

縱向資料分析 HW2

410978002 統計三 謝元皓

1.(a)

i. Write down the hypothesis.

H_0 : The rabbits don't have improvement in CV when treated by TH.

H_a : The rabbits have improvement in CV when treated by TH.

ii. Establish the ANOVA model to test the program effect

ANOVA 程序					
應變數: y1					
來源	DF	平方和	均方	F 值	Pr > F
模型	2	2854.998462	1427.499231	10.73	0.0002
誤差	36	4790.289231	133.063590		
已校正的總計	38	7645.287692			

R 平方	變異係數	根 MSE	y1 平均值
0.373432	16.26808	11.53532	70.90769

來源	DF	Anova SS	均方	F 值	Pr > F
time	2	2854.998462	1427.499231	10.73	0.0002

iii. Write down the result and make inference

在本題中，虛無假設為經過一段時間後，兔子的狀況沒有任何的改善。而因為本題的條件是各組間資料獨立，所以轉至資料後直接進行 anova 檢定即可。根據報表內容。首先，應變數 y1 為 CV(傳導速率)，模型自由度為 2 是因為共有三組資料分別為 baseline, time 1 and time 2。透過 f 檢定後得知，p 值=0.02 檢定結果為非常顯著，所以會 reject H_0 。因此，我們可以得出結論: The rabbits have improvement in CV when treated by TH.

1.(b)

i. Write down the hypothesis.

H_0 : The rabbits don't have improvement in CV when treated by TH.

H_a : The rabbits have improvement in CV when treated by TH.

ii. Establish the ANOVA model to test the program effect

來源	DF	平方和	均方	F 值	Pr > F
模型	14	7374.779487	526.769963	46.74	<.0001
誤差	24	270.508205	11.271175		
已校正的總計	38	7645.287692			

R 平方	變異係數	根 MSE	CV1 平均值
0.964618	4.734687	3.357257	70.90769

來源	DF	類型 I SS	均方	F 值	Pr > F
ID	12	4519.781026	376.648419	33.42	<.0001
time	2	2854.998462	1427.499231	126.65	<.0001

來源	DF	類型 III SS	均方	F 值	Pr > F
ID	12	4519.781026	376.648419	33.42	<.0001
time	2	2854.998462	1427.499231	126.65	<.0001

iii. Write down the result and make inference

在本題中，虛無假設為經過一段時間後，兔子的狀況沒有任何的改善。而因為本題的條件是各組間資料不獨立，所以在 sas 中必須用 proc glm 非 proc anova。根據報表內容。首先，應變數 y1 為 CV(傳導速率)，模型自由度為 2 是因為共有三組資料分別為 baseline, time 1 and time 2。透過 f 檢定後得知，p 值<0.0001 檢定結果為非常顯著，所以會 reject H_0 。因此，我們可以得出結論: The rabbits have improvement in CV when treated by TH.

2. (a)

i. Print out the table of expected mean squares.

應變數: CST CST						
來源	DF	類型 III SS	均方	F 值	Pr > F	
* TIMES	3	2771.767045	923.922348	145.37	<.0001	
TIMES*group	3	509.194107	169.731369	26.71	<.0001	
id(group)	86	3008.675740	34.984602	5.50	<.0001	
誤差: MS(誤差)	254	1614.288848	6.355468			
* 此檢定假設有一或多個其他固定效果為零。						
來源	DF	類型 III SS	均方	F 值	Pr > F	
* group	1	819.794403	819.794403	23.65	<.0001	
誤差	86.355	2993.336040	34.663098			
誤差: $0.9888 * MS(id(group)) + 0.0112 * MS(誤差)$						
* 此檢定假設有一或多個其他固定效果為零。						

