MATH 531T-A : Exam 3

Louis Bensard

June 19, 2018

Problem 1:

Let O = Oil temperature, C = Percentage of Carbon, S = Steel temperature be the three factors and the response y = Percentage of Non-cracked Springs (we want to maximize the response). Let's use the OFAT approach:

- Assume O is the important factor and we fix $C = 0.5$ and $S = 1450$, then to maximize y, we need $O = 70$
- Assume C is the second most important factor and fix $O = 70$ and $S = 1450$, then to maximize y, we need $C = 0.50$
- Now fix $O = 70$ and $C = 0.50$, then to maximize y, we need $S = 1600$

Therefore, the OFAT approach gives us the setting $O = 70$, $C = 0.50$, $S = 1600$ with response $y = 79$, but it is clear that the optimal setting here is $O = 120$, $C = 0.50$, $S = 1600$ with response $y = 90$. Therefore this approach missed the optimal setting in this case. As a result, the OFAT approach can miss the optimal settings obtained from analyzing the factorial effects.

Problem 2:

(a)

Let T = Task factor and y the response. T has 2 levels, Task Y (low) = T- and Task Z (high) = T+. The main effect of T is:

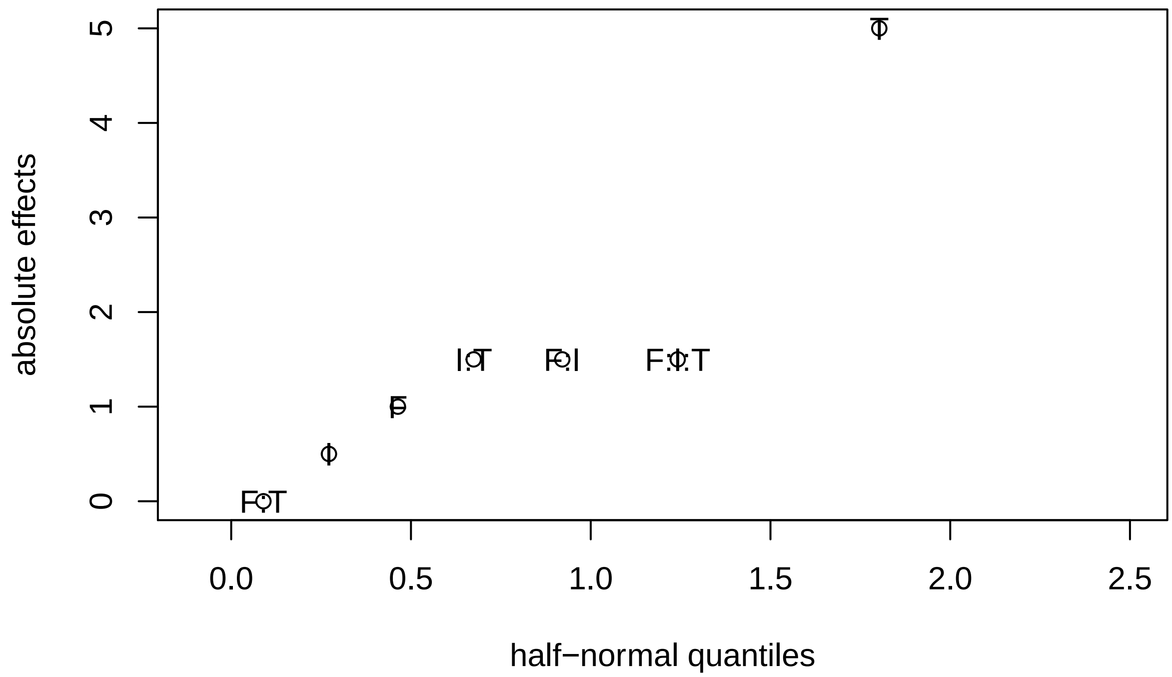
$$
\begin{aligned}
ME(T) &= \bar{y}(T+) - \bar{y}(T-) \\
&= \frac{16 + 14 + 15 + 19}{4} - \frac{11 + 12 + 10 + 11}{4} \\
&= 5
\end{aligned}
$$

(b)

Let's fit a linear model with all the factors and their interactions and use a half-normal plot to see if some of them are significant. F is the Flasher type and I is the Inertia of Lever, each of those factors have 2 levels. Here is the model:

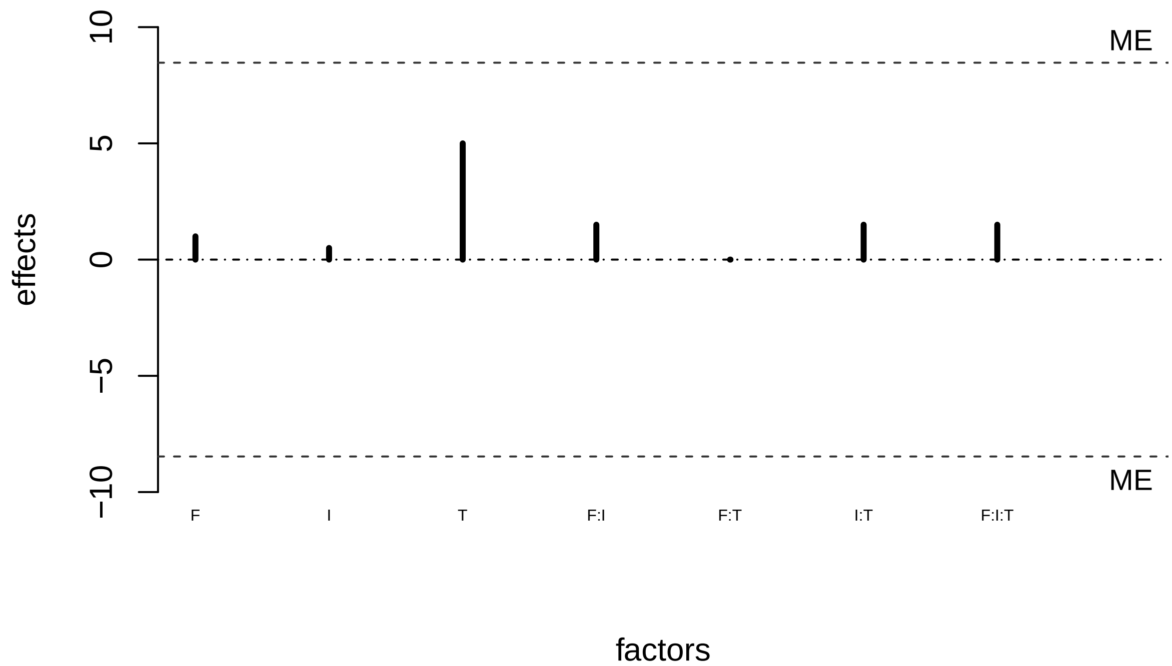
```
model = lm(y~F*I*T)
```

Here is the corresponding Half-Normal plot:



T clearly stands out of the line that we can draw through all of the remaining main effects and interactions, therefore T is significant (according to Half-Plot diagnosis).

Now let's use Lenth's method with $\alpha = 0.05$ for a more formal analysis of significance of factorial effects:




```
##  alpha    PSE     ME    SME
##  0.050  2.250  8.469 20.269
```

We can clearly see that all factorial effects are within the margin of error (ME), therefore, at $\alpha = 0.05$, Lenth's method disagrees with the visual diagnosis of the Half-Normal plot and does not detect any significant factorial effect.

(c)

Since the total number of observation is $N = 8$ and $\sigma_1 = 1$ and $\sigma_2 = 4$ are respecively the standard deviation on day 1 and day 2, we have:

$$Var(ME(T)) = \frac{\sigma_1^2}{N/2} + \frac{\sigma_2^2}{N/2} = \frac{1}{4} + \frac{16}{4} = 4.25$$

Problem 3:

(a)

All the interactions that are confounded with block effects are :

$$B_1 = 123, B_2 = 456, B_3 = 167$$
$$B_4 = B_1B_2 = 123456, B_5 = B_1B_3 = 2367$$
$$B_6 = B_2B_3 = 1457, B_7 = B_1B_2B_3 = 23457$$

Thus the blocking scheme is: {123,456,167,1457,2367,23457,123456} and the confounding pattern is: (0,0,3,2,1,1,0).

As a result, $g_2(b) = 0$, $g_3(b) = 3$,$g_4(b) = 2$, $g_5(b) = 1$, $g_6(b) = 1$ and $g_7(b) = 0$. Thus the order of estimability is $e_1 = 2$.

