Stat453_Assignment04

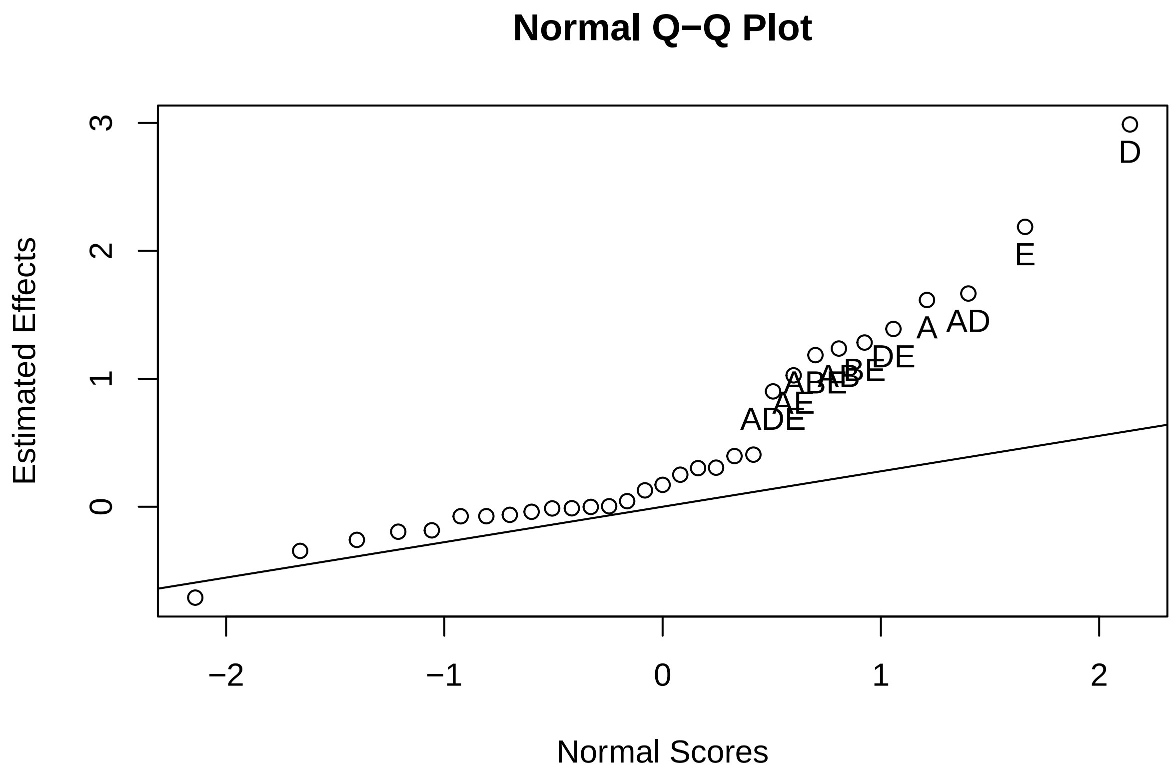
Belina Jang V00924202

2023-03-25

Questions

Question 1. Consider the full 2^5 factorial design Question 5 of Assignment 3. Suppose that this experiment had been run in two blocks with ABCDE confounded with the blocks. Set up the blocked design and perform the analysis. Compare your results with the results obtained for the completely randomized design in Question 5 Assignment 3.

```
A = rep(c(-1,1), 16)
B = rep(c(-1,-1,1,1),8)
C = rep(c(rep(-1,4),rep(1,4)),4)
D = rep(c(rep(-1,8),rep(1,8)),2)
E = c(rep(-1,16),rep(1,16))
ABCDE = A*B*C*D*E

y = c(8.11,5.56,5.77,5.82,9.17,7.8,3.23,5.69,8.82,14.23,9.2,8.94,8.68,11.49,6.25,9.12,
      7.93,5,7.47,12,9.86,3.65,6.4,11.61,12.43,17.55,8.87,25.38,13.06,18.85,11.78,26.05)
q5_data = data.frame(A,B,C,D,E,ABCDE,y)

# if ABCDE = -1, Block1 = 1, else Block1 = 2
q5_data$Block1 = ifelse(q5_data$ABCDE == -1, 1, 2)
q5_data = subset(q5_data, select = -ABCDE) # remove ABCDE column

# Option1: Confouding ABCDE with blocks: Not dropping C
res.lm1=lm(y~A*B*C*D*E-A:B:C:D:E+factor(Block1),data=q5_data)
fullnormal(na.omit(coef(res.lm1)[-1]),alpha=.025)
```




```r
# Option2: We can drop C since it's insignificant
res.aov2=aov(y~A*B*D*E-A:B:D:E+factor(Block1),data=q5_data)
summary(res.aov2)
```

```
##                 Df Sum Sq Mean Sq F value   Pr(>F)    
## A                1  83.56   83.56  54.291 1.59e-06 ***
## B                1   0.06    0.06   0.039 0.845490    
## D                1 285.78  285.78 185.681 3.20e-10 ***
## E                1 153.17  153.17  99.518 2.84e-08 ***
## factor(Block1)   1   4.04    4.04   2.625 0.124743    
## A:B              1  48.93   48.93  31.792 3.70e-05 ***
## A:D              1  88.88   88.88  57.746 1.07e-06 ***
## B:D              1   0.01    0.01   0.004 0.951902    
## A:E              1  33.76   33.76  21.937 0.000249 ***
## B:E              1  52.71   52.71  34.248 2.45e-05 ***
## D:E              1  61.80   61.80  40.153 9.88e-06 ***
## A:B:D            1   3.82    3.82   2.479 0.134928    
## A:B:E            1  44.96   44.96  29.211 5.84e-05 ***
## A:D:E            1  26.01   26.01  16.899 0.000818 ***
## B:D:E            1   0.05    0.05   0.033 0.858668    
## Residuals       16  24.63    1.54                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
res.lm2=lm(y~A*B*D*E-A:B:D:E+factor(Block1), data=q5_data)
summary(res.lm2)
```

```
##
## Call:
## lm(formula = y ~ A * B * D * E - A:B:D:E + factor(Block1), data = q5_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.03250 -0.67984  0.05281  0.69375  1.52688
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     10.53562    0.31015  33.969 2.42e-16 ***
## A                1.61594    0.21931   7.368 1.59e-06 ***
## B                0.04344    0.21931   0.198 0.845490
## D                2.98844    0.21931  13.626 3.20e-10 ***
## E                2.18781    0.21931   9.976 2.84e-08 ***
## factor(Block1)2 -0.71062    0.43862  -1.620 0.124743
## A:B              1.23656    0.21931   5.638 3.70e-05 ***
## A:D              1.66656    0.21931   7.599 1.07e-06 ***
## B:D             -0.01344    0.21931  -0.061 0.951902
## A:E              1.02719    0.21931   4.684 0.000249 ***
## B:E              1.28344    0.21931   5.852 2.45e-05 ***
## D:E              1.38969    0.21931   6.337 9.88e-06 ***
## A:B:D           -0.34531    0.21931  -1.575 0.134928
## A:B:E            1.18531    0.21931   5.405 5.84e-05 ***
## A:D:E            0.90156    0.21931   4.111 0.000818 ***
## B:D:E           -0.03969    0.21931  -0.181 0.858668
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1.241 on 16 degrees of freedom
## Multiple R-squared:  0.973,  Adjusted R-squared:  0.9477
## F-statistic: 38.44 on 15 and 16 DF,  p-value: 1.07e-09
```

