Journal of Quality Engineering and Production Optimization

Vol. 7, No. 2, Winter & Spring 2022

http://jqepo.shahed.ac.ir

Research Paper

DOI: 10.22070/JQEPO.2022.5744.1167

Estimating step and linear drift change point in contingency tables

Mona Ayoubi[1*], Maedeh Ebadi[1]

[1] *Department of Industrial Engineering, College of Engineering, West Tehran Branch, Islamic Azad University, Tehran, Iran*

**** Corresponding Author:** Mona Ayoubi (Email: ayoubi.m@wtiau.ac.ir)*

***Abstract** – In many practical cases, product or process quality is defined by frequency table of two or more qualitative variables. This frequency table is called contingency table. Monitoring the contingency tables is an area in statistical process control with many applications in industrial and service units.*
On the other hand, reducing quality costs is the most fundamental issue preoccupying the minds of managers. It is clear that a quicker diagnosis of the assignable causes can reduce the quality costs. Estimating change point by limiting the probable interval of change, reduces the cost and time of detecting assignable causes. In this research, using maximum likelihood approach, the step and linear drift change points estimators are proposed for multivariate multi-nominal contingency tables. After the change point, parameters are estimated with making the average in the proposed step estimator, and using the linear regression in the proposed linear drift estimator. Results of the simulations demonstrated that the proposed step change point estimator carries out very well in all shift types and shift magnitudes from small to large. Furthermore, the proposed estimator of the linear drift change point has relatively good performance in moderate changes. Finally, the proposed estimators' performance is assessed by a numerical example.

***Keywords**– change point estimation, contingency table, statistical process control, step and linear drift change, maximum likelihood approach.*

I. INTRODUCTION

In some practical situations, the quality of a product or a process is defined with frequency table between two or more qualitative variables entitled contingency table. Such tables are expressed as r×c tables in some cases, in which *r* and *c* stand for the number of rows and columns, respectively. Monitoring contingency table evaluates the equality of sample proportions at column levels for all societies (i.e., all levels of the row variable). As there are many cases in practice where qualitative characteristics are defined as multivariate discrete variables in contingency table, some approaches have been presented for monitoring contingency tables in the statistical process control literature. In order to monitor the contingency tables, the stability of the multivariate discrete variables in each sample of the process should be examined. In general, monitoring a process can be considered in Phase I and Phase II. Agresti (2002) has provided some information about the contingency table and its structure. In statistical process control (SPC), contingency tables are used in monitoring multivariate multinomial processes, simultaneously. Yashchin (2012) applied contingency table for monitoring production line air-quality control with two multinomial variables named particle types and an air filtration system. Processes with contingency table quality characteristics, can be formulated with linear logarithmic model. For example,

Manuscript Received: 20- July -2022 & Revised: 3- Sept -2022 & Accepted: 29- Oct -2022
ISSN:2423-3781

Zou and Tsung (2011) used Exponentially Weighted Moving Average (EWMA) control chart for monitoring the model parameters. Besides, Yashchin (2012) utilized the likelihood ratio test (LRT) for developing control chart to monitor multinomial observations. Furthermore, he has conversed about the parameter estimation of multivariate data with a sudden change. Li et al. (2012) presented a LRT to monitor contingency table via multivariate binomial/multinomial distribution through linear logarithmic model in Phase II. They used Generalized LRT (GLRT) and EWMA statistics so that the GLRT control chart was strengthened under small parameter changes in linear logarithmic model. Also, Li et al. (2013) suggested a multivariate non-parametric control chart for monitoring model parameters by integrating multivariate and EWMA control charts. A control chart for monitoring a multivariate multinomial / binomial process was introduced by Li et al. (2014). They provided linear logarithmic model to show the relationships between multinomial variables following multivariate multinomial / binomial distributions.

Kamranrad et al. (2017a) suggested two generalized linear test (GLT) and EWMA-GLT control charts based on GLT to monitor contingency tables. Besides, a diagnostic approach was proposed to the parameters that triggered out-of-control alarms. Kamranrad et al. (2017b) proposed two control charts using WALD and STUART statistics to monitor contingency tables in Phase II. Also, they proposed EWMA-WALD and EWMA-STUART control charts to improve the efficiency of Shewhart-type control charts for identifying small and medium changes in the parameters of contingency tables. Kamranrad et al. (2019) introdused three methods including Hoteling's T^2, standardized LRT (SLRT), and F for monitoring ordinal multivariate contingency table in Phase I. Furthermore, they developed an estimator of the step change point using SLRT in Phase I. Perry (2020) also proposed an EWMA control chart in order to categorical process monitoring. Hakimi et al. (2019) suggested multivariate ordinal-normal statistic to design their new control chart for monitoring ordinal contingency tables in Phase II. Bersimis and Sachlas (2019) introduced one-sided approach for monitoring contingency tables. Xiang et al. (2021) considered monitoring sparse contingency table of a multivariate categorical process.

An issue in SPC is that the time at which the control charts indicate an out of control signal, T, is not the real time of change and change in the process has occurred before out of control signal. The unknown real time of change is named the change point, τ. Estimating change point through limiting the probable change period makes it easier to search for the assignable causes and speeds it up. Searching for assignable causes requires a lot of time and cost. If quality control team spend less time to search for assignable causes, the relevant cost will reduce. Also, by estimating change point and finding assignable causes sooner, the out-of-control process comebacks to the in-control situation sooner. Hence, less defective products are produced in the out-of-control situation leading to less cost of product reworks and rejects. Hence, Thus, estimating the change point has a significant effect on reducing quality cost. In the literature of change point estimation, there are various methods for estimation of change point. Overall, the approaches used to find change points so far are the internal estimators approach of EWMA and CUSUM control charts, the maximum likelihood approach, clustering and Artificial Neural Networks (ANNs). Amiri and Allahyari (2012) made a comprehensive review of the literature on the field of change point estimation. Page (1954) has defined the internal estimator of CUSUM control chart. Nishina (1992) proposed an internal EWMA estimator to estimate the change point. Using the maximum likelihood Estimation (MLE) approach is a common approach for estimating change point, and many researchers use this approach. Given the existence of step change in process parameters, extensive studies have been done in the literature. For instance, Samuel et al. (1998), Samuel and Pignatiello (1998), Pignatiello and Samuel (2001), Perry and Pignatiello (2010), Noorossana et al. (2009), Nedumaran et al. (2000), Perry and Pignatiello (2011), Niaki and Khedmati (2014a), Steward and Rigdon (2017) estimated the step change point using MLE method in some processes.

