Chapter 10 假設檢定

李水彬

Chien Hsin University of Science and Technology

2014.02

愛因斯坦: 一個從未犯錯的人是因爲他不曾嘗試新鮮事物。

Anyone who has never made a mistake has never tried anything new.

- 統計假設檢定 (hypothesis testing) 是研究者將母體參數空間劃分成兩類, 並陳述他認爲眞正的參數是屬於這兩類其中一, 做出假設性的陳述;
- 利用隨機抽樣方式收集本資訊進行驗證其陳述是否爲眞的一個過程,
- 因爲其判定是基於樣本資訊, 故稱此過程爲統計檢定。

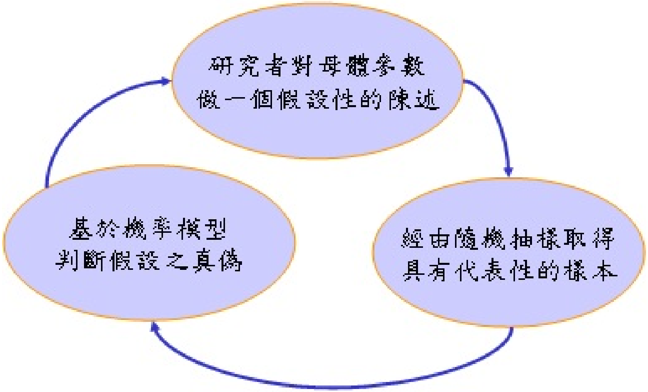

圖: 統計假設檢定是一個循環的過程。

- 一些生產因素可能隨時間推移而改變, 例如, 原物料品質, 機器老化等。品管人員擔心某一台機器, 或某一條生產線的生產狀況可能會偶而發生不穩定的情形, 所以, 定期檢測產品的品質特性變數 (quality characteristic variable) 以監控 (monitor) 製程的穩定性。(參數空間爲穩定和不穩定, 或品質特性變數的範圍)
- 爲了確保供應商的元件品質, 品管部門訂定抽樣計畫用以判定供應商所供應的元件批量是否可以允收。(參數空間爲合格和不合格, $0 \leq p \leq 1$)

- 藥廠開發新的降血壓藥品, 他們相信這新的藥比舊藥更爲有效, 在時效上與降幅皆是如此, 然而這需要一些臨床數據支持, 才能建議醫師開立新藥取代舊藥, 改善病人的生活品質。
- 小學生中, 老師普遍認爲男學生比較好動, 在課堂上無法專心聽講, 是以有人認爲男學生的學習成效較女學生爲低。但事實是否如此, 果眞男學生比較好動, 與學習成效是否有關呢? 這並不容易得到一個普遍性的結論。

- 星期五接近週末，公司員工是否受到預期放假的影響，使得產量低於其它時間呢？這可以分析生產數據得到一些佐證。
- 某公司爲了增加訂單，對外宣稱其工廠生產的成品良率可達 95%，高於一般同業水準。

太極生兩儀, 天下本爲渾沌, 一個假設性的陳述雖僅講述一面『比較好』, 另一面卻同時存在。

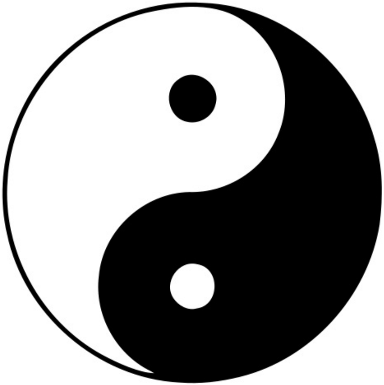

圖: 一個陳述表現出兩種假設。

- 假設檢定必須在兩者選一。兩儀爲陰陽二爻，假設也有兩個。**虛無假設** (null hypothesis, H_0)和**對立假設** (alternative hypothesis, H_1)。
- 何者爲虛無，何者爲對立，並非全然確定，這關乎**研究者的立場與目的**。基本上，我們想要證實爲眞的那個面向，那個假設稱爲對立假設，我們想要否定的，稱爲虛無假設。
- 角度不同，立場不同，背景不同，對同樣的問題在虛無與對立假設的設定上就有可能不同。

Example

假設某手機業者宣稱其手機之平均待機時間爲96小時, 請問消費者欲檢定此手機業者之宣稱是否爲真, 請問該如何假設?

♠

令 μ 爲此業者手機的平均待機時間, 這兩個假設可以是

$$(a)\begin{cases} \text{虛無假設} & H_0 : \mu \geq 96 \\ \text{對立假設} & H_1 : \mu < 96 \end{cases}$$

也可以是

$$(b)\begin{cases} \text{虛無假設} & H_0 : \mu \leq 96 \\ \text{對立假設} & H_1 : \mu > 96 \end{cases}$$

假設設立的原則, 提出以下幾點:

- 意欲否定的假設爲虛無假設。
- 意欲證實爲眞的假設爲對立假設。一般而言, 我們總是希望發現差異的存在, 或證實經過一些努力後情況和過去不同。
- 以問題陳述的反面主張爲虛無假設。反面指保持原狀沒有改變之意, 無差別和沒有差異, 正面表是進步提升, 有所改變之意。例如, 新藥與舊藥的療效沒有差別, 現在的製程狀態和過去一樣。

- 錯誤拒絕某一假設後果較爲嚴重，此一假設爲虛無假設。
- 他人的主張爲虛無假設。

- 關於第四點, 在一個民主法治國家, 強調無罪推定。一個嫌疑犯在被判定有罪之前, 都被認爲是無罪的, 有罪是需要證實的。
- 關於第五點, 他人就是有別於己。研究者斷無利用自己資源去證實別人陳述爲眞的道理。對於他人, 我們只有想要去否定他。意欲否定者, 就是虛無假設。

