

# Comparative Study of Shewhart Control Chart and CUSUM Control Chart

Lin Jing

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming Yunnan  
Email: jinglinscnu@126.com

Received: Jul. 4<sup>th</sup>, 2016; accepted: Jul. 24<sup>th</sup>, 2016; published: Jul. 27<sup>th</sup>, 2016

Copyright © 2016 by author and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

This paper firstly presents a brief introduction to the emergence and development for the figure of statistical process control, and analyzes the advantages and disadvantages between Shewhart control chart and the CUSUM control chart, and then gives the Shewhart control chart estimation method and the theoretical basis and construction method for the CUSUM control chart. Finally, we compare the average run length in the process of controlled and uncontrolled under the two kinds of situations with Shewhart control chart and the CUSUM control chart. It is concluded that the CUSUM control chart for monitoring small process mean shifts is more sensitive than the Shewhart control chart.

## Keywords

Shewhart Control Chart, CUSUM Control Chart, Average Run Length

---

# Shewhart控制图与CUSUM控制图的比较研究

敬 林

云南财经大学统计与数学学院, 云南 昆明  
Email: jinglinscnu@126.com

收稿日期: 2016年7月4日; 录用日期: 2016年7月24日; 发布日期: 2016年7月27日

---

## 摘 要

本文首先简单介绍了统计过程控制图的产生及其发展, 分析了一般的Shewhart控制图和CUSUM控制图

的优缺点, 其次给出了Shewhart控制图的估计方法和CUSUM控制图的理论基础及其构造方法。最后, 我们比较了在过程受控与失控两种情况下的一般的Shewhart控制图和CUSUM控制图的平均运行长度, 得出了CUSUM控制图对监测过程均值小漂移比一般的Shewhart控制图更灵敏的结论。

## 关键词

Shewhart控制图, CUSUM控制图, 平均运行长度

## 1. 引言

现代质量管理要求以统计技术的应用来实现过程控制, 因此称为“统计过程控制”, 统计过程控制强调预防为主的原则, 因而在质量管理中受到管理者的青睐。实现统计过程控制的过程称其为处于“稳定受控状态”或“正常状态”或“统计稳态”。控制图是生产过程中控制过程质量的一种有效工具, 它能直观地反映出加工过程质量情况。它形象直观, 图上有用实线绘制的中心线和虚线绘制的上下控制线, 图中有按时间先后顺序抽取的样本的点列, 各点之间用直线段连接, 可以清晰直观地看出变化趋势。控制图就相当于通常的正态分布旋转 90 度后形成的。将  $\mu$ ,  $\mu+3\sigma$  以及  $\mu-3\sigma$  分别标注为中心线, 上控制线, 下控制线, 这样就得到了一个基本的控制图[1]-[3]。Shewhart 控制图理论基础是“3 $\sigma$ ”原理, 也就是说在 1000 个产品中有不超过 2.7 个不合格品出现, 就认为该过程的波动属于正常波动; 否则, 就认为该过程的波动属于异常波动。Shewhart 常规控制图只是利用了过程中的当前数据, 而没有充分利用整个样本中的数据。Shewhart 控制图对当前数据的信息进行了充分的利用, 但是一旦“点出界就报警”, 并且所有历史数据的权重都为 0。由于 Shewhart 控制图是建立在数理统计的原理之上, 因此在质量控制过程中需要满足样本容量要大, 样本服从或接近正态分布, 样本相互独立等条件。而在实际控制过程中这些条件很难被满足, 从而造成 Shewhart 控制图在实际运用中的各种不适应性。针对这种情况人们做了大量的研究, Page 基于序贯分析原理于 1954 年提出了累积和控制图简称为 CUSUM 控制图。将一系列点子(数据在控制图中的表达)的微弱信息累加起来, 就对过程中的微小波动也变得灵敏了。CUSUM 控制图不仅拓宽了 Shewhart 控制图的应用范围, 而且弥补 Shewhart 控制图对中小漂移和异常点的检测灵敏性。控制图方法因具有捕捉异常先兆的功能, 因而成为实现统计过程控制的首选方法。在统计过程控制中, 应用统计技术实现过程中各个阶段的监控和诊断, 捕捉到异常先兆后立即采取纠正和预防措施, 将异常消灭在萌芽状态, 消灭在过程之中。因而有效减少过程中的质量损失, 降低成本, 对保证产品质量和增加企业的经济效益都起到了非常重要的作用。

