Journal of Statistical Modelling: Theory and Applications
Vol. 4, No. 2, 2023, pp. 119-130
Yazd University Press 2023

JSMTA

Research Paper

A multivariate *p*-chart with a suggestion for measurement error correction

Shohreh Enami[1]*, Hamzeh Torabi[2]
[1]Department of Statistics, Payame Noor University, Tehran, Iran
[2]Department of Statistics, Yazd University, Yazd, Iran

Received: April 02, 2024/ Revised: August 05, 2024/ Accepted: September 02, 2024

Abstract: This study aims to introduce a new multivariate *p*-chart using a structure of the multivariate Bernoulli distribution. In this structure, we consider a class of dependent Bernoulli variables where the conditional success probability is a linear combination of the last few trials and the original success probability. The efficiency of the proposed control chart is investigated in terms of the average run length criterion with two approaches. Also, the issue of misclassification and measurement error in binary random variables is discussed, and another control chart based on the corrected proportion of non-conforming products is proposed.

Keywords: Average run length; Misclassification; Multivariate Bernoulli distribution; Multivariate *p*-chart; Monte Carlo simulation.
Mathematics Subject Classification (2010): 62Hxx; 62P30

1 Introduction

Control charts are important in statistical quality control to monitor and improve production processes. Control charts are divided into attributes and variable control charts, depending on whether the quality characteristic is attributable or measurable. In some processes, quality characteristics cannot be measured numerically. In such cases, each inspected item is usually classified as either conforming or non-conforming to the specifications of that quality characteristic. Quality characteristics of this type are called attributes. The most common attribute control charts are the p-chart and np-chart (for binomially distributed processes) and the c and u control charts (for Poisson

*Corresponding author: enami1387@yahoo.com

distributed processes). For more discussion, one may refer to Aslam et al. (2018), Falahnezhad and Oulia (2010), Chukhrova and Johannssen (2019), David (2002), Kooli and Limam (2011), and Shu and Wu (2010).

In some situations, researchers are interested in the investigation of more than one correlated quality characteristic simultaneously. In this case, the process is called a multivariate process, and it would be necessary to use a multivariate control chart.The use of multivariate control schemes is significantly widespread nowadays. With modern data-acquisition equipment, sensors and online computers, it is very common to monitor several quality characteristics simultaneously. In this paper, we focused on a multivariate attribute control chart. In most of the monitoring processes, the quality characteristics cannot be measured on a continuous scale, such as manufacturing processes from industrial settings, healthcare processes or processes from service industries, and environments of non-manufacturing quality improvement. There are many works performed in the literature in the field of control charts for multivariate attribute processes. For more discussion, one may refer to Cozzucoli and Marozzi (2018), Enami and Torabi (2019), Ali Raza and Aslam (2018), Pascual and Akhundjanov (2020), Makhdoom and Basikhasteh (2022), Li et al. (2019), and Haridy et al. (2014). Also, several control charts are provided for multivariate Binomial distribution with different dependent structures, Li et al. (2014); Topalidou and Psarakis (2009). In this paper, we consider a class of dependent Bernoulli variables that first introduced by Yang and Lixin (2017). In this structure, we consider a class of dependent Bernoulli variables where the conditional success probability is a linear combination of the last few trials and the original success probability. Then, we propose a control chart for this structure.

We should note that to obtain p-chart or Mp-chart (multivariate p-chart), there is an essential concern in using these charts: In the product line of the factory, sometimes products may be falsely detected as confirmed (or non-confirmed) due to imprecise measurement equipment or human-made mistakes. Because of such a misclassification, the observed status is different from what it should be. Chen and Yang (2022) discussed p-chart with measurement error correction. To the best of the authors' knowledge, there is no research on the multivariate p-chart with measurement error correction.

The remainder of this paper is organized as follows: In Section 2, introducing a new statistic D, which is the sum of the counts of nonconforming units of all the quality characteristics in a sample, the control limits for a structure of the multivariate Bernoulli distribution are proposed. In Section 3, the efficiency of the proposed control chart is investigated in terms of the average run length (ARL) criterion with two approaches. In Section 4, the issue of misclassification and measurement error in binary random variables is discussed, and then a control chart based on it is proposed. Finally, some conclusions and suggestions for future research are presented in Section 5.

2 Structure of the proposed control chart based on approximation

In a manufacturing process, assume that there are K quality characteristics. Let t denote the monitoring time. For each time $t = 1, 2, ..., T$, there are n samples $(i = 1, \ldots, n)$. Let vector $\boldsymbol{X}_{it} = (X_{it1}, X_{it2}, ..., X_{itK})$ follows a multivariate Bernoulli distribution, so that X_{itj} is the number of defects or non-conformities with respect to quality characteristic j, $j = 1, 2, ..., K$, where $X_{itj} = 1$ represents a non-conforming product, while $X_{itj} = 0$ demonstrates a conforming product. So

$$X_{itj} \sim ber(p_j), \qquad j = 1, 2, \ldots, K,$$

where p_j is the proportion of the non-conforming products and $\boldsymbol{p} = (p_1, p_2, ..., p_K)$ is a vector of proportion of the non-conforming products.

