Defects Control Charts for High-Quality Processes

Surath Aebtarm

A Thesis

in

The Concordia Institute

for

Information Systems Engineering

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Applied Science (Quality Systems Engineering) at
Concordia University
Montréal, Québec, Canada

November 2008

Your file Votre référence
ISBN: 978-0-494-63320-5
Our file Notre référence
ISBN: 978-0-494-63320-5

ABSTRACT

Defects Control Charts for High-Quality Processes

Surath Aebtarm

The traditional C-chart by Shewhart has been widely applied for monitoring count data in industrial and nonindustrial processes. However, using C-chart always experiences an excessive amount of false alarms, since control limits of traditional C-chart are defined by impractical normal assumption. Specially, when we monitor two or more correlated characteristics of defects, C-chart becomes unsuitable. Thus, monitoring a process by traditional C-chart leads to the increase of unnecessary costs of inspection.

There are many works that have attempted to improve C-charts. In this thesis, 11 selected improved versions of C-chart are presented. The performances of improved C-charts are evaluated in term of numerical results to demonstrate the sensitivity of the charts and costs of inspections.

We also propose an optimal bivariate Poisson field chart to monitor two correlated characteristics of defects. Our chart is based on the optimization of bivariate Poisson confidence interval and projection of bivariate Poisson data in Poisson field. The detailed description of our proposed algorithm is presented by numerical data. The experimental results demonstrate improved performances regarding user-friendly visualization and false alarm rate.

Furthermore, we propose an optimal diagonal inflated bivariate Poisson field chart to monitor two over/under dispersed correlated count data. The detailed description of our chart will be presented. The experimental results demonstrate improved performances according to loss function and false alarm rate compared to other methods.

Acknowledgements

I would like to express my sincere appreciation, Dr. N. Bouguila, for his continuous guidance and support through this thesis. Without his mentor, I might not have discovered some parts of myself in this field. I would also like to thank all my colleagues and friends who have been very supportive. People who also deserve my acknowledgment are my professors in CIISE. Especially, Dr. A. Ben Hamza for statistical knowledge and Matlab programming from his course.

I would like to thank my parents and my lovely sister for unconditional support throughout my studies. My special thank go to Miss Prim for her encouragement. Without her, all of my inspiration might not come out through this thesis.

Last, but not least, I dedicate my thesis to Mr. Pongsek , my ME SIIT friend who passed away from brain tumor. Without him, I might not really learn “to live as if I do not have tomorrow”.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Manufacturing of products always deals with the variation in the production. The variations can be due to common and special causes [26] [48]. When a process contains only common causes of variation, it is in control. The average level of events, errors, or defects per unit can be used to calculate the process capability. When a process contains special causes, the process is out of control.

To deal with process variation, control charts are effective tools that are widely used for quality inspection. The main purpose of a control chart is continually monitoring a process by illustrating its behavior [9]. Figure 1.1 shows an example of control chart. Control charts illustrate data on time series axis which provide the historical characteristics of a process. The maximum expected variation is shown as the upper and lower control limits. When a sample is excessive from maximum expected variation, it is indicated as out-of-control which can be refereed to "assignable cause" [9]. The application of control chart are defining process capability, benchmarking processes, and evaluating pilot state.

In a process, data can be continuous or discrete. Attribute charts have been widely used to monitor discrete data. Since attribute data can be gathered from every process or even transformed from continuous data, attribute control charts are widely used in many fields to monitor both manufacturing and non-manufacturing processes [62]. For instance, a control chart of number of defects can be used for manufacturing purposes, and control chart of number of accidents per week is used in non-manufacturing issue [26]. Although an attribute chart is not always as effective tool as continuous control charts to find root problems and solutions, it is an economical tool to collect and analyze the process characteristics before continuous charts

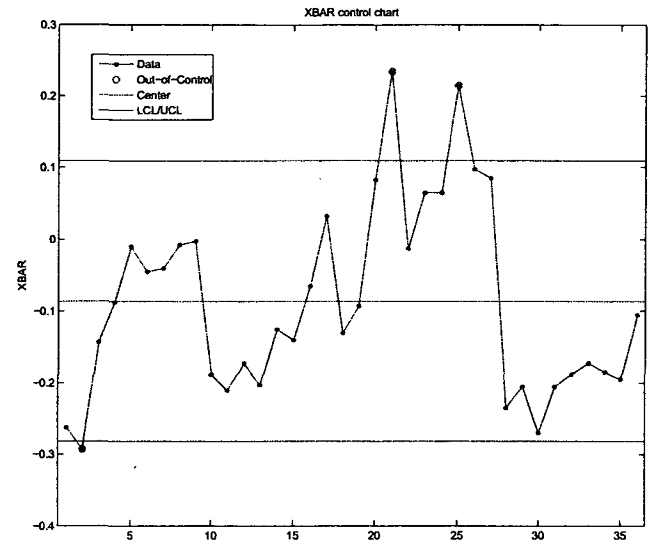


Figure 1.1: Control Chart.

can be applied [32]. Besides, attribute charts are more practical in some cases. For example, monitoring number of survived patients per year is more practical than monitoring how long patient can be survival which usually uses continuous control chart [62].

There are two types of attribute control charts, defective and defects charts. The term "defective" is referred as the item that has one or more defects. The term "nonconforming" is also used instead of "defective" [25]. These types of charts are used with the number or proportion of defective units. For instance, P-charts and np-charts are used for evaluating proportion and number of defectives, respectively, which are produced by a process. Examples of defective data are number of defective items per proportion, number of rejected invoices per 30 inspections, and number of surgical operations that went smoothly [62]. The term "defect" in quality field is referred as a single non-conforming quality characteristic of an item. It's possible that a nonconforming unit has many non-conformities [58] [66]. C-chart is a defect control chart that has been widely used in statistical process control (SPC). C-chart is used to monitor the actual total number of defects per unit. For example, number of defects per item and number of patients in a hospital per day [32]. Constructing C-chart is inexpensive since the plotted data are count data which does not require measurement, and can be collected from daily report in many cases. Furthermore, C-chart is used for plotting numbers of defects.

Thus, it's simpler to plot C-chart than any other control chart by just plotting raw data without necessity of transformation [62].

However, there are some points that need to be considered before attribute charts are applied. Attribute control charts can be biased if an inspector misjudges a product to be a defective [26]. For measuring small changes of variables, attributes are not as sensitive as continuous charts to represent the process, since attribute chart plots only in term of acceptable or not, instead of exact value of data. The result of using attribute chart is sometimes out of reality because in some cases small variation cannot count as a defective in reality [26].

1.1 Background

1.1.1 The Classic Shewhart's C-chart

The classic C-chart by Shewhart is widely used in SPC when a process provides count data. Poisson distribution is in general used to model count data and is given as the following [9]:

$$P(x = k) = \frac{e^{-\lambda}\lambda^k}{k!} \tag{1}$$

where λ represents at the same time the mean and the variance and k is the number of occurrences. The usual approach to obtain control limits of Shewhart's attribute control chart is to use plus and minus three standard deviation limits (under the assumption of normal approximation [9]):

$$\bar{c} \pm 3\sqrt{\bar{c}} \tag{2}$$

where $\bar{c}$ is sample mean. To construct the traditional Shewhart C-chart [9] [26] [32], the upper control limit (UCL) and the lower control limit (LCL) can be calculated by

$$UCL = \bar{c} + 3\sqrt{\bar{c}} \tag{3}$$

$$LCL = \bar{c} - 3\sqrt{\bar{c}} \tag{4}$$

When LCL is negative in case of low mean of Poisson variable, LCL will be considered as zero because number of defects cannot be negative [9].

The average number of defects can be calculated by the estimated mean of Poisson distribution as follows

$$\lambda = \frac{1}{n}\sum_{i=1}^{n} k_i. \tag{5}$$

where n is the total number of variables, and can be derived by maximum likelihood estimation (MLE) [33]:

$$\begin{aligned} L(\lambda) &= \ln\prod_{i=1}^{n} f(k_i \mid \lambda) = \sum_{i=1}^{n}\ln\left(\frac{e^{-\lambda}\lambda^{k_i}}{k_i!}\right) \\ &= -n\lambda + \left(\sum_{i=1}^{n} k_i\right)\ln(\lambda) - \sum_{i=1}^{n}\ln(k_i!) \end{aligned} \tag{6}$$

Take the derivative of L with respect to λ and equate it to zero:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}L(\lambda) = 0 \iff -n + \left(\sum_{i=1}^{n} k_i\right)\frac{1}{\lambda} = 0 \tag{7}$$

Solving for λ yields the maximum-likelihood estimate of λ:

$$\widehat{\lambda}_{\mathrm{MLE}} = \frac{1}{n}\sum_{i=1}^{n} k_i. \tag{8}$$

For the classic Shewhart's C-chart, there are two main drawbacks: (1) imprecise control limits by normal approximation and (2) the mean and variance may be unequal for small sample size.

Generally, the normal approximation performs poorly due to its rigidity (always symmetric). Thus, the normal approximation can not provide the efficient control limits for C-chart as it was discussed in [58]. Especially, when the mean of Poisson distribution is small, the shape of count data distribution tails to infinity on one side. Thus, count data should not be plotted in normal approximation control limits.

The unequal mean and variance for small sample size also cause imprecise control limits as it was explained in [37] where the author noticed that *"for the small probability such as Poisson distribution, the sample size needs to be large enough in order to give the good distribution fit* ". This issue was explained by using divergence coefficient:

$$Q^2 = \frac{s^2}{\bar{x}} \tag{9}$$

where s^2 and $\bar{x}$ are the sample mean and variance, respectively. Q^2 is the square of the "divergence coefficient". By Poisson definition, mean and variance should be equal. Thus, divergence coefficient should always be equal to one. However, the divergence coefficient is not always one if sample size is small. Thus, if sample size is not large enough, divergence coefficient will not be unity [37], which supports the reason why normal approximation control limit of C-chart cannot provide high standard of false alarm.

Moreover, mean of Poisson population that is acquired by using MLE or taking the average directly from the data is not as precise as the exact mean of Poisson population [33]. Shenton and Bowman pointed out this issue in the case of small sample size (which follows Poisson distribution) and mentioned that *"although MLE is good choice to use as estimator, it can mislead to the asymptotic variance which occur when the sample sizes are small, and leads to error of estimation"* [33]. As a result, using MLE to approximate the mean of sample leads to construct imprecise control limits. The constructed chart will not be able to satisfy three standard deviation confidential interval and also lead to create excessive false alarm, and then waste, time, and more expensive efforts. Thus, a challenging problem is to improve such an effective and flexible tool to reach high standard false alarm rate.

1.1.2 Key Parameters of Control Chart's Performances

Since in many circumstances, C-chart does not perform efficiently, classic C-chart still leaves the room for improvement. In order to assess the efficiency and performance of a given control chart, three key parameters which are plotting illustration, loss function, and average run length will be used.

Although quantitative analysis is the main reliable approach to analyze a statistical algorithm, in practise, the main concerns are not only flexibility and accuracy, but control charts also need to be handy and visual. Handiness and visualization can not be shown by numerical results. Thus, the control chart plotting illustration should be considered. The characteristics of a robust control chart that are mentioned in many research papers are the ability to plot raw data in the control chart, and plot many different data in the same control chart. Since the main purpose of control chart is illustration of process behavior, Ryan mentioned that an efficient control chart needs to be able to show the raw data in order to be able to illustrate if something goes right or wrong [60]. Moreover, Quesenberry also explained that we need to be able to plot different data in the same control chart in order to help an inspector to compare two or more different data at the same time [7]. Thus, selecting the proper control chart depends on the actual process needs. The actual data would be the main requirement of charts if the charts are needed to illustrate the real data. In the other hand, a chart that can plot many different types of data at the same time would be needed to monitor different types of defects. In this

Table 1.1: Nominal value of ARL for each state

	In control state	Shift = 0.5σ	Shift = 1σ	Shift = 1.5σ	Shift = 2σ
Both tails	370.40	33.42	4.50	1.57	1.07
Each tail	740.80	66.84	9.00	3.14	2.14

research, the data illustration of each tested chart will be presented.

Second, Average Run Length (ARL) is the vital parameter which is used for evaluating the performance of control charts and making decision to choose sample size and sampling frequency [9]. In statistical process control, ARL can be explained as the average number of samples that can be inspected before a point is indicated out of control. In other words, ARL can be described as how often that alarm will be signaled when a process is out of control. There are many contradictions among researchers to describe the ARL of the high performance control chart. Nevertheless, the majority of researchers agree that *"The high value of ARL will be desired when the process is in control but the short value of ARL will be desired when the mean of the process is shifted or been out of control"* [69]. ARL can describe sensitivity of control charts whenever the mean of the process changed [35]. ARL can be calculated by the average total number of samples that can be plotted in the chart before the point that is out of control appeared [9]. For independent variable, ARL is given by:

$$ARL = \frac{1}{p(\text{sample point plots out of control})} \tag{10}$$

For three standard deviation level, during in-control state, the probability of out of control points is 0.0027. Thus, since p is the probability of out of control, ARL is equal to $1/p = 370$, which means that if the process is in control, an out of control product will appear in every 370 products. Thus, during the control state, preferable ARL value will be 370. Since many statisticians always compare the level of ARL to three standard deviation level, it's interesting to compare defects chart to three standard level. In [12] the author presented the run length for three standard deviation $\bar{x}$ due to infinite sample size when means of samples are shifted which can be referenced as ARL of three standard deviation.

Moreover, in some processes, it's very interesting to consider ARL of lower and higher control limits separately in order to monitor characteristics of each control limit. Because Poisson distribution has two tails, the probability of out of control point can be divided by two [42]. ARL of each tail will be equal to $1/(p/2) = 740.4$. In table 1.1, we show ARL values of each state. The nominal ARL values of both tails and single tail can be shown in form of variation of state such as in-control state and mean shifted states.

Third, loss function is also a key measure that can be considered in order to monitor the level of cost that can be gained in the control charts. Loss function illustrates cost, which can be caused by excessive false alarm [42]. In practice, loss function shows how much a process creates cost of inspection, and waste. The loss function is defined as [60]:

$$Loss.function = (\frac{1}{LTA} - \frac{1}{0.00135} + \frac{1}{UTA} - \frac{1}{0.00135}) \tag{11}$$

where $LTA = \frac{1}{ARL_L}$ and $UTA = \frac{1}{ARL_U}$. A lower limit total area (LTA) is the total area of out of control of the lower control limit, and an upper limit total area (UTA) is total area of out of control of the upper control limit. They can be calculated by ARL of lower and upper control limits. According to the previous equation, LTA and UTA are subtracted by nominal three standard deviation confidential interval ($ARL = 740.7$) to calculate how much the ARL of the control chart gains over three standard deviation level. Thus, an inspector will be able to locate how much each control limit produces false alarm rate more than three standard deviation confidence interval. To compare the performances of each approach, it's important to calculate the loss function that illustrates the cost that will be created by using each method. To have an effective control chart, the loss function must be minimized.

1.2 Contributions

The contributions of this thesis are as follows:

- **An Empirical Evaluation of Selected Defects charts:** We investigate the performances of 11 selected defects charts. To compare the performances of each chart, we tested each chart and considered some key factors such as low mean responsiveness, loss function, and mean shifting sensitivity. To simplify the comparison of each control chart performances, we categorize all control chart into three groups, transforming data, standardizing data, and optimizing control limits.

- **An Optimal Bivariate Poisson Field Chart for High-quality Manufacturing processes:** We propose an optimal bivariate Poisson field chart to monitor correlated defects. This chart is improved by optimizing bivariate Poisson confidence interval and illustrate bivariate Poisson data in Poisson field. The results show that our control chart provides excellent rate of false alarms, and enhanced visual-interface of correlated characteristics on a single chart.

☞ **An Optimal Diagonal Inflated Bivariate Poisson Field Chart:** We propose an optimal bivariate Poisson field chart to monitor bivariate over/under dispersed count data. This chart is also improved by optimization of confidence interval and illustration of Poisson field. However, to deal with over/under-dispersion, the diagonal inflated bivariate Poisson model is used instead of usual bivariate Poisson model. The proposed chart presents excellent rate of false alarms, and high sensitivity to handle over/under-dispersed count data. The various simulated data demonstrate the enhanced performances of our control chart compared to other previously proposed charts.

1.3 Thesis Overview

The organization of this thesis is as follows:

- ❑ The first Chapter introduced Shewhart C-chart, and reviewed some keys parameters to assess the performances of control chart.

- ❑ In Chapter 2, we present 11 selected improved versions of C-charts. The performances of improved C-charts are evaluated in term of numerical results to demonstrate the sensitivity of the charts and costs of inspections.

- ❑ In Chapter 3, we propose an optimal bivariate Poisson field chart to monitor two correlated defects. The detailed description of our chart will be presented. The comparative results and case study demonstrate improved performances according to data visualization and false alarm rate compared to other methods.

- ❑ In Chapter 4, we propose an optimal diagonal inflated bivariate Poisson field chart to monitor two over/under dispersed correlated count data. The detailed description of our chart will be presented. The experimental results demonstrate improved performances according to loss function and false alarm rate compared to other methods.

- ❑ In the **Conclusions**, we summarize our contributions.

Chapter 2

An Empirical Evaluation of Selected Defects charts

The traditional C-chart by Shewhart has been widely applied for monitoring count data in industrial and nonindustrial processes. However, using C-chart always experiences an excessive amount of false alarms, since control limits of traditional C-chart are defined by impractical normal assumption. Thus, monitoring a process by traditional C-chart leads to increase unnecessary costs of inspection. There are many works that have attempted to improve C-charts. In this chapter, 11 selected improved versions of C-charts are presented. The basic concepts and detailed description of all charts are discussed. The performances of improved C-charts are evaluated in term of numerical results to demonstrate the sensitivity of the charts and costs of inspections.

2.1 Introduction

Many approaches have been proposed to improve attribute control charts. However, it's possible to categorize them into three major groups of approaches which are: (1) the transforming data approach, (2) the standardizing data approach, and (3) the optimizing control limits approach. To introduce all control charts, the theoretical ideas behind each chart and its construction will be presented. In order to visualize each control chart and investigate its ability to illustrate data, we will use the dataset in table 2.1 [32]. For some control charts that can be applied with different sample sizes, the charts will be illustrated by the data in table 2.2 which was used in [8].

