Hans汉斯

带有误分类修正的DEWMA p控制图

卿 晶,王凯明,胡小红,王佳颖,宋学力*

长安大学理学院,陕西 西安

收稿日期: 2024年7月13日; 录用日期: 2024年8月7日; 发布日期: 2024年8月14日

摘要

在质量过程控制中,不合格品率小漂移的检测会受到样本的误分类以及小样本量的影响而产生严重的质量误判。本文首先针对样本的分类误差建立误分类修正数据模型实现样本观测值误分类修正,再结合 DEWMA p控制图对参数小漂移的敏感性和对样本量的宽容特性,开发了带有误分类修正的DEWMA p控制图(MisC-DEWMA p)。实验结果表明,在对不合格品率的检测,特别是基于小样本的相应检测中, MisC-DEWMA p控制图对不合格品率小而持久的漂移具有更高的敏感性和精确性。

关键词

误分类,小样本, DEWMA p控制图, 平均运行长度

DEWMA p Control Chart with Misclassification Correction

Jing Qing, Kaiming Wang, Xiaohong Hu, Jiaying Wang, Xueli Song*

School of Science, Changan University, Xi'an Shaanxi

Received: Jul. 13th, 2024; accepted: Aug. 7th, 2024; published: Aug. 14th, 2024

Abstract

In the quality process control, the monitoring of potential small shift at nonconforming proportion will suffer from misclassification and the small-size of sampled observations and thus present kind of a misjudge. In this paper, a data-modified model is firstly established to correct the sample data carrying misclassification, and then DEWMA p Control Chart with Misclassification Correction (MisC-DEWMA p) is proposed, equipped with the strong small-shift-sensitivity and small-samplekindness of DEWMA p control chart. The numerical and real data experiment results show that, for

*通讯作者。

a process with small persistent shift of nonconforming proportion, MisC-DEWMA p control chart provides a higher sensitivity and accuracy.

Keywords

Misclassification, Small Sample, DEWMA p Control Chart, Average Run Length

Copyright © 2024 by author(s) and Hans Publishers Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). http://creativecommons.org/licenses/by/4.0/

CC O Open Access

1. 引言

在对产品质量的统计过程控制(Statistical Process Control, SPC)中,由于仪器误差和采样误差[1] [2]等的存在,过程指标样本数据的测量误差(Measurement Error)是不可避免的。特别地,过程指标变量是类别变量时相应的测量误差被称为误分类(Misclassification)(对类别变量的测量带有测量误差[3])。相比于普通的测量误差,误分类的影响几乎存在于所有定性质量控制问题中,如医疗领域对疾病的误诊会引发健康灾难,金融风险误分类会引发信用危机等。

控制图是实现 SPC 的关键技术,属性控制图[4]-[8]则是针对类别变量设计的定性质量监测控制图, 所以误分类是设计属性控制图需要解决的首要问题。文献[9] [10]研究了误分类对于属性控制图监控策略 的影响,但是模型修正精度依赖于样本容量,且对于过程的微小变化并不敏感。众所周知,SPC 第 II 阶 段监控面临两个重要问题:(1) 采样代价限制,导致小样本常常是必须的,而小样本相对成本低廉但携带 信息有限,所以对模型的要求严苛。(2) 过程变化微小,所以过程参数可能的漂移小而持久是监测任务的 主要特点。因此,开发基于小样本数据的带有误分类修正的 p 控制图,制定精准监控策略是一个极具挑 战性的问题。

带有误分类修正的 EWMA p 控制图是小样本意义下的参数小尺度漂移的监控策略[11],但是由于 EWMA 对大/小偏移的敏感性依赖于平滑参数 λ 的选取,很难选取合适的 λ 既保证监测效率又兼顾到监 测灵敏度。Zhang [12]通过对历史与当前样本信息的两次指数加权平均构造 DEWMA p 图统计量,降低了 图统计量对于平滑参数取值(也就是参数漂移)的依赖性。然而,该图是基于精确测量的假设前提下开发的,并未考虑样本数据的误分类以及因误分类带来的负面影响,致使实际控制结果失真。

基于此,针对过程质量检测中观测值存在误分类的问题,本文首先对误分类样本数据建模,修正样本数据,再结合 DEWMA 对参数小漂移的敏感性以及 DEWMA p 图对样本量的宽容特性,构建带有误分 类修正的 DEWMA p (DEWMA p with Misclassification Correction, MisC-DEWMA p)控制图,以实现对过 程参数小漂移的更精准检测,最后通过模拟仿真和实际案例来验证所提控制图的有效性和优越性。

本文结构安排如下:第2节,基于误分类数据的 DEWMA p (Mis-DEWMA p)控制图建模;第3节,基于误分类数据修正的 DEWMA p (MisC-DEWMA p)控制图建模;第4节控制图性能分析;第5节应用 实例。

2. 带有误分类数据建模的 DEWMA p 控制图及其设计

本节首先利用 Buonaccorsi [13]的方法进行误分类数据建模,再设计带有误分类的 DEWMA p 统计量 以及控制限,构建 Mis-DEWMA p 控制图。

在某质量检测的 SPC 第II阶段,以二值随机变量 *X* 为检测指标, *X* = 0 (表示产品合格)或 1 (表示产品 不合格), $X_i = (X_{1t}, X_{2t}, \dots, X_{nt})'$ 为 *X* 在 *t* 时刻的简单样本,样本容量为 *n* (*n* ≥ 1, *i* = 1, 2, …, *n*, *t* = 1, 2, …*T*)。 $p_j \triangleq p(X_{it} = 1)$ 表示过程受控(in control, IC, *j* = 0)或者过程失控(out of control, OC, *j* = 1)时过程中不合格 品的比率, $q_j \triangleq 1 - p_j$ 为合格品比率, $\hat{p}_{jt} \triangleq \frac{1}{n} \sum_{i=1}^{n} X_{it}$ 为*t* 时刻样本的不合格品率的无偏估计。