It is rational to assume that the non-conformity of a product in terms of a quality characteristic, may affect its non-conformity in another quality characteristic. Therefore, let $p_1 = P(X_{it1} = 1)$ be probability of non-conformity of a chosen product in terms of the first quality characteristic. By observing X_{it1}, it is reasonable to consider that probability of non-conformity of that chosen product in terms of the second quality characteristic, is a weighted mean between p_1 and X_{it1}(this weight is indicated by θ). We can assume this pattern up to Kth quality characteristic. Hence, we consider the following simple model:

$$\begin{aligned} p_1 &= P(X_{it1} = 1), \\ p_j &= P(X_{itj} = 1|\mathcal{F}_{j-1}) = \theta X_{it(j-1)} + (1-\theta)p_1, \quad j = 2, 3, \ldots, K, \end{aligned} \tag{1}$$

where $\theta \in [0, 1]$, and $\mathcal{F}_K = \sigma\{X_1, \ldots, X_K\}$ is the σ-field generated by the random variables $X_1, \ldots, X_K$. In other words, the probability of success for each trial conditional on all previous trials is a linear combination of the previous trial and the probability of success of the first trial. Define the statistic D as follows

$$D = \sum_{t=1}^{T} D_t, \tag{2}$$

where $D_t = \sum_{j=1}^{K} \sum_{i=1}^{n} X_{itj}$. In fact, D is the number of all non-conforming products.

Theorem 2.1. *Suppose that $\boldsymbol{X}_{it}$ follows a multivariate Bernoulli distribution with structure defined in* (1) *and D is the proposed statistic in* (2)*. Then,*

$$\frac{D - KTnp_1}{\sqrt{KTn}} \xrightarrow{\text{d}} N\left(0, \frac{p_1(1-p_1)(1+\theta)}{1-\theta}\right),$$

where $\xrightarrow{\text{d}}$ denotes the convergence in distribution.

Proof. According to Corollary 2 of Yang and Lixin (2017), when $K \longrightarrow \infty$, we have

$$\frac{\sum_{j=1}^{K} X_{itj} - Kp_1}{\sqrt{K}} \xrightarrow{\text{d}} N\left(0, \frac{p_1(1-p_1)(1+\theta)}{1-\theta}\right), \quad \forall\ i = 1, \ldots, n,\ \forall\ t = 1, \ldots, T.$$

On the other hand, it is clear that for each $K = 1, 2, ...$, the random variables $\sum_{j=1}^{K} X_{11j}$, $\ldots, \sum_{j=1}^{K} X_{nTj}$ are independent for each n and T. Thus,

$$\frac{D - KTnp_1}{\sqrt{KTn}} = \frac{\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{K} X_{itj} - KTnp_1}{\sqrt{KTn}} \xrightarrow{\text{d}} N\left(0, \frac{p_1(1-p_1)(1+\theta)}{1-\theta}\right).$$

□

In the following, we suppose that when the process is in-control

$$\begin{aligned} p_{10} &= P(X_{it1} = 1), \\ p_{j0} &= P(X_{itj} = 1|\mathcal{F}_{j-1}) = \theta X_{it(j-1)} + (1-\theta)p_{10}, \quad j = 2, 3, \ldots K. \end{aligned} \tag{3}$$

Using the above normal approximation, the traditional Shewhart-type control limits can be written

$$\begin{aligned} \text{LCL} &= \mu_{D_{in-control}} - 3\sigma_{D_{in-control}} = KTnp_{10} - 3\sqrt{KTn\frac{p_{10}(1-p_{10})(1+\theta)}{1-\theta}}, \\ \text{UCL} &= \mu_{D_{in-control}} + 3\sigma_{D_{in-control}} = KTnp_{10} + 3\sqrt{KTn\frac{p_{10}(1-p_{10})(1+\theta)}{1-\theta}}, \end{aligned} \tag{4}$$

where $\mu_{D_{in-control}}$ and $\sigma_{D_{in-control}}$ are the mean and the standard deviation of D, when the process is in-control. Note that, if LCL < 0, then we suppose that LCL $= 0$.

