



國立高雄科技大學

National Kaohsiung University of Science and Technology

分析化學 (Analytical Chemistry)

Instructor : 潘俊仁(Chun-Jern Pan)

E-mail: ppan@nkust.edu.tw

Office: 化材館 709



CH 3 數學工具與統計



- ~~3-1~~ 有效數字
- ~~3-2~~ 有效數字的運算
- 3-3 誤差的形式
- 3-4 不準度的擴張
- 3-5 高斯分布
- 3-6 以 F 檢定比較標準差



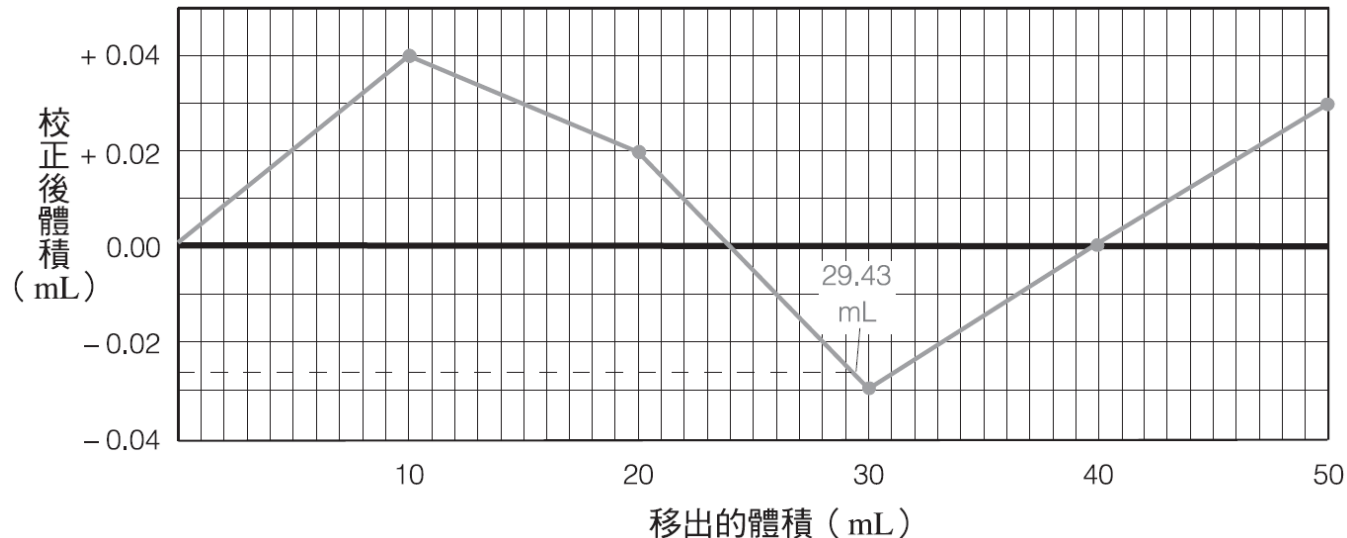
3-3 誤差的形式



- 每一個測量都包含一定程度的不準度，我們稱為實驗誤差（experimental error）。
- 實驗誤差可分成系統性（systematic）與隨機性（random）兩種。

系統誤差

- 系統誤差（systematic error）也稱為**確定誤差**（determinate error）：測量的過程中，以相同的方式進行測量時會具有再現性的誤差。**可修正**
- (ex) 沒有經過校正的滴定管。建立一條校正曲線（圖 3-2）可以修正這樣的誤差。





3-3 誤差的形式



隨機誤差

隨機誤差（random error）也稱為未定誤差（indeterminate error），源自於儀器本身測量能力的限制與測量本身的隨機變動。

精密度與準確度

精密度(**precision**):測量結果的再現性；

準確度(**accuracy**):測量值與真實值接近的程度

不確定度：在一組測量數據中的變化性



3-3 誤差的形式



- 絕對不準度（absolute uncertainty）：測量中不確定值的範圍。
- 相對不準度（relative uncertainty）：絕對不準度與所對應測量值比較而來。