不拒絕 H_0： 另說無法證實 H_1 爲真，原有的虛無假設需要修正。就資料的顯示，無充分的證據與信心認爲虛無假設是不對的 (do not reject H_0)。

拒絕 H_0： 另說證實 H_1 爲真，也就是研究者證實自己想要的結果。就資料顯示，已經有充分的證據與信心認爲虛無假設是不對的 (reject H_0)。

統計檢定是企圖用樣本統計量拒絕 H_0 的假設, 然而當證據無法證實 H_0 爲假時, 我們通常說無法拒絕 H_0。然而, 無法拒絕 H_0 並不是代表 H_0 爲眞, 只是就目前資料而言, 我們沒有充分的證據他人有關 H_0 的陳述。

定義

檢定統計量是一個做爲判定是否拒絕 H_0 的統計量。利用樣本 $X_1, \cdots, X_n$ 計算而得的數值稱爲統計量,此值做爲檢定之依據故稱爲檢定統計量。

♠

定義

拒絕域是一個集合 *(範圍)* C, 若檢定統計量落在 C 則拒絕虛無假設, 反之, 則無法拒絕 *(接受)* 虛無假設。

♠

Example

- 擲一個的銅板 4 次, 請問如何判定這個骰子是否公平?
- 銅板實驗的參數爲出現正面的機率, 以 p 表示, 其參數空間爲 $0 \leq p \leq 1$,
- 當 $p = 0.5$ 時, 認定這是一個公平的銅板。

♠

1. 兩個假設

$$\begin{cases} \text{虛無假設} & H_0 : \text{銅板是公平的} \\ \text{對立假設} & H_1 : \text{銅板不是公平的} \end{cases}$$

亦可寫成

$$\begin{cases} \text{虛無假設} & H_0 : p = 0.5 \\ \text{對立假設} & H_1 : p \neq 0.5 \end{cases}$$

2 令 X_i 代表第 i 次投擲的結果, 正面爲 1, 反面爲 0。此一實驗的樣本空間

$$\begin{aligned}\Omega = \; & \{0000, 1000, 0100, 0010, 0001, \\ & 1100, 1010, 1001, 0110, 0101, 0011, \\ & 1110, 1101, 1011, 0111, 1111\}\end{aligned}$$

3 令正面出現的次數

$$X = X_1 + X_2 + X_3 + X_4$$

爲檢定統計量。

4 拒絕域

$$C_1 = \{0, 1, 3, 4\}$$

或

$$C_2 = \{0, 4\}$$

有何差異?

Example

- 若 $\bar{X} = 94.7$ 是否可以否定了業者的宣稱?
- 拒絕域 $C_1 = \{\bar{X} < 95\}$
- 拒絕域改爲 $C_2 = \{\bar{X} < 94\}$。

Table: 統計決策的結果

決策	真實情況	
	H_0	H_1
拒絕 H_0	**型 I 錯誤**	正確
不拒絕 H_0	正確	**型 II 錯誤**

兩種錯誤, 分述如下:

- **型 I 錯誤:** 當 H_0 爲眞時, 我們拒絕 H_0 之宣稱, 此時發生的錯誤, 稱之爲型 I 錯誤。定義如下:
- 通常用希臘字母 α 表示型 I 錯誤的機率, 稱之爲顯著水準。

$$\alpha = P\left(\text{拒絕 } H_0 \mid H_0 \text{ 爲眞}\right) \tag{1}$$

- 品管檢驗稱爲型 I 錯誤的機率爲**生產者風險**。

Example

例如, 供應商的貨品實際上是合格的, 但檢驗時因爲抽到比較多的不良品造成誤判, 認爲這批貨不合格而退貨, 造成供應商的損失, 故稱爲生產者風險。

- **型 II 錯誤:** 當 H_1 爲眞時, 我們拒絕 H_1 之宣稱, 此時發生的錯誤, 稱爲型 II 錯誤。
- 通常用希臘字母 β 表示型 II 錯誤的機率。定義如下:

$$\beta = P\,(\text{不拒絕 } H_0 \mid H_1 \text{ 爲眞}) \tag{2}$$

- 品管檢驗稱爲型 II 錯誤的機率爲**消費者風險**。

Example

供應商的貨品實際上是不合格的，但檢驗時因爲抽到比較少的不良品造成誤判，認爲這批貨合格而進貨，造成消費者的損失，故稱爲消費者風險。

Example

已知常態母體之變異數為 16, 今對此母體平均數作以下之假設

$$H_0 : \mu = 5 \text{ vs. } H_1 : \mu \neq 5$$

並決定其拒絕域為

$$C = \{\bar{X} > 5.8 \text{ 或 } \bar{X} < 4.2\}$$

請問以一樣本個數為 25 之樣本平均數來檢定母體平均數, 求型 I 錯誤之機率 α 為何?

樣本平均數的抽樣分配爲

$$\bar{X} \sim N\left(5, \frac{4^2}{25}\right)$$

$$\begin{aligned}
\alpha &= P\left(\bar{X} > 5.8 \text{ 或 } \bar{X} < 4.2 \mid H_0\right) \\
&= 1 - P\left(4.8 \leq \bar{X} \leq 5.8 \mid H_0\right) \\
&= 1 - P\left(-1 \leq Z \leq 1\right) \\
&= 1 - (\Phi(1) - \Phi(-1)) = 2(1 - \Phi(1)) \\
&= 2(1 - 0.8417) = 2 \times 0.1587 = 0.3174
\end{aligned}$$

標準分數的公式:

$$z = \frac{\sqrt{n}(x-\mu)}{\sigma}$$

左側 4.2 的標準分數

$$z_L = \frac{\sqrt{25}(4.2-5)}{4} = -1$$

右側 5.8 的標準分數

$$z_U = \frac{\sqrt{25}(5.8-5)}{4} = 1$$

Example

接續例 9

若 $\mu = 6.6$ 時, 求型 II 錯誤之機率 β 爲何?

