

作物試驗之一般統計分析方法

楊嘉凌

摘要

農業研究者處理作物試驗資料，可利用如Microsoft Office Excel套裝軟體統計工具，對資料進行簡單之初步分析。然而較複雜之複因子試驗資料則多利用如SAS、SPSS、STATISTICA等功能強大的統計分析軟體，惟對一般使用者而言具有考驗。作物試驗資料分析，一般先將資料進行變方分析以判斷處理效應是否顯著，再分析解釋各效應性質或進一步進行平均值多重比較，惟多數的農業試驗研究，若再以迴歸分析進行估計反應曲線或反應曲面之結果與訊息較理想。

前言

作物試驗經常對蒐集資料進行各種統計分析，研究者處理資料大多使用較簡單之統計描述，繪製各種圖表，或進行t檢定、變方分析、相關分析、迴歸分析等較不複雜的方法。利用現有之Microsoft Office套裝軟體工具，其簡明快捷且易學易用，若精通其資料管理及分析技巧(譬如t檢定、F檢定、變異數分析、相關及迴歸分析)，並注意應用上的限制，便足以應付試驗資料的基本統計分析，惟其僅略為解決部份的資料分析。較複雜的複因子試驗或多變量分析，Excel分析工具較不易有精確度。部分統計分析軟體，如SAS、SPSS、STATISTICA等軟體之統計分析功能強大，惟非短時間能學會。本文即介紹EXCEL與SAS分析軟體之簡單操作流程，提供本場同仁參考。