2.1. 误分类数据建模

设 $X_{i_t}^*$ 为 X_{i_t} 的观测值,根据文献[13],记

$$\pi_{kl} = p\left(X_{it}^* = k \left| X_{it} = l \right),$$
(1)

其中k,l=0,1。显然, π_{00} 和 π_{11} 为正确分类概率, π_{10} 和 π_{01} 为误分类概率, $\pi_{11} + \pi_{01} = 1$, $\pi_{00} + \pi_{10} = 1$ 。 误分类存在时, π_{10} 和 π_{01} 不全为零。

令 $p_j^* = p(X_{ii}^* = 1)$, $q_j^* = 1 - p_j^* = p(X_{ii}^* = 0)$, p_j 它们分别为质量检测过程中和 q_j 的观测值, 由全 概率公式可得

$$p_j^* = \pi_{11} p_j + \pi_{10} q_j \tag{2}$$

$$q_j^* = \pi_{01} p_j + \pi_{00} q_j \tag{3}$$

整理式(2)和(3)得到

$$\begin{pmatrix} p_j^* \\ q_j^* \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{10} \\ \pi_{01} & \pi_{00} \end{pmatrix} \begin{pmatrix} p_j \\ q_j \end{pmatrix} = \prod \begin{pmatrix} p_j \\ q_j \end{pmatrix},$$
(4)

其中, $\Pi = \begin{pmatrix} \pi_{11} & \pi_{10} \\ \pi_{01} & \pi_{00} \end{pmatrix}$ 为误分类矩阵。

由公式(4)可知,如果观测数据中存在误分类(即 π_{10} 或 $\pi_{01} \neq 0$),那么 p_j 和 q_j 不等于其观测值 p_j^* 和 q_j^* 。 且 π_{10} 或 π_{01} 越大时, p_j , q_j 与 p_j^* , q_j^* 的差异就越大。此外,带有误分类的不合格品率无偏估计为 $\hat{p}_{j,t}^* \triangleq \frac{1}{n} \sum_{i,t}^n X_{it}^*$, $\hat{q}_{j,t}^* = 1 - \hat{p}_{j,t}^*$ 。

2.2. 带有误分类数据建模的 DEWMA p 控制图

2.2.1. DEWMA p 控制图

根据文献[12] [14]可知二值型过程指标变量的 DEWMA p 图统计量如下:

$$\begin{cases} y_{j,t} = \lambda_1 \hat{p}_{j,t} + \lambda_2 y_{j,t-1}, t \ge 1 \\ y_{j,0} = p_j \\ z_{j,t} = \lambda_1 y_{j,t} + \lambda_2 z_{j,t-1}, t \ge 1, \\ z_{j,0} = p_j \end{cases}$$
(5)

其中 $t=1,2,\dots,\infty$, j=0,1; p_j 是不合格品的比率, t时刻样本的不合格品率 p_j 的无偏估计为 $\hat{p}_{j,t}$, 平滑 参数 $\lambda_1 \in (0,1]$, $\lambda_2 = 1 - \lambda_1$ 。图统计量 $z_{i,t}$ 的均值和方差分别为

$$E(z_{j,t}) = p_j, (6)$$

$$Var(z_{j,t}) = \frac{p_j (1 - p_j) \lambda_1^4 \left[1 + \lambda_2^2 - (t + 1)^2 \lambda_2^{2t} + (2t^2 + 2t - 1) \lambda_2^{2t+2} - t^2 \lambda_2^{2t+4} \right]}{n (1 - \lambda_2^2)^3}.$$
(7)

DOI: 10.12677/aam.2024.138358

如果过程处于 OC 状态,不合格品比率会向上漂移,高于 IC 状态,因此可将控制图的下控制限(LCL) 设定为 0,上控制限(UCL)设为

$$UCL = p_{j} + \rho_{\sqrt{\frac{p_{j} \left(1 - p_{j}\right) \lambda_{1}^{4} \left[1 + \lambda_{2}^{2} - \left(t + 1\right)^{2} \lambda_{2}^{2t} + \left(2t^{2} + 2t - 1\right) \lambda_{2}^{2t+2} - t^{2} \lambda_{2}^{2t+4}\right]}{n \left(1 - \lambda_{2}^{2}\right)^{3}},$$
(8)

其中, ρ表示控制限系数。

本文称上述监控策略为 DEWMA p 控制图, 当 $z_{j,t}$ 超出控制上限 UCL 时, DEWMA p 控制图发出失 控信号。

2.2.2. 基于误分类数据建模的 DEWMA p 控制图

当质量过程的样本数据带有误分类时,记带有误分类的样本为 X^{*},在公式(5)~(8)所示的 DEWMA p 控制策略中加入误分类的影响,构造带有误分类数据建模的 DEWMA p 统计量如下:

$$\begin{cases} y_{j,t}^{*} = \lambda_{1} \hat{p}_{j,t}^{*} + \lambda_{2} y_{j,t-1}^{*}, t \ge 1 \\ y_{j,0}^{*} = p_{j}^{*} \\ z_{j,t}^{*} = \lambda_{1} y_{j,t}^{*} + \lambda_{2} z_{j,t-1}^{*}, t \ge 1, \\ z_{j,0}^{*} = p_{j}^{*} \end{cases}$$
(9)

