SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

Statistical Process Monitoring and the Zero Inflated Poisson Model

By

Marianthi T. Axiomakarou

A THESIS
Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
September 2019


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

Στατιστικος Έλεγχος Διαδικασιών και το Zero Inflated Poisson μοντέλο

Μαριάνθη Θ. Αξιομακάρου

ΔΙΑΤΡΙΒΗ
Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα, Ελλάδα

Σεπτέμβριος 2019


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

DEDICATION

...to my beloved grandma


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ACKNOWLEDGMENTS

I would, first, like to thank my supervisor, professor Stelios Psarakis for his advice, guidance and understanding. I am also grateful to all the people who have contributed to this effort in their own way.

Special mention should be given to my parents who always support me and help me gain my goals. This accomplishment would not have been possible without them. Thank you!


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ABSTRACT

Marianthi Axiomakarou

Statistical Process Monitoring and the Zero Inflated Poisson Model

September, 2019

Statistical Process Control (SPC) is a collection of statistical techniques that provide a rational management of industrial and nonindustrial processes. It can be a powerful tool to characterize a process in both normal (in control) and abnormal (out-of-control) conditions. When the monitoring of a high-yield process is of interest, the common practice is to classify each inspected item (or unit) as either conforming (non defective) or non-conforming (defective) according to the specifications of the quality characteristic. High-yield process is characterized by the excessive number of zero-defect counts. In such a situation, the Zero Inflated Poisson distribution is more appropriate to be used than a standard Poisson distribution.

In this thesis, a thorough review of the literature concerning the Zero Inflated Poisson model in statistical process monitoring will be presented. We test one of the mentioned works via simulation analysis presenting an alternative approach to obtain each time the desired results.


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΠΕΡΙΛΗΨΗ

Μαριάνθη Αξιομακάρου

Στατιστικος Έλεγχος Διαδικασιών και το Zero Inflated Poisson Model

Σεπτέμβριος, 2019

Ο Στατιστικος Έλεγχος Διαδικασιών είναι μια συλλογή από στατιστικές τεχνικές που βοηθούν στον εντοπισμό και στη διαχείριση ασυνήθιστων συμπεριφορών μιας βιομηχανικής ή μη διαδικασίας. Είναι ένα ισχυρό εργαλείο το οποίο μπορεί να χαρακτηρίσει μια διαδικασία υπό έλεγχο ή εκτός ελέγχου. Όταν κάτω από το μικροσκόπιο βρίσκεται η παρακολούθηση μιας διαδικασίας υψηλής αποδόσεως, η συνήθης πρακτική που ακολουθείται είναι να ταξινομήσουμε κάθε επιθεωρούμενο στοιχείο ως ελαττωματικό ή μη ελαττωματικό σύμφωνα με τις προδιαγραφές ποιότητας του στοιχείου. Η διαδικασία υψηλής αποδόσεως χαρακτηρίζεται από τον μεγάλο αριθμό μη ελαττωματικών μετρήσεων. Σε τέτοιες καταστάσεις, είναι καταλληλότερο να χρησιμοποιηθεί Zero Inflated Poisson κατανομή αντί της απλής Poisson.

Στη διπλωματική αυτή εργασία παρουσιάζεται μια αναλυτική ανασκόπηση της βιβλιογραφίας σχετικά με το Zero Inflated Poisson μοντέλο και το Στατιστικό Έλεγχο Διαδικασιών. Μελετάται μέσω προσομοίωσης μια από τις αναφερθείσες θεωρίες-τεχνικές και παρουσιάζεται μια εναλλακτική προσέγγιση για την απόκτηση των επιθυμητών αποτελεσμάτων.


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Contents


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

List of Figures


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

List of Tables


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Chapter 1

Introduction

1.1 What is Statistical Process Control

Quality has become one of the most prominent factors that determines the consumer's choice of products and services. Therefore, controlling and improving quality has been marked out as the key factor of business strategy for many companies, financial services organizations and government agencies.

Quality is a competitive factor leading to success. Since a company continually improves the quality of the products it produces, is able to dominate in the market.

The quality of a product is evaluated in several ways,but the most widespread is "fitness of use". According to Garvin (1988) quality has eight components or dimensions:

Performance · Reliability · Durability · Serviceability · Aesthetics · Features · Perceived Quality · Conformance to Standards

Improvement in quality is widely admitted as one of the key elements in the success of a product and consequently of a business. In this direction Statistical Quality Control (SQC) has great contribution. SQC is a field of statistics widely applied in industry in order to maintain the production processes under control. This field includes the areas of design experiments, acceptance sampling and **Statistical Process Control (SPC)**. The last one mentioned, verifies if a process is functioning properly or not. Specifically, a random sample of some products is chosen, inspected and then decided if the process is producing products with characteristics that fall into an already determined range.

In practice, the way that Statistics are used in quality control is through the construction of control charts. A control chart is one of the most basic techniques of SPC in the effort to achieve and maintain stability in industrial or other processes. A typical control chart is shown in figure 1.1.

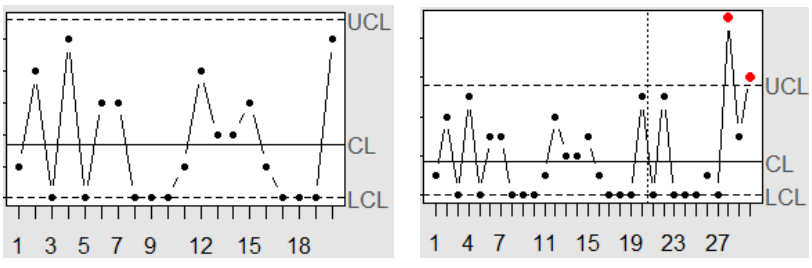

Figure 1.1: Left: An in control process. Right: An out of control process with an alarm at the 28^{th} point.

This chart plots the average number of measurements of a quality characteristic in samples versus time or the sample number. The chart has a center line (CL) that represents the average value of the quality characteristic based on the in-control state (i.e where the process characteristic should fall if there are no non-conformities) and some other horizontal lines named by the Upper Control Limit (UCL) and the Lower Control Limit (LCL) and they are determined from some statistical considerations. If the process is in control all the sample points fall between them.

In the design of a control chart two discrete and separate phases take place:

1. Phase I: A set of process data is gathered and analyzed in order to determine if the process has been in control when the data where collected and to see if dependable control limits can be established to monitor future production. Generally, in this phase, the control chart is the tool helping to bring the process into a state of statistical control.

2. Phase II: After the set of‘ “clean” process data were gathered under stable conditions, in phase II the control chart is used to check whether the process remains in control when current and future samples are taken. Also, the charts are used in order to monitor the process for any change to an out of control state. The change from an in control to an out of control state is determined by shifts in process parameters.

It is desirable that if there is a change in the process, a control chart should detect it as early as possible and give an out of control signal in order to make adjustments to bring the process back to an in control state. On the other hand, when the process is in control, the chart is expected to signal infrequently. Obviously, when the process experience a change, the quicker the detection and the signal, the more efficient the chart is.

The most popular and most used measure of chart performance, concerning the previous two objectives, is the Average Run Length (ARL). The ARL is actually the average number of observations required until a control chart signals. Alternatively, the ARL could be expressed as the average time to signal (ATS) or the average number of observations to signal (ANOS). ATS is the average time needed for a control chart to signal and it is actually a product of the ARL and the sampling interval used in the case of fixed sampling (Montgomery, 2009).

ARL is divided in two similar but quite different -in their use- measures: ARL_0 and ARL_1. The in-control average run length, namely ARL_0, is the average number of observations needed until a control chart signals falsely when the process is actually in-control. The out-of-control average run length, namely ARL_1, is the average number of observations needed until a control chart signals when the process is actually out of control. In general, a reasonably large ARL_0 value is desirable in order to have a low false alarm rate, and a small ARL_1 value is required to quickly detect the changes in the process.

