連串長度分佈與百分位：參數未知之舒瓦茲 $\bar{x}$ 管制圖

◎李麗女

一般公眾關心的事為有關一個管制圖的績效，但可能太過於注重平均連串長度(ARL)而錯過其重要的資訊，因為平均連串長度通常為高右偏的分佈而使得這樣的情況特別地真實。整個連串長度分佈必需加以診察，以得到一個較完整的管制圖績效研究，這可藉由對數個代表性的百分位包括中位數、及某些百分位函數如四分位差等加以解釋。Khoo(2004)研究舒瓦茲x管制圖百分位，當製程的平均值與變異數是明確的x管制圖(所謂的”統計量已知”之情況)；本文中我們仔細檢視連串長度分佈及舒瓦茲x管制圖百分位，於製程的平均值與變異數是未知的(所謂的”統計量未知”之情況)加以推估；評估並繪出精確的連串長度累積分佈函數，包括次群組組數(m)、一個次群組的樣本大小(n)5及名目誤警報率(FAR)為0.0027；在統計量已知之情況下，連串長度累積分佈函數已知為幾何分佈，也被納入比較；此外計算一些特定的百分位並與那些統計量已知之情況相互比較，可以發現對於小至中的m值，連串長度分佈既不會管制幾何分佈亦不被幾何分佈管制。當參數是被推估出來且製程是在管制內的，百分位之順序是小於並接近0.82者，對小至中的m值而言，其先期的連串之累積機率會較大；但對百分位超越那個數值者，當在幾何分佈下其末期的連串之累積機率會較小。在不受管制的情況下(對一個階級位移0.5)一個類似的現象於百分位順序接近0.62者是可被理解的。

對m=500及n=5的情況下，連串長度分佈在統計量未知之情況是收斂於統計量已知之情況所謂的幾何分佈。在此根據管制內的中位數連串長度提出一個替代的管制圖設計準則。

1.前言

在確定一個製程是否為統計管制內的管制圖是相當有用的，實務上一個管制圖的績效是一項重要的考量，而且典型上依特定的度量或特性隨同其連串長度分佈來做判定。連串長度是一個隨機的變數定義為次群組的數量，必需蒐集(或等同地為必需被繪製的統計量之個數)直到從管制內的製程中觀察到第一個信號的改變。就如Radson及Boyd(2005)和其他人數十年來所提到的，一個管制圖的績效之傳統度量方式為連串長度分佈之”平均數”或”平均值”或”期望值”，亦即所謂的平均連串長度(ARL)。ARL如其所訴求的是具有吸引性的，平均而言在取得一個管制外的信號前，必需要等待多久或是需要繪製多少個統計量圖；無論如何，按照定義連串長度隨機變數只能採用正整數，而且考慮到事實的分佈形狀是明顯地右偏，研究人員致力於使用其它較具代表性的度量，以評估管制圖的績效；此類的一個度量為連串長度分佈之某一個百分位。在統計製程管制圖中百分位的概念可以追溯到Barnard(1959)，百分位提供有關分佈的重要資訊，也提供一個管制圖的績效，績效並非由平均數提供的。在敘述性統計中用於描述一個分佈所選定的綜合度量值是一個重要的主題；已知在一個右偏的的分佈中，平均值是大於中位數，因此平均值通常並不能公正地代表一個”典型”或”中心”。在這樣的觀念，中位數是一個較佳的”中央趨勢”的度量值，也因此中位數連串長度(MRL)在典型的管制圖績效是一個較準確的指標。例如，假設MRL

等於250，我們可得知第一次的管制圖信號在第250個樣本或之前出現的機率至少是50%；此外，中位數是一個穩健的度量值，很少被偏離值或異常點所影響。

對中心而言此兩個綜合度量值即為平均值及中位數間的選擇，並另有其它的隱含之意以描述其分散性、變異性或廣度性。例如，假設挑選的是平均數做為”典型”，則使用標準差以度量與中心的變異性，但假設挑選的是中位數，則以第三個四分位與第一個四分位間的差即四分位差(IQR)；或雙尾順序百分位的差，如第95百分位和第5百分位以度量分散性或廣度性。一般而言，當一個分佈是偏移的，建議使用百分位以描述中央性及分散性兩者。Khoo(2004)曾研究$\bar{x}$管制圖之連串長度分佈百分位；Radson和Boyd(2005)推荐使用統計製程管制圖的百分位，他們並提出一個百分位分佈(PD)圖，透過第5百分位、第25百分位、第50百分位、第75百分位和第95百分位描述連串長度分佈；Shmueli及Cohen(2003)也致力於使用連串長度百分位。所有這些論文報告只考慮到統計量已知的情況(例如Montgomery，2004)，亦即當製程參數是已知的或已被指定的情況下之研究。

