UiO : Universitetet i Oslo

CANDIDATE

184110

TEST

MAE4011 1 Principles of Measurement

Subject code	MAE4011
Evaluation type	Individuell skriftlig prøve
Test opening time	20.12.2022 09:00
End time	20.12.2022 13:00
Grade deadline	--
PDF created	26.12.2022 19:33
Created by	Tony Clifford Austin Tan

9 SR1H22

A scale to measure depression severity was developed and data were collected from a large group of students, along with the scores of an existing scale for satisfaction with life.

You observed the following covariance matrix for the scores of the two scales, where X denotes the depression severity scale scores and Y denotes the satisfaction with life scale scores:

$$\boldsymbol{\Sigma} = \begin{pmatrix} 10 & -7 \\ -7 & 10 \end{pmatrix}.$$

Based on these observations, how would you characterize the relationship between depression severity and satisfaction with life?

State the assumptions made in the interpretations of the relationship.

Fill in your answer here

For the interpretation, i will first look at the correlation between the depression severity and satisfaction with life

$Corr = \frac{Cov(X,Y)}{\delta X \delta Y}$

$= \frac{-7}{10}$

-0.70

Which indicates that the variables tend to differ the opposite way, e.g: When scoring high on satisfaction with life (that you are satisfied with life) you would score low on depression severity. That they tend to opposite.

Correlation linear relationship

Words: 63

Answered.

11 SR3H22

X and Y are two random variables where $\mathbf{Var}(X) = 2$, $\mathbf{Var}(Y) = 3$ and $\mathbf{Cov}(X, Y) = 1$.

1. Calculate $\mathbf{Var}(Z)$, where $Z = X - Y$. Show your work.
2. Calculate $\mathbf{Var}(U)$, where $U = X + 2Y$. Show your work.

Fill in your answer here

1.
$Var(Z) = Var(X) + Var(Y) - 2Cov(X, Y)$
=2+3-2*1
=3
Var(Z)=3

2.
$Var(U) = Var(X) + 2(Y) + 2Cov(X, Y)$
=2+2^2*3+2*1
=2+12+3
=17

Words: 10

Answered.

12 SR4H22

Let m be the number of items on a test. For a five-item test, the common factor loading λ was 1 and the variance of the sum score Y was 10. Compute coefficient alpha

$\alpha = m\frac{\lambda^2}{\mathrm{Var}(Y)}$

and interpret it. State the assumptions underlying the interpretation.

Fill in your answer here

assuming a single factor model used.
Since the the factor loadings was the same, and the common factor loading was 1, the coefficient alpha is the reliability of the sum score.

$\alpha = 5\frac{1}{10}$

The estimated coefficient alpha/cronbachs alpha was 0.50
which is interpreted as a low/moderate estimate.
Showing the internal consistency is moderate

Words: 53

Answered.

13 SR5H22

The *Standards for Educational and Psychological Testing* (2014) state that it is useful to consider ways in which the test scores can be influenced by either (1) too much or (2) too little.

A three-domain test is administered for the purpose of measuring Norwegian 15-year-olds' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges. The test is a low-stakes test for the respondents since individual assessment is not of interest.

Provide **one example** of a way in which the test-scores might be influenced by too much, and **one example** of how the test-scores might be influenced by too little.

Fill in your answer here

1. to much content coverage
2. to little time, not enough time to do the test.

Words: 16

Answered.

14 SR6H22

For two tests of reading comprehension, X and Y, the linear equating function was estimated to be $\mathbf{eq}(Y) = \mathbf{1.2}X + \mathbf{6}$. The cut score for passing test Y was determined to be 30.
Give the cut score for pass in terms of the test X scores, based on the estimated equating function. Present and explain how the result was obtained.

Fill in your answer here

eq(Y)=1.2X+6
Y=30
30=1,2X+6
24=1.2X, 20*1.2=24
The cut score pass in terms of the test X scores should be 20.
The cut score was found by estimating what X would be when y was 30.

Words: 34

Answered.

15 SR7H22

Item scores on a test of mathematics and a test of interest in mathematics were given to the same group of students. A two-factor model with correlated factors (one factor measured by the mathematics test items and the other by the interest in mathematics items) was estimated, yielding the model fit indices:

GFI	0.95
RMSEA	0.05
SRMR	0.06

The correlation between the sum scores of the respective tests was 0.2 while the estimated factor correlation was 0.5. Explain why there is a difference in the factor correlation and the sum score correlation in this context.

