HW1_answer

Guojing Wu | UNI: gw2383

9/23/2019

Part A

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

Set up the model as: $Y = XB + E$, and we assume:

- $y_{s_i} \sim MVN(\mu_{s_i}, \Sigma)$

- each group has its own μ_{s_i}

- Σ is the same across groups

The null hypothesis has the general form as $ABC = D$, where

$$B = \begin{pmatrix} \mu_{11} & \mu_{12} & \mu_{13} & \mu_{14} & \mu_{15} & \mu_{16} \\ \mu_{21} & \mu_{22} & \mu_{23} & \mu_{24} & \mu_{25} & \mu_{26} \end{pmatrix}$$

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

$$A = I_2$$
$$C = \begin{pmatrix} I_5 \\ -1_5' \end{pmatrix}$$
$$D = \mathbf{0}_{2 \times 5}$$

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

$$A = \begin{pmatrix} 1 & -1 \end{pmatrix}$$
$$C = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & 0 & -1 \end{pmatrix}$$
$$D = \mathbf{0}_{1 \times 5}$$

(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.

We need to test whether $\mu_{15} - \mu_{25} = \mu_{16} - \mu_{26}$

$$A = \begin{pmatrix} 1 & -1 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

$$D = \begin{pmatrix} 0 \end{pmatrix}$$

Part B

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.
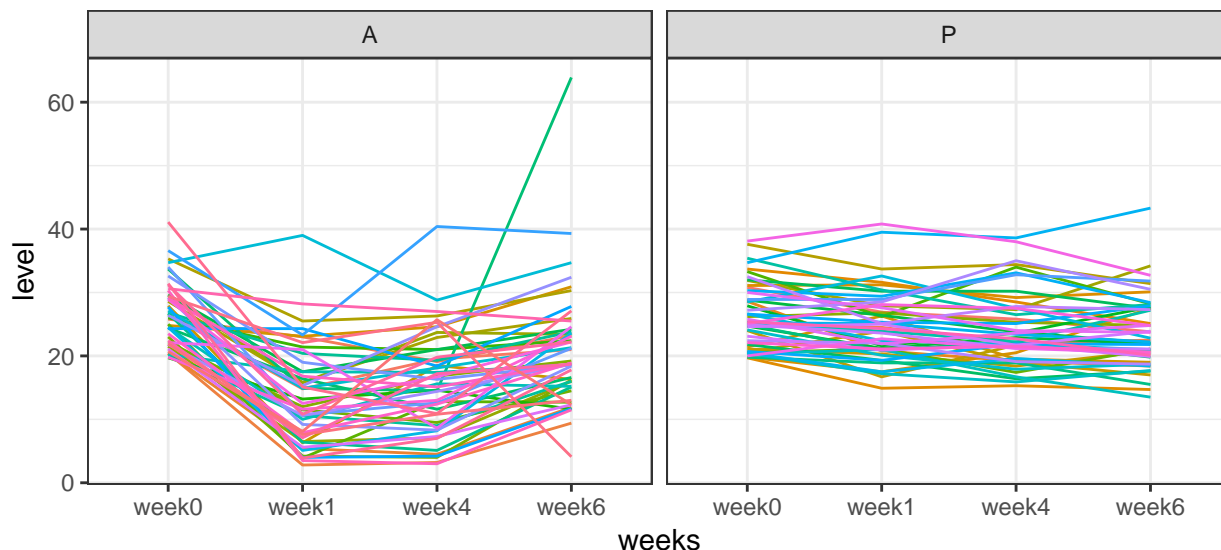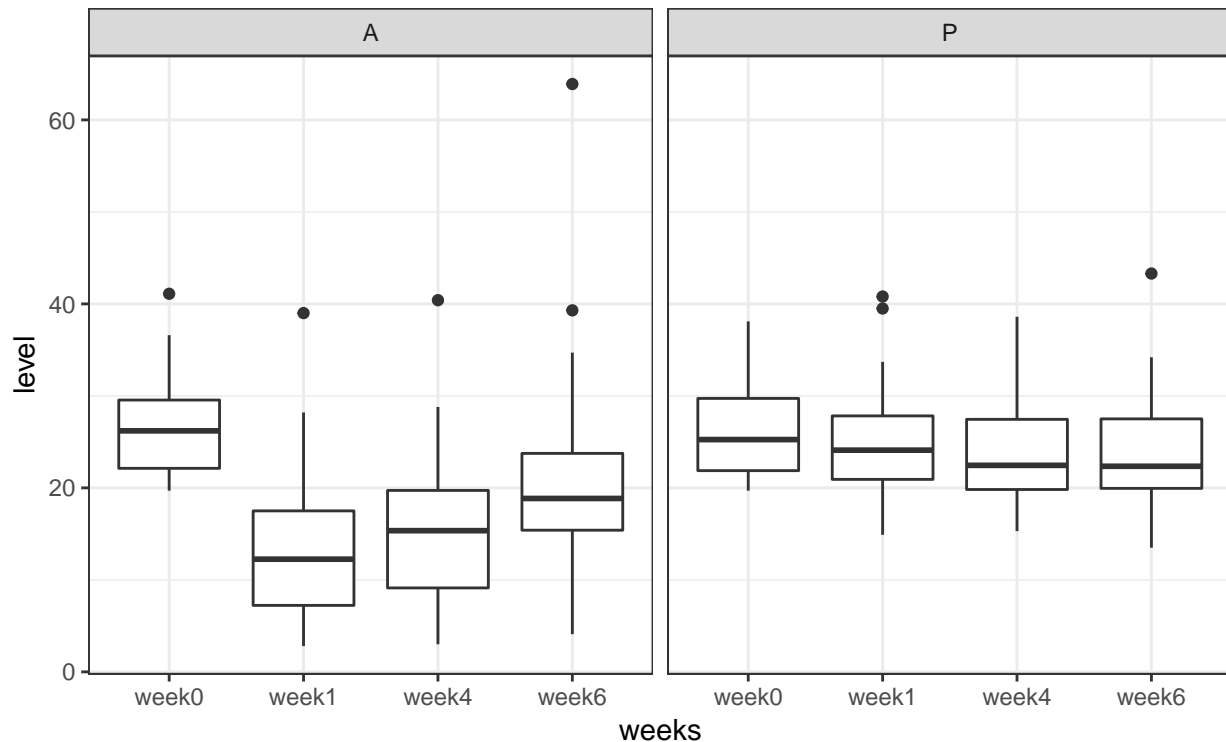
Figure1. spaghetti plot for TLC data

Figure2. boxplot for TLC data mean and s.d.

Appendix

```r
knitr::opts_chunk$set(echo = FALSE, message = FALSE, warning = FALSE, comment = "")
library(tidyverse)
options(knitr.table.format = "latex")
theme_set(theme_bw())
# EDA for TLC data
TLC = read.table("TLC.dat", header = F) %>%
  rename(id = V1, assignment = V2, week0 = V3, week1 = V4, week4 = V5, week6 = V6) %>%
  mutate(id = as.factor(id)) %>%
  as.tibble()


# scatterplot
TLC %>% gather(week0:week6, key = "weeks", value = "level") %>%
  ggplot(aes(x = weeks, y = level, group = id, color = id)) +
  geom_line() +
  facet_grid(. ~ assignment) +
  theme(legend.position = "none")
# boxplot about mean and sd
TLC %>% gather(week0:week6, key = "weeks", value = "level") %>%
  ggplot(aes(x = weeks, y = level, group = weeks)) +
  geom_boxplot() +
  facet_grid(. ~ assignment) +
  theme(legend.position = "none")
```