

107 農業試驗設計與統計分析研習會
(2017/10/23~24)

SAS-EG7.1 統計分析軟體 相關與回歸分析

主講: 呂椿棠

tang@tari.gov.tw

助教: 楊滿霞

ymh@tari.gov.tw

農業試驗所

◆ 試驗資料範例：

SAS-EG Trainging data - corr-reg.xls

內含練習用工作表：

multi-corr (4個變數資料相關分析)

reg-1&2 (回歸分析一次&二次)

multi-reg (複回歸分析)

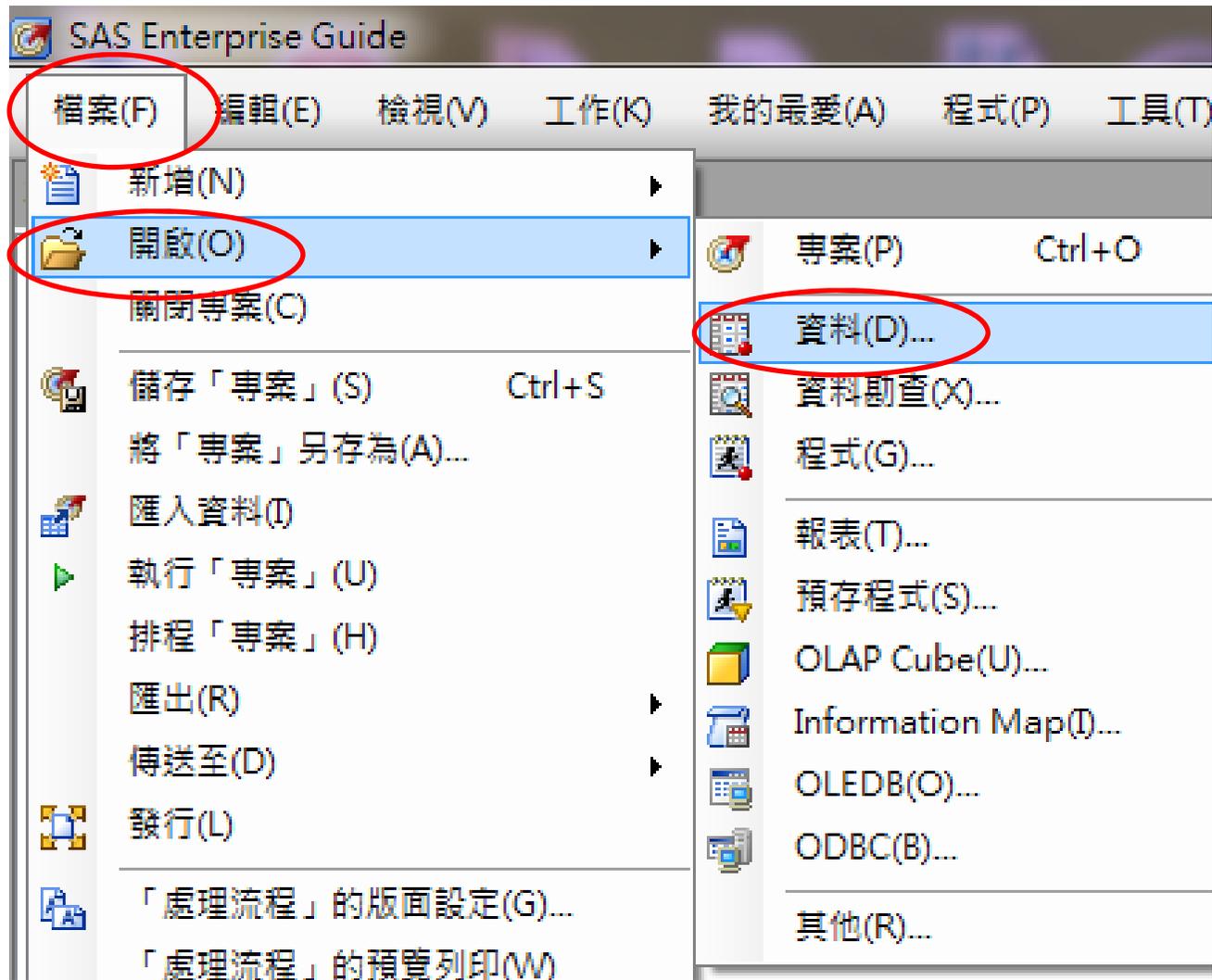
◆ SAS輸入格式：逐欄輸入處理標記、
各性狀變數數值

資料分析前的一般步驟

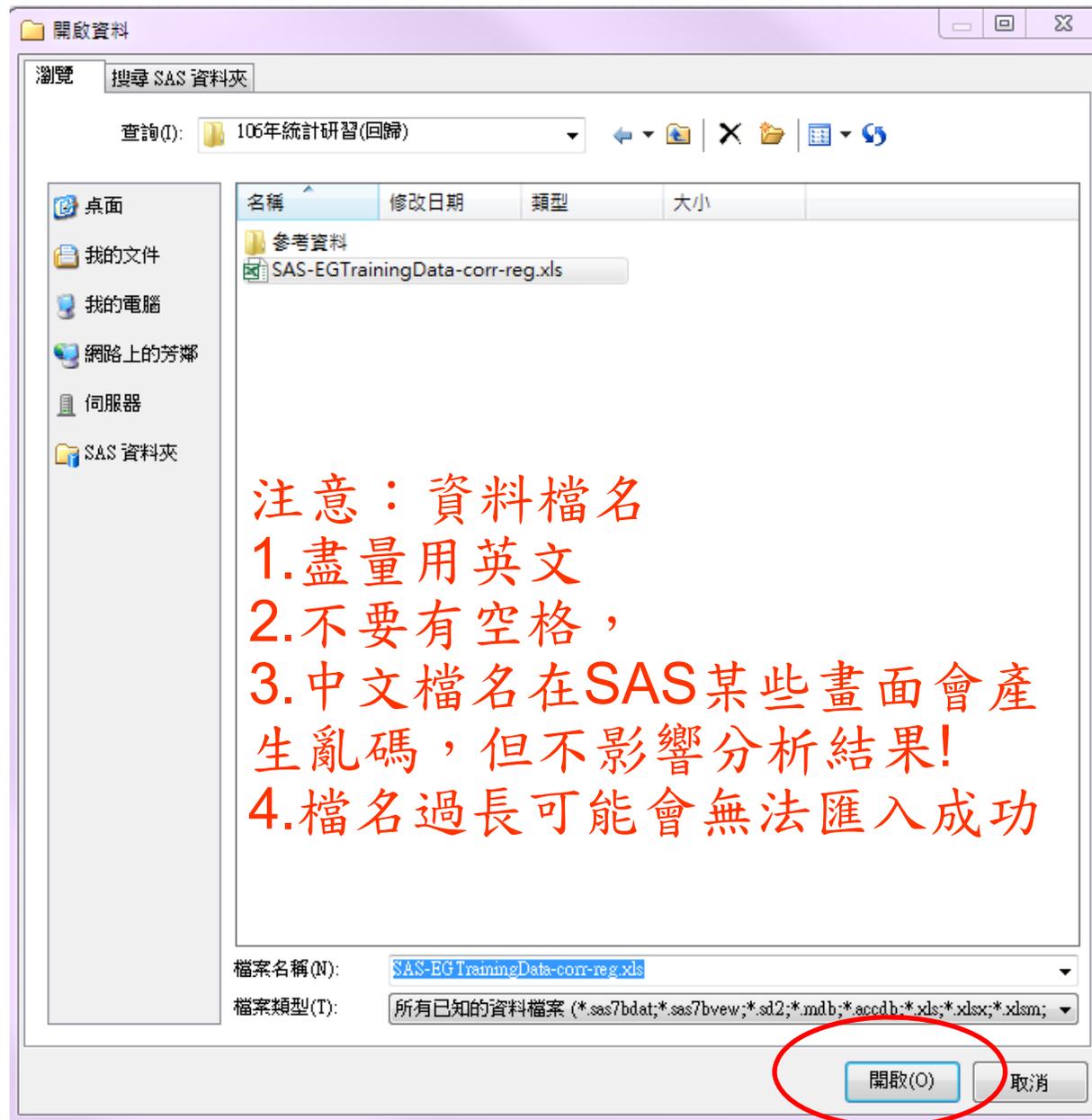
開啟SASEG7.1時的歡迎畫面



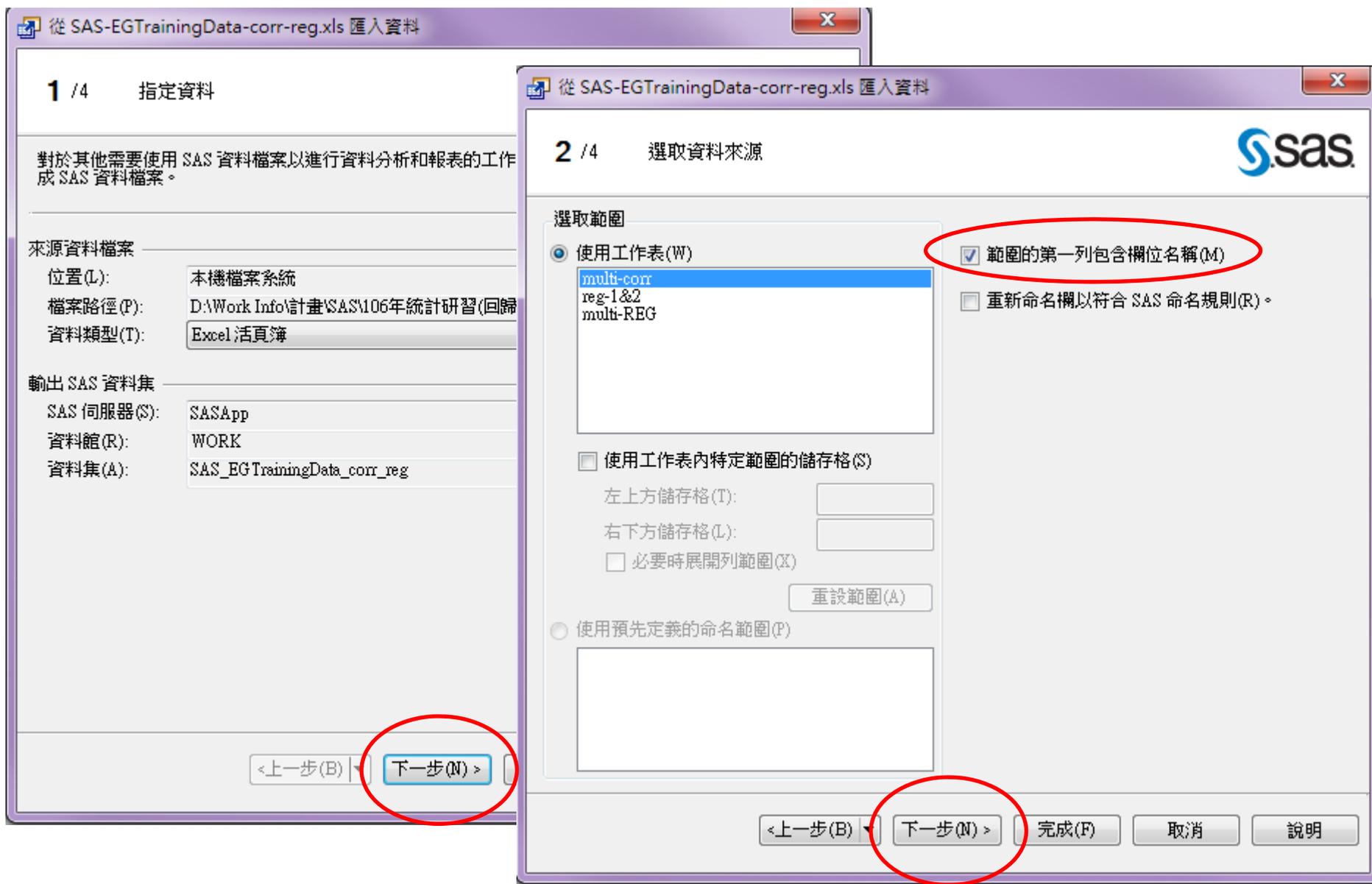
一、讀取資料檔-開啟已建的Excel檔： 檔案/開啟/資料



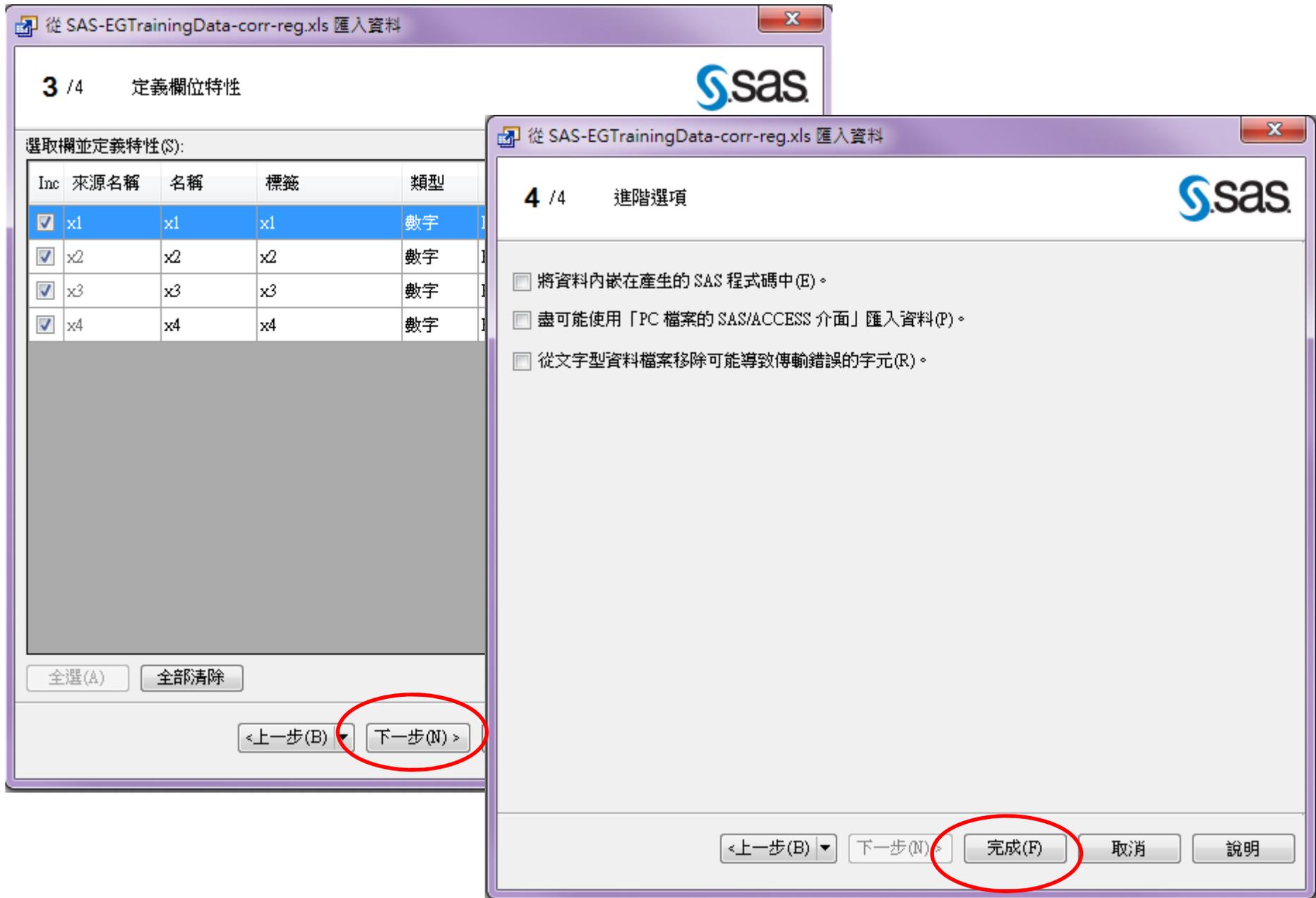
1.指定檔案所在位置



2. 勾選資料所在的工作表



3. 指定檔案開啟選項



4.指定工作表資料預覽

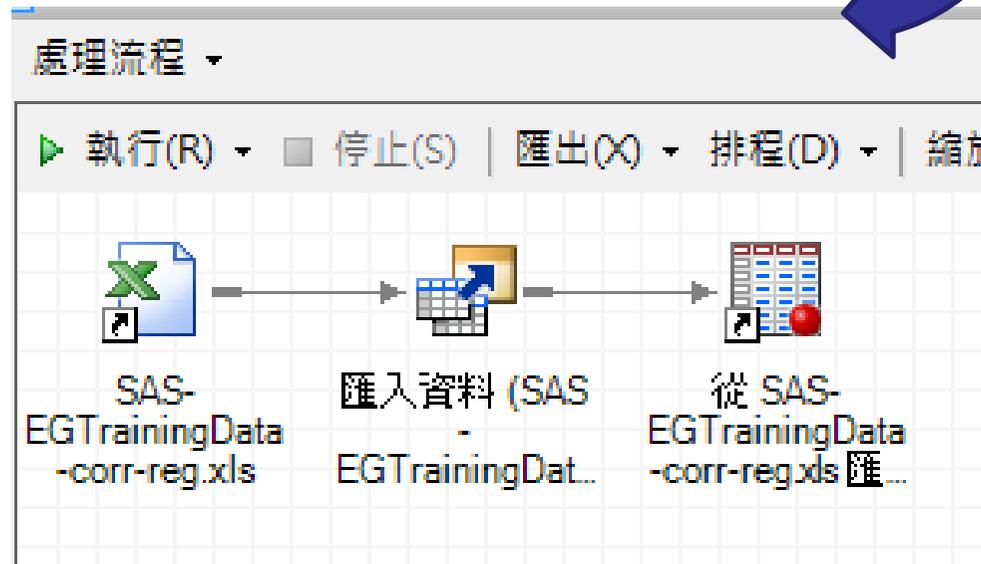
	x1	x2	x3	x4
1	12	30	4	54
2	14	34	6	60
3	20	30	5	70
4	25	35	6	80
5	30	40	8	100
6	35	50	10	120
7	40	60	12	134
8	50	76	15	150
9	55	80	16	145
10	60	90	18	152

