

无先验信息的小批量生产质量控制图动态优化

孙 群¹, 赵 颖¹, 潘全科²

(1. 聊城大学 汽车与交通工程学院, 山东 聊城 252059; 2. 聊城大学 计算机学院, 山东 聊城 252059)

摘 要: 针对日益发展的小批量生产模式的质量控制, 在样本数据无先验信息条件下, 提出了一种基于随机加权法的质量控制图参数优化方法. 采用随机加权法对样本数据进行再抽样, 估计样本数据的分布参数(样本均值和方差), 经假设检验, 动态优化质量控制图的控制参数弥补了小样本条件下传统控制图的不足, 实现了无先验信息下小批量生产过程的质量控制. 实验表明, 优化后的控制图的控制参数更加逼近于理论值, 优于传统控制图的控制参数, 适合用于小批量生产环境下的质量控制.

关键词: 小批量; 控制图; 随机加权方法; 质量控制

中图分类号: TH 165

文献标志码: A

文章编号: 0254-0037(2011)08-1148-07

质量控制是产品在制造阶段的质量保证, 其主要方法之一便是控制图. 控制图是按照经典的统计理论建立起来的, 要求质量特征数据为大样本(至少 20 组、样本容量大于 5)时才能有效^[1]. 然而, 目前的生产方式已经由单品种大规模生产向多品种小批量生产转变. 在小批量生产环境下, 采用传统的控制图方法, 导致发生第一类错误和第二类错误的概率增加, 无法满足小批量生产的质量过程控制需求. 对于小批量生产下的柔性制造方式, 国内外学者提出了许多实施小批量生产质量过程控制的方法, 如 Bayes 方法、概率变换方法等^[2-7]. 其中, Bayes 方法是受关注和易于接受的一种方法. Bayes 方法利用历史信息作为先验信息, 融合样本数据, 适合用于小批量问题^[6-7]. 然而, 对于无先验信息下的小批量质量控制问题却研究较少. F. S. Hiller 等提出的小批量 $\bar{X}$ - R 控制图、 $\bar{X}$ - S 控制图及 C. P. Quesenberry 提出的小批量 Q 图诚然是解决该问题的途径^[8-9], 但是由于它们避免了样本数据的参数估计, 因此它们对异常的检出率远远低于样本数据分布参数已知的控制图.

针对这一问题, 本文以多数生产过程服从的正态分布为例, 提出一种无先验信息下的小批量控制图建立方法, 通过自助(Bootstrap)方法和随机加权法充分利用样本数据的自身特征, 估计样本数据的分布参数, 使之逼近于理论值, 从而优化控制图的控制参数.

1 质量控制图参数优化思想

1.1 样本数据预处理

假设质量特征样本数据有 m 组, 每组样本容量为 n , 其计量特性服从正态分布, 可以表示为

$$X^{(0)} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{pmatrix} \quad (i = 1, 2, \cdots, m; j = 1, 2, \cdots, n) \quad (1)$$

收稿日期: 2009-03-26.

基金项目: 国家自然科学基金资助项目(60874075).

作者简介: 孙 群(1979—), 男, 山东淄博人, 博士, 副教授.

1) 计算组内算术平均值

$$\bar{x}_i = \frac{1}{n}(x_{i1} + x_{i2} + \cdots + x_{ij} + \cdots + x_{in}) \quad (2)$$

式中, $\bar{x}_i$ 为第 i 组平均值; n 为组内样本容量; x_{ij} 为第 i 组数据中的第 j 个数据.

2) 计算组内极差

$$r_i = (x_{ij\max} - x_{ij\min}) \Big|_{j=1}^n \quad (3)$$

式中 r_i 为第 i 组极差.

3) 计算组内标准差

$$s_i = \left(\frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2 \right)^{\frac{1}{2}} \quad (4)$$

式中 s_i 为第 i 组组内标准差.

经过以上的数据预处理,得到新样本: $X^{(1)} = \{\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_i, \cdots, \bar{x}_m\}$, $\bar{x}_i \sim N(\mu, \sigma^2/n)$, $i = 1, 2, \cdots, m$, μ 为均值, σ^2 为方差.