相對
不準度：

$$\text{相對不準度} = \frac{\text{絕對不準度}}{\text{測量值}} \quad (3-2)$$

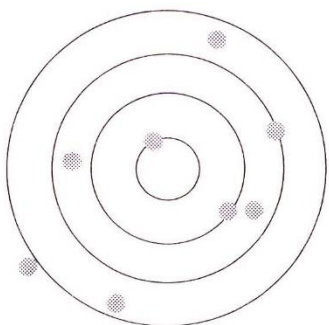
$$= \frac{0.02\text{mL}}{12.35\text{mL}} = 0.002$$

$$\text{百分相對不準度} = 100 \times \text{相對不準度} \quad (3-3)$$

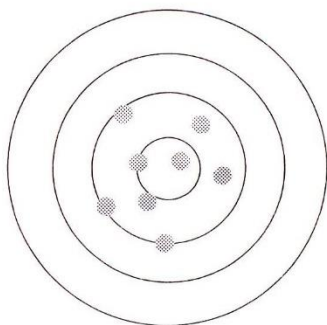
$$= 100 \times 0.002 = 0.2\%$$

3-3 誤差的形式

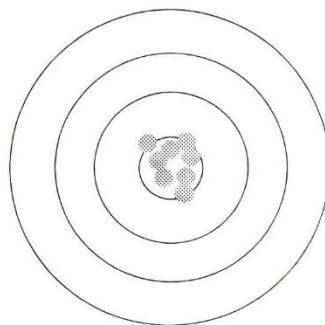
請將這些結果與四個描述配對。



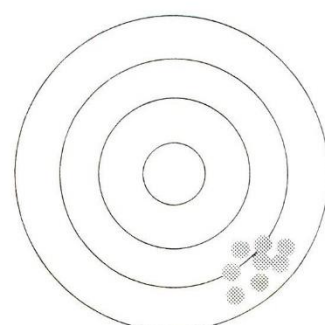
綺兒



辛西雅



卡門



雀斯蒂

(a) 準確又精密

(b) 準確但是不精密

(c) 精密但是不準確

(d) 不準確也不精密



3-4 不準度的擴張



- 大部分的實驗都需要將數個數據經過數學運算的整理來得到結果，而每個數據都有隨機誤差。
- 所得結果的隨機誤差並不是將各數值的隨機誤差相加而已，因為這些數據的隨機誤差可能有些是正偏差，有些是負偏差。
- 如何由各數值的不準度得到運算答案的不準度，這個過程稱為不準度擴張(propagation of uncertainty)。



3-4 不準度的擴張



加減運算



3-4 不準度的擴張



範例 2 滴定管讀數的不準度

由滴定管流出的體積一般是滴定的最終體積讀數與未滴定时起始體積讀數的差值。假設每一個讀數的不準度為 $\pm 0.02 \text{ mL}$ ，那麼流出體積的不準度為多少？