In practice, usually, the proportion of non-conforming products is not known and they must be estimated. The vector of the proportion of the non-conforming products is estimated as follows $\bar{\boldsymbol{p}} = (\bar{p}_1, \bar{p}_2, ..., \bar{p}_K)$, where

$$\begin{aligned} \bar{p}_1 &= \sum_{t=1}^{T}\sum_{i=1}^{n} \frac{X_{it1}}{nT}, \\ \bar{p}_j &= \theta X_{j-1} + (1-\theta)\bar{p}_1, \quad j = 2, 3, \ldots K. \end{aligned}$$

So, in this case, the control limits are obtained as follows

$$\begin{aligned} \text{LCL} &= \mu_{D_{in-control}} - 3\sigma_{D_{in-control}} = KTn\bar{p}_{10} - 3\sqrt{KTn\frac{\bar{p}_{10}(1-\bar{p}_{10})(1+\theta)}{1-\theta}}, \\ \text{UCL} &= \mu_{D_{in-control}} + 3\sigma_{D_{in-control}} = KTn\bar{p}_{10} + 3\sqrt{KTn\frac{\bar{p}_{10}(1-\bar{p}_{10})(1+\theta)}{1-\theta}}, \end{aligned} \tag{5}$$

where $\bar{p}_{10}$ is the estimate of p_1 when the process is in-control.

3 The efficiency evaluation of the chart by the ARL criterion

In this section, the efficiency of the proposed control chart is evaluated using the ARL criterion. There are two types of ARLs namely the in control (ARL_0) and the out-of-control (ARL_1) ARLs Montgomery (2007), which are defined as follows:

$$\text{ARL}_0 = \frac{1}{\alpha}, \qquad \text{ARL}_1 = \frac{1}{1-\beta},$$

where α and β are Type I error and Type II error, respectively.

When the process is in-control (i.e., $p_j = p_{j0}$), but the control chart shows an out-of-control sample, then a Type I error occurs. The probability of Type I error can be obtained as follows:

$$\alpha = P(D < \text{LCL}|p_j = p_{j0}) + P(D > \text{UCL}|p_j = p_{j0}).$$

On the other hand, when the parameters shift from the in-control values to any other undesirable values such as p_{1j}, but the control chart shows a sample inside the control limits, then a Type II error occurs. The probability of the Type II error can be obtained as follows:

$$\beta = P(\text{LCL} < D < \text{UCL}|p_j = p_{j1}).$$

Here, the ARL_1s of the proposed chart are examined for a 5-variable process and a 10-variable process (i.e., $K = 5$ and $K = 10$) when ARL_0 is fixed at 370. We consider different values for p_{10} that are specified as $0.1, 0.4, 0.5, 0.7$. For out-of-control state, we consider $p_{11} = (1 + \delta)p_{10}$, where δ is a shift value.

In this study, two values for the monitoring time (i.e., $T = 10, 30$) are assumed, and four values for the sample size (i.e., $n = 5, 10, 15, 20$). Tables 1-6 present the performance of the proposed control chart in terms of ARL_1 for different shifts (Table 1 and 3, for $\theta = 0.1$, Table 2 and 4, for $\theta = 0.5$ and Table 3 and 6, for $\theta = 0.8$). In these Tables, the values of ARL_1 are calculated in two cases: i. by using Theorem 2.1 and ii. by using Monte Carlo simulation (values in brackets) and by using the bellow algorithm.

Algorithm 3.1. *Given k, T, n, p_{10}, θ and δ,*
(i) Calculate LCL and UCL by (4),
(ii) Generate the random vector $(X_{it1}, X_{it2}, ..., X_{itK})$ *for* $i = 1, \ldots, n$ *and* $t = 1, 2, ..., T$ *from* (1) *with* $p_1 = (1 + \theta)p_{10}$,
(iii) Compute $D = \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{K} X_{itj}$ *based on* (2),
(iv) Repeat steps (ii)- (iii) for M *times to obtain* $D_1, \ldots, D_M$,
(v) Compute $\beta = \frac{1}{M} \sum_{i=1}^{M} I(LCL < D_i < UCL)$,
(vi) Compute $ARL_1 = \frac{1}{1-\beta}$.

These results are obtained using `R` software. The conclusions made based on the results in Tables 1-6 are summarized as follows

i. The obtained results show that for$\theta < 0.5$ (especially for $\delta > 0.03$) both approaches, almost have the same results, even for small values of k.In the other words, for $\theta < 0.5$ we can calculate control limits by using approximately normal for every value of k.

ii. Our results showed that for $\theta > 0.5$, the two approaches have not the same results. Therefore, for $\theta > 0.5$ and small values of k, we recommend to use the Monte Carlo approach.

iii. For the constant values of n, p, T, δ and θ with increasing K the performance of the proposed chart is better. For example, if $\theta = 0.1, \delta = 0.03, n = 5, T = 30, p = 0.1$ and $K = 5$ then $ARL_1 = 252$ and if $\theta = 0.1, \delta = 0.03, n = 5, T = 30, p = 0.1$ and $K = 10$

then $ARL_1 = 203$.
iv-For the constant values of n, p, T, k and δ with decreasing θ the performance of the proposed chart is better. For example, if $K = 5, \delta = 0.03, n = 5, T = 30, p = 0.7$ and $\theta = 0.1$ then $ARL_1 = 35$ and if $K = 5, \delta = 0.03, n = 5, T = 30, p = 0.7$ and $\theta = 0.5$ then $ARL_1 = 99$.
v-In some cases, the out-of-control process may not be detected ($ARL_1 > ARL_0$). These modes have been indicated by the expression "...".

