第5.4節 信賴區間
Confidence Intervals

5.4-1 何謂信賴區間 (Confidence Intervals)

5.4-2 標準偏差的不確定性：t 分佈 (t -Distributions)

5.4-3 平均值的不確定性 (Uncertainty of Averages)

5.4-4 因子反應值的信賴區間 (CI for Factor Responses)

5.4-5 預測值的信賴區間 (CI for Predictive Values)

5.4-6 確認實驗計算值的信賴區間 (CI for Confirmation)

5.4-7 預測值與確認實驗計算值的比較 (Prediction and Confirmation)

5.4-8 t 分佈與F分佈的關係 (t - and F -Distributions)

信賴區間

- 本節主要的目的是：
  - 評估實驗模式參數值（亦即____________）的誤差。
  - 評估預測值及確認實驗計算值的誤差，並判定兩者是否「　　　　　」。
- 這些工作背後的主要觀念就是「　　　　　」及其相對應的「　　　　　」。

5.4-1 何謂信賴區間？（Confidence Intervals）

- 任何量測值或計算值都可以標註其________，代表這個量的精確程度。
- 標準偏差又稱為__________。
- 「　　　　　　」的觀念是標準誤差的延伸，標準誤差只是_________一種特殊的表示方式而已。
- Example：經抽樣10個瓷磚並量測厚度如下：

9.984	9.971	10.113	10.022	9.972
9.956	10.036	10.047	9.986	9.975

- 這10個數據的平均值 $\bar{y}$ 是10.0062，標準偏差 S 是0.0483。
- 如何描述這個製程的誤差？

描述誤差的方式

以前例而言，

- 誤差是__________。
- 標準誤差是___________。
- 但是所謂「標準」是何意義？
- 更具體的描述方式是：
  - 這個製程的瓷磚厚度有 α% 的機率會在 a 到 b 之間。
  - α% 稱為____________（或稱為信心水準，confidence level）。
  - a 到 b 的區間稱為_____________（confidence interval, CI）。
  - a 與 b 稱為_____________（confidence limits, CL）。

假設

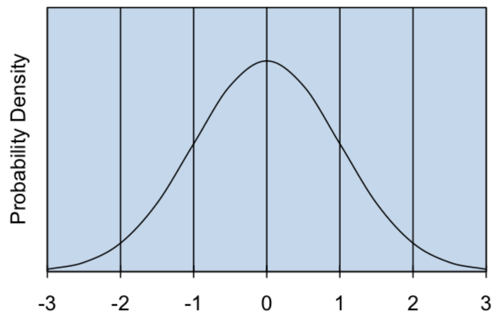
- 瓷磚厚度是呈______________。
- 抽樣檢測是__________。
- 這10個抽樣數據的標準偏差及平均值可以代表_________標準偏差及平均值。
  - 以這麼_________樣本所計算的標準偏差及平均值，來代表_________標準偏差及平均值似乎有點冒險：
  - 樣本標準偏差與母體標準偏差__________的機會是存在著。
  - 樣本平均值與母體平均值__________的機會也是存在著。
- 以下的討論先忽略這些不確定性，之後再分別針對標準偏差及平均值的不確定性來修正。

瓷磚厚度的分佈曲線

- 在上述假設下，這個製程的量測數據y會呈現以$\bar{y}$（10.0062）為中心，標準誤差為S（0.0483）的常態分佈。
- 或用另一種相同的說法：

（5.4-1式）

會呈現標準常態分佈。



- 因為標準偏差S的不確定性，5.4-1式會呈現________而非標準常態分佈。

以信賴水準/信賴區間來描述誤差

- 常態分佈曲線下，涵蓋 1 個標準偏差區間的面積是________，而涵蓋 2 個標準偏差區間的面積是________。
- 因此，我們可以描述製程誤差如下：「這個製程的瓷磚厚度有_______的機率會在_______之間，或「這個製程的瓷磚厚度有________的機率會在_________之間」。
- 一般而言，在α% 的信賴水準下，製程的信賴界限 CL（亦即誤差範圍）是：