It is obvious that ARL_0 and ARL_1 are directly connected with type I and type II errors respectively. The false alarm rate is the probability that a chart signals a process change when in fact there is no change.This is similar to a probability of a Type I error, i.e $\alpha = P\,(\text{the process is out of control}|\text{no shift has occurred})$. Also, the probability that a chart shows no change in the process while the process is out of control, namely type Type II error, is $\beta = P\,(\text{the process is in control}|\text{a shift has occurred})$, and thus the probability that this out of control condition will be detected on any succeeding sample is $1 - \beta$, which is the power of the chart.

Lastly, the Operating Characteristic Curve (OC) has to be mentioned. This is a measure of the ability of a control chart to detect the changes in process parameters. This means that is a graph that constructed in order to show how changes in the sample size, affect the probability of making a type II error.

1.2 Some relative control charts

There are many processes in which quality characteristics cannot be easily measured on a numerical scale. In such cases, each inspected item is usually classified as either conforming or non-conforming to the specifications on that quality characteristic. The terminology "defective" or "non-defective" is used frequently in order to identify these two classifications of a product, or the terminology "non-conforming" and "conforming", accordingly. Quality characteristics of these type are called attributes. The most usual attributes control charts are the p, np, c and u control charts. The four of them are widely known as Shewhart attribute control charts.

Firstly, a c-chart is designed to deal with the number of defects or nonconformities that are observed in an inspection unit of product. The number of nonconformities or defects of a repetitive production process is denoted by the random variable D. Assume that the observations $D_1, D_2, ...,$ are i.i.d random variables with some mean $E(D)$ and some variance $Var(D)$. Therefore, a control chart for nonconformities, or c chart with three-sigma limits would be defined as follows:

$$\begin{aligned} UCL &= E(D) + 3\sqrt{Var(D)} \\ CL &= E(D) \\ LCL &= E(D) - 3\sqrt{Var(D)} \end{aligned}$$

If the calculation of LCL leads to a negative value, then LCL is set equal to 0, i.e. $LCL = 0$. Also, if the true values of mean and variance are not known, an estimation of them is applied.

Two very effective alternatives to the Shewhart control chart may be used when small process shifts are of interest: the cumulative sum (CUSUM) control chart, and the exponentially weighted moving average (EWMA) control chart.

To begin with, **CUSUM** charts use all the available past sequentially accumulated data in order to detect out-of-control conditions in a process. The shift size of a parameter which is desired to be detected by a CUSUM chart is pre-determined. In the context of this thesis it is useful to be mentioned that is possible to develop CUSUM charts for sample statistics such as fractions nonconforming, and defects. Assume that the observations $D_1, D_2, ...,$ are i.i.d random variables with some mean μ_0, where μ_0 is the value of the in-control process mean. The process is deemed out-of-control when the mean shifts apart from the in-control value μ_0, i.e. $\mu_1 > \mu_0$.

A CUSUM chart for sample statistics such as defects, is operated by accumulating the differences between observed values D_i and a reference value k at each sampling time t. An out-of-control alarm is triggered the first time accumulation equals or exceeds the control limit h. If there is interest in detecting increases in the mean of counts of nonconformities or defects ($\mu_1 > \mu_0$), the cumulative sum statistic is determined by

$$C_i = \max(D_i - k + C_{i-1}, 0)$$

where C_i is the cumulative sum of the differences of the previous samples, D_i is the number of nonconformities in the i^{th} sample, and k is called the reference value. Because C_i can never be negative, the process remains in control if the samples fall between zero and the decision interval limit h. A standard CUSUM procedure has a starting value, $C_0 = 0$. The control limit h for the CUSUM procedure is determined in combination with the reference value k on the basis of ARLs. After k is determined, h is selected to give an appropriately large ARL_0 when the process is in control and an appropriately small ARL_1 when the process goes out of control (Montgomery, 2009).

The performance of the **EWMA** control chart is approximately equivalent to that of the CUSUM control chart, and in some ways it is easier to set up and operate. As with the CUSUM, the EWMA

is typically used with individual observations. This method uses a weighted average of the past observations. So, the exponentially weighted moving average is defined as:

$$z_t = (1-\kappa)z_{t-1} + \kappa y_t$$

where $\kappa \in (0,1]$ is a constant and the starting value $z_0 = \mu_0$. To illustrate that the EWMA z_t is a weighted average of all previous sample means, z_{t-1} must be recursively substituted on the right-hand side of the above equation to obtain

$$z_t = (1-\kappa)^t z_0 + \kappa \sum_{i=0}^{t-1} (1-\kappa)^i y_{t-i}.$$

For $\kappa = 1$, the EWMA places all of its weight on the most recent observation, as does the c- chart. For κ close to 0, the most recent observation receives little weight, and the EWMA seems like the CUSUM chart.

The weights $\kappa(1-\kappa)^i$ decrease geometrically with the oldness of the sample mean. Furthermore, the weights sum to unity, since

$$\kappa \sum_{i=0}^{t-1} (1-\kappa)^i = \kappa[\frac{1-(1-\kappa)^t}{1-(1-\kappa)}] = 1-(1-\kappa)^t.$$

If the observations y_t are independent random variables with variance σ^2 , then the variance of z_t is

$$\sigma^2_{z_t} = (\frac{\kappa}{2-\kappa})[1-(1-\kappa)^{2t}] \cdot \sigma^2.$$

Therefore, the EWMA control chart would be constructed by plotting z_t versus the sample number or time. The center line and control limits for the EWMA control chart are as follows:

$$\begin{aligned} UCL &= \mu_0 + L \cdot \sigma \sqrt{\frac{\kappa}{2-\kappa}[1-(1-\kappa)^{2t}]} \\ CL &= \mu_0 \\ LCL &= \mu_0 - L \cdot \sigma \sqrt{\frac{\kappa}{2-\kappa}[1-(1-\kappa)^{2t}]} \end{aligned}$$

The term $[1-(1-\kappa)^{2t}]$ approaches unity as t gets larger. This means that after the EWMA control chart has been running for several time periods, the control limits will approach steady-state values given by

$$UCL = \mu_0 + L \cdot \sigma \sqrt{\frac{\kappa}{2-\kappa}} \quad \text{and} \quad LCL = \mu_0 - L \cdot \sigma \sqrt{\frac{\kappa}{2-\kappa}} \tag{1.1}$$

The design parameters of the chart are the multiple of sigma used in the control limits (L) and the value of κ. It is possible to choose these parameters to give a desired ARL performance for the EWMA control chart (Montgomery, 2009).

Lastly, another type of time-weighted control chart based on a simple, unweighted **moving average** (MA) is going to be mentioned above. Suppose that individual observations $y_1, y_2, ...$ have been collected. The moving average of width w at time i is defined as

$$M_i = \frac{y_i + y_{i-1} + ... + y_{i-w+1}}{w} = \frac{1}{w} \sum_{j=i-w+1}^{i} y_j \quad , \; i \geq w \tag{1.2}$$

For periods $i < w$, there are not w observations to calculate a moving average of width w. For these periods, the average of all observations up to period i defines the moving average. Hence,

$$M_i = \frac{1}{i}\sum_{j=1}^{i} y_j \quad , 1 \leq i < w \tag{1.3}$$

Thus, if the target value for the process mean is μ_0, then, for periods $i \geq w$, the center line and $L \cdot \sigma$ control limits are given by

$$UCL = \mu_0 + L \cdot \frac{\sigma}{\sqrt{w}} \tag{1.4}$$

$$CL = \mu_0 \tag{1.5}$$

$$LCL = \mu_0 - L \cdot \frac{\sigma}{\sqrt{w}} \tag{1.6}$$

whereas for periods $i < w$, $\frac{\sigma}{\sqrt{w}}$ is replaced with $\frac{\sigma}{\sqrt{i}}$. Here, σ is the in-control standard deviation of the process (Boon Chong, 2004).

