中國統計學報
第 42 卷 第 4 期
九十三年十二月
349-364 頁

測量設計之評介

林彩玉　　廖振鐸
台灣大學農藝所生物統計組

摘　要

"測量"(*measurement*)被廣爲使用於一般的生活中，其主要目的爲估計未知品(*unknowns*)的特徵值(*characteristic*)。然而，在測量的過程中通常會受到誤差(*errors*)的干擾。因此，在實驗室中或工業產品的產製過程之測量，通常會利用標準品(*standards*)來估計誤差，進而來校準未知品的測量值。本文主要由實驗設計(*experimental designs*)的角度來介紹有關測量過程的研究。首先，介紹測量過程中常見的兩種統計模型：累加性模型(*additive models*)及線性校準模型(*linear calibration models*)。再針對不同的模型，介紹相關的實驗設計之研究。實驗設計的主要議題(*issues*)包含量測過程中標準品與未知品的配置(*allocation*)及排列方式(*arrangement*)。最後提出一些關於測量設計未來可能的研究方向。

關鍵詞：系統誤差，隨機誤差，校準模型，A-最適準則，非二元設計。

美國數學會分類索引：主要 62K99; 次要 62P30。

1. 簡介

"測量"(measurement) 這個動作不但被廣泛地運用於一般的生活中, 對於實驗室裡儀器的操作與工業上產品的製程方面, 更是一項不可或缺的行動。測量的主要目的在於估計樣品 (specimens) 或稱未知品 (unknowns) 的特徵值 (characteristic)。此處所稱的樣品或未知品是一般通用的說法, 並非僅限於實驗室裡的樣品。例如: 某人對西瓜的重量感興趣, 則此例的樣品是西瓜, 而重量就是特徵值。

測量的過程中通常會受到誤差的干擾。一般而言, 誤差來源可分爲系統誤差 (systematic error) 與隨機誤差 (random error) 兩類。亦即當我們進行測量時, 可能會受到隨機誤差或兩種誤差同時的影響。其中, 隨機誤差定義爲期望值 (expected value) 爲 0 的隨機變數 (random variable), 而系統誤差則定義爲量測過程中所產生的偏差 (bias), 由未知的參數 (parameters) 來描述。

在實驗室或工業產品產製過程進行測量時, 通常會經由校準 (calibration) 的過程來校正在測量過程中所產生之系統誤差, 並估計隨機誤差的大小。因此, 在測量的過程裡, 除了上述所提及的未知品外, 亦包括標準品 (standards) 或稱校準品 (calibrates)。標準品的眞值 (true value) 是已知 (known) 的, 一旦量測了標準品, 則測量過程中的誤差即可被觀測到。換言之, 適當地放置標準品於測量的過程可以觀控 (monitor) 測量誤差的大小。標準品在測量過程中, 扮演著校正誤差的角色。亦即, 我們不僅可以藉由校準過程對系統誤差做估計, 進而亦可修正系統誤差對未知品的眞值估計所造成的偏差。

過去文獻關於測量的研究大都著重在估計 (estimation) 方面, 主要的研究結果可參閱 Fuller (1987) 與 Brown (1993)。然而, 從實驗設計 (experimental designs) 的角度著眼的文獻並不多。這裡所謂的實驗設計, 可歸納爲量測過程中標準品與未知品的配置 (allocation) 及排列方式 (arrangement) 等問題。一般實驗設計的研究大多數都建構在統計模型 (statistical models) 上, 因此我們將文獻上的研究, 依研究者所使用的統計模型, 區分成兩部份來討論: (1) 累加性模型 (additive models) 方面有 Pepper (1973) 與 Liao, Taylor and Iyer (2000) 的研究; (2) 線性校準模型 (linear calibration models) 的研究, 則參考 Perng and Tong (1977) 及 Lin and Liao (2004)。

下一節, 我們先對測量過程中常用的統計模型做介紹。第三節及第四節, 將

針對量測過程中標準品與未知品的配置與排列方式這兩個問題, 分別進行說明與討論。最後, 提出一些有關測量設計之未來可能的研究方向。

2. 測量模型介紹

我們針對過去文獻中較常見的兩種測量模型: 累加性模型及線性校準模型, 分別介紹如下。

2.1 累加性模型

一般而言, 對於某個受測品 (未知品或標準品) 的測量值 (measured value) 或稱觀測值 (observed value) 可利用下列簡單的累加性模型來表示:

$$y_i = \mu_i + \tau + \epsilon_i. \tag{2.1}$$

其中 μ_i, τ, ϵ_i 分別代表系統誤差, 受測品的眞值及隨機誤差。

Pepper (1973) 考慮一個不包含系統誤差的累加性模型:

$$y_i = \tau + (b_i + \eta_i). \tag{2.2}$$

式中 $(b_i + \eta_i)$ 爲隨機誤差, b_i 是來自隨機漫步過程 (random walk process), $b_i = b_{i-1} + \epsilon_i$, 且 $\epsilon_i \sim N(0, \sigma_\epsilon^2)$, η_i 爲平均數 0 且有共同變異數 σ_η^2 的常態隨機變數, 又隨機變數間假設爲相互獨立。另外, Liao, Taylor and Iyer (2000) 則考慮在 N 個總測量次數下的單種標準品與 m 種未知品的累加性模型:

$$y_i = \mu + \delta_{i0}\tau_0 + \sum_{j=1}^{m} \delta_{ij}\tau_j + \epsilon_i, \qquad i = 1, 2, \ldots, N. \tag{2.3}$$