1.2 质量控制图参数优化思想

传统控制图理论认为,当生产过程不存在系统误差时,样本数据服从均值为 μ 、方差为 σ^2 的正态分布 $N(\mu, \sigma^2)$, 样本落在区间 $(\mu - 3\sigma, \mu + 3\sigma)$ 中的概率为 0.997 3, 据此确定控制图的控制参数. 通常用样本均值 $\bar{X}$ 和样本标准差 S 分别代替未知参数 μ 和 σ^2 , 当小样本数据时, $\bar{X}$ 与 μ 、 S 与 σ^2 之间的差异过大, 导致控制参数的偏移过大. 因此, 应该采用合适的方法估计样本数据的分布参数 μ 和 σ^2 , 进而优化控制图的控制参数.

2 质量控制图参数优化方法

斯坦福大学的 B. Efron 教授提出 Bootstrap 方法, 利用自身样本数据通过再抽样进行统计推断, 在此基础上又发展出随机加权法, 为小样本数据且无先验信息的质量控制问题提供了一条有益的途径.

2.1 Bootstrap 方法的基本步骤

Bootstrap 方法估计样本数据分布参数的基本步骤如下所述^[10-11].

1) 由 F_n 产生 N 组自助估计样本, $\{X^*(1), X^*(2), \cdots, X^*(j), \cdots, X^*(N)\}$ ($j = 1, 2, \cdots, N$).

2) 计算参数 θ 估计误差的自助统计量 $R_n(j)$ ($j = 1, 2, \cdots, N$).

记均值 μ 的估计误差为 $T_n^{(1)}(j) = \bar{X}(j) - \mu$, 方差 σ^2 的估计误差为 $T_n^{(2)}(j) = \frac{n}{n-1}S^2(j) - \sigma^2$, 得到自助估计统计量

$$\begin{cases} R_n^{(1)*}(j) = \bar{X}^*(j) - \bar{X} \\ R_n^{(2)*}(j) = \frac{n}{n-1}[S^{*2}(j) - S^2] \end{cases} \quad (5)$$

式中, $\bar{X}^*(j) = \frac{1}{n} \sum_{i=1}^n X_i^*(j)$; $S^{*2}(j) = \frac{n}{n-1} \sum_{i=1}^n [X_i^*(j) - \bar{X}^*(j)]^2$.

$X_i^*(j)$ ($i = 1, 2, \cdots, n$) 为取自 F_n 的再生样本. $R_n^{(1)*}$ 和 $R_n^{(2)*}$ 具有如下的统计特性: $E[R_n^{(1)*}] = E[T_n^{(1)}] = 0$; $E[R_n^{(2)*}] = E[T_n^{(2)}] = 0$.

3) 以 $P_n(j)$ ($j = 1, 2, \cdots, N$) 作为 T_n 的估计, 于是得到关于 θ 的一组估计值, 并以此确定 θ 的分布密度. 得到均值 μ 和方差 σ^2 的估计

$$\hat{\mu} = \frac{1}{N} \sum_{j=1}^N [\bar{X} - R_n^{(1)*}(j)] = \bar{X} - \bar{R}_n^{(1)} \quad (6)$$

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{j=1}^N [S^2 - R_n^{(2)*}(j)] = S^2 - \bar{R}_n^{(2)} \quad (7)$$

N 可以任意大, 则 $\bar{R}_n^{(k)}$ ($k=1, 2$) 为 $E[R_n]$ 的无偏一致估计.

2.2 随机加权方法的基本步骤

假设 $X = (X_1, X_2, \dots, X_n)$ 为来自分布 $N(\mu, \sigma^2)$ 的独立同分布样本, 采用随机加权法对未知均值 μ 和方差 σ^2 进行估计的步骤如下^[10-11].

1) 产生 N 组具有 Dirichlet 分布 $D(1, 1, \dots, 1)$ 的随机向量序列 $V(1), V(2), \dots, V(N)$.

每一组 $V(j) = (V_1(j), V_2(j), \dots, V_n(j))$ ($j=1, 2, \dots, N$) 的生成过程为: 设 $v_1, v_2, \dots, v_{n-1}$ 是 $[0, 1]$ 上均匀分布的随机变量 v 的独立同分布子样, 按照由小到大的次序重新排序, 记为: $v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(n-1)}$. 令 $v_{(0)} = 0, v_{(n)} = 1$, 那么 $V_i(j) = v_{(i)} - v_{(i-1)}$ ($i=1, 2, \dots, n$) 的联合分布就是所需要的 $D_n(1, 1, \dots, 1)$ 随机向量.

