

R與實驗設計應用 (Design Of Experiments with R)

大數據分析

- R/Python/Julia/SQL程式設計與應用
(R/Python/Julia/SQL Programming and Application)
- 資料視覺化 (Data Visualization)
- 機器學習 (Machine Learning)
- 統計品管 (Statistical Quality Control)
- 最佳化 (Optimization)



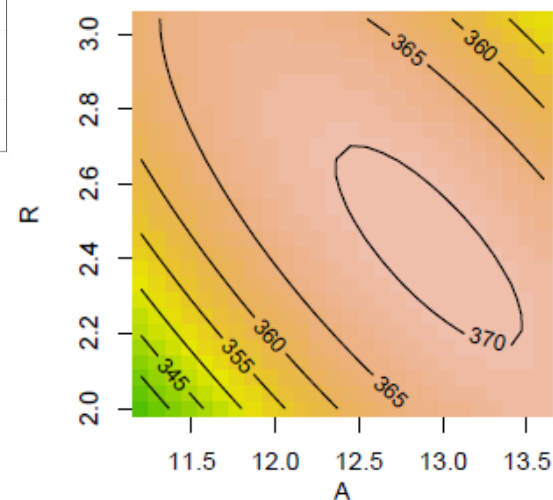
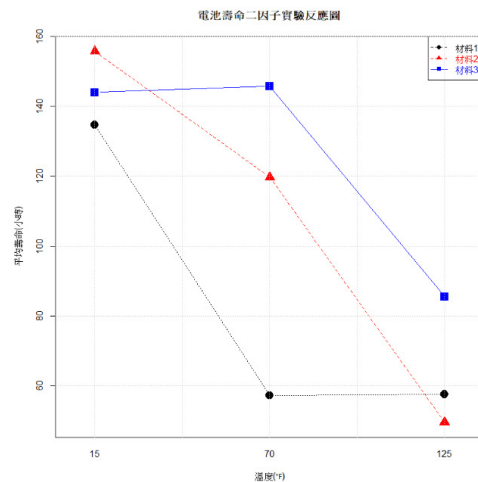
李明昌博士

alan9956@gmail.com

@RWEPA

大綱

1. 實驗設計簡介
2. R, RStudio簡介
3. 一因子實驗-變異數分析
4. 隨機化完全區集設計
5. 因子設計
6. 2^3 因子設計
7. 反應曲面法

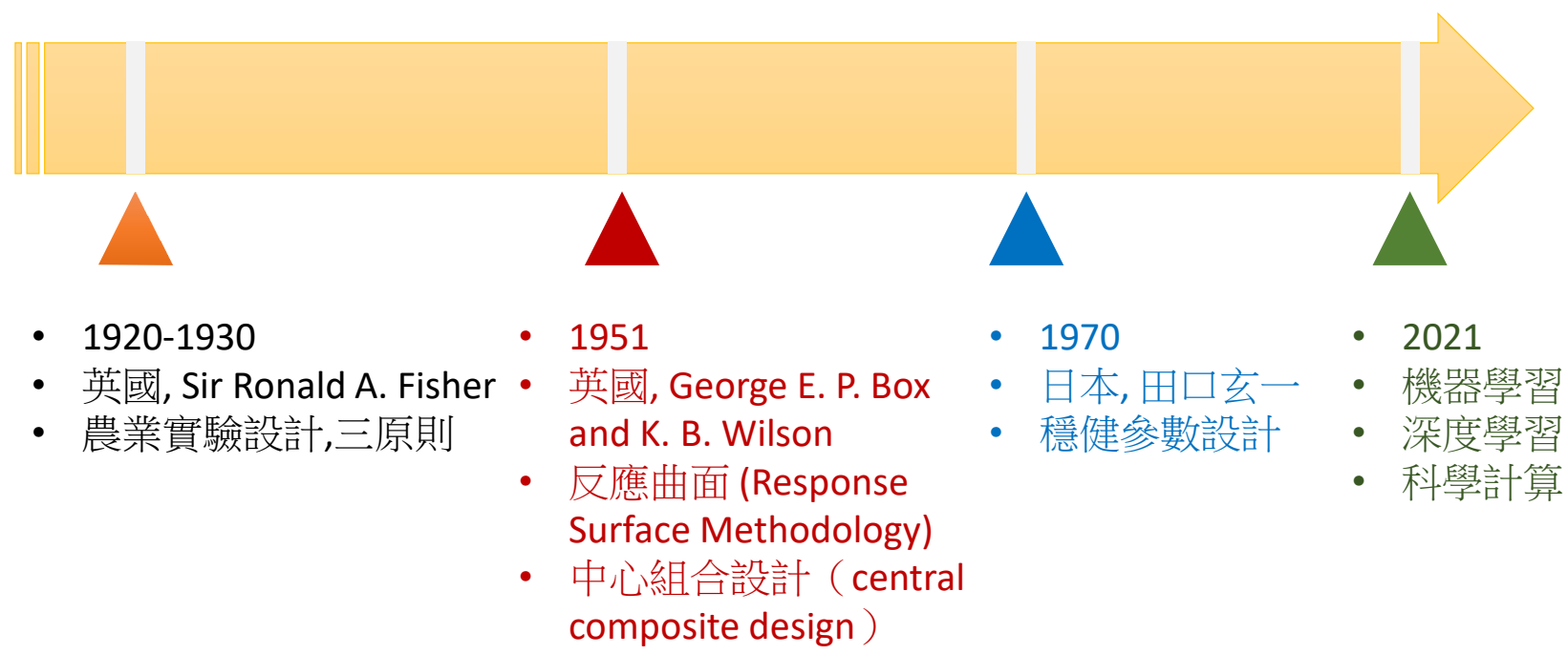


1. 實驗設計簡介

實驗設計

- 實驗設計(Design Of Experiments, DOE)，又稱為試驗設計，屬於科學探究的一種規劃設計，其目的是獲得特定的過程或是系統的相關訊息。
- 實驗中盡可能排除實驗變數以外的影響和干擾，以提高實驗效果的精準度。
- 實驗目標：
 - 製程良率的提昇
 - 找出最大或最小反應值的產品設計參數
 - 比較在不同可控制變數條件下所獲得的反應值
 - 建立數學模型以預測未來的反應值

實驗設計發展



實驗設計的用語

- **實驗 (Experiment)** 是一種行動，是實驗者改變至少一個正在研究的變數，然後觀察此變數的影響。一般被動收集資料不是實驗。
- **實驗單位 (Experimental unit)**：接受研究的對象，如特定商場(單位)、餐廳或特定製程、產品、消費者。
- **因子 (Factor, Treatment factor, Independent variable)**：能夠控制或調整的變數。
- **水準 (Level)**：一個因子的各種不同程度。例如：溫度水準為 {30、40、50 °C}。
- **處理 (Treatment)**：因子中一個或以上水準的組合。溫度30 °C、溫度40 °C 處理。
- **反應變數 (Response variable)**：實驗欲研究的目標變數。

實驗設計的用語 (續)

- **背景變數(Background variable, Lurking variable 潛在變數)** 是實驗者未知或無法控制的變數，它可能對實驗結果產生影響。在穩健規劃的實驗設計中，這些潛在變數的影響應該平衡，以避免改變研究結論。
- **效果 (Effect)**是由因子或自變數的改變而引起的反應變化，一般使用反應變數加以估算。
 - 在進行實驗設計後，可以通過觀察到的反應數據計算來估算效果。該估計值稱為計算效果 (Calculated effect)。
 - 在進行實驗之前，研究人員可能知道效果應該有多大才能具有實際重要性。這稱為實務效果(Practical effect)。
- **交絡 (Confounding)**是指實驗過程中，操控的獨立變數與被忽略的其他獨立變項之間，存在一些相互干擾的現象，而影響到獨立變數與依變數之間因果關係的認定。
 - 在這種情況下，無法確定哪個因素導致響應或因變量出現任何觀察到的變化。
 - 例:金屬試驗在50°C油水與100 °C鹽水實驗中，溫度與溶液效果產生交絡現象。

實驗設計的三個基本原理

- 隨機化(Randomization)
 - 實驗材料的配置與實驗進行的順序都要隨機化。但有些如化工產業隨機化進行較困難，此時可使用隨機效果模型(Random effects models)。
- 重複化(Replication)
 - 獨立反覆每一種因子組合的實驗，重複可以估計實驗誤差的變異數，並獲得更精確的估計值
- 區集化(Blocking)
 - 區集可以減少或消除干擾因子的變異。

實驗設計標準流程

1. 問題之認知
2. 實驗單位確認
3. 選擇可以測量的反應變數(Y)
4. 選擇因子(X)與可控制水準
5. 選擇適當之實驗設計 (隨機化，重複化，區集化)
6. 執行實驗(記錄完整步驟)
7. 資料分析
8. 結論、建議與佈署