Fill in your answer here

The GFI (godness-of fit) was estimated to be 0.95, which is "good".

The RMSEA = 0.05, which indicates a acceptable fit.

SRMR is ok, but should be 0.

The difference could be that it was fitted a Two factor model, and that it should be a single factor model, that mathematics and interest in matemathics is same construct, and is unidimensional.

the parameters of the model fit is acceptable, but could be better, and that can be an indicator that a singlefactor model is a better fit.

Words: 87

Answered.

16 SR8H22

A bifactor model with one general factor and two subfactors (all factors independent) was estimated for an Norwegian test with two subdomains (reading and writing), yielding the following factor loading estimates:

Item	General	Reading	Writing
1	3	0.5	0
2	1	0.5	0
3	2	1	0
4	1	0	1
5	1	0	0.5
6	1	0	0.5

The model fit was judged to be satisfactory.

In a previous study, the sum score was used. Based on the estimated factor loadings, would you recommend doing this? Justify your answer.

Fill in your answer here

According to the estimated factor loadings it doesn't seem the right model that is estimated, since the general factor has the highest factor loadings, which is indicative that the items highly measures one attribute.

It seems that reading and writing in some sense could fall under one umbrella, and that they correlate highly. The recommendation would be to conduct a single factor model, to see if it maybe is a unidimensional construct, the factor loadings from the general factor is indicative that the items measures one factor.

Words: 87

Answered.

17 LR1H22

You have been asked to assist a group of teachers of Norwegian as a foreign language to find the appropriate cut-score for a test of Norwegian reading proficiency.

As part of the process, the test was piloted with a representative sample of the intended population and the results are available to you. In addition, an established framework describes the expected level of Norwegian reading proficiency.

Give a brief outline of how a standard-setting procedure could be used to find the cut-score for pass/fail on the Norwegian reading proficiency test.

Fill in your answer here

A procedure could be done by equating, transforming scores, and comparing to a another test for Norwegian reading proficiency that has been validated.
Could also get a expert-panel to judge to what extent the cut score should be.
A standard setting procedure, could be to look at the mean of the test, which can be seen as the item difficulty, and conduct a confidence interval, and set a cut- score from this, the item difficulty would also be as an indicator if the test is to hard or to easy.

Words: 90

Answered.

18 LR2H22

A scale is being developed to measure satisfaction with life with the intended purpose to use the scale in national survey to identify which factors are associated with high satisfaction of life in the population. The scale consists of Likert items. According to the underlying theory of satisfaction with life, it is a unidimensional attribute. The theory also states that satisfaction with life is expected to have differences based on gender.

With this information in mind, do the following:

- Describe what evidence sources you want to consider in order to evaluate the validity of the scale scores for their intended purpose
- Describe the data you would like to collect to conduct the validity study
- Describe the analyses you would do in the validity study
- Outline what results you would consider as evidence supporting the validity of using the scale scores in the national survey

Fill in your answer here

According to the assumption that it is expected to have differences based on gender, it would be appropriate ti get a sample with "equal" proportion of men and women.

The construct validity, that the measure survey covers what is important, and measures what it is intended to measure.

With convergent validity, to see the correlation between a already validated construct that is supposed to measure the same, to see if the intended scale is good.

Since it states that it is unidimensional, would conduct a single factor model to see the model fit, and be open that it could also be a multidimensional construct that measures different factors of satisfaction with life.

For the evidence, after the items is done, look at the factor loading and see which items contributes the most to the underlying construct and the factor.

The evidence supporting the validity of using the scale scores in the national survey would be if the scale showing that there is differences based on gender, and that it is a unidimensional.

Words: 172

Answered.

19 LR3H22

The following output was obtained from estimating a single factor model to five 4-category Likert scale items from a scale measuring the environmental awareness of 15-year olds in Norway.

