

運用 ISO 13528 於能力試驗結果統計之效益

財團法人台灣電子檢驗中心/黃凱斌博士、實驗室認證處二處/鄭嘉芬

能力試驗活動主要藉由實驗室間的比對來評估參加者(如實驗室或檢驗機構等)於特定測試/校正範圍的技術能力表現。實驗室除可藉由參加能力試驗評估實驗室的技術能力、監控實驗室的持續表現、鑑別問題、建立方法的有效性、鑑別實驗室間差異以及提供顧客更大的信心，同時亦能滿足認證組織或權責機關對於實驗室展現其技術能力的要求。然而，一個優質的能力試驗活動，多數仰賴一個具有能力的能力試驗執行機構(Proficiency Testing Provider)的辦理，而能力試驗執行機構於提供具有信賴結果的能力試驗，則需仰賴一個可靠、適用及嚴格理論基礎的統計設計與分析技術。

ISO 13528:2015 Statistical methods for use in proficiency testing by interlaboratory comparison 在實驗室間比對用於能力試驗的統計方法，此國際標準是能力試驗執行機構作為規劃能力試驗活動過程，評估待測件均勻性與穩定性評估、指定值評價以及評估參加者於結果表現的國際共識標準。能力試驗執行機構藉由 ISO 13528 內統計方法的建議，讓測試樣品的均勻性與穩定性的評估更具代表性，指定值的選擇更適合此能力試驗的目的，同時也提供參加者於結果表現的判定，更客觀精準。

對於能力試驗執行機構運作的要求，國際間另有一份國際標準 ISO/IEC 17043:2010 Conformity assessment -- General requirements for proficiency testing 能力試驗執行機構一般要求，其內容中即直接引用 ISO 13528，作為能力試驗執行機構符合 ISO/IEC 17043:2010 標準之能力試驗的統計分析的要求。隨著 ISO 13528 於 2015 年更新與改版，瞭解此份標準對於我國能力試驗執行機構於推動能力試驗活動與提供更具有統計意義的能力試驗活動尤為重要。因此，本會於 2016 年即召集國內相關能力試驗執行機構專家、權責機構代表、技術專業團體，共同成立工作小組探討此份文件，並於 2017 至 2018 年陸續完成 ISO 13528 內涵討論與中文化作業，以期讓能力試驗執行機構認證更滿足使用者需求與符合國際標準。



講師 黃凱斌博士



活動照片(10.12.新竹)

為協助我國能力試驗執行機構對於 ISO 13528:2015 更為熟悉，本會於 2018 年 10 月 12 日假竹科集思會議中心辦理一場「2018 年能力試驗執行機構主管暨評審員/技術專家在職訓練」，並邀請黃凱斌博士(作者之一)講授 ISO 13528 能力試驗之統計方法與應用。課程中除說明此標準的原則內容，亦加入測試領域及校正領域的能力試驗案例討論，協助參與的能力試驗執行機構主管及能力試驗執行機構評審員/技術專家，更為了解此標準的主要改變內容，期望由說明與演練的方式，提升本會能力試驗執行機構評審員/技術專家與能力試驗執行機構之能力，更藉由共同與會的方式達到參與者對於此標準應用的一致性，以共同提升認可能力試驗執行機構於未來辦理能力試驗活動的品質。本次活動共有能力試驗執行機構代表、能力試驗執行機構之評審員、技術專家以及本會同仁計 64 人踴躍的參與。由當天出席者熱烈的討論與各領域的能力試驗執行機構與專家提出不同的看法，已達到同儕間互相學習與瞭解的成效。

以下為

ISO 13528 實驗室能力試驗計畫之統計方法摘要介紹

(註:ISO 13528:2015 版權為 ISO 國際標準組織所有，本會不會提供此文件或相關中文文化翻譯，請有興趣者可自行至 ISO 官網購買下載。)