2.R, RStudio簡介

R , RStudio

- <http://rwepa.blogspot.com/>
- 步驟一: R下載
- 步驟二: RStudio下載

RWEPA 搜尋此網誌 (例: task)

 搜尋

- [GitHub DataDemo](#)
- [iPAS-R-tutorial](#)
- [iPAS-Python-tutorial](#)
- [★★★★R入門資料分析與視覺化\(付費,中文字幕\)](#)
- [★★★★R商業預測與應用\(付費,中文字幕\)](#)
- [R教學-基礎篇/程式碼\(免費\)](#)
- [Python程式設計PDF\(免費\)](#)
- [★R 4.1.0-Windows下載](#) 1
- [★RStudio-1.4.1106下載](#) 2
- [R-bloggers](#)

R 基礎篇

- R 基礎篇 - 國立台北商業技術學院上課教材(238頁, 2011.7.4)
- <http://rwepa.blogspot.com/2013/01/r-201174.html>

2013年1月4日 星期五

R 基礎篇 - 國立台北商業技術學院上課教材(238頁, 2011.7.4)

整合下列五大主題：

1. Basic R
2. Preparing Data
3. Graphics
4. Applied Statistics
5. Application

檔案：`aMarvelousR_Lee(pp238).pdf`

頁數：238

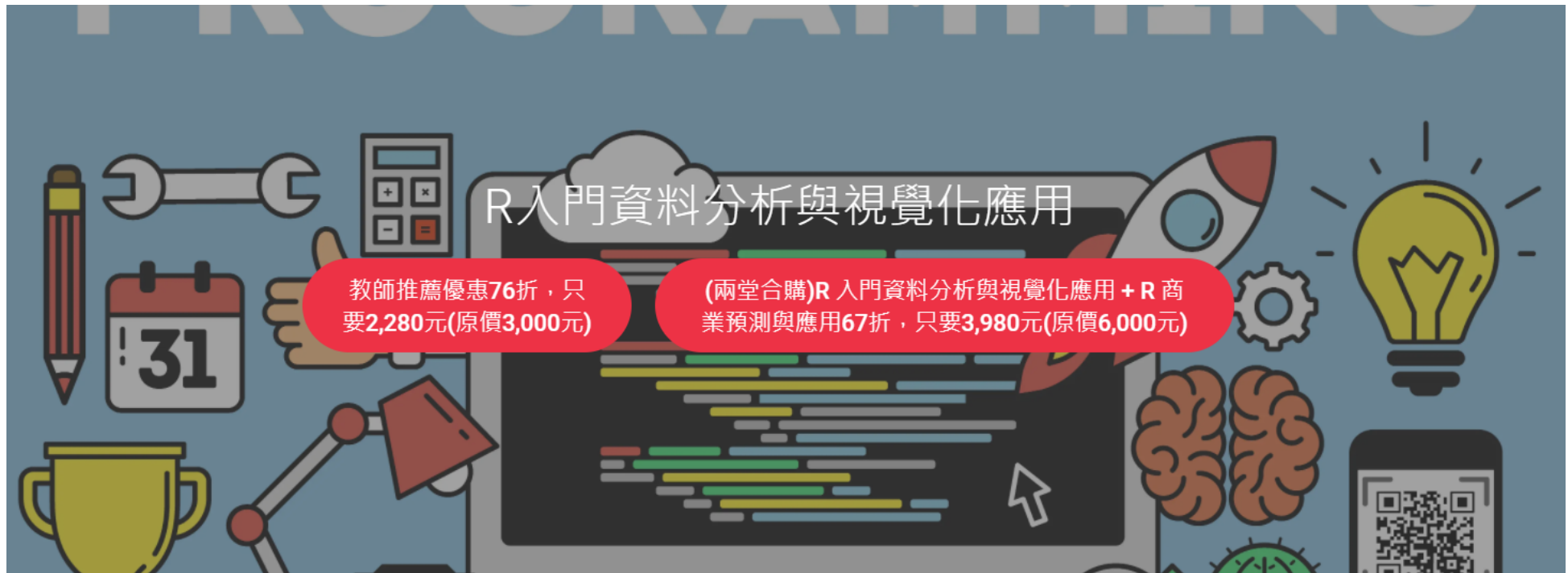
大小：9.86MB

R程式碼：`aMarvelousR_Lee(pp238).R`

更新日期：2020.6.3 (903行)

R入門資料分析與視覺化應用(中文,字幕)

- <https://courses.mastertalks.tw/courses/R-teacher>



R商業預測與應用(中文,字幕)

- <https://courses.mastertalks.tw/courses/R-2-teacher>



3.一因子實驗-變異數分析

(One way Analysis of Variance) or
(Single factor Analysis of Variance)

參考: Douglas C. Montgomery, Design and Analysis of Experiments, 10th Edition, Wiley, 2020.

ANOVA資料結構

處理	觀測值				總和	平均值
1	y_{11}	y_{12}		y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
...				...		
a	y_{a1}	y_{a2}		y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

- 平均值模型(Means model)

- $y_{ij} = \mu_i + \varepsilon_{ij}, i = 1, 2, \dots, a; j = 1, 2, \dots, n$

- 效果模型(Effects model)

- 考慮 $\mu_i = \mu + \tau_i, i = 1, 2, \dots, a$

- $y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \mu$ 總平均, τ_i 第 i 個處理效果

- 處理有 a 個(或因子有 a 個水準)
- y_{ij} : 第 i 個處理, 第 j 個觀測值
- μ_i : 第 i 處理的平均數, 即 $\bar{y}_{i.}$
- ε_{ij} : 隨機誤差, $E(\varepsilon_{ij}) = 0$, 即 $E(y_{ij}) = \mu_i$

假設檢定

- 考慮隨機誤差 ε_{ij} 獨立且服從常態隨機變數.
- ε_{ij} 平均數為0, 變異數為 σ^2
- y_{ij} 為彼此獨立 $y_{ij} \sim N(\mu + \tau_i, \sigma^2)$

資料模型

- $y_{i.} = \sum_{j=1}^n y_{ij}, i = 1, 2, \dots, a$
- $\bar{y}_{i.} = \frac{y_{i.}}{n}$
- $y_{..} = \sum_{i=1}^a \sum_{j=1}^n y_{ij}$
- $\bar{y}_{..} = \frac{y_{..}}{N}, N = an$ 總觀測值個數
- 檢定 a 個處理的平均值是否相等
 - $H_0: \mu_1 = \mu_2 = \dots = \mu_n$
 - $H_1: \mu_i \neq \mu_j$ 至少有一對 (i, j)

處理	觀測值				總和	平均值
1	y_{11}	y_{12}		y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
...				...		
a	y_{a1}	y_{a2}		y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

$$\because \mu = \frac{\sum_{i=1}^a \mu_i}{a} = \frac{\sum_{i=1}^a (\mu + \tau_i)}{a} = \frac{a\mu + \sum_{i=1}^a \tau_i}{a} = \mu + \frac{\sum_{i=1}^a \tau_i}{a}$$

$$\because \sum_{i=1}^a \tau_i = 0, \text{ 其中 } \tau_i = \mu_i - \mu$$

- $H_0: \tau_1 = \tau_2 = \dots = \tau_a = 0$
- $H_1: \text{至少有一個 } i \text{ 使得 } \tau_i \neq 0$



總平方和分解法

- 總平方和 (Total Sum of Squares)

$$(a + b)^2 = a^2 + 2ab + b^2$$

$$SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2 = \sum_{i=1}^a \sum_{j=1}^n [(\bar{y}_{i.} - \bar{y}_{..}) + (y_{ij} - \bar{y}_{i.})]^2$$

$$= n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + 2 \sum_{i=1}^a \sum_{j=1}^n (\bar{y}_{i.} - \bar{y}_{..}) (y_{ij} - \bar{y}_{i.}) + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$$

$$= n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$$

$$\sum_{j=1}^n (y_{ij} - \bar{y}_{i.}) = y_{i.} - n\bar{y}_{i.} = y_{i.} - n \frac{y_{i.}}{n} = 0$$

= 平均與總平均差的平方和+處理內觀測值與處理平均的差平方和

= 處理平均間差異+隨機誤差

- $SS_T = SS_{Treatment} + SS_E$ = 處理平方和(處理間)+誤差平方和(處理內)

變異數分析表

變異來源	平方和	自由度	均方	F_0
處理間	$SS_{Treatment}$	$a - 1$	$MS_{Treatment} = \frac{SS_{Treatment}}{a - 1}$	$F_0 = \frac{MS_{Treatment}}{MS_E}$
誤差	SS_E	$N - a$	$MS_E = \frac{SS_E}{N - a}$	
總和	SS_T	$N - 1$		

註:

1. 總樣本數 $N = a \times n$
2. 處理間平方和 $SS_{Treatment} = n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2$
3. 誤差平方和 $SS_E = SS_T - SS_{Treatment}$
4. 總平方和 $SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$
5. 誤差項自由度 $(N - 1) - (a - 1) = N - a$
6. 如果 $F_0 \geq F_\alpha(a - 1, N - a)$, 則表示 **p值** $< \alpha$, 即拒絕 H_0 , 接受不同處理平均數有不同.

