Scientia Iranica E (2013) 20(6), 2238–2246

SCIENTIA IRANICA

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
www.scientiairanica.com

Phase II monitoring of binary response profiles

M.E. Soleymanian, M. Khedmati* **and H. Mahlooji**

Department of Industrial Engineering, Sharif University of Technology, Tehran, P.O. Box 11155-9414, Iran.

Received 26 August 2012; received in revised form 12 December 2012; accepted 20 April 2013

KEYWORDS
Profile monitoring;
Binary response;
Hotelling T^2;
Multivariate Exponentially Weighted Moving Average (MEWMA);
Likelihood ratio test.

Abstract. In many situations, the quality of a process can be characterized better by a relationship, known as a profile, between a response variable and one or more predictors. Almost all research efforts assume that response variable is continuous and follows a normal distribution, while there are instances in which the response is a binary variable, and methods such as logistic regression are commonly used. In this paper, four control schemes namely Hotelling T^2, MEWMA, Likelihood Ratio Test (LRT) and LRT/EWMA are proposed to monitor binary response profiles in phase II. The performance of the proposed control charts is evaluated and compared by simulation experiments for different shift values in the parameters of the profile in terms of the Average Run Length (ARL) criterion. The results show that all methods work well in the sense that they can effectively detect shifts in the process parameters. Based on the results, MEWMA and LRT/EWMA methods display a better performance for small to moderate and large shift values, respectively.

1. Introduction

In numerous situations, based on the univariate or multivariate distribution of the quality characteristics of a process, univariate and multivariate control charts are used to monitor the quality of the process. However, in some cases, the quality of a process is characterized by a relationship between a response variable and one or more explanatory variables. Such a relationship is commonly referred to as a '*profile*'. There exist various types of profiles based on whether the relationship assumes the form of a simple linear, multiple linear, polynomial or nonlinear regression. Simple linear regression profiles are mostly used in calibration applications and have been widely studied in the literature for both phases I and II. The purpose of the phase-I analysis is to evaluate the stability of a process and to estimate the process parameters, while in phase-II analysis, one is interested in detecting shifts in the process parameters as quickly as possible. Kang and Albin [1] proposed a multivariate T^2 and an EWMA/R control chart to monitor the parameters of a simple linear profile. Kim et al. [2] coded the values of the predictor variable first, such that the average of the coded values become zero, and then proposed three EWMA control charts to monitor the parameters and error variance of simple linear profiles. Mahmoud and Woodall [3] developed a monitoring scheme and an F test-based on indicator variables and a control chart in line with one of the charts proposed by Kim et al. [2] to monitor the stability of a linear regression profile. Saghaei et al. [4] developed a cumulative sum (CUSUM) control chart, and compared its performance with some existing methods. Other related works in this area include Zhang et al. [5], Mahmoud et al. [6], Soleimani et al. [7] and Hossienifard et al. [8].

In some cases, more complicated models such as multiple linear and polynomial regression are needed

*. *Corresponding author. Tel.: +98 21 66165717;*
Fax: +98 21 66022702
E-mail addresses: Emad.soleimanian@gmail.com (M.E. Soleymanian); majid.khedmati@yahoo.com (M. Khedmati); mahlooji@sharif.edu (H. Mahlooji)

to represent the relationship. Kazemzadeh et al. [9] proposed three methods to monitor a polynomial profile in phase I. Zou et al. [10] proposed a Multivariate Exponentially Weighted Moving Average (MEWMA) control chart for monitoring general linear profiles in phase II. Amiri et al. [11] proposed a dimension reduction method to cope with the dimensionality problem of existing methods in monitoring multiple linear regression profiles, and compared it with other methods (see also [12,13]).

All the references cited above assume that the response variable is continuous, whereas in many applications, the response variable of a profile can be better considered as a discrete variable. One application arises when a process or product should be classified as defective or non-defective, or when the quality of a process would be either acceptable or non-acceptable. For instance, the quality of surface-to-air anti aircraft missiles is inspected by test-firing at targets of varying speed. The outcome for this test will be either a hit or a miss [14]. Hosmer and Lemeshow [15] provided an example in which the relationship between age (predictor) and presence or absence of evidence of significant coronary heart disease (response) for 100 subjects selected to participate in a study is investigated (see [14,15] for more motivating examples). The response in these cases will be a binary (Bernoulli) variable. Despite the fact that there exist many real applications for binary response profiles, only few research efforts have concentrated on this issue. Yeh et al. [16] provided the only research work on phase I binary response profile monitoring in which logistic regression is used to model the relationship between a binary response and one or more continuous explanatory variables. They developed five Hotelling T^2 control charts, and compared them in terms of signal probabilities. In a more recent research work, Shang et al. [17] considered the binary response profile monitoring problem and developed three control charts for phase II monitoring purposes. Considering a random explanatory variable, they compared the proposed control charts based on the Average Run Length (ARL).

In this paper, we assume that the explanatory variable values are fixed and constant from profile to profile, and develop some phase II control chart schemes in order to identify shifts in the parameters of the profile in a quick manner. We use logistic regression to model the relationship. Since Koosha and Amiri [18] compared different link functions including Probit, Logit, Log-Log and Comp log log and concluded that Logit outperforms other functions for both in-control and out-of-control conditions, we use Logit link function in this paper as well. The proposed control charts are compared through simulation experiment studies in terms of ARL criterion.

The rest of this paper is organized as follows: In the next section, the profile model is described; the proposed control chart schemes for phase II monitoring of binary response profiles are presented in Section 3; the performance of the proposed control charts is studied in Section 4 through simulation experiments; an illustrative example is presented in Section 5; and finally, the concluding remarks are presented in Section 6.

2. The profile modeling

The Generalized Linear Model (GLM) is widely used to model profiles with discrete responses. However, in case of a binary response variable, logistic regression is the most common model used for this purpose.

Assuming n independent settings, the p explanatory variables and the corresponding response in each setting i are shown by vector $\mathbf{X}_i = (X_{i1}, X_{i2}, \cdots, X_{ip})^T$ and variable z_i, $i = 1, 2, \cdots, n$, respectively. We assume z_i is a Bernoulli variable with a probability of success π_i for which $E(z_i) = \pi_i$ and $\mathrm{Var}(z_i) = \pi_i(1-\pi_i)$ hold. Denoting the Logit link function by $g(\pi_i)$, we have:

$$g(\pi_i) = \log\left(\frac{\pi_i}{1-\pi_i}\right) = \mathbf{X}_i^T\boldsymbol{\beta}$$
$$= \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip}, \tag{1}$$

in which $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_p)^T$ represents the parameter vector of the model. It is customary to set $X_{i1} = 1$ for β_1 to show the intercept. According to this model, the probability of success will be:

$$\pi_i = \frac{\exp(\mathbf{X}_i^T\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i^T\boldsymbol{\beta})}. \tag{2}$$

We assume there exist m_i independent observations in each setting of the explanatory variables, and as such $M = \sum_{i=1}^n m_i$ denotes the total number of observations. Based on this model, $y_i = \sum_{j=1}^{m_i} z_{ij}$ represents the total number of successes in the ith setting, and consequently follows a binomial distribution with parameters (m_i, π_i) in which z_{ij} represents the jth observation in the ith setting. Now we have $E(y_i) = m_i\pi_i$ and:

$$\mathrm{Var}(y_i) = m_i\pi_i(1-\pi_i)$$
$$= m_i \times \frac{\exp(\mathbf{X}_i^T\boldsymbol{\beta})}{1+\exp(\mathbf{X}_i^T\boldsymbol{\beta})} \times \frac{1}{1+\exp(\mathbf{X}_i^T\boldsymbol{\beta})}. \tag{3}$$

Moreover, the likelihood function for $y_1, y_2, \cdots, y_n$ can be written as:

$$L(\boldsymbol{\pi}, \mathbf{y}) = \prod_{i=1}^{n} \binom{m_i}{y_i} \pi_i^{y_i}(1-\pi_i)^{m_i-y_i}, \tag{4}$$

whose logarithm can be written as follows:

$$\log L(\boldsymbol{\pi},\mathbf{y}) = \sum_{i=1}^{n} \log \binom{m_i}{y_i} + \sum_{i=1}^{n} y_i(\mathbf{X}_i^T\boldsymbol{\beta}) - \sum_{i=1}^{n} m_i \log(1+\exp(\mathbf{X}_i^T\boldsymbol{\beta})), \tag{5}$$

in which $\boldsymbol{\pi} = (\pi_1, \pi_2, \cdots, \pi_n)^T$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)^T$. Taking the partial derivative of the log-likelihood function with respect to $\boldsymbol{\beta}$, we have:

$$\frac{\partial \log L(\boldsymbol{\pi},\mathbf{y})}{\partial \boldsymbol{\beta}} = \sum_{i=1}^{n} y_i\mathbf{X}_i^T - \sum_{i=1}^{n} m_i \frac{\exp(\mathbf{X}_i^T\beta)}{1+\exp(\mathbf{X}_i^T\boldsymbol{\beta})}\mathbf{X}_i^T = \mathbf{X}^T\mathbf{y} - \sum_{i=1}^{n} m_i\pi_i\mathbf{X}_i^T = \mathbf{X}^T(\mathbf{y}-\boldsymbol{\mu}), \tag{6}$$

where:

$$\boldsymbol{\mu} = (\mu_1, \mu_2, \cdots, \mu_n)^T = E(\mathbf{y}) = (m_1\pi_1, m_2\pi_2, \cdots, m_n\pi_n)^T,$$

and:

$$\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n)^T,$$

is an $n \times p$ matrix. The Maximum Likelihood Estimator (MLE) of $\boldsymbol{\beta}$ will be obtained by solving the equation $\mathbf{X}^T(\mathbf{y}-\boldsymbol{\mu}) = \mathbf{0}_p$. In this paper, the Iterative Weighted Least Squares (IWLS) estimation method is used to approximate the MLE of $\boldsymbol{\beta}$, denoted by $\hat{\boldsymbol{\beta}}$ (see [16,19] for more details). According to this model, $\hat{\boldsymbol{\beta}}$ asymptotically follows a p-dimensional Normal distribution, $N_p(\boldsymbol{\beta}, (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1})$, in which:

$$\mathbf{W} = \mathrm{diag}\{m_1\pi_1(1-\pi_1), m_2\pi_2(1-\pi_2), \cdots, m_n\pi_n(1-\pi_n)\}.$$

3. The proposed control charts for phase II

In this section, four control charts are proposed to monitor a binary response profile in phase II, in order to detect shifts in parameters in a quick manner.

3.1. Hotelling T^2 control chart

The Hotelling T^2 control chart is used by Kang and Albin [1] for linear profiles in phase II. Also, Yeh et al. [16] have developed five Hotelling T^2 charts for phase I of binary profiles. Considering the asymptotic normal distribution of the parameters of a binary response profile, the Hotelling T^2 control chart can be modified and used for monitoring a binary profile in phase II. The T^2 statistic in this method for the jth profile is calculated as:

$$T_j^2 = \left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right)^T \boldsymbol{\Sigma}_0^{-1} \left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right), \qquad j = 1, 2, \cdots \tag{7}$$

where $\hat{\boldsymbol{\beta}}_j = (\hat{\beta}_1, \cdots, \hat{\beta}_p)^T$ is the estimated parameters of the logistic regression and the $n \times p$ matrix $\boldsymbol{\beta}_0$ represents the in-control parameters. Moreover, $\boldsymbol{\Sigma}_0$ denotes the in-control variance-covariance matrix of the parameters of the logistic regression model, which is estimated as $\hat{\boldsymbol{\Sigma}}_0 = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}$. Now, the T^2 statistic can be re-written as:

$$T_j^2 = \left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right)^T \mathbf{X}^T\mathbf{W}\mathbf{X} \left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right). \tag{8}$$

As long as a statistic is below the Upper Control Limit (UCL) of a proposed control chart, the process is assumed in control. However , when $T_j^2 >$ UCL, the process will be out of control. We resort to simulation to calculate the UCL of the control chart, yielding the specified in-control ARL.

3.2. The MEWMA control chart

In this section, an MEWMA control chart is proposed to monitor binary profiles. This kind of chart was first developed by Zou et al. [10] for generalized linear profiles. In this method, we first define the following variable:

$$\mathbf{Z}_j = \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2} \left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right), \tag{9}$$

where $\hat{\boldsymbol{\beta}}$ is the vector of estimated logistic regression parameters in the jth profile. It is straightforward to show that the vector $\mathbf{Z}_j$ follows a multivariate normal distribution with mean $\mathbf{0}$ and identity variance when the process is in-control and n is sufficiently large, i.e.:

$$E(\mathbf{Z}_j) = E\left(\left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2}\left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right)\right) = \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2} E\left(\left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right)\right) = 0, \tag{10}$$

$$\begin{aligned}\mathrm{Var}(\mathbf{Z}_j &= \mathrm{Var}\left(\left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2}\left(\hat{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_0\right)\right) \\ &= \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2} \mathrm{Var}\left(\hat{\boldsymbol{\beta}}_j\right) \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2} \\ &= \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2} \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{-1} \left(\mathbf{X}^T\mathbf{W}\mathbf{X}\right)^{1/2} \\ &= \mathbf{I}.\end{aligned} \tag{11}$$

Now, the statistic for the MEWMA chart is proposed as:

$$\mathbf{W}_j = \theta\mathbf{Z}_j + (1-\theta)\mathbf{W}_{j-1}, \qquad j = 1, 2, \cdots \tag{12}$$

where θ is the smoothing parameter and $\mathbf{W}_0 = \mathbf{0}$. In this approach, the control chart generates a signal due to the out-of-control state of profile when:

$$U_j = \mathbf{W}_j^T\mathbf{W}_j > L\frac{\theta}{2-\theta}. \tag{13}$$

The parameter L is calculated by simulation method in such a way that the specified in-control ARL is achieved.

3.3. The likelihood ratio test approach

In this section, based on Niaki et al. [20] we propose an approach to monitor binary profiles in phase II. Since the parameters of the profile are known in phase II, the objective is to test the null hypothesis 'H_0: all parameters of the regression are equal to in-control values' versus the alternative hypothesis 'H_1: some parameters are significantly different from the in-control values', i.e.:

$$\begin{array}{ll} H_0: & \boldsymbol{\beta}=\boldsymbol{\beta}_0 \\ H_1: & \boldsymbol{\beta}\neq\boldsymbol{\beta}_0 \end{array} \tag{14}$$

Niaki et al. [20] proposed a control chart based on the generalized linear test to monitor linear profiles where the concept of full and reduced models is used to develop their control chart. Therefore, we resort to a similar concept, and the partial deviance measure is applied here to test the null hypothesis. The deviance is minus twice the log-likelihood statistic. Rewriting the null and alternative hypotheses given in Relations (14) for logistic regression, we have:

$$\begin{array}{ll} H_0: & \beta_1=\beta_{01},\beta_2=\beta_{02},\cdots,\beta_p=\beta_{0p} \\ H_1: & \text{not all } \beta_k \text{ in } H_0 \text{ are equal to } \beta_{0k} \end{array} \tag{15}$$

in which β_k denotes the parameter of the logistic regression model, and β_{0k} stands for the in-control parameter of the model.

Accordingly, for the full model we have:

$$\pi_i=\frac{\exp\left(\mathbf{X}_i^T\boldsymbol{\beta}_F\right)}{1+\exp\left(\mathbf{X}_i^T\boldsymbol{\beta}_F\right)}, \tag{16}$$

where:

$$\mathbf{X}_i^T\boldsymbol{\beta}_F=\beta_1X_{i1}+\beta_2X_{i2}+\cdots+\beta_pX_{ip}.$$

The parameter vector and the deviance of the full model are represented by $\boldsymbol{\beta}_F$ and DEV($\boldsymbol{\beta}_f$), respectively. Considering the null hypothesis and denoting the parameter vector and the deviance of the reduced model by $\boldsymbol{\beta}_R$ and DEV($\boldsymbol{\beta}_R$), the deviance of the reduced model is calculated as follows:

$$\mathrm{DEV}\left(\boldsymbol{\beta}_R\right)=-2\left[\sum_{i=1}^{n}\log\binom{m_i}{y_i})+\sum_{i=1}^{n}y_i\left(\mathbf{X}_i^T\boldsymbol{\beta}_0\right)-\sum_{i=1}^{n}m_i\log\left(1+\exp\left(\mathbf{X}_i^T\boldsymbol{\beta}_0\right)\right)\right]. \tag{17}$$

Now, the partial deviance of the hypothesis testing for the jth profile can be obtained by:

$$\mathrm{DEV}_j\left(\boldsymbol{\beta}_R|\boldsymbol{\beta}_F\right)=\mathrm{DEV}_j\left(\boldsymbol{\beta}_R\right)-\mathrm{DEV}_j\left(\boldsymbol{\beta}_F\right),\quad j=1,2,\cdots \tag{18}$$

For the case of large sample sizes when the null hypothesis is accepted, the partial deviance follows a chi-square distribution with p degrees of freedom (χ_p^2). Accordingly, if $\mathrm{DEV}_j(\boldsymbol{\beta}_R|\boldsymbol{\beta}_F)\leq\chi^2_{1-\alpha,p}$, then the null hypothesis is accepted and the jth profile is ruled to be in-control.

3.4. The LRT/EWMA approach

In this section, we develop another approach by combining the LRT method with the EWMA approach to sensitize and strengthen the LRT method in detecting small shifts as well as large shifts. In this approach, based on the statistic of the proposed LRT method, an EWMA control chart is proposed to monitor a binary profile in phase II.

As stated in the previous section, the partial deviance of the hypothesis testing follows a χ_p^2 distribution. Accordingly, the partial deviance, as specified in Eq. (18), only assumes positive values. In preparation for feeding the partial deviance values to the EWMA chart, we propose to normalize the values of the partial deviance for the jth profile as shown below:

$$d_j=\frac{\left(DEV_j\left(\boldsymbol{\beta}_R|\boldsymbol{\beta}_F\right)-\chi^2_{0.5,p}\right)}{\sigma_{\mathrm{DEV}}},\qquad j=1,2,\cdots \tag{19}$$

We then calculate the statistic of EWMA control chart, using this variable as:

$$W_j=\theta d_j+(1-\theta)W_{j-1},\qquad j=1,2,\cdots \tag{20}$$

where θ represents the smoothing parameter and $W_0=0$. The proposed control chart generates a signal when $W_j>L\frac{\theta}{2-\theta}$. Again, the parameter L is calculated by simulation experiments in such a way that the specified in-control ARL is achieved.

4. Performance evaluation

In this section, the performance of the proposed control charts for phase II monitoring of a binary profile is evaluated through simulation experiments for various shifts in the parameters of the profile. All control charts are designed to have in-control ARL of approximately 200, and the smoothing parameter in both MEWMA and LRT/EWMA approaches is set equal to 0.2.

To evaluate the performances of the proposed control charts, the following profile model with one explanatory variable is considered (Montgomery et al. [14], pp. 479). In this model, the compressive strength of an alloy fastener used in aircraft construction is studied. Ten levels of the explanatory variable, loading strength measured in pounds per square inch (psi), are selected over the range 2500 – 4300 psi, and a number of fasteners are tested at each of these levels. Then, the number of fasteners failing at each level of the loading strength is recorded as the response variable. Replacing the explanatory variable x with $\log(x)$, the matrix $\mathbf{X}$ will be as follows:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ \log(2500) & \log(2700) & \log(2900) & \cdots & \log(4100) & \log(4300) \end{pmatrix}^T. \quad (21)$$

Based on this model, the in-control values of the parameters and variance-covariance matrix of logistic regression are:

$$\boldsymbol{\beta}_0 = (-42.1110,\ 5.1772)^T,$$

and:

$$\boldsymbol{\Sigma}_0 = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1} = \begin{pmatrix} 18.5689 & -2.2833 \\ -2.2833 & 0.2809 \end{pmatrix},$$

respectively.

This model is used to evaluate the performance of the proposed control chart schemes. In this study, five different values of $m_i = 5, 10, 15, 25, 50$ are considered and the upper control limit for each method and for each value of m_i is calculated by 10000 simulation runs, to achieve an in-control ARL of roughly 200. The resulting UCLs are reported in Table 1.

According to the results in Table 1, the UCL of the control charts decreases as m_i increases. However, the difference in UCL of LRT and LRT/EWMA schemes, for various values of m_i, is negligible and one can conclude that these methods are robust with respect to different values of m_i.

The out-of-control ARL values of the proposed control chart schemes for different shifts in β_1 and β_2 are summarized in Tables 2 and 3, respectively. The Standard Deviations of Run Length (SDRL) values are also reported in these tables to provide an estimate of the precision of simulation results. As stated in [21], run length values follow a geometric distribution in cases where the parameters are known. Therefore, the standard deviation of run length will be close to the mean, whenever type I error is small (see also [22]).

Based on the results in Table 2, all methods perform relatively well, and the performance of all methods improves as the value of m_i increases. The MEWMA control chart outperforms other competing methods for small to moderate shifts in β_1. LRT/EWMA method performs better than T^2 and LRT control chart schemes for shifts of almost all magnitudes. Moreover, this method performs better than MEWMA method for large shifts. The T^2 control chart displays a rather poor performance.

Table 1. Upper control limits for the proposed control charts.

m_i	T^2	MEWMA	LRT	LRT/EWMA
5	19.86	1.70	10.74	1.37
10	13.85	1.35	10.72	1.36
15	12.60	1.24	10.71	1.35
25	11.87	1.18	10.70	1.34
50	11.35	1.12	10.70	1.33

Table 3 shows that MEWMA control chart has a better performance than all other methods for small to moderate shift values. However, the performance of both LRT and LRT/EWMA methods is better than MEWMA method for larger shifts. Also, the LRT/EWMA control chart outperforms T^2 and LRT methods for almost all shifts. Again, the T^2 control chart has a poor performance in comparison to other methods.

In summary, MEWMA method performs better than the other competing control chart schemes for detecting small and moderate shift values in the parameters of the logistic regression model. However, LRT/EWMA method outperforms T^2 and LRT control charts for shifts of almost any magnitude, and outperforms MEWMA method for larger shifts. In addition, the LRT and LRT/EWMA methods utilize the robustness of the statistic for different sample sizes. This is a practical advantage that differentiates them from other methods. This characteristic for profile monitoring is significant and practical, because in real world applications, the sample size for each profile may change based on various conditions. Thus, developing a method whose statistic behaves consistently in the face of varying sample sizes in each profile will be helpful.

5. Illustrative example

In this section, the application of the proposed methods is illustrated by an example which we borrow from Montgomery et al. [14]. In this example, the effectiveness of a price discount coupon on the purchase of a two-liter beverage product is investigated by market research department of a soft drink manufacturer. The predictor variable (x) is price discount ranging from 5 to 25 cents, and the response variable (y) is the number of coupons in each price discount category redeemed after one month. Applying the logistic regression model and replacing x with $\log(x)$, the matrix $\mathbf{X}$ will be:

$$\mathbf{X} = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 & \\ \log(5) & \log(7) & \log(9) & \cdots & \log(21) & \log(23) & \log(25) \end{pmatrix}^T.$$

The in-control values of the parameters and variance-covariance matrix of logistic regression are estimated

Table 2. ARL and SDRL (in parentheses) comparisons between four proposed approaches in detecting various shifts in parameter β_1.

m_i	Chart	Shift									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
5	T^2	195.23	160.24	97.45	65.97	40.65	25.34	14.79	10.02	6.63	4.35
		(194.07)	(157.95)	(95.14)	(63.79)	(38.53)	(24.74)	(14.08)	(9.69)	(5.96)	(3.87)
	MEWMA	121.33	40.65	17.11	9.63	6.59	5.03	4.08	3.46	3.02	2.69
		(114.64)	(35.56)	(11.97)	(5.22)	(2.99)	(1.96)	(1.42)	(1.12)	(0.90)	(0.77)
	LRT	142.87	82.21	44.43	23.66	13.05	7.89	5.05	3.38	2.52	1.92
		(141.76)	(81.87)	(44.01)	(22.45)	(12.70)	(7.49)	(4.56)	(2.81)	(1.96)	(1.33)
	LRT/EWMA	144.59	64.33	28.70	13.99	8.26	5.34	3.95	3.08	2.43	2.10
		(140.95)	(60.06)	(24.55)	(10.59)	(5.54)	(3.18)	(2.14)	(1.55)	(1.17)	(0.93)
10	T^2	145.47	75.05	33.72	15.34	7.61	4.25	2.62	1.85	1.44	1.21
		(146.04)	(74.38)	(32.80)	(14.89)	(7.10)	(3.76)	(2.06)	(1.25)	(0.80)	(0.51)
	MEWMA	59.96	15.17	7.37	4.78	3.59	2.92	2.47	2.19	1.98	1.83
		(54.19)	(10.22)	(3.65)	(1.83)	(1.18)	(0.86)	(0.66)	(0.53)	(0.46)	(0.44)
	LRT	125.28	52.43	21.04	9.40	4.86	2.89	1.94	1.47	1.23	1.09
		(122.70)	(51.25)	(20.10)	(8.78)	(4.38)	(2.30)	(1.35)	(0.84)	(0.52)	(0.32)
	LRT/EWMA	112.53	32.92	12.16	6.12	3.81	2.70	2.07	1.66	1.39	1.23
		(107.57)	(28.85)	(9.00)	(3.80)	(2.05)	(1.31)	(0.94)	(0.71)	(0.56)	(0.44)
15	T^2	120.37	45.37	16.24	6.78	3.36	2.02	1.44	1.17	1.06	1.02
		(119.46)	(44.37)	(15.62)	(6.22)	(2.81)	(1.42)	(0.80)	(0.45)	(0.26)	(0.14)
	MEWMA	35.24	9.63	5.10	3.53	2.75	2.28	2.01	1.82	1.64	1.45
		(30.06)	(5.46)	(2.12)	(1.17)	(0.78)	(0.58)	(0.45)	(0.44)	(0.48)	(0.50)
	LRT	103.35	32.80	11.74	6.95	2.60	1.67	1.28	1.11	1.03	1.01
		(104.13)	(32.69)	(11.42)	(4.47)	(2.02)	(1.07)	(0.61)	(0.35)	(0.18)	(0.09)
	LRT/EWMA	86.54	19.85	7.36	3.94	2.59	1.91	1.51	1.25	1.11	1.04
		(82.91)	(15.82)	(4.80)	(2.16)	(1.25)	(0.84)	(0.62)	(0.45)	(0.31)	(0.20)
25	T^2	91.11	21.66	6.64	2.75	1.58	1.18	1.04	1.01	1.00	1.00
		(90.22)	(20.76)	(6.17)	(2.20)	(0.96)	(0.50)	(0.22)	(0.10)	(0.02)	(0.00)
	MEWMA	20.57	6.19	3.58	2.61	2.12	1.83	1.60	1.34	1.13	1.04
		(15.59)	(2.82)	(1.21)	(0.72)	(0.49)	(0.43)	(0.49)	(0.47)	(0.34)	(0.19)
	LRT	78.97	17.86	5.55	2.40	1.46	1.14	1.03	1.00	1.00	1.00
		(77.29)	(17.14)	(5.11)	(1.80)	(0.81)	(0.40)	(0.18)	(0.07)	(0.02)	(0.00)
	LRT/EWMA	54.25	10.39	4.18	2.42	1.67	1.29	1.10	1.02	1.00	1.00
		(49.14)	(7.40)	(2.32)	(1.15)	(0.72)	(0.49)	(0.30)	(0.15)	(0.06)	(0.01)
50	T^2	48.98	7.55	2.23	1.24	1.03	1.01	1.00	1.00	1.00	1.00
		(49.31)	(6.87)	(1.67)	(0.54)	(0.18)	(0.04)	(0.01)	(0.00)	(0.00)	(0.00)
	MEWMA	10.22	3.75	2.39	1.89	1.53	1.18	1.03	1.00	1.00	1.00
		(6.15)	(1.32)	(0.62)	(0.43)	(0.50)	(0.38)	(0.17)	(0.05)	(0.02)	(0.00)
	LRT	40.87	6.60	2.06	1.20	1.03	1.00	1.00	1.00	1.00	1.00
		(40.27)	(6.02)	(1.48)	(0.49)	(0.16)	(0.05)	(0.01)	(0.00)	(0.00)	(0.00)
	LRT/EWMA	24.30	4.72	2.17	1.38	1.08	1.01	1.00	1.00	1.00	1.00
		(20.22)	(2.70)	(0.99)	(0.55)	(0.28)	(0.09)	(0.01)	(0.00)	(0.00)	(0.00)

Table 3. ARL and SDRL (in parentheses) comparisons between four proposed approaches in detecting various shifts in parameter β_2.

m_i	Chart	Shift									
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**	**1.0**
5	T^2	194.72	168.37	132.75	87.58	57.35	39.42	28.75	18.76	13.49	8.78
		(181.82)	(170.81)	(132.87)	(85.87)	(57.06)	(40.87)	(28.86)	(16.88)	(13.15)	(8.92)
	MEWMA	145.26	59.31	26.31	14.32	9.46	6.83	5.44	4.48	3.86	3.38
		(140.31)	(53.82)	(20.99)	(9.37)	(5.23)	(3.20)	(2.23)	(1.68)	(1.31)	(1.09)
	LRT	154.01	105.67	63.22	37.99	22.84	14.47	9.33	6.34	4.47	3.39
		(154.86)	(105.42)	(61.94)	(37.13)	(22.84)	(13.92)	(8.80)	(5.77)	(3.94)	(2.84)
	LRT/EWMA	160.03	90.65	44.62	23.45	13.27	8.56	6.14	4.59	3.70	3.02
		(151.52)	(85.04)	(40.53)	(19.37)	(9.80)	(5.79)	(3.78)	(2.55)	(1.94)	(1.49)
10	T^2	152.56	97.72	52.09	26.28	14.01	8.09	4.94	3.26	2.29	1.77
		(152.38)	(96.88)	(51.94)	(25.52)	(13.75)	(7.58)	(4.44)	(2.70)	(1.73)	(1.15)
	MEWMA	80.94	22.33	10.10	6.41	4.69	3.69	3.08	2.67	2.37	2.15
		(75.45)	(16.91)	(6.19)	(2.88)	(1.80)	(1.24)	(0.93)	(0.76)	(0.62)	(0.52)
	LRT	144.14	72.89	34.26	17.01	8.93	5.39	3.36	2.38	1.76	1.43
		(142.19)	(72.80)	(33.18)	(16.45)	(8.40)	(4.84)	(2.83)	(1.85)	(1.17)	(0.79)
	LRT/EWMA	138.01	52.26	20.59	9.86	5.92	4.03	2.99	2.35	1.91	1.62
		(133.62)	(47.29)	(16.99)	(6.89)	(3.64)	(2.19)	(1.51)	(1.12)	(0.86)	(0.68)
15	T^2	136.16	66.44	27.79	12.97	6.29	3.58	2.30	1.68	1.32	1.16
		(133.39)	(65.39)	(27.18)	(12.49)	(5.81)	(3.06)	(1.76)	(1.07)	(0.66)	(0.43)
	MEWMA	50.62	13.71	6.97	4.58	3.47	2.82	2.42	2.13	1.94	1.79
		(44.76)	(9.16)	(3.40)	(1.75)	(1.14)	(0.83)	(0.64)	(0.51)	(0.44)	(0.45)
	LRT	124.06	51.14	20.46	9.18	4.77	2.81	1.91	1.45	1.21	1.10
		(122.71)	(50.66)	(19.94)	(8.76)	(4.29)	(2.25)	(1.34)	(0.80)	(0.51)	(0.34)
	LRT/EWMA	113.03	32.59	12.15	6.16	3.86	2.70	2.07	1.67	1.42	1.25
		(108.32)	(28.98)	(8.92)	(3.78)	(2.12)	(1.32)	(0.93)	(0.71)	(0.56)	(0.45)
25	T^2	114.14	36.65	12.41	5.16	2.61	1.66	1.27	1.09	1.03	1.01
		(114.26)	(36.13)	(12.07)	(4.65)	(2.07)	(1.04)	(0.59)	(0.32)	(0.17)	(0.08)
	MEWMA	30.18	8.48	4.63	3.26	2.56	2.17	1.91	1.71	1.50	1.29
		(24.83)	(4.66)	(1.80)	(1.03)	(0.70)	(0.51)	(0.43)	(0.47)	(0.50)	(0.46)
	LRT	99.91	30.71	10.36	4.38	2.34	1.54	1.21	1.06	1.02	1.00
		(101.14)	(30.51)	(9.99)	(3.79)	(1.78)	(0.92)	(0.51)	(0.27)	(0.14)	(0.07)
	LRT/EWMA	79.69	17.59	6.57	3.56	2.35	1.74	1.39	1.18	1.07	1.02
		(74.18)	(14.01)	(4.16)	(1.90)	(1.12)	(0.75)	(0.55)	(0.39)	(0.26)	(0.14)
50	T^2	71.09	14.20	4.12	1.84	1.22	1.04	1.01	1.00	1.00	1.00
		(71.17)	(13.09)	(3.59)	(1.25)	(0.51)	(0.21)	(0.07)	(0.01)	(0.00)	(0.00)
	MEWMA	14.54	4.86	2.98	2.23	1.88	1.58	1.28	1.08	1.01	1.00
		(9.88)	(1.96)	(0.89)	(0.54)	(0.42)	(0.49)	(0.45)	(0.27)	(0.12)	(0.03)
	LRT	62.08	12.29	3.74	1.73	1.18	1.03	1.00	1.00	1.00	1.00
		(61.46)	(11.79)	(3.25)	(1.12)	(0.47)	(0.19)	(0.07)	(0.02)	(0.00)	(0.00)
	LRT/EWMA	39.67	7.39	3.12	1.89	1.35	1.10	1.02	1.00	1.00	1.00
		(36.11)	(4.86)	(1.61)	(0.84)	(0.52)	(0.31)	(0.14)	(0.04)	(0.00)	(0.00)

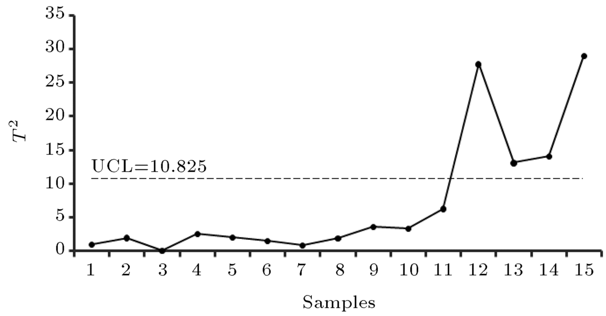

Figure 1. T^2 control chart for the example with shift from $\beta_1 = -4.5986$ to $\beta_1 = -4.4986$ at sample 11.

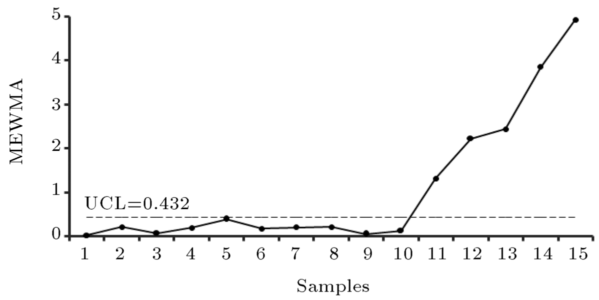

Figure 2. MEWMA control chart for the example with shift from $\beta_1 = -4.5986$ to $\beta_1 = -4.4986$ at sample 11.

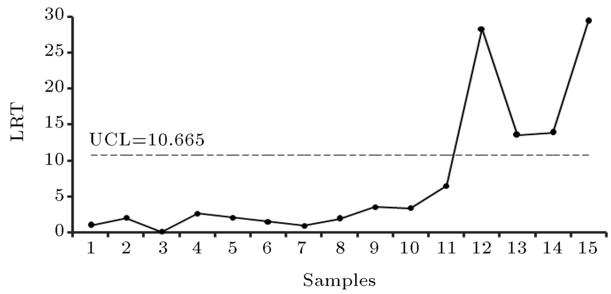

Figure 3. LRT control chart for the example with shift from $\beta_1 = -4.5986$ to $\beta_1 = -4.4986$ at sample 11.

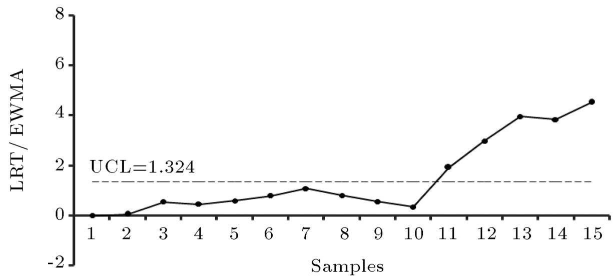

Figure 4. LRT/EWMA control chart for the example with shift from $\beta_1 = -4.5986$ to $\beta_1 = -4.4986$ at sample 11.

as $\boldsymbol{\beta}_0 = (-4.5986,\ 1.7397)^T$ and $\boldsymbol{\Sigma}_0 = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1} = \begin{pmatrix} 0.0315 & -0.0116 \\ -0.0116 & 0.0044 \end{pmatrix}$, respectively. Based on the in-control values of the parameters, 10 samples were first generated from the in-control process, and then 5 samples were generated from out-of-control process in which the value of the parameter β_1 was shifted from the in-control value of -4.5986 to -4.4986. The UCL of the proposed control charts are obtained as 10.825, 0.432, 10.665 and 1.324 to achieve the in-control ARL of roughly 200. The control charts are presented in Figures 1-4. According to these figures, T^2 and LRT generated a signal two samples after the occurrence of the shift, while both MEWMA and LRT/EWMA detected the shift correctly at sample 11.

6. Concluding remarks

Despite the fact that in many situations the response variable is discrete and in most cases a binary variable, only few research efforts have focused on this area. In this paper, we proposed four control chart schemes namely Hotelling T^2, Multivariate Exponentially Weighted Moving Average (MEWMA), likelihood ratio test and LRT/EWMA control chart for monitoring a binary response profile in phase II. We then evaluated the performance of the proposed methods in terms of Average Run Length (ARL) criterion. The results of simulation show that all methods perform relatively well. Also, the performance of the methods improves as the sample size increases. According to the results, MEWMA method outperforms other methods for small and moderate shifts in the parameters of the logistic regression model, while LRT/EWMA method performs better than MEWMA for large shift values. In addition, LRT/EWMA control chart outperforms T^2 and LRT methods for shifts of almost any magnitude and all values of sample size. Developing some control schemes for monitoring multinomial logistic regression profiles in both phases I and II would be an interesting area for future research. In addition, one may consider some control schemes for monitoring binary response profiles in the presence of profile autocorrelation for further studies.

Acknowledgment

The authors are thankful to the anonymous reviewers for their valuable comments which significantly improved the presentation of this paper.

References

1. Kang, L. and Albin, S.L. "On-line monitoring when the process yields a linear profile", *Journal of Quality Technology*, **32**, pp. 418-426 (2000).
2. Kim, K., Mahmoud, M.A. and Woodall, W.H. "On the monitoring of linear profiles", *Journal of Quality Technology*, **35**, pp. 317-328 (2003).
3. Mahmoud, M.A. and Woodall, W.H. "Phase I analysis

of linear profiles with calibration applications", *Technometrics*, **46**, pp. 380-391 (2004).

4. Saghaei, A., Mehrjoo, M. and Amiri, A. "A CUSUM-based method for monitoring simple linear profiles", *The International Journal of Advanced Manufacturing Technology*, **45**, pp. 1252-1260 (2009).
5. Zhang, J., Li, Z. and Wang, Z. "Control chart based on likelihood ratio for monitoring linear profiles", *Computational Statistics and Data Analysis*, **53**, pp. 1440-1448 (2009).
6. Mahmoud, M.A., Morgan, J.P. and Woodall, W.H. "The monitoring of simple linear regression profiles with two observations per sample", *Journal of Applied Statistics*, **37**, pp. 1249-1263 (2010).
7. Soleimani, P., Noorossana, R. and Amiri, A. "Simple linear profiles monitoring in the presence of within profile autocorrelation", *Computers & Industrial Engineering*, **57**, pp. 1015-1021 (2009).
8. Hosseinifard, S.Z., Abdollahian, M. and Zeephongsekul, P. "Application of artificial neural networks in linear profile monitoring", *Expert Systems with Applications*, **38**, pp. 4920-4928 (2011).
9. Kazemzadeh, R.B., Noorossana, R. and Amiri, A. "Phase I monitoring of polynomial profiles", *Communications in Statistics-Theory and Methods*, **37**, pp. 1671-1686 (2008).
10. Zou, C., Tsung, F. and Wang, Z. "Monitoring general linear profiles using multivariate exponentially weighted moving average schemes", *Technometrics*, **49**, pp. 395-408 (2007).
11. Amiri, A., Eyvazian, M., Zou, C. and Noorossana, R. "A parameters reduction method for monitoring multiple linear regression profiles", *The International Journal of Advanced Manufacturing Technology*, **58**, pp. 621-629 (2012).
12. Mahmoud, M.A. "Phase I analysis of multiple linear regression profiles", *Communications in Statistics-Simulation and Computation*, **37**, pp. 2106-2130 (2008).
13. Jensen, W.A., Birch, J.B. and Woodall, W.H. "Monitoring correlation within linear profiles using mixed models", *Journal of Quality Technology*, **40**, pp. 167-183 (2008).
14. Montgomery, D.C., Peck, E.A. and Vining, G.G., *Introduction to Linear Regression Analysis*, 3rd Edn., John Wiley & Sons, New York (2001).
15. Hosmer, D.W. and Lemeshow, S., *Applied Logistic Regression*, 2nd Edn., John Wiley & Sons, New Jersey (2000).
16. Yeh, A.B., Huwang, L. and Li, Y.M. "Profile monitoring for binary response", *IIE Transactions*, **41**, pp. 931-941 (2009).
17. Shang, Y., Tsung, F. and Zou, C. "Profile monitoring with binary data and random predictors", *Journal of Quality Technology*, **43**, pp. 196-208 (2011).
18. Koosha, M. and Amiri, A. "The effect of link function on the monitoring of logistic regression profiles", *Proceedings of the World Congress on Engineering*, pp. 6-8 (2011).
19. McCullagh, P. and Nelder, J.A., *Generalized Linear Models*, 2nd Edn., Chapman & Hall, London (1989).
20. Niaki, S.T.A., Abbasi, B. and Arkat, J. "A generalized linear statistical model approach to monitor profiles", *International Journal of Engineering Transactions A: Basics*, **20**, pp. 233-242 (2007).
21. Ryan, T.P., *Statistical Methods for Quality Improvement*, 3rd Edn., John Wiley & Sons, New Jersey (2011).
22. Kazemzadeh, R.B., Noorossana, R. and Amiri, A. "Phase II monitoring of autocorrelated polynomial profiles in AR(1) processes", *Scientia Iranica*, **17**, pp. 12-24 (2010).

Biographies

Mir Emad Soleymanian is a PhD degree student in Statistics at University of Florida. He received his BS and MS degrees both in Industrial Engineering from Sharif University of Technology. His research interests are in the areas of applied statistics and statistical quality control.

Majid Khedmati is a PhD degree student in Industrial Engineering at Sharif University of Technology. He received his BS and MS degrees both in Industrial Engineering from Iran University of Science and Technology and Sharif University of Technology in 2010 and 2012, respectively. His research interests are in the areas of multivariate quality control, applied statistics and operations research.

Hashem Mahlooji is a Professor of Industrial Engineering at Sharif University of Technology. His research interests include discrete event systems simulation, Monte Carlo and applied statistics.