TMA4255 Applied Statistics
Solution to Exercise 6

Problem 1

a) Main effect of z_1 = expected average response when z_1 is on the high level minus the expected average response when z_1 is on the low level.

$$\begin{aligned}\frac{y_4+y_2}{2}-\frac{y_3+y_1}{2} &= \frac{(\beta_0+\beta_1+\beta_2+\beta_{12})+(\beta_0+\beta_1-\beta_2-\beta_{12})}{2} \\ &- \frac{(\beta_0-\beta_1+\beta_2-\beta_{12})+(\beta_0-\beta_1-\beta_2+\beta_{12})}{2} = 2\beta_1\end{aligned}$$

Main effect of z_2 = expected average response when z_2 is on the high level minus the expected average response when z_2 is on the low level.

$$\begin{aligned}\frac{y_4+y_3}{2}-\frac{y_2+y_1}{2} &= \frac{(\beta_0+\beta_1+\beta_2+\beta_{12})+(\beta_0-\beta_1+\beta_2-\beta_{12})}{2} \\ &- \frac{(\beta_0+\beta_1-\beta_2-\beta_{12})+(\beta_0-\beta_1-\beta_2+\beta_{12})}{2} = 2\beta_2\end{aligned}$$

The interaction between z_1 and z_2: 1) half the main effect of z_1 when z_2 is on the high level minus 2) half the main effect of z_1 when z_2 is on the low level.

1) the main effect of z_1 when z_2 is on the high level:

$$\frac{(\beta_0+\beta_1+\beta_2+\beta_{12})-(\beta_0-\beta_1+\beta_2-\beta_{12})}{1} = 2\beta_1+2\beta_{12}$$

2) the main effect of z_1 when z_2 is on the low level:

$$\frac{(\beta_0+\beta_1-\beta_2-\beta_{12})-(\beta_0-\beta_1-\beta_2+\beta_{12})}{1} = 2\beta_1-2\beta_{12}$$

Then, the interaction between z_1 and z_2:

$$\frac{2\beta_1+2\beta_{12}}{2}-\frac{2\beta_1-2\beta_{12}}{2} = 2\beta_{12}$$

b) Main effect of z_1 while keeping z_2 at low level (we have already calculated above):

$$\frac{(\beta_0+\beta_1-\beta_2-\beta_{12})-(\beta_0-\beta_1-\beta_2+\beta_{12})}{1} = 2\beta_1-2\beta_{12}$$

Main effect of z_1 while keeping z_2 at high level (we have already calculated above):

$$\frac{(\beta_0+\beta_1+\beta_2+\beta_{12})-(\beta_0-\beta_1+\beta_2-\beta_{12})}{1} = 2\beta_1+2\beta_{12}$$

c) Based on the results in **b** we see that main effect of z_1 when z_2 is at its low level is $2\beta_1 - 2\beta_{12}$, and main effect of z_1 when z_2 is at its high level is $2\beta_1 + 2\beta_{12}$. If the interaction between z_1 and z_2 is zero we can find the main effect of z_1 by fixing the other factors at a given level. But, when there are interactions present, just fixing one factor at a given level will not give us estimate of the main effect, but the main effect and the interaction effect (as shown above). Therefore, we do not fix one factor at one level and vary the other factor in DOE!

Problem 2

a) We do the analysis in MINITAB:

```
Estimated Effects and Coefficients for C9 (coded units)

Term       Effect    Coef
Constant           17,544
A           8,837   4,419
B          -2,512  -1,256
C          -1,087  -0,544
D           0,112   0,056
A*B        -0,762  -0,381
A*C         1,013   0,506
A*D         0,212   0,106
B*C         1,012   0,506
B*D         0,262   0,131
C*D        -0,162  -0,081
A*B*C       0,213   0,106
A*B*D      -0,038  -0,019
A*C*D       1,387   0,694
B*C*D       0,288   0,144
A*B*C*D    -0,263  -0,131


S = *   PRESS = *


Analysis of Variance for C9 (coded units)

Source              DF   Seq SS   Adj SS   Adj MS  F  P
Main Effects         4  342,437  342,437  85,6094  *  *
2-Way Interactions   6   11,089   11,089   1,8481  *  *
3-Way Interactions   4    8,217    8,217   2,0544  *  *
4-Way Interactions   1    0,276    0,276   0,2756  *  *
Residual Error       0        *        *        *
Total               15  362,019
```

