P8157 - Analysis of Longitudinal Data, Fall 2019
Homework - 1
Due : September 24, 2019

September 10, 2019

1. Plasma inorganic phosphate measurements obtained from 13 control and 20 obese patients 0, 0.5, 1, 1.5, 2, and 3 hours after an oral glucose challenge. The investigators intend to test the following hypotheses using Hotelling's T^2 statistic :

   (a) To test the null hypothesis that the group means are the same at all six measurement times.

   (b) To test whether the profiles in the two groups are parallel.

   (c) To test whether the differences in means at 2 and 3 hours after an oral glucose challenge are different between the control and obese patients.

   Set up suitable model clearly stating the assumptions. Suggest appropriate contrast matrices for the hypothesis tests proposed by the investigators.

Part B [1]

For each of the following data sets perform *Exploratory Data Analysis (EDA)*, to get insights into the data to eventually perform an appropriate longitudinal data analysis.

1. Exposure to lead can produce cognitive impairment, especially among young children and infants. Interventions known as *chelation treatments* can help a child to excrete the lead that has been ingested. A chelating agent known as Succimer can be administered orally leading to urinary excretion of lead, unlike previous treatments which required injections and hospitalization. The Treatment of Lead-Exposed Children (TLC) trial was a placebo-controlled, randomized study of succimer (a chelating agent) in children with blood lead levels of 20-44 micrograms/dL. These data (**TLC.dat**) consist of subject id, assignment to chelation treatment with succimer or placebo and four repeated measurements of blood lead levels obtained at baseline (or week 0), week 1, week 4, and week 6 on 100 children who were randomly assigned to chelation treatment with succimer or placebo.

2. The data (**dental.dat**) are from a study of dental growth measurements of the distance (mm) from the center of the pituitary gland to the pteryomaxillary fissure were obtained on 11 girls and 16 boys at ages 8, 10, 12, and 14. The variables consist of ID, Gender, and responses at ages 8,10,12 and 14.

3. The data for the Plasma inorganic phosphate measurements from 13 control and 20 obese patients describe in question 2 of part A are available in the file *ZERBE2.dat*. Carry out the Hotelling's T^2 test to test the hypotheses proposed above in part A. Perform EDA and provide any insights available from exploration. The data has the following information

[1] Always explore the options available in the commands in your software

```
Column 1 : Group (control=1,Obese=2)
Column 2 : Subject id
Columns 2-8 : Plasma inorganic phosphate measurements 0, 0.5, 1, 1.5, 2, and 3
              hours after an oral glucose challenge
```

4. The pigs data set (**PIGS.txt**) and CD4+ count data set (**MACS.txt**), demonstrated in class are provided. Try to repeat the EDA for these data sets as shown in class. (It need not be identical, but you should be able to attain similar insights from the data using your methods). **Do not submit this part of the homework**. The pigs data consists of 9 weeks of data for each of the 48 pigs. The CD4+ data has the following information

```
1) time since seroconversion 2) CD4 count
3) age (relative to arbitrary origin) 4) packs of cigarettes smoked per day
5) recreational drug use yes/no 6) number of sexual partners
7) cesd (mental illness score) 8) subject ID
```