Table 2.1: Defects data set

Sample No.	Number of Defects (c)
1.00	8.00
2.00	7.00
3.00	6.00
4.00	4.00
5.00	3.00
6.00	9.00
7.00	1.00
8.00	5.00
9.00	0.00
10.00	0.00
11.00	23.00*
12.00	3.00
13.00	15.00*
14.00	8.00
15.00	5.00
16.00	7.00
17.00	3.00
18.00	0.00
19.00	12.00
20.00	3.00
21.00	4.00
22.00	18.00*
23.00	7.00
24.00	4.00
25.00	4.00
"*Denotes sample above upper limit"	Total defects =159

Table 2.2: Defect data set for different sample sizes

Sample No.	Size of sample	Number of Defects
1.00	4.00	4.00
2.00	4.00	3.00
3.00	4.00	3.00
4.00	4.00	7.00
5.00	4.00	6.00
6.00	4.00	10.00
7.00	4.00	7.00
8.00	4.00	6.00
9.00	4.00	7.00
10.00	4.00	4.00
11.00	4.50	7.00
12.00	4.50	10.00
13.00	4.50	10.00
14.00	4.50	6.00
15.00	4.50	9.00
16.00	4.50	5.00
17.00	4.50	9.00
18.00	4.50	5.00
19.00	4.50	11.00
20.00	4.50	3.00
21.00	5.00	8.00
22.00	5.00	11.00
23.00	5.00	12.00
24.00	5.00	5.00
25.00	5.00	7.00
26.00	5.00	8.00
27.00	5.00	8.00
28.00	5.00	10.00
29.00	5.00	5.00
30.00	5.00	14.00
31.00	5.00	7.00
32.00	5.00	9.00
33.00	5.00	14.00
34.00	5.00	16.00
35.00	5.00	17.00
36.00	5.00	15.00
37.00	5.00	15.00
38.00	5.00	9.00
39.00	5.00	13.00
40.00	5.00	11.00
41.00	4.40	20.00
42.00	4.40	15.00
43.00	4.40	11.00
44.00	4.40	18.00
45.00	4.40	11.00
46.00	4.40	6.00
47.00	4.40	8.00
48.00	4.40	15.00
49.00	4.40	11.00
50.00	4.40	10.00

Table 2.3: Summary of data transformation equation

Author	Transforming equation
Bartlett [39] [38]	$y = 2\sqrt{c}$
Anscombe [15]	$y = 2\sqrt{c + \frac{3}{8}}$
Freeman and Tukey [36]	$y = \sqrt{c} + \sqrt{c+1}$

2.1.1 The Transforming Data Approach

The transforming data approach is the first option to improve the performance of attribute control charts. This approach is based on approximating Poisson distribution by a normal distribution after transforming count data.

The idea of transforming data approach is related to transform asymmetrical distribution to almost symmetric one. There are many approaches that statisticians have applied in order to acquire the perfect normal distribution transformation. Ryan presented three interesting transformations in his book which are Bartlett transformation model, Anscombe's transformation model, Freeman and Tukey transformation model, and illustrated the construction of each control chart [58]. Another transformation approach has been proposed in [63]. In the following, we discuss in details these approaches.

Bartlett, Anscombe, and Freeman and Tukey Control Charts

Three main data transformation approaches have been proposed by Bartlett [39] [38], Anscombe [15], and Freeman and Tukey [36], and presented in table 2.3. In this table, c is the number of defects (original data) and y is the transformed data [58]. By using these three transformations, control limits are given by the following:

$$CL = \bar{y} \tag{1}$$

$$LCL = \bar{y} - 3 \tag{2}$$

$$UCL = \bar{y} + 3 \tag{3}$$

where $\bar{y}$ is the mean of transformed data. To illustrate each transforming data control chart, the data set in table 2.1 is transformed by using the equations in table 2.3. Figures 2.1, 2.2, and 2.3 show the resulted control charts using equation 1, 2, and 3 to compute the control limits.

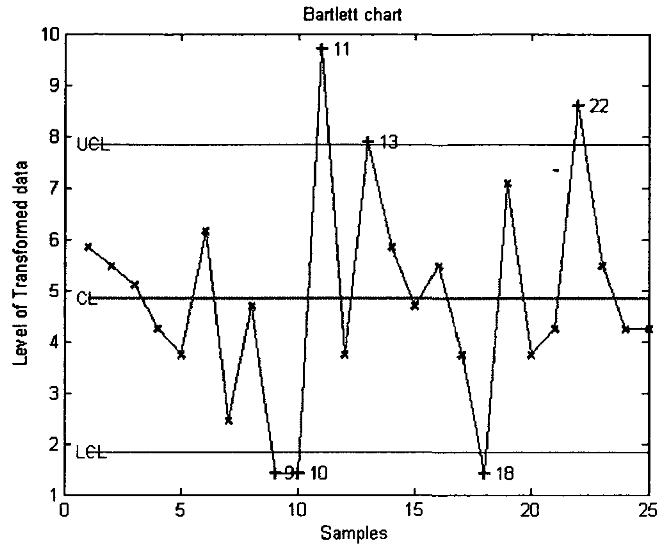


Figure 2.1: Bartlett chart

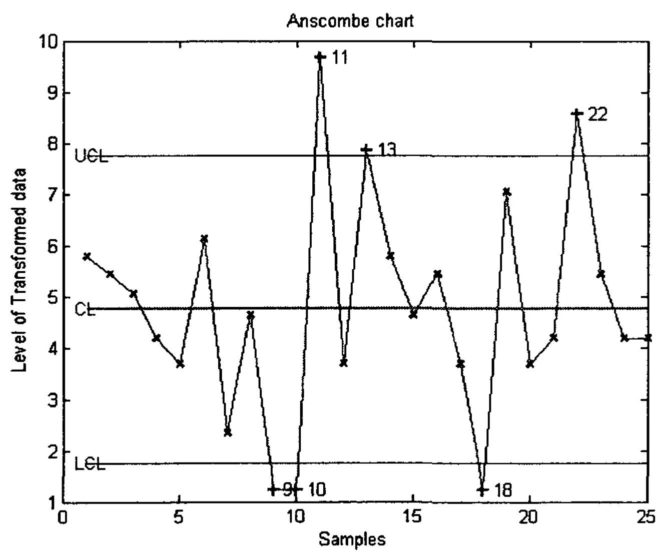


Figure 2.2: Anscombe chart

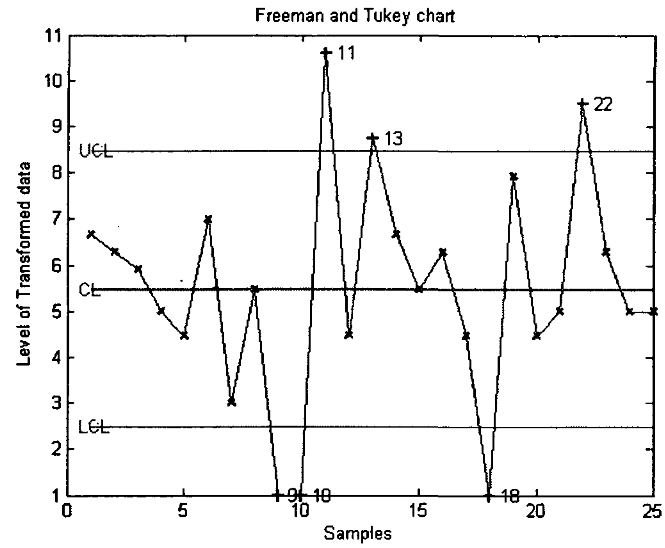


Figure 2.3: Freeman and Tukey chart

Improved Square Root Transformation Control Chart

Alternative attribute control chart based on improved square root transformation (ISRT) was proposed in [63]. The authors used square root of data for transforming Poisson distribution into symmetric distribution. Although their work mainly show the result of transformation to P-chart [63], they emphasized that is is also applicable to C-chart. They explained that ISRT chart can be constructed by the following equations:

$$CL = \sqrt{\bar{c}} \tag{4}$$

$$LCL = \sqrt{\bar{c}} - \frac{3}{2} - \frac{9}{8}(\frac{1}{\sqrt{\bar{c}}}) \tag{5}$$

$$UCL = \sqrt{\bar{c}} + \frac{3}{2} - \frac{1}{2}(\frac{1}{\sqrt{\bar{c}}}) \tag{6}$$

where $\bar{c}$ is the samples mean. Figure 2.4 shows the actual plotting of ISRT control chart using the data set in table 2.1.

Despite transforming data approaches can provide good results, they have some weaknesses. For example, the transforming data approaches require sophisticated inspectors and time in order to transform data. Besides, since data that can be plotted in this type of chart

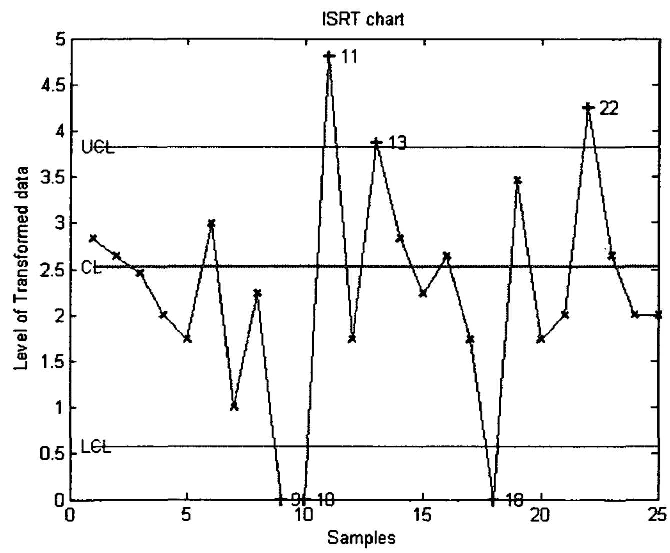


Figure 2.4: ISRT chart

needs to be transformed, the transformed charts can not illustrate the exact level of defects in a process.

2.1.2 The Standardizing Data Approach

Similarly to the transforming approaches, the standardizing approach is based on data transformation. However, in the case of standardizing control chart, after data are transformed, the data will be plotted in -3 to 3 control limits where they will be monitored in form of number of standard deviation. Similarly to the transforming data approach, the original data cannot be plotted on the control chart, and illustrate process characteristics. However, the main advantages of standardizing chart in term of data illustration is that the different types of defects can be plotted in the same standardized chart.

Since the standardized data are plotted in -3 and 3 control limits, the control limits are defined as the following:

$$CL = 0 \quad (7)$$

$$LCL = -3 \quad (8)$$

$$UCL = 3 \quad (9)$$

In 1991, Quesenberry presented his Q-chart along with other charts [8]. His work is one of the first standardizing charts that inspired many other researchers. Q-chart can monitor both continuous and attribute data with constant or non-constant sample size [7]. Moreover, Q-chart can be used in the case of both known and unknown parameters such as mean and standard deviation. In this research, Q-chart and other charts such as standardization C-chart and W transformation chart which were presented along Quesenberry's work will be investigated.

The Standardization C-Chart (Defects Z-chart)

The standardization Poisson transformation (Z-chart) transforms random variables by relocating the data, and spreading of the distribution [8]. This method is based on subtracting the mean of data ($\bar{c} \times n$) and dividing it by its standard deviation:

$$z = \frac{c - n\bar{c}}{\sqrt{n\bar{c}}} \quad (10)$$

Then, the data that are transformed can be plotted in -3 and 3 control limits as usual standardized type of control chart, which can show the level of out of mean data in term of numbers of standard deviation. Figure 2.5 shows ISRT chart using the data in table 2.1.

The Square Root (W) Transformation Chart

The square root transformation chart is a common approach that is applied in count data monitoring and defined by the following equation [8]:

$$W = 2\sqrt{c} - 2\sqrt{n\bar{c}} \quad (11)$$

With this transformation, data can fit in -3 to 3 control limits, and center line is zero. Figure 2.6 shows W transforming chart using the data in table 2.1.

Q-chart for a Poisson Parameter

Q-chart proposed by Quesenberry is a well-known classic standardizing data chart. This chart is based on approximately normalized control chart [8]. Quesenberry explained that his chart can monitor the count data even if the mean of data is unknown with variable sample

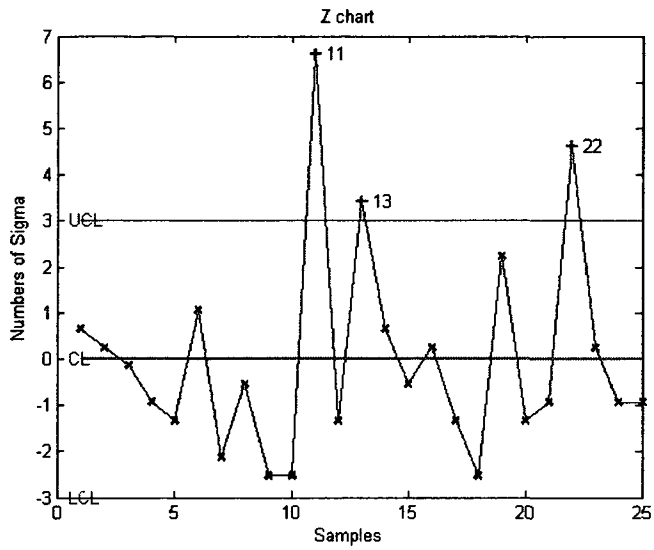


Figure 2.5: Z-chart

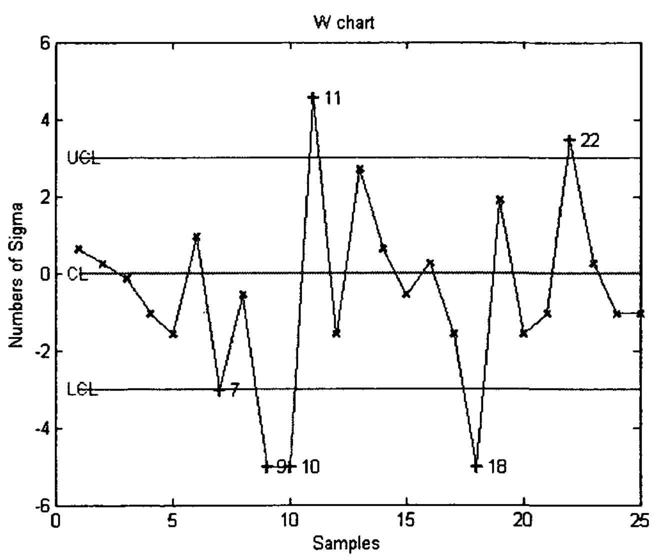


Figure 2.6: W-chart

sizes. For the design of Q-chart for Poisson parameter, one should suppose that y_i is the number of defects on a sampling, and λ is the average rate at which defects occur on a standard size inspection unit. Quesenberry presented Q-chart for Poisson parameter in 2 cases with known and unknown mean.

- **Q-chart for Known λ**

For known λ, the count data will be transformed in order to fit them into -3 and 3 control limits. The data will be transformed from Poisson distribution to Q distribution by the Poisson cumulative distribution function and the inverse of the standard normal distribution function according to these equations:

$$u_i = F(c_i; n_i\lambda) \tag{12}$$

$$Q_i = \Phi^{-1}(u_i) \tag{13}$$

for i = 1,2,...

Each data point of the Poisson distribution is transformed by Poisson cumulative distribution function which has mean λ. Then, the transformed data from Poisson cumulative distribution function will be transformed again by the inverse of the standard normal distribution. After that, Q-distribution will be acquired.

After all transformation steps, the data will be plotted in -3 and 3 control limits which is similar to any other standardized methods. The known λ Q-chart corresponding to the data in table 2.1 is shown in figure 2.7.

- **Q-chart for Unknown λ**

For unknown λ, the count data will be transformed by another way. By uniform minimum variance unbiased (UMVU) estimating distribution function, Lehmann explained that Binomial distribution function can be used for estimation when the probability of occurrence is low which can be referred to Poisson distribution [13]. Thus, in order to transform unknown λ Poisson distribution, data will be transformed by the Binomial cumulative distribution function and the inverse of the standard normal distribution function as the following:

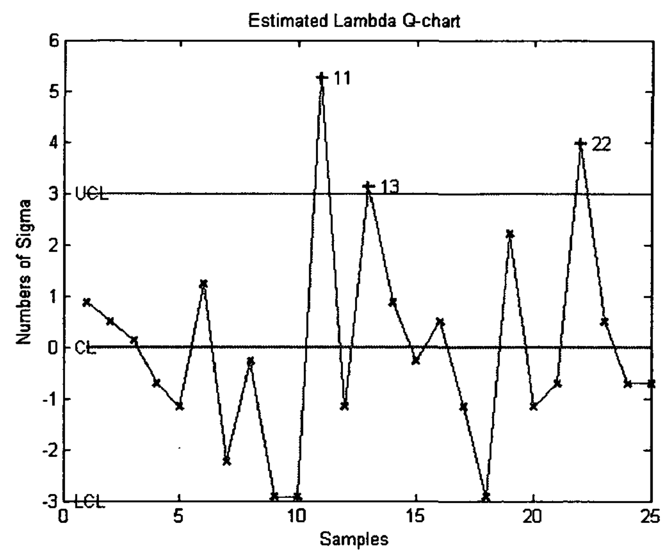


Figure 2.7: Estimated known λ Q-chart

$$u_i = B(c_i; t_i, \frac{n_i}{N_i}) \quad (14)$$

$$Q_i = \Phi^{-1}(u_i) \quad (15)$$

for i = 1,2,...

Where n_i is size of subgroup of sample, and N_i is the total size of sample. For transforming unknown λ Poisson distribution, the data will be transformed by binomial cumulative function, then, similarly to known λ case, the data will be transformed again by inverse of standard normal distribution.

After all transformation steps, the data will be plotted in -3 and 3 control limit. To illustrate unknown λ Q-chart, the actual plotting of unknown λ Q-chart from the data in table 2.1 is given in figure 2.8.

- **The Masking Shifted Effect of Q-chart**

Masking of shifts has been presented as a problem in Q-chart [16]. Indeed, Quesenberry noticed that unknown λ Q-chart is not always accurate if the process mean is shifted at the beginning [7]. This issue was explained in [16] by the fact that, to approximate mean of population, unknown λ Q-chart uses the past observations to update Q-statistic (pool of collected data). Then, if the mean of

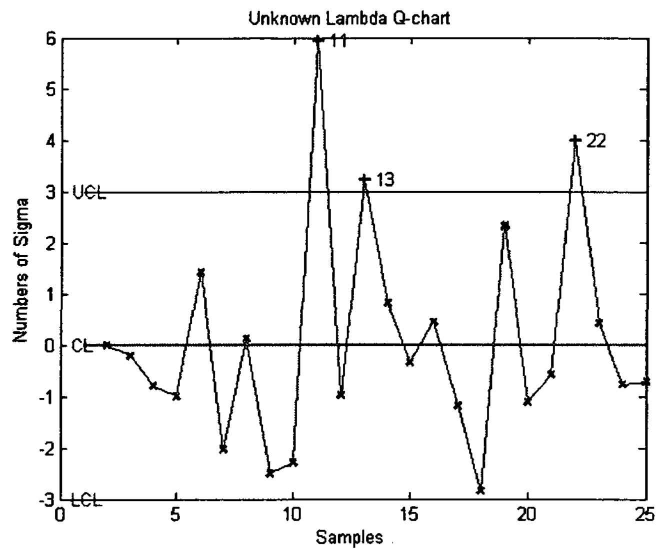


Figure 2.8: Unknown λ Q-chart

population is shifted in early state, the out of control data will be contaminated which leads to inaccurate Q-chart. An example is shown in figure 2.9. In order to present the masking shifted effect, the data set in table 2.1 is augmented by 25 out of control data points which are provided by table 2.4. The plotted data between 1 to 25 are in control state. However, the last 25 data are out of control. In figure 2.9, instead of showing out of control signal after point 25, unknown λ Q-chart presents that out of control data are gradually rebounded back in control when the process reaches point 37. Thus, leaking the defects to the process can be caused by masking shifted effect which creates visibly in-control even out of control data still persisted [16].