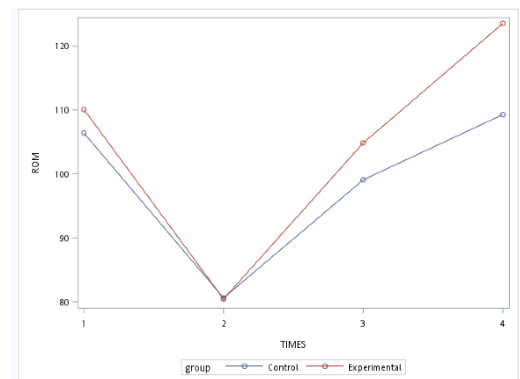
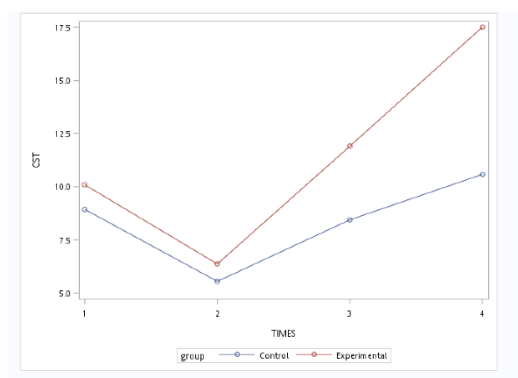
應變數: ROM ROM						
來源	DF	類型 III SS	均方	F 值	Pr > F	
* TIMES	3	60476	20159	193.82	<.0001	
TIMES*group	3	2604.608615	868.202872	8.35	<.0001	
id(group)	86	32666	379.840860	3.65	<.0001	
誤差: MS(誤差)	254	26418	104.008428			
* 此檢定假設有一或多個其他固定效果為零。						
來源	DF	類型 III SS	均方	F 值	Pr > F	
* group	1	3253.965116	3253.965116	8.64	0.0042	
誤差	86.535	32602	376.743276			
誤差: $0.9888 * MS(id(group)) + 0.0112 * MS(誤差)$						
* 此檢定假設有一或多個其他固定效果為零。						

ii. Explaing your finding.

根據上左表，Times 這項因子的 f 檢定小於 0.001，檢定結果為非常顯著。代表的結果是 cst 這項因子在每段時間下有非常顯著的差異；接著往下看 group 這項因子的 F 檢定依然<0.001，代表的結果是 cst 這項因子在不同的組別(實驗組，對照組)下也有顯著的差異；Times*group 這項為時間與組別間的交互作用項，檢定結果依然為非常顯著，代表存在著非常顯著的交互作用，透過下左圖可得知，cst 這項因子分成兩組(控制組，實驗組)時，到時間 4(第 12 周時)兩線明顯斜率不同，故存在強烈的交互作用。

根據上右表，Times 這項因子的 f 檢定小於 0.001，檢定結果為非常顯著。代表的結果是 rom 這項因子在每段時間下有非常顯著的

差異；接著往下看 **group** 這項因子的 F 檢定依然 <0.001 ，代表的結果是 **rom** 這項因子在不同的組別(實驗組，對照組)下也有顯著的差異；**Times*group** 這項為時間與組別間的交互作用項，檢定結果依然為非常顯著，代表存在著非常顯著的交互作用，透過下右圖可得知，**rom** 這項因子分成兩組(控制組，實驗組)時，兩線明顯斜率不同，故存在強烈的交互作用。



2. (b)

i. Use the polynomial transformation.

time_N 代表 time 的第 n 度多項式對比 說明轉換變數的 M 矩陣				
	y1	y2	y3	y4
time_1	-.6708203932	-.2236067977	0.2236067977	0.6708203932
time_2	0.5000000000	-.5000000000	-.5000000000	0.5000000000
time_3	-.2236067977	0.6708203932	-.6708203932	0.2236067977

Time 1 代表線性

time2 代表二次多項式

time3 代表三次多項式

ii. Make inference about the association between mean outcomes versus time and group.

對比變數: time_3

來源	DF	類型 III SS	均方	F 值	Pr > F
Mean	1	8691.57674	8691.57674	139.62	<.0001
group	1	56.79186	56.79186	0.91	0.3409
誤差	170	10582.83140	62.25195		

看這個報表時必須由 Time3 往上看，根據 time3 這個報表可知， h_0 為實驗組與控制組的三次多項式係數是否相等，即為 group 這項因子跟時間的趨勢是否需要到三次多項式；然而 group do not reject h_0 ，因此組別跟時間的趨勢並不需要用到三次多項式。因為 group 在 time 3 檢定結果為不顯著，所以需往下看 time 2。

