Bootstrap 製程能力指標C_p, C_{pk} 及C_{pm} 的下限信賴區間

S. Balamurali 與M. Kalyanasundaram
印度Bharathiar 大學數學系

November 23, 2024

Abstract

能力指標被品質專業人員廣泛使用作為製程能力的估計。已提出並發展了許多製程指標,其中C_p、C_{pk} 和C_{pm} 最為廣泛使用。近期,已發展出技術來建構每個指標的95%下限信賴區間。這些技術基於底層製程為常態分配的假設。利用了非參數但運算密集的Bootstrap方法,並計算這些指標的Bootstrap信賴區間。使用三種分配(常態、對數常態和卡方)進行了模擬,並比較了Bootstrap和參數估計的表現。

1 簡介

近年來,製程能力指標(PCI)在品質保證和統計文獻中受到重視。製程位置、製程變異和目標位置的量化對於理解製程產品的品質至關重要。製程能力指標被產業界統計製程控制的許多倡導者視為實用工具。它們用於判定製造製程是否能在指定公差範圍內生產。能力指標是基於製程參數和製程規格的無量綱測度,旨在以簡單易懂的方式量化製程表現。製程指標C_p 和C_{pk} 作為無單位測度而普及,它們關聯自然製程公差(6σ)、上下規格限(見Kane (1986))。近期Chan 等人(1988) 開發了新指標C_{pm},其納入製程目標值(Hsiang 和Taguchi (1985) 及Taguchi (1986))。Chou 等人(1990) 提供了構建C_p 和C_{pk} 的95%下限信賴區間的表格。然而,他們對C_{pk} 的限制表格較為保守,建議採用Bissel (1990) 提出的近似值(見Franklin 和Wasserman (1992a) 及Kushler 和Hurley (1992))。最後,Boyles (1991) 提供了尋找C_{pm} 下限信賴區間的近似方法。

所有這些下限信賴區間的計算都假設製程為常態分配,而如Gunter (1989) 所指出,許多現實世界的製程並非常態分配,且這種偏離常態可能難以檢測。這可能影響指標的估計值和基於這些估計值的下限信賴區間。Efron (1979, 1981, 1982, 1985) 引入並發展了非參數但運算密集的Bootstrap估計方法。特別是Efron 和Gong (1983) 以及Efron 和Tibshirani (1986) 進一步發展了三種Bootstrap信賴區間,即標準Bootstrap(SB)信賴區間、百分位Bootstrap(PB)信賴區間和偏差校正百分位Bootstrap(BCPB)信賴區間。Franklin 和Wasserman (1991) 對這三種Bootstrap信賴區間在C_{pk} 上的性質進行了初步研究。Franklin 和Wasserman (1992b) 構建了一些基本能力指標的信賴限。

2 C_p、C_{pk} 和C_{pm} 的定義

若USL和LSL分別為上下規格限,且σ 為製程標準差,則

$$C_p = \frac{USL - LSL}{6\sigma}$$

當製程變異未知時,使用無偏樣本變異數S^2 (見Kane (1986))以獲得估計的能力指標

$$\hat{C}_p = \frac{USL - LSL}{6S}$$

由於指標C_p 未考慮製程平均值(μ)的位置,因此提出了指標

$$C_{pk} = \text{Min}\left(\frac{USL - \mu}{3\sigma}, \frac{\mu - LSL}{3\sigma}\right)$$

通常μ 和σ 均未知,典型上以$\bar{X}$ 和S 估計(見Kane (1986)),得到指標估計量

$$\hat{C}_{pk} = \text{Min}\left(\frac{USL - \bar{X}}{3S}, \frac{\bar{X} - LSL}{3S}\right)$$

由於C_{pk} 在製程居中於規格限之間時達到最大值(無論變異性大小為何),因此自然傾向於調整製程直到μ 恰好位於中點;然而,這可能不是最佳位置。因此,田口等人提出了指標

$$C_{pm} = \frac{USL - LSL}{6\sigma'}$$

其中$\sigma'^2 = E[(X-T)^2]$ 是製程X 圍繞期望目標值T 的變異。我們將使用估計量

$$\hat{C}_{pm} = \frac{USL - LSL}{6\hat{\sigma}'}$$

其中

$$\hat{\sigma}' = \sqrt{\frac{\sum (X_i - T)^2}{n}}$$

此估計量由Hsiang 和Taguchi (1985) 提出。Pearn 等人(1992) 對C_{pm} 的估計量進行了討論。

3 用於構建製程能力指標信賴區間的Bootstrap方法

令$X_1, X_2, ..., X_n$ 為大小為n 的隨機樣本,即一序列n 個獨立同分配的隨機變數,且$X_1 = x_1, X_2 = x_2, ..., X_n = x_n$。則定義一個在$\{x_1, x_2, ..., x_n\}$ 中取值的均勻分配隨機變數X^*:

$$P(X^* = x_i^*) = 1/n, \quad i = 1, 2, ..., n$$

Bootstrap估計基於從X^* 抽取的重複獨立樣本。共有n^n 個這樣的重複樣本。在我們的情況下,這些重複樣本將用於計算n^n 個C_p^*, C_{pk}^* 和C_{pm}^* 值。每一個這些值將是C_p^*, C_{pk}^* 和C_{pm}^* (分別)的估計,整個集合將構成$\hat{C}_p$, $\hat{C}_{pk}$ 和$\hat{C}_{pm}$ 的(完整)Bootstrap分配。

