

用 Excel 解決農業研究資料統計 分析的方法 (七) — 誤差線

▲ 農試所技服組 呂秀英

▲ 農試所作物組 魏夢麗 呂椿棠

一、前言

在統計或科學計數資料中，通常會以資料點或長條形來代表個體的數值或平均值，同時也在該資料點或資料標記上加上一個誤差線，用來表達出該資料點或資料標記的潛在誤差量或不確定度。這種圖形表達方式相當重要，可顯示資料的可信賴程度。然而統計圖中的誤差線可以是SD (standard deviation)、SE (standard error of mean) 或 CI (confidence intervals)，它們的計算方式和表達意義明顯不同，但很多研究人員對於究竟應該採用哪一種，往往不清楚。本篇目的在釐清這三種誤差線的差異，同時說明如何利用Excel快速計算以及在圖上繪製出誤差線。

二、SD、SE和CI的意義

SD、SE 和 CI的計算公式和應用場合之區分，整理如表一。必須注意的是，只有當資料是常態分布或近似於常態分布時，這三種誤差量才具意義。

(一)SD：將變方開平方根後的值，用來表達樣本各資料點分布的離散或變異情形，簡稱「標準差」。

(二)SE：將SD除以觀測值數目開平方根 (\sqrt{n}) 的值，用來測量樣本平均值的估計準確度，是平均值的SD，通常在正式學術報告中多以SEM或SE of mean 稱謂 (即多了一個mean，以強調是平均數的標準差，而非觀測值的標準差)，因此確實的名稱應該為「平均值的標準差」，但常有人將它簡稱為標準誤，這種稱法過於簡略，很容易造成與SD混淆不清。

(三)CI：當我們想了解在95%或99%可信賴度下平均數的真值所落在範圍是多少時，這個範圍就是CI，稱為「信賴區間」，平均值之95%CI的公式為 $\text{mean} \pm (1.96 \times \text{SE})$ ，而平均值之99%CI的公式為 $\text{mean} \pm (2.58 \times \text{SE})$ ；由於通常設定在5%或1%統計顯著水準下 (即要求95%或99%信賴度) 來檢定各樣本平均值間的差異性，因此這種CI常被拿來作為處理間比較分析之用。

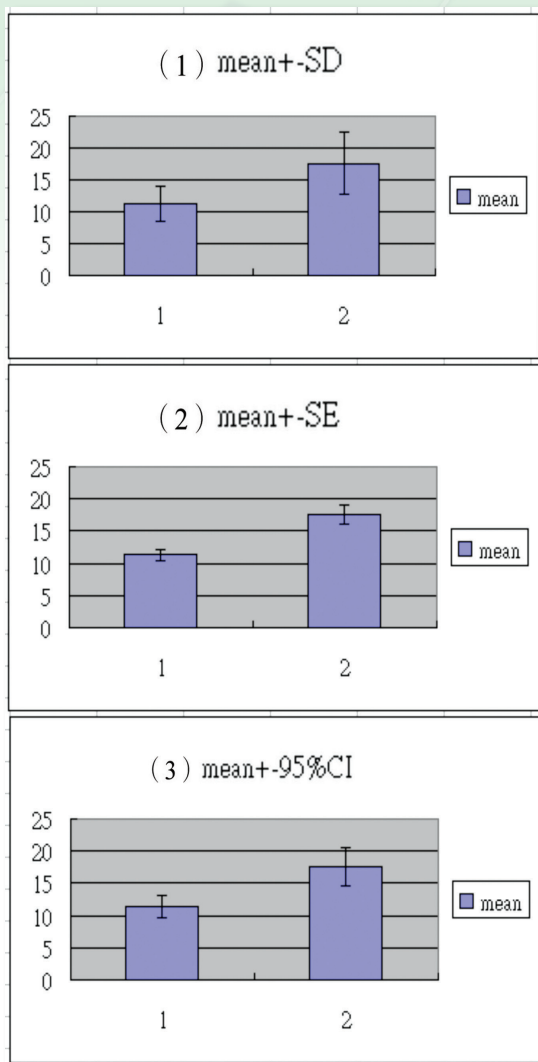
當我們在描述資料時，應該使用 $\text{mean} \pm \text{SD}$ ，讓讀者可清楚地感覺到資料觀測點的分布狀況，例如圖一-(1)實例，兩處理資料各為 11.3 ± 2.8 及 17.6 ± 4.8 ，表示處理1的10個資料點之分布變異較處理2為小 (此解讀僅限測量單位相同的兩組資料)。而在使用 $\text{mean} \pm \text{SEM}$ 時，圖一-(2)的兩處理資料各為 11.3 ± 0.9 及 17.6 ± 1.5 ，表示我們可以推測處理1之平均值的準確度較處理2為高。至於95%或