(b)

All the interactions that are confounded with block effects are :

$$B_1 = 1234, B_2 = 1256, B_3 = 1357$$
$$B_4 = B_1B_2 = 3456, B_5 = B_1B_3 = 2457$$
$$B_6 = B_2B_3 = 2367, B_7 = B_1B_2B_3 = 1467$$

Thus the blocking scheme is: {1234,1256,1357,3456,2457,2367,1467} and the confounding pattern is: (0,0,0,7,0,0,0).

As a result, $g_2(b) = 0$, $g_3(b) = 0$,$g_4(b) = 7$, $g_5(b) = 0$, $g_6(b) = 0$ and $g_7(b) = 0$. Thus the order of estimability is $e_2 = 3$.

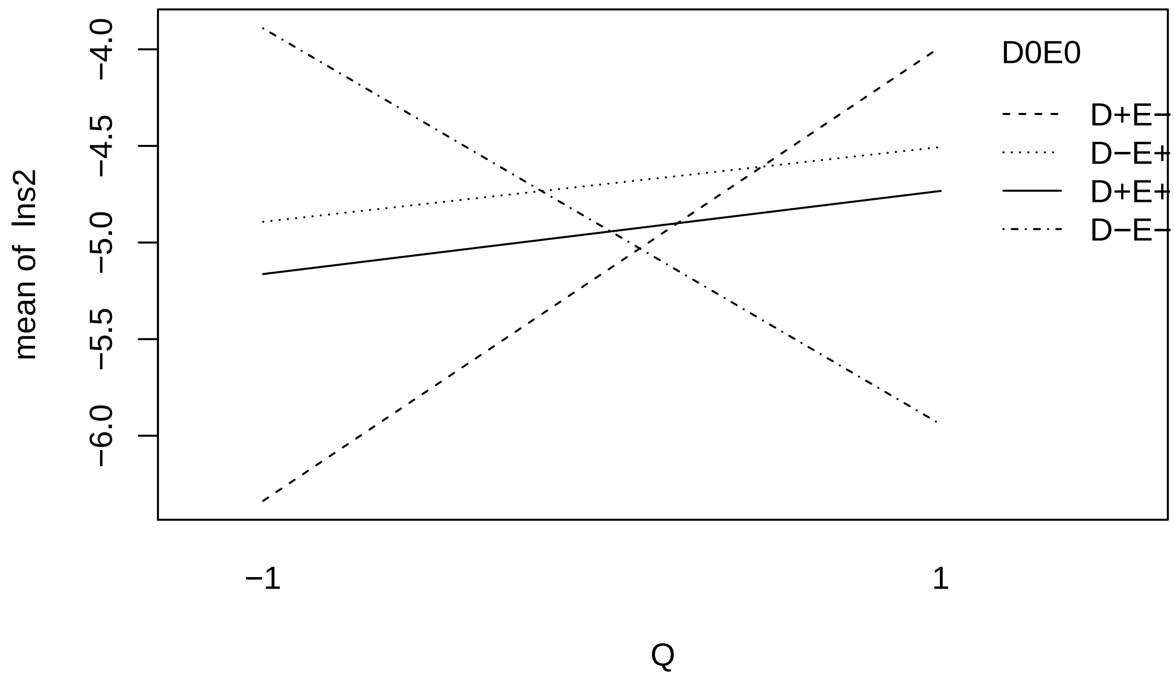
(c)

Scheme 2 is better than scheme 1 by the minimum aberration critera of the blocking scheme. The clear advantages of the second scheme is that all Main Effects, 2fi's and 3fi's are estimable and not sacrificed while in scheme 1, only the Main Effects and 2fi's are estimable.

Problem 4:

(a)

Here is the D x E x Q interaction plot (code in Appendix 1):



E at low level minimizes dispersion (along with D at high level and Q at low level).

(b)

Step (i):

$$ln(s^2) = -4.9313 + 0.9455x_B + 0.5556x_Dx_Q - 0.5445x_Dx_Ex_Q$$

To minimise s^2, B clearly needs to be at low level ($x_B = -1$). In (a), we found that E should be at low level ($x_E = -1$), D at high level ($x_D = 1$)and Q at low level ($x_Q = -1$).

(c)

$$\hat{y} = 7.6360 + 0.1106x_B + 0.0519x_E + 0.0881x_C - 0.1298x_Q + 0.0423x_Bx_Q - 0.0827x_Cx_Q$$

The only factor not in (i) is C, thus the adjustment factor is C.