1.3 About Zero Inflated Poisson

1.3.1 Univariate Zero Inflated Poisson

It is known that the Poisson distribution is used to model the numbers of events occurring over a given time period. Events occur independently. That means, if an event occurs, it does not affect the probability that a second event will occur in the same time period. Moreover, the rate at which events happen is constant - does not change based on time. The formula of the probability mass function (pmf) is :

$$P(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, ...$$

where $\lambda > 0$ is the shape parameter which indicates the average number of events which are happened within the given time period. Also, the mean and the variance of the Poisson distribution are $E(X) = \lambda$ and $Var(X) = \lambda$.

The cumulative distribution function (cdf) for a Poisson random variable is:

$$F(x;\lambda) = e^{-\lambda}\sum_{i=0}^{x}\frac{\lambda^i}{i!}, x = 0, 1, 2, ...$$

However, in case there is a large or an "excessive" number of zeros contained in the data, the relationship of mean-variance of the Poisson distribution is violated. In fact, the Poisson model often underestimates the observed dispersion. The result of this in terms of a statistical quality control point of view is that the calculated control limits are inappropriately narrow. That has a consequence of an increased rate of false alarms than the expected and the control charts cannot be used efficiently.

In such cases, a special mixture model is more appropriate to be used. This model is known as Zero Inflated Poisson (ZIP). Generally, Zero Inflated models are capable of dealing with excess zero counts. In particular, a ZIP process is different from a standard Poisson process. A ZIP distribution has two sources of zeros. Zeros may come from both the point mass and from the count component which is the Poisson distribution. As far as statistical quality control is concerned ZIP is often used to describe a near zero-defect (or in other words, an almost perfect) process with occasional occurrences

of non-conformity (or defective) products. It's distribution is constructed as a mixture of a Poisson distribution and a distribution that is generated at zero.
The probability mass function (pmf) of the ZIP process is the following:

$$P(y;p,\lambda)=\begin{cases}p+(1-p)e^{-\lambda} & \text{for } y=0\\ \\ (1-p)\frac{e^{-\lambda}\lambda^{y}}{y!} & \text{for } y=1,2,...\end{cases} \tag{1.7}$$

where $\lambda>0$ and $0\leq p\leq 1$.
p: the probability for including extra zeros than those permitted by the standard Poisson distribution or the probability a process is in a near perfect state (with almost zero defects or non-conformities).
1-p: the probability a process is in an imperfect state where the number of defects (or non-conformities) has a Poisson distribution with mean λ.

More thoroughly, the first part of the above distribution is a process that is governed by a binary distribution that generates structural zeros (the outcome is always a zero count). The second part has to do with a Poisson distribution that generates counts which some of those may be zero. When $p=0$ the ZIP distribution is diminished to the Poisson distribution. The parameter λ is the mean of the defects (or non-conformities) in a sample unit. Also, the mean and the variance of the ZIP distribution are $E(Y)=\lambda(1-p)$ and $Var(Y)=\lambda(1-p)(1+p\lambda)$.

At this point let's mention and the ZIP cumulative distribution function (cdf):

$$F_{ZIP}(y;p,\lambda)=\begin{cases}p+(1-p)e^{-\lambda} & \text{for } y=0\\ \\ p+(1-p)e^{-\lambda}\sum_{t=0}^{y}\frac{\lambda^{t}}{t!} & \text{for } y=1,2,...\end{cases}$$

Lastly, it is useful to refer to the likelihood function, and specifically the log-likelihood of the ZIP distribution and the maximum likelihood estimators (MLEs), as well as the moment method estimators (MMEs) for the ZIP parameters. Considering a sample $y_1,y_2,...,y_n$ of n counts coming from a zero inflated process with probability function 1.7 , the likelihood function is:

$$L(p,\lambda)=[p+(1-p)e^{-\lambda}]^{n_0}(1-p)^{n}\prod_{i=1}^{n}\frac{e^{-\lambda}\lambda^{y_i}}{y_i!},$$

where n_0 is the number of zeros in the sample and n is the number of nonzeros in the sample such that $N=n+n_0$. The log-likelihood function now becomes:

$$lnL(p,\lambda)=n_0 ln[p+(1-p)e^{-\lambda}]+nln(1-p)-n\lambda+ln\lambda\sum_{i=1}^{n}y_i-\sum_{i=1}^{n}lny_i!.$$

To obtain the maximum likelihood estimators (MLEs), the derivatives of lnL are taken equal to zero. Hence, the MLEs of the parameters λ and p in the ZIP model are given as:

$$\hat{\lambda}=\bar{y}^{+}(1-e^{-\hat{\lambda}}) \qquad \text{and} \qquad \hat{p}=1-\frac{\bar{y}}{\hat{\lambda}},$$

where $\bar{y}=\frac{1}{N}\sum_{i=1}^{N}y_i$ is the sample mean and $\bar{y}^{+}=\frac{1}{n}\sum_{i=1}^{n}y_i$ is the mean of the n nonzero observations (Clifford Cohen, 1991) .

In order, now, to obtain the MMEs, sample moments should be equal with theoretical moments. Thus, $E(Y^k)$ is the k^{th} (theoretical) moment and $\frac{1}{n}\sum_{i=1}^{n}Y_i^k$ is the k^{th} sample moment of the ZIP

distribution respectively. Because of the parameters of the ZIP distribution, which are two, the first two moments are going to be utilized. Hence, the MMEs of the parameters λ and p are the following:

$$\begin{cases} E(Y) = \frac{1}{n}\sum_{i=1}^{n} Y_i \\ E(Y^2) = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 \end{cases} \Rightarrow \begin{cases} E(Y) = \frac{1}{n}\sum_{i=1}^{n} Y_i \\ Var(Y) + (EY)^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 \end{cases} \Rightarrow \begin{cases} \lambda(1-p) = \frac{1}{n}\sum_{i=1}^{n} Y_i \\ \lambda(1-p)(\lambda+1) = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 \end{cases}$$

$$\Rightarrow \hat{\lambda} = \frac{\sum_{i=1}^{n}(Y_i^2 - Y_i)}{\sum_{i=1}^{n} Y_i} \quad \text{and} \quad 1 - \hat{p} = \frac{1}{n}\frac{(\sum_{i=1}^{n} Y_i)^2}{\sum_{i=1}^{n}(Y_i^2 - Y_i)}$$

Now, with everything taken into account , we are trying to give a graphical representation of the Poisson and ZIP distribution. Using the programming language R, we generate data from a Poisson and a ZIP distribution with $\lambda = \frac{5}{2}$ and $p = \frac{1}{3}$. Hence, a random variable Y takes the value 0 with probability $p = \frac{1}{3}$ or Y follows a Poisson distribution with mean $\lambda = \frac{5}{2}$. The results of that are shown in the graph bellow:

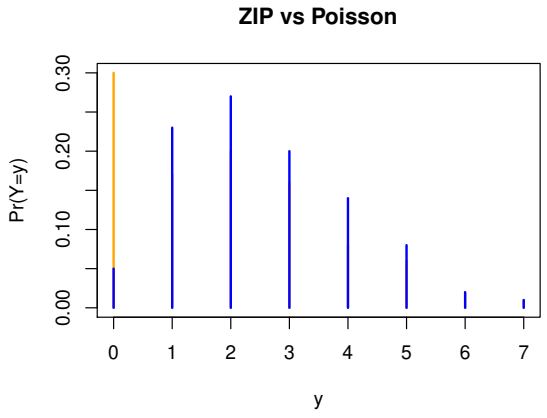

Figure 1.2: A graphical representation of a ZIP vs Poisson distribution

As we can see in figure 1.2 the probability P(Y=0) has two components. The blue part is from the underlying Poisson distribution, the orange part is due to zero-inflation.