ISO 13528 實驗室能力試驗計畫之統計方法摘要介紹

財團法人台灣電子檢驗中心/黃凱斌博士

ISO 13528 編寫計畫最初於 1997 年被提出，並於 2005 年正式發行，為執行能力試驗活動過程中須使用統計技術時，如均勻性測試、穩定性測試、計算指定值(assigned value)、計算數據之離散程度及最後決定實驗室量測能力的表現統計量(performance statistic)，可依據的需求文件。ISO 13528 亦被 ILAC 會員國家或地區所廣泛採用。目前 ISO 13528 最新版本為 2015 年版(ISO 13528:2015)，相較於 2005 年版本的標準，主要有以下變更：

- 滿足 ISO/IEC 17043：2010 的新要求
 - 能力試驗計畫設計
 - 定性分析 (Qualitative analysis)
 - 應用範圍擴大 (inspection, sampling, individual)
- 重新安排本文及範例
- 增加指引 (guidance)，減少要求 (requirements)
- 刪除過時或無效的方法
- 新的符號 (New Symbols)

ISO 13528:2015 內容主要有以下部分：

- 能力試驗計畫統計設計原則
- 能力試驗待測件與結果初步審查準則
- 判定指定值與其標準不確定度
- 表現評估判定準則
- 計算表現統計量
- 表現分數說明圖示法

章節安排原則上乃依據能力試驗執行過程中需使用統計方法的工作依序排列，包含能力試驗待測件傳遞前後的均勻性及穩定性試驗、數據結果回傳後的指定值及其對應之不確定度計算、回傳數據的離散趨勢計算，最後結合指定值、不確定度及離散趨勢估計值(能力試驗標準差)、計算最後的表現統計量，如 $z - score$ 及 E_n 值等，最後則輔以圖表以更直覺確認各參加者的數據狀況，例如以尤頓圖(Youden plot)判斷量測誤差來自隨機誤差或系統誤差。

ISO 13528 主要的統計方法應用程序：

一、 能力試驗待測件之均勻性測試程序

1. 均勻性測試之一般程序及分析方式

對於大量製備之能力試驗待測件，一開始先決定所欲評估待測件之一項或多項特性(property or properties)或受測量(measurand(s))並且選擇一實驗室(通常是能力試驗執行機構之中心實驗室)使用規定之量測方法執行樣品均勻性判定。而利用所選之量測方法進行量測時，應確保量測之重複性標準差(s_r)需充分的小，IUPAC Harmonized Protocol 建議重複性標準差與能力試驗評定標準差(σ_{pt})之比值應小於0.5以確保所使用之量測方法精密度及確保後續計算之樣品間(Between-sample)標準差值較不會受樣品內差異的影響，進而可忠實反映樣品間之差異。而當上述之條件無法滿足時，應藉由增加重複量測次數以達到所需之精密度。另外，在製備能力試驗待測件時，應確保所製備之待測件數量足夠滿足均勻性測試及發送給各參加實驗室。而若可取得上一回合之能力試驗計畫中，適合的均勻性資料(如使用相同程序製備之特性相似待測件所測得的數據)，則可減少在當次能力試驗計畫之均勻性數據，也就是可減少用來執行均勻性測試之待測件數目。

而在均勻性測試之統計設計方面，首先從被包裝成最終形式之待測件中，隨機選取 g 個待測件($g \geq 10$)，接者將每個待測件各製備成 m ($m \geq 2$)個子樣品，要注意的是，需採用適當的方法使每個待測件的 m 個子樣品間的差異達到最小。接下來將 $g \times m$ 個樣品做隨機化排序，並在重複條件下量測每個子樣品以得到量測結果值，有了 g 組不同待測件及 m 次重複之量測結果，便可計算一般平均數($\bar{X}$)、樣品內標準差(within – sample standard deviation, S_w)及樣品間標準差(between – sample standard deviation, S_s)，其中 S_w 及 S_s 之計算乃採用變異數分析(Analysis of Variance, ANOVA)的概念計算，詳細計算流程如下：