In addition, considering a linear drift change in process parameters using maximum likelihood method, Perry and Pignatiello (2006), Perry et al. (2006), Fahmy and Elsayed (2006), Kazemzadeh et al. (2015) carried out some researches. Amiri and Khosravi (2013), Perry et al. (2007), Noorossana and Shadman (2009), Niaki and Khedmati (2014b), Movafagh and Amiri (2013), Ashuri and Amiri (2016) considered monotonic change point estimation. Venegas et al. (2016) also focused in change point estimation.

Some researchers used the clustering approach for estimating the change point. For exapmle, Ghazanfari et al. (2008), Alaeddini et al. (2009), Zarandi and Alaeddini (2010), Kazemi et al. (2014) suggested some approaches based on clustering methods. Kazemzadeh et al. (2015), Ayoubi et al. (2014) and Ayoubi et al. (2016) considered estimating change point in multivariate linear profiles' mean.Via artificial neural networks, Atashgar and Noorossana (2011), Noorossana et al. (2011) and Atashgar and Noorossana (2012), Amiri et al. (2015), Yeganeh et al. (2021) and Ghazizadeh et al. (2021) considered estimating the change point. He et al. (2019) also focused on detecting change point in multivariate categorical processes based on the hierarchical log-linear model. Maleki et al. (2018) estimated change point of poisson profiles in the presence of autocorrelation. Dette, et al. (2022) Estimated change point of high-dimensional covariance matrices.

In this paper, step and drift change points in contingency tables are estimated using maximum likelihood approach. After the step change point, the unknown parameters are also estimated by averaging the calculated probability of each cell, and the unknown shift slopes after the drift change point are approximated using forced regression to pass the origin of the axis of coordinates. The structure of the paper is such a way that in section II the contingency tables model based on the probability of each cell is described. The applied monitoring approach of EWMA-WALD control chart is explained in section III. The proposed step change point estimator is derived in section IV. Section V presents the calculations of the proposed drift change point estimator. Section VI shows the simulation results of performance evaluation. Confidence set of the proposed estimators is constructed in section VII. Section VIII reports the consequences of a numerical example. The last section consists of concluding remarks.

II. UNDERLYING MODEL OF CONTINGENCY TABLE BY USING THE PROBABILITY OF EACH CELL

The contingency table is one of the most efficient tools to analyze multivariate multinomial categorical variables. According to the number of categorical variables, two broad categories are defined which are Tow-Way Contingency Table (with two categorical variables), and Multi-Way Contingency Table (with more than two) categorical variables, Agresti (2002). Now considering two-way contingency table with I and J classes for both row and column variables, Kamranrad et al. (2017) defined the contingency table parameters as in Table I. When $\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij} = N$, and $\sum_{i=1}^{I}\sum_{j=1}^{J} \pi_{ij} = 1$. The observations of n_{ij} follow multivariate multinomial distribution with sample size N and probability vector of π.

TABEL I. NOMENCLATURE OF CONTINGENCY TABLE PARAMETERS

N	***sample size***	$\pi_{+j} = \frac{n_{+j}}{N}$	***probability of jth column***
n_{ij}	Number of observations in cell (i, j) for $i = 1,2,\dots,I,\quad j = 1,2,\dots,J$	$\hat{\pi}_{ij} = \frac{\hat{n}_{ij}}{N}$	estimated probability of (i,j)th cell
$n_{i+} = \sum_j n_{ij}$	sample size in ith row $\forall_i = 1\,.\,2\,.\,\dots.\,I$	$\hat{\pi}_{i+} = \frac{\hat{n}_{i+}}{N}$	estimated probability of ith row
$n_{+j} = \sum_i n_{ij}$	sample size in jth column $\forall_j = 1\,.\,2\,.\,\dots.J$	$\hat{\pi}_{+j} = \frac{\hat{n}_{+j}}{N}$	estimated probability of jth column
π	probabilities vector of contingency table	$\mu_{ij} = N\pi_{ij}$	expected observation value for (i,j)th cell
$\pi_{ij} = \frac{n_{ij}}{N}$	probability of (i,j)th cell	$\mu_{i+} = N\pi_{i+}$	expected observation value for ith row
$\pi_{i+} = \frac{n_{i+}}{N}$	probability of ith row	$\mu_{+j} = N\pi_{+j}$	expected observation value for jth column,

III. MONITORING APPROACH: EWMA-WALD CONTROL CHART

Kamranrad et al. (2017b) introduced the WALD statistic for monitoring contingency tables in Phase II. For a contingency table with two categorical variables in the row and column with respectively i ($i = 1,2,\ldots,I$) and j ($j = 1,2,\ldots,J$) levels, the WALD statistic is defined as follows:

$$W = Nd'\hat{V}^{-1}d. \tag{1}$$

Thus, W has a chi-square distribution with $I-1$ degrees of freedom. In addition, $d' = (d_1,\ldots,d_{I-1})$ whose values are calculated as follows:

$$d_i = \pi_{+i} - \pi_{i+} \tag{2}$$

where, and $\hat{V}$ is a variance-covariance matrix whose entries can be calculated as follows:

$$\begin{aligned}\hat{V}_{ij} &= -(\pi_{ij} + \pi_{ji}) - (\pi_{+i} - \pi_{i+})(\pi_{+j} - \pi_{j+}), \\ \hat{V}_{ij} &= \pi_{+i} + \pi_{i+} - 2\pi_{ii} - (\pi_{+i} - \pi_{i+})^2,\end{aligned} \qquad \text{i} \neq \text{j , i=j} \tag{3}$$

The main problem in this statistic is that the sample variance-covariance matrix inverse cannot be calculated because its determinant is near zero. Thus, Singular Value Decomposition (SVD) algorithm used by Kamranrad et al. (2017b), is considered in this paper. EWMA-WALD statistic is calculated using the following equation:

$$Z_t = \lambda W_t + (1-\lambda)Z_{t-1} \tag{4}$$

where at time t, the WALD statistic is shown by $W_t.\lambda$ is smoothing constant. Z_0 is the initial value of EWMA which is set equal to WALD average (as WALD has chi square distribution with I-1 degrees of freedom, the average of this statistic is I-1 and the value $Z_0 = I-1$ is considered.)

