

**Lab Assignment No. 7:**  
**Principal Components and Factor Analysis**  
PSGA7816 Multivariate Statistical Analysis

In completing your assignment you are expected to be specific and analytical in your responses. Simply producing tables or output is **not sufficient** and if not accompanied by any explanation or discussion it will not be given **any** credit.

Please append the full syntax you used to compute your answers. Your typed assignment answers should include the products of your computations (estimates, tables and/or charts) pasted into the body of your paper with appropriate remarks and comments (preferably in APA style).

No e-mail submissions will be accepted. Late assignments will be penalized 2 points per day. The dataset for the assignment (track.txt, consisting of men's national track records for 55 countries in 1984 for the following distances: 100m, 200m, 300m, 500m, 1000m, 1500m, 2000m, 2500m) will be e-mailed to students. Homework 7 is due **April 23, 2009 at 1:30 PM**.

- 1) How many principal components account for at least 80% of the data? How many principal components would you retain based on the usual eigenvalue criterion?
- 2) Provide a correlation matrix of all principal components and items analyzed. What information can you gather from the correlation matrix?
- 3) Conduct a factor analysis (using the principal components method) extracting the maximum number of factors. Explain the utility and meaning (in your own words) of the eigenvalues, eigenvectors and factor loadings.
- 4) Run a factor analysis (using the maximum likelihood method) using the eigenvalue criterion as an initial guide for the number of factors to be extracted. What are your initial conclusions? What corrective steps would you have to take?
- 5) Report and discuss whether the correlation matrix is adequate to be used in this factor analysis. Report the measures of sampling adequacy and the appropriate Chi-square statistic.
- 6) Fit a four factor (maximum likelihood) model. Is this model a better fit than the one in question 4?
- 7) Which distances load on which factors in the four factor model?
- 8) Discuss the two rotations (varimax and promax) in a 2 factor model. Which one is a better fit? Why?
- 9) Briefly describe the differences and similarities between principal components and factor analysis. How do these two approaches provide different information in the analysis you have conducted?
- 10) What are some of your final remarks/observations regarding your analysis?