

Study Questions For Predict 410

Topic: Multivariate Statistical Techniques

Our learning format requires that you complete the assigned readings efficiently and "intelligently". In order to help you focus your attention on important concepts in the course reading, we have constructed a list of study questions for each topic covered in PREDICT 410. You should preview each list of study questions before you begin your reading, and then answer the questions in a notebook while you are performing your reading. If you cannot answer a question, then you should look up the answer. If you cannot find the answer, then you should post a question in your Blackboard course shell.

Principal Components Analysis

- (37) Principal Components Analysis (PCA): What is principal components analysis? How does PCA eliminate the problem of multicollinearity? What does it mean for X_1 and X_2 to be orthogonal? In order to better understand orthogonality, take the building prices data set and perform these steps:
- (a) Perform a PROC CORR on X_1 - X_9 .
 - (b) Create nine orthogonal predictor variables using PCA. Call these variables Z_1 - Z_9 .
 - (c) Perform a PROC CORR on Z_1 - Z_9 .
- (38) Principal Components Analysis is described as a method of 'dimension reduction'. How does PCA reduce the dimension of a statistical problem? How do you select the reduced dimension for your problem.

Factor Analysis

- (39) Are the factor scores always orthogonal? Are they orthogonal after a rotation?
- (40) If two analysts perform a factor analysis, are they likely to arrive at the same result? If the same two analysts perform a principal components analysis, are they likely to get the same result?
- (41) What is the first step in performing a factor analysis?
- (42) In the context of factor analysis, what is the *communality* of factors?

Cluster Analysis

- (43) What is the difference between hierarchical and non-hierarchical clustering?
- (44) What is *linkage*? What types of linkage are there?
- (45) How do we examine the goodness-of-fit of a cluster analysis or two comparative cluster analyses?
- (46) Do the data need to be treated before we perform a cluster analysis?