Bootstrap抽樣等價於從經驗機率分配函數中抽樣(有放回)。因此, $\hat{C}_p$, $\hat{C}_{pk}$ 和$\hat{C}_{pm}$ 的Bootstrap分配分別是C_p, C_{pk} 和C_{pm} 分配的估計量。實際上,通常只抽取n^n 個可能重複樣本中的一個隨機樣本,計算每個樣本的統計量,所得的經驗分配稱為該統計量的Bootstrap分配。Efron 和Tibshirani (1986) 指出通常1,000個Bootstrap重複樣本足以計算合理精確的信賴區間估計。

理論工作已發展出使用Bootstrap技術構建信賴區間的三種可能方法。在本討論中,假設假設取$B = 1,000$ 個Bootstrap重複樣本並計算C_p、C_{pk} 和C_{pm} 的$B = 1,000$ 個Bootstrap估計值(每個)並從小到大排序。將使用通用記號$\hat{C}$ 和$\hat{C}^*(i)$ (其中i 表示迭代次數)來表示能力指標的估計量和相關的有序Bootstrap估計值。例如,$\hat{C}_p(1)$ 是C_p 的1,000個Bootstrap估計值中最小的。將描述雙尾$(1-2\alpha)100$ 百分比信賴區間的構建,但通過僅使用下限可獲得單尾$(1-\alpha)100$ 百分比信賴限。

3.1 標準Bootstrap(SB)信賴區間的計算

從1,000個Bootstrap估計值$\hat{C}^*(i)$ 中,計算樣本平均值
$\hat{C}^* = (1/1,000)\sum_{i=1}^{1,000} \hat{C}^*(i)$
和樣本標準差
$S^* = \sqrt{\frac{1}{999}\sum_{i=1}^{1,000}[\hat{C}^*(i) - \hat{C}^*]^2}$
量S^* 實際上是C 標準差的估計量,因此(若C 的分配近似常態)C 的$(1-2\alpha)100$ 百分比SB信賴區間為
$\hat{C} \pm Z_\alpha S^*$
其中Z_α 是標準常態分配的上α 分位數。

3.2 百分位Bootstrap(PB)信賴區間

從$\hat{C}^*(i)$ 的有序集合中使用α 百分位和$(1-\alpha)$ 百分位點以獲得
$[\hat{C}^*(\alpha B), \hat{C}^*((1-\alpha)B)]$
作為C 的$(1-2\alpha)100$ 百分比PB信賴區間。對於$B = 1,000$ 的90
$[\hat{C}^*(50), \hat{C}^*(950)]$

3.3 偏差校正百分位Bootstrap(BCPB)信賴區間

僅使用完整Bootstrap分配的樣本所獲得的Bootstrap分配可能比預期偏高或偏低(即有偏分配)。因此,開發了第三種方法以校正這種潛在偏差(見Efron(1982)完整說明這方法)。首先,使用C^* 的(有序)分配,計算機率

$P_0 = P_r(\hat{C}^* \leq \hat{C})$

其次計算

$Z_0 = \Phi^{-1}(P_0)$ $P_L = \Phi(2Z_0 - Z_\alpha)$ $P_U = \Phi(2Z_0 + Z_\alpha)$

其中$\Phi(.)$ 是標準常態累積分配函數。最後,BCPB信賴區間由下式給出

$[\hat{C}^*(P_L B), \hat{C}^*(P_U B)]$

4 模擬

為比較Bootstrap下限信賴限與基於常態製程假設的信賴限的表現,進行了一系列模擬。所有模擬均使用$USL = 61$, $LSL = 40$ 和$T = 49$ (與Franklin 和Wasserman (1992b) 使用的相同值)。對每個使用的分配,對$\mu = 50$ 或52 且$\sigma = 2$、3 或3.7 的組合執行了六種不同的模擬。這產生了表I 中所示的每個製程指標的六個不同值。這些值的選擇旨在表示從"非常有能力"(即指標為1.5及以上)到"無能力"(即指標小於1.0)的製程。

