

Process Watch：對自己信賴區間充滿自信

作者: Douglas G. Sutherland 和 David W. Price

作者按語: Process Watch 系列連載探討了半導體產業製程管制（缺陷檢測與量測）的基本概念。上一集連載探討了製程管制 [10 大根本法則](#)，這一全新的系列文章連載著重說明製程管制的其他趨勢，包括積體電路生產的成功實施策略及益處。

在 1900 年代初期效力於愛爾蘭都柏林 Guinness® 啤酒釀酒廠期間，William Sealy Gosset 發現了被稱為 T 檢驗¹ 的統計算法。Gosset 採用這種算法確定產量最高的大麥品種，最大限度為其雇主降低成本，但為了保護 Guinness 的知識產權，他以 Student「（「學生」）為筆名發表他的發現和理論。我們今天使用的 T 檢驗版本是經過 Gosset 在牛津大學的同事 Ronald Fisher 爵士發揚光大的，但通常仍被稱為學生 T 檢驗。本文並不闡述 T 檢驗本身的數學理論，而是討論持續達到 T 檢驗結果中的 95% 的信賴區間所需要的資料數量。

T 檢驗是一種統計算法，用於確定兩個樣本是否屬於同一個母群體。它並不明確解決問題，而只是計算兩個樣本同屬一個母群體的概率。例如，如果我們開發出一種清潔蝕刻室的新方法，則我們可能想要證明它能夠減少掉落的顆粒。利用晶圓檢測系統，我們可以測量採用舊的清潔製程後蝕刻室內晶圓上的顆粒數量，然後測量採用新的清潔製程後的顆粒數量。我們隨後可以用 T 檢驗來證明兩者之間的差異是否具有統計意義，抑或僅僅是隨機波動的結果。T 檢驗可以回答以下這個問題：兩種樣本同屬一個母群體群的概率有多大？

但如圖 1 所示，T 檢驗可能會產生兩種偽結果：偽陽性或偽陰性。若要證實實驗資料與基線確實不一樣，T 檢驗的結果通常必須小於 5%（即偽陽性的概率小於 5%）。但如果 T 檢驗的結果大於 5%（陰性結果），它並不能說明有多少概率這種結果是假的。偽陰性的概率是由測量次數決定的。因此始終存在兩種檢驗標準：（1）我的實驗是否通過了 T 檢驗？（2）我測量的次數是否足以讓我對結果有信心？在本文中我們嘗試闡述後一個問題。

		實際	
		無變化	變化
T 檢驗結論	變化	偽陽性	
	無變化		偽陰性

圖 1. 「真值表」強調 T 檢驗得出錯誤結果的兩種可能。

改變半導體製造製程是個昂貴的工程議題。實施沒有任何意義（假陽性）的變動不僅浪費時間，而且還可能有害。不實施可能有益（假陰性）的變動則可能會因喪失機會而付出數千萬美元的代價。重要的是，您需要對結果具有適當程度的信心，而要做到這一點，您使用的樣本量必須與您試圖達到的變化程度相匹配。以蝕刻清潔製程為例，這就意味著需要收集足夠數量的晶圓檢測資料，以確定新的清潔製程能否真正減少了顆粒的數量。

一般來說，兩件事物之間的差異越大，就越容易區分它們。區分紅色和藍色比較容易，但是區分兩種不同深淺度的紅色或兩種不同深淺度的藍色則比較困難。同理，樣本自身的變異性越小，就越容易看出變化²。在統計學中，變異性（有時稱為雜訊）通常以標準差(σ)為單位來衡量。通常兩個樣本之間的平均差也很方便地以 σ 為單位表示（例如，實驗結果均值低於基線均值 1σ ）。其優勢在於，它用共同計量單位 (σ) 來標準化呈現結果。僅僅用一些絕對數值來說明兩個均值之間的差異不夠充分（例如，A 的平均值比 B 的平均值大 42）。然而，如果我們能夠以標準偏差為單位表示這個絕對值，則能夠直接說明問題的背景，讓人很容易理解這兩個均值之間的相對差異（例如，A 的平均值比 B 的平均值大 1 個標準偏差）。

圖 2 顯示了變動前後兩種資料集的範例。我們可以將其視作之前討論過的蝕刻室清潔實驗。基線資料是採用新清潔製程之前的每個晶圓上的顆粒數量，結果資料是採用新清潔製程之後的每個晶圓上的顆粒數量。圖 2A 的結果顯示出資料集均值的變化較小，但標準偏差高，圖 2B 中均值的變化量與圖 2A 相同，但雜訊較低（標準偏差較低）。相比圖 2B，您將需要更多的資料（例如檢測更多的晶圓）來證實圖 2A 中的變化，這是因為，儘管兩者的絕對變化相同，但圖 2A 中的信噪比較低。