♠

樣本平均數的抽樣分配為

$$\bar{X} \sim N\left(6.6, \frac{4^2}{25}\right)$$

$$\begin{aligned}
\beta &= P\left(4.2 \le \bar{X} \le 5.8 \mid \mu = 6.6\right) \\
&= P\left(-3 \le Z \le -1\right) \\
&= \Phi(-1) - \Phi(-3)) = \Phi(3) - \Phi(1) \\
&= 0.9985 - 0.8417 = 0.1568
\end{aligned}$$

左側 4.2 的標準分數

$$z_L = \frac{\sqrt{25}(4.2 - 6.6)}{4} = -3$$

右側 5.8 的標準分數

$$z_U = \frac{\sqrt{25}(5.8 - 6.6)}{4} = -1$$

Example

接續例9
若拒絕域改爲

$$C = \{\bar{X} > 6.6 \text{ 或 } \bar{X} < 3.4\}$$

求型 I 錯誤之機率 α 爲何?

♠

樣本平均數的抽樣分配爲

$$\bar{X} \sim N\left(5, \frac{4^2}{25}\right)$$

$$\begin{aligned}
\alpha &= P\left(\bar{X} > 6.6 \text{ 或 } \bar{X} < 3.4 \mid H_0\right) \\
&= 1 - P\left(3.4 \leq \bar{X} \leq 6.6 \mid H_0\right) \\
&= 1 - P\left(-2 \leq Z \leq 2\right) \\
&= 1 - (\Phi(2) - \Phi(-2)) = 2(1 - \Phi(2)) \\
&= 2(1 - 0.9772) = 2 \times 0.0228 = 0.0456
\end{aligned}$$

左側 3.4 的標準分數

$$z_L = \frac{\sqrt{25}(3.4 - 5)}{4} = -2$$

右側 6.6 的標準分數

$$z_U = \frac{\sqrt{25}(6.6 - 5)}{4} = 2$$

而型 II 錯誤之機率 β 爲何?

樣本平均數的抽樣分配為

$$\bar{X} \sim N\left(6.6, \frac{4^2}{25}\right)$$

$$\begin{aligned}
\beta &= P\left(3.4 \leq \bar{X} \leq 6.6 \mid \mu = 6.6\right) \\
&= P\left(-4 \leq Z \leq 0\right) \\
&= \Phi(0) - \Phi(-4) = \Phi(4) - \Phi(0) \\
&= 1 - 0.5 = 0.5
\end{aligned}$$

左側 3.4 的標準分數

$$z_L = \frac{\sqrt{25}(3.4 - 6.6)}{4} = -4$$

右側 6.6 的標準分數

$$z_U = \frac{\sqrt{25}(6.6 - 6.6)}{4} = 0$$

關於型 I 錯誤的機率與型 II 錯誤的機率的一些討論。

1. 決策犯錯的無可避免的, 有正確判斷的可能, 但也有犯錯的機會。當然, 決策者於當時都相信他的決策是正確的。
2. 我們希望犯錯的機率愈小愈好, 即 α 與 β 愈接近0愈好, 但這需要成本, 統計上的成本就是樣本大小, 樣本愈大成本愈高, 推論更爲精確而使 α 與 β 變小。

1. 然而在樣本大小 (n) 給定的情況下, α 愈小則 β 愈大, 反之亦然, 兩者不可兼得。
2. 在假設設立原則, 錯誤拒絕虛無假設的後果較爲嚴重, 是以執行統計檢定必須控制型 I 錯誤的機率。通常的做法爲設定 α 後, 然後再決定拒絕域。稱 α 爲一個統計檢定顯著水準。

一個好的檢定統計量能在 H_0 爲錯誤時，拒絕 H_0，達成研究者的目的。一個檢定統計量的檢定力爲正確拒絕 H_0 的能力，在 H_1 爲眞的條件下，檢定統計量證實其爲眞的能力。在固定 α 下，檢定力爲判定檢定統計量良窳的指標。

定義

統計量的檢定力的定義:

$$\begin{aligned} 1-\beta = & \; 1-P\,(\text{不拒絕 } H_0 \mid H_1 \text{ 爲眞}) \\ = & \; P\,(\text{拒絕 } H_0 \mid H_1 \text{ 爲眞}) \end{aligned} \qquad (3)$$

♠

假設兩個檢定方法 C_1 和 C_2 有相同的型 I 錯誤的機率, 若 C_1 的檢定力 β_1 大於 C_2 的檢定力 β_2, 則檢定方法 C_1 較優。

Example

接續例 2, 我們以試驗排列情形定義以下兩個拒絕域

$$C_1 = \{0000, 1000, 0100, 0010, 0001, 1111\}$$

和

$$C_2 = \{0000, 1000, 0001, 1110, 0111, 1111\}$$

它們型 I 錯誤的機率皆為 $3/8$, 假設 $p = 0.6$, 請問兩個檢定方法的檢定力。

♠

Table: 兩個假設, 拒絕域, 顯著水準的關係表

	雙尾檢定	左尾檢定	右尾檢定
虛無假設 H_0 的符號	$=$	$\geq$	$\leq$
對立假設 H_1 的符號	$\neq$	$<$	$>$
拒絕域	在左右兩尾	在左尾	在右尾
α	$\alpha/2$	α	α

統計假設檢定的程序與第 **??** 章信賴區間類似, 都是基於點估計的方法與其抽樣分配, 分成以下四個步驟:

1. 決定合適點估計統計量
2. 推導檢定統計量在 H_0 爲眞下的的抽樣分配
3. 決定拒絕域型態: 以**左尾檢定**爲例, 配合對立假設拒絕域的型態爲

$$C = \{\bar{X} \leq c\}$$

4. 計算 c 值。

母體平均數檢定的三種類型:

- 右尾 (right-tailed) $\begin{cases} H_0: & \mu \leq \mu_0 \\ H_1: & \mu > \mu_0 \end{cases}$
- 左尾 (left-tailed) $\begin{cases} H_0: & \mu \geq \mu_0 \\ H_1: & \mu < \mu_0 \end{cases}$
- 雙尾檢定 (two-sided test) $\begin{cases} H_0: & \mu = \mu_0 \\ H_1: & \mu \neq \mu_0 \end{cases}$

1. 決定合適點估計統計量: $\bar{X}$
2. 推導檢定統計量 $\bar{X}$ 在 H_0 爲眞下的的抽樣分配:
   $$\bar{X} \sim N\left(\mu_0, \frac{\sigma^2}{n}\right)$$

3. 決定拒絕域型態: 以左尾檢定為例, 配合對立假設拒絕域的型態為

$$C = \{\bar{X} \leq c\}$$

   $\bar{x}$ 愈小愈傾向拒絕 H_0。

4. 計算 c 值。

因爲

$$\alpha = P(C \mid H_0) = P\left(\bar{X} \leq c \mid H_0\right)$$

所以,

$$\alpha = P\left(\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \leq \frac{c - \mu_0}{\sigma/\sqrt{n}}\right)$$

根據 z_α 的定義, 可得

$$\frac{c - \mu_0}{\sigma/\sqrt{n}} = -z_\alpha$$

因此,

$$c = \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}} \tag{4}$$

母體平均數左尾檢定的拒絕域爲

$$C = \left\{ \bar{X} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}} \right\} \tag{5}$$

Table: 臨界值法 (三種檢定類型的拒絕域)

	雙尾檢定	左尾檢定	右尾檢定
H_0	$H_0 : \mu = \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu \leq \mu_0$
H_1	$\neq$	$<$	$>$
拒絕域	$\begin{cases} \bar{X} < \mu_0 - z_{\alpha/2}\frac{\sigma}{\sqrt{n}} & \text{或} \\ \bar{X} > \mu_0 + z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \end{cases}$	$\bar{X} < \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}}$	$\bar{X} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}}$
α	$\alpha/2$	α	α

Example

假設某公司生產之巧克力重量的標準差爲 4 公克。該公司宣稱其巧克力淨重平均重量超過 100 公克, 今於販賣場隨機抽取該產品 25 件, 得其平均數爲 98.4 公克, 請以顯著水準 $\alpha = 0.05$ 檢定此廠商宣稱是否爲眞? (假設母體具常態分配)

♠

1. 設立兩個假設。

$$\begin{cases} H_0: & \mu \geq 100 \\ H_1: & \mu < 100 \end{cases}$$

2. 選擇檢定統計量。

$$\bar{X} = 98.4$$

3 決定拒絕域。

$$
\begin{aligned}
c &= \mu_0 - z_\alpha \frac{\sigma}{\sqrt{n}} \\
&= 100 - 1.645 \times \frac{4}{\sqrt{25}} = 98.684
\end{aligned}
$$

$$C = \left\{\bar{X} < 98.684\right\}$$

4 結論。
因爲 $\bar{X} = 98.4$ 落在 C 內, 故拒絕 H_0, 即證據足以否定廠商的宣稱。

樣本平均值標準化, 其分配爲 $N(0,1)$, 即

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \tag{6}$$

式5的標準化爲

$$C = \left\{ \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} < -z_\alpha \right\} = \{Z < -z_\alpha\} \tag{7}$$

僅要比較樣本平均數的標準值與標準常態的臨界值就可以進行統計檢定了。

Table: Z 值法 (三種檢定類型的拒絕域)

	雙尾檢定	左尾檢定	右尾檢定
H_0	$H_0 : \mu = \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu \leq \mu_0$
H_1	$\neq$	$<$	$>$
拒絕域	$Z < -z_{\alpha/2}$ 或 $Z > z_{\alpha/2}$	$Z < -z_\alpha$	$Z > z_\alpha$
α	$\alpha/2$	α	α

Example

接續例10
改用 Z 值法。

♠

1. 設立兩個假設。

$$\begin{cases} H_0: & \mu \geq 100 \\ H_1: & \mu < 100 \end{cases}$$

2. 選擇檢定統計量。(計算 $\bar{X}$ 的標準分數)

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} = \frac{98.4 - 100}{4/\sqrt{25}} = -2$$

3. 決定拒絕域。

$$\alpha = 0.05, z_{0.05} = 1.645$$

$$C = \{Z < -1.645\}$$

4. 結論。
因爲 $Z = -2$ 落在 C 內, 故拒絕 H_0, 即證據足以否定廠商的宣稱。

定義

令檢定統計量 Z 值爲 z_0 其 P_{value} 的定義如下

$$\begin{array}{ll} \text{右尾:} & P_{value} = P\left(Z > z_0\right) \\ \text{左尾:} & P_{value} = P\left(Z < z_0\right) \\ \text{雙尾:} & P_{value} = 2P\left(Z >\mid z_0 \mid\right) \end{array} \tag{8}$$

♠

以右尾檢定爲例, 已知 $P(Z > z_\alpha) = \alpha$, 若 $z_0 > z_\alpha$ 拒絕 H_0, 則若且唯若 $P_{value} = P(Z > z_0) < \alpha$。其他兩種類型的檢定, 亦可以相同的結論。

定理

若檢定統計量的 $P_{value} < \alpha$ 則若且唯若拒絕 H_0。我們稱此法爲 P 值法。

♠

Example

接續例11
改用 P 值法。

♠

這是一個左尾檢定, 因爲 $z_0 = -2$, 所以,

$$P_{value} = P\left(Z < -2\right) = 0.0228$$

因爲 $P_{value} < 0.05$, 故拒絕 H_0, 即證據足以否定廠商的宣稱。

任何一種檢定包含往後介紹的方法都有對應的 P 值法, 判斷準則都以統計量的 P_{value} 是否小於顯著水準 α。所以, 電腦統計軟體都會輸出一個檢定的 P_{value}, 只要 $P_{value} < \alpha$ 我們就可以拒絕 H_0。