變異數分析-範例

- 考慮蝕刻製程，研究 RF功率(X)與蝕刻率(Y)的關係
- RF功率有4個水準，160,180,200,220W
- 每個RF功率測試5片晶圓
- 反應變數是蝕刻率
- 實驗設計計劃
 - 水準 $a = 4$
 - 重複 $n = 5$
 - 只考慮RF功率，因此是一因子實驗
 - 全部實驗次數20

步驟1.進行隨機化實驗

```
> # 步驟1. 進行隨機化實驗
> rf <- c("160", "180", "200", "220")
>
> rfPower <- gl(4, 5, label=rf)
> rfPower
[1] 160 160 160 160 160 180 180 180 180 180 200 200 200 200 200 220 220 220 220 220
Levels: 160 180 200 220
>
> # 隨機化
> set.seed(168)
> ind <- sample(4*5, 20)
> rfPower <- rfPower[ind]
> head(rfPower)
[1] 200 220 220 200 160 180
Levels: 160 180 200 220
> # [1] 200 220 220 200 160 180
> # Levels: 160 180 200 220
> # 第1筆試驗為200w, 第2筆為220w
```

步驟2.實驗結果

蝕刻率資料表 (Å/min)					
RF功率 (W)	晶圓1	晶圓2	晶圓3	晶圓4	晶圓5
160	575	542	530	539	570
180	565	593	590	579	610
200	600	651	610	637	629
220	725	700	715	685	710

- 埃格斯特朗 (Ångström, 簡稱埃, 符號Å)
- Å是一個長度計量單位。
- $1 \text{ Å} = 10^{-10} \text{ 米} = 0.1 \text{ 奈米}$ 。

```
> # 步驟2.實驗結果
> EtchRate <- c(575, 542, 530, 539, 570,
+              565, 593, 590, 579, 610,
+              600, 651, 610, 637, 629,
+              725, 700, 715, 685, 710)
>
> plasmaEtch <- data.frame(rfPower = as.factor(rfPower), EtchRate)
> plasmaEtch
  rfPower EtchRate
1     160      575
2     160      542
3     160      530
4     160      539
5     160      570
6     180      565
7     180      593
8     180      590
9     180      579
10    180      610
11    200      600
12    200      651
13    200      610
14    200      637
15    200      629
16    220      725
17    220      700
18    220      715
19    220      685
20    220      710
>
```


步驟3.ANOVA分析

```
> y_dot_dot <- mean(plasmaEtch$EtchRate) # 617.75
> sum(plasmaEtch$EtchRate - y_dot_dot)^2 # ERROR
[1] 0
> SST <- sum((plasmaEtch$EtchRate - y_dot_dot)^2) # 72209.75
>
> y_i_dot <- aggregate(EtchRate ~ rfPower, data = plasmaEtch, FUN = mean)
> y_i_dot
```

	rfPower	EtchRate	1	2	3
1	160	551.2			
2	180	587.4			
3	200	625.4			
4	220	707.0			

```
> SSTreatment <- 5*sum((y_i_dot$EtchRate - y_dot_dot)^2)
> SSTreatment # 66870.55
[1] 66870.55
>
> SSE <- SST - SSTreatment # 5339.2
> SSE
[1] 5339.2
>
```

變異數分析表

變異來源	平方和	自由度	均方	F_0
處理間	$SS_{Treatment} = 66870.55$	$a - 1 = 4 - 1 = 3$	$MS_{Treatment} = \frac{66870.55}{3}$ $= 22290.18$	$F_0 = \frac{22290.18}{333.7} = 66.8$
誤差	$SS_E = 5339.2$	$N - a = 20 - 4 = 16$	$MS_E = \frac{5339.2}{16}$ $= 333.7$	
總和	$SS_T = 72209.75$	$N - 1 = 4 \times 5 - 1 = 19$		

變異數分析表 (續)

```
> MSTreatment <- 66870.55/3 # 22290.18
> MSE <- 5339.2/16 # 333.7
>
> F0 <- MSTreatment/MSE # 66.8
>
> # a.using critical values
> F_critical <- qf(p = 0.95, df1 = 3, df2 = 16) # 3.238872
> ifelse(F0 >= F_critical, "Reject Ho", "Accept H0") # "Reject Ho"
[1] "Reject Ho"
>
> # b.using pvalue
> pvalue <- pf(q = F0, df1 = 3, df2 = 16, lower.tail = FALSE) # 2.882866e-09, p值很小
> ifelse(pvalue < 0.05, "Reject Ho", "Accept H0") # "Reject Ho"
[1] "Reject Ho"
>
```

1

2

3

步驟4.使用 aov {stats}

```
> # ANOVA分析-aov
> plasmaEtchANOVA <- aov(EtchRate ~ rfPower, data=plasmaEtch)
```

```
>
> # 模型結果
> plasmaEtchANOVA
```

Call:

```
aov(formula = EtchRate ~ rfPower, data = plasmaEtch)
```

Terms:

	rfPower	Residuals
Sum of Squares	66870.55	5339.20
Deg. of Freedom	3	16

Residual standard error: 18.26746

Estimated effects are balanced

```
>
```

```
> summary(plasmaEtchANOVA)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
rfPower	3	66871	22290	66.8	2.88e-09 ***
Residuals	16	5339	334		

```
---
```

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```
>
```

- p 值很小, 拒絕 H_0 ,
- 結果與變異數分析表相同

4. 隨機化完全區集設計

隨機化完全區集設計

Randomized Complete Block Design

- 在實驗中，干擾因子之一是未知且無法控制，可以使用隨機化降低此情形。
- 在實驗中，干擾因子之一是已知且無法控制，即可觀察其值，可以使用共變數分析(Analysis of covariance, ANCOVE)。
- 在實驗中，干擾因子之一是已知且可以控制，可以使用區集化(blocking)設計，以消除處理間統計比較的影響。

隨機化完全區集設計使用時機

- 考慮4種不同的錐頭在硬度測試機台是否會有不同影響，即4種處理。4種金屬物品須測量，每個錐頭實驗4次，合計實驗16次，即實驗單位為16。
- 考慮避免因金屬物品之間的差異而影響，使用隨機化完全區集設計(Randomized Complete Block Design, RCBD)，完全表示每個區集包括所有的處理，區集內的處理順序的隨機的。

金屬物品1(區集1)	金屬物品2(區集2)	金屬物品3(區集3)	金屬物品4(區集4)
錐頭2	錐頭2	錐頭1	錐頭4
錐頭1	錐頭3	錐頭2	錐頭3
錐頭4	錐頭1	錐頭4	錐頭2
錐頭3	錐頭4	錐頭3	錐頭1

RCBD資料結構

- $y_{ij} = \mu + \tau_i + \beta_j + \varepsilon_{ij}, i = 1, 2, \dots, a, b = 1, 2, \dots, b$
- μ 總平均, τ_i 第 i 個處理效果, β_j 第 j 舍區集效果
- ε_{ij} 隨機誤差項 $\varepsilon_{ij} \sim iid.N(0, \sigma^2)$
- $\sum_{i=1}^a \tau_i = 0, \sum_{j=1}^b \beta_j = 0$
- $SS_T = \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{..})^2$
 $= b \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + a \sum_{j=1}^b (\bar{y}_{.j} - \bar{y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^b (y_{ij} - \bar{y}_{.j} - \bar{y}_{i.} + \bar{y}_{..})^2$
- $SS_T = SS_{Treatment} + SS_{Blocks} + SS_E$

區集1	區集2	...	區集b
y_{11}	y_{12}		y_{1b}
y_{21}	y_{22}		y_{2b}
...
y_{a1}	y_{a2}		y_{ab}

RCBD變異數分析表

變異來源	平方和	自由度	均方	F_0
處理	$SS_{Treatment}$	$a - 1$	$MS_{Treatment} = \frac{SS_{Treatment}}{a - 1}$	$F_0 = \frac{MS_{Treatment}}{MS_E}$
區集	SS_{Blocks}	$b - 1$	$MS_{Blocks} = \frac{SS_{Blocks}}{b - 1}$	
誤差	SS_E	$(a - 1)(b - 1)$	$MS_E = \frac{SS_E}{(a - 1)(b - 1)}$	
總和	SS_T	$N - 1$		

註:

1. 總樣本數 $N = a \times b$
2. 如果 $F_0 \geq F_{\alpha}(a - 1, (a - 1)(b - 1))$, 則表示 $p\text{值} < \alpha$, 即拒絕 H_0 , 接受不同處理平均數有不同.