其中 $t=1,2,\dots,\infty$, j=0,1; p_j^* 是误分类样本不合格品的比率, $p_j^* \oplus t$ 时刻的无偏估计为 $\hat{p}_{j,t}^*$, 平滑参数 $\lambda_t \in (0,1]$, $\lambda_2 = 1 - \lambda_1$ 。此时, 图统计量 $z_{i,t}^*$ 的均值和方差分别为

$$E(z_{j,t}^{*}) = p_{j}^{*},$$
 (10)

$$Var\left(z_{j,t}^{*}\right) = \frac{p_{j}^{*}\left(1-p_{j}^{*}\right)\lambda_{1}^{4}\left[1+\lambda_{2}^{2}-\left(t+1\right)^{2}\lambda_{2}^{2t}+\left(2t^{2}+2t-1\right)\lambda_{2}^{2t+2}-t^{2}\lambda_{2}^{2t+4}\right]}{n\left(1-\lambda_{2}^{2}\right)^{3}}.$$
(11)

对应的上下控制限为

$$\text{UCL}^{*} = p_{j}^{*} + \rho^{*} \sqrt{\frac{p_{j}^{*} \left(1 - p_{j}^{*}\right) \lambda_{1}^{4} \left[1 + \lambda_{2}^{2} - \left(t + 1\right)^{2} \lambda_{2}^{2t} + \left(2t^{2} + 2t - 1\right) \lambda_{2}^{2t+2} - t^{2} \lambda_{2}^{2t+4}\right]}{n \left(1 - \lambda_{2}^{2}\right)^{3}},$$
 (12)

$$LCL^* = 0.$$
 (13)

其中, ρ *为控制限系数,本文称此图为带有误分类数据建模的(Mis-DEWMA p)控制图,当 $z_{j,t}^*$ 超出控制 上限 UCL^{*}时,Mis-DEWMA p 控制图发出失控信号。

通过比较控制图 Mis-DEWMA p 与 DEWMA p 的控制限(计算方法见公式(12)与(8))可知,如果生产过程存在误分类,那么在不合格品率 $p < \frac{1}{2}$ 时, $UCL^* > UCL$,这正是原本应处于 OC 状态的样本可能被错判为处于 IC 状态的原因,这里对不合格品率 $p < \frac{1}{2}$ 的要求显然是合理的。所以有必要对误分类样本数据按照误分类概率进行修正,再作为控制图的输入,判断过程是否处于受控状态。

3. 带有误分类修正的 DEWMA p 控制图及其设计

本节首先根据文献[12]中的方法,利用误分类矩阵∏ (公式(4))修正样本数据,进而开发带有误分类 修正的 DEWMA p 控制图。

3.1. 误分类修正模型

3.1.1. 误分类矩阵的估计

在公式(2)和(3)中,误分类矩阵 Π 的元素涉及到过程指标变量 X,因此通常是未知的。现有文献中 Π 的元素一般可利用辅助信息估计得到,其中一种常见的辅助信息是外部验证数据[15][16]。这里我们利用外部验证数据 $r_i \ r_o$,也就是误分类的相对比率来确定误分类概率,估计 Π 中各元素的取值。

当过程指标变量 $X_{ii}=1$ 时,误分类相对比率为

$$r_1 = \frac{\pi_{11}}{1 - \pi_{11}},\tag{14}$$

显然 $r_1 \in [0,\infty)$, r_1 越大, 说明产品被正确分类的概率越大。在给定 r_1 后, 可得 $\pi_{11} = \frac{r_1}{1+r_1}$, $\pi_{01} = \frac{1}{1+r_1}$ 。 同理, 当 $X_{ii} = 0$ 时, $r_0 = \frac{\pi_{00}}{1-\pi_{00}}$, $r_0 \in [0,\infty)$, 给定 r_0 , 可得 $\pi_{00} = \frac{r_0}{1+r_0}$, $\pi_{10} = \frac{1}{1+r_0}$ 。所以, 可得到误分 类矩阵的估计, 为方便起见仍记为 [],

$$\Pi = \begin{pmatrix} \frac{r_1}{1+r_1} & \frac{1}{1+r_0} \\ \frac{1}{1+r_1} & \frac{r_0}{1+r_0} \end{pmatrix}.$$
(15)

到此,我们利用误分类概率量化了误分类,并借助外部验证数据计算了误分类矩阵Ⅱ,接下来将应 用误分类矩阵进行误分类数据修正[13]。

3.1.2. 误分类修正模型

显然∏与其估计值(公式(15))都是可逆矩阵,

$$\begin{pmatrix} \pi_{11} & \pi_{10} \\ \pi_{01} & \pi_{00} \end{pmatrix}^{-1} = \frac{1}{\pi_{11}\pi_{00} - \pi_{10}\pi_{01}} \begin{pmatrix} \pi_{00} & -\pi_{10} \\ -\pi_{01} & \pi_{11} \end{pmatrix},$$
(16)

所以公式(4)等价于

$$\begin{pmatrix} p_{j} \\ q_{j} \end{pmatrix} = \begin{pmatrix} \pi_{11} & \pi_{10} \\ \pi_{01} & \pi_{00} \end{pmatrix}^{-1} \begin{pmatrix} p_{j}^{*} \\ q_{j}^{*} \end{pmatrix},$$
(17)

