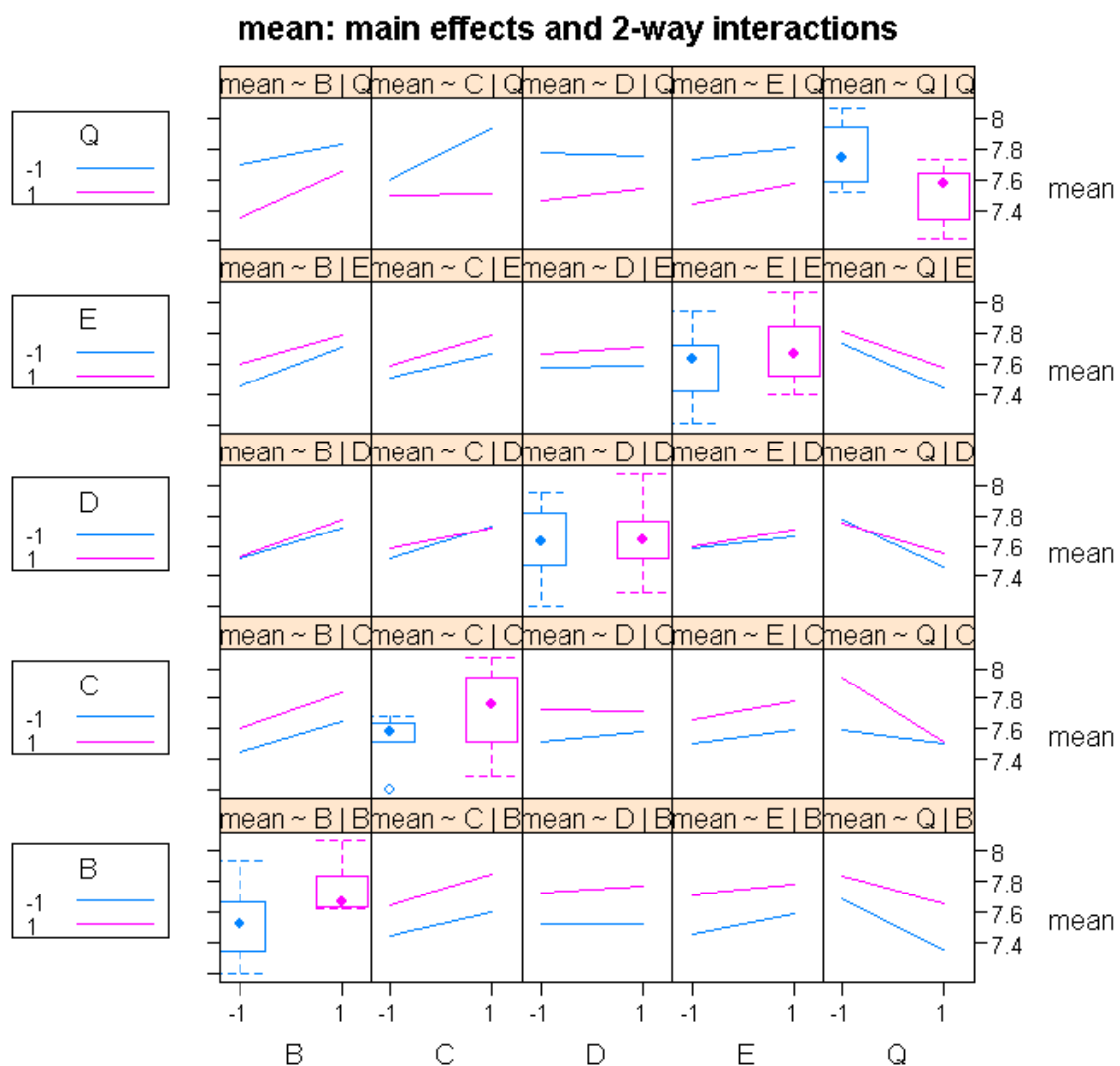


5.

