Started on	Wednesday, 16 November 2022, 10:29 PM
State	Finished
Completed on	Monday, 21 November 2022, 1:59 PM
Time taken	4 days 15 hours
Grade	Not yet graded

Information

Declaration

The University regards plagiarism as a form of academic misconduct, and has very strict rules regarding plagiarism. See UNSW policies, penalties, and infor

By submitting this assessment item, you declare that this assessment item is your own work, except where acknowledged, and acknowledge that the asses

1. Reproduce this assessment item and provide a copy to another member of the University; and/or
2. Communicate a copy of this assessment item to a plagiarism checking service (which may then retain a copy of the assessment item on its database

You certify that you have read and understood the University Rules with respect to Student Conduct.

Information

Software

You may use any software of your choice to help you in this assessment. Use of R and computer algebra systems is permitted and encouraged. You are not

However, any code you use to get your results must be submitted as a part of your working.

Information

Unless otherwise stated, multiple choice and True/False questions in the following section have a *guessing penalty*. This means that a correct answer gains

Do **not** let it prevent you from answering questions, because unless you are completely randomly guessing, you are far more likely to gain marks by answerir

True or False? A factor analysis model is a kind of structural equation model.

Select one:

- True
- False

Your answer is correct.

The correct answer is: True

Question 2

Complete

Marked out of 3.00

In your own words, and using plain (non-technical) language, write at most two sentences explaining the main difference between principal component analy

In FA, a latent factor is an unobserved phenomenon that is acting on observed variables. By contrast, in PCA, a component is simply an aggregate of corre

Information

Instructions: Please read carefully

The following instructions apply to all CodeRunner questions.

1. Into the space provided, write R code that (ultimately) assigns to each of these variables the answer to the question.
   - You may either implement your answer in the space provided, or you may work elsewhere, and only provide your final answers as constants.
   - Note that not all R packages may be available in the testing environment.
2. Click "Precheck" to check that your answers are formatted correctly, and address any issues.
   - Passing tests get marked with ✔, failing tests with ✘.
   - **IMPORTANT:** If your code fails the first pre-check (i.e., code produces errors when executing), your assessment **cannot** be marked, and you may rec
3. The code will be submitted when you submit the quiz.
4. When you submit the quiz, you **may** get a weird error message along the lines of "Expected 0 test results, got 1." Please disregard it.

(a)

Consider a random vector $\boldsymbol{X} \in \mathbb{R}^4$ that mean $\boldsymbol{\mu}$ and variance-covariance matrix Σ.

Suppose that a random sample of size 30 was taken, resulting in $\hat{\Sigma} = \begin{bmatrix} 177 & 40 & -14 & 4 \\ 40 & 98 & -37 & -4 \\ -14 & -37 & 314 & 5 \\ 4 & -4 & 5 & 1 \end{bmatrix}$.

For your convenience, you can construct it in R by running the following code:

Sigma <- matrix(c(177, 40, -14, 4, 40, 98, -37, -4, -14, -37, 314, 5, 4, -4, 5, 1), 4, 4)

Assuming multivariate normality, test the null hypothesis that the population variance-covariance matrix of $\boldsymbol{X}$ is diagonal. That is, that

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{bmatrix}$$

for appropriate values of $\sigma_i, i = 1, \dots, 4$. Store the test statistic in variable `ans_a_chisq`, its degrees of freedom in `ans_a_df`, and the p-value in `ans_a_`

Answer: (penalty regime: 0 %)

Reset answer

```
Sigma <- matrix(c(177, 40, -14, 4, 40, 98, -37, -4, -14, -37, 314,

ans_a_chisq <- 19.61531

ans_a_df <- 6

ans_a_pval <- 0.003241336

```

	Criterion	Result	Mark	
✔	Quiz results available?	Yes	0.00/0.00	✔
✔	(a) chi-squared statistic is correct	Yes	1.00/1.00	✔
✔	(a) degrees of freedom are correct	Yes	1.00/1.00	✔
✔	(a) p-value is correct or consistent with the above	Yes	1.00/1.00	✔

Passed all tests! ✔

Question author's solution (R):

```
# NOTE: These are for validation purposes only. These are *not* t

Sigma <- matrix(c(177, 40, -14, 4, 40, 98, -37, -4, -14, -37, 314

ans_a_chisq <- 1

ans_a_df <- 1

ans_a_pval <- 1

```

▲ Back to Top

Throughout this question, you may assume multivariate normality. Suppose that we have observed 59 cases of 9 variables, which we believe to be determin

```
X <-
matrix(c(-0.28,0.02,0.45,0.55,-0.24,-0.92,0.46,-0.25,-0.46,-0.65,-0.09,-0.17,-2.66,-0.53,-0.37,0.01,0.97,0.32,0.26,
nrow=59, ncol=9, byrow=TRUE, dimnames=list(NULL,paste0("X",1:9)))
```

