Chapter 9 估計

李水彬

Chien Hsin University of Science and Technology

2014.02

尼曼(Neyman) 於1937年首先發表信賴區間的觀念。

- 統計學的目的為探討全體研究對象在某一項量化特徵 (characteristic) 上的分配情形。
- 分配就是描述此特徵的各種可能結果佔全體的相對次數 (relative frequency)。

Example

某班有 N 同學, 同學成績分配 (成績爲一種量化特徵) 就是描述每一個分數 (各種可能結果) 上的人次占全體同學的比例 (相對次數)。爲了更爲明確說明這意義, 令 x 代表同學的成績, $n(x)$ 代表得到這個分數的學生人數。我們定義 **相對次數**爲 (參見第2章敍述統計)

$$f(x) = \frac{n(x)}{N} \tag{1}$$

- 在這個班上抽樣，就是獲取統計資料的過程，記錄這些抽樣結果稱爲**隨機變數**。
- 在古典機率觀點下，我們隨機從班上抽出一個人，每位同學被抽到的機會被假設是相同的。
- 這位同學的成績爲 x 的機率就是 $f(x)$，我們稱 $f(x)$ 爲這班學生成績的分佈，換言之隨機變數的分配與母體分配皆是 $f(x)$。

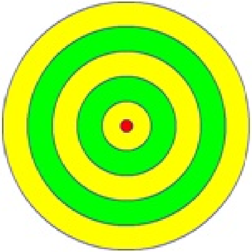
當我們知道這個分配 $f(x)$，就可以回答有關這班同學成績的任何問題。

Example

- 投擲一個銅板的實驗, 出現正面與反面的機率是否相等呢?
- 這個實驗的重要參數爲出現正面的機率 (p), 它是否公平 $(p = 0.5)$, 卻只有上帝可知。
- 最困難的地方是定義『公平』一詞。

- 歷史上記載有人曾經紀錄 4000 次丟擲的結果
- 18世紀的法國人**布豐伯爵 (Count Buffon)** 投擲 4040 次銅板, 共有 2048 次正面,
- 英國統計學家**卡爾皮爾生 (Karl Peason)**更有投擲 24000 次的紀錄, 其中有 12012 次正面, 企圖用多次實驗中銅板出現的相對次數做爲銅板出現正面的機率值。

關於母體參數的估計, 如射箭選手面對射箭場上的靶心一樣, 期望他射出的箭能夠命中靶心。靶心 (中心紅點) 爲未知的母體參數, 估計是箭在靶上的落點, 一位射箭選手就是一種估計方法。



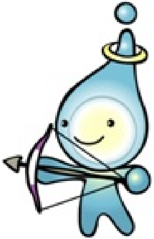


圖: 母體參數就像是靶上的靶心, 一個好的估計式就像一個訓練有素的射箭選手。

- 我們通常不會用一個落點來評估射箭選手的優劣，我們期待射箭選手在每一次射擊中有很好的表現，落點與靶心的距離越短，就是表現越好。
- 當一位射箭選手射出 100 箭的落點，以這些落點與靶心的距離評估選手的能力。
- 同樣地，統計上好的估計就是使估計值都能夠接近未知母體參數。

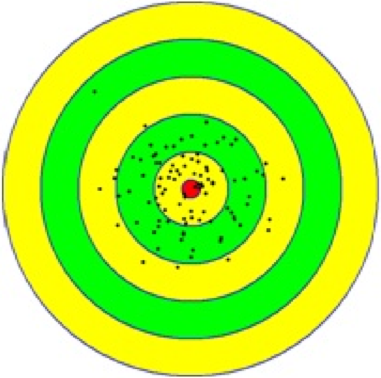


圖: 箭的落點視爲估計式的一個估計值。

Example

假設一個銅板是公平的, 它的意思就是投擲一個銅板它出現正面 (H) 的機會與反面 (T) 的機會都是相同的, 即出現的機率各 $1/2$。令 X 代表投擲一個銅板的結果, 它的分配爲

$$\begin{aligned} P(X = H) &= 0.5 \\ P(X = T) &= 0.5 \end{aligned} \tag{2}$$

♠

讓1爲正面, 0 爲反面。但事實上, 銅板出現正面與反面的機會並不一定相同, 我們可以假設爲 (p 代表成功的機率) 以下是實驗20次的結果, 請問你覺得這銅板公平嗎? 你認爲 p 應該爲何?

$$1\ 0\ 0\ 1\ 1\ 1\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ 0\ 0\ 1$$

Example

某便利商店雇用2位工讀生, 每小時可以服務 40 位顧客。假設便利商店每小時的顧客人數是一個卜瓦松分配, 過去資料顯示每小時平均顧客人數爲 30 人次, 假設 X 爲一小時的顧客人數, 我們可以假設

$$X \sim P(30)$$

♠

上述分配代表每小時顧客人數的分配情形, 據此可以計算超出服務能量的機率。所謂超出服務能量就是顧客人數超過40人, 其機率爲

$$P(X > 40) = \sum_{x=41}^{\infty} \frac{e^{-30} 30^x}{x!}$$

在常態近似如下

$$
\begin{aligned}
P(X > 40) &= P(X \geq 40.5) \\
&= P\left(\frac{X-30}{\sqrt{30}} \geq \frac{40.5-30}{\sqrt{30}}\right) \\
&\approx 1 - \Phi\left(\frac{40.5-30}{\sqrt{30}}\right) \\
&= 1 - \Phi(1.92) = 0.028
\end{aligned}
$$

換言之, 約每100小時約有2.8個時段發生顧客人數超出服務能量的機會。但若每小時的平均顧客人數升爲35人, 則超出服務能量的機率爲

$$\begin{aligned} P(X > 40) &= P(X \geq 40.5) \\ &= P\left(\frac{X-35}{\sqrt{35}} \geq \frac{40.5-35}{\sqrt{35}}\right) \\ &\approx 1 - \Phi(\frac{40.5-35}{\sqrt{35}}) \\ &= 1 - \Phi(1.004) \\ &= 1 - 0,8423 = 0.1576 \end{aligned}$$

約每10小時就有1.6個時段發生顧客人數超出服務能量的情形。這裡說明, 正確推估平均到客人數的重要, 如果平均到客人數超過35人, 就容易產生超出服務能量的問題, 降低服務品質。

Example

- 假設有一個公平的骰子, 換言之, 擲一次骰子出現1 6的機會皆相同, 且機率值等於1/6。如果擲兩次骰子, 則骰子出現的組合數共有36種, 每一種組合出現的機率也相同, 其機率值等於1/36。
- 令隨機變數 X 代表兩次的和, 其可能值和組合情形如下表。但事實上, 我們無從得知一個銅板是否是公平的, 甚至定義公平都非常困難。而我們有如何從實際投擲的結果推論那未知的機率呢?