RCDB範例

- 考慮人工血管RCDB實驗
- 產品射出壓力 (Extrusion Pressure)
 - 處理4個水準:{8500,8700,8900,9100PSI}
- 樹脂批號 (Batch of Resin)
 - 區集6個
- 反應變數是良率

RCDB實驗結果

人工血管製造實驗

射出壓力(PSI)	樹脂批號1	樹脂批號2	樹脂批號3	樹脂批號4	樹脂批號5	樹脂批號6
8500	90.3	89.2	98.2	93.9	87.4	97.9
8700	92.5	89.5	90.6	94.7	87	95.8
8900	85.5	90.8	89.6	86.2	88	93.4
9100	82.5	89.5	85.6	87.4	78.9	90.7

步驟1. 將資料輸入Excel

- 依照資料集輸入資料
- 第1列是標題列
- 資料筆數為24
- 欄位個數為3
- 儲存檔名:
C:/rdata/R-rcbd.xlsx

	A	B	C
1	extrusionPressure	batch	yield
2	8500	1	90.3
3	8700	1	92.5
4	8900	1	85.5
5	9100	1	82.5
6	8500	2	89.2
7	8700	2	89.5
8	8900	2	90.8
9	9100	2	89.5
10	8500	3	98.2
11	8700	3	90.6
12	8900	3	89.6
13	9100	3	85.6
14	8500	4	93.9
15	8700	4	94.7
16	8900	4	86.2
17	9100	4	87.4
18	8500	5	87.4
19	8700	5	87.0
20	8900	5	88.0
21	9100	5	78.9
22	8500	6	97.9
23	8700	6	95.8
24	8900	6	93.4
25	9100	6	90.7

人工血管製造實驗						
射出壓力(PSI)	樹脂批號1	樹脂批號2	樹脂批號3	樹脂批號4	樹脂批號5	樹脂批號6
8500	90.3	89.2	98.2	93.9	87.4	97.9
8700	92.5	89.5	90.6	94.7	87	95.8
8900	85.5	90.8	89.6	86.2	88	93.4
9100	82.5	89.5	85.6	87.4	78.9	90.7

資料輸入: 由上至下, 由左至右

步驟2. 使用 readxl 套件匯入Excel

- 使用 readxl 套件
- 使用 dplyr
- 匯入 Excel: read_excel

```
> library(readxl) # read_excel
> library(dplyr) # v.1.0.6 mutate
```

載入套件：'dplyr'

下列物件被遮斷自 'package:stats':

filter, lag

下列物件被遮斷自 'package:base':

intersect, setdiff, setequal, union

```
> artificialVein <- read_excel("C:/rdata/R-rcbd.xlsx")
> artificialVein
```

A tibble: 24 x 3

	extrusionPressure	batch	yield
	<dbl>	<dbl>	<dbl>
1	8500	1	90.3
2	8700	1	92.5
3	8900	1	85.5
4	9100	1	82.5
5	8500	2	89.2
6	8700	2	89.5
7	8900	2	90.8
8	9100	2	89.5
9	8500	3	98.2
10	8700	3	90.6

... with 14 more rows

```
>
```

步驟3.將前2欄轉換為factor – mutate {dplyr}

```
> # 將前2欄轉換為factor
> artificialVein <- artificialVein %>%
+   mutate(across(c(extrusionPressure, batch),
+                   factor))
>
> artificialVein
# A tibble: 24 x 3
  extrusionPressure batch yield
  <fct>            <fct> <dbl>
1 8500              1    90.3
2 8700              1    92.5
3 8900              1    85.5
4 9100              1    82.5
5 8500              2    89.2
6 8700              2    89.5
7 8900              2    90.8
8 9100              2    89.5
9 8500              3    98.2
10 8700             3    90.6
# ... with 14 more rows
>
```

步驟4.建立 RCDB模型

```
> # 步驟4.建立 RCDB模型
> artificialVeinRCBD <- aov(yield ~ extrusionPressure + batch, data=artificialVein)
>
> artificialVeinRCBD
Call:
  aov(formula = yield ~ extrusionPressure + batch, data = artificialVein)

Terms:
              extrusionPressure      batch Residuals
Sum of Squares          178.1712 192.2521  109.8863
Deg. of Freedom              3         5         15

Residual standard error: 2.706612
Estimated effects may be unbalanced
>
> summary(artificialVeinRCBD)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
extrusionPressure	3	178.2	59.39	8.107	0.00192	**
batch	5	192.2	38.45	5.249	0.00553	**
Residuals	15	109.9	7.33			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

變異來源	自由度
處理	$a - 1 = 4 - 1 = 3$
區集	$b - 1 = 6 - 1 = 5$
誤差	$(a - 1)(b - 1) = 3 \times 5 = 15$
總和	$N - 1 = 4 \times 6 - 1 = 23$

5.因子設計

(Factorial Design)

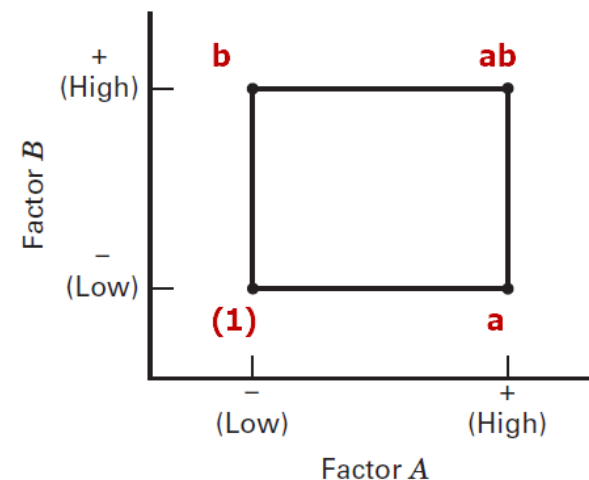
因子設計

- 考慮兩個或兩個以上因子實驗時，可採用因子設計。
- 因子設計表示每次實驗包括所有可能因子水準組合，因子水準組合可採取重複或不重複的實驗。
- 評量各實驗變數對實驗結果的個別效果稱為主效果(main effect)，如果是高低二個水準，則主效果 = 高水準平均值 - 低水準平均值。
- 將各種實驗水準的配對組合重複實驗時，可以評估兩個實驗變數之間交互作用(interaction effects)或交叉效果(cross effects)。
- 交互作用是一個實驗變數與結果之關係型態，會受另一個實驗變數於不同水準而產生顯著的改變之影響。

二因子因子實驗

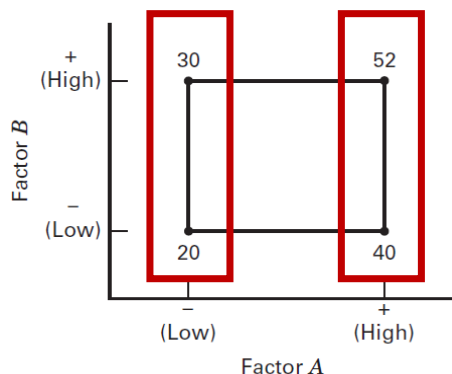
- 因子A主效果 $\frac{ab+a}{2} - \frac{b+(1)}{2} = \frac{1}{2}(ab + a - b - (1))$
- 因子B主效果 $\frac{ab+b}{2} - \frac{a+(1)}{2} = \frac{1}{2}(ab + b - a - (1))$
- 因子AB交互效果: B在高水準時A的效果與B在低水準時A的效果的平均差異。

$$\frac{ab - b}{2} - \frac{a - (1)}{2} = \frac{1}{2}(ab + (1) - a - b)$$



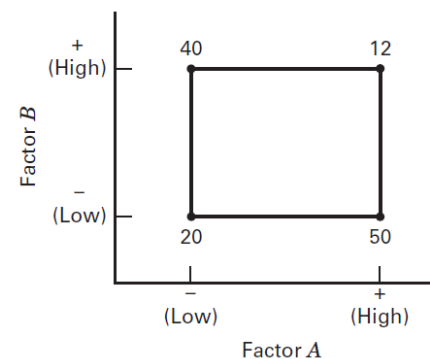
二因子因子實驗-範例

- 二因子(A,B)且無交互作用



- 因子A主效果 $\frac{1}{2}(52 + 40 - 30 - 20) = 21$
- 因子B主效果 $\frac{1}{2}(52 + 30 - 40 - 20) = 11$
- 交互作用AB效果 $\frac{1}{2}(52 + 20 - 40 - 30) = 1$

- 二因子(A,B)且有交互作用



- 因子A主效果 $\frac{1}{2}(12 + 50 - 40 - 20) = 1$
- 因子B主效果 $\frac{1}{2}(12 + 40 - 50 - 20) = -9$
- 交互作用AB效果 $\frac{1}{2}(12 + 20 - 50 - 40) = -29$

二因子因子實驗模型

- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$
- y 是反應變數
- β_0 表示常數項, $\hat{\beta}_0 = \frac{(1)+a+b+ab}{4}$
- x_1 表示因子A的變數, $\hat{\beta}_1 = \frac{x_1 \text{主效果}}{2}$
- x_2 表示因子B的變數, $\hat{\beta}_2 = \frac{x_2 \text{主效果}}{2}$
- $x_1 x_2$ 表示因子A,B之間的交互作用, $\hat{\beta}_{12} = \frac{x_1 x_2 \text{交互作用}}{2}$
- ε 隨機誤差項

二因子因子實驗模型-估計

考慮二因子(A,B)且無交互作用 $y = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_{12}x_1x_2 + \varepsilon$

- 常數項 $\hat{\beta}_0 = \frac{(1)+a+b+ab}{4} = \frac{20+40+30+52}{4} = 35.5$
- 因子A主效果 21, $\hat{\beta}_1 = \frac{21}{2} = 10.5$
- 因子B主效果 11, $\hat{\beta}_2 = \frac{11}{2} = 5.5$
- 交互作用AB效果 1, $\hat{\beta}_{12} = \frac{1}{2} = 0.5$
- 二因子因子實驗模型 $\hat{y} = 35.5 + 10.5x_1 + 5.5x_2 + 0.5x_1x_2$
- 考慮交作作用很小忽略 $\hat{y} = 35.5 + 10.5x_1 + 5.5x_2$

二因子因子實驗變異數分析表

變異來源	平方和	自由度	均方	F_0
A處理	SS_A	$a - 1$	$MS_A = \frac{SS_A}{a - 1}$	$F_A = \frac{MS_A}{MS_E}$
B處理	SS_B	$b - 1$	$MS_B = \frac{SS_B}{b - 1}$	$F_B = \frac{MS_B}{MS_E}$
交互作用	SS_{AB}	$(a - 1)(b - 1)$	$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)}$	$F_{AB} = \frac{MS_{AB}}{MS_E}$
誤差	SS_E	$ab(n - 1)$	$MS_E = \frac{SS_E}{ab(n - 1)}$	
總和	SS_T	$N - 1$		

註：

- 1. 總樣本數 $N = a \times b \times n$, 其中 n 表示重複次數.
- 2. 如果 $F_A \geq F_\alpha(a - 1, ab(n - 1))$, 則表示 $p\text{值} < \alpha$, 即拒絕 H_0 , 接受不同A處理平均數有不同.

二因子因子實驗-範例

- 考慮電池壽命(小時)實驗
- 電池極板材料有3個水準 $a = 3$
- 外在環境溫度有3個水準: 15, 70, 125°F, $b = 3$
- 每次水準組合重複4次實驗 $n = 4$
- 反應變數是電池壽命(小時)

二因子因子實驗結果

電池壽命實驗			
材料	溫度 15°F	溫度 70°F	溫度 125°F
1	130 155	34 40	20 70
	74 180	80 75	82 58
2	150 188	136 122	25 70
	159 126	106 115	58 45
3	138 110	174 120	96 104
	168 160	150 139	82 60

步驟1. 將資料輸入Excel

- 依照資料集輸入資料
- 第1列是標題列
- 資料筆數為 $3 \times 3 \times 4 = 36$
- 欄位個數為4
- 儲存檔名:
C:/rdata/R-fd.xlsx

電池壽命實驗			
材料	溫度 15°F	溫度 70°F	溫度 125°F
1	130 155 74 180	34 40 80 75	20 70 82 58
2	150 188 159 126	136 122 106 115	25 70 58 45
3	138 110 168 160	174 120 150 155	96 104 82 60

	A	B	C	D
1	material	temperature	replication	life
2	1	15	1	130
3	1	15	2	155
4	1	15	3	74
5	1	15	4	180
6	1	70	1	34
7	1	70	2	40
8	1	70	3	80
9	1	70	4	75
10	1	125	1	20
11	1	125	2	70
12	1	125	3	82
13	1	125	4	58
14	2	15	1	150

步驟2. 使用 readxl 套件匯入Excel

- 使用 readxl 套件
- 使用 dplyr
- 匯入 Excel: read_excel

