UiO : Universitetet i Oslo

CANDIDATE

184112

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

Correlations between depression severity and satisfaction with life= (-7/10)= - 0.70
This shows a high negative correlation between between depression severity and satisfaction with life with the assumptions that there is linear relationship and there can be measurement error.

Words: 39

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(X-Y)= Var(X) + Var(Y) - 2COV(X,Y)
= 2 + 3 - 2*1
= 5-2
=3
So, Var (Z) = 3

2. Var(X-Y)= Var(X) + Var(2Y) + 2 Cov (X,2Y)
= 2 + (2*2)*3+2*2*1
= 2+12+4
= 18
So, Var (U) = 18

Words: 43

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

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

and interpret it. State the assumptions underlying the interpretation.

Fill in your answer here

Coefficient alpha= 5 * [(1*1) / 10)
= 5* 0.10
= 0.5

Reliability coefficient of sumscore is 0.5. This score is not acceptable. We take a acceptable reliability coefficient more than or equal 0.8. Coefficient alpha can be said as a reliability coefficient when factor loadings are equal.

Assumptions: This is a single factor model and their is no sampling variability.

Words: 62

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: Individuals fear regarding exam or test can affect or their anxiety level can influence
Too little: These three domain may not be enough to measure their ability to meet real life challenges. There can be other factors that needs to be added.

Words: 44

Answered.

14 SR6H22

For two tests of reading comprehension, X and Y, the linear equating function was estimated to be $\mathrm{eq}(Y) = 1.2X + 6$. 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

Y= 1.2 X + 6
or, 30 = 1.2 X + 6
or, 1.2 X = 30 - 6
or, X = (24 / 1.2)
or, X = 20

So, cut score for passing the liner equating function Y= 1.2X + 6, in terms of test X is 20.

Words: 49

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

If we analysis the outcome of confirmatory factor analysis, we can say the model fits very well with goodness of fit = 0.95, RMSEA = 0.05 and SRMR = 0.06. All of these are indicator of a good model fit with these two factors, i.e. mathematics and interest in mathematics. Correlation between sumscore showing very low as there should be measurement error.

Words: 62

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

We can check how much variance the general factor and those two subfactors can explain to compare:
General: (3*3)+(1*1)+(2*2)+(1*1)+(1*1)+(1*1) = 9+1+4+1+1+1=17
Reading and writing: (0.5*0.5)+(0.5*0.5)+(1*1)+(1*1)+(0.5*0.5) + (0.5*0.5) = 3

Reliability= 3/17

The reliability is very low, much less than the acceptable level of 0.7

based on this, i can not recommend this as a multidimensional construct is more proffered.

Words: 59

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

I will follow following procedure:
1. First I will check item mapping so that each sets of items provide specific idea regarding the proficiency.
2. After item mapping is done, a team with experts will set anchor for different level of proficiency.
This will include cut score for pass fail also.
3. A committee will be made to decide the cut score for pass fail.
4. It will be checked if there is any bias through the differences of group mean score and sum score.
5. From the mean and after checking if there is any bias, the committee will decide the cut score.

Words: 104

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

1. To evaluate the validly of the scale scores I will use construct validity, internal structure and relationship with other variable. I will evaluate established theories, compare with those theories and check for unidimensionality.

2. I will collect random data covering information from different genders.

3. I will conduct confirmatory factor analysis to check the model fit and unidimensionality. To be unidimensional model should fit the data very well. I would expect GFI more than or equale 0.95, RMSEA less than or equal 0.5 and SRMR less than or equal 0.6. I will check what items are contributing mostly and the items not contributing to measure the attribute.

4. I will compare the factors with thier standardised correlation with other established factors that measures high satisfaction of life. Convergent validity evidence suggest that, I will get high correlation with those factors.That can be used as validity evidence.

Words: 147

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. Content validity and construct validity.

2. As ordinal data have been collected with likert scale, single factor model is not appropriate.From residual correlation matrix we can see, errors are correlated which volates the assumption of single factor model.

3. Contribution of the items can be found by deviding squared factor loading with by error variance. From this analysis, item 2 contributes most (4.5) and item 3 contributes least (0.25).

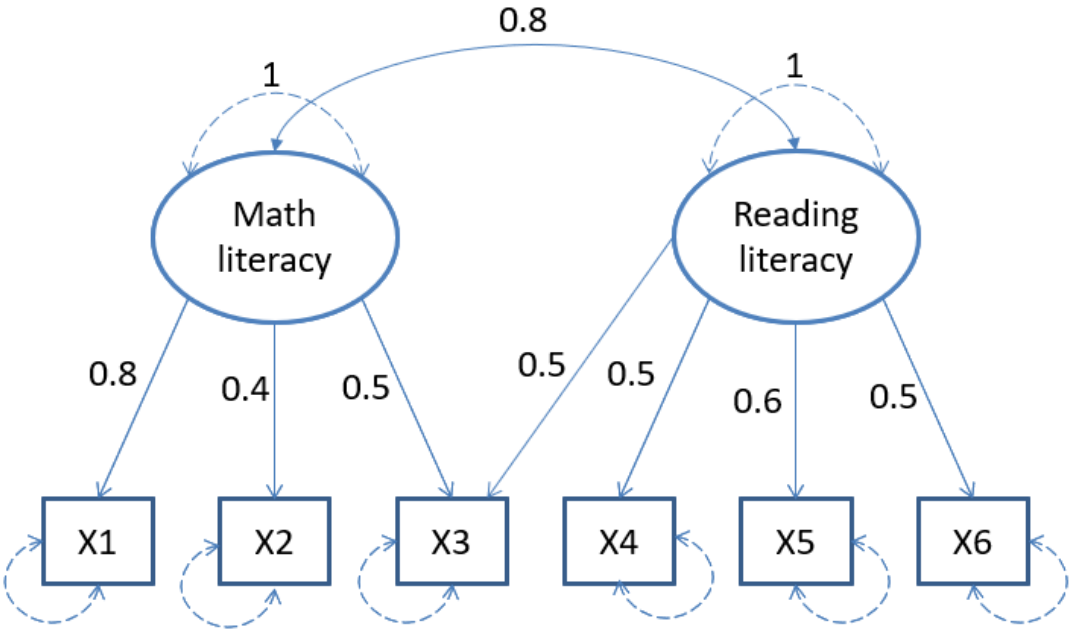
4. Linear factor model assume no sampling variance which can be difficult to meet here.

Words: 85

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. X3= meu3 + Lemda3M + Lemda3R+E3
Here, meu3=Item difficulty,
Lemda 3 is the factor loading for factor Math literacy. Change in X3 due to 1 unit change in math literacy equals to Lemda3. M, factor Math Literacy
Lemda 3 is the factor loading for factor Reading literacy. Change in X3 due to 1 unit change in Reading literacy equals to Lemda3. R, factor Reading literacy
E3, unique variance for X3.

2. Cov (X3,X4) = Cov (lemda3M+lemda3R+E3, lemda4R,E4)
= Cov(lemda3M+lemda3R+E3, lemda4R) + Cov(lemda3M+lemda3R+E3, E4)
= lemda3.lemda4.Cov(M,R) + lemda3.lemda4.Cov(R,R)+0+0+0+0
=0.5*0.5*0.8 + 0.5*0.8
= 0.6 (answer)

Words: 94

Answered.