♠

Table: x 值和其對應之機率值就構成隨機變數 X 的機率分配

x	組合情形	組合數	機率
2	(1,1)	1	1/36
3	(1,2) (1,3)	2	2/36
4	(1,3) (2,2) (3,1)	3	3/36
5	(1,4) (2,3) (3,2) (4,1)	4	4/36
6	(1,5) (2,4) (3,3) (4,2) (5,1)	5	5/36
7	(1,6) (2,5) (3,4) (4,3) (5,2) (6,1)	6	6/36
8	(2,6) (3,5) (4,4) (5,3) (6,2)	5	5/36
9	(3,6) (4,5) (5,4) (6,3)	4	4/36
10	(4,6) (5,5) (6,4)	3	3/36
11	(5,6) (6,5)	2	2/36
12	(6,6)	1	1/36

Example

假設本校學生身高 (X) 具有個常態分配, 平均值 $\mu = 170$, 變異數 $\sigma^2 = 25$, 即

$$X \sim N(170, 5^2)$$

今我們隨機抽出10位同學, 令 $\bar{X}$ 爲同學的平均身高, 根據chapter 9抽樣分配的理論, 我們知

$$\bar{X} \sim N(170, \frac{5^2}{10})$$

『樣本平均數的期望值與母體平均數相等, 而樣本平均數的變異數為母體變異數的 $1/n$ 』。

實際上, 當我們在一個常態母體中抽樣, 母體平均數與變異數是未知的, 我們如何用隨機抽樣得到的樣本推論母體平均數呢? 。

Example

假設一組來自常態母體的隨機樣本, 即

$$X_i \sim N(\mu, \sigma^2), i = 1, \cdots, n$$

習慣上以樣本平均值 $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ 做爲 μ 的估計式, 表示成

$$\hat{\mu} = \bar{X}$$

♠

若 $n = 3$,

$$X_1 = 2, X_2 = 3, X_3 = 4$$

則 $\bar{X} = \dfrac{2+3+4}{3} = 3$ 稱爲 μ 之估計值。

- 估計 (estimation) 是利用所知的樣本資訊推論母體參數的方法。
- 對於母體參數的估計, 若是以單一數值作為估計, 我們稱此方法為點估計。
- 若是考慮到估計的準確性, 用區間來描述參數的可能值範圍, 稱此為區間估計。

- 估計方法就如射箭選手, 有很多射箭選手各有特色與差異, 他們彼此競爭,
- 需要一些規則和計分方式協助區別選手之間的能力差異, 通常用落點與靶心的距離作爲評估方式。
- 相對的, 參數的估計也存有無數的方法, 這些估計方法就有差別, 估計的結果也不同。

Example

假設我們從一個常態母體中隨機抽出 5 個觀察值如下:

$$X_1 = 105, X_2 = 101, X_3 = 100, X_4 = 102, X_5 = 104$$

我們知道常態分配的平均值與中位數是相同的, 所以樣本平均數 $\bar{X} = 102.4$ 和中位數 102 都可以作爲母體平均數的估計。

- 他們之間的差異爲何? 哪一個估計方法比較好呢? 以下我們將介紹幾種評估點估計方法好壞的一些準則。
- 爲了討論方便, 令 θ 爲泛指一般未知的母體參數, 他的估計式通常用 $\hat{\theta}$ 表示之。例如, $\hat{\mu}$ 代表母體參數 μ 的估計, 而 $\hat{\sigma}$ 代表母體參數 σ 的估計。

定義

若 θ 的估計式 $\hat{\theta}$ 的期望值等於 θ, 即

$$E(\hat{\theta}) = \theta \tag{3}$$

則稱估計式 $\hat{\theta}$ 爲 θ 不偏估計式。反之, 若上式等號不成立, 則稱 $\hat{\theta}$ 爲有偏估計式。偏誤量爲

$$bias(\hat{\theta}) = E(\hat{\theta}) - \theta \tag{4}$$

Example

假設 $X_i \sim N(\mu, \sigma^2), i = 1, \cdots, 3$, 請說明以下四種母體平均數估計式的不偏性

$$
\begin{aligned}
\hat{\mu}_1 &= \frac{X_1 + X_2 + X_3}{3} \\
\hat{\mu}_2 &= \frac{X_1 + 2X_2 + X_3}{4} \\
\hat{\mu}_3 &= \frac{1.3X_1 + 1.2X_2 + 0.5X_3}{3} \\
\hat{\mu}_4 &= \frac{0.3X_1 + 1.2X_2 + 0.5X_3}{3}
\end{aligned}
$$

♠

因爲

$$
\begin{aligned}
E(\hat{\mu}_1) &= E\left(\frac{X_1+X_2+X_3}{3}\right) = \frac{\mu+\mu+\mu}{3} = \mu \\
E(\hat{\mu}_2) &= E\left(\frac{X_1+2X_2+X_3}{4}\right) = \frac{\mu+2\mu+\mu}{4} = \mu \\
E(\hat{\mu}_3) &= E\left(\frac{1.3X_1+1.2X_2+0.5X_3}{3}\right) \\
&= \frac{1.3\mu+1.2\mu+0.5\mu}{3} = \mu \\
E(\hat{\mu}_4) &= E\left(\frac{0.3X_1+1.2X_2+0.5X_3}{3}\right) \\
&= \frac{0.3\mu+1.2\mu+0.5\mu}{3} = \tfrac{2}{3}\mu
\end{aligned}
$$

所以, $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3$ 爲 μ 的不偏估計式, 但 $\hat{\mu}_4$ 爲有偏估計, 偏差量爲 $\frac{1}{3}\mu$。
在統計學這門課中, 我們僅討論幾種常用參數(成功的比率, 母體平均數和母體變異數) 的估計, 估計式和它的偏性如表 2 。

Table: 常用的不偏估計式

估計式	表示式	不偏性
$\bar{X}$ 估計 μ	$\hat{\mu} = \bar{X}$	不偏
S^2 估計 σ^2	$\hat{\sigma^2} = S^2$	不偏
S 估計 σ	$\hat{\sigma} = S$	有偏
$\bar{p} = \frac{X}{n}$ 估計 p	$\hat{p} = \frac{X}{n}$	不偏

- 隨機變數 $\hat{\theta}$ 的期望值就是估計值的平均值，就表示有很多可能結果的平均之意，而非單一數值。
- 這些數量夠多的落點可以代表一個估計式的所有可能結果，所以，期望值就如是這些點的平均值。
- 我們要從這些落點來評估射箭者的射箭能力，而評估方式不是以某一次的落點爲依據，而是用很多次的落點所呈現的整體趨勢。

一個落點並無法解釋這位選手的能力, 通常需要有很多射擊的結果才能反映出這個事實。下圖是100箭的落點, 這些點的中心可以用來代表這位射箭選手的瞄準點。一個簡單的判定方式就是看射擊選手的瞄準點是否落在靶心上。

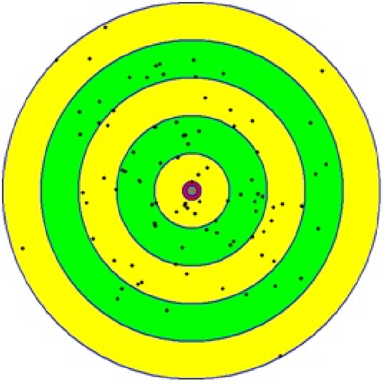


圖: 一個瞄準點與靶心相同的落點分配。

- 相對於這個射擊例子, 估計式是一種估計的方法, 結果稱爲估計值, 也就是靶上某一個落點。就如我們用 100 個落點評估射箭選手的能力, 確保代表國家參加比賽能夠有優異的表現。
- 請注意到**這關乎一個估計式的好壞, 而非某一個估計值是否正確。**
- 一些這估計式的估計值 (落點) 的平均值與被估計參數的差異, 就可以用來衡量這估計式的優劣。

- 雖然可以用射箭來類比參數估計，然其間還是存有一個重要的差異，射箭者是知道靶心的位置，但**估計卻不知道真正的參數為何?**
- 只有好的估計式才能讓我們確信一個 (唯一的) 估計值是可靠的。