```
> # 步驟1. 將資料輸入Excel
> # 路徑 C:/rdata/R-fd.xlsx
>
> # 步驟2. 使用 readxl 套件匯入Excel
> library(readxl) # read_excel
> library(dplyr) # v.1.0.6 mutate

載入套件: 'dplyr'

下列物件被遮斷自 'package:stats':

    filter, lag

下列物件被遮斷自 'package:base':

    intersect, setdiff, setequal, union

>
> batteryLife <- read_excel("C:/rdata/R-fd.xlsx")
> batteryLife
# A tibble: 36 x 4
   material temperature replication  life
  <dbl>      <dbl>      <dbl> <dbl>
1       1         15           1    130
2       1         15           2    155
3       1         15           3     74
4       1         15           4    180
5       1         70           1     34
6       1         70           2     40
7       1         70           3     80
8       1         70           4     75
9       1        125           1     20
10      1        125           2     70
# ... with 26 more rows
>
```

步驟3.將前3欄轉換為factor – mutate {dplyr}

```
> # 步驟3.將前3欄轉換為factor – mutate {dplyr}
> batteryLife <- batteryLife %>%
+   mutate(across(c(material:replication),
+                 factor))
> batteryLife
# A tibble: 36 x 4
  material temperature replication  life
  <fct>      <fct>      <fct>    <dbl>
1 1         15         1        130
2 1         15         2        155
3 1         15         3         74
4 1         15         4        180
5 1         70         1         34
6 1         70         2         40
7 1         70         3         80
8 1         70         4         75
9 1        125         1         20
10 1        125         2         70
# ... with 26 more rows
>
```

步驟4.建立二因子因子實驗模型

使用乘號 ×: 包括交互作用

```
> # 步驟4.建立 二因子模型
> batteryLifeFD <- aov(life ~ material*temperature, data=batteryLife)
> batteryLifeFD
Call:
  aov(formula = life ~ material * temperature, data = batteryLife)

Terms:
 material temperature material:temperature Residuals
Sum of Squares 10683.72    39118.72          9613.78   18230.75
Deg. of Freedom      2          2              4       27

Residual standard error: 25.98486
Estimated effects may be unbalanced
> summary(batteryLifeFD)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
material	2	10684	5342	7.911	0.00198	**
temperature	2	39119	19559	28.968	1.91e-07	***
material:temperature	4	9614	2403	3.560	0.01861	*
Residuals	27	18231	675			

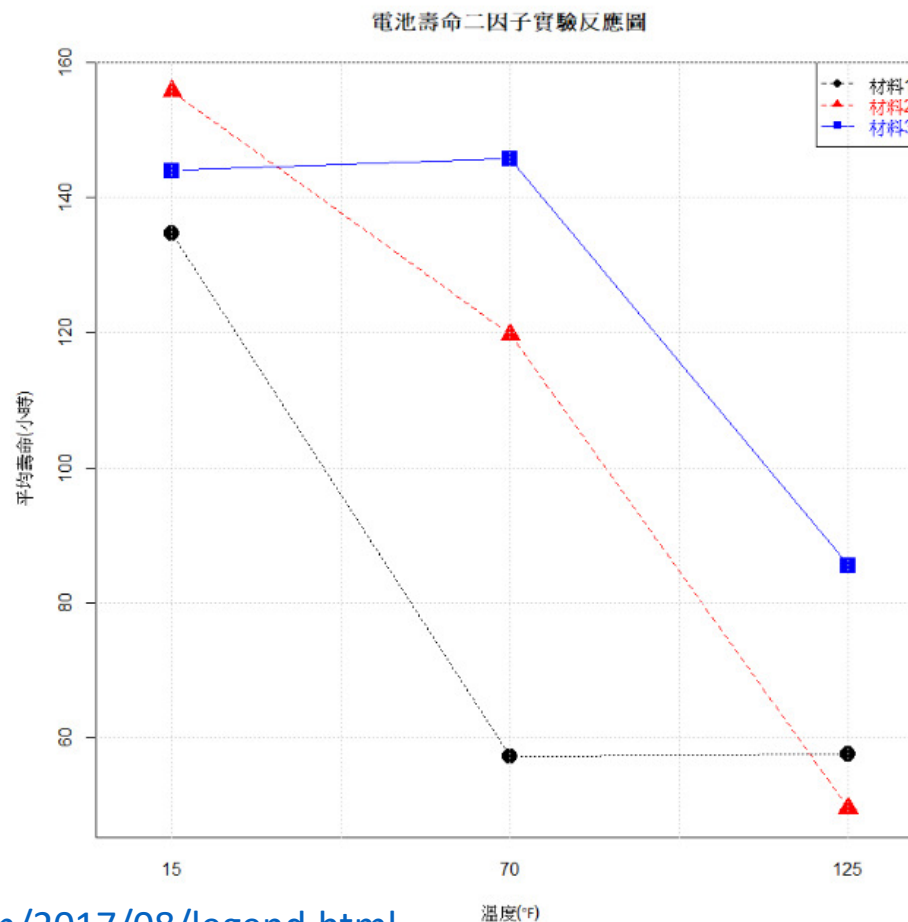
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