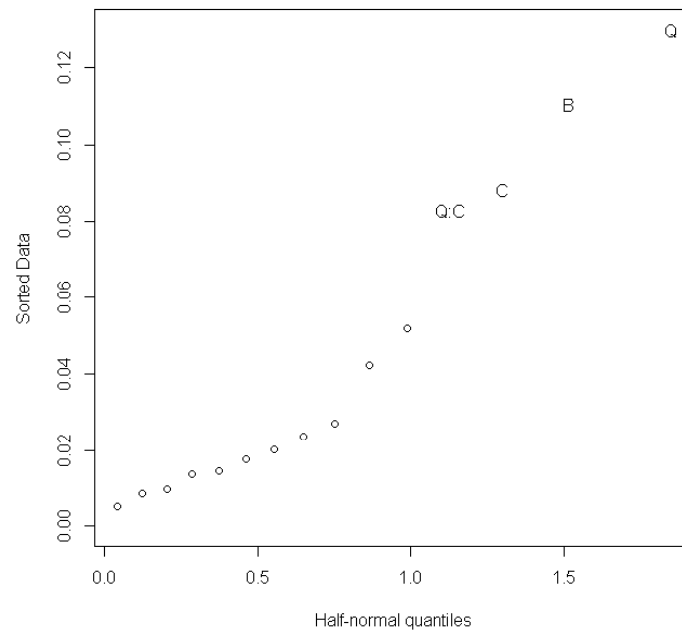
將每組設定的三個重複觀測值分別取平均與 $\log(\text{變異數})$ 後如下：

	B	C	D	E	Q	mean	lnS2
1	-1	1	1	-1	-1	7.790000	-8.111728
2	1	1	1	1	-1	8.070000	-3.600869
3	-1	-1	1	1	-1	7.520000	-6.725434
4	1	-1	1	-1	-1	7.633333	-4.562749
5	-1	1	-1	1	-1	7.940000	-5.626821
6	1	1	-1	-1	-1	7.946667	-3.003093
7	-1	-1	-1	-1	-1	7.540000	-4.779524
8	1	-1	-1	1	-1	7.686667	-4.158350
9	-1	1	1	-1	1	7.290000	-3.288762
10	1	1	1	1	1	7.733333	-2.740573
11	-1	-1	1	1	1	7.520000	-6.725434
12	1	-1	1	-1	1	7.646667	-4.684935
13	-1	1	-1	1	1	7.400000	-5.339139
14	1	1	-1	-1	1	7.623333	-5.464766
15	-1	-1	-1	-1	1	7.203333	-6.417132
16	1	-1	-1	1	1	7.633333	-3.671695

此設計為一 2^{5-1} design，其 defining relation 為 $I=BCDE$ ，因此每個效應都和某一個效應完全混淆。首先對 mean 作分析。觀察 ME plot 與 2fi's plot (下圖)發現效應 $B=CDE, C=BDE, Q=BCDEQ$ 與 $CQ=BDEQ$ 的趨勢較明顯，根據 effect ordering principle in parameter design，這些趨勢歸因於(B,C,Q,CQ)造成的影響。



有了圖形上觀察到的現象後，接著用數值分析來找出顯著的效應。由於這是一個正規設計且無多餘自由度來估計 **error variance**，故使用 **half-normal plot** 來挑選重要效應。



從圖上可觀察到(B,C,Q,CQ)明顯偏離其餘效應形成的群，故取此 4 個效應來配適模型如下：

```
Call:
lm(formula = mean ~ B + C + Q + C:Q, data = data2)

Residuals:
    Min       1Q   Median       3Q      Max
-0.18688 -0.04510  0.01146  0.04063  0.12979

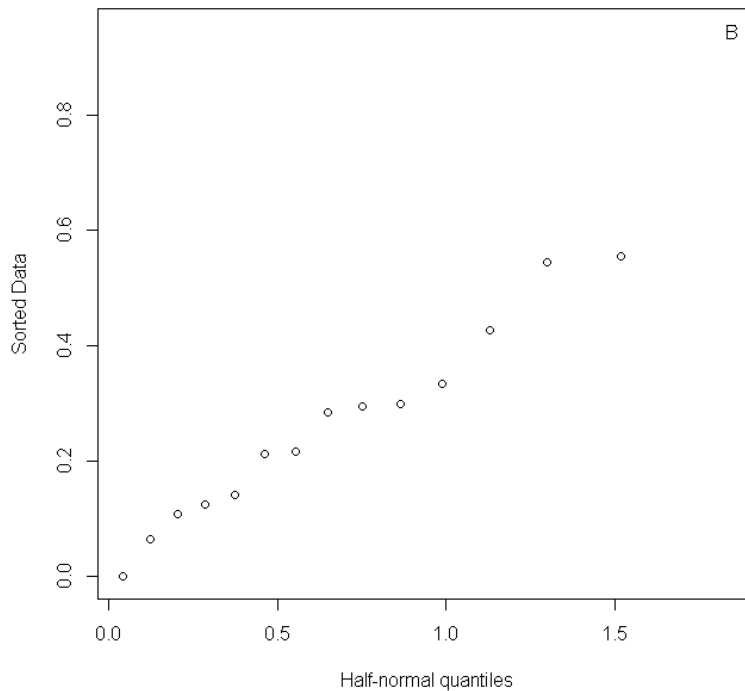
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.63604     0.02534 301.327  < 2e-16 ***
B              0.11063     0.02534   4.365 0.001126 **
C              0.08813     0.02534   3.478 0.005171 **
Q             -0.12979     0.02534  -5.122 0.000333 ***
C:Q           -0.08271     0.02534  -3.264 0.007548 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1014 on 11 degrees of freedom
Multiple R-squared:  0.8608,    Adjusted R-squared:  0.8102
F-statistic: 17.01 on 4 and 11 DF,  p-value: 0.0001117
```