作者：呂組長秀英
連絡電話：04-23302301-7450

99%CI，可以讓我們知道平均值的95%或99%可信賴範圍，如圖一-(3)的兩處理資料各為 11.3 ± 1.7 及 17.6 ± 3.0 ，表示該兩處理平均值的95%可信賴範圍分別落在(9.6,13.0)和(14.6, 20.6)。由此可知，統計圖中的誤差線可以是SD、SE或CI，因此必須在圖中加註標明，以免混淆。其中只有從CI誤差線的重疊與否，可直接判斷出兩樣本平均值間的差異是否顯著或極顯著存在，例如圖一-(3)在使用95%CI誤差線時，當平均值較大的處理2之誤差下限與平均值較小的處理1之誤差上限沒有重疊時，代表在5%顯著水準下該兩處理平均值之間存在顯著差異；倘若重疊，則兩樣本平均值無顯著差異。不過必須注意的是，若比較的處理超過兩個以上時，由於我們的目標是進行兩兩處理之間的比較，利用兩兩CI誤差線彼此之間是否重疊來判斷多個處理平均值間的差異顯著性，可能不是那麼精準，但無論如何，相較於SD或SE誤差線，透過CI誤差線仍有助於從圖示目視約略判斷出處理間的差異性 (Streiner, 1996)。

三、利用EXCEL快速計算誤差量的方法

活用Excel的函數和公式，可容易地計算出樣本資料的平均值和三種誤差



圖一、用Excel繪製出(1)SD、(2)SE和(3)95%CI三種不同類型的誤差線圖(兩處理之資料如圖二)。

表一、SD、SE及CI三種誤差量的計算公式和意義

誤差量	簡稱	計算公式	意義
標準差 (standard deviation)	SD	$SD = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}}$	樣本各資料點分布的離散或變異情形
平均值的標準差 (standard error of mean)	SE或SEM	$SE = \frac{SD}{\sqrt{n}}$	樣本平均值的估計準確度
信賴區間 (confidence intervals)	CI	95%CI=1.96×SE 99%CI=2.58×SE	樣本平均值的95%或99%信賴區間

註：上述公式中 x_i 為資料觀測值， \bar{x} 為平均值， n 為觀測值個數。

量。以圖二之資料為範例，處理1資料集所在位址為B2:B11，在指定位址輸入=AVERAGE(B2:B11)，即計算出平均值；輸入=STDEVA(B2:B11)，可得SD；進而利用於B13位址的已知SD，輸入=B13/SQRT(COUNT(B2:B11))，可得SE（其中SQRT為開平方根函數，COUNT為計算觀測值個數之函數）；至於95% CI可再利用於B14位址的已知SE乘以1.96求得，即輸入=1.96*B14。另一個處理2，則直接利用複製B12:B15範圍位址後貼上C12:C15，即可快速求得該四種統計量的估計值。除了直接輸入函數和公式的方法外，平均值、SD和SE也可經由Excel在功能表上之「工具」中的「資料分析」，選擇「敘述統計」，可針對多個變數性狀快速準確地一次計算出各種常用的敘述統計量；但CI仍必須經由上述輸入公式始能求得。

四、在Excel圖中加入、修改和取消誤差線的方法

利用Excel所繪製出的區域圖、橫條圖、直條圖、折線圖或XY散佈圖上，可加上誤差線。由於農業研究上常以處理平均值作為資料點來進行差異性比較，因此本篇以圖二之資料為例，說明如何利用Excel繪製出以平均值為數列且含誤差線的長條圖，以及產生後如何修改誤差線類型與取消誤差線（長條圖以外其他類型圖的誤差線之加入、修改和取消的方法皆相同，請參照以下(二)~(四)步驟）。