$\therefore$ Comparing with the original full model (not projected), same factors and interactions (A, D, E, AB, AD, AE, BE, DE, ABE and ADE) are significant and SS of them stayed the same except SS of A:B:C:D:E (which was 4.04) disappeared since we confound it with blocks and the SS of block appeared now as 4.04 instead. All the estimates didn't change except (Intercept) changed, ABCDE disappeared and block appeared due to confounding with blocks. Now the sum of estimates for (Intercept) and ABCDE from the original model (10.180312 -0.355312 = 9.825) is the same as the sum of new estimates for (Intercept) and block (10.53562 -0.71062 = 9.825).

Question 2. Consider the data in Example 7.2 showed on page 17 of the slides of Chapter 7. Suppose that all the observations in block 2 are increased by 20. Analyze the data that would result. Estimate the block effect. Can you explain its magnitude? Do blocks now appear to be an important factor? Are any other effect estimates impacted by the change you made in the data?

Block 1	Block 2
(1) = 25	a = 71
ab = 45	b = 48
ac = 40	c = 68
bc = 60	d = 43
ad = 80	abc = 65
bd = 25	bcd = 70
cd = 55	acd = 86
$abcd$ = 76	abd = 104

(*b*) Assignment of the 16 runs to two blocks

```
# Before
A=rep(c(-1,1),8)
B=rep(c(rep(-1,2),rep(1,2)),4)
C=rep(c(rep(-1,4),rep(1,4)),2)
D=c(rep(-1,8),rep(1,8))

FiltrationRate=c(25,71,48,45,68,40,60,65,43,80,25,104,55,86,70,76)
Block1=c(1,2,2,1,2,1,1,2,2,1,1,2,1,2,2,1)
FiltrationRate_Data = data.frame(FiltrationRate,A,B,C,D,Block1)

res.aov1=aov(FiltrationRate~A*B*C*D-A:B:C:D+Block1,data=FiltrationRate_Data)
summary(res.aov1)
```