結果如以下畫面：

左側小視窗



右側小視窗



二、散佈圖繪製

資料範例說明

【試驗內容】

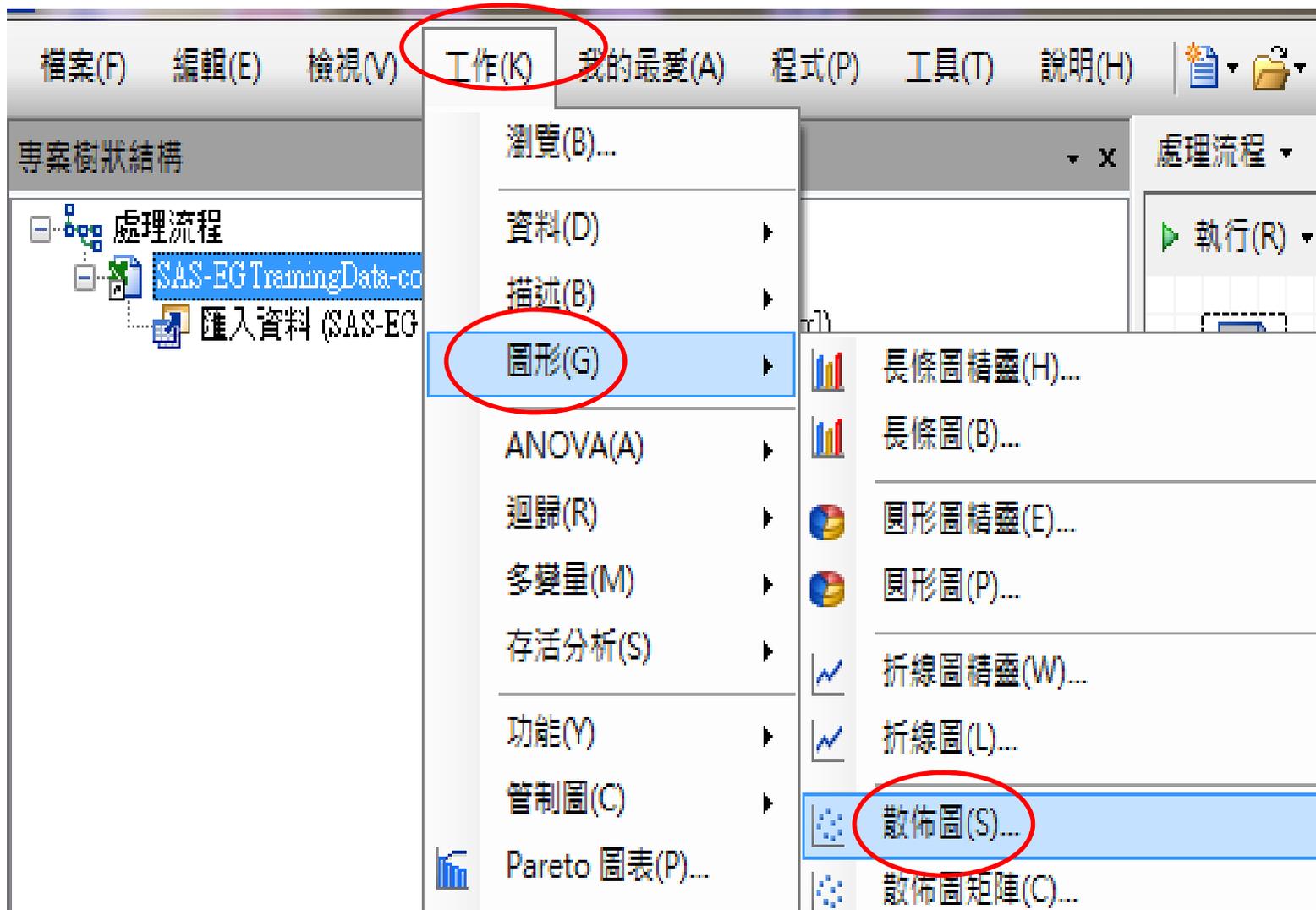
四個變數 X_1 、 X_2 、 X_3 、 X_4 的調查資料

【工作表】 multi-corr

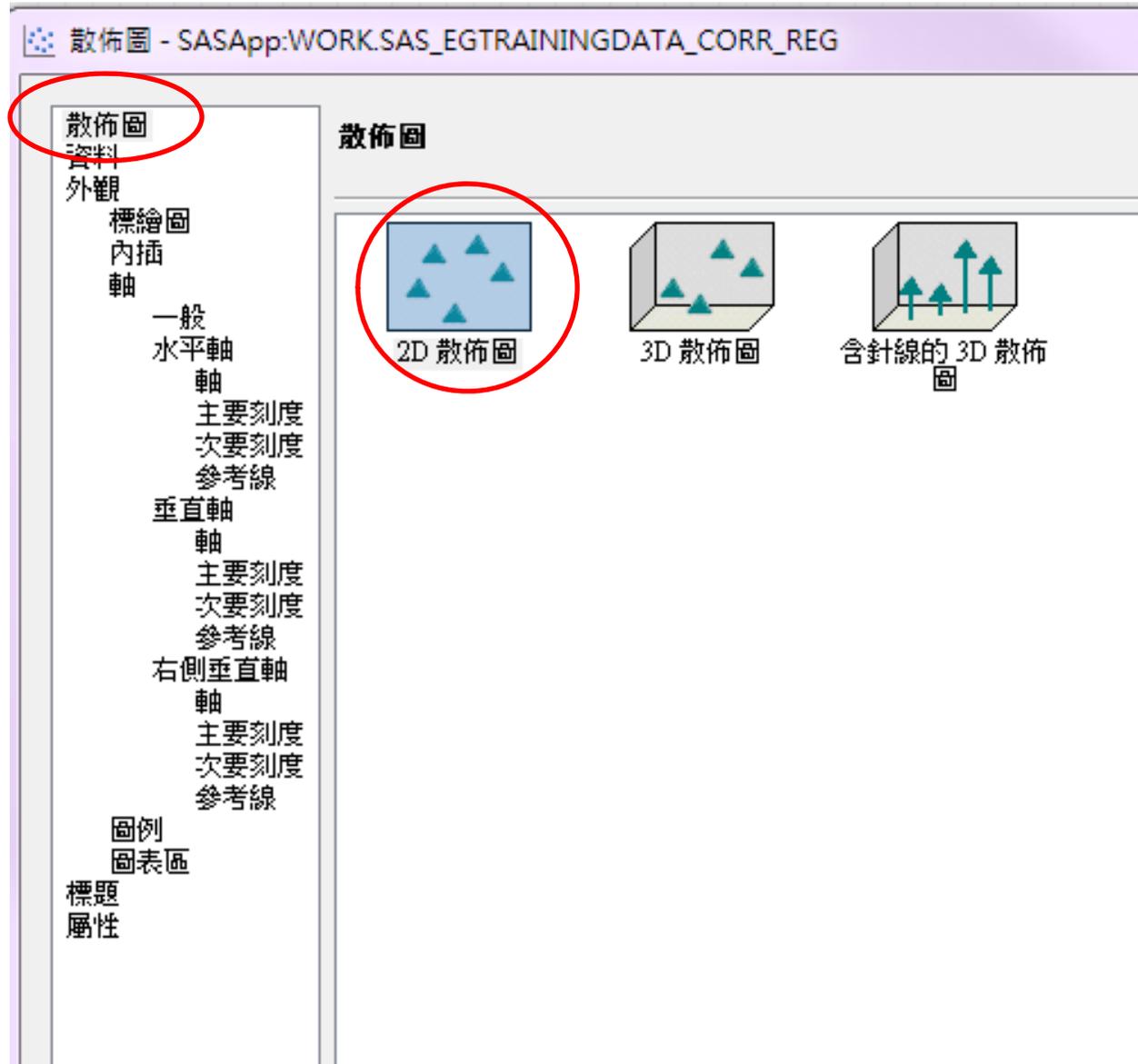
【分析內容】

將此4個變數進行兩兩之間的散佈圖繪製

1. 由散佈圖檢視兩變數間的關係



2. 選擇繪製散佈圖的類別



在「工作角色」畫面內指定變數的角色

(1) 散佈圖
資料
外觀
標繪圖
內插
軸
一般
水平軸
軸
主要刻度
次要刻度
參考線
垂直軸
軸
主要刻度
次要刻度
參考線
右側垂直軸
**

資料來源: SASApp:WORK.SAS_EGTRAININGDATA_CORR_REG
工作篩選: 無

要指派的欄(A): (2)

名稱
x1
x2
x3
x4

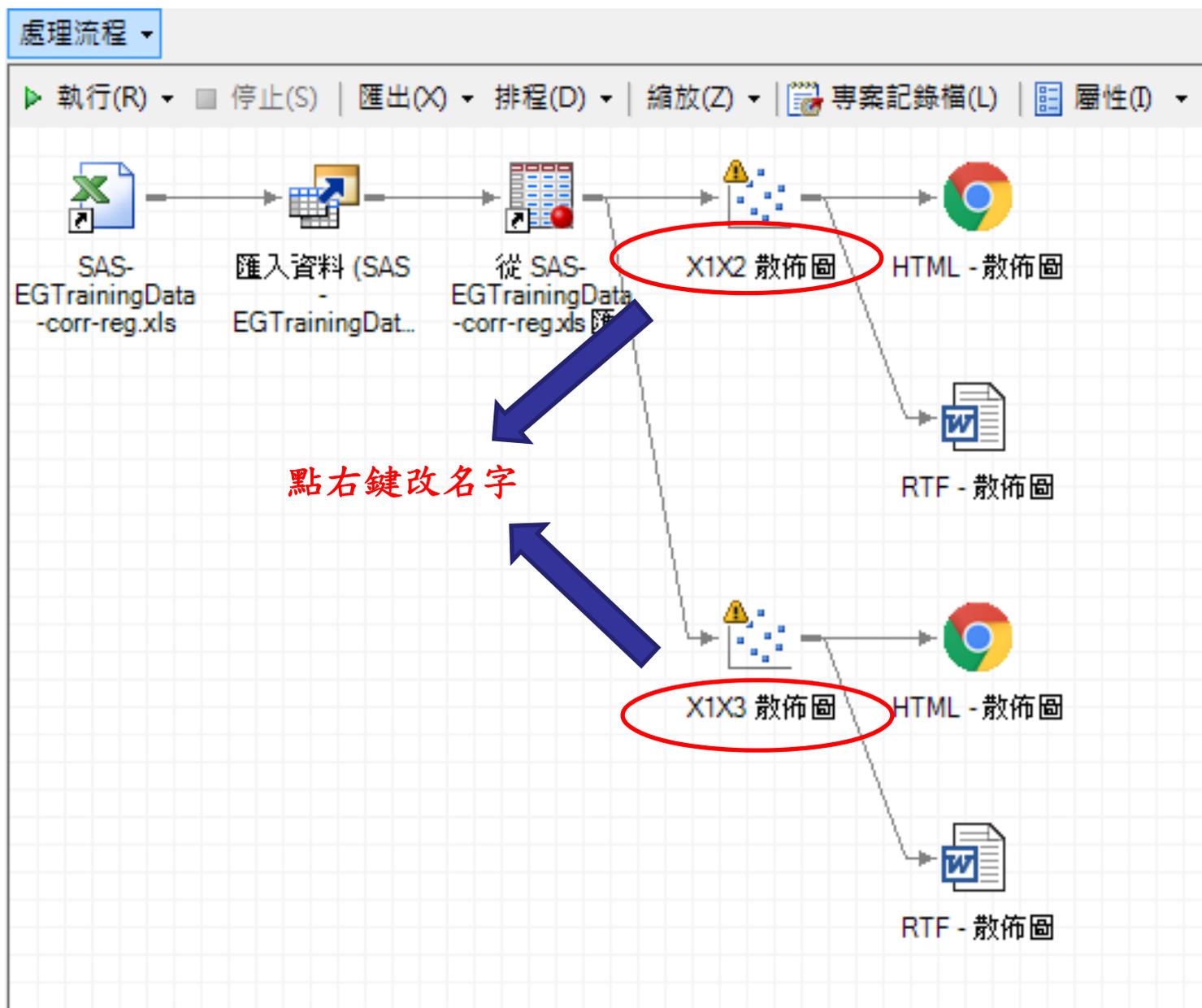
工作角色(T):

- 水平 (限制: 1)
- 垂直 (限制: 1)
- 垂直 (靠右) (限制: 1)
- 圖表群組依據

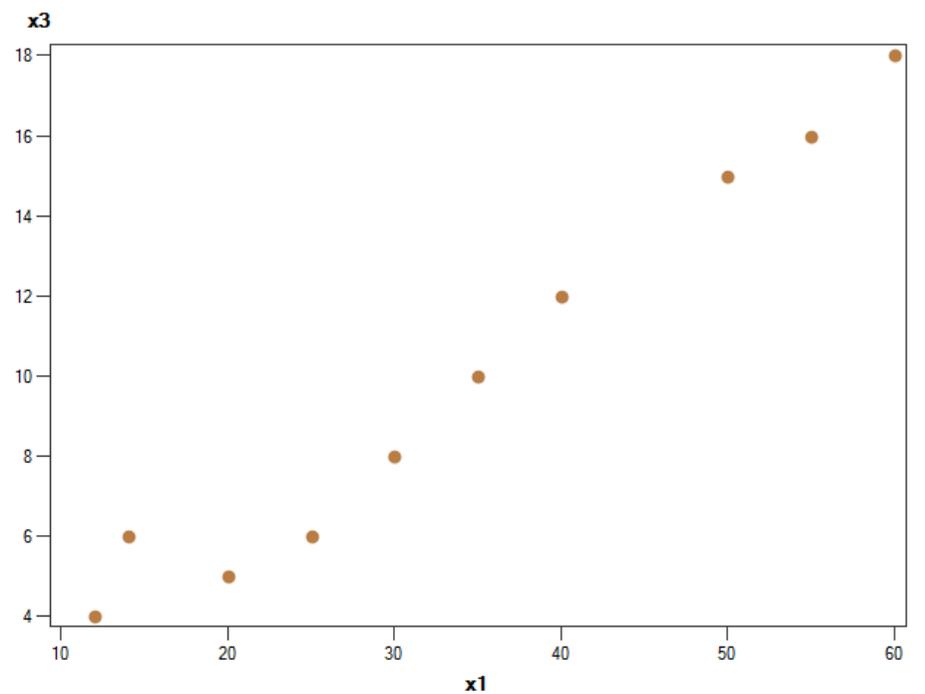
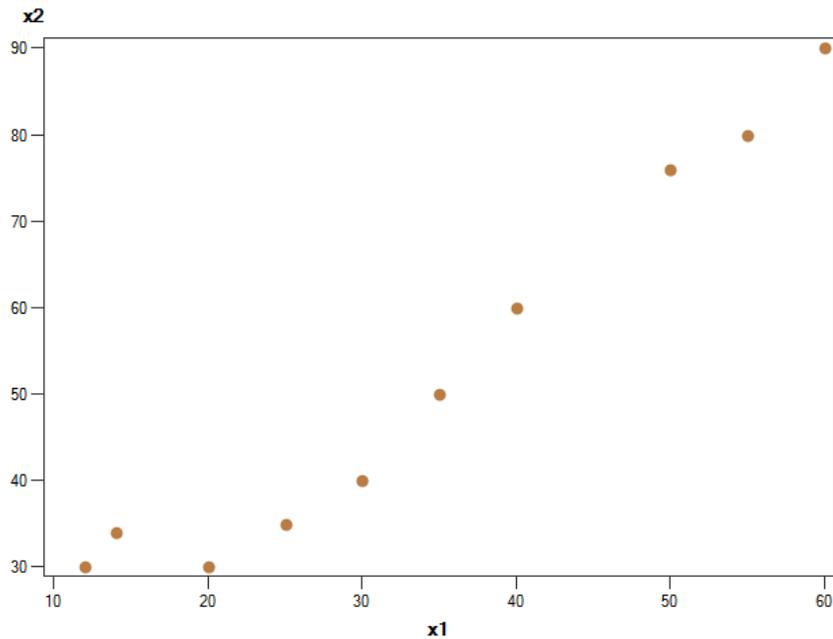
Red circles highlight 'x1' and 'x2' in the 'Variables to Assign' list. Red arrows point from 'x1' to '水平 (限制: 1)' and from 'x2' to '垂直 (限制: 1)' in the 'Work Roles' list.