- 材料p值小,有顯著材料影響
- 溫度p值小,有顯著溫度作用
- 交互作用p值小,有顯著交互作用

變異來源	自由度
材料	$a - 1 = 3 - 1 = 2$
溫度	$b - 1 = 3 - 1 = 2$
交互作用	$(a - 1)(b - 1) = 2 \times 2 = 4$
誤差	$ab(n - 1) = 3 \times 3 \times (4 - 1) = 27$
總和	$N - 1 = 3 \times 3 \times 4 - 1 = 35$

反應圖

- 反應圖沒有明顯平行線，因此有交互作用。
- 無論材料為何，低溫會有較長壽命。
- 低溫(15°F)轉中溫(70°F)時，材料3壽命有增加，但材料1,2降低。
- 中溫(70 °F)轉高溫(150°F)時，材料1壽命無影響，但材料2,3會降低。



legend 使用參考: <http://rwepa.blogspot.com/2017/08/legend.html>

反應圖(續)

反應圖

反應圖

```
interaction.plot(x.factor      = batteryLife$temperature,  
                 trace.factor = batteryLife$material,  
                 response     = batteryLife$life,  
                 fun = mean,  
                 type = "b",  
                 col = c("black", "red", "blue"),  
                 pch = c(19, 17, 15),  
                 fixed = TRUE,  
                 legend = FALSE,  
                 cex=2,  
                 xlab = "溫度(°F)",  
                 ylab = "平均壽命(小時)",  
                 main = "電池壽命二因子實驗反應圖")
```

圖例

```
legend("topright",  
       legend = c("材料1", "材料2", "材料3"),  
       pch = c(19, 17, 15), lty = c(2, 2, 1),  
       col = c("black", "red", "blue"),  
       text.col = c("black", "red", "blue"))
```

網格線

```
grid()
```

6. 2^3 因子設計

(2^3 factorial design)

2^3 因子設計

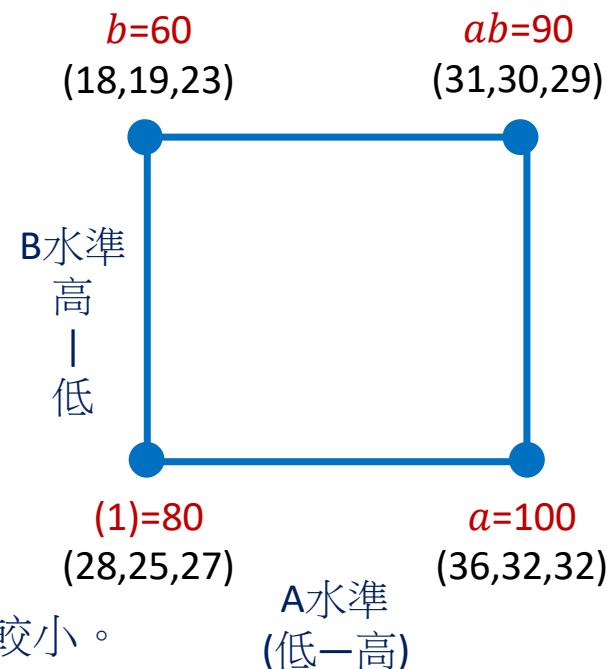
- 因子設計是研究多個因子對一個反應變數的常用方法。
- 考慮3個因子，且每個因子有2個水準，則稱為 2^3 因子設計。
- 同理， 2^k 因子設計亦常用於因子篩選實驗
- 常用因子設計方法
 - 2^2 因子設計 (2^2 factorial design)
 - 2^3 因子設計 (2^3 factorial design)
 - 2^k 因子設計 (2^k factorial design)
 - 2^k 加中心點因子設計

2²因子設計

- 2²因子設計延伸先上一章的因子實驗， n 為處理的實驗次數。
- 考慮良率實驗，每一個處理組合實驗3次，合計12($2^2 \times 3 = 4 \times 3$)次試驗。
 - (1)表示二個因子皆低水準
 - a 表示A高水準與B低水準
 - b 表示A低水準與B高水準
 - ab 表示二個因子皆高水準

對比(Contrast): $ab + a - b - (1)$

- $A = \frac{1}{2n}(ab + a - b - (1)) = \frac{1}{2 \times 3}(90 + 100 - 60 - 80) = 8.33$
- $B = \frac{1}{2n}(ab + b - a - (1)) = \frac{1}{2 \times 3}(90 + 60 - 100 - 80) = -5$
- $AB = \frac{1}{2n}(ab + (1) - a - b) = \frac{1}{2 \times 3}(90 + 80 - 100 - 60) = 1.67$
- A效果為正數，表示從低水準到高水準會增加良率。
- B效果為負數，表示從低水準到高水準會降低良率。AB效果影響較小。



2²因子設計(續)

- $SS_A = \frac{Contrast_A^2}{4n} = \frac{(ab+a-b-(1))^2}{4n} = \frac{50^2}{4 \times 3} = 208.33$
- $SS_B = \frac{(ab+b-a-(1))^2}{4n} = 75$
- $SS_{AB} = \frac{(ab+(1)-a-b)^2}{4n} = 8.33$
- $SS_T = \sum_{i=1}^2 \sum_{j=1}^2 \sum_{k=1}^3 y_{ijk}^2 - \frac{y_{...}^2}{4n}$
 $= (28^2 + 25^2 + \dots + 29^2) - \frac{(80 + 100 + 60 + 90)^2}{4 \times 33} = 323$
- $SS_E = SS_T - SS_A - SS_B - SS_{AB} = 31.34$

2²因子設計變異數分析表

變異來源	平方和	自由度	均方	F_0
A處理	$SS_A = 208.33$	$2 - 1 = 1$	$MS_A = \frac{SS_A}{a - 1} = 208.33$	$F_A = \frac{MS_A}{MS_E} = 53.15$
B處理	$SS_B = 75$	$2 - 1 = 1$	$MS_B = \frac{SS_B}{b - 1} = 75$	$F_B = \frac{MS_B}{MS_E} = 19.13$
交互作用	$SS_{AB} = 8.33$	$(2 - 1)(2 - 1) = 1$	$MS_{AB} = \frac{SS_{AB}}{(a - 1)(b - 1)} = 8.33$	$F_{AB} = \frac{MS_{AB}}{MS_E} = 2.13$
誤差	SS_E	$ab(n - 1) = 2 \times 2 \times (3 - 1) = 8$	$MS_E = \frac{SS_E}{ab(n - 1)} = 3.92$	
總和	$SS_T = 323$	$N - 1 = (2 \times 2 \times 3) - 1 = 11$		

- A, B處理均有統計上顯著效果
- AB因子間沒有顯著交互作用

```
> pf(q=53.15, df1 = 1, df2 = 8, lower.tail = FALSE)
[1] 8.466994e-05
> pf(q=19.13, df1 = 1, df2 = 8, lower.tail = FALSE)
[1] 0.002368432
> pf(q=2.13, df1 = 1, df2 = 8, lower.tail = FALSE)
[1] 0.182562
```

2²因子設計對比係數矩陣

- $Contrast_A = ab + a - b - (1)$
- $Contrast_B = ab + b - a - (1)$
- $Contrast_{AB} = ab + (1) - a - b$

2²因子設計對比係數矩陣 (contrast coefficients matrix)

效果	(1)	<i>a</i>	<i>b</i>	<i>ab</i>
A	-1	+1	-1	+1
B	-1	-1	+1	+1
AB	+1	-1	-1	+1

2²因子設計代數符號表

- 將2²因子設計對比係數矩陣轉置，最左側加上 I ，表示整個實驗的總和或平均。
- A, B, AB 之間任二個相乘為零，因此稱為直交設計(orthogonal design)→2^k
- 例: $A \cdot B = (-1, +1, -1, +1) \cdot (-1, -1, +1, +1) = 0$

