PSYC 7433

Fall 2021

Problem Set 5

**NOTE: For all problems you are expected to “show your work” if I ask for computations with statistical software this includes RELEVENT code and output, if I ask for hand computations, this includes writing out the steps you took.**

1. In the file pfs.csv is data derived from *N* = 229 parents enrolled in community-based family support programs in Texas and Kansas. The data are from a preliminary version of the Protective Factors Survey (PFS), a scale intended to capture individual differences in **constructs thought to inhibit child abuse.** Each question was answered on a 7-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree." The items are:

|  |  |
| --- | --- |
| Variable | Item |
| PFS1 | I have neighbors, friends or relatives that help me when I need it |
| PFS2 | My family members feel closer to people outside the family than to our own family members |
| PFS3 | I know where to go in my community to get help with family needs |
| PFS4 | More bad things happen to my family than to other families |
| PFS5 | My family enjoys spending time together |
| PFS6 | When I am worried about my children, I have someone to talk to |
| PFS7 | I don't think my family can survive if another problem hits us |
| PFS8 | I praise my children when they behave well |
| PFS9 | My family shows each other love and affection |
| PFS10 | My family is able to solve our problems |
| PFS11 | When we have disagreements, family members listen to both sides of the story |
| PFS12 | When I discipline my children, I have a hard time keeping my feelings under control |
| PFS13 | I try to comfort my children when something is bothering them |
| PFS14 | My family members discuss problems with each other |
| PFS15 | In my family, we take time to listen to each other |
| PFS16 | I try to take a break when I am frustrated by my children's behavior |
| PFS17 | In my family, we support one another when something goes wrong |
| PFS18 | When my child misbehaves, I... use time-out |
| PFS19 | When my child misbehaves, I... spank |
| PFS20 | When my child misbehaves, I... hit |
| PFS21 | When my child misbehaves, I... ground |
| PFS22 | When my child misbehaves, I... take away privileges |

**Table 1.**

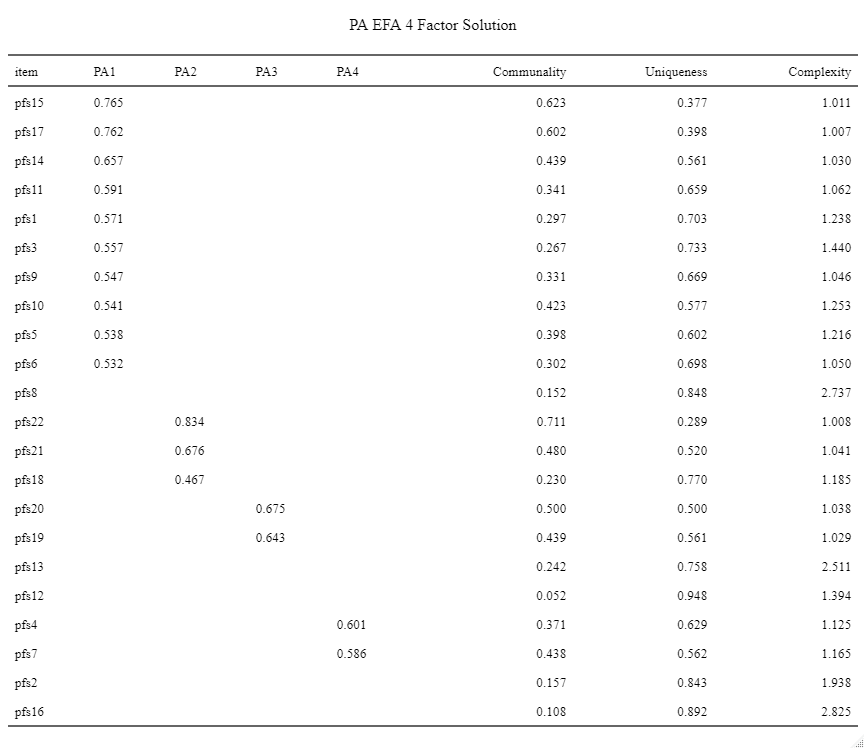
*SS Loadings and Variance Explained*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Support | Behavioral Disc. | Corporal Disc. | Neg. Outlook |
| SS Loadings | 4.03 | 1.51 | 1.32 | 1.03 |
| Proportion Var. | 0.18 | 0.07 | 0.06 | 0.05 |
| Cumulative Var. | 0.18 | 0.25 | 0.31 | 0.36 |
| Proportion Exp. | 0.51 | 0.19 | 0.17 | 0.13 |
| Cum. Prop. Exp. | 0.51 | 0.70 | 0.87 | 1.00 |
|  |  |  | *RMSEA* = 0.052 | |
|  |  |  | *TLI* = 0.869 | |