$$
\begin{aligned}
\hat{A} &= 8.84 \\
\hat{B} &= -2.51 \\
\hat{C} &= -1.09 \\
\hat{D} &= 0.11 \\
\vdots \;&\; \vdots \\
\widehat{ABCD} &= -0.262
\end{aligned}
$$

From the normal plot in figure(1) it looks like A and B are the most important factors.

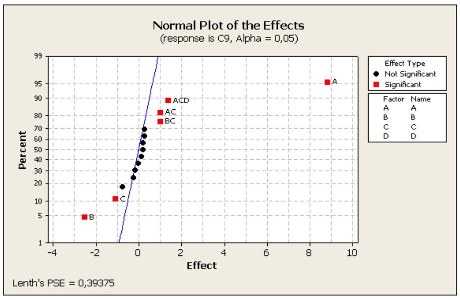

Figure 1: Normal plot a)

b) The corresponding regression model is

$$
\begin{aligned}
Y = \;&\beta_0 + \beta_1 z_1 + \beta_2 z_2 + \beta_3 z_3 + \beta_4 z_4 && (1)\\
&+ \beta_{12} z_1 z_2 + \beta_{13} z_1 z_3 + \beta_{14} z_1 z_4 && (2)\\
&+ \beta_{23} z_2 z_3 + \beta_{24} z_2 z_4 + \beta_{34} z_3 z_4 && (3)\\
&+ \beta_{123} z_1 z_2 z_3 + \beta_{124} z_1 z_2 z_4 + \beta_{134} z_1 z_3 z_4 && (4)\\
&+ \beta_{234} z_2 z_3 z_4 + \beta_{1234} z_1 z_2 z_3 z_4 + \epsilon && (5)
\end{aligned}
$$

And the estimated effects are of the kind

$$\hat{A} = 2b_1 \tag{6}$$

where b_1 is the least squares estimator of β_1. Same goes for the other effects.

c) In the analysis in a) we have 16 equations and 16 coefficients to estimate. Therefore there are no degrees of freedom left to estimate the variance. If we assume that the variance is known it is possible to make inference about the effects. For factor A we have:

$$
\left.
\begin{aligned}
&\hat{A} = \tfrac{1}{8}(-Y_1 + Y_2 - \cdots - Y_{15} + Y_{16}) \\
&\mathrm{Var}(\hat{A}) = \tfrac{1}{64} 16\sigma^2 = \tfrac{\sigma^2}{4}
\end{aligned}
\right\} \Rightarrow (\hat{A} \sim N(\mu_A, \frac{\sigma^2}{4})
$$

95 % confidence interval for μ_A:

$$\hat{A} \pm z_{0.025} \frac{\sigma}{2} = (6.88, 10.80)$$

95 % confidence interval for μ_B:

$$\hat{B} \pm z_{0.025}\frac{\sigma}{2} = (-4.47, -0.5)$$

d) If there are good reasons to assume that the 3- and 4-factor interactions are 0, we have enough degrees of freedom to estimate the variance.
From MINITAB we get:

```
Fractional Factorial Fit

Estimated Effects and Coefficients for Response (coded units)

Term         Effect     Coef  StDev Coef       T      P
Constant              17,544      0,3258   53,84  0,000
A             8,837    4,419      0,3258   13,56  0,000
B            -2,512   -1,256      0,3258   -3,86  0,012
C            -1,087   -0,544      0,3258   -1,67  0,156
D             0,112    0,056      0,3258    0,17  0,870
A*B          -0,762   -0,381      0,3258   -1,17  0,295
A*C           1,012    0,506      0,3258    1,55  0,181
A*D           0,212    0,106      0,3258    0,33  0,758
B*C           1,012    0,506      0,3258    1,55  0,181
B*D           0,262    0,131      0,3258    0,40  0,704
C*D          -0,162   -0,081      0,3258   -0,25  0,813

Analysis of Variance for Response (coded units)

Source                 DF      Seq SS     Adj SS      Adj MS      F      P
Main Effects            4     342,437    342,437      85,609  50,40  0,000
2-Way Interactions      6      11,089     11,089       1,848   1,09  0,473
Residual Error          5       8,493      8,493       1,699
Total                  15     362,019
```

