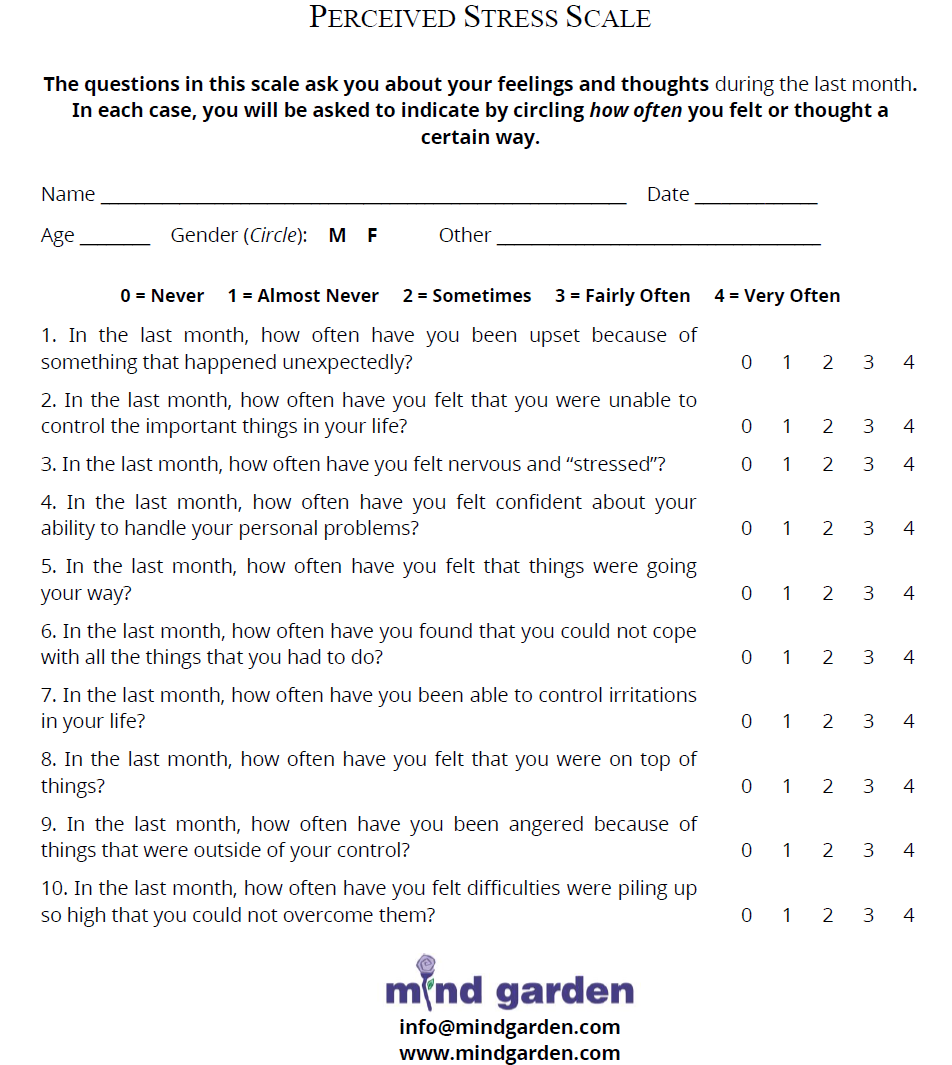
STAT 335: HW #4 Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_  
Fall 2018  
Points: 40

*Note: Whenever possible you can use an R package and related functions to compute all quantities. You should provide a copy of your R output for each question.*

For the first part of this homework assignment, you will be computing internal consistency reliability measures for outcomes from a survey instrument used to measure stress, i.e. the Perceived Stress Scale survey instrument.



Scoring for this instrument:

* Overall Score
* Two sub-scores
  + Perceived Helplessness – Questions 4, 5, 7, and 8
  + Perceived Self-Efficacy – Questions 1, 2, 3, 6, 9, and 10

Note: The perceived helplessness items are reserve coded on this survey. Thus, when an overall score is computed, adjustments to these scores must be made.

Questions

1. What is the average inter-item correlation for the overall score? What is this value for each of the sub-scores? Provide an interpretation of each of these values? (4 pts)
2. Which items appear to have the strongest correlation with each other? Consider the wording for these items, does it make sense that these two items have the strongest correlation? Discuss. (3 pts)
3. Use the split-half reliability approach to compute the Spearman-Brown corrected internal consistency correlation for the overall score. Interpret this value. (3 pts)
4. Compute Cronbach’s alpha for the overall score. Provide an interpretation of this value. Compute Cronbach’s alpha for each of the sub-scores. Again, provide an interpretation of these values. (3 pts)

Next, consider a second survey instrument that is intended to measures the “well-being” of an individual, i.e. the Linear Analog Self Assessment (LASA) survey instrument. LASA was given to the same study participants as above, i.e. those that gave outcomes from the Perceived Stress Scale survey instrument.