## 2. Shewhart 控制图的区间估计

### 2.1. 过程总体均值和方差已知的情况

假定一系列样本  $X_i$ ,  $i=1,2,\dots$  来自一生产过程, 它们独立服从均值  $\mu$ , 标准差为  $\sigma$  的正态分布, 即  $X_i \sim N(\mu, \sigma^2)$ 。  $X_{i_1}, X_{i_2}, \dots, X_{i_n}$  ( $i=1,2,\dots,t$ ) 表示在时刻  $t$  从生产过程中抽取的  $n_i$  个样本, 假定每个时点抽取的样本数量恒为  $n$ , 则容量为  $n$  的样本均值  $\bar{X}_i \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ 。当  $\mu$  和  $\sigma$  都已知时,  $\bar{X}_i$  以  $1-\alpha$  的概率落在区间内, 其控制区间为

$$\left[ \mu - U_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \mu + U_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \quad (1)$$

在实际操作中, 取  $U_{1-\alpha/2}$  为 3, 即实际生产中的  $3\sigma$  控制线。以上我们对样本  $X_i, i=1, 2, \dots, n$  作了正态性假设, 但是根据中心极限定理, 当样本不符合正态性假定时, 只要样本数量较大一般大于 50 即可, 那么样本也近似服从正态分布, 故 Shewhart 控制图的结果仍然可用。

## 2.2. 过程总体均值和方差未知的情况

当  $\mu$  和  $\sigma$  都未知时, 假设我们有  $m$  个样本容量为  $n$  的样本, 则均值的估计值为

$$\hat{\mu} = \bar{\bar{X}} = \frac{\bar{X}_1 + \bar{X}_2 + \dots + \bar{X}_m}{m} \quad (2)$$

又总体标准差的无偏估计为  $\hat{\sigma} = \frac{\bar{R}}{d_1}$ , 设  $R_1, R_2, \dots, R_m$  是  $m$  组样本的极差,  $\bar{R}$  是这些极差的平均值, 且总体均值的估计值为  $\bar{\bar{X}}$ 。因此, 每组样本的均值  $\bar{X}_i$  以  $1-\alpha$  的概率落在区间

$$\left[ \hat{\mu} - u_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}}, \hat{\mu} + u_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}} \right] \quad (3)$$

之内。所以在总体均值与标准差的估计值均已知的前提下, 我们代入  $\mu$  和  $\sigma$  的估计值就得到具体的检测过程均值的 Shewhart 控制图的

上控制线为:

$$\bar{\bar{X}} + u_{1-\alpha/2} \frac{\bar{R}}{d_1 \sqrt{n}} \quad (4)$$

下控制线为:

$$\bar{\bar{X}} - u_{1-\alpha/2} \frac{\bar{R}}{d_1 \sqrt{n}} \quad (5)$$

## 3. CUSUM 控制图的构造

### 3.1. CUSUM 控制图的理论基础: 序贯概率比检验

CUSUM 控制图通过观测值与目标值之差的累积和来画图, 充分利用了整个观测样本的信息, 其理论基础就是序贯分析原理中的序贯概率比检验(Sequential Probability Ratio Test, 简称 SPRT), 这是一种基本的序贯检验法[4]。

对一个简单零假设和备择假设, 假定我们有一列样本量为  $n$  的独立观测值  $\{X_i\}$ , 而且假定与每个假设分别联系在一起的概率密度函数是:  $f_0(x), f_1(x)$ , 记此时的似然比统计量为

$$\lambda_n = \frac{\prod_{i=1}^n f_1(x_i)}{\prod_{i=1}^n f_0(x_i)} \quad (6)$$

对于事先给定的两个常数  $k, l$ , 我们逐一抽取样本, 如在抽得的样本  $X_1, X_2, \dots, X_{n-1}$  抽样尚不能完成, 则继续抽  $X_n$ , 并计算  $\lambda_n$ 。

当  $\lambda_n \geq l$  时, 停止抽样, 拒绝零假设; 当  $\lambda_n \leq k$  时, 停止抽样, 拒绝备择假设。

当  $k \leq \lambda_n \leq l$  时, 则抽取第  $n+1$  个样本, 并计算  $\lambda_{n+1}$ , 重复以上的步骤, 依次进行。我们称这样的检验就为序贯概率比检验, 简记为  $s(k, l)$ 。

