Summary

Design of experiments (DOE) The maximal score will be 60 points. The DOE report should be no longer than 8 pages. You will work individually since this assignment replaces the exam to a large extent, making up 60% of your final grade. The R code you use (including your collected observations) should be added to the report.

1. Use DOE to plan an experiment with at least 3 factors each at 2 levels. You need to do a minimum of 16 single experiments (2^3 in duplicate, 2^4 or fractions of 2^5 or higher). The response needs to be continuous so that the experiment can be analysed with linear regression.
2. Plan how the collection of data can be done, focus on genuine run replicates and randomization. Discuss if you need blocking.
3. Collect your data.
4. Analyse the data.
5. Report on findings.

About the exercise

The theme for the exercise is design of experiments (DOE). The purpose is to provide insight and training in planning, performing and analysing a statistical experiment, as well as to report the results.

The task

Carry out a *k*-factor two-level experiment where the goal is to determine how the various factors influence a response. You should yourself decide what kind of experiment to perform. This may be a laboratory experiment or be from a problem in your daily life.

Alternatively, you may do a different statistical analysis, using multiple linear regression or another suitable method, using your own data. In this case you should present a brief sketch to the course teacher before the project starts.

Important: You are strongly encouraged to collect at least 16 observations, for example in a full design with 4 factors, or do two repetitions of 3 factors. Why? With as few as 8 observations we will seldom see significant effects, and that would make it difficult for you to write about in the report.

Keywords

1. **Issues to be addressed:**
    - Describe the problem you want to study.
    - Why is this interesting?
    - What prior knowledge do you have?
    - What do you want to achieve?
2. **Selection of factors and levels:**
    - Which factors do you think are relevant to the problem described above?
    - Do you expect an interaction between some of the factors?
    - Which levels should be used, and why do you think these are reasonable?

- How can you control that the factors really are at the desired level?

3. **Selection of response variable:**
    - Which response variable will provide information about the problem described above?
    - Are there several response variables of interest?
    - How should the response be measured?
    - What can you say about the accuracy of these measurements?
4. **Choice of design:**
    - 2^k factorial.
    - 2^{k-p} fractional factorial or other design?
    - Desired resolution of the design?
    - Is it necessary or desirable to use a blocked design?
    - Is it necessary or desirable with replicates?
5. **Implementation of the experiment:**
    - Randomization.
    - Describe any problems with the implementation (maybe the randomization was not followed?).
    - Is each experiments a genuine run replicate, that is reflects the total variability of the experiment? (Each trial should be a performed independently and constitute a full trial.)
6. **Analysis of data**
    - Calculation of effects and assessment of statistical significance.
    - Check the assumptions. Important: residual plots!
7. **Conclusion and recommendations:**
    - Which conclusions can you draw from the experiment?
    - Interpretation of significant effects, main and interaction plots.
    - Remember that plots are illustrative and very useful for demonstrations.

Examples

List of the projects done in 2011. List of some of the projects in 2012 from the TMA4255 course. More examples from a US course. You may of course perform the same experiment as listed here!

Below are given some examples of problems that have been examined by former students:

- Baking of pie: Importance of the type of flour, type of berries and appearance (with or without cover), regarding taste, consistency and experience – rated by a taste panel.

- Treatment of welding to avoid fatigue: The importance of voltage, frequency and hammering on the life of a welding.

- Corona-based electric field probe: Importance of air flow across and through a probe that measures the electrical field, with application in warning systems for lightning in helicopters.

- Sound level for fireworks: Importance of quantity and fill volume of gunpowder, different mix ratio of gunpowder and wall thickness of the rockets, regarding sound level (dB).

- Maximum performance in weightlifting: Importance of kreatin-intake, physical exhaustion, position and grip on the number of lifts.

- Yield of pyridin: Importance of pH, % methanol, the number of equivalents $NaHSO_3$ and temperature on the yield of pyridine.

- Optimizing the use of heat pump in propane–propene distillation: Importance of pressure in the column, reflux ratio and temperature change over the vessel on energy costs.

- Perfect 'popping' of popcorn: Importance of the type of popcorn, the type and amount of oil in comparison to the amount of popcorn.

- Reaction speed in an S_N2 reaction between 1-brompropan and NaOH: Importance of reaction temperature, amount of solvent and start concentration of reactant on the reaction speed.

- Purification of wastewater: Meaning of focculation time, focculation intensity and sedimentation time for the remaining concentration of small particles in the cleaned water.

- Basketball shots: Importance of distance, type of defence and shot position on the number of points scored in basketball.

- Study of line widths in photoresist: Importance of exposure, development and baking on the line width.

- Development of fish food for use in conditioning of fish: Significance of the amount of alginat, the concentration of calcium chloride, stirring time after adding alginat and curing temperature on the consistency of fish.

- Economic aspects of burning candles: Importance of price, style and colour of the burning time.

- Running performance: effects of time of day, once speed and inclination is taken into account.