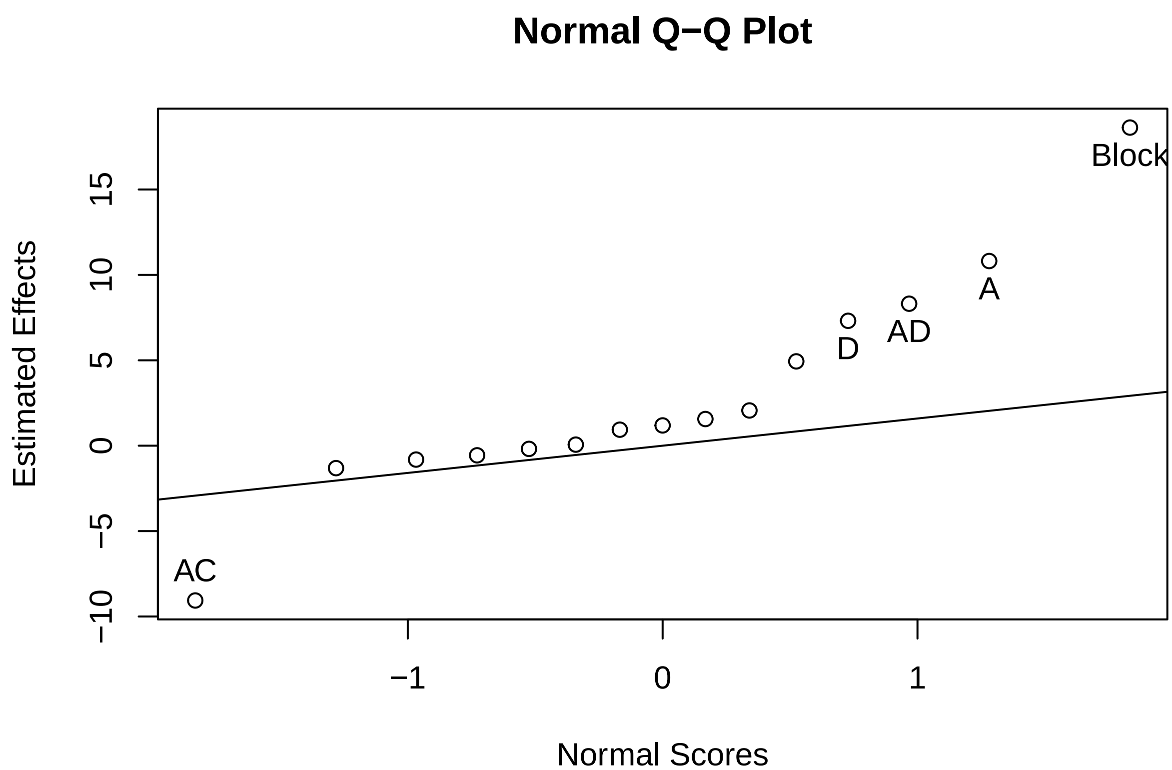
```
##              Df Sum Sq Mean Sq
## A             1 1870.6  1870.6
## B             1   39.1    39.1
## C             1  390.1   390.1
## D             1  855.6   855.6
## Block1        1 1387.6  1387.6
## A:B           1    0.1     0.1
## A:C           1 1314.1  1314.1
## B:C           1   22.6    22.6
## A:D           1 1105.6  1105.6
## B:D           1    0.6     0.6
## C:D           1    5.1     5.1
## A:B:C         1   14.1    14.1
## A:B:D         1   68.1    68.1
## A:C:D         1   10.6    10.6
## B:C:D         1   27.6    27.6
```

```
res.aov1$coefficients # R automatically uses Block2-Block1 (Block2 as referance) instead.
```

```
## (Intercept)           A           B           C           D      Block1
##     32.1250     10.8125      1.5625      4.9375      7.3125     18.6250
```

```
##        A:B        A:C        B:C        A:D        B:D        C:D
##     0.0625    -9.0625     1.1875     8.3125    -0.1875    -0.5625
##      A:B:C      A:B:D      A:C:D      B:C:D
##     0.9375     2.0625    -0.8125    -1.3125
```

```
res.lm1=lm(FiltrationRate~A*B*C*D-A:B:C:D+Block1, data=FiltrationRate_Data)
fullnormal(na.omit(coef(res.lm1)[-1]),alpha=.025)
```




```
# B not significant, drop B and insignificant interactions
res.aov2=aov(FiltrationRate~A*C*D-C:D-A:C:D+Block1,data=FiltrationRate_Data)
summary(res.aov2)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## A            1 1870.6  1870.6   89.76 5.60e-06 ***
## C            1  390.1   390.1   18.72 0.001915 **
## D            1  855.6   855.6   41.05 0.000124 ***
## Block1       1 1387.6  1387.6   66.58 1.89e-05 ***
## A:C          1 1314.1  1314.1   63.05 2.35e-05 ***
## A:D          1 1105.6  1105.6   53.05 4.65e-05 ***
## Residuals    9  187.6    20.8
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
res.aov2$coefficients # R automatically uses Block2-Block1 (Block2 as referance) instead.
```

```
## (Intercept)           A           C           D      Block1         A:C 
##     32.1250     10.8125      4.9375      7.3125     18.6250     -9.0625 
##         A:D 
##      8.3125
```

```r
# Calculating block effect
# select block 1
FiltrationRate_Data1 = subset(FiltrationRate_Data, Block1 == 1)
sum1 = sum(FiltrationRate_Data1$FiltrationRate)
# select block 2
FiltrationRate_Data2 = subset(FiltrationRate_Data, Block1 == 2)
sum2 = sum(FiltrationRate_Data2$FiltrationRate)
# Block effect
BE = sum(FiltrationRate_Data1$FiltrationRate)/8-sum(FiltrationRate_Data2$FiltrationRate)/8
BE # -18.625 based on Block1-Block2 (Block1 as referance)
```

```
## [1] -18.625
```