Known λ Q-chart will be only tested in the case which can be called "Estimated Q-chart" since the mean is estimated by MLE. However, in the case of unknown λ Q-chart, the masking shifted effect gains the unusual low rate of out-of control signals of the unknown λ Q-chart since the mean shift will rebound back after the points are out-of control. Thus, in this chapter, the comparison of Q-chart with other C-charts will be performed only in the case of estimated λ Q-chart.

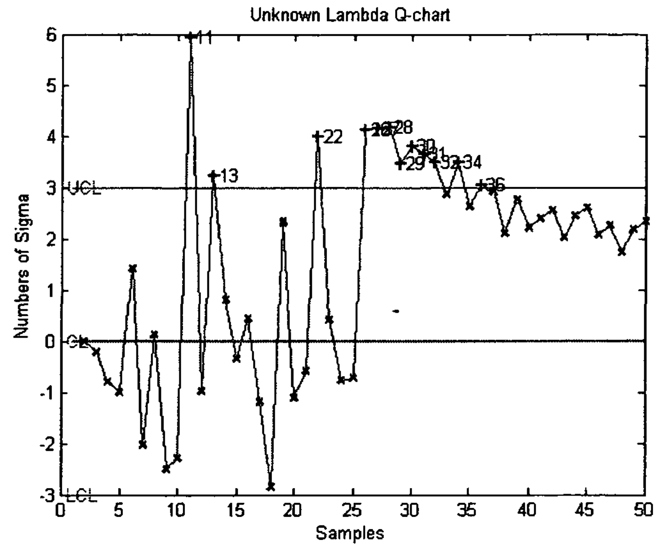


Figure 2.9: Masking shifted effect of Q-chart

Table 2.4: out of control data set for masking shifted effect

Sample No.	Number of Defects (c)
26.00	19.00
27.00	20.00
28.00	21.00
29.00	19.00
30.00	21.00
31.00	21.00
32.00	21.00
33.00	19.00
34.00	22.00
35.00	19.00
36.00	21.00
37.00	21.00
38.00	18.00
39.00	21.00
40.00	19.00
41.00	20.00
42.00	21.00
43.00	19.00
44.00	21.00
45.00	22.00
46.00	20.00
47.00	21.00
48.00	19.00
49.00	21.00
50.00	22.00

2.1.3 The Optimizing Control Limits Approach

Mainly, the optimizing control limits approaches try to define the exact control limits in order to construct the absolute three standard deviation attribute control charts.

With these approaches, the control limits are defined in term of tables or equations which are convenient to apply in practice. Without transforming data, there is no needs of high working skills and a lot of time to construct control charts. There are some works which have been developed by the optimizing control limits approach. For instance, Ryan proposed

"The optimal control limit C-chart" [60]. This chart is the first approach that was proposed by optimizing C-chart control limit. Then, followed works such as "Winterbottom's control chart" [2] and "the (Almost) Exact Control Limits for a C-chart" [50]. In the following, we present these charts.

Optimal Control Limit C-chart

The basic concept of the optimal control limit C-chart is defining C-chart limits that provide satisfied rate of out of control signals. The steps to construct this chart are finding the optimal limits, obtaining table of means and limits, and applying linear regression (See [60] and [42]). As other optimizing control limit charts, the actual data can be plotted. Thus, an inspector can easily monitor the process. The actual number of defects will be shown without transformation to any form. Ryan explained that the optimal control limit C-chart can provide the control limits in both table and equation which are flexible in real manufacturing processes. By following Ryan's works, the optimal control limit C-chart can be constructed by the following equations:

$$CL = \bar{c} \tag{16}$$

$$UCL = 0.6182 + 0.9996\bar{c} + 3.0303\sqrt{\bar{c}} \tag{17}$$

$$LCL = 1.5307 + 1.0212\bar{c} - 3.2197\sqrt{\bar{c}} \tag{18}$$

where $\bar{c}$ is the mean of sample. Using data set in table 2.1, the optimal control limits C-chart is illustrated in figure 2.10.

Winterbottom Control Chart

Winterbottom presented his attribute control chart by using Cornish and Fisher expansions transformation [2]. He explained that Cornish and Fisher expansions transformation can be used for defining the control limits of defects chart. Winterbottom's control chart (Winter chart) can be constructed as the following:

$$CL = \bar{c} \tag{19}$$

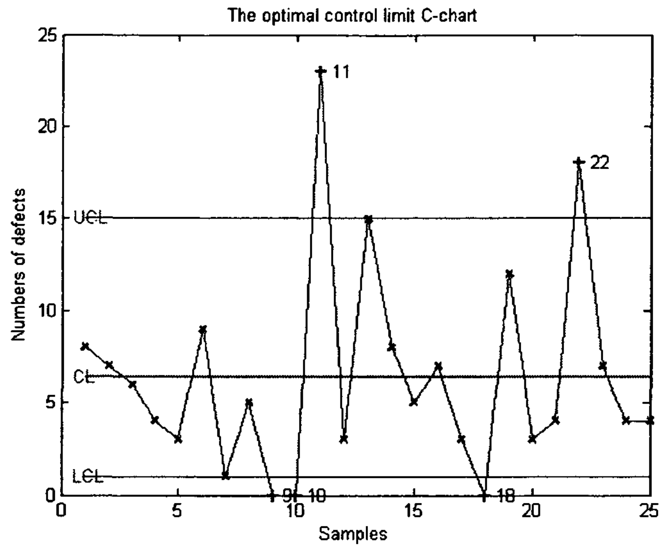


Figure 2.10: The optimal control limit C-chart

$$LCL = \bar{c} - 3\sqrt{\bar{c}} + 4/3 \tag{20}$$

$$UCL = \bar{c} + 3\sqrt{\bar{c}} + 4/3 \tag{21}$$

To illustrate Winterbottom chart, the data in table 2.1 is used (see figure 2.11).

Furthermore, Winterbottom chart is only the optimized control limits chart that can be used with variable sample sizes. Winterbottom provided the control limits in case of variable sample sizes by the following:

$$CL = \bar{\bar{c}} \tag{22}$$

$$LCL = \bar{\bar{c}} - 3\sqrt{\bar{\bar{c}}/n} + 4/3n \tag{23}$$

$$UCL = \bar{\bar{c}} + 3\sqrt{\bar{\bar{c}}/n} + 4/3n \tag{24}$$

where n is a subgroup sample size, and $\bar{\bar{c}}$ is the center line. An example of Winterbottom chart in the case of different sample sizes using the data in table 2.2 is shown in figure 2.12.

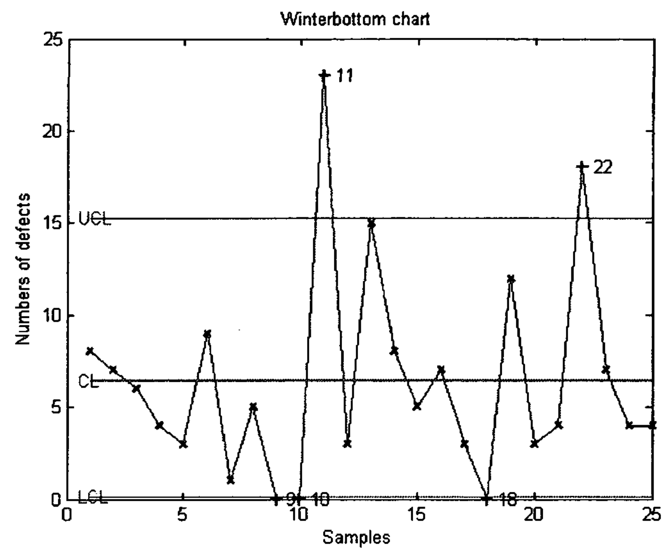


Figure 2.11: Winterbottom's C-chart

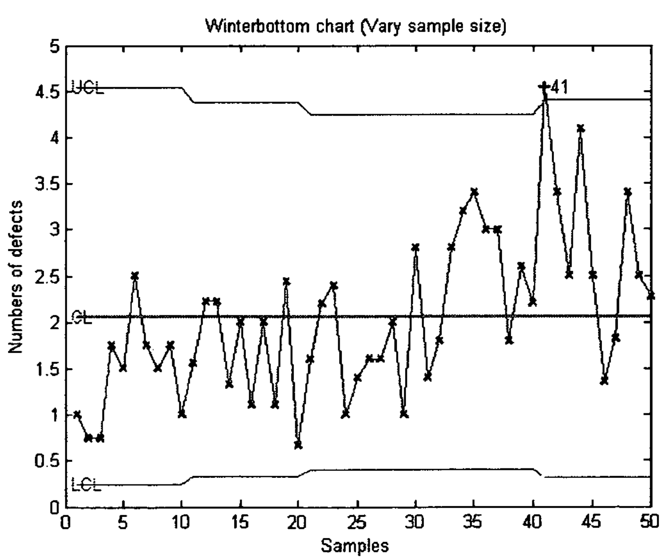


Figure 2.12: Winterbottom's C-chart for vary sample sizes

The (Almost) Exact Control Limits for a C-chart

In general, to improve attribute control chart, many researches introduce many approaches for transforming attribute data from Poisson distribution to nearly normal distribution. However, instead of transforming data, Rudolf proposed his control chart which achieves three standard deviation limits of defects chart by transforming normal three standard control limits [50].

To optimize control limits, the author presented the basic equation of his control limits as following [50]:

$$c \simeq [\hat{\mu}_y \pm z\hat{\sigma}_y]^{\frac{3}{2}} \tag{25}$$

Where z is usually equal to 3 (three standard deviation) and c can be upper or lower control limit. Then, in his separated work [49], he determined the expected mean and standard deviation on transformation scale as:

$$\hat{\mu}_y \simeq (\bar{c} + \frac{1}{12})^{\frac{2}{3}} \tag{26}$$

$$\hat{\sigma}_y \simeq (\frac{2}{3})(\bar{c})^{\frac{1}{6}} \tag{27}$$

Haldane's work shows that 2/3 power can provide symmetric Poisson transformation [21]. Moreover, before 2/3 power transformation, Read and Cressie work also suggested that the constant (1/4) from Anscombe's Poisson transformation is needed to be added [61]. Therefore, Rudolf gave the transforming control limits equation given by:

$$Y = (c + \frac{1}{4})^{\frac{2}{3}} \tag{28}$$

By substituting estimated mean and variance into the basic form, and then substituting again into the transforming control limits equation, Rudolf's control limits are given by:

$$CL = \bar{c} \tag{29}$$

$$LCL = [(\bar{c} + \frac{1}{12})^{\frac{2}{3}} - 3(\frac{2}{3})(\bar{c})^{\frac{1}{6}}]^{\frac{3}{2}} + \frac{1}{4} \tag{30}$$

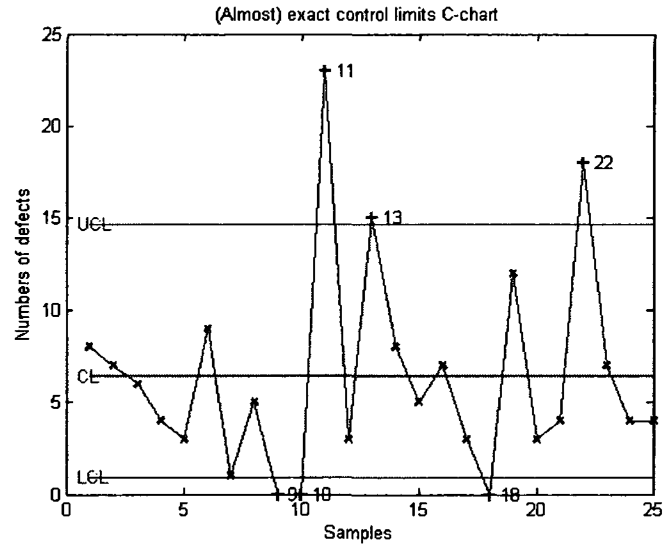


Figure 2.13: The (Almost) Exact Control Limits for a C-chart

$$UCL = [(\bar{c} + \frac{1}{12})^{\frac{2}{3}} + 3(\frac{2}{3})(\bar{c})^{\frac{1}{6}}]^{\frac{3}{2}} - \frac{3}{4} \tag{31}$$

By equation 29, 30 and 31, the (Almost) Exact Control Limits (ECL) can be constructed. To illustrate the chart, the data in table 2.1 is plotted and shown in figure 2.13.

2.2 Experimental Results: Comparison of the Different Defects Charts Performances

2.2.1 Objectives and Methodology

The main objectives of this section are:

- To compare control charts by categories, and find the best chart in each class.
- To find the lowest mean that the lower control limit of each chart can provide.
- To find the control chart that provides the lowest loss due to in-control state.
- To find the control chart that provides the highest sensitivity due to mean shifting.

The data sets used to test the different control charts were generated by Poisson random number generation using MATLAB 7 build in function as follows:

- Each data set contains 100,000 data.

Table 2.5: List of transforming equation and control limits of defects chart

	Transforming Equation	CL	LCL	UCL
Bartlett	$y = 2\sqrt{c}$	$\bar{y}$	$\bar{y} - 3$	$\bar{y} + 3$
Anscombe	$y = 2\sqrt{c + \frac{3}{8}}$	$\bar{y}$	$\bar{y} - 3$	$\bar{y} + 3$
Freeman&Tukey	$y = \sqrt{c} + \sqrt{c+1}$	$\bar{y}$	$\bar{y} - 3$	$\bar{y} + 3$
ISRT	$y = \sqrt{c}$	$\sqrt{\bar{c}}$	$\sqrt{\bar{c}} - \frac{3}{2} - \frac{9}{8}(\frac{1}{\sqrt{\bar{c}}})$	$\sqrt{\bar{c}} + \frac{3}{2} - \frac{1}{2}(\frac{1}{\sqrt{\bar{c}}})$
Z	$z = \frac{y=n\lambda}{\sqrt{n\lambda}}$	0	-3	3
W	$W = 2\sqrt{y} - 2\sqrt{n\lambda}$	0	-3	3
Q	CDF of Poisson and inverse Guassian	0	-3	3
C	None	$\bar{c}$	$\bar{c} - 3\sqrt{\bar{c}}$	$\bar{c} + 3\sqrt{\bar{c}}$
Ryan&Schwertman	None	$\bar{c}$	$1.5307 + 1.0212\bar{c} - 3.2197\sqrt{\bar{c}}$	$0.6182 + 0.9996\bar{c} + 3.0303\sqrt{\bar{c}}$
Winterbottom	None	$\bar{c}$	$\bar{c} - 3\sqrt{\bar{c}} + 4/3$	$\bar{c} + 3\sqrt{\bar{c}} + 4/3$
ECL	None	$\bar{c}$	$[(\bar{c} + \frac{1}{12})^{\frac{2}{3}} - 3(\frac{2}{3})(\bar{c})^{\frac{1}{6}}]^{\frac{3}{2}} + \frac{1}{4}$	$[(\bar{c} + \frac{1}{12})^{\frac{2}{3}} + 3(\frac{2}{3})(\bar{c})^{\frac{1}{6}}]^{\frac{3}{2}} - \frac{3}{4}$

- 50 data sets were generated by different means from 1 to 50.
- The mean shifts were set to 0.5σ, σ, 1.5σ, and 2σ.

After generating Poisson random numbers, we define LCL and UCL of each control chart. In some cases such as transformed control charts and standardized charts, data were transformed before control limits were defined. All formula used for transformation and the computation of control limits are summarized in table 2.5. After transforming or standardizing data and defining the control limits, the data were plotted and the out of control points were counted and used for calculating ARL values. Then, all ARL values were recorded and plotted to compare the performances.

2.2.2 Performances Comparison Based on responsiveness of lower control limit to low mean samples

Ryan explained that the classic C-chart can not be performed when the mean of defects is low [58]. Indeed, when a control chart is constructed by low mean samples, the lower control limit is always negative. Thus, the lower control limit can not be shown in the control chart since numbers of defects can not be lower than zero. He pointed out that the control chart that its lower control limit can respond to the lowest mean can be considered as high performing control chart in term of monitoring the data at the low mean which can be necessary in high yield process where numbers of defects are very small [58].

Figure 2.14 shows the lowest mean that the lower control limit of each control chart can use for monitoring a process. In this figure, the lowest peak corresponds to the best chart in term of detecting a defect in low mean population. From this figure, we can point out noticeable results:

- Except Z-chart, other charts can perform better than Classic C-chart in this test.

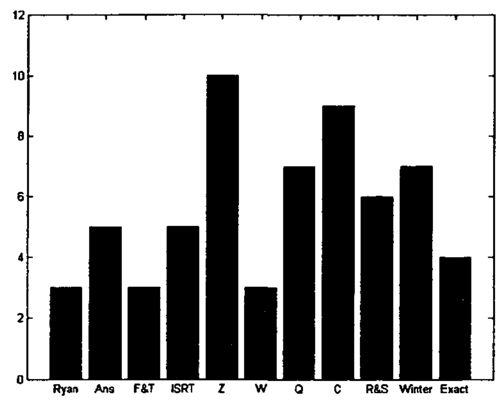
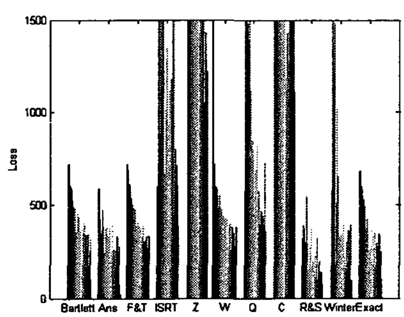
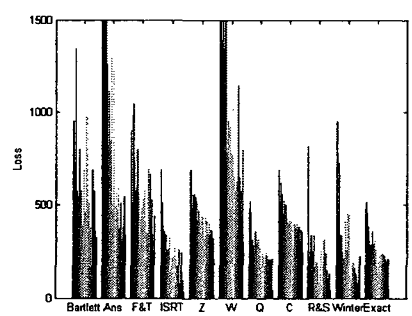


Figure 2.14: The lowest level of mean that the lower control limit of each control chart can monitor in the case of out of control signal. The names of each control chart on X-axis, and the level of mean is shown in the Y-axis.

- The transformation data approach category performed the best in this test. The charts in this category can be applied even though mean of samples is as low as 3 to 5.
- For standardized charts and optimal control limit charts, W transformation control chart and (Almost) Exact control limit chart (ECL) are the best in each category, respectively.



(a)



(b)

Figure 2.15: Loss function of (a) lower control limit and (b) upper control limit of each control chart due to in-control state. The names of each control chart on X-axis, and the level of loss gain is shown in the Y-axis.