3-4 不準度的擴張



乘除運算



3-4 不準度的擴張



範例 3 科學記號與不準度的擴張

請將下面兩個運算式結果以絕對不準度來表示。

(a)
$$\frac{3.43 (\pm 0.08) \times 10^{-8}}{2.11 (\pm 0.04) \times 10^{-3}}$$

(b)
$$[3.43 (\pm 0.08) \times 10^{-8}] + [2.11 (\pm 0.04) \times 10^{-7}]$$



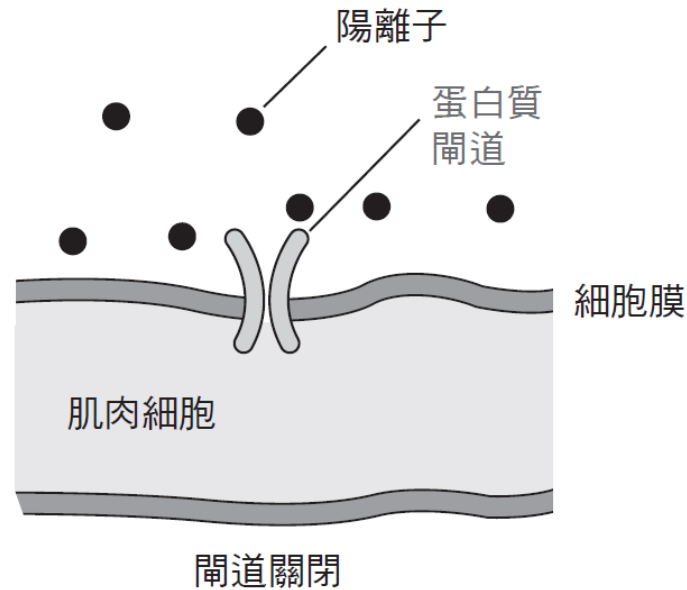
3-4 不準度的擴張



混合運算

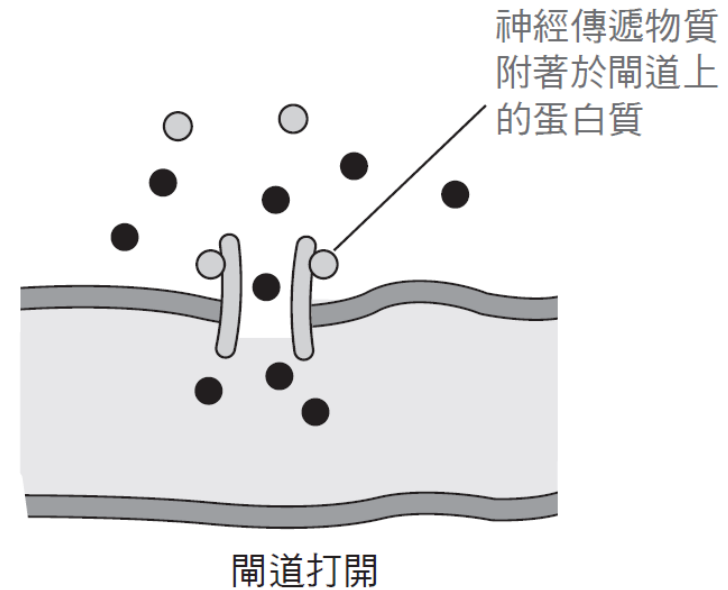
3-5 高斯分布

沙克曼（B. Sakmann）與尼爾（E. Neher）在 1991 年共同獲得諾貝爾獎，表彰他們在肌肉神經接合處有關訊息傳遞研究上的卓越成就。



(a)