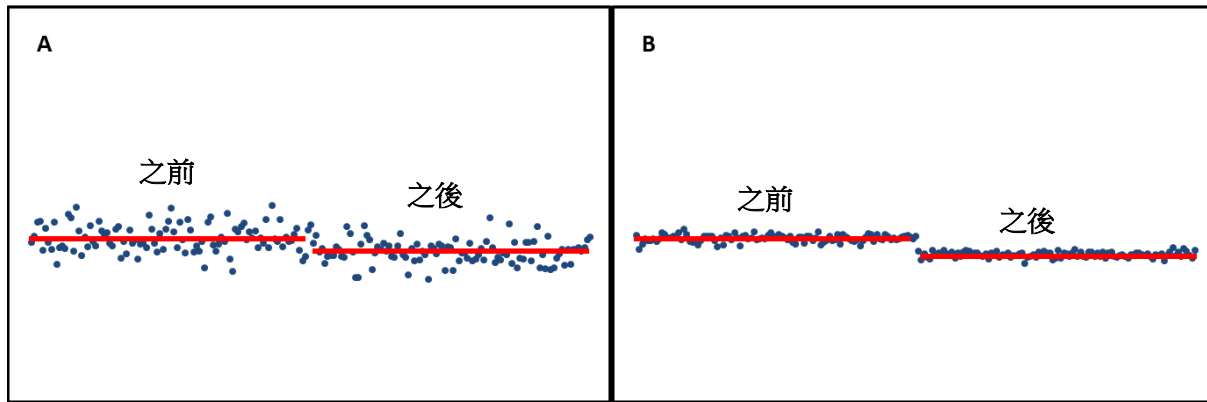


圖 2. 兩張圖顯示之前和之後的絕對變化相同，但圖 2B（右）的標準偏差低很多。如圖 2A（左）所示，如果相對於標準偏差，變化很小時，則需要更多的資料證實。

問題在於：我們需要多少資料才能充滿自信地證明這些差異？當我們用標準誤差（SE）來描繪資料時，就能直觀地看到這一點。SE 可以被看作是計算平均值時的誤差（例如，平均值為 $\bar{X} \pm SE$ ）。SE 與 σ/\sqrt{n} 成正比，其中 n 是樣本量。圖 3 顯示了兩個不同樣本的 SE 作為測量次數 n 的函數。

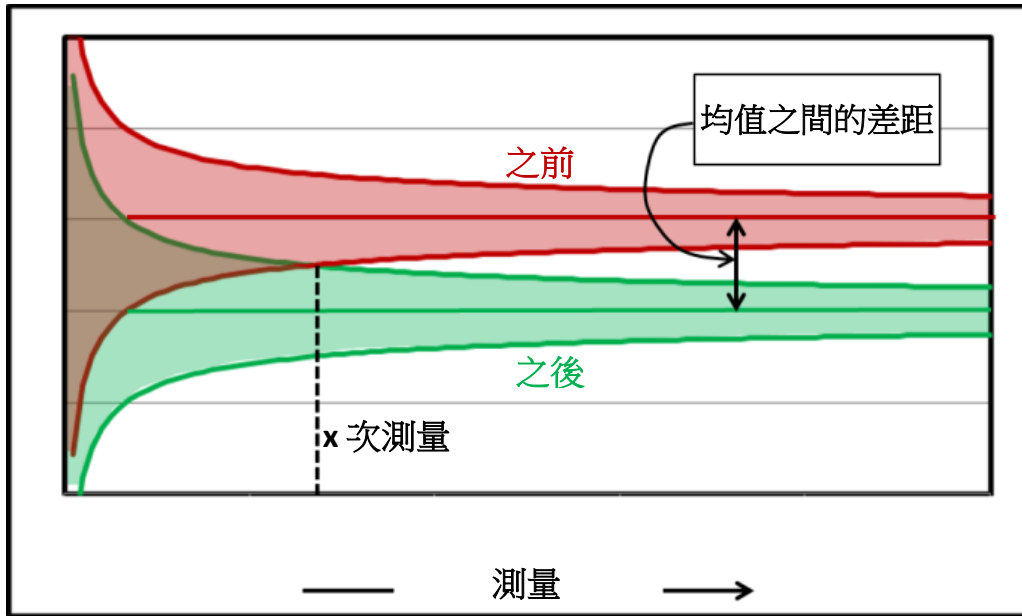


圖 3. 具有不同均值的兩個樣本平均值中的標準誤差 (SE)。在本例中，兩個資料集中的標準差是相同的，但並不需要這樣。測量次數超過 x 時，誤差線不再重疊，因此可以 95% 以上的可信度說兩組資料是不相同的。