1.3.2 Multivariate Zero Inflated Poisson

The univariate ZIP distribution is useful for modeling counts which have large number of zeros. For example, there is a minus number of defect products in a manufacturing process. But, when there are several types of defects in a process, the ZIP distribution is no longer a suitable one. In that case, the multivariate ZIP (MZIP) model can be used to track down different kind of problems in a process. So, when events involve different types of defects in a process near its perfect state, the MZIP model is the appropriate one to handle the situation.

To begin with, we are going to describe the Bivariate ZIP (BZIP) distribution and after that we are going to generalize for the MZIP. First of all, it is necessary to mention the distribution of the Bivariate Poisson and, by extension, the Multivariate Poisson distribution. So, the formula of the probability mass function (pmf) of the Bivariate Poisson is based on the joint distribution of the variables $X_1 = Y_1 + Y_{12}$ and $X_2 = Y_2 + Y_{12}$ where Y_1, Y_2 and Y_{12} are mutually independent Poisson random variables with respect means θ_1, θ_2 and θ_{12} (Johnson et al., 1996). The joint probability

mass function (pmf) is:

$$P\left(X_1 = x_1, X_2 = x_2\right) = P\left(x_1, x_2\right) = e^{-(\theta_1+\theta_2+\theta_{12})} \sum_{i=0}^{min(x_1,x_2)} \frac{\theta_1^{x_1-i}\theta_2^{x_2-i}\theta_{12}^{i}}{(x_1-i)!(x_2-i)!i!}.$$

Now, we have all the equipment to proceed with the Bivariate ZIP model. A model like this can be constructed as a mixture of a point mass at (0,0), two univariate Poisson distributions and a Bivariate Poisson. Hence, we have:

$$\begin{aligned}
(Y_1, Y_2) &\sim (0,0) \text{ with probaability } \ p_0 \\
&\sim \left(\text{Poisson}\left(\lambda_1\right), 0\right) \text{ with probability } \ p_1 \\
&\sim \left(0, \text{Poisson}\left(\lambda_2\right)\right) \text{ with probability } \ p_2 \\
&\sim \text{Bivariate Poisson}\left(\theta_1, \theta_2, \theta_{12}\right) \text{ with probability } \ p_{12}
\end{aligned}$$

where $p_0 + p_1 + p_2 + p_{12} = 1$. Thus, the probability distribution of the BZIP, which is described above, is the following:

$$\begin{aligned}
P\left(Y_1 = 0, Y_2 = 0\right) &= p_0 + p_1 e^{-\lambda_1} + p_2 e^{-\lambda_2} + p_{12} e^{-\theta} \\
P\left(Y_1 = y_1, Y_2 = 0\right) &= \frac{p_1 \lambda_1^{y_1} e^{-\lambda_1} + p_{12}\theta_1^{y_1} e^{-\theta}}{y_1!} \\
P\left(Y_1 = 0, Y_2 = y_2\right) &= \frac{p_2 \lambda_2^{y_2} e^{-\lambda_2} + p_{12}\theta_2^{y_2} e^{-\theta}}{y_2!} \\
P\left(Y_1 = y_1, Y_2 = y_2\right) &= p_{12} \cdot e^{-\theta} \sum_{i=0}^{min(y_1,y_2)} \frac{\theta_1^{y_1-i}\theta_2^{y_2-i}\theta_{12}^{i}}{(y_1-i)!(y_2-i)!i!}
\end{aligned} \tag{1.8}$$

for $y_1, y_2 = 1, 2, ...$ and where $\theta = \theta_1 + \theta_2 + \theta_{12}, \lambda_1 = \theta_1 + \theta_{12}$ and $\lambda_2 = \theta_2 + \theta_{12}$ (Li et al., 1999) . The covariance and non-negative correlation coefficient between the two count variables Y_1, Y_2 are: $Cov\left(Y_1, Y_2\right) = \theta_{12}$ and $Corr\left(Y_1, Y_2\right) = \frac{\theta_{12}}{\sqrt{(\theta_1+\theta_{12})(\theta_2+\theta_{12})}}$ respectively. If $\theta_{12} \neq 0$, there is correlation between Y_1, Y_2 and BZIP distributed data cannot be represented by two independent ZIP models.

By the same token, we can expand the BZIP probability distribution into the MZIP one. Therefore, we mention the Multivariate Poisson pmf :

$$P\left(x_1, x_2, ..., x_m\right) = e^{-(\theta_1+\theta_2+...+\theta_m+\theta_{12})} \sum_{i=0}^{min(x_1,x_2,...,x_m)} \frac{\theta_1^{x_1-i}\theta_2^{x_2-i}\cdots\theta_m^{x_m-i}\theta_{12}^{i}}{(x_1-i)!(x_2-i)!\cdots(x_m-i)!i!}.$$

As far as MZIP is concerned we have to stress that is a mixture of m+2 components of m-dimentional discrete distributions with mixing probabilities $(p_0, p_1, ..., p_m, p_{12})$ where $\sum_{i=0}^{m} p_i + p_{12} = 1$. More specifically, this model is constructed as a mixture of a point mass at (0,0,...,0), m distributions each with univariate Poisson for one defect and m-1 zeros and a m-dimensional Poisson for all m defects. Hence, we have the following probability distribution of the MZIP:

$$P\left(Y_1=0,Y_2=0,...,Y_m=0\right)=p_0+p_1e^{-\lambda_1}+p_2e^{-\lambda_2}+...+p_me^{-\lambda_m}+p_{12}e^{-\theta}$$

$$P\left(Y_1=y_1,Y_2=0,...,Y_m=0\right)=\frac{p_1\lambda_1^{y_1}e^{-\lambda_1}+p_{12}\theta_1^{y_1}e^{-\theta}}{y_1!}$$

$$\vdots$$

$$P\left(Y_1=0,Y_2=0,...,Y_m=y_m\right)=\frac{p_m\lambda_m^{y_m}e^{-\lambda_m}+p_{12}\theta_m^{y_m}e^{-\theta}}{y_m!}$$

$$P\left(Y_1=y_1,Y_2=y_2,...,Y_m=y_m\right)=p_{12}\cdot e^{-\theta}\sum_{i=0}^{min(y_1,...,y_m)}\frac{\theta_1^{y_1-i}\theta_2^{y_2-i}\cdot\theta_m^{y_m-i}\theta_{12}^i}{(y_1-i)!(y_2-i)!\cdot(y_m-i)!i!},$$

for the cases in which at least two of the $y_i's$ are not zero and $\theta=\theta_1+\theta_2+...+\theta_m+\theta_{12}$ (Li et al., 1999).

So, the first thing that someone has to do is to check out whenever is necessary to use the ZIP instead of the standard Poisson model. Xie et al. (2001) dealt with this matter.

To begin with, this article refers to a number of different tests of Poisson against ZIP distribution. The reason why is that necessary is that a vast number of zero-counts may not lead to the choice of a ZIP model and a test should be applied before someone proceed to the selection of the model. We consider that is a good idea to mention the six tests which are reported in the article in order to choose each time the appropriate model for the data. The hypothesis test which has to be examined is: $H_0:p=0$ against $H_1:p\neq 0$. In case that the null hypothesis cannot be rejected we come to the conclusion that there is no need of the ZIP model.

So, before the tests are mentioned, a set of observations $\{Y_1,Y_2,...,Y_n\}$ with sample size n and n_i the number of count i in the sample are considered. Specifically, n_0 is the number of zeros in the sample. Also, $\overline{y}=\sum_{i=1}^{n}\frac{y_i}{n}$ is the mean of the observations.

Hence, the tests that are mentioned in this paper are the following:

1. The score test

This test proposed by Vandenbroek (1995) for the hypothesis which is mentioned above. The score statistic can be written as:

$$S_1=\frac{(n_0-np_0)}{np_0(1-p_0)-n\overline{y}p_0^2},$$

where $p_0=P\left(y=0\right)=e^{-\hat{\lambda_0}}$ and $\hat{\lambda_0}$ is the estimate of the Poisson parameter under the null hypothesis (van den Broek, 1995).