注意：只能一次指定一個變數給水平或垂直的工作角色，目前有4個變數，要兩兩互相比較，則要重複6次的散佈圖繪製工作。

3. 產生散佈圖的處理流程圖



由x1與x2、x3的散佈圖發現x2、x3的值，有隨x1的增加而上升的趨勢



三、簡單相關分析

資料範例說明

【試驗內容】

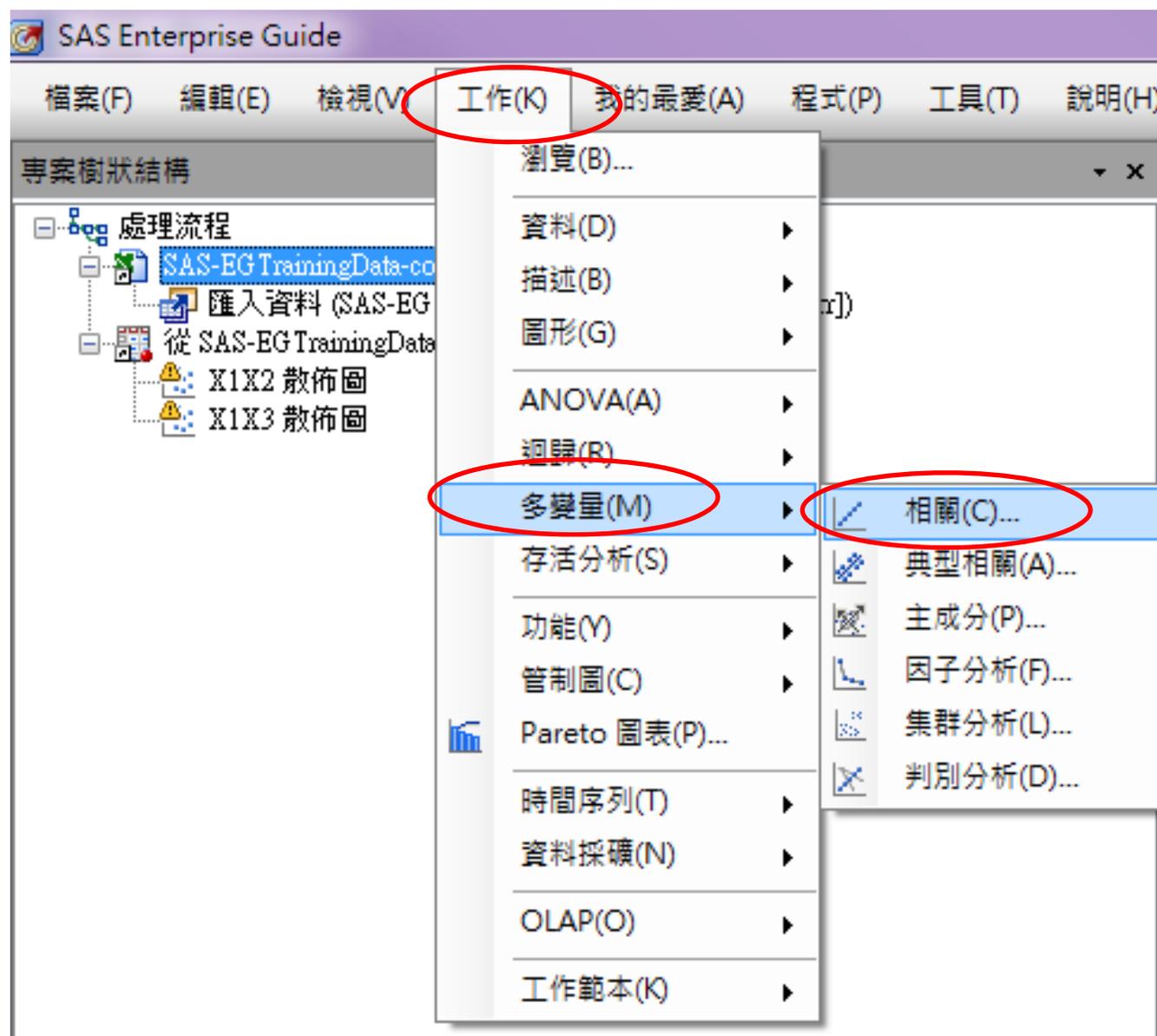
四個變數X1、X2、X3、X4的調查資料

【工作表】 multi-corr

【分析內容】

將此4個變數進行相關分析

1. 簡單相關分析: 工作 / 多變量 / 相關



在「資料」畫面內直接用滑鼠拖曳或按箭頭按鈕將要指派的變數拉到右側工作角色框內的“分析變數”(同時指派多個分析變數)

資料
選項
結果
輸出資料
標題
屬性

資料

資料來源: SASApp:WORK.SAS_EGTRAININGDATA_CORR_REG
工作篩選: 無

要指派的變數(A):

名稱
x1
x2
x3
x4

(1)

工作角色(T):

- 分析變數
 - x1
 - x2
 - x3
 - x4
- 相關對象
- 分析群組依據
- 次數計數 (限制: 1)
- 偏變數
- 相對加權 (限制: 1)

- 資料
- 選項**
- 結果
- 輸出資料
- 標題
- 屬性

選項 (2)

相關類型

- Pearson(P)
- Hoeffding(H)
- Kendall(K)
- Spearman(E)

(2)在「選項」選單挑選相關類型中的"Pearson(P)"
(3)在「結果」選單挑選下圖打勾的項目
(4)按"執行"

- 資料
- 選項
- 結果**
- 輸出資料
- 標題
- 屬性

結果 (3)

標繪圖

- 建立每個相關配對的散佈圖(T)

要顯示的結果(D)

顯示每個變數的統計值(H)

顯示與相關有關聯的顯著機率(W)

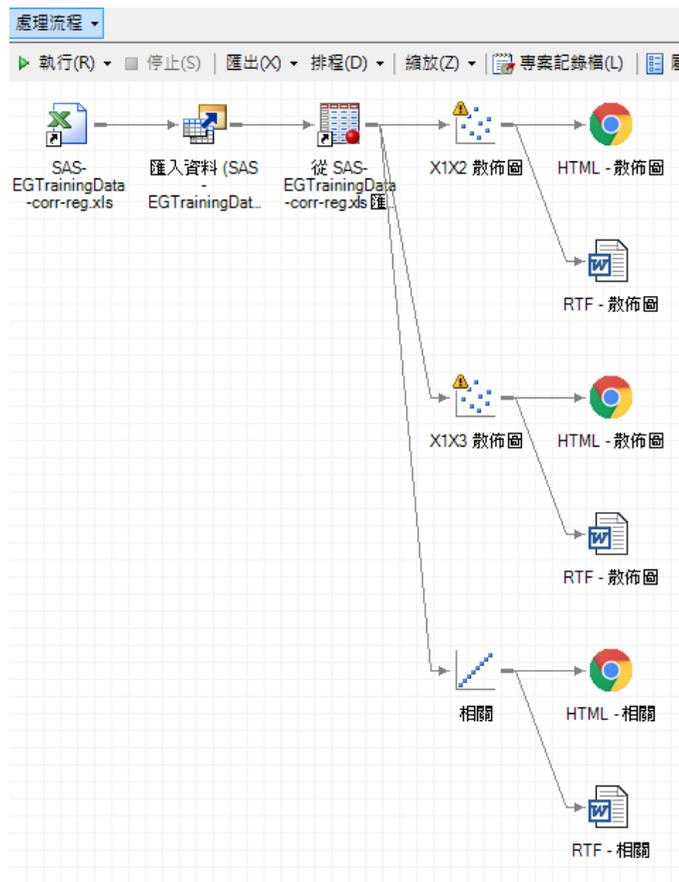
以遞減的強度順序顯示相關(O)

顯示每列

(4)

2. 分析結果解讀

完成相關分析後的處理流程圖



4個變數相關分析後的相關矩陣圖，包括倆倆變數的相關係數與其相關係數機率值。

Pearson 相關係數, N = 10
Prob > |r| (位於 H0 底下): Rho=0

	x1	x2	x3	x4
x1	1.00000	0.97417 <.0001	0.98350 <.0001	0.97341 <.0001
x2	0.97417 <.0001	1.00000	0.99510 <.0001	0.93972 <.0001
x3	0.98350 <.0001	0.99510 <.0001	1.00000	0.96041 <.0001
x4	0.97341 <.0001	0.93972 <.0001	0.96041 <.0001	1.00000

3. 相關分析結果-4個變數的基本統計量

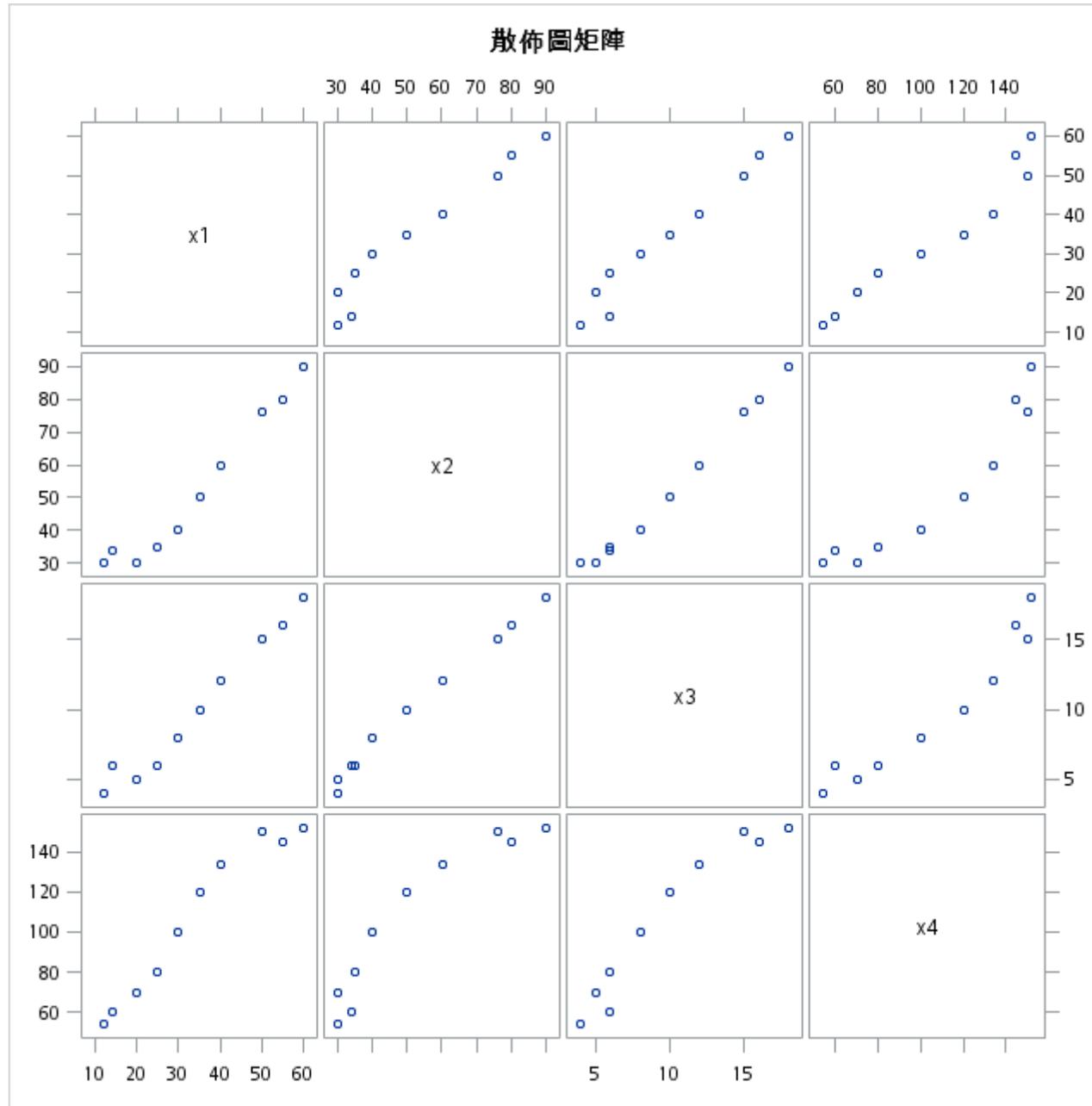
相關分析

CORR 程序

4 變數: x1 x2 x3 x4

簡單統計值						
變數	N	平均值	標準差	總和	最小值	最大值
x1	10	34.10000	16.95386	341.00000	12.00000	60.00000
x2	10	52.50000	22.59425	525.00000	30.00000	90.00000
x3	10	10.00000	5.01110	100.00000	4.00000	18.00000
x4	10	106.50000	38.55516	1065	54.00000	152.00000