得 final fitted model 為

$$y = 7.636 + 0.110 * B + 0.088 * C - 0.129 * Q - 0.082 * CQ$$

接著對 $\log(\text{變異數})$ 做分析，觀察 half-normal plot(下圖)



由 half-normal plot 可知只有 B 顯著，故配適模型如下：

```
Call:
lm(formula = lnS2 ~ B, data = data2)

Residuals:
    Min       1Q   Median       3Q      Max
-2.23498 -0.73646  0.03873  0.64890  2.58798

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -4.9313     0.3046 -16.190 1.85e-10 ***
B              0.9454     0.3046   3.104  0.00777 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.218 on 14 degrees of freedom
Multiple R-squared:  0.4076,    Adjusted R-squared:  0.3653
F-statistic: 9.634 on 1 and 14 DF,  p-value: 0.007772
```

Final fitted model 爲

$$\log(S^2) = -4.931 + 0.945 * B$$

令 Q 爲滿足 $P(Q=-1)=P(Q=1)=0.5$ 的隨機變數，則

$$E(y) = 7.636 + 0.110 * B + 0.088 * C$$

從 CQ 的 interaction plot 可知當 $C=-1$ 時 Q 的變異較小，故設 $C=-1$ ，再由 $\text{target}=8$ 可得 $B=4.109$ 。但此爲外插設定值，故取 $B=1$ 。此時 $\text{bias}=0.342$ 。第二種方法是直接計算 transmitted variance model 如下：

$$\text{var}(y) = (0.129 + 0.082 * C)^2$$

當 $C=-1$ 時可讓 $\text{var}(y)$ 最小，與 interaction plot 的資訊一致。因此最佳設定爲 $(B,C)=(1,-1)$ 、 (D,E) 可任意。注意到 $C=-1$ 的結論與使用 location-dispersion modeling 一致。

9.

(a)

此 control array 是一 2^{8-4} design，其 defining contrast subgroup 為

$$I = -ABCD = ABEF = ACEG = BCEH = -CDEF = -BDEG = -ADEH = BCFG = ACFH = ABGH = -ADFG = -BDFH = -CDGH = EFGH = -ABCDEFGH$$

首先要注意『eligible』、『clear』、『strongly clear』是定義在 ME 與 2fi's 上。

由於所有的 defining word 長度都至少為 4 且 A~H 每個字母都有出現在某個 defining word，因此 control factor 的 ME 全部都是 clear 但非 strongly clear。而由於所有 control factor 的 ME 都是 clear，因此所有 control factor 的 2fi's 都是 eligible。再觀察 defining contrast subgroup 可發現任意兩個 control factor 的交互作用都有出現在某個 defining word 內，因此所有 control factor 的 2fi's 都不是 clear。

由於 noise array 是一個 2x4 的 full factorial design，因此 noise factor 的 ME 與 2fi's 都是 strongly clear。

因為整個 design 是一個 cross array，control factor 的 ME 是 clear 且 noise factor 的 ME 是 strongly clear，故所有 control-by-noise 的交互作用都是 strongly clear。

(b)

因子 M 為 4-level，其三個主效應可 model 成 M_I, M_q, M_c (see Wu and Hamada, 2009)，且一個 4-level 因子可由 2 個 2-level 因子形成，故原實驗可視為一 11 個 2-level 因子的設計。取一設計滿足 $G=ABC, H=ABD, L=ACDE, M_I=ACDF, M_q=ABEF$ ，此為一 2^{11-5}_{IV} 的 minimum aberration 設計 (Table 5A.5 in Wu and Hamada, 2009)。所有的 ME (control & noise factor) 皆為 clear，包含 M_q 的所有 10 個 2fi's 為 clear，且 (A,B,C,D,G,H) 任一因子與 (E,F,L, M_I) 任一因子形成的 2fi's 為 clear。故此設計滿足『all control factor MEs are clear』、『some CxN are clear』與『some CxC are clear』。

