Assignment 2

SEM 1: Confirmatory Factor Analysis

Please hand in a `.pdf` file containing your report and the `.csv` file that has been generated. The deadline of this assignment is Tuesday April 16 13:00.

Assignment

Download the file `data.csv` from canvas, which I obtained from `https://openpsychometrics.org/_rawdata/`. You can load the data in R as follows:

```
FullData <- na.omit(read.csv("data.csv"))
```

Next, run the following code in R (replace ... with your student number, or your birthday (YYYMMDD) if you do not have a student number):

```
studentNumber <- ...
set.seed(studentNumber)
Data <- FullData[sample(1:nrow(FullData), 500), ]
write.csv(Data,paste0("conspiracyData_",studentNumber,".csv"), row.names = FALSE)
```

This code created an object `Data`, which is also stored as a csv file in your working directory. If you do not use R, let someone run this code for you and send you the csv file (do use your own student number). This dataset contains your personal $n = 500$ subset of a larger dataset in which the Generic Conspiracist Beliefs Scale (Brotherton, French, & Pickering, 2013) was administered. This scale aims to measure "government malfeasance" (GM), "extraterrestrial cover-up" (ET), "malevolent global conspiracies" (MG), "personal wellbeing" (PW), and "control of information" (CI). The items are:

Item	Item description	Factor
Q1	The government is involved in the murder of innocent citizens and/or well-known public figures, and keeps this a secret	GM
Q2	The power held by heads of state is second to that of small unknown groups who really control world politics	MG
Q3	Secret organizations communicate with extraterrestrials, but keep this fact from the public	ET
Q4	The spread of certain viruses and/or diseases is the result of the deliberate, concealed efforts of some organization	PW
Q5	Groups of scientists manipulate, fabricate, or suppress evidence in order to deceive the public	CI
Q6	The government permits or perpetrates acts of terrorism on its own soil, disguising its involvement	GM
Q7	A small, secret group of people is responsible for making all major world decisions, such as going to war	MG
Q8	Evidence of alien contact is being concealed from the public	ET
Q9	Technology with mind-control capacities is used on people without their knowledge	PW
Q10	New and advanced technology which would harm current industry is being suppressed	CI
Q11	The government uses people as patsies to hide its involvement in criminal activity	GM
Q12	Certain significant events have been the result of the activity of a small group who secretly manipulate world events	MG
Q13	Some UFO sightings and rumors are planned or staged in order to distract the public from real alien contact	ET
Q14	Experiments involving new drugs or technologies are routinely carried out on the public without their knowledge or consent	PW
Q15	A lot of important information is deliberately concealed from the public out of self-interest	CI

Question 1 (1 point) Without fitting the model to data yet, give the number of degrees of freedom for the 5-factor model for these items. ∎

Question 2 (1 point) The paper reports RMSEA = 0.06 in a confirmatory factor analysis. Taking this value as prior guess for the true RMSEA score, what is the sample size required to reject the test for not-close fit of RMSEA > 0.08 with a power of 0.80 and $\alpha = 0.05$? ∎

Question 3 (3 points) Fit the theoretical 5-factor model to your dataset and write a paragraph (max 200 words) judging the fit of your model. Include in your paragraph a test for exact fit, a test for not-close fit, and a broad overview of fit measures and their interpretation. You will be graded based on correctness of the estimated model (1 point), correctness of the reported indices and their interpretation (1 point), and writing style and clarity (1 point). ∎

Question 4 (2 points) Perform three model modifications to your model to improve fit (1 point), describe why you chose these particular model modifications (0.5 point), and perform model comparisons between your original model and your final model (0.5 point). You may use modification indices to inform on which model modifications to make, but also describe if they make sense to you from a conceptual point of view. Do not just give the model comparison values, but also interpret them. ∎

Question 5 (2 points) Using any software or method you like, make clear and interpretable path diagrams for both your original model and your modified model. The path diagrams should make visual comparison between your two models easy. You may choose to either include or omit parameter estimates in the path diagram if you wish. You will be graded on the correctness of the path diagrams (1 point) and their clarity (1 point). ∎

Question 6 (1 point) Are the following statements true or false?

- When comparing two models, the *simpler* model is the model in which there are *more* degrees of freedom.
- If a χ^2 difference test between two nested models is *significant*, we would prefer to select the *simpler* model.
- If the AIC is lower for model A than for model B, then so is the BIC.
- We would always prefer the model with the lowest *RMSEA*.

∎

Challenge questions

You may choose to complete one of the challenge questions below for a bonus point. Note, completing several questions does not lead to more bonus points.

Challenge question 1

If you followed the network analysis class, you may try to estimate a *latent network model* on your conspiracy dataset (Epskamp, Rhemtulla, & Borsboom, 2017), either using the *lvnet* package:

- `http://sachaepskamp.com/files/BFI_example.R`

Or the newer *psychonetrics* package:

- https://github.com/SachaEpskamp/psychonetrics
- https://www.youtube.com/watch?v=Hdu5z-fwuk8

Plot the estimated latent network structure and interpret the results. ∎

Challenge question 2

Use your conspiracy dataset to estimate a *regularized* CFA model using the *regSem* package (https://CRAN.R-project.org/package=regsem, https://arxiv.org/pdf/1703.08489.pdf). Do you retrieve the same factors as according to the theoretical model? ∎

Challenge question 3

In this assignment, you fitted two models: a CFA model (model A) and an adjusted CFA model with three extra parameters (model B). Set up a simulation study that investigates the *power* to reject the χ^2 test at $n = 500$ in the case we fit model A, but model B actually generated the data. One method for generating such data is to extract Σ from your fitted model B, generate random multivariate normal data using the *rmvnorm* function in the *mvtnorm* package, and subsequently fit model A to this new dataset. Repeat this process 100 times to estimate the power. ∎

References

Brotherton, R., French, C. C., & Pickering, A. D. (2013). Measuring belief in conspiracy theories: The generic conspiracist beliefs scale. *Frontiers in psychology*, *4*, 279.

Epskamp, S., Rhemtulla, M., & Borsboom, D. (2017). Generalized network psychometrics: Combining network and latent variable models. *Psychometrika*, 1–24. Retrieved from http://dx.doi.org/10.1007/s11336-017-9557-x doi: 10.1007/s11336-017-9557-x