因爲

$$\frac{\sqrt{n}(\bar{X}-\mu_0)}{S} \sim t(n-1) \tag{9}$$

將標準常態臨界值 z_α 或 $z_{\alpha/2}$, 改成 $t_\alpha(n-1)$ 或 $t_{\alpha/2}(n-1)$ 即可。整理在表 5。

Table: 母體變異數未知的臨界值法 ($t_\alpha = t_\alpha(n-1)$)

	雙尾檢定	左尾檢定	右尾檢定
虛無假設 H_0	$H_0 : \mu = \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu \leq \mu_0$
對立假設 H_1	$\neq$	$<$	$>$
拒絕域	$\begin{cases} \bar{X} < \mu_0 - t_{\alpha/2}\dfrac{S}{\sqrt{n}} \text{ 或} \\ \bar{X} > \mu_0 + t_{\alpha/2}\dfrac{S}{\sqrt{n}} \end{cases}$	$\bar{X} < \mu_0 - t_\alpha \dfrac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + t_\alpha \dfrac{S}{\sqrt{n}}$
α	$\alpha/2$	α	α

樣本平均數標準化得

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t(n-1) \tag{10}$$

將表4轉成表6。

Table: T 值法 $t_\alpha = t_\alpha(n-1)$

	雙尾檢定	左尾檢定	右尾檢定
虛無假設 H_0	$H_0 : \mu = \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu \leq \mu_0$
對立假設 H_1	$\neq$	$<$	$>$
拒絕域	$\begin{cases} T < -t_{\alpha/2} \text{ 或} \\ T > t_{\alpha/2} \end{cases}$	$T < -t_\alpha$	$T > t_\alpha$
α	$\alpha/2$	α	α

Example

在半導體製造的 Si-Nitride 爐管擴散製程, 沉積厚度爲品質特性變數爲常態隨機變數廠商宣稱製程平均值爲 1650。今抽樣量測 5 片晶片的 Si-Nitride 沉積厚度爲

1623, 1670,1625, 1633,1649

請問證據是否支持廠商的宣稱? $\alpha = 0.05$

♠

1. 設設立兩個假設。

$$\begin{cases} H_0: & \mu = 1650 \\ H_1: & \mu \neq 1650 \end{cases}$$

2. 選擇檢定統計量。
先計算基本統計量

$$\bar{X} = \frac{1623 + 1670 + 1625 + 1633 + 1649}{5} = 1640$$

$$S^2 = \frac{(-17)^2 + 30^2 + (-15)^2 + (-7)^2 + 9^2}{4} = \frac{1544}{4}$$

$$s = \sqrt{386} = 19.65$$

所以, 檢定統計量爲

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\bar{X} - 1650}{19.65/\sqrt{5}} = -1.14$$

3 決定拒絕域。

$$\alpha = 0.05, t_{0.025}(4) = 2.776$$

$$C = \{T < -2.776 \text{ 或 } T > 2.776\}$$

4 結論。
因爲 $T = -1.14$ 不落在 C 內, 故無法拒絕 H_0, 即證據不足以否定廠商的宣稱。

檢定統計量的 P 值爲

$$P_{value} = 2P\left(T >| -1.14 |\right) = 2P\left(T > 1.14\right) = 0.31$$

因爲 $P_{value} > 0.05$, 故無法拒絕 H_0。

當樣本大小 $n \geq 30$, 根據中央極限定理, 樣本平均數的抽樣分配近似常態。即當 $\mu = \mu_0$ 時,

$$\bar{X} \sim N\left(\mu_0, \sigma^2/n\right) \tag{11}$$

當變異數已知時,

$$\frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1) \tag{12}$$

當變異數未知時, 亦有

$$\frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim N(0,1) \tag{13}$$

Table: 大樣本之臨界值法 (三種檢定類型的拒絕域)

	雙尾檢定	左尾檢定	右尾檢定
虛無假設 H_0	$H_0 : \mu = \mu_0$	$H_0 : \mu \geq \mu_0$	$H_0 : \mu \leq \mu_0$
對立假設 H_1	$\neq$	$<$	$>$
拒絕域	$\begin{cases} \bar{X} < \mu_0 - z_{\alpha/2}\dfrac{S}{\sqrt{n}} \text{ 或} \\ \bar{X} > \mu_0 + z_{\alpha/2}\dfrac{S}{\sqrt{n}} \end{cases}$	$\bar{X} < \mu_0 - z_{\alpha}\dfrac{S}{\sqrt{n}}$	$\bar{X} > \mu_0 + z_{\alpha}\dfrac{S}{\sqrt{n}}$
α	$\alpha/2$	α	α

Table: 大樣本 Z 值法 (三種檢定類型的拒絕域)

	雙尾檢定	左尾檢定	右尾檢定
虛無假設 H_0	$H_0: \mu = \mu_0$	$H_0: \mu \geq \mu_0$	$H_0: \mu \leq \mu_0$
對立假設 H_1	$\neq$	$<$	$>$
拒絕域	$Z < -z_{\alpha/2}$ 或 $Z > z_{\alpha/2}$	$Z < -z_\alpha$	$Z > z_\alpha$
α	$\alpha/2$	α	α

Example

接續例 13

今抽樣量測 36 片晶片, 得到相同樣本平均值與標準差, 即 $\bar{X} = 1640, s = 19.65$。請問證據是否支持廠商的宣稱? $\alpha = 0.05$

♠

1. 設設立兩個假設。

$$\begin{cases} H_0: & \mu = 1650 \\ H_1: & \mu \neq 1650 \end{cases}$$

2. 檢定統計量為

$$Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{\bar{X} - 1650}{19.65/\sqrt{36}} = -3.05$$

