Factorial Experiments

Example

The connection between yield of a chemical process and the two factors temperature and concentration is to be investigated. Four experiments are conducted, where two values of each factor are used. This gives 4 possible level combinations of the two factors to investigate the yield. The experiment is given in the table below, where the observed responses (yield) are also given:

Experiment no.	Temperature	Concentration	Yield
1	160	20	60
2	180	20	72
3	160	40	54
4	180	40	68
	x_1	x_2	y

The appropriate linear regression model is

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon,$$

The design matrix X of this model is obviously:

$$X = \begin{bmatrix} 1 & 160 & 20 & 3200 \\ 1 & 180 & 20 & 3600 \\ 1 & 160 & 40 & 6400 \\ 1 & 180 & 40 & 7200 \end{bmatrix}$$

MINITAB fits the following model:

```
Regression Analysis: y versus x1; x2; x1x2

The regression equation is
y = - 14,0 + 0,500 x1 - 1,10 x2 + 0,00500 x1x2

Predictor        Coef
Constant     -14,0000
x1           0,500000
x2           -1,10000
x1x2       0,00500000
```


Let us now recode the factors by introducing new independent variables

$$\begin{aligned} z_1 &= \frac{x_1 - 170}{10} \\ z_2 &= \frac{x_2 - 30}{10} \\ z_{12} &= z_1 \cdot z_2 \end{aligned}$$

The regression model is now

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_{12} z_{12} + \epsilon$$

with design matrix

$$X = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$


```
Regression Analysis: y versus z1; z2; z12

The regression equation is
y = 63,5 + 6,50 z1 - 2,50 z2 + 0,500 z12

Predictor      Coef
Constant    63,5000
z1          6,50000
z2         -2,50000
z12        0,500000
```

To see that we have the same fitted model, we can substitute the expressions for z_1, z_2, z_{12} in terms of the x_1, x_2, to get:

$$\begin{aligned} \hat{y} &= 63.5 + 6.5 \cdot \frac{x_1 - 170}{10} - 2.5 \cdot \frac{x_2 - 30}{10} + 0.5 \cdot \frac{x_1 - 170}{10} \cdot \frac{x_2 - 30}{10} \\ &= -14 + 0.5 x_1 - 1.1 x_2 + 0.005 x_1 x_2 \end{aligned}$$

Design of Experiments (DOE) terminology

In the example we consider two *factors*, A=temperature, B=concentration, and the response y=yield.

Each factor has two levels:

Factor	low	high
A	160^o (-1)	180^o (+1)
B	20^o (-1)	40^o (+1)

We have thus 2 factors which each can be on 2 levels, making $2^2 = 4$ possible combinations. The following is standard notation of such an experiment, a so called 2^2 experiment:

A	B	AB	Level code	Response
-1	-1	1	1	y_1
1	-1	-1	a	y_2
-1	1	-1	b	y_3
1	1	1	ab	y_4
z_1	z_2	z_{12}		

Multivariate regression with orthogonal design matrix X (Chapter 12.7 in book)

Consider the vector/matrix setup $y = X\beta + \epsilon$, or written out,

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{k1} \\ 1 & x_{12} & x_{22} & \cdots & x_{k2} \\ \vdots & & & & \\ 1 & x_{1n} & x_{2n} & \cdots & x_{kn} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{bmatrix}$$

We say that X has orthogonal columns if the product-sum of any two columns is 0. This means here that:

$$\sum_{i=1}^{n} x_{ji}x_{\ell i} = 0 \text{ when } j \neq \ell \ (j, \ell = 1, \ldots, k)$$

$$\sum_{i=1}^{n} x_{\ell i} = 0 \text{ for } \ell = 1, \ldots, k$$

A remarkable fact about the estimated regression coefficients in the above model is that each b_j depends on X only via the corresponding column for x_j, and that the estimated coefficients hence do not change when we look at submodels (i.e. take out variables from the model). The formulas are:

$$\begin{aligned} b_0 &= \bar{y} \\ b_j &= \frac{\sum_{i=1}^{n} x_{ji} y_i}{\sum_{i=1}^{n} x_{ji}^2} \text{ for } j = 1, 2, \ldots, k \end{aligned} \quad (3)$$

from which we get in particular

$$Var(b_j) = \frac{\sigma^2}{\sum_{i=1}^{n} x_{ji}^2} \quad \text{(prove it!)}$$

We also have:

$$SSR = b_1^2 \sum_{i=1}^{n} x_{1i}^2 + b_2^2 \sum_{i=1}^{n} x_{2i}^2 + \cdots + b_k^2 \sum_{i=1}^{n} x_{ki}^2 \quad (4)$$

Back to two-factor experiment

The regression model is now

$$y = \beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_{12} z_{12} + \epsilon$$

with design matrix

$$X = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

We get, using the formula in (3):

$$\begin{aligned} b_0 &= \frac{y_1 + y_2 + y_3 + y_4}{4} = 63.5 \\ b_1 &= \frac{-y_1 + y_2 - y_3 + y_4}{4} = \frac{y_2 + y_4}{4} - \frac{y_1 + y_3}{4} = 6.5 \\ b_2 &= \frac{-y_1 - y_2 + y_3 + y_4}{4} = \frac{y_3 + y_4}{4} - \frac{y_1 + y_2}{4} = -2.5 \\ b_{12} &= \frac{y_1 - y_2 - y_3 + y_4}{4} = \frac{y_4 - y_3}{4} - \frac{y_2 - y_1}{4} = 0.5 \end{aligned}$$

DOE terminology – main effects:

$$\begin{aligned}\hat{A} &= 2b_1 \\ &= \frac{y_2 + y_4}{2} - \frac{y_1 + y_3}{2} \\ &= \text{mean response when A is high} - \text{mean response when A is low}\end{aligned}$$

Similarly, the estimated effect of B is:

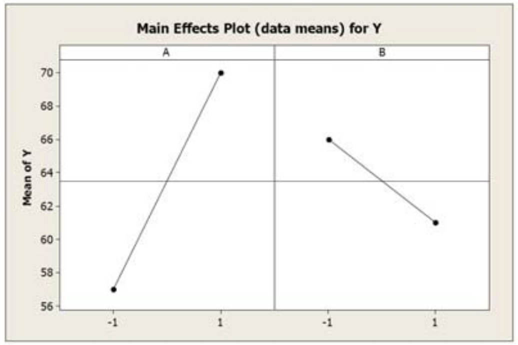
$$\begin{aligned}\hat{B} &= 2b_2 \\ &= \frac{y_3 + y_4}{2} - \frac{y_1 + y_2}{2} \\ &= \text{mean response when B is high} - \text{mean response when B is low}\end{aligned}$$

