is estimated using the delta method.

## Reliability

For an observed endogenous variable,  $y$ , the reliability may be specified as  $p$  or  $100 \times p\%$ . The variance of  $e.y$  is then constrained to  $(1 - p)$  times the observed variance of  $y$ .

## Postestimation

### Model framework

`estat framework` reports the fitted parameters in their individual matrix forms as introduced in *Fitted parameters*.

### Goodness of fit

`estat gof` reports the following goodness-of-fit statistics.

Let the degrees of freedom for the specified model be denoted by  $df_m$ . In addition to the specified model, `sem` fits saturated and baseline models corresponding to the observed variables in the specified model. The saturated model fits a full covariance matrix for the observed variables and has degrees of freedom

$$df_s = \binom{p + q + 1}{2} + p + q$$

where  $p$  is the number of observed endogenous variables and  $q$  is the number of observed exogenous variables in the model. The baseline model fits a reduced covariance matrix for the observed variables depending on the presence of endogenous variables. If there are no endogenous variables, all variables are uncorrelated in the baseline model; otherwise, only exogenous variables are correlated in the baseline model. The degrees of freedom for the baseline model is

$$df_b = \begin{cases} 2q, & \text{if } p = 0 \\ 2p + q + \binom{q + 1}{2}, & \text{if } p > 0 \end{cases}$$

For `method(m1)` and `method(mlmv)`, let the saturated log likelihood be denoted by  $\log L_s$  and the baseline log likelihood be denoted by  $\log L_b$ . The likelihood-ratio test of the baseline versus saturated models is computed as

$$\chi_{bs}^2 = 2(\log L_s - \log L_b)$$

with degrees of freedom  $df_{bs} = df_s - df_b$ . The likelihood-ratio test of the specified model versus the saturated model is computed as

$$\chi_{ms}^2 = 2\{\log L_s - \log L(\hat{\theta})\}$$

with degrees of freedom  $df_{ms} = df_s - df_m$ .

The Satorra–Bentler (1994) scaled  $\chi^2$  test is computed as

$$\chi_{sb,ms}^2 = \chi_{ms}^2 / c$$

where  $c = \text{trace}(UW/df_{ms})$ .  $U$  is computed as

$$U = H - H\Delta(\Delta'H\Delta)^{-1}\Delta'H$$

where  $\Delta$  and  $H$  are defined as in *Satorra–Bentler variance estimation* above.

The Satorra–Bentler scaled  $\chi^2$  test for the baseline versus saturated test,  $\chi_{sb,bs}^2$ , is defined as above, but  $\Delta$  is replaced for the specified model with  $\Delta$  for the baseline model.

For `method(adf)`, the saturated criterion value is 0,  $F_s = 0$ . Let the baseline criterion value be denoted by  $F_b$ . The  $\chi^2$  test of the baseline versus saturated models is computed as

$$\chi_{bs}^2 = NF_b$$

with degrees of freedom  $df_{bs} = df_s - df_b$ . The  $\chi^2$  test of the specified model versus the saturated model is computed as

$$\chi_{ms}^2 = NF_{WLS}(\hat{\theta})$$

with degrees of freedom  $df_{ms} = df_s - df_m$ .

The Akaike information criterion (Akaike 1974) is defined as

$$AIC = -2 \log L(\hat{\theta}) + 2df_m$$

The Bayesian information criterion (Schwarz 1978) is defined as

$$BIC = -2 \log L(\hat{\theta}) + Ndf_m$$

See [R] **IC note** for additional information on calculating and interpreting the BIC.

The overall coefficient of determination is computed as

$$CD = 1 - \frac{\det(\hat{\Psi})}{\det(\hat{\Sigma}_{YY})}$$

This value is also referred to as the overall  $R^2$  in `estat eqgof` (see [SEM] `estat eqgof`).