2. The likelihood ratio test

This test compares the goodness of fit of two statistical models based on the likelihood function. In this case, the statistical models are the standard Poisson and the ZIP distribution with likelihood functions $L_{Poisson}$ and L_{ZIP} respectively. The above test can be written as follows: $LRT=\frac{L_{Poisson}}{L_{ZIP}}$. According to El-Shaarawi (1985) the test statistic of likelihood ratio test can be calculated from the formula above:

$$-2lnLRT=2\{n_0ln(\frac{n_0}{n})+(n-n_0)[ln(\frac{\overline{y}}{\hat{\lambda_1}})-\hat{\lambda_1}]+n\overline{y}(ln\hat{\lambda_1}+1-ln\overline{y})\},$$

where $\hat{\lambda_1}$ is the maximum likelihood estimate of λ. Under the null hypothesis, the above test statistic approximatelly follows a X_1^2 distribution (H el Shaarawi, 1985) .

3. The Chi-square test

The third test in row is the Chi-squared test. This test is used to decide whether there is any significant difference between the observed frequencies and the expected ones in one or more categories. Suppose that n observations in a random sample from a population are classified into c mutually exclusive classes with respective observed numbers f_i , for $i = 1, 2, ..., c$ and a null hypothesis gives the probability p_i that an observation falls into the i^{th} class. So we have the expected numbers $m_i = np_i$ for all i's. So, the chi-squared statistic can be computed as follows:

$$X^2 = \sum_{i=1}^{c} \frac{(f_i - m_i)^2}{m_i},$$

where c denotes the number of classes decided for a given data set. Hence, when the null hypothesis cannot be rejected and $n \to \infty$ the chi-squared statistic follows a X^2_{c-1} distribution (Xie et al., 2001) .

4. A test based on a confidence interval of p

When a variable Y follows a ZIP distribution with parameters p and λ , then the mean and the variance of the sample mean $\bar{Y}$ are $E\left(\bar{Y}\right) = E\left(Y\right) = \lambda(1-p)$ and $Var\left(\bar{Y}\right) = \frac{1}{n}Var\left(Y\right) = \frac{1}{n}[\lambda(1-p)(1+p\lambda)]$, accordingly. Based on the central limit theorem, the distribution of the limiting form of

$$Z = \frac{\bar{Y} - E\left(Y\right)}{\sqrt{\frac{Var(Y)}{n}}} = \frac{\bar{Y} - \lambda(1-p)}{\sqrt{\frac{\lambda(1-p)+\lambda(1-p)(\lambda-(1-p)\lambda)}{n}}} \sim N\left(0,1\right)$$

as $n \to \infty$. Hence, a $100(1-a)\%$ asymptotic confidence interval of $1-p$ can be taken as :

$$\frac{\bar{y} + Z_{\frac{a}{2}}\sqrt{[\bar{y} + \bar{y}(\hat{\lambda_1} - \bar{y})]/n}}{\hat{\lambda_1}} \leq 1 - p \leq \frac{\bar{y} - Z_{\frac{a}{2}}\sqrt{[\bar{y} + \bar{y}(\hat{\lambda_1} - \bar{y})]/n}}{\hat{\lambda_1}},$$

where practically in a given data set $\bar{y}$ substitutes the $E\left(Y\right)$ and $\hat{\lambda_1}$ is the maximum likelihood estimate of λ. Thus, a test statistic on the confidence interval for $1-p$ can be written as :

$$S_2 = \frac{\bar{y} + Z_a\sqrt{[\bar{y} + \bar{y}(\hat{\lambda_1} - \bar{y})]/n}}{\hat{\lambda_1}}.$$

When, $S_2 < 1$, which is the critical region, the null hypothesis is rejected at α level of significance and the ZIP model is more appropriate than the standard Poisson to fit the data. But, when $S_2 \geq 1$ the reverse is happening (Xie et al., 2001) .

5. The Cochran test

Cochran (1954) proposed a statistic for comparing the observed and expected frequencies of a single outcome from a Poisson distribution. The test statistic to compare the observed and expected zero-frequencies is the following :

$$C = \frac{n_0 - ne^{-\bar{y}}}{[ne^{-\bar{y}}(1 - e^{-\bar{y}} - \bar{y}e^{-\bar{y}})]^{\frac{1}{2}}}.$$

When the null hypothesis is valid, C approximately follows a standard normal distribution (Xie et al., 2001; van den Broek, 1995; Cochran, 1954) .

6. The Rao-Chakravarti test

The last one proposed test in this article about Poisson against ZIP model is the one of Rao and Chakravarti (1956). The formula of this test statistic is the above:

$$R = \frac{n_0 - n(\frac{n-1}{n})^{n\bar{y}}}{\sqrt{n(\frac{n-1}{n})^{n\bar{y}} - n^2(\frac{n-1}{n})^{2n\bar{y}} + n(n-1)(\frac{n-2}{n})^{n\bar{y}}}}.$$

Under the null hypothesis, the above test statistic is approximately normally distributed with zero mean and variance equals to one (Rao and Chakravarti, 1956) . Hence, if the test statistic R is larger than a normal deviate, indicates significance in the frequency of zero which means a ZIP model is more appropriate to be used (H_0 is rejected).

The next move for the authors was to conduct a simulation study in which they compared the power of the forenamed tests. The data were simulated from a ZIP distribution with known parameters λ and p. According to the paper, 1000 samples were taken for three sample sizes 10, 20 and 50, respectively.

Based on the results, the mentioned tests are all proper for the testing of Poisson versus the ZIP distribution. Some remarks about the test outcomes are the following:

- The confidence interval test is weaker than the others due to the fact that the null hypothesis has a strong protection.
- The powers of those tests are also good when parameter p is close to one (i.e 1-p is close to zero). For example, when sample size is 50, $\lambda = 10$ and $p = 0.9$ the test powers are larger than 0.99.

This article, also, discusses the statistical control limits of the ZIP model. The lower control limit of that model is not exist, which is common as regards attribute charts. Hence, the authors deal with the upper control limit, n_u, for a control chart based on the above:

$$P\,(n_u \text{ or more "defects" in a sample }) \leq a_L,$$

where a_L is the predetermined false alarm probability for n_u. So, the conclusion is that when the distribution is the ZIP one, there should not be any alarm for the values less than or equal to n_u .

Moreover, a study about the sensitivity of a control chart takes place in this paper. A single control limit is scrutinized here based on the change of the two parameters of the ZIP model. Here, the operating-characteristic function (OCF) and the average run length (ARL) have been utilized. The OCF is linked with type II error β as follows:

$$\beta = P\,(D < UCL|p, \lambda, a_L) = p + (1-p)e^{-\lambda} + \sum_{y=1}^{UCL-1}(1-p)\frac{\lambda^y e^{-\lambda}}{y!},$$

where D is the number of defects in the sample (i.e $D \sim ZIP\,(\lambda, p)$) and a_L is the type I error probability.

Furthermore, the ARL for the ZIP model is the above :

$$ARL = \frac{1}{P\,(\text{ point is out of control })} = \frac{1}{\sum_{y=UCL}^{\infty}(1-p)\frac{\lambda^y e^{-\lambda}}{y!}}.$$

Hence, based on the numerical values which are shown in the paper, the conclusions are the ones referred above:

- The procedure is strongly affected by the increase of p and that is because of the relation of p and the occurrence of nonconformities.
- When p decreases, there are more units with nonconformities, so the average number of items to be checked when an alarm happened will decrease very quickly.
- The same happens when λ increases, but when it has increased beyond a certain value the ARL will not decrease any further depending on the value of p.
- The above situation is not desirable, but is expected. The reason is that p is related to the probability of extra zeros (i.e less nonconformities). So, when p is large the chance to happen a nonconformity is small regardless of value λ.