對於既定的均值差和標準差，我們可以計算出使這兩個測量分佈的標準誤差不相重疊所需要的測量次數 x （以既定的信心水準）。

確定 T 檢驗中的正確樣本量的實際方程式如下：

$$n = \frac{2(Z_{\beta} + Z_{1-\alpha/2})^2}{\text{Delta}^2} \quad \text{公式 1}$$

其中 n 是所需要的樣本量，「Delta」是測量出來的以標準偏差 (σ) 為單位兩個均值之間的差異， Z_x 是概率為 x 的 T 分佈下的面積。如果 $\alpha=0.05$ （偽陽性的概率為 5%）、 $\beta=0.95$ （偽陰性的概率為 5%），則 $Z_{1-\alpha/2}$ 和 Z_{β} 分別等於 1.960 和 1.645（其他 α 和 β 值的 Z 值請參閱大多數統計教科書、Microsoft® Excel® 或網路）。如圖 3 所示和公式 1 的數學計算所指出，隨著兩組資料之間的差異 (Delta) 變小，區分兩者所需要的測量次數將呈指數級增加。圖 4 顯示了所需要的樣本量作為以 σ 為單位的均值差 Delta 的函數。正如預期，對於大於 3σ 的大變化，95% 的時候可以用很少的資料通過 T 檢驗。

隨著 Delta 變小，則需要更多的測量來持續證實變化。只有 1 個標準偏差的變化需要之前和之後進行 26 次測量，而 0.5σ 的變化則需要超過 100 次測量。