Table 1: The ARL_1 of the proposed control chart, when $K = 5$, $\theta = 0.1$ and $\text{ARL}_0 = 370$.

		(p_{10}, T)							
δ	n	(0.1, 10)	(0.1, 30)	(0.4, 10)	(0.4, 30)	(0.5, 10)	(0.5, 30)	(0.7, 10)	(0.7, 30)
0.03	5	297(....)	252(....)	217(....)	114(....)	187(....)	83(....)	119(....)	35(40)
	10	273(....)	203(....)	152(160)	60(64)	118(120)	39(40)	58(62)	13(13)
	15	252(261)	168(175)	114(120)	38(42)	83(85)	23(25)	35(42)	7(7)
	20	233(245)	143(150)	89(92)	26(28)	62(66)	15(16)	24(28)	4(4)
0.07	5	184(187)	104(106)	65(65)	17(18)	44(48)	10(10)	17(20)	3(3)
	10	134(136)	58(58)	29(31)	6(6)	18(20)	3(3)	5(5)	1(1)
	15	104(106)	38(42)	17(18)	3(3)	10(10)	2(2)	3(3)	1(1)
	20	83(86)	28(29)	11(12)	2(2)	6(7)	1(1)	2(2)	1(1)
0.1	5	121(124)	53(55)	28(28)	6(6)	17(19)	3(3)	5(6)	1(1)
	10	76(77)	25(26)	11(11)	2(2)	6(7)	1(1)	2(2)	1(1)
	15	53(54)	15(16)	6(6)	1(1)	3(3)	1(1)	1(1)	1(1)
	20	40(42)	10(13)	1(3)	2(2)	1(1)	3(2)	1(1)	1(1)
0.15	5	62(60)	9(9)	2(2)	5(5)	1(1)	1(1)	2(1)	1(1)
	10	32(31)	8(8)	3(8)	1(1)	2(2)	1(1)	1(1)	1(1)
	15	20(22)	5(5)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
	20	14(16)	3(3)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
0.2	5	34(35)	9(10)	4(4)	1(1)	2(2)	1(1)	1(1)	1(1)
	10	16(16)	4(4)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
	15	9(10)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
	20	6(7)	1(2)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)

4 Misclassification

In practice, variables are often mismeasured. That is, we may encounter measurement errors in continuous variables or misclassification in discrete variables. To explain more about the concept of misclassification, we suppose that X_{itj} is the true situation of the product, which is unknown and X^*_{itj} is the observed status (or surrogate version of X_{itj}) of the product that is recorded by factory's staff. The relationship between X^*_{itj} and X_{itj} can be characterized as follows

$$\pi_{kl} = P(X^*_{itj} = k | X_{itj} = l), \qquad k, l = 0, 1.$$

This is clear that, π_{01} (or π_{10}) indicates the probability that a non-conforming (or conforming) product is falsely recorded as conforming (or non-conforming). According to Chen and Yi (2021) π_{11} and π_{00}, are called classification probability and π_{01} and π_{10} are called misclassification probability.

Table 2: The ARL_1 of the proposed control chart, when $K = 5$, $\theta = 0.5$ and $ARL_0 = 370$.

		(p_{10}, T)							
δ	n	(0.1, 10)	(0.1, 30)	(0.4, 10)	(0.4, 30)	(0.5, 10)	(0.5, 30)	(0.7, 10)	(0.7, 30)
0.03	5	314(....)	291(....)	283(....)	199(....)	269(....)	163(....)	227(....)	99(....)
	10	302(....)	263(....)	235(....)	133(....)	207(....)	100(....)	142(....)	46(....)
	15	291(....)	239(....)	199(....)	97(....)	163(184)	68(75)	99(143)	27(52)
	20	281(293)	219(224)	171(180)	74(86)	137(142)	50(61)	74(80)	18(24)
0.07	5	231(258)	170(182)	137(140)	53(56)	110(115)	35(36)	62(62)	12(12)
	10	197(200)	119(121)	79(81)	23(24)	56(58)	13(13)	23(23)	4(4)
	15	170(173)	89(92)	52(55)	13(13)	35(35)	7(7)	12(13)	2(2)
	20	149(150)	70(70)	38(38)	8(8)	24(24)	5(5)	7(7)	1(1)
0.1	5	177(185)	108(119)	76(80)	22(25)	55(56)	13(13)	25(25)	4(4)
	10	135(141)	64(67)	36(40)	8(10)	23(25)	4(4)	8(8)	1(1)
	15	108(112)	44(49)	22(27)	4(5)	13(13)	2(2)	4(4)	1(1)
	20	89(93)	32(35)	15(18)	3(3)	8(10)	2(2)	2(2)	1(1)
0.15	5	112(116)	52(54)	31(31)	7(8)	20(24)	4(4)	7(8)	1(1)
	10	73(75)	26(27)	12(12)	2(2)	7(9)	1(1)	2(2)	1(1)
	15	52(53)	16(16)	7(7)	1(1)	4(4)	1(1)	1(1)	1(1)
	20	40(41)	11(11)	4(4)	1(1)	2(2)	1(1)	1(1)	1(1)
0.2	5	71(74)	27(28)	14(15)	3(3)	9(10)	2(2)	2(2)	1(1)
	10	41(43)	12(14)	56(57)	1(1)	2(2)	1(1)	1(1)	1(1)
	15	27(29)	7(8)	3(3)	1(1)	2(2)	1(1)	1(1)	1(1)
	20	20(20)	5(5)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)