每個相關配對的散佈圖



四、簡單直線回歸分析

資料範例說明

【試驗內容】

欲探討水梨果園的六個氮肥施用量 $x(\text{kg})$ 和可溶性固形物含量 $y(\%)$ 間之關係。

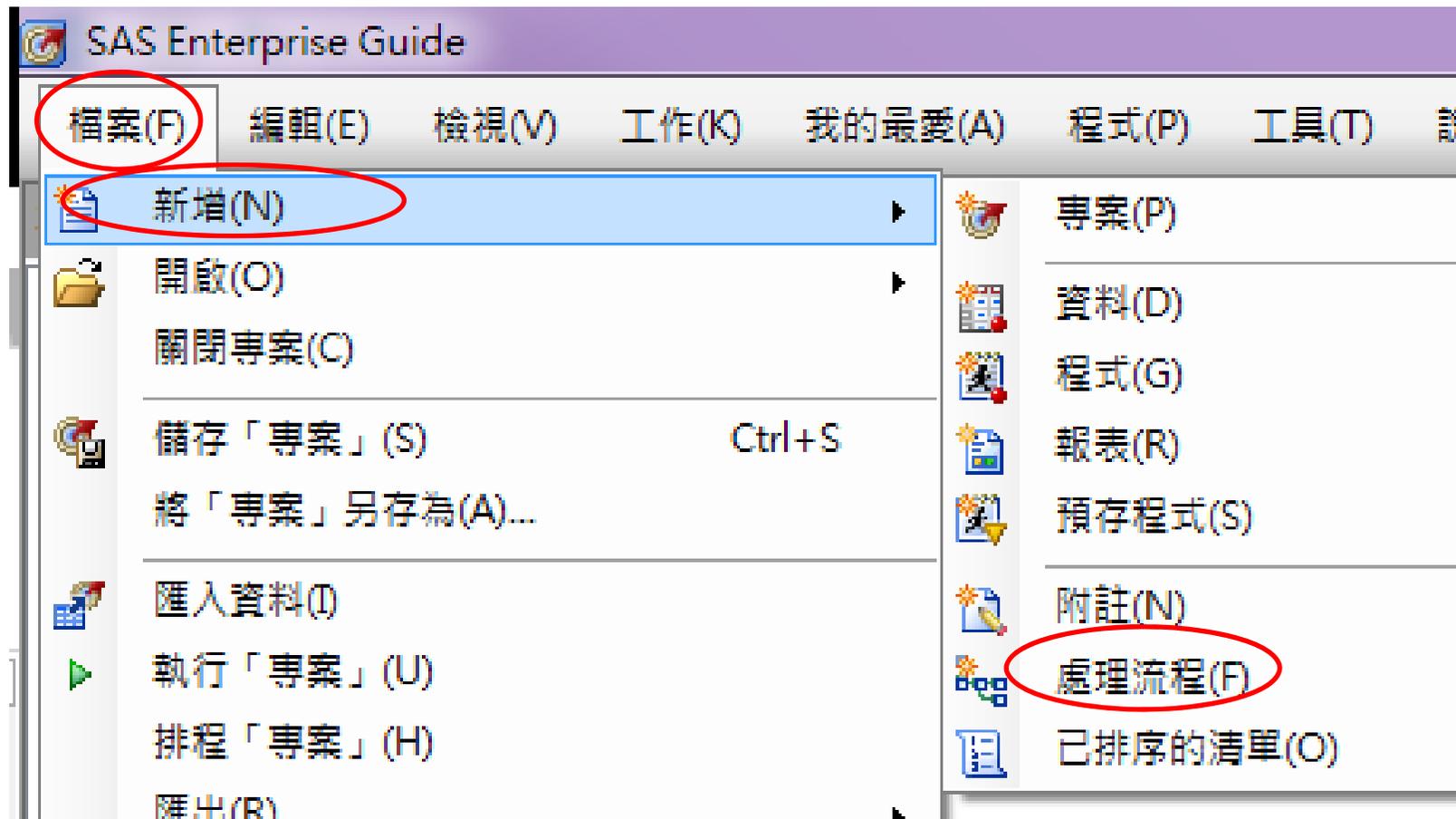
【工作表】 reg-1&2

【分析內容】

簡單直線回歸

$$Y_i = \alpha + \beta X_i + \varepsilon_i, i = 1, \dots, n$$

新增「處理流程」(如下圖)



載入資料檔，記得先點選載入的Excel資料工作表reg-1&2。(如投影片第4-7頁)

reg-1&2 工作表預覽

	固形物含量(%)	施氮量(kg)	施氮量-二次
1	16.2	0	0
2	15.7	5	25
3	15.4	10	100
4	15.2	15	225
5	14.3	20	400
6	13.9	25	625



1. 常態分布檢測：工作/描述/分配分析



(1)在「資料」畫面內將調查性狀指派到”分析變數” (可同時指派多個分析變數)

資料
分配
摘要
常態
對數常態
指數
Weibull
Beta
Gamma
核
標繪圖
外觀
插頁
表格
標題
屬性

資料

資料來源: SASApp:WORK.SAS_EGTRAININGDATA_CORR
工作篩選: 無

要指派的變數(A):

名稱
固形物含量(%)
施氮量(kg)
施氮量-二次

工作角色(T):

- 分析變數
- 固形物含量(%)
- 分析群組依據
- 次數計數 (限制: 1)
- 相對加權 (限制: 1)
- 分類變數 (限制: 2)

注意：資料的分布測驗是看整套資料，因此無須指定分類變數!!

(2) 在「常態」畫面內勾選常態

資料
分配

- 摘要
- 常態**
- 對數常態
- 指數
- Weibull
- Beta
- Gamma
- 核

標繪圖

- 外觀
- 插頁

表格

- 標題
- 屬性

分配 > 常態

常態 (N)

套用分配到所有變數 (A)

分析變數 (V):

平均值 (Mu) (M)

標準差 (Sigma) (G)

(3) 在「標繪圖-外觀」畫面內勾選直方圖和機率圖

資料分配

- 摘要
- 常態
- 對數常態
- 指數
- Weibull
- Beta
- Gamma
- 核
- 標繪圖**
- 外觀
- 插頁
- 表格
- 標題
- 屬性

標繪圖 > 外觀

附註: 插頁僅適用於直方圖、機率圖和分位數-分位數圖。

	軸顏色:	背景色彩:	軸寬度:
<input checked="" type="checkbox"/> 直方圖(H)			1
<input checked="" type="checkbox"/> 機率圖(P)			1
<input type="checkbox"/> 分位數圖(Q)			1
<input type="checkbox"/> 盒形圖(B)			1
<input type="checkbox"/> 文字為基礎的標繪圖(L)	產生莖葉圖或長條圖 (視觀測值的數目而定)、盒形圖及常態機率圖。產生方塊圖 (如果有依據變數的話)。		

建立連續理論分配的機率圖，並疊置指定或估計位置參數的對應參考線，及理論分配的縮放參數。

預覽程式碼(C) **執行(R)** 儲存(S) 取消 說明

常態性測驗之結果解讀-配適度測驗 (原始資料)

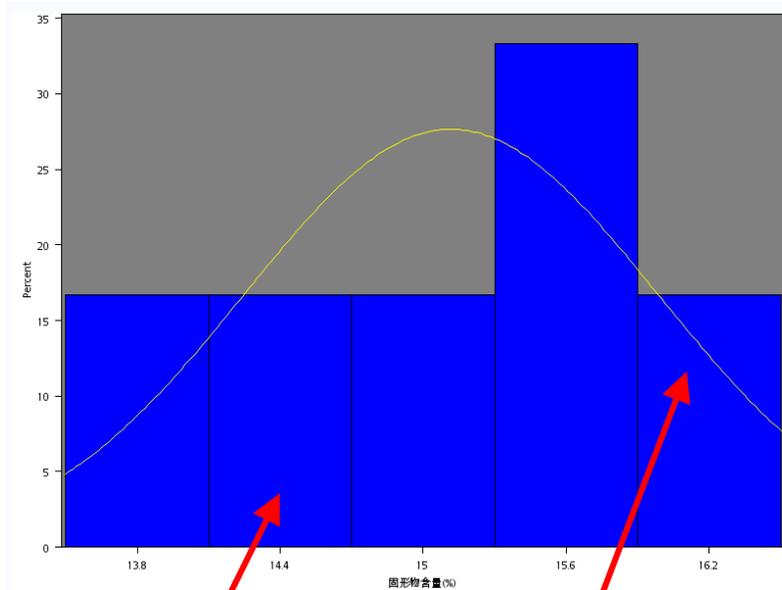
Parameters for Normal Distribution		
參數	符號	估計值
Mean	Mu	15.11667
Std Dev	Sigma	0.865833

判斷準則：
p value > 0.05
即不顯著，
表示符合常態

Goodness-of-Fit Tests for Normal Distribution				
檢定	統計值		p 值	
Kolmogorov-Smirnov	D	0.20500424	Pr > D	>0.150
Cramer-von Mises	W-Sq	0.03717239	Pr > W-Sq	>0.250
Anderson-Darling	A-Sq	0.22478918	Pr > A-Sq	>0.250

常態性測驗之結果解讀-圖示 (原始資料)

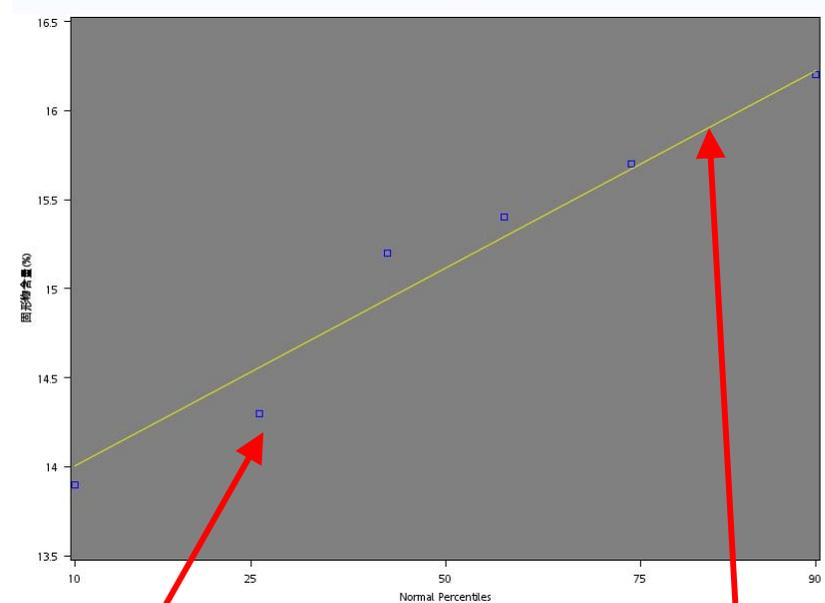
直方圖



實際觀測頻度
(長條)

常態理論頻度
(平滑鐘形曲線)

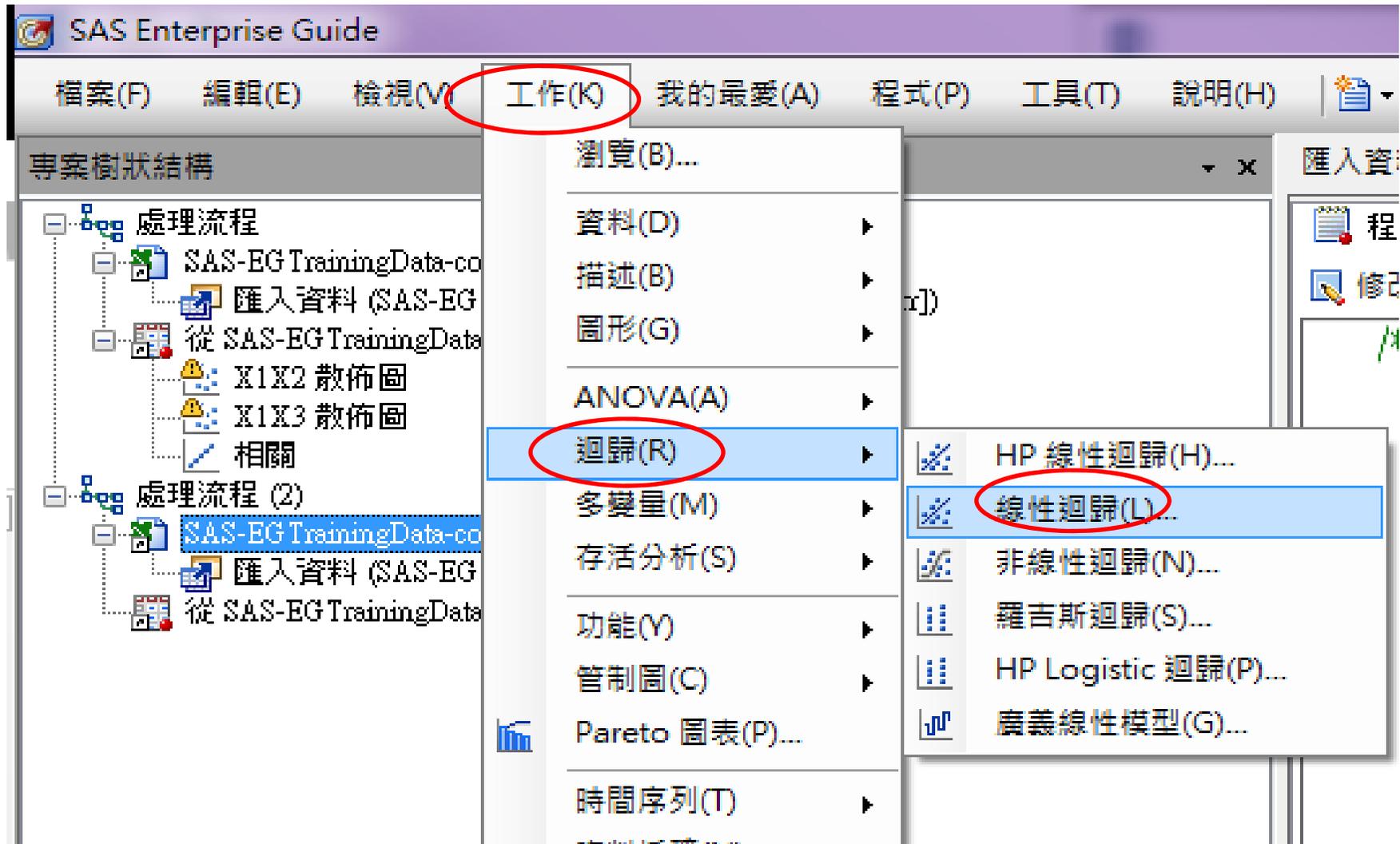
機率圖



實際觀測機率值
(點)

理論常態機率值
(對角直線)

2. 簡單直線回歸分析: 工作/回歸/線性回歸



在「工作角色」內用滑鼠拖曳(或按箭頭按鈕)將固形物含量和施氮肥分別指派到右側工作角色框內的"應變數"和"解釋變數"

資料 (1)

資料來源: SASApp:WORK.SAS_EGTRAININGDATA_CORR_REG_0005
工作篩選: 無

要指派的變數(A):

名稱
<input checked="" type="checkbox"/> 固形物含量(%)
<input checked="" type="checkbox"/> 施氮量(kg)
<input type="checkbox"/> 施氮量-二次

工作角色(T):

- 應變數 (限制: 1)
- 固形物含量(%)
- 解釋變數
- 施氮量(kg)
- 分析群組依據
- 次數計數 (限制: 1)
- 相對加權 (限制: 1)

模型 (2)

模型選取法(M): 符合的完整模型 (沒有選取)

顯著水準

進入模型(N): 0.5

保留在模型中(T): 0.1

統計值 (3)

