Classical Test Theory

Harry Ganzeboom

RESMA Data Analysis & Report #3

February 12 2009

Reliability

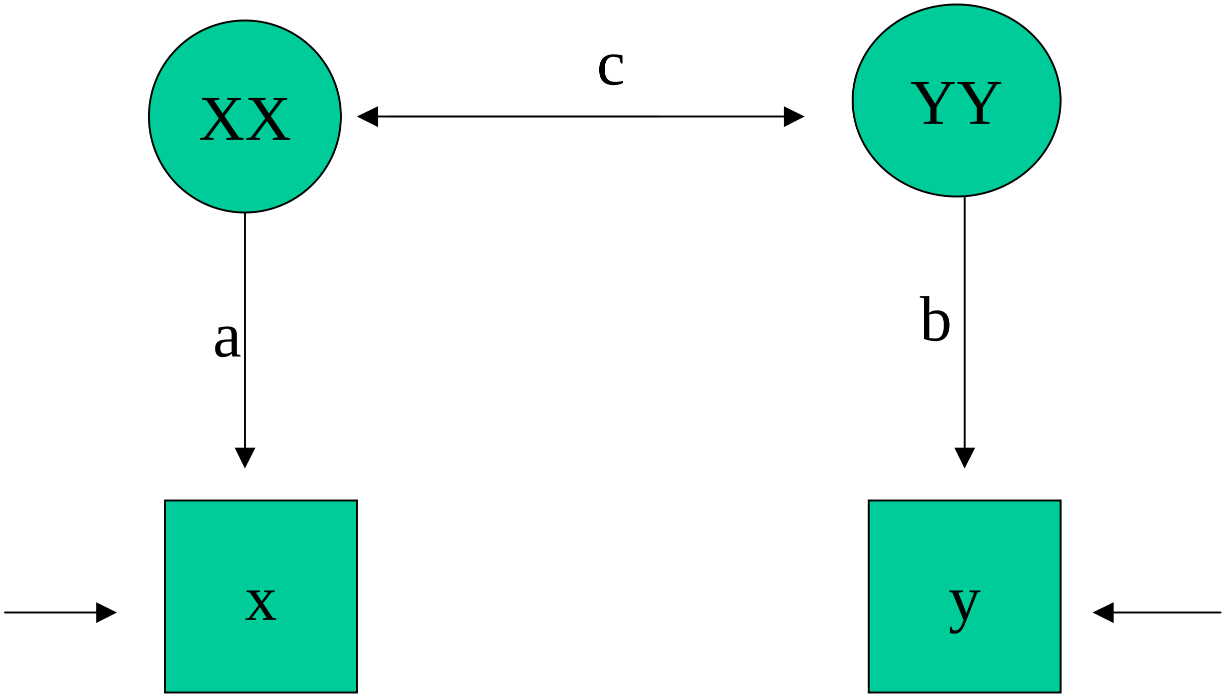
- Reliability = stability: a measure is said to be reliable if it has the same results if the object did not change.
- Disregard uniform (systematic) bias: reliability = correlation between two measurements.
- Reliability refers to <u>random</u> measurement error:
  - Expected(error) = zero
  - No correlation of errors with true score or any other score.

True scores & reliability designs

- Observed score = true score + random error.
- We can estimate reliability by using the measure (at least) twice.
  - Parallel: Simply repeat the same questions.
  - Test-retest: repeat the questions with some interval (that is long enough to forget the errors in the previous respons).
  - Alternative forms: ask the questions is a different format that will make respondents forget their errors instantaneously.
  - Split-half: test-retest using two halves of the indicators.
  - Internal consistency: parallel measures using all possible halves of the indicators.