2) 计算参数 θ 估计误差的随机加权统计量 $D_n(j)$ ($j=1, 2, \dots, N$).

记均值 μ 的估计误差为 $T_n^{(1)}(j) = \bar{X}(j) - \mu$, 方差 σ^2 的估计误差为 $T_n^{(2)}(j) = S^2(j) - \sigma^2$, 得到随机加权统计量

$$\begin{cases} D_n^{(1)}(j) = \sum_{i=1}^n V_i(j) X_i - \bar{X} \\ D_n^{(2)}(j) = \frac{n}{n-1} \sum_{i=1}^n V_i(j) (X_i - \bar{X})^2 - S^2 \end{cases} \quad (8)$$

$D_n^{(1)}$ 和 $D_n^{(2)}$ 有如下的统计特性: $E[D_n^{(1)}] = E[T_n^{(1)}] = 0$; $E[D_n^{(2)}] = E[T_n^{(2)}] = 0$.

3) 以 $D_n(j)$ 作为误差的 T_n 估计, 得到 θ 的一组估计, 并以此确定 θ 的分布密度.

得到均值 μ 和方差 σ^2 的估计

$$\hat{\mu} = \frac{1}{N} \sum_{j=1}^N [\bar{X} - D_n^{(1)}(j)] = \bar{X} - \bar{D}_n^{(1)} \quad (9)$$

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{j=1}^N [S^2 - D_n^{(2)}(j)] = S^2 - \bar{D}_n^{(2)} \quad (10)$$

N 可以任意大, 则 $\bar{D}_n^{(k)}$ ($k=1, 2$) 为 $E[D_n]$ 的无偏一致估计.

2.3 小样本数据中 2 种方法的比较

定理 1 假设分布 F 具有有限 4 阶矩, 其峰度为

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 \geq 0.5 \quad (11)$$

式中 μ_2 和 μ_4 分别为 F 的 2 阶矩和 4 阶矩. 那么, 对于 $n \geq 1$, 有

$$E(\text{var } D_n - \text{var } T_n)^2 \leq E(\text{var } R_n - \text{var } T_n)^2 \quad (12)$$

定理 1 说明了 Bootstrap 方法和随机加权方法模拟效果的差别, 即以模拟变量方差接近于被模拟变量方差的程度作为模拟效果好坏的一个标志^[12]. 可以验证在小样本相同样本容量下, 随机加权法优于 Bootstrap 方法^[12].

3 质量控制图参数的动态确定

3.1 估计值的假设检验^[13-14]

3.1.1 估计均值的假设检验

方差 σ^2 未知时, 在显著性水平 α 下, 对估计均值 $\hat{\mu}$ 进行如下的假设检验:

$$H_0: \mu = \hat{\mu}, H_1: \mu \neq \hat{\mu} \quad (13)$$

由于 S^2 是 σ^2 的无偏估计, 故选取检验统计量

$$t = \frac{\bar{x} - \hat{\mu}}{S/\sqrt{n}} \quad (14)$$

当 $|t|$ 过分大时便拒绝 H_0 , 拒绝域的形式为

$$|t| = \left| \frac{\bar{x} - \hat{\mu}}{S/\sqrt{n}} \right| \geq k \quad (15)$$

当 H_0 成立时, 统计量 $t \sim t(n-1)$. 由 $P\{\text{拒绝 } H_0 | H_0 \text{ 为真}\} = P(\hat{\mu})\{|t| \geq k\} = \alpha$, 可以得到 $k = t_{\alpha/2}(n-1)$, 即得拒绝域为

$$|t| = \left| \frac{\bar{x} - \hat{\mu}}{S/\sqrt{n}} \right| \geq t_{\alpha/2}(n-1) \quad (16)$$

