Factorial designs (Chapter 5 in the book)

Ex: We are interested in what affects pH in a liquide.

- pH is the response variable
- Choose the factors that affect

  amount of soda

  air flow ...
- Choose the number of levels of amount of soda and air flow, in the simplest case two levels for each factor.
- Run experiments for all possible combinations of the factor levels

  $\rightarrow$ factorial design

Two-level factorial design

Factors can be quantitative (two temperatures) or qualitative (presence or absence of some entity)

Two levels for each factor, experiments are run with all possible combinations of the factor levels.

Two-level factorial designs

- require relatively few runs per factor
- observations that are produced are quite easy to interpret by common sense, easy arithmetics or computer graphics
- can be used to determine directions for further experiments

Example: 2^2-design

Two factors are believed to control pH (response varaible) in the production of some drug

- S: amount of soda
- F: air flow

We have the following table

Factor	low (-)	high (+)
S	48kg	52kg
F	1m/s	2m/s

Measurements:

Points	S	F	y_i (pH)
1	-	-	8.65
2	+	-	8.81
3	-	+	8.80
4	+	+	8.94

How can we compute the effect of S and F on pH?

- Main effect: the effect a single factor has on the response variable.
- Interaction effect: there is interaction between two factors if the effect of increased value of one depends on the level of the other.

2^3-design

What if we believe that even temperature T affects pH and make a new design with three factors?

Points	S	F	T	pH (y_i)
1	-	-	-	8.64
2	+	-	-	8.78
3	-	+	-	8.80
4	+	+	-	8.95
5	-	-	+	8.64
6	+	-	+	8.82
7	-	+	+	8.76
8	+	+	+	8.97

Table of contrasts (design matrix)

(Mean)	S	F	T	SF	ST	FT	SFT	pH
+	-	-	-	+	+	+	-	y_1
+	+	-	-	-	-	+	+	y_2
+	-	+	-	-	+	-	+	y_3
+	+	+	-	+	-	-	-	y_4
+	-	-	+	+	-	-	+	y_5
+	+	-	+	-	+	-	-	y_6
+	-	+	+	-	-	+	-	y_7
+	+	+	+	+	+	+	+	y_8
8	4	4	4	4	4	4	4	

Estimated effects

- Main effects: $L_S = 0.17$, $L_F = 0.15$, and $L_T = 0.005$
- Two factor interaction effects: $L_{SF} = 0.01$, $L_{ST} = 0.025$, and $L_{FT} = -0.015$
- Three factor interaction effect: $L_{SFT} = 0.005$

Interpretation of the results

- The main effect should be individually interpreted only if there is no evidence that the factor interacts with other factors
  $\rightarrow$ the interactive factors should be considered jointly
- Use the graphical representation to see and compare the main effects and interaction effects
- If replications, the variance of the effects can be estimated. Sometimes the variance is estimated by considering the higher order interaction terms noise.

Normal plot (Ch 5.14)

Two problems in assessment from unreplicated factorial designs

a) sometimes meaningful higher order interactions

b) it is necessary to allow for selection

$\rightarrow$ Normal probability plot to decide which effects are real and which are a consequence of the process noise

How to plot a normal plot?

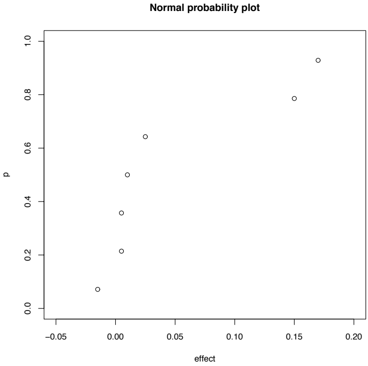
1) Order the effects from the smallest to the largest
2) Calculate

$$p_i = \frac{i - 0.5}{k} \cdot 100\%, \quad i = 1, ..., k,$$

where i is the "rank" number and k is the number of effects

3) Plot the effects against p_i in a normal plot.
4) Draw a line concentrating mostly on the points in the middle. The points that are far away from the line are exceptional (statistically significant)

Normal plot for the pH example



Blocking

Sometimes it is hard to maintain the same conditions in all experiments

- the experiment takes a long time to perform
- different patches of raw material
- different machines used
- different people used

These factors cannot be ignored (for example, one person may always obtain higher measurements than another even though the same tool was used).

Two ways to solve the problem

- Randomize the order of the experiments
  $\rightarrow$ difference between the block contributes the random effect
- Use blocking and randomize within blocks
  $\rightarrow$ the block effect will be balanced

Ex. Car factory with two lines. We have a 2^3-design. The 8 experiments are divided into 2 blocks (lines) of 4 experiments so that the "line-effect" (block-effect) does not influence the result. Three factors A, B and C, each having 2 levels

Design matrix

Run	A	B	C	AB	AC	BC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

Problem 5.12

a) Why do we block experimental designs?

b) Write a 2^3 factorial design in two blocks of four runs each such that no main effect or two-factor interaction is confounded with block differences.

c) Write a 2^3 factorial design in four blocks of two runs each such that main effects are not confounded with block effects.

Homework: estimating standard error

1) Genuine replicates (Ch 5.7, p.183-185)

2) Using higher order interactions (p.201, Table 5.11)

Estimating variation by using genuine replicates

A 2^3 factorial design with

- response variable "yield"
- three factors: quantitative factors temperature T and concentration C, and a qualitative factor type of catalyst K.
- two runs (replicates) of each experiment (combination of levels)
- order randomized

Interpretation of the results (Table 5.4)

- The main effects of T and C, and the interaction effect of T and K are large
- Effect of concentration C significant, changing from "low" to "high" reduces the yield by five units independently of the tested levels of the other factors T and K.
- The effect of temperature T and catalyst K cannot be interpreted separately because of the significant interaction effect. The effect of T is much higher with catalyst B (high level) than with catalyst A low level.

Estimating variation by ignoring higher order interactions

A 2^4 factorial design: process development study

- response variable "percent conversion"
- four factors: catalyst charge **1**, temperature **2**, pressure **3**, and concentration of one of the reactants **4**
- no replicated runs

Estimate variation by regarding the higher order (three and higher) interactions as noise.

Interpretation of the results (Table 5.11)

- The main effects of **1**, **2**, and **4**, and the interaction effect of **2** and **4** are large (significant)
- An increase in catalyst charge, factor **1**, from "low" to "high" reduces the conversion independently of the levels of the other factors.
- Effects of factors **2** and **4** have to be considered jointly due to the interaction: an increase in concentration, factor **4**, reduces conversion at the lower temperature, factor **2**, but does not give any effect at the higher temperature.