Item	Factor loading	Error variance
1	2.00	4.00
2	3.00	2.00
3	1.00	4.00
4	2.00	5.00
5	2.00	1.00

The residual correlation matrix was

$$\boldsymbol{\Sigma}_{\text{res}} = \begin{pmatrix} \mathbf{0.000} & & & & \\ \mathbf{0.026} & \mathbf{0.000} & & & \\ \mathbf{0.017} & \mathbf{-0.035} & \mathbf{0.000} & & \\ \mathbf{-0.014} & \mathbf{0.072} & \mathbf{-0.019} & \mathbf{0.000} & \\ \mathbf{-0.025} & \mathbf{-0.039} & \mathbf{0.020} & \mathbf{0.009} & \mathbf{0.000} \end{pmatrix}.$$

Address the following in your response:

1. What validity evidence categories from the Standards for Educational and Psychological Testing are relevant in this analysis? (1p)
2. Based on your appraisal, does the single factor model fit well?
3. Assume that a single factor model is appropriate for the analysis of the five item scores. Which item contributes the most to the reliability of the sum score and which item contributes the least? Justify your answers. (1p)
4. From the description of the items above and the results of the estimated model, give **one reservation** against the use of the linear factor model in this case. (1p)

Fill in your answer here

1.
That the scores are used as a validation argument.

2. the residual correlation matrix, if the singlefactor model is a good fit, that the estimates should be under 0.1, which the different residual correlations is. All the items also has a "high" factor loading, which is indicative that the single factor model was a good fit.

3.
to find out which items contributes the least and most.

Squared factor loading/ error variance for each item.
Item 1: 4/4
Item 2: 9/2
Item 3:1/4
Item 4: 4/5
Item 5:4/1

Item 2 and 5 contributes the most, and item 3 contributes the least
This is by conducting the squared factor loading / error variance for each item, and the factor loading tells something about how much the item contributes to the factor. Where item 2 and 5 both have the highest factor loading and the lowest error variance, and for item 3 which has the lowest factor loading and highest error variance.

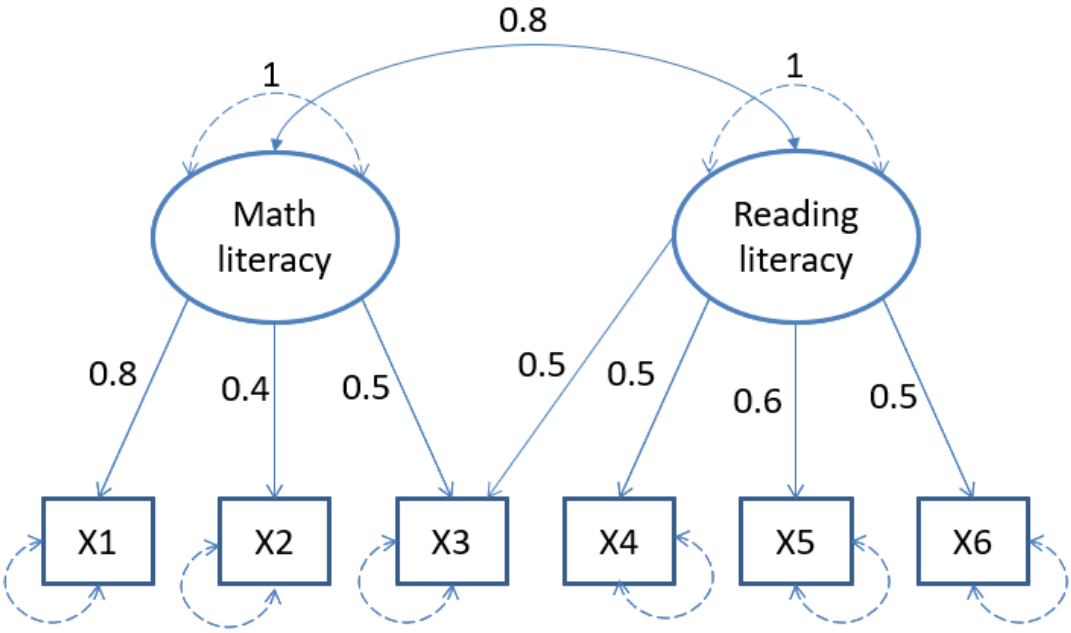
4. to few/ to many items.

Words: 167

Answered.

20 LR4H22

A multiple factor model is illustrated in the graph below. The latent variables and the observed variables are all standardized.



Answer the following questions based on the graph.

1. What is the equation which describes the model for the item score **X3**? Write down the equation with an explanation of the parameters and variables included. (2p)
2. What is the covariance between item scores **X3** and **X4** according to the model? Show your work and explain the steps taken. (2p)

Fill in your answer here

1.

$\lambda_3 M + \lambda_3 R + E_3$

The factor loading for item x3 on math literacy, the factor Math literacy (M), the factor loading for reading literacy, the factor for reading literacy (R) and the error. And the factor correlation.

2.

Cov(X3,X4)=0.4

0.5*0.8*0.5*0.5+

0.5*0.5*0.8

Words: 39

Answered.