**Table 2**

*Factor Correlations*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Support | Behavioral Disc. | Corporal Disc. | Neg. Outlook |
| Support | 1.00 |  |  |  |
| Behavioral Disc. | 0.14 | 1.00 |  |  |
| Corporal Disc. | 0.26 | -.022 | 1.00 |  |
| Neg. Outlook | 0.33 | 0.06 | 0.21 | 1.00 |



Perform an exploratory factor analysis from start to finish. EFA involves several key decisions on the part of the researcher. These involve:

* Software to use
* Estimation method
* Number of factors
* Rotation method (orthogonal vs. oblique)
* Interpretation of the results

...as well as other items not listed here (e.g., additional software options, naming the factors, assessing model fit...). Provide a reasoned justification for every choice you make. Provide a complete interpretation of the results. Comment on what you perceive as weak points in the scale and offer suggestions for improvement.

The evaluation criteria for this assignment are intentionally open-ended to reflect the subjective component that should be part of every good factor analysis. All of you may end up with different—but equally legitimate—results. **(25 points)**

1. Find an article published in your area of research using cluster analysis/mixture modeling and write a short review of the article focusing on the use of cluster analysis/mixture modeling. Some questions you may want to address are: What variable(s) were used in the model? What conclusions did they draw from the model? Was this an appropriate use of the technique? Why or why not? **(10 points)**

**See below for details. Highlighted sections are relevant to this question. Basically, they clustered participants based on risk behavior frequency (cigs, marijuana, alcohol, sex) into ‘at-risk’ or ‘safe’ groups. Group/cluster membership was then used as a dependent variable in a logistic regression model that predicted ‘at-risk’ membership from cyberbullying victimization, while controlling for demographic variables and measures of traditional bullying and self-control. Results indicated that cyberbullying victimization significantly increased the odds of a student belonging to the ‘at-risk’ cluster, with the effect remaining consistent when controlling for exposure to physical bullying, a proxy measure of self-control, and various demographic variables**. **This was methodologically appropriate (discussed further below, highlighted pink).**

Wood Jr, F. R., & Graham, R. (2020). “Safe” and “At-Risk”: Cyberbullying victimization and deviant health risk behaviors in youth. *Youth & Society, 52*(3), 449-468. <https://doi.org/10.1177%2F0044118X18810943>

This non-longitudinal study investigated the links between cyberbullying (CB) victimization and a set of health risk behaviors associated with juvenile delinquency (cigarette smoking, marijuana and alcohol usage, and sexual frequency), while controlling for self-control and deviant peer associations. The links between CB victimization and specified deviant health risk behaviors were examined within a theoretical framework aligned with General Strain Theory, in which the presence of unwanted stimuli (i.e. hostile messages intended to inflict harm or discomfort), are hypothesized to produce negative emotions that can lead to deviant behavior (Agnew, 2001). Data were collected from the 2015 *Youth Risk Behavior Survey*, which consisted of a representative sample of 9-12th graders in the U.S. Specifically, this study focused upon the items in which respondents were asked to self-report whether or not they had experienced CB and/or traditional bullying victimization within the previous 12 months, the number of recent sexual partners, and the frequency of recent substance abuse (marijuana and/or alcohol). Using cluster analysis, respondents with complete data observations (*N* = 9,122) were categorized into two groups: “safe” students who reported, on average, no engagement in the behaviors measured, and “at-risk” students who reported, on average, moderate to high levels of engagement in the health risk behaviors outlined above. Following the cluster analysis, the data were also subset into participants who were not physically bullied and those who were bullied.

Logistic regression models were then used to predict group membership (‘at-risk’ vs ‘safe’) from CB victimization, controlling for gender, age, race, and sexual orientation. Two models were fit on each physical bullying subset, one of which included a self-control measure and one of which did not, for a total of four separate models (see figure below). Results indicated that CB victimization significantly increases the odds of a student belonging to the ‘at-risk’ cluster, with the effect remaining consistent when controlling for exposure to physical bullying, a proxy measure of self-control, and various demographic variables. Key limitations of this study include the reliance upon a measure of grades/academic outcomes as an indirect measure of self-control, as well as the reliance upon self-report data (memory, social desirability bias). Despite the other limitations, this was an appropriate use of both cluster analysis and logistic regression given that the cluster analysis was used to form a dichotomous grouping of ‘safe’ vs ‘at-risk’ participants. Future studies should incorporate more precise measures of strain, peer associations, and self-control, as well as utilize a mixed-methods approach to examine qualitative data as well.

**Extra Credit (3 points).** Create a meme about how any topic in multivariate statistics makes you feel. There are lots and lots of meme generators online no need to break out Photoshop skills.