2.2.3 Performances Comparison Based on Loss Function

As explained in section 1.1.2, Loss function indicates how much the process signals false alarm when the process is in control. The preferable control charts will be judged by the lowest loss creation which can be shown separately into lower and upper control limits. The Loss function of lower and upper control limits of each control chart are shown in figures 2.15.a and 2.15.b, respectively. The lowest peak corresponds to the best chart in term of loss gain. Each bar represents the loss gain of each chart. The top of each bar is not flat since each bar is contained with loss gaining of each mean samples from 1 to 50. The chart that has logarithmic-like top shows the changed level of loss gain due to the mean changes from 1 to 50. There are some noticeable points which can be pointed out according to lower control limit performance:

- The result of lower control limit shows that the optimized control limit category performs the best among other classes. The optimal control limit C-chart has the best performance since the loss function is laid on the bottom of the graph. The other two optimized control limit charts which are the Winterbottom's chart and the ECL chart also gain low loss.
- Transformed data category has the second performance which shows slightly higher loss function than the optimized control limit category, except ISRT chart that creates very high cost function.
- For the standardized control charts, W transformation can perform at low loss function, and Q-chart can perform low loss only in case of high mean of samples.
- In term of improvement compared to classic C-chart, there are only ISRT and Z-charts that provide unsatisfied result.

There are also some interesting points that can be mentioned according to upper control limit performances:

- The optimizing control limit category performed the best according to this test. The result of upper control limit shows that the optimal control limit can perform excellently among defects charts. In this category, Winterbottom's chart also gains the low loss function due to upper control limit except in the case of low mean (slightly high loss function), and ECL chart draws slightly higher loss than the two other charts.

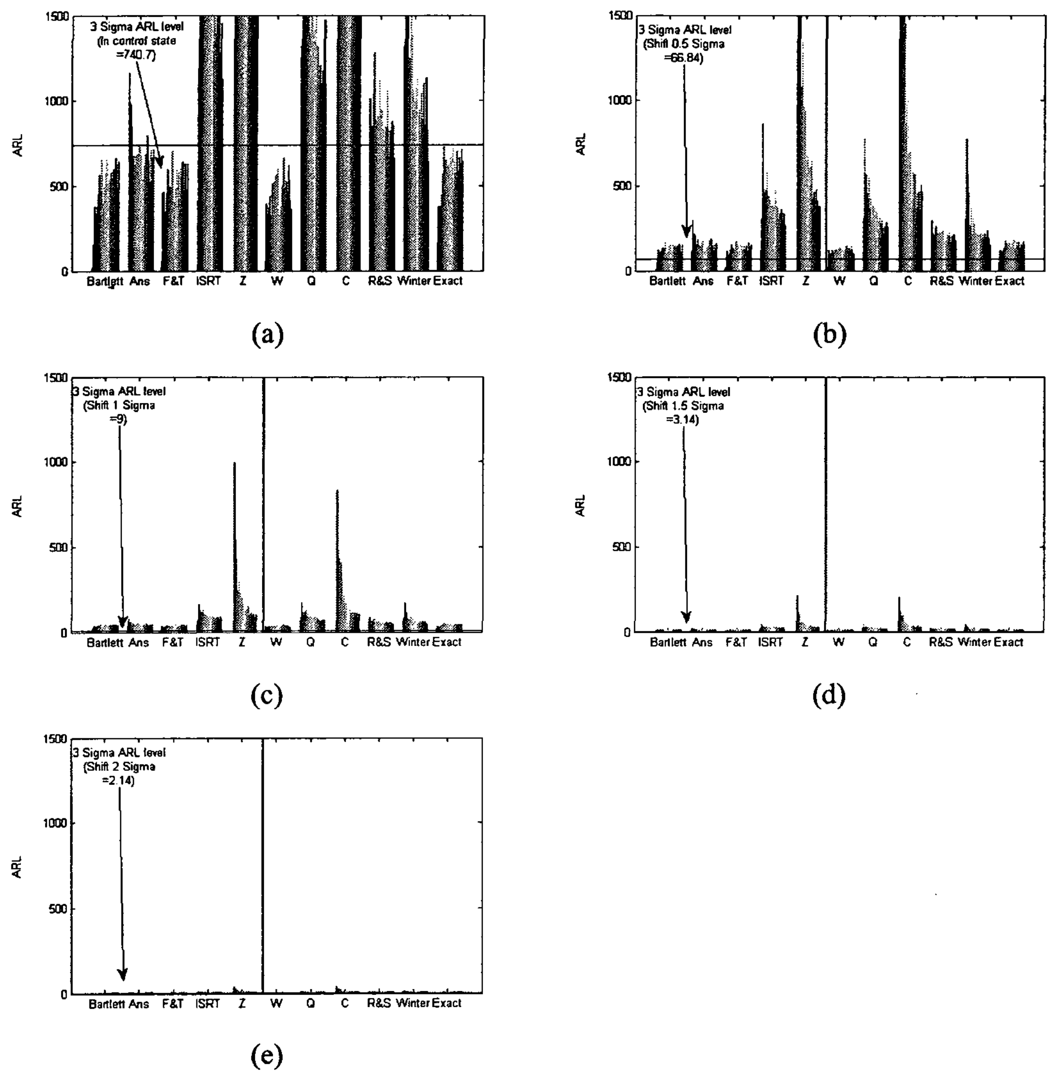


Figure 2.16: ARL_L of each control control chart in different states (a) no shifting (b) 0.5σ (c) σ (d) 1.5σ (e) 2σ

- For groups of transformed control charts and standardized control charts, ISRT chart and Q-chart can perform at the low wastes as closed as the level of optimizing control limit category does.
- Bartlett, Anscombe, Freeman and Tukey, and W transformation charts create even higher cost than classic C-chart.

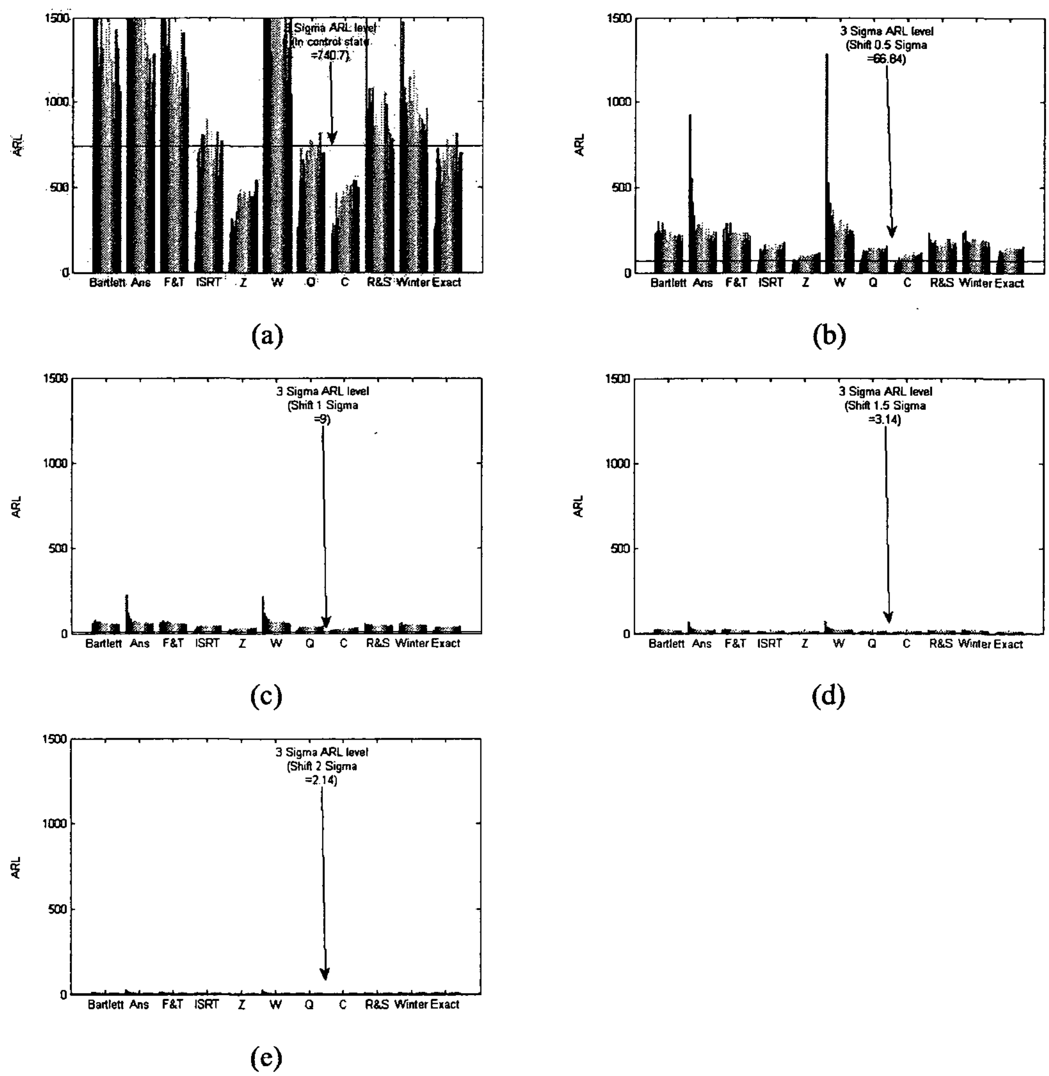


Figure 2.17: ARL_U of each control control chart in different states (a) no shifting (b) 0.5σ (c) σ (d) 1.5σ (e) 2σ

2.2.4 Performance Comparisons Based on Mean Shifting Sensitivity

In order to define the effective mean shifting detection of each chart, the best performance chart will be judged by ARL interpretation (see section 1.1.2). The preferable level of ARL during in control state should be high and close to in control nominal ARL level (see table 1.1). For the preferable level of ARL during the out-of-control state, ARL level should be as low as the out-of-control ARL nominal values. Furthermore, to compare the sensitivity of each chart, the performance will be evaluated by separated control limits.

In figure 2.16, the graphs show the ARL levels of lower control limit due to shifting of σ level. Figure 2.16.a shows the ARL level of each chart during the in control state, figure 2.16.b, 2.16.c, 2.16.d, and 2.16.e show the ARL levels during means of samples shifting by 0.5σ, σ, 1.5σ, and 2σ, respectively. From figure 2.16, we can point out noticeable remarks:

- All ARL levels of the different charts from in-control state to out-of-control state show significant changes.
- The optimizing control limit category has the highest mean shifting sensitivity compared to the other categories. The optimal control limit C-chart and Winterbottom's chart have the best performance.
- There are three charts (Bartlett, Anscombe, Freeman and Tukey charts) in transforming data category, and ECL chart have the second high performance due to mean shifting sensitivity.
- For the last group which is standardizing data approach, Q-chart has the highest mean shifting sensitivity and show improved performance compared to the classic C-chart.
- Z-chart and ISRT chart have almost the same result as classic C-chart, and W transformed chart performs less efficient than the classic C-chart.

In figure 2.17, the graphs shows the ARL levels of upper control limit due to shifting of σ level. Figure 2.17.a shows the ARL level of each chart during the in control state. Figures 2.17.b, 2.17.c, 2.17.d, and 2.17.e show the ARL levels during mean of simple shifting by 0.5σ, σ, 1.5σ, and 2σ. From figure 2.17, we can point out the following remarks:

- All ARL levels of upper control limit of the different charts from in-control state to out of control state show significant changes.
- The optimizing control limit category has the best performance. The optimal control limit C-chart and Winterbottom's chart have the highest mean shifting sensitivity.
- The second high sensitivity is given by ISRT chart, Q-chart, and ECL chart. ISRT chart and Q-chart have the best sensitivity in their own category.
- For other charts from the transforming data category, the results show an acceptable level compared to the Classic C-chart.
- Z-chart and W transformed chart poorly perform due to mean shifting toward upper control limit.

At this point, we have tested all charts according to different considerations. To illustrate the performances of all charts, table 2.6 shows a summary of the different defect control charts performances and give the strong and weak aspects of each chart.

Table 2.6: Full comparison of all defects charts

	Vary sample size	Plotted data	Min λ**	LCL Loss gain		UCL Loss gain		λ Shifting Sensitivity	
				λ = 1 to 25	λ = 26 to 50	λ = 1 to 25	λ = 26 to 50	LCL	UCL
Bartlett	No	Transformed	3	Medium	Low	Medium	Medium	High	Low
Anscombe	No	Transformed	5	Low	Very low*	Very high	Medium	High	Low
Ft	No	Transformed	3*	Medium	Low	Medium	Medium	High	Low
ISRT	No	Transformed	5	High	High	Low	Very low*	Medium	Very high
Z-chart	No	Standardized	10	Very high	Very high	Medium	Low	Medium	Very Low
W chart	No	Standardized	3*	Medium	Low	Very high	Medium	Very Low	Medium
Q-chart	Yes*	Standardized	7	High	Medium	Low	Low	Low	High
C-chart	No	Original	9	Very high	Very high	Medium	Low	Medium	Very Low
Optimal	No	Original	6	Very low*	Very low*	Low	Very low*	Very high*	Very high*
Winter	Yes*	Original	7	High	Low	Low	Very low*	Very high*	Very high*
ECL	No	Original	4	Medium	Low	Low	Low	High	High

*Preferred chart. ** Minimal λ that LCL can respond to out-of-control point. Low value are preferred.

2.3 Conclusion

In this chapter, we compared 11 defects charts that have been proposed for statistical process control. The results shows that optimizing control limits approach is the best approach to acquire a defects chart. By this approach, there is no requirement of transforming or standardizing data, therefore, plotted data still maintain their original meaning in control chart. From experimental results, optimal C-chart shows outstanding results. By lowest loss gain and highest mean shifting sensitivity, we concluded that this chart is the best defects chart to replace traditional C-chart in attribute monitoring.

Chapter 3

An Optimal Bivariate Poisson Field Chart

Shewhart C-Chart is a widely accepted control chart for monitoring number of defects per unit. However, in high-quality process, where normal assumption is impractical, and characteristics are correlated, C-chart becomes unsuitable. In this chapter, we propose an optimal bivariate Poisson field chart to monitor two correlated characteristics of count data for both industrial and non-industrial purposes. This chart is based on optimization of bivariate Poisson confidence interval and illustration of bivariate Poisson data in Poisson field. The detailed description of our proposed algorithm is presented by numerical data. The comparative results present 3-dimensional visualization and improved false alarm rate of our proposed algorithm compared to existing approaches. The performances of our proposed algorithm is presented by both real case study and simulation. The experimental results demonstrate improved performances regarding visualization and false alarm rate.

3.1 Introduction

To deal with number of defects, C-chart is the most widely used tool in statistical process control. There are many works that have improved C-chart [58] [60] [42] [8] [50] [2] [63]. However, when numbers of defects are very low such as in high quality processes, traditional C-chart become unsuitable tool. Instead of focusing on numbers of defects or fraction of nonconforming items, counting numbers of conforming items between the occurrences of non-conforming items is introduced, and refereed to “interevent counts”. In [59], Goh introduced

CCC-chart which can be used for monitoring interevent counts, by pointing out the effect of low fraction of nonconforming (i.e where small probability of nonconforming occurred). Indeed, in this case, normal approximation becomes out of reality. By using actual numerical examples, Xie and Goh presented application of CCC-chart, and suggested some methods for decision making in high yield processes [41] [40]. In [56], numbers of interevent counts are transformed by simple power transformation from exponential distribution to Weibull distribution. Then, normal approximation is applied to construct control limits. In [4], authors applied generalized Poisson distribution to model over-dispersed data, and suggested the use of CCC-chart for high quality process monitoring. Furthermore, in [17] [4], authors agreed that hypothesis test or histogram should be conducted in initial state of constructing the charts for high yield process. However, although monitoring interevent counts is preferable, to observe various types of defects simultaneously, multiple C-charts and CCC-charts are needed.

Monitoring two or more types of correlated characteristics in high quality process still leave room for improvement. Lowery and Montgomery pointed out that multivariate control charts perform better to signal out of control alarms than univariate charts, since correlation between variables is taken into account [5]. They also suggested that univariate charts are only suitable for diagnosing process behavior. In [53], the authors pointed out four conditions that every control chart needs to satisfy: "Is the process in-control?", "Is out-of control state pointed out?","Is relationship between two or more variables taken into account?", and "what is the problem that out-of-control signal actually tells?". According to [53], there are many alternative charts which are based on improving χ^2 and T^2 charts for continuous data. However, for discrete variables, few multivariate attribute charts such as [20], [24], [30], and [67] have been proposed. In [20], Patel presented his multivariate control chart for both binomial and Poisson data. For multiple defects, he presented multivariate Hotelling-like chart where time dependency between variables is considered. However, this chart is not practical to apply in nearly zero defect processes, since it considers normal assumption, and requires complicated steps to construct the chart [57] [22]. In [67], to deal with multi-attribute variables, improved Mnp-chart is presented by considering correlation between characteristics [67]. Not only this chart shows improved results compared with univariate p-chart, but it is also simple. Moreover, Joel have shown another enhanced Mnp-chart for multiple independent discrete variables by proposing simple designing of optimal Mnp-chart [24]. Nonetheless, both Mnp-charts are not practical for high quality processes, since these chart are constructed under normal assumption. Skinner et al. have suggested to use generalized linear model (GLM) to construct attribute control chart for multiple counts where input and output variables are measurable. By observing the residuals, Generalized linear model-based control charts are more effective to monitor multi-count data than C-chart. Furthermore, the results show effective performance in the case of overdispersion. However, inputs and outputs are not measurable in every process. Besides, GLM based charts require multiple charts to observe multivariate variables. In [57], the authors suggested two transformations for multivariate Poisson distribution. For the first transformation, they applied bisection method to find the proper power of the root transformation of each attribute characteristics. The second transformation is Normal distribution To Anything (NORTA) inverse transformation method. After acquiring almost zero skew distribution from both transformations, χ^2 control chart is applied. According to this paper, NORTA inverse transformation method shows robust performance when dealing with correlated multivariate Poisson data. Moreover, it needs less complex steps than other charts. In [22], the authors presented the use of multivariate Poisson sum probability density function to define the control limits of multivariate Poisson sum chart (MPSUM chart). By their chart, monitoring multiple attribute characteristics can be done in single chart. However, in high quality processes, numbers of defects are very low, and correlation between pairwises of two characteristics is crucial in some processes. According to our knowledge, there are no works that have provided a chart which robustly monitors correlated characteristics. Moreover, none of the charts is mainly concerned with the illustration of how pairwises of characteristic spread which can reflect process behavior.

In this chapter, we propose an optimal bivariate Poisson field chart for monitoring two correlated characteristics of defects. The basic concept is defining the optimal limit of bivariate Poisson distribution and illustrating data in Poisson field. This chart provides satisfactory rate of false alarms, and illustrate original values of two attribute characteristics and changes of correlation between them.

In Section 3.2, the basic concept of bivariate Poisson distribution is briefly discussed. In Section 3.3, the basic principals of an optimal bivariate Poisson field chart are introduced. Finally, real case study and simulations are presented to illustrate the effectiveness of our control chart in Section 3.4, and this chapter is concluded in Section 3.5.

