實驗設計與分析 HW6 solution

Problem 1.

Do Problem 4 in Chapter 5 of textbook.

Prove that, for a resolution R design, its projection onto any $R-1$ factors is a full factorial. (Hint: If the projection is not a full factorial, it must be a fractional factorial with some defining relation among the $R-1$ factors.)

假設一個 reolution R 實驗其投影至 $R-1$ 個因子上並非 full factorial，則其為 fractional factorial，表示這 $R-1$ 個因子間存在某個 defining relation L。由於此 $R-1$ 個因子所形成的設計矩陣必定為整個設計的子矩陣，因此 defining relation L 也必定為整個實驗的 defining relation，但因為 L 是由 $R-1$ 個因子所構成，其 resolution 必小於 R，和"此為 resolution R 的實驗相互矛盾"，故得證。

made by 陳冠霖, 宇晉賢 助教

Problem 2.

Do Problem 13(a)(b) in Chapter 5 of textbook.

(a) Design an eight-run experiment to study the effect of the following five factors on yield: temperature (160 or $180^{\circ}F$), concentration (30 or 40%), catalyst (A or B), stirring rate (60 or 100 rpm), pH (low or high) with the prior knowledge that the combinations (180°F, 40%, 100 rpm) of (temperature, concentration, stirring rate) and (180°F, B, high) of (temperature, catalyst, pH) may lead to disastrous results and should be avoided. Write down the factor levels for each of the eight runs in a planning matrix.

根據題目此實驗為 2^{5-2} 設計，五個因子與其對應水準如下

Table 1: Experiment Table

		+	-
A	temperature	180	160
B	concentration	40	30
C	catalyst	A	B
D	stirring rate	100	60
E	pH	high	low

根據題目可知此實驗需避免 $(A, B, D) = (+, +, +)$ 及 $(A, C, E) = (+, -, +)$，因此取 defining relation

$$\boldsymbol{I} = -ABD = ACE$$

其設計可先透過 A、B、C 建構全因子設計，其後 D、E 由 $D = -AB$ 及 $E = AC$ 生成，其 design matrix 及 planing matrix 如下

Table 2: Design Matrix

run	A	B	C	D	E
1	-1	-1	-1	-1	1
2	1	-1	-1	1	-1
3	-1	1	-1	1	1
4	1	1	-1	-1	-1
5	-1	-1	1	-1	-1
6	1	-1	1	1	1
7	-1	1	1	1	-1
8	1	1	1	-1	1

Table 3: Planning Matrix

run	temperature	concentration	catalyst	stirring rate	pH
1	160	30	B	60	high
2	180	30	B	100	low
3	160	40	B	100	high
4	180	40	B	60	low
5	160	30	A	60	low
6	180	30	A	100	high

run	temperature	concentration	catalyst	stirring rate	pH
7	160	40	A	100	low
8	180	40	A	60	high

(b) For the five factors in (a), find an eight-run design such that the catalyst-by-temperature interaction and the catalyst-by-concentration interaction are neither aliased with the main effects nor with each other.

根據題目所述，為使 AC、BC 不會與任何 main effect alias 或相互 alias，可取以下 defining relation

$$\boldsymbol{I} = ABD = CDE = ABCE$$

其中 AC、BC 的 alias set 如下

$$AC = BCD = ADE = BE$$
$$BC = ACD = BDE = AE$$

故符合題目需求。其設計可先透過 A、B、C 建構全因子設計，其後 D、E 由 $D = AB$ 及 $E = ABC$ 生成，其 design matrix 及 planing matrix 如下

Table 4: Design Matrix

run	A	B	C	D	E
1	-1	-1	-1	1	-1
2	1	-1	-1	-1	1
3	-1	1	-1	-1	1
4	1	1	-1	1	-1
5	-1	-1	1	1	1
6	1	-1	1	-1	-1
7	-1	1	1	-1	-1
8	1	1	1	1	1

Table 5: Planning Matrix

run	temperature	concentration	catalyst	stirring rate	pH
1	160	30	B	100	low
2	180	30	B	60	high
3	160	40	B	60	high
4	180	40	B	100	low
5	160	30	A	100	high
6	180	30	A	60	low
7	160	40	A	60	low
8	180	40	A	100	high

Problem 3.