對比變數: time_2

來源	DF	類型 III SS	均方	F 值	Pr > F
Mean	1	26675.37209	26675.37209	191.80	<.0001
group	1	692.00581	692.00581	4.98	0.0270
誤差	170	23643.62209	139.08013		

根據 time2 這個報表可知， h_0 為實驗組與控制組的二次多項式係數是否相等，即為 group 這項因子跟時間的趨勢是否需要到二次多項式；檢定結果為 reject h_0 ，因此組別跟時間的趨勢需要用到二次

多項式。因為 group 在 time 2 檢定結果為顯著，所以就不需往下看

time 1。結論：在 time2 的時候，實驗組的趨勢與對照組不同。

2. (C)

H = 類型 III SSCP 矩陣 - 針對 group				
	y1	y2	y3	y4
y1	194.70348837	7.4476744186	422.38953488	969.26162791
y2	7.4476744186	0.2848837209	16.156976744	37.075581395
y3	422.38953488	16.156976744	916.33139535	2102.7151163
y4	969.26162791	37.075581395	2102.7151163	4825.122093

無整體 group 效果假設的 MANOVA 檢定準則和精準 F 統計值 H = 類型 III SSCP 矩陣 - 針對 group E = 誤差 SSCP 矩陣 S=1 M=1 N=82.5					
統計值	值	F 值	分子自由度	分母自由度	Pr > F
Wilks 的 Lambda	0.79204353	10.96	4	167	<.0001
Pillai 追蹤	0.20795647	10.96	4	167	<.0001
Hotelling-Lawley 追蹤	0.26255686	10.96	4	167	<.0001
Roy 最大根	0.26255686	10.96	4	167	<.0001

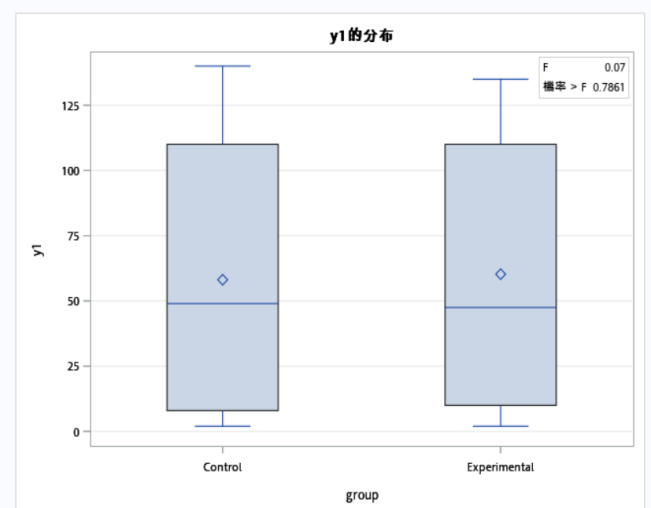
SAS 系統

GLM 程序

應變數: y1

來源	DF	平方和	均方	F 值	Pr > F
模型	1	194.7035	194.7035	0.07	0.7861
誤差	170	447851.1744	2634.4187		
已校正的總計	171	448045.8779			