计算可得

$$p_j = \frac{p_j^* - \pi_{10}}{1 - \pi_{10} - \pi_{01}},\tag{18}$$

其中 j = 0,1,式(18)给出了不合格品率 p_i 的"修正"值,记为 p_i^{**} , X_{ii} 对应的修正观察值为

 $X_{it}^{**} = \frac{X_{it}^{*} - \pi_{10}}{1 - \pi_{10} - \pi_{01}} \circ \text{ is } p_{j}^{**} \triangleq t \text{ by in π in$

3.2. 带有误分类修正的 DEWMA p 控制图

3.2.1. 带有误分类修正的 EWMA p 控制图的介绍

2022 年 Chen 和 Yang (2022) [11]提出了带有误分类修正的 EWMA p 控制图,用 $\hat{p}_{j,t}^{**}$ 修正 $\hat{p}_{j,t}$, 对应的 EWMA p 修正统计量为

$$EWMA_{i,t}^{**} = \lambda \hat{p}_{i,t}^{**} + (1 - \lambda) EWMA_{i,t-1}^{**},$$
(19)

其中, $t=1,2,...,\infty$, j=0,1; 平滑参数 $\lambda \in (0,1]$ 。图统计量 EWMA^{**}_{j,i} 的初始值 EWMA^{**}_{j,0} = p_j^{**} , 上下控 制限分别记为 UCL'和 LCL',

UCL' =
$$p_j^{**} + \rho_{\sqrt{\frac{p_j^* (1 - p_j^*) \lambda \left[1 - (1 - \lambda)^{2t}\right]}{n(1 - \pi_{10} - \pi_{01})^2 (2 - \lambda)}}},$$
 (20)

$$LCL' = 0,$$
 (21)

其中, ρ 表示控制限系数。我们把监控策略(19~21)称为 MisC-EWMA p 控制图,当 EWMA^{**}_{j,t} 超出控制上限 UCL'时,MisC-EWMA p 控制图发出失控信号。

如前所述,现有的 DEWMA p 控制图具有小飘移敏感性和小样本的宽容特性,但会受到样本数据误 分类的影响;而 MisC-EWMA p 控制图修正了误分类,但由于 EWMA 控制图的共有特性对于过程的微小 变化没有 DEWMA 敏感。因此,当样本数据带有误分类时,针对不合格品向上的微小漂移的监测问题, 受文献的[11] [12]启发,我们将结合 DEWMA p 与 MisC-EWMA p 控制图的优点开发带有误分类修正的 DEWMA p 控制图。

3.2.2. 带有误分类修正的 DEWMA p 控制图

根据公式(9)(12)(13)(18),构建基于误分类修正的 DEWMAp 统计量为

$$\begin{cases} y_{j,t}^{**} = \lambda_1 \hat{p}_{j,t}^{**} + \lambda_2 y_{j,t-1}^{*}, t \ge 1 \\ y_{j,0}^{**} = p_j^{**} \\ z_{j,t}^{**} = \lambda_1 y_{j,t}^{**} + \lambda_2 z_{j,t-1}^{**}, t \ge 1 \\ z_{j,0}^{**} = p_j^{**} \end{cases}$$
(22)

其中, $t=1,2,...,\infty$, j=0,1; 平滑参数 $\lambda_1 \in (0,1]$, $\lambda_2 = 1-\lambda_1$ 。图统计量 $z_{i,t}^*$ 的均值和方差分别为

$$E(z_{j,t}^{**}) = p_j^{**},$$
 (23)

$$Var\left(z_{j,t}^{**}\right) = \frac{p_{j}^{*}\left(1-p_{j}^{*}\right)\lambda_{1}^{4}\left[1+\lambda_{2}^{2}-\left(t+1\right)^{2}\lambda_{2}^{2t}+\left(2t^{2}+2t-1\right)\lambda_{2}^{2t+2}-t^{2}\lambda_{2}^{2t+4}\right]}{n\left(1-\pi_{10}-\pi_{01}\right)^{2}\left(1-\lambda_{2}^{2}\right)^{3}}.$$
(24)

上下控制限分别为

$$\text{UCL}^{**} = p_{j}^{**} + \rho^{**} \sqrt{\frac{p_{j}^{*} \left(1 - p_{j}^{*}\right) \lambda_{1}^{4} \left[1 + \lambda_{2}^{2} - \left(t + 1\right)^{2} \lambda_{2}^{2t} + \left(2t^{2} + 2t - 1\right) \lambda_{2}^{2t+2} - t^{2} \lambda_{2}^{2t+4}\right]}{n \left(1 - \pi_{10} - \pi_{01}\right)^{2} \left(1 - \lambda_{2}^{2}\right)^{3}},$$
(25)

$$LCL^{**} = 0,$$
 (26)

其中, ρ^* 表示控制限系数 ρ 的修正值。

本文称公式(22) (25) (26)为 MisC-DEWMA p 控制图模型,当 *z*^{**}_{*j*,*i*} 超出控制上限 UCL^{**} 时, MisC-EWMA p 控制图发出失控信号。

4. 控制图性能分析

平均运行长度(ARL)和运行长度标准差(SDRL)是控制图性能的主要评价指标,它们分别表示检测到 控制图发出报警信号所需采集的样本数量的均值和方差[17]。一个有效的控制图应当在 IC 状态下具有较 大的 ARL (ARL₀),以降低虚假警报的发生频率;相反,在 OC 状态下应当具有较小的 ARL (ARL₁),以 便及时检测到任何异常情况。为了评估 MisC-DEWMA p 控制图的 OC 检测性能,我们选取 DEWMA p 控制图, Mis-DEWMA p 控制图以及 MisC-EWMA p 控制图作为比较。

