Estimation of means of latent variables: mean difference in verbal ability

Since a latent variable is unobservable, it does not have an intrinsic scale. Neither he origin nor the unit of measurement are defined. In a single population the origin is fixed by assuming that all observed variables are measured in deviations from their means and that the means of all latent variables are zero. The unit of measurement of each latent variable is usually fixed either by assuming that it is a standardized variable with variance 1 or by fixing a non-zero loading for a reference variable.

In multi-group studies, these restrictions can be relaxed somewhat by assuming that the latent variables are on the same scale in all groups. The common scale may be defined by assuming that the means of the latent variables are zero in one group and that the loadings of the observed variable on the latent variables are invariant over groups, with one loading for ach latent variable fixed for a reference variable. Under these assumptions it is possible to estimate the means and covariance matrices of the latent variables relative to this common scale. To illustrate this we will first consider a small example based on the same data used in previous examples, and then consider a larger example.

Using the variables and data in the table below, take READING5 and WRITING5 to be indicators of a latent variable Verbal5 (Verbal Ability at Grade 5) and estimate the mean difference in Verbal5 between groups.

Table 1: Covariance matrices for SAT Verbal and Math sections

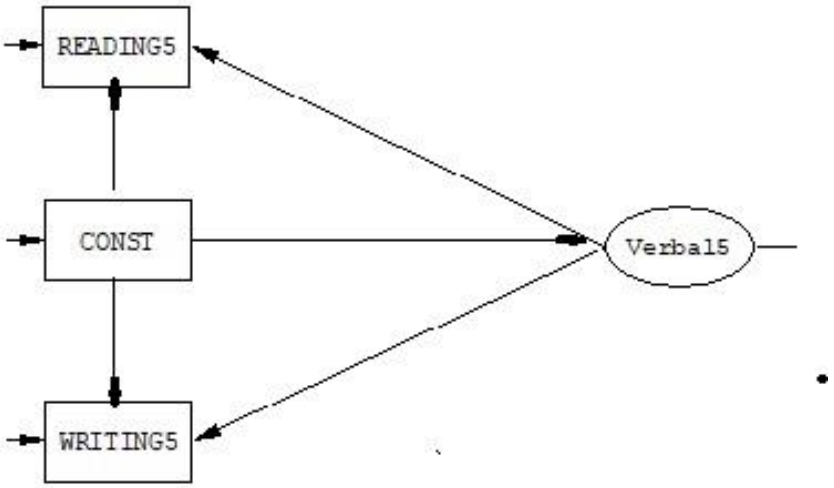
Boys Academic (*N* = 373)				
STEP Reading, Grade 5	281.349			
STEP Writing, Grade 5	184.219	182.821		
STEP Reading, Grade 7	216.739	171.699	283.289	
STEP Writing, Grade 7	198.376	153.201	208.837	246.069
Means	262.236	258.788	275.630	269.075
Boys Non-academic (*N* = 249)				
STEP Reading, Grade 5	174.485			
STEP Writing, Grade 5	134.468	161.869		
STEP Reading, Grade 7	129.840	118.836	228.449	
STEP Writing, Grade 7	102.194	97.767	136.058	180.460
Means	248.675	246.896	258.546	253.349

The measurement model for READING5 and WRITING5 is:

$$\text{READING5} = v_1 + l_1\text{Verbal5} + \text{D1}$$
$$\text{WRITING5} = v_2 + l_2\text{Verbal5} + \text{D2}$$

Note that these relationships now have intercept terms v_1 and v_2 . The scale for Verbal5 is fixed by assuming that the mean is zero in Group BA and that λ_1 =1. It is further assumed that $v_1, v_2, \lambda_1, \lambda_2$ are invariant over groups. We can then estimate the mean of Verbal5 of Group NBA as well as all the other parameters. The mean of Verbal5 is interpreted as the mean difference in verbal ability between the groups.

A path diagram is shown below. Note that the path from CONST to READING5 corresponds to v_1, the path from CONST to WRITING5 corresponds to v_2, and the path from CONST to Verbal5 corresponds to the mean of Verbal5.



The input file for this model is **EX13A.SPL** (see the **Simplis Examples** folder):

```
Group BA: STEP Reading and Writing, Grades 5 and 7
Observed Variables: READING5 WRITING5 READING7 WRITING7
Covariance Matrix from File EX12.DAT
Means from File EX12.DAT
Sample Size: 373
Latent Variable: Verbal5
Relationships:
   READING5 = CONST + 1*Verbal5
   WRITING5 = CONST + (1)*Verbal5

Group BNA: STEP Reading and Writing, Grades 5 and 7
Covariance Matrix from File EX12.DAT
Means from File EX12.DAT
Sample Size: 249
Relationship: Verbal5 = CONST
Set the Error Variances of READING5 - WRITING5 free
Set the Variance of Verbal5 free
Path Diagram
End of Problem
```

The estimated group means, variances and covariances of the two latent variables are shown in the table below. It can be seen that the non-academic group mean is below the mean of the academic group both in Grade 5 and in Grade 7. Furthermore, the non-academic group has a less favorable development from Grade 5 to Grade 7; its mean has decreased compared to the academic group.

Table: Estimated means and covariance matrices of Verbal5 and Verbal7

Boys Academic (N = 373)		
	Verbal5	Verbal7
Verbal5	220.06	
Verbal7	212.11	233.59
Means	0.00	0.00
Boys Non-academic (N = 249)		
	Verbal5	Verbal7
Verbal5	156.34	
Verbal7	126.96	153.73
Means	-13.80	-17.31

One way to describe the group differences graphically is to draw 95 percent confidence regions. These confidence regions are ellipses, such that 95 percent of the population is located within the ellipse. The ellipses can differ in origin, shape and orientation. For any score on Verbal5, one can see the most likely ranges of Verbal7for each group. When this is done for this example, the slope of the regression lines of Verbal7on Verbal5are similar. This leads to the problem of estimating these regression lines and testing whether they are the same, a topic that will be considered in a next example.