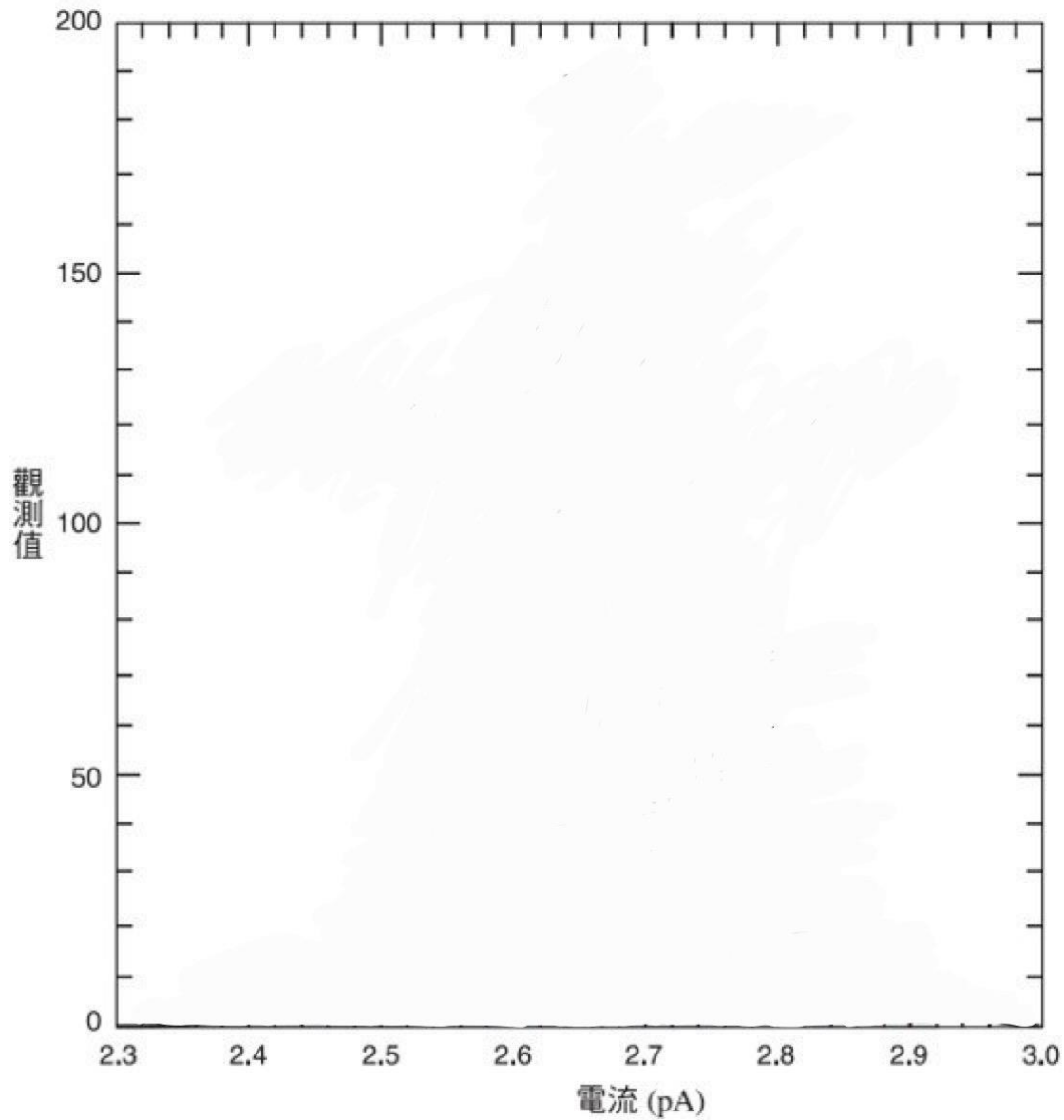
(a) 缺乏神經傳遞物質時離子閘道是關閉的，陽離子無法進入肌肉細胞。



(b)

(b) 神經傳遞物質存在時，閘道會打開且陽離子可進入細胞，啟動了肌肉運動。

3-5 高斯分布





3-5 高斯分布



- 高斯分布可以由**平均值**與**標準差**來描述。
- **平均值**（mean；也稱 average） \bar{x} 是所有測量值的總和除以測量的次數。
- **標準差**（standard deviation） s 表示分布曲線的寬度，標準差愈小，數據分布的範圍就愈窄。