對每個製程平均值和製程標準差的組合,抽取大小為$n = 20$、40或70的樣本,並從該單一樣本中抽取$B = 1,000$ 個Bootstrap重複樣本(每個大小為n)。用三種方法分別為三個指標構建95%Bootstrap下限信賴限。確定計算的Bootstrap限是否實際小於指標的真值。此外,對該樣本,使用Chou 等人(1990) 的表值(需內插)計算C_p 的常態導出95%下限信賴限,使用Bissell (1990) 建議的近似值計算C_{pk} 的信賴限,使用Boyles (1991) 建議的近似值計算C_{pm} 的信賴限。確定這些值是否實際小於指標的真值。

Table 1: 模擬研究中使用的每個製程指標的六個值

μ	σ	C_p	C_{pk}	C_{pm}
50	2.0	1.75	1.67	1.57
52	2.0	1.75	1.50	0.97
50	3.0	1.17	1.11	1.11
52	3.0	1.17	1.00	0.83
50	3.7	0.95	0.90	0.91
52	3.7	0.95	0.81	0.74

此單一模擬然後重複$N = 1,000$ 次。因此,我們能計算三種Bootstrap下限和常態導出下限小於相應指標的比例。這個"實際涵蓋比例"(將下限L 視為單側信賴區間(L, ∞))可與預期值0.95比較。對每個特定的n、μ 和σ 組合,使用三種可能的製程分配執行完整模擬:常態、對數常態和卡方分配。選擇偏態分配代表是因為,如Gunter(1989)所述,它們在實務中經常出現且特別麻煩。然而,選擇這些特定分配是因為它們易於產生和數學操作。後一特徵很重要,因為每個分配都經過平移和縮放以獲得表I所示的平均值和標準差。

5 模擬結果

5.1 常態分配製程

對於常態分配製程,基於常態性的方法和SB方法對所有製程指標都產生接近預期值0.95的比例(見表II)。下限的涵蓋頻率是參數為$p = 0.95$ 和$N = 1,000$ 的二項隨機變數。因此,涵蓋比例的99%信賴區間為

$0.95 \pm 2.576\sqrt{(0.95)(0.05)/1,000} = 0.95 \pm 0.0178 = (0.9322, 0.9678)$

因此,可以99%確信"真實95%下限信賴限"的涵蓋比例在0.932到0.968之間。事實上,108次中沒有一次超出這些限制(完全在預期的模擬表現之內)。這些結果傾向於驗證模擬(因為基於常態的限制表現如預期),也驗證了SB方法在底層製程為常態時對所有三個指標的涵蓋表現等效。

相比之下,PB和BCPB限制的涵蓋比例顯著低於0.95(見表II)。對於PB,54個可能的限制全部,對於BCPB,54個可能限制中的48個低於0.932。在54個情況中的每一個,BCPB限制都有比PB限制更高的涵蓋比例。

5.2 對數常態製程

對於對數常態製程,常態限制從未達到至少0.932的涵蓋率,而其他方法對任何指標都達到了至少0.932的期望涵蓋率(見表III)。此外,將基於常態性的限制與SB限制比較發現,SB限制表現顯著更好,對所有製程指標的涵蓋比例普遍更高。

關於另外兩個Bootstrap限制,BCPB限制比PB限制具有持續更高的涵蓋比例。對於$n = 20$,

BCPB限制的涵蓋率通常與基於SB方法的限制相當或更低。但對於$n = 70$,BCPB限制的涵蓋率通常更高(在$\alpha = 0.01$ 時經常有統計顯著性高於)於基於SB方法的限制。最後,所有四種方法的涵蓋比例都趨向於隨著n 增加而增加到預期值0.95。對基於常態性的方法和SB方法而言這種趨勢非常緩慢,對PB和BCPB方法而言則是中等緩慢。

5.3 卡方製程

對於自由度為4的卡方分配數據,四種可能的95%下限信賴限都從未達到期望涵蓋水準至少達0.932(見表IV)。此外,與對數常態分配相比,所有四種方法的涵蓋比例(對所有指標和樣本大小)都普遍較低。此外,基於常態性的限制和SB限制的涵蓋比例通常比較穩定,增加趨勢不明顯,而PB和BCPB限制只表現出緩慢增加到預期值0.95的趨勢。

然而,很明顯對於卡方分配,BCPB方法獲得的涵蓋比例與其他三種方法相當且經常顯著更大。具體而言,在54個可能的值對中,BCPB方法54次都有顯著大於(在$\alpha = 0.01$ 時)基於常態性方法的涵蓋比例。這種涵蓋差異出現在所有指標中,在$n = 70$ 時甚至大到0.091。基於常態性的限制從未顯著大於SB限制。此外,BCPB限制的涵蓋比例始終高於PB或常態性方法。

關於另外兩個Bootstrap限制,SB方法再次比PB方法有持續更高的涵蓋比例。對於$n = 20$,SB方法的涵蓋率通常與基於BCPB的方法相當或更低。