

# TMA4267 Linear Statistical Models V2017 (L20)

Part 4: Design of Experiments  
Fractional factorial designs  
Quiz with Kahoot!

Mette Langaas

Department of Mathematical Sciences, NTNU

To be lectured: March 30, 2017

## What did we learn last lesson?

- ▶ Why don't we want to perform a full factorial experiment, but instead a fractional factorial? (If we have many factors we maybe not need to be able to estimate all possible interactions, and may accept that effects are confounded.)
- ▶ What is the easiest way to design a half-fraction of a  $2^k$  factorial experiment? (Perform all the experiments where the highest order interaction = -1 or +1. E.g. for  $k=4$  we may do 16 different experiments, and now we only do the 8 possible experiments where  $ABCD=+1$ =defining relation. This is the same as thinking that  $D=ABC$ =generator).
- ▶ New words:
  - ▶ *generator(s)*=how to generate the design,
  - ▶ *defining relation(s)*, found from the generators,
  - ▶ *resolution*=length of shortest defining relation,
  - ▶ *alias structure*=confounding pattern, found by multiplying each effect of interest with the defining relation.
- ▶ Today: more on interpreting "confounding", interpreting "resolution" and more fractional factorial experiments

## Box, Hunter, Hunter: Reactor example

- ▶ A=feed rate (liters/min).
- ▶ B=Catalyst (%).
- ▶ C=Agitation rate (rpm).
- ▶ D=Temperature (deg C).
- ▶ E=Concentration (%).
- ▶ Response= (%) reacted.

Full factorial with  $2^5 = 32$  experiments.

From Box, Hunter, Hunter (1978, 2005): "Statistics for Experimenters", Ch.12.2.

## Half fraction with reactor example

- ▶ Instead of running a full factorial with  $2^5 = 32$  experiments,
- ▶ we suggest running a half-fraction.
- ▶ We choose  $I = ABCDE$  as the defining relation.

## Reactor data: answer in groups

	A	B	C	D	E	y						
1	-1	-1	-1	-1	-1	61	17	-1	-1	-1	-1	1 56
2	1	-1	-1	-1	-1	53	18	1	-1	-1	-1	1 63
3	-1	1	-1	-1	-1	63	19	-1	1	-1	-1	1 70
4	1	1	-1	-1	-1	61	20	1	1	-1	-1	1 65
5	-1	-1	1	-1	-1	53	21	-1	-1	1	-1	1 59
6	1	-1	1	-1	-1	56	22	1	-1	1	-1	1 55
7	-1	1	1	-1	-1	54	23	-1	1	1	-1	1 67
8	1	1	1	-1	-1	61	24	1	1	1	-1	1 65
9	-1	-1	-1	1	-1	69	25	-1	-1	-1	1	1 44
10	1	-1	-1	1	-1	61	26	1	-1	-1	1	1 45
11	-1	1	-1	1	-1	94	27	-1	1	-1	1	1 78
12	1	1	-1	1	-1	93	28	1	1	-1	1	1 77
13	-1	-1	1	1	-1	66	29	-1	-1	1	1	1 49
14	1	-1	1	1	-1	60	30	1	-1	1	1	1 42
15	-1	1	1	1	-1	95	31	-1	1	1	1	1 81
16	1	1	1	1	-1	98	32	1	1	1	1	1 82

- ▶ Which of the 32 experiments should be performed when  $I = ABCDE$  is the defining relation? What is then the generator?
- ▶ What is the resolution for this design?
- ▶ Write down the aliasing pattern.

## Resolution

*A design is said to be of resolution  $R$  if no  $p$ -factor effect is aliased with an effect containing less than  $R-p$  factors.*

A design of resolution

- III does not confound main effects with one another, but does confound main effects with two-factor interactions.
- IV does not confound main effects and two-factor interactions, but does confound two-factor interactions with other two-factor interactions.
- V does not confound main effects and two-factor interactions with each other, but does confound two-factor interactions with three-factor interactions and so on.

In general the resolution of a two-level fractional design is *the length of the shortest word in the defining relation.*

## Half fraction with reactor example: generator and defining relation

- ▶ Instead of running a full factorial with  $2^5 = 32$  experiments,
- ▶ we suggest running a half-fraction.
- ▶ We choose  $I = ABCDE$  as the defining relation.
- ▶ Alternative thinking:
  - ▶ Construct a full  $2^4$  design for A, B, C and D.
  - ▶ The column of signs for the ABCD interaction is written and used to define the levels for factor E.
  - ▶ This means  $E = ABCD$  is the generator for the design, and  $I = ABCDE$  is the defining relation.