模型中的 μ 是系統誤差; τ_0 表標準品的已知眞值; $\tau_1, \tau_2, \ldots, \tau_m$ 分別爲 m 種未知品的眞值; ϵ_i 則是來自平穩型一階自我迴歸過程 (stationary first order autoregressive process, AR(1)), $\epsilon_i = \phi\epsilon_{i-1} + z_i$, 其中 z_i 假設爲平均數 0 且同質變異數 σ^2 的互不相關 (uncorrelated) 的隨機變數, ϕ 爲 AR(1) 的參數; 當第 i 次測量爲標準品, 則指示變數 (indicator variable) δ_{i0} 爲 1, 否則爲 0; 當第 i 次測量爲第 j 種未知品, 則指示變數 δ_{ij} 爲 1, 否則爲 0。他們採用使未知品的眞值之最佳線性不偏估計式 (best linear unbiased estimators, BLUE) 的平均變異數達到最小的 A-最

適 (A-optimality) 準則來進行分析, 以得到最適平衡測量設計 (optimal balanced measurement designs)。值得注意的是, 他們所討論的問題同等於在一無區集的 (unblocked) 實驗設計中, 即單一區集中包含所有的測量, 比較實驗組與對照組實驗 (test-control experiments) 的問題。我們可以將 (2.3) 重新參數化 (reparameterization) 如下: 定義 $\theta_0 = \mu$ 與 $\theta_j = \mu + \tau_j$, $j = 1, 2, \ldots, m$, 且令 $z_i = y_i - \delta_{i0}\tau_0$。則 (2.3) 可重新表達爲

$$z_i = \sum_{j=0}^{m} \delta_{ij}\theta_j + \epsilon_i, \qquad i = 1, 2, \ldots, N. \tag{2.4}$$

由模型 (2.4) 可看出其爲一簡單之 $m+1$ 個處理的單向分類模型 (one-way classification model)。其中, 處理的平均分別爲 $\theta_0, \theta_1, \ldots, \theta_m$, 又 θ_0 可看成是對照組的平均, 而 $\theta_j - \theta_0$ 就是 τ_j, $j = 1, 2, \ldots, m$。然而, 相對於過去多數是針對二元區集設計 (binary block design) 進行之實驗組與對照組實驗的研究, 討論此種實驗設計並加入 AR(1) 到隨機誤差中的研究可參閱 Cutler (1993), 此處所論及的測量設計, 則是利用單一區集的非二元設計 (nonbinary designs)。

2.2 線性校準模型

另一種常看到的測量現象是線性系統誤差, 描述這類型測量過程最常見的統計模型如下:

$$y_i = \alpha + \beta\tau + \epsilon_i. \tag{2.5}$$

模型中的 y_i 爲測量值, α 和 β 皆表示系統誤差的參數, τ 是受測品的眞值, ϵ_i 則爲隨機誤差。例 2.1 (Hunter and Lamboy(1981)) 可說明這種現象的發生。

例 2.1. 某一化學家想建構一校準統計模型, 因此測量一組標準品中鉬 (Molybdenum) 的含量。實際的資料如表 2.1。

表 2.1 校準過程中鉬的實際資料

鉬的眞值	1.0	1.0	2.0	2.0	3.0	3.0	4.0	4.0	5.0	5.0
鉬的測量值	1.8	1.6	3.1	2.6	3.6	3.4	4.9	4.2	6.0	5.9
鉬的眞值	6.0	6.0	7.0	7.0	8.0	8.0	9.0	9.0	10.0	10.0
鉬的測量值	6.8	6.9	8.2	7.3	8.8	8.5	9.5	9.5	10.6	10.6