Table 3: The ARL_1 of the proposed control chart, when $K = 5$, $\theta = 0.8$ and $ARL_0 = 370$.

		(p_{10}, T)							
δ	n	(0.1, 10)	(0.1, 30)	(0.4, 10)	(0.4, 30)	(0.5, 10)	(0.5, 30)	(0.7, 10)	(0.7, 30)
0.03	5	322(....)	313(....)	327(....)	283(....)	330(....)	268(....)	347(....)	227(....)
	10	317(....)	302(....)	304(....)	234(....)	296(....)	206(....)	276(....)	141(....)
	15	313(....)	291(....)	283(....)	199(....)	268(....)	166(....)	227(....)	99(....)
	20	309(....)	281(....)	265(....)	171(....)	245(....)	173(....)	190(....)	74(....)
0.07	5	260(....)	231(....)	232(....)	136(....)	220(....)	110(....)	196(....)	62(....)
	10	245(....)	197(....)	174(....)	78(....)	149(....)	56(....)	100(....)	23(....)
	15	231(....)	170(....)	136(....)	52(....)	109(....)	35(....)	62(....)	12(86)
	20	218(....)	149(....)	111(....)	38(....)	85(....)	24(....)	42(....)	7(25)
0.1	5	220(365)	177(358)	168(347)	76(312)	150(324)	55(307)	118(232)	25(96)
	10	196(324)	136(309)	107(268)	36(291)	85(279)	23(289)	47(195)	8(25)
	15	177(301)	108(294)	67(259)	22(250)	55(223)	13(90)	25(157)	4(6)
	20	161(270)	89(254)	57(242)	15(101)	39(207)	8(33)	15(121)	2(3)
0.15	5	164(350)	112(349)	95(327)	31(198)	78(307)	20(227)	52(199)	7(20)
	10	133(319)	73(302)	49(298)	12(70)	35(278)	7(22)	15(165)	2(2)
	15	112(283)	52(277)	31(265)	7(21)	20(263)	4(6)	7(20)	1(1)
	20	95(268)	40(257)	21(253)	4(8)	13(99)	2(3)	4(7)	1(1)
0.2	5	121(315)	71(309)	55(269)	14(25)	42(186)	9(37)	24(65)	2(3)
	10	90(221)	41(296)	24(254)	5(8)	16(136)	3(4)	5(14)	1(1)
	15	71(213)	27(273)	14(105)	3(4)	7(35)	2(2)	2(2)	1(1)
	20	58(201)	20(172)	9(37)	2(2)	5(13)	1(1)	2(1)	1(1)

Table 4: The ARL_1 of the proposed control chart, when $K = 10$, $\theta = 0.1$ and $ARL_0 = 370$.

		(p_{10}, T)							
δ	n	$(0.1, 10)$	$(0.1, 30)$	$(0.4, 10)$	$(0.4, 30)$	$(0.5, 10)$	$(0.5, 30)$	$(0.7, 10)$	$(0.7, 30)$
0.03	5	273(232)	203(190)	152(161)	60(58)	118(132)	39(38)	58(59)	12(13)
	10	233(198)	143(130)	89(93)	26(27)	62(68)	15(16)	24(28)	4(4)
	15	203(159)	108(93)	60(58)	15(15)	39(39)	8(8)	13(13)	2(2)
	20	179(147)	85(81)	44(43)	10(10)	27(27)	5(5)	8(8)	2(2)
0.07	5	134(110)	59(58)	29(30)	6(6)	18(19)	3(3)	5(5)	1(1)
	10	63(74)	27(27)	11(11)	2(2)	6(6)	1(1)	2(2)	1(1)
	15	60(53)	16(17)	6(6)	1(1)	3(3)	1(1)	1(1)	1(1)
	20	43(40)	11(11)	4(4)	1(1)	2(2)	1(1)	1(1)	1(1)
0.1	5	76(74)	25(25)	11(11)	2(2)	6(7)	1(1)	2(2)	1(1)
	10	40(36)	10(10)	4(4)	1(1)	2(2)	1(1)	1(1)	1(1)
	15	25(24)	6(5)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
	20	18(17)	4(4)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
0.15	5	32(32)	8(8)	3(3)	1(1)	2(2)	1(1)	1(1)	1(1)
	10	14(14)	3(3)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
	15	8(8)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
	20	5(5)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
0.2	5	16(16)	4(4)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
	10	6(6)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
	15	4(4)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
	20	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)