R-code on course [www-page](#).

## Interpretation of confounding: example

Suppose there are three factors,  $A$ ,  $B$ ,  $C$ , for which we know the true effects and interaction effects:

$$A = 8$$

$$B = 20$$

$$C = 2$$

$$AB = 4$$

$$AC = 2$$

$$BC = 6$$

$$ABC = 4$$

Also is known that average response is 70.



## True regression model

The corresponding regression model is:

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_{12} z_{12} + \beta_{13} z_{13} + \beta_{23} z_{23} + \beta_{123} z_{123} + \epsilon$$

where  $z_{12} = z_1 z_2$ ,  $z_{13} = z_1 z_3$ ,  $z_{23} = z_2 z_3$ ,  $z_{123} = z_1 z_2 z_3$ , and where the coefficients  $\beta$  are half the corresponding effects, while  $\beta_0 = 70$ .

The regression model is hence

$$y = 70 + 4z_1 + 10z_2 + z_3 + 2z_{12} + z_{13} + 3z_{23} + 2z_{123} + \epsilon$$

In the following we shall also for simplicity assume that the errors  $\epsilon$  are 0. This makes it possible to compute the responses for any experiment for which the levels of A, B, C are specified.

## Confounding example (cont.)

Assume now that a  $2^{3-1}$  experiment is performed, with generator  $C = AB$ . And responses are computed using the true regression model (check!).

St. order		A	B	C=AB	AB	AC	BC	ABC	y
1	+	-	-	+	+	-	-	+	57
2	+	+	-	-	-	-	+	+	65
3	+	-	+	-	-	+	-	+	73
4	+	+	+	+	+	+	+	+	93
Coeff.	Const. 70	$z_1$ 4	$z_2$ 10	$z_3$ 1	$z_{12}$ 2	$z_{13}$ 1	$z_{23}$ 3	$z_{123}$ 2	

## Confounding example (cont.)

It is now seen that in all of these 4 experiments are

$$\text{Const.} = z_{123}$$

$$z_1 = z_{23}$$

$$z_2 = z_{13}$$

$$z_3 = z_{12}$$

so for the performed experiment we may as well write the model as

$$y = (\beta_0 + \beta_{123}) + (\beta_1 + \beta_{23})z_1 + (\beta_2 + \beta_{13})z_2 + (\beta_3 + \beta_{12})z_3$$

Using that we know the values of the coefficients, the true model for the data is thus

$$\begin{aligned} y &= (70 + 2) + (4 + 3)z_1 + (10 + 1)z_2 + (1 + 2)z_3 \\ &= 72 + 7z_1 + 11z_2 + 3z_3 \end{aligned}$$

## Confounding example (cont.)

- ▶ Suppose now that we try to compute the main effect of A from our data. Apparently this will be

$$l_A = \frac{65 + 93}{2} - \frac{57 + 73}{2} = 79 - 65 = 14$$

which is also found as twice the coefficient before  $z_1$  in the regression model above.

- ▶ Similarly, the apparent interaction effect of B and C would be computed as

$$l_{BC} = \frac{-57 + 65 - 73 + 93}{2} = 14$$

The truth (which is known to us) is, however, that  $A = 8$  and  $BC = 6$ , so that it is the sum of A and BC which is 14.

This is what is meant by saying that the main effect of A and the interaction effect between B and C are *confounded* (mixed). The confounded effects are listed in R as the *alias structure*.

Factorial Fit: y versus A; B; C

Estimated Effects and Coefficients for y (coded units)

Term	Effect	Coef
Constant		72,000
A	14,000	7,000
B	22,000	11,000
C	6,000	3,000

Alias Structure

I + A\*B\*C

A + B\*C

B + A\*C

C + A\*B

# The bicycle example

**TABLE 12.5.** An eight-run experimental design for studying how time to cycle up a hill is affected by seven variables (I = 124, I = 135, I = 236, I = 1237).