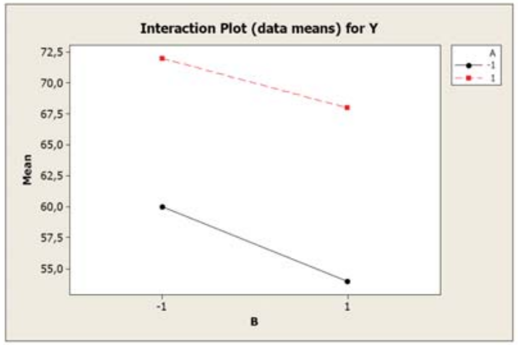
Interaction effects

Now what is the DOE interpretation corresponding to b_{12}? The answer is that $2b_{12}$ is denoted $\widehat{AB}$ and called the *estimated interaction effect between A and B*. We have the following motivation for this, where the last line is the general definition of a two-factor interaction:

$$\begin{aligned}\widehat{AB} &= 2b_1 \\ &= \frac{y_4 - y_3}{2} - \frac{y_2 - y_1}{2} \\ &= \frac{\text{estimated main effect of A when B is high}}{2} \\ &- \frac{\text{estimated main effect of A when B is low}}{2}\end{aligned}$$

$$\begin{aligned}\hat{A} &= \frac{72+68}{2} - \frac{60+54}{2} = 13 \\ \hat{B} &= \frac{54+68}{2} - \frac{60+72}{2} = -5 \\ \widehat{AB} &= \frac{68-54}{2} - \frac{72-60}{2} = 1\end{aligned}$$





Three factors

A	B	C	AB	AC	BC	ABC	Level code	Response
-	-	-	+	+	+	-	1	60
+	-	-	-	-	+	+	a	72
-	+	-	-	+	-	+	b	54
+	+	-	+	-	-	-	ab	68
-	-	+	+	-	-	+	c	52
+	-	+	-	+	-	-	ac	83
-	+	+	-	-	+	-	bc	45
+	+	+	+	+	+	+	abc	80
z_1	z_2	z_3	z_{12}	z_{13}	z_{23}	z_{123}		

The corresponding regression model is:

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

$$\begin{aligned}\hat{A} &= \frac{72+68+83+80}{4} - \frac{60+54+52+45}{4} = 23 \\ \hat{B} &= \frac{54+68+45+80}{4} - \frac{60+72+52+83}{4} = -5 \\ \hat{C} &= \frac{52+83+45+80}{4} - \frac{60+72+54+68}{4} = 1.5\end{aligned}$$

Two-factor interaction

$$\begin{aligned}
\widehat{AB} &= 2b_{12} \\
&= \frac{\text{estimated main effect of A when B is high}}{2} \\
&- \frac{\text{estimated main effect of A when B is low}}{2} \\
&= \frac{\frac{68+80}{2} - \frac{45+54}{2}}{2} - \frac{\frac{83+72}{2} - \frac{52+60}{2}}{2} \\
&= 1.5
\end{aligned}$$

Three-factor interaction

$$\begin{aligned}
\widehat{ABC} &= 2b_{123} \\
&= \frac{\text{estimated interaction between A and B when C is high}}{2} \\
&- \frac{\text{estimated interaction between A and B when C is low}}{2}
\end{aligned}$$

Or: using + and – in columns:

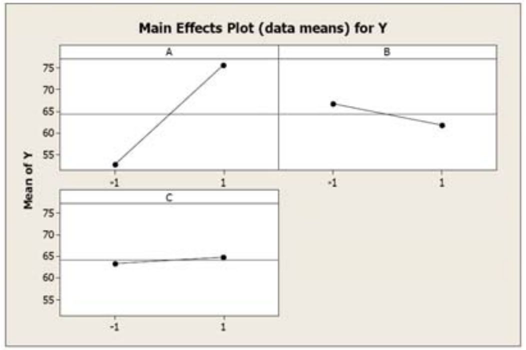
A	B	C	AB	AC	BC	ABC	Level code	Response
-	-	-	+	+	+	-	1	60
+	-	-	-	-	+	+	a	72
-	+	-	-	+	-	+	b	54
+	+	-	+	-	-	-	ab	68
-	-	+	+	-	-	+	c	52
+	-	+	-	+	-	-	ac	83
-	+	+	-	-	+	-	bc	45
+	+	+	+	+	+	+	abc	80

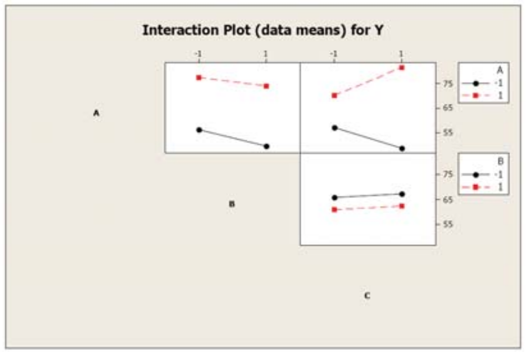
$$A\hat{B} = \frac{60+68+52+80}{4} - \frac{72+54+83+45}{4} = 1.5$$

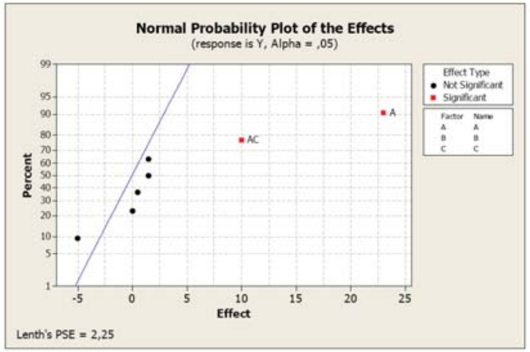
$$A\hat{C} = \frac{60+54+83+80}{4} - \frac{72+68+52+45}{4} = 10$$

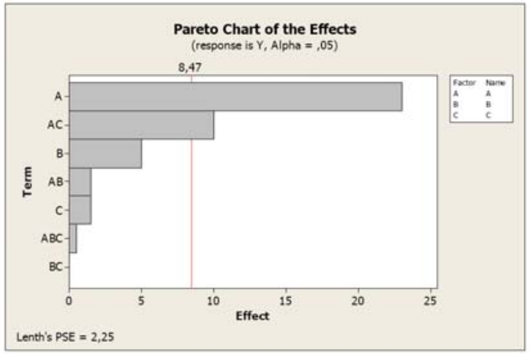
$$B\hat{C} = \frac{45+80+60+72}{4} - \frac{83+52++68+54}{4} = 0$$

