HUDM5124 SPRING 2020

MIDTERM REVIEW QUESTIONS

1. Three cases of the Minkowski "p-metric" distance model that we have discussed are: A. the city-block metric (p=1), B. the Euclidean distance metric (p=2), and C. the maximum metric (p=). Give one example of psychological proximity data where each type of distance model might be most appropriate. Support your answer with good arguments and examples, and/or references to the literature.

2. Assume we had data on how 50 students score on 15 test items. The scores on the 15 items are dichotomous, =1 if correct and =0 if incorrect. We wish to find out if there are common patterns of test performance among subjects. Two data analysis methods we might use are 1) to compute correlations among SUBJECTS, giving a 50 x 50 correlation matrix, and attempt to run a PCA of this correlation matrix or 2) to compute Euclidean distances among all pairs of subjects, and run an MDS to find salient dimension that distinguish subjects.

Discuss how each of the methods might work, problems or issues you might encounter in running the analyses and interpreting the results, and which approach you might expect to work better to discover patterns of common test performance among students.

3. Answer Problem 6.6 in the Modern MDS text (parts a and b ONLY)