run	seat up/down 1	dynamo off/on 2	handlebars up/down 3	gear low/medium 4	raincoat on/off 5	breakfast yes/no 6	tires hard/soft 7	time to climb hill (sec) y
	12	13	23	123				
1	-	-	-	+	+	+	-	69
2	+	-	-	-	-	+	+	52
3	-	+	-	-	+	-	+	60
4	+	+	-	+	-	-	-	83
5	-	-	+	+	-	-	+	71
6	+	-	+	-	+	-	-	50
7	-	+	+	-	-	+	-	59
8	+	+	+	+	+	+	+	88

From Box, Hunter, Hunter (1978, 2005): "Statistics for Experimenters", Ch.12.25

## The bicycle example

- ▶ Set up a full factorial design in the three variables A, B, C.
- ▶ Use the generators:  $D=AB$ ,  $E=AC$ ,  $F=BC$ ,  $G=ABC$ .
- ▶ Defining relations:  $I=ABD=ACE=BCF=ABCG$ .
- ▶ The design is of resolution III.
- ▶ It is a  $1/16$  fraction of the full  $2^7$ , and thus called  $2^{7-4}_{III}$ .
- ▶ A design where every available contrast is associated with a factor is called a *saturated design*.

## Using FrF2 in R, see file L20.R

```
> plan <- FrF2(nruns=8,nfactors=7,
generators=c("AB","AC","BC","ABC"),alias.info=2,randomize=FALSE)
> plan
  A B C D E F G
1 -1 -1 -1 1 1 1 -1
2 1 -1 -1 -1 -1 1 1
3 -1 1 -1 -1 1 -1 1
4 1 1 -1 1 -1 -1 -1
5 -1 -1 1 1 -1 -1 1
6 1 -1 1 -1 1 -1 -1
7 -1 1 1 -1 -1 1 -1
8 1 1 1 1 1 1 1
class=design, type= FrF2.generators
> summary(plan)
Call:
FrF2(nruns = 8, nfactors = 7, generators = c("AB", "AC", "BC",
      "ABC"), alias.info = 2, randomize = FALSE)
Experimental design of type FrF2.generators
8 runs
Factor settings (scale ends):
  A B C D E F G
1 -1 -1 -1 -1 -1 -1 -1
2 1 1 1 1 1 1 1
Design generating information:
$legend
[1] A=A B=B C=C D=D E=E F=F G=G
$generators
[1] D=AB E=AC F=BC G=ABC
Alias structure:
$main
[1] A=BD=CE=FG B=AD=CF=EG C=AE=BF=DG D=AB=CG=EF E=AC=BG=DF F=AG=BC=DE G=AF=BE=CD
```



## Exam question on fractional factorials (K2014)

In a pilot study with four factors A, B, C and D, the 8 experiments listed below were run.

	A	B	C	D
1	-1	-1	-1	1
2	1	-1	-1	-1
3	-1	1	-1	-1
4	1	1	-1	1
5	-1	-1	1	1
6	1	-1	1	-1
7	-1	1	1	-1
8	1	1	1	1

What type of experiment is this?

What is the generator and the defining relation for the experiment?

What is the resolution of the experiment?

Write down the alias structure of the experiment.

# Not covered: Response Surface Methods

Dates back to the 1950s, with popular book by Box and Draper.

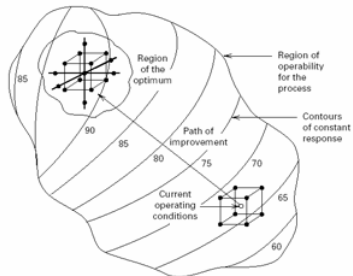


Figure 11-3 The sequential nature of RSM.

- ▶ The method performs sequential optimization, and can deal with several responses simultaneously.
- ▶ Central Composite Designs (CCD) and Box-Behnken Designs are two popular methods.
- ▶ John Tyssedal supervises 5th year project and master thesis in DOE.

<https://onlinecourses.science.psu.edu/stat503/node/57>

## Final word about the DOE Compulsory Exercise 4

- ▶ If you want to have 4 factors and perform 16 runs see R-code named <https://www.math.ntnu.no/emner/TMA4267/2017v/RscriptDOEtreadmill.R>
- ▶ If you want to have 3 factors, but need a block effect - look at this code <https://www.math.ntnu.no/emner/TMA4267/2017v/DOE2in3withrepl.R>, because it is best to code the block with effect coding - FrFr use treatment coding - and then we don't have orthogonal columns and everything becomes difficult...

## Summing up with Kahoot! quiz



[kahoot.it](https://kahoot.it)