估計式的有效性是以估計式的均方差(mean square error) 來衡量, 若其平均平方誤差越小, 則代表估計式的有效性越高 (誤差越小)。

定義

估計式 $\hat{\theta}$ 的均方差定義爲

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2 \tag{5}$$

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + bias^2(\hat{\theta}) \tag{6}$$

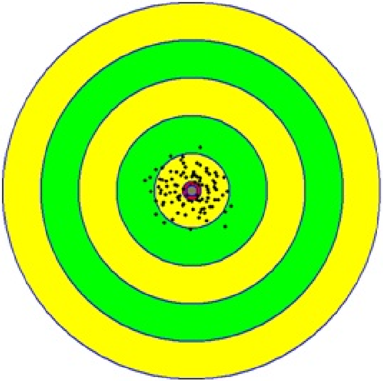


圖: 另一個瞄準點與靶心相同的落點分配。

- 均方差計算每一個落點與靶心的距離平分方和, 這距離就是一種估計的誤差。
- 在數學計算上, 誤差 $\hat{\theta}-\theta$ 具有方向性, 其值可能爲正或爲負值。將誤差平方避免負值。
- 分解公式:

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + bias^2(\hat{\theta}) \tag{7}$$

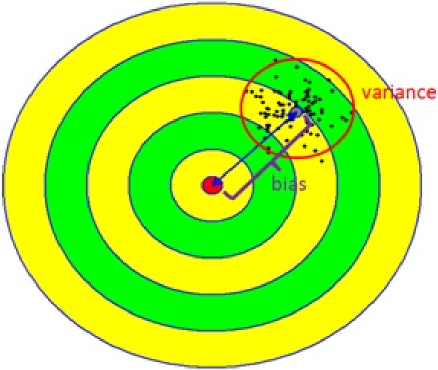


圖: 變異數與偏誤量的幾何意義。

- 估計式的均方差分解成估計式的變異 $Var(\hat{\theta})$ 與偏差量平方 $bias^2(\hat{\theta})$ 的總和。
- 估計式的變異代表自我內部的分散程度，也就是這些落點的散佈範圍，與靶心無關。
- 偏差量就是瞄準點與靶心的距離，代表射箭者的系統性變異。
- **均分差為自我內部差異與系統性變異的總和。**

定義

若 $\hat{\theta}_1$ 和 $\hat{\theta}_2$ 爲 θ 的兩個估計式, 若

$$r(\hat{\theta}_1, \hat{\theta}_2) = \frac{MSE(\hat{\theta}_1)}{MSE(\hat{\theta}_2)} < 1 \tag{8}$$

我們稱 $\hat{\theta}_1$ 相對於 $\hat{\theta}_2$, 在估計 θ 具有相對有效性 *(relative efficiency)*。相反地, 若是 $r(\hat{\theta}_1, \hat{\theta}_2) > 1$, 則 $\hat{\theta}_2$ 具有相對有效性。

假設隨機變數 $X_i, i = 1, \cdots, n$ 相互獨立且有共同的變異數 σ^2, 則

$$Var\left(\sum_{i=1}^{n} \omega_i X_i\right) = \left(\sum_{i=1}^{n} \omega_i^2\right) \sigma^2 \qquad (9)$$

其中 $\omega_i, i = 1, \cdots, n$ 爲任意常數係數。

Example

令 $W = 0.2X_1 - 3X_2 + 1.2X_3$, 即
$\omega_1 = 0.2, \omega_2 = -3, \omega_3 = 1.2$, 所以

$$\begin{aligned} Var(W) &= Var(0.2X_1 - 3X_2 + 1.2X_3) \\ &= (0.2^2 + (-3)^2 + 1.2^2)\sigma^2 \\ &= (0.04 + 9 + 1.14)\sigma^2 = 10.48\sigma^2 \end{aligned}$$

♠

Example

接續例 9, 討論 $\hat{\mu}_1, \hat{\mu}_2, \hat{\mu}_3, \hat{\mu}_4$ 的有效性。

$$
\begin{aligned}
MSE(\hat{\mu}_1) &= Var\left(\hat{\mu}_1\right) + bias^2\left(\hat{\mu}_1\right) \\
&= Var\left(\frac{X_1 + X_2 + X_3}{3}\right) + 0^2 \\
&= \frac{Var(X_1) + Var(X_2) + Var(X_3)}{9} \\
&= \frac{\sigma^2 + \sigma^2 + \sigma^2}{9} = \frac{\sigma^2}{3} \\
MSE(\hat{\mu}_2) &= Var\left(\hat{\mu}_2\right) + bias^2\left(\hat{\mu}_2\right) \\
&= Var\left(\frac{X_1 + 2X_2 + X_3}{4}\right) + 0^2 \\
&= \frac{Var(X_1) + 4Var(X_2) + Var(X_3)}{16} \\
&= \frac{\sigma^2 + 4\sigma^2 + \sigma^2}{16} = \frac{3\sigma^2}{8}
\end{aligned}
$$

$$
\begin{aligned}
MSE(\hat{\mu}_3) &= Var\left(\hat{\mu}_3\right) + bias^2\left(\hat{\mu}_3\right) \\
&= Var\left(\frac{1.3X_1 + 1.2X_2 + 0.5X_3}{3}\right) + 0^2 \\
&= \frac{1.69Var(X_1) + 1.44Var(X_2) + 0.25Var(X_3)}{9} \\
&= \frac{1.69\sigma^2 + 1.44\sigma^2 + 0.25\sigma^2}{9} = \frac{3.38\sigma^2}{9} \\
MSE(\hat{\mu}_4) &= Var\left(\hat{\mu}_4\right) + bias^2\left(\hat{\mu}_4\right) \\
&= Var\left(\frac{0.3X_1 + 1.2X_2 + 0.5X_3}{3}\right) + \left(\frac{\mu}{3}\right)^2 \\
&= \frac{0.09Var(X_1) + 1.44Var(X_2) + 0.25Var(X_3)}{9} \\
&+ \frac{\mu^2}{9} \\
&= \frac{0.09\sigma^2 + 1.44\sigma^2 + 0.25\sigma^2}{9} = \frac{1.78\sigma^2}{9} + \frac{\mu^2}{9}
\end{aligned}
$$

因爲

$$\begin{aligned} MSE(\hat{\mu}_1) &= \frac{\sigma^2}{3} < MSE(\hat{\mu}_2) = \\ \frac{3\sigma^2}{8} &< MSE(\hat{\mu}_3) = \frac{3.38\sigma^2}{9} \end{aligned}$$

所以, $\hat{\mu}_1$ 爲這3個不偏估計式中最爲有效。但 $\hat{\mu}_1$ 和 $\hat{\mu}_4$ 比較, 則有些情況下 $\hat{\mu}_1$ 相對有效, 相反的條件下, 則是 $\hat{\mu}_4$ 相對有效。

一致性 (consistency) 爲描述一個估計式的大樣本性質。

定義

令 $\hat{\theta}_n$ 爲樣本大小爲 n 的估計式, 若對任意 $\epsilon > 0$

$$\lim_{n\to\infty} P\left(\mid \hat{\theta}_n - \theta \mid < \epsilon\right) = 1 \qquad (10)$$