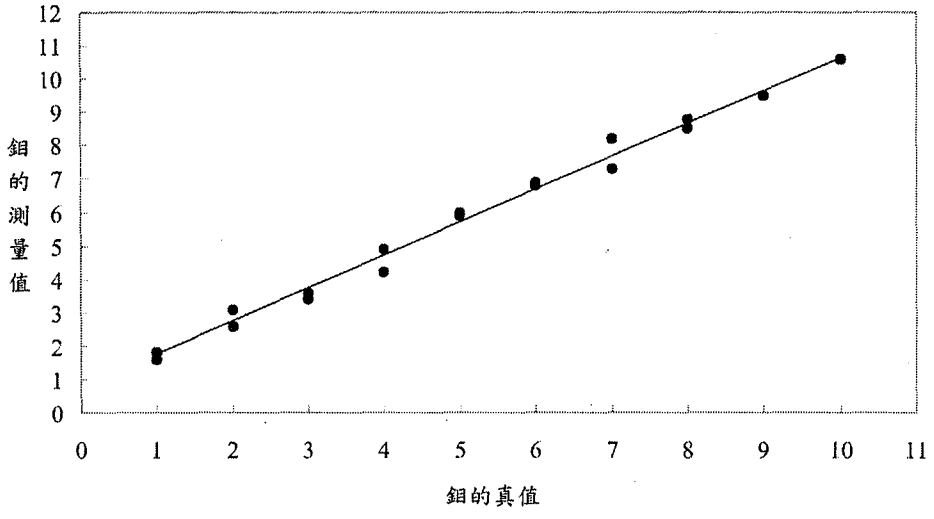

圖 2.1 鉬的眞值與測量值之散佈圖, 其中線性校準模型是利用最小平方估計(least square estimation) 所獲得。

由圖 2.1 可發現鉬的眞值與測量值間有線性關係存在, 此即反應了模型 (2.5)。

Perng and Tong (1977) 考慮線性校準模型 (2.5), 在一序貫 (sequential) 的測量過程中, 總測量次數 N 固定又足夠多 (large enough) 的條件下, 決定每次測量中該量測未知品或標準品, 以使得未知品的眞值之區間估計 (confidence intervals) 達到最高的精確度 (precision)。最近 Lin and Liao (2004) 也考慮模型 (2.5), 並架構在 A-最適準則的基礎上, 討論兩種不同標準品和 m 種未知品的最適配置問題, 他們的模型可以表達如下:

$$y_i = \alpha + \beta(\delta_i^{S_0} x_0 + \delta_i^{S_1} x_1 + \sum_{j=1}^{m} \delta_i^{U_j} \tau_j) + \epsilon_i, \qquad i = 1, 2, \ldots, N. \tag{2.6}$$

模型中的 x_0 與 x_1 分別爲兩種不同標準品 S_0 與 S_1 的眞值; τ_j 表未知品 U_j 的眞值, $j = 1, 2, \ldots, m$; 當第 i 次測量爲 S_k, 則指示變數 $\delta_i^{S_k}$ 爲 1, 否則爲 0, $k = 0, 1$; 當第 i 次測量爲未知品 U_j, 則指示變數 $\delta_i^{U_j}$ 爲 1, 否則爲 0, $j = 1, 2, \ldots, m$; 又隨機誤差 $\epsilon_i \overset{iid}{\sim} N(0, \sigma^2)$。

值得注意的是, 即使由最簡單的線性校準模型 (2.5) 來看, 實際上就已經是一個"非線性 (non-linear)"的問題, 即有兩個未知參數的乘積項 ($\beta\tau$) 在模型中。因此, 對於線性校準模型在參數的估計及設計的建構上, 相對地都比累加性模

型來得複雜。

3. 測量品的配置問題

本文所談及測量設計的配置問題, 其實和抽樣方法 (sampling) 中的樣本配置有異曲同工之妙。回顧抽樣調查設計的一個主要目的, 在限定成本的考量下, 對於有興趣估計的參數提供一最小變異數的估計式。當總樣本數確定後, 就有很多不同的方式進行樣本的配置。且每一種不同的配置方式, 會導致我們感興趣的估計式之變異數有所差異。同樣架構在這個想法下, 若考慮模型 (2.3) 且假設 $\epsilon_i \overset{iid}{\sim} N(0,\sigma^2)$ 的前提時, 則

$$Var(\hat{\tau}_j) = (\frac{1}{a_0} + \frac{1}{t_j})\sigma^2, \qquad j = 1, 2, \ldots, m. \tag{3.1}$$

其中, a_0 是標準品 S_0 的測量數, t_j 表第 j 種未知品 U_j 的測量數, 總測量數 $N = a_0 + \sum_{j=1}^{m} t_j$。由 (3.1) 得知, 當 N 固定時, 標準品與未知品的配置將會影響 τ_j 估值的精確度。其 A-最適測量設計爲

$$\begin{aligned} a_0^* &= \frac{N}{1+\sqrt{m}}, \\ t_j^* &= \frac{N}{\sqrt{m}+m}, \qquad j = 1, 2, \ldots, m. \end{aligned}$$

實際應用時, 通常會取 $\lfloor a_0^* \rfloor$ 及 $\lfloor t_j^* \rfloor$ 或 $\lfloor t_j^* \rfloor + 1$, 當成標準品與未知品的測量數。其中,$\lfloor a \rfloor$ 表示小於或等於 a 的最大整數。

若進一步考慮異質變異數的前提假設, 例如: 當測量標準品 S_0 時, $\epsilon_i \overset{iid}{\sim} N(0,\sigma_0^2)$; 當測量未知品 U_j 時, $\epsilon_i \overset{iid}{\sim} N(0,\sigma_j^2)$, 則

$$Var(\hat{\tau}_j) = (\frac{\sigma_0^2}{a_0} + \frac{\sigma_j^2}{t_j}), \qquad j = 1, 2, \ldots, m, \tag{3.2}$$