EWMA-WALD statistical control limit can be also calculated as follows:

$$\begin{aligned}UCL_{EWMA-WALD} &= \mu_w + L.\sigma_w.\sqrt{\tfrac{\lambda}{2-\lambda}} \\ CL_{EWMA-WALD} &= \mu_w \\ LCL_{EWMA-WALD} &= \mu_w - L.\sigma_w.\sqrt{\tfrac{\lambda}{2-\lambda}}.\end{aligned} \tag{5}$$

Here, L stands for the coefficient of the control limit and is set to obtain a desired value of the in-control average run length (ARL). In addition, $\mu_w$is the mean and σ_w is the standard deviation of the WALD statistic, which follows chi square distribution and it has $(I-1)$ degrees of freedom; in other words, $\mu_w = I-1$ and $\sigma_w = \sqrt{2(I-1)}$

IV. THE PROPOSED STEP CHANGE POINT MAXIMUM LIKELIHOOD ESTIMATOR

As the focus of this study is on the estimate of the change point in Phase II, it is assumed that the in-control vector of π_0 is definite and known. In order to obtain the proposed maximum likelihood estimator, it is also assumed that before the actual change point (τ) the process is in-control and there is no changes in the probability vector (so, $\pi = \pi_0$), but after the change point, the step change in the parameters is created according to the following equation:

$$\pi_1 = \pi_0 + s \tag{6}$$

Here, s is the vector whose elements show the magnitudes of step shifts in model parameters. s has the identical dimension to π_0. The density function of a multivariate multinomial distribution is as follows (for two variables with levels I and J):

$$f(n_{ij}) = \frac{N!}{\prod_{i=1}^{I}\prod_{j=1}^{J} n_{ij}!}\prod_{i=1}^{I}\prod_{j=1}^{J}\hat{\pi}_{ij}^{n_{ij}}. \tag{7}$$

Hence, the likelihood function can be defined as follows: (assuming there are two variables which have I and J levels, respectively)

$$L = \prod_{k=1}^{\tau}\frac{N!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\boldsymbol{n}_{ij}!}\prod_{i=1}^{I}\prod_{j=1}^{J}(\boldsymbol{\pi}_0)_{ij}^{\mathbf{n}_{ij}} \times \prod_{k=\tau+1}^{T}\frac{N!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\boldsymbol{n}_{ij}!}\prod_{i=1}^{I}\prod_{j=1}^{J}(\boldsymbol{\pi}_1)_{ij}^{\mathbf{n}_{ij}} \quad , \tag{8}$$

where, $\sum_{i=1}^{I}\sum_{j=1}^{J}\pi_{ij} = 1$and also at sample T , the out of control signal is triggered. In addition, k is the sample number. After taking natural logarithm, the following equation is obtained:

$$ln\, L = \sum_{k=1}^{\tau} ln\, N\,! - \sum_{k=1}^{\tau}\sum_{u=1}^{IJ} ln(n_{ku}!) + \sum_{k=1}^{\tau}\sum_{u=1}^{IJ} n_{ku}\, ln(\pi_0)_{ku} + \sum_{k=\tau+1}^{T} ln\, N\,! - \sum_{k=\tau+1}^{T}\sum_{u=1}^{IJ} ln(n!_{ku}) + \sum_{k=\tau+1}^{T}\sum_{u=1}^{IJ} n_{ku}\, ln(\pi_1)_{ku} \tag{9}$$

In the above equation, $\sum_{k=1}^{\tau} ln\, N\,! + \sum_{k=\tau+1}^{T} ln\, N\,!$ equals $\sum_{k=1}^{T} ln\, N\,!$ is a constant number, and because of its large quantity, it is eliminated for better results. vector of $\boldsymbol{\pi}_0$ is known in Phase II, but vector of $\boldsymbol{\pi}_1$ have to be estimated. Hence, the proposed estimator of the step change point is as the following equation:

$$\hat{\tau}_{Step\quad change} = \underset{0\le t\le T-1}{arg\; max}\{ - \sum_{k=1}^{t}\sum_{u=1}^{IJ} ln(n_{ku}!) + \sum_{k=1}^{t}\sum_{u=1}^{IJ} n_{ku}\, ln(\pi_0)_{ku} - \sum_{k=t+1}^{T}\sum_{u=1}^{IJ} ln(n!_{ku}) + \sum_{k=t+1}^{T}\sum_{u=1}^{IJ} n_{ku}\, ln(\hat{\pi}_1)_{ku}\}, \tag{10}$$

To estimate vector of $\hat{\pi}_1$, we use the averaging method using all samples after the change point. For this purpose, Probability vector of sample k is first obtained using $\hat{\pi}_k = \frac{n_k}{N}$, and then the vector of $\hat{\pi}_1$ is calculated by averaging the calculated vectors using $\hat{\boldsymbol{\pi}}_1 = \frac{\sum_{k=t+1}^{\mathrm{T}}\hat{\boldsymbol{\pi}}_k}{\mathrm{T}-t}$.

V. THE PROPOSED LINEAR DRIFT CHANGE POINT MAXIMUM LIKELIHOOD ESTIMATOR

As the focus of this paper is on estimating the change point in Phase II, it is assumed that vector of $\boldsymbol{\pi}_0$ has known elements. In order to obtain linear drift maximum likelihood estimator, it is assumed that before the actual change point of τ, the process is in-control and there is no changes in the probability vector (thus, $\pi_k = \pi_0$), but after the change point, the linear drift shift is exposed to the process parameters according to the following equation:

$$\pi_k = \pi_0 + (k-\tau)\beta, \pi_k = \pi_0 + (k-\tau)\beta, \tag{11}$$

Here, β is a vector whose elements show the slope of linear drift changes in its corresponding parameters of π_0. Using the density function of Equation (7), the likelihood function with linear drift change in multivariate multinomial contingency tables (assuming there are two variables with I and J levels) is defined as:

$$L=\prod_{k=1}^{\tau}\frac{N!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\boldsymbol{n}_{ij}!}\prod_{i=1}^{I}\prod_{j=1}^{J}(\boldsymbol{\pi}_0)_{ij}^{\mathbf{n}_{ij}} \times\prod_{k=\tau+1}^{T}\frac{N!}{\prod_{i=1}^{I}\prod_{j=1}^{J}\boldsymbol{n}_{ij}!}\prod_{i=1}^{I}\prod_{j=1}^{J}\left(\boldsymbol{\pi}_0+\boldsymbol{\beta}(k-\tau)\right)_{ij}^{\mathbf{n}_{ij}}, \tag{12}$$

where, $\sum_{i=1}^{I}\sum_{j=1}^{J}\pi_{ij}=1$, and T is the time of out of control signal time of the EWMA-WALD control chart and τ is the change point. In addition, k is the sample number.