首先定義：

$X_{t,k}$ ：執行均勻性測試之數據

t ：表示能力試驗待測件， $t = 1, 2, \dots, g$

k ：表示各能力試驗待測件之子樣品， $k = 1, 2, \dots, m$

另外定義各能力試驗待測件之平均數及變異數如下式(1)及式(2)

$$\bar{X}_t = \frac{\sum_{k=1}^m X_{t,k}}{m} \quad (1)$$

$$S_t^2 = \frac{\sum_{k=1}^m (X_{t,k} - \bar{X}_t)^2}{m} \quad (2)$$

每一待測件之組內變異數估計值則可表示如下式(3)：

$$w_t^2 = \frac{\sum_{k=1}^m (X_{t,k} - \bar{X}_t)^2}{m-1} \quad (3)$$

由式(1)可求得所有樣品之總平均值如下式(4)

$$\bar{X} = \frac{\sum_{t=1}^g \bar{X}_t}{g} \quad (4)$$

樣品平均之變異數估計量及樣本內變異數(within-sample variance)，則分別如(5)及(6)所示

$$S_X^2 = \frac{\sum_{t=1}^g (\bar{X}_t - \bar{X})^2}{g-1} \quad (5)$$

$$S_w^2 = \frac{\sum_{t=1}^g S_t^2}{g} \quad (6)$$

S_s 及 S_w 之組合變異數估計量則如(7)所示

$$S_{s,w}^2 = \frac{\sum_{t=1}^g (\bar{X}_t - \bar{X})^2}{g-1} + \left(1 - \frac{1}{m}\right) S_w^2 = S_s^2 + S_w^2 \quad (7)$$

最後，可求得樣本間之變異數估計量如(8)所示：

$$S_s^2 = S_{s,w}^2 - S_w^2 = \frac{\sum_{t=1}^g (\bar{X}_t - \bar{X})^2}{g-1} - \frac{1}{m} S_w^2 \quad (8)$$

要注意的是，在某些情況下若 $S_s^2 < 0$ ，則式(8)應直接使用 $S_s^2 = 0$ 。

2. 均勻性測試之評估法則

2-1 均勻性分析有效性之評定法則

以下三個均勻性評定法則可用來確保當執行均勻性分析時，測試資料對於分析方法是有效的：

- 將量測結果排序後，檢視各待測件下，各子樣品之量測結果是否有趨勢產生，若量測結果有出現明顯的趨勢或飄移，則需對量測方法採取適當矯正措施，或是在判讀結果上須謹慎。
- 依生產順序檢視能力試驗待測件結果之平均數；如果有嚴重趨勢造成能力試驗待測件超出如下第(9)式所述準則，或者阻礙能力試驗待測件之使用時，則(i)為各項能力試驗待測件指定個別值；或(ii)廢棄受到嚴重影響的能力試驗待測件子集，重測剩餘部分以確保具備足夠均勻性；或(iii)該趨勢影響到全體能力試驗待測件時，依第 B. 2. 4 節條款進行。
- 比較重複測試結果之間的差異(重複測試次數為 2 次以上時，則使用全距進行比較)，必要時，可使用 Cochran 檢定來檢視重複測試間是否有統計上的顯著差異(參考 ISO 5725-2)。倘若檢定結果發現有任何一對重複測試間有明顯大的差異，則應從專業技術的觀點上解釋為何造成該差異，若情況許可，在執行分析時，排除該組離群數據。若重複次數 $m > 2$ 且因單一離群值造成整體數據有較大的變異數時，則將該離群點從數據中移除，但要注意的是，從原有一組數據中移除單一離群結果，後續計算 S_w^2 與 S_s^2 時，需要考量據此產生的數據不平衡性，也就是每一組待測件最後的重複次數將不全相同。

2-2 均勻性判定準則

以下第(9)式的判定準則，乃藉由比較樣本間標準差 S_s 及能力試驗標準差 σ_{pt} 以評估能力試驗待測件的均勻性，若第(9)式成立，則可判定待測件據均勻性：

$$S_s < 0.3\sigma_{pt} \quad (9)$$

2-3 無法事先得知能力試驗標準差 σ_{pt} 之做法

在執行能力試驗計畫時，有些情況無法事先得知能力試驗標準差 σ_{pt} ，例如當 σ_{pt} 為由參加者回饋之結果所計算之穩健標準差時，此時能力試驗執行機構應選擇其他準則，在待測件發送或傳遞前，判定是否具備足夠均勻性。此類程序應包括：

