Experimental Design
Homework 2
Due June, 5, 2023

- Discussion with fellow students during the preparation of the assignments does stimulate statistical thinking as well as its implementation. It is allowed to work in pairs, however, clearly indicate to whom the work belongs.
- Write your answer to each exercise in clear language. Include ONLY key R-output, such as tables, figures and analysis, necessary to answer the question. Also include key R-programming that you use to find your answer. Any output that is included need to be accompanied by a written interpretation.
- Deliver the pdf of your work into the dropbox of the BrightSpace course.
- Remark: The number of points are indicated in a box; ten are free.

1. In an experiment the resistance to corrosion is measured depending on heats and coating. Use the `corrosion.data` data from the `BHH2` library. Load it by

   ```
   data(corrosion.data)
   ```

   (a) $\boxed{2}$ Compute the mean and the SD for each heats by coating combination.

   (b) $\boxed{2}$ Analyse the data by two-factor analysis of variance and use F-testing to investigate for the presence of main and interaction effects. What are the conclusions?

   (c) $\boxed{2}$ Investigate the diagnostics of the fit of the best model for the data.

   (d) $\boxed{2}$ Use Tukey's LSD method to analyse the paired differences in the levels for heats.

   (e) $\boxed{2}$ Construct a mean with confidence intervals plot for each of the levels of heats.

2. A data set was obtained in the IC industry in which the thickness of the oxide coating layer was measured on three randomly selected sites in each of three wafers from each of eight lots. There are two nested grouping levels in this experiment: lot and wafer within lot. The data `Oxide` are available from the `nlme` library.
   Hint: Transform the data into a data frame to facilitate standard programming.

   ```
   dat <- as.data.frame(Oxide)
   ```

   (a) $\boxed{5}$ Test for the fixed effect of Source and Site.
   Hint: Use the `anova` function.

   (b) $\boxed{3}$ Estimate the model with the constant and only random lot effects and wafer within lot effects. Compute the confidence intervals for the standard deviations of the random effects.
   Hint: Use

   ```
   random = ~ 1 | Lot/Wafer
   ```

in the lme function from the `lme4` library to arrange for the nested properties in the design.

(c) [2] Construct a plot of the Thickness measurements by lot per wafer including a line representing the predicted values from the model. Is there from your observations a lot with the largest thickness?
Hint: Use `xyplot` from `lattice` library.

3. In line with a saying of Feynman: "If I cannot construct, I cannot understand", or to paraphrase a little more positively: "If I can construct, I understand".

   (a) [10] Construct a model of data with significant quadratic and interaction effect such that the yield is maximized.
   Hint 1: Use three levels for two factors an transform these into numeric.
   Hint 2: Construct the true part of the model with the quadratic and the interaction effects in such a way that the Hessian matrix become negative definite.

   (b) [5] Add a contour plot and a response perspective plot to illustrate your findings.
   Hint: It is convenient to combine the two figures in a single penal of plots.

4. Use the `VSGFS` data from the library `DoE.base` with the data from a Tageugi design.

   (a) [3] What do the numbers of the generalized words of lengths 3 and 4 indicate?

   (b) [2] Construct a correlation plot of the design. Are the main effects confounded with the 2-way interaction effects?

   (c) [5] Find a model for the most important main and two factor interactions for the outcome Yield. What is the fit of the model? Does inspection of model diagnostics indicate any presence of outliers?

5. The injection molding example from Box, Hunter and Hunter (1978), see also Box and Meyer (1993), is a 2^{8-4} fractional factorial design.

```
data(BM93.e3.data, package="BsMD")
```

   (a) [5] Compute a full 2 factorial model with a linear blocking effect and extract the model matrix X. Give the dimension of X. Use the inner product matrix X^TX to investigate identical columns.

   (b) [3] Extract all non-diagonal pairs from the X^TX matrix with value 20. Note that the are equal (up to sign).

   (c) [8] Compute a half normal plot with the effects and construct a model on the basis of this. Are there main or interaction effects? How well does the model fit the data? Do model diagnostics indicate any outliers?

   (d) [4] Provide two points of discussion on the approach.

SEE NEXT PAGE

6. Constructing and analysing designs.

   (a) [5] Construct a full factorial design 2^4 and compute its generalized words of lengths. What do you conclude?
   Hint: Use the library `DoE.base`.

   (b) [5] Construct a full factorial design 2^4 with 4 blocks and compute its generalized words of lengths. What do you conclude?

   (c) [5] Compute the three 2^{7-2}_{IV} designs in Chapter 8 Table 8.13 (Edition 8) and compare their lengths of their word patterns to verify Montgomery's conclusions.
   Hint: Use the `FrF2` function of the `FrF2` library.

   (d) [5] Construct a 2^{10-5} Design A with
   generators = c("ABCD","ABCE","ABDE","ACDE","BCDE"),
   and a 2^{10-6} Design B with
   generators = c("ABC","BCD","ACD","ABD","ABCD","AB").
   Give their properties. If estimation of main effect is deemed import, which of the would you prefer?

   (e) [5] Construct a 64 runs design with all (A+B+C)*(D+E+F+G+H) interactions estimable.