After taking natural logarithm, the following equation is derived:

$$ln\,L=\sum_{k=1}^{\tau} ln\,N\,!-\sum_{k=1}^{\tau}\sum_{u=1}^{IJ} ln(n_{ku}!)+\sum_{k=1}^{\tau}\sum_{u=1}^{IJ} n_{ku}\,ln(\pi_0)_{ku} \sum_{k=\tau+1}^{T} ln\,N+!-\sum_{k=\tau+1}^{T}\sum_{u=1}^{IJ} ln(n!_{ku})\quad+\sum_{k=\tau+1}^{T}\sum_{u=1}^{IJ} n_{ku}\,ln\big(\pi_0+\beta(k-\tau)\big)_{ku}, \tag{13}$$

In the above equation, $\sum_{k=1}^{t} ln\,N\,!+\sum_{k=t+1}^{T} ln\,N\,!$ equals $\sum_{k=1}^{T} ln\,N\,!$ is constant, and because of its large value, it is eliminated for better results. In the above equation, β is an unknown vector and have to be estimated. Hence, the proposed estimator of the linear drift change point is derived as:

$$\hat{\tau}_{Drift\quad change}=\underset{0\le t\le T-1}{arg}\quad max\{-\sum_{k=1}^{t}\sum_{u=1}^{IJ} ln(n_{ku}!)+\sum_{k=1}^{t}\sum_{u=1}^{IJ} n_{ku}\,ln(\pi_0)_{ku} -\sum_{k=t+1}^{T}\sum_{u=1}^{IJ} ln(n!_{ku})+\sum_{k=t+1}^{T}\sum_{u=1}^{IJ} n_{ku}\,ln\left(\pi_0+\hat{\beta}(k-t)\right)_{ku}\}. \tag{14}$$

To estimate the slope vector of β, we use the forced regression to pass the origin of the axis of coordinates using calculated probability vectors of all samples after the change point, i.e. $\hat{\pi}_k=\frac{n_k}{N}$. Then considering the linear drift model of $\hat{\pi}_k=\pi_0+\beta(k-\tau)$, one can obtain $\hat{\pi}_k-\pi_0=\beta(k-\tau)$. So, Considering $(k-\tau)$ as an independent variable and $\hat{\pi}_k-\pi_0$ as a dependent variable, the vector of $\hat{\beta}$ is calculated as follows:

$$\hat{\boldsymbol{\beta}}=\frac{\sum x_i y_i}{\sum x_i^2}=\frac{\sum_{\mathrm{k}=t+1}^{\mathrm{T}}(\hat{\boldsymbol{\pi}}_{\mathrm{k}}-\boldsymbol{\pi}_0)\times(\mathrm{k}-\mathrm{t})}{\sum_{\mathrm{k}=t+1}^{\mathrm{T}}(\mathrm{k}-\mathrm{t})^2}. \tag{15}$$

VI. PERFORMANCE EVALUATION OF THE PROPOSED STEP AND LINEAR DRIFT ESTIMATORS OF THE CHANGE POINT

In this section, proposed estimators' performance is assessed using 5000 simulation runs. In each iteration, when the EWMA-WALD control chart issues an out-of-control warning, the corresponding estimator of the change point is used. Value of $\lambda=0.2$ is used for the simulations. Ayoubi et al. (2014) investigated the smoothing constant effect on the change point estimators. Generally, while λ increases, the proposed change point estimators' performance becomes worse in small shifts and better in large shifts. Also, when λ decreases, change point estimators performs better in small shifts and also worse in large shifts.

In addition, the in-control ARL of 200 is considered for the EWMA-WALD control chart. In doing so, the value of 4.075 is chosen for L to reach the in-control ARL of 200. In order to deal with the false alarms, as EWMA-WALD statistics consist of the previous samples information, it is not possible to delete the only last sample in which the control chart falsely issues an alarm. Thus, all samples up to the false alarm are removed, also the first sample after the false alarm is then fixed as the process first sample. Hence, the change point changes in each iteration. For instance, it is assumed in this paper that $\tau=25$. Hence, 25 in-control samples are generated in each iteration. If false alarm occurs at sample 15, so the first 15 samples are deleted and sample 16 is considered as the first sample of the process. If we do not have false alarms in the rest of the samples, 10 other samples are generated $\tau = 10$ is considered for the corresponding iteration. The underlying in-control table used by Kamranrad et al. (2017b) is shown in Table II.

Table II. Assumed in-control contingency table

x_1	x_2			
	1	*2*	*3*	*4*
1	65	20	39	20
2	29	67	28	21
3	36	35	10	30
4	12	32	27	33

The simulation results are summarized in Tables III and IV for the step estimator, and in Tables V and VI for the drift estimator. In the accuracy sections, the mean of square error of estimates is reported in parenthesis.

For monitoring contingency tables, sum of all elements in the vector of π must be equal to one due to the condition of $\sum_{i=1}^{I}\sum_{j=1}^{J}\pi_{ij} = 1$. Hence, to impose shifts to the parameters, increasing and decreasing shifts with the same size must be occurred together to satisfy $\sum_{i=1}^{I}\sum_{j=1}^{J}\pi_{ij} = 1$. For example in Table III, the element of π_{11}has decreasing shift of $\pi_{11} - s_{11}$and the element of π_{34} has increasing shift of $\pi_{34} + s_{34}$ in which $s_{11} = s_{34}$, i.e. if an increasing shift occurs in a parameter, a decreasing shift with the same size must be occurred in other parameters.

Table III. Proposed step estimator's accuracy and precision under step shifts in $\pi_{11} - s_{11}$ and $\pi_{34} + s_{34}$.

	$s_{11} = s_{34}$	*0.01*	*0.02*	*0.03*	*0.04*	*0.05*	*0.07*	*0.09*	*0.1*
Accuracy Performances	$\widehat{ARL}$	72.776	24.5236	10.3354	5.9052	3.837	2.237	1.6026	1.417
	$\bar{\hat{\tau}}_{step}$	25.3418 (35.5222)	24.8428 (4.9684)	24.9078 (1.507)	24.934 (0.724)	24.9758 (0.2022)	24.996 (0.0048)	25 (0)	25 (0)
Precision Performances	$\hat{p}(\lvert\hat{\tau} - \tau\rvert = 0)$	0.2468	0.604	0.819	0.933	0.9762	0.9964	1	1
	$\hat{p}(\lvert\hat{\tau} - \tau\rvert \le 1)$	0.4648	0.8364	0.9542	0.9842	0.9978	0.9996	1	1
	$\hat{p}(\lvert\hat{\tau} - \tau\rvert \le 3)$	0.681	0.9496	0.99	0.9956	0.9988	1	1	1
	$\hat{p}(\lvert\hat{\tau} - \tau\rvert \le 5)$	0.7914	0.976	0.9944	0.9964	0.9992	1	1	1
	$\hat{p}(\lvert\hat{\tau} - \tau\rvert \le 7)$	0.8604	0.987	0.9962	0.9978	0.9992	1	1	1
	$\hat{p}(\lvert\hat{\tau} - \tau\rvert \le 10)$	0.9134	0.9918	0.9972	0.9988	0.9992	1	1	1

Table IV. Proposed step estimator's accuracy and precision under step shifts in $\pi_{22} - s_{22}$ and$\pi_{34} + s_{34}$.