$$AB\hat{C} = \frac{80+52+54+72}{4} - \frac{83+45+60+68}{4} = 0.5$$









Cube plot

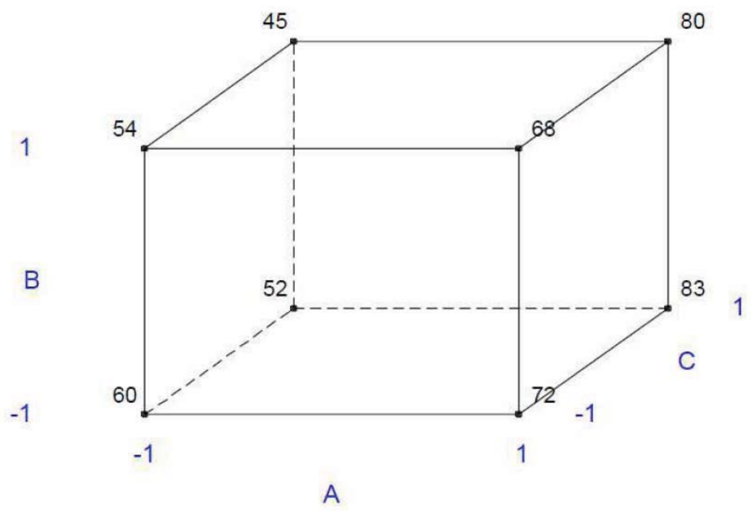
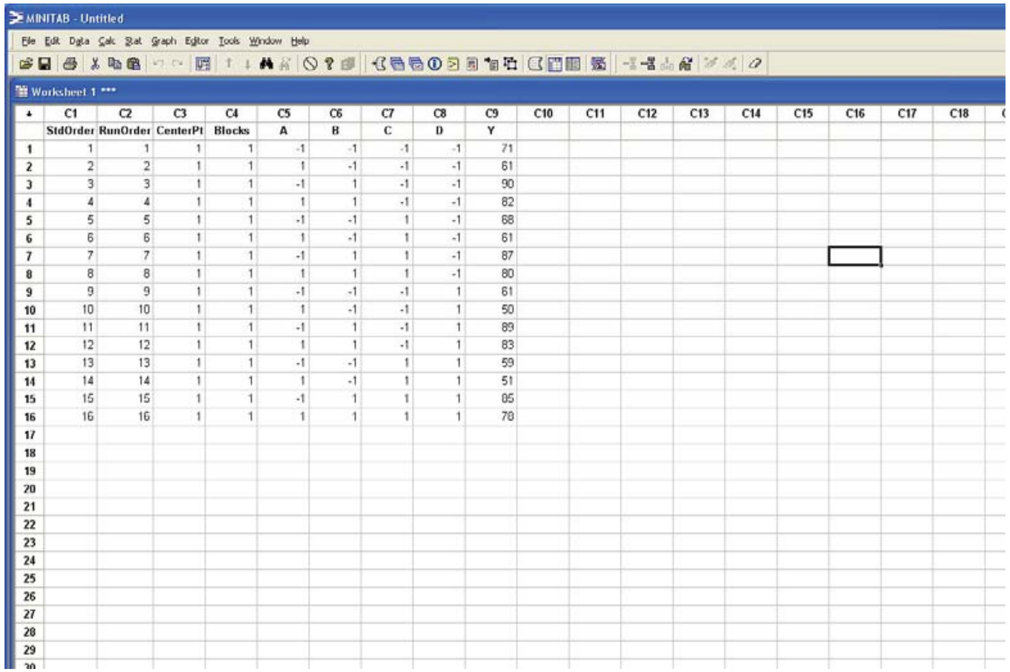


Figure 3: Cube plot of data in the table for Three factors

Four factors – example



	C1 StdOrder	C2 RunOrder	C3 CenterPt	C4 Blocks	C5 A	C6 B	C7 C	C8 D	C9 Y
1	1	1	1	1	-1	-1	-1	-1	71
2	2	2	1	1	1	-1	-1	-1	61
3	3	3	1	1	-1	1	-1	-1	90
4	4	4	1	1	1	1	-1	-1	82
5	5	5	1	1	-1	-1	1	-1	68
6	6	6	1	1	1	-1	1	-1	61
7	7	7	1	1	-1	1	1	-1	87
8	8	8	1	1	1	1	1	-1	80
9	9	9	1	1	-1	-1	-1	1	61
10	10	10	1	1	1	-1	-1	1	50
11	11	11	1	1	-1	1	-1	1	89
12	12	12	1	1	1	1	-1	1	83
13	13	13	1	1	-1	-1	1	1	59
14	14	14	1	1	1	-1	1	1	51
15	15	15	1	1	-1	1	1	1	85
16	16	16	1	1	1	1	1	1	78

Full Factorial Design

```
Factors:   4   Base Design:
4; 16
Runs:     16   Replicates:
1
Blocks:    1   Center pts
(total):      0


All terms are free from
aliasing.
```

Factorial Fit: Y versus A; B; C; D

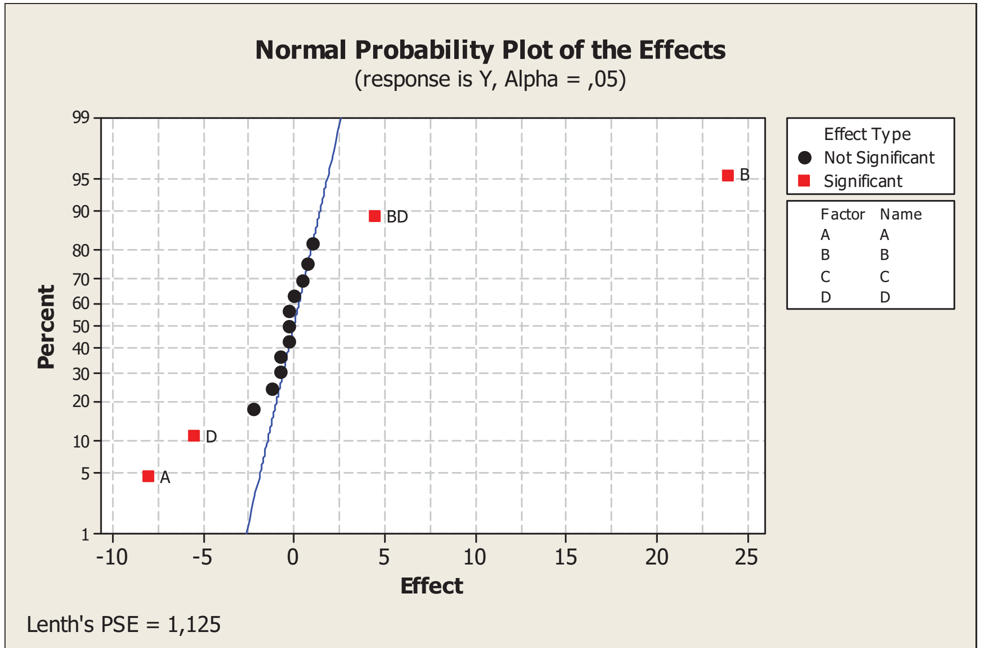
Estimated Effects and Coefficients for Y (coded units)

