Blocking

An experiment should always be randomized an each factor should be reset for each experimental run. In order to reduce variation caused by external factors, we can block our experiments. Randomization is then performed in each block.

A 2^3 experiment in two blocks

	A	B	C	ABC
1	−	−	−	−
2	+	−	−	+
3	−	+	−	+
4	+	+	−	−
5	−	−	+	+
6	+	−	+	−
7	−	+	+	−
8	+	+	+	+

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

If a number h is added to all runs in block 2, the calculated estimates of main-effects and two-factor interactions remain unchanged.

The 2^3 experiment can also be blocked in 4 blocks, each of size 2.

Suppose we block according to the sign combinations in AB og BC

Block 1	Block 2	Block 3	Block 4
(− −)	− +	+ −	+ +

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

AB, BC and AC are confounded with the block effect

A, B, C and ABC are not.

<u>How should we block</u>

Priority to be able to estimate main-effects and low-order interactions.

Notice $I = AA = BB = CC \ldots$

If we blocked using $D = ABC$ and $E = AC$. Their interaction $DE = ABC \cdot AC = B$ would also be confounded with the block effect, i tillegg til ABC og AC

Bloking in general

Assume we are going to block a 2^6 experiment in 8 blocks using $B_1 = ACE$, $B_2 = ABEF$, $B^3 = ABCD$

The following sign patterns will determine the blocks.

Block	1	2	3	4	5	6	7	8
	− − −	+ − −	− + −	+ + −	− − +	+ − +	− + +	+ + +

$B_1 B_2 = ACE\ ABEF = BCF$

$B_1 B_3 = ACE\ ABCD = BDE$

$B_2 B_3 = ABEF\ ABCD = CDEF$

$B_1 B_2 B_3 = ACE\ ABEF\ ABCD = ADF$

which together with ACE, ABEF and ABCD will be confounded by the block effect.

An analysis of variance table for 2^p experiment.

In a 2^p experiment we have $k = 2^p - 1$ orthogonal effect columns. Conciding these as columns for regression variables we get.

$$SS_R = b_1^2 \sum_{i=1}^{n} x_{1i}^2 + b_2^2 \sum_{i=1}^{n} x_{2i}^2 + \cdots + b_k^2 \sum_{i=1}^{n} x_{ki}^2$$

For an experiment with n runs we have $\sum_{i=1}^{n} x_{ji}^2 = n$

$j = 1, 2, \ldots, k$

and $b_j = \dfrac{\text{effect}_j}{2}$

Hence $SS_R = \frac{\hat{A}^2 \cdot m}{4} + \frac{\hat{B}^2 \cdot m}{4} + \dots + \frac{\hat{ABC}^2 \cdots m}{4}$.

For a 2^3 experiment we get.

Source	SS	DF
A	$SS_A = 2\hat{A}^2$	1
B	$SS_B = 2\hat{B}^2$	1
C	$SS_C = 2\hat{C}^2$	1
AB	$SS_{AB} = 2\hat{AB}^2$	1
AC	$\vdots$	$\vdots$
BC		
ABC	$SS_{ABC} = 2\hat{ABC}^2$	1
Total	$SS_T = SS_R = \sum_{i=1}^{m}(y_i - \bar{y})^2$	7

These sum of squares tell how the total variation in the data can be explained by the variation caused by each effect.

Suppose we block the experiment using ABC as the block factor. The block factor has has to levels and the variation caused by the the block factor is ~~SS Block~~

$SS_{Block} = SS_{ABC}$. So we substitute SS_{Block} for SS_{ABC} in the Analysis of variance table

What about a 2^3 experiment in 4 blocks using AB and AC as blocking factors.

Then also BC is confounded with blocks and

$SS_{Block} = SS_{AB} + SS_{AC} + SS_{BC}$

Analysis of variance table

Source	SS	DF
A	$SS_A = 2\hat{A}^2$	1
B	$SS_B = 2\hat{B}^2$	1
C	$SS_C = 2\hat{C}^2$	1
Block	$SS_{Block} = 2\hat{AB}^2 + 2\hat{AC}^2 + 2\hat{BC}^2$	3
ABC	$SS_{ABC} = 2\hat{ABC}^2$	1
Total	$SS_T = \sum_{i=1}^{n} (y_i - \bar{y})^2$	7

Fractional Factorial Two-level designs

Can we investigate more than p factors i 2^p runs

For instance 3 factors in 4 runs

The 2^3 experiment

Run	A	B	AB	C	AC	BC	ABC	
1	−	−	+	−	+	+	−	y_1
2	+	−	−	−	−	+	+	y_2
3	−	+	−	−	+	−	+	y_3
4	+	+	+	−	−	−	−	y_4
5	−	−	+	+	−	−	+	y_5
6	+	−	−	+	+	−	−	y_6
7	−	+	−	+	−	+	−	y_7
8	+	+	+	+	+	+	+	y_8

Let us pick the 4 runs with a + in ABC

Run	A	B	AB	C	AC	BC
2	+	−	−	−	−	+
3	−	+	−	−	+	−
5	−	−	+	+	−	−
8	+	+	+	+	+	+

For these 4 runs, $I = ABC \rightarrow$ A and BC have the same signs.

Also B and AC and C and AB

From the 2^3 experiment

$$\hat{A} = \frac{y_2 + y_4 + y_6 + y_8 - (y_1 + y_3 + y_5 + y_7)}{4}$$

$$\hat{BC} = \frac{y_1 + y_2 + y_7 + y_8 - (y_3 + y_4 + y_5 + y_6)}{4}$$

From the 4 run ~~off~~ in the 2^{3-1} experiment we get

$$l_A = \frac{y_2+y_8}{2} - \frac{(y_3+y_5)}{2} = \hat{A} + \hat{BC}$$

In the same way

$$l_B = \hat{B} + \hat{AC}$$

$$l_C = \hat{C} + \hat{AB}$$

such that A is aliased with BC and B with AC and C with AB

$$l_I = \frac{1}{4}\sum_{ABC+} y_i = \frac{1}{8}\sum_{i=1}^{8} y_i + \frac{1}{8}\sum_{ABC+} y_i - \frac{1}{8}\sum_{ABC-} y_i$$

$$= \bar{y} + \frac{ABC}{2}$$

Generator and defining relation

The 2^{3-1} design can be generated by first constructing a 2^2 experiment in A and B and thereafter letting $C = AB$.

$C = AB$ is called the generator for the design.

$C = AB \Leftrightarrow CC = I = ABC$. I is called the defining relation. To find out which effects that are aliased (have the same signs) we can multiply by the defining relation.

$$AI = AABC = BC \Rightarrow A \equiv BC$$

$$BI = BABC = AC \Rightarrow B \equiv AC$$

$$CI = CABC = AB \Rightarrow C \equiv AB.$$