	$s_{22} = s_{34}$	*0.01*	*0.02*	*0.03*	*0.04*	*0.05*	*0.07*	*0.09*	*0.1*
Accuracy Performances	$\widehat{ARL}$	73.788	24.0236	10.2288	5.7458	3.8812	2.231	1.5998	1.4126
	$\bar{\hat{\tau}}_{step}$	25.438 (37.5796)	24.8608 (4.6972)	24.9104 (1.1872)	24.9386 (0.5494)	24.9732 (0.2328)	24.9972 (0.0032)	25 (0)	24.9998 (0.0002)

Precision Performances	$\hat{p}(\lvert\hat{\tau}-\tau\rvert=0)$	0.2438	0.5946	0.8152	0.9228	0.977	0.9974	1	0.9998
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 1)$	0.4516	0.8302	0.9544	0.9828	0.9944	0.9998	1	1
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 3)$	0.6678	0.949	0.9912	0.9956	0.9988	1	1	1
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 5)$	0.7762	0.9798	0.9952	0.997	0.9994	1	1	1
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 7)$	0.8452	0.9884	0.9966	0.998	0.9996	1	1	1
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 10)$	0.9078	0.9928	0.998	0.9992	0.9996	1	1	1

The results of Tables III and IV show that estimating the proposed step change point in all small to large shifts has a very good performance in terms of both accuracy and precision of estimates. Its performance also improves with increasing the shift size.

Table V. Proposed drift estimator's accuracy and precision under drift shifts in $\pi_{11}-\beta_{11}(k-\tau)$ and $\pi_{34}+\beta_{34}(k-\tau)$.

	$\beta_{11}=\beta_{34}$	*0.001*	*0.002*	*0.003*	*0.004*	*0.005*	*0.007*	*0.009*	*0.015*
Accuracy Performances	$\widehat{ARL}$	26.1064	16.5814	12.817	10.5946	9.0564	7.3634	6.2174	4.4826
	$\bar{\hat{\tau}}_{drift}$	46.5494 (555.1494)	37.642 (187.7084)	34.181 (99.8818)	31.9878 (58.6446)	30.8662 (41.9782)	28.8562 (24.6386)	27.4056 (28.1412)	22.716 (97.6584)
Precision Performances	$\hat{p}(\lvert\hat{\tau}-\tau\rvert=0)$	0.0046	0.0072	0.0084	0.0166	0.0152	0.0322	0.0594	0.1552
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 1)$	0.0174	0.0206	0.0312	0.04	0.0532	0.1142	0.2018	0.416
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 3)$	0.0456	0.0538	0.0812	0.1232	0.1782	0.4192	0.5906	0.7068
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 5)$	0.0698	0.0938	0.1634	0.3	0.4382	0.7432	0.8542	0.8406
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 7)$	0.0912	0.16	0.314	0.5632	0.729	0.9328	0.9562	0.848
	$\hat{p}(\lvert\hat{\tau}-\tau\rvert\leq 10)$	0.1326	0.3212	0.6184	0.8754	0.9572	0.993	0.9758	0.849

Table VI. Proposed DRIFT ESTIMATOR'S ACCURACY AND PRECISION UNDER DRIFT SHIFTS IN $\pi_{22}-\beta_{22}(k-\tau)$ AND $\pi_{34}+\beta_{34}(k-\tau)$.

	$\beta_{22}=\beta_{34}$	*0.001*	*0.002*	*0.003*	*0.004*	*0.005*	*0.007*	*0.009*	*0.015*
Accuracy Performances	$\widehat{ARL}$	25.8772	16.6722	12.778	10.5546	9.1254	7.2762	6.2174	4.4718
	$\bar{\hat{\tau}}_{drift}$	46.1484 (538.9512)	37.717 (190.0382)	34.2384 (100.5716)	32.0476 (59.7896)	30.7274 (40.2886)	29.3344 (26.3272)	28.0334 (24.2122)	23.768 (72.4424)
Precision Performances	$\hat{p}(\|\hat{\tau}-\tau\|=0)$	0.0068	0.0084	0.0076	0.0158	0.0164	0.026	0.0364	0.168
	$\hat{p}(\|\hat{\tau}-\tau\|\leq 1)$	0.0192	0.0244	0.028	0.0414	0.0542	0.087	0.1398	0.45
	$\hat{p}(\|\hat{\tau}-\tau\|\leq 3)$	0.0464	0.0564	0.0808	0.1284	0.1946	0.3342	0.5198	0.7546
	$\hat{p}(\|\hat{\tau}-\tau\|\leq 5)$	0.0704	0.0934	0.1596	0.2948	0.4564	0.6902	0.8396	0.884
	$\hat{p}(\|\hat{\tau}-\tau\|\leq 7)$	0.0928	0.1552	0.3052	0.5414	0.7472	0.9136	0.9666	0.8898
	$\hat{p}(\|\hat{\tau}-\tau\|\leq 10)$	0.1388	0.304	0.6126	0.8718	0.9622	0.9958	0.9856	0.8898

The results of Tables V and VI indicate that linear drift change point estimator has relatively good performance considering estimations accuracy and precision in moderate shift sizes. Nevertheless, in the case of very small and very large shifts, its performance is weak. Overall, the estimated performance of the proposed step estimator for the step change point estimation is better than the performance of the proposed linear drift estimator under the drift shifts. Simulation results confirm that proposed estimators make it easier to search for the assignable causes and speed it up. for example, consider the shift size of 0.01 in Table III. The actual change point is $\tau=25$but the EWMA-WALD control chart issues an out-of-control warning averagely at the time of 72.776+25=97.776 which is far from the actual change point. The proposed step estimator also identifies the change point averagely at 25.3418 which is close to the true change point of 25 leading to find assignable causes sooner, because practitioners search for assignable causes only around the estimated change point.