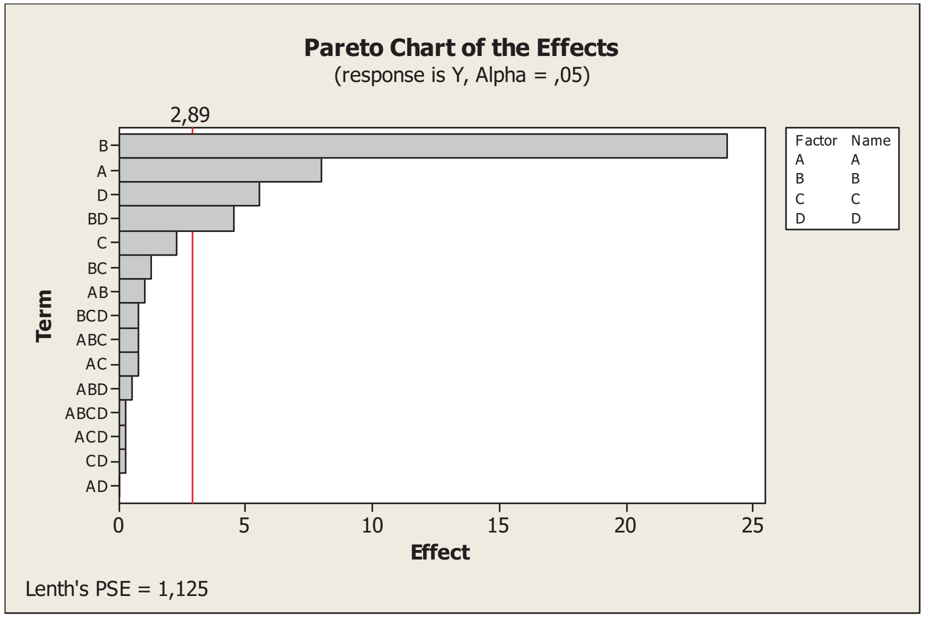
Term	Effect	Coef
Constant		72,250
A	-8,000	-4,000
B	24,000	12,000
C	-2,250	-1,125
D	-5,500	-2,750
A*B	1,000	0,500
A*C	0,750	0,375
A*D	-0,000	-0,000
B*C	-1,250	-0,625
B*D	4,500	2,250
C*D	-0,250	-0,125
A*B*C	-0,750	-0,375
A*B*D	0,500	0,250
A*C*D	-0,250	-0,125
B*C*D	-0,750	-0,375
A*B*C*D	-0,250	-0,125

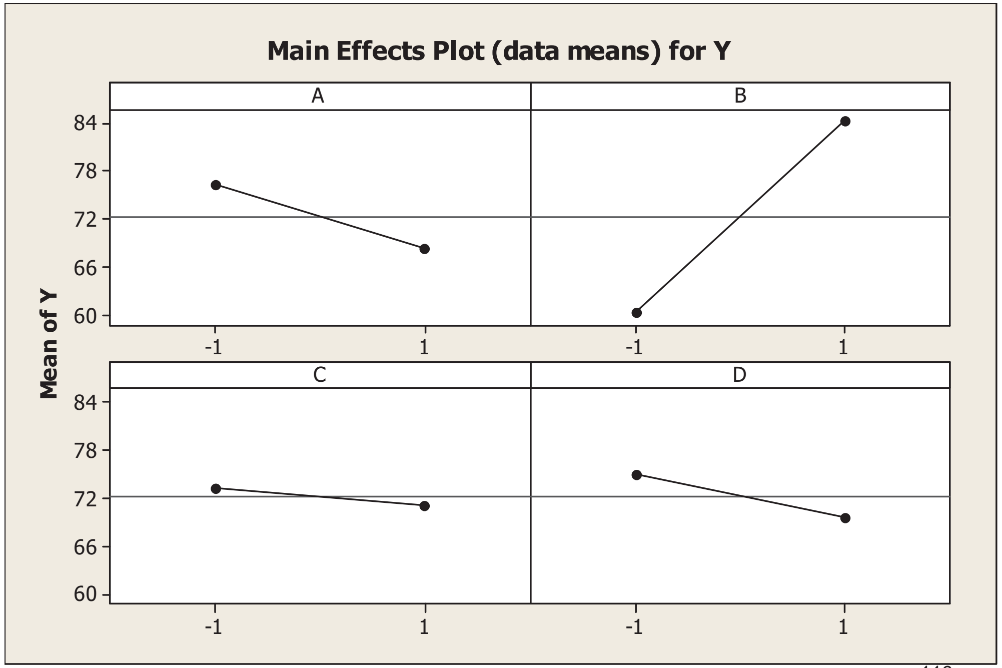
S = *

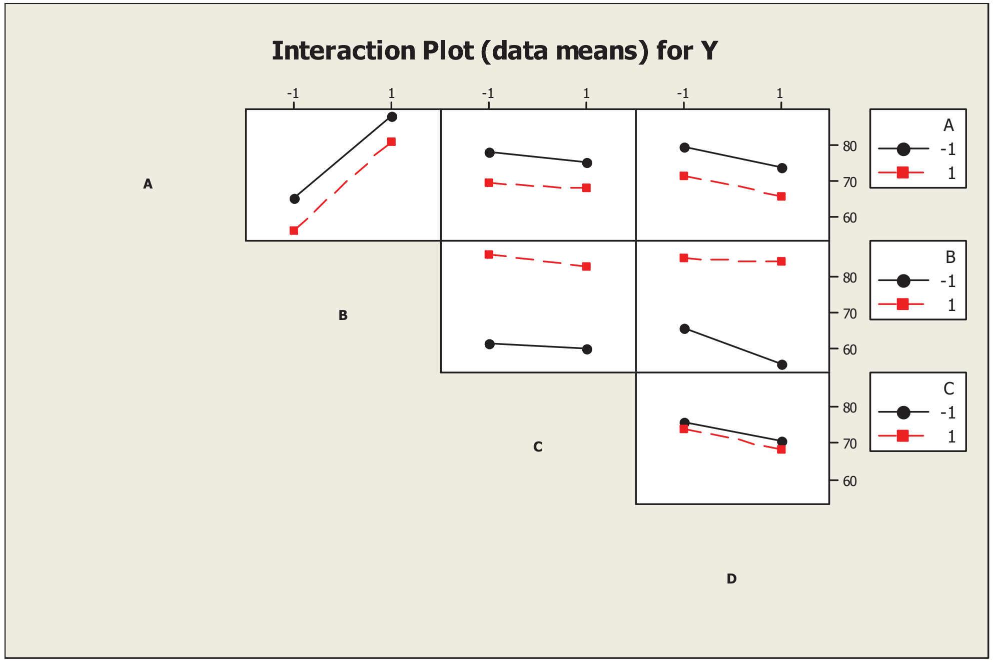
Analysis of Variance for Y (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	2701,25	2701,25	675,313	*	*
2-Way Interactions	6	93,75	93,75	15,625	*	*
3-Way Interactions	4	5,75	5,75	1,438	*	*
4-Way Interactions	1	0,25	0,25	0,250	*	*
Residual Error	0	*	*	*		
Total	15	2801,00				

