UiO : Universitetet i Oslo

CANDIDATE

184122

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

correlation *r* =Cov(X,Y)/$\sqrt{Var\ (X) * Var(Y)}$
-7/$\sqrt{10}$ $* \sqrt{10}$
-7/10
$-0 \cdot 7$
The Pearson correlation coefficient *r equals* $-0 \cdot 7$ which suggests there is a strong but negative relationship between depression severity and satisfaction with life. The correlation assumes a linear relationship between the two measures i.e., X and Y. Moreover, the correlation is affected by measurement errors, and adjusting it for attenuation will increase the absolute value of the correlation.

Words: 69

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 of Z
Var(Z)=Var(X)+Var(Y)-2Cov(X,Y)
=2+3-2*1
=5-2
=3

2, Var of U
Var(U)=Var(X)+Var(2Y)+2Cov(X,2Y)
=Var(X)+2^2Var(Y)+2*2Cov(X,Y)
=2+4*3+4*1
=2+12+4
=18

Words: 20

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

=5(1^2)/10
=5/10
=0.5
The coefficient alpha can be interpreted as the reliability of the sum score when all the factor loadings are equal which is the case here. Again we are assuming the factor model is appropriate and there is no sampling variability. Under these conditions 0.5 is considered low reliability.

Words: 51

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

Too Much, i.e., construct irrelevancy e.g., cheating, anxiety or construction work near the exam venue.

Too little, i.e., construct underrepresentation e.g., insufficient sampling of items from target domain for instance only including the easy items or the difficulty items.

Words: 39

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

30=1.2X+6
30-6=1.2X
24=1.2X
20=X

the cut score for pass in terms of test X scores is 20 obtained by substituting values in the linear equating function.

Words: 26

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 fact that the factor model fits the data well (GFI $\geq$0.95, RMSEA<0.06,SRMR<0.08) means that the factor correlation can be interpreted to represent the underlying factors.
The sum score correlation will be lower because it is affected by measurement errors.

Words: 40

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

compute the explained common variance (ECV) i.e., compare the variance explained by the general factor to the variance explained by the sub-factors and see if the general factor explains the largest variance.

ECV= variance of the general factor /total variance

variance explained by the general factor = $\sum_{i=1}^{6} \lambda^2$ = 9+1+4+1+1+1=17

variance explained the sub-factor reading =$\sum_{i=1}^{3} \lambda^2$ =0.25+0.25+1=1.5

variance explained the sub-factor writing=$\sum_{i=1}^{3} \lambda^2$ =1+0.25+0.25=1.5

ECV=17/17+1.5+1.5

=17/20

=0.85

a rule of thumb/threshold is if ECV $\geq$0.70, then we can assume unidimensionality. In our case we have ECV greater than 0.7 hence the use of sums score is justified.

Words: 98

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

Using the bookmark method, determine the difficulty of the items e.g., using the response probability value and rank items by order of difficulty starting with the least difficulty to the most difficulty. Assign scores to the categories, then discuss and determine the threshold examinee and the items they are likely to answer correctly and place a cut point. Place a bookmark at a point where the threshold examinee will be able to answer correctly with a likelihood greater than or equal to the cut-point for the item before the bookmark and a likelihood less than the cut-point for the item after the bookmark.

Words: 103

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

Evidence sources

1. Content of the items
2. Relations with other variables
3. Internal structure

Data description

Get a random sample from the target population

Give the scale to this individuals

Record their gender

Get a panel of experts to evaluate the items based on content

Analysis

Test for unidimensionality since the theory posits that it is a unidimensional attribute. Run a factor model and see if the data fits well. A good fit of the data is indication of unidimensionality e.g., (GFI $\geq$0.95, RMSEA<0.06,SRMR<0.08).

Estimate reliability coefficient omega. Can be done using a factor model

T test or Multi group factor analysis to determine gender differentials in line with theory.

Results to consider as evidence supporting validity of using the scale scores

Good fit of the data e.g., if the factor model fits the data well then the sum score is reflective of the true reliability

Check the coefficient omega, if it is 0.8 and above this acceptable reliability

Experts evaluation consistent with theory

Difference in the means of item scores for men and women

Words: 173

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. Internal Structure

2. No. it does not fit the data well. There are some high residual values like 0,07. There will be discrepancies between the observed covariance matrix and the model implied covariance matrix.