95%信賴水準下的信賴區間

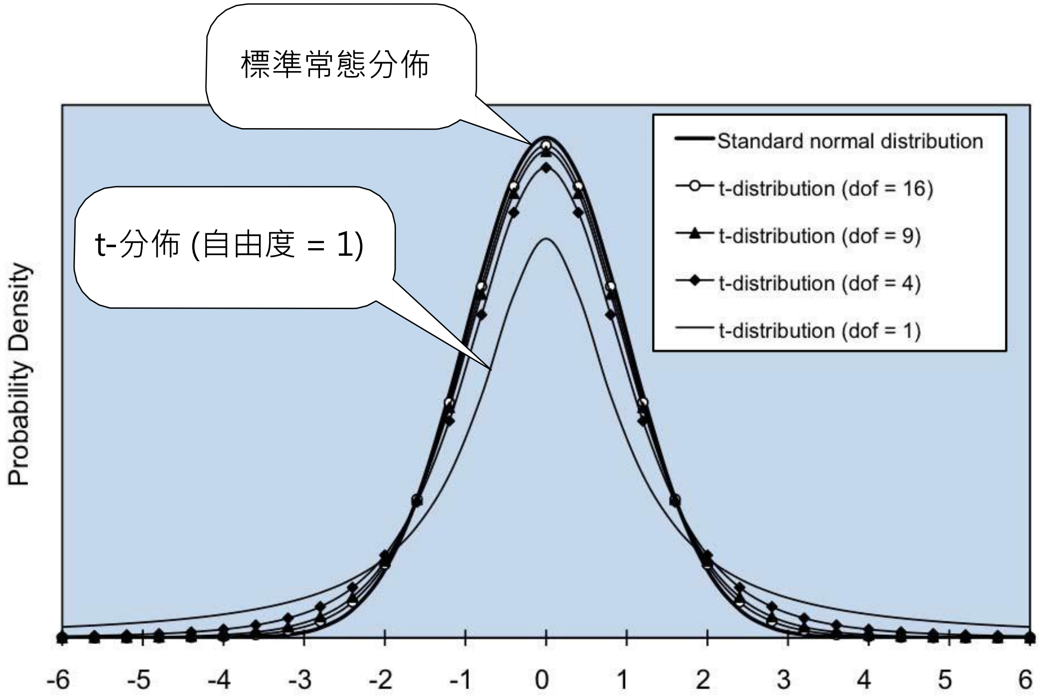
- 工程應用上最常使用的信賴水準是95%，在此情況下5.4-2式可以寫成：在95% 的信賴水準下，此製程的信賴界限是

- 換句話說，此製程瓷磚厚度有95%的機率會在__________________之間。
- 因為_______________，所以95%信賴區間又稱為二倍標準偏差信賴區間（two- standard-deviation confidence interval）；亦即，在95%的信賴水準下，信賴界限為

5.4-2 標準偏差的不確定性：t分佈（t-Distributions）

- 由於標準偏差估計值的不確定性，5.4-1式會呈現_______而非標準常態分佈。
- t分佈是考慮標準偏差估計值的不確定性而對標準常態分佈修正的結果。
- 當標準偏差估計值很可靠時，t分佈和標準常態分佈是一樣的；可是當標準偏差估計值並不可靠時，t分佈和標準常態分佈會開始偏離。
- 標準偏差估計值的可靠性取決於_____________，___________則越可靠。
- Excel有兩個與 t分佈有關的函數：

t 分佈


標準常態分佈
t-分佈 (自由度 = 1)
Standard normal distribution
t-distribution (dof = 16)
t-distribution (dof = 9)
t-distribution (dof = 4)
t-distribution (dof = 1)
Probability Density
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6

信賴區間的修正

- 考慮標準偏差的不確定後，5.4-2式必須加以修正，亦即以 t分佈來取代標準常態分佈：在 α%的信賴水準下，製程的信賴界限 CL是

其中 dof是計算標準偏差 S時使用到的資料（偏離量）的自由度。

信賴區間的修正

- 在本例中dof = __________，其中 n (=______) 是樣本數目。
- 在95%的信賴水準下，信賴界限是：

- 換句話說，此製程的瓷磚厚度有95%的機率會在___________________之間。
- 與5.4-3式比較，95%信賴區間由10.0062±0.0947擴大到上述範圍，這是因為考慮標準偏差的_____________，使得信賴區間____________，亦即變得______________。

信賴水準與信賴區間關係

- 信賴水準與信賴區間的關係是：信賴水準_______則信賴區間就_______。
- 若考慮平均值的不確定性，則信賴區間會散佈_______（____________）。