Do Problem 15 in Chapter 5 of textbook.

(a) What is the resolution of each of the fractional factorial designs indicated below? Which design do you prefer? Justify your answers.

$$\text{(i)} \quad 2^{6-2} \quad \text{with} \quad \mathbf{5 = 1234}, \quad \mathbf{6 = 124},$$
$$\text{(ii)} \quad 2^{6-2} \quad \text{with} \quad \mathbf{5 = 123}, \quad \mathbf{6 = 124},$$

根據題目我們能得到

$$\text{Defining contrast subgroup of (i)} : \boldsymbol{I} = 12345 = 1246 = 356,$$

因此 (i) 為 resolution III，

$$\text{Defining contrast subgroup of (ii)} : \boldsymbol{I} = 1235 = 1246 = 3456,$$

因此 (ii) 為 resolution IV，因此 design (ii) 較 (i) 好。

(b) For the design in (ii), if we further know that any two-factor interaction involving factor 6 (i.e., $\mathbf{16}, \mathbf{26}, \mathbf{36}, \mathbf{46}, \mathbf{56}$) is negligible, which two-factor interactions are estimable under the usual assumptions that three-factor and higher interactions are negligible?

根據題目所述，首先忽略所有包含 6 的 2fi，我們能得到剩下的 2fi 有

$$\text{2-factor interaction} : 12, 13, 14, 15, 23, 24, 25, 34, 35,$$

然而因為

$$\mathbf{12} = \mathbf{35}, \quad \mathbf{13} = \mathbf{25}, \quad \mathbf{15} = \mathbf{23}$$

這幾項 2fi 相互 alias，因此剩下 14、24、34、45 可被估計。

(c) Under the same assumptions as in (b), find a scheme to arrange the design in (ii) in two blocks each of size 8. Explain why your choice is the best.

根據 (b) 小題的假設及 design (ii) 能發現有兩組 alias set 由 3fi 組成，分別為

$$\mathbf{134} = \mathbf{245} = \mathbf{236} = \mathbf{156},$$
$$\mathbf{136} = \mathbf{256} = \mathbf{234} = \mathbf{145},$$

考慮以下兩種 blocking scheme

$$\boldsymbol{B}_1 = \mathbf{134} (= \mathbf{245} = \mathbf{236} = \mathbf{156}),$$
$$\boldsymbol{B}_2 = \mathbf{136} (= \mathbf{256} = \mathbf{234} = \mathbf{145}),$$

此時所有 main effect 及 14、24、34、45 皆為 clear，且沒有 2fi 與 block confounded。

Problem 4.

Do Problem 16 in Chapter 5 of textbook.

Two choices of generators for a 2^{6-2} design are being considered:

$$A : \mathbf{5 = 1234}, \quad \mathbf{6 = 123,}$$
$$B : \mathbf{5 = 123}, \quad \mathbf{6 = 234.}$$

(a) Which design, A or B, would you recommend? Why?

根據題目我們能得到

$$\text{Defining contrast subgroup of A} : \boldsymbol{I} = 12345 = 1236 = 456,$$

因此 A 為 resolution III，

$$\text{Defining contrast subgroup of B} : \boldsymbol{I} = 1235 = 2346 = 1456,$$

因此 B 為 resolution IV，因此 design B 較 A 好。

(b) Show that it is impossible to have a 2_V^{6-2} design. (Hint: Count the degrees of freedom.)