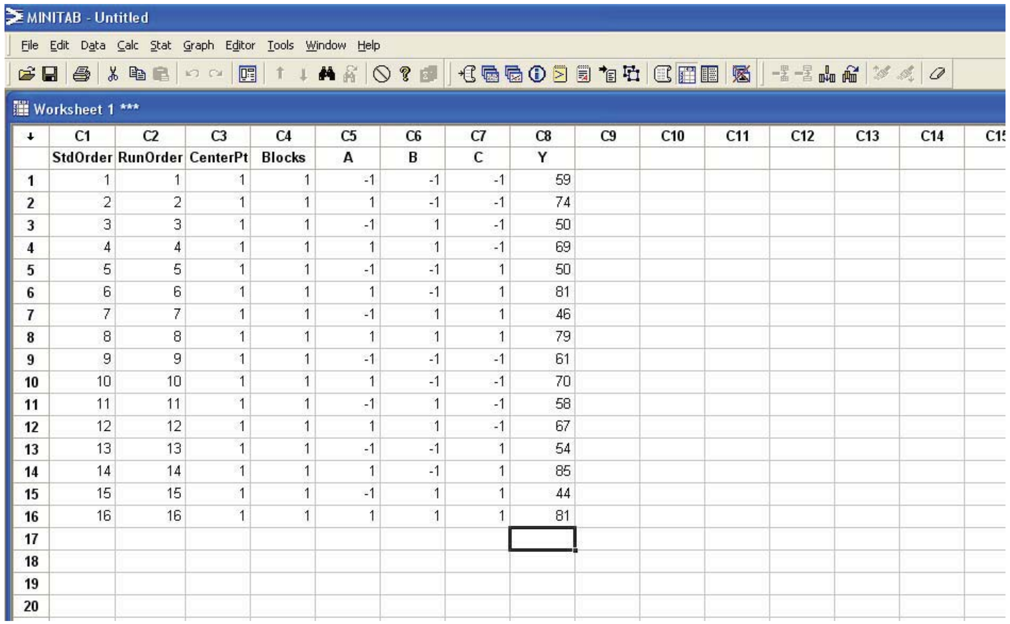








Example: Three factors and replicate



	C1	C2	C3	C4	C5	C6	C7	C8
	StdOrder	RunOrder	CenterPt	Blocks	A	B	C	Y
1	1	1	1	1	-1	-1	-1	59
2	2	2	1	1	1	-1	-1	74
3	3	3	1	1	-1	1	-1	50
4	4	4	1	1	1	1	-1	69
5	5	5	1	1	-1	-1	1	50
6	6	6	1	1	1	-1	1	81
7	7	7	1	1	-1	1	1	46
8	8	8	1	1	1	1	1	79
9	9	9	1	1	-1	-1	-1	61
10	10	10	1	1	1	-1	-1	70
11	11	11	1	1	-1	1	-1	58
12	12	12	1	1	1	1	-1	67
13	13	13	1	1	-1	-1	1	54
14	14	14	1	1	1	-1	1	85
15	15	15	1	1	-1	1	1	44
16	16	16	1	1	1	1	1	81

Factorial Fit: Y versus A; B; C

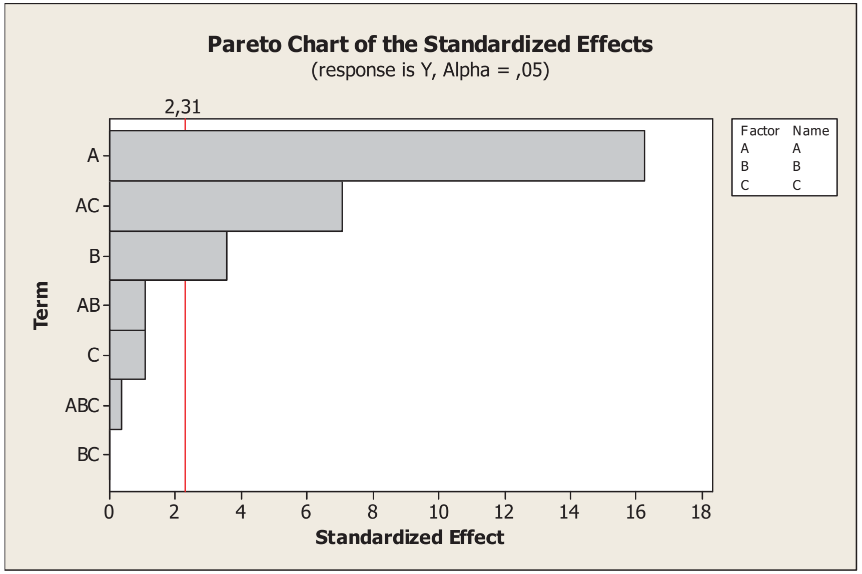
Estimated Effects and Coefficients for Y (coded units)

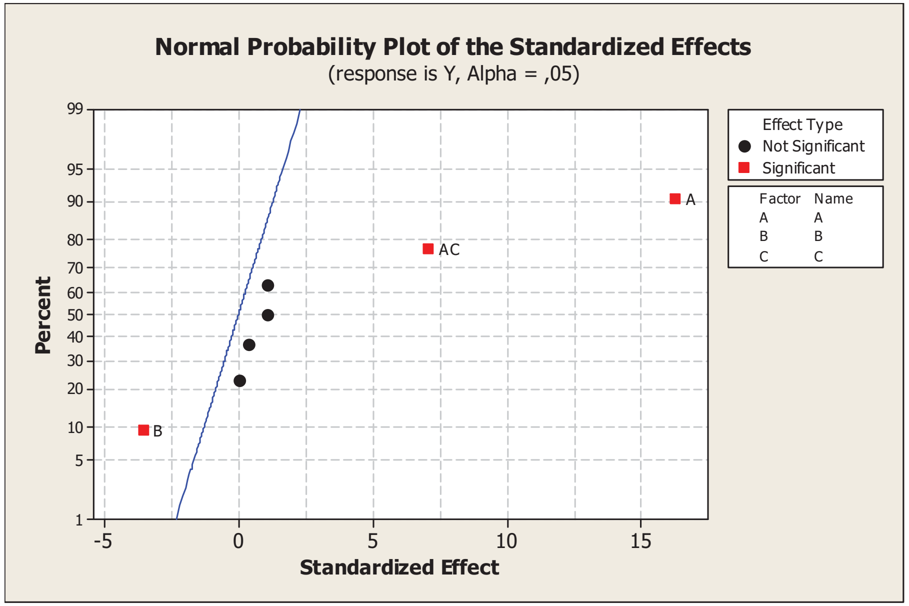
Term	Effect	Coef	SE Coef	T	P
Constant		64,250	0,7071	90,86	0,000
A	23,000	11,500	0,7071	16,26	0,000
B	-5,000	-2,500	0,7071	-3,54	0,008
C	1,500	0,750	0,7071	1,06	0,320
A*B	1,500	0,750	0,7071	1,06	0,320
A*C	10,000	5,000	0,7071	7,07	0,000
B*C	0,000	0,000	0,7071	0,00	1,000
A*B*C	0,500	0,250	0,7071	0,35	0,733

S = 2,82843   R-Sq = 97,63%   R-Sq(adj) = 95,55%

Analysis of Variance for Y (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	2225,00	2225,00	741,667	92,71	0,000
2-Way Interactions	3	409,00	409,00	136,333	17,04	0,001
3-Way Interactions	1	1,00	1,00	1,000	0,13	0,733
Residual Error	8	64,00	64,00	8,000		
Pure Error	8	64,00	64,00	8,000		
Total	15	2699,00				





Blocking in 2^k experiments

Full experiment:

StdO	A	B	C	AB	AC	BC	ABC
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

Two blocks:

Use column ABC as generator, i.e.