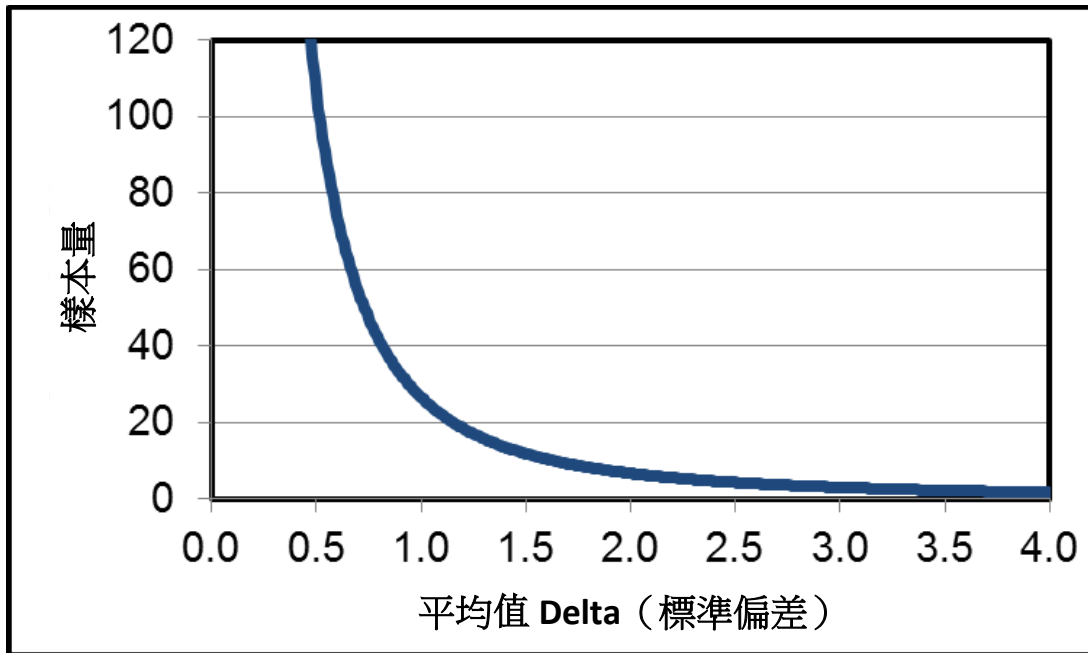


圖 4. 證實 5% 偽陽性和 5% 為陰性的兩組數據的既定平均值變化所需要的樣本量

這種變化量與檢測變化量所需要的最少測量次數之間的關係，對如何選擇檢測或量測工具的類型有深遠影響，以證實所期望的變化。採用圖 4 中的資料，圖 5 顯示了用不同類型的工具證實既定變化所需的時間。在這個例子中，樣本量是以晶圓數量衡量的。對於快速的檢測工具（高產出，例如雷射掃描晶圓檢測系統），證實相對較小的製程改進 ($<0.5\sigma$) 是可行的，因為它們可在相對較短的時間內進行 200 次必要的測量（之前 100 次和之後 100 次）。諸如電子束檢測系統之類的慢速檢測工具只限於檢測重大製程變化，如且製程的改進大於 2σ 。即使這樣，僅測量時間一項就意味著需要數週時間才能證實一個陽性結果。以蝕刻室清潔為例，需要迅速確定清潔過程中的變化結果，以保證蝕刻工具可以重新投入生產。確定顆粒數量變化的最佳檢測系統是高產出系統，它可以檢測需要關注的顆粒，且晶圓與晶圓之間的變異性很低。

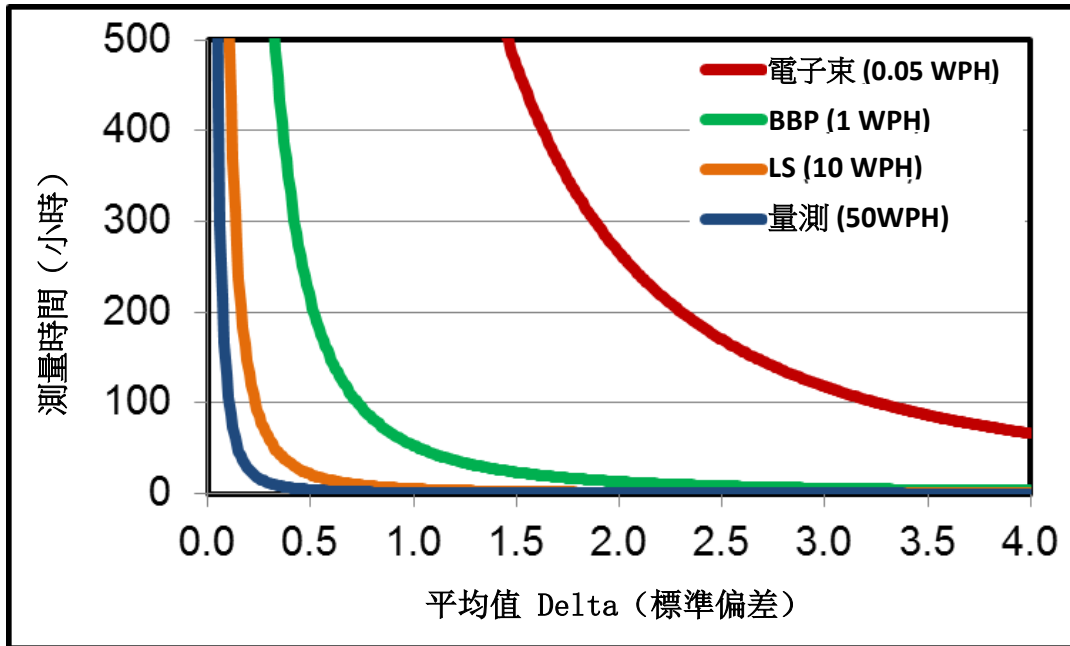


圖 5. 確定四種不同產出的製程管制工具的特定變化所需要的測量時間（電子束、寬頻電漿、雷射散射和量測）

實驗的運行成本很高。如果實驗結果為偽陽性，就會浪費時間和資源，如果實驗結果為偽陰性，則可能因為喪失機會而損失數百萬美元。若要對您的結果具有適當的信心水準，您必須使用與您想要達到的變化量相對應的正確樣本量（因此要使用適當的工具）。

參考文獻:

- 1) https://en.wikipedia.org/wiki/William_Sealy_Gosset
- 2) Process Watch: [Know Your Enemy \(瞭解你的敵人\)](#), *Solid State Technology (固態技術)*, 2015 年 3 月

作者簡介:

David W. Price 博士是 KLA-Tencor 公司的資深總監。Douglas Sutherland 是 KLA-Tencor 公司的首席科學家。在過去 10 年間，Price 博士和 Sutherland 博士一直與 50 多家半導體積體電路製造商直接合作，協助他們優化整體檢測策略，以實現最低總成本。此系列文章對他們在這些工作中觀察到的一些普遍經驗進行了總結。