$n - 1$ 稱為自由度（degree of freedom）。

3-5 高斯分布

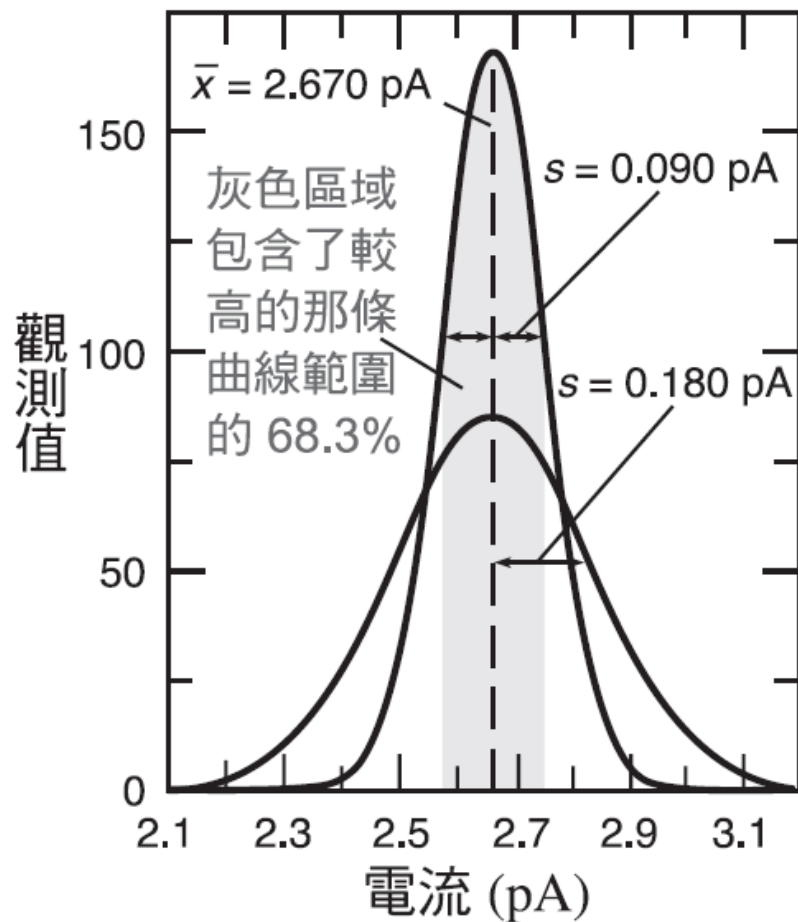
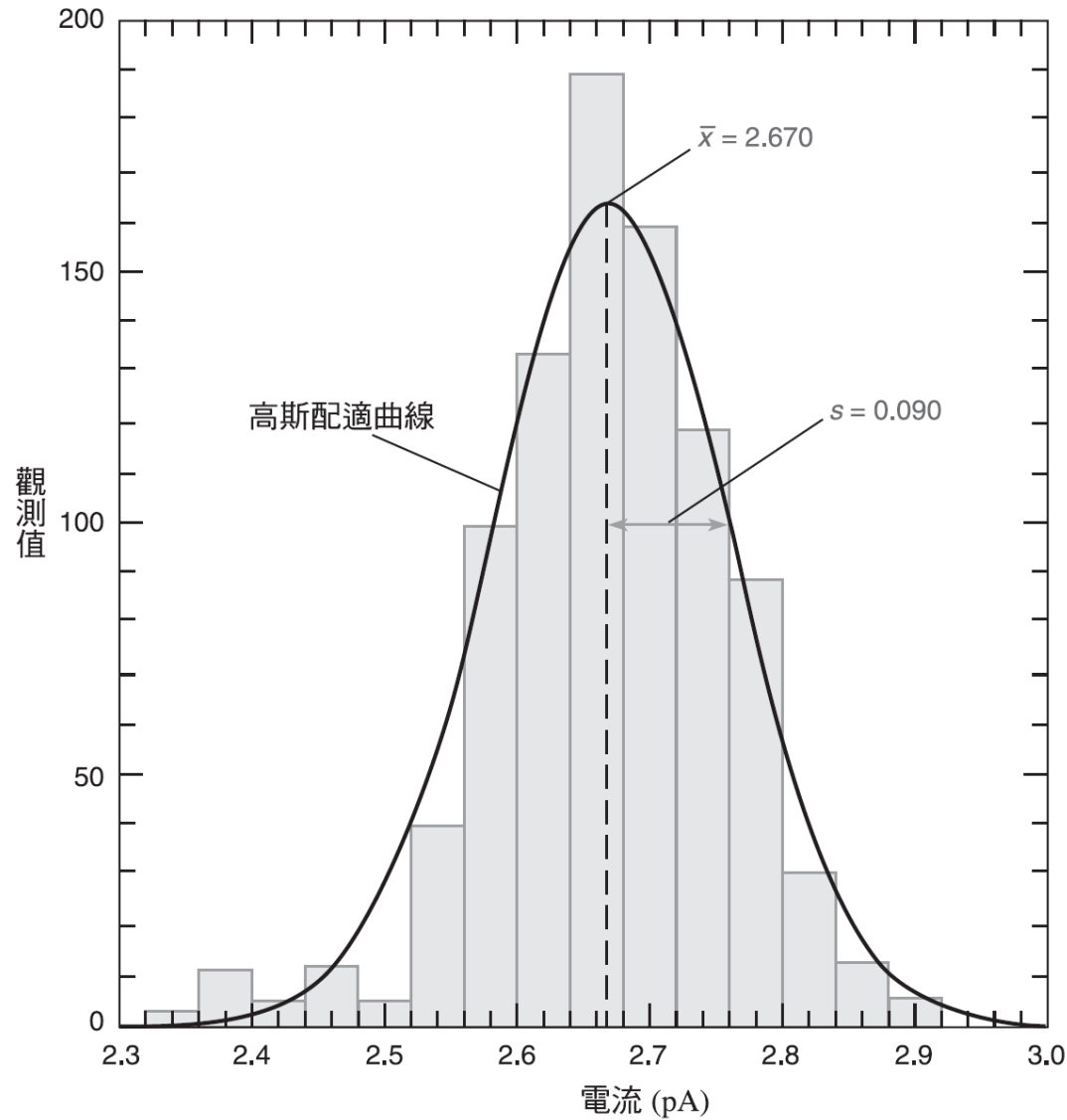