3.1.2 估计方差的假设检验

均值 μ 未知时, 在显著性水平 α 下, 进行估计方差 σ^2 的假设检验

$$H_0: \sigma^2 \leq \hat{\sigma}^2, H_1: \sigma^2 > \hat{\sigma}^2 \quad (17)$$

选择检验统计量

$$\chi^2 = \frac{(n-1)S^2}{\hat{\sigma}^2} \quad (18)$$

查 χ^2 分布表, 得临界值 $\chi_{1-\alpha/2}^2(n-1)$ 和 $\chi_{\alpha/2}^2(n-1)$, 使下面 2 个式子成立:

$$P(\hat{\sigma}^2) \{\chi^2 \leq \chi_{1-\alpha/2}^2(n-1)\} = \frac{\alpha}{2} \quad (19)$$

$$P(\hat{\sigma}^2) \{\chi^2 \geq \chi_{\alpha/2}^2(n-1)\} = \frac{\alpha}{2} \quad (20)$$

则可得拒绝域

$$\chi^2 \leq \chi_{1-\alpha/2}^2(n-1) \text{ 或 } \chi^2 \geq \chi_{\alpha/2}^2(n-1) \quad (21)$$

3.2 控制图参数的动态确定

根据求得样本数据的分布参数, 按照常规控制图的 3σ 方式, 便可得到 $\bar{X}-R$ 图的控制参数.

$\bar{X}$ 图

$$\begin{cases} \text{UCL}(i) = \hat{\mu}(i) + 3\hat{\sigma}(i) \\ \text{CL}(i) = \hat{\mu}(i) \\ \text{LCL}(i) = \hat{\mu}(i) - 3\hat{\sigma}(i) \end{cases} \quad (22)$$

R 图

$$\begin{cases} \text{UCL}(i) = 3.69\hat{\sigma}(i) \\ \text{CL}(i) = 1.13\hat{\sigma}(i) \\ \text{LCL}(i) = 0 \end{cases} \quad (23)$$

随着样本数量 i 的依次获取, 便可以动态优化控制图的控制参数.

4 实验

某手表厂为了提高手表的质量, 应用控制图的方法对装配作业中的螺栓扭矩进行过程控制^[6], 总计获得 25 组数据, 每组数据的样本容量为 5. 为了体现小样本特征, 本文选用前 12 组数据作为研究对象, 如表 1 所示.

1) 传统休哈特控制图

根据样本数据计算出: $\bar{X} = 164.02$, $\bar{R} = 13.33$. 由 $n = 5$, 查表得到: $A_2 = 0.577$, $D_4 = 2.115$, $D_3 = 0$, 得到

控制图的控制参数. 均值控制图($\bar{X}$ 图): $UCL = 171.71$; $CL = 164.02$; $LCL = 156.33$. 极差控制图(R 图): $UCL = 28.19$; $CL = 13.33$; $LCL = 0$.

表 1 螺栓扭矩抽样数据表
Table 1 List of bolt torque sample data

组号	x_{i1}	x_{i2}	x_{i3}	x_{i4}	x_{i5}	均值 $\bar{x}_i$	极差 R
1	154	174	164	166	162	164.0	20
2	166	170	162	166	164	165.6	8
3	168	166	160	162	160	163.2	8
4	168	164	170	164	166	166.4	6
5	153	165	162	165	167	162.4	14
6	164	158	162	172	168	164.8	14
7	167	169	159	175	165	167.0	16
8	158	160	162	164	166	162.0	8
9	156	162	164	152	164	159.6	12
10	174	162	162	156	174	165.6	18
11	168	174	166	160	166	166.8	14
12	148	160	162	164	170	160.8	22

2) 随机加权法改进的控制图

取抽样次数 $N = 10\,000$, 计算样本数据分布参数的估计值: $\hat{\mu} = 163.50$, $\hat{\sigma} = 2.16$, 得到控制图控制参数. 均值控制图($\bar{X}$ 图): $UCL1 = 169.98$; $CL1 = 163.50$; $LCL1 = 157.02$. 极差控制图(R 图): $UCL1 = 7.97$; $CL1 = 2.44$; $LCL1 = 0$.

3) Bootstrap 方法改进的控制图

取抽样次数 $N = 10\,000$, 计算样本数据分布参数的估计值: $\hat{\mu} = 160.80$, $\hat{\sigma} = 2.32$, 得到控制图的控制参数. 均值控制图($\bar{X}$ 图): $UCL2 = 167.76$; $CL2 = 160.80$; $LCL2 = 153.84$. 极差控制图(R 图): $UCL2 = 6.96$; $CL2 = 2.32$; $LCL2 = 0$.

前 12 组样本数据的均值控制图($\bar{X}$ 图)如图 1 所示, 极差控制图(R 图)如图 2 所示.