我們稱 $\hat{\theta}$ 爲 θ 的一致估計式。

表示樣本大小增加, 估計值與參數的差 $\mid \hat{\theta}_n - \theta \mid$ 會變得越來越小, 趨近於零。

- 參數 θ 的區間估計 (interval estimation) 是用它的估計量 $\hat{\theta}$ 決定一個區間 $\left[\hat{\theta}_L, \hat{\theta}_U\right]$,
- 認爲眞正的母體參數應該落在這個區間之內, 其中 $\hat{\theta}_L$ 爲區間下界與 $\hat{\theta}_U$ 爲區間上界, 他們都點估計量 $\hat{\theta}$ 的函數。

- 眞正的母體參數有可能落在區間之外, 所以, 區間估計除了提供此一個區間說明參數的範圍, 仍須說明參數落在這個區間的機率。
- 此一機率與採用的統計量有關, 我們必須利用它的抽樣分配計算區間估計 $\left[\hat{\theta}_L, \hat{\theta}_U\right]$ 包含參數 θ 的機率值。

- 實務上, 我們會先給定參數落在區間之外的機率值 α, 再求適當的區間估計式 $\left[\hat{\theta}_L, \hat{\theta}_U\right]$。
- 實際上, 滿足落在區間之外的機率值爲 α 的區間可能有無限多種可能, 較好的區間爲何? 以及如何找到它? 超出我們的課程範圍在此略過不細究。
- 當我決定好一個區間估計後, 將估計量代 $\hat{\theta}$ 入 $\left[\hat{\theta}_L, \hat{\theta}_U\right]$ 函數中, 得出參數的區間估計。

區間估計不同於『點估計』, 它以區間長度與包含參數的機率值 $1-\alpha$ 表現區間估計的精確度 (precision)。

推導參數的區間估計之步驟如下:

1. 找出母體參數 θ 的點估計 $\hat{\theta}$。
2. 導出 $\hat{\theta}$ 的抽樣分配。(參見第 **??** 章)
3. 給定機率值 α, 利用2的抽樣分配決定區間 $\left[\hat{\theta}_L, \hat{\theta}_U\right]$, 使得。

$$P\left(\hat{\theta}_L \leq \theta \leq \hat{\theta}_U\right) = 1 - \alpha \tag{11}$$

4. 進行隨機抽樣, 得樣本 $X_1, \cdots, X_n$, 計算估計量 $\hat{\theta}$ 代入估計式 $\hat{\theta}_L$ 和 $\hat{\theta}_U$ 計算區間上下界之值。

Table: 一些常見的母體參數與它的抽樣分配

參數 θ	估計式 $\hat{\theta}$	抽樣分配
μ	$\hat{\mu} = \bar{X}$	$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
σ^2	$\hat{\sigma^2} = S^2$	$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$
p	$\hat{p} = \frac{X}{n}$	$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right)$

我們將所得樣本 $x_1, \cdots, x_n$ 代入, 得

$$\hat{\theta}_L = \hat{\theta}_L(x_1, \cdots, x_n), \hat{\theta}_U = \hat{\theta}_U(x_1, \cdots, x_n) \quad (12)$$

我們稱 $(\hat{\theta}_L, \hat{\theta}_U)$ 爲參數 θ 的 $(1-\alpha)\times 100\%$ 信賴區間 (confidence interval, CI)。其中 $(1-\alpha)\times 100\%$ 爲信心水準 (confidence level) 或信賴係數 (confidence coefficient)。$\hat{\theta}_L$ 和 $\hat{\theta}_U$ 分別稱爲信賴下界與信賴上界。

一組來自目標母體的樣本 $X_1, \cdots, X_n$, 我們想要估計這個母體的平均數 μ, 點估計的方法就是用樣本平均數估計

$$\hat{\mu} = \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{13}$$

假設這組樣本是抽自一個常態母體, 即 $X_1, \cdots, X_n \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$。樣本平均數的抽樣分配爲

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \tag{14}$$

當參數 μ, σ^2 已知時, 上述分配可以協助我們瞭解樣本平均數的分配。

- 我們只有這組樣本 $X_1, \cdots, X_n$, 如何推論平均數 μ 呢?
- 應該是落在 $\bar{X}$ 附近, 所以我們可以令平均數 μ 的區間估計的上下界分別爲

$$\hat{\mu}_L = \bar{X} - a, \hat{\mu}_U = \bar{X} + a \tag{15}$$

其中 a 爲一個正數。

平均數 μ 落在這個區間的機率為

$$
\begin{aligned}
P(\bar{X}-a\leq\ & \mu\leq\bar{X}+a)=P(\mu-a\leq\bar{X}\leq\mu+a)\\
&=P(-a\leq\bar{X}-\mu\leq a)\\
&=P(\frac{-a}{\sigma/\sqrt{n}}\leq Z\leq\frac{a}{\sigma/\sqrt{n}})\\
&=\Phi\left(\frac{a}{\sigma/\sqrt{n}}\right)-\Phi\left(-\frac{a}{\sigma/\sqrt{n}}\right)\\
&=2\Phi\left(\frac{a}{\sigma/\sqrt{n}}\right)-1
\end{aligned}
\tag{16}
$$

當 $\bar{X} = \bar{x}$ 時, 得平均數 μ 的信賴區間爲

$$[\bar{x} - a, \bar{x} + a]$$

信心水準爲

$$2\Phi\left(\frac{a}{\sigma/\sqrt{n}}\right) - 1$$

令 $w = \hat{\theta}_U - \hat{\theta}_L$ 代表區間長度, 圖 6 爲區間長度與信心水準的函數關係。

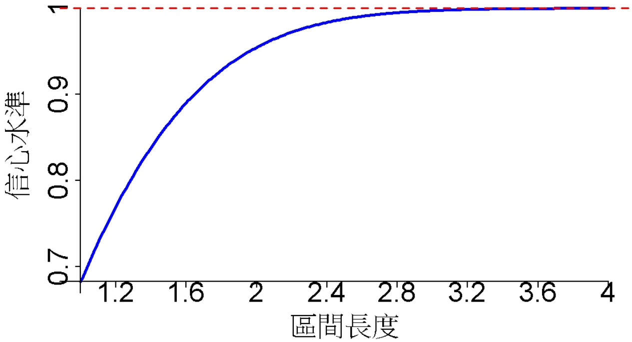


圖: 區間長度與信心水準的關係, 橫軸爲區間長度, 縱軸爲信賴水準。

1. 當 α 愈小, 信心水準愈高, 區間長度愈長。
2. 假設兩個區間估計 $(\hat{\theta}_{1,l}, \hat{\theta}_{1,u})$ 和 $(\hat{\theta}_{2,l}, \hat{\theta}_{2,u})$ 包含母體參數 θ 的機率相同。若

$$\hat{\theta}_{1,u} - \hat{\theta}_{1,l} < \hat{\theta}_{2,u} - \hat{\theta}_{2,l}$$

即第一個信賴區間長度較短爲較優的選擇。

Neyman 於1937 提出信賴區間的觀念，其實也不是一下就為大家所能接受，甚至有人認為這是一種欺騙。

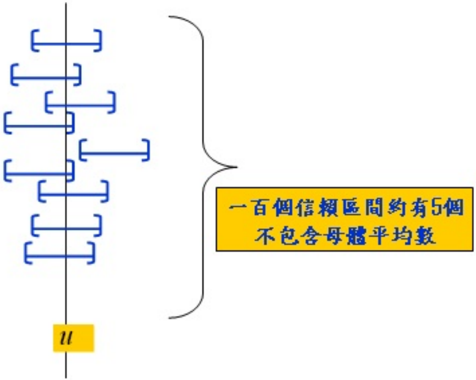


圖: 95%信心水準的意義。

以下我們以建立95%($\alpha = 0.05$) 信心水準的信賴區間為例。假設 $X_1, \cdots, X_n \overset{iid}{\sim} N(\mu, \sigma^2)$, 樣本平均數的抽樣分配爲 $\bar{X} = N\left(\mu, \dfrac{\sigma^2}{n}\right)$。