Block 1 consists of experiments with ABC = -1

Block 2 consists of experiments with ABC = 1

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

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

The interaction ABC is confounded (”mixed”) with the block effect. This means that the value of the estimated coefficient of ABC can be due to both interaction effect and block-effect.

Suppose all Y in block 2 are increased by a value h. Then the estimated effect of ABC will increase by h. But one cannot know from observations whether this is due to the interaction ABC or the block effect.

On the other hand, the estimated main effects A,B,C and the two-factor interactions AB,AC,BC are not changed by the h. These are of most importance to estimate, so the choice of blocking seems reasonable.

Four blocks in 2^3 experiment

Need two columns of +/- to define 4 blocks. Turns out that the best option is to use two two-factor interactions, e.g. AB and AC (which is default in MINITAB

Block 1: Experiments where AB = AC = -1
Block 2: Experiments where AB = -1, AC = 1
Block 3: Experiments where AB = 1, AC = -1
Block 4: Experiments where AB = AC = 1

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

Block structure is as follows:

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

Interaction effects AB and AC are confounded with the block effect, since they are generators for the blocks. In addition, their product AB*AC = AABC = BC is confounded with the block effect (Note: the BC column is constant within each block.

Adding h2 to block 2, h3 to block 3, h4 to block 4 does not change estimated effects of A,B,C, and also does not change the third order interaction ABC. However, e.g. AB will change by 2h3+2h4-2h2 and we do not know whether this is due to an interaction effect or blocking effect: This is CONFOUNDING.

How to determine which columns to use for blocking?

Idea: Try to leave estimates for main effects and low order interactions unchanged by blocking.

Note: I = AA = BB = CC where I is a column of 1's

Find the blocks for a 2^3 experiment using columns ABC and AC.
The interaction between ABC and AC is

ABC*AC = AA*B*CC = B

which is a main effect, which hence is confounded with the block effect (in addition to ABC and AC)

Generalisering

Gå ut i frå at vi skal dele eit 2^6 forsøk opp i 8 blokker etter blokkfaktorane $B_1 = ACE$ $B_2 = ABEF$ og $B_3 = ABCD$. Blokkinndelinga følgjer då følgjande mønster:

Blokk 1	Blokk 2	Blokk 3	Blokk 4	Blokk 5	Blokk 6	Blokk 7	Blokk 8
(- - -)	(+ - -)	(- + -)	(+ + -)	(- - +)	(+ - +)	(- + +)	(+ + +)

Vi får:

$$B_1B_2 = ACEABEF = BCF$$
$$B_1B_3 = ACEABCD = BDE$$
$$B_2B_3 = ABEFABCD = CDEF$$
$$B_1B_2B_3 = ACEABEFABCD = ADF$$

Som saman med $B_1 = ACE$, $B_2 = ABEF$ og $B_3 = ABCD$ blir konfundert med blokkeffekten.

Example obligatory project

"From a seed to a nice plant"

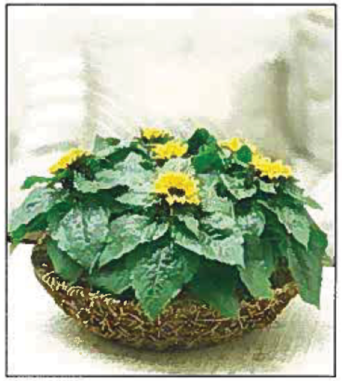


Figure 1.1 Box and 16 containers with the seeds: during the experience all the glasses were put inside the green box which was covered with a plastic film on the top to guarantee proper humidity conditions.

Factor	-	+
Seeds (A)	Broccoli Decicco	Sunflowers
Watering fluid (B)	Coffee	Water
Growth medium (C)	Soil	Cotton
Additional nutrients (D)	Without	With

StdOrder	RunOrder	CenterPt	Blocks	Seeds	Watering fluid	Growth medium	Additional nutrients	Length (response variable)
5	1	1	1	-1	-1	1	-1	0.1
2	2	1	1	1	-1	-1	-1	20.3
16	3	1	1	1	1	1	1	0.9
9	4	1	1	-1	-1	-1	1	0.2
15	5	1	1	-1	1	1	1	0.0
12	6	1	1	1	1	-1	1	6.9
6	7	1	1	1	-1	1	-1	1.1
1	8	1	1	-1	-1	-1	-1	11.7
10	9	1	1	1	-1	-1	1	5.9
13	10	1	1	-1	-1	1	1	0.0
4	11	1	1	1	1	-1	-1	23.3
8	12	1	1	1	1	1	-1	4.5
7	13	1	1	-1	1	1	-1	9.1
3	14	1	1	-1	1	-1	-1	12.2
14	15	1	1	1	-1	1	1	1.5
11	16	1	1	-1	1	-1	1	2.9

Table 3.1 Matrix of the design of experiments.


```
Estimated Effects and Coefficients for length (coded units)

Term        Effect     Coef
Constant               6,287
A            3,525    1,763
B            2,375    1,187
C           -8,275   -4,138
D           -8,000   -4,000
A*B         -0,675   -0,337
A*C         -3,825   -1,913
A*D         -0,500   -0,250
B*C          0,575    0,287
B*D         -1,600   -0,800
C*D          4,900    2,450
A*B*C       -0,875   -0,438
A*B*D        0,100    0,050
A*C*D        2,000    1,000
B*C*D       -1,650   -0,825
A*B*C*D      1,150    0,575
```

MINITAB plots

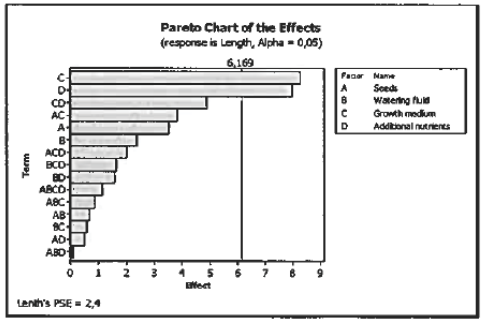


Figure 5.2 Pareto-chart of the effects with terms up to 4th order.