且 A-最適測量設計爲

$$\begin{aligned} a_0^* &= \frac{\sqrt{m}N\sigma_0}{\sqrt{m}\sigma_0 + \sum_{j=1}^{m}\sigma_j}, \\ t_j^* &= \frac{N\sigma_j}{\sqrt{m}\sigma_0 + \sum_{j=1}^{m}\sigma_j}, \qquad j = 1, 2, \ldots, m. \end{aligned}$$

因此, 測量品的配置會受到變異數 $\sigma_0^2, \sigma_1^2, \sigma_2^2, \cdots, \sigma_m^2$ 的影響。

另一方面, 就線性校準模型 (2.6) 來說, Lin and Liao (2004) 指出 A-最適配置的結果除了受到標準品與未知品的測量數影響外, 還會受到未知品的眞值 (τ_j) 影響, 此乃因爲線性校準模型本身是一個非線性模型。由模型 (2.6) 在同質變異數的前提假設下, 可求得未知品的眞值之最大概似估計式 (MLE) 的漸近變異數 (asymptotic variance) 如下:

$$Var(\hat{\tau}_j) = \frac{\sigma^2}{\beta^2}[\frac{1}{t_j} + \frac{a_0(x_0-\tau_j)^2 + a_1(x_1-\tau_j)^2}{a_0 a_1 (x_0-x_1)^2}].$$

其中, a_0 與 a_1 分別爲兩種不同標準品 S_0 與 S_1 的測量次數, t_j 則表示未知品 U_j 的測量數, $j = 1, 2, \ldots, m$, 且 $a_0 + a_1 + \sum_{j=1}^{m} t_j = N$。當 τ_j 固定時, 則局部 (locally) A-最適測量設計爲

$$\begin{aligned} a_0^* &= \frac{N\sqrt{\theta_1}}{\sqrt{\theta_0}+\sqrt{\theta_1}+m}, \\ a_1^* &= \frac{N\sqrt{\theta_0}}{\sqrt{\theta_0}+\sqrt{\theta_1}+m}, \\ t_j^* &= \frac{N}{\sqrt{\theta_0}+\sqrt{\theta_1}+m}, \qquad j = 1, 2, \ldots, m. \end{aligned}$$

其中, $\theta_0 = \sum_{j=1}^{m} \frac{(x_0-\tau_j)^2}{(x_0-x_1)^2}$ 及 $\theta_1 = \sum_{j=1}^{m} \frac{(x_1-\tau_j)^2}{(x_0-x_1)^2}$。此外, 他們還訴諸於貝氏 (Bayesian) A-最適測量設計來處理這個模型下測量品的配置問題。

4. 測量品的排列問題

在一般的應用上, 隨機誤差通常假設成互不相關的隨機變數。然而, 我們也常常遇到隨機誤差間存在著某種的相關性 (correlation) 結構。特別在每一次測量的間隔很小的測量過程中, 我們常會發現隨機誤差間存在著某種程度的相關性。例 4.1 的資料是由 Stewart Environmental Company in Fort Collins, Colorado, USA 所提供, 可支持上述的說法。

例 4.1. 水中含砷 (Arsenic) 濃度的測量實驗

這是一組藉由 AAS (Atomic Absorption Spectroscopy) 測量不同來源 (地區) 的地下水樣品中, 有毒重金屬物砷的含量。表 4.1 爲砷濃度的測量資料。其中, 測量型態裡的 CCV (Continuing Calibration Verification) 表示眞值爲 20 $\mu g/L$ 的標準品, 其它則爲未知品。由表 4.1 的資料繪製圖 4.1。進一步爲了想了解此測量過程誤差發生的情況, 因此, 僅將標準品的測量值繪製成圖 4.2。

表 4.1 砷濃度的測量資料

測量順序	測量型態	濃度($\mu g/L$)	測量順序	測量型態	濃度($\mu g/L$)
1	*CCV*	19.500	14	962206	83.660
2	962180	9.527	15	962207	85.190
3	962181	14.680	16	*CCV*	21.160
4	962182	86.010	17	962202	-8.131
5	962183	40.580	18	962203	-0.699
6	962183	45.690	19	962204	-2.263
7	962183	52.150	20	*CCV*	21.080
8	962184	74.110	21	962253	2.970
9	962185	50.530	22	962254	2.129
10	*CCV*	20.750	23	962255	2.500
11	962186	7.781	24	962256	2.441
12	962187	107.700	25	*CCV*	22.400
13	962205	68.420			