3. use the information to rank items i.e., the ration between squared factor loadings and unique variance
item1= $2^2/4=4/4$
item2=$3^2/2=9/2$
item3=$1^2/4=1/4$
item4=$2^2/5=4/5$
item5=$2^2/1=4/1$
Items according to order from most to least in contribution is Item2, Item5,Item1,Item4 and Item3
so Item2 contributes the most and Item3 contributes the least

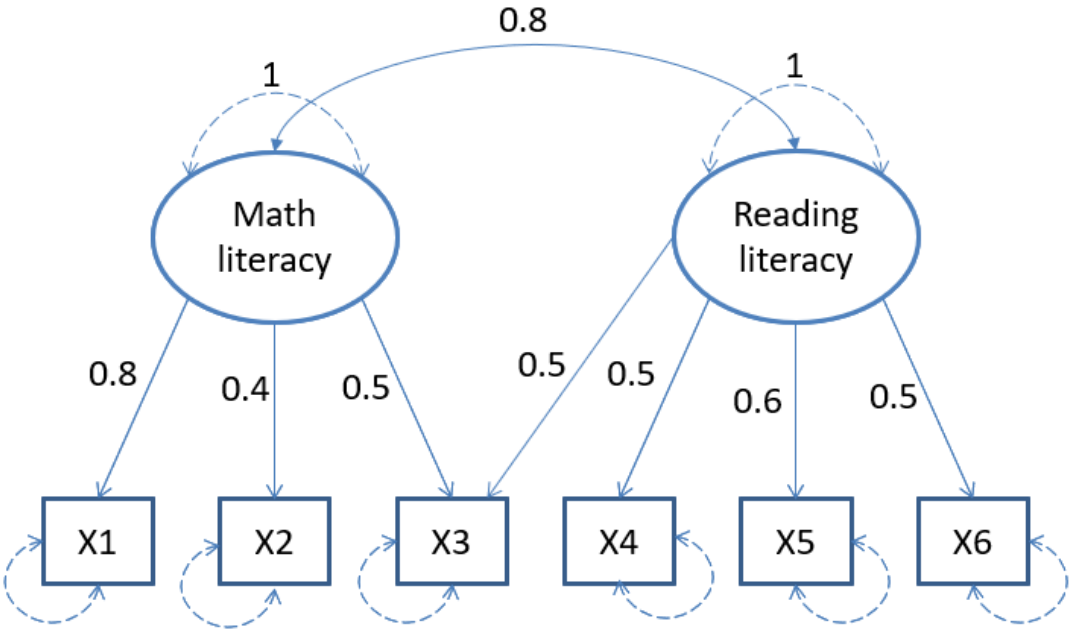
4. Likert scale items are ordinal data which strictly speaking makes the linear factor model inappropriate.

Words: 99

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. equation for X3

$X3=\lambda_{m3}M+\lambda_{m4}R+E3$

M is the factor for Math literacy, R is the factor for Reading literacy, E3 is the error term associated with item 3, theses are all random variables. λ_{m3} is the factor loading for Math literacy factor, λ_{m4} is the factor loading for the Reading literacy latent variable. The factor loadings together with the variance of the error term are the parameters for the model. λ_{m3} says how sensitive the item is in measuring the latent variable Math literacy, and λ_{m4} says how sensitive the item is in measuring the latent variable Reading literacy. The error term quantifies the random error associated with item 3.

2. covariance X3X4

=0.5*0.8*0.5+0.5*1*0.5

=0.45

$X4=\lambda_{m5}R+E4$

$Cov(X3,X4)=Cov(\lambda_{m3}M+\lambda_{m4}R+E3)+Cov(\lambda_{m5}R+E4)$

$= Cov(\lambda_{m3}M,\lambda_{m5}R)+Cov(\lambda_{m3}M,E4)+ Cov(\lambda_{m4}R, \lambda_{m5}R)+Cov(\lambda_{m5}R,E4)+Cov(E3, \lambda_{m5}R)+Cov(E3,E5)$

$=\lambda_{m5}\lambda_{m3}Cov(M,R)+\lambda_{m3}Cov(M,E4)+ \lambda_{m4}\lambda_{m5}Cov(R, R)+\lambda_{m5}Cov(R,E4)+\lambda_{m5}Cov(E3, R)+Cov(E3,E5)$

=0.5*0.5*0.8+0+0.5*0.5*1=0.45

Words: 136

Answered.