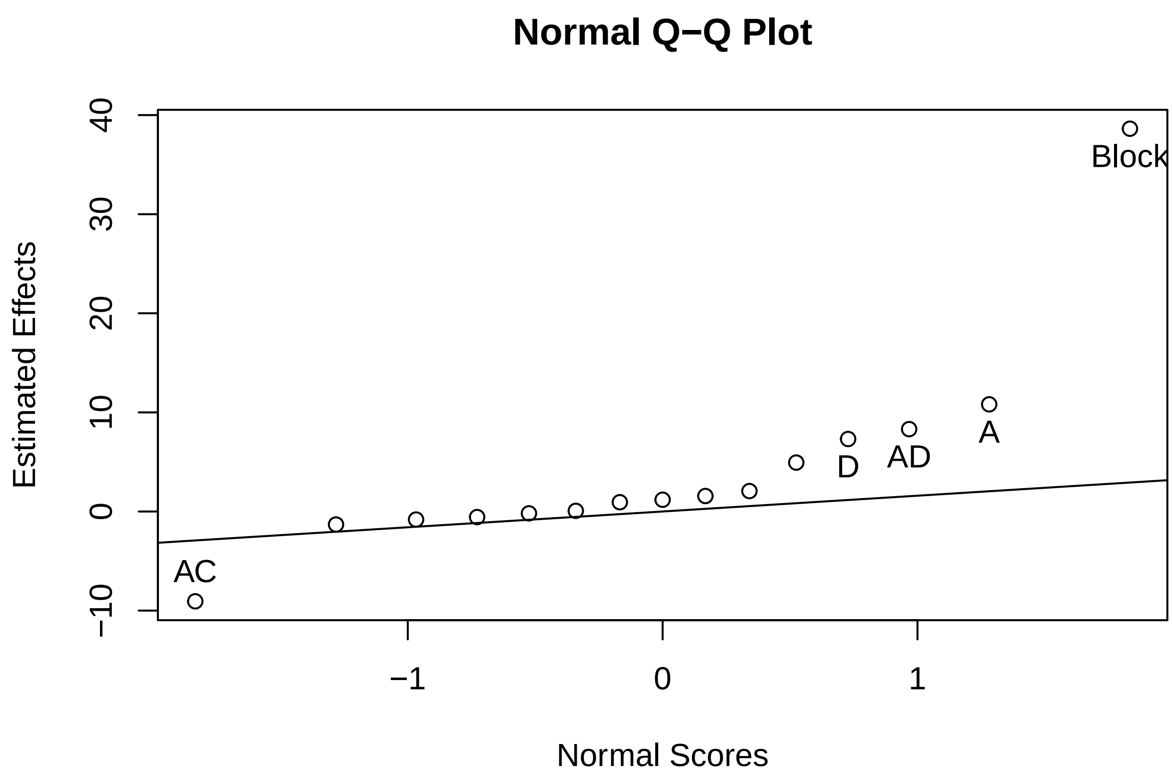
```r
# All the observations in block 2 are increased by 20
FiltrationRate[Block1==2]=FiltrationRate[Block1==2]+20
FiltrationRate_Data = data.frame(FiltrationRate,A,B,C,D,Block1)

# Method1: Confound ABCD with blocks
res.aov3=aov(FiltrationRate~A*B*C*D-A:B:C:D+Block1,data=FiltrationRate_Data)
summary(res.aov3)
```

```
##             Df Sum Sq Mean Sq
## A            1   1871    1871
## B            1     39      39
## C            1    390     390
## D            1    856     856
## Block1       1   5968    5968
## A:B          1      0       0
## A:C          1   1314    1314
## B:C          1     23      23
## A:D          1   1106    1106
## B:D          1      1       1
## C:D          1      5       5
## A:B:C        1     14      14
## A:B:D        1     68      68
## A:C:D        1     11      11
## B:C:D        1     28      28
```

```r
res.aov3$coefficients # R automatically calculates effects using Block2 as the reference
```

```
## (Intercept)           A           B           C           D      Block1 
##     12.1250     10.8125      1.5625      4.9375      7.3125     38.6250 
##         A:B         A:C         B:C         A:D         B:D         C:D 
##      0.0625     -9.0625      1.1875      8.3125     -0.1875     -0.5625 
##       A:B:C       A:B:D       A:C:D       B:C:D 
##      0.9375      2.0625     -0.8125     -1.3125
```

```r
res.lm3=lm(FiltrationRate~A*B*C*D-A:B:C:D+Block1,data=FiltrationRate_Data)
fullnormal(na.omit(coef(res.lm3)[-1]),alpha=.025)
```




```r
# Method2: Confound ABCD with blocks (dropping B since it's not significant)
res.aov4=aov(FiltrationRate~A*C*D-C:D-A:C:D+Block1,data=FiltrationRate_Data)
summary(res.aov4)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)    
## A            1   1871    1871   89.76 5.60e-06 ***
## C            1    390     390   18.72 0.001915 ** 
## D            1    856     856   41.05 0.000124 ***
## Block1       1   5968    5968  286.35 3.94e-08 ***
## A:C          1   1314    1314   63.05 2.35e-05 ***
## A:D          1   1106    1106   53.05 4.65e-05 ***
## Residuals    9    188      21                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
res.aov4$coefficients # R automatically calculates effects using Block2 as the reference
```

```
## (Intercept)           A           C           D      Block1         A:C 
##     12.1250     10.8125      4.9375      7.3125     38.6250     -9.0625 
##         A:D 
##      8.3125
```

```r
# Calculating block effect
# select block 1
FiltrationRate_Data1 = subset(FiltrationRate_Data, Block1 == 1)
sum1 = sum(FiltrationRate_Data1$FiltrationRate)
# select block 2
FiltrationRate_Data2 = subset(FiltrationRate_Data, Block1 == 2)
sum2 = sum(FiltrationRate_Data2$FiltrationRate)
# SSblock
SSblock = (sum1^2+sum2^2)/8-(sum(FiltrationRate))^2/16
SSblock
```

```
## [1] 5967.562
```

```r
# Block effect
BE = sum(FiltrationRate_Data1$FiltrationRate)/8-sum(FiltrationRate_Data2$FiltrationRate)/8
BE # -38.625 based on Block1-Block2 (Block1 as referance)
```

```
## [1] -38.625
```

∴ The previous block effect for (Block1-Block2: Block 1 as reference level) was -18.625. And since we increased all the observations in Block2 by 20 and we want Block1 to be the reference level, we need to subtract 20 from the previous block effect. Then the new block effect is -18.625 - 20 = -38.625. Therefore, the magnitude of the block effect is larger than before. Blocks is still an important factor. The other effect estimates (other than the intercept and Block1) are not impacted by the change we made in the data.