在本論文報告中，當製程參數是被推估的(統計量未知之情況)，研究連串長度分佈及其百分位時，我們考慮用雙邊的舒瓦茲$\bar{x}$管制圖平均值。雙邊管制圖追蹤平均值在兩個方向的高低變化，也可以運用相似的想法，以研究其它可用的管制圖之連串長度分佈，如針對連續的變數及類別的變數兩者之累計和(CUSUM)及指數加權移動均值(EWMA)。但是這裡使用$\bar{x}$管制圖，因為實務上它是最受歡迎的圖，不只有可用的公式而且計算也相對地簡單容易。如同往常，假設過程不是常態性的分佈就是滿足中央極限定理的運用條件。在統計量已知的情況下，舒瓦茲$\bar{x}$管制圖的連串長度分佈已知是為幾何分佈(Montgomery，2004)，當製程是管制內的，該分佈的成功機率(信號)是等於誤警報率(FAR)α(典型上為0.0027)。但是，舒瓦茲$\bar{x}$管制圖的很多運用裡，製程參數如平均值與變異數是未知的或未被指定的，在這樣的情況下未知參數首先是由階段1的分析中獲取一組管制內的或參考的資料，由此推估出來的；實務上管制圖的特性極限之推估效應是顯著的，最重要的是連串長度分佈不再是幾何性的(Quesenberry，1993)。因此，如一個或多個參數是未知的，則一個3-sigma管制圖的FAR是逼近0.0027(或ARL_0將接近370)，此並非是真實的。除非有大量的資料可用，事實上已有論文顯示出(Quesenberry，1993及Chakraborti，2000)真實的FAR(或ARL_0)可能在統計量已知的情況下其相對應值是截然不同的。

數位作者已致力於推估的參數值下之管制圖領域，包括Ghosh, Reynolds及Van Hui(1981)、Quesenberry(1993)、Del Castillo(1996)、Chen(1997)、Chakraborti(2000)和其他人。參數推估本質上會影響連串長度分佈現在已廣為大家所接受，也因此對管制圖績效有重要的影響性；在設計一個管制圖時就需考慮到參數推估的效應。在所有人中，當不管是一個或兩個製程參數－平均值與標準差是被推估出來的，Chakraborti(2000)提供一個精確的連串長度分佈表示法。由Jensen et al.(2006)中也發現到各種管制圖特性對參數推估效應之回顧；雖然有數位作者建議對百分位做研究，特別當參數是推估的，但一個精確的、徹底的分析整個連串長度分佈及其百分位，現行是不可得的。本文的目的是要弭平部份的分歧，以展現當一個或較多個的參數是被推估出來的，仍可精確的計算出舒瓦茲$\bar{x}$管

制圖的連串長度分佈之百分位，因此將Khoo(2004)的著作列入屬於推估的參數之領域。如前所提及，一個百分位如MRL是一個較佳的典型管制圖績效指標；其它的百分位如四分位，及截尾百分位如第5個百分位與第95個百分位，提供有關連串長度分佈之廣度性資訊。我們的成果是當參數是推估的，提供在實務上對舒瓦茲x管制圖之績效較具洞察力，這對實務者及研究者都極具吸引性。

2.連串長度分佈

2.1.平均值及變異數兩者是確定的(統計量已知)

