

# Statistics 579

## Applied Multivariate Methods

### Exam 1

The file Softdrink.txt gives data from a survey undertaken to determine consumer perceptions of six competing brands of soft drinks. The brands rated were as follows: (1) Pepsi Cola (regular); (2) Coca-Cola (regular); (3) Gatorade; (4) Allsport; (5) Lipton original tea; (6) Nestea.

Respondents used a 7-point Likert scale (1=strongly disagree to 7=strongly agree) to indicate their level of agreement with the following 10 statements:

1. Brand X has a refreshing taste.
2. I prefer Brand X because it has fewer calories than other drinks.
3. Brand X quenches my thirst immediately.
4. I like the sweet taste of Brand X.
5. I prefer drinking Brand X after workouts and sports because it gives me energy.
6. I prefer Brand X because it comes in environment friendly packaging.
7. Brand X has minerals and vitamins that help quench my deep down body thirst.
8. Brand X has a unique flavor of its own.
9. Brand X has the right mix of minerals and vitamins that are healthy for my body.
10. I prefer to drink Brand X when I am really thirsty.

1. Determine if there are any outliers in these data (but do not eliminate any observations in any of the subsequent analyses).
2. Determine if the 10 variables (i.e., responses to the 10 statements) in this data set can be regarded as arising from a multivariate normal distribution.
3. Using Fisher's method with a bias adjustment, find a 95% confidence interval for the true population correlation between the responses to Statement # 2 and the responses to Statement # 5.
4. Test whether the responses to the 10 statements can be regarded as statistically independent.
5. Perform a Principal Components Analysis on the correlation matrix using all 10 variables, and determine the actual dimensionality of the space in which the data fall (i.e., determine the number of principal components that would adequately describe these data).
6. Perform a Factor Analysis on the correlation matrix using all 10 variables. After determining the number of factors needed, compare the results produced from using both the iterated Principal Factor Method and the Maximum Likelihood method. Use a Varimax Rotation with both methods. Specifically, which variables load on which factors, and do the 2 methods produce the same results? Also, provide an interpretation of each factor that you identify.