Question 3. (Example 6.6. of your textbook). An article in the International Journal of Research in Marketing ("Experimental Design on the Front Lines of Marketing: Testing New Ideas to Increase Direct Mail Sales," 2006, Vol. 23, pp. 309–319) describes an experiment to test new ideas to increase credit card division of a financial services company. They want to improve the response rate to its credit card offers. They know from experience that the interest rates are an important factor in attracting potential customers, so they have decided to focus on factors involving both interest rates and fees. They want to test changes in both introductory and long-term rates, as well as the effects of adding an account-opening fee and lowering the annual fee. The factors tested in the experiment are as follows:

Table 1: The 2^4 Factorial Design Used in the Credit Card Marketing Experiment

Run	A	B	C	D	y	label
1	-	-	-	-	2.45	(1)
2	+	-	-	-	3.36	a
3	-	+	-	-	2.16	b
4	+	+	-	-	2.29	ab
5	-	-	+	-	2.49	c
6	+	-	+	-	3.39	ac
7	-	+	+	-	2.32	bc
8	+	+	+	-	2.44	abc
9	-	-	-	+	1.84	d
10	+	-	-	+	2.24	ad
11	-	+	-	+	1.69	bd
12	+	+	-	+	1.87	abd
13	-	-	+	+	2.29	cd
14	+	-	+	+	2.92	acd
15	-	+	+	+	2.04	bcd
16	+	+	+	+	2.03	abcd

Factor	(-1) Control	(+1) New Idea
A: Annual Fee	Current	Lower
B: Account-opening fee	No	Yes
C: Initial interest rate	Current	Lower
D: Long-term interest rate	Low	High

```
A=rep(c(-1,1),8)
B=rep(c(rep(-1,2),rep(1,2)),4)
C=rep(c(rep(-1,4),rep(1,4)),2)
D=c(rep(-1,8),rep(1,8))
y=c(2.45,3.36,2.16,2.29,2.49,3.39,2.32,2.44,1.84,2.24,1.69,1.87,2.29,2.92,2.04,2.03)
q3_data = data.frame(A,B,C,D,y)
```

(a) Analyze the data and determine which factor is not significant.

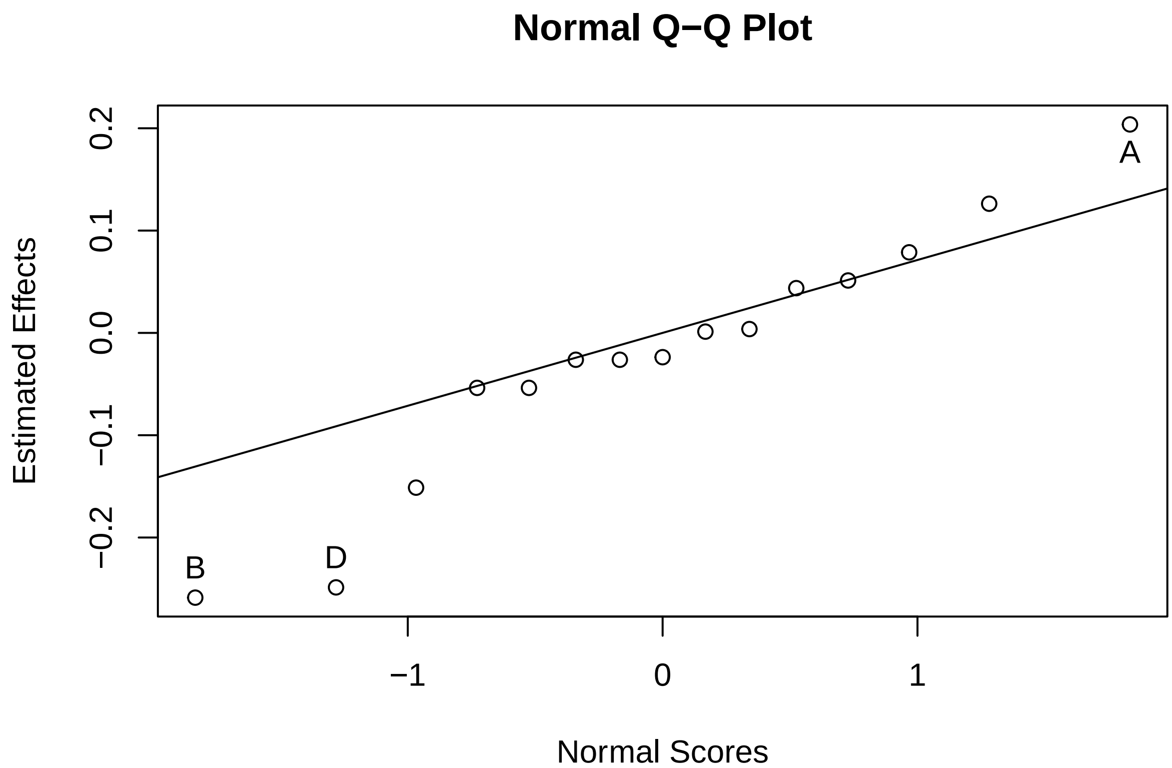
```
res.aov = aov(y~A*B*C*D,data=q3_data)
summary(res.aov) #original model automatically based on Block2 - Block1 (Block2 as reference)
```

```
##             Df Sum Sq Mean Sq
## A            1 0.6642  0.6642
## B            1 1.0712  1.0712
## C            1 0.2550  0.2550
## D            1 0.9900  0.9900
## A:B          1 0.3660  0.3660
## A:C          1 0.0000  0.0000
## B:C          1 0.0090  0.0090
## A:D          1 0.0462  0.0462
## B:D          1 0.0420  0.0420
## C:D          1 0.0992  0.0992
## A:B:C        1 0.0110  0.0110
## A:B:D        1 0.0306  0.0306
## A:C:D        1 0.0002  0.0002
## B:C:D        1 0.0462  0.0462
## A:B:C:D      1 0.0110  0.0110
```

```
res.aov$coefficients
```

```
## (Intercept)           A           B           C           D         A:B
##     2.36375     0.20375    -0.25875     0.12625    -0.24875    -0.15125
##         A:C         B:C         A:D         B:D         C:D       A:B:C
##     0.00125    -0.02375    -0.05375     0.05125     0.07875    -0.02625
##       A:B:D       A:C:D       B:C:D     A:B:C:D
##     0.04375     0.00375    -0.05375    -0.02625
```

```
res.lm=lm(y~A*B*C*D, data=q3_data)
fullnormal(na.omit(coef(res.lm)[-1]),alpha=.05) #assume a=0.05 then A,B,D are significant
```