3. 決定拒絕域。

$$\alpha = 0.05, z_{0.025} = 1.96$$

$$C = \{Z < -1.96 \text{ 或 } Z > 1.96\}$$

4. 結論。
因爲 $Z = -3.05$ 落在 C 內, 故拒絕 H_0, 即證據足以否定廠商的宣稱。

註: 少數樣本反應的事實不夠顯著, 同樣的統計量若是來自大量樣本, 則有強化此一事實的意義, 故容易產生拒絕 H_0 的結果。

在大樣本下 $n \geq 30$, 根據中央極限定理樣本比例 $\hat{p}$ 具有近似常態分配, 但近似效果與 p 值有關。當 (n, p) 滿足

$$np \geq 5 \text{且} n(1-p) \geq 5 \tag{14}$$

時, $\hat{p}$ 近似常態分配方可成立。

樣本比例的近似分配爲

$$\hat{p} \sim N\left(p, \frac{\sigma_p^2}{n}\right) \tag{15}$$

其中$\sigma_p^2 = p(1-p)$。以左尾檢定

$$\begin{cases} H_0: & p \geq p_0 \\ H_1: & p < p_0 \end{cases} \tag{16}$$

爲例, 在 H_0 爲眞時 $(p = p_0)$, 樣本比例的抽樣分配爲

$$\hat{p} \sim N\left(p_0, \frac{p_0(1-p_0)}{n}\right) \tag{17}$$

將式 (4) 中的 σ 用 $\sigma_p = \sqrt{p_0(1-p_0)}$ 代入, 得

$$c = p_0 - z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}} \tag{18}$$

母體比例之左尾檢定的拒絕域爲

$$C = \left\{ \hat{p} < p_0 - z_\alpha \sqrt{\frac{p_0(1-p_0)}{n}} \right\} \tag{19}$$

Table: 母體比例之臨界值法拒絕域

	雙尾檢定	左尾檢定	右尾檢定
H_0	$H_0: p = p_0$	$H_0: p \geq p_0$	$H_0: p \leq p_0$
H_1	$\neq$	$<$	$>$
拒絕域	$\begin{cases} \hat{p} < p_0 - z_{\alpha/2}\dfrac{\sigma_{p_0}}{\sqrt{n}} \\ \text{或} \\ \hat{p} > p_0 + z_{\alpha/2}\dfrac{\sigma_{p_0}}{\sqrt{n}} \end{cases}$	$\hat{p} < p_0 - z_\alpha \dfrac{\sigma_{p_0}}{\sqrt{n}}$	$\hat{p} > p_0 + z_\alpha \dfrac{\sigma_{p_0}}{\sqrt{n}}$
α	$\alpha/2$	α	α

其中 $\sigma_{p_0} = \sqrt{p_0(1-p_0)}$

我們將樣本比例標準化得

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} \sim N(0,1) \tag{20}$$

將式 (19) 左尾檢定的拒絕域改爲如下 Z 值法的拒絕域

$$C = \{Z < -z_\alpha\} \tag{21}$$

因此, 用表 4 進行母體比例的假設檢定, 僅將虛無假設的 μ 和 μ_0 分別用 p 和 p_0 取代即可, 得表 10。

Table: 母體比例之 Z 值法拒絕域

	雙尾檢定	左尾檢定	右尾檢定
虛無假設 H_0	$H_0 : p = p_0$	$H_0 : p \geq p_0$	$H_0 : p \leq p_0$
對立假設 H_1	$\neq$	$<$	$>$
拒絕域	$Z < -z_{\alpha/2}$ 或$Z > z_{\alpha/2}$	$Z < -z_{\alpha}$	$Z > z_{\alpha}$
α	$\alpha/2$	α	α

Example

某民意調查機構欲瞭解某次中壢市市長選舉各候選人之支持度, 於是隨機抽取 300 位選民作調查, 發現支持現任市長者有 75 位, 請以顯著水準 $\alpha = 0.05$ 檢定現任市長的支持度是否低於 30% ?

1. 設立兩個假設。

$$\begin{cases} H_0: & p \geq 0.3 \\ H_1: & p < 0.3 \end{cases}$$

2 選擇檢定統計量。

$$\hat{p} = \frac{X}{n} = \frac{75}{300} = 0.25,$$

$$\frac{\sigma_p}{\sqrt{n}} = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0.3 \times 0.7}{300}} = 0.0265$$

$$Z = \frac{\hat{p} - p_0}{\frac{\sigma_p}{\sqrt{n}}} = \frac{0.25 - 0.3}{0.0265} = -1.87$$

3. 決定拒絕域。

$$\alpha = 0.05, z_{0.05} = 1.645$$

$$C = \{Z < -1.645\}$$

4. 結論。
因為 $Z = -1.87$ 落在 C 內, 故拒絕 H_0, 即證實現任市長的支持度是低於30%。

Example

在一次桃園縣政府之滿意度調查中, 有效樣本爲 1068 位, 共有 522 縣民不滿意縣政, 請以顯著水準 $\alpha = 0.05$ 檢定縣民對桃園縣政府的不滿意度是否等於 50%?

♠

1. 設立兩個假設。

$$\begin{cases} H_0: & p = 0.5 \\ H_1: & p \neq 0.5 \end{cases}$$

2. 選擇檢定統計量。

$$\hat{p} = \frac{X}{n} = \frac{522}{1068} = 0.489,$$

$$\frac{\sigma_p}{\sqrt{n}} = \sqrt{\frac{p_0(1-p_0)}{n}} = \sqrt{\frac{0.5 \times 0.5}{1068}} = 0.0153$$

$$Z = \frac{\hat{p} - p_0}{\frac{\sigma_p}{\sqrt{n}}} = \frac{0.489 - 0.5}{0.0153} = -0.72$$

3. 決定拒絕域。

$$\alpha = 0.05, z_{0.025} = 1.96$$

$$C = \{Z < -1.96 \text{ 或 } Z > 1.96\}$$

4. 結論。
因爲 $Z = -0.72$ 不落在 C 內, 故無法拒絕 H_0, 可說縣民對桃園縣政府的不滿意度達到50%。

Example

大高公司欲證實其高階產品良率是否已經達到量產階段的 75%, 抽測 200 個產品, 有 160 個良品, 請問該公司高階產品良率是否達到 75%?