Suppose, when the process is in-control, $p^*_{j0} = P(X^*_{itj} = 1)$, $j = 1, 2, ..., K$, and $q^*_{j0} = 1\text{-}p^*_{j0}$. Using the technique of the law of total probability, they can be written as

$$\begin{aligned} p^*_{j0} &= \pi_{11} p_{j0} + \pi_{10} q_{j0}, \\ q^*_{j0} &= \pi_{01} p_{j0} + \pi_{00} q_{j0}, \end{aligned} \tag{6}$$

where p_{j0}, $j = 1, 2, ..., K$, is introduced in (3), and $q_{j0} = 1\text{-}p_{j0}$. The matrix form of (6) is given by

$$\begin{bmatrix} p^*_{j0} \\ q^*_{j0} \end{bmatrix} = \Pi \begin{bmatrix} p_{j0} \\ q_{j0} \end{bmatrix}, \tag{7}$$

where $\Pi = \begin{bmatrix} \pi_{11} & \pi_{10} \\ \pi_{01} & \pi_{00} \end{bmatrix}$ is the 2×2 (mis)classification matrix and $\pi_{11} + \pi_{10} = 1$ and $\pi_{01} + \pi_{00} = 1$. Note that the case of no misclassification corresponds to having $\Pi = I$, the identity matrix.

We assume that Π has the spectral decomposition $\Pi = \Omega M \Omega^{-1}$, where M is the diagonal matrix with diagonal elements being the eigenvalues of Π, and Ω is the corresponding matrix of eigenvectors (Chen and Yi, 2021).

Since our main target is to monitor p_{10}, from the equations (6), we observe that when $\pi_{10} \neq 0$ or $\pi_{01} \neq 0$, p^*_{10} and q^*_{10} are different from p_{10} and q_{10}. Therefore, with the availability of X^*_{it1}s, the estimate of sample proportion $\bar{p}^*_{10} = \frac{\sum_{t=1}^{T} \sum_{i=1}^{n} X^*_{it1}}{nT}$ has bias for $\bar{p}_{10}$ and the corresponding control limits, determined by (5) with $\bar{p}^*_{10}$ replaced by $\bar{p}_{10}$ may incur wrong detection. From (7) have

$$\Pi^{-1} \begin{bmatrix} p^*_{j0} \\ q^*_{j0} \end{bmatrix} = \begin{bmatrix} p_{j0} \\ q_{j0} \end{bmatrix}.$$

Table 5: The ARL_1 of the proposed control chart, when $K = 10$, $\theta = 0.5$ and $ARL_0 = 370$.

		(p_{10}, T)							
δ	n	$(0.1, 10)$	$(0.1, 30)$	$(0.4, 10)$	$(0.4, 30)$	$(0.5, 10)$	$(0.5, 30)$	$(0.7, 10)$	$(0.7, 30)$
0.03	5	302(369)	263(363)	235(315)	133(217)	208(335)	100(163)	142(360)	46(72)
	10	281(358)	219(322)	171(256)	74(134)	137(238)	50(70)	74(130)	18(24)
	15	263(333)	186(227)	133(219)	48(70)	100(148)	30(42)	46(75)	9(11)
	20	247(331)	161(208)	107(180)	34(51)	76(119)	21(27)	31(43)	6(6)
0.07	5	170(202)	119(135)	79(124)	23(28)	56(92)	13(16)	23(36)	4(4)
	10	150(160)	70(73)	38(40)	8(9)	24(26)	5(5)	7(8)	2(2)
	15	119(120)	47(49)	23(23)	5(5)	13(13)	3(3)	4(4)	1(1)
	20	98(101)	35(35)	15(15)	3(3)	9(9)	1(1)	2(2)	1(1)
0.1	5	135(137)	65(66)	36(36)	8(8)	23(23)	4(4)	8(8)	1(1)
	10	89(91)	32(32)	15(15)	3(3)	8(8)	2(2)	2(2)	1(1)
	15	64(66)	20(21)	8(8)	2(2)	4(4)	1(1)	1(1)	1(1)
	20	49(50)	13(13)	5(5)	1(1)	3(2)	3(3)	1(1)	1(1)
0.15	5	73(75)	26(27)	12(13)	2(2)	7(7)	1(1)	2(2)	1(1)
	10	40(41)	11(12)	4(4)	1(1)	2(2)	1(1)	1(1)	1(1)
	15	26(28)	6(7)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
	20	18(18)	4(4)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
0.2	5	41(42)	1212()	5(5)	1(1)	3(3)	1(1)	1(1)	1(1)
	10	20(21)	5(5)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)
	15	12(12)	3(3)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)
	20	8(8)	2(2)	1(1)	1(1)	1(1)	1(1)	1(1)	1(1)

Table 6: The ARL_1 of the proposed control chart, when $K = 10$, $\theta = 0.8$ and $ARL_0 = 370$.

		(p_{10}, T)							
δ	n	$(0.1, 10)$	$(0.1, 30)$	$(0.4, 10)$	$(0.4, 30)$	$(0.5, 10)$	$(0.5, 30)$	$(0.7, 10)$	$(0.7, 30)$
0.03	5	317(....)	302(....)	303(....)	235(....)	297(....)	207(....)	276(....)	143(....)
	10	309(....)	281(....)	265(....)	171(....)	245(....)	137(....)	190(....)	74(....)
	15	302(....)	263(....)	234(....)	132(....)	207(....)	100(....)	142(....)	45(241)
	20	295(....)	247(....)	209(....)	106(....)	178(....)	76(....)	110(....)	31(135)
0.07	5	245(....)	196(....)	174(....)	79(....)	149(295)	56(257)	100(131)	23(86)
	10	218(....)	149(....)	111(209)	38(141)	85(136)	24(79)	42(68)	7(13)
	15	197(....)	19(....)	79(196)	23(74)	56(115)	13(30)	23(28)	4(5)
	20	178(....)	98(....)	59(226)	15(36)	40(95)	7(17)	14(21)	2(3)
0.1	5	196(367)	135(348)	107(315)	36(168)	85(264)	23(80)	47(201)	8(14)
	10	161(340)	89(293)	57(307)	14(37)	39(182)	8(15)	15(45)	2(2)
	15	135(299)	64(260)	36(147)	8(14)	23(73)	4(6)	8(13)	1(1)
	20	116(248)	49(227)	25(78)	5(8)	15(39)	3(3)	5(6)	1(1)
0.15	5	133(305)	73(256)	49(89)	12(28)	35(58)	7(12)	15(54)	2(2)
	10	95(298)	40(112)	21(64)	4(6)	13(31)	2(3)	4(5)	1(1)
	15	73(251)	26(77)	12(27)	2(3)	7(12)	1(1)	2(2)	1(1)
	20	58(208)	18(44)	8(15)	2(2)	4(6)	1(1)	1(1)	1(1)
0.2	5	90(259)	41(126)	25(86)	5(8)	16(43)	3(3)	5(9)	1(1)
	10	58(200)	20(52)	9(19)	2(8)	5(8)	1(1)	2(1)	1(1)
	15	41(86)	12(26)	5(7)	1(1)	3(8)	1(1)	1(1)	1(1)
	20	31(92)	8(14)	3(4)	1(1)	2(2)	1(1)	1(1)	1(1)

Consequently

$$p_{10} = \frac{\pi_{00} p_{10}^* - \pi_{10} q_{10}^*}{\pi_{11}\pi_{00} - \pi_{10}\pi_{01}} = \frac{(1-\pi_{01}) p_{10}^* - \pi_{10} q_{10}^*}{(1-\pi_{10})(1-\pi_{01}) - \pi_{10}\pi_{01}}.$$

Suppose for a particular state, $\pi_{00} = \pi_{11} = \pi$ and $\pi_{10} = \pi_{01} = 1 - \pi$, then

$$p_{10} = \frac{p_{10}^{*} - \pi_{10}}{1 - \pi_{10}\pi_{01}}.$$

So, the "corrected" proportion of non-conforming products, denoted as p_{10}^{**}, is defined as follows

$$p_{10}^{**} = \frac{p_{10}^{*} - \pi_{10}}{1 - \pi_{10}\pi_{01}}. \tag{8}$$

Under (4), the corrected LCL and UCL are given by

$$\begin{aligned}
\text{LCL}^{**} &= \mu_{D_{in-control}} - 3\sigma_{D_{in-control}} = KTnp_{10}^{**} - 3\sqrt{KTn\frac{p_{10}^{**}(1-p_{10}^{**})(1+\theta)}{1-\theta}},\\
\text{UCL}^{**} &= \mu_{D_{in-control}} + 3\sigma_{D_{in-control}} = KTnp_{10}^{**} + 3\sqrt{KTn\frac{p_{10}^{**}(1-p_{10}^{**})(1+\theta)}{1-\theta}}.
\end{aligned} \tag{9}$$

From Chen and Yang (2022) we can write the "corrected" random variable

$$X_{it1}^{**} = \frac{X_{it1}^{*} - \pi_{10}}{1 - \pi_{10}\pi_{01}}.$$

Thus, the "corrected" sample proportion is calculated as follows

$$\bar{p}_{10}^{**} = \frac{\sum_{t=1}^{T}\sum_{i=1}^{n} X_{it1}^{**}}{nT}.$$

In this case, the control limits are estimated as follows:

$$\begin{aligned}
\text{LCL}^{**} &= KTn\bar{p}_{10}^{**} - 3\sqrt{KTn\frac{\bar{p}_{10}^{**}(1-\bar{p}_{10}^{**})(1+\theta)}{1-\theta}},\\
\text{UCL}^{**} &= KTn\bar{p}_{10}^{**} + 3\sqrt{KTn\frac{\bar{p}_{10}^{**}(1-\bar{p}_{10}^{**})(1+\theta)}{1-\theta}}.
\end{aligned} \tag{10}$$