∴ Factor C is not significant.

(b) Project the 2^4 design into two replicates of a 2^3 on the significant factors. The new design table should include the runs, factors, responses, and labels

```
# Drop C; Projected model
res.aov = aov(y~A*B*D,data=q3_data)
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value  Pr(>F)
## A            1 0.6642  0.6642  12.306 0.00798 **
## B            1 1.0712  1.0712  19.847 0.00213 **
## D            1 0.9900  0.9900  18.342 0.00268 **
## A:B          1 0.3660  0.3660   6.781 0.03142 *
## A:D          1 0.0462  0.0462   0.856 0.38181
## B:D          1 0.0420  0.0420   0.779 0.40330
## A:B:D        1 0.0306  0.0306   0.567 0.47288
## Residuals    8 0.4318  0.0540
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# Each replicate separate entry
runs = c(1:16)
labels = c("(1)","a","b","ab","(1)","a","b","ab","d","ad","bd","abd","d","ad","bd","abd")
```

```r
d2_3_1 = data.frame(runs,labels,A,B,D,y)
d2_3_1
```

```
##    runs labels  A  B  D    y
## 1     1    (1) -1 -1 -1 2.45
## 2     2      a  1 -1 -1 3.36
## 3     3      b -1  1 -1 2.16
## 4     4     ab  1  1 -1 2.29
## 5     5    (1) -1 -1 -1 2.49
## 6     6      a  1 -1 -1 3.39
## 7     7      b -1  1 -1 2.32
## 8     8     ab  1  1 -1 2.44
## 9     9      d -1 -1  1 1.84
## 10   10     ad  1 -1  1 2.24
## 11   11     bd -1  1  1 1.69
## 12   12    abd  1  1  1 1.87
## 13   13      d -1 -1  1 2.29
## 14   14     ad  1 -1  1 2.92
## 15   15     bd -1  1  1 2.04
## 16   16    abd  1  1  1 2.03
```

```r
# 2 replicates in a single entry
A_2=rep(c(-1,1),4)
B_2=rep(c(rep(-1,2),rep(1,2)),2)
D_2=rep(c(rep(-1,4),rep(1,4)),1)
runs_2 = c(1:8)
labels_2 = c("(1)","a","b","ab","d","ad","bd","abd")
rep1 = c(2.45,3.36,2.16,2.29,1.84,2.24,1.69,1.87)
rep2 = c(2.49,3.39,2.32,2.44,2.29,2.92,2.04,2.03)
d2_3_2 = data.frame(runs_2,labels_2,A_2,B_2,D_2,rep1,rep2)
d2_3_2
```

```
##   runs_2 labels_2 A_2 B_2 D_2 rep1 rep2
## 1      1      (1)  -1  -1  -1 2.45 2.49
## 2      2        a   1  -1  -1 3.36 3.39
## 3      3        b  -1   1  -1 2.16 2.32
## 4      4       ab   1   1  -1 2.29 2.44
## 5      5        d  -1  -1   1 1.84 2.29
## 6      6       ad   1  -1   1 2.24 2.92
## 7      7       bd  -1   1   1 1.69 2.04
## 8      8      abd   1   1   1 1.87 2.03
```

(c) In the projected design, what is the estimated effect of the account-opening fee in the response rate?

```r
res.aov = aov(y~A*B*D,data=d2_3_1) # projected design
estimated_effect = res.aov$coefficients*2 # coefficient*2
estimated_effect
```

```
## (Intercept)           A           B           D         A:B         A:D
```

```
##      4.7275      0.4075     -0.5175     -0.4975     -0.3025     -0.1075
##         B:D       A:B:D
##      0.1025      0.0875
```

∴ The estimated effect of the account-opening fee (B) is -0.5175.

(d) Using the projected design, is the account-opening fee significant?

```
summary(res.aov)
```

```
##             Df Sum Sq Mean Sq F value  Pr(>F)
## A            1 0.6642  0.6642  12.306 0.00798 **
## B            1 1.0712  1.0712  19.847 0.00213 **
## D            1 0.9900  0.9900  18.342 0.00268 **
## A:B          1 0.3660  0.3660   6.781 0.03142 *
## A:D          1 0.0462  0.0462   0.856 0.38181
## B:D          1 0.0420  0.0420   0.779 0.40330
## A:B:D        1 0.0306  0.0306   0.567 0.47288
## Residuals    8 0.4318  0.0540
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

∴ The p-value is smaller than 0.05, so the effect of account-opening fee (B) is significant.

(e) Confound the projected design with blocks using the highest order interaction as a confounding. Write down the runs for both blocks and estimate the block effect. What is the block effect really estimating in this case?