(c)

Array (a) 與 Array (b) 都滿足『all control factor MEs are clear』。然而在 2fi's 的部份，Array (a) 使得所有 CxC 皆非 clear，但 Array (b) 可使得某些 CxC clear。另一方面 Array (a) 可使得所有 CxN strongly clear，但 Array (b) 只能讓某些 CxN clear。值得注意的是除了 estimation capacity 外，Array (a) 的 run size 是 Array (b) 的兩倍。

22.

(a)

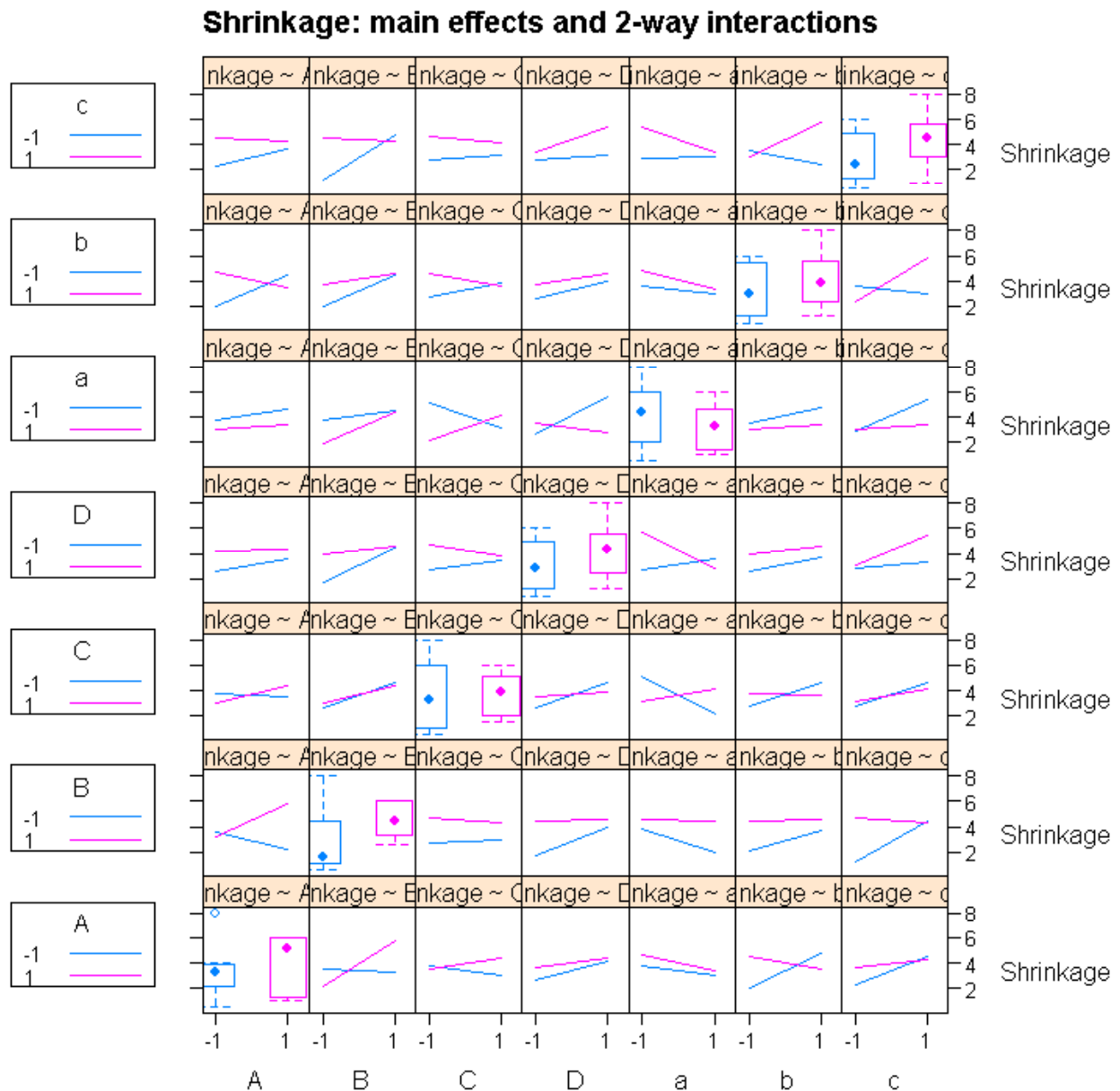
此設計的 defining contrast subgroup 為

$$I = ABCa = BCDB = ACDc = ADab = BDac = ABbc = Cabc$$

所有的 defining word 長度都是 4 且任意兩個因子的交互作用都有出現在某個 defining word 內，故無 clear 2fi's。

(b)