- 進行變異數分析(Analysis of Variance, ANOVA)，以 F -檢定檢定能力試驗待測件間是否有統計上的顯著差異，ANOVA 檢定過程不需事先得知能力試驗標準差 σ_{pt} 的資訊。但要留意，使用 F -檢定可能會產生不敏感(incentive test)會過於敏感(too insensitive)之檢定的問題，例如若樣本內標準差過大時，可能會使得較大的樣本間差異無法被偵測出來，或是當樣本內標準差過小時，則會將樣本間標準差小的情況判定為不具均勻性。上述情況將造成檢定結果與實際情況不一致。
- 使用先前能力試驗計畫的資訊來估計 σ_{pt} 。
- 使用由精密實驗得來的資料(例如 ISO 5725-2 所提的再現性標準差)。
- 接受因發送均勻性不足之能力試驗待測件所伴隨的風險，並且需在後續量測完成，且能力試驗標準差 σ_{pt} 已計算出後，利用第(9)式評估待測件是否具均勻性。

2-4 當均勻性判定結果為不具均勻性時之做法

當發現均勻性判定結果不滿足均勻性準則(第(9)式)時，能力試驗執行機構應考慮採取下列任一行動。

- 依照方程式(B.3)計算 σ'_{pt} ，進而在能力評定標準差納入樣本間標準差。請留意這需要向參加者充分說明。

$$\sigma'_{pt} = \sqrt{\sigma_{pt}^2 + S_s^2} \quad (10)$$

- $S_s < 0.3\sigma_{pt}$ 將 S_s 放入指定值不確定度，並使用 z' 來評定表現。
- σ_{pt} 是參加者結果穩健標準差時，將能力試驗待測件間不均勻性放入 σ_{pt} ，放寬均勻性接受基準，惟需要特別留意。

上述 a)至 c)點皆不適用時，則廢棄能力試驗待測件，矯正造成不均勻性之原因後，重新製備樣品。

若從先前實驗性研究、經驗或認知得以合理保證，不可能發生不穩定情況時，穩定性實驗性檢查應限制在檢查能力試驗循環期間的顯著變化，而且在循環期間與結束後進行。至於其他情況，運輸效應與標準能力試驗循環期間穩定性研究，可在發送能力試驗待測件之前，針對各循環或初期規劃及可行性研究期間，採取規劃研究形式進行，藉此制訂一致性運輸與儲放條件。能力試驗執行機構亦可檢查回報結果，是否浮現符合量測日期的趨勢，從中檢查不穩定性證據。

二、 能力試驗待測件之穩定性測試程序

倘若從先前實驗性研究、經驗或認知得以合理保證，不可能發生不穩定情況時，穩定性實驗性檢查應限制在檢查能力試驗循環期間的顯著變化，而且在循環期間與結束後進行。至於其他情況，運輸效應與標準能力試驗循環期間穩定性研究，可在發送能力試驗待測件之前，針對各循環或初期規劃及可行性研究期間，採取規劃研究形式進行，藉此制訂一致性運輸與儲放條件。能力試驗執行機構亦可檢查回報結果，是否浮現符合量測日期的趨勢，從中檢查不穩定性證據。

三、 能力試驗指定值及不確定度的決定

在能力試驗計畫中，通常會由實驗室之數據結果與指定值之比較的概念出發，再透過適當的公式計算出評估能力試驗結果表現之統計量。而由於指定值為一估算出的值，因此通常也需估算其不確定度以作為當使用此指定值時，對於其變異程度的了解。底下為求取能力試驗指定值及其不確定度的方法。

1. 公式方法

1-1 公式方法指定值之決定

所謂的公式方法指的是測試物質乃由特定比例的混合成分所組成，或是由添加特定比例的成分至某一基質中得到。此法中的指定值乃由所使用的各成分之量值(例如混凝土成分的測試，可先計算水、泥砂及水等各成分之重量)所計算而得。

