UiO : Universitetet i Oslo

CANDIDATE

184103

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:

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

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

Calculation correlation from the covariance matrix.

Cor(X,Y) = $\frac{Cov(X,Y)}{\sqrt{\sigma_X}\sqrt{\sigma_Y}}$ = $\frac{-7}{\sqrt{10}\sqrt{10}}$= $\frac{-7}{10}$=-0.7

The relationship between depression and satisfaction with life is negative and strong.

The assumptions are that correlation assumes a linear relationship between the variables. Further, relationship between the variables can be effected by measurement error, as well sampling variation.

Words: 51

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

2. Var(U) = Var(X+2Y) = Var(X)+Var(2Y)+2Cov(X,Y) = Var(X)+2^2Var(Y)+2Cov(X,Y) = 2+(4*3)+(2*1) = 16

Words: 22

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

$\alpha = m\frac{\lambda^2}{Var(Y)} = 5\frac{1^2}{10} = 5\frac{1}{10} = 0.5$

The coefficient alpha is 0.5. The coefficient is moderate. As the factor loadings must be equal when using this formula to get the coefficient alpha, and if a single factor model is true for the data, the coefficient alpha in this instance is the reliability coefficient of the items.

Words: 55

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

Influenced by too little

If only two out of the three domains of the test are measured using this test, the test score would reflect to little. For example if the test only measures reading and mathematics knowledge and skills, not science.

Influenced by too much

If the test measures the individuals motivation as well as the three domains, the test scores will be influenced by too much.

Words: 68

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

As the cut score for Y is 30 i insert that into the equation.

30 = 1.2X+6
I then subtract 6 from both sides.

24 = 1.2X
I then divide both sides by 1.2

20 = X

The cut score of X after using the equating function is 20.

Words: 52

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 the factor model used is not perfectly fitting the data, it means that there is measurement error. As measurement error is a consequence of the test not the factors, the correlation between the factors will only be the same as the correlation of the sum score if there is no measurement error.

Words: 53

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

General factor loadings = 3+1+2+1+1+1 = 9

Whole model = 9+0.5+0.5+1+0.5+0.5+0.5= 12.5

$\frac{9}{12.5} \approx 0.71$

As the ratio of the sum of the factor loadings of only the general factor to the sum of the factors loadings of the full model is above 0.7, I would recommend using the sum score, as the model can fit a single factor model, meaning the test can measure one construct, the general factor.

Words: 69

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

The cut-score for a pass/fail Norwegian reading proficiency test can be found by first evaluation based on the framework made what should be expected as a satisfactory Norwegian reading level for those who fit the test. Based on the sample the percentile that most fit this levels score can be used as a cut off point. The level that is satisfactory should be evaluated by experts of Norwegian reading proficiency.

Words: 70

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
The evidence sources for validation that I would like to consider is evidence based on internal structure.

Data
If would like to collect response data from a representative sample of the population of interest, the population in this case in those who take part in a national survey.

Analysis
As the theory specifies a unidimentional construct, I wish to analyze this through a confirmatory factor analysis with a single factor model.

Evidence
Results that would support the validity of the use of the scale scores in the national survey would be that the single factor model had a good fit, which would indicate a unidimentional construct, confirming the theory the scale is based on. Further, results that indicated differences between men and women would also confirm the theoretical assumptions for the scale, and therefore strengthens the validation argument for the inclusion of the satisfaction with life scale in the national survey.

Words: 153

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. The evidence category from the Standards that is relevant in this analysis is validity evidence based on internal structure.

2. 0.000+0.026+0.000+0.017+-0.035+0.00+-0.014+0.072+-0.019+0.00+-0.025+-0.039+0.02+0.009+0.00/15= 0.00

The mean of the residuals is 0 which is good, but to evaluate the model properly more information is needed.

