UiO : Universitetet i Oslo

CANDIDATE

184107

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

Cov(X,Y) = -7
Var(X) = 10
Var(Y) = 10

Depression severity scale scores and satisfaction with life scores has equal variances.
Depression severity scale scores and satisfaction with life scores are negatively correlated with each other. For example, student with low satisfaction with life has high depression severity and vice versa.
Correlation = Cov(X,Y) / sdX*sdY
Thus, negative Cov(X,Y) means negative correlation between 2 variables.

Words: 65

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

Var(X-Y) = Var(X) + Var(Y) - 2*Cov(X,Y) = 2 + 3 - 2*1 = 3
Var(X+2Y) = Var(X) + Var(2Y) + 2*Cov(X,Y) = 2 + 2^2*3 + 2*1 = 16

Words: 30

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

alpha = 5 * 1^2/10 = 1/2 = 0,5
Coefficient alpha is equal to the reliability coefficient in this case. Because the factor loadings are equal across 5 items. Reliability 0,5 means that precision of the measurement of this test is not good.

Words: 43

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 much:
influenced by too little: difficult to compare real-life challenges with school subjects

Words: 16

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

30 = 1,2*X + 6
24 = 1,2X
X = 20
Cut score for pass in terms of the test X is 20.

This means that difficulty level of the test X was higher than the test Y.
One who scores 30 on test Y can be said that have same achievement with one who scores 20 on the test X.

Words: 61

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

Correlation of sum scores affected by measurement errors. Thus, after fitting the factor model, factor correlation is higher than the sum score correlation due to measurement error.

- Factor model, in general, is good fit. GFI = 0,95, RMSEA < 0,06 or (RMSEA < 0,05) and SRMR < 0,08

Words: 48

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

Since the factor loadings are not entirely identical for all factors and the sum score was used in a previous study, this bifactor model would recommended to use.

Words: 28

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

1. compare expected level of Norwegian reading proficiency with the construct of the test.

Words: 13

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 based on internal structure + Evidence based on relation to other variables
- Same number of men and women respondents that are collected from the different socio-economic status and have different income level, and in same age interval.
- Since the purpose of the survey is to identify which factors are associated with high satisfaction of life, and to see the relationship between factors, Multiple or Bifactor model is appropriate to analyze relationship between factors and items in between, and also relationship between items and general factor, high satisfaction of life
- Correlation between income factor and satisfaction of life and correlation between socio-economic factor and satisfaction of life

Words: 107

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. Evidence based on internal structure

2. The residual correlation matrix shows that residuals of items are uncorrelated with each other

3. reliability of item1 = $\frac{\lambda^2}{\psi^2}$ = 2^2/4^2 = 1/4 = 0,25

item2 = 3^2/2^2 = 9/4 = 1,25

item3 = 1^2/4^2 = 1/16

item4 = 2^2/5^2 = 4/25

item5 = 2^2/1^2 = 4

Item 5 contributes the most to the reliability of the sum score and item 3 contributes the least.

4. Coefficient omega is lower than 1/3.

$$\omega = \frac{\left(\sum_{j=1}^{m} \lambda\right)^2}{\left(\sum_{j=1}^{m} \lambda\right)^2 + \left(\sum_{j=1}^{m} \psi\right)^2}$$ = 10^2/ 10^2+16^2 < 1/3

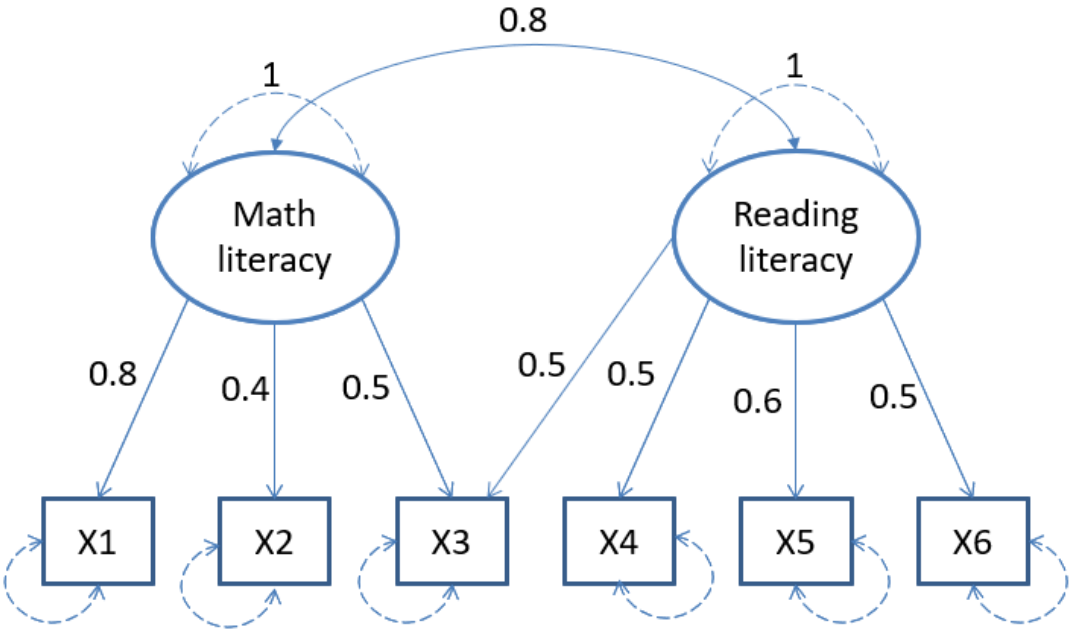
Thus, the reliability of the test is lower than 1/3. Test is not validated in terms of the evidence based on internal structure.

Words: 112

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 = $\mu_3 + \lambda_{m3}M + \lambda_{r3}R + \epsilon_3$

- μ and λ are parameter (item difficulty and factor loading)
- M and R are factor corresponding Math literacy and Reading literacy
- ϵ_3 is an error term of item scores X3

2. X3 = $\mu_3 + \lambda_{m3}M + \lambda_{r3}R + \epsilon_3$ and X4 = $\mu_4 + \lambda_{r4}R + \epsilon_4$
Cov(X3, X4) = Cov($\mu_3 + \lambda_{m3}M + \lambda_{r3}R + \epsilon_3$, $\mu_4 + \lambda_{r4}R + \epsilon_4$)
= Cov($\mu_3 + \lambda_{m3}M + \lambda_{r3}R + \epsilon_3$, μ_4) + Cov($\mu_3 + \lambda_{m3}M + \lambda_{r3}R + \epsilon_3$, $\lambda_{r4}R$) + Cov($\mu_3 + \lambda_{m3}M + \lambda_{r3}R + \epsilon_3$, ϵ_4)
= 0 + Cov($\lambda_{m3}M + \lambda_{r3}R$) + 0 + 0 + 0 + Cov($\lambda_{m3}M + \lambda_{r3}R$, $\lambda_{r4}R$) 0 + Cov($\lambda_{m3}M + \lambda_{r3}R$) + 0+0+0
= $\lambda_{m3}\lambda_{r3}\lambda_{r4}$ Cov(M,R) = 0,5*0,5*0,5*0,8 = 0,1

In addition to covariance: For the multiple factor, Cov between 2 items = product of the factor loadings of each item and covariance between 2 factor. In this case, item scores X3 has relation to both factors.
$\lambda_{m3}\lambda_{r3}\lambda_{r4}$ Cov(M,R) = 0,5*0,5*0,5*0,8 = 0,1

Words: 190

Answered.