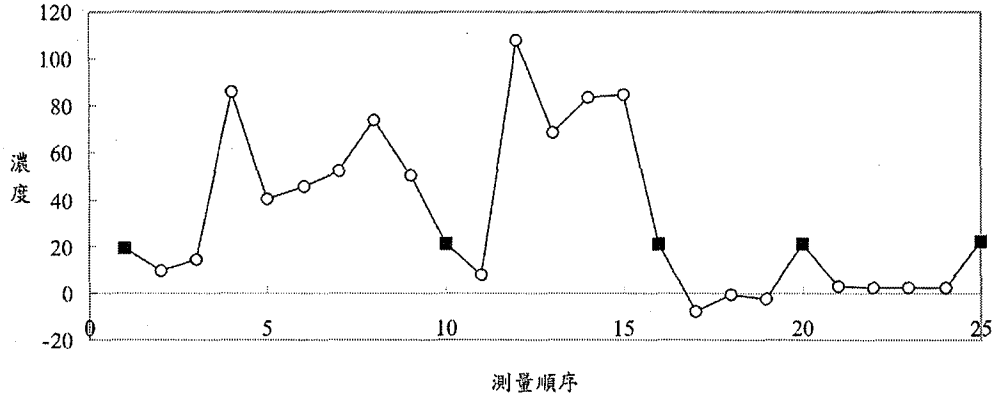

圖 4.1 砷的未知品(以空心圓表之) 及標準品(以實心矩形表之) 之測量值

圖 4.1 可看到測量的順序及砷濃度量測間的關係。其中, 共量測了 20 次未知品及 5 次的標準品。

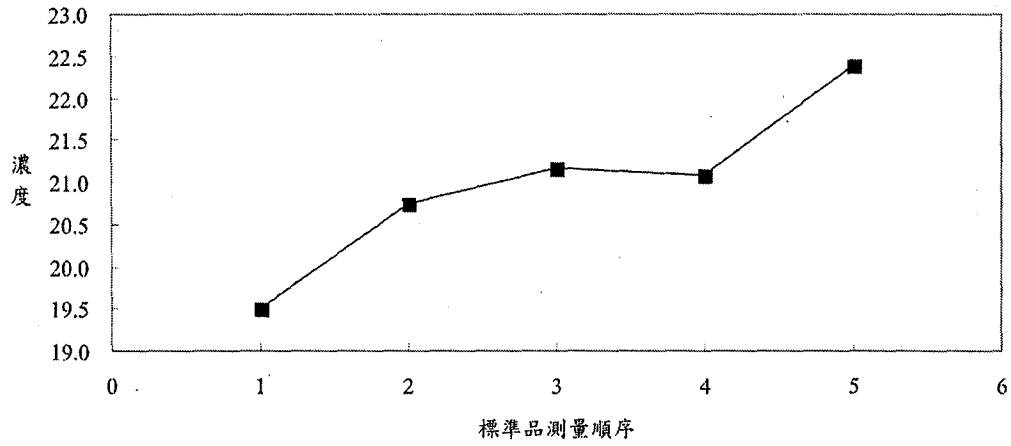

圖 4.2 標準品(眞值爲 20 $\mu g/L$) 之測量值

再由圖 4.2 所觀測到的 5 次標準品中發現, 測量的過程中濃度有上升的趨勢, 這意味著可能有某種相關性存在測量值之間。

因此, 不僅此種相關性結構的估計問題受到重視, 對於測量品的排列方式亦受到研究者的注意。假若這樣的相關性可以利用統計模型來描述, 並且藉由標準品的測量得到非常精確的估計, 或甚至先假設其爲某種已知的相關性結構, 則我們可以藉助標準品與未知品的適當排列方式, 來增加未知品估計的精確度。例 4.2 可以用來說明這樣的想法。

例 4.2. 在某測量過程中, 1 個未知品測量 1 次與 1 個標準品測量 2 次。由僅包含隨機誤差的累加性模型 $y_i = \tau + \epsilon_i,\ i = 1,2,3$ 來描述, 且三個測量品的隨機誤差有如下的分佈:

$$\begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.00 & 0.50 & 0.25 \\ 0.50 & 1.00 & 0.50 \\ 0.25 & 0.50 & 1.00 \end{bmatrix} \right).$$

當此三個測量品的量測順序爲 USS, 其中 U 表未知品, S 爲標準品, 則未知品的眞值估計如下:

$$\hat{\tau} = y_1 - E(\epsilon_1 \mid \epsilon_2, \epsilon_3) = y_1 - 0.5\epsilon_2,$$

且此估值的變異數爲

$$Var(\hat{\tau}) = 0.75.$$

若不考慮相關性存在, 直接用 y_1 估計時, 則 $Var(\hat{\tau}) = Var(y_1) = 1.00$。再者, 當此三個測量品的量測順序爲 SUS 時, 則未知品的眞值估計如下:

$$\hat{\tau} = y_2 - E(\epsilon_2 \mid \epsilon_1, \epsilon_3) = y_2 - 0.4\epsilon_1 - 0.4\epsilon_3,$$

此估值的變異數爲

$$Var(\hat{\tau}) = 0.60.$$

同樣的, 若不考慮相關性存在, 直接用 y_2 估計時, 則 $Var(\hat{\tau}) = Var(y_2) = 1.00$。由上述兩種排列方式結果得知, 當考慮相關性到模型中時, 不論何種排法, 其皆可較精確地估計未知品的眞值。進一步比較不同的排列方式時, 則 SUS 明顯較 USS 來的佳。