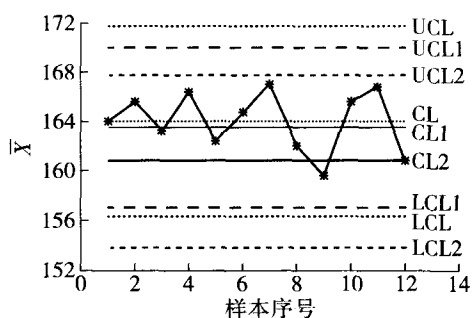


图 1 12 组样本数据的均值控制图

Fig. 1 $\bar{X}$ control chart of the 12 groups of sample data

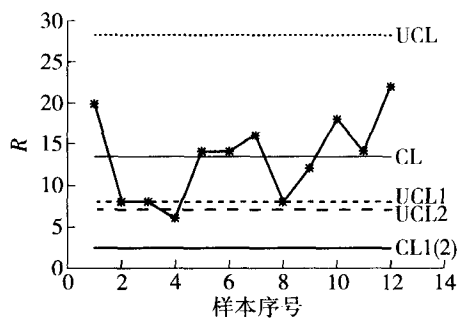


图 2 12 组样本数据的极差控制图

Fig. 2 R control chart of the 12 groups of sample data

实际应用过程中, 一般只采用均值控制图进行质量控制. 为了方便比较, 对全部 25 组数据采用传统控制图方法, 计算出: $\bar{X} = 163.27$, $\bar{R} = 14.28$, 得到均值控制图($\bar{X}$ 图)的控制参数: $UCL = 171.51$; $CL = 163.27$; $LCL = 155.03$. 将采用前 12 组样本数据利用随机加权法改进的控制图和采用 25 组样本数据改进的传统控制图画到一起, 如图 3 所示.

实验分析: 全部 25 组样本数据的均值为 163.27, 可以作为理论参考均值. 12 组样本数据分析中, 采用随机加权法估计的均值为 163.50, 与理论参考均值最为接近, 明显优于传统休哈特控制图的估计均值

164.02. 而 Bootstrap 方法估计的控制图估计均值有所偏离, 验证了定理 1 的正确性. 由于样本数据的极差较大, 因此随机加权法估计的控制图控制参数偏小, 说明生产过程中同组样本具有一定的波动. 图 3 中, 随机加权法的控制图中心线与传统控制图的中心线基本重合, 表明了小样本数据特征下该方法的有效性. 改进后的控制图控制限变窄, 有效地抑制了第二类错误的发生.

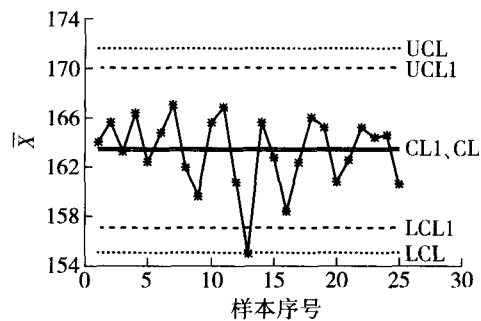


图 3 25 组样本数据的均值控制图

Fig. 3 $\bar{x}$ control chart of the 25 groups of sample data

5 结束语

对于无先验信息的小批量生产的质量控制问题, 利用随机加权法通过对样本数据自身信息的提携, 对样本数据的分布参数进行估计, 从而改进了传统控制图的控制参数. 随着样本数据的不断积累, 控制图参数动态优化, 为小批量生产环境下的质量控制提供了一种有效的途径.

致谢: 本文写作过程中得到了北京工业大学机械工程与应用电子技术学院林家春副教授的指导和帮助, 在此表示诚挚的感谢!