圖 3-4 中， $s = 0.090 \text{ pA}$ 。圖 3-5 顯示標準差變兩倍時，分布曲線會變得較短且較寬。

圖 3-5 較寬曲線的標準差是另一個分布曲線的兩倍，兩條曲線的觀測值數目是相同的。

3-5 高斯分布





3-5 高斯分布

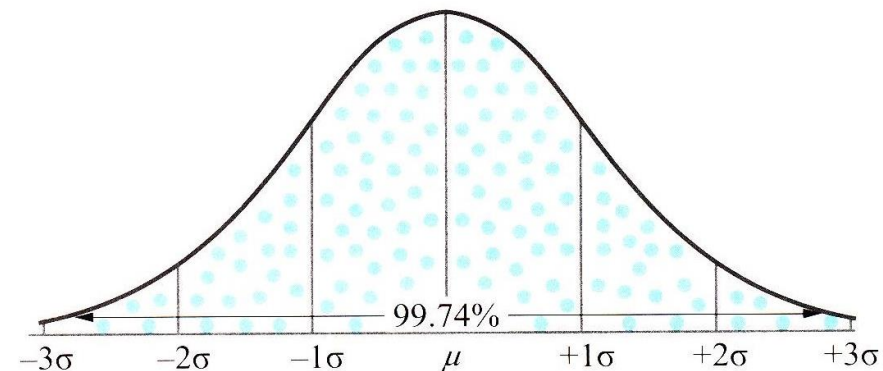
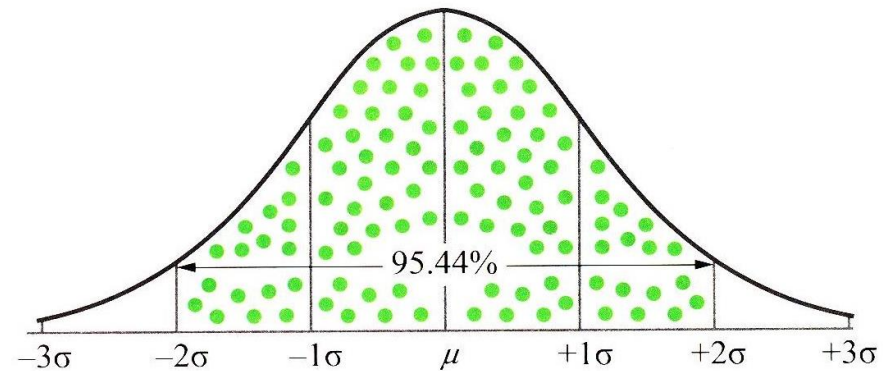
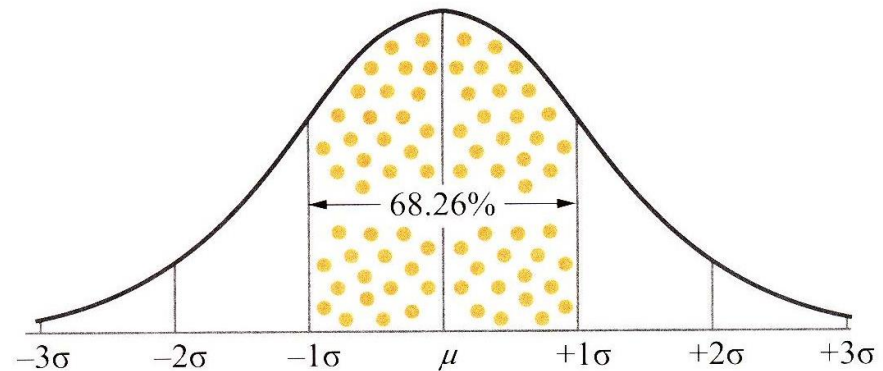


- 相對標準差（relative standard deviation）是標準差除以平均值（ s/\bar{x} ），一般以百分比表示。
- 母體平均(**population mean**)，母體標準差
- 變異數（variance）常常使用在許多的應用中，變異數是標準差的平方（ s^2 ）。
- 中位數（**median**）：中位數是指在一系列的測量值中，將結果從低到高依序排列後，排在中間的測量值。
- 全距（**range**）則是最大與最小值的差距。

3-5 高斯分布(對稱分布)

表 3-1 高斯分布中的觀測值百分比

範圍	高斯分布	圖 3-4 的觀測值
$\mu \pm 1\sigma$	68.3%	71.0%
$\mu \pm 2\sigma$	95.5	95.6
$\mu \pm 3\sigma$	99.7	98.5





3-5 高斯分布



範例 5 平均值與標準差

試求出 (7, 18, 10, 15) 這組數字的平均值，標準差與相對標準差。



信賴區間(confidence interval)



