Multivariate Statistical Techniques
CLPS2908, Spring 2019
Prof. Bertram Malle

Take-home Midterm Exam

Rules of Engagement

You may use any material you have available to answer the questions (including books, notes, web pages), but you must write your answers in your own words. If you use outside sources, reference them (on a separate page). Most important, you must not collaborate with any other person, inside or outside this class.

If you have questions of clarification, please address them to me (BFM).

Each of your answers must be on 1 page or less (measured at 1.5 line spacing, 1" margins, 12-pt Times New Roman font). Exceeding the length restriction leads to point deductions.

Number the subparts of your answers (a, b, etc.) clearly, but do not include the question text.

The completed exam is due on Thursday, April 4, 2019, at 5:00 p.m. Please submit it through the assignment module in canvas as a .doc(x) or .pdf file.

Before you turn in the exam, sign or type your name on the line below, which indicates that you have read the rules of engagement and have abided by them throughout your work on this exam. Then add this page to your submitted document.

Signature:

1. What simple but powerful form of data reduction and integration underlies any multiple regression equation, the forming of principal component/factor scores, and contrasts in ANOVA? (a) Name it, (b) describe what it does mathematically, (c) describe what it means conceptually, and (d) briefly explain what role it plays in each of the three abovementioned statistical applications (multiple regression, PCA, and contrasts).

2. Reliability and validity come in several forms. (a) Describe two types of reliability and (b) provide examples (real or invented) of conditions/situations/applications for which you would use each. (c) Then describe two types validity and (d) provide examples of situations or applications in which you would use each. (Note that this material was covered only partially in the lecture itself. Refer to the additional lecture handout pages, the readings, and use other resources, if appropriate, to answer the question.)

3. Some people claim that exploratory data analysis (EDA) is like cheating; they argue that looking at your data before running your significance tests biases your testing strategy and therefore the interpretation of significance tests. Write a critical analysis of this claim, both (a) discussing what may be correct or incorrect about it and (b) making a counterargument by pointing to the strengths of EDA.

4. (a) Name the type of matrix that PCA's $\mathbf{\Lambda}$ matrix is. (b) Specify what all the elements (entries) of the matrix mean conceptually. (c) Explain (verbally or formally) how SPSS, or any other statistical program, moves from R_{xx} to $\mathbf{\Lambda}$.

5. Write a dialogue (feel free to be funny) between a fanatic proponent of PCA and a fanatic proponent of factor analysis. Let each person state at least *two clearly distinct* arguments for the supremacy of their preferred technique and finish with your own (fanatic or not) conclusion.