```
# Confound with the highest order interaction ABD
# Each replicate in a separate entry
d2_3_C = d2_3_1
d2_3_C$Block1 = A*B*D

# if ABD = -1, Block1 = 1, else Block1 = 2
d2_3_C$Block1 = ifelse(d2_3_C$Block1 == -1, 1, 2)
d2_3_C
```

```
##    runs labels  A  B  D    y Block1
## 1     1    (1) -1 -1 -1 2.45      1
## 2     2      a  1 -1 -1 3.36      2
## 3     3      b -1  1 -1 2.16      2
## 4     4     ab  1  1 -1 2.29      1
## 5     5    (1) -1 -1 -1 2.49      1
## 6     6      a  1 -1 -1 3.39      2
## 7     7      b -1  1 -1 2.32      2
## 8     8     ab  1  1 -1 2.44      1
## 9     9      d -1 -1  1 1.84      2
## 10   10     ad  1 -1  1 2.24      1
## 11   11     bd -1  1  1 1.69      1
```

```
## 12   12    abd  1  1  1 1.87      2
## 13   13      d -1 -1  1 2.29      2
## 14   14     ad  1 -1  1 2.92      1
## 15   15     bd -1  1  1 2.04      1
## 16   16    abd  1  1  1 2.03      2
```

```r
# Block effect based on Block1 - Block2 (Block1 as the reference)
block_effect = sum(d2_3_C$y[d2_3_C$Block1 == 1])/8 - sum(d2_3_C$y[d2_3_C$Block1 == 2])/8 # -0.0875

res.lm = lm(y~A*B*D+factor(Block1),data=d2_3_C)
summary(res.lm) # R automatically estimates using Block2-Block1 (Block2 as the reference)
```

```
##
## Call:
## lm(formula = y ~ A * B * D + factor(Block1), data = d2_3_C)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -0.34  -0.08   0.00   0.08   0.34
##
## Coefficients: (1 not defined because of singularities)
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.32000    0.08214  28.245 2.67e-09 ***
## A                0.20375    0.05808   3.508  0.00798 **
## B               -0.25875    0.05808  -4.455  0.00213 **
## D               -0.24875    0.05808  -4.283  0.00268 **
## factor(Block1)2  0.08750    0.11616   0.753  0.47288
## A:B             -0.15125    0.05808  -2.604  0.03142 *
## A:D             -0.05375    0.05808  -0.925  0.38181
## B:D              0.05125    0.05808   0.882  0.40330
## A:B:D                 NA         NA      NA       NA
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2323 on 8 degrees of freedom
## Multiple R-squared:  0.8814, Adjusted R-squared:  0.7777
## F-statistic: 8.497 on 7 and 8 DF,  p-value: 0.003616
```

$\therefore$ The estimate for the block effect (Block1 as reference) is -0.0875. What the block effect really estimating is Block + ABD.

Question 4. In the previous example, a 2^4 factorial design was used to improve the response rate to a credit card marketing offer. Suppose that the researchers had used the 2^(4-1) fraction factorial design with I=ABCD instead. Set up the design and select the responses for the runs from the full factorial data in Example 6.6. Analyze the data and draw conclusions. Compare your findings with those from the full factorial in Example 6.6.

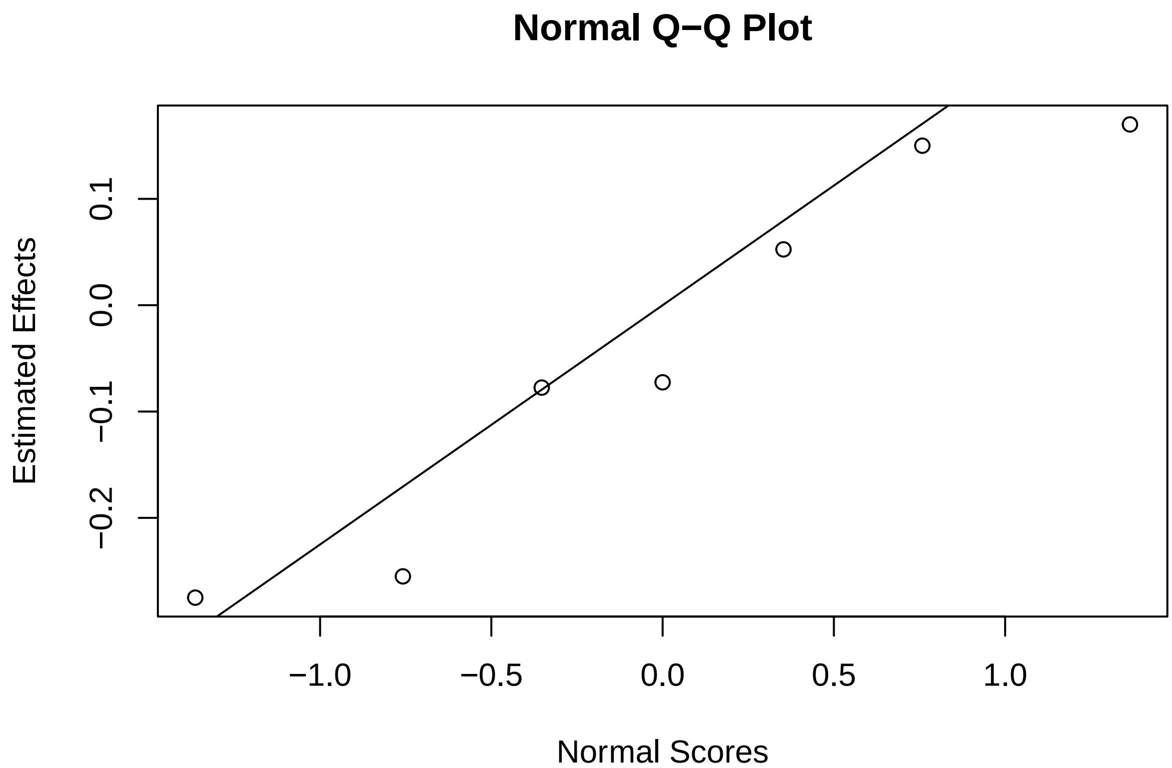
Run	A	B	C	D=ABC	Treatment	Rate
1	-	-	-	-	(1)	2.45
2	+	-	-	+	ad	2.24
3	-	+	-	+	bd	1.69
4	+	+	-	-	ab	2.29
5	-	-	+	+	cd	2.29
6	+	-	+	-	ac	3.39
7	-	+	+	-	bc	2.32
8	+	+	+	+	abcd	2.03