估計值的詳細資料

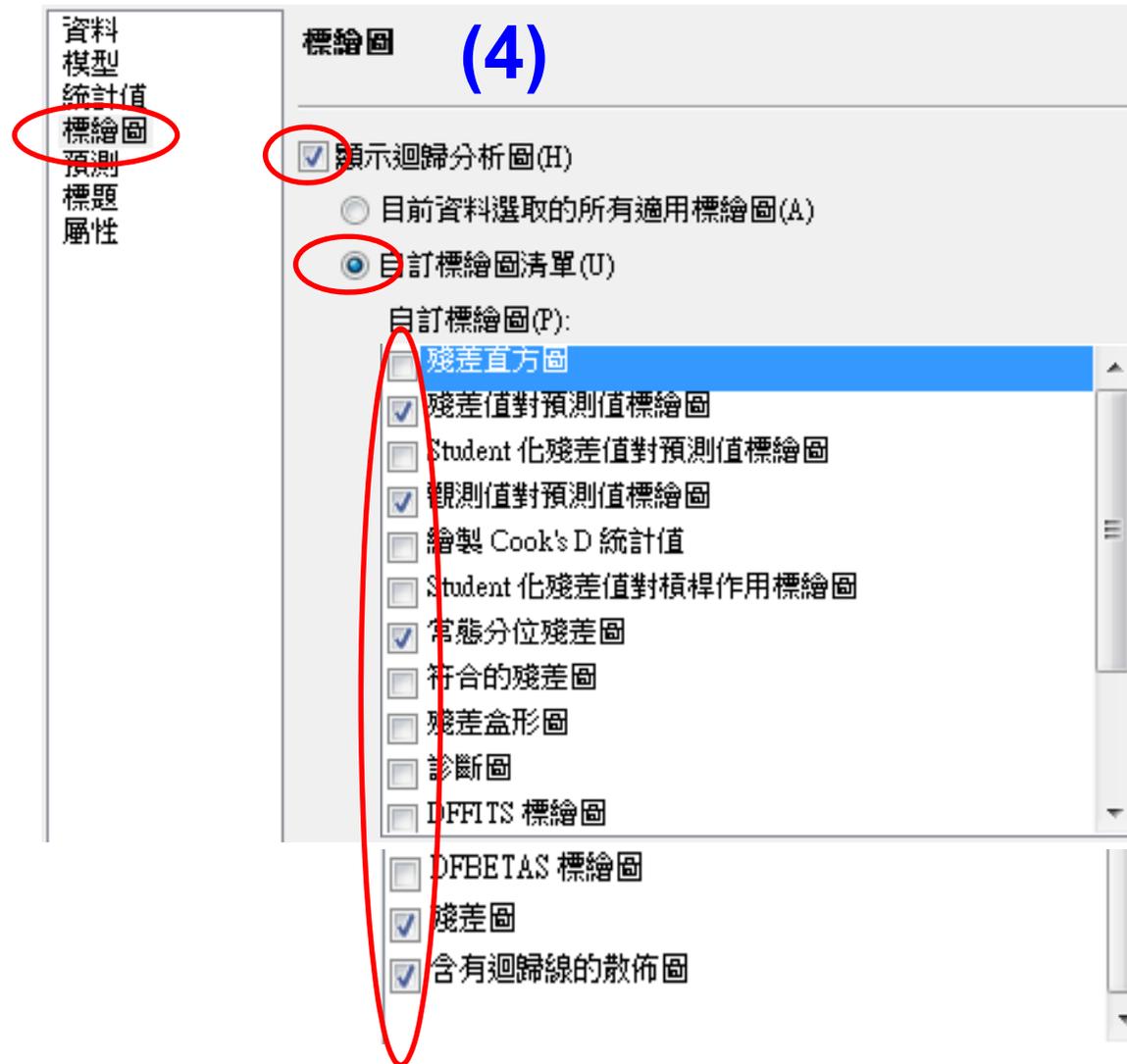
- 標準化迴歸係數(D)
- 平方和, 型一(1)
- 平方和, 型二(2)
- 估計值的相關矩陣(L)
- 估計值的共變異數矩陣(V)
- 參數估計值的信賴界限(F)

信賴水準(N): 95%

診斷

- 共線性分析(O)
- 不含截距的共線性分析(O)
- 估計值的允差值(T)
- 變異數膨脹值(I)
- 不等變異性檢定(H)
- 漸近共變異數矩陣(A)
- Durbin-Watson 統計值(B)

繪出變數間的散佈圖與殘差圖，可做回歸模式的檢測參考。



利用求得之直線回歸方程式進行 反應變數預測值之估計

資料
模型
統計值
標繪圖
預測
標題
屬性

預測 (5)

要預測的資料

原始樣本(O)
 其他資料(A)

瀏覽(B)...

儲存輸出資料

預測(P)
 診斷統計值(G)

SASApp:WORK.PREDLinReg 瀏覽(E)...

其他統計值

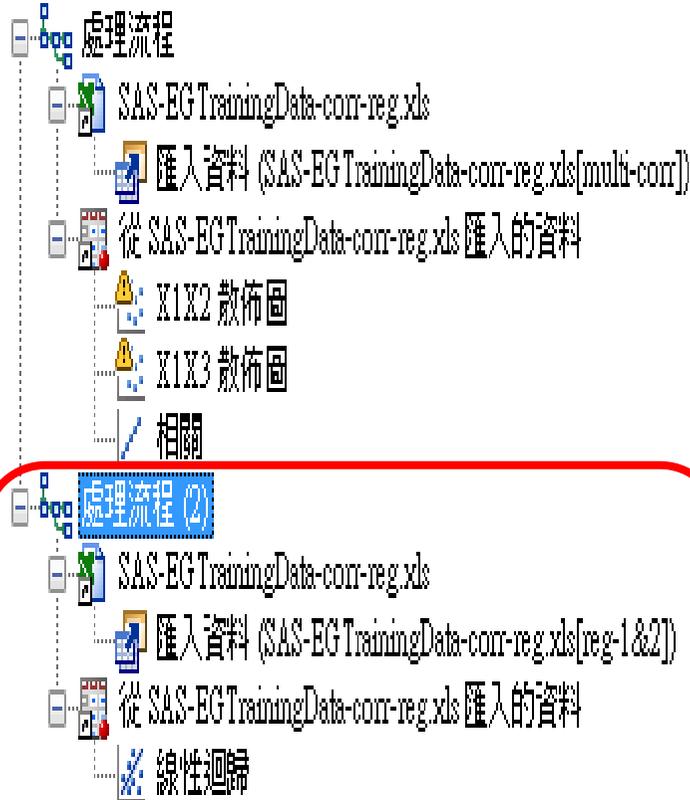
殘差值(U)
 預測界限(L)

顯示輸出和標繪圖(D)
 顯示預測值(W)

3. 直線回歸分析結果解讀-ANOVA

直線回歸分析的 處理流程圖

專案樹狀結構



回歸變方分析表：
假說 $H_0: \beta=0$ $H_1: \beta \neq 0$

讀取的觀測值數目 6
使用的觀測值數目 6

$P < 0.01$
表示接受 H_1 ，
此回歸式存在

來源	自由度	平方和	平均值平方	F 值	Pr > F
模型	1	3.61157	3.61157	105.63	0.0005
誤差	4	0.13676	0.03419		
已校正的總計	5	3.74833			

根 MSE	0.18491	R 平方	0.9635
應變平均值	15.11667	調整 R 平方	0.9544
變異係數	1.22320		

決定係數為 0.96

直線回歸方程式的截距與回歸係數及其 t檢定的顯著性值

回歸方程式的截距		參數估計值				
變數	自由度	參數估計值	標準誤差	t 值	Pr > t	變異數膨脹
Intercept	1	16.25238	0.13383	121.44	<.0001	0
施氮量(kg)	1	-0.09086	0.00884	-10.28	0.0005	1.00000

$H_0: \beta = \beta_0$, (β_0 為一常數) $H_1: \beta \neq \beta_0$

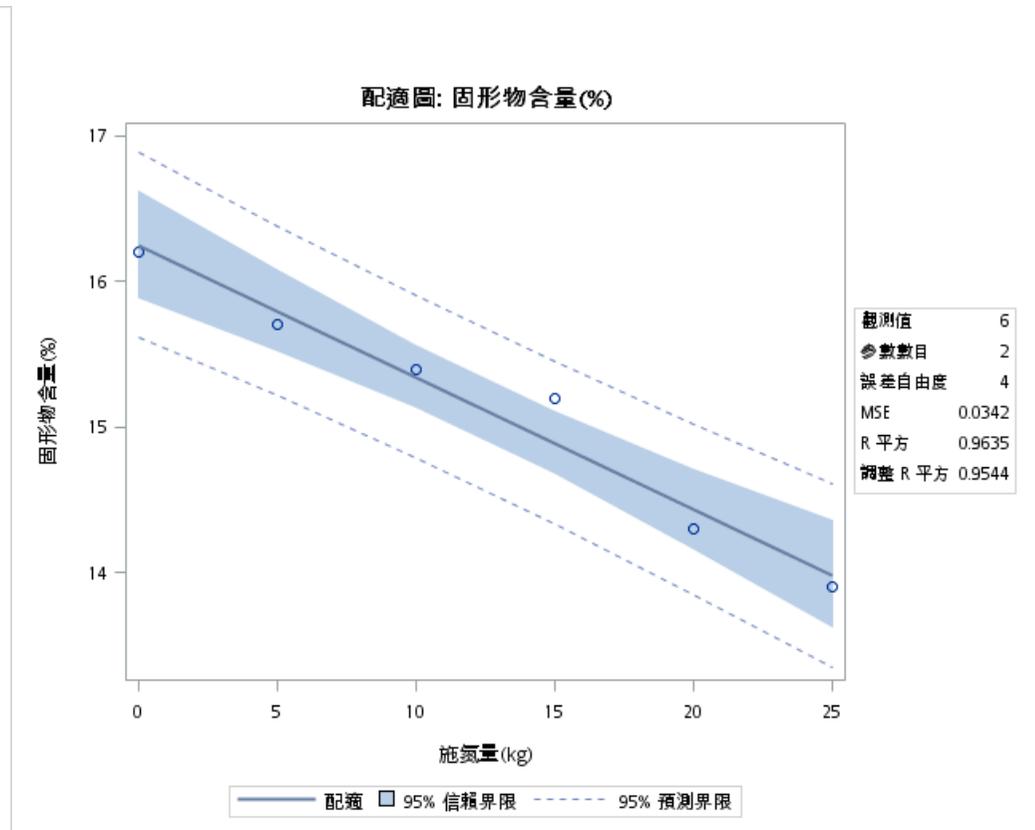
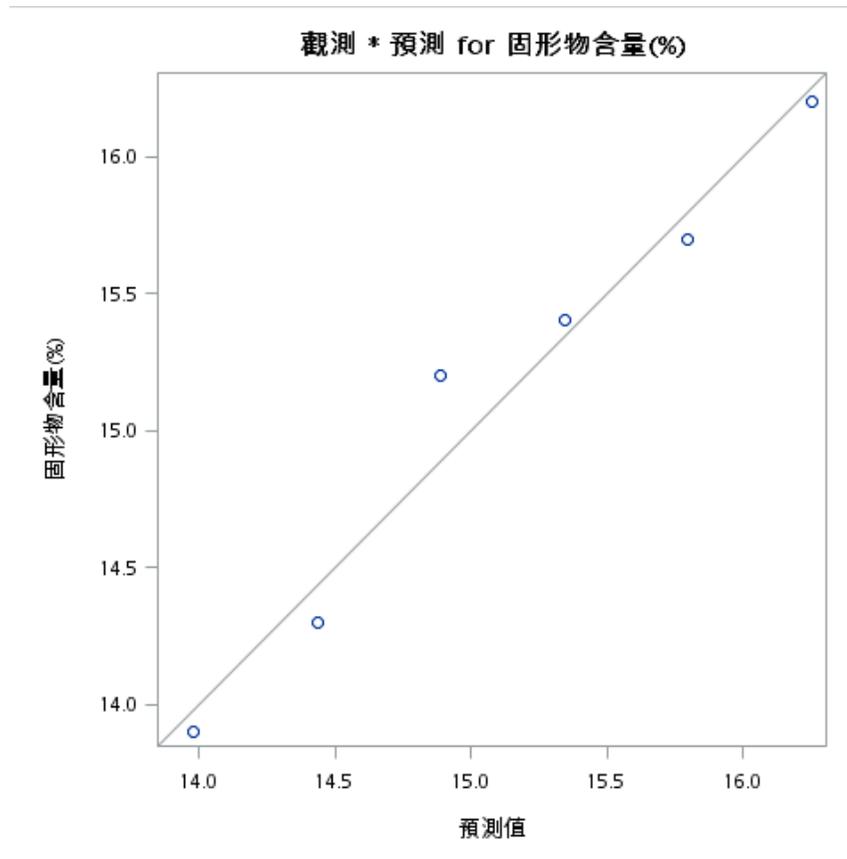
$$|t| = \frac{|b - \beta_0|}{SE_b} > t_{(n-2, \alpha/2)}$$

回歸方程式的回歸係數與t檢定的顯著性值，P值小於0.05，表示此回歸係數顯著存在。

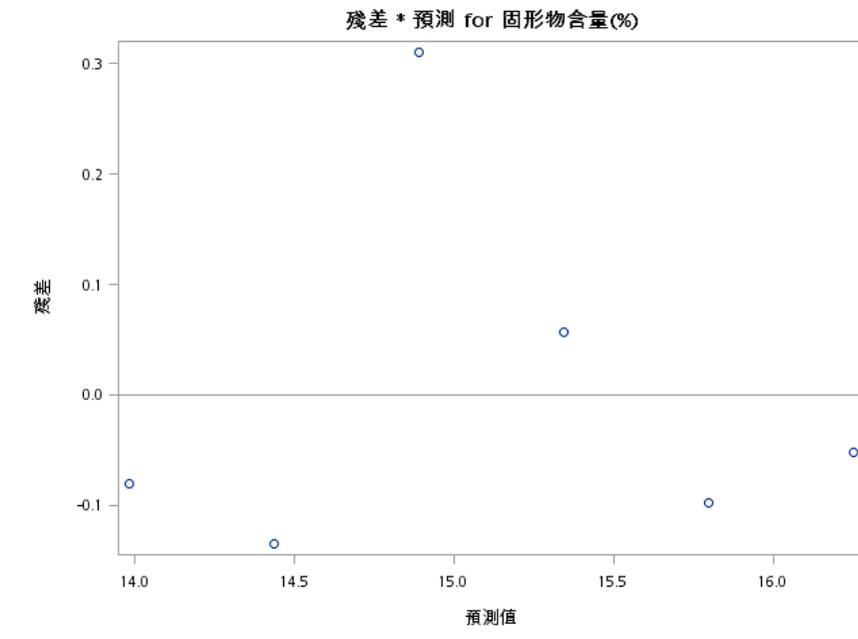
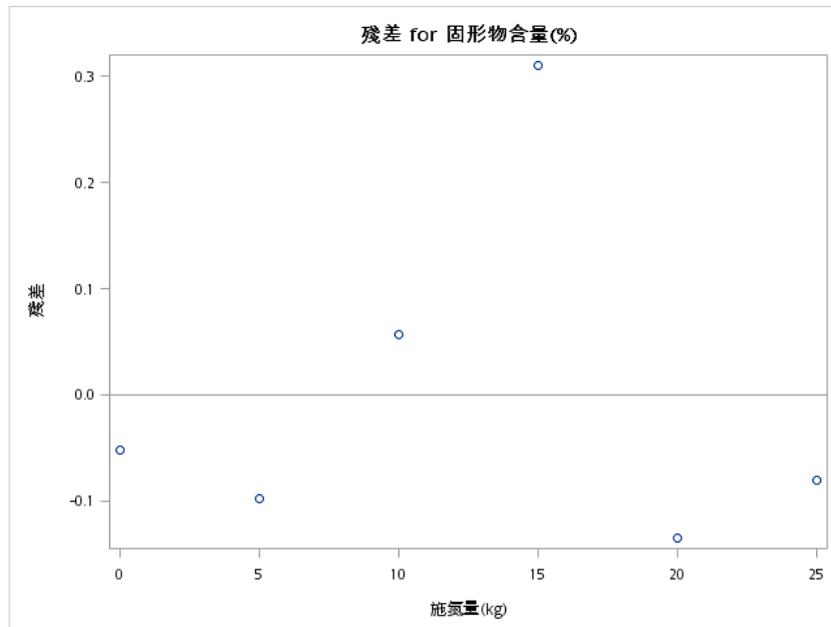
當 $t > t_{\alpha/2, n-2}$ 拒絕 H_0 ，表示該回歸係數顯著存在。

回歸方程式 $Y = 16.25 - 0.09^{**}X$

反應變數、解釋變數與反應變數預測值的繪圖



殘差分析：當殘差值圖形以 $e=0$ 為中心線兩邊上下的區域對稱分布，表示該回歸線與資料的配合程度。



H_0 : 資料均值
 H_1 : 資料不均值

第一和第二動差規格的檢定

自由度	卡方	Pr > ChiSq
2	1.49	0.4748

P-value > 0.05 ,
接受均值性的
假設

利用此回歸方程式計算在各解釋變數
對應點上的反應變數預測值。

固形物含量(%)	predicted_固形物含量(%)
16.2	16.2524
15.7	15.7981
15.4	15.3438
15.2	14.8895
14.3	14.4352
13.9	13.9810