當平均值(μ_0)和變異數(σ_0^2)為已知，想想一個母體之平均值(μ)的舒瓦茲x管制圖，典型的雙邊舒瓦茲x管制圖使用於此個案中為已知$\mu_0 \pm a_n \sigma_0$，其中$a_n = z_{\alpha/2}/n^{1/2}$,統計量$z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$為對應到名目的(指定的)FAR($\alpha$)之製圖常數，而Φ代表標準常態分佈之累積分佈函數，當α為0.0027時製圖常數大約等於3，我們也可得到典型的3-sigma管制圖。令N為連串長度亦即直到第一個信號在圖上觀察到所需次群組的個數，N的分佈為有興趣於研究管制圖之功效。在統計量已知的情況下，它是相當知名的(如Montgomery，2004)，N屬於幾何分佈，因此，管制圖的所有特性可經由幾何分佈的結果推得出來。例如，N的累積分佈函數已知為$P(N \le a) = 1-(1-\beta)^a \quad a=1,2,\ldots(1)$，其中$\beta$代表一個信號的機率(或信號性機率)，該幾何分佈是屬於右偏的，而且通常在SPC所面對的小數值的信號性機率β時偏離性較顯著。在管制內的情況下β等於FAR、於很多應用中α則取0.0027。為了描繪管制圖的績效，通常是看連串長度分佈的平均值亦即是眾所周知的ARL。但如我們先前的解釋，因為連串長度分佈的長右尾，ARL無法對管制圖的績效做很好的詮釋；因此雖然是管制內的ARL(ARL_0)在統計量已知的情況下，可以很容易的計算出來(在此$ARL_0 = 1/\alpha$)，卻不被做為管制圖的管制內績效之唯一度量值。

2.1.1. 連串長度百分位

百分位是一個有益的及穩健的管制圖績效度量值，第100p($0 < p < 1$)百分位定義為最小的整數t，為累積分佈函數在t時至少等於p，因此利用等式(1)第100p百分位t可由$1-(1-\beta)^t \ge p$求得，其減化了找尋最小正整數t，因此$t \ge \ln(1-p)/\ln(1-\beta)\ldots(2)$，其中$\beta$為一個信號的機率；例如，第50百分位或中位數連串長度(MRL)，設p=0.5可由等式(2)中求得。因此利用$\beta = \alpha = 0.0027$及$p = 0.5$，一個管制內MRL值，對FAR=0.0027其ARL_0為257，但ARL_0為$1/\alpha = 370.4$，其比MRL_0大了31%。該詮釋說明也是相當的不同，ARL_0象徵在一個管制內的製程下，平均而言每370個樣本將會觀察到一個信號(一個錯誤的警訊)；另一方面，MRL_0象徵至少有50%的情況下，在最先的257個樣本中將會觀察到一個錯誤的警訊。正如另一個例子，第5個連串長度百分位為19，因此在最先的19個樣本內並不會發生錯誤的警訊，其機率是小於95%。因為觀察到連串長度分佈是離散型的，因此累積分佈函數是一個階梯函數，通常的慣例是定義百分位以使累積分佈函數之百分位至少(非準確地)等於100p%。整個連串長度分佈提供對管制圖績效相當有用的資訊，而且所選定的百分位也有助於彙整該資訊。

2.2.平均值及變異數兩者皆未知的(統計量未知)

當製程參數未知時，一直到他們的值被推估出來才能計算出管制界限；通常當

製程是管制內的(這樣的資料被視為是階段I或參考資料)，而且得到的管制圖適用於預期的(階段II)監控，則分別從m個次群組中每組大小為n的資料以推估出參數(以及界限)。眾所週知(Quesenberry，1993及Chakraborti，2000)只要一使用這些推估的管制界限，管制圖的特性基本上就已改變了；例如，N的連串長度分佈不再是幾何分佈的。當製程的平均值及變異數兩者未知的情況下，以說明我們研究舒瓦茲$\bar{x}$管制圖的連串長度分佈；其它的參數推估，例如當平均值已知但變異數未知等情況下也可以以相類似的方式處理。

假設階段I或參考資料的大平均值$\bar{x}$，為m組、每組大小n之管制內樣本之平均值，δ用於推估製程平均值μ，而且$\hat{\sigma} = S = \sqrt{\frac{1}{m} \sum_{i=1}^{m} S_i^2}$ 使用於推估製程標準差σ其中s_i^2為第i個參考樣本的變異數；這些估計值的選定在基本的製程屬常態性之假設是根深柢固的。大平均值是最佳的(均一的最小不偏變異數)平均值估計值，而且即使估計值σ(Montgomery，2004)有數個選擇，Derman及Ross(1995)和其他的人已證明從均方差(MSE)判斷，s較合意而且從統計的觀點看亦較方便於使用。舒瓦茲$\bar{x}$管制圖之推估的平均值與變異數為$\bar{\bar{x}} \pm a_n s$，其中$a_n = z_{\alpha/2}/n^{1/2}$。