1-2 公式方法指定值不確定度之決定

當指定值是由測試物質的配方(formulation)計算而得時，其標準不確定度之估算，乃依據 ISO/IEC Guide 98-3:2008 所描述的方法，將所有不確定度分量結合起來計算而得。例如，在化學分析中，不確定度通常與重量及體積等不確定度分量有關。

2. 使用驗證指定值

2-1 驗證指定值之決定

若能力試驗計畫中，使用到的物質為驗證參考物質(Certified Reference Material, CRM)，則該參考物質的驗證指定值即可作為指定值。

2-2 驗證指定值不確定度之決定

驗證參考物質的證書會提供此參考物質之不確定度相關資訊，因此若使用驗證參考物質做為測試物質，其標準不確定度可由此驗證參考物質之證書的資訊得到。

3. 專家實驗室的共識值

3-1 利用專家實驗室共識值作為指定值之決定

此法需先準備好多份用來做為參考物質(Reference Material, RM)之測試物質的樣品(此樣品預備用來做為能力試驗之傳遞樣品)，接者從這些樣品中隨機抽樣選取一些樣品，並由一組專家實驗室對這些樣品進行分析，這組專家實驗室可以是某一回合能力試驗計畫的參與者，在這一回合之能力試驗活動結束後，由此回合所得到之實驗室數據，確定指定值及其不確定度。指定值之不確定度的推導方式可由所有實驗室測試結果所計算出的穩健平均數得到，關於穩健平均數的計算方式可參考 ISO 13528 附錄 C 所介紹的演算法 A。除了上述的演算法 A，也可使用其他方法代替演算法 A，只要該方法有可靠的

統計理論基礎並須在能力試驗報告中描述其使用方法即可，例如使用一組數據之中位數。

3-2 利用專家實驗室共識值作為指定值不確定度之決定

當蒐集到由 n 個專家實驗室所提供之測試物質的量測值 X_i 及量測標準不確定度 u_i 的估計值後，其指定值 X 可由 *ISO 13528* 附錄 C 所介紹的演算法 A 求得，指定值的標準不確定度則可由下式求得：

$$u_X = \frac{1.25}{n} \times \sqrt{\sum_{i=1}^n u_i^2}$$

使用專家實驗室的共識值作為指定值方法的局限在於，專家實驗室的結果數據可能會存在未知的偏誤，且其所宣告的不確定度可能是不可靠的。

4. 參加者的共識值

4-1 利用參加者共識值作為指定值之決定

此法是將某一回合能力試驗計畫中所有參加者的結果數據之穩健平均數做為所使用測試物質的指定值。穩健平均數的計算方式可參考 *ISO 13528* 附錄 C 所介紹的演算法 A。除了上述的演算法 A，也可使用其他方法代替演算法 A，只要該方法具可靠的統計理論基礎並須在能力試驗報告中描述其使用方法即可。此方法若在量測方法已標準化的情況下使用將會特別有幫助。

4-2 利用參加者共識值作為指定值不確定度之決定

當指定值 X 是由 *ISO 13528* 附錄 C 所介紹的演算法 A 求得的穩健平均值，則指定值的標準不確定度則可由下式求得：

$$u_X = 1.25 \times \frac{S^*}{\sqrt{n}}$$

其中， S^* 乃利用 *ISO 13528* 之附錄 C 所介紹的演算法 A 求得。

須注意的是，當使用參加者的共識值作為指定值的方法時會有下列的侷限：

1. 參加者中也許並沒有真正的共識值。
2. 因量測方法本身有誤差而造成共識值會有偏差，而此偏差並不會反應在共識值中。

五、能力試驗評估標準差的方法

標準差可用來衡量一組數據的變異性(分散性)，利用標準差可用來衡量一組數據的不確定性及幫助判定數據中是否存在離群值。標準差也是能力試驗中用來計算參加者表現指標的一項重要因子。在此部分，我們將介紹五種決定能力試驗中，參加者數據之標準差的方法。