The root mean squared error of approximation (Browne and Cudeck 1993) is computed as

$$RMSEA = \left\{ \frac{(\chi_{ms}^2 - df_{ms})G}{(N - 1)df_{ms}} \right\}^{1/2}$$

The 90% confidence interval for RMSEA is

$$\left( \sqrt{\frac{G\lambda_L}{(N - 1)df_{ms}}}, \sqrt{\frac{G\lambda_U}{(N - 1)df_{ms}}} \right)$$

where  $\lambda_L$  and  $\lambda_U$  are the noncentrality parameters corresponding to a noncentral  $\chi^2$  distribution with  $df_{ms}$  degrees of freedom in which the noncentral  $\chi^2$  random variable has a cumulative distribution function equal to 0.95 and 0.05, respectively.

The Browne and Cudeck (1993)  $p$ -value for the test of close fit with null hypothesis

$$H_0: \text{RMSEA} \leq 0.05$$

is computed as

$$p = 1 - \Pr(\chi^2 < \chi_{\text{ms}}^2 | \lambda, \text{df}_{\text{ms}})$$

where  $\chi^2$  is distributed noncentral  $\chi^2$  with noncentrality parameter  $\lambda = (0.05)^2(N - 1)\text{df}_{\text{ms}}$  and  $\text{df}_{\text{ms}}$  degrees of freedom. This  $p$ -value is not computed when the `group()` option is specified.

The comparative fit index (Bentler 1990) is computed as

$$\text{CFI} = 1 - \left[ \frac{(\chi_{\text{ms}}^2 - \text{df}_{\text{ms}})}{\max\{(\chi_{\text{bs}}^2 - \text{df}_{\text{bs}}), (\chi_{\text{ms}}^2 - \text{df}_{\text{ms}})\}} \right]$$

The Tucker–Lewis index (Bentler 1990) is computed as

$$\text{TLI} = \frac{(\chi_{\text{bs}}^2/\text{df}_{\text{bs}}) - (\chi_{\text{ms}}^2/\text{df}_{\text{ms}})}{(\chi_{\text{bs}}^2/\text{df}_{\text{bs}}) - 1}$$

Let  $k$  be the number of observed variables in the model. The standardized root mean squared residual is computed as

$$\text{SRMR} = \left\{ \frac{2 \sum_{i=1}^k \sum_{j \leq i} r_{ij}^2}{k(k+1)G} \right\}^{1/2}$$

where  $r_{ij}$  is the standardized covariance residual

$$r_{ij} = \frac{s_{ij}}{\sqrt{s_{ii}s_{jj}}} - \frac{\hat{\sigma}_{ij}}{\sqrt{\hat{\sigma}_{ii}\hat{\sigma}_{jj}}}$$

These standardized residuals are not the same as those reported by `estat residuals`; see [Residuals](#) below.

The RMSEA, CFI, and TLI are also reported using Satorra–Bentler scaled statistics when the model is fit using `vce(sbentler)`. In the formulas above,  $\chi_{\text{ms}}^2$  and  $\chi_{\text{bs}}^2$  are replaced, respectively, with  $\chi_{\text{sb,ms}}^2$  and  $\chi_{\text{sb,bs}}^2$ .

## Group goodness of fit

`estat ggof` reports CD, SRMR, and model versus saturated  $\chi^2$  values for each group separately. The group-level formulas are the same as those computed for a single group analysis; see [Goodness of fit](#) above.

## Equation-level goodness of fit

`estat eqgof` reports goodness-of-fit statistics for each endogenous variable in the specified model. The coefficient of determination for the  $i$ th endogenous variable is computed as

$$R_i^2 = 1 - \frac{\hat{\psi}_{ii}}{\hat{\sigma}_{ii}}$$



The Bentler–Raykov (Bentler and Raykov 2000) squared multiple correlation for the  $i$ th endogenous variable is computed as

$$mc_i^2 = \frac{\widehat{\text{Cov}}(y_i, \hat{y}_i)}{\sqrt{\hat{\sigma}_{ii} \widehat{\text{Var}}(\hat{y}_i)}}$$