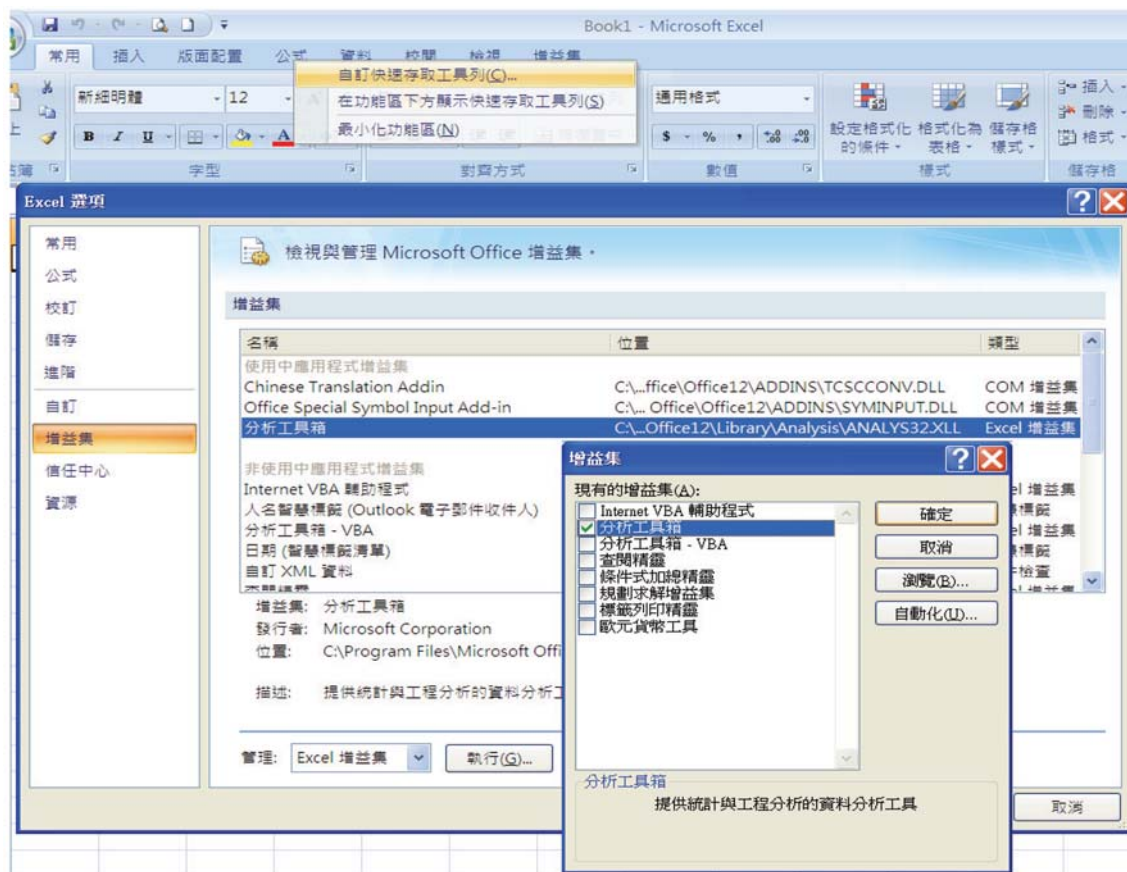
內容

一、Excel視窗下進行資料分析：

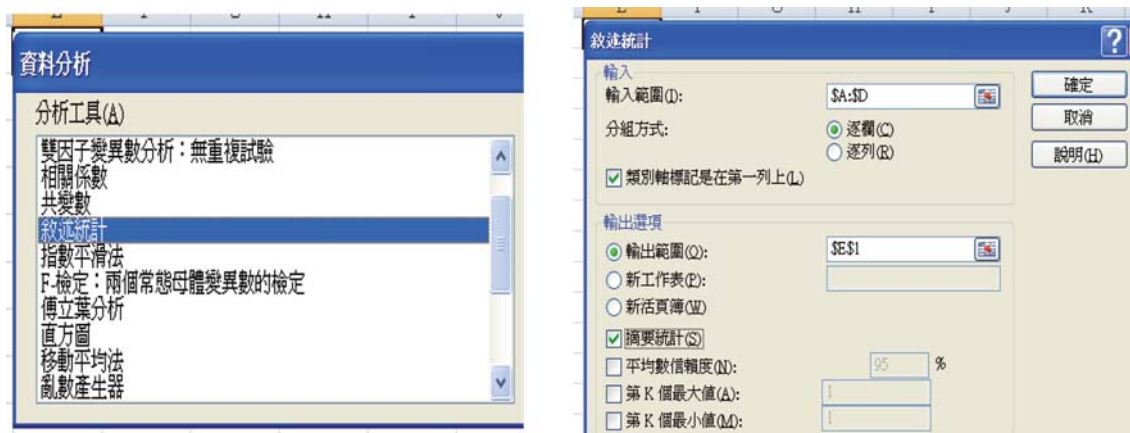
當進行資料分析之前，於Office Excel視窗下進行「分析工具箱」之安裝步驟如下，以便日後統計分析之使用。

(一) Microsoft Office 2003版本：Excel畫面上方快捷鍵，選擇「工具」/「增益集」，「增益功能」勾選「分析工具箱」。

(二) Office 2007版本：Excel畫面，上方快捷鍵-按右鍵，選按「自訂快速存取工具列」，左欄選取「增益集」，勾選「分析工具箱」。如下圖所示。



Microsoft Office EXCEL之分析工具(如下圖所示)，可對試驗資料進行變異數分析(變方分析，ANOVA)、相關係數、敘述統計、F-檢定、t檢定、迴歸分析等。



EXCEL之變異數分析針對單因子試驗資料為主，其分析工具之「單因子變異數分析」係分析CRD(完全逢機設計)且不等重複之資料，「雙因子變異數分析：

重複試驗」係分析CRD且等重複之資料，而「雙因子變異數分析：無重複試驗」可分析單因子之RCRD(逢機完全區集設計)資料，惟皆未提供變方分析後處理間之差異顯著性分析(呂等，2005)。由EXCEL之「敘述統計」，可得到1組試驗資料之平均數、標準誤、變異數、標準差等基本統計量。

「F-檢定」係測驗2組試驗資料之變方差異性。「t檢定」可進行2組試驗資料平均值之差異性測驗，惟需注意其資料分布之特性進行。2組試驗資料若來自同一來源、相同族群或互有關連，例如同一批作物植株於噴灑農藥前後之病斑面積差異比較、同一批土壤樣本分送2個不同化驗室進行PH值測定等，由於樣本成對，因此採用「t檢定：成對母體平均數差異檢定」；若2組試驗資料互為獨立無關連，譬如2種不同肥料對作物產量差異性比較、兩個不同品種間之產量差異比較等，則採用「t檢定：兩個母體平均數差的檢定」，惟2組樣本之觀測值個數未必相同，須先進行F-檢定2組樣本之變方是否相等，F-檢定若未顯著採用「t檢定：兩個母體平均數差的檢定，假設變異數相等」；反之，F-檢定若顯著，表示2組樣本之變方具差異性，則採用「t檢定：兩個母體平均數差的檢定，假設變異數不相等」，依此進行兩組資料之平均值差異性檢定(沈，1993)。