五、二次效應回歸分析

資料範例說明

【試驗內容】

欲探討水梨果園的六個氮肥施用量 $x(\text{kg})$ 和可溶性固形物含量 $y(\%)$ 間之關係，現要探討是否存在氮肥的二次效應。

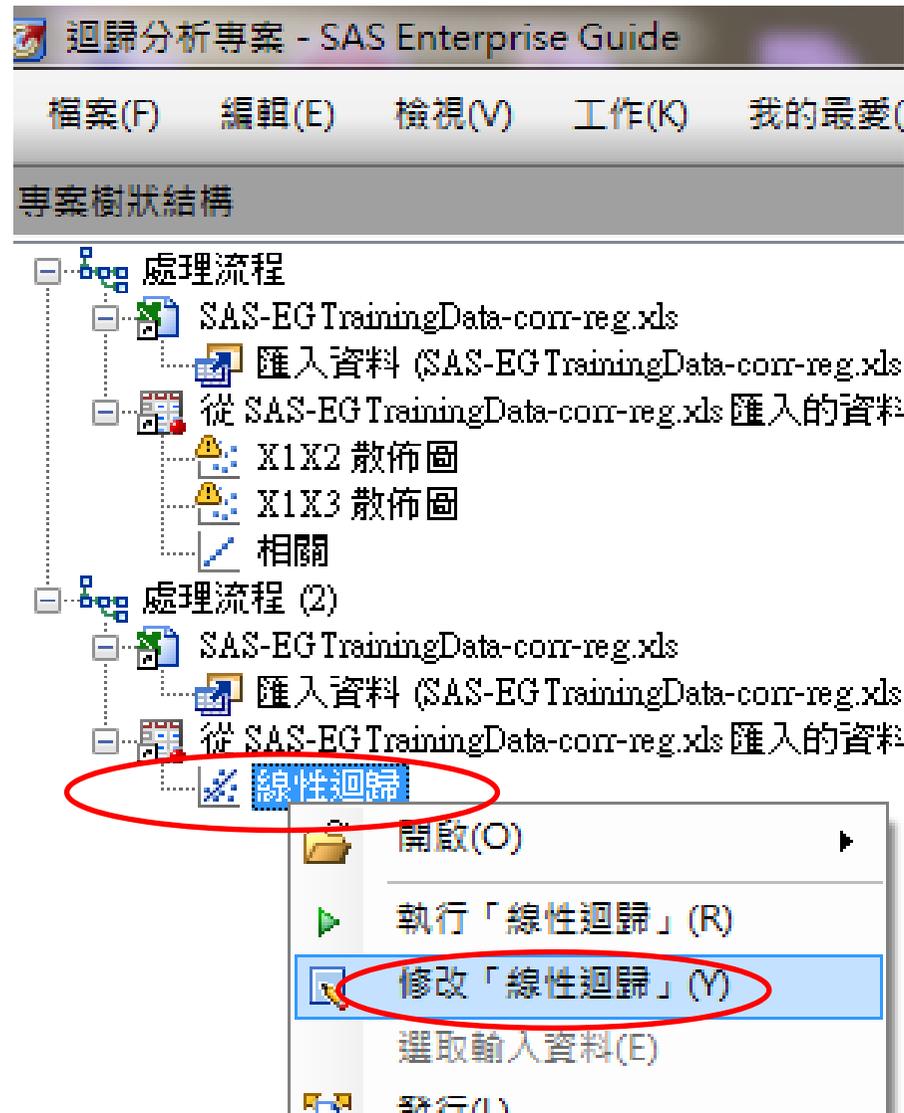
【工作表】 reg-1&2

【分析內容】

二次效應回歸分析

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i}^2 + \varepsilon_i \quad , i = 1..n$$

1. 處理流程>線性迴歸 按右鍵「修改線性迴歸」

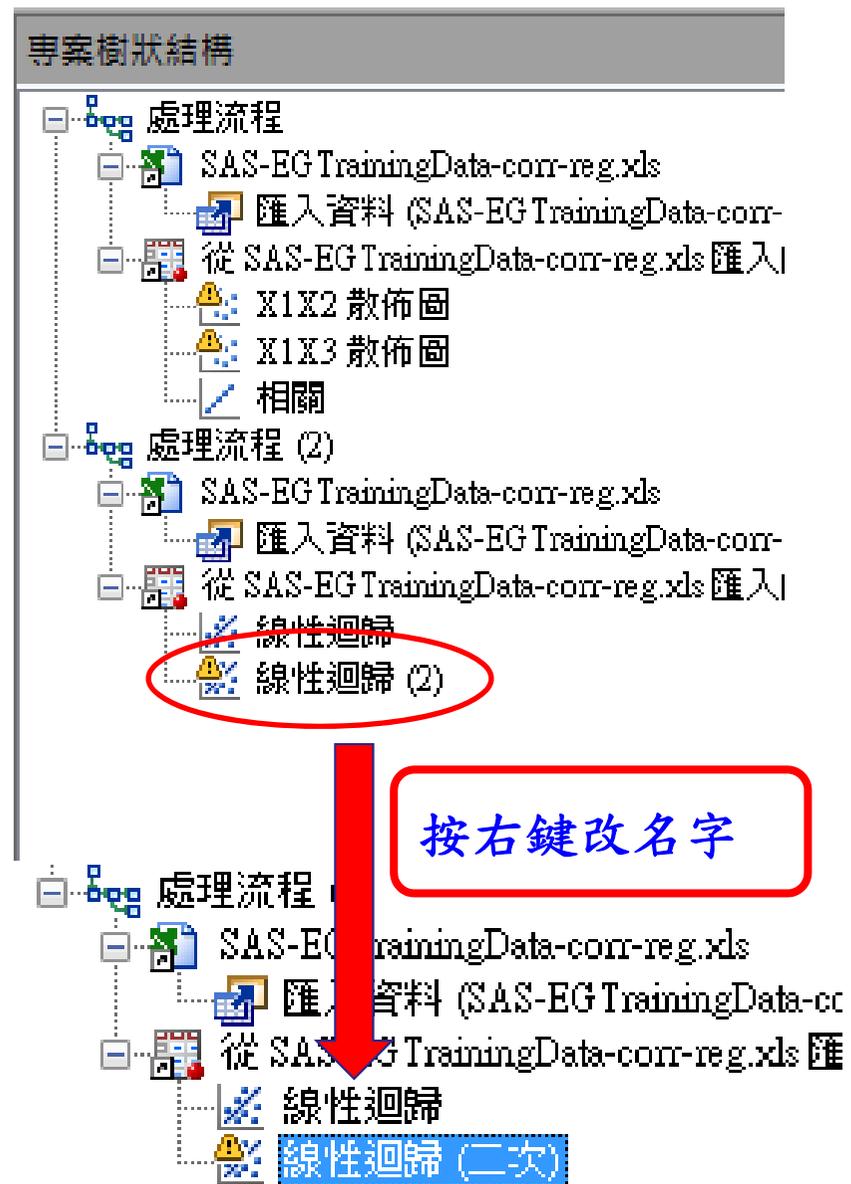
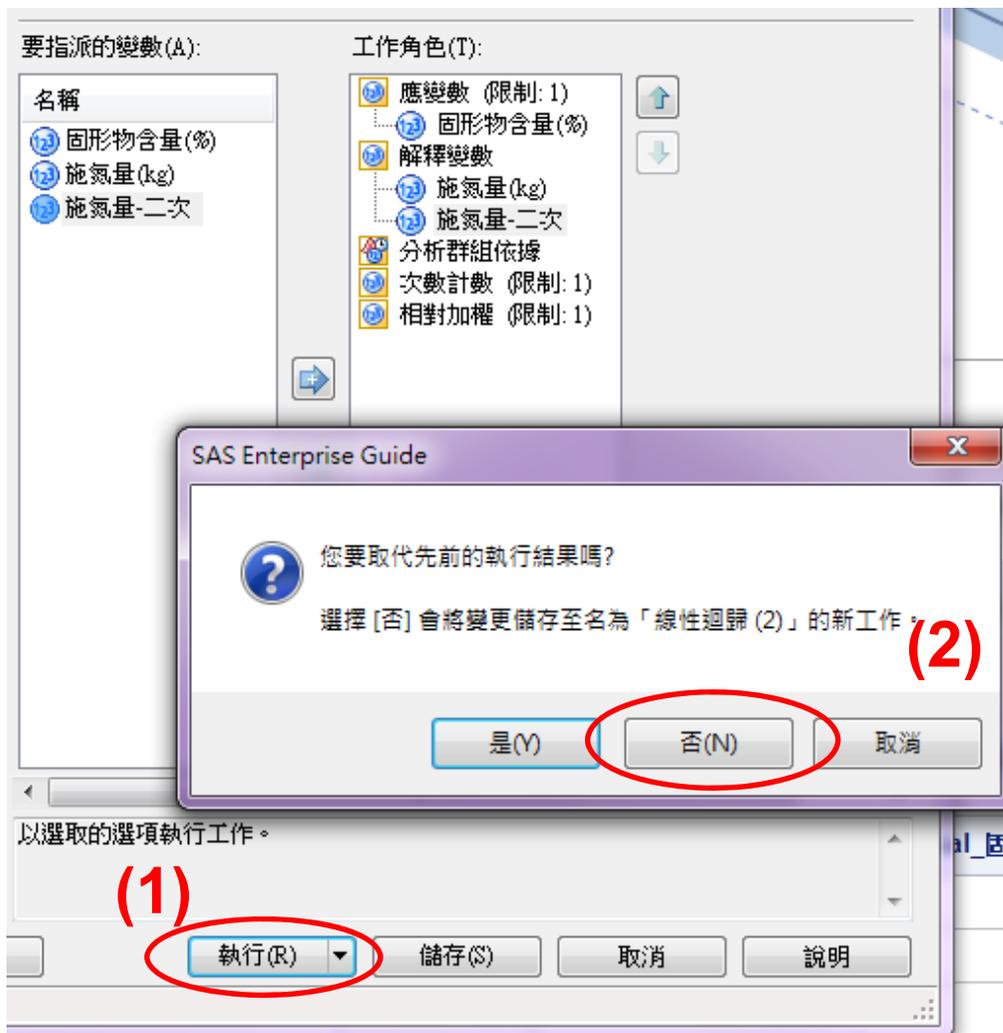


在「資料」內用滑鼠拖曳(或按箭頭按鈕)將「施氮量-二次」指派到右側工作角色框內的「解釋變數」。

The screenshot displays the SAS software interface for variable assignment. On the left, a vertical menu lists options: 資料 (highlighted with a red circle), 模型, 統計值, 標繪圖, 預測, 標題, and 屬性. The main area is titled '資料' and shows '資料來源: SASApp:WORK.SAS_EGTRAININGDATA_CORE' and '工作篩選: 無'. Below this, two panels are shown: '要指派的變數(A):' and '工作角色(T):'. The '要指派的變數(A):' panel lists variables: 名稱, 固形物含量(%), 施氮量(kg), and 施氮量-二次 (highlighted with a red circle). The '工作角色(T):' panel lists roles: 應變數 (限制: 1), 固形物含量(%), 解釋變數, 施氮量(kg), 施氮量-二次 (highlighted with a blue dashed border), 分析群組依據, 次數計數 (限制: 1), and 相對加權 (限制: 1). A red arrow points from the '施氮量-二次' variable in the left panel to the '施氮量-二次' role in the right panel.

其餘設定內容與直線回歸分析相同

- (1) 點選「執行」
- (2) 點選「否」



2. 二次效應回歸分析結果解讀-ANOVA

回歸變方分析表：假說 $H_0: \beta_1 = \beta_2 = 0$

H_1 : 任一個 $\neq 0$

變異數分析					
來源	自由度	平方和	平均值平方	F 值	Pr > F
模型	2	3.65455	1.82727	58.45	0.0040
誤差	3	0.09379	0.03126		
已校正的總計	5	3.74833			

$P < 0.01$
表示接受 H_1 ，
此回歸式存在

根 MSE	0.17681	R 平方	0.9750
應變平均值	15.11667	調整 R 平方	0.9583
變異係數	1.16964		

決定係數為0.96

二次回歸方程式的截距與回歸係數及其t檢定的顯著性值

回歸方程式的截距		參數估計值			
變數	自由度	參數估計值	標準誤差	t 值	Pr > t
Intercept	1	16.13929	0.16025	100.71	<.0001
施氮量(kg)	1	-0.05693	0.03015	-1.89	0.1554
施氮量-二次	1	-0.00136	0.00116	-1.17	0.3256

回歸方程式的回歸係數與t檢定的顯著性值，此兩個回歸係數之P值皆大於0.05，表示此兩個回歸係數不存在，所以此回歸方程式不適用，即此試驗的氮肥二次效應不存在。

六、複回歸分析

資料範例說明

【試驗內容】

研究身高大致相同的男人血管收縮壓(y)與他們的身高(x1)、腰圍(x2)、體重(x3)和年齡(x4)的關係，共有13個對象的資料

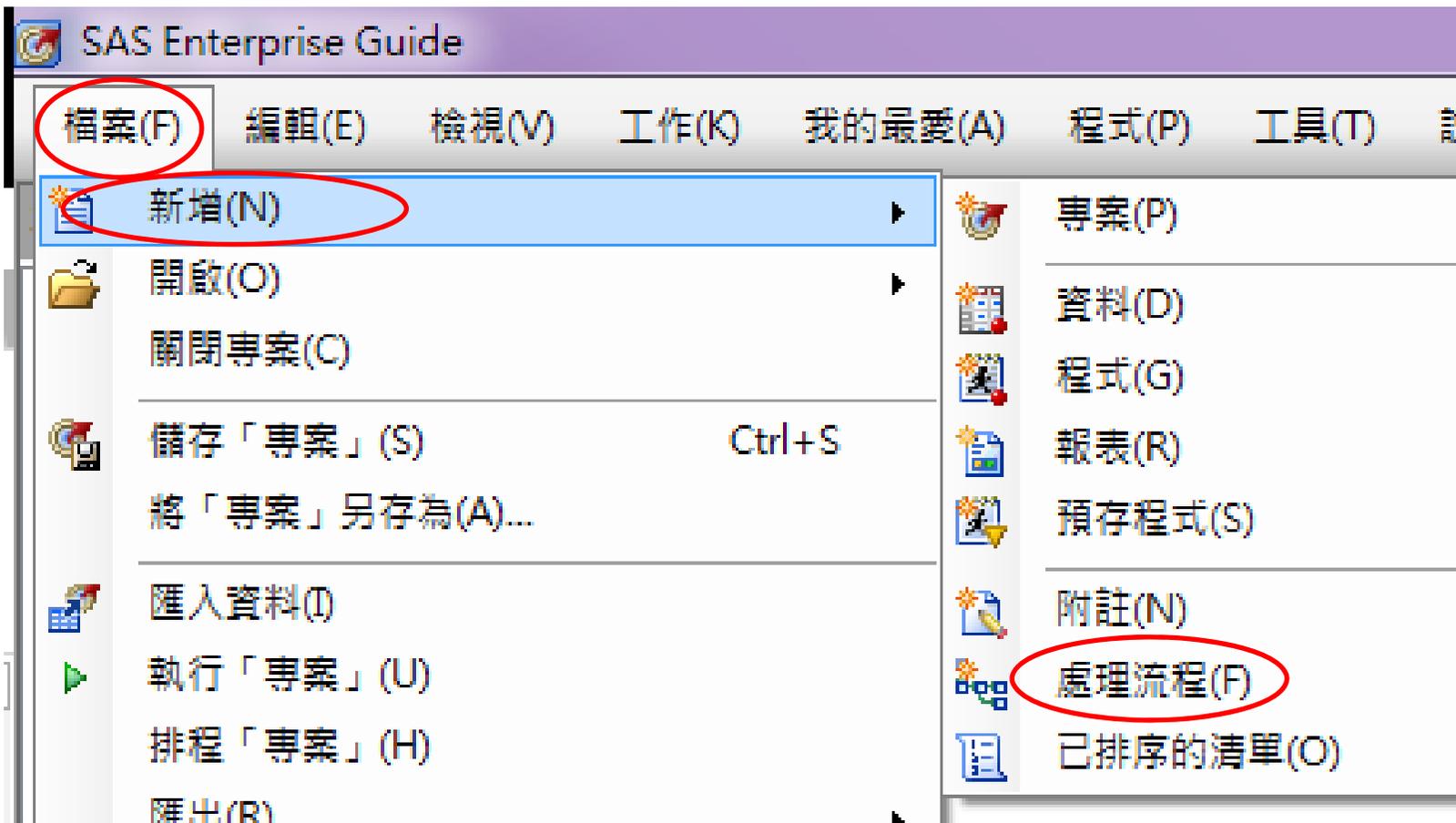
【工作表】 multi-reg

【分析內容】

這4個解釋變數是否都能用來預測反應變數血管收縮壓(y)，若否，應保留哪幾個解釋變數。

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \varepsilon_i$$

新增「處理流程」(如下圖)