5 Conclusions and future research

In the present article, a new control chart for a structure of the multivariate Bernoulli distribution was first introduced (multivariate p-chart), then the efficiency of this chart was evaluated in terms of out-of-control ARL (when ARL_0 remains constant). In applications, measurement error exists due to imprecise operation systems or human-made mistakes, and ignoring measurement error effects may cause the wrong detection. To the best of the authors' knowledge, there is no research on the multivariate p-chart with measurement error correction. In this paper, the issue of misclassification and measurement error in binary random variables was discussed and then, a control chart based on it was proposed. We should note that measurement errors are parameters, and in application, they have to be estimated. We suggest that in future research, these parameters will be estimated, and then will be evaluated the efficiency of the proposed control chart. Also, the correlation between monitoring times can be the next research project in the future.

Acknowledgments

We would like to thank and express our deep gratitude to the editor, anonymous referees, and Dr. Hossin Nadeb for their useful suggestions and comments, which helped us to improve the manuscript greatly.

References

Ali Raza, M. and Aslam, M. (2018). Design of control charts for multivariate Poisson distribution using generalized multiple dependent state sampling. *Quality Technology and Quantitative Management*, **16**(6):629–650.

Aslam, M., Azam, M., Kim., K.J. and Jun, Ch.H. (2018). Designing of an attribute control chart for two-stage process. *Measurement and Control*, **51**(7-8):285–292.

Chen, L.P. and Yi, G.Y. (2021). Analysis of noisy survival data with graphical proportional hazards measurement error models. *Biometrics*, **77**(3):956–969.

Chen, L.P. and Yang, S.F. (2022), A new p-control chart with measurement error correction. *Quality and Reliability Engineering International*, **39**(1):81–98.

Chukhrova, N. and Johannssen, A. (2019). Improved control charts for fraction nonconforming based on hypergeometric distribution. *Computers and Industrial Engineering*, **128**:795–806.

Cozzucoli, P.C. and Marozzi, M. (2018). Monitoring multivariate Poisson processes: A review and some new results. *Quality Technology and Quantitative Management*, **15**(1):53–68.

David, L. (2002). Improved control charts for attributes. *Quality Engineering*, **14**(4):531–537.

Enami, Sh. and Torabi, H. (2019). A repetitive sampling-based control chart for multivariate weighed Poisson distribution with two different indices, *Journal of Statistical Research of Iran*, **16**(1):245–254.

Falahnezhad, Ms. and Oulia, Ms. (2010). A cumulative binomial chart for uni-variate process control, *Scientia Iranica*, **17**(2):111–119.

Haridy, S., Wu, Z., Abhary, K., Castagliola, Ph. and Shamsuzzaman, M. (2014). Development of a multiattribute synthetic-np chart. *Journal of Statistical Computation and Simulation*, **84**(9):1884–1903.

Kooli, I. and Limam, M. (2011). Economic design of an attribute np control chart using a variable sample size.*Sequential Analysis*, **30**(2):145–159.

Li, C., Pan, J. and Huang, M.H. (2019). A new demerit control chart for monitoring the quality of multivariate Poisson processes. *Journal of Applied Statistics*, **46**(4):680–699.

Li, J., Tsung, F. and Zou, C. (2014). Multivariate binomial/multinomial control chart. *IIE Transactions*, **46**(5):526–542.

Makhdoom, I. and Basikhasteh, M. (2022). Location charts based on concomitants of ranked set samples from Morgenstern family. ,*Statistical Computation and Simulation*, **92**(5):911–927.

Montgomery, D.C. (2007). *Introduction to Statistical Quality Control*. United States of America: John Wiley & Sons.

Pascual, F.G. and Akhundjanov, S.B. (2020). Copula-based control charts for monitoring multivariate Poisson processes with application to hepatitis C counts. *Journal of Quality Technology*, **52**(2):128–144.

Shu, M.H. and Wu, H.C. (2010). Monitoring imprecise fraction of nonconforming items using p control charts. *Journal of Applied Statistics*, **37**(8):1283–1297.

Topalidou, E. and Psarakis, S. (2009). Review of multinomial and multiattribute quality control charts. *Quality and Reliability Engineering International*, **25**(7):773–804.

Yang, Z.H and Lixin, Z.H. (2017). Limit theorems for dependent Bernoulli variables with statistical inference. *Communications in Statistics-Theory and Methods*, **46**(4):1551–1559.