Microsoft Office EXCEL雖簡單方便，但仍有相當之限制，譬如變方分析僅限於單因子試驗，且分析後之多重比較並無直接分析工具，另須配合函數與公式；相關分析無法跨欄指定不同變數，顯著性測驗得配合函數與公式之步驟較繁瑣；迴歸分析之自變數欄位須連續，且僅能分析至16個自變數亦不可有缺值，其輸出結果之標準化殘差並不正確等。

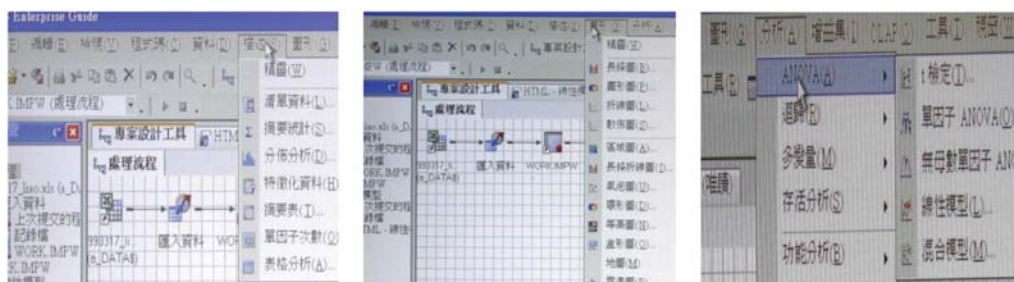
二、SAS-EG(Enterprise Guide)視窗下進行統計分析：

本場係透過農委會資訊中心申請SAS使用帳號及密碼，安裝程式後以連結到農委會SAS伺服器進行統計分析之操作。SAS-EG版本使用視窗操作介面，讓使用者充分運用其強大功能且以拖、拉、點、選方式即可進行統計分析及統計報表之製作，而不用再編寫SAS程式碼。

利用SAS-EG進行統計分析，先將蒐集之試驗資料建立成EXCEL檔案格式較為便利。一般操作流程於開啟SAS-EG畫面後，按取「檔案」下拉對話框之「匯入資料」，開啟已建立於「本機電腦」之EXCEL檔案的某個工作表，即可匯入SAS畫面進行分析(如下圖所示)。惟SAS係依據EXCEL工作表之各欄位進行資料讀取，EXCEL工作表即應將試驗資料建立完備，「各欄位」應清楚表示試驗之各別處理、區集或性狀資料等，欄位title盡量勿以中文輸入避免SAS讀取錯誤或產生亂碼。



匯入EXCEL資料後，即可於SAS視窗進行分析工作(如下圖)。按取「描述」可分別進行「摘要統計」得出資料之基本平均值、變方、標準差等；「分佈分析」顯示常態分布型資料等。按取「圖形」可對資料建立長條圖、圓形圖、雷達圖及曲面圖等圖形。按取「分析」可進行「ANOVA」得出資料之t檢定、單因子變方分析、線性與混合模型(複因子變方分析)等結果，亦可對資料進行「迴歸」、「多變量」等較複雜之分析。



SAS工具雖具有拖、拉、點、選之便利性，惟使用者仍應先具備試驗資料經統計分析結果之瞭解程度，即必須先了解試驗進行之目的與田間設計之佈置情形，由得到之試驗資料經SAS分析流程得到良好的變方分析模型結果(表一)，才具有相當可信度。

表一、一般試驗設計變方分析之變因及自由度

單因子RCBD		雙因子RCBD		裂區RCBD	
變因	自由度(df)	變因	自由度(df)	變因	自由度(df)
區集(n)	n-1	區集(n)	n-1	區集(n)	n-1
處理(a)	a-1	處理(a)	a-1	處理(a)	a-1
		處理(b)	b-1	機差(Ea)	(n-1)(a-1)
		交互(a*b)	(a-1)(b-1)	處理(b)	b-1
				交互(a*b)	(a-1)(b-1)

本文就一般單因子試驗之逢機完全區集設計(RCBD)，茲舉一例介紹SAS-EG之變方分析(ANOVA)流程。下圖所示係進行a、b、c、d等4個處理(tre)共4個區集(rep)的性狀(percent)資料。