where  $\hat{\sigma}_{ii}$  is a diagonal element of  $\widehat{\Sigma}$ ,  $\widehat{\text{Var}}(\hat{y}_i)$  is a diagonal element of

$$\widehat{\text{Var}}(\hat{Y}) = (I - \widehat{B})^{-1} \widehat{\Gamma} \widehat{\Phi} \widehat{\Gamma}' \left\{ (I - \widehat{B})^{-1} \right\}' + \left\{ (I - \widehat{B})^{-1} - I \right\} \widehat{\Psi} \left\{ (I - \widehat{B})^{-1} - I \right\}'$$

and  $\widehat{\text{Cov}}(y_i, \hat{y}_i)$  is a diagonal element of

$$\widehat{\text{Cov}}(Y, \hat{Y}) = (I - \widehat{B})^{-1} \widehat{\Gamma} \widehat{\Phi} \widehat{\Gamma}' \left\{ (I - \widehat{B})^{-1} \right\}' + (I - \widehat{B})^{-1} \widehat{\Psi} \left\{ (I - \widehat{B})^{-1} - I \right\}'$$

### Wald tests

`estat eqtest` performs Wald tests on the coefficients for each endogenous equation in the model. `estat ginvariant` computes a Wald test of group invariance for each model parameter that is free to vary across all groups. See [R] [test](#).

### Score tests

`estat mindices` computes modification indices for each constrained parameter in the model, including paths and covariances that were not even part of the model specification. Modification indices are score tests, which are also known as Lagrange multiplier tests. `estat scoretests` performs a score test for each user-specified linear constraint. `estat ginvariant` performs a score test of group invariance for each model parameter that is constrained to be equal across all groups.

A score test compares a constrained model fit to the same model without one or more constraints. The score test is computed as

$$\chi^2 = g(\hat{\theta})' V(\hat{\theta}) g(\hat{\theta})$$

where  $\hat{\theta}$  is the fitted parameter vector from the constrained model,  $g(\theta)$  is the gradient vector function for the unconstrained model, and  $V(\theta)$  is the variance matrix function computed from the expected information matrix function for the unconstrained model. For `method(ml)` and `method(mlmv)`,

$$g(\theta) = \frac{\partial \log L^*(\theta)}{\partial \theta}$$

$$V(\theta) = \left[ E \left\{ - \frac{\partial^2 \log L^*(\theta)}{\partial \theta \partial \theta'} \right\} \right]^{-1}$$

where  $\log L^*(\theta)$  is the log-likelihood function for the unconstrained model. For `method(adf)`,

$$g(\theta) = - \frac{\partial F_{\text{WLS}}^*(\theta)}{\partial \theta}$$

$$V(\theta) = \left[ E \left\{ \frac{\partial^2 F_{\text{WLS}}^*(\theta)}{\partial \theta \partial \theta'} \right\} \right]^{-1}$$

where  $F^*(\theta)$  is the WLS criterion function for the unconstrained model.

The score test is computed as described in [Wooldridge \(2010\)](#) when `vce(robust)` or `vce(cluster clustvar)` is specified.

## Residuals

`estat residuals` reports raw, normalized, and standardized residuals for means and covariances of the observed variables.

The raw residual for the mean of the  $i$ th observed variable is

$$\bar{z}_i - \hat{\mu}_i$$

The raw residual for the covariance between the  $i$ th and  $j$ th observed variables is

$$S_{ij} - \hat{\sigma}_{ij}$$

The normalized residual for the mean of the  $i$ th observed variable is

$$\frac{\bar{z}_i - \hat{\mu}_i}{\sqrt{\widehat{\text{Var}}(\bar{z}_i)}}$$

where

$$\widehat{\text{Var}}(\bar{z}_i) = \begin{cases} \frac{S_{ii}}{N}, & \text{if the sample option is specified} \\ \frac{\hat{\sigma}_{ii}}{N}, & \text{otherwise} \end{cases}$$

The normalized residual for the covariance between the  $i$ th and  $j$ th observed variables is

$$\frac{S_{ij} - \hat{\sigma}_{ij}}{\sqrt{\widehat{\text{Var}}(S_{ij})}}$$

where

$$\widehat{\text{Var}}(S_{ij}) = \begin{cases} \frac{S_{ii}S_{jj} + S_{ij}^2}{N}, & \text{if the sample option is specified} \\ \frac{\hat{\sigma}_{ii}\hat{\sigma}_{jj} + \hat{\sigma}_{ij}^2}{N}, & \text{otherwise} \end{cases}$$

If the `nm1` option is specified, the denominator in the variance estimates is  $N - 1$  instead of  $N$ .