觀察 ME plot 與 2fi's plot 如下：



較明顯的 ME 為 A,B,a,b,c，較明顯的 2fi's 為 $AB=Ca=bc$ 、 $AC=Ba=Dc$ 、 $CD=Bb=Ac$ 、 $Da=Bc=Ab$ 、 $Cb=BD=ac$ ，要注意每一張圖都呈現一組 joint effect。接著作數值分析來進一步觀察：

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.4750     0.2168  16.029 8.86e-05 ***
A              0.3000     0.2424   1.238  0.2835
B              0.8250     0.2424   3.404  0.0272 *
C              0.0125     0.2424   0.052  0.9613
D              0.5750     0.2424   2.372  0.0766 .
A:B            0.9875     0.2424   4.074  0.0152 *
A:C            0.4250     0.2424   1.753  0.1544
B:C           -0.1500     0.2424  -0.619  0.5695
A:D           -0.2125     0.2424  -0.877  0.4301
B:D           -0.5125     0.2424  -2.114  0.1020
C:D           -0.4000     0.2424  -1.650  0.1742
A:B:C         -0.4875     0.2424  -2.011  0.1146
A:B:D         -0.0500     0.2424  -0.206  0.8466
A:C:D          0.7125     0.2424   2.940  0.0424 *
B:C:D          0.4125     0.2424   1.702  0.1640
A:B:C:D       -0.9250     0.2424  -3.816  0.0188 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9695 on 4 degrees of freedom
Multiple R-squared:  0.9506,    Adjusted R-squared:  0.7653
F-statistic:  5.13 on 15 and 4 DF,  p-value: 0.06276

```

由於這 16 個效應是直交的，故可直接保留顯著的效應得到以下模型：

```

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   3.4750     0.2835  12.258 7.12e-09 ***
B              0.8250     0.3170   2.603  0.02086 *
D              0.5750     0.3170   1.814  0.09114 .
B:A            0.9875     0.3170   3.116  0.00759 **
D:A:C          0.7125     0.3170   2.248  0.04121 *
B:D:A:C       -0.9250     0.3170  -2.918  0.01123 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.268 on 14 degrees of freedom
Multiple R-squared:  0.7043,    Adjusted R-squared:  0.5987
F-statistic:  6.669 on 5 and 14 DF,  p-value: 0.002248

```

因 $AB=Ca=bc$ ，由 effect ordering principle in parameter design 可將此顯著性歸因於 Ca 。同理將 ACD 歸因於 c ，而 $ABCD=Da=Ab=Bc$ 且 model 中已有主效應 B, c ，故將之歸因於 Bc (effect heredity principle)。Final fitted model 如下：

$$y = 3.475 + 0.825 * B + 0.575 * D + 0.987 * Ca + 0.712 * c - 0.925 * Bc$$

令 (a,b,c) 為獨立且滿足 $P(a=-1)=P(a=1)=P(b=-1)=P(b=1)=P(c=-1)=P(c=1)=0.5$ 的隨機變數。則

$$E(y) = 3.475 + 0.825 * B + 0.575 * D$$

$$\text{var}(y) = (0.712 - 0.925 * B)^2 + 0.987^2$$

$B=1$ 時 $\text{var}(y)$ 較 $B=-1$ 小，此現象與 interaction plot 一致：



故取 $B=1$ 可降低 noise factor 造成的變異。

made by 張明中 助教

(c)

4 個中心點對估計交互作用無任何幫助，這點可由 model matrix 上對應到交互作用的行看出。加入中心點的目的通常是要瞭解 error variance 或是 joint quadratic effect。若是要瞭解 error variance，則需要『constant variance』的假設成立，如此中心點造成的變異才能適用在整個實驗區間上。

(d)

原始設計會造成所有 2fi's 不 clear，若採用 $a=AB, b=BC, c=AC$ ，則

$$I = ABa = BCb = ACc = ACab = BCac = ABbc = abc$$

雖然 resolution 降成 3，但與 D 相關的所有 2fi's 都為 clear。當實驗者要對 D 放比較多資源可考慮使用。