At last, the conclusion is that many zero-counts in a data set may not indicate that a ZIP model has to be used, and a test should be performed to be sure. In addition, if a Poisson model had been used, when a ZIP model was more appropriate, the result would have been too many false alarms because of the narrower control limits. That would lead to frequent unnecessary inspections of the manufacturing process which is costly. So, ZIP model gives a wider, but more proper upper control limit to avoid the above undesirable situations.

Chapter 2

Shewhart Type control charts and the Zero Inflated Poisson model

In 2003, He et al. (2003) investigate for the first time the effect of estimation error concerning the ZIP model. So, in order to estimate someone the control limits, a historical Phase I sample is been utilized. Also, in this paper, He et al. (2003) investigate the sufficient number of samples in order to have the estimated control limits close enough to the true ones.

Given that the random variable $D \sim ZIP(p, \lambda_0)$ and α is a given false alarm probability, the UCL could be determined as:

$$1 - \alpha = P\left(D \leq UCL | p, \lambda_0, \alpha\right) = p - (1-p) \sum_{y=0}^{UCL} \frac{\lambda_0^y e^{-\lambda_0}}{y!} \tag{2.1}$$

He et al. (2003) used the MLE method to estimate the parameters of the ZIP model. The authors assume that the parameter p is known and they analyze the effect of sample size on the ZIP model charts through the estimation of parameter λ_0. Hence, $\hat{\lambda_0} = \frac{X}{m(1-p)}$, where X is the total number of defects in a sample of size m. Clearly, $X \sim Poisson\left(m\lambda_0(1-p)\right)$ and based on 2.1 the estimated UCL can be obtained as:

$$\widehat{UCL} = \text{Poisson Quantile}\{\frac{X}{m(1-p)}, \frac{(1-\alpha)-p}{1-p}\} .$$

Afterwards, the writers define the event $E_i = \{Y^i > \hat{UCL}\left(X\right)\}$, where Y^i is a future observation from a process with random shocks, possibly shifted from λ_0 to λ. The $P\left(E_i\right)$ is the actual alarm rate (AR) and when $\lambda = \lambda_0$ is named by false alarm rate (FAR). The $P\left(E_i\right)$ can be computed by using a conditional argument,

$$P\left(E_i\right) = E_X\{P\left(E_i | X = n\right)\} = \sum_{n=0}^{\infty} P\left(E_i | X = n\right) P\left(X = n\right)$$

(for details see He et al. (2003)).

So, according to the numerical results of FAR for selected values of m is ascertained that the FARs deviate from the ideal rates when the process average number of defects, $\hat{\lambda_0}$, is estimated from samples of size m. However, as the sample size m increases, the deviations from the ideal rates get smaller and that is because of the smaller variations in the estimations of λ_0. Hence, when m gets to a certain level with respect to λ_0, FAR is almost equal to the true rate and the estimation effects on

the FAR of the ZIP chart can be considered nonexistent.

In the case of the AR values when the process shifts from $\lambda_0 = 5$, the following are stressed:

- For small sample sizes, the AR is higher.
- When λ shifts from 1 to 4 the AR is (very) low because the number of defects decreases. When λ shifts from 6 to 10 the AR gets higher.

Next, the distribution of the Run Length is given. Let L be the number of points plotted on the ZIP chart until an out-of-control signal is happened. Denotind $P(E_i|X)$ by $p(X)$ th conditional distribution of L given X is geometric with parameter $p(X)$. Hence, the unconditional distribution of L is:

$$P(\ell, \lambda_0, \lambda)) = \sum_{n=0}^{\infty} p(n) [1 - p(n)]^{\ell-1} \cdot Poisson(m\lambda_0(1-p)).$$

Depending on the similar conditional arguments, the average run length (ARL) and the standard deviation of the run length (SDRL) have the above form:

$$ARL = E_X[\tfrac{1}{p(X)}] \quad \text{and} \quad SDRL = \sqrt{Var[\tfrac{1}{p(X)}] + E_X[\tfrac{1-p(X)}{p^2(X)}]}$$

(for details see He et al. (2003)).

Furthermore, the ARL and SDRL with known control limits are computed by:

$$ARL_0 = \tfrac{1}{P(T_i)} \quad \text{and} \quad SDRL_0 = \sqrt{\tfrac{1-P(T_i)}{P(T_i)}} = \sqrt{ARL_0(1 - ARL_0)}$$

The events $\{T_i\}$ are defined similarly as $\{E_i\}$, but corresponding to the known control limits. Hence, the sample size of m is large enough for $\widehat{UCL}$ to be essentially equal to UCL when $ARL \simeq SDRL$.

In order to study the effect of the sample size m on the ARL and SDRL when $\lambda_0 = 5$, the authors calculated the latter values for a range of values of m and λ. The results where the following:

- As the sample size increases, the difference between the estimated ARL and SDRL decreases and approaches the true values.
- When the sample size m is small, the estimated SDRLs are larger than the corresponding ARLs. However, as m increases, this differences get smaller and smaller until eventually ARL exceeds SDRL.

Finally, this research study tried to point out the importance and usefulness of the right sample size so as to achieve desired performance of the control chart. To do so, the false alarm probability equation has been given. Lastly, ARL and SDRL values were computed to show the effectiveness of the chart.

In 2008, Sim and Lim (2008) extended the findings of Xie et al. (2001); He et al. (2003) on the c-chart. He's He et al. (2003) proposal based on the fact that increases in the value of the parameters $1-p$ and λ of the ZIP model would lead to large number of defects. However, the problem with this chart was that increases in $1-p$ would not lead to large numbers of defects necessarily. The increases might concern only the none-zero Poisson count. So, the current paper is trying to define the upper control limit (UCL) of the c-chart by focusing on the one-sided Jeffreys prior interval for the ZIP parameter λ. Also, an analogous study which consider the zero inflated Binomial distribution take place in the paper, but it's not going to be included here. We focus only on the ZIP model.

In general, Jeffreys prior plays a special role in the Bayesian analysis. The form of Jeffreys prior is defined in terms of the Fisher information [1] as

$$\pi_j(\lambda) \propto \sqrt{I(\lambda)},$$

where the Fisher Information, $I(\lambda)$, is given by $I(\lambda) = -E_\lambda[\frac{\partial^2}{\partial\lambda^2}\log f(y|\lambda)]$ and $\log f(y|\lambda)$ is the log-likelihood of the ZIP distribution. Hence, if $Y = y$ is observed, then the $(1-\alpha)100\%$ one-sided Jeffreys prior interval for λ is defined as follows:

$$CI_J^\lambda(y) = [G(\alpha; y+0.5, 1), \infty),$$

where $G(\alpha; a, b)$ denotes the $100\alpha^{th}$ percentile of a Gamma distribution (Cai, 2005; Sim and Lim, 2008) . One advantage of Jeffreys prior interval is that the coverage probability is closer to the desired one than the confidence interval which is taken under the normal approximation.

The c-chart which is based on an upper Jeffreys prior interval with Poisson parameter λ estimated from the ZIP model is named as c_J-chart. The upper control limit of the c_J-chart is defined as the largest Poisson count y such that the estimated value, $\hat{\lambda}$, of λ (estimated from the MLEs) falls inside the $(1-\alpha)100\%$ interval $CI_J^\lambda(y)$. Hence,

$$UCL_{c_J}\left(\hat{\lambda}\right) = max[y|\hat{\lambda} > G(\alpha; y+0.5, 1)].$$

So, a Poisson count that is plotted above the $UCL_{c_J}\left(\hat{\lambda}\right)$ is evidence that the process is out of control.