Since, $\bar{\hat{\tau}}$ of the proposed estimators is closer to the true change point of 25 than average signal time of the control chart, i.e. $E(T)=\widehat{ARL}+25$, we can conclude that our proposed estimators have an acceptable performance in detecting change point.

VII. CARDINALITY AND COVERAGE PROBABILITY OF CONFIDENCE SETS

Perry et al. (2006) recommended calculating cardinality and coverage probability of the change point estimators by constructing confidence set that helps quality engineers to start the search for assignable causes by a set of probable change points.

Confidence sets for the proposed change point estimators satisfying Equation (16), are constructed in this section. Note that the natural logarithms of the likelihood function for the proposed drift estimator are negative, hence the absolute value of the $Ln \quad L(\hat{\tau}) - Ln \quad L(t)$is considered as follows:

$$Confidence \quad set = \{t: \quad |Ln \quad L(\hat{\tau}) - Ln \quad L(t)| < D\} \tag{16}$$

In the above equation, $Ln \quad L(\hat{\tau})$ is the maximum value of the natural logarithms of the likelihood function over all possible change point *t*. Cardinality is defined as the number of elements in confidence set if the confidence set contains the actual change point. The simulation is repeated 5000 times and the mean cardinality and coverage probability are calculated. Coverage probability is the probability that the actual change point is in the confidence set. The values of *D* are calculated using try and error to obtain coverage probability between all range of 0 to 1. If *D* is chosen less, coverage probability is near zero and if *D* is chosen more, the coverage probability is near one. Tables VII and VIII report the coverage probability and mean cardinality of the proposed step and drift change point estimators, respectively.

For example, Table VII shows that if quality practitioners want to have maximum mean cardinality of almost 15 and coverage probability more than 0.8 to detect small shift size of 0.01, they should choose the value of *D*=5 for the proposed estimator of the step change point. Also with the value of *D*=55, the coverage probability for all shift sizes is more than 0.9 that is because of increase in mean cardinality.

Results of Table VIII for the proposed drift change point estimator is also the same as Table VII. It is clear that for each shift size, the coverage probability increases when mean cardinality increases.

Table VII. Coverage probability and mean cardinality of the proposed step estimator under step shifts in $\pi_{11} - s_{11}$ and $\pi_{34} + s_{34}$.

$s_{11} = s_{34}$						
0.1	0.07	0.05	0.03	0.01	D	
0.0010	0.0038	0.0174	0.1018	0.284	1	***Coverage probability***
0.0016	0.0198	0.1154	0.4772	0.8502	5	
0.0022	0.0952	0.4138	0.8462	0.9434	10	
0.0136	0.2938	0.7322	0.899	0.9748	15	
0.1486	0.768	0.8966	0.9064	0.9818	25	
0.5268	0.886	0.897	0.9104	0.9818	35	
0.7884	0.9006	0.893	0.915	0.9832	45	
0.8706	0.8944	0.9066	0.9172	0.9832	55	
0.001	0.0061	0.0298	0.1986	1.1061	1	***Mean cardinality***
0.0025	0.0415	0.2998	1.7316	15.3436	5	
0.0038	0.2283	1.4124	5.577	32.737	10	
0.0283	0.8598	3.5578	9.5714	44.0876	15	
0.3384	4.2234	8.9058	17.3696	56.1998	25	
1.6163	8.7322	14.4062	23.2576	64.7098	35	
4.5122	13.0262	18.6348	27.0322	71.141	45	
9.476	16.8226	22.181	29.2632	76.2638	55	

Table VIII. Coverage probability and mean cardinality of the proposed drift estimator under drift shifts in $\pi_{22} - \beta_{22}(k - \tau)$ and $\pi_{34} + \beta_{34}(k - \tau)$.

$\beta_{22} = \beta_{34}$							
0.015	0.009	0.007	0.005	0.003	0.001	D	
0.5336	0.4124	0.2948	0.2346	0.1074	0.0078	150	***Coverage probability***
0.6534	0.5696	0.523	0.4718	0.3652	0.0236	200	
0.699	0.6632	0.6456	0.5854	0.591	0.185	250	
0.7228	0.7176	0.6876	0.6964	0.6886	0.5074	300	
0.7376	0.7336	0.7246	0.724	0.7276	0.6764	350	
0.75	0.7544	0.7758	0.7514	0.7636	0.7514	400	
0.7646	0.7562	0.7536	0.7654	0.7786	0.8066	450	
0.7598	0.764	0.7756	0.7772	0.792	0.8268	500	
0.7864	0.8194	0.8092	0.808	0.8402	0.9054	1000	
14.5347	11.915	8.6102	7.0076	3.2735	0.0277	150	***Mean cardinality***
17.9716	16.6832	15.8085	14.7159	12.5703	0.5934	200	
19.8474	19.612	19.9986	18.9594	20.6276	8.4153	250	
20.8736	21.7268	21.591	22.7423	24.5586	24.043	300	
21.4515	22.4460	22.9248	24.018	26.317	32.1004	350	
21.904	23.1964	24.4894	25.0792	28.0482	36.4337	400	
22.459	23.3922	24.1406	25.7885	28.8286	39.4169	450	
22.3388	23.6826	24.9196	26.2654	29.3301	40.3462	500	
23.1238	25.4832	26.0974	27.4167	31.3186	45.2882	1000	

VIII. A NUMERICAL CASE

A numerical example is now illustrated based on the real case of kidney patients used by Kamranrad et al. (2019) to show the application of the proposed estimators. For this purpose, we chose to use age and disease type as the two variables. For each variable three levels are considered. categories of age are 20-40, 41-60 and 61-80. Three categories of disease type are also chosen from existing categories which are kidney stone (KS), hydronephrosis (HN), kidney transplant (KT). Kamranrad et al. (2019) reported 7 samples of phase I. In this section to obtain an underlying model for Phase II, we add data for all 7 samples and for both male and female. Hence the underlying contingency table model is shown in Table IX.

After receiving out-of-control warning from the EWMA-WALD control chart with the parameters of $\lambda = 0.2$and $L = 145.9$to reach the in-control ARL of 200, the proposed estimators are used to estimate the change point.

Table IX. The Underlying Contingency Table Model

		Disease type		
		KS	HN	KT
Age	20-40	107	62	29
	41-60	150	95	31
	61-80	95	69	26

The step changes are imposed to the parameters after the actual change point of $\tau = 25$ such that changes are $\pi_{12} - 0.03$ and $\pi_{33} + 0.03$. Results are summarized in Fig. 1.

