Interactive Assignment 7 - Factor Analysis

Factor analysis

This interactive assignment is going to cover the idea of factor analysis. Factor analysis is a process by which we take many variables and try to examine which underlying variables, or **factors** are present in the data.

One example of this is the Big 5 Personality Factors, or the Five Factor Inventory. This research examined how people rated themselves on many, many different personality attributes. Then the researchers wanted to see whether the variables are related to one another. For example, people who rate themselves highly on "outgoing" would probably rate themselves high on "social" and low on "quiet". These adjectives are related to one another and thus can be a factor.

If we want to examine a psychological construct, like a personality variable, we want to examine whether it is a single construct or several constructs, and what it is related to.

Initial data summary

For this project, you will work with two dataframes, called "IA7factoranalysis.csv" and "IA7summary.csv".

Step 1: Set up your data and notebook in a folder. Download both dataframes into the folder. load the "IA7factoranalysis.csv" as a dataframe named "spdry" and load "IA7summary.csv" with the name "summary".

This project investigates two dataframes. The one called spdry, was a questionnaire where participants answered 33 questions about how they felt regarding spiritual dryness, or how connected a person felt spiritually with God or religion. Higher numbers indicate more spiritual dryness and questions which may indicate low spiritual dryness are reverse coded. The variables are named "Response.ID" which is the participant number and Q1-Q33, representing the questions. The questions are given as the last page of the assignment.

The second file is called summary and that is a file which contains each participant's result in several other established measures of religiosity. Those include: intellect, public, private, and experience from Huber and Huber's religion scales, external, internal, and transactional religion from Batson's scales, Allport's extrinsic and intrinsic scales, and Worthington's intrapersonal and interpersonal scales.

Step 2: For this project, we will need to install and load the "psych" library. Go ahead and install this by clicking on the top menu and selecting Tools -> Install Packages. Load the "psych" and "tidyverse" libraries.

Step 3: First, we will look at how the items are intercorrelated. Here is an easier way to view this. Type the following R code. This code uses the `cor` command to create a correlation matrix and then sends it to the round command, to round the variables to 2 decimal places.

```
x = round(cor(spdry[,2:34]), 2)
```

A way we can visualize this is by using the RStudio viewer. In the environment tab, click on the x under data. Then you can select each of the variables and sort them by their correlation.

Step 4:: Answer the following in your annotation in your R notebook.

1. Which variables don't seem to correlate with the rest? I found at least two.

2. Are there any variables very correlated with each other?

Factor analysis.

In this part, we will run a factor analysis to see if our questionnaire has any factors. There are two types of factor analysis, exploratory and confirmatory. Exploratory means that we don't know how many factors to expect, whereas confirmatory means that we want to test whether our data has a certain number of factors. In this case, we don't know how many factors we want so we will do exploratory factor analysis. R has a command that tells us how many factors to expect.

The psych package has a command called `fa.parallel()` which does a concept called *parallel analysis*. This uses the data to get an idea of how many factors and how many components to select from the data. Components are similar to factors though they use a slightly different type of math to calculate them, a concept called Principal Component Analysis. So you may get a different answer for the number of factors versus the number of components.

Step 5: Type the following as a code chunk to see how many factors and compoents. How many factors does this suggest? How many components?

```
fa.parallel(spdry[,2:34])
```

```
## Parallel analysis suggests that the number of factors =  2  and the number of components =  2
```

Now we are going to do the factor analysis. I will look for two factors, despite the answer for the question above, just because it is easier to visualize.

Step 6: Type the code below to run the factor analysis. There are two lines here. The first line uses the `factanal` command to do the factor analysis. It has the option `factors = 2` which sets how many factors we want. Note that I am giving it the spdry dataframe with columns 2-34 because we don't want the first column which contains subject numbers. The second line prints out these results and cuts off any factors which are less than .5.

```
f = factanal(spdry[,2:34], factors = 2,rotation = 'varimax')
print(f, sort=TRUE, cutoff = .5)
```

This runs the factor analysis and prints the results. The idea is that if a question fits on a factor with a cutoff value of .5 or better, then it is grouped in a factor. If it does not, it does not get a factor analysis value at all.

Step 7: Look at the output of this question and compare it to what the questions themselves say. Does it seem like the questions that are in factor 1 seem to relate to one concept? What about the questions in factor 2? What about the questions that don't make the cutoff? Are these including many of the same items you answered in question 1?

Step 8: Now we can re-run the command reducing the cutoff.

```
f = factanal(spdry[,2:34], factors = 2,rotation = 'varimax')
print(f, sort=TRUE, cutoff = .3)
```

NOtice that when we reduce the cutoff, many questions fit in both factors. This can get kind of messy.

Step 9: We can graph this to better visualize the factors too.

```
fa.plot(f)
```

This command visualizes the factors on a plot. The X axis represents how much a given question is related to factor 1 and the y axis represents how mch a given question is related to factor 2. Notice that some factors are very high in one and low in the other, some are middle in both, and a few are low in both.

Step 10: Based on these factor analysis and the graphs, choose which questions you would select for a questionnaire to examine spiritual dryness. You should only choose 5-10 questions for each factor and omit questions which do not load highly on either factor. Also, explain what construct you think is best captured by factor 1 and by factor 2

Spiritual Dryness Questions

1. Even when I don't "feel" God, I know He is there
2. I believe that God really loves me and knows my needs
3. I find it difficult to allow myself to depend on God
4. I find it relatively easy to get close to God
5. I often feel disconnected from God
6. I rarely worry about God abandoning me
7. I sometimes feel abandoned by God
8. I sometimes worry that God doesnít really love me
9. I worry a lot about my relationship with God
10. I'm afraid that I might lose God's love
11. It helps to turn to God in times of need
12. Sometimes I just feel empty inside, like God isnít even there
13. There are times in my life when I felt really connected to God
14. I can see God working in my life
15. I don't feel like Scripture talks to me
16. I feel close to God when I pray
17. I feel close to God when I worship
18. I feel connected when I pray to God
19. I feel God answers my prayers
20. I feel God has a plan for me
21. I feel like God doesnít hear me when I pray
22. I feel like God has forgotten about me
23. I feel like God hears other people more than me
24. I feel like God is active in my life
25. I feel like God is distant
26. I feel like God is personally involved in my life
27. I feel like God knows me as an individual
28. I feel like God talks to me
29. I feel like other people are more connected to God than I am
30. I feel spiritually dry
31. I feel very connected to God
32. I feel very connected to my religious community
33. I have trouble believing God has a plan for me individually