3.2 Bivariate Poisson distribution and its estimation

Bivariate Poisson distribution (BP) is often used for modelling pairwises of correlated count data. Bivariate Poisson distribution was firstly introduced by Campbell [27]. Let random variables X_1, X_2, X_3 are unobserved variables which follow independent Poisson distribution with parameters $\lambda_1, \lambda_2, \lambda_3$. Then, $X = X_1 + X_3$ and $Y = X_2 + X_3$ are observed pairwises which follow jointly a bivariate Poisson distribution $BP(\lambda_1, \lambda_2, \lambda_3)$ with joint probability function [44] [54] [43]:

$$P_{BP}(X = x, Y = y|\lambda_1, \lambda_2, \lambda_3) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \frac{\lambda_1^x}{x!} \frac{\lambda_2^y}{y!} \sum_{i=0}^{min(x,y)} \binom{x}{i} \binom{y}{i} i! (\frac{\lambda_3}{\lambda_1 \lambda_2})^i \quad (1)$$

for $i = 1, 2, ...$

where

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \quad and \quad \bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \quad (2)$$

Where n is total number of samples. The marginal distribution of X and Y with mean $\lambda_1 + \lambda_3$ and $\lambda_2 + \lambda_3$ are also following recurrence relations [44]:

$$xP(x, y) = \lambda_1 P(x - 1, y) + \lambda_3 P(x - 1, y - 1)) \quad (3)$$

$$yP(x, y) = \lambda_2 P(x, y - 1) + \lambda_3 P(x - 1, y - 1) \quad (4)$$

For maximum likelihood estimation, if equation 1 is differentiated with respect to parameters λ_1, λ_2, and λ_3, from recurrence relation in 3 and 4, the differential-different equations are given by [44]:

$$\frac{\partial P(x, y)}{\partial \lambda_1} = P(x - 1, y) - P(x, y) \quad (5)$$

$$\frac{\partial P(x, y)}{\partial \lambda_2} = P(x, y - 1) - P(x, y) \quad (6)$$

$$\frac{\partial P(x, y)}{\partial \lambda_3} = P(x, y) - P(x - 1, y) - P(x, y - 1) + P(x - 1, y - 1) \quad (7)$$

Using the above and the recurrence relations, Holgate also showed that [44]:

$$\sum \frac{1}{P}\frac{\partial P}{\partial \lambda_1} = \sum \frac{1}{P}\frac{\partial P}{\partial \lambda_2} = \sum \frac{1}{P}\frac{\partial P}{\partial \lambda_3} = 0 \tag{8}$$

Which can be written as:

$$\frac{\bar{x}}{\lambda_1} - \frac{\lambda_3}{\lambda_1}\bar{R} - 1 = 0 \tag{9}$$

$$\frac{\bar{y}}{\lambda_2} - \frac{\lambda_3}{\lambda_2}\bar{R} - 1 = 0 \tag{10}$$

$$\frac{\bar{x}}{\lambda_1} + \frac{\bar{y}}{\lambda_2} - (1 + \frac{\lambda_3}{\lambda_1} + \frac{\lambda_3}{\lambda_2})\bar{R} - 1 = 0 \tag{11}$$

Where

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad and \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{12}$$

and

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} \frac{P(x_i - 1, y_i - 1}{P(x_i, y_i)}) \tag{13}$$

Therefore,

$$\bar{x} = \lambda_1 + \lambda_3 \quad and \quad \bar{y} = \lambda_2 + \lambda_3 \tag{14}$$

and

$$\bar{R} = 1 \tag{15}$$

Since parameters $\vec{\lambda} = (\lambda_1, \lambda_2, \lambda_3)$ of bivariate Poisson distribution in equation 14 are shown in decomposed form, the method to estimate these three parameters have received attentions [44] [54] [43] [34] [28]. Correlation between x and y can be calculated by [54] [43]:

$$\rho_{xy} = \frac{\lambda_3}{\sqrt{(\lambda_1 + \lambda_3)(\lambda_2 + \lambda_3)}} \tag{16}$$

Correlation (ρ) of bivariate Poisson has value between 0 and 1 [54]. When ρ is zero, λ_3 become zero and bivariate Poisson distribution can be referred to double Poisson distribution [54] [10]. Furthermore, from the mean decomposition in equation 14, the covariance variance matrix of bivariate Poisson distribution provides only non-negative λ_3 [54]. Thus, this distribution permits only non-negative correlation. For this issue, [45] pointed out that there are few cases of bivariate Poisson distribution that can provide the negative correlation.

3.3 Proposed Method

An optimal bivariate Poisson field chart is based on the probabilistic optimization of Bivariate Poisson confidence interval and projection of bivariate Poisson histogram or Poisson field. Suppose two correlated characteristics of defects are pairwises of (x, y) with parameters $\vec{\lambda} = (\lambda_1,\lambda_2,\lambda_3)$. First, $\vec{\lambda}$ is estimated from complete samples $(x_i, y_i), i = 1, 2, 3, \ldots, n$, and the expected type I error rate (α_0) is specified.

Second, instead of presenting number of frequencies of bivariate Poisson pairwise as in bivariate field, our poisson field presents probability of each pairwise. By equation 1 and estimated parameters ($\vec{\lambda}$), probabilities of all pairwises ($p_{BP}(r, s|\vec{\lambda})$) are calculated from pair $(0, 0)$ until all numbers on row and column are equal to zeros as shown in table 3.1.

Table 3.1: Probabilistic Poisson field

		y							
		0	1	2	3	.	.	.	s_{max}
	0	p_{00}	p_{01}	p_{02}	p_{03}	.	.	.	0
	1	p_{10}	p_{11}	p_{12}	p_{13}	.	.	.	0
	2	p_{20}	p_{21}	p_{22}	p_{23}	.	.	.	0
x	3	p_{30}	p_{31}	p_{32}	p_{33}	.	.	.	0
	.	.	.	.	.	.	.	.	0
	.	.	.	.	.	.	.	.	0
	.	.	.	.	.	.	.	.	0
	r_{max}	0	0	0	0	0	0	0	0

Where $p_{00}, p_{10}, p_{01}, p_{11}, \ldots, p_{r_{max}s_{max}}$ are probabilities of each pairwise on probabilistic Poisson field, and r_{max} and s_{max} are numbers of row and column where all probabilities are equal to zero.

The optimization of control limits is a widely used approach to acquire a robust control chart. According to optimal C-chart by Ryan [60] [42], the optimal limits can be defined by closeness of alpha rate to nominal value ($\alpha = 0.0027$ or $ARL_0 = \frac{1}{\alpha_0} = \frac{1}{0.0027} = 370$). To obtain expected alpha rate, Ryan shifts lower control limit from zero and upper control limit from mean until having the expected rate of α.

In the third step, to minimize confidence area, we shift the control limit of our chart from $min(p_{rs})$ to $max(p_{rs})$. Since total probability in Poisson field is equal to 1, and in order to optimize our chart control limit, let $\min(p_{00}, p_{10}, p_{01}, ..., p_{j_{max}k_{max}}) = 0$. Then, the total probability of Poisson field is calculated by:

$$\frac{1}{ARL_{total}} = 1 - \sum_{r=0}^{r_{max}} \sum_{s=0}^{s_{max}} p_{rs} \tag{17}$$

To acquire the optimal Poisson field, the pairwises that have the lowest probability in Poisson probabilistic field are removed until total probability in Poisson field is satisfied:

$$ARL_{total} \geq ARL_0 \tag{18}$$

Where ARL_0 is the expected average run length or nominal average run length ($ARL_0 = \frac{1}{\alpha_0}$). After the total probability in Poisson field is satisfied given condition in equation 10, the optimal Bivariate Poisson control limit can be obtained.

3.4 Example and Numerical Results

The purpose of this section is to present the effectiveness of our chart based on actual count data with known assignable causes, and compare performances of our chart to previously proposed charts such as NORTA chart and MPSUM chart by testing each chart with simulated datasets.

In [57], NORTA chart is based on transformation of multi-attribute data into almost symmetric distributions. By this approach, every discrete random vector will be transformed by Q-transformation from Poisson distribution to Q-distribution. Let x_{ij} is numbers of defects per unit for $i = 1, 2, 3, \ldots, n$ and $j = 1, 2, 3, \ldots, p$. Where n is number of total samples and p is numbers of characteristics. For Q-transformation, initially, $u_{ij} = F(x_{ij}; \lambda_j)$ is the Poisson cumulative distribution function ($F(.)$) for transforming Poisson variable to percentile. Then, $Q_{ij} = \Phi^{-1}(u_{ij})$ for transforming percentile to Q-statistic variable. Then, T^2 control chart will be applied to transformed vectors (following normal distribution) in order to plot them in the chart. Let $Q_i = (Q_{i1}, Q_{i2}, \ldots, Q_{ij})$ be the transformed vector. T^2-statistics can be calculated by:

$$T_i^2 = (Q_i - \bar{Q})^T S^{-1} (Q_i - \bar{Q}) \tag{19}$$

Where S is an estimated population covariance matrix, which is constant in the process. The lower control limit of T^2 control chart is always zero and the upper control limit of this chart can be calculated by Chi-square distribution with p degrees of freedom.

In [22], Chiu and Kuo presented their chart which is based on the exact distribution of sum of multivariate Poisson variables. By Hermit distribution which was presented as the special case of Poisson "doublet" variables in [6], the exact distribution of sum of Bivariate Poisson variables (SUMMTP) is:

$$P_{SUMMTP}(d|\vec{\lambda}) = e^{-(\lambda_1+\lambda_2+\lambda_3)} \sum_{i=0}^{d/2} \frac{(\lambda_1+\lambda_2)^{d-2i}(\lambda_3)^i}{(d-2i)!(i)!} \tag{20}$$

Where $d_i = x_i + y_i$. By giving α_0 and mean parameters $\vec{\lambda}$, the upper control limit is defined by $P(d > UCL|\vec{\lambda}) \leq \alpha_0$ and the lower limit by $P(d < LCL)|\vec{\lambda}) \leq \alpha_0$. Then, the control limits of their control chart can be defined by probabilistic approach to acquire the optimal limit of the exact distribution of sum of multivariate Poisson distribution. By this chart, multiple-type of count data can be monitored in single chart.

3.4.1 Example

Consider the situation where counting numbers of people flow in and out of the building are monitored in order to predict an event such as a conference in the building. These data which are called "CalIt2" data were collected over 15 weeks for every half hour from the main door of the CalIt2 building at University of California (UCI) [1]. The dataset contains 10080 samples (5040 samples for each in and outflow) from 07/24/05 to 11/05/05, and schedule of events (See table 3.2).

Table 3.2: The schedule of events in Calit2 building from 07/24/05 to 11/05/05.

Dates of events (Month/Day/Year)	Starting time (Hour:Minute)	Finishing time (Hour:Minute)
7/26/2005	11:00	14:00
7/29/2005	8:00	11:00
8/2/2005	15:30	16:30
8/4/2005	16:30	17:30
8/5/2005	8:00	11:00
8/9/2005	11:00	14:00
8/9/2005	8:00	16:00
8/10/2005	8:00	16:00
8/12/2005	8:00	11:00
8/16/2005	11:00	14:00
8/18/2005	8:00	17:00
8/18/2005	18:00	20:30
8/19/2005	8:00	11:00
8/23/2005	11:00	14:00
08/26/05	08:00	11:00
08/30/05	16:00	18:00
09/01/05	14:00	16:30
09/15/05	08:30	10:00
09/21/05	09:00	14:00
09/22/05	14:00	14:30
10/03/05	15:30	17:00
10/04/05	12:00	15:00
10/07/05	09:00	10:30
10/10/05	16:30	19:00
10/14/05	09:00	10:30
10/19/05	22:00	23:30
10/21/05	09:00	10:30
10/23/05	21:00	22:30
10/24/05	08:00	12:00
10/24/05	16:00	21:00

In this example, to construct our control chart for detecting events, bivariate Poisson Parameters were estimated from the data containing only non-event periods. Let (x, y) represent pairwises of in flow and out flow of people, respectively. By maximum likelihood estimation, $\vec{\lambda} = (1.56, 1.76, 2.03)$. Then, defining probability of each pairwise by using equation 1 provides Poisson field in table 3.3.

Table 3.3: $p(x, y|1.56, 1.76, 2.03)$in Poisson Field

		y													
	Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13
	0	0.0047	0.0084	0.0074	0.0043	0.0019	0.0007	0.0002	0	0	0	0	0	0	0
	1	0.0074	0.0227	0.0284	0.0217	0.0117	0.0049	0.0017	0.0005	0.0001	0	0	0	0	0
	2	0.0058	0.0252	0.0452	0.0458	0.0311	0.0157	0.0063	0.0021	0.0006	0.0001	0	0	0	0
	3	0.003	0.017	0.0405	0.0544	0.0472	0.0293	0.0139	0.0053	0.0017	0.0005	0.0001	0	0	0
	4	0.0012	0.0081	0.0244	0.0418	0.046	0.0354	0.0203	0.0091	0.0034	0.001	0.0003	0.0001	0	0
x	5	0.0004	0.003	0.0109	0.0229	0.0313	0.0297	0.0207	0.0111	0.0048	0.0017	0.0005	0.0001	0	0
	6	0.0001	0.0009	0.0039	0.0096	0.0159	0.0183	0.0154	0.0099	0.005	0.0021	0.0007	0.0002	0.0001	0
	7	0	0.0002	0.0011	0.0033	0.0063	0.0087	0.0088	0.0067	0.004	0.0019	0.0008	0.0003	0.0001	0
	8	0	0	0.0003	0.0009	0.0021	0.0033	0.0039	0.0035	0.0025	0.0014	0.0006	0.0002	0.0001	0
	9	0	0	0.0001	0.0002	0.0006	0.001	0.0014	0.0015	0.0012	0.0008	0.0004	0.0002	0.0001	0
	10	0	0	0	0	0.0001	0.0003	0.0004	0.0005	0.0005	0.0004	0.0002	0.0001	0	0
	11	0	0	0	0	0	0.0001	0.0001	0.0002	0.0002	0.0001	0.0001	0.0001	0	0
	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0

To optimize control area, the low probabilities in the field are removed until total sum of probabilities in Poisson field provides the expected average run length (ARL_0). This optimal control area is presented in table 3.4.

Table 3.4: The optimal area for Callt2 data

		y													
	Number	0	1	2	3	4	5	6	7	8	9	10	11	12	13
	0	0.0047	0.0084	0.0074	0.0043	0.0019	0.0007	0.0002	0	0	0	0	0	0	0
	1	0.0074	0.0227	0.0284	0.0217	0.0117	0.0049	0.0017	0.0005	0	0	0	0	0	0
	2	0.0058	0.0252	0.0452	0.0458	0.0311	0.0157	0.0063	0.0021	0.0006	0	0	0	0	0
	3	0.003	0.017	0.0405	0.0544	0.0472	0.0293	0.0139	0.0053	0.0017	0.0005	0	0	0	0
	4	0.0012	0.0081	0.0244	0.0418	0.046	0.0354	0.0203	0.0091	0.0034	0.001	0.0003	0	0	0
x	5	0.0004	0.003	0.0109	0.0229	0.0313	0.0297	0.0207	0.0111	0.0048	0.0017	0.0005	0	0	0
	6	0	0.0009	0.0039	0.0096	0.0159	0.0183	0.0154	0.0099	0.005	0.0021	0.0007	0.0002	0	0
	7	0	0.0002	0.0011	0.0033	0.0063	0.0087	0.0088	0.0067	0.004	0.0019	0.0008	0.0003	0	0
	8	0	0	0.0003	0.0009	0.0021	0.0033	0.0039	0.0035	0.0025	0.0014	0.0006	0.0002	0	0
	9	0	0	0	0.0002	0.0006	0.001	0.0014	0.0015	0.0012	0.0008	0.0004	0.0002	0	0
	10	0	0	0	0	0	0.0003	0.0004	0.0005	0.0005	0.0004	0.0002	0	0	0
	11	0	0	0	0	0	0	0	0	0.0002	0	0	0	0	0
	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0

To compare our optimal bivariate Poisson field chart with previously proposed charts such as NORTA chart and MPSUM chart, ARL_0 of each chart needs to be equalized. In this experiment, ARL_0 for all control charts were equally set to 370 (Nominal value). The accuracy of detecting out of control signal can be measured by ability to detect unusual high numbers of people due to an event. Furthermore, since information about dates of events are given, there are two types of error that can be determined. Type I error (α) when a control chart indicates out of control but there is no event at the time, and Type II error (β) when a control chart indicates in-control but there are some events at the time.

