Inleiding multivariate statistiek/Introduction to Multivariate Statistics

Assignment 2 (0.5 points)

Deadline: October 11, 9:00 a.m. (submit on Canvas in pdf format).
Submissions after the deadline (send by mail) get 0.2 points penalty.

Exercise 1

We have the data file data.txt. It is a three dimensional dataset.

1. Test the hypothesis that $\boldsymbol{\mu} = (0, 0, 0)^T$. Use a 5% significance level for the test.
2. Calculate the axes for the 95% joint confidence region for $\boldsymbol{\mu}$.
3. Calculate the 95% simultaneously valid confidence intervals for components of $\boldsymbol{\mu}$.
4. Calculate the 95% confidence intervals for components of $\boldsymbol{\mu}$ using Bonferroni correction.

Exercise 2

We have the data (file MarkVal.txt) (Data about various companies and their economic and financial characteristics) dataset which contains four variables: *Assets* (in millions of dollars), *Sales* (in millions of dollars), *Market_Value* (in millions of dollars), *employees* (in thousands of people). Note that, the dataset is quite old, so some numbers might seem implausible now. You do not need the sector variable.

1. Using quantile-quantile plots, evaluate the normality assumption for different characteristics and assess joint multivariate normality.
2. Consider the transformation of all the variables by natural logarithm. Assess the normality of the transformed data.

Report the outputs and figures to justify your answers. You can use R or Matlab.