You can write it out with the following command:

```
write.csv(as.data.frame(X), file="X.csv")
```

(a)

(i)

Fit a 1-factor factor analysis model using maximum likelihood, and test whether 1 factor is adequate to explain the correlations in the data. Store the test sta

(ii)

Fit a 2-factor factor analysis model using maximum likelihood, and test whether 2 factors are adequate to explain the correlations in the data. Store the test

(iii)

Fit a 3-factor factor analysis model using maximum likelihood, and test whether 3 factors are adequate to explain the correlations in the data. Store the test

(iv)

Based on your answer to the previous three parts alone, how many factors do we require? Store your result in variable: `ans_a_iv`.

Note: If your conclusion that we require more than 3 factors, then please answer "4" (without the quotes).

(b)

Suppose that the substantive theory suggests that:

- the true model has 2 factors
- there are 3 variables associated with the one factor, 3 with the other factor, and the remaining 3 variables have no communality.

(i)

Using an appropriate model, with maximum likelihood estimation and varimax rotation, identify which three variables are most likely to belong to the first fac

Answer in the following format: suppose that you believe that (X_1, X_2, X_3) belong to one factor, (X_4, X_5, X_6) to the other, and the rest three to neither.

```
ans_b_i_f1 <- c(1,2,3)
ans_b_i_f2 <- c(4,5,6)
```

or

```
ans_b_f1 <- c(4,5,6)
ans_b_f2 <- c(1,2,3)
```

(ii)

Based on the estimated Λ and Σ_e (with 2 factors), what is the predicted correlation matrix of $\boldsymbol{X}$? Store your result in variable `ans_b_ii` as a 9 by 9 matrix.

Hints: `factanal()` and `fa()` both compute loadings and uniquenesses on the **correlation** matrix rather than covariance matrix; if `z` is the result, `z$loadi`

Answer: (penalty regime: 0 %)

Reset answer

```
1, -0.27467543011208373, -0.60301237870208468, 0.14456719230299
-0.092384105256179294, 0.91676070172309476, 0.99009509661955 43
0.97256129174122963, 0.44372933735973991, 0.081384239840726624
0.99999999999999978, 0.87033767354461611, 0.23169699294466967,
-0.13365196522608966, 1.0850515277458004, 0.98343025340795531,
0.98312102851105487, 0.49126083355868533, -0.0083550593211870 74
1.1089353229727035, 0.99999999999999989, 0.18902504710504203,
-0.085532050551642286, 1.0938344773461006, 1.0772194724043251,
0.82758484958369949, -0.22399523068468552, 0.684051398869047 87
0.41512152955791154, 0.13385193429019665, 1, -0.9476502470780 08
1.0981364433322198, 0.90376599661803958, 1.1284817602583364,
1.187379071033082, 0.4929415160291199, 0.095613986106399179,
-0.094863748647240703, -0.90180883235876153, 1, 0.84591577505 02
1.0011526179668755, 0.95886462185715882, 0.39043007795448104,
0.54461061475146588, 0.35684177082136104, 0.12702707099357413,
```

▲ Back to Top

	Criterion	Result
✔	Quiz results available?	Yes
✔	(a)i chi-squared test statistic is correct	Yes
✔	(a)i degrees of freedom are correct	Yes
✔	(a)i p-value is correct or consistent with the above	Yes
✔	(a)ii chi-squared test statistic is correct	Yes
✔	(a)ii degrees of freedom are correct	Yes
✔	(a)ii p-value is correct or consistent with the above	Yes
✔	(a)iii chi-squared test statistic is correct	Yes
✔	(a)iii degrees of freedom are correct	Yes
✔	(a)iii p-value is correct or consistent with the above	Yes
✔	(a)iv is consistent with the previous answers	Yes
✔	(b)i is at least partially correct	Yes
✔	(b)i is completely correct	Yes
✘	(b)ii is correct	Uniquenesses were not converted to a diagonal matrix. Notice that this produces an asymmetric correlation matrix, which cannot be corr

Question author's solution (R):

```
X <- matrix(c(-0.28,0.02,0.45,0.55,-0.24,-0.92,0.46,-0.25,-0.46,-

ans_a_i_teststat <-
ans_a_i_df <-
ans_a_i_pval <-

ans_a_ii_teststat <-
ans_a_ii_df <-
ans_a_ii_pval <-

ans_a_iii_teststat <-
ans_a_iii_df <-
ans_a_iii_pval <-

ans_a_iv <- 1

ans_b_i_f1 <- c(1,2,3)
ans_b_i_f2 <- c(4,5,6)

ans_b_ii <- diag(9)
```

Partially correct

Marks for this submission: 5.00/7.00.