We can see from figures 3.1 to 3.21 that all charts show high number of type I error, since

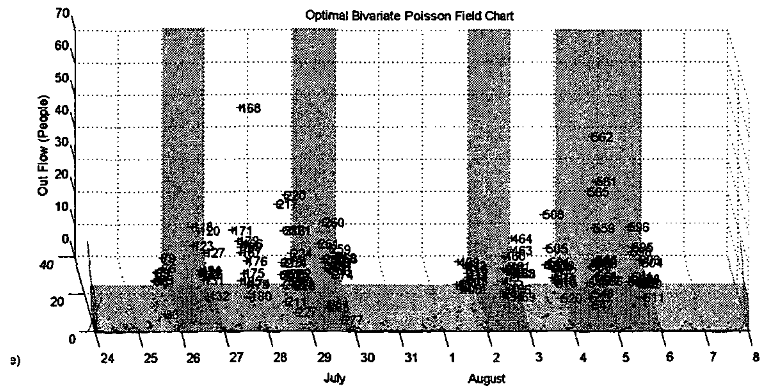


Figure 3.1: An Optimal Bivariate Poisson field chart from samples in 2005 July 23 to August 7

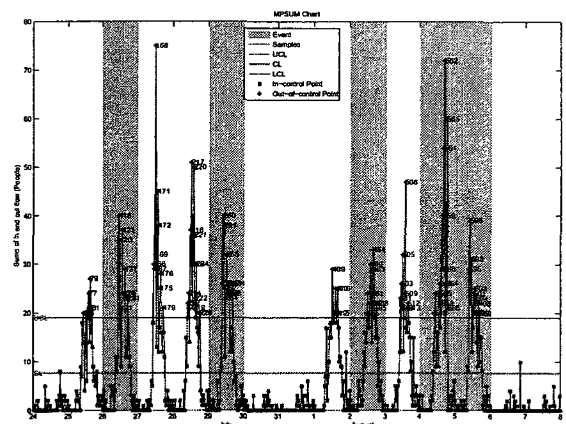


Figure 3.2: MPSUM chart from samples in 2005 July 23 to August 7

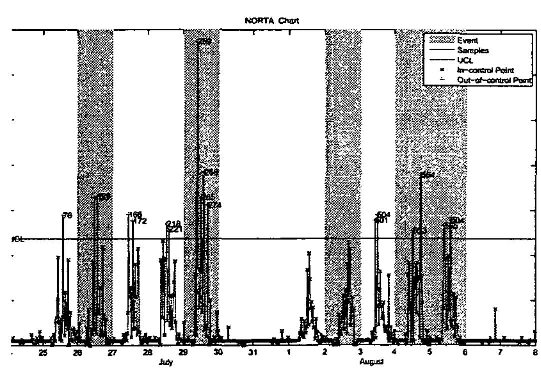


Figure 3.3: NORTA chart from samples in 2005 July 23 to August 7

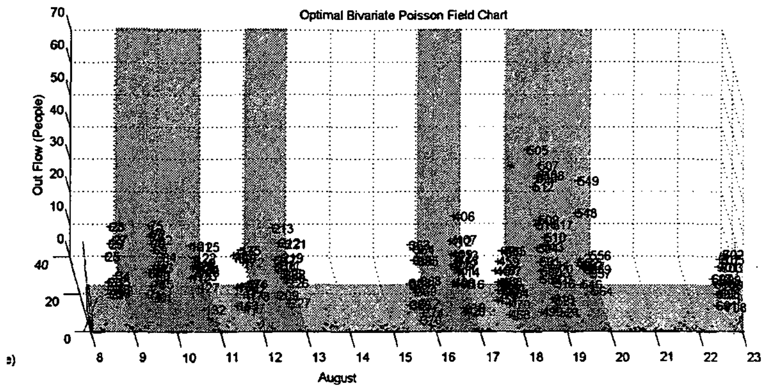


Figure 3.4: An Optimal Bivariate Poisson field chart from samples in 2005 August 8 to 22

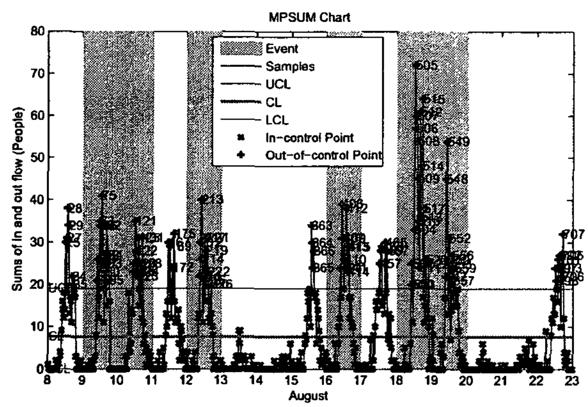


Figure 3.5: MPSUM chart from samples in 2005 August 8 to 22

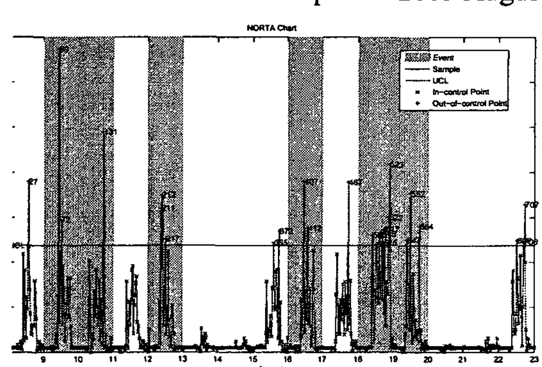


Figure 3.6: NORTA chart from samples in 2005 August 8 to 22

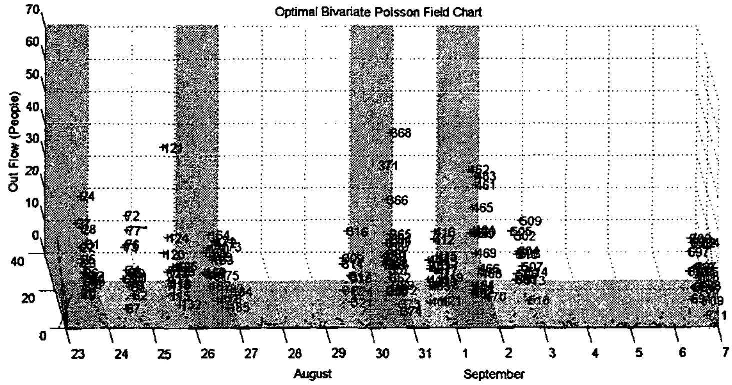


Figure 3.7: An Optimal Bivariate Poisson field chart from samples in 2005 August 23 to September 6

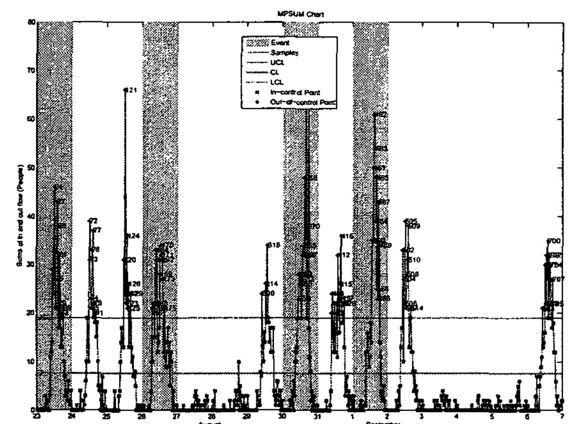


Figure 3.8: MPSUM chart from samples in 2005 August 23 to September 6

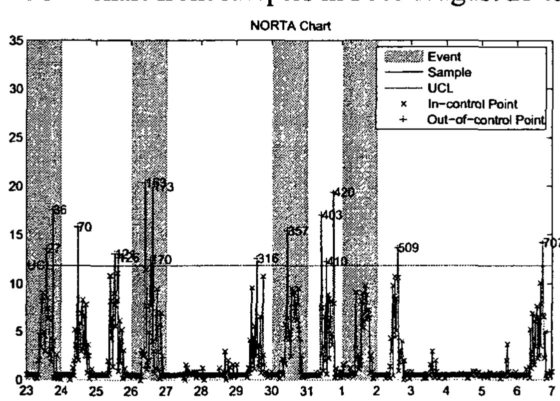


Figure 3.9: NORTA chart from samples in 2005 August 23 to September 6

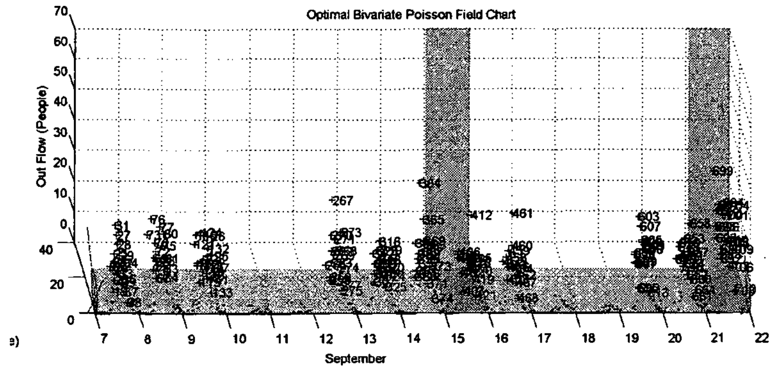


Figure 3.10: An Optimal Bivariate Poisson field chart from samples in 2005 September 7 to 21

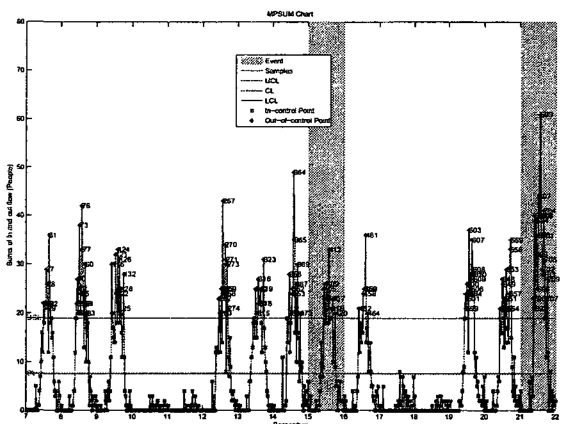


Figure 3.11: MPSUM chart from samples in 2005 September 7 to 21

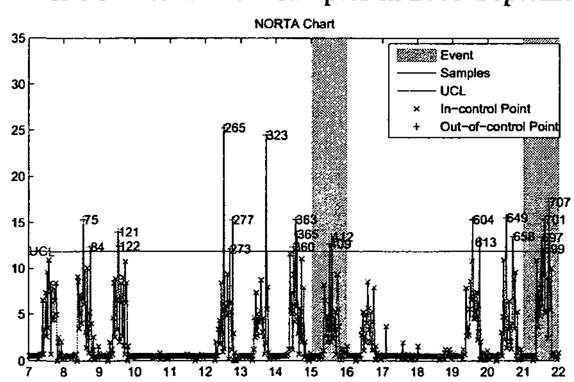


Figure 3.12: NORTA chart from samples in 2005 September 7 to 21

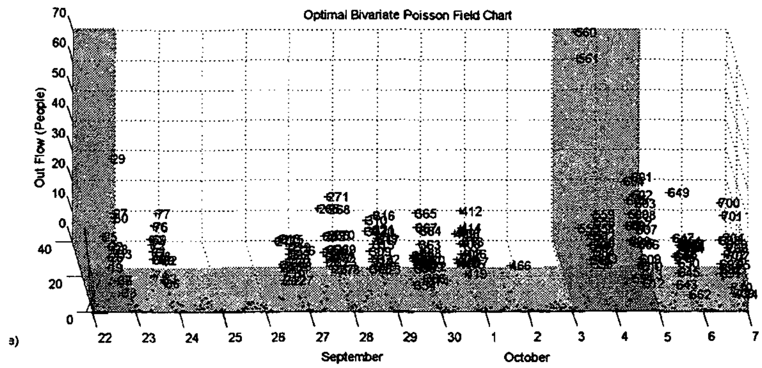


Figure 3.13: An Optimal Bivariate Poisson field chart from samples in 2005 September 22 to October 6

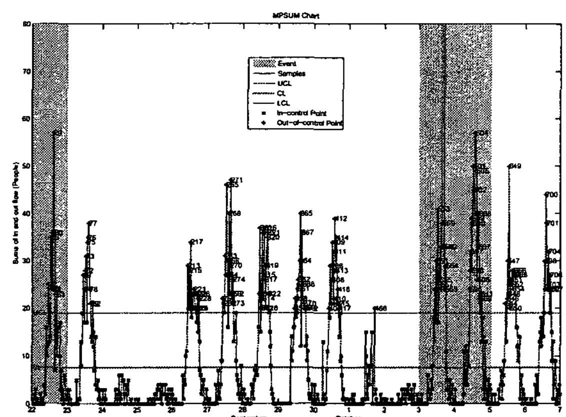


Figure 3.14: MPSUM chart from samples in 2005 September 22 to October 6

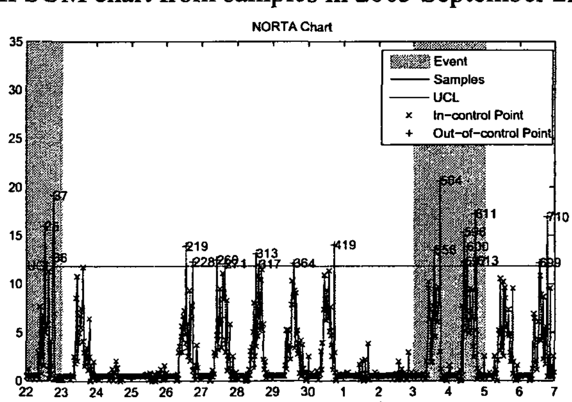


Figure 3.15: NORTA chart from samples in 2005 September 22 to October 6

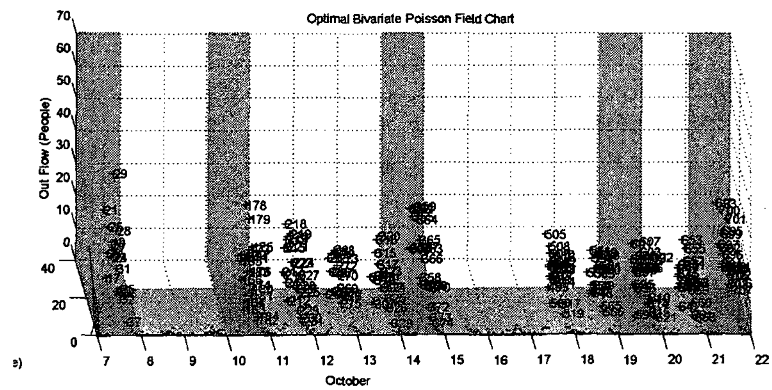


Figure 3.16: An Optimal Bivariate Poisson field chart from samples in 2005 October 7 to 21

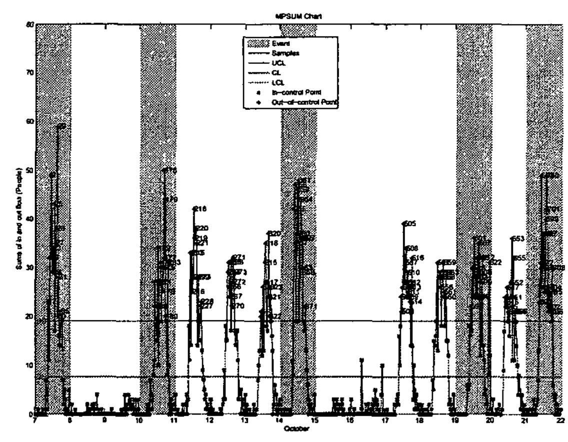


Figure 3.17: MPSUM chart from samples in 2005 October 7 to 21

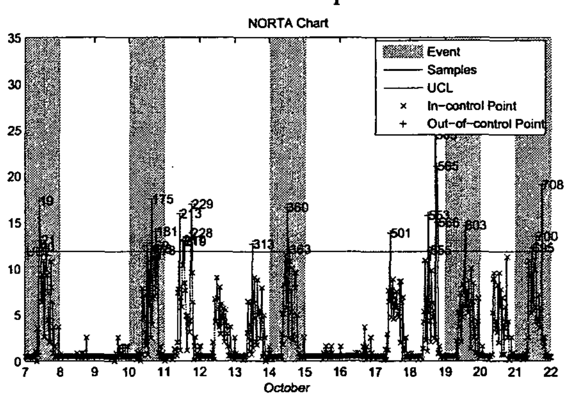


Figure 3.18: NORTA chart from samples in 2005 October 7 to 21

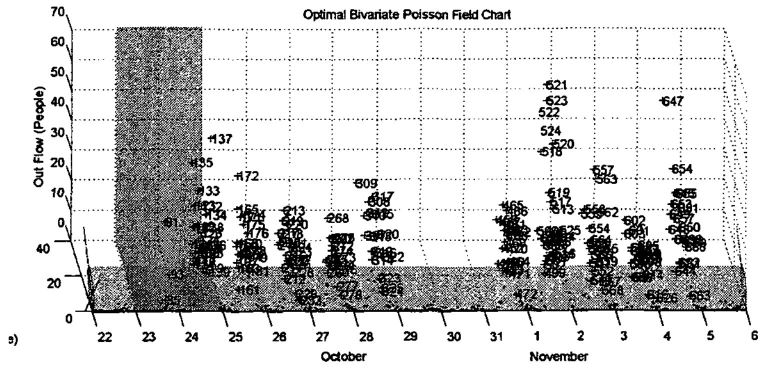


Figure 3.19: An Optimal Bivariate Poisson field chart from samples in 2005 October 22 to November 5

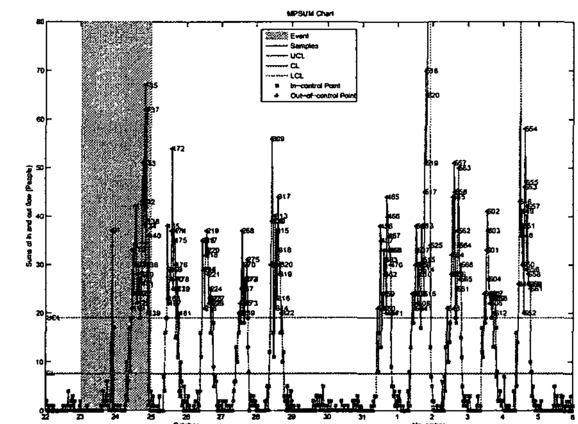


Figure 3.20: MPSUM chart from samples in 2005 October 22 to November 5

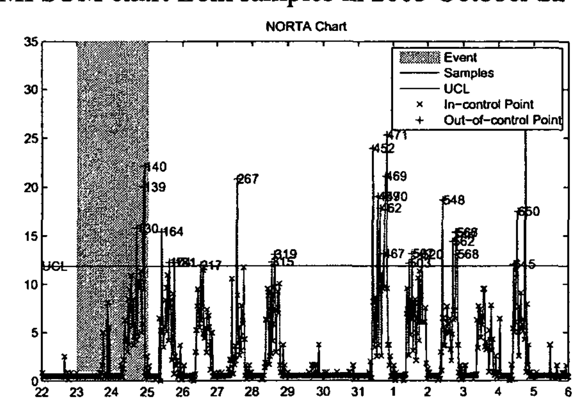


Figure 3.21: NORTA chart from samples in 2005 October 22 to November 5

Table 3.5: Simulation Cases

Case	Means	Correlation	λ_1	λ_2	λ_3
A	Equal	0.8	0.4		1.5
B	Equal	0.5	4		4
C	Equal	0.91	4		40
D	Unequal	0.8	0.1	0.2	1.5
E	Unequal	0.5	1	2	4
F	Unequal	0.91	1	2	40

number of people can be high on the days before events.

Figures 3.1, 3.4, 3.7, 3.10, 3.13, 3.16, and 3.19 clearly show that our proposed control chart can efficiently indicate every unusual high number of in and out flow people due to periods of events.

Figure 3.2, 3.5, 3.8, 3.11, 3.14, 3.17, and 3.20 show that MPSUM chart can also detect events during considered time. However, NORTA chart in figures 3.3, 3.9, and 3.21 shows undetected events on date 8/2/2005, 9/1/2005, and 10/23/2005 (Type II error).

3.4.2 ARL performance

The performance of our proposed control chart is compared to other previously proposed charts by using different simulated datasets. The steps to generate bivariate Poisson random numbers are [68]:

- Generate 2-dimensional normal vector (x_1, x_2) with zero mean and unit variance, and desired correlation (ρ).
- For each vector, calculate the normal cumulative distribution function (CDF).

$$z_i = \Phi(x_i) \tag{21}$$

- For each vector (z_1, z_2), calculate the Poisson inverse cumulative distribution function with desired λ.

$$b_i = F^{-1}(z_i; \hat{\lambda}_i) \tag{22}$$

Where vectors (b_1, b_2) are bivariate Poisson data with desired rates of means ($\hat{\lambda}_1 = \lambda_1 + \lambda_3, \hat{\lambda}_2 = \lambda_2 + \lambda_3$). To present our control chart in various scenarios, the details of the experiment can be summarized as following:

- Datasets are generated by means in table 3.5.
- The expected ARL (ARL_0) was set to 370.

- The mean shift was set to 0.5σ, σ, 1.5σ, 2σ, 2.5σ, and 3σ.