- 在有限的測量中，我們無法得到真實的平均值 μ ，或真實的標準差 σ 。我們所得的是關於這些樣品的平均值 \bar{x} 與標準差 s 。

生物學家想知道  的平均體重

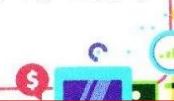
母體資料：全部海洋中的每隻  重量。（母體稀少）

因為除非把海水全部抽光，再撈起全部 ，秤完重量再放回，再注入海水，此行為簡直是天方夜譚。所以僅能隨機抽樣，由樣本  的平均重量回推母體  的平均體重。

教育部長想知道當今大學生的平均 IQ

母體資料：當今所有具大學學籍生的 IQ 數值。（母體量大）

因為收集要花大量測驗時間、人力，所以隨機抽樣部份大學生做測驗，由樣本平均 IQ 回推母體平均 IQ。





信賴區間



- **信賴區間**（confidence interval）是指在某一個範圍內發現真實平均值的機率。

考慮信賴區間時：

$$\mu = \bar{x} \pm \frac{ts}{\sqrt{n}} \quad (3-12)$$

司徒頓 t （Student's t ）是一種用來表達**信賴區間**的統計工具。



信賴區間



表 3-4 t 值

自由度	信賴區間 (%)						
	50	90	95	98	99	99.5	99.9
1	1.000	6.314	12.706	31.821	63.656	127.321	636.578
2	0.816	2.920	4.303	6.965	9.925	14.089	31.598
3	0.765	2.353	3.182	4.541	5.841	7.453	12.924
4	0.741	2.132	2.776	3.747	4.604	5.598	8.610
5	0.727	2.015	2.571	3.365	4.032	4.773	6.869
6	0.718	1.943	2.447	3.143	3.707	4.317	5.959
7	0.711	1.895	2.365	2.998	3.500	4.029	5.408
8	0.706	1.860	2.306	2.896	3.355	3.832	5.041
9	0.703	1.833	2.262	2.821	3.250	3.690	4.781
10	0.700	1.812	2.228	2.764	3.169	3.581	4.587
15	0.691	1.753	2.131	2.602	2.947	3.252	4.073
20	0.687	1.725	2.086	2.528	2.845	3.153	3.850
25	0.684	1.708	2.060	2.485	2.787	3.078	3.725
30	0.683	1.697	2.042	2.457	2.750	3.030	3.646
40	0.681	1.684	2.021	2.423	2.704	2.971	3.551
60	0.679	1.671	2.000	2.390	2.660	2.915	3.460
120	0.677	1.658	1.980	2.358	2.617	2.860	3.373
∞	0.674	1.645	1.960	2.326	2.576	2.807	3.291

計算信賴區間時，如果用某種特別的方法能得到足夠大量的實驗數據，我們可用 σ 取代 3-12 式中的 s 來進行計算，而得到真正的分布與標準差。當利用 σ 進行計算時，3-12 式中的 t 值來自於表格中最底下的一行。



範例 7 計算信賴區間

經過重複多次測定醣蛋白中的醣含量，所測得的值分別為每 100 g 蛋白質中含有 12.6, 11.9, 13.0, 12.7 與 12.5 g 醣，請求出 50% 與 90% 信賴區間。



準確度:信賴區間與司徒頓t



考慮信賴區間時：

$$\mu = \bar{x} \pm \frac{ts}{\sqrt{n}}$$

(3-12)

關於量測品質...



3-5 F-test for 精密度



- 兩組不同測量值的平均值在統計上是否不同???從精密度與準確度來看。
- 比較二組相同測定之精密度: **F-test**。

表 3-2 馬血液中 HCO_3^- 含量的測定值 ^a

	原機型	新機型
平均值 (\bar{x} , mM)	36.14	36.20
標準差 (s , mM)	0.28	0.47
測定次數 (n)	10	4

a. 數據來自 M. Jarrett, D. B. Hibbert, R. Osborne and E. B. Young, *Anal. Bioanal. Chem.* **2010**, 397, 717.

新機器的標準差 (s) 是否比舊機型「顯著」來的大？



3-5 F-test



表 3-2 馬血液中 HCO_3^- 含量的測定值 ^a

	原機型	新機型
平均值 (\bar{x} , mM)	36.14	36.20
標準差 (s , mM)	0.28	0.47
測定次數 (n)	10	4

a. 數據來自 M. Jarrett, D. B. Hibbert, R. Osborne and E. B. Young, *Anal. Bioanal. Chem.* **2010**, 397, 717.