(一)長條圖之繪製

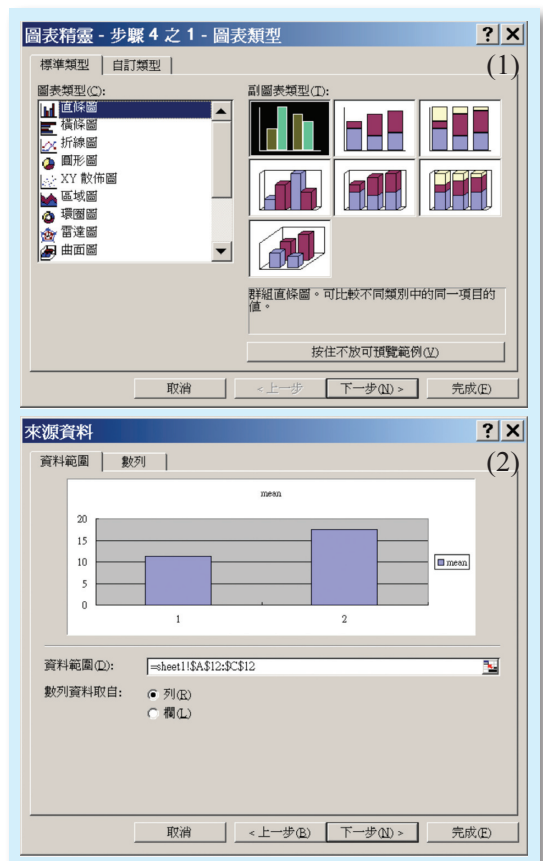
1. 進入圖表精靈，指定直條圖類型(圖三-(1))。
2. 按下一步，進入「來源資料」。
3. 然後按一下「資料範圍」索引標籤。

4. 在對話方塊中的「資料範圍」區域，指定平均值所在位址（圖三-(2)），由於本例同時將命名為mean的資料

	A	B	C
1		處理1	處理2
2		10	15
3		11	15
4		10	22
5		15	15
6		8	22
7		8	25
8		15	13
9		10	12
10		11	14
11		15	23
12	mean	=AVERAGE(B2:B11)	=AVERAGE(C2:C11)
13	SD	=STDEVA(B2:B11)	=STDEVA(C2:C11)
14	SE	=B13/SQRT(COUNT(B2:B11))	=C13/SQRT(COUNT(C2:C11))
15	95%CI	=1.96*B14	=1.96*C14

12	mean	11.3	17.6
13	SD	2.750757471	4.812021982
14	SE	0.86986589	1.521694961
15	95%CI	1.704937144	2.982522124

圖二、利用Excel函數和公式計算平均值、SD、SE及CI的方法。



圖三、Excel「圖表精靈」中(1)指定圖表類型，(2)來源資料中「資料範圍」索引標籤的方塊選項。

標記連同平均值一起選入作為資料範圍(A12:C12)，所以「mean」文字可在圖例中顯示出來，倘沒有將標記選入作為資料範圍而僅指定平均值位址(B12:C12)，則圖例文字僅以內設的「數列」文字顯示。

5. 在對話方塊中的「數列資料取自」，由於本例的兩個平均值是逐列呈現，因此仍維持系統內設的「列」(若數列資料是逐欄呈現，則此項必須勾選「欄」)。最後，按「下一步」或「完成」。

(二)加入誤差線

1. 在圖上之資料點按一下，以選取該資料數列。
2. 按右鍵或從功能表選擇「格式」，進入「資料數列格式」。
3. 然後按一下「Y誤差線」索引標籤。
4. 在對話方塊中的「顯示」區域，按一下所需要的誤差線類型，如圖四所示。誤差線能顯示實際資料點的正差、負差或兩者，若選擇無，則取消誤差線。
5. 在對話方塊中的「誤差量」區域，選項有定值、百分比(資料點的數值)、標準差、標準誤差或自訂(手

動指定數量)，可選擇其中一種來計算誤差線描繪出的數量。但此處請務必勾選「自訂」，並指定要選擇的誤差量所在範圍位址，如本例，若要繪製SD誤差線，則在+和-區域內都選擇SD數值所在位址B13:C13(圖四)，按「確定」後，即在長條圖中新增SD誤差線，如圖一-(1)所示。若繪製SE或CI誤差線，同樣地在「自訂」的+和-區域內改成其數值所在位址即可，結果分別列於圖一-(2)、(3)。注意，無論使用SD或SE時，都切勿勾選方塊區域內的標準差或標準誤差，否則會出錯。

(三)修改誤差線類型

1. 在圖上的誤差線按一下，以選取該誤差線。
2. 按右鍵，進入「誤差線格式」。
3. 然後按一下「Y誤差線」索引標籤。在「自訂」的+和-區域內，重新指定其他誤差量所在範圍位址，然後按「確定」。

(四)取消誤差線

1. 在圖上的誤差線按一下，以選取該誤差線。
2. 按右鍵，選擇「清除」。



圖四、Excel「資料數列格式」中之「Y誤差線」索引標籤的方塊選項(圖中「顯示」區域選擇兩者，並在「誤差量」區域內點選自訂後指定自訂值所在之範圍位址)。

五、參考文獻

- 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(一)－敘述統計量。技術服務 63：33-36。
- Streiner, D. L. 1996. Maintaining standards: differences between the standard deviation and standard error, and when to use each. *Can. J. Psychiatry* 41:498-502.