1. 專家實驗室之認知

能力評定之最大允許誤差或標準差，可設定為主管機關、認證機構或能力試驗執行機構技術專家公認對參加者合理之對應到表現水準的特定值。此法的優點在於標準差乃直接與量測方法之適用性的聲明直接關聯。

2. 根據能力試驗計畫先前循環之經驗

參加者使用相容之量測程序時，能力試驗評定標準差(σ_{pt})與最大允許誤差(δ_E)，可透過先前能力試驗活動中使用相同受測量及相似特性值的經驗來決定。當專家對適用性沒有共識時，此法尤為實用。此法的優點如下：

- 基於合理的預期表現統計量來評估結果。
- 評估準則不會隨著每一能力試驗回合中，參加者結果構成的母體中的隨機變異或參加者母體的改變而變化。
- 倘若測試或校正領域，有兩家以上的合格能力試驗執行機構時，評估準則不會隨能力試驗執行機構而異。

此外，在檢視先前能力試驗計畫回合時，應包括考量從合格參加者所得到的表現成果，而且不受新進參加者、群組規模較小、或特定能力試驗回合的罕見因素所引起的隨機變異性影響。可藉由檢視先前能力試驗計畫回合的一致性做主觀判定或藉用平均數或調整受測量之值的迴歸模型客觀判定。迴歸方程式則可能是直線或曲線。標準差與相對標準差之選擇，應考量何者在適當的受測量水準範圍內較具一致性。合適的最大允許誤差亦能以此方式取得。

而當依據能力試驗計畫先前試驗的共識統計量作表現評估標準時，應使用標準差的穩健估計量，其計算可參考 *ISO 13528* 附錄 C 所介紹的演算法 A。除了演算法 A，也可使用其他方法代替演算法 A，只要該方法具可靠的統計理論基礎並須在能力試驗報告中描述其使用方法即可。需注意的是，利用此法會使得在每一回合之能力試驗計畫可能會有顯著不同的標準差，因此若要利用實驗室之 Z-Score 來觀察幾個回合中的能力試驗計畫之可能趨勢時會有一定的困難度，但此缺點可利用在數個回合之能力試驗計畫中所得到的標準差的穩健聯合值 (Robust Pooled Value) 來解決，穩健聯合值可由 *ISO 13528* 附錄 C 所介紹的演算法 S 求得。

3. 由一般模型決定

能力試驗評定的標準差可由某一量測方法的再現性之一般模型求出。而因為一般模型假設再現性只依賴於被測量的水準，而不依賴於被測量本身、量測過程或樣本數，因此會造成某特定量測方法的真實再現性可能與一般模型所得到的數據有顯著差別。以下為一個由一般模型決定標準差的例子。

$$\sigma_R = \begin{cases} 0.22c, & c < 1.2 \times 10^{-7} \\ 0.22c^{0.8495}, & 1.2 \times 10^{-7} \leq c \leq 0.138 \\ 0.01c^{0.5}, & c > 0.138 \end{cases}$$

Horwitz 提出了一個化學分析方法之再現性的模型，此方法可調用以下再現性標準差的數學公式：

$$\sigma_R^2 = 0.02c^{0.3495}$$

其中 c 為以百分比表示的待測化學成分之濃度。

4. 從先前精密量測法之協同研究使用重複性與再現性標準差

當能力試驗計畫使用以標準化的量測方法，且該方法之重複性與再現性為可獲得時，則能力試驗評定的標準差可以計算如下：

已知：

σ_R ：再現性標準差，

及

σ_r ：重複性標準差，

則可得能力試驗評定的標準差公式如下：

$$\sigma_{pt} = \sqrt{\sigma_R^2 - \sigma_r^2(1 - 1/m)}$$

其中 m 為每一能力試驗計畫執行回合中，每一參與實驗室的重複量測次數。

六、能力試驗評估數據表現之統計量

此部分將介紹，在能力試驗活動中，用來評估參加者表現的一些常用的指標。這些指標中，將會使用到前面所介紹的各種統計量。

1. 偏差 (bias)