The standardized residual for the mean of the  $i$ th observed variable is

$$\frac{\bar{z}_i - \hat{\mu}_i}{\sqrt{\widehat{\text{Var}}(\bar{z}_i - \hat{\mu}_i)}}$$

where

$$\widehat{\text{Var}}(\bar{z}_i - \hat{\mu}_i) = \widehat{\text{Var}}(\bar{z}_i) - \widehat{\text{Var}}(\hat{\mu}_i)$$

and  $\widehat{\text{Var}}(\hat{\mu}_i)$  is computed using the delta method. Missing values are reported when the computed value of  $\widehat{\text{Var}}(\bar{z}_i)$  is less than  $\widehat{\text{Var}}(\hat{\mu}_i)$ . The standardized residual for the covariance between the  $i$ th and  $j$ th observed variables is

$$\frac{S_{ij} - \hat{\sigma}_{ij}}{\sqrt{\widehat{\text{Var}}(S_{ij} - \hat{\sigma}_{ij})}}$$

where

$$\widehat{\text{Var}}(S_{ij} - \widehat{\sigma}_{ij}) = \widehat{\text{Var}}(S_{ij}) - \widehat{\text{Var}}(\widehat{\sigma}_{ij})$$

and  $\widehat{\text{Var}}(\widehat{\sigma}_{ij})$  is computed using the delta method. Missing values are reported when the computed value of  $\widehat{\text{Var}}(S_{ij})$  is less than  $\widehat{\text{Var}}(\widehat{\sigma}_{ij})$ . The variances of the raw residuals used in the standardized residual calculations are derived in [Hausman \(1978\)](#).

## Testing standardized parameters

`estat stdize` provides access to tests on the standardized parameter estimates. `estat stdize` can be used as a prefix to `lincom` (see [\[R\] lincom](#)), `nlcom` (see [\[R\] nlcom](#)), `test` (see [\[R\] test](#)), and `testnl` (see [\[R\] testnl](#)).

## Stability of nonrecursive systems

`estat stable` reports a stability index for nonrecursive systems. The stability index is calculated as the maximum of the modulus of the eigenvalues of  $\mathbf{B}$ . The nonrecursive system is considered stable if the stability index is less than 1.

## Direct, indirect, and total effects

`estat teffects` reports direct, indirect, and total effects for the fitted model. The direct effects are

$$\mathbf{E}_d = [\widehat{\mathbf{B}} \quad \widehat{\mathbf{\Gamma}}]$$

the total effects are

$$\mathbf{E}_t = [(\mathbf{I} - \widehat{\mathbf{B}})^{-1} - \mathbf{I} \quad , \quad (\mathbf{I} - \widehat{\mathbf{B}})^{-1}\widehat{\mathbf{\Gamma}}]$$

and the indirect effects are  $\mathbf{E}_i = \mathbf{E}_t - \mathbf{E}_d$ . The standard errors of the effects are computed using the delta method.

Let  $\mathbf{D}$  be the diagonal matrix whose elements are the square roots of the diagonal elements of  $\widehat{\mathbf{\Sigma}}$ , and let  $\mathbf{D}_Y$  be the submatrix of  $\mathbf{D}$  associated with the endogenous variables. Then the standardized effects are

$$\widetilde{\mathbf{E}}_d = \mathbf{D}_Y^{-1} \mathbf{E}_d \mathbf{D}$$

$$\widetilde{\mathbf{E}}_i = \mathbf{D}_Y^{-1} \mathbf{E}_i \mathbf{D}$$

$$\widetilde{\mathbf{E}}_t = \mathbf{D}_Y^{-1} \mathbf{E}_t \mathbf{D}$$

## Predictions

`predict` computes factor scores and linear predictions.