We see that the estimator for σ^2 is now:

$$s^2 = MS_E = \frac{s_{ABC} + \cdots + s_{BCD} + s_{ABCD}}{5} = \frac{8.22 + 0.28}{5} = 1.70$$

where 8.22 is 3-way Seq SS, and 0.28 is 4-way Seq SS from the full analysis in section a. The variance of the effects is thus estimated by

$$s^2_{effect} = \frac{4s^2}{n} = 0.425$$

We can also obtain this estimate of σ^2_{effect} directly by using the estimated effects

$$s^2_{effect} = \frac{\widehat{ABC}^2 + \cdots + \widehat{BCD}^2 + \widehat{ABCD}^2}{5} = \frac{0.213^2 + 0.038^2 + 1.387^2 + 0.288^2 + 0.263^2}{5} = 0.425$$

Now we can do a T-test or an equivalent F-test to decide which of the effects are significant. We use the results and do an F-test:

$$F_A = \frac{MS_A}{MS_E} = \frac{s^2_A}{1.7} = \frac{(\hat{A}2^k/2)^2/2^k}{1.7} = \frac{312.37}{1.7} = 183.74$$

or alternatively, with $n = 2^k$,

$$F_A = \frac{MS_A}{MS_E} = \frac{\beta_A^2 n}{1.7} = \frac{(\hat{A}/2)^2 n}{1.7} = \frac{312.37}{1.7} = 183.74$$

$$F_B = \frac{MS_B}{MS_E} = \frac{s_B^2}{1.7} = \frac{25.26}{1.7} = 14.86,$$

and get the p-values:

$$p = P(F_{1,5} > 183.74) = 2P(T_5 > 13.56) \approx 0$$

$$p = P(F_{1,5} > 14.86) = 2P(T_5 > 3.85) = 0.012$$

Use that

$$\boxed{F_{1,\nu} = T_\nu^2}$$

): A has effect and B is significant at all levels > 0.012.

e) From MINITAB we get

```
Full Factorial Design

Factors:   4   Base Design:         4; 16   Resolution with blocks:  V
Runs:     16   Replicates:              1
Blocks:    2   Center pts (total):      0


Block Generators: ABCD


Alias Structure

I

Blk = ABCD

A
B
C
D
AB
AC
AD
BC
BD
CD
ABC
ABD
ACD
BCD


Design Table

Run  Block  A  B  C  D
  1      1  +  -  -  -
  2      1  -  +  -  -
  3      1  -  -  +  -
  4      1  +  +  +  -
  5      1  -  -  -  +
```

6	1	+	+	−	+
7	1	+	−	+	+
8	1	−	+	+	+
9	2	−	−	−	−
10	2	+	+	−	−
11	2	+	−	+	−
12	2	−	+	+	−
13	2	+	−	−	+
14	2	−	+	−	+
15	2	−	−	+	+
16	2	+	+	+	+

We see that ABCD is the only effect confounded with the block effect

f) To perform the experiment in two blocks, we need two generators. Choosing ABC and AD as generators gives

$$ABC \cdot AD = BCD \tag{7}$$

$$ABC \cdot BCD = AD \tag{8}$$

$$AD \cdot BCD = ABC \tag{9}$$

And we see that the effects confounded with the blocks are ABC, BCD and AD. This design avoids main efffects being confounded with the block effect.