2_V^{6-2} design 共有 $2^4 = 16$ run，因此自由度為 16-1=15，若為 resolution V，代表所有 main effect 及 2fi 皆為 clear，然而在此情況下所需的自由度為 $6 + \binom{6}{2} = 6 + 15 = 21$，故不可能有此設計。

Problem 5.

Do Problem 28 in Chapter 5 of textbook.

In a resistance spot welding experiment, five factors were chosen to study their effects on the tensile strength, which is the maximum load a weld can sustain in a tensile test. The five factors are: button diameter (A), welding time (B), holding time (C), electrode force (D), and machine type (E), each at two levels. The last factor is qualitative, while the others are quantitative. A 2_V^{5-1} design with $\boldsymbol{I} = -ABCDE$ was used for the experiment. Each run has V three replicates. The data are given in Table 5.13.

Table 5.13 Design Matrix and Tensile Strength Data, Spot Welding Experiment

	Factor							
Run	A	B	C	D	E	Tensile Strength		
1	−	−	−	−	−	1330	1330	1165
2	+	+	−	−	−	1935	1935	1880
3	+	−	+	−	−	1770	1770	1770
4	−	+	+	−	−	1275	1275	1275
5	+	−	−	+	−	1880	1935	1880
6	−	+	−	+	−	1385	1440	1495
7	−	−	+	+	−	1220	1165	1440
8	+	+	+	+	−	2155	2100	2100
9	+	−	−	−	+	1715	1715	1660
10	−	+	−	−	+	1385	1550	1550
11	−	−	+	−	+	1000	1165	1495
12	+	+	+	−	+	1990	1990	1990
13	−	−	−	+	+	1275	1660	1550
14	+	+	−	+	+	1660	1605	1660
15	+	−	+	+	+	1880	1935	1935
16	−	+	+	+	+	1275	1220	1275

Figure 1: Table 5.13

(a) What effects are clear and strongly clear in this experiment?

因為 2_V^{5-1} design with $\boldsymbol{I} = -ABCDE$ 可以得出 main effect 為 $A = -BCDE, B = -ACDE, \cdots$，都只和 4fi alias，而 2fi 為 $AB = -CDE, AC = -BDE, \cdots$ 都只和 3fi alias，因此所有 2fi 為 clear，所有 main effect 為 strongly clear。

(b) Analyze the location and dispersion effects separately, including the fitted models for each.

首先計算各個組合的 $\bar{y}$、s^2、$\ln s^2$，其中將部分 $s^2 = 0$ 的以 0.001 代替

```
dat5 <- as.matrix(dat5)
dat5 <- cbind(unique(dat5[,-6]),matrix(dat5[,6],ncol=3,byrow=T))
dat5[which(dat5=="-",arr.ind=TRUE)]=-1
dat5[which(dat5=="+",arr.ind=TRUE)]=1
dat5 <- apply(dat5,2,as.numeric)
colnames(dat5)[1:5]=c("A","B","C","D","E")
dat5 <- as.data.frame(dat5)
ybar = apply(dat5[,6:8], 1, mean)
s2 = apply(dat5[,6:8],1,var)
s2[which(s2==0)]=0.001
lns2 = log(s2)
```

```
dat5_tab = cbind(dat5[,1:5],ybar,s2,lns2)
kable(dat5_tab, row.names = F, format = 'markdown',
      col.names = c("A","B","C","D","E","$\\bar{y}$","$s^2$","ln$s^2$"),
      caption = "Design Matrix",
      align = c(rep("r",8)),
      digits = 4)
```