Fig. 1 shows that estimated change point using the proposed step and drift estimator are 25 and 22, respectively. Hence, the performance of the estimators is acceptable.

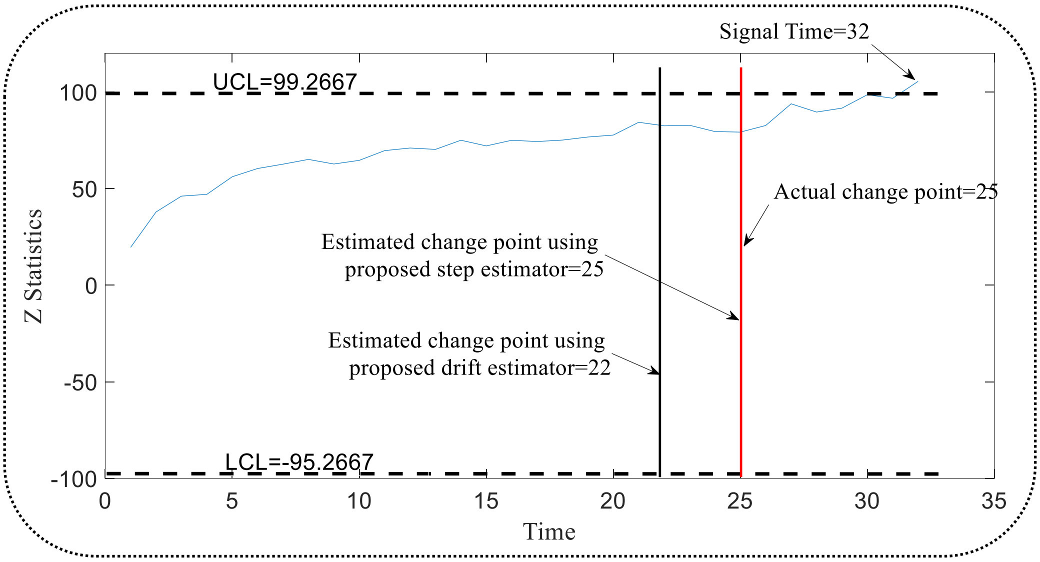

Fig 1. The proposed estimators' performance in numerical example

IX. CONCLUSION AND FUTURE RESEARCH

In this study, maximum likelihood estimators were developed for estimating the step and drift change points in contingency table-based processes. The averaging method was used for calculating unknown parameters in the step change, and for estimation unknown shift slopes in linear drift change, the forced regression to pass the origin of the axis of coordinates was used. Simulation results reported the excellent performance of the proposed step estimator. However, the proposed drift change point performed better in moderate shifts and weak in very small and large shift sizes. In addition, the proposed step estimator was more accurate and more precise than the proposed estimator of the linear drift change in all shifts from small to large. Cardinality and coverage probability of the confidence sets were also calculated. Finally, performance of the proposed estimators assessed by a real example.

In the area of change point estimation in contingency tables, estimating monotonic change point and also estimating sporadic change point in Phase II, can be considered as future research.

REFERENCES

Agresti, A. (2002). Categorical data analysis. Department of Statistics University of Florida Gainesville, Florida: John Wiley & Sons, Inc, Hoboken, New Jersey.

Alaeddini, A., Ghazanfari, M., & Nayeri, M. A., (2009). A hybrid fuzzy-statistical clustering approach for estimating the time of changes in fixed and variable sampling control charts. Information Sciences, 179(11), 1769-1784.

Amiri, A., & Allahyari, S., (2012). Change point estimation methods for control chart postsignal diagnostics: a literature review. Quality and Reliability Engineering International, 28(7), 673-685.

Amiri, A., & Khosravi, R., (2013). Identifying time of a monotonic change in the fraction nonconforming of a high-quality process. The International Journal of Advanced Manufacturing Technology, 68(1-4), 547-555.

Amiri, A., Niaki, S. T. A., & Moghadam, A. T., (2015). A probabilistic artificial neural network-based procedure for variance change point estimation. Soft Computing, 19(3), 691-700.

Ashuri, A., & Amiri, A., (2016). Estimating monotonic change in the rate and dependence parameters of INAR (1) process (Case study: IP counts data). Communications in Statistics-Simulation and Computation, 46(7), 5023-5053.

Atashgar, K., & Noorossana, R., (2011). An integrating approach to root cause analysis of a bivariate mean vector with a linear trend disturbance. The International Journal of Advanced Manufacturing Technology, 52(1-4), 407-420.

Atashgar, K., & Noorossana, R., (2012). Diagnosing the source (s) of a monotonic change in the process mean vector. The International Journal of Advanced Manufacturing Technology, 60(9-12), 1175-1183.

Ayoubi, M., Kazemzadeh, R. B., & Noorossana, R. (2014). Estimating multivariate linear profiles change point with a monotonic change in the mean of response variables. The International Journal of Advanced Manufacturing Technology, 75(9), 1537-1556.

Ayoubi, M., Kazemzadeh, R. B., & Noorossana, R. (2016). Change point estimation in the mean of multivariate linear profiles with no change type assumption via dynamic linear model. Quality and Reliability Engineering International, 32(2), 403-433.

Bersimis, S., & Sachlas, A. (2019). A one-sided procedure for monitoring variables defined on contingency tables. Journal of Quality Technology, 51(2), 125-142.

Dette, H., Pan, G., & Yang, Q. (2022). Estimating a change point in a sequence of very high-dimensional covariance matrices. Journal of the American Statistical Association, 117(537), 444-454.

Fahmy, H. M., & Elsayed, E. A., (2006). Detection of linear trends in process mean. International Journal of Production Research, 44(3), 487-504.

Ghazanfari, M., Alaeddini, A., Niaki, S.T.A, & Aryanezhad, M.B., (2008). A clustering approach to identify the time of a step change in shewhart control charts. Quality and Reliability Engineering International, 24(7), 765-778.

Ghazizadeh, A., Sarani, M., Hamid, M., & Ghasemkhani, A. (2021). Detecting and estimating the time of a single-step change in nonlinear profiles using artificial neural networks. International Journal of System Assurance Engineering and Management, DOI: https://doi.org/10.1007/s13198-021-01121-y.

Hakimi, A., Farughi, H., Amiri, A., & Arkat, J. (2019). New phase II control chart for monitoring ordinal contingency table based processes. Journal of Industrial and Systems Engineering, 12, (Special issue on Statistical Processes and Statistical Modelling), 15-34.