In this experiment, we tested each control chart by 10,000 pairs of simulated data generated according to parameters in table 3.5. To measure the sensitivity of each chart, we tested each in two cases which are shifting means of single and all variables. The results of single counts shifts are given in tables 3.6, 3.7, and 3.8, and the results of all counts shifts are in table 3.9, 3.10, and 3.11.

Table 3.6: ARL levels of NORTA chart for different single mean shifts

Case	Levels of shifting (σ)						
	0	0.5	1	1.5	2	2.5	3
A	10000	33.3333	16.1031	9.9305	6.8074	5.176	4.1911
B	10000	80.6452	47.3934	27.248	20.6612	14.3885	10.8814
C	10000	16.3132	5.4945	2.9257	2.0202	1.5969	1.345
D	10000	9.2593	5.6085	3.8745	3.127	2.6267	2.291
E	10000	14.6413	9.5057	6.8074	5.4201	4.415	3.8314
F	10000	3.076	1.8464	1.4514	1.2367	1.1442	1.0903

Table 3.7: ARL levels of MPSUM chart for different single mean shifts

Case	Levels of shifting (σ)						
	0	0.5	1	1.5	2	2.5	3
A	1428.571	357.1429	208.3333	156.25	114.9425	92.5926	81.3008
B	384.6154	135.1351	120.4819	79.3651	67.5676	46.5116	38.7597
C	103.0928	53.4759	45.8716	32.7869	27.5482	21.978	20.7469
D	1428.571	666.6667	322.5806	147.0588	120.4819	101.0101	80
E	526.3158	526.3158	303.0303	200	119.0476	94.3396	81.3008
F	100	70.922	55.2486	39.8406	31.1526	25.641	24.4499

Table 3.8: ARL levels of the optimal Poisson field chart for different single mean shifts

Case	Levels of shifting (σ)						
	0	0.5	1	1.5	2	2.5	3
A	303.0303	34.965	16.0256	10.1937	7.0423	5.4377	4.4111
B	400	63.6943	36.1011	20.4499	15.5039	11.1235	8.7951
C	454.5455	15.748	5.5586	2.9197	2.0292	1.5926	1.3503
D	434.7826	9.0416	5.5066	3.8212	3.0845	2.5974	2.2696
E	400	16.6945	10.3627	7.5586	5.8445	4.7059	4.0371
F	312.5	3.2927	1.9275	1.4863	1.2579	1.1574	1.0983

Table 3.9: ARL levels of NORTA chart for different both means shifts

Case	Levels of shifting (σ)						
	0	0.5	1	1.5	2	2.5	3
A	10000	55.5556	29.0698	19.1571	14.43	11.4679	9.1827
B	10000	80.6452	46.5116	25.7069	18.9036	13.9276	10.1523
C	10000	114.9425	61.7284	34.7222	24.1546	17.5747	14.245
D	10000	15.3846	11.3636	8.9526	7.4129	6.4433	5.6338
E	10000	21.0084	15.8479	12.1803	9.8232	8.1633	7.0472
F	10000	11.274	9.5602	7.7821	6.7797	5.9809	5.3476

Table 3.10: ARL levels of MPSUM chart for different both means shifts

Case	Levels of shifting (σ)						
	0	0.5	1	1.5	2	2.5	3
A	1428.571	103.0928	50.7614	31.1526	22.2717	17.1233	13.1062
B	384.6154	54.0541	28.3286	17.5439	12.5313	9.1659	6.9686
C	103.0928	25.7069	15.2439	10	7.1736	5.777	4.7237
D	1428.571	142.8571	66.6667	35.2113	24.0385	18.797	14.4509
E	526.3158	125	59.8802	33.7838	22.2717	15.1286	12.1951
F	100	31.25	18.018	11.274	8.0451	6.4392	5.3505

Table 3.11: ARL levels of the optimal Poisson field chart for different both means shifts

Case	Levels of shifting (σ)						
	0	0.5	1	1.5	2	2.5	3
A	303.0303	47.3934	25.3807	17.452	12.9366	10.4712	8.4034
B	400	52.0833	31.8471	17.6678	13.459	10.2041	7.758
C	454.5455	88.4956	48.5437	28.8184	19.4175	14.7929	11.7647
D	434.7826	13.0039	9.2421	7.2046	6.0753	5.2219	4.7059
E	400	17.762	13.369	10.2459	8.2102	6.8166	5.9524
F	312.5	10.3306	8.6505	7.0077	6.1652	5.4496	4.9116

Similarly to bivariate Poisson control charts, the results in table 3.6, 3.7, 3.8, 3.9, 3.10, and 3.11 also show outstanding performances of our optimal bivariate Poisson field chart in different mean samples and mean shifts. In case of in-control state, the proposed algorithm outperforms other charts since ARL levels in different mean samples are close to ARL_0. NORTA chart shows high rate of ARL than ARL_0 that may lead to type II error. Besides, MPSUM chart provides shortage false alarms for monitoring low mean samples, and excessive false alarms for high mean samples. Only, medium mean samples can provide ARL level close to ARL_0. The results of detecting mean shifts also show that optimal bivariate Poisson field chart is more sensitive than the other control charts since ARL levels of our chart effectively respond to mean shifts by decreasing ARL level to low values.

3.5 Conclusion

In this chapter, we have proposed an optimal bivariate Poisson field chart for monitoring two correlated characteristics of defects simultaneously. The basic concept is defining the optimal limits of bivariate Poisson distribution and illustrating Poisson pairwises in Poisson field. By using practical example, our optimal bivariate Poisson field chart shows robustness of detecting any assignable cause in the process. Furthermore, by testing our chart using different generated datasets, we show clearly the improved performances of the proposed algorithm compared to others.

Chapter 4

An Optimal Diagonal Inflated Bivariate Poisson Field Chart

In this chapter, we present an optimal diagonal inflated bivariate Poisson field chart to over/under dispersed data. This chart is based on our optimal bivariate Poisson field chart which is based to two steps: optimization of confidence interval and illustration of Poisson field. However, to deal with over/under-dispersion, the diagonal inflated bivariate Poisson model is used instead of usual bivariate Poisson model. The proposed chart presents excellent rate of false alarms, and high sensitivity to handle over/under-dispersed count data. The various simulated data demonstrate the enhanced performances of our control chart compared to other previously proposed charts.

4.1 Introduction

In this chapter, we present our optimal bivariate Poisson field chart with diagonal inflated model to monitor two over/under-dispersed count data. To deal with this type of data, the diagonal inflated bivariate Poisson model is used.

A common problem when modelling two count data by using bivariate Poisson distribution is that variation of count data are sometimes higher than expectation. Mean and variance are not always equal and estimation of bivariate Poisson parameters is not always precise [31]. Thus, it is vital to consider the problem of over/under-dispersion of bivariate Poisson data when a control chart is designed.

To deal with over-dispersed count data, [18] [51] [3] proposed the use of generalized linear model (GLM). This model shows flexibility to model over/under-dispersed count data. In [30] [31], deviance residual chart is proposed to observe count data. This chart is based on monitoring residuals from GLM. By numerical results, this chart shows effectiveness for monitoring over/under-dispersed count data. However, this chart presents process monitoring in form of residuals plot. Thus, meaningful values of raw data may be distorted by monitoring residuals.

In [10], authors presented estimation of bivariate Poisson and diagonal inflated bivariate Poisson regression models. By using an expectation-Maximization (EM) algorithm, parameters of both models can be estimated. In case of diagonal inflated bivariate Poisson model, the authors proposed a general model based on a mixture of three independent Poison distributions and additional distribution for diagonal data. Based on extension of the simple zero-inflated model [23] for only an excessive variables of pair $(0,0)$, the jointly distribution of diagonal inflated bivariate Poisson distribution is given by [10]:

$$P_{IBP}(X=x,Y=y)=\begin{cases}(1-p)P_{BP}(x,y|\vec{\lambda}), & x\neq y\\ (1-p)P_{BP}(x,y|\vec{\lambda})+pP_D(x|\theta), & x=y\end{cases} \tag{1}$$

and

$$COV_{IBP}(X=x,Y=y)=(1-p)\{\lambda_3+(\lambda_1+\lambda_3)(\lambda_2+\lambda_3)\}+pE_D(X^2) \tag{2}$$

$$-(1-p)^2(\lambda_1+\lambda_3)(\lambda_2+\lambda_3) \tag{3}$$

$$-(1-p)pE_D(X)(\lambda_1+\lambda_2+2\lambda_3)-p^2\{E_D(X)^2\} \tag{4}$$

Where $P_{BP}(x,y|\vec{\lambda})$ is bivariate Poisson probability density function, and $P_D(x|\theta)$ is the probability function of diagonal distribution with parameter θ, and p is inflation proportion. $E_D(X)$ is expected value of diagonal distribution. Diagonal distribution can be presented by Poisson, discrete or geometric distribution. Diagonal distribution is inflated, when p is relatively high. There are two properties of this diagonal inflated bivariate Poisson model. First, the marginal distribution of a diagonal inflated model for X is given by [10]:

$$P_{IBP}(X=x)=(1-p)P_0(x|\lambda_1+\lambda_2)+pP_D(x|\theta) \tag{5}$$

Where $P_0(x|\lambda_1 + \lambda_2)$ is Poisson probability function with mean $(\lambda_1 + \lambda_2)$. It is clearly shown that this marginal is two mixture distributions which are univariate Poison and geometric distribution. Therefore, mean and variance of this marginal distribution are given by [10]:

$$E(X) = (1-p)(\lambda_1 + \lambda_3) + pE_D(X) \tag{6}$$

and

$$VAR(X) = (1-p)\{(\lambda_1+\lambda_3)^2+(\lambda_1+\lambda_3)\}+pE_D(X^2)-\{(1-p)(\lambda_1+\lambda_3)+E_D(X)\}^2 \tag{7}$$

The marginal distribution of diagonal inflated model is not individual Poisson distribution. Thus, the distribution can be over or under dispersed depending on distribution $D(x,\theta)$, and p. Furthermore, in case of usual bivariate Poisson, when $\lambda_3 = 0$, all pairs of bivariate poisson count become independent. Nonetheless, when λ_3 of diagonal inflated bivariate Poisson model is zero, all pairs are still dependent by diagonal inflated distribution $D(x,\theta)$. Let $\lambda_3 = 0$, the covariance of diagonal inflated bivariate Poisson with $\lambda_3 = 0$ is given by:

$$COV_{IBP}(X,Y) = p(1-p)\lambda_1\lambda_2 + pE_D(X^2) - p(1-p)E_D(X)(\lambda_1+\lambda_2) - p^2E_D(X)^2 \tag{8}$$

This equation also shows that covariance between X and Y can be negative by distribution $D(x,\theta)$, and p. By using this model, Karlis and Ntzoufras provided many detailed numerical simulations and case studies to illustrate effectiveness of fitting over and under dispersed count data with this model [10].

In this chapter, we propose an optimal diagonal inflated bivariate Poisson field chart for over/under-dispersed count data. The basic concept is based on optimization of confidence interval of diagonal bivariate Poisson model, and illustration of Poisson field. The various simulated datasets present the robustness of monitoring over/under-dispersed data by our control chart.

The rest of the chapter is organized as follows. In Section 4.2, the basic principals of an optimal diagonal inflated bivariate Poisson field chart are introduced. Finally, simulations are presented to illustrate the effectiveness of our control chart in Section 4.3, and this chapter is concluded in Section 4.4.

4.2 Proposed Method

Similarly to our previously proposed chart, the optimal diagonal inflated bivariate Poisson field chart is based on the probabilistic optimization and projection of bivariate Poisson histogram or Poisson field. However, this chart uses diagonal inflated Poisson distribution dealing with two correlated over/under-dispersed count data. Suppose two over/under-dispersed correlated characteristics of defects are pairwises of (x, y) with parameters $\vec{\lambda} = (\lambda_1,\lambda_2,\lambda_3)$, and unknown distribution on diagonal samples $(P_D(x = y, \theta))$. First, $\vec{\lambda}$ is estimated from complete samples $(x_i, y_i), i = 1, 2, 3, \ldots, n$, and the expected type I error rate (α_0) is specified. Moreover, Unknown parameters (θ and p) of additional distribution on diagonal samples are also estimated.

Second, by equation 1 and estimated parameters ($\vec{\lambda}$), probabilities of all pairwises $(p_{IBP}(r, s|\vec{\lambda}, \theta, p))$ are calculated from pair $(0, 0)$ until all numbers on row and column are equal to zeros as shown in table 4.1.

Table 4.1: Probabilistic Poisson field

		y							
		0	1	2	3	.	.	.	s_{max}
	0	p_{00}	p_{01}	p_{02}	p_{03}	.	.	.	0
	1	p_{10}	p_{11}	p_{12}	p_{13}	.	.	.	0
	2	p_{20}	p_{21}	p_{22}	p_{23}	.	.	.	0
x	3	p_{30}	p_{31}	p_{32}	p_{33}	.	.	.	0
	.	.	.	.	.	.	.	.	0
	.	.	.	.	.	.	.	.	0
	.	.	.	.	.	.	.	.	0
	r_{max}	0	0	0	0	0	0	0	0

Where $p_{00}, p_{10}, p_{01}, p_{11}, \ldots, p_{r_{max}s_{max}}$ are probabilities of each pairwise on probabilistic Poisson field, and r_{max} and s_{max} are numbers of the row and the column where all probabilities are equal to zero.

In the third step, similarly to the case of bivariate Poisson distirubtion (see section 3.2), we shift the control limit of our chart from $min(p_{rs})$ to $max(p_{rs})$. Since total probability in Poisson field is equal to 1, and in order to optimize our chart control limit, let $\min(p_{00}, p_{10}, p_{01}, ..., p_{j_{max}k_{max}}) = 0$. Then, the total probability of Poisson field is calculated by:

$$\frac{1}{ARL_{total}} = 1 - \sum_{r=0}^{r_{max}} \sum_{s=0}^{s_{max}} p_{rs} \quad (9)$$

To acquire the optimal diagonal inflated bivariate Poisson field, the pairwises that have the lowest probability in Poisson probabilistic field are removed until total probability in Poisson field is satisfied:

$$ARL_{total} \geq ARL_0 \quad (10)$$

Where ARL_0 is the expected average run length or nominal average run length ($ARL_0 = \frac{1}{\alpha_0}$). After the total probability in Poisson field is satisfied given condition in equation 10, the optimal diagonal inflated Bivariate Poisson control limit can be obtained.

4.3 Experimental Results

In this section, simulations to compare performances of our proposed chart and others are conducted. In this experiment, we tested each control chart by 10,000 pairs of simulated data generated according to parameters in table 4.2. To measure the sensitivity of each chart, we shifted mean parameters ($\vec{\lambda}$, θ) by σ, 2σ, and 3σ. The results are given in tables 4.3, 4.4, and 4.5.

These tables show outstanding performances of the optimal diagonal inflated bivariate Poisson field chart in different mean samples and mean shifts. For in-control state, the proposed algorithm outperforms other charts since ARL levels in different mean samples are close to ARL_0. NORTA chart shows high rate of ARL than ARL_0 that relates to type II error. MP-SUM chart provides higher ARL in case of low mean and ARL shortage in case of medium to high mean. In figure 4.1, it is clear that our proposed chart gained the lowest loss function since the levels of loss function are drawn in the bottom of the graph.

In case of detecting mean shifts, the optimal diagonal inflated bivariate Poisson field chart is more sensitive than the other control charts since ARL levels of our chart effectively respond to mean shifts by decreasing ARL level to low values.

Table 4.2: Simulation cases

Case	λ_1,λ_2	λ_3	p	θ
A	0.4	1.5	0.25	1.5
B	0.4	1.5	0.5	1.5
C	0.4	1.5	0.75	1.5
D	4	4	0.25	5
E	4	4	0.5	5
F	4	4	0.75	5
G	4	40	0.25	50
H	4	40	0.5	50
I	4	40	0.75	50

Table 4.3: ARL Levels of NORTA chart for different mean shifts

		Diagonal Shifts			λ_1 and λ_2Shifts			λ_3 Shifts			All Shifts		
Case	No Shift	σ	2σ	3σ	σ	2σ	3σ	σ	2σ	3σ	σ	2σ	3σ
A	10000	196.07	68.02	24.15	33.67	13.83	8.3	57.47	13.94	5.35	15.22	3.87	1.98
B	10000	72.46	29.94	11.32	11.8	10.1	6.85	47.84	15.89	7.27	12.97	4.3	2.4
C	10000	31.05	15.31	6.83	4.5	9.66	7.81	27.24	16.15	10.24	10.98	5.62	3.44
D	10000	243.90	185.18	80.64	45.87	17.33	7.17	111.11	30.76	9.57	22.83	3.96	1.78
E	10000	85.47	63.69	26.45	31.94	13.85	6.97	62.5	25.18	9.78	18.86	4.6	2.35
F	10000	33	25.31	10.63	18.18	11.82	7.56	27.24	17.15	10.53	13.83	5.78	3.43
G	10000	4.24	8.6	41.32	69.44	33.22	20	136.98	26.04	6.77	3.87	4.01	2.51
H	10000	54.34	17.48	5.42	39.84	22.22	17.54	142.85	50.5	15.82	34.96	7.55	2.55
I	10000	32.46	12.61	4.44	22.83	16.75	14.32	46.51	32.25	21.64	23.09	8.3	3.06

Table 4.4: ARL Levels of MPSUM chart for different mean shifts

		Diagonal Shifts			λ_1 and λ_2Shifts			λ_3 Shifts			All Shifts		
Case	No Shift	σ	2σ	3σ	σ	2σ	3σ	σ	2σ	3σ	σ	2σ	3σ
A	588.23	153.84	36.63	15.45	200	81.3	30.67	50	11.12	4.54	20.74	4.21	1.98
B	476.19	99	18.45	7.47	105.26	60.97	26.38	41.15	11.18	5.39	18.34	4.27	2.14
C	555.55	50.5	12.31	5	125	112.35	60.6	85.47	21.36	10.82	24.5	5.77	2.93
D	76.33	217.39	99	27.47	37.73	11.75	4.74	31.74	9.68	4.24	11.96	2.53	1.45
E	128.2	111.11	28.9	9.3	28.73	9.77	4.3	24.44	8.27	4.23	9.24	2.72	1.76
F	83.33	45.24	12.51	4.41	29.49	11.96	6.53	25.9	11.37	6.77	11.06	3.65	2.23
G	54.34	3.89	4.48	9.74	56.17	45.66	29.94	23.69	6.32	2.6	3.3	1.99	1.47
H	43.29	10.53	4.07	2.47	63.29	59.17	42.01	42.91	12.53	4.67	9.56	2.56	1.33
I	45.04	11.13	3.53	1.87	75.75	81.96	73.52	72.99	26.17	10.83	11.64	2.95	1.45

Table 4.5: ARL Levels of the optimal diagonal inflated bivariate Poisson chart for different mean shifts