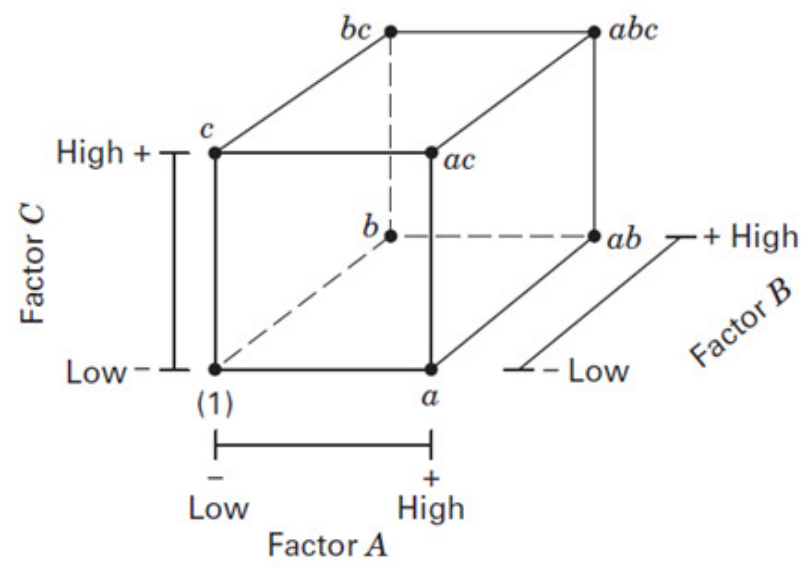
2²因子設計對比係數矩陣 (contrast coefficients matrix)

效果	(1)	a	b	ab
A	-1	+1	-1	+1
B	-1	-1	+1	+1
AB	+1	-1	-1	+1

2²因子設計代數符號表

處理組合	I	A	B	AB
(1)	+	-	-	+
a	+	+	-	-
b	+	-	+	-
ab	+	+	+	+

2³因子設計



2³幾何檢視圖

直交編碼表

試驗	A	B	C	字母標示	A	B	C
1	-	-	-	(1)	0	0	0
2	+	-	-	a	1	0	0
3	-	+	-	b	0	1	0
4	+	+	-	ab	1	1	0
5	-	-	+	c	0	0	1
6	+	-	+	ac	1	0	1
7	-	+	+	bc	0	1	1
8	+	+	+	abc	1	1	1

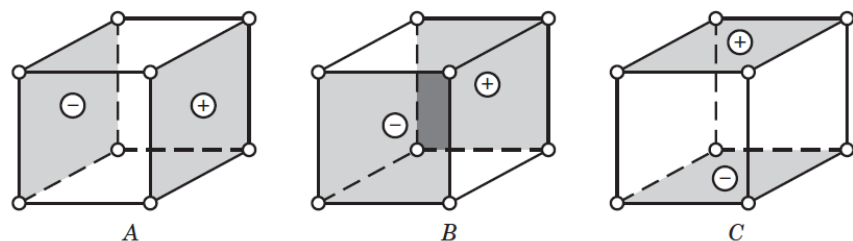


使用正負號

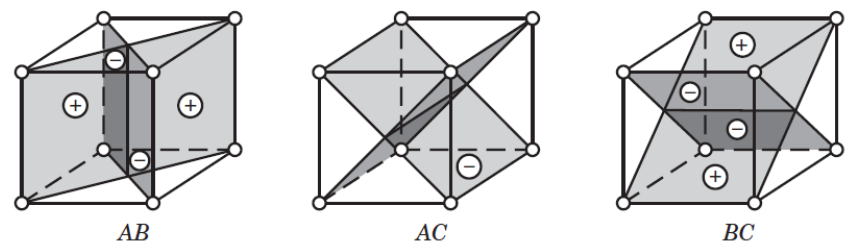


使用0,1表示

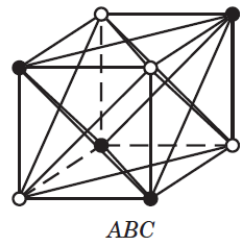
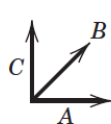
2³因子設計-對比圖



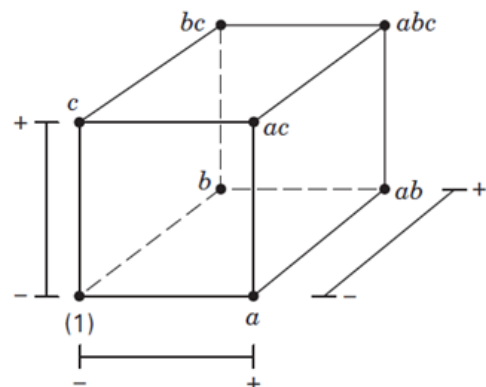
(a) Main effects



(b) Two-factor interaction



● = + runs
○ = - runs



直交編碼表	
字母標示	A
(1)	0
a	1
b	0
ab	1
c	0
ac	1
bc	0
abc	1

$$A \text{ 效果} = \bar{y}_{A^+} - \bar{y}_{A^-}$$

$$= \frac{a + ab + ac + abc}{4n} - \frac{(1) + b + c + bc}{4n}$$

結果與直交編碼表相同

7.反應曲面法

(Response Surface Methodology)

反應曲面法

- 反應曲面法為結合數學與統計而延生出的方法，其目標是最佳化反應變數(y)且會有多個解釋變數(x_1, x_2, \dots)。
- 1951年，Box 和 Wilson 共同進行數學模式的建立與推導，而後普遍應用於電子、機械、農業、化學工業、生物科技、材料科學、食品科學及工業製程改善等各項研究領域中。
- 考慮製程良率(y)是由溫度(x_1)和壓力(x_2)所影響，則模型為：
 - Max: $y = f(x_1, x_2) + \epsilon$
 - 考慮 $E(y) = f(x_1, x_2) = \eta$ 表示期望的反應，則 η 稱為反應曲面。

反應曲面法模型

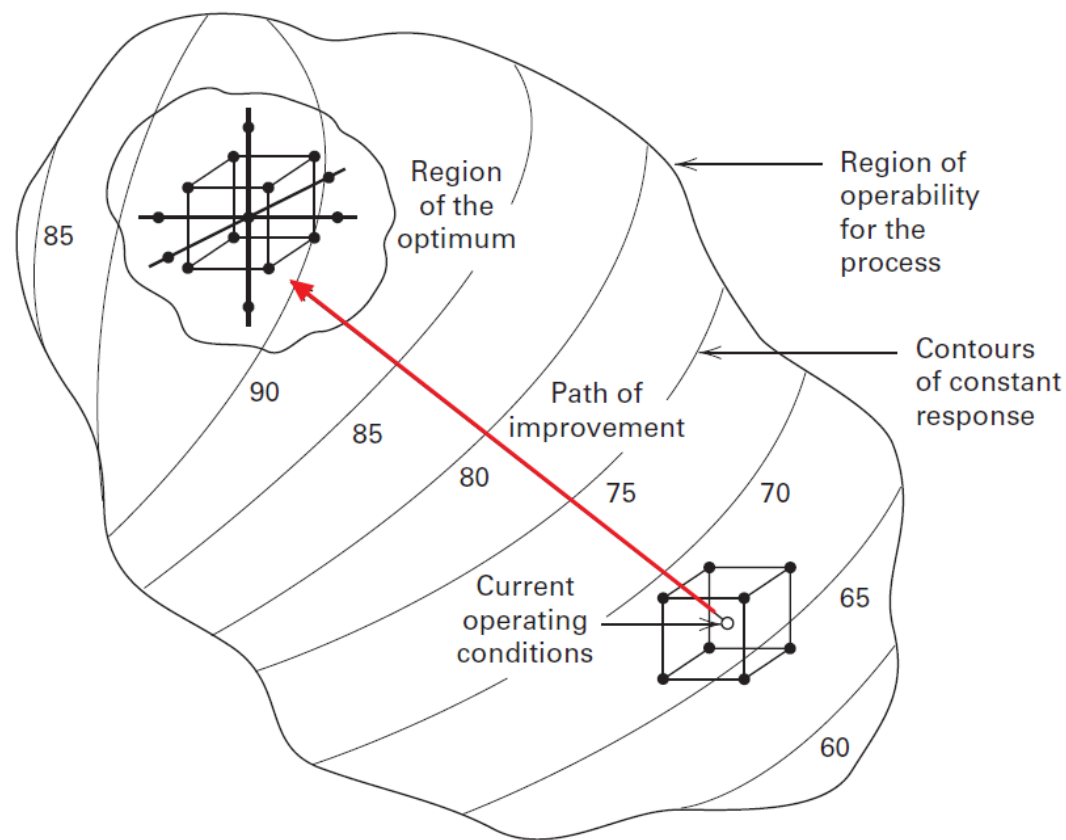
- 一階模型 (First-order model)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + \varepsilon$$

- 二階模型 (Second-order model)

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^k \sum_{j=1, i < j}^k \beta_{ij} x_i x_j + \varepsilon$$

逐次程序 (Sequential procedure)



R- rsm 套件


rsm: Response-Surface Analysis

Provides functions to generate response-surface designs, **fit first-** and **second-order response-surface models**, make surface plots, obtain the path of steepest ascent, and do canonical analysis. A good reference on these methods is Chapter 10 of Wu, C-F J and Hamada, M (2009) "Experiments: Planning, Analysis, and Parameter Design Optimization" ISBN 978-0-471-69946-0.