假设检测时间 T = 5000, 样本大小 n = 5,10, IC 阶段不合格品率 $p_0 = 0.12(0.04)0.4$, OC 阶段不合格品率为 $p_1 \triangleq p_0 + \delta p_0$, 其中 $\delta = 0.05$ 。此外,对于公式(4)所示的误分类模型, 令 $\pi_{00} = \pi_{11} = \pi$, $\pi_{10} = \pi_{01} = 1 - \pi$,其中 $\pi = 0.95$ 。

对于 Mis-DEWMA p 控制图,用公式(2)计算出 p_j^* ,并利用 p_j^* 生成带有误分类的过程指标变量 X^* , OC 状态中带有误分类的 ARL 和 SDRL 分别用 ARL₁^{*}和 SDRL^{*}表示;对于 MisC-DEWMA p 控制图,用公 式(19)计算修正 IC 和 OC 修正的概率 p_0^{**} 和 p_1^{**} ,用 ARL₁^{**}和 SDRL^{**}分别表示 OC 状态中修正的 ARL 和 SDRL;此外,为了便于进行区分和比较,对于 MisC-EWMA p 控制图,用 ρ' 表示其控制限系数 ρ ,用 ARL₁['] 和 SDRL' 分别表示其 OC 阶段的 ARL 和 SDRL。下面我们给定 $ARL_0 = 370$, $\lambda_1 = \lambda = 0.1, 0.2$,利用 Monte Carlo 模拟 10,000 次,详细的模拟结果对比见表 1~6。

Table 1. When ARL₀ = 370 and n = 5, different p_0 values correspond to the control limit coefficient ρ 表 1. 当 ARL₀ = 370, n = 5时,不同 p_0 值对应的控制限系数 ρ

p_0		$\lambda =$	0.1		$\lambda = 0.2$				
	ρ	$ ho^*$	$ ho^{**}$	ho'	ρ	$ ho^{*}$	$ ho^{**}$	ho'	
0.12	5.137	5.110	4.105	2.134	5.804	5.730	4.643	2.428	
0.16	5.119	5.093	4.282	2.201	5.724	5.653	4.776	2.446	
0.20	5.109	5.111	4.368	2.213	5.661	5.623	4.834	2.442	
0.24	5.050	5.036	4.411	2.225	5.586	5.576	4.876	2.435	
0.28	5.037	5.035	4.427	2.213	5.570	5.567	4.900	2.431	
0.32	5.029	5.036	4.464	2.218	5.554	5.510	4.923	2.416	
0.36	5.046	5.053	4.502	2.212	5.493	5.472	4.903	2.395	
0.40	4.981	5.000	4.476	2.204	5.455	5.445	4.898	2.363	

Table 2. When $ARL_0 = 370$ and n = 10, different p_0 values correspond to the control limit coefficient ρ 表 2. 当 $ARL_0 = 370$, n = 10 时,不同 p_0 值对应的控制限系数 ρ

		$\lambda =$	0.1			$\lambda = 0.2$			
P_0	ρ	$ ho^*$	$ ho^{**}$	ho'	ρ	$ ho^{*}$	$ ho^{**}$	ho'	
0.12	5.088	5.100	4.070	2.096	5.702	5.627	4.559	2.326	
0.16	5.092	5.048	4.260	2.147	5.620	5.591	4.700	2.373	
0.20	5.047	5.057	4.333	2.171	5.594	5.557	4.774	2.387	
0.24	5.042	5.027	4.386	2.190	5.542	5.545	4.828	2.397	
0.28	5.030	5.033	4.419	2.196	5.521	5.502	4.856	2.390	

续表								
0.32	5.006	4.993	4.440	2.196	5.479	5.481	4.874	2.392
0.36	4.996	5.005	4.458	2.202	5.458	5.451	4.873	2.376
0.40	5.014	4.974	4.498	2.190	5.450	5.439	4.891	2.369

如表 1、表 2 所示, Mis-DEWMA p 图的控制限系数 ρ^* 总大于 MisC-DEWMA p 图的控制限系数 ρ^{**} 。 此外,对于上述四种控制图,给定 ARL₀后, ρ^{**} 和 λ 具有相同的递减规律。

Table 3. When $ARL_0 = 370$ and n = 5, different p_0 values correspond to the control limit coefficient UCL **表 3.** 当 $ARL_0 = 370$, n = 5 时,不同 p_0 值对应的控制限系数 UCL

n		$\lambda =$	0.1			$\lambda = 0.2$				
p_0	UCL	UCL^*	UCL**	UCL'	UCL	UCL^*	UCL ^{***}	UCL'		
0.12	0.166	0.210	0.166	0.209	0.203	0.250	0.203	0.267		
0.16	0.212	0.250	0.212	0.259	0.252	0.292	0.252	0.320		
0.20	0.257	0.290	0.257	0.306	0.299	0.334	0.299	0.370		
0.24	0.300	0.328	0.300	0.352	0.345	0.374	0.345	0.418		
0.28	0.343	0.366	0.343	0.396	0.390	0.414	0.390	0.465		
0.32	0.385	0.404	0.385	0.440	0.434	0.452	0.434	0.509		
0.36	0.427	0.442	0.427	0.482	0.476	0.490	0.476	0.552		
0.40	0.468	0.478	0.468	0.524	0.517	0.528	0.518	0.593		