2.2.1. 連串長度分佈

當平均值與變異數以$\bar{\bar{x}}$及s推估時，N的分佈不再是呈幾何分佈，可以證明N的確切累積分佈函數已知為(Chakraborti，2000)

p=(N≤a)

$$=1-\int_{\infty}^{\infty}\int_{0}^{\infty}\left[\Phi\left(-\delta\sqrt{n}+\frac{Z_{\alpha/2}}{\sqrt{v}}\sqrt{y}+\frac{z}{\sqrt{m}}\right)-\Phi\left(-\delta\sqrt{n}-\frac{Z_{\alpha/2}}{\sqrt{v}}\sqrt{y}+\frac{z}{\sqrt{m}}\right)\right]^{a}\times f_{\chi_v^2}(y)\,dy\,\phi(z)\,dz$$

，a=1,2,...(3)，其中$\delta > 0$為平均值依照標準差σ所做的位移，$\nu = m(n-1)$為自由度，α為名目的FAR，$f_{x_v^2}$為ν個自由度的卡方分佈之機率密度函數，而φ為標準常態分佈之機率密度函數，附錄提供了詳細的說明。因此，對舒瓦茲$\bar{x}$管制圖之推估的平均值及變異數而言，真實的(得到的)FAR值即使是以$z_{\alpha/2}=3.0$作為管制圖常數也已不再是名目的0.0027。事實上，除非資料點的總數mn非常大，如大約為1,000，3-sigma($z_{\alpha/2}=3.0$)管制圖之真實的FAR值，實際上可能大於0.0027(Quesenberry，1993及Chakraborti，2000)。

等式(3)中的累積分佈函數完整地指定了連串長度分佈及管制圖的績效，例如，當a=2我們可用它以找出第一個信號在第二個或之前的樣本出現的機率，兩個連續累積機率在所謂a及(a-1)時之相差值，剛好是第一個信號在第a個樣本觀察到的機率。在管制內的情況下$\delta=0$，否則δ等於某個正數值代表製程平均值位移了δ，可使用如Mathematica或Mathcad軟體以計算所需的積分值，本文使用後一個軟體。

2.2.2. 連串長度百分位

等式(3)之連串長度累積分佈函數可用於研究管制圖的統計特性，其包括各種不同的績效特性。例如，從定義中最小的正整數t可以計算出第100p個百分位，因此累積分佈函數在t時至少等於p。管制內的百分位，在累積分佈函數中可用$\delta=0$替代以算出來；但是，管制外的情況下，百分位可用某個正數值δ替代以算出來。

3.例子

為了說明，當參數是推估的，我們評估並繪製連串長度分佈之累積分佈函數的舒瓦茲$\bar{x}$管制圖，圖1及圖2分別是當位移為$\delta=0$及$\delta=0.5$，其m=30、100、500，n=5及$\alpha=0.0027$；我們也顯示幾何分佈之累積分佈函數，在統計量已知之連串長度分佈情況下，每一個圖之一個信號機

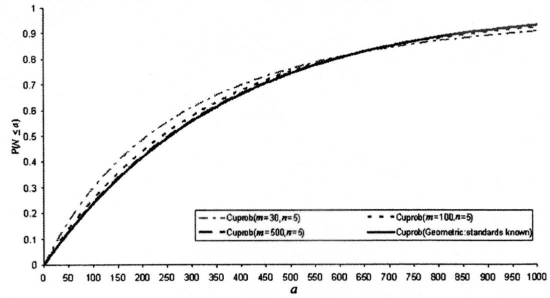

圖1 In-Control Run Length c.d.f. of the Shewhart x̄ Chart in the Standards Unknown Case for m=30,100,500, n=5, α =0.0027

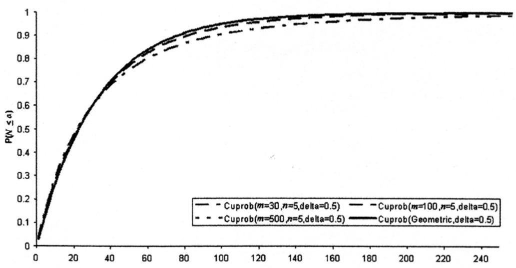

圖2 Out-of-Control Run Length c.d.f. of the Shewhart x̄ Chart in the Standards Unknown Case for m=30,100,500, n=5, α =0.0027 and δ =0.5