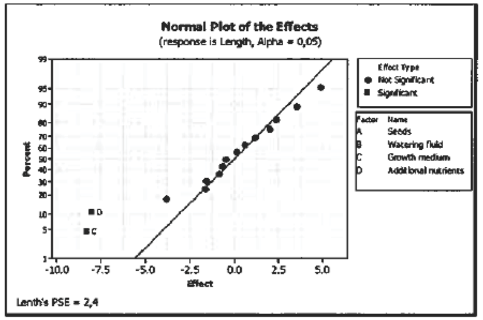


Figure 5.3 Normal plot of the effects with terms up to 4th order.

Assuming third and fourth order interactions are 0

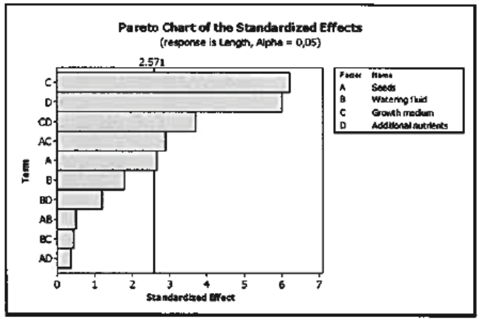


Figure 5.6 Pareto-chart of the effects with terms up to 2nd order.

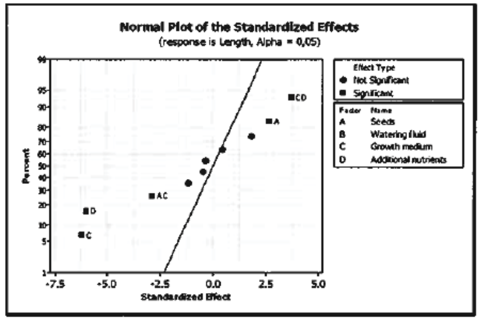


Figure 5.7 Normal plot of the effects with terms up to 2nd order.

Interaction plots

The plots of the interactions CD and AC are the following:

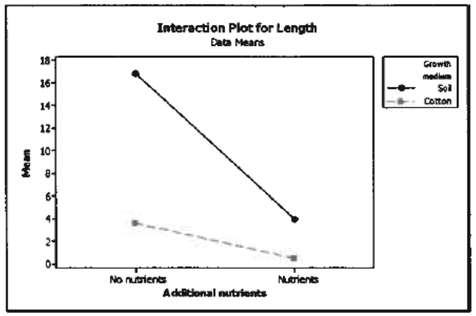


Figure 6.1 Interaction plot between growth medium and additional nutrients (CD).

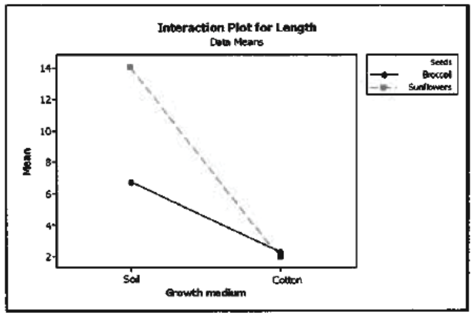


Figure 6.2 Interaction plot between seeds and growth medium (AC).

Fractional Factorial Designs at Two Levels

12.1. REDUNDANCY

Consider a two-level design in seven variables. A complete factorial arrangement requires $2^7 = 128$ runs. From these runs 128 statistics can be calculated, which estimate the following effects:

		interactions					
average	main effects	2-factor	3-factor	4-factor	5-factor	6-factor	7-factor
1	7	21	35	35	21	7	1

Now the fact that all these effects can be estimated does not imply that they all are of appreciable size. There tends to be a certain hierarchy. In terms of absolute magnitude, main effects tend to be larger than two-factor interactions, which in turn tend to be larger than three-factor interactions, and so on. This fact relates directly to the properties of smoothness and similarity discussed earlier. (In particular, for quantitative variables the main effects and interactions can be associated with the terms of a Taylor series expansion of a response function. Ignoring, say, three-factor interactions corresponds to ignoring terms of third order in the Taylor expansion.)

Fractional Factorial Design

Reactor Example i BHH kap. 12

```
Factors:   5   Base Design:         5; 16   Resolution:    V
Runs:     16   Replicates:              1   Fraction:    1/2
Blocks:    1   Center pts (total):      0


Design Generators: E = ABCD


Defining Relation:  I = ABCDE


Alias Structure

I + ABCDE

A + BCDE
B + ACDE
C + ABDE
D + ABCE
E + ABCD
AB + CDE
AC + BDE
AD + BCE
AE + BCD
BC + ADE
BD + ACE
BE + ACD
CD + ABE
CE + ABD
DE + ABC
```


Data Display

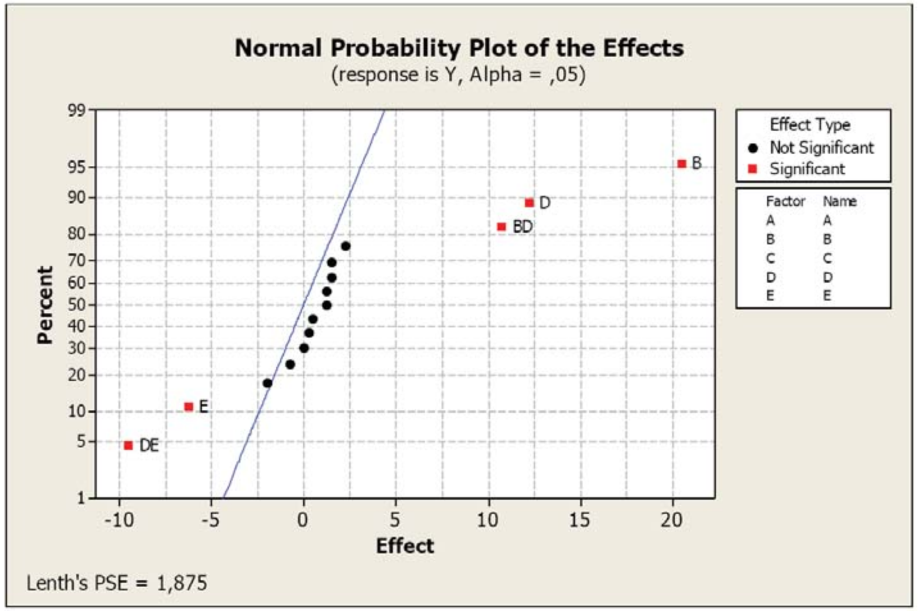
Row	A	B	C	D	E	Y
1	-1	-1	-1	-1	1	56
2	1	-1	-1	-1	-1	53
3	-1	1	-1	-1	-1	63
4	1	1	-1	-1	1	65
5	-1	-1	1	-1	-1	53
6	1	-1	1	-1	1	55
7	-1	1	1	-1	1	67
8	1	1	1	-1	-1	61
9	-1	-1	-1	1	-1	69
10	1	-1	-1	1	1	45
11	-1	1	-1	1	1	78
12	1	1	-1	1	-1	93
13	-1	-1	1	1	1	49
14	1	-1	1	1	-1	60
15	-1	1	1	1	-1	95
16	1	1	1	1	1	82