表5.4-1 信賴水準與信賴區間的關係

Confidence level	Confidence limits	Excel expression
75%	10.0062 ± 0.0483×1.2297 = 10.0062 ± 0.0594	= 10.0062 ± 0.0483×TINV(25%, 9)
90%	10.0062 ± 0.0483×1.8331 = 10.0062 ± 0.0885	= 10.0062 ± 0.0483×TINV(10%, 9)
95%	10.0062 ± 0.0483×2.2622 = 10.0062 ± 0.1093	= 10.0062 ± 0.0483×TINV(5%, 9)
99%	10.0062 ± 0.0483×3.2498 = 10.0062 ± 0.1570	= 10.0062 ± 0.0483×TINV(1%, 9)
99.9%	10.0062 ± 0.0483×4.7809 = 10.0062 ± 0.2309	= 10.0062 ± 0.0483×TINV(0.1%, 9)

不良率重新計算

- t分佈和標準常態分佈在兩尾端有明顯差異，這在計算不良率時會影響很大。
- 比較表2.1-3與表5.4-2的不良率（由 0.207%變成1.291%），我們獲得一個 結論：在樣本太少的情況下，若考慮標 準偏差的不確定性，所計算的不良率也__________（_____）。
- 若再加上平均值的不確定性，不良率會_______。
- 在2.1-5小節的例子中，樣本共有200個，因此所估計的標準偏差及平均值應該都________，不良率的估計（0.126%）應該是________。

表5.4-2 不良率的計算

Average of 10 data	=	10.0062
Standard Deviation	=	0.0483
Target	=	10.00
Tolerance	=	0.15
LSL	=	9.85
USL	=	10.15
% below *LSL*	=	0.514%
% above *USL*	=	0.777%
Total defect rate	=	1.291%

5.4-3 平均值的不確定性
（Uncertainty of Averages）

平均值的信賴區間

- 平均值$\bar{y}$（10.0062）是由n（10）個量測數據取平均而得的。
- 根據中間極限值定理，在隨機抽樣的假設下，此平均值的標準偏差估計值是________，其中S 是__________的標準偏差（0.0483）。
- 因此，真正（亦即母體）的平均值有α% 的機率會在以下的區間內：

平均值的信賴區間

- 我們將結論再整理如下：若平均值 $\bar{y}$是由 n個數據取平均而來的，而此 n個數據的標準偏差是S（其自由度為 dof），則 $\bar{y}$的 α%信賴界限是

- 譬如，瓷磚厚度的平均值（10.0062）的95%信賴界限是

5.4-4 因子反應值的信賴區間

- 在瓷磚實例中，A因子在水準1時的反應值是 10.016 mm，如何計算此反應值的 α% 信賴區間？
- 5.4-8式可以應用在因子反應值的信賴區間的計算，但是修改為下列形式更符合符號的使用：

其中 η 為_______________，dof 是計算標準偏差 S 時的自由度，而 m 是計算此反應值時的___________。

A, B 因子反應值的95%信賴區間

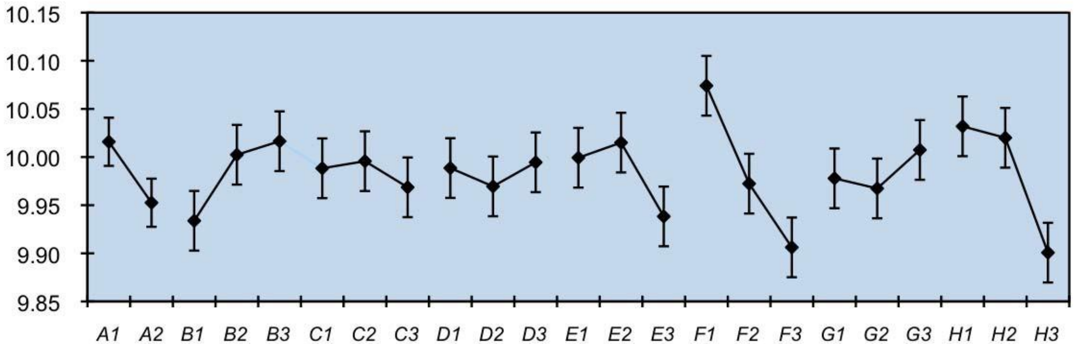
- 反應值10.016是由_______個數據中的______個數據取平均而得到的，所以 $m =$______。
- 這些實驗數據的標準偏差估計值是 $S = 0.1002$，自由度是 dof $= 116$（表5.3-2），故此反應值的95%信賴界限是

- B因子在水準1時的反應值（9.934, $m =$______）的95%信賴區間是

因子反應值的95%誤差

表5.4-3 瓷磚實例中因子反應值的95%誤差值

Level	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
1	10.016	9.934	9.988	9.989	9.999	10.074	9.978	10.032
2	9.953	10.002	9.996	9.970	10.015	9.972	9.967	10.020
3		10.016	9.969	9.995	9.938	9.906	10.007	9.901
95% Error	0.025	0.031	0.031	0.031	0.031	0.031	0.031	0.031



For each factor, why the 95% errors of its levels are the same?