假設 x 表示某一回合之能力試驗計畫中參加者對測試物質之某種特性的報告結果，則當實驗室量測該特性時，其量測偏差之估計值可以如下公式表示：

$$D = x - X$$

其中 X 為指定值。

2. Z分數 (Z-Score)

Z分數可用來評估每一實驗室之結果表現，其公式如下：

$$Z = \frac{(x - X)}{\hat{\sigma}}$$

其中，

x = 個別實驗室的測試結果

X = 指定值

$\hat{\sigma}$ = 能力試驗評定之標準差

要判定參與之實驗室是否為異常實驗室，可由以下Z值之判定準則來決定：

$|Z| \geq 3$ 實驗室之數據結果為異常，應立即採取異常處理行動。

$2 < |Z| < 3$ 實驗室之數據結果雖無顯著異常，但已出現警告訊號。

$|Z| \leq 2$ 實驗室之數據結果正常。

要特別說明的是，在 NATA(澳洲國際測試機構協會)所介紹之用於能力試驗的相關統計方法中，其使用Z分數時，乃將參與實驗室所回饋之數據之中位數做為指定值，標準差則使用標準化四分位數差(normalized IQR)來估計。標準化四分位數差之公式如下：

$$\text{normalized IQR} = (Q_3 - Q_1) * 0.7413$$

其中

Q_3 = 第三四分位數，

Q_1 = 第一四分位數。

3. E_n 值 (E_n Ratio or E_n Number)

E_n 值之 E_n 為 *Error normalised* 之縮寫，其公式如下：

$$E_n = \frac{(x - X)}{\sqrt{U_{lab}^2 + U_{ref}^2}}$$

其中，

x = 個別實驗室的測試結果

X = 指定值 (參考實驗室之值)

U_{lab} = 參與實驗室之不確定度

U_{ref} = 指定值之不確定度

要判定參與之實驗室是否為異常實驗室，可由以下 E_n 值之判定準則來決定：

$|E_n| \geq 1$ 實驗室之數據結果可能與指定值不一致。

$|E_n| < 1$ 實驗室之數據結果與指定值一致。

如同 Z 分數， E_n 值亦可用來評估實驗室之結果表現，不同的是， Z 分數衡量的是實驗室之量測數據與指定值的接近程度，其通常用於測試領域，而 E_n 值通常用於校正領域。 E_n 值指出了一般實驗室之數據與參考實驗室之數據的差值是否小於其兩者之不確定度的特定組合。因為 E_n 值並非全然衡量實驗室之數據與指定值的差異程度，因此有較低不確定度的實驗室，其 E_n 值也可能會與有較大不確定度的實驗室相似。當 $|E_n| \geq 1$ 時，實驗室可能須檢視其所宣告不確定度之合理性或需修正量測問題。

結論

本文簡單介紹並分析 ISO 13528:2015 所介紹之用於能力試驗的相關統計方法，主要著重於用來評估能力試驗之實驗室表現統計量的介紹。能力試驗活動的目的為藉由實驗室間的相互比對，讓實驗室能夠了解本身技術能力與其他實驗室是否有差異。而用來評估實驗室之能力的指標，則涉及了相關統計方法的應用，因此能力試驗執行機構應了解對於不同性質的能力試驗，其所適用的統計方法為何。另外，透過本文也提到能力試驗指定值及不確定度的決定、能力試驗評估標準差的方法及表現統計量的各式統計分析方法。ISO 13528 除可讓能力試驗舉辦機構能正確使用適當的統計方法於不同之特性的樣品，並正確估算出能力試驗評估數據表現之統計量，而參與能力試驗活動之實驗室對於能力試驗報告中之統計結果也能具備正確判讀之能力。

參考文獻

1. ISO 13528:2015 Statistical methods for use in proficiency testing by interlaboratory comparisons.
2. GUIDE TO NATA PROFICIENCY TESTING (2004) Ver.1
3. ISO/IEC Guide 98-3:2008 Uncertainty of measurement -- Part 3: Guide to the expression of uncertainty in measurement (GUM:1995)

4. ISO 3534-2:2006 Statistics -- Vocabulary and symbols -- Part 2: Applied statistics.