率p=0.0027以做為參考。雖然一個離散的隨機變數之累積分佈函數，按照定義為階梯函數，但由於繪製了大量的點，因此圖形看起來似乎是連續的；該圖顯示除非m夠大到大約是500，否則連串長度分佈可能相當地不同於幾何分佈。在管制內的情況下，當m=30和100時，連串長度之累積分佈函數落在幾何分佈高達(接近)第82個百分位之上；但是，除了第82個百分位外，連串長度之累積分佈函數亦落在幾何分佈之下。在非管制內的情況下也可看到相同的現象，但是大約是在接近第62個百分位。因此，當參數是推估的而且m的大小是小至中，則連串長度分佈並未優於、也未劣於統計量已知之連串長度分佈，因為累積分佈函數已決定了結果。這個顯示了連串長度分佈並非隨機取決於幾何分佈；當一個隨機變數的累積分佈函數整個落在另一個隨機變數的累積分佈函數之上(之下)，則第一個隨機變數是比第二個隨機變數(參看Gibbons及Chakraborti，2003)為隨機過程地小(大)。因此，假使一個隨機變數是隨機過程地大於另一個，則它的百分位及平均值皆較大。另一方面，對m=500在統計量已知情形下之雙尾分佈，連串長度累積分佈函數收斂於此；無論如何，順序百分位小於等於0.82(0.62)者遞增，而順序百分位大於等於0.82(0.62)者遞減，其分別地極相像於在統計量已知情形下之管制內與管制外的情形。相類似的結論可以由表1中推論得出，並將於後面再討論。

表1 Percentiles and the Average Run Lengths of the Shewhart $\bar{x}$ Chart, When Process Meean and Variance are Unknown, Given The Number of Subgroups (m) and Shift Size (δ) With n=5 and α=0.0027

			Percentiles of the run length distribution: standards unknown												
m	Shift δ	ARL	5	10	20	25	30	40	50 (MRL)	60	70	75	80	90	95
20	0.0	422.29	12	25	54	71	90	136	194	274	390	473	582	997	1540
	0.2	250.84	7	14	30	39	49	75	108	154	223	272	338	592	931
	0.6	27.25	1	2	4	6	7	10	14	19	26	32	38	63	95
	1.0	5.14	1	1	1	2	2	2	3	4	5	6	7	11	16
30	0.0	398.77	14	29	62	81	101	150	211	291	405	482	584	947	1390
	0.2	224.52	8	15	33	43	54	80	113	157	221	265	323	535	799
	0.6	24.65	1	2	5	6	7	10	14	19	26	30	36	57	83
	1.0	4.91	1	1	1	2	1	2	3	4	5	6	7	11	15
50	0.0	384.19	16	32	69	90	112	163	227	308	418	492	586	908	1274
	0.2	205.03	8	17	35	46	57	84	117	159	218	258	309	486	693
	0.6	22.86	1	1	5	6	7	10	14	19	25	30	35	53	73
	1.0	4.73	1	1	1	2	1	2	3	4	5	6	7	10	14
100	0.0	375.91	18	36	76	98	122	175	241	322	431	501	590	879	1190
	0.2	191.10	9	18	37	48	60	87	120	161	216	252	298	448	613
	0.6	21.66	1	2	5	6	7	11	14	19	25	29	34	50	67
	1.0	4.61	1	1	1	2	2	3	3	4	5	6	7	10	13
Standards Known	0.0	370.37	19	39	83	107	132	189	257	339	446	513	596	852	1109
	0.2	177.72	10	19	40	40	64	91	123	163	214	214	286	409	531
	0.6	20.56	2	3	5	5	8	11	14	19	25	25	33	47	61
	1.0	4.49	1	1	1	1	2	3	3	5	5	6	7	10	12

當計算連串長度機率時，也令人關注到連串長度分佈之最高機率發生在a=1時，因此連串長度分佈是單一型的眾數發生在1；舒瓦茲$\bar{x}$管制圖中，下一個抽樣樣品能偵測到一個持續不變的位移最大機率值，是發生在位移一發生之後；事實上，此為舒瓦茲管制圖的一項優勢(Quesenberry，1993)。

針對樣本數m=20,30,50,100，位移大小δ=0,0.2,0.6,1.0及樣本大小n=5，我們計算管制內及管制外兩者之一些連串長度分佈之百分位，取名目FAR值α為0.0027，但在統計量未知情形下，沒有明確的公式(如等式(2)為統計量已知情形)，而是利用累積分佈函數及定義來計算百分位，結果展露於表1中。為了比較的目的，納入統計量已知(幾何分佈)情形下的計算結果。在此，非0位移的結果與Khoo(2004)的結果非常的不同，只因為在此定義位移為$\delta=(\mu_1-\mu_0)/\sigma$，但是Khoo定義為$\delta\sqrt{5}$。透過等式(3)評估累積分佈函數需要多重積分，在此我們使用Mathcad以進行所有的計算。