因爲

$$P(\bar{X} - a \leq \mu \leq \bar{X} + a) = 0.95$$

也就是

$$2\Phi\left(\frac{a}{\sigma/\sqrt{n}}\right) - 1 = 0.95$$

標準化得

$$\frac{a}{\sigma/\sqrt{n}} = 1.96$$

所以,

$$a = \frac{1.96\sigma}{\sqrt{n}}$$

因此, 從式15 可得平均數區間估計的上下界為

$$\begin{aligned} \hat{\mu}_L(\bar{X}) &= \bar{X} - \frac{1.96\sigma}{\sqrt{n}} \\ \hat{\mu}_U(\bar{X}) &= \bar{X} + \frac{1.96\sigma}{\sqrt{n}} \end{aligned} \tag{17}$$

我們稱

$$\left(\bar{X}-\frac{1.96\sigma}{\sqrt{n}},\bar{X}+\frac{1.96\sigma}{\sqrt{n}}\right)$$

爲母體平均數的區間估計, 其包含 μ 的機率爲0.95。當我們得到一組樣本, 計算樣本平均數得 $\bar{X}=\bar{x}$ 時, (註 $\bar{X}$ 爲隨機變數, $\bar{x}$ 爲數値。) 代入此一區間可得

$$\left(\bar{x}-\frac{1.96\sigma}{\sqrt{n}},\bar{x}+\frac{1.96\sigma}{\sqrt{n}}\right)$$

稱此爲母體平均數 μ 的95%信賴區間。考慮重複100次的抽樣, 可得100組信賴區間, 其中約有95組包含母體平均數, 故稱每次的信賴區間的信心水準爲95%。

建立母體平均數信賴區間的步驟:

分成如下四個步驟:

1. 決定點估計的方法: 母體平均數 μ 的點估計爲 $\hat{\mu} = \bar{X}$.
2. 推導估計式的抽樣分配: $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$, 所以

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

3 區間估計的上下界: 依據 α 值, 利用標準常態機率分配表查 $z_{\alpha/2}$。

$$\begin{aligned}\hat{\mu}_L(\bar{X}) &= \bar{X} - \frac{z_{\alpha/2}\sigma}{\sqrt{n}} \\ \hat{\mu}_U(\bar{X}) &= \bar{X} + \frac{z_{\alpha/2}\sigma}{\sqrt{n}}\end{aligned} \tag{18}$$

4 寫下信賴區間: 代入 $\bar{x}$, 我們得母體平均數 μ 的 $(1-\alpha)\times 100\%$ 信賴區間爲

$$\left(\bar{x}-\frac{z_{\alpha/2}\sigma}{\sqrt{n}},\bar{x}+\frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right)$$

母體平均數 μ 的信賴區間的解題步驟 (常態母體, 變異數已知):

1. 計算點估計統計量值。這個步驟要將基本統計量準備好, 通常就是樣本平均數與樣本標準差。

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, s = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}}$$

2. 根據信心水準 $(1-\alpha) \times 100\%$ 查 $z_{\alpha/2}$。

3. 計算抽樣誤差 $E = z_{\alpha/2}\dfrac{\sigma}{\sqrt{n}}$。

   抽樣誤差=查表 $\times$ 估計式的標準差。

4. 結論: 包含須寫清楚兩個部份 (a) 信賴區間與 (b) 信心水準。

Table: 常用的 $z_{\alpha/2}$ 值表

α	0.2	0.1	0.05	0.025	0.02	0.01
z_α		1.282	1.645	1.96	2.054	2.326
$\alpha/2$	0.1	0.05	0.025	0.0125	0.01	0.005
$z_{\alpha/2}$	1.282	1.645	1.96	2.241	2.326	2.576

Example

假設產品重量分配爲 $X \sim N(\mu, 5^2)$, 今抽出 15 個, 得樣本平均值爲 $\bar{X} = 60$。

1. 請建立 μ 的 95% 信賴區間。
2. 請建立 μ 的 99% 信賴區間。

♠

95%信賴區間

1. 計算基本統計量: $\bar{X} = 60$。
2. 查表: $\alpha = 0.05, z_{0.025} = 1.96$。
3. 計算抽樣誤差: $E = 1.96 \times \frac{5}{\sqrt{15}} = 2.53$。
4. 結論: μ 的95%信賴區間爲 60 ± 2.45 或 $(57.47, 62.53)$。

99%信賴區間

1. 計算基本統計量: $\bar{X} = 60$。
2. 查表: $\alpha = 0.01, z_{0.005} = 2.576$。
3. 計算抽樣誤差: $E = 2.576 \times \frac{5}{\sqrt{15}} = 3.33$。
4. 結論: μ 的99%信賴區間爲 60 ± 3.33。

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 $X \sim N(\mu, 50^2)$。今抽樣量測 5 片晶片的 Si-Nitride 沉積厚度爲

1623, 1670,1625, 1633,1649

請問平均沉積厚度的 95% 信賴區間?

♠

1. 計算基本統計量: $\bar{X} = \dfrac{1623 + 1670 + 1625 + 1633 + 1649}{5} = 1640$。
2. 查表: $\alpha = 0.05$, $z_{0.025} = 1.96$。
3. 計算抽樣誤差: $E = 1.96 \times \dfrac{50}{\sqrt{5}} = 43.827$。
4. 結論: 平均沉積厚度 μ 的95%信賴區間爲 1640 ± 43.827。

Example

假設健行圖書館 (雲書館) 每天使用的人數爲常態分配 $X \sim N(\mu, 10^2)$。觀察6天的人數爲:

312, 293, 274, 323, 301, 300

請問每天平均人數 μ 的90%信賴區間。

♠

1. 計算基本統計量: $\bar{X} =$ ____________________ 。
2. 查表: ____________________ 。
3. 計算抽樣誤差: ____________________ 。
4. 結論: 每天平均人數 μ 的 90% 信賴區間爲 ____________________ 。

Example

假設麥香紅茶飲料重量 (公克) 爲常態分配 $X \sim N(\mu, 5^2)$。在賣場上抽驗10瓶麥香紅茶飲料的重量如下:

352, 354, 354, 351, 352, 346, 346, 347, 348, 350

請問麥香紅茶飲料平均重量 μ 的99%信賴區間。

1. 計算基本統計量: $\bar{X} =$ ____________________。
2. 查表: ____________________。
3. 計算抽樣誤差: ____________________。
4. 結論: 麥香紅茶飲料平均重量 μ 的99%信賴區間爲: ____________________。

在抽樣前研究者應該先根據**精確度**決定抽取的樣本大小, 影響精確度的因素有

- 抽樣誤差的上界 E^* 和
- 最低信心水準。

研究者要求**實際的抽樣誤差**$\dfrac{z_{\alpha/2}\sigma}{\sqrt{n}}$ 應該不大於抽樣誤差上界E^*, 即

$$\frac{z_{\alpha/2}\sigma}{\sqrt{n}} \leq E^* \tag{19}$$

可得

$$n \geq \left(\frac{z_{\alpha/2}\sigma}{E^*}\right)^2 \tag{20}$$

影響樣本大小的因素有

1. **母體變異數**: σ^2 越大，需要樣本大小 n 越大。亦即對於變異數較大的母體，參數的估計較難正確，需要提高樣本大小已達相同的精確度。
2. **信心水準**:信心水準高，表示 α 值小，n 值大。對於估計結果要有比較高的信心，無疑就需要更多的證據，在統計上唯一的證據就是樣本。
3. **抽樣誤差上界** E^*:可以容忍的抽樣誤差大，表是估計較爲不精確，所以就可以用比較少的樣本。

Example

鮪魚重量 (公斤) 分配爲常態 $X \sim N(\mu, 25^2)$。爲推估鮪魚平均重量 μ 的 95% 信賴區間, 使抽樣誤差低於5公斤, 請問所需樣本大小爲何?