(d)

Step (ii):

If we plug in the (-1/1) coded levels of the factors found in part (b) into $\hat{y}$, and set $\hat{y} = 8$ (the target value), we get $x_C = 2.0749$ which gives us (after decoding the value with $C+ = 25$ and $C- = 23$), C = 26.0749.

Note: The value of C is outside of the of its experimental range, C represents the heating time, so a greater value than the range most likely make sense but the predicted variance value for this setting may be too optimistic.

(e)

The values of $\hat{\sigma}^2$ are exactly the same in both analysis of the experiment as they have the same expression and the chosen levels are the same. Both predicted variances are $\hat{\sigma}^2 = e^{-4.9313-0.9455-0.5556-0.5445} = 0.00093$ (indeed very small, probably too optimistic). Also, since we solve the factors for $\hat{y} = 8$, both values of $\hat{y}$ are very close from 8. In the class analysis, $\hat{y} = 8.0001$ while here, $\hat{y} = 7.99999$.

In the class analysis, we found $E = 3.77$ while the range of E is 2-3, therefore the value we obtained was ((3.77-3)/1 =) 0.77 range higher than the top of the range. Here, we found $C = 26.0749$ while the range of C is 23-25, therefore, the value obtained was ((26.0749-25)/2 =) 0.54 range higher than the top of the range.

As a result, we could say that the degree of extrapolation of the adjustment factor in the second analysis is better than in the first one. However, we don't know how accurate each of those 2 adjustment factors need to be, so we cannot really conclude on which analysis is best here.

(f)

If the range of C become 21-27, the value of $C = 26.0749$ falls in the range as opposed to the value of E as mentionned above. Therefore, the second version of the two step procedure would be preffered.

Problem 5:

(a)

(i)

2^{6-2} design with 5=1234 and 6=124, then we have:

$$DCS = \{I, 356, 1246, 12345\}$$

Thus, the resolution of the design is $r_1 = 3$.

(ii)

2^{6-2} design with 5=123 and 6=124, then we have:

$$DCS = \{I, 1235, 1246, 3456\}$$

Thus, the resolution of the design is $r_2 = 4$.

I prefer the second design because of the Maximum Resolution Criterion (For a fixed k and p, choose the 2^{k-p} design with maximum resolution).

(b)

Here are all the 2-factor interactions from design 2:

$$\begin{aligned}
12 &= 35 = 46 = 123456 \\
13 &= 25 = 2346 = 1456 \\
14 &= 26 = 2345 = 1356 \\
15 &= 23 = 2456 = 1346 \\
16 &= 24 = 2356 = 1345 \\
34 &= 56 = 1245 = 1236 \\
36 &= 45 = 1256 = 1234
\end{aligned}$$

The 2-factor interactions that are not estimable are the ones that are aliased with any main effect or any 2-factor interaction not involving 6. Thus, the remaining 2-fi's that are estimable are:

$$14, 24, 34, 45$$

(c)

We now want a 2^{6-2} design in a 2^1 blocks of size 8. My blocking scheme is $B_1 = 134$. I picked this scheme because 134 is the only factorial interaction (along with 136) to be aliased with no interaction lower than a 3-factor interactions (134=245=236=156). Therefore, we most likely won't lose a significant interaction by confounding the block factor and those interactions. That is why this scheme (or 136) is the best.

Problem 6:

I would recommend a $2^{8-2} = 64$ fractional factorial design in $2^2 = 4$ blocks of size 16. From the table of optimum designs from the book, I would choose $7 = 1234$ and $8 = 1256$ as treatment generator and $B_1 = 135$, $B_2 = 246$ as block generators. All main effects and all 28 2-fi's are clear.

Appendix 1:

```r
data <- read.table("http://www2.isye.gatech.edu/~jeffwu/book/data/LeafSpring.dat", header = FALSE)

as.level2=function(x){
    y = rep(0, length(x))
    y[ x=="+" ] = 1
    y[ x=="-" ] = -1
    y
}

B = as.level2(data[,1])
C = as.level2(data[,2])
D = as.level2(data[,3])
E = as.level2(data[,4])
Q = as.level2(data[,5])
```

```r
ybar = apply(data[,6:8], 1, mean)
lns2 = log(apply(data[,6:8], 1, var))

D0=paste("D", data[,3], sep="")
E0=paste("E", data[,4], sep="")
D0E0=paste(D0,E0, sep="")
interaction.plot(Q, D0E0, lns2) # DEQ interaction plot
```