S/N反應值的信賴區間

- A因子在水準1時的反應值43.10 dB是由18個數據中的_____個數據取平均而得到的，所以m =______。
- 此18個數據的標準偏差估計值是S = 1.47 dB ，自由度是 dof = 8（表5.3-6），故此反應值的 95%信賴界限是

$$
\begin{aligned}
CL &= 43.10 \pm \frac{1.47}{\sqrt{9}} \times \mathrm{TINV}(5\%, 8) \\
&= 43.10 \pm 0.490 \times 2.306 \\
&= 43.10 \pm 1.130 \text{ dB}
\end{aligned}
$$

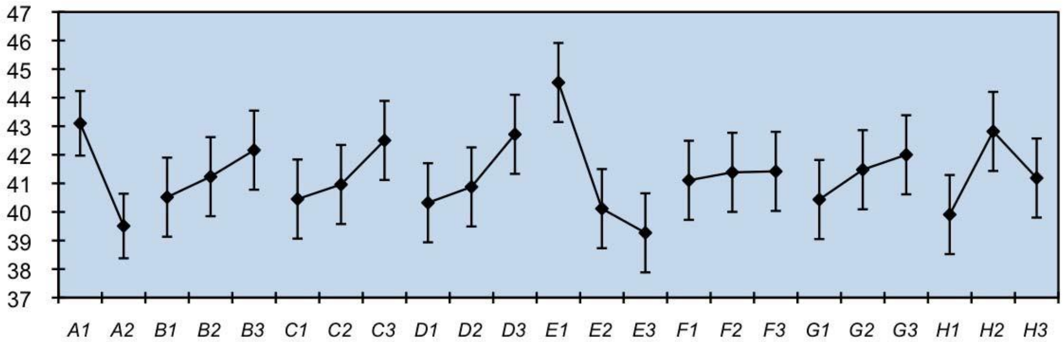
- B因子在水準1時的反應值（40.52 dB, m =_____）的95%信賴區間

$$
\begin{aligned}
CL &= 40.52 \pm \frac{1.47}{\sqrt{6}} \times \mathrm{TINV}(5\%, 8) \\
&= 40.52 \pm 0.600 \times 2.306 \\
&= 40.52 \pm 1.384 \text{ dB}
\end{aligned}
$$

因子反應值的95%誤差（S/N）

表5.4-4 瓷磚實例中*S/N*反應值的95%誤差值

Level	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
1	43.10	40.52	40.45	40.32	44.53	41.11	40.44	39.91
2	39.51	41.24	40.96	40.88	40.12	41.39	41.48	42.82
3		42.16	42.50	42.72	39.27	41.42	42.00	41.19
95% Error	1.130	1.384	1.384	1.384	1.384	1.384	1.384	1.384



5.4-5 預測值的信賴區間

- 在瓷磚實例中，我們預測在最佳製程組合下（A1 C3 D3 E1 H2，只考慮重要因子），其 S/N 比是50.4 dB。如何計算此預測值的 α% 信賴區間？
- 5.4-9中的樣本數目m在此處沒有具體數目可以代入。
- 觀察此預測值的計算方式：

- 上式中等號右邊的每一個數值都是某幾個 S/N比的平均值。
- 利用上述概念去計算一個「等效」樣本數目m_e：

預測值的信賴區間

- 在瓷磚實例中

  亦即， $m_e =$________ 。

- 以m_e取代5.4-9式中的m，則預測值的 α%信賴界限是

預測值的信賴區間

- 在瓷磚實例中，最佳設計下預測值的95%信賴界限是

- 同理，在原始製程組合下（A2 C2 D2 E2 H2，只考慮重要因子），其S/N 比的預測值是39.1 dB；此預測值的95%信賴界限是

5.4-6 確認實驗計算值的信賴區間

- 在瓷磚實例中，在最佳製程組合下的確認實驗，總共14個數據的S/N 比計算值是50.1 dB（表1.2-9）。
- 這個計算值的信賴區間可以使用5.4-9式，但是其中樣本數目m要改為實驗重複次數m_r。
- 在本例中，每次量測7個數據，共重複2次，所以$m_r = 2$；此確認值的α% 信賴界限是