對於相關性誤差的測量問題, 早期是由 Pepper (1973) 所討論。基於模型 (2.2), 他的研究有兩個主要目的: 首先, 他試圖尋找關於隨機誤差參數的估計及校準的方法; 再者, 他關心在測量過程中的標準品及未知品的量測次數及排列順序。因此, 他考慮以下三種不同的排列方式進行分析: (1) On Line 測量, 例如:

$$S\ U_1\ U_2\ U_3\ U_4\ S\ U_5\ U_6\ U_7\ U_8\ S \cdots .$$

其中 S 表標準品, U_j 爲第 j 種未知品; (2) On Stream 測量, 例如:

$$S\ U_1\ U_1\ U_1\ S\ U_2\ U_2\ U_2\ S;$$

(3) $m-1$ 種不同未知品的重複測量, 例如:

$$\begin{array}{ccccccc} S & U_1 & U_2 & \cdots & U_{m-1} \\ S & U_1 & U_2 & \cdots & U_{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ S & U_1 & U_2 & \cdots & U_{m-1} & S. \end{array}$$

最後, 他建議使用排列方式 (3), 因爲這種方式有兩種好處: (i) 就實際應用來說, 這種排列方式適用於自動化儀器的連續測量; (ii) 這種排列方式最接近最適設計, 即可使未知品的眞值之 MLE 的漸近變異數達到最小的排列方式。然而, Pepper (1973) 的研究, 並沒有提出一套系統性的方法來建構最適測量設計。

Liao, Taylor and Iyer (2000) 考慮累加性模型 (2.3) 且隨機誤差 $\epsilon_i \sim AR(1)$ 的情況, 則 $\boldsymbol{\tau} = [\tau_1, \tau_2, \ldots, \tau_m]'$ 的廣義最小平方估計式 (generalized least square

estimators, GLSE) 的變異數矩陣如下:

$$Var(\hat{\tau}) = \sigma^2 \mathbf{C}^{-1}.$$

其中,

$$\mathbf{C} = \mathbf{X}_2'[\mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}_1(\mathbf{X}_1'\mathbf{V}^{-1}\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{V}^{-1}]\mathbf{X}_2, \tag{4.1}$$

又 $\mathbf{X}_1$ 是元素全爲 1 之行向量, $\mathbf{X}_2$ 則是由 δ_{ij} 所組成, 令 $\mathbf{V}^{-1} = (a_{ij})_{i,j=1,2,\ldots,N}$, 則

$$a_{i,j} = \begin{cases} 1 & \text{if } i = j = 1 \text{ or } i = j = N, \\ 1+\phi^2 & \text{if } i = j \text{ and } 1 < j < N, \\ -\phi & \text{if } \mid i-j \mid = 1, \\ 0 & \text{otherwise.} \end{cases}$$

矩陣 $\mathbf{C}$ 稱爲 τ 的訊息矩陣 (information matrix)。在這種情況下, τ_j 的估計不只受到標準品與未知品的測量數影響, 還會受制於 AR(1) 的參數 ϕ 之正負號的干擾。爲了描述訊息矩陣 $\mathbf{C}$ 與標準品及未知品的排列關係, 他們引進下列設計參數 (design parameters) b_0, e_j, q_j, $r_{jj'}$。b_0 表示連續兩次測量標準品的次數加上第 1 次或最後 1 次測量爲標準品的次數; e_j 表未知品 U_j 於第 1 次或最後 1 次測量的次數, 即 $e_j = \delta_{1,j} + \delta_{N,j}$, 其中 $\sum_{j=1}^{m} e_j = 0, 1,$ 或 2; q_j 表未知品 U_j 連續兩次出現的次數, 即 $q_j = \sum_{k=1}^{N-1} \delta_{k,j}\delta_{k+1,j}$; $r_{j,j'}$ 表未知品 U_j 與 $U_{j'}$ 互爲相鄰位置的次數。以下針對上述的定義, 舉一實例說明。

例 4.3. 某測量過程中, 共包含 1 種標準品與 3 種未知品。我們利用參數 b_0, e_j, q_j 及 $r_{jj'}$ 來說明以下的排列方式。

$$U_1U_3SU_1U_3U_3U_3U_2U_1U_2U_2SS$$

$b_0 = 2$ (標準品S在第12與第13次爲連續測量, 且最後1次亦爲標準品的測量)
$e_1 = 1$ (第 1 次測量爲未知品 U_1)
$e_2 = 0$ (未知品 U_2 非於第 1 次或最後 1 次測量)
$e_3 = 0$ (未知品 U_3 非於第 1 次或最後 1 次測量)
$q_1 = 0$ (未知品 U_1 沒有連續兩次的測量)
$q_2 = 1$ (未知品 U_2 在第 10 與第 11 次爲連續測量)
$q_3 = 2$ (未知品 U_3 在第 5 與第 6 次且第 6 與第 7 次皆爲連續測量)
$r_{1,2} = 2$ (未知品 U_1 與 U_2 在第 8 與第 9 次且第 9 與第 10 次皆爲相鄰測量)