对于一个 SPRT  $s(k, l)$  而言, 其临界值与犯第一、二类错误概率  $\alpha, \beta$  之间的关系为

$$\alpha \leq \frac{1-\beta}{l} \quad (7)$$

$$\beta \leq k(1-\alpha) \quad (8)$$

由以上式子可以看出, 对于事先给定的第一二类错误概率  $\alpha, \beta$ , 我们可以将常数  $k, l$  近似选定为  $k = \frac{\beta}{1-\alpha}, l = \frac{1-\beta}{\alpha}$ 。

### 3.2. CUSUM 控制图参数的选择与平均运行长度

我们以检测向上漂移为例进行说明, 即已知漂移量  $\delta > 0$ , 则基于似然比统计量为  $C_n = \max \sum_i^n \left( x_i - \frac{\delta}{2} \right)$ , 若事先选定阈值  $h > 0$ , 如果  $C_i \leq h, i = 1, 2, \dots, n$  说明到时刻  $n$  为止的过程中表明过程未进入失控状态, 那么继续检测如果直到某个  $C_i > h$  则报警, 认为过程在时刻  $n$  以前有某一个点发生了漂移。在实际应用中,  $\delta$  是未知的, 因此我们常常用一个待定的参数  $k$  代替从而可得到一般的 CUSUM 统计量  $C_n = \max \{0, C_{n-1} + x_n - k\}$  这样 CUSUM 就由两个参数  $(k, h)$  所决定。

参数的选择的关键原则是用平均运行长度  $ARL$ , 一般情况是保证  $ARL_0$  满足条件的情况下选择使  $ARL_1$  尽可能小的  $k$  和  $h$ , 这里

$ARL_0 = E(T^* | \text{过程均值为} 0)$ : 表示过程开始时处于受控状态。

$ARL_1 = E(T^* | \text{过程均值为} \delta)$ : 表示过程开始时处于失控状态。

其中  $T^* = \min \{n : n \geq 1, C_n > h\}$  表示报警所需的运行时间。

我们用上述  $ARL_1$  定义失控状态下的平均运行长度是为了研究的简化, 事实上 CUSUM 具有无后效性即是如果某个  $C_n = 0$ , 则从  $n+1$  开始与从 1 开始的过程在检测方面没有区别。

我们考察向上漂移的检测问题, 以  $L_H(u)$  表示受控状态下让  $C_n$  从  $C_0 = \mu$  开始进行检测所需的平均运行长度, 即

$$\begin{aligned} C_n &= \max \{0, C_{n-1} + x_n - k\} \\ C_0 &= u \\ T(u) &= \min \{n, C_n > h\} \\ L_H(u) &= E(T(u)) \end{aligned} \quad (9)$$

其中样本  $x_1, x_2, \dots, x_n$  为来自标准正态分布的观测值, 濮晓龙[5]运用 Page 于 1954 年导出的积分方程:

$$L_H(u) = 1 + L_H(0)P\{X \leq k - u\} + \int_0^h L_H(x)f_X(x+k-u)dx \quad (10)$$

其中  $X$  与  $X_1, X_2, \dots, X_n$  同分布,  $f_X(\cdot)$  表示  $X$  的密度函数, 那么  $L_H(0)$  就是对向上漂移进行检测的平均运行长度  $ARL_0$ 。

对向下漂移的检测有类似的积分方程:

$$L_L(u) = 1 + L_L(0)P(X \geq -k - u) + \int_0^h L_L(x)f_X(x-k-u)dx \quad (11)$$

对双向检验的问题, Kemp 证明了

$$\frac{1}{ARL_0} = \frac{1}{L_H(0)} + \frac{1}{L_L(0)} \quad (12)$$

如果观测样本并非来自标准正态分布, 而是来自一列发生了漂移量为  $\delta$  的样本, 那么此时的  $ARL_1$  不仅依赖于  $k, h$  还与  $\delta$  有关系。

