Title

sem ssd options — Options for use with summary statistics data

Description Syntax Options Remarks and examples Also see

Description

Data are sometimes available in summary statistics form only. These summary statistics include means, standard deviations or variances, and correlations or covariances. These summary statistics can be used by `sem` in place of the underlying raw data.

Syntax

`sem` *paths* `...,` `...` *ssd_options*

ssd_options	Description
`select()`	alternative to `if` *exp* for SSD
`forcecorrelations`	allow groups and pooling of SSD correlations

Options

`select()` is an alternative to `if` *exp* when you are using summary statistics data (SSD). Where you might usually type

```
. sem ... if agegrp==1 | agegrp==3 | agegrp==5, ...
```

with SSD in memory, you type

```
. sem ..., ... select(1 3 5)
```

See [SEM] **sem option select()** and [SEM] **Intro 11**.

`forcecorrelations` tells `sem` that it may make calculations that would usually be considered suspicious with SSD that contain only a subset of means, variances (standard deviations), and covariances (correlations). Do not specify this option unless you appreciate the statistical issues that we are about to discuss. There are two cases where `forcecorrelations` is relevant.

In the first case, `sem` is unwilling to produce `group()` estimates if one or more (usually all) of the groups have correlations only defined. You can override that by specifying `forcecorrelations`, and `sem` will assume unit variances for the group or groups that have correlations only. Doing this is suspect unless you make `ginvariant()` all parameters that are dependent on covariances or unless you truly know that the variances are indeed 1.

In the second case, `sem` is unwilling to pool across groups unless you have provided means and covariances (or means and correlations and standard deviations or variances). Without that information, should the need for pooling arise, `sem` issues an error message. The `forcecorrelations` option specifies that `sem` ignore its rule and pool correlation matrices, treating correlations as if they were covariances when variances are not defined and treating means as if they were 0 when means are not defined. The only justification for making the calculation in this way is that variances truly are 1 and means truly are 0.

Understand that there is nothing wrong with using pure correlation data, or covariance data without the means, so long as you fit models for individual groups. Doing anything across groups basically requires that `sem` have the covariance and mean information.

Remarks and examples

stata.com

See [SEM] **Intro 11**.

Also see

[SEM] **sem** — Structural equation model estimation command

[SEM] **Intro 11** — Fitting models with summary statistics data (sem only)

[SEM] **ssd** — Making summary statistics data (sem only)

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. 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.