Factor scores are computed with a linear regression by using the mean vector and variance matrix from the fitted model. For notational convenience, let

$$\mathbf{Z} = \begin{pmatrix} z \\ l \end{pmatrix}$$

where

$$z = \begin{pmatrix} y \\ x \end{pmatrix}$$

and

$$l = \begin{pmatrix} \eta \\ \xi \end{pmatrix}$$

The fitted mean of  $Z$  is

$$\hat{\mu}_Z = \begin{pmatrix} \hat{\mu}_z \\ \hat{\mu}_l \end{pmatrix}$$

and fitted variance of  $Z$  is

$$\hat{\Sigma}_Z = \begin{pmatrix} \hat{\Sigma}'_{zz} & \hat{\Sigma}'_{zl} \\ \hat{\Sigma}'_{zl} & \hat{\Sigma}'_{ll} \end{pmatrix}$$

The factor scores are computed as

$$\tilde{l} = \begin{pmatrix} \tilde{\eta} \\ \tilde{\xi} \end{pmatrix} = \hat{\Sigma}'_{zl} \hat{\Sigma}_{zz} \hat{\mu}_z + \hat{\mu}_l$$

The linear prediction for the endogenous variables in the  $t$ th observation is computed as

$$\hat{Y}_t = \hat{B} \tilde{Y}_t + \hat{\Gamma} \tilde{X}_t + \hat{\alpha}$$

where

$$\tilde{Y}_t = \begin{pmatrix} y_t \\ \tilde{\eta} \end{pmatrix}$$

and

$$\tilde{X}_t = \begin{pmatrix} x_t \\ \xi \end{pmatrix}$$

## References

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19: 716–723. <https://doi.org/10.1109/TAC.1974.1100705>.
- Bentler, P. M. 1990. Comparative fit indexes in structural models. *Psychological Bulletin* 107: 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>.
- Bentler, P. M., and T. Raykov. 2000. On measures of explained variance in nonrecursive structural equation models. *Journal of Applied Psychology* 85: 125–131. <https://doi.org/10.1037/0021-9010.85.1.125>.
- Bollen, K. A. 1989. *Structural Equations with Latent Variables*. New York: Wiley.
- Browne, M. W., and R. Cudeck. 1993. Alternative ways of assessing model fit. Reprinted in *Testing Structural Equation Models*, ed. K. A. Bollen and J. S. Long, pp. 136–162. Newbury Park, CA: Sage.
- Hausman, J. A. 1978. Specification tests in econometrics. *Econometrica* 46: 1251–1271. <https://doi.org/10.2307/1913827>.
- Satorra, A., and P. M. Bentler. 1994. Corrections to test statistics and standard errors in covariance structure analysis. In *Latent Variables Analysis: Applications for Developmental Research*, ed. A. von Eye and C. C. Clogg, 399–419. Thousand Oaks, CA: Sage.
- Schwarz, G. 1978. Estimating the dimension of a model. *Annals of Statistics* 6: 461–464. <https://doi.org/10.1214/aos/1176344136>.
- Wooldridge, J. M. 2010. *Econometric Analysis of Cross Section and Panel Data*. 2nd ed. Cambridge, MA: MIT Press.

## Also see

[SEM] [sem](#) — Structural equation model estimation command

Stata, Stata Press, and Mata are registered trademarks of StataCorp LLC. Stata and Stata Press are registered trademarks with the World Intellectual Property Organization of the United Nations. StataNow and NetCourseNow are trademarks of StataCorp LLC. Other brand and product names are registered trademarks or trademarks of their respective companies. Copyright © 1985–2023 StataCorp LLC, College Station, TX, USA. All rights reserved.



For suggested citations, see the [FAQ on citing Stata documentation](#).