研究這些百分位後浮現出一個較清晰的圖像，例如管制內($\delta=0$)，對m=20,30,50及100的第50個百分位(MRL_0)分別為194,211,227及241，這些數值與管制內統計量已知之MRL_0=257比較，可理解到當參數是推估的MRL_0較短，當m大小適中意味其比名義上統計量已知α=0.0027的情況有較多的錯誤警訊；事實上，相對應的FAR值分別顯示為0.0044,0.0038,0.0033及0.0030；另一方面ARL_0分別為422.29,398.77,384.19及375.91，其彰顯穩定的連串長度分佈之右偏形狀。對小至中的m值而言，連串長度分佈是較右偏的；這可由比較m=20,20,50

及100時之百分位看出。再者，當討論到累積分佈函數，可觀察到當參數是推估的，連串長度分佈的尾端表現得迥異於幾何分佈；因此，對$p \leq 0.82$(接近)的百分位是較那些統計量已知的小，而對$p > 0.82$的百分位則較大。例如，在管制內於m=30及n=5時，當參數是推估的第90和第95百分位分別為947和1390，當統計量已知時則分別為852和1109；另外在相同的情況下，於參數是推估的和統計量已知時，第40百分位分別為150和189。當m=500時，將有500×5=2,500個參考資料點用於推估平均值及變異數，連串長度分佈及百分位皆收斂到統計量已知時的數字。因此對大量的階段I資料，推估的管制圖行為尤如統計量已知的情形；當使用更多的資料估計，中位數(MRL)的行為更不同於平均數(ARL)；在統計量已知下MRL_0增加到它們的對應值，而ARL_0下降。在任何情況下，中位數與平均數保持相當的分離，顯示即使階段I資料很大時連串長度分佈仍維持右偏。

五個百分位的揭露可能有所幫助：如表1所示的第5個, 第25個, 第50個, 第75個和第95個百分位，在一個PD(Probability of Default，違約機率)圖(Radson及Boyd,2005)中更易顯現與存取整個連串長度分佈。相鄰的PD圖將有助於比較兩個以上的管制圖。順著這條線索，我們提出使用一個管制內的百分位如中位數(MRL_0)，以設計一個管制圖，雖然現行產業界實務上是使用ARL_0。例如，假設m=50和n=5，欲使95%管制內的連串大於100，解決等式"第5個管制內百分位=100"之管制圖係數，得到$z_{\alpha/2}=3.7409$而且相對應到一個FAR，$\alpha=0.000183$。很明顯地，此一管制圖是相當的寬廣；可能反而較務實的是尋找至少50%管制內的連串大於300，如此其亦闡明了$MRL_0=300$，這產生了一個管制圖係數3.2876其具有FAR為0.0010；應用確定的管制內百分位如MRL_0以設計一個圖可以引導出一個較穩健的管制圖。

4.總結與建議

傳統的度量值如ARL和SDRL並未對一個統計圖的績效提供一幅完整的圖像，因為連串長度分佈是高度的右偏；當一個或更多個的參數是推估出來的，以此推估值建立適當資料量的管制圖，此典型上廣為產業界所使用，但對一個極小機率的錯誤警訊將使得分佈右偏更加嚴重。連串長度分佈百分位包括中位數連串長度，其提供了一個較佳的管制圖績效指標。再者，度量值乃根據百分位函數算出來的，如IQR或第95個和第5個百分位間的相差值，它們對提供連串長度分佈的散佈情形是相當有用的。在例行性的應用上，特別是舒瓦茲$\bar{x}$管制圖，我們建議計算和研究五個百分位：第5、第25、第50、第75和第95個百分位。