He, Q., Liu, W., & Wang, J. (2019). Change-Point Detection in Multivariate Categorical Processes with Variable Selection. IEEE Access, 7, 73065-73075.

Kamranrad, R., Amiri, A., & Niaki, S. T. A. (2019). Phase-I monitoring of log-linear model-based processes (a case study in health care: Kidney patients). Quality and Reliability Engineering International, 35(6), 1766-1788.

Kamranrad, R., Amiri, A., & Niaki, S. T. A., (2017a). Phase-II monitoring and diagnosing of multivariate categorical processes using generalized linear test-based control charts. Communications in Statistics-Simulation and Computation, 46(8), 5951-5980.

Kamranrad, R., Amiri, A., & Niaki, S. T. A., (2017b). New Approaches in Monitoring Multivariate Categorical Processes based on Contingency Tables in Phase II. Quality and Reliability Engineering International, 33(5), 1105-1129.

Kazemi, M. S., Bazargan, H., & Yaghoobi, M. A., (2014). Estimating the drift time for processes subject to linear trend disturbance using fuzzy statistical clustering. International Journal of Production Research, 52(11), 3317-3330.

Kazemzadeh, R. B., Noorossana, R., Ayoubi, M., (2015). Change point estimation of multivariate linear profiles under linear drift. Communications in Statististics-Simulation and Computation, 44(6), 1570-1599.

Li, J., Tsung, F., and Zou, C., (2012). Directional control schemes for multivariate categorical processes. Journal of Quality Technology, 44 (2), 136-154.

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

Li, J., Zou, C., Wang, Z., and Huwang, L. A., (2013). Multivariate sign chart for monitoring process shape parameters. Journal of Quality Technology, 45 (2), 149-165.

Maleki, M. R., Amiri, A., Taheriyoun, A. R., & Castagliola, P. (2018). Phase I monitoring and change point estimation of autocorrelated poisson regression profiles. Communications in statistics-Theory and Methods, 47(24), 5885-5903.

Movaffagh, A., & Amiri, A., (2013). Monotonic change point estimation in the mean vector of a multivariate normal process. The International Journal of Advanced Manufacturing Technology, 69(5-8), 1895-1906.

Nedumaran, G., Pignatiello Jr, J. J., & Calvin, J. A., (2000). Identifying the time of a step-change with χ^2 control charts. Quality Engineering, 13(2), 153-159.

Niaki, S. T. A., & Khedmati, M. (2014a). Step change-point estimation of multivariate binomial processes. International Journal of Quality & Reliability Management, 31(5), 566-587.

Niaki, S. T. A., & Khedmati, M., (2014b). Monotonic change-point estimation of multivariate Poisson processes using a multi-attribute control chart and MLE. International Journal of Production Research, 52(10), 2954-2982.

Nishina, K., (1992). A comparison of control charts from the viewpoint of change point estimation. Quality and Reliability Engineering International, 8(6), 537–541.

Noorosana, R., Saghaei, A., Paynabar, K., & Abdi, S., (2009). Identifying the period of a step change in High yield processes. Quality and Reliability Engineering International, 25(7), 875-883.

Noorossana, R., & Shadman, A., (2009). Estimating the change point of a normal process mean with a monotonic change. Quality and Reliability Engineering International, 25(1), 79-90.

Noorossana, R., Atashgar, K., & Saghaei, A., (2011). An integrated supervised learning solution for monitoring process mean vector. The International Journal of Advanced Manufacturing Technology, 56(5-8), 755-765.

Page, E.S., (1954). Continuous inspection schemes. Biometrika, 41(1-2), 100–115.

Perry, M. B. (2020). An EWMA control chart for categorical processes with applications to social network monitoring. Journal of Quality Technology, 52(2), 182-197.

Perry, M. B., & Pignatiello Jr, J. J., (2011). Estimating the time of step change with Poisson CUSUM and EWMA control charts. International Journal of Production Research, 49(10), 2857-2871.

Perry, M. B., & Pignatiello Jr, J. J., (2006). Estimation of the change point of a normal process mean with a linear trend disturbance in SPC. Quality Technology and Quantitative Management, 3(3), 325-334.

Perry, M. B., & Pignatiello Jr, J. J., (2010). Identifying the time of step change in the mean of autocorrelated processes. Journal of Applied Statistics, 37(1), 119-136.

Perry, M. B., Pignatiello Jr, J. J., & Simpson J. R., (2006). Estimating the change point of a poisson rate parameter with a linear trend disturbance. Quality and Reliability Engineering International, 22(4), 371-384.

Perry, M. B., Pignatiello Jr, J. J., & Simpson, J. R., (2007). Change point estimation for monotonically changing Poisson rates in SPC. International Journal of Production Research, 45(8), 1791-1813.

Pignatiello Jr, J. J., & Samuel, T. R., (2001). Identifying the time of a step change in the process fraction nonconforming. Quality Engineering, 13(3), 357-365.

Samuel, T. R., & Pignatiello Jr, J. J., (1998). Identifying the time of a change in a poisson rate parameter. Quality Engineering, 10(4), 673-681.

Samuel, T. R., Pignatiello Jr, J. J., & Calvin, J. A., (1998). Identifying the time of a step change in a normal process variance. Quality Engineering, 10(3), 529-538.

Steward, R. M., & Rigdon, S. E., (2017). Risk adjusted Monitoring of Healthcare Quality: Model Selection and Change Point Estimation. Quality and Reliability Engineering International, 33(5), 979-992.

Venegas, J., Tercero-Gomez, V., Cordero Franco, A., Temblador-Pérez, M., & Beruvides, M., (2016). An evaluation of change-point estimators for a sequence of normal observations with unknown parameters. Communications in Statistics-Simulation and Computation, 46(6), 4297-4317.

Xiang, D., Pu, X., Ding, D., & Liang, W. (2021). An efficient charting scheme for multivariate categorical process with a sparse contingency table. Journal of Quality Technology, 53(1), 88-105.

Yashchin, E. (2012). On detection of changes in categorical data. Quality Technology & Quantitative Management, 9(1), 79-96.

Yeganeh, A., Pourpanah, F., & Shadman, A. (2021). An ANN-based ensemble model for change point estimation in control charts. Applied Soft Computing, 110, 107604.

Zarandi, M. H. F. & Alaeddini, A., (2010). A general fuzzy-statistical clustering approach for estimating the time of change in variable sampling control charts. Information Sciences, 180(16), 3033-3044.

Zou, C., and Tsung, F.A., (2011). Multivariate sign EWMA control chart. Technometrics, 53 (1), 84-97.