R 平方	變異係數	根 MSE	y1 平均值
0.000435	86.69521	51.32659	59.20349



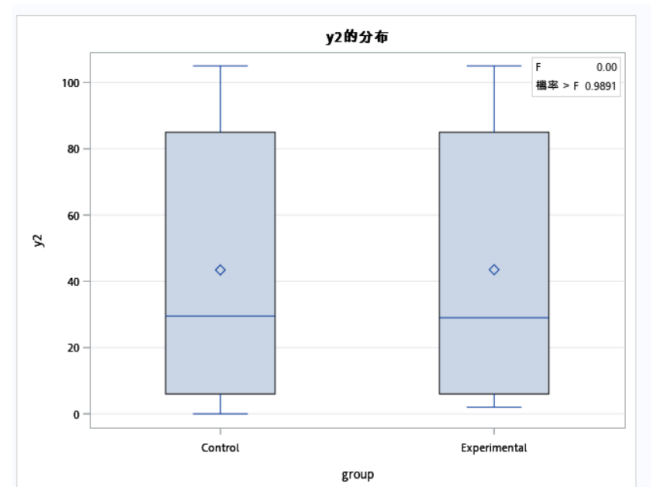
SAS 系統

GLM 程序

應變數: y2

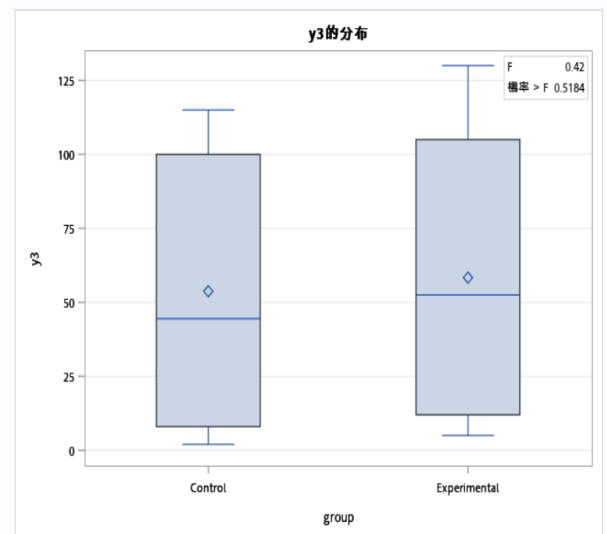
來源	DF	平方和	均方	F 值	Pr > F
模型	1	0.2849	0.2849	0.00	0.9891
誤差	170	259454.4302	1526.2025		
已校正的總計	171	259454.7151			

R 平方	變異係數	根 MSE	y2 平均值
0.000001	89.89247	39.06664	43.45930



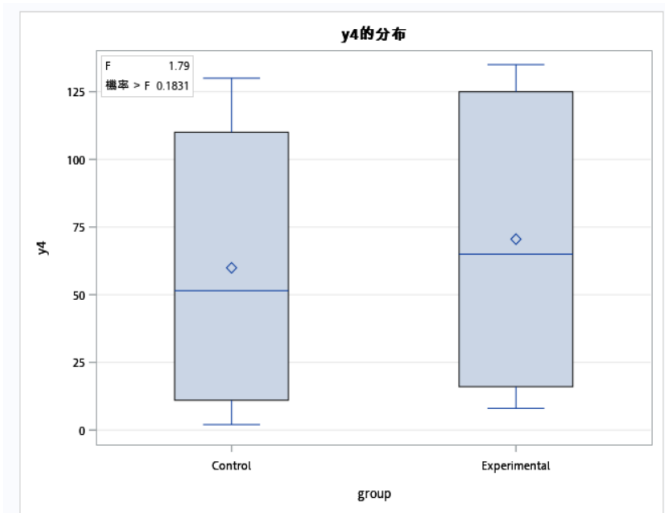
SAS 系統					
GLM 程序					
應變數: y3					
來源	DF	平方和	均方	F 值	Pr > F
模型	1	916.3314	916.3314	0.42	0.5184
誤差	170	371905.9651	2187.6821		
已校正的總計	171	372822.2965			

R 平方	變異係數	根 MSE	y3 平均值
0.002458	83.42734	46.77266	56.06395



SAS 系統					
GLM 程序					
應變數: y4					
來源	DF	平方和	均方	F 值	Pr > F
模型	1	4825.1221	4825.1221	1.79	0.1831
誤差	170	459123.0349	2700.7237		
已校正的總計	171	463948.1570			

R 平方	變異係數	根 MSE	y4 平均值
0.010400	79.67359	51.96849	65.22674



(2) 從最上表格可以發現，4 個 p 值皆極小，意即在這 4 個時間點的平均值，控制組和實驗組有顯著差異。接著 y1 時間(手術前)的 p 值為 0.7861 大於 $\alpha(0.05)$ ，意即控制組和實驗組無顯著差異。y2 時間(術後第 1 週)的 p 值為 0.9891 大於 $\alpha(0.05)$ ，意即控制組和實驗組無顯著差異。接著 y3 時間(術後第 6 週)的 p 值為 0.5184 大於 α

(0.05)，意即控制組和實驗組無顯著差異。

接著 y4 時間(術後第 12 週)的 p 值為 0.1831 大於 α

(0.05)，意即控制組和實驗組無顯著差異。雖然 4 個時間點控制組和實驗組均無顯著差異，但 p 值隨著時間逐漸減少，即控制組和實驗組差異逐漸明顯。