♠

因爲

$$
\begin{aligned}
n \geq & \left(\frac{z_{\alpha/2}\sigma}{E^*}\right)^2 \\
= & \left(\frac{1.96 \times 25}{5}\right)^2 = 96.04
\end{aligned}
$$

所以, 至少要抽 $n = 97$ 條鮪魚。

Example

假設某批塑膠元件的寬度 (mm) 爲常態分配 $X \sim N(\mu, 0.05^2)$。爲了估計這批貨的平均寬度 μ 的 99% 信賴區間, 抽樣誤差低於 0.01 mm 下, 請問所需樣本大小爲何?

♠

因爲

$$
\begin{aligned}
n \geq & \left(\frac{z_{\alpha/2}\sigma}{E^*}\right)^2 \\
= & \left(\frac{2.576 \times 0.05}{0.01}\right)^2 = 165.8944
\end{aligned}
$$

所以, 至少要測量 $n = 166$ 塑膠元件的寬度。

Example

接續例16, 若 $X \sim N(\mu, 40^2)$。請問所需樣本大小爲何?

♠

因爲

$$\begin{aligned} n \geq & \left(\frac{z_{\alpha/2}\sigma}{E^*}\right)^2 \\ = & \left(\frac{1.96 \times 40}{5}\right)^2 = 245.86 \end{aligned}$$

所以, 至少要抽 $n = 246$ 條鮪魚。

Example

接續例13, 設定抽樣誤差低於20, 請問所需樣本大小為何?

♠

因為

$$\begin{aligned} n \geq & \left(\frac{z_{\alpha/2}\sigma}{E^*}\right)^2 \\ = & \left(\frac{1.96 \times 50}{20}\right)^2 = 24.01 \end{aligned}$$

所以, 至少要抽 $n = 25$ 片晶片。

Example

接續例15, 設定抽樣誤差低於1.5, 請問所需樣本大小為何?

♠

因為

$$
\begin{aligned}
n &\geq \left(\frac{z_{\alpha/2}\sigma}{E^*}\right)^2 \\
&= \left(\frac{\qquad\times\qquad}{\qquad\qquad\qquad}\right)^2 = \underline{\qquad\qquad}
\end{aligned}
$$

所以, 至少要抽 $n = \underline{\qquad\qquad}$ 瓶麥香紅茶飲料。

Example

接續例17, 設定抽樣誤差低於 0.02 mm, 請問所需樣本大小爲何?

♠

因爲

$$
\begin{aligned}
n \geq & \left(\frac{z_{\alpha/2}\sigma}{E^*}\right)^2 \\
= & \left(\frac{\qquad\times\qquad}{\qquad\qquad}\right)^2 = \underline{\qquad\qquad}
\end{aligned}
$$

所以, 至少要測量 $n = \underline{\qquad\qquad}$ 塑膠元件的寬度。

當 σ 未知時, 用樣本標準差 S 取代 σ, 可得

$$T_{n-1} = \frac{\sqrt{n}(\bar{X} - \mu)}{S} \sim t(n-1) \tag{21}$$

令臨界值 $t_\alpha(n-1)$ 滿足

$$P(T_{n-1} > t_\alpha(n-1)) = 1 - \alpha \tag{22}$$

所以,

$$P\left(-t_{\alpha/2}(n-1) < \frac{\sqrt{n}(\bar{X}-\mu)}{S} \leq t_{\alpha/2}(n-1)\right)$$

$$P\left(\bar{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}} < \mu \leq \bar{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right) \tag{23}$$

當母體分配爲常態, 且變異數未知時, 我們稱

$$\left(\bar{X} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}} < \mu \leq \bar{X} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$$

爲母體平均數的區間估計, 包含 μ 的機率爲 $(1-\alpha) \times 100\%$。

假設樣本平均數為 $\bar{X} = \bar{x}$, 代入此一區間可得

$$\left(\bar{x} - t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}} < \mu \leq \bar{x} + t_{\alpha/2}(n-1)\frac{S}{\sqrt{n}}\right)$$

稱此為母體平均數 μ 的 $(1-\alpha)\times 100\%$ 信賴區間。

母體平均數 μ 的信賴區間的解題步驟 (常態母體, 變異數未知):

1. 計算點估計統計量值。這個步驟要將基本統計量準備好, 通常就是樣本平均數與樣本標準差。

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}, s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

2. 自由度 $n-1$, 根據信心水準 $(1-\alpha) \times 100\%$ 查 $t_{\alpha/2}(n-1)$。

3. 計算抽樣誤差 $E = t_{\alpha/2}\frac{s}{\sqrt{n}}$。
4. 結論: 包含須寫清楚兩個部份 (a) 信賴區間與 (b) 信心水準。

Example

接續例13
假設沈積厚度的變異數未知, 則其分配爲
$X \sim N(\mu, \sigma^2)$, 請問平均沈積厚度 的99%信賴區間?

♠

1. 計算基本統計量:

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

$$\begin{aligned} S^2 &= \frac{(-17)^2 + 30^2 + (-15)^2 + (-7)^2 + 9^2}{4} \\ &= \frac{1544}{4} = 386 \end{aligned}$$

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

2. 自由度 $n - 1 = 4$, 查表:
$\alpha = 0.01, t_{0.005}(4) = 4.604$。

3. 計算抽樣誤差: $E = 4.604 \times \frac{19.65}{\sqrt{5}} = 40.45$。
4. 結論: 平均沉積厚度 μ 的95%信賴區間爲 1640 ± 40.45。

Example

接續例 15

假設麥香紅茶飲料重量 (公克) 的變異數未知, 其分配為 $X \sim N(\mu, \sigma^2)$。請問麥香紅茶飲料平均重量 μ 的 95% 信賴區間。

♠

1. 計算基本統計量:
   樣本平均值 $\bar{X}$ = ______________________ 。
   樣本變異數 S^2 = ______________________ 。
   樣本標準差 S = ______________________ 。
2. 自由度 $n-1$ = _____, 查表:
   ______________________ 。
3. 計算抽樣誤差: ______________________ 。
4. 結論: 麥香紅茶飲料平均重量 μ 的95%信賴區間為
   ______________________ 。

Example

本校每年舉行教師體適能檢測, 抽測 15 位教師踩階休息 3 分鐘後的心跳

$$54,63,58,72,60,92,70,73,69,104,48,66,80,64,77$$

假設教師心跳分配爲常態, 請求心跳平均值的90%信賴區間。

♠

1. 計算基本統計量:

$$\bar{X} = \frac{54 + 63 + \cdots + 77}{15} = \frac{1050}{15} = 70$$

$$S^2 = \frac{(-16)^2 + (-7)^2 + \cdots + 7^2}{14} = 206.29$$

$$s = \sqrt{206.29} = 14.36$$

2. 自由度 $n - 1 = 4$, 查表:
$\alpha = 0.1, t_{0.05}(14) = 1.7613$。
3. 計算抽樣誤差: $E = 1.7613 \times \frac{14.36}{\sqrt{15}} = 6.53$。
4. 結論: 心跳平均值 μ 的90%信賴區間爲
70 ± 6.53。