附錄

1.統計量已知

當事件有一個信號，亦即當$\frac{|\bar{x}_i-\mu_0|}{\sigma_0} \geq a_n$或等同於當$\sqrt{n}|\bar{x}_i-\mu_0| \geq \sigma_0 z_{\alpha/2}$ 即所謂的一個信號事件。在統計量已知下，信號事件對$i=1,2,\ldots$是獨立的且已知，N服從幾何分佈其成功機率（亦即一個信號）為$P(\sqrt{n}|\bar{x}_i-\mu_0|) \geq \sigma_0 z_{\alpha/2}$。後者的機率等於

$$1-\{\Phi(Z_{\alpha/2}-\delta\sqrt{n}\}-\Phi(-Z_{\alpha/2}-\delta\sqrt{n})\}=\beta(\delta,n)\ldots(4)$$

，其中$\delta=(\mu_1-\mu_0)/\sigma$且$\Phi$為標準常態分佈的累積分佈函數。

再者從幾何分佈的特性得知，N

的累積分佈函數為P(N≤α)=1-(1-β(δ,n))[a]其中α=1,2,3,…(5)，連串長度分佈的第p個百分位k_p為最小的整數，使得P(N≤k_p)≥p…(6)，方程式(6)中的不等式起因於連串長度分佈為離散的事實，同時，N分佈的平均值等於ARL(n,δ)=[β(δ,n)]$^{-1}$…(7)。當製程是管制內的，δ=0之信號機率或FAR等於β(0,n)=P(｜Zi｜＞$Z_{\alpha/2}$)=α…(8)，而且管制內的ARL亦即管制內的連串長度分佈之期望值等於ARL(n,0)=ARL_0=1/α…(9)。

2.統計量未知

為了方便獲得連串長度分佈，在這個個案中乃取決於$\bar{\bar{x}}$和S，這兩個統計量附有條件，使得連串長度隨機變數N服從幾何分佈信號機率等於如下

$$1-\Phi\left(-\delta\sqrt{n}+z_{\alpha/2}\frac{S}{\sigma}+\frac{\sqrt{n}\left(\bar{\bar{X}}-\mu\right)}{\sigma}\right)+\Phi\left(-\delta\sqrt{n}-z_{\alpha/2}\frac{S}{\sigma}+\frac{\sqrt{n}\left(\bar{\bar{X}}-\mu\right)}{\sigma}\right)\ldots(10),$$

重寫方程式(10)是相當有幫助的，其為

$$1-\Phi\left(-\delta\sqrt{n}+\frac{z_{\alpha/2}}{\sqrt{\nu}}\sqrt{Y}+\frac{Z}{\sqrt{m}}\right)+\Phi\left(-\delta\sqrt{n}-\frac{z_{\alpha/2}}{\sqrt{\nu}}\sqrt{Y}+\frac{Z}{\sqrt{m}}\right)=\beta\left(\delta,m,n,Y,Z\right)\ldots(11)$$

，其中 $Z=\sqrt{mn}\left(\bar{\bar{X}}-\mu\right)/\sigma$且$Y=\nu S^2/\sigma^2$ 。

現在利用下列4個事實，可以計算連串長度分佈，(i)Z服從標準常態分佈，(ii)Y服從 χ^2_ν 分佈，自由度為ν=m(n-1)，(iii)隨機變數Z和Y是在正常下其隨機過程相互獨立的，而且(iv)附有條件的Z和Y，連串長度變數N為幾何分佈，成功機率β(δ,m,n,Y,Z)，利用(iv)和幾何分佈的特性，首先得出，條件連串長度累積分佈函數為P(N≤a｜Z,Y)=1-{1-β(δ,m,n,Y,Z)}^{a}…(12)，之後，針對條件累積分佈函數積分，以得出非條件連串長度累積分佈函數，這即如下面的式子

$$P(N\le a)=1-\iint\left\{1-\beta\left(\delta,m,n,Y,Z\right)\right\}^a\times f_Y(y)f_Z(z)dydz$$

$$=1-\int_{\infty}^{\infty}\int_{0}^{\infty}\left[\Phi\left(-\delta\sqrt{n}+\frac{Z_{\alpha/2}}{\sqrt{v}}\sqrt{y}+\frac{z}{\sqrt{m}}\right)-\Phi\left(-\delta\sqrt{n}-\frac{Z_{\alpha/2}}{\sqrt{v}}\sqrt{y}+\frac{z}{\sqrt{m}}\right)\right]^a\times f_{\chi^2_v}(y)dy\varphi(z)dz\ldots(13)$$

若欲知更詳細的資料可參看Chakraborti(2000)。

(資料來源：譯自S. Chakraborti (2007) Run Length Distribution and Percentiles: The Shewhart Chart with Unknown Parameters Quality Engineering 19 (2): 119-127)