## 4. 实际比较分析

### 4.1. 判断控制图性能的标准

控制图从检测开始到它发生产出现问题的警号为止所抽取的平均样本数，称为平均链长  $ARL$ 。在实际研究中，我们固定过程受控时的  $ARL$ ，记为  $ARL_0$ ，然后比较不同控制图在过程失控时的  $ARL$  记为  $ARL_1$ ，过程失控时  $ARL_1$  越小，表示该控制图检测漂移越有效。当过程受控时我们希望平均运行长度尽可能的大，这样控制图误报的几率会小一些；当过程失控时，我们希望平均运行长度尽可能的小，这样控制图检测到异常点的概率就会大一些。所以，判别控制图好坏的标准就相当于假设检验中的第一二类错误概率。一般地我们取， $ARL_0 = \frac{1}{\alpha}$   $ARL_1 = \frac{1}{1-\beta}$ 。

### 4.2. 过程受控时两类控制图的平均运行长度：

对于 Shewhart  $\bar{X}$  控制图，当  $u_{1-\alpha/2}$  取为 3 时，根据  $3\sigma$  原则知道，标准正态分布样本落在 3 以外的概率约为 0.27%，即  $\alpha = 0.27\%$ ，所以，过程受控时其平均运行长度为

$$ARL_0 = \frac{1}{0.0027} \approx 370 \quad (13)$$

利用计算机随机产生一组服从标准正态分布，样本容量为 64 的随机数，计算其均值  $\bar{X}_i$ ，判断  $\bar{X}_i$  是否在  $\left[-\frac{3}{\sqrt{n}}, \frac{3}{\sqrt{n}}\right]$  范围内，即  $|\bar{X}_i|$  是否大于 0.375。若  $\bar{X}_i$  落在上下控制线之间，则说明该过程受控，否则继续抽样，计算均值，再判断，直到过程  $\bar{X}_i$  落在上下线之外，记下此时抽样的次数，即为该控制图的运行长度。因一次实验存在误差，故将上述实验重复 10000 次，可得到 Shewhart  $\bar{X}$  控制图的平均运行长度。对上述实验运行 5 次，分别得到过程受控时的平均运行长度  $ARL_0$  (代码见附录)如下表 1。

由以上运行结果可知生产过程受控时 Shewhart 控制图的平均运行长度约为 370。

对于 CUSUM 控制图，计算平均运行长度，关键是确定参数  $k$  和  $h$ 。我们以上单侧为例进行说明，根据参数  $k$  和  $h$  的选择原则，我们一般选择  $k$  为  $\frac{\delta}{2}$ ，然后不断调整控制线  $h$  直到 CUSUM 控制图过程受控时的平均运行长度约为 370 为止。首先随机产生一组服从标准正态分布，样本容量为 1 的随机数，然后计算出其 CUSUM 统计量  $C_1$ ，判断是否超出控制线。如果超出事先设定的上控制线，则说明该过程失控，其运行长度即为 1；如果在均值和上控制线之间，说明该过程受控，则继续产生一个随机数，再计算出其统计量判断是否在均值和上控制线之间，如果超出事先设定的上控制线，则停止，记下此时取样的次数，即为该控制图的运行长度；由于误差的存在，我们通过重复实验次数来减少误差，运行 10,000 次即可得到 CUSUM 控制图的平均运行长度。

通过不断调整控制限  $h$ ，保证  $ARL_0$  大致为 370，发现当  $h$  大致为 2.5。

### 4.3. 过程失控时二者平均运行长度的对比分析

(1) 当漂移量较小时： $\delta = 0.5$  时，对于过程失控时的 Shewhart  $\bar{X}$  控制图的平均运行长度结果见下表 2：

对于过程失控时的 CUSUM 控制图平均运行长度，操作步骤基本同过程受控时，此时我们假设漂移量  $\delta = 0.5$ ，濮晓龙[5]指出如果控制线  $h = 6$  时平均运行长度为 21 左右，通过计算机模拟也证实了这点，结果见下表 3。

由上述两张表的结果我们可以得出结论当偏移量 0.5 时，CUSUM 控制图的平均运行长度远小于 Shewhart  $\bar{X}$  控制图的平均运行长度，由于两种控制图过程受控时的平均运行长度  $ARL_0$  一样，所以失控时的平均运行长度  $ARL_1$  越小越好，从表中数据则足以说明 CUSUM 控制图检测中小漂移更加灵敏。