The authors give an example in which data are generated from an in control ZIP process with unknown process mean λ. The MLEs of λ and p were taken from the likelihood of the ZIP distribution. Also, the 99.73% $CI_J^\lambda(y)$ were computed. Then, The upper control limit was calculated based on:

1. Jeffreys prior interval.
2. $z_{0.0027}$-sigma control limit and the parameter λ which was estimated by using the ZIP distribution.
3. $z_{0.0027}$-sigma control limit and the parameter λ which was estimated by using the Poisson distribution.

Only the first approach showed that all the observed counts fell into the control limits.

Furthermore, a performance of the c_J-chart via its ARL was took place in the current paper. So, the ARL of the c_J-chart is the following:

$$ARL(\Delta) = \frac{1}{P_{c_J}(\Delta)},$$

where Δ is a shift of the Poisson process mean from λ_0 to $\lambda_1 = \lambda_0 + \Delta, \Delta \geq 0$.

$$P_{c_J}(\Delta) = P(\text{point out of control}) = P\left(X > UCL_{c_J}(\lambda_0)\,|\lambda = \lambda_1\right) = \sum_{x=UCL_{c_J}(\lambda_0)+1}^{\infty} \frac{\lambda_1^x e^{-\lambda_1}}{x!}.$$

Given the ARL values regarding the classical c-chart (with and without the ZIP model) and the c_J-chart with $p = 0.4$ and $\lambda = 4, 4.5, 5, 5.5$ the following conclusions are made:

[1] It measures the expected amount of information given by a random variable Y for a parameter λ of interest.

- ARL_0 values are bigger as far as c_J-chart is concerned for all values of parameter λ. The value of ARL_0 of the c-chart (not based on the ZIP model) is unacceptably small for all values of parameter λ.
- ARL_1 values are decreasing when the shift of the parameter λ get bigger for all values of parameter λ.
- The c-chart (without taking into consideration the ZIP model) has the smallest ARL_0 in comparison with the others for all values of parameter λ.

In addition, a simple two-of-two control rule [2] (Sim and Lim, 2008) is also recommended for further improvement according to the performance of the proposed c_J-chart. The c_J-chart based on the 99.73 % Jeffreys prior interval yields an $ARL_0 = 352.1$ and the c_J-chart based on the 99.73 % Jeffreys prior interval and the two-of-two control rule gives an $ARL_0 = 402.013$. It's not only that, the two-of-two control rule gives also smaller ARL_1 values. However, the latter does not hold in general. For instance, when $\lambda = 4.5$ ARL_0 of c_J-chart and ARL_0 of c_J-chart with two-of-two control rule are 415.906 and 641.876, respectively. The effect is that the latter one fails to reduce the ARL_1 value because of the over coverage of Jeffreys interval at $\alpha = 0.0533$. For this reason, the writers of the paper provide a guideline which tells us when the control rule should be used (Sim and Lim, 2008) .

Considering all the above, this article propose Jeffreys interval which is a confidence interval with good coverage probability . Moreover, depends on the study which has preceded, a two-of-two control rule can be used in order to improve the performance of the concerning chart.

In 2012, Katemee and Mayureesawan (2012a) write a paper in which three new control charts of nonconformities are evolved based on a ZIP distribution. These modified versions of the c-chart are going to be compared with the c-charts which are mentioned to previous papers.

First of all, the authors develop an approximation of the ZIP distribution, based on Kolmogorov-Smirnov test (Gibbons and Chakraborti, 2003), by using a Non-central Chi-square distribution with parameter λ_{Chi} . The probability function of the Non-central Chi-square distribution is the following:

$$f\left(x|n,\lambda_{Chi}\right) = \sum_{k=0}^{\infty} \overbrace{\frac{e^{-\frac{\lambda_{Chi}}{2}}\left(\frac{\lambda_{Chi}}{2}\right)^k}{k!}}^{\text{Poisson}\left(\frac{\lambda_{Chi}}{2}\right)} \overbrace{\frac{e^{-\frac{x}{2}}x^{\frac{n+2k}{2}-1}}{2^{\frac{n+2k}{2}}\Gamma\left(\frac{n+2k}{2}\right)}}^{X^2_{n+2k}},$$

where $x > 0, n > 0$ and $\lambda_{Chi} > 0$. More specifically, n here is the number of degrees of freedom and $\lambda_{Chi} = \sum_{i=0}^{k}(\frac{\mu_i}{\sigma_i})^2$. Also, $E\left(X\right) = n + \lambda_{Chi}$ and $V\left(X\right) = 2(n + 2\lambda_{Chi})$ are the mean and variance of the above distribution, respectively (Katemee and Mayureesawan, 2012a; Krishnamoorthy, 2006).

Hence, the three new charts for the number of nonconformities are the above:

1. c_{Chi}**-Chart** : The control limit of this chart is eventuated by replacing c with λ_{Chi}, i.e

$$UCL = \lambda_{Chi} + 3\sqrt{\lambda_{Chi}} \quad \text{and} \quad LCL = 0$$

[2] The two-of-two control rule states the presence of assignable causes if two points in row fall above (or bellow) the UCL (or LCL).

2. c_{CChi}**-Chart** : The control limit of this chart is eventuated by replacing the estimated values of the mean and variance in the UCL with the estimators of the mean and variance of the concerning distribution respectively, with $n = 0$, i.e

$$UCL = \lambda_{Chi} + 3\sqrt{4\lambda_{Chi}} \quad \text{and} \quad LCL = 0$$

3. c_{MChi}**-Chart** : The control limit of this chart is eventuated by replacing the estimated value of the mean with the estimator of the mean of the concerning distribution with $n = 0$ and the estimated value of the variance is replaced with the interquartile range $(IQR\,(c) = Q_3 - Q_1)$ which is a measure of variability, based on dividing a data set into quartiles, i.e

$$UCL = \lambda_{Chi} + 3\sqrt{Q_3 - Q_1} \quad \text{and} \quad LCL = 0$$

Next, the authors proceed to a simulation study in order to document the results of the three new charts and compare them with the previous ones. To begin with, they simulate data from a ZIP distribution with different values of the parameters. For each of them they calculate the λ_{Chi} which gives a best fit between the ZIP distribution and a Non Central Chi square with $n = 0$. In addition, a Kolmogorov-Smirnov test is occurred 20000 times to see if the Non Central Chi square distribution with $n = 0$ and λ_{Chi} fits the data which are simulated by a specific ZIP distribution. Lastly, 100000 replications are accomplished to calculate the averaged control limits for the c-charts in interest and a new set of 100000 replications then used to compute the ARL and the ACP (Average Coverage Probability) values for each of the charts.

The main conclusions of the authors, after the simulation results, are the above:

- The requested λ_{Chi} which gives the best fit, varies and depends on the values of p and λ of the ZIP distribution. When $p = 0.7 - 0.9$, then λ_{Chi} gets the constant value of 1.44 regardless of the value of λ.
- For $p = 0.3 - 0.5$ the c_{CChi}-chart has the highest ARL_0 values for all levels of λ. But when $p = 0.6 - 0.9$ the c_{ZIP} and c_J are more suitable charts depends on the ARL_0 (Katemee and Mayureesawan, 2012a).
- Also, when $p = 0.3 - 0.5$ the ACP of the c_{CChi}, c_{ZIP}, c_J charts are very high, i.e the ACP values are close to the target level of 0.9973. However, for $p = 0.6 - 0.9$ the c_{CChi}- chart does not give so desired ACP values as compared with the two others.
- When the condition is **out of control**, the charts detect shifts slowly as $p = 0.8$ or 0.9. As far as ACP is concerned, when $p = 0.3 - 0.7$ the c_{MChi}-chart has the highest value of ACP for all the tested shifts. But, when $p = 0.8, 0.9$ the ACP of the c_{Chi}-chart has a value closer to the target value of 0.9973 (Katemee and Mayureesawan, 2012a).

In the same year, the same authors present something analogous. They use a Geometric distribution with parameter p_g to approximate a ZIP process.