Table 4. When $ARL_0 = 370$ and n = 10, different p_0 values correspond to the control limit coefficient UCL **表 4.** 当 $ARL_0 = 370$, n = 10 时, 不同 p_0 值对应的控制限系数 UCL

n		$\lambda =$	0.1			$\lambda = 0.2$			
P_0	UCL	UCL*	UCL**	UCL'	UCL	UCL*	UCL**	UCL'	
0.12	0.152	0.194	0.152	0.182	0.178	0.222	0.177	0.219	
0.16	0.197	0.233	0.197	0.228	0.224	0.263	0.224	0.270	
0.20	0.240	0.272	0.240	0.274	0.269	0.303	0.269	0.318	
0.24	0.282	0.310	0.282	0.318	0.314	0.342	0.314	0.364	
0.28	0.324	0.347	0.324	0.361	0.357	0.380	0.357	0.408	
0.32	0.366	0.385	0.366	0.404	0.399	0.419	0.400	0.453	
0.36	0.407	0.422	0.407	0.446	0.441	0.456	0.441	0.495	
0.40	0.448	0.458	0.448	0.487	0.483	0.493	0.483	0.536	

如表 3、表 4 所示,在相同参数条件下,MisC-DEWMA p 图的控制限 UCL** 的值和 UCL 的值几乎相

等,这表明 MisC-DEWMA p 控制图监控策略能够有效地调整控制限; Mis-DEWMA p 图的控制限 UCL^{*} 的 值比 UCL、 UCL^{**} 的值大,这表明当采集到的样本数据中存在误分类时,变量确定的控制范围更广。

		$\lambda =$	= 0.1		$\lambda = 0.2$				
p_1	ARL_1 (SDRL)	ARL_{1}^{*} $(SDRL^{*})$	ARL_{1}^{**} $(SDRL^{**})$	$\begin{array}{c} ARL_{1}'\\ \left(SDRL'\right) \end{array}$	ARL_1 (SDRL)	ARL_{1}^{*} $(SDRL^{*})$	ARL_{1}^{**} $(SDRL^{**})$	$\begin{array}{c} ARL_{1}'\\ \left(SDRL'\right) \end{array}$	
0.126	250.1	279.4	247.3	256.9	261.7	289.2	259.8	270.0	
	(239.6)	(269.5)	(237.6)	(253.8)	(256.1)	(285.6)	(254.7)	(273.5)	
0.168	231.5	259.6	233.7	240.4	246.5	264.9	246.5	254.7	
	(224.8)	(256.0)	(226.3)	(239.4)	(242.2)	(258.7)	(242.2)	(261.8)	
0.210	220.0	244.3	219.8	224.0	233.6	250.8	231.4	243.5	
	(213.5)	(239.4)	(213.3)	(216.8)	(223.8)	(249.7)	(222.0)	(241.9)	
0.252	200.6	215.2	203.5	211.7	215.8	230.2	218.7	233.6	
	(193.9)	210.4	(196.9)	(210.6)	(212.9)	(228.6)	(216.0)	(232.3)	
0.294	188.0	215.2	187.3	197.9	206.0	219.8	205.3	216.4	
	(181.7)	(210.4)	(181.1)	(194.6)	(201.5)	(215.7)	(200.3)	(216.3)	
0.336	174.6	200.5	174.7	185.6	197.8	212.8	197.0	211.1	
	(167.4)	(192.1)	(167.5)	(181.0)	(192.7)	(208.2)	(192.0)	(208.1)	
0.378	165.3	196.9	165.0	174.1	184.1	202.2	183.9	199.9	
	(159.3)	(190.7)	(159.1)	(170.0)	(177.9)	(196.5)	(177.8)	(196.6)	
0.420	152.2	169.2	153.4	165.9	173.6	186.7	174.9	192.3	
	(149.1)	(165.5)	(150.1)	(162.6)	(169.0)	(181.9)	(170.4)	(190.3)	

Table 5. When $ARL_0 = 370$ and n = 5, the ARL_1 and SDRL values under different p_1 values **表 5.** 当 $ARL_0 = 370$, n = 5 时, 不同 p_1 值下的 ARL_1 和 SDRL值

Table 6. When $ARL_0 = 370$ and n = 10, the ARL_1 and SDRL values under different p_1 values **表 6.** 当 $ARL_0 = 370$, n = 10 时, 不同 p_1 值下的 ARL_1 和 SDRL 值

		$\lambda = 0.1$				$\lambda = 0.2$				
p_1	ARL_1 (SDRL)	ARL_{1}^{*} $(SDRL^{*})$	$\begin{array}{c} ARL_{1}^{**} \\ \left(SDRL^{**} \right) \end{array}$	ARL' ₁ (SDRL')	ARL_1 (SDRL)	$\begin{array}{c} ARL_{1}^{*} \\ \left(SDRL^{*} \right) \end{array}$	$\frac{ARL_{1}^{**}}{\left(SDRL^{**}\right)}$	ARL' ₁ (SDRL')		
0.126	210.8	250.3	209.5	221.6	227.7	255.8	225.1	238.8		
	(206.4)	(241.7)	(205.0)	(222.0)	(224.6)	(251.0)	(221.5)	(253.1)		
0.168	190.7	217.1	192.7	201.5	205.9	232.8	208.1	223.0		
	(183.9)	(211.6)	(185.8)	(198.7)	(200.7)	(234.0)	(203.5)	(230.3)		
0.210	170.9	195.1	173.1	182.5	192.8	210.3	190.5	207.2		
	(164.0)	(187.0)	(166.4)	(182.0)	(189.8)	(205.6)	(187.5)	(216.5)		
0.252	157.3	171.5	157.3	170.0	174.4	191.8	175.5	194.9		
	(151.8)	(164.0)	(151.9)	(168.1)	(170.1)	(186.0)	(171.5)	(197.8)		