Table 6: Design Matrix

A	B	C	D	E	$\bar{y}$	s^2	lns^2
-1	-1	-1	-1	-1	1275.000	9075.000	9.1133
1	1	-1	-1	-1	1916.667	1008.333	6.9161
1	-1	1	-1	-1	1770.000	0.001	-6.9078
-1	1	1	-1	-1	1275.000	0.001	-6.9078
1	-1	-1	1	-1	1898.333	1008.333	6.9161
-1	1	-1	1	-1	1440.000	3025.000	8.0147
-1	-1	1	1	-1	1275.000	21175.000	9.9606
1	1	1	1	-1	2118.333	1008.333	6.9161
1	-1	-1	-1	1	1696.667	1008.333	6.9161
-1	1	-1	-1	1	1495.000	9075.000	9.1133
-1	-1	1	-1	1	1220.000	63525.000	11.0592
1	1	1	-1	1	1990.000	0.001	-6.9078
-1	-1	-1	1	1	1495.000	39325.000	10.5796
1	1	-1	1	1	1641.667	1008.333	6.9161
1	-1	1	1	1	1916.667	1008.333	6.9161
-1	1	1	1	1	1256.667	1008.333	6.9161

接著對 location and disperation effect 估計

```
M = model.matrix(ybar~A*B*C*D*E,data = dat5)[,2:16]
y.effect = round(t(M)%*%ybar/8,4)
s.effect = round(t(M)%*%lns2/8,4)
effect = c("A=-BCDE","B=-ACDE","C=-ABDE","D=-ABCE","E=-ABCD",
           "AB=-CDE","AC=-BDE","BC=-ADE","AD=-BCE","BD=-ACE",
           "CD=-ABE","AE=-BCD","BE=-ACD","CE=-ABD","DE=-ABC")
tab = data.frame(effect,y.effect,s.effect)
kable(tab,row.names = F,format = 'markdown',
      col.names = c("Effect","$\\bar{y}$","$\\ln s^2$"),
      caption = "Effect Estimation Table",
      align = c("l",rep("r",2)),
      digits = 4)
```