The general model

Model with two latent variables: r(xy)=a*b*c

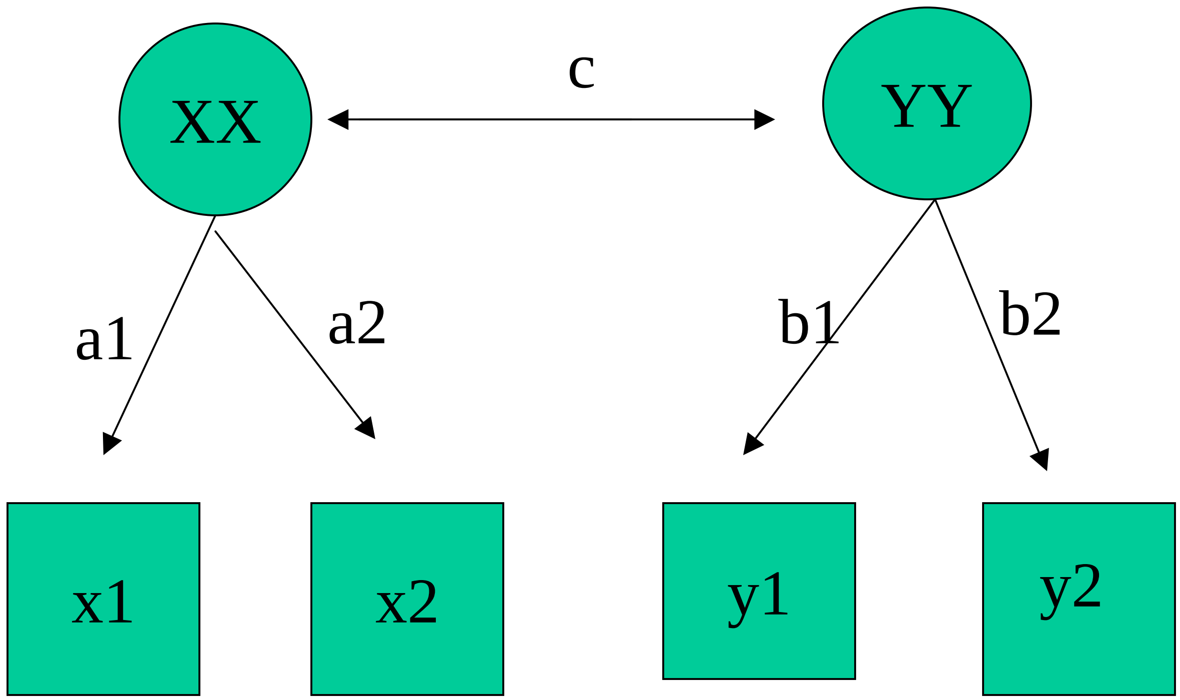


Identification

- The general model is not identified: 3 parameters with 1 equation.
- Add another measure for XX en YY: 5 parameters with 6 equations: overdetermined = identified!
- Note that a crucial assumption is that the error terms are uncorrelated, both within en between latent variables!
- Note that the model with two measures identifies both true score reliability and true score stability!

Model with double indicators

Model with two latent variables:



Model implications

- R(x1,x2) = a1*a2
- R(x1,y1) = a1*c*b1
- R(x1,y2) = a1*c*b2
- R(x2,y1) = a2*c*b1
- R(x2,y2) = a2*c*b2
- R(x1,y2) = b1*b2
- Six equations with five unknows: overdetermined = identified!

Common Factor Analysis

- Common Factor Analysis (SPSS: Principal Axis Factoring = PAF) can estimate this model from data.
- Let’s look at some simple simulations.
- Note that the model does not use a reliability coefficient – it builds it into the model.

Principal Component Analysis

- PCA is the default option in SPSS Factor, but in fact it is not (common) factor analysis at all.
- PCA: how can I create a sum-score from a set of indicators that has maximal variance.
- VAR(a+b)=VAR(a)+VAR(b)+2*COVAR(a,b).
- PCA: largest weights for variables with strong correlations.
- While PCA and PAF ask very different questions, the answers are likely to be very similar.

PAF and PCA

- PCA leads directly to component scores, PAF leads to estimated correlations in a latent variables model.
- PCA is computationally stable, PAF may run into problems if the model does not apply.
- PAF fits our common sense measurement models very well, PCA is harder to understand. In particular (oblique) rotation is hard to interpret in PCA, but easy to understand in PAF.

PAF, PCA and Cronbach's alpha

- Cronbach alpha takes a middling ground between PAF and PCA.
  - PCA: Cronbach's alpha is optimal when variation of the sum-score is maximal.
  - PAF: alpha estimates reliability when all indicators are equally correlated (= have same amount of random error). Loading = $\sqrt{\alpha}$.

Systematic error

- Not all error can be assumed to be random!
- Systematic error = correlated error = when error arise in similar indicators in the same way!
- Latent variable models (but not in SPSS) can estimate (and control) this type of error, provided the error is repeated!
- MTMM: multiple measures of multiple constructs (traits).

LISREL

- LISREL = Linear Structural Relations
- Karl Jöreskog & Dag Sörbom
- Related programs: AMOS, EQS, Mplus
- Lisrel is little else but a computer program to solve an (overdetermined) set of linear equations.