HUDM 6122
Spring 2020

Homework # 8

Due on 04/08/2020 at 1:00 pm

1. The data on national track record for women for Helsinki 2005 are in the file `TrackWomen.dat`. Note that the file is *tab-delimited* and has a header with the variable names. Note also that the times for the events 100 m to 400 m are measured in seconds, while 800 m to Marathon events times are measured in minutes.

   a) Perform a factor analysis on the *correlation* matrix with two factors using the principal component method without rotation. Report the loadings and interpret the factors.

   b) Compute the factor scores from part a) and plot them on a scatterplot to check for outliers.

   c) Repeat part a) with varimax rotation.

   d) Repeat part a) but this time use *three* factors and the maximum likelihood method without rotation.

   e) Repeat part d) with varimax rotation.

   f) Test formally with a hypothesis if three factors are enough in part d). Report the p-value and conclusion.

   g) Repeat parts d) & e) but this time use the *covariance* matrix of the data converted to speeds, like in HW 7.

   h) Which analysis do you prefer? Why?