To start with, the probability function of the Geometric distribution is given by:

$$P\,(X = x) = (1 - p_g)^k p_g \ , \quad k = 0, 1, 2, ...,$$

where X is the number of failures until the occurrence of a success and p_g is the probability of success on each event. The mean and variance of the distribution is $E\,(X) = \frac{1-p_g}{p_g}$ and $V\,(X) = \frac{1-p_g}{p_g^2}$, accordingly.

So, in developing the charts, the number of nonconformities is modeled by a Geometric(p_g). The estimated p_g, $\hat{p_g}$, gives the best fit between the Geometric and ZIP distribution. Hence, the three new charts for the ZIP process are defined as follows:

1. c_g**-Chart** : The control limit of this chart is obtained by using the estimated p_g, $\hat{p_g}$. The authors used the MLE of p_g, which is $\hat{p_g} = (1 + \frac{1}{n}\sum_{i=0}^{n} k_i)^{-1}$ and $k = k_1, k_2, ..., k_n$ is a sample where $k_i \geq 1$ for $i = 1, 2, ..., n$. Therefore, $E(X) = \frac{1-\hat{p_g}}{\hat{p_g}}$ and $V(X) = \frac{1-\hat{p_g}}{\hat{p_g}^2}$. Hence,

$$UCL = \frac{1-\hat{p_g}}{\hat{p_g}} + 3\sqrt{\frac{1-\hat{p_g}}{\hat{p_g}^2}} \quad \text{and} \quad LCL = 0$$

2. c_{mg}**-Chart** : The control limit of this chart is eventuated by replacing both the mean and the variance in the UCL of the Shewhart c-chart with the estimator of the mean of the Geometric distribution, i.e

$$UCL = \frac{1-\hat{p_g}}{\hat{p_g}} + 3\sqrt{\frac{1-\hat{p_g}}{\hat{p_g}}} \quad \text{and} \quad LCL = 0$$

3. c_{me}**-Chart** : The control limit of this chart is eventuated by replacing the mean with the median of the Geometric distribution. The form of the concerning median is $M = [\frac{-1}{log_2(1-\hat{p_g})}] - 1$. Hence,

$$UCL = M + 3\sqrt{\frac{1-\hat{p_g}}{\hat{p_g}}} \quad \text{and} \quad LCL = 0$$

(Katemee and Mayureesawan, 2012b).

Next, the authors proceed to a simulation study in order to document the results of the above three new charts and compare them with all the c-charts which have mentioned previously. To begin with, the authors simulate data from a ZIP distribution with different values of the parameters. For each of them they calculate the p_g which gives a best fit between the ZIP distribution and the Geometric. In addition, a Kolmogorov-Smirnov test is occurred 20000 times to see if the Geometric(p_g) fits the data which are simulated by a specific ZIP distribution. Lastly, 100000 replications are accomplished to calculate the averaged control limits for the c-charts in interest and a new set of 100000 replications then used to compute the ARL and the ACP values for each chart.

After the simulations, the authors come to the following conclusions:

- The requested $\hat{p_g}$, which gives the best fit, varies and depends on the value of p and λ of the ZIP distribution. However, when $\lambda = 1$ and $p = 0.3 - 0.9$, then $\hat{p_g}$ gets the constant value of 0.53. The same thing happens for $\lambda = 2$, $p = 0.7 - 0.9$ and $\lambda = 3$, $p = 0.8 - 0.9$ respectively (Katemee and Mayureesawan, 2012b).

- For $\lambda = 1$ and all levels of p, c_g-chart and c_J-chart have almost the similar highest ARL_0 values. Hence, they are the most preferable charts because of the slow shift detection. Also, when $\lambda = 2-4$ and $p = 0.3, 0.4$ the c_g-chart has the highest ARL_0 value, but when $p = 0.5-0.9$ c_J-chart is more appropriate (Katemee and Mayureesawan, 2012b).

- When $p = 0.3, 0.4$ the c_g and c_J charts have in general similar high ACP for all values of λ, but when $p = 0.5 - 0.9$ the c_{ZIP} and c_J charts gives ACP values close to the target level of 0.9973. So, when both ARL_0 and ACP values are considered, the c_g and c_J charts will be more preferable for $p = 0.3, 0.4$ and $\lambda = 1$.

- When the condition is **out of control**, the charts detect shifts slowly as $p = 0.8$ or 0.9. As far as ACP is concerned, the c_{mg} and c_{me} charts have the highest values close to the target value of 0.9973 for all values of λ and for all the tested shifts. When both ARL_1 and ACP values are considered, for the performance of the charts, c_{mg} and c_{me} will be the preferred ones for all shifts, as $\lambda = 1$ and $p = 0.3 - 0.6$. However, for $p = 0.7 - 0.9$ none of the charts are acceptable. Lastly, for all values of p, for all the tested shifts and $\lambda = 2-4$, c_{mg} and c_{me} charts are preferred (Katemee and Mayureesawan, 2012b).

Almost 12 years later, Rakitzis and Castagliola (2015) write a paper which refers again to the effect of the estimated parameters on the performance of one-sided Shewhart control chart for ZIP process and give guidance on how to design the examined chart when the sample size is predetermined. Until the writing of this article, only He et al. (2003) had dealt with this mater.

The article begins by giving the, already known, upper control chart limit of the upper-sided ZIP Shewhart control chart for ZIP with known parameters, which is, the rounded down integer of $UCL_{ZIP} = E(Y) + K\sqrt{VAR(Y)}$, where $Y \sim ZIP(p, \lambda_0)$ and $K > 0$ is a constant that plays the role of chart's design parameter. Also, the probability

$$\beta = P(Y_i > UCL_{ZIP} | \lambda = \lambda_1) = 1 - F_{ZIP}(UCL_{ZIP} | p, \lambda_1)$$

.

In this article, Rakitzis and Castagliola (2015) symbolize the pmf and cdf of the Run Length as

$$f_L(\ell) = \beta(1-\beta)^{\ell-1} \quad \text{and} \quad F_L(\ell) = 1 - (1-\beta)^{\ell} \ , \ \ell = 1, 2, ... \quad \text{respectively.}$$

Therefore, the ARL and SDRL of the ZIP Shewhart control chart with known parameters, are equal to

$$ARL = \tfrac{1}{\beta} \quad \text{and} \quad SDRL = \tfrac{\sqrt{1-\beta}}{\beta} \ .$$

Afterwards, the upper-sided ZIP Shewhart control chart with estimated λ_0 is of interest. So, assuming that a Phase I data set exists and it is composed by m independent random variables which follow a $ZIP(p, \lambda_0)$, the MLE, $\tilde{\lambda_0}$, can be taken from

$$\tilde{\lambda_0} = \frac{1}{m(1-p)} \sum_{i=1}^{m} Y_i = \frac{W}{m(1-p)} \ .$$

So, the upper control chart limit of the upper-sided ZIP Shewhart control chart for ZIP with unknown parameters is the rounded down integer of

$$\widetilde{UCL}_{ZIP} = \tilde{\lambda_0}(1-p) + K\sqrt{\tilde{\lambda_0}(1+\tilde{\lambda_0}p)(1-p)}$$

$W = \sum_{i=1}^{m} Y_i$ is not a ZIP random variable and its pmf can be numerically evaluated via its probability generation function (Rakitzis and Castagliola, 2015).

By replacing $\tilde{\lambda_0}$ with $\frac{W}{m(1-p)}$, the probability

$$\tilde{\beta} = P\left(Y_i > \widetilde{UCL}_{ZIP} | W = w, \lambda = \lambda_1\right) = 1 - F_{ZIP}\left(\frac{w}{m} + K\sqrt{\frac{w}{m}(1 + \frac{wp}{m(1-p)})} | p, \lambda_1\right)$$

Furthermore, the unconditional ARL of the run length L of the upper-sided ZIP-Shewