Research Note

Monitoring multivariate-attribute quality characteristics in two-stage processes using discriminant analysis based control charts

S. Zolfaghari and A. Amiri*

Department of Industrial Engineering, Faculty of Engineering, Shahed University, P.O. Box 18151-159, Tehran, Iran.

Abstract. In this paper, we focus specifically on a two-stage process with multivariate-attribute quality characteristics in the second stage. The main purpose of this study is extending Discriminant Analysis (DA) based control charts to monitor a two-stage process. We propose three methods including EWMA (DA), integrated EWMA (DA), and P-value (DA), and integrated Multivariate Exponentially Weighted Moving Average (MEWMA) and T^2 charts based on the DA approach to monitor the multivariate-attribute quality characteristic in a two-stage process. The performance of the proposed methods is evaluated through simulation studies as well as a real case.

1. Introduction

Statistical process control techniques are widely used for monitoring and improving quality characteristics. Control charts are usually used to identify the unnatural patterns in manufacturing processes. Control charts are an important tool for distinguishing between common causes and assignable causes of variation. These unnatural patterns are often related to assignable causes that lead to generation of out-of-control signals. Afterwards, corrective actions are necessary to restore the process to in-control state.

In most industries, a product is usually manufactured in more than one stage, referred to as multistage process. In these processes, the quality in the current stage is affected by the previous stages. This feature is mentioned as cascade property. As a result, quality of the final product depends not only on the current stage, but also on quality characteristics of the preceding stages. So, using the common approaches can be misleading. Different studies have been done in this area. Using separate control charts to monitor the quality characteristics in each stage is one of the common approaches in the literature. This approach does not consider the correlation propagation through different stages. The second approach is application of multivariate control charts to monitor all stages, simultaneously. In this approach, the interpretation of out-of-control signal is complicated. The next approach is developing the cause-selecting control chart. This approach removes the effect of preceding stages from the current stage. Zhang [1] suggested cause-selecting control chart based on the residuals. In this approach, the relation between stages is defined and then the values of residuals are calculated. The effect of the previous stages is removed from the computed residuals. Hawkins [2] proposed model-void method. This type of regression adjustment includes the state to which all or some of the quality characteristics are related, but a change in one quality characteristic does not affect the others. Moreover, Hawkins [3] introduced model-fix procedure. However, a shift in

*. Corresponding author. Tel.: +98 21 51212065
E-mail addresses: s.zolfaghari@shahed.ac.ir (S. Zolfaghari);
amiri@shahed.ac.ir (A. Amiri)

any quality characteristic will affect some or all of the following quality characteristics. Generally, the output quality characteristics are not always normally distributed. In this case, the Generalized Linear Model (GLM) is applied. This technique relates the mean of response variable to input variables. Jearkpaporn et al. [4] considered multiple gamma-distributed output variables and proposed a likelihood ratio statistic to monitor the process. Skinner et al. [5] extended control chart based on deviance residuals for monitoring Poisson distributed response variables. Jearkpaporn et al. [6] discussed the multistage process in which output variables included normal and non-normal variables. Deviance residuals were calculated with regard to GLM. The performance of the proposed method was compared with Hotelling T^2 chart based on U statistic and conventional Shewhart control chart. Aghaie et al. [7] considered a two-stage process in which the output quality characteristic is affected by input quality characteristic. They assumed that quality characteristics follow Bernoulli and Poisson distributions in the first and second stages, respectively. They explored several monitoring schemes based on generalized Poisson distribution. A process with two dependent stages is considered by Yang and Yeh [8] in which quality characteristics in each stage follow a bivariate binary distribution. The performance of the proposed method based on cause-selecting control chart was compared with those of Shewhart attribute control chart and bivariate binomial control region. The result of the proposed method overshadowed the other methods. Asgari et al. [9] developed a new link function which combined the two log and square root link functions for Poisson quality characteristic of the second stage. Afterwards, the Shewhart and EWMA control charts were used to monitor the computed residuals. Asadzadeh et al. [10] considered a two-stage process with the assumption that historical data include outliers. They established the relation between two stages through extending robust fitting approach by applying compound estimator. Then, the robust monitoring procedure was utilized to control the process. The performance of the proposed method was evaluated through average run length criterion. Liu [11] provided a review of the explored methodologies with regard to variation reduction scheme in multistage manufacturing processes. He acknowledged the effect of recent development in the stream of variation and statistical process control methods on the variation reduction. The two mentioned methodologies are evaluated through a real case to conclude their significant influence on variation reduction.

A multistage process with binary data is investigated by Shang et al. [12]. They used a binary state space model and proposed monitoring and diagnosis schemes based on a hierarchical likelihood method.

They showed that the proposed schemes perform better than the χ^2 control chart. Ghahyazi et al. [13] considered a two-stage process with simple linear profile quality characteristic in each stage. They first showed the effect of cascade property on the statistical performance of T^2 control chart and then proposed a control chart based on the U statistic to monitor the process. Some researchers such as Asadzadeh and Aghaie [14], and Asadzadeh et al. [15-17] considered multistage processes with reliability data in the second stage. A literature review on cause-selecting control chart is given by Asadzadeh et al. [18]. Li [19] considered a multistage process under multiple faults. He used the state space model to define the process and Bayesian theory to propose a new control chart for monitoring the process. Li and Tsung [20] investigated changes in the covariance matrix of multistage processes and extended an EWMA-based control chart. Amiri et al. [21] assumed a two-stage process with Poisson output quality characteristic. They developed a cause-selecting control chart using standardized residual of the generalized linear model.

In some processes, quality of a process or product is described by correlated variable and it attributes quality characteristics. Most of the research has been done in multivariate or multi-attribute processes. Chiu and Kuo [22] considered bivariate binomial processes. They developed a new control chart with the ability to determine the source of out-of-control state. Moreover, their proposed control chart was not sensitive to transform multi attribute quality characteristics to multivariate normal quality characteristics. Li and Tsung [23] proposed the multiple binomial and Poisson CUSUM control charts. They utilized the concept of False Discovery Rate (FDR) control techniques to construct new control chart to enhance the power of this type of control charts. They proved their claim through Mont Carlo simulation. Butte and Tang [24] suggested a new procedure to enhance the ability of multivariate control charts such as T^2, MEWMA, and CUSUM to determine the variable responsible for out-of-control state. McCracken and Chakraborti [25] provided an overview on control charts proposed for monitoring mean and variance of normal quality characteristics, individually or simultaneously. However, there is little research on monitoring multivariate and multi-attribute quality characteristics, simultaneously. Doroudyan and Amiri [26] applied four transformation techniques for monitoring multivariate-attribute processes. They applied skewness reduction and correlation elimination techniques in the developed methods. Amiri et al. [27] explored a process in which a correlated simple linear profile and multivariate quality characteristics were monitored over time. They proposed a method based on MEWMA control chart.

Multivariate statistical methods like discriminant

analysis are also used in statistical process control. The fundamental concept of DA is to detect the optimal discriminant function. DA uses both normal and abnormal data to find the optimal direction for classification. He et al. [28] proposed fault diagnose approach based on discriminant analysis. The procedure includes three stages. The first step is the analysis that categorizes historical data to normal and fault separate clusters by k-means method. The next step is fault visualization that allocates data in two clusters by discriminant analysis and the last step is fault diagnosis. Pei et al. [29] applied DA to detect discriminant direction, then used X/MR control chart to monitor a chemical process. Bazdar and Kazemzadeh [30] introduced discriminant function to distinguish the source of variation in multistage processes.

In this paper, a two-stage process with combined variable-attribute quality characteristics at the second stage is considered. We propose some new control charts based on the discriminant analysis and evaluate performance of the proposed control charts with the traditional ones in the literature. The result of simulation studies shows the better performance of the proposed DA-based control chart. Also, using the DA-based control charts reduces the dimension of quality characteristics to univariate quality characteristic, which is easily monitored by univariate control charts.

This paper is organized as follows: In the next section, a two-stage process is modeled using a simple linear regression and a generalized linear model. In Section 3, the proposed control charts are developed. In Section 4, the performance of the proposed control chart is compared through simulation studies with the traditional control charts in terms of average run length criterion. The performance of the best proposed control charts is also evaluated in a real case in Section 5. Conclusion and future research are given in the final section.

2. Modeling a two-stage process

The process is assumed to include two dependent stages. The quality characteristic of the first stage is normally distributed and the quality characteristics of the second stage are combined correlated normal and Poisson distribution. Due to the correlation structure and dependent stages, monitoring quality characteristics in each stage with separate control charts leads to misleading results. Any shift in the mean of Poisson distribution affects the variance of Poisson distribution as well. Hence, without loss of generality, we assume that the correlation matrix of the quality characteristics in the second stage is fixed. The process is illustrated in Figure 1.

In two-stage processes, quality characteristics of
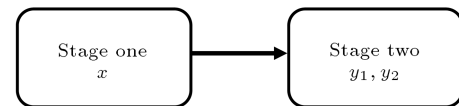
Figure 1. A two-stage process with variable-attribute quality characteristic in the second stage.

the second stage are linked to quality characteristics in the first stage through link function. When the quality characteristics follow normal distribution, least square method is applied. However, the least square method does not perform well in non-normal quality characteristics. The alternative procedure for linking the mean of quality characteristics in the second stage to quality characteristic in the first stage is Generalized Linear Model (GLM). The GLM method is used for Exponential, Gamma, Poisson, and Binomial distributions by specific link function. The simple linear regression model is given by:

$$y_1 = \alpha \mathbf{x^T} + \varepsilon. \tag{1}$$

The Poisson log link is written as follows:

$$\log \mu_{y_2} = \mathbf{x^T} \boldsymbol{\beta}. \tag{2}$$

Therefore, the mean of Poisson distribution is defined as:

$$\mu_{y_2} = \exp\left(\mathbf{x^T} \boldsymbol{\beta}\right), \tag{3}$$

where x and $\mathbf{y} = (y_1, y_2)$ are the quality characteristics in the first and second stages. $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are the vectors of regression parameters in simple linear regression and GLM, respectively. The vector of $\mathbf{x^T}$ includes unit vector and quality characteristics vector in stage one, $\mathbf{x^T} = (1, x)$.

3. Proposed methods

Since the quality characteristics in the second stage are a combination of normal and non-normal quality characteristics, skewness of non-normal quality characteristics affects the performance of traditional multivariate control charts. Hence, various transformation techniques are suggested by many authors to decrease the skewness of observations. Here, the root transformation technique is applied to reduce the skewness. The transformed data becomes zero skewed and follows approximately normal distribution. In this case, the vector of power in the root transformation method is specified by a bisection technique. The values of all elements of the vector are between 0 and 1. This vector, as the power of original variables, generates new variables by zero skewness. For more information about the root transformation method refer to Niaki and Abbasi [31]. In the root transformation method, a numerical algorithm, such as bisection method, is

used. In each iteration, one of the boundary intervals is replaced by the middle of an interval. The procedure of root finding $f(x) = 0$ in the interval of (a_0, b_0) is given in Appendix C. The idea of skewness reduction is finding an argument of the skewness function, $f(r) = 0$, such that it becomes close to zero. Then, this value is used as power of the original data and transforms the data set to the ones with symmetric distribution.

Since shift in the mean of Poisson distribution changes the variance of this distribution, the observations are standardized to prevent the shift in variance of the Poisson distribution.

As mentioned before, the quality characteristics in the second stage are dependent on quality characteristics in the first stage. To overcome the effect of cascade property, residual values are calculated by using Eq. (4) as follows:

$$e_{ji} = y_{ji} - \widehat{y}_{ji}, \qquad (4)$$

where y_{ij} is the jth quality characteristic in time i, and the value of $\hat{y}_{ji}$ is conditional expected value, $\hat{y}_{ji} = E(y_{ji}|x_i)$. Residuals are normally distributed with mean 0 and standard deviation σ_y, $e_{ji} \sim N(0, \sigma_y^2)$.

For monitoring the two-stage processes discussed in Section 2, three methods are proposed hereunder.

3.1. EWMA control chart based on the DA statistic

Discriminant analysis defines a new axis that provides the maximum isolation between normal and abnormal observations and divides the observations into two groups. The new variables are a linear combination of the initial variables. Hence, discriminant analysis is described as dimensionality reduction method by projection of initial variables on the new axis. The objective is to obtain optimal discriminant direction by maximizing the following criterion:

$$\vartheta = \frac{SS_B}{SS_W}, \qquad (5)$$

where SS_B represents the variation between the groups and SS_W is the variation within the groups for new variables. The maximum value of ϑ reflects the best homogenization in groups. It means the groups are formed such that in-control observations are allocated in one group and the out-of-control observations are dedicated to another group. The objective of the DA method is providing maximum isolation between two groups, so θ is calculated such that the ϑ ratio of new observations (obtained from projection) is maximized. It implies that new observations, ω_i, provide the maximum distinction between two groups and maximum homogeneity within the groups.

Generally, the values of SS_B and SS_W are calculated from Eq. (6):

$$SS_B = \sum_{j=1}^{2} n_j \left(\bar{y}_{j.} - \bar{y}_{..} \right)^2,$$

$$SS_W = \sum_{j=1}^{2} \sum_{i=1}^{n} n_j \left(y_{ji} - \bar{y}_{j.} \right)^2, \qquad (6)$$

where n_j is the number of observations in each group, $\bar{y}_{j.}$ is the mean of observations within the groups, and $\bar{y}_{..}$ is the overall mean.

In the two dimensional spaces, with the axis of y_1 and axis of y_2, consider a new axis such as ω, where the angle between the new axis and axis y_1 is equal to θ. The projection of initial data on the new axis is obtained by Eq. (7). Since the multistage process is studied, the primary quality characteristics, y_1 and y_2, are replaced by residual values, e_1 and e_2, to eliminate the effect of quality characteristic in the first stage on the quality characteristics in the second stage. For more information about two-dimensional discriminant analysis, refer to [32].

With the purpose of dimensionality reduction and generating new variables, the following statistic is achieved:

$$\omega_i = \cos\theta e_{1i} + \sin\theta e_{2i}, \qquad (7)$$

where ω_i is the projection of the ith observation on the axis ω, and e_{1i} and e_{2i} are the residual values. ω is a linear combination of the residuals. Since the residuals are normally distributed, ω is normally distributed as:

$$\omega_i \sim N\left(0, (\cos\theta\sigma_1 + \sin\theta\sigma_2)^2 - \sin(2\theta)\sigma_1\sigma_2(1-\rho) \right),$$

where ρ is the correlation coefficient. Since the study is conducted in Phase II, the value of ρ is estimated from historical observations in Phase I. The proofs for mean and variance of the ω statistic are given in Appendix A.

The EWMA control charts carry out the better detection small and moderate shifts rather than Shewhart control charts. The EWMA control charts use all the previous data by allocating a weight to the last observation and to the sum of all prior values. The importance of historical data in calculating the EWMA statistic is defined by weight value. The large value indicates the less importance of historical data. The EWMA statistic is defined in Eq. (8):

$$u_i = \lambda\omega_i + (1-\lambda)u_{i-1}. \qquad (8)$$

λ is smoothing parameter between 0 and 1 and z_0 is equal to zero. The EWMA control limits are derived as:

$$\begin{aligned} UCL \\ &= \mu_u \pm L\sigma_\omega\sqrt{\tfrac{\lambda}{2-\lambda}}. \\ LCL \end{aligned} \qquad (9)$$

L is computed with simulation at the target value of ARL$_0$.

3.2. Integrated EWMA (DA) and P-value (DA) control charts

In this subsection, another method, called P-value (DA), is proposed and applied in combination with the EWMA (DA) method. In the P-value (DA) method, first, we standardized the DA statistic explained in the previous section as follows:

$$z_i = \frac{\omega_i - \mu_\omega}{\sigma_\omega}. \tag{10}$$

In Eq. (10), ω_i is the DA statistic of the ith sample and z_i is standardized DA statistic. Since the DA statistic follows normal distribution, the standardized DA statistic follows normal distribution as well. Hence, the P-value is calculated by the following Equation:

$$P - \text{value} = 2p(z > z_i) = 2\left(1 - \phi(|z_i|)\right), \tag{11}$$

where $\phi(0)$ is cumulative normal distribution function. The P-value statistic proposed for monitoring the process is compared with predefined significant level, α. If the P-value exceeds α, the process is in-control; otherwise it is in out-of-control state.

The EWMA (DA) control chart, explained in the previous subsection, is used in combination with the proposed P-value (DA) method.

The overall probability of Type I error (α_{overall}) in two control charts with equivalent probability of Type I error, α, is determined as follows:

$$\alpha_{\text{overall}} = 1 - \left(1 - \alpha\right)^2. \tag{12}$$

Consequently:

$$\alpha = 1 - \sqrt{1 - \alpha_{\text{overall}}}. \tag{13}$$

3.3. Integrated MEWMA and T^2 control charts based on DA

Lowry et al. [33] extended the univariate EWMA statistic to multivariate case. The statistic of the MEWMA control chart is given by:

$$w_i = \boldsymbol{\gamma}_i' \sum_{\boldsymbol{\gamma}_i}^{-1} \boldsymbol{\gamma}_i. \tag{14}$$

The $\boldsymbol{\gamma}_i$ vector is calculated from recursive equation as follows:

$$\boldsymbol{\gamma}_i = \lambda \mathbf{y}_i + (1 - \lambda)\boldsymbol{\gamma}_i - 1, \tag{15}$$

where $\mathbf{y}_i$ is the vector of observations in time of i and $\boldsymbol{\mu}$ is the mean vector of observations. In the proposed method, the vector of observations, $\mathbf{y}_i$, is replaced by vector of residuals, $\mathbf{e}_i$, obtained from Eq. (4). z_0 is equal to zero vector and the covariance matrix of $\boldsymbol{\gamma}_i$ is derived as:

$$\sum_{\boldsymbol{\gamma}_i} = \frac{\lambda}{n(2 - \lambda)} \sum_{\mathbf{e}}. \tag{16}$$

The upper control limit is determined such that the desired in-control average run length (ARL_0) is achieved.

To improve the sensitivity of MEWMA control chart in detecting large shifts, a T^2 control chart is used simultaneously. The T^2 statistic is calculated by the following Equation:

$$T_i^2 = (\mathbf{e}_i)' \sum_{\mathbf{e}}^{-1} (\mathbf{e}_i). \tag{17}$$

A new method is proposed based on discriminant analysis. DA method is applied on MEWMA and T^2 statistics. The DA method reduces the multivariate space to univariate space through DA statistic. The plotting statistic is given by:

$$\zeta_i = \cos\theta(T_i^2) + \sin\theta(w_i), \tag{18}$$

ζ_i is the integrated T^2 and MEWMA statistic based on DA, T_i^2 and w_i are the T^2 Hotelling statistic and the MEWMA statistics of the ith sample, respectively. Since ζ_i is a linear combination of two statistics with upper control limit, we only set the UCL for ζ_i. The UCL is obtained by simulation study to achieve the desired in-control ARL.

Generally, the steps of the proposed methods as well as competing methods are summarized in the flowchart shown in Figure 2.

4. Simulation studies

In this section, the performance of the proposed methods is evaluated through a numerical example. To compare the performances of the proposed methods, control limits are calculated such that the in-control Average Run Length (ARL_0) equal to 200 is obtained. Then, the out-of control Average Run Length (ARL_1) criterion is computed by using simulation through 5000 replication runs.

In the numerical example, a two-stage process is considered. Quality characteristic of the first stage is normally distributed with mean 2 and standard deviation 1. Quality characteristics in the second stage are illustrated by a vector of quality characteristics as $\mathbf{y} = (y_1, y_2)$. y_1 and y_2 are quality characteristics with normal and Poisson distributions, respectively. Due to the cascade property, the mean of quality characteristic y_1 through the simple linear regression model and the mean of y_2 through log link function are linked to the quality characteristic in the first stage. When the process is in-control, the coefficients vectors are defined as $\boldsymbol{\beta} = (1, 0.5)$ and $\boldsymbol{\gamma} = (0.5, 0.5)$ for simple linear regression model and generalized linear model, respectively. To evaluate performance of the proposed methods, a random vector of correlated variables is generated using Gaussian copula function. The root transformation technique is implemented
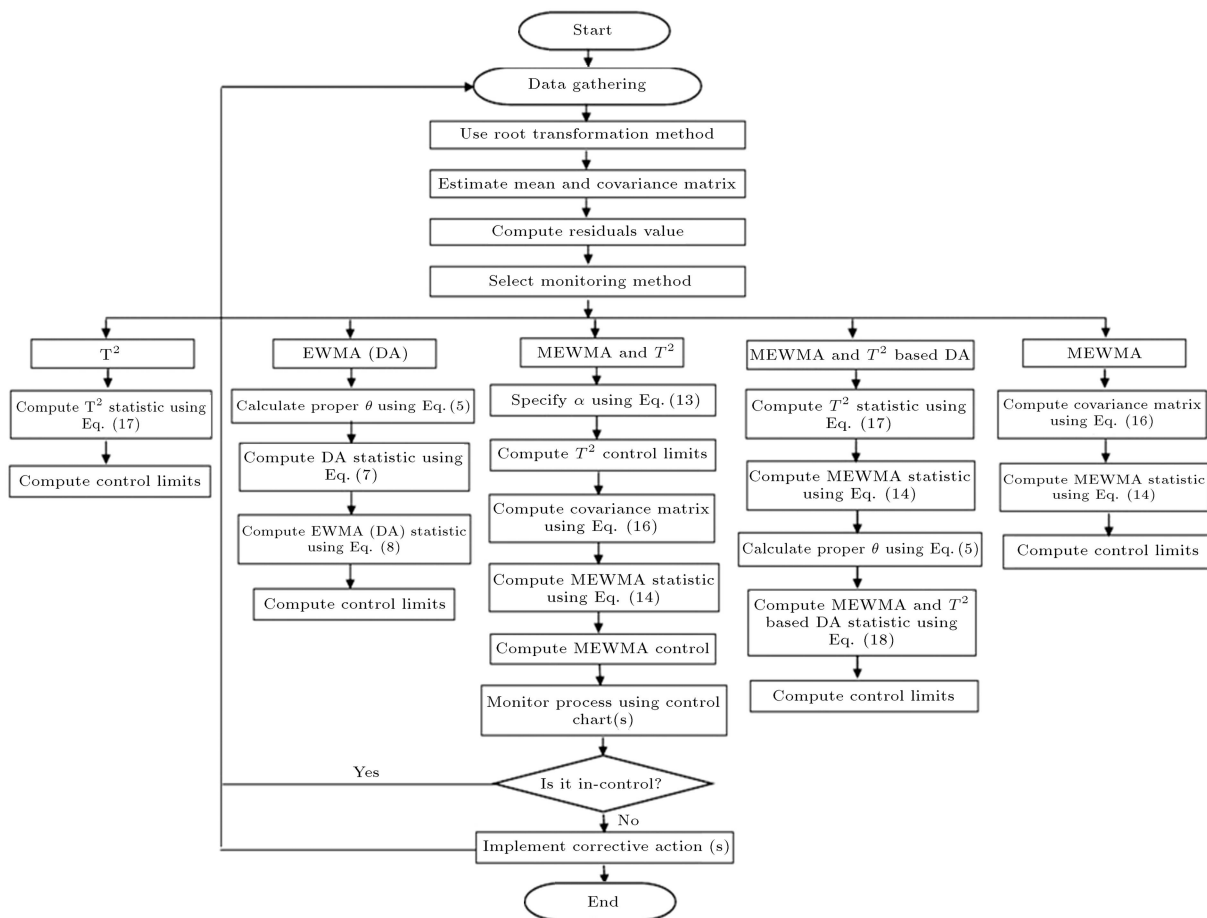
Figure 2. The steps of the proposed methods as well as competing methods.

on the original data; so, the corresponding vector of power is $\gamma = (1, 0.7754)$. After transformation, the new data follows bivariate normal distribution. Due to the cascade property in multistage processes, residual values are calculated to eliminate effect of the quality characteristic in the first stage on the quality characteristics in the second stage.

4.1. Parameters in the proposed and competing methods

4.1.1. Hotelling T^2 control chart
The statistic of T^2 control chart is obtained by Eq. (17). $\sum$ is the covariance matrix of residuals, which is equal to $\sum_e = \begin{pmatrix} 0.9903 & 0.2355 \\ 0.2355 & 0.4475 \end{pmatrix}$ and it is estimated by simulation based on a dataset with 10000 historical values. The UCL equal to 10.5966 is calculated based on χ^2 distribution with probability of Type I error equal to 0.005 and degrees of freedom 2 (number of quality characteristics).

4.1.2. EWMA (DA) control chart
The EWMA (DA) statistic is calculated with value of $\theta = 35.6014$ by Eq. (8). For calculating θ, data generation is done based on the following procedure.

First, 900 in-control observations are generated. Next, 300 out-of-control observations is procreated by applying shift in coefficient x with magnitude of 1 in the mean of normal distribution. Afterwards, 300 out-of-control observations is obtained by applying shift in Poisson mean with the same procedure. At the end, 300 out-of-control observations is produced by simultaneous shifts in normal and Poisson means. Finally, the residual values are calculated using Eq. (4) and the maximum value of ϑ is computed based on the calculated residuals using Eq. (5).

According to Eq. (9), to set the EWMA control limits based on $\lambda = 0.2$, the mean of u statistic and standard deviation of ω statistic are required. The aforementioned parameters are computed as -0.0225 and 1.0184 with 100000 replications, respectively. To achieve the desired $ARL_0 = 200$, L in Eq. (9) is defined equal to 2.633.

4.1.3. Integrated EWMA (DA) and P-value (DA) control chart
In integrated control charts, the first step is determining the probability of Type I error for each control chart, separately. As mentioned before, ARL_0 equals to 200 and it represents the overall probability of Type I

error equal to 0.005. Therefore, the probability of Type I error for each control chart is calculated by applying Eq. (13) such that the probability of Type I error equal to 0.0025 or, similarly, ARL_0 equal to 400 is obtained.

Designing the EWMA (DA) is similar to the one in the previous subsection, except for which is equal to 2.89 and leads to the desired in-control ARL equal to 400.

In the P-value (DA) control chart, to compute the in Eq. (10), the mean value of ω is specified as -0.0225 by 100000 replication runs. Based on the concept of P-value, the P-value larger than 0.0025 shows in-control state.

4.1.4. Integrated T^2 and MEWMA control charts

In this method, two multivariate control charts monitor the process simultaneously. Designing the T^2 control chart is similar to that in subsection 4.1.1. As mentioned previously, the overall probability of Type I error is equal to 0.005. By applying Eq. (13), the probability of Type I error for each control chart is equal to 0.0025; hence, the UCL is obtained equal to 11.829.

The MEWMA statistic is calculated by using Eq. (14). The covariance matrix of $\sum_e$ is equal to $\begin{bmatrix} 0.9903 & 0.2355 \\ 0.2355 & 0.4475 \end{bmatrix}$ and UCL equals 11.16 to achieve the desired $ARL_0 = 400$ via the UCL = 9.83 in the MEWMA control chart.

4.1.5. Integrated T^2 and MEWMA control charts based on DA

The corresponding MEWMA covariance matrix is calculated using Eq. (16) by replacing the residual covariance matrix mentioned in Subsection 4.1.1. The statistic based on DA approach is computed by using Eq. (18) with using the simulated $\theta = 66.26$. The upper control limit is achieved 11.661 by 5000 replications.

The results of the average run length and standard error under different magnitudes of shift are summarized in Table 1. All of the proposed methods are compared with the competing T^2, MEWMA and integrated T^2, and MEWMA control charts through 5000 replications. The results show the perfect performance of the integrated MEWMA and T^2 control charts based on the DA under different magnitudes of shift.

5. Case study

In this section, the case study by Jearkpaporn et al. [6] is used. As shown in Figure 3, a two-stage semiconductor manufacturing process is considered. The first stage is an oxide deposition process and the second stage is spin on glass coat process. The averages of the oxide thickness measurements and particle counts are two quality characteristics in each stage with normal and Poisson distribution (y_{ai}), respectively.
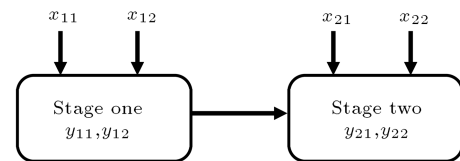
Figure 3. An illustration of a two-stage process in a real case.

The applied link functions between quality characteristics in two stages are simple linear regression model and log link for normal and Poisson quality characteristics, respectively. y_{ai} indicates the ith quality characteristic in the ath stage. Each stage includes two input variables with uniform distribution in the interval of [0,1]. These input variables are gas flow and volume dispensed in both stages. The distribution function and parameters of quality characteristics are given in Appendix B. Jearkpaporn et al. [6] proposed Deviance Residual (DR) based on Generalized Linear Model (GLM) to monitor the process. By considering probability of Type I error equal to 0.005, the performances of the DR method and the integrated T^2 and MEWMA based on the DA, EWMA (DA) and integrated EWMA (DA), and P-value (DA) are compared. In the DR method, separate control charts are utilized to monitor each quality characteristic; so, the probability of Type I error is calculated from Eq. (13) to design the control chart limits. Shift in x_{2i} and y_{1i} coefficients with magnitude of 0.1 is evaluated to detect out-of-control condition. x_{2i} shows the ith input variable in the second stage and y_{1i} shows the ith output quality characteristic in the first stage. The simulation results with 5000 replications are illustrated in Figure 4. The horizontal and vertical axes represent output quality characteristics and out-of control average run length (ARL_1) in the second stage, respectively. The results show the satisfied performance of the proposed methods.

6. Conclusion and future researches

Most of the processes include two or more stages that produce service or products. Frequently, quality of a product or service is represented by several correlated quality characteristics. In this paper, we considered a two-stage process with multivariate-attribute quality characteristics. The root transformation technique was utilized to transform attribute quality characteristics. We proposed three methods to monitor multivariate-attribute processes in a two-stage process. These methods were (i) EWMA (DA), (ii) integrated EWMA (DA) and P-value (DA), and (iii) integrated MEWMA and T^2 control chart based on the DA. The performance of the proposed methods was compared with those of conventional competing methods, including Hotelling T^2, MEWMA, and the integrated T^2 and

Table 1. ARL_1 values of all the proposed and competing methods under different magnitudes of shift.

Parameter	Magnitude shift	T^2		EWMA (DA)		MEWMA and T^2		EWMA (DA) and P-value (DA)		MEWMA and T^2 based DA		MEWMA	
		ARL_1	Std-error	ARL_1	Std-error	ARL_1	Std-error	ARL_1	Std-error	ARL_1	Std-error	ARL_1	Std-error
$\beta = (\beta_1, \beta_2)$	(0.1,0)	186.3050	5.4085	177.277	2.4366	170.0972	2.3432	183.6342	2.4849	1603174	2.2289	160.9062	2.1194
	(0.2,0)	174.154	5.2675	124.2722	1.6771	121.4740	1.6363	142.786	1.9302	108.3592	1.4518	105.8134	1.3846
	(0.3,0)	154.0720	4.8119	83.1056	1.0825	80.0002	1.0266	99.6884	1.3466	70.9382	0.9035	67.4618	0.8514
	(0.5,0)	100.278	3.2046	40.7750	0.4963	35.1814	0.4181	48.1068	0.6157	32.0780	0.3698	30.5438	0.3412
	(1,0)	34.472	1.0211	13.2036	0.1167	10.8446	0.0838	14.7722	0.1441	10.3350	0.0769	9.9068	0.0715
	(1.5,0)	11.921	0.373	7.3862	0.0479	6.1234	0.0364	7.7002	0.0586	5.9540	0.0320	5.923	0.0287
	(2,0)	5.3050	0.1509	5.2206	0.0254	4.3594	0.0215	5.0802	0.0349	4.3362	0.0191	4.4254	0.0161
	(3,0)	1.8026	0.1072	3.6422	0.0121	2.7654	0.0126	2.7968	0.0210	2.9546	0.0107	3.2646	0.0077
	(4,0)	1.1314	0.0056	3.0054	0.0075	2.1622	0.0058	1.6738	0.0132	2.3456	0.0071	2.8056	0.0063
	(0,0.1)	176.7226	2.4994	121.7178	1.6387	122.8610	1.6397	138.7046	1.9383	107.3878	1.4376	103.2996	1.3739
	(0,0.2)	115.725	1.5995	54.2736	0.6984	51.0478	0.6357	65.0806	0.8393	44.3272	0.5526	42.5052	0.5093
	(0,0.3)	68.3132	0.9607	28.2992	0.3284	24.5696	0.2651	34.2374	0.4203	22.73	0.244	20.8170	0.2129
	(0,0.5)	22.7490	0.3173	12.4564	0.1148	10.1740	0.0820	13.6662	0.1383	9.7908	0.0767	9.5684	0.0703
	(0,1)	3.7050	0.0451	5.3198	0.0312	4.1952	0.0262	4.6306	0.039	4.2708	0.0242	4.566	0.0217
	(0,1.5)	1.855	0.0176	3.718	0.0178	2.9226	0.0161	2.5596	0.0234	3.039	0.0153	3.365	0.0143
	(0,2)	1.4054	0.0108	3.0856	0.0136	2.4884	0.0114	1.8006	0.0157	2.5762	0.0111	2.8524	0.0117
	(0,3)	1.1878	0.0067	2.552	0.0103	2.2056	0.0069	1.3038	0.0087	2.2544	0.0074	2.4035	0.0089
	(0,4)	1.111	0.0049	2.3106	0.0078	2.115	0.005	1.1768	0.0064	2.1488	0.0057	2.2368	0.0072
	(0.25,-0.25)	145.722	4.6788	94.1842	1.2489	102.6468	1.4206	109.0700	1.4726	92.6998	1.2201	95.8686	1.2414
	(0.5,-0.5)	61.52	1.8917	34.7622	0.4055	31.1874	0.3764	40.4306	0.5125	29.2758	0.3404	29.4796	0.3404
	(1,-1)	11.2030	0.1484	11.6776	0.1049	9.0054	0.0835	11.1144	0.1145	9.0334	0.0823	9.489	0.0802
	(2,-2)	2.7578	0.0311	5.3032	0.0384	3.7476	0.0271	3.8766	0.0406	4.0774	0.0297	4.679	0.0323
$\gamma = (\gamma_1, \gamma_2)$	(0.1,0)	200.353	6.3529	198.5202	2.7379	203.4636	2.7874	203.1578	2.7711	200.7912	3.273	208.8306	2.8480
	(0.2,0)	187.027	2.62356	175.987	2.40272	159.148	2.16796	185.3588	2.5997	144.1626	1.9813	144.8824	1.9326S
	(0.3,0)	168.202	2.30500	149.049	2.06892	105.272	1.43374	169.463	2.3633	92.2834	1.2002	87.6338	1.1601
	(0.5,0)	110.833	1.55854	93.4474	1.2146	39.4662	0.4598	113.7020	1.5995	32.7252	0.3704	31.419	0.3507
	(1,0)	26.9036	0.382	28.3946	0.3143	7.9934	0.0475	35.928	0.4325	7.8010	0.0472	7.4796	0.0426
	(1.5,0)	4.9752	0.0642	12.344	40.1065	4.1408	0.0193	14.0698	0.131	4.1204	0.0168	4.2648	0.0146
	(2,0)	1.633	0.0141	6.9978	0.0422	2.6728	0.0116	7.3604	0.0533	2.8626	0.0102	3.1796	0.0074
	(3,0)	1.0066	0.0011	3.759	0.0129	2.0098	0.0014	2.9084	0.0221	2.0356	0.0026	2.199	0.0056
	(4,0)	1.0002	0.0002	2.6888	0.0076	2	0	1.2922	0.0084	2.0002	0.0002	2.001	0.0004
	(0,0.1)	181.382	2.56422	170.106	2.30198	140.342	1.87220	180.325	2.4845	127.2802	1.7112	124.9786	1.6842
	(0,0.2)	95.9090	1.3635	100.2768	1.3602	42.2216	0.5050	118.9156	1.6024	37.6328	0.4512	35.0264	0.4097
	(0,0.3)	32.3598	0.4519	55.293	0.6978	15.6932	0.1582	68.4414	0.9175	14.8360	0.1467	14.5788	0.1346
	(0,0.5)	6.3402	0.0819	19.892	0.2158	6.0012	0.0478	21.9802	0.2564	5.9508	0.0432	6.229	0.0415
	(0,1)	1.783	0.0371	5.5792	0.0398	2.8502	0.0163	4.5752	0.0457	2.9726	0.0168	3.1776	0.0136
	(0,1.5)	1.3526	0.0098	3.5084	0.0206	2.3898	0.0101	2.3382	0.230	2.425	0.0101	2.5804	0.0111
	(0,2)	1.2158	0.0072	2.8394	0.0145	2.2386	0.0075	1.6802	0.0144	2.2488	0.0075	2.3502	0.0089
	(0,3)	1.1248	0.0053	2.3904	0.0095	2.122	0.0052	1.2942	0.0086	2.1454	0.0055	2.1838	0.0063
	(0,4)	1.0844	0.0043	2.2206	0.007	2.786	0.0042	1.1806	0.0065	2.0854	0.0043	2.1168	0.0049
	(0.25,-0.25)	116.892	1.6603	139.1778	1.9283	58.0754	0.7293	155.4454	2.1424	52.6624	0.6726	50.9274	0.64
	(0.5,-0.5)	35.4914	0.4946	84.9876	1.1295	20.3998	0.2233	92.9712	1.2969	19.4598	0.2094	19.3836	0.2005
	(1,-1)	8.7742	0.1148	39.4292	0.4828	8.2898	0.0781	44.6294	0.5703	8.4570	0.0732	9.066	0.0764
	(2,-2)	3.6724	0.0439	20.8964	0.2223	2.9538	0.0185	23.4628	0.2727	3.0408	0.019	5.7336	0.0418

MEWMA control charts through a numerical example. Simulation results demonstrated superiority of the integrated MEWMA and T^2 control chart based on DA in detecting all separate and simultaneous shifts. This method intensified the power of the proposed control chart, rather than the other evaluated control charts. Although the MEWMA control chart outperformed the proposed method in some small shifts, the significant feature of the proposed control chart is its capability in detecting shifts with different directions in small or large shifts in normal and non-normal quality characteristics rather than the MEWMA control chart. The performance of the proposed methods was also evaluated through a real case in the literature. The limitation of the proposed methods, especially MEWMA and T^2 based DA methods, is inability of the
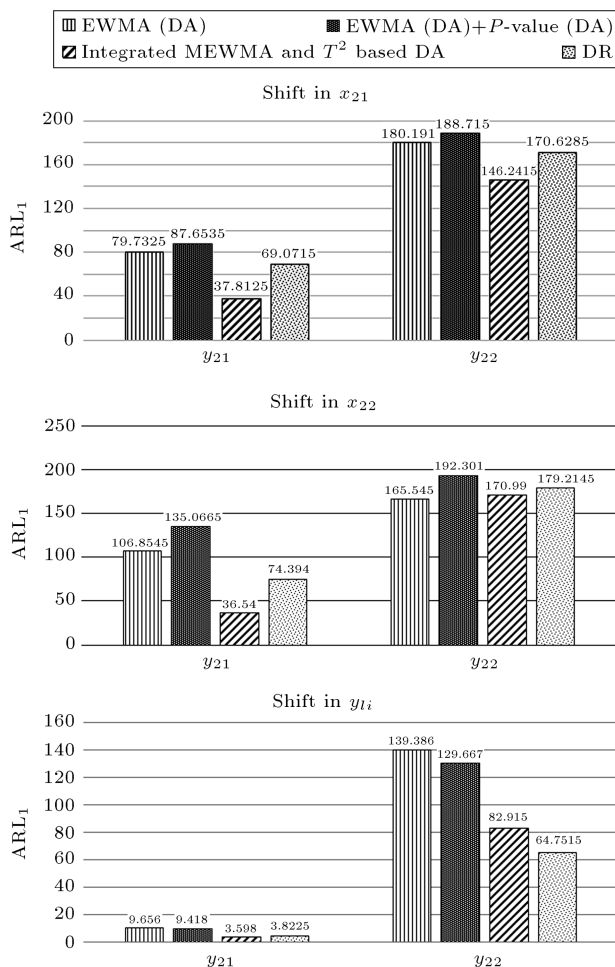
Figure 4. Simulation results of multivariate-attribute real case.

control chart in diagnosing the quality characteristics responsible for the out-of-control signal. Therefore, a new control chart or procedure to enhance this property is suggested for future research. Moreover, extension of the proposed methods in Phase I can be a fruitful area for future research.

Acknowledgment

References

1. Zhang, G.X. "A new type of control charts and theory of diagnosis with control charts", In *ASQC Annual Quality Congress Transactions*, Milwaukee, WI: American Society for Quality Control, pp. 75-85 (1984).

2. Hawkins, D.M. "Multivariate quality control based on regression adjusted variables", *Technometrics*, **33**(1), pp. 61-75 (1991).

3. Hawkins, D.M. "Regression adjustment for variables in multivariate quality control", *J. Qual. Technol.*, **25**(3), pp. 170-182 (1993).

4. Jearkpaporn, D., Montgomery, D.C., Runger, G.C. and Borror, C.M. "Process monitoring for correlated gamma-distributed data using generalized-linear model-based control charts", *Qual. Reliab. Eng. Int.*, **19**(6), pp. 477-491 (2003).

5. Skinner, K.R., Montgomery, D.C. and Runger, G.C. "Process monitoring for multiple count data using generalized-linear model-based control charts", *Int. J. Prod. Res.*, **41**(6), pp. 1167-1180 (2003).

6. Jearkpaporn, D., Borror, C.M., Runger, G.C. and Montgomery, D.C. "Process monitoring for mean shifts for multiple stage processes", *Int. J. Prod. Res.*, **45**(23), pp. 5547-5570 (2007).

7. Aghaie, A., Samimi, Y. and Asadzadeh, S. "Monitoring and diagnosing a two-stage production process with attribute characteristics", *Iran. J. Oper. Res.*, **2**(1), pp.1-16 (2010).

8. Yang, S.F. and Yeh, J.T. "Using cause selecting control charts to monitor dependent process stages with attributes data", *Expert. Syst. Appl.*, **38**(1), pp. 667-672 (2011).

9. Asgari, A., Amiri, A. and Niaki, S.T.A. "A new link function in GLM-based control charts to improve monitoring of two-stage processes with Poisson response", *Int. J. Adv. Manuf. Technol.*, **72**(9-12), pp. 1243-1256 (2014).

10. Asadzadeh, Sh., Zerehsaz, Y., Saghaei, A. and Aghaie A. "Compound-estimator based cause-selecting control chart for monitoring multistage processes", *Commun. Stat.-Simul. C.*, **40**(3), pp. 322-344 (2011).

11. Liu, J. "Variation reduction for multistage manufacturing processes: A comparison survey of statistical process control vs stream of variation methodologies", *Qual. Reliab. Eng. Int.*, **26**(7), pp. 645-661 (2010).

12. Shang, Y., Tsung, F. and Zou, C. "Statistical process control for multistage processes with binary outputs statistical process control for multistage processes with binary outputs", *IIE Trans.*, **45**(9), pp. 1008-1023 (2013).

13. Ghahyazi, M.E., Niaki, S.T.A. and Soleimani, P. "On the monitoring of linear profiles in multistage processes", *Published online in Qual. Reliab. Eng. Int.*, **30**(7), pp. 1035-1047 (2014).

14. Asadzadeh, S. and Aghaie, A. "Improving the product reliability in multistage manufacturing and service operations", *Qual. Reliab. Eng. Int.*, **28**(4), pp. 397-407 (2012).

15. Asadzadeh, S., Aghaie, A. and Shahriari, H. "Using frailty models to account for heterogeneity in mul-

tistage manufacturing and service processes", *Qual. Quant.*, **48**(2), pp. 593-604 (2014).

16. Asadzadeh, S., Aghaie, A. and Niaki, S.T.A. "AFT regression-adjusted monitoring of reliability data in cascade processes", *Qual. Quant.*, **47**(6), pp. 3349-3362 (2013).

17. Asadzadeh, S., Aghaie, A. and Shahriari, H. "Cause-selecting charts based on proportional hazards and binary frailty models", *Int. J. Ind. Eng. Prod. Res.*, **24**(2), pp. 107-112 (2013).

18. Asadzadeh, S., Aghaie, A. and Yang, S.F. "Monitoring and diagnosing multistage processes: A review of cause selecting control charts", *J. Ind. Sys. Eng.*, **2**(3), pp. 214-235 (2008).