確認實驗計算值的信賴區間

- 在最佳製程組合下，確認實驗的S/N 比（50.1 dB）的95%信賴界限是

- 同理，在原始製程組合下確認實驗的S/N 比（38.6 dB）的95%信賴界限是

5.4-7 預測值與確認實驗計算值的比較

- 如果預測值與確認實驗值相當一致，我們就可以認定實驗模式是可靠的。
- 我們以5.4-13式的預測值（50.4 dB）與5.4-15式的確認實驗值（50.1 dB）為例來說明。
- 這兩個數值各有其不確定性（其95%誤差分別為2.53 dB及2.40 dB）。
- 當兩個具不確定性的數值相加減時，其變異數（variance）必須疊加起來。
- 預測值的標準偏差是__________（5.4-12式），而確認實驗值的標準偏差是_________（5.4-14式）。
- 因此，預測值與確認實驗值差異的變異數是

預測值與確認實驗計算值的比較

- 將此變異數取代5.4-9式中的變異數部份，我們可以得到「預測值與確認實驗值差異」的 α%容許誤差值，或 α%信賴區間：

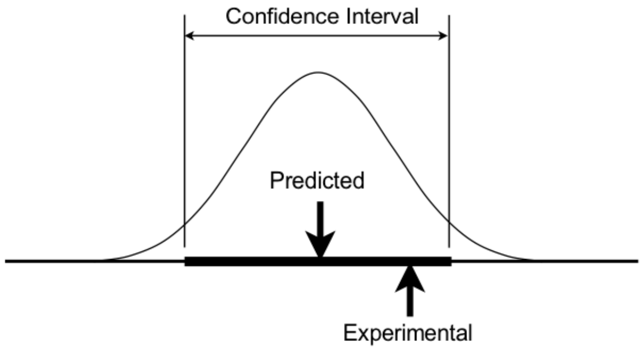
- 若預測值與確認實驗值差異________此容許值時，我們就必須懷疑實驗模式是否足夠精確；若預測值與確認實驗值差異________此容許值，我們就可以認為實驗模式是夠精確的。

預測值與確認實驗計算值的比較

- 本例中預測值與確認實驗值的差異是0.3 dB，而95%容許誤差值是

$$CI = \pm\sqrt{2.53^2 + 2.40^2} = \pm 3.49 \text{ dB}$$

- 所以兩個的差異可以認為是__________，實驗模式是______________。
- 在原始製程組合下，預測值（39.1 dB）與確認實驗值（38.6 dB）是足夠接近的。
- 若以預測值為中心，標出信賴區間（容許誤差值），若確認實驗值在此區間內，則我們可以認定兩個數值是足夠接近的。



5.4-8 t分佈與 F分佈的關係

- 在田口博士的原著中，所有計算信賴界限的公式全部寫成F 分佈的形式：田口博士以 $\sqrt{\mathrm{FINV}(1-\alpha\%, 1, \mathrm{dof})}$ 來代替$\mathrm{TINV}(1-\alpha\%, \mathrm{dof})$，譬如，將5.4-9式寫成

$$CL = \bar{\eta} \pm \frac{S}{\sqrt{m}} \times \sqrt{\mathrm{FINV}(1-\alpha\%, 1,\ dof)} \qquad (5.4\text{-}17式)$$

- 可以利用Excel驗証下列 t分佈與 F分佈的關係

$$\mathrm{TINV}(1-\alpha\%, dof) = \sqrt{\mathrm{FINV}(1-\alpha\%, 1,\ dof)} \qquad (5.4\text{-}18式)$$

- 用$\sqrt{\mathrm{FINV}(1-\alpha\%, 1, \mathrm{dof})}$ 來代替$\mathrm{TINV}(1-\alpha\%, \mathrm{dof})$ 的最大好處是可以利用F分佈數值表（附錄B.3）而不用再另外製作 t分佈數值表。
- 本書不沿用田口博士的公式，因為使用 t分佈比較容易理解其統計上的意義。

t分佈與常態分佈的關係

- 此外，有些教科書以標準常態分佈$\mathrm{NORMSINV}\left((1-\alpha\%)/2\right)$來代替 t分佈 $\mathrm{TINV}(1-\alpha\%,\mathrm{dof})$。
- 這是因為一方面不要讓太多統計課題困擾了讀者，另一方面其誤差常常是可以接受的。
- 前面提過，當 dof很大時，兩者會相等，亦即

$$\mathrm{TINV}(1-\alpha\%,dof)\approx \mathrm{NORMSINV}\left(\frac{1-\alpha\%}{2}\right) \qquad (5.4\text{-}19式)$$

- 甚至在dof 不是很大的情況下，其誤差在工程實務上也不至於太大。