参考文献:

- [1] 杨跃进. 统计过程控制技术[M]. 北京: 航空工业出版社, 2003.
- [2] 苗瑞, 孙小明, 李树刚, 等. 基于小批量生产的统计过程质量控制研究[J]. 计算机集成制造系统, 2005, 11(11): 1633-1635.
MIAO Rui, SUN Xiao-ming, LI Shu-gang, et al. Research on statistical process quality control based on low-volume manufacturing[J]. Computer Integrated Manufacturing Systems, 2005, 11(11): 1633-1635. (in Chinese)
- [3] 杨旭, 马玉林, 杨晓慧. 基于小批量生产的统计质量控制[J]. 计算机集成制造系统, 2001, 7(12): 62-64.
YANG Xu, MA Yu-lin, YANG Xiao-hui. Statistical quality control based on low-volume production[J]. Computer Integrated Manufacturing Systems, 2001, 7(12): 62-64. (in Chinese)
- [4] 杨晓慧, 宋士龙, 谷德义. 面向柔性制造系统的统计质量控制[J]. 系统工程理论与实践, 2003(2): 91-94.
YANG Xiao-hui, SONG Shi-long, GU De-yi. Statistical quality control oriented to flexible manufacturing system[J]. System Engineering Theory and Practice, 2003(2): 91-94. (in Chinese)
- [5] 胡兴才, 叶文华. 小批量生产条件下的统计过程控制研究[J]. 机械研究与应用, 2006, 19(1): 36-38.
HU Xing-cai, YE Wen-hua. Research on the statistical process control in small batch production[J]. Mechanical Research and Application, 2006, 19(1): 36-38. (in Chinese)
- [6] 张公绪. 两种质量诊断理论及其应用[M]. 北京: 科学出版社, 2001.
- [7] 朱慧明, 韩玉启. 基于随机参数的贝叶斯统计质量控制模型[J]. 南京理工大学学报, 2004, 28(4): 445-448.
ZHU Hui-ming, HAN Yu-qi. Bayesian statistical quality control models based on the random parameters[J]. Journal of Nanjing University of Science and Technology, 2004, 28(4): 445-448. (in Chinese)
- [8] HILLIER F S. $\bar{x}$ and R chart control limits based on a small number of subgroups[J]. Journal of Quality Technology, 1969(1): 17-26.
- [9] QUESENBERY C P. SPC $\bar{Q}$ charts for start-up processes and short or long runs [J]. Journal of Quality Technology, 1991, 23(3): 213-224.
- [10] 张金槐. 精度评估中的自助方法[J]. 飞行器测控学报, 2000, 19(3): 61-66.
ZHANG Jin-huai. Bootstrap method of precision assessment[J]. Journal of Spacecraft TT&C Technology, 2000, 19(3): 61-66. (in Chinese)
- [11] 张湘平. 小子样统计推断与融合理论在武器系统评估中的应用研究[D]. 长沙: 国防科技大学机电工程与自动化学院, 2003.
ZHANG Xiang-ping. Researches on the small sample statical inference and fusion theory and its application to the assessment

of weapon system[D]. Changsha: College of Electromechanical Engineering and Automation, National Univeristy of Defence Technology, 2003. (in Chinese)

[12] 郑国忠. 随机加权法[J]. 应用数学学报, 1987, 10(2): 247-253.

ZHENG Guo-zhong. Stochastic weighted method[J]. Acta Mathematicae Applicate Sinica, 1987, 10(2): 247-253. (in Chinese)

[13] 王中宇, 夏新涛, 朱坚民. 非统计原理及其工程应用[M]. 北京: 科学出版社, 2005.

[14] 盛骤, 谢式千, 潘承毅. 概率论与数理统计[M]. 3 版. 北京: 高等教育出版社, 2001.

Dynamic Optimum Method of Control Charts for Small Batch Manufacturing

SUN Qun¹, ZHAO Ying¹, PAN Quan-ke²

(1. School of Automobile and Transportation Engineering, Liaocheng University, Liaocheng 252059, Shandong, China;

2. School of Computer Science, Liaocheng University, Liaocheng 252059, Shandong, China)

Abstract: For quality control of the developed small batch manufacturing mode, a method based on the stochastic weighted theory is proposed to optimize control parameters of control charts. The small sample data is sampled again through the stochastic weighted method to obtain more information. The sample data's distribution parameters are then estimated. After the estimated distribution parameters are tested through the supposed test method, control parameters of control charts can be dynamically optimized. The method improves traditional control charts under low-volume production and realizes the quality control of small batch of manufacturing. Experiments show that these optimized control parameters approach theory true values and are better than that of the traditional control charts. Therefore, the proposed method is suitable for the quality control of small batch of manufacturing.

Key words: small batch manufacturing; control charts; stochastic weighted method; quality control

(责任编辑 吕小红)