$r_{1,3}=2$ (未知品 U_1 與 U_3 在第 1 與第 2 次且第 4 與第 5 次皆爲相鄰測量)
$r_{2,3}=1$ (未知品 U_2 與 U_3 在第 7 與第 8 次爲相鄰測量)
進一步, Liao 等人考慮在 $Var(\hat{\tau}_j)$ 皆相等及 $Cov(\hat{\tau}_j,\hat{\tau}_{j'})$, $1\le j\ne j'\le m$, 皆相等的條件下, 換言之, 即 (4.1) 的訊息矩陣 $\mathbf{C}$ 是一個完全對稱 (completely symmetric) 的前提時, 來尋找 A-最適設計。他們並稱滿足這種條件的測量設計爲平衡 (balanced) 測量設計。對於一個平衡測量設計的充分必要條件爲 $t_j=t$, $e_j=e, q_j=q, r_{j,j'}=r$, 對於所有 j 及 j', 並且

$$t=\frac{1}{m}\sum_{j=1}^{m}t_j,\quad e=\frac{1}{m}\sum_{j=1}^{m}e_j,\quad q=\frac{1}{m}\sum_{j=1}^{m}q_j,\quad r=\frac{1}{\binom{m}{2}}\sum_{j<j'}r_{j,j'}.$$

而 A-最適平衡測量設計, 即設計參數 b_0^*,t^*,e^*,q^*,r^* 可使得目標函數 (objective function)

$$Trace(\mathbf{C}^{-1})=\frac{m-1}{t(1-\phi)^2+\frac{2(a_0+1-b_0)\phi}{m}+mr\phi}+\frac{1}{t(1-\phi)^2+\frac{2(a_0+1-b_0)\phi}{m}-\frac{mt^2(1-\phi)^3}{2+(N-2)(1-\phi)}} \tag{4.2}$$

達到最小的排列方式, 式中 a_0 爲標準品的測量數。(4.2) 雖然與 AR(1) 參數 ϕ 有關, 但 Liao 等人証明, 事實上 A-最適平衡測量設計僅與 ϕ 的正負號有關, 即 $-1<\phi<0$ 僅存在一個 A-最適平衡測量設計, 不會因爲 ϕ 值的不同而有所差異。同樣的, 當 $0<\phi<1$ 時, 亦有相同的結果。例 4.4 即爲一個 A-最適平衡測量設計的排列結果。

例 4.4. 當 $0<\phi<1$, $N=60$, $m=4$, 則 $a_0^*=20$, $t^*=10$, $q^*=0$, $r^*=4$ 及 $b_0^*=5$ 滿足 A-最適平衡測量設計, 且其排列方式如下 (測量順序由左至右, 由上而下):

$$\begin{array}{cccccccccccccc}
S & U_4 & U_1 & U_3 & S & U_3 & U_2 & U_3 & S & U_3 & U_1 & U_4 & S & U_4 & S\\
S & U_1 & U_2 & U_4 & S & U_4 & U_3 & U_4 & S & U_4 & U_2 & U_1 & S & U_1 & S\\
S & U_3 & U_4 & U_2 & S & U_2 & U_1 & U_2 & S & U_2 & U_4 & U_3 & S & U_3 & S\\
S & U_2 & U_3 & U_1 & S & U_1 & U_4 & U_1 & S & U_1 & U_3 & U_2 & S & U_2 & S
\end{array}$$

Liao 等人所建構 A-最適平衡測量設計的步驟, 主要是藉由 Kiefer and Wynn (1981) 及 Cheng (1983) 研究相鄰次數相等 (equi-neighbored) 之區集設計中所提出的一組特殊拉丁方設計 (latin squares) 來建構。

5. 未來可能的研究方向

整體而言, 測量品的配置問題是在有限成本的條件下, 以有興趣的參數估計之精確度作爲考量的基礎。一旦考慮隨機誤差間有相關性存在時, 則測量品的排列問題會隨之出現。通常在測量設計的問題裡, 排列問題相對於配置問題來的繁雜且困難許多。綜合以上的介紹, 我們認爲有以下幾方面值得未來有興趣者繼續研究。

5.1 測量品的配置問題

配置問題可以針對更廣義的線性校準模型來研究。例如 Rocke and Lorenzato (1995) 所考慮的模型:

$$y_i = \alpha + \beta\tau e^{\eta} + \epsilon. \tag{5.1}$$

式中 y_i 爲測量值; τ 爲受測品的眞值; α 和 β 皆爲系統誤差, 其中 α 又稱爲系統誤差中的累加性誤差, β 則稱爲系統誤差中的比例性 (proportional) 誤差或相乘性 (multiplicative) 誤差; η 和 ϵ 皆表示隨機誤差, η 是隨機誤差中的比例性誤差, ϵ 則是隨機誤差中的累加性誤差。此模型不僅被使用來描述化學品的濃度測量, 近來亦用來描述生物晶片 (microarrays) 的基因表現量 (gene expression) 之測量值, 可參閱 Rocke and Durbin (2001)。因此, 針對此模型測量品的配置問題, 值得研究者深入探討。