19. Li, Y. "Statistical process control of multistage processes with Bayesian sequential bifurcation", *18th Int. Conf. on Ind. Eng. Eng. Manage.*, Changchun, China, pp. 1061-1065 (2011).

20. Li, Y. and Tsung, F. "Detecting and diagnosing covariance matrix changes in multistage processes", *IIE Trans.*, **43**(4), pp. 259-274 (2011).

21. Amiri, A., Asgari, A. and Zerehsaz, Y. "Developing a cause selecting control chart for monitoring two-stage processes with Poisson quality characteristic", *Int. J. Ind. Eng. Prod. Manage.*, **24**(2), pp. 191-202 (2013).

22. Chiu, J. and Kuo, T. "Control charts for fraction nonconforming in a bivariate binomial process", *J. Appl. Stat.*, **37**(10), pp. 1717-1728 (2010).

23. Li, Y. and Tsung, F. "Multiple attribute control charts with false discovery rate control", *Published online in Qual. Reliab. Eng. Int.*, **28**(8), pp. 857-871 (2012).

24. Butte, V.K. and Tang, L.C. "Multivariate charting techniques: A review and a line-column approach", *Qual. Reliab. Eng. Int.*, **26**(5), pp. 443-451 (2010).

25. McCracken, A.K. and Chakraborti, S. "Control chart for joint monitoring of mean and variance: An overview", *Qual. Technol. Quant. Manage.*, **10**(1), pp. 17-36 (2013).

26. Doroudyan, M.H. and Amiri, A. "Monitoring multivariate-attribute processes based on transformation techniques", *Int. J. Adv. Manuf. Technol.*, **69**(9-12), pp. 2161-2172 (2013).

27. Amiri, A., Zou, C. and Doroudyan, M.H. "Monitoring correlated profile and multivariate quality characteristics", *Qual. Reliab. Eng. Int.*, **30**(1), pp. 133-142 (2014).

28. He, Q.P., Joe Qin, S.J. and Wang, J. "A new fault diagnosis method using fault directions in fisher discriminant analysis", *Am. Inst. Chem. Eng. J.*, **51**(2), pp. 555-571 (2005).

29. Pei, X., Yamashita, Y., Yoshidam, M. and Matsumoto, S. "Fault detection in chemical processes using discrim-inant analysis and control chart", *J. Chem. Eng. of Jpn.*, **41**(1), pp. 25-31 (2008).

30. Bazdar, A.A. and Kazemzadeh, R.B. "A novel approach for variation source identification of multi-stage manufacturing processes based on discriminant analysis", *9th Int. Conf. on Ind. Eng.*, Tehran, Iran (2013).

31. Niaki, S.T.A. and Abbasi, B. " Skewness reduction approach in multi-attribute process monitoring", *Commun. Stat. Theor.*, **36**(12), pp. 2313-2325 (2007).

32. Sharma, S., *Applied Multivariate Techniques*, 1st Edn., pp. 237-286, John Wiley & Sons, Inc., New York, USA (1996).

33. Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. "A multivariate exponentially weighted moving average control chart", *Technometrics*, **34**(1), pp. 46-53 (1992).

Appendix A

The mean and variance of ω:

$$E(\omega) = \cos\theta^* E(e_1) + \sin\theta^* E(e_2) = 0, \qquad (A.1)$$

$$\begin{aligned}
var(\omega) &= \cos^2\theta var(e_1) + \sin^2\theta var(e_2) \\
&\quad + 2cov\left(e_1\cos\theta, e_2\sin\theta\right) = \cos^2\theta var(e_1) \\
&\quad + \sin^2\theta var(e_2) + 2\sin\theta\cos\theta cov(e_1, e_2) \\
&= \cos^2\theta var(e_1) + \sin^2\theta var(e_2) \\
&\quad + \sin(2\theta)std(e_1)std(e_1)\rho(e_1, e_2) \\
&= \cos^2\theta var(e_1) + \sin^2\theta var(e_2) \\
&\quad + 2\sin\theta\cos\theta std(e_1)std(e_1) \\
&\quad - 2\sin\theta\cos\theta std(e_1) + \sin(2\theta)std(e_1)std(e_1) \\
\rho(e_1, e_2) &= \left(\cos\theta\sigma_1 + \sin\theta\sigma_2\right)^2 \\
&\quad - 2\sin\theta\cos\theta\sigma_1\sigma_2\left(1 - \rho(e_1, e_2)\right). \qquad (A.2)
\end{aligned}$$

Appendix B

Distribution of quality characteristics and parameters in the link functions in the real case:

First stage:

$$y_{11} \sim N\left(\mu = 1 + x_{11} + x_{12}, \sigma^2 = 0.01\right), \qquad (B.1)$$

$$y_{12} \sim \text{Poisson}\left(\mu = \exp(-1.11 + 2.3x_{11} + 2x_{12})\right). \qquad (B.2)$$

Second stage:

$$y_{21} \sim N\left(\mu = x_{21}, x_{22} + y_{11}, \sigma^2 = 0.01\right), \qquad \text{(B.3)}$$

$$y_{22} \sim \text{Poisson}\left(\mu_{y_{22}} = \exp(-0.35 + 2x_{21} - 2x_{22}\right.$$
$$\left. + \ln(y_{12}))\right) \text{ with the mean of } 4.00. \qquad \text{(B.4)}$$

Appendix C

The algorithm of the bisection method used to find the value which makes the skewness function equal to zero:

$k = 0$

while $|f(x_{k+1})| > \varepsilon$

$$x_{k+1} = \frac{a_k + b_k}{2}$$

If $(f(x_{k+1})f(a_k) < 0)$,

Then

$a_{k+1} = a_k$ and $b_{k+1} = x_{k+1}$,

Else

$b_{k+1} = b_k$ and $a_{k+1} = x_k$

End if

$k = k + 1$

End while $\qquad \text{(C.1)}$

ε is a very small value. The above algorithm is general form, but in skewness reduction, $f(r)$ equal to zero is applied to find r in the interval of $(0,1)$ (r is the power of non-zero root skewed data).

Biographies

Samaneh Zolfaghari holds MSc degree in Industrial Engineering at Shahed University of Iran. She received her BS in Industrial Engineering from Azad University, South Tehran Branch in 2011. Her research interests include statistical quality control and multistage processes.

Amirhossein Amiri is an Associate Professor at Shahed University in Iran. He holds a BS, an MS, and a PhD in Industrial Engineering from Khajeh Nasir University of Technology, Iran University of Science and Technology, and Tarbiat Modares University, respectively. He is a member of the Iranian Statistical Association. His research interests are statistical quality control, profile monitoring, and Six Sigma.