LASA is copyright protected; thus, one should not openly share the exact wording of these survey items. There are only five items on LASA and the focus of each item is presented in the table below. The scale on LASA goes from 0 (bad) through 10 (good). None of the items on LASA are reverse coded.

|  |  |
| --- | --- |
| Items | Focus of Item |
| 1 | Physical well-being |
| 2 | Emotional well-being |
| 3 | Spiritual well-being |
| 4 | Intellectual well-being |
| 5 | Overall well-being |

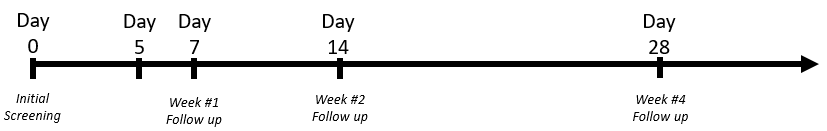
1. Compute the inter-item correlations, i.e. pair-wise correlations, for all five items. Which type of well-being (physical, emotional, spiritual or intellectual) appears to be most correlated with item #5 -- overall well-being? Discuss. (3 pts)
2. Compute an average well-being score using items 1-4. What is the correlation between this average and item #5 - overall well-being? Consider this correlation value, what would be your advice on whether or not Item #5 should be removed from this survey instrument? Discuss. (4 pts)

|  |  |
| --- | --- |
| The Suicide Prevention Resource Center provides a list of resources that can be used to evaluate the risk of suicide for an individual.  Source: <https://zerosuicide.sprc.org/toolkit/identify/screening-and-assessing-suicide-risk> |  |

Suppose you work for the Suicide Prevention Clinic. One of the questions on a commonly used tool to evaluate the risk of suicide is, “Over the past month, how often have you considered harming yourself?”. This standard question allow respondents to select either “Not Often” or “Often”. You believe that too often respondents tend to pick “Not Often” and you’re considering changing the way in which people respond to this question. Your two alternatives are shown here.

|  |  |
| --- | --- |
| Standard Question |  |
| Alternative #1  *Add an option for Sometimes* |  |
| Alternative #2  *Sliding scale:*  *Not Often = 0.0*  *Often = 10.0* |  |

As stated above, the standard question is commonly used in practice. Your supervisor suggests that before changing the tool used by your clinic to evaluate risk of suicide, you must first establish that one or both of your alternative versions are *better* than the standard question. The usual timeline for contact with patients in your suicide prevention clinic is Day 0 (Initial contact with subject), Day 5, Day 7, Day 14, Day 28. Contact with subjects beyond 28 days is done on a subject-by-subject basis.



You decide to run two separate studies to evaluate the alternative versions of this question

* In Study A, subjects were be given the Standard Question and Alternative #1 on Day 5 and again on Day 7.
* In Study B, subjects were given the Standard Question and Alternative #2 on Day 5 and again on Day 7.

Days 5 and 7 were used as one risk of suicide is not likely to change over a 2 day period. For each study, the collection of data was over a 60 day period. Study A included 108 subjects and Study B included 166 subjects. The dataset include information for Study A and Study B (the first set of rows are from Study A and the next set is from Study B).

1. Cohen’s Kappa statistic can be used to measure the Test / Retest Reliability of an item. Compute Cohen’s Kappa statistic for the Standard Question from Study A. Would you consider this item to have good test/retest reliability? Discuss. (4 pts)

Note: I used the following to read in the data and get the rows for Study A

SuicideData <- read.csv( file.choose() )

StudyA <- filter(SuicideData, Study == ‘A’)

1. Next, compute the Cohen’s Kappa statistic for Alternative #1 from Study A. You will need to use the weighted version of Cohen’s Kappa statistic here as Alternative #1 includes three levels. You can use either the Cicchetti-Allison or Fleiss-Cohen weights for this calculation. See page 9 of Handout #7 for details. (5 pts)
2. Compute Cohen’s Kappa statistic for the Standard Question from Study B. (3 pts)
3. Next, compute the intraclass correlation coefficient (ICC) as a measure of test/retest reliability for Alternative #2 from Study B. Would you consider Alternative #2 to have good test/retest reliability? (4 pts)
4. Which version (Standard, Alternative #1, Alternative #2) would you suggest using on this suicide risk evaluation tool? How did you make this determination? Discuss. (4 pts)