♠

1. 設立兩個假設。

$$\begin{cases} H_0: & p \leq 0.75 \\ H_1: & p > 0.75 \end{cases}$$

2. 選擇檢定統計量。

$\hat{p} = \dfrac{X}{n} = \dfrac{160}{200} = 0.8,$

$\dfrac{\sigma_p}{\sqrt{n}} = \sqrt{\dfrac{p_0(1-p_0)}{n}} = \sqrt{\dfrac{0.75 \times 0.25}{200}} =$

$0.0306 \;\; Z = \dfrac{\hat{p} - p_0}{\frac{\sigma_p}{\sqrt{n}}} = \dfrac{0.8 - 0.75}{0.0306} = 1.634$

3. 決定拒絕域。

$$\alpha = 0.05, z_{0.05} = 1.645$$

$$C = \{Z > 1.645\}$$

另外,

$$p_{value} = P(Z > 1.634) = 0.051$$

4. 結論。
因爲 $Z = 1.636$ 不落在 C 內 (或者 $p_{value} = 0.051 > 0.05$), 故無法拒絕 H_0, 可說檢驗結果不足以證實該公司高階產品良率已經達到 75%。

推論的方式與第 **??** 章相同, 假設 $X_1, \cdots, X_n \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$, 則

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \tag{22}$$

變異數 (標準差) 假設檢定的類型有以下三種。

1. 左尾檢定
$$\begin{cases} H_0: & \sigma \geq \sigma_0 \\ H_1: & \sigma < \sigma_0 \end{cases} \Leftrightarrow \begin{cases} H_0: & \sigma^2 \geq \sigma_0^2 \\ H_1: & \sigma^2 < \sigma_0^2 \end{cases}$$
2. 右尾檢定
$$\begin{cases} H_0: & \sigma \leq \sigma_0 \\ H_1: & \sigma > \sigma_0 \end{cases} \Leftrightarrow \begin{cases} H_0: & \sigma^2 \leq \sigma_0^2 \\ H_1: & \sigma^2 > \sigma_0^2 \end{cases}$$
3. 雙尾檢定
$$\begin{cases} H_0: & \sigma = \sigma_0 \\ H_1: & \sigma \neq \sigma_0 \end{cases} \Leftrightarrow \begin{cases} H_0: & \sigma^2 = \sigma_0^2 \\ H_1: & \sigma^2 \neq \sigma_0^2 \end{cases}$$

我們以右尾檢定爲例, 推導檢定的拒絕域:

1. 決定合適點估計統計量:
$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$ 爲 σ^2 的不偏估計式。
2. 推導檢定統計量: 由式 (22), 在 H_0 爲眞的假設下, 我們令檢定統計量爲

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} \tag{23}$$

3. 拒絕域型態。配合對立假設拒絕域的型態爲

$$C = \{\chi_0^2 > c\} \tag{24}$$

4. 計算決定 c 值。因爲 c 值 滿足

$$P\left(\chi_0^2 > c \mid H_0\right) = \alpha \tag{25}$$

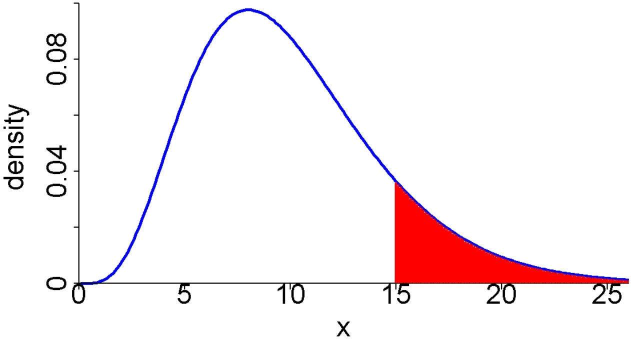

圖: 卡方分配的臨界值

根據卡方分配臨界值的定義, 如圖 3 紅色面積爲 α, 紅色面積左側橫軸端點爲 $c = \chi^2_\alpha(n-1)$ 拒絕域爲

$$C = \left\{\chi^2_0 > \chi^2_\alpha(n-1)\right\} \tag{26}$$

同理可得出其他兩種類型的拒絕域, 列於表11。

Table: 母體變異數檢定之拒絕域

	雙尾檢定	左尾檢定	右尾檢定
虛無假設 H_0	$H_0 : \sigma = \sigma_0$	$H_0 : \sigma \geq \sigma_0$	$H_0 : \sigma \leq \sigma_0$
對立假設 H_1	$\neq$	$<$	$>$
拒絕域	$\chi_0 < \chi^2_{1-\alpha/2}$ 或$\chi^2_0 > \chi^2_{\alpha/2}$	$\chi^2_0 < \chi^2_{1-\alpha}$	$\chi^2_0 > \chi^2_{\alpha}$
α	$\alpha/2$	α	α

定義

假設 χ_0^2 爲檢定統計量, 變異數假設檢定的 P_{value} 定義如下:

- 左尾檢定: $P_{value} = P\left(X < \chi_0^2\right)$
- 右尾檢定: $P_{value} = P\left(X > \chi_0^2\right)$
- 雙尾檢定:
  $P_{value} = 2\min\left\{P\left(X > \chi_0^2\right), P\left(X < \chi_0^2\right)\right\}$

♠

Example

假設某一製程在穩定狀況下其所生產產品標準差爲 0.34 公分, 今隨機抽出 5 個產品測得樣本標準差爲 0.8 公分。假設該產品品質特性爲常態分布, 請問製程是否偏離穩定狀況? $(\alpha = 0.05)$