		Diagonal Shifts			λ_1 and λ_2Shifts			λ_3 Shifts			All Shifts		
Case	No Shift	σ	2σ	3σ	σ	2σ	3σ	σ	2σ	3σ	σ	2σ	3σ
A	263.15	172.41	69.93	24.09	38.46	13.64	7.73	39.06	11.93	5.04	12.78	3.463	1.87
B	250	175.43	37.45	12.37	27.7	16.83	9.25	54.34	16.05	7.28	16.18	4.53	2.43
C	217.39	112.35	24.27	8.23	16.58	23.2	15.17	70.92	21.73	11.52	21.55	6.6	3.71
D	400	263.15	204.08	61.34	48.54	13.6	5.52	81.3	16.83	6.09	15.62	3	1.58
E	400	294.11	92.59	21.5	68.02	17	6.84	78.12	19.26	7.66	19.92	4.05	2.15
F	285.71	188.67	71.94	14.81	60.6	22.27	9.56	88.49	26.04	12.16	25.51	6.28	3.46
G	434.78	4.26	8.68	44.44	103.09	45.45	22.02	71.94	12.46	4.31	3.67	3.04	1.97
H	400	312.5	48.07	10.03	128.2	43.85	26.04	86.95	14.94	5.58	31.84	5.49	2.26
I	400	200	39.06	8.1	128.2	63.69	34.36	86.2069	19.8	9.09	41.66	8.31	3.2

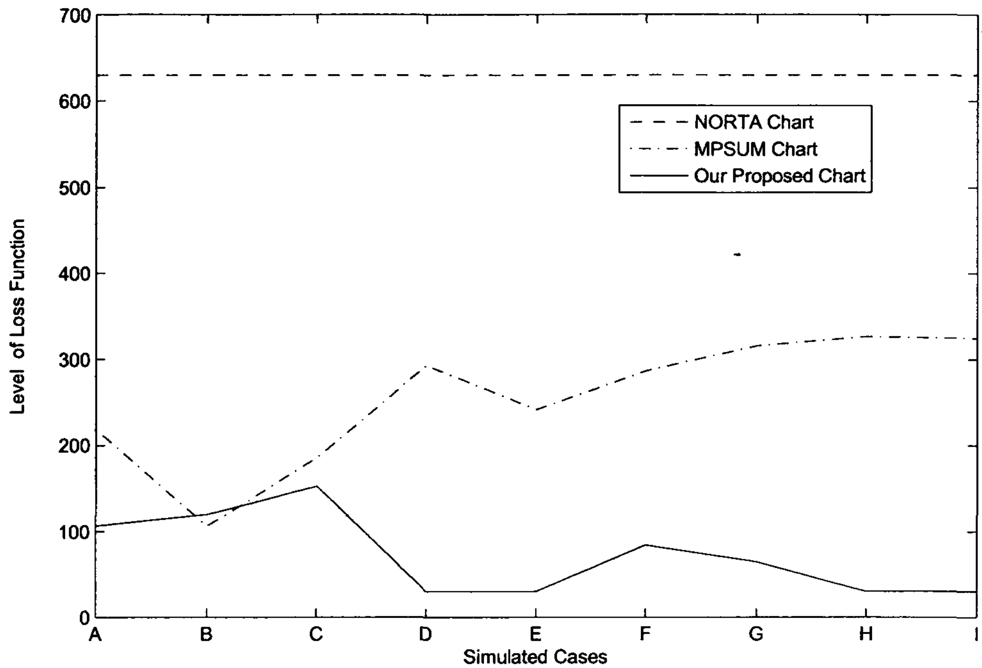


Figure 4.1: Loss function of each charts

4.4 Conclusion

In this chapter, we propose a new optimal diagonal inflated bivariate chart. The basic concept of this chart is based on optimization of diagonal inflated bivariate Poisson confidence interval, and projection of data in Poisson field. The numerical results show good performance. Clearly, our optimal diagonal inflated bivariate chart outperforms other perviously proposed charts.

Chapter 5

Conclusions

This thesis has presented several defect charts to monitor both univariate and bivariate discrete data. In addition, we also proposed a new bivariate defects chart. For univariate case, we compared 11 defect charts that have been proposed for statistical process control. To compare the performances of each chart, we tested each chart and considered some key factors such as low mean responsiveness, loss function, and mean shifting sensitivity. Moreover, to simplify the comparison of control chart performances, we categorized all control charts into three groups, transforming data, standardizing data, and optimizing control limits.

Transforming data approach shows effective performance to construct defect charts for very low mean samples. Lower control limit of charts by Poisson transformation of Bartlett, Anscomebe, and Freeman and Tukey are shown to be precise for indicating out-of-control signal at low mean samples. Thus, this approach is suitable to indicate process improvement. However, this approach is ineffective for signaling out-control alarm since upper control limit of transformed charts always draws high loss function. Furthermore, after transforming data, transformed chart can not provide original value of each data point.

Standardizing data approach is more practical than transforming data approach. Although the charts in this group are not effective to deal with small mean samples, they provide meaningful illustration in -3 and 3 control limits. Especially, Q-chart has the advantage of illustrating different types of defects simultaneously within single chart, and also provides improved results compared to classic C-chart.

Optimizing control limits approach is the best approach to acquire a defect chart. By this approach, there is no requirement of transforming or standardizing data, therefore, plotted data still maintain their original meaning in control chart. From experimental results, Optimal C-chart shows outstanding results. By lowest loss gain and highest mean shifting sensitivity, we concluded that this chart is the best defect chart to replace traditional C-chart in attribute monitoring. However, we do not claim that optimal C-chart is the only defect chart that should be applied in defect monitoring. In practice, selection of charts also depends on suitability of a particular process. Factors such as samples size, mean, UCL, LCL, undetected defect cost, and inspection cost may also influence defect chart selection.

We have also proposed an optimal bivariate defects chart for two correlated characteristics. The basic concept is defining the optimal limit of bivariate Poisson distribution and illustrating data in Poisson field. The proposed control chart shows excellent performances in both practical dataset and various simulations. Our experimental results show improved rate of average run length and robust detection of means shifts compared with other charts. The proposed chart is an effectively applicable chart, especially, in high quality processes.

Furthermore, as a major issue of bivariate Poisson count data is over-dispersion of count variables, designing a defects chart for observing over/under-dispersed Poisson samples becomes vital. We have also extended our optimal bivariate Poisson field chart into an optimal diagonal inflated bivariate Poisson field chart to deal with this issue. The core idea is based on both confidence interval optimization and Poisson field projection. Besides, the diagonal inflated bivariate Poisson model is used for overcoming over/under-dispersion of data. By various simulations, our diagonal inflated bivariate Poisson field chart still outperforms other existed charts in terms of loss function and mean shifting sensitivity.

Future work can be devoted to propose new multivariate defects charts. Multivariate Poisson model is the model that can be used when the monitoring of many characteristics is necessary [43]. The advantage of using multivariate Poisson model is that correlation between variables are taken into account. Moreover, there is no need of transformation which allows us to keep meaningful information. Furthermore, bivariate and multivariate Poisson generalized models are flexible to handle over-dispersed Poisson variable. As shown in [30] and [31], monitoring the deviance residuals from generalized linear model can outperform C-chart. Thus, monitoring the deviance residual from bivariate generalized Poisson distribution [14] can be interesting investigation.

List of References

[1] A. Asuncion and D. J. Newman (2007), UCI Machine Learning Repository [http://www.ics.uci.edu/ mlearn/MLRepository.html]. Irvine, CA: University of California, School of Information and Computer Science.

[2] A. Winterbottom (1993), Simple Adjustments to Improve Control Limits on Attribute Charts, Quality and Reliability Engineering International, Vol. 9 , pp. 105–109.

[3] A. Yesilova and A. Yilmaz (2007), The Application of Overdispersion and Generalized Estimating Equations in Repeated Categorical Data Related to the Sexual Behaviour Traits of Farm Animal, Journal of Applied Sciences, Vol 9, No. 12, pp. 1762–1767.

[4] B. He, M. Xie, T. N. Goh, and K. L. Tsui (2006), On Control Charts Based on the Generalized Poisson Model, Quality Technology and Quantitative Management, Vol. 3, No. 4, pp. 383–400.

[5] C. A. Lowery and D. C. Montgomery (1995), A Review of Multivariate Control Charts, IIE Trans, Vol.27 pp. 800–810.

[6] C. D. Kemp and W. Kemp (1965), Some Properties of the "Hermite" Distribution, Biometrika, Vol. 52, No. 3/4, pp. 381–394.

[7] C. P. Quesenberry (1991), SPC Q Charts for Start-up Processes and Short and Long Runs, Journal of Quality Technology, Vol. 23, No.3, pp. 213–224.

[8] C. P. Quesenberry (1991), SPC Q charts for a Poisson Parameter λ: Short or Long Runs, Journal of Quality Technology, Vol. 23, No.4, pp. 296–303.

[9] D. C. Montgomery (1985), Introduction Statistical Quality Control, John Wiley and Sons, First Edition, pp. 160.

[10] D. Karlis and I. Ntzoufras (2005), Bivariate Poisson and Diagonal Inflated Poisson Regression Models in R, Journal of Statistical Software Vol. 10, No. 10.

[11] E. A. Cornish and R. A. Fisher (1938), Moments and Cumulants in the Specification of Distributions Revue de L'institut International de Statistique / Review of the International Statistical Institute, Vol. 5, No. 4, pp. 307–320.

[12] E. D. Castillo (1996), Run Length Distributios and Economic Design of $\bar{X}$ Charts with Unknown Process Variance, Metrika, Vol. 43, pp. 189–201.

[13] E. L. Lehmann (1917), Theory of Point Estimation, New York: Wiley.

[14] F. Famoye and P. C. Consul (1995), Bivariate Generalized Poisson Distribution with some Applications, Metrika, Vol. 42, pp. 127–138.

[15] F. J. Anscombe (1948), The transformation of Poisson, binomial and Negative Binomial Data, Biometrika, Vol. 35, pp. 246-254.

[16] F. He, W. Jiang, and L. Shu (2007), Improved Self-Starting Control Charts for Short Runs, Quality Technology and Quantitative Management.

[17] F. C Kaminsky, J. C. Benneyan, R. D. Davis ,and R. J. Burke (1992), Statistical Control Chart Based on a Geometric Distribution, Journal of Quality Technology, Vol. 24, No. 2, pp. 63–69.

[18] G. J. McLachlan (1997), On the EM Algorithm for Overdispersed Count Data, Statistical Methods in Medical Research, Vol. 6, pp. 76–98.

[19] H. Hotelling (1947), Multivariate Quality ControlIllustrated by the Air Testing of Sample Bombsights, Techniques of Statistical Analysis, Eisenhart C, Hastay MW, Wallis WA (eds.). McGraw-Hill: New York, pp. 111-184.

[20] H. Patel (1973), Quality Control Methods for Multivariate Binomial and Poisson Distributions, Technometrics, Vol. 15, pp. 103-112.

[21] J. B. S. Haldane (1938), The Approximate Normalization of a Class of Frequency Distributions, Biometrika, Vol. 29, pp. 392–404.

[22] J. Chiu and T. Kuo (2008), Attribute Control Chart for Multivariate Poisson Distribution, Communications in Statistics - Theory and Methods, Vol. 37, Issue 1, pp. 146–158.

[23] J. F. Walhin (2001), Bivariate ZIP Model, Biometrical Journal, Vol. 43, Issue 2, pp. 147–160.

[24] J. Jolayemi (1999), A Statistical Model for the Design of Multiattribute Control Charts, The Indian Journal of Statistics, Vol. 61, pp. 351-365.

[25] J. R. Evans and William M. Lindsay (2005), An Introduction to Six Sigma and Improvement, Thomson South-Western.

[26] J. S. Oakland (2003), Statistical Process Control, Butterworth Heinemann, Fifth Edition.

[27] J. T. Campbell (1934), The Poisson Correlation Function, Proc. Edinb. Math. Soc. Ser. 2, pp. 18–26.

[28] K. Kawamura (1984), Direct Calculation of Maximum Likelihood Estimator for the Bivariate Poisson distribution, Kodai Math. J. Vol. 7, Issue 2 , pp. 211–221.

[29] K. R. Skinner (2002), Multivariate Process Control for Discrete Data , Doctor Thesis of Philosophy, Arizona State University.

[30] K. R. Skinner, D. C. Montgomery, and G. C. Runger (2003), Process monitoring for multiple count data using generalized linear model-based control charts, International Journal of Production Research, Vol. 41, pp. 1167-1180.

[31] K. R. Skinner, D. C. Montgomery, and G. C. Runger (2004), Generalized Linear Model-based Control Charts for Discrete Semiconductor Process Data, Quality and Reliability Engineering International, Vol. 20, pp. 777–786.

[32] L. A. Doty (1996), Statistical Process Control, Industrial Press INC., Second Edition.

[33] L. R. Shenton and K. O. Bowman (1977), Maximum Likelihood Estimation in Small Samples, Griffin's statistical Monographs and Courses, No. 38, Series Editor: Professor Alan Stuart, D.Sc. (Econ).

[34] M. A. Hamdan and H. Hilow (1993), Estimation and Testing for the Parameters of the Bivariate Poisson Population Based on a 2 x 2 Sample, International Statistical Review / Revue Internationale de Statistique, Vol. 61, Issue 3, pp. 361–367.

[35] M. B. C. Khoo (2004), Poisson Moving Average Versus c Chart for Nonconformities, Quality Engineering, Vol. 16, No.4, pp. 525–534.

[36] M. F. Freeman and J. W. Tukey (1950), Transformations Related to the Angular and the Square Root, Annals of Mathematical Statistics, Vol. 21, pp. 607–611.

[37] M. P. Quine and E. Seneta (1987), Bortkiewicz's Data and the Law of Small Numbers, International Statistical Review / Revue Internationale de Statistique ,International Statistical Institute (ISI).

[38] M. S. Bartlett (1936), The Use of Transformations, Biometrics, Vol. 3, No. 1, pp. 39–52.

[39] M. S. Bartlett (1936), The Square Root Transformation in Analysis of Variance, the Journal of the Royal Statistical Society, Vol. 3, No. 1, pp. 68–78.

[40] M. Xie and T. N. Goh (1992), Some Procedures for Decision Making in Controlling High Yield Processes, Quality and Reliability engineering international, Vol. 8, pp. 335–360.

[41] M. Xie and T. N. Goh (1993), SPC of a Near Zero-Defect Process Subject to Random Shocks, Quality and Reliability Engineering International, Vol. 9, pp. 89–93.

[42] N. C. Schwertman and T. P. Ryan (1997), Implementing Optimal Attributes Control Charts, Journal of Quality Technology, Vol. 29, pp. 99–104.

[43] N. L. Johnson, S. Kotz, N. Balakrishnan (1997), Discrete Multivariate Distribution, Wiley Serie in Probability and Statistics.

[44] P. Holgate (1964), Estimation for the Bivariate Poisson Distribution, Biometrika, Vol. 51, pp. 241–245.

[45] R. C. Griffiths, R. K. Milne, R. Wood (1979), Aspects of Correlation in Bivariate Poisson Distribution and Processes, Australian & New Zealand Journal of Statistics, Vol. 21, Issue 3, pp. 238–255.

[46] R. W. Hamming (1991), The Art of Probability for Scientists and Engineers, Redwood City, California: Addison-Wesley.

[47] H. Gitlow, A. Opperheim, and R. Oppenheim (1995), Quality Management: Tool and Methods for Improvement, Irwin, Second Edition.

[48] H. J. Mittag and H. Rinne (1984), Statistical methods of quality Assurance, Chapman and Hall.

[49] R. G. Jr. Kittlitz (2003), Transforming the Poisson Distribution To Symmetry For SPC Applications and Other Statistical Analysis, MS Capstone Project for University of Alabama at Huntsville.

[50] R. G. Jr. Kittlitz (2006), Calculating the (Almost) Exact Control Limits for a C-Chart, Quality Engineering, Vol. 18, No. 3, pp. 359–366.

[51] S. Mitra and S. Washington (2007), On the Nature of Over-Dispersion in Motor Vehicle Crash Prediction Models, Accident Analysis and Prevention, Vol. 39, pp. 459–468.

[52] S. Bersimis and S. Psarakis (2005), Multivariate Statistical Process Control Charts and the Problem of Interpretation: A Short Overview and Some Applications in Industry, Proceedings of the 7th Hellenic European Conference on Computer Mathematics and its Applications, Athens, Greece.

[53] S. Bersimis, S. Psarakis, and J. Panaretos (2006), Multivariate Statistical Process Control: Over view, Munich Personal RePEc Archive, Issue 6399, posted 19.

[54] S. Kocherlakota and K. Kocherlakota (1992), Bivariate Discrete Distribution, Vol. 132, pp. 87–119.

[55] S. A. Niaki and B. Abbasi (2007), On the Monitoring of Multi-Attributes High-Quality Production Processes, Springer.

[56] S. L. Nelson (1994). A Control Chart for Parts-per-Million Nonconforming Items, Journal of Quality Technology, Vol 24, No. 2, pp. 63–69.

[57] S. T. Akhavan Niaki and B. Abbasi (2006), On the Monitoring of Multi-Attributes High-Quality Production Processes, Metrika, Vol. 66, Issue 3.

[58] T. P. Ryan (1989), Statistical Methods for Quality Improvement, New York : Wiley.

[59] T. N. Goh (1987), A Control Chart for Very High Yield Processes, Quality Assurance, Vol. 13, No. 1, pp. 18–22.

[60] T. P. Ryan and N. C. Schwertman (1997), Optimal Limits for Attributes Control Charts, Journal of Quality Technology, Vol. 29, pp. 86-98.

[61] T. R. C. Read and N. A. C. Cressie (1988), Goodness-of-Fit Statistics for Discrete Multivariate Data, New York: Springer-Verlag, pp. 86.

[62] T. Stapenhurst (2005), Mastering Statistical Process Control, Elsevier butterworth heinemann, First edition.

[63] T. Tsai, C. Lin, and S. Wu (2006), Alternative Attribute Control Charts Based on Improved Square Root Transformation, Tamsui Oxford Journal of Mathematical Sciences, Vol. 22, pp. 61–72.

[64] T. T. Allen (2006), Introduction to Engineering Statistics and Six Sigma, Springer, First Edition.

[65] W. H. Woodall (1997), Control Charting Based on Attribute Data: Bibliography and Review, Journal of Quality Technology, Vol. 29, pp. 172–183.

[66] W. S. Messina (1987), Statistical Quality Control for Manufacturing managers, Wiley Series in Engineering Management.

[67] X. Lu, M. Xie, T. Goh, and C. Lai (1998), Control Charts for Multivariate Attribute processes, International Journal of Production Research, Vol. 36, pp. 3477–3489.

[68] Y. Inbal and S. Galit (2008), An Elegant Method for Generating Multivariate Poisson Random Variable, eprint arXiv:0710.5670.

[69] Z. Wang and W. X. Ma (2003), Design of an Optimum Adaptive Control Chart for Attributes, Proceedings of the 2003 Systems and Information Engineering Design Symposium Matthew H. Jones, Barbara E. Tawney, and X. Preston White, Jr.