DOI: 10.12677/aam.2024.138358

续表								
0.294	142.5	158.1	141.6	154.4	163.4	175.9	163.1	181.2
	(137.6)	(152.5)	(136.4)	(152.6)	(159.7)	(173.0)	(159.6)	(191.5)
0.336	130.1	152.9	129.9	144.7	150.1	170.6	151.3	169.3
	(126.2)	(147.1)	(126.2)	(140.5)	(146.7)	(165.8)	(147.8)	(169.5)
0.378	118.6	139.5	118.4	132.2	139.4	154.2	139.3	157.0
	(113.9)	(134.3)	(113.6)	(130.2)	(136.1)	(149.6)	(135.9)	(162.6)
0.420	108.6	121.7	108.9	121.2	130.3	141.9	130.8	146.4
	(104.9)	(116.2)	(105.0)	(119.2)	(126.8)	(135.7)	(127.1)	(144.9)

如表 5、表 6 所示,对于上述四种控制图,增加 p_1 和 n 中任意一个值,都会使 ARL₁和 SDRL 的值减 小;其次,在相同参数条件下,MisC-DEWMA p 图的 ARL^{**}和 SDRL^{**}的值非常接近 ARL₁和 SDRL 的值, Mis-DEWMA p 图的 ARL^{*}₁和 SDRL^{*}的值分别大于 ARL^{**}和 SDRL^{**}、ARL₁和 SDRL 的值,且随着 n 和 λ 的 增加,这种差异会越来越大,说明 Mis-DEWMA p 控制图不适合用来检测 OC 状态;此外,在相同参数 条件下,MisC-EWMA p 图的 ARL[']₁和 SDRL[']的值分别大于 ARL^{**}和 SDRL^{**}的值、小于 ARL^{*}₁和 SDRL^{**}的 值,这说明 MisC-EWMA p 图可以降低误分类的影响,但 MisC-DEWMA p 图检测效果更好。

5. 案例应用

考虑某工厂生产印刷电路板的过程,数据集来源于 Montgomery [2],具体数据可见表 7。表 7 中第 1 至 26 号样本记录了 OC 状态下连续采集的 26 个样本中的不合格品数量; 27 至 46 号样本数据记录了过程 由 OC 转为 IC 状态后的 20 个新样本对应的不合格品数,样本容量 n = 100。因此,可取 OC 状态运行时间 长度为 T = 26, IC 的为 20。又假设样本数据存在误分类(即将合格(不合格)印刷电路板质量特征被错误地记 录为不合格(合格)的),所以由 OC 状态数据(样本号 1-26)计算的不合格品率 $p_0^* = 0.183$ 是带有误分类的。

样本号	不合格品数	样本号	不合格品数	样本号	不合格品数	样本号	不合格品数
1	21	13	16	25	17	37	18
2	24	14	19	26	15	38	21
3	16	15	10	27	16	39	16
4	12	16	17	28	18	40	22
5	15	17	13	29	12	41	19
6	5	18	22	30	15	42	12
7	28	19	18	31	24	43	14
8	20	20	39	32	21	44	9
9	31	21	30	33	28	45	16
10	25	22	24	34	20	46	21
11	20	23	16	35	25		
12	24	24	19	36	19		

 Table 7. The number of defective printed circuit boards before and after repair

 表 7. 维修前和维修后的不合格印刷电路板数

选取平滑参数 $\lambda = 0.2$, ARL₀ = 370, 用公式(4)来表征误分类对不合格品率的影响, 规定 $\pi_{00} = \pi_{11} = \pi$, $\pi_{10} = \pi_{01} = 1 - \pi$, 其中 $\pi = 0.95$, 1 (易知, π 值越小误分类问题越严重, 当 $\pi = 1$ 时, $X_{ii}^* = X_{ii}^{**} = X_{ii}$)。利用 Monte Carlo 方法, 模拟计算出上述三种控制图相应的控制限和控制限系数, 具体结果见表 8。

MisC-DEWMA p MisC-EWMA p Mis-DEWMA p λ ρ^{**} ρ UCL* UCL** ρ' UCL' 0.2 5.464 0.204 4.492 0.167 0.215 2.219

 Table 8. Simulation results of control limit parameters ρ and control limit values UCL

 表 8. 控制限参数 ρ 和控制限 UCL 数值模拟结果

当系统处于 IC 状态时,我们计算三种控制图的图统计量样本观测值,如图 1 所示,三种控制图均作 出系统处于 IC 状态的推断,而且 Mis-DEWMA p 图和 MisC-DEWMA p 图的图统计量观测值呈现相似的 波动模式,说明 MisC-DEWMA p 控制图对 IC 阶段具有鲁棒性,并不影响对 IC 状态的判定。



Figure 1. Comparison of the application of three control charts in the IC stage 图 1. 三种控制图 IC 阶段的应用对比

而当系统处于 OC 阶段时,如图 2 所示,MisC-DEWMA p 控制图在第 21 个观察点处检测到失控信 号,Mis-DEWMA p 控制图在第 22 个观察点处检测到失控信号,MisC-EWMA p 控制图则未发出报警信 号。这说明在 MisC-EWMA p 和 Mis-DEWMA p 监控策略下,误分类会淹没或者延迟失控信号,而 MisC-DEWMA p 控制图监控策略能够有效改善这种不足。