3. $Item\ 1 = \frac{2^2}{4} = 0$
$Item\ 2 = \frac{3^2}{2} = 4.5$
$Item\ 3 = \frac{1^2}{4} = 0.25$
$Item\ 4 = \frac{2^2}{5} = 2$
$Item\ 5 = \frac{2^2}{1} = 4$

To estimate the items that contribute the most you square the factor loading and divide it by the unique error variance of the items. The one with the highest value using this formula, is the item that contributes the most. In this item set the one that contributes the most is item 2.

4. One reservation against using the linear factor model in this case is that the Likert-scale items are measured in the ordinal level, and is therefore not appropriate to use in a factor

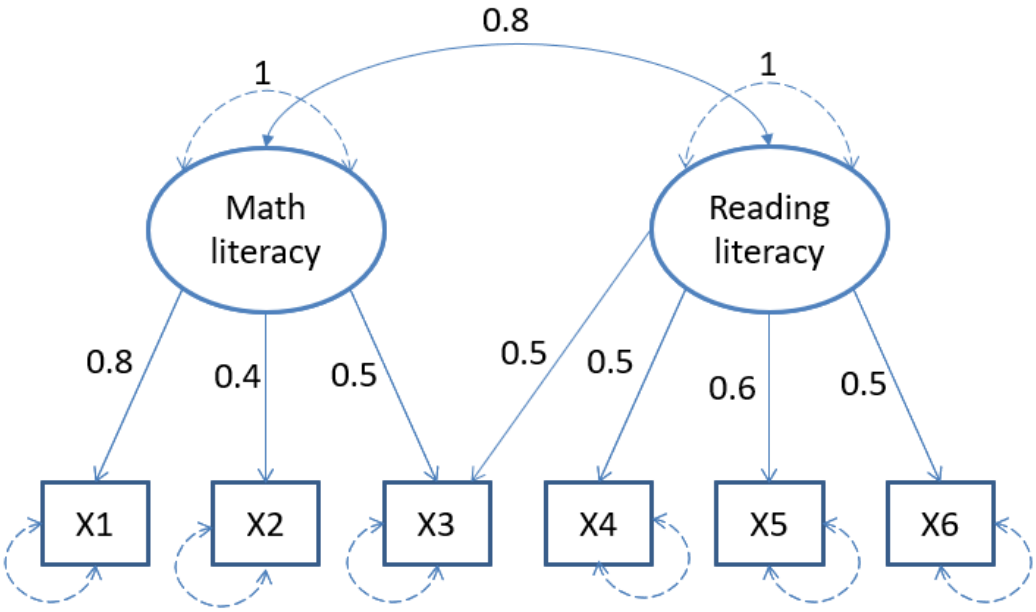
model.

Words: 146

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. $X_3 = \lambda_{M3}M + \lambda_{R3}R + \epsilon_3$
The lambda M3 represents the extent to which item 3 measures the factor Math literacy. The M is the factor of Math literacy. The lambda R3 is the extent to which the item measures the other factor, which is represented by the R standing for reading literacy. E3 is the unique error of the item.

2.
$Cov(X_3, X_4) = Cov(\lambda_{M3}M + \lambda_{R3}R + \epsilon_3,\ \lambda_{R4}R + \epsilon_4)$
I write out the equation for the two items.
$= Cov(\lambda_{M3}M, \lambda_{R4} + \epsilon_4) + Cov(\lambda_{R3}R, \lambda_{R4}R + \epsilon_4) + Cov(\epsilon_3, \lambda_{R4}R + \epsilon_4)$
I look at the covariation between each part of the first item and the full second item.
$=Cov(\lambda_{M3}M, \lambda_{R4}R) + 0 + Cov(\lambda_{R3}R, \lambda_{R4}R) + 0 + 0 + 0$
As the error of all the items in a factor model is uncorrelated with the factors, the errors are zero.
$=\lambda_{M3}\lambda_{R4}Cov(M, R) + \lambda_{R3}\lambda_{R4}Var(R, R)$
I take the factor loading outside the brackets, the covariation of the second part is turned to variance.
=0.5*0.5*0.8+0.5*0.5*1= 0.45
I insert the values from the model and find that the covariance between item 3 and 4 is 0.45.

Words: 152

Answered.