載入資料檔，記得先點選載入的Excel資料工作表 **multi-reg**。(如投影片第4-7頁)

multi-reg 工作表預覽

	⑫ 身高(cm)	⑫ 腰圍(CM)	⑫ 體重(Kg)	⑫ 年齡(歲)	⑫ 血壓(mm/Hg)Y
1	170	80	52	50	120
2	155	99	83	21	141
3	170	95	71	20	124
4	168	80	65	30	126
5	170	75	58	32	117
6	176	70	61	50	129
7	168	65	49	60	123
8	165	75	58	50	125
9	172	83	70	40	132
10	178	66	53	55	123
11	166	74	64	45	132
12	158	95	90	40	155
13	166	90	85	22	147

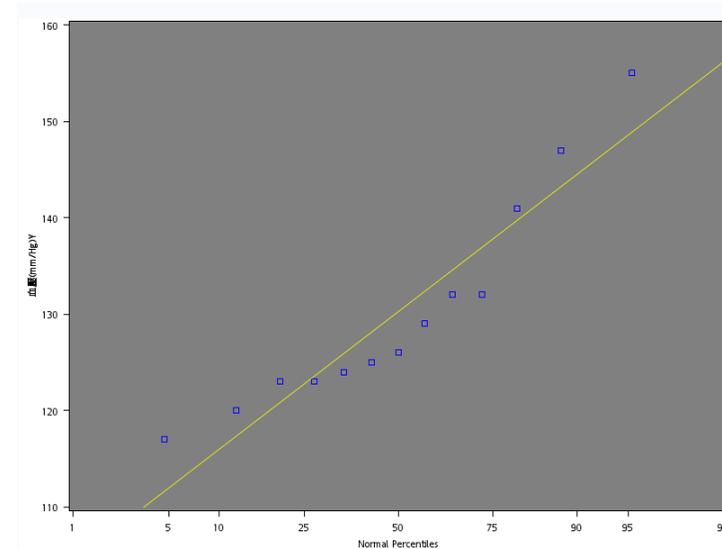
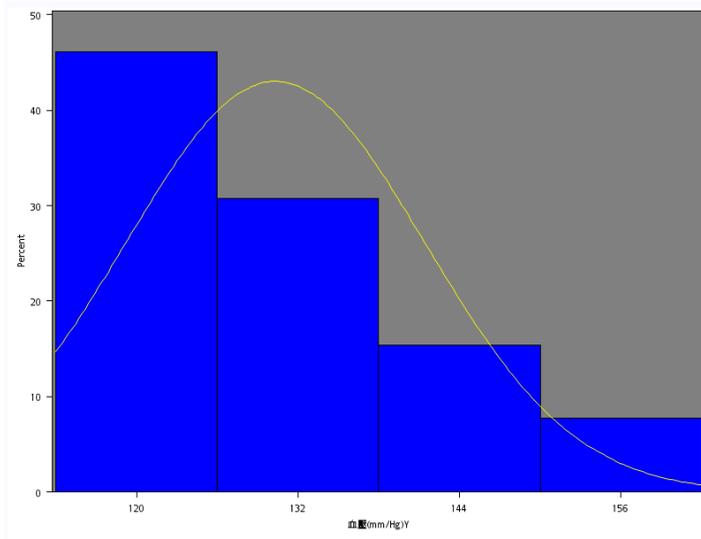
1. 常態分布檢測: 參考投影片第26~31頁

Parameters for Normal Distribution

參數	符號	估計值
Mean	Mu	130.3077
Std Dev	Sigma	11.13092

Goodness-of-Fit Tests for Normal Distribution

檢定	統計值		p 值	
Kolmogorov-Smirnov	D	0.20880980	Pr > D	0.122
Cramer-von Mises	W-Sq	0.11472668	Pr > W-Sq	0.066
Anderson-Darling	A-Sq	0.64288758	Pr > A-Sq	0.076



2. 複回歸分析：工作/回歸/線性回歸



在「資料」內用滑鼠拖曳(或按箭頭按鈕)將血壓和身高、腰圍、體重、年齡分別指派到右側工作角色框內的"應變數"和"解釋變數"

- 資料
- 模型
- 統計值
- 標繪圖
- 預測
- 標題
- 屬性

資料 (1)

資料來源: SASApp.WORK.SAS_EGTRAININGDATA_CORP
工作篩選: 無

要指派的變數()

名稱
身高(cm)
腰圍(CM)
體重(Kg)
年齡(歲)
血壓(mm/Hg)Y

工作角色(T):

- 應變數 (限制: 1)
 - 血壓(mm/Hg)Y
- 解釋變數
 - 身高(cm)
 - 腰圍(CM)
 - 體重(Kg)
 - 年齡(歲)
- 分析群組依據

- 資料
- 模型
- 統計值
- 標繪圖
- 預測
- 標題
- 屬性

模型 (2) 選擇完整模式

模型選取法(M):

符合的完整模型 (沒有選取)

顯著水準

(3)

資料
模型
統計值
標繪圖
預測
標題
屬性

統計值

估計值的詳細資料

- 標準化迴歸係數(D)
- 平方和，型一(1)
- 平方和，型二(2)
- 估計值的相關矩陣(L)
- 估計值的共變異數矩陣(V)
- 參數估計值的信賴界限(F)

信賴水準(N): 95%

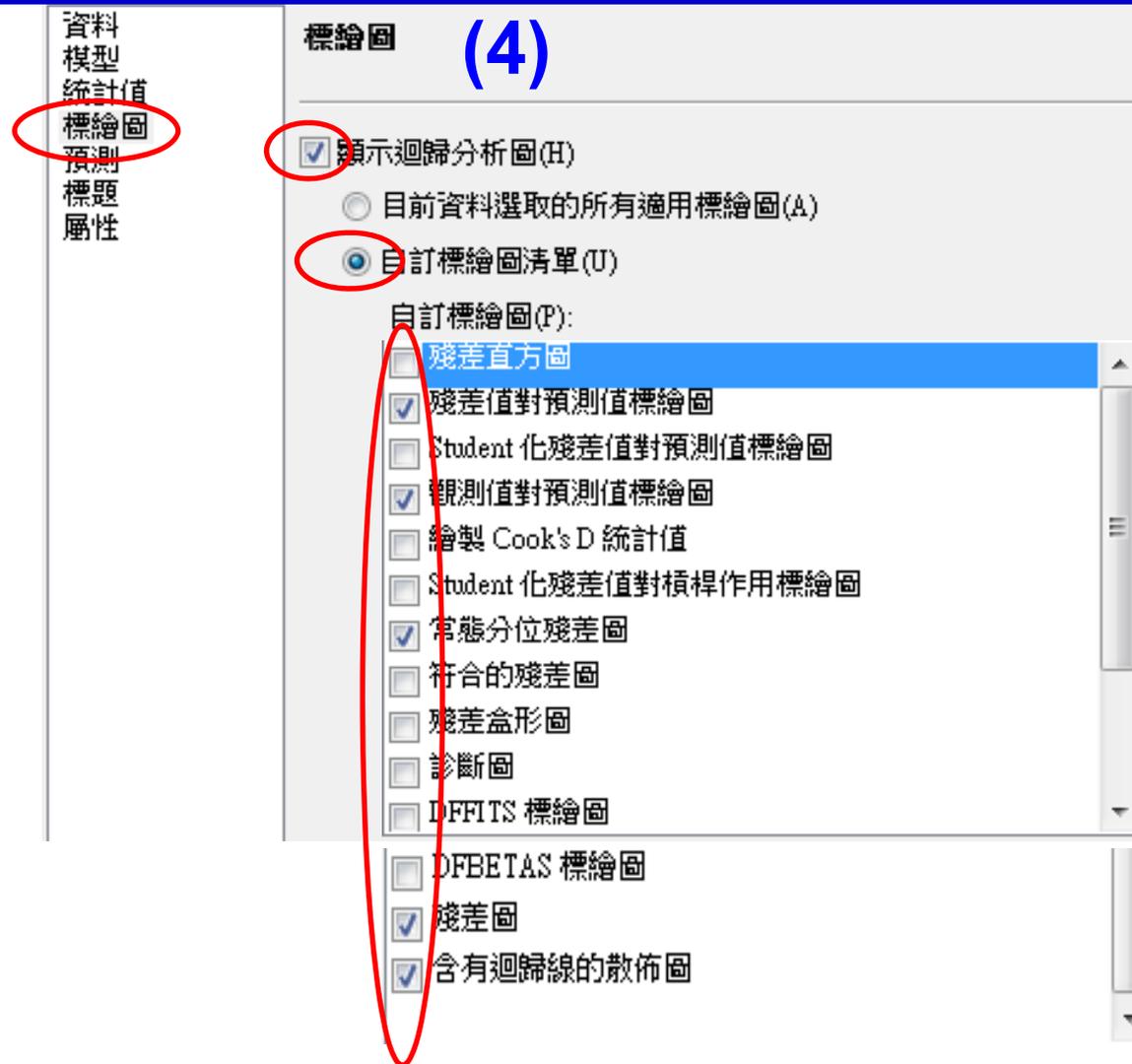
診斷

- 共線性分析(O)
- 不含截距的共線性分析(W)
- 估計值的允差值(T)
- 變異數膨脹值(I)
- 不等變異性檢定(H)
- 漸近共變異數矩陣(A)
- Durbin-Watson 統計值(B)

VIF

進行回歸模式之解釋變數的
共線性與自我相關的檢測

繪出變數間的散佈圖與殘差圖，可做回歸模式的檢測參考。



利用求得之直線回歸方程式進行 反應變數預測值之估計

資料
模型
統計值
標繪圖
預測
標題
屬性

預測 (5)

要預測的資料

原始樣本(O)
 其他資料(A)

瀏覽(B)...

儲存輸出資料

預測(P)
 診斷統計值(G)

SASApp:WORK.PREDLinReg 瀏覽(E)...

其他統計值

殘差值(U)
 預測界限(L)

顯示輸出和標繪圖(D)
 顯示預測值(W)

3. 複回歸分析結果解讀-ANOVA

回歸變方分析表：假說 $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$ H_1 : 其中一個 $\neq 0$

讀取的觀測值數目	13
使用的觀測值數目	13

變異數分析					
來源	自由度	平方和	平均值平方	F 值	$P_r > F$
模型	4	1455.88878	363.97220	94.29	<.0001
誤差	8	30.88045	3.86006		
已校正的總計	12	1486.76923			

$P < 0.01$ ，表示接受 H_1 ，此回歸式存在

根 MSE	1.96470	R 平方	0.9792
應變平均值	130.30769	調整 R 平方	0.9688
變異係數	1.50774		

決定係數為0.97

複回歸方程式的截距與各回歸係數及其t檢定的顯著性值

回歸方程式的截距		參數估計值				變異數膨脹
變數	自由度	參數估計值	標準誤差	t 值	Pr > t	
Intercept	1	88.64612	29.84946	2.97	0.0179	0
身高(cm)	1	-0.17578	0.13065	-1.35	0.2154	2.11232
腰圍(CM)	1	-0.24351	0.13294	-1.83	0.1044	6.97350
體重(Kg)	1	1.15366	0.08783	13.13	<.0001	4.15482
年齡(歲)	1	0.36721	0.07551	4.86	0.0013	3.24512

回歸方程式的回歸係數與t檢定的顯著性值，P值小於0.05，表示此回歸係數顯著存在，若>0.05則此回歸係數不存在。

VIF診斷結果，通常>10會考慮剔除此變數

共線性診斷							
數目	特徵值	條件索引	變異的比例				
			Intercept	身高(cm)	腰圍(CM)	體重(Kg)	年齡(歲)
1	4.85582	1.00000	0.00001411	0.00002606	0.00010598	0.00034918	0.00116
2	0.13198	6.06560	9.300107E-7	0.00000571	0.00237	0.01335	0.15165
3	0.00955	22.55338	0.00478	0.01991	0.00085563	0.37952	0.39371
4	0.00244	44.64628	0.00208	0.03971	0.67255	0.58222	0.21169
5	0.00021537	150.15598	0.99312	0.94034	0.32411	0.02456	0.24180