♠

1. 設立兩個假設

$$\begin{cases} H_0: & \sigma = 0.34 \\ H_1: & \sigma \neq 0.34 \end{cases}$$

2. 檢定統計量

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{4 \times 0.8^2}{0.34^2} = 22.14$$

3. 決定拒絕域

$$\alpha = 0.05, \chi^2_{0.025}(4) = 11.14, \chi^2_{0.975}(4) = 0.4844$$

$$C = \{\chi^2_0 > 11.14 \text{ 或 } \chi^2_0 < 0.4844\}$$

4. 結論: 因爲 $\chi^2_0 = 22.14$ 落在 C 內, 故拒絕 H_0, 表示製程已經偏離穩定狀況。

Example

某工廠進螺帽一批，供應商宣稱此批貨的標準差小於 0.05。抽驗 10 個，測量每一個螺帽的直徑，如下：

1.20, 1.25, 1.27, 1.17, 1.30, 1.18, 1.23, 1.22, 1.33, 1.32

假設螺帽直徑爲常態分布，請問是否否定供應商宣稱？ ($\alpha = 0.05$)

♠

1 設立兩個假設

$$\begin{cases} H_0: & \sigma \le 0.05 \\ H_1: & \sigma > 0.05 \end{cases}$$

2 檢定統計量

$$S^2 = 0.00325 = 3.25 \times 10^{-3}$$

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{9 \times 3.25 \times 10^{-3}}{0.05^2} = 11.7$$

3. 決定拒絕域

$$\alpha = 0.05, \chi^2_{0.05}(9) = 16.92$$

$$C = \{\chi^2_0 > 16.92\}$$

4. 結論: 因爲 $\chi^2_0 = 11.7$ 不落在 C 內, 故無法拒絕 H_0, 證據不足否定供應商宣稱。

或者, 利用 P 値法

$$P_{value} = P\left(X > 11.7\right) = 0.264$$

因爲 $P_{value} > 0.05$, 故無法拒絕 H_0。

Example

圖 4 爲本系品質管理實驗室切割器構造圖, 可以用來切割圓條狀物品, 如電線、吸管等。現在設定機器長度爲 20 公厘, 使用連續裁切吸管 20 段, 使用圖 5 游標卡尺測得吸管長度紀錄如下, 假設裁切長度分布爲常態, 請問裁切長度是否偏離目標值 (20公厘)。

20.16, 20.01, 19.68, 19.70, 19.88,
20.00, 20.12, 19.86, 19.60, 19.84,
20.07, 20.03, 19.99, 19.73, 19.96,
19.83, 19.71, 19.85, 20.10, 20.15

♠

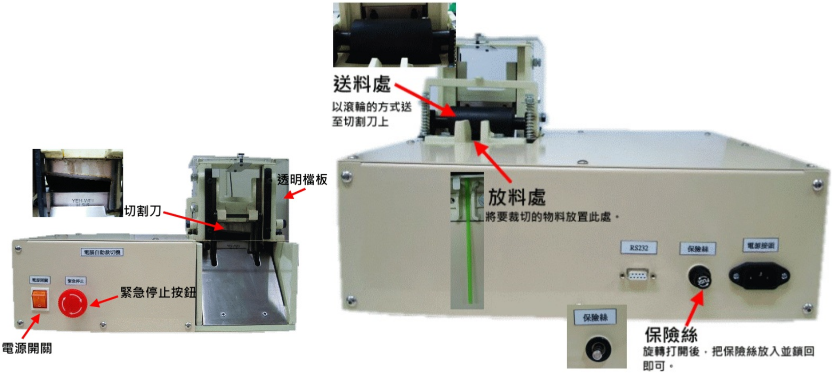

圖: 切割機的正面與反面圖

1 設立兩個假設。

$$\begin{cases} H_0: & \mu = 20 \\ H_1: & \mu \neq 20 \end{cases}$$

2 選擇檢定統計量。

$$\begin{aligned} \bar{X} &= \frac{20.16 + 20.01 + \cdots + 20.10 + 20.15}{20} \\ &= \frac{398.27}{20} = 19.9135 \end{aligned}$$

$$\begin{aligned} S^2 &= \frac{20.16^2 + \cdots + 20.15^2 - 20 \times 19.9135^2}{19} \\ &= 0.02899 \end{aligned}$$

$$s = \sqrt{0.02899} = 0.17$$

所以, 檢定統計量爲

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{19.9135 - 20}{0.17/\sqrt{20}} = -2.27$$

3 決定拒絕域。

$$\alpha = 0.05, t_{0.025}(19) = 2.09$$

$$C = \{T < -2.09 \text{ 或 } T > 2.09\}$$

4 結論。
因爲 $T = -2.27$ 落在 C 內, 故拒絕 H_0, 表示裁切長度已經偏離目標值。

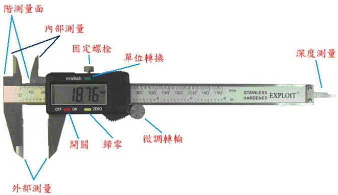

圖: 數位游標卡尺

Example

接續例20
假設要求切割機裁切吸管的標準差不得超過 0.15 公厘，請問這組資料是否足以證實標準差超過0.15公厘。

♠

1. 設立兩個假設

$$\begin{cases} H_0: & \sigma \leq 0.15 \\ H_1: & \sigma > 0.15 \end{cases}$$

2. 檢定統計量

$$S^2 = 0.0289$$

$$\chi_0^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{19 \times 0.0289}{0.15^2} = 24.4$$

3. 決定拒絕域

$$\alpha = 0.05, \chi^2_{0.05}(19) = 30.142$$

$$C = \{\chi^2_0 > 30.14\}$$

4. 結論: 因爲 $\chi^2_0 = 24.4$ 不落在 C 內, 故無法拒絕 H_0, 證據不足以證實標準差超過0.15公厘。