```r
A = rep(c(-1,1),4)
B = rep(c(rep(-1,2),rep(1,2)),2)
C = rep(c(rep(-1,4),rep(1,4)),1)
D = A*B*C
treatment = c("(1)","ad","bd","ab","cd","ac","bc","abcd")
rate = c(2.45, 2.24, 1.69, 2.29, 2.29, 3.39, 2.32, 2.03)
q4_data = data.frame(A,B,C,D,treatment,rate)
```

```r
res.lm=lm(rate~A*B*C*D, data=q4_data)
summary(res.lm)
```

```
##
## Call:
## lm(formula = rate ~ A * B * C * D, data = q4_data)
##
## Residuals:
## ALL 8 residuals are 0: no residual degrees of freedom!
##
## Coefficients: (8 not defined because of singularities)
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.3375        NaN     NaN      NaN
## A             0.1500        NaN     NaN      NaN
## B            -0.2550        NaN     NaN      NaN
## C             0.1700        NaN     NaN      NaN
## D            -0.2750        NaN     NaN      NaN
## A:B          -0.0725        NaN     NaN      NaN
## A:C           0.0525        NaN     NaN      NaN
## B:C          -0.0775        NaN     NaN      NaN
## A:D               NA         NA      NA       NA
## B:D               NA         NA      NA       NA
## C:D               NA         NA      NA       NA
## A:B:C             NA         NA      NA       NA
## A:B:D             NA         NA      NA       NA
## A:C:D             NA         NA      NA       NA
## B:C:D             NA         NA      NA       NA
## A:B:C:D           NA         NA      NA       NA
##
## Residual standard error: NaN on 0 degrees of freedom
## Multiple R-squared:      1,  Adjusted R-squared:    NaN
## F-statistic:   NaN on 7 and 0 DF,  p-value: NA
```

```r
fullnormal(na.omit(coef(res.lm)[-1]), alpha = 0.05)
```



∴ Assuming a = 0.05, no effect is significant (A,B,D are no longer significant in this fractional design). The estimates for each factor/interaction are changed based on the aliased structure of the factor/interaction using the effect estimates from the previous full design.
A is aliased with BCD, so [A] = A(full) + BCD(full) = 0.20375 - 0.05375 = 0.15 (fraction)
B is aliased with ACD, so [B] = B(full) + ACD(full) = -0.25875 + 0.00375 = -0.255 (fraction)
C is aliased with ABD, so [C] = C(full) + ABD(full) = 0.12625 + 0.04375 = 0.17 (fraction)
D is aliased with ABC, so [D] = D(full) + ABC(full) = -0.24875 - 0.02625 = -0.275 (fraction)
AB is alised with CD, so [AB] = AB(full) + CD(full) = -0.15125 + 0.07875 = -0.0725 (fraction)
AC is alised with BD, so [AC] = AC(full) + BD(full) = 0.00125 + 0.05125 = 0.0525 (fraction)
BC is alised with AD, so [BC] = BC(full) + AD(full) = -0.02375 - 0.05375 = -0.0775 (fraction)

Question 5.

(a) Construct 2^(5-2) design using the generator D=+AB and E=-AC.

```
runs = c(1:8)
A = rep(c(-1,1),4)
B = rep(c(rep(-1,2),rep(1,2)),2)
C = rep(c(rep(-1,4),rep(1,4)),1)
D = A*B
E = -A*C
labels = c("d","ae","b","abde","cde","ac","bce","abcd")
q4_data = data.frame(runs,labels,A,B,C,D,E)
q4_data
```

```
##   runs labels  A  B  C  D  E
## 1    1      d -1 -1 -1  1 -1
## 2    2     ae  1 -1 -1 -1  1
## 3    3      b -1  1 -1 -1 -1
## 4    4   abde  1  1 -1  1  1
## 5    5    cde -1 -1  1  1  1
## 6    6     ac  1 -1  1 -1 -1
## 7    7    bce -1  1  1 -1  1
## 8    8   abcd  1  1  1  1 -1
```

(b) What is the complete defining relation for this design?

∴ I = ABD = -ACE = -BCDE

(c) What is the alias structure for the effects of A, B, C, D, and E? What does this alias structure tell us about the resolution of this design?

A = BD = -CE = -ABCDE
B = AD = -ABCE = -CDE
C = ABCD = -AE = -BDE
D = AB = -ACDE = -BCE
E = ABDE = -AC = -BCD

∴ The resolution for this design is 3.