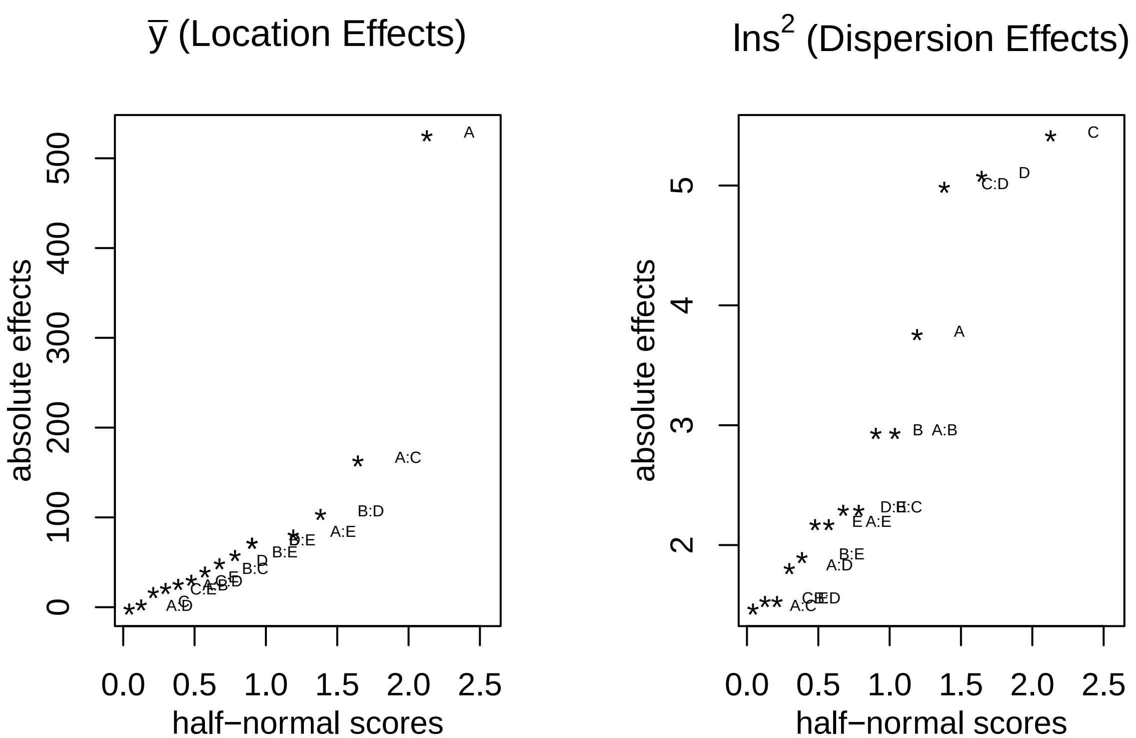
Table 7: Effect Estimation Table

Effect	$\bar{y}$	$\ln s^2$
A=-BCDE	527.0833	-3.7710
B=-ACDE	73.3333	-2.9471
C=-ABDE	-4.5833	-5.4300
D=-ABCE	50.4167	5.0926
E=-ABCD	-32.0833	2.1859
AB=-CDE	22.9167	2.9471
AC=-BDE	165.0000	-1.4819

Effect	$\bar{y}$	$\ln s^2$
BC=-ADE	41.2500	-2.3058
AD=-BCE	0.0000	1.8193
BD=-ACE	-105.4167	1.5447
CD=-ABE	27.5000	5.0006
AE=-BCD	-82.5000	-2.1859
BE=-ACD	-59.5833	-1.9113
CE=-ABD	18.3333	1.5447
DE=-ABC	-73.3333	-2.3058

由上表可知所有 2fi 為 clear，所有 main effect 為 strongly clear。接著做 half-normal plot

```
M1 = lm(ybar~A*B*C*D*E,data = dat5)
M2 = lm(lns2~A*B*C*D*E,data = dat5)
par(mfrow=c(1,2))
DanielPlot(M1,half = T,datax = F,autolab = F,
           main =TeX("$\\bar{y}$ (Location Effects)"),cex.fac = .5)
DanielPlot(M2,half = T,datax = F,autolab = F,
           main =TeX("$\\ln s^2$ (Dispersion Effects)"),cex.fac = .5)
```



根據上圖發現 location effect 中 A 為顯著效應，在 dispersion effect 中 C、D、CD 為顯著效應，因此我們能建立以下模型

$$\hat{\bar{y}} = 1605 + 263.5A,$$
$$ln\hat{s}^2 = 5.3456 - 2.715C + 2.5463D + 2.5003CD.$$

(c) For the dispersion effect model in (b), interpret the meaning of the significant effects. Use this model or an interaction effects plot to select optimal factor settings.

從 dispersion model 可以得知顯著效應有 C、D、CD，為使實驗變異 $ln\hat{s}^2$ 達到最小，optimal factor settings 為 $(C, D) = (+, -)$。

(d) Based on the location effect model in (b), select optimal factor settings in terms of maximizing the tensile strength.

從 location model 可以得知顯著效應有 A，為使實驗的 mean $\hat{\bar{y}}$ 達到最大，optimal factor settings 為 $(A) = (+)$。

(e) Is there any adjustment factor? Based on the results in (c) and (d), can you devise a two-step procedure for design optimization? (Hint: Because the response is a larger-the-better characteristic, you should consider the two-step procedure in Section 6.2.)

在 two-step procedure for design optimization，(c) 對應 dispersion effect，(d) 對應 location effect，在 larger-the-better problems 中，因為 (c) 中的 factor C 及 factor D 沒有出現在 (d) 中，調整 factor C 及 factor D 時不會影響 step 1 的 location model，故 C 及 D 為 adjustment factor。

Two-step procedure for larger-the-better problems.

1. $(A) = (+)$ 實驗的 mean 達到最大，$\hat{\bar{y}} = 1605 + 263.5 = 1868.5$。
2. $(C, D) = (+, -)$ 使實驗變異達到最小，$s^2 = exp(5.3456 - 2.715 - 2.5463 - 2.5003) = 0.08927802$。