假設母體分配爲 F_θ, 其中 θ 代表未知的母體參數。亦即

$$X_1, \cdots, X_n \overset{iid}{\sim} F_\theta \tag{24}$$

若 F_θ 分配的平均值爲 μ 和標準差爲 σ。根據中央極限定理, 當 $n \geq 30$ 時, 我們可得樣本平均數的抽樣分配

$$\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \sim N(0, 1) \tag{25}$$

因此, 我們得母體平均數 μ 的 $(1-\alpha)\times 100\%$ 信賴區間爲

$$\left(\bar{x}-\frac{z_{\alpha/2}\sigma}{\sqrt{n}}, \bar{x}+\frac{z_{\alpha/2}\sigma}{\sqrt{n}}\right)$$

當上式中的 σ 爲未知, 用樣本標準差 s 取代之, 改寫爲

$$\left(\bar{x}-\frac{z_{\alpha/2}s}{\sqrt{n}}, \bar{x}+\frac{z_{\alpha/2}s}{\sqrt{n}}\right)$$

Example

調查本系50位畢業四年內校友的薪資 (月薪), 資料如表5 (萬), 請問本系畢業四年內校友平均月薪的95%信賴區間。

Table: 常校友的薪資 (月薪, 萬)

2.9	3.8	3.3	4.3	4.1	5.4	4	5.3	4.5	2.3
3.1	4.4	3.9	4.2	4.6	3.6	4.2	3.8	4.8	5.7
4	3.4	2.7	6.3	2.9	4.6	3.4	2.5	5.2	4.6
4.2	2.3	3.1	5	3.3	3.4	3.2	2.8	3.9	4.5
3.6	2.4	2.6	5.4	5.3	5.5	3.6	5.4	4.3	5

1 計算基本統計量:

$$\bar{X} = \frac{2.9 + \cdots + 5}{50} = 4.012$$

$$\begin{aligned} S^2 &= \frac{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}{n-1} \\ &= \frac{(2.9)^2 + \cdots + 5^2 - 50 \times 4.012^2}{49} = 1.0 \end{aligned}$$

$$s = \sqrt{1} = 1$$

2 查表: $\alpha = 0.05$, $z_{0.025} = 1.96$。

3. 計算抽樣誤差: $E = 1.96 \times \dfrac{1}{\sqrt{50}} = 0.277$。
4. 結論: 母體平均值 μ 的95%信賴區間爲 4.012 ± 0.277。

推導參數信賴區間公式的流程:

(a) 決定此參數的點估計方法
(b) 推導估計式的抽樣分配
(c) 根據抽樣分配求出區間估計的上下界
(d) 寫下信賴區間

(a) **點估計式**
直觀的作法就是利用樣本變異數估計 σ^2,
它是 σ^2 不偏估計。

(b) **抽樣分配**: 假設
$X_1, \cdots, X_n \overset{iid}{\sim} N\left(\mu, \sigma^2\right)$, 則

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

(c) 區間估計式: 因爲

$$P\left(\chi^2_{1-\alpha/2}(n-1) \le \frac{(n-1)S^2}{\sigma^2} \le \chi^2_{\alpha/2}(n-1)\right) = 1-\alpha \quad (26)$$

可得

$$P\left(\frac{1}{\chi^2_{1-\alpha/2}(n-1)} \ge \frac{\sigma^2}{(n-1)S^2} \ge \frac{1}{\chi^2_{\alpha/2}(n-1)}\right) = 1-\alpha$$

$$P\left(\frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} \ge \sigma^2 \ge \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)}\right) = 1-\alpha$$

$$P\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)}\right) = 1-\alpha$$

所以, σ^2 的區間估計上下界為

$$\begin{aligned} \hat{\sigma^2}_L(S^2) &= \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \\ \hat{\sigma^2}_U(S^2) &= \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} \end{aligned} \tag{27}$$

其中 $\chi^2_{\alpha/2}(n-1)$ 為自由度 $n-1$ 卡方分配的臨界值, 即對於 $X \sim \chi^2(n-1)$, 滿足

$$P(X > \chi^2_{\alpha/2}(n-1)) = \alpha$$

(d) **信賴區間**

令 $S^2 = s^2$, 代入此一區間可得 σ^2 的 $(1-\alpha)\times 100\%$ 信賴區間爲

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right) \quad (28)$$

母體變異數 σ^2 的 $(1-\alpha)\times 100\%$ 信賴區間的解題步驟:

1. 計算樣本變異數 S^2
2. 查表 $\chi^2_{\alpha/2}(n-1), \chi^2_{1-\alpha/2}(n-1)$
3. 計算區間上下界
4. 結論

Example

若某公司想估計其產品之變異數, 今隨機抽檢該公司 20 件產品, 其平均重量 70 公克, 標準差5公克。假設此公司生產之產品重量具有常態分配, 請問此公司產品重量變異數之 90% 信賴區間爲何?

1. 計算樣本變異數 $S^2 = 25$
2. 查表 $\alpha = 0.1$,
   $\chi^2_{0.05}(19) = 30.114, \chi^2_{0.95}(19) = 10.117$

3 計算區間上下界

$$\hat{\sigma}^2{}_L = \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} = \frac{19\times 25}{30.114} = 15.76$$

$$\hat{\sigma}^2{}_U = \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} = \frac{19\times 25}{10.117} = 46.95$$

4 結論: 產品重量變異數之 90% 信賴區間爲

$$(15.76, 46.95)$$

Example

某工廠進螺帽一批, 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配, 請問這批貨變異數的 90% 信賴區間。

1. 計算樣本變異數 $S^2 = 0.00325 = 3.25 \times 10^{-3}$
2. 查表 $\alpha = 0.1$,
   $\chi^2_{0.05}(9) = 16.917, \chi^2_{0.95}(9) = 3.334$

3 計算區間上下界

$$\begin{aligned}
\hat{\sigma}^2{}_L &= \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} \\
&= \frac{9 \times 3.25 \times 10^{-3}}{16.917} = 1.72 \times 10^{-3} \\
\hat{\sigma}^2{}_U &= \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} \\
&= \frac{9 \times 3.25 \times 10^{-3}}{3.334} = 8.77 \times 10^{-3}
\end{aligned}$$

4 結論: 這批貨變異數的 90% 信賴區間爲

$$(1.72 \times 10^{-3}, 8.77 \times 10^{-3})$$

Example

接續例 13
請求出 Si-Nitride 沉積厚度變異數的 95% 信賴區間。
$(S^2 = 386)$

1. 計算樣本變異數 $S^2 =$ ____________________
2. 查表 $\alpha = 0.05$, $\chi^2_{0.025}(4) =$ ____________________,
   $\chi^2_{0.975}(4) =$ ____________________

3 計算區間上下界

$$\hat{\sigma^2}_L = \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} = \underline{\hspace{6cm}}$$

$$\hat{\sigma^2}_U = \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} = \underline{\hspace{6cm}}$$