Factorial Fit: Y versus A; B; C; D; E

Estimated Effects and Coefficients for Y (coded units)

Term	Effect	Coef	"Fasit" fra fullt forsøk
Constant		65,250	65,5
A	-2,000	-1,000	-1,375
B	20,500	10,250	19,5
C	0,000	0,000	-0,625
D	12,250	6,125	10,75
E	-6,250	-3,125	-6,25
A*B	1,500	0,750	1,375
A*C	0,500	0,250	0,75
A*D	-0,750	-0,375	0,875
A*E	1,250	0,625	0,125
B*C	1,500	0,750	0,875
B*D	10,750	5,375	13,25
B*E	1,250	0,625	2,0
C*D	0,250	0,125	2,125
C*E	2,250	1,125	0,875
D*E	-9,500	-4,750	-11,0

S = *



From Exam in TMA4260 Industrial Statistics, december 2003, Exercise 2

A company decides to investigate the hardening process of a ballbearing production.
The following four factors are chosen:

- A: content of added carbon
- B: Hardening temperature
- C: Hardening time
- D: Cooling temperature.

Row	StdOrder	A	B	C	D	Hardhet
1	1	-1	-1	-1	1	15.32
2	2	1	-1	-1	-1	18.24
3	3	-1	1	-1	-1	17.18
4	4	1	1	-1	1	16.90
5	5	-1	-1	1	-1	15.95
6	6	1	-1	1	1	17.52
7	7	-1	1	1	1	14.26
8	8	1	1	1	-1	18.59

a) What is the generator and the defining relation of the design, and what is the design's resolution? Write down the alias structure.

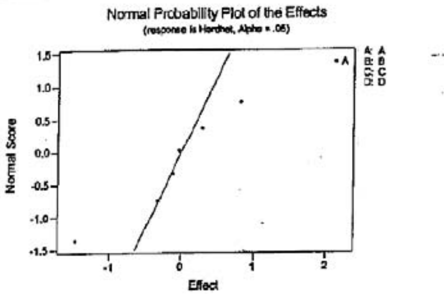
Find the estimates of the main effect of A and the interaction effect AC.

b) What is the variance of the main effect A and the interaction AC?

Assume that the st deviation sigma has been estimated from other experiments, by s = 0.312 with 9 degrees of freedom (in the exam, this had been done in Ex 1.) Use this estimate to find out whether the interaction AC is significantly different from 0 (i.e. "active") Use 5% significance level. What is the conclusion of the experiment so far?

Estimated Effects and Coefficients for Hardhet (coded units)

Term	Effect	Coef
Constant		16.7450
A	2.1350	1.0675
B	-0.0250	-0.0125
C	-0.3300	-0.1650
D	-1.4900	-0.7450
A*B	-0.1100	-0.0550
A*C	0.8150	0.4075
A*D	0.2850	0.1425



Analysis of Variance for Hardhet (coded units)

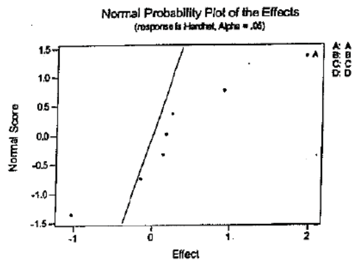
Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	13.776	13.776	3.4439	*	*
2-Way Interactions	3	1.515	1.515	0.5050	*	*
Residual Error	0	0.000	0.000	0.0000		
Total	7	15.291				

The company is well satisfied with the results so far and they decide to carry out also the other half fraction. The result of the other half fraction is given below.

Row	StdOrder	A	B	C	D	Hardhet
1	1	-1	-1	-1	-1	16.57
2	2	1	-1	-1	1	16.72
3	3	-1	1	-1	1	15.76
4	4	1	1	-1	-1	17.69
5	5	-1	-1	1	1	14.59
6	6	1	-1	1	-1	18.63
7	7	-1	1	1	-1	16.18
8	8	1	1	1	1	17.86

Estimated Effects and Coefficients for Hardhet (coded units)

Term	Effect	Coef
Constant		16.7500
A	1.9500	0.9750
B	0.2450	0.1225
C	0.1300	0.0650
D	-1.0350	-0.5175
A*B	-0.1450	-0.0725
A*C	0.9100	0.4550
A*D	0.1650	0.0825



Analysis of Variance for Hardhet (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	4	9.901	9.901	2.4753	*	*
2-Way Interactions	3	1.753	1.753	0.5842	*	*
Residual Error	0	0.000	0.000	0.0000		
Total	7	11.654				

Use this to find unconfounded estimates for the main effects and the two-factor interactions.

Assume that one would like to estimate the variance of the effects from the higher order interactions. Explain how this can be done, and find the estimate. Is it wise to include the four-factor interaction in this calculation? Why (not)?

Later, one of the operators that participated in the experiments asked whether one could have carried out the first half fraction in (a) in two blocks. This would, he said, have simplified considerably the performance of the experiments. What answer would you give to the operator?

From Exam in SIF 5066 Experimental design and…, May 2003, Exercise 1

A company making ballbearings experienced problems with the lifetimes of the products. In an experiments that they carried out they considered the factors

A: type of ball – standard (-) or modified (+)
B: type of cage - standard (-) or modified (+)
C: type of lubricate - standard (-) or modified (+)
D: quantity of lubricate – normal (-) or large (+)

The repsonse was the lifetime of the ballbearing in an accelerated life testing experiment. The results are given on the next page.

Forsøk	A	B	C	D	Y
1	-	-	-	-	0.31
2	-	+	+	-	0.92
3	+	+	+	+	2.57
4	+	-	-	+	1.38
5	+	+	-	-	2.17
6	-	+	-	+	0.73
7	-	-	+	+	0.95
8	+	-	+	-	1.37

A: type of ball
B: type of cage
C: type of lubricate
D: quantity of lubricate

a) What type of experiment is this? What is the defining relation? What is the resolution? Calculate estimates of the main effect of A and the two-factor interaction AB.

b) Estimated contrasts for B,C,D,AC,AD are, respectively, 0.60, 0.31, 0.22, -0.11, -0.01. What can you say about the estimated effects for CD, BD, BC. BCD, ACD, ABD,ABC?
Assume that factors C and D do not influence the response. Explain why this is then a 2^2 experiment with replicate. Calculate an estimate for the variance of the effects, and find out whether A, B and AB are now significant.

c) Give an interpretation of the results. The experiment was in fact carried out in two blocks, where experiments 1-4 was one block and 5-8 the other. How is this blocking constructed? How will we need to modify the analysis of significance in (b)? (Assume again that C,D do not influence the response)