6. 结论

本文主要研究数据存在误分类的小样本过程质量监控问题,首先通过构建误分类修正模型实现误分 类的修正,再结合 DEWMA p 图对过程参数小漂移和小样本量的敏感性,设计了 MisC-DEWMA p 控制 图。模拟仿真和实际应用结果表明: MisC-DEWMA p 控制策略显著降低了误分类对控制图检测性能的负 面影响,提高了对不合格品率小漂移检测的准确性,减少了检测失控信号所需的次品数量,降低了质检 成本,具有重要的应用价值;且相较于 Mis-DEWMA p 控制图和 MisC-EWMA p 控制图, MisC-DEWMA p 控制图具有更高的准确性和灵感度。

另外,本文假定过程参数漂移大小已知,在实际生产过程中,过程参数漂移大小往往是未知的,后续研究工作可以探讨带有误分类修正的自适应控制图策略,以适应实际生产环境的复杂性。

基金项目

陕西省自然科学基础研究计划资助项目(2024JC-ZDXM-23);长安大学中央高校基本科研业务费专项 资金资助项目(310812163504)。

参考文献

- Javaid, A., Noor-Ul-Amin, M. and Hanif, M. (2020) Performance of Max-EWMA Control Chart for Joint Monitoring of Mean and Variance with Measurement Error. *Communications in Statistics-Simulation and Computation*, **52**, 1-26. <u>https://doi.org/10.1080/03610918.2020.1842886</u>
- [2] Lee, M.H., Lee, Z.Y., Teoh, W.L., Tan, V.M., Kong, M. and Chew, X. (2023) The Effect of Measurement Error on the Performance of the S Chart. *International Conference on Mathematical and Statistical Physics, Computational Science, Education and Communication (ICMSCE* 2023), Istanbul, 19 December 2023. <u>https://doi.org/10.1117/12.3011408</u>
- [3] Zhang, Q. and Yi, G.Y. (2022) Generalized Network Structured Models with Mixed Responses Subject to Measurement Error and Misclassification. *Biometrics*, 79, 1073-1088. <u>https://doi.org/10.1111/biom.13623</u>
- [4] Lee Ho, L. and Quinino, R.C. (2013) An Attribute Control Chart for Monitoring the Variability of a Process. International Journal of Production Economics, 145, 263-267. <u>https://doi.org/10.1016/j.ijpe.2013.04.046</u>
- [5] Yang, B., He, Y. and Yin, H. (2020). Research on Data Analysis and Quality Control Based on P Control Chart. 2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Chengdu, 17-19 October 2020, 1098-1102. <u>https://doi.org/10.1109/cisp-bmei51763.2020.9263537</u>
- [6] Aslam, M., Balamurali, S., Periyasamypandian, J. and Khan, N. (2019) Designing of an Attribute Control Chart Based on Modified Multiple Dependent State Sampling Using Accelerated Life Test under Weibull Distribution. *Communications in Statistics-Simulation and Computation*, **50**, 902-916. <u>https://doi.org/10.1080/03610918.2019.1571606</u>
- [7] Arafah, M. (2022) Using the Laney P' Control Chart for Monitoring COVID-19 Cases in Jordan. *Journal of Healthcare Engineering*, 2022, 1-18. <u>https://doi.org/10.1155/2022/6711592</u>
- [8] Silva, L.A.d., Ho, L.L. and Quinino, R.C. (2022) An Approach for Improvements on the Attribute npx Control Chart. *Quality and Reliability Engineering International*, 38, 3342-3356. <u>https://doi.org/10.1002/qre.3121</u>
- Shin, W.S. and Lingayat, S. (1992) Design of Acceptance Sampling Plans under Varying Inspection Error. *IIE Transactions*, 24, 111-120. <u>https://doi.org/10.1080/07408179208964208</u>
- [10] Ho, L.L., Quinino, R.d.C. and Trindade, A.L.G. (2011) An np-Control Chart for Inspection Errors and Repeated Classifications. Quality and Reliability Engineering International, 27, 1087-1093. <u>https://doi.org/10.1002/qre.1197</u>
- [11] Chen, L. and Yang, S. (2022) A New p-Control Chart with Measurement Error Correction. Quality and Reliability Engineering International, 39, 81-98. <u>https://doi.org/10.1002/qre.3219</u>
- [12] Zhang, L.Y. (2002) EWMA Control Charts and Extended EWMA Control Charts. University of Regina.
- [13] Buonaccorsi, J.P. (2010) Measurement Error: Models, Methods, and Applications. CRC Press.
- [14] Shamma, S.E. and Shamma, A.K. (1992) Development and Evaluation of Control Charts Using Double Exponentially Weighted Moving Averages. *International Journal of Quality & Reliability Management*, 9. https://doi.org/10.1108/02656719210018570
- [15] Carroll, R.J., Ruppert, D., Stefanski, L.A., et al. (2006) Measurement Error in Nonlinear Models: A Modern Perspective. Chapman and Hall/CRC.
- [16] Chen, L.P. (2020) Semiparametric Estimation for the Transformation Model with Length-Biased Data and Covariate Measurement Error. *Journal of Statistical Computation and Simulation*, **90**, 420-442. https://doi.org/10.1080/00949655.2019.1687700
- [17] 刘利平, 吴敏鑫, 胡雪龙. 考虑双分量测量误差的 CUSUM 控制图设计[J]. 工业工程与管理, 2016, 21(2): 81-85.