4 結論: Si-Nitride 沉積厚度變異數的 95% 信賴區間爲

$$(\underline{\hspace{3.5cm}}, \underline{\hspace{3.5cm}})$$

Example

接續例 15
請求出麥香紅茶飲料重量變異數的 95% 信賴區間。
$(S^2 = 9.56)$

1. 計算樣本變異數 $S^2 =$ ____________________
2. 查表 $\alpha = 0.05$,
   $\chi^2_{0.025}(\underline{\quad\quad}) =$ ____________________,
   $\chi^2_{0.975}(\underline{\quad\quad}) =$ ____________________

3 計算區間上下界

$$\hat{\sigma}^2{}_L = \frac{(n-1)S^2}{\chi^2_{\alpha/2}(n-1)} = \rule{6cm}{0.4pt}$$

$$\hat{\sigma}^2{}_U = \frac{(n-1)S^2}{\chi^2_{1-\alpha/2}(n-1)} = \rule{6cm}{0.4pt}$$

4 結論: 麥香紅茶飲料重量變異數的 95% 信賴區間為

$$(\rule{3.5cm}{0.4pt}, \rule{3.5cm}{0.4pt})$$

標準差 σ 的 $(1-\alpha)\times 100\%$ 信賴區間

步驟1 先求變異數的信賴區間

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}, \frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}\right)$$

步驟2 將上式區間上下界分別開 根號, 得

$$\left(\sqrt{\frac{(n-1)s^2}{\chi^2_{\alpha/2}(n-1)}}, \sqrt{\frac{(n-1)s^2}{\chi^2_{1-\alpha/2}(n-1)}}\right)$$

Example

接續例27

螺帽標準差的 90% 信賴區間爲

$$(\sqrt{1.72\times10^{-3}}, \sqrt{8.77\times10^{-3}}) = (0.042, 0.094)$$

♠

母體比例參數是母體平均數的一種特例, 其資料是來自白努利隨機實驗, 每一個觀察值僅有 0 或 1 的可能。讓 X 代表 n 次實驗的總和, X 具有二項分配。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

$$np \geq 5且n(1-p) \geq 5 \tag{29}$$

$\hat{p}$ 近似常態分配方可成立。當 p 值過小或過大 $(1-p$ 過小) 屬稀有事件, 可以用卜瓦松分配近似求得機率值, 亦可以建立 p 的信賴區間。

本節以常態近似方法建構母體比例的信賴區間。

- 在民意調查調查中，一個議題或是一位候選人的支持度 p 通常不會趨於極端，介於 $0.3 \sim 0.7$ 之間。根據式 (29) 的要求，受訪者人數超過 17 人即可以用常態分配近似。
- 但是在工廠產品檢驗的問題上，檢驗出不良品屬於稀有事件，不良 率 p 往往低於 0.1，因此爲滿足式 (29) 的要求，抽樣個數至少要 50 個以上才能得到好的近似結果。

這裡以產品檢驗爲例, 將相關使用到的符號說明如下:

- n : 檢驗個數;
- X : 被檢驗產品中的不良品總數;
- p : 製程不良率;
- $\hat{p} = \dfrac{X}{n}$ 爲樣本比例, 母體比例 p 的不偏估計。

根據中央極限定理知, 樣本比例的近似分配爲

$$\hat{p} \sim N\left(p, \frac{p(1-p)}{n}\right) \tag{30}$$

類似母體平均數信賴區間的推導方式, 我們可以得到下面結果

$$P\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} \leq p \leq \hat{p} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) = 1-\alpha \tag{31}$$

母體比例 p 的區間估計爲

$$\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}\right) \quad (32)$$

此區間包含未知母體比例 p, 通常有兩種作法得到 p 的信賴區間:

1. 在上式中用 $\hat{p}$ 代替母體比例數 p, 可得母體比例數 p 的 $(1-\alpha)\times 100\%$ 信賴區間爲
$$\left(\hat{p}-z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}},\hat{p}+z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) \tag{33}$$
2. 因爲 $p(1-p)\leq\frac{1}{4}$, 當 $\hat{p}\approx 0.5$ 時, 用 $\frac{1}{4}$ 取代 $p(1-p)$, 得到一個較寬, 保守的區間爲
$$\left(\hat{p}-\frac{z_{\alpha/2}}{2\sqrt{n}},\hat{p}+\frac{z_{\alpha/2}}{2\sqrt{n}}\right) \tag{34}$$

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, 請問實際支持現任市長之比例值之95% 信賴區間爲何?

1. 計算基本統計量: $\hat{p} = \dfrac{X}{n} = \dfrac{75}{300} = 0.25$, $s = \sqrt{\hat{p}(1-\hat{p})} = \sqrt{0.25 \times 0.75} = 0.433$。
2. 查表: $\alpha = 0.05$, $z_{0.025} = 1.96$。
3. 計算抽樣誤差: $E = 1.96 \times \dfrac{0.433}{\sqrt{300}} = 1.96 \times 0.025 = 0.049$。
4. 結論: 支持現任市長之比例值之95% 信賴區間爲

$$0.25 \pm 0.049$$

Example

某公司進行最終產品檢驗, 共抽測 100 個產品, 有 10 個瑕疵品, 試求瑕疵率 p 的 95% 信賴區間。

♠

1. 計算基本統計量: $\bar{X} =$ ____________________。
2. 查表: ____________________。
3. 計算抽樣誤差: ____________________。
4. 結論: 瑕疵率 p 的95%信賴區間爲____________________。

抽樣誤差的上界 E^* 和最低信心水準, 樣本大小滿足

$$z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq E^* \tag{35}$$

據此決定應該抽取的樣本大小。根據式 (35), 我們可得

$$n \geq \frac{z_{\alpha/2}^2\hat{p}(1-\hat{p})}{E^{*2}} \tag{36}$$

但式 (36) 的 $\hat{p}$ 在抽樣前是未知的, 通常有兩種做法:

1. 因爲 $\hat{p}(1-\hat{p}) \leq \frac{1}{4}$, 我們可以取樣本大小爲

$$n \geq \frac{z_{\alpha/2}^2}{4E^{*2}} \tag{37}$$

這個方法得到的 n 會比實際上需要的還大, 在 $0.3 < p < 0.7$ 時是還算合適的, 但在 $p < 0.3$ 或 $p > 0.7$ 時, 這個方法過於保守, 發生 n 值過大浪費抽樣成本的情形。

2. 兩個階段方法, 第一階段先抽少量樣本 n_1 得到 $\hat{p}$, 代入式 (36), 得到 n 值。第二階段抽 $n_2 = n - n_1$。

Example

若某公司欲估計其生產線產品之不良 率, 今隨機抽取 50 個產品, 發現有5個不良 品, 請問需再抽取幾個產品作檢查, 方可使得估計誤差不超過 0.05?
$(\alpha = 0.05)$

♠

Example

在一次市政府之滿意度調查中, 若想以樣本比例值來估計市政府之滿意度, 請問在 95% 信心水準下, 欲使誤差在 0.03 之內, 至少需抽取多少位市民作調查?

♠

Example

2004總統大選在即, 多家民意調查機構不斷發布有關各組候選人的支持度, 一般而言, 這些民調結果顯示目前兩組候選人的支持度皆在誤差範圍內。

1. 假設某調查機構隨機成功電話訪問 1000 位合格公民, 試問在 95% 的信心水準下, 候選人支持度的誤差範圍爲多少?
2. 由於選情膠著, 爲降低抽樣誤差的大小, 以利區別候選人的支持度, 今此一調查機構欲使誤差範圍爲 1.5%, 試問應調查訪問多少人?

♠