3-5 F-test



表 3-3 在 95% 信賴區間時 $F = s_1^2 / s_2^2$ 的臨界值

s_2 的自由度	s_1 的自由度													
	2	3	4	5	6	7	8	9	10	12	15	20	30	∞
2	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5
3	9.55	9.28	9.12	9.01	8.94	8.89	8.84	8.81	8.79	8.74	8.70	8.66	8.62	8.53
4	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.75	5.63
5	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.50	4.36
6	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.81	3.67
7	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.58	3.51	3.44	3.38	3.23
8	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.08	2.93
9	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.86	2.71
10	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.84	2.77	2.70	2.54
11	3.98	3.59	3.36	3.20	3.10	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.57	2.40
12	3.88	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.47	2.30
15	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.25	2.07
20	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.04	1.84
30	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.84	1.62
∞	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.46	1.00

對於 n 個測量值，自由度 = $n - 1$ ，實驗所得的 F 有 5% 機率大於表列值。

你可以利用 Excel 函數 FINV(probability, Deg_freedom 1, Deg_freedom 2) 選擇適合的信賴區間得到 F 值，鍵入「= FINV(0.05,7,6)」會得到如表列的 $F = 4.21$ 。



t-test



- 英國科學家雷利（Lord Rayleigh，發現惰性氣體氬氣而於 1904 年獲得諾貝爾獎）
- 在雷利那個年代，認為乾燥的空氣中有 $\sim 1/5$ 的 O_2 與 $\sim 4/5$ 的 N_2 。

真的嗎？

實驗一：

高熱的銅將乾燥的空氣中的氧移開

$Cu(s) + 3O_2(g) \rightarrow CuO(s)$ ，然後收集

剩下的氣體，並且在固定壓力與固

定體積之下測定它的密度。

實驗二：

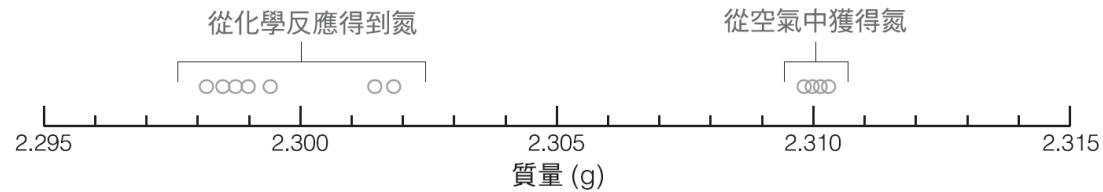
分解一氧化二氮(N_2O)與一氧化氮(NO)

或是亞硝酸銨(NH_4NO_3)的方式製備同體

積的氮氣。

表 3-5 雷利所得到的氮
氣質量^a

從空氣 中獲得	從化學分解 中獲得
2.31017	2.30143
2.30986	2.29890
2.31010	2.29816
2.31001	2.30182
2.31024	2.29869
2.31010	2.29940
2.31028	2.29849
—	2.29889
平均值	
2.31010 ₉	2.29947 ₂
標準差	
0.00014 ₃	0.00137 ₉



實驗誤差???

化學反應中得到的平均質量大 0.46%。



案例 A：標準差無顯著差異時

- 信賴如果 F 檢定告訴我們兩者的標準差沒有顯著差異，那麼對於分別有 n_1 與 n_2 個測量值的兩組數據（平均值分別為 \bar{x}_1 與 \bar{x}_2 ），可以利用以下公式計算 t 值：
- s_{pooled} 是兩組數據的綜合標準差。

如果 $t_{\text{計算值}} > t_{\text{臨界值}} (95\%)$ ，則差異是顯著的。



案例 B：標準差有顯著差異時



t-test



範例 8 雷利從空氣中所得的氮氣比從反應中得到的重嗎？

表 3-5 中，從空氣所得到的氮氣平均質量 $\bar{x}_1 = 2.31010_9 \text{ g}$ ，標準差 $s_1 = 0.00014_3$ ($n_1 = 7$ 個測量值)，而從化學反應所得到的平均質量 $\bar{x}_2 = 2.29947_2 \text{ g}$ ，標準差 $s_2 = 0.00137_9$ ($n_2 = 8$ 個測量值)。請問這兩者的質量是否顯著不同？



t-test





t-test





t-test





t-test