Version: 2.10.2
Imports: [estimability](#)
Suggests: [emmeans](#) (> 1.3.5.1), [Vdgraph](#), [conf.design](#), [DoE.base](#), [FrF2](#)
Published: 2020-09-04
Author: Russell Lenth [aut, cre]
Maintainer: Russell Lenth <russell-lenth@uiowa.edu>
License: [GPL-2](#) | [GPL-3](#) [expanded from: GPL (≥ 2)]
NeedsCompilation: no
Citation: [rsm citation info](#)
Materials: [NEWS](#)
In views: [ExperimentalDesign](#)
CRAN checks: [rsm results](#)

Downloads:

Reference manual: [rsm.pdf](#)
Vignettes: [Response-Surface Illustration](#)
[Surface Plots in the rsm Package](#)
[Response-Surface Methods in R, Using rsm](#)



步驟1.載入套件 rsm, 步驟2.載入資料集

```
> # 步驟1.載入套件
```

```
> library(rsm)
```

```
>
```

```
> # 步驟2.載入資料集
```

```
> ChemReact
```

	Time	Temp	Block	Yield
1	80.00	170.00	B1	80.5
2	80.00	180.00	B1	81.5
3	90.00	170.00	B1	82.0
4	90.00	180.00	B1	83.5
5	85.00	175.00	B1	83.9
6	85.00	175.00	B1	84.3
7	85.00	175.00	B1	84.0
8	85.00	175.00	B2	79.7
9	85.00	175.00	B2	79.8
10	85.00	175.00	B2	79.5
11	92.07	175.00	B2	78.4
12	77.93	175.00	B2	75.6
13	85.00	182.07	B2	78.5
14	85.00	167.93	B2	77.0

```
>
```

```
> ChemReact1 # Time: 85±5, Temp: 175±5
```

	Time	Temp	Yield
1	80	170	80.5
2	80	180	81.5
3	90	170	82.0
4	90	180	83.5
5	85	175	83.9
6	85	175	84.3
7	85	175	84.0

```
>
```

```
>
```

```
>
```

```
>
```

```
>
```

```
>
```

```
>
```

```
>
```

```
> ChemReact2
```

	Time	Temp	Yield
8	85.00	175.00	79.7
9	85.00	175.00	79.8
10	85.00	175.00	79.5
11	92.07	175.00	78.4
12	77.93	175.00	75.6
13	85.00	182.07	78.5
14	85.00	167.93	77.0

```
>
```

步驟3.編碼變數(coded variable)

```
> # 步驟3.編碼變數(coded variable)
> # (variable - center)/(half of width)
> # Time: 中心點=80, 寬度的一半=5
> # Temp: 中心點=175, 寬度的一半=5
>
> CR1 <- coded.data(ChemReact1, x1 ~ (Time - 85)/5, x2 ~ (Temp - 175)/5)
> CR1
```

	Time	Temp	Yield
1	80	170	80.5
2	80	180	81.5
3	90	170	82.0
4	90	180	83.5
5	85	175	83.9
6	85	175	84.3
7	85	175	84.0

Data are stored in coded form using these coding formulas ...

```
x1 ~ (Time - 85)/5
x2 ~ (Temp - 175)/5
```

編碼變數轉換為data.frame

設計矩陣

```
> # 編碼變數轉換為data.frame
> as.data.frame(CR1)
```

	x1	x2	Yield
1	-1	-1	80.5
2	-1	1	81.5
3	1	-1	82.0
4	1	1	83.5
5	0	0	83.9
6	0	0	84.3
7	0	0	84.0

```
>
```

步驟4.建立反應曲面設計 - ccd

```
> # 步驟4.建立反應曲面設計
> # ccd: central-composite designs(CCD, 中心組合設計)
> # 中心組合設計包括:(1)  $2^k$ 要因實驗: $2^k$ 個,主軸實驗: $2k$ 個,中心點實驗: $n_{\text{center}}$ 
> # bbd: Box-Behnken design (for 3 to 7 factors)
>
> # bbd(k, n0, coding)
> # k: 變數個數
> # n0: 每個區集的中心點個數
> # coding: 編碼公式
>
> # 建立5個變數A,B,C,D,E-CCD模型
> des1 <- ccd(y1 + y2 ~ A + B + C + D, generators = E ~ - A * B * C * D, n0 = c(6, 1))
```


建立3個變數 x_1, x_2, x_3 - bbd模型

```
> # 建立3個變數x1,x2,x3-BBD模型  
> des2 <- bbd(3, n0 = 2, coding = list(x1 ~ (Force - 20)/3, x2 ~ (Rate - 50)/10, x3 ~ Polish - 4))  
> des2
```

	run.order	std.order	Force	Rate	Polish
1	1	14	20	50	4
2	2	8	23	50	5
3	3	7	17	50	5
4	4	12	20	60	5
5	5	4	23	60	4
6	6	3	17	60	4
7	7	13	20	50	4
8	8	10	20	60	3
9	9	2	23	40	4
10	10	5	17	50	3
11	11	9	20	40	3
12	12	6	23	50	3
13	13	1	17	40	4
14	14	11	20	40	5

Data are stored in coded form using these coding formulas ...

```
x1 ~ (Force - 20)/3  
x2 ~ (Rate - 50)/10  
x3 ~ Polish - 4
```

```
>
```

步驟5.建立反應曲面模型

- 使用 `rsm()`

```
> # 步驟5.建立反應曲面模型
>
> # FO - first-order
> # TWI - two-way interaction
> # PQ - pure quadratic
> # SO - second-order
>
> CR1.rsm <- rsm(Yield ~ FO(x1, x2), data = CR1)
> summary(CR1.rsm) # Lack of fit: p值=0.01034 不顯著
```

Call:

```
rsm(formula = Yield ~ FO(x1, x2), data = CR1)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	82.81429	0.54719	151.3456	1.143e-08 ***
x1	0.87500	0.72386	1.2088	0.2933
x2	0.62500	0.72386	0.8634	0.4366

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Multiple R-squared: 0.3555, Adjusted R-squared: 0.0333
F-statistic: 1.103 on 2 and 4 DF, p-value: 0.4153

Analysis of Variance Table

Response: Yield

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
FO(x1, x2)	2	4.6250	2.3125	1.1033	0.41534
Residuals	4	8.3836	2.0959		
Lack of fit	2	8.2969	4.1485	95.7335	0.01034
Pure error	2	0.0867	0.0433		

Direction of steepest ascent (at radius 1):

x1	x2
0.8137335	0.5812382

Corresponding increment in original units:

Time	Temp
4.068667	2.906191

步驟6.優化反應曲面模型- 加入 two-way interactions

- 使用 `update()`

```
> CR1.rsmi <- update(CR1.rsm, . ~ . + TWI(x1, x2))
> summary(CR1.rsmi) # # Lack of fit: p值~0.005221, 已經減小.
```

Call:
rsm(formula = Yield ~ FO(x1, x2) + TWI(x1, x2), data = CR1)

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	82.81429	0.62948	131.5604	9.683e-07 ***
x1	0.87500	0.83272	1.0508	0.3705
x2	0.62500	0.83272	0.7506	0.5074
x1:x2	0.12500	0.83272	0.1501	0.8902

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Multiple R-squared: 0.3603, Adjusted R-squared: -0.2793
F-statistic: 0.5633 on 3 and 3 DF, p-value: 0.6755

Analysis of Variance Table

Response: Yield

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
FO(x1, x2)	2	4.6250	2.3125	0.8337	0.515302
TWI(x1, x2)	1	0.0625	0.0625	0.0225	0.890202
Residuals	3	8.3211	2.7737		
Lack of fit	1	8.2344	8.2344	190.0247	0.005221
Pure error	2	0.0867	0.0433		

參考資料

1. R 基礎篇 - 國立台北商業技術學院上課教材(238頁, 2011.7.4),
<http://rwepa.blogspot.com/2013/01/r-201174.html>
2. R入門資料分析與視覺化應用(中文,字幕),
<https://courses.mastertalks.tw/courses/R-teacher>
3. R商業預測與應用(中文,字幕),
<https://courses.mastertalks.tw/courses/R-2-teacher>
4. 圖例中顯示點線混合符號legend,
<http://rwepa.blogspot.com/2017/08/legend.html>
5. Douglas C. Montgomery, Design and Analysis of Experiments, 10th Edition, Wiley, 2020.
<https://www.amazon.com/Design-Analysis-Experiments-Douglas-Montgomery-ebook/dp/B07ZN7SRT5>

THANK YOU!

MING-CHANG LEE

Email: alan9956@gmail.com

URL: <http://rwepa.blogspot.tw/>

WEB: @RWPEA