(2) 当漂移量较大时， $\delta = 2$  时，类似的通过计算机模拟得到如下结果：过程失控时 CUSUM 控制图

**Table 1. Average run length of Shewhart control chart during process control**  
**表 1. 过程受控时 Shewhart 控制图的平均运行长度**

序号	1	2	3	4	5
ARL0	369.47	367.08	370.53	373.51	374.57

**Table 2. Average run length of Shewhart control chart out of process control**  
**表 2. 过程失控时 Shewhart 控制图的平均运行长度**

序号	1	2	3	4	5
ARL1	155.2458	156.2054	153.1586	153.2689	154.2716

**Table 3. Average run length of CUSUM control chart out of process control**  
**表 3. 过程失控时 CUSUM 控制图的平均运行长度**

序号	1	2	3	4	5
ARL1	21.9066	21.8475	21.906	21.8391	22.0378

平均运行长度为 4，过程失控时的 Shewhart  $\bar{X}$  控制图的平均运行长度为 1。由平均运行长度我们可以看出对于漂移量较大的过程检测，Shewhart  $\bar{X}$  控制图的平均运行长度较小，说明 Shewhart  $\bar{X}$  控制图对检测大漂移更加灵敏有效。

## 5. 结论

Shewhart  $\bar{X}$  控制图对过程均值的监控是以最后一次观测结果为依据的，它对检测小漂移不灵敏。而 CUSUM 控制图对过程的监控是以所有观测结果为依据的，它对检测小漂移比较有效。而在实际的质量控制中，如果我们联合使用累积和控制图和休哈特控制图，那么就能弥补各自的不足，达到真正的对企业质量的控制和提升企业的经济效益。

## 参考文献 (References)

- [1] 周纪芾, 茆诗松. 质量管理统计方法[M]. 北京: 中国统计出版社, 1999.
- [2] 张公绪, 孙静. 常规控制图标准及其应用[M]. 北京: 中国标准出版社, 2000.
- [3] 王兆军, 主编. 统计质量控制图理论和方法[M]. 天津: 南开大学出版社, 2011.
- [4] 王敏华, 张维铭. 累积和控制图用的方法和应用[J]. 中国质量, 2000(5): 39-41.
- [5] 濮晓龙. 关于累积和检验的改进[J]. 应用数学报, 2003, 26(2): 225-241.

## 附 录

1.Shewhart 控制图过程受控时的平均运行长度

```
l=[];
fori=1:5
    for m=1:10000 %循环 1 万次
        k=0;
        n=1;
    while k==0
        x=randn(1,64); %产生 64 个服从正态分布的随机数
        y=mean(x);
        if abs(y)>0.375 %判断样本均值是否在控制线以内
            k=n;
        end
        n=n+1;
    end
    l(1,m)=k;
end
ARL=mean(l); %平均运行长度
r(1,i)=ARL;
end
r
```

2.Shewhart 控制图过程受控时的平均运行长度

```
l=[];%控制图每次运行的长度
for i=1:5%运行 5 次
    m=1;
    while m<=10000%循环 1 万次
        k=0;
        n=1;
    while k==0
        x=randn(1,25)+2;%产生 25 个服从正态分布的随机数
        y=mean(x);%计算 25 个样本的均值
        if abs(y)>0.6%判断样本均值是否在控制线以内
            k=n;
        end
        n=n+1;
    end
    l(1,m)=k;
    m=m+1;
end
ARL=mean(l);%平均运行长度
```

```
s(1,i)=ARL;
end
s%运行 5 次，每次的平均运行长度
3.CUSUM 控制图失控时的平均运行长度
l=[]; %控制图每次运行的长度
k=0.25;%参数值
n=1;
    c=[];%S0
for i=1:5%运行 5 次
    m=1;
while m<=20000%循环 2 万次
    p=0;
    n=1;
while p==0
c(1,1)=0;
x(1,n+1)=randn(1,1)+0.5;
    c(1,n+1)=max(0,c(1,n)+x(1,n)-k)
if c(1,n+1)>6
p=n;
end
n=n+1;
end
l(1,m)=p;
    m=m+1;
end
ARL=mean(l);%平均运行长度
s(1,i)=ARL;
end
s
```

**期刊投稿者将享受如下服务：**

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击：<http://www.hanspub.org/Submission.aspx>