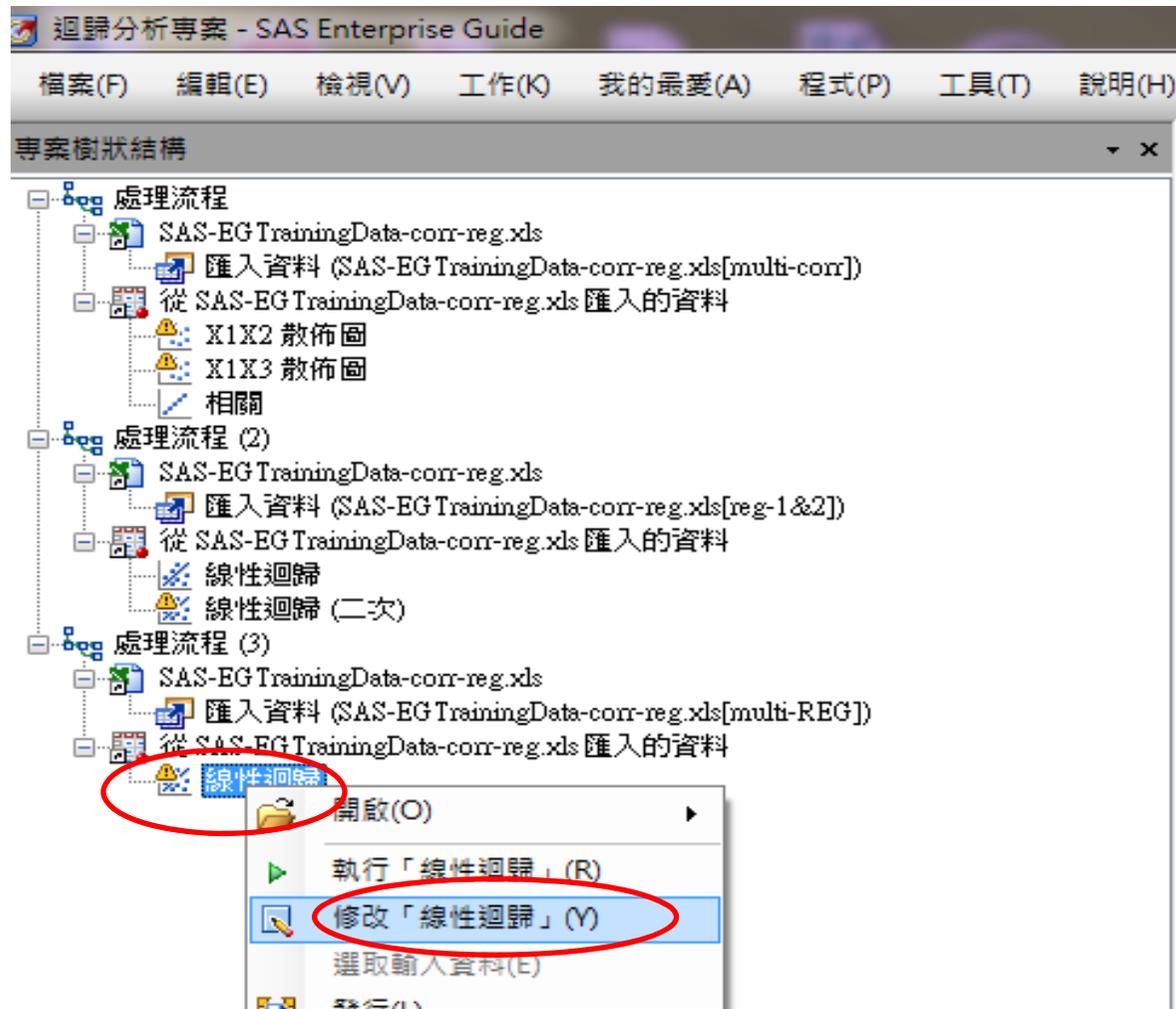
共線性診斷條件索引>100，代表此變數共線性大，應移除回歸式

第一和第二動差規格的檢定		
自由度	卡方	Pr > ChiSq
13	12.87	0.4579

代表資料是否具均質性？

七、複回歸分析:逐步回歸法

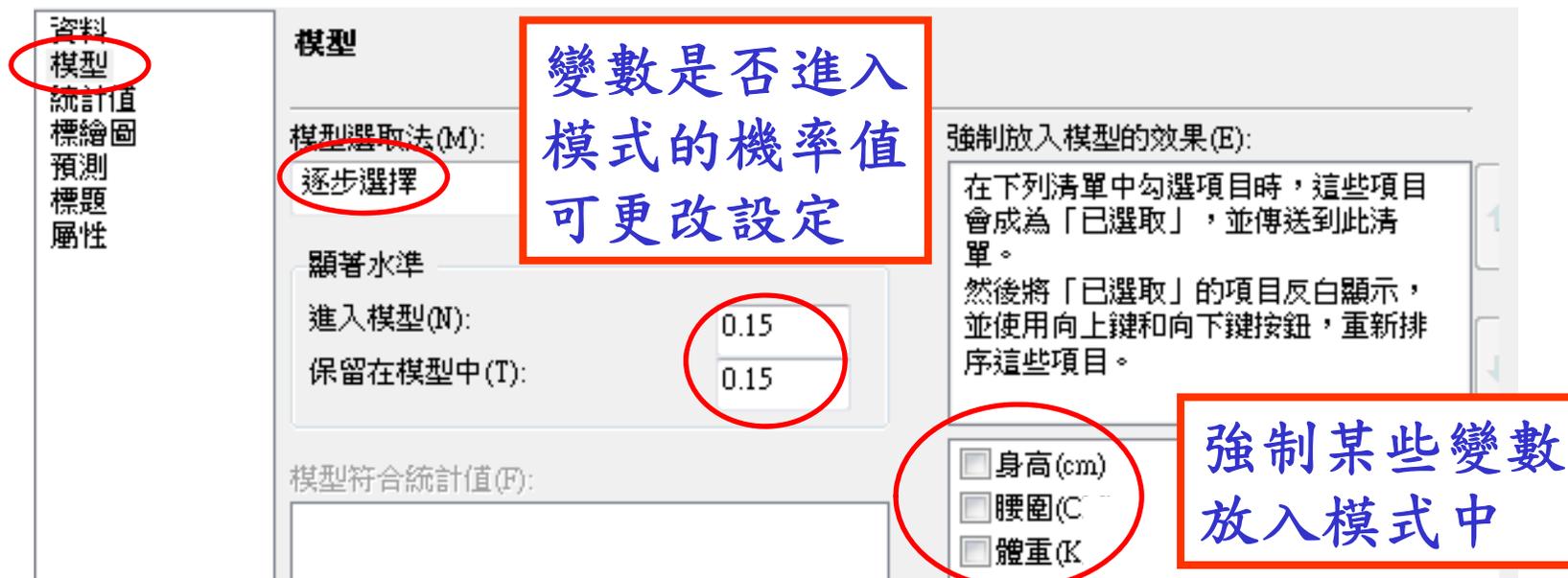
處理流程>線性回歸 按右鍵「修改線性回歸」



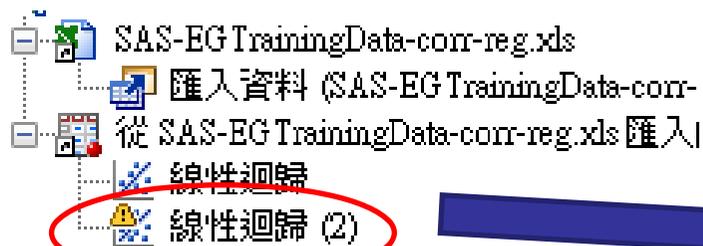
1. 在「資料」內用滑鼠拖曳(或按箭頭按鈕)將「年齡」拉離解釋變數



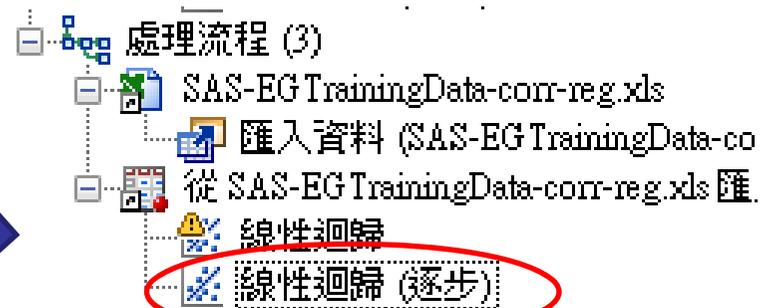
選擇「逐步選擇」模式



(1) 點選「執行」
(2) 點選「否」



按右鍵改名字



2. 逐步回歸法之變數選取過程

第一個解釋變數體重進入模式中，且模式成立。

逐步選擇: 步驟 1

已輸入變數 體重(Kg): R 平方 = 0.8145 和 C(p) = 11.3143

變異數分析					
來源	自由度	平方和	平均值平方	F 值	Pr > F
模型	1	1211.00800	1211.00800	48.31	<u><.0001</u>
誤差	11	275.76123	25.06920		
已校正的總計	12	1486.76923			

變數	參數估計值	標準誤差	第二型 SS	F 值	Pr > F
Intercept	79.87593	7.38775	2930.54498	116.90	<.0001
體重(Kg)	0.76323	0.10981	1211.00800	48.31	<u><.0001</u>

第二個解釋變數腰圍再進入模式中，且模式成立。

逐步選擇: 步驟 2

已輸入變數 腰圍(CM) R 平方 = 0.8956 和 C(p) = 4.4344

變異數分析					
來源	自由度	平方和	平均值平方	F 值	Pr > F
模型	2	1331.55083	665.77541	42.89	<u><.0001</u>
誤差	10	155.21840	15.52184		
已校正的總計	12	1486.76923			

變數	參數估計值	標準誤差	第二型 SS	F 值	Pr > F
Intercept	97.58628	8.61286	1992.62757	128.38	<.0001
腰圍(CM)	-0.55975	0.20086	120.54283	7.77	0.0192
體重(Kg)	1.17746	0.17193	727.97456	46.90	<.0001

當全部變數都依設定之機率值0.15為標準篩選，最後只保留二個解釋變數體重與腰圍在模式中。

留在模型中的所有變數都是顯著於 0.1500 層級。

沒有其他變數符合輸入至模型中的 0.1500 顯著層級。

逐步選擇的摘要								
步驟	輸入的變數	移除的變數	數目 Vars In	偏 R 平方	模型 R 平方	C(p)	F 值	Pr > F
1	體重(Kg)		1	0.8145	0.8145	11.3143	48.31	<.0001
2	腰圍(CM)		2	0.0811	0.8956	4.4344	7.77	0.0192

3. 逐步回歸分析結果解讀-ANOVA

來源	自由度	平方和	平均值平方	F 值	Pr > F
模型	2	1331.55083	665.77541	42.89	<.0001
誤差	10	155.21840	15.52184		
已校正的總計	12	1486.76923			

$P < 0.01$ ，表示接受 H_1 ，此回歸式存在

根 MSE	3.93978	R 平方	0.8956
應變平均值	130.30769	調整 R 平方	0.8747
變異係數	3.02344		

回歸方程式

$$Y = 97.59 + 1.18^{**}\text{體重} - 0.56^{**}\text{腰圍}$$

$$R^2 = 0.87$$

截距與回歸係數

參數估計值

變數	自由度	參數估計值	標準誤差	t 值	Pr > t	變異數膨脹
Intercept	1	97.58628	8.61286	11.33	<.0001	0
腰圍(CM)	1	-0.55975	0.20086	-2.79	0.0192	3.95926
體重(Kg)	1	1.17746	0.17193	6.85	<.0001	3.95926

VIF

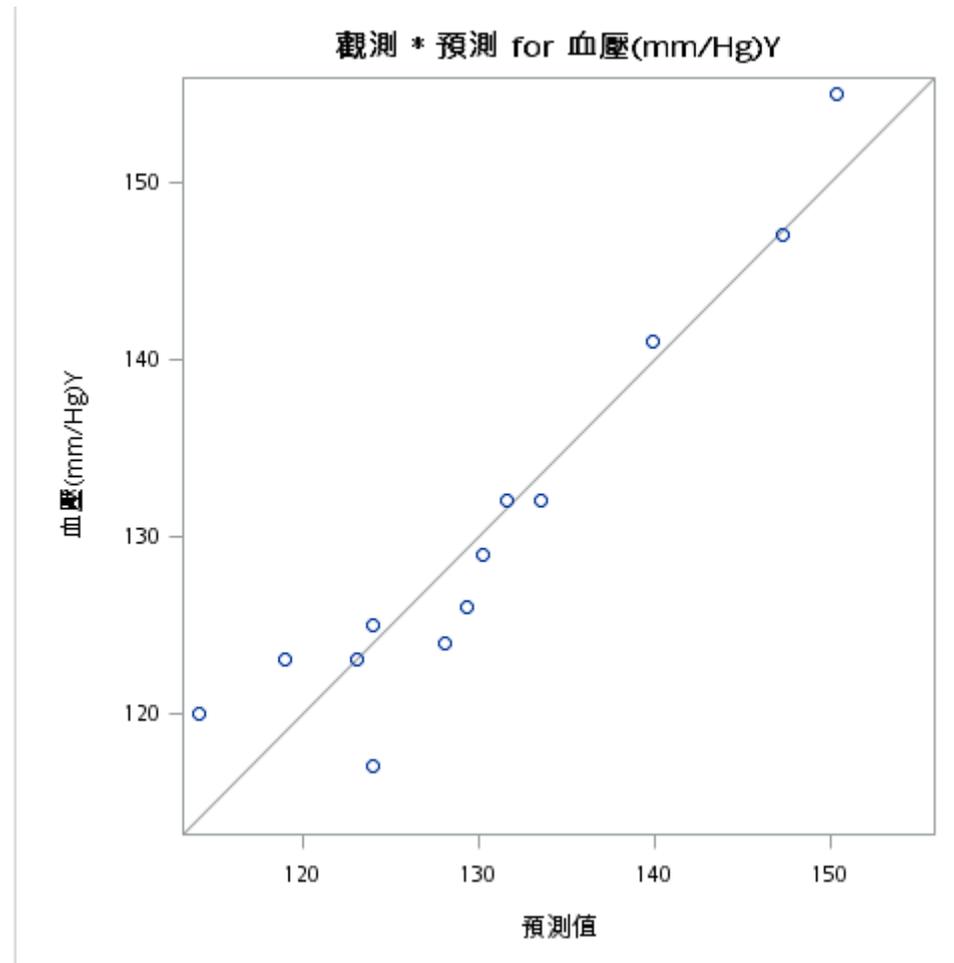
共線性診斷					
數目	特徵值	條件索引	變異的比例		
			Intercept	腰圍(CM)	體重(Kg)
1	2.97890	1.00000	0.00179	0.00050393	0.00099748
2	0.01818	12.80141	0.51344	0.00378	0.19873
3	0.00292	31.91729	0.48477	0.99571	0.80027

共線性診斷條件索引皆<100

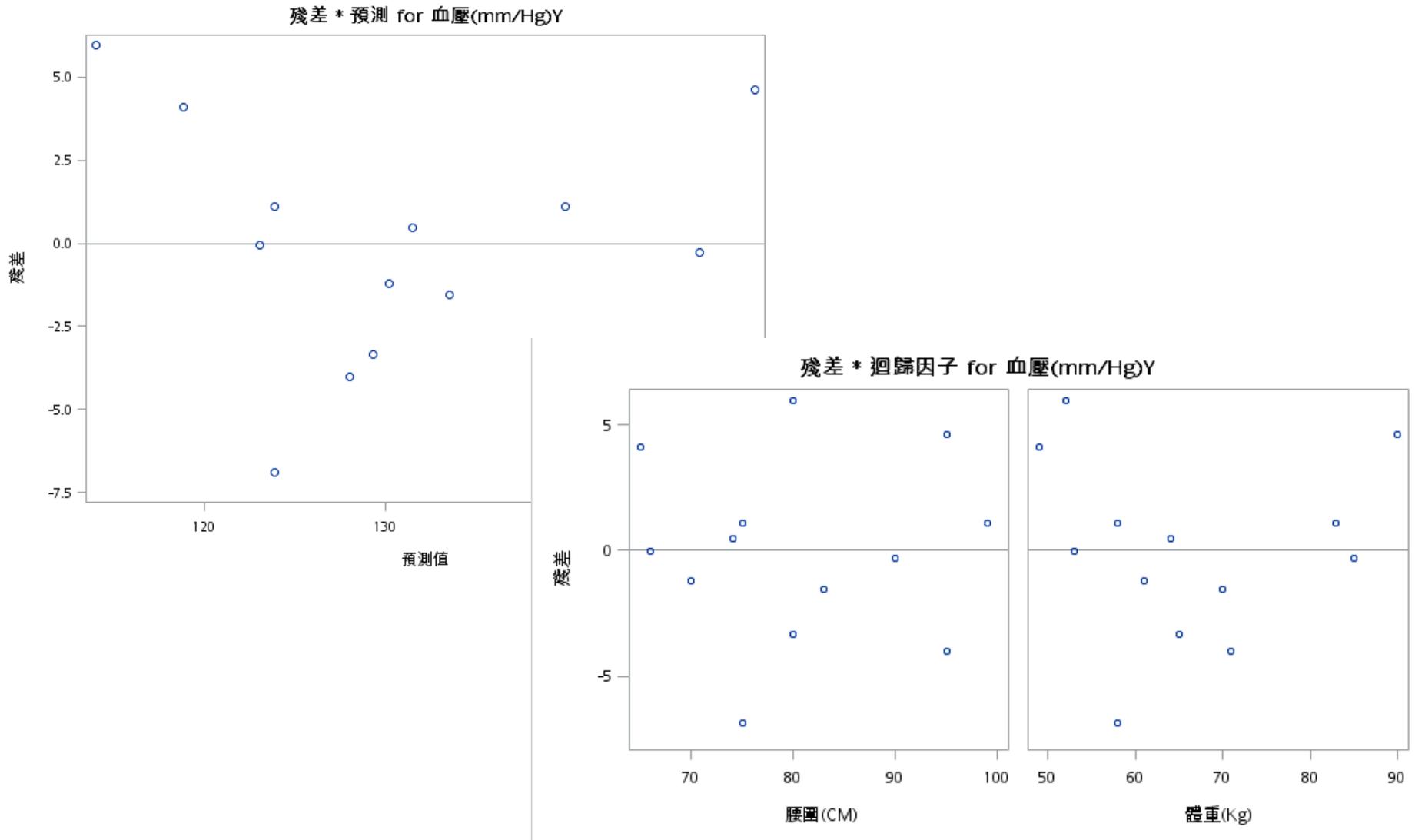
第一和第二動差規格的檢定		
自由度	卡方	Pr > ChiSq
5	4.11	0.5341

代表資料是否具均質性?

反應變數與反應變數預測值的散佈圖



殘差分析：當殘差值圖形以 $e=0$ 為中心線兩邊上下的區域對稱分布，表示該回歸線與資料的配合程度。



利用求得之回歸方程式計算在各解釋變數對應點上的反應變數預測值。

血壓(mm/Hg)Y	predicted_血壓(mm/Hg)Y
120	114.034
141	139.900
124	128.010
126	129.341
117	123.898
129	130.229
123	118.898
125	123.898
132	133.549
123	123.048
132	131.522
155	150.381
147	147.293

回歸方程式

$$Y=97.59 + 1.18*體重 -0.56**腰圍$$

原解釋變數值之反應變數預測值

$$Y=97.59 + 1.18X52-0.56X80=114.15$$

其他解釋變數之反應變數預測值

體重=80kg, 腰圍=90cm

$$Y=97.59 + 1.18X80-0.56X90= 141.59$$

感謝聆聽