Design of experiments: Lab 2

Exercise 1. [From Montgomery (2009).] During cooking, doughnuts absorb fat in various amounts. A scientist wished to learn if the amount absorbed depends on the type of fat used. For each of four fats, six batches of doughnuts were prepared. The data in the table below are the grams of fat absorbed per batch, coded by deducting 100g to give simpler figures.

T1	T2	T3	T4
64	78	75	55
72	91	93	66
68	97	78	49
77	82	71	64
56	85	63	70
95	77	76	68

a) Read the data into `R` and plot it. By looking at your plot, do you think that there will be significant differences between the types of fat?

b) Check that the assumptions of the one-way ANOVA model are satisfied.

c) Fit a one-way ANOVA model with the sum-to-zero constraint. Find point estimates and confidence intervals for the grand mean and the treatment effects. Comment on the results.

d) Is there evidence to claim that there are differences between types of fat at the $\alpha = 0.01$ significance level?

e) Run `TukeyHSD` to perform pairwise comparisons and comment on the results.

Exercise 2. In this exercise, we will work with three datasets where the assumptions of the one-way ANOVA model are not satisfied. Here is the code that reads the datasets:

```
dat1 = read.csv("http://vicpena.github.io/doe/dat1.csv")
dat2 = read.csv("http://vicpena.github.io/doe/dat2.csv")
dat3 = read.csv("http://vicpena.github.io/doe/dat3.csv")
```

In `dat3`, the experiments were performed sequentially in an order given by the variable `order`. Identify the assumptions that are not satisfied in each case.

Exercise 3. In the process of turning wood into paper sheets, phenol waste is produced. A group of technicians working at a paper company want to compare the average phenol waste produced by 3 different species of trees. The code below reads in the data, which contains the percentage of phenol waste produced in 15 experiments.

```
phenol = read.csv("http://vicpena.github.io/doe/phenol.csv")
```

a) Are there differences between tree species?

b) What assumptions did you make to draw your conclusions in part a)? Are they reasonable in this context?

c) The technicians ran all the experiments for type A first, then those for type B, and then those for type C. Do you think that's reasonable, or would you design the experiment differently?

d) Actually, the design was originally balanced, but the technicians excluded three observations from the analysis (with phenol waste percentages 3.6, 4.1, and 3.5). Do you think this is reasonable?

Exercise 4. [Ott (1973)] In the process of manufacturing an electronic circuit, an engineer noticed that there was more variability than expected between circuits. Having studied the problem, the engineer concluded that one of the sources of variability could be a flat piece of insulating ceramic that is placed in the circuits.

This flat piece is obtained from ceramic sheets that are purchased from an outside supplier. Each sheet is cut into small pieces, each of which is placed in a circuit. The engineer suspects that there is variation between sheets and that this difference affects the outcome.

You can read in the data with the following command:

```
ott = read.csv2("http://vicpena.github.io/doe/Ott_Case_13_1.csv")
```

To analyze whether the suspicion was founded, 6 sheets were taken and cut into 7 pieces each. Then, a circuit was assembled in each of the 42 pieces.

a) Plot the data and draw some preliminary conclusions about the influence of the sheet on the outcome.

b) Does the ceramic sheet influence the outcome?

c) Verify that the assumptions of the model are satisfied.

Exercise 5. [Montgomery (2009)] An experiment is conducted to investigate the effectiveness of 5 insulating materials. Four test pieces of each material are taken and subjected to a high potential difference to accelerate their time to failure. The times to failures (in minutes) are shown in the following table:

Material	Time to failure			
1	110	157	194	178
2	1	2	4	18
3	880	1256	5276	4355
4	495	7040	5307	10050
5	7	5	29	2

a) Plot the data and comment on what you see.

b) Do the materials have the same effect on the average time to failure?

c) Check the model assumptions. Do you think they are satisfied?

d) Try out taking a logarithmic transformation of the time to failure and running the analysis again. Comment on the results.

Exercise 6. A textile factory has 5 looms. Each loom is assumed to provide the same amount of fabric per minute. To investigate this hypothesis, the company recorded the Kg of fabric processed in one minute at the looms at different times. The data obtained are the following:

Loom	Output [Kg/min]				
1	14.0	14.1	14.2	14.0	14.1
2	13.9	13.8	13.9	14.0	14.0
3	14.1	14.2	14.1	14.0	13.9
4	13.6	13.8	14.0	13.9	13.7
5	13.8	13.6	13.9	13.8	14.0

a) Plot the data and fit a one-way ANOVA model. Check the assumptions of the model.

b) Is there evidence to reject the hypothesis that the looms produce the same amount of fabric?

c) Find an estimate of the variability between looms and compare it to the variability within looms.