5.2 測量品的排列問題

針對累加性模型 (2.3), 可以考慮加入區集因子 (blocking factor), 此區集因子可能意指不同的操作者 (operators) 或不同的測量儀器 (instruments) 等。因此, 模型可重新表達如下:

$$y_{ik} = \mu + \delta_{i0}\tau_0 + \sum_{j=1}^{m}\delta_{ij}\tau_j + \beta_k + \epsilon_{ik} \qquad \begin{cases} i = 1, 2, \ldots, N_k \\ k = 1, 2, \ldots, b \end{cases} \tag{5.2}$$

其中, β_k 爲區集 k 的效應。對於每個區集 $k = 1, 2, \ldots, b$, 則 $\epsilon_{1k}, \epsilon_{2k}, \ldots, \epsilon_{N_k,k}$ 遵循 AR(1)。此問題等同於區集內實驗單位 (experimental units) 存在 AR(1) 的相關性趨勢 (trend), 實驗組與對照組的非二元 (nonbinary) 區集設計之問題。

另外, 對於累加性模型, 亦可考慮其它相關性結構, 例如: AR(2) 或更高階 AR(p) 的隨機過程。同樣對於線性校準模型 (2.6), 亦可以考慮加入 AR(1) 過程來討論測量品的排列方式。

5.3 穩健性 (robust) 測量設計

此問題主要是針對在測量過程中可能發生不同的相關性結構, 如 AR(1), AR(2), MA(1), MA(2) 等。因此, 有必要使用對於不同相關性結構具有穩健性的測量設計, 即對於未知品的眞值估計之精確度, 不會因爲不同的相關性結構而受到劇烈的影響。最近 Zhou(2001) 提出一個新的衡量設計穩健性的準則 (robust criterion), 值得被應用在測量設計上。本文作者現在正從事這方面的研究。

參考文獻

Brown, P. J. (1993). *Measurement, regression and calibration.* Oxford University Press, New York.

Cheng, C. S. (1983). Construction of optimal balanced incomplete block designs for correlated observations. *Ann. Stat.* **11**, 204-209.

Cutler, D. R. (1993). Efficient block designs for comparing test treatments to a control when the errors are correlated. *J. Stat. Plann. Inference* **36**, 107-125.

Fuller, W. A. (1987). *Measurement error models.* John Wiley and Sons, New York.

Hunter, W. G. and Lamboy, W. F. (1981). A Bayesian analysis of the linear calibration problem. *Technometrics* **23**, 323-328.

Kiefer, J. and Wynn, H. P. (1981). Optimal balanced block and latin square designs for correlated observations. *Ann. Stat.* **9**, 737-757.

Liao, C. T., Taylor, C. H. and Iyer, H. K. (2000). Optimal measurement designs when errors are correlated. *J. Stat. Plann. Inference* **84**, 295-321.

Lin, T. Y. and Liao, C. T. (2004). Optimal allocation of measurement in a linear calibration process. *Metrika* (in press).

Pepper, M. P. G. (1973). A calibration of instruments with non-random errors. *Technometrics* **15**, 587-599.

Perng, S. K. and Tong, Y. L. (1977). Optimal allocation of observations in inverse linear regression. *Ann. Stat.* **5**, 191-196.

Rocke, D. M. and Lorenzato, S. (1995). A two-component model for measurement error in analytical chemistry. *Technometrics* **37**, 176-184.

Rocke, D. M. and Durbin, B. (2001). A model for measurement error for gene expression arrays. *J. Comput. Biol.* **8**, 557-569.

Zhou, J. (2001). A robust criterion for experimental designs for serially correlated observations. *Technometrics* **43**, 462-467.

[民國 93 年 10 月收稿,民國 93 年 12 月接受。]

AN INTRODUCTION TO MEASUREMENT DESIGNS

Tsai-Yu Lin and Chen-Tuo Liao
Division of Biometry, Department of Agronomy,
National Taiwan University

ABSTRACT

Measurement making is an activity for most people. The purpose of the measurements is to estimate the degree to which some characteristic of a specimen is present. A measurement process is usually subject to errors which may be classified as random errors only or a combination of both random errors and systematic errors. An experimenter may monitor the errors by measuring standards at appropriate intervals during a measurement process in a laboratory or a product process in industry. This is because that the true values of standards are known, so the errors can be observed whenever standards are measured. Hence, one may use standard measurements to estimate the parameters associated with the errors, and obtain a precise estimate for the true value of unknown specimen by calibrating its observed values. This article discusses some literature concerning two important design issues of a measurement process, including the allocation and the arrangement of order of measurements for standards and unknown specimens. Some possible future research problems are also presented.

Key words and phrases: Systematic errors, random errors, calibration models, A-optimality criterion, nonbinary design.

AMS 2000 subject classifications: Primary 62K99; secondary 62P30.