1. 由SAS-EG視窗點按「檔案」之「匯入資料」，讀取並匯入EXCEL檔案，點按2次出現下圖左方之資料工作表，查閱輸入資料是否正確。
2. 點選「分析」功能表「ANOVA」之「線性模型」，出現下圖右方之「線性模型」對話框，依次進行工作角色、模型、多重檢定、繪圖等功能，以完成「分析」程序。
3. 工作角色：將指派變數tre及rep點按拉至分類變數(即試驗處理項)之下，指派percent至應變數(即試驗之性狀資料)之下。



4. 模型：點選「模型」出現下圖左方對話框，同時點選tre及rep拉至(點按)主要效果，交叉表示複因子試驗之因子交感項效果，本文例子僅單因子試驗，不用再點按交叉、巢狀或因子等效果。
5. Post Hoc檢定：變方分析後進行處理間平均質差異性比較，點選Post Hoc檢定下之「算術」，出現下圖右方對話框，在「要估計的效果」下方先點按新增，會在「平均值檢定的選項」出現對話框，點按「要使用的類別效果」下之tre，將原設定之「False」點按為「True」，點按「比較」之比較法，將原設定之「預設值」點選「成對檢定(即LSD比較)」，其他選項於本例子可省略跳過。



6. 基本上「線性模型」對話框進行至此，已具有變方分析及平均值比較之結果，接著之標繪圖、預測與標題等選項視使用者需要與否，最後點按右下方之儲存或執行。點按執行後，出現下圖左方之分析結果。
7. 至於變方分析結果之輸出格式，可先於SAS-EG主畫面之「工具」選單下方的「選項」，開啟如下圖右方之「選項」視窗，選取「結果」選單勾選HTML、PDF等格式，一般以可複製於Office Word並編輯的HTML格式較佳。

The screenshot displays the SAS-EG interface. The main window shows the results of a GLM procedure for the variable 'percen'. The results include a table of sources (Model, Error, Corrected Total) and an R-Square value of 0.473149. The 'Options' dialog box is open, showing the 'Results' section with 'HTML' and 'PDF' selected under 'Output Formats'. The 'Management' section has 'Do not replace' selected for 'Replace' and 'Generate SAS log when running' checked.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	6	0.6666750	0.11111250	1.85	0.3299
Error	6	0.7428828	0.12381380		
Corrected Total	12	1.4095578			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
ire	5	0.4731494	0.09462988	1.5	0.1704
rep	5	0.1934074	0.03868148	0.79	0.6324

結語

由作物試驗資料進行統計分析，一般係先將資料進行變方分析以判斷處理效應是否顯著，然後再進一步分析資料，以解釋各效應的性質(Steel and Torrie, 1980)。根據統計教科書或期刊雜誌，若變方分析之處理效應顯著，即選擇一種多重比較方法來判斷那幾個處理平均值間有顯著差異；其中最常見 Fisher的最小顯著差異法(LSD)及Duncan的多變域測驗法(MRT)。其實這種分析程序只適用各處理變級間無特定關係，即定性處理試驗資料，如品種或除草劑篩選試驗，該等試驗目的係於其中選取較佳處理。有關多重比較被誤用問題，國外已有許多報告提出警示，並提供正確可行的方法。然而反觀國內，多重比較被誤用情形仍屢見不鮮，對試驗結果的解釋影響甚大(呂，1990；蔣，2006)。其實大多數的農業試驗研究，其他統計方法要比多重比較法可提供更正確而豐富的資訊(魏等，1997)。過去認為可以使用Duncan多變域比較法的場合，反不如利用迴歸分析估得反應曲線或反應曲面較理想，因為由迴歸分析得到的訊息絕對要比Duncan多變域比較法為多。

參考文獻

1. 呂秀英 1990 多重比較方法之使用與誤用 科學農業 38:177-181。

2. 呂秀英、魏夢麗、呂椿棠 2005 用Excel解決農業研究資料統計分析的方法
(二) t及F檢定 農業試驗所技術服務 64:30-33。
3. 沈明來 1993 試驗設計學 Pp617 九州圖書文物有限公司出版。
4. 蔣國司 2006 多重比較法在植物保護研究上之使用 植物保護學會會刊
48:259-268。
5. 魏夢麗、呂椿棠、呂秀英 1997 處理平均值間之差異也可以用圖示表達 農
業試驗所技術服務季刊 8(4):20-23。
6. 魏夢麗、陳烈夫、呂秀英 2003 利用Excel進行農業試驗資料之統計分析的應
用及限制 科學農業 51:63-70。
7. Steel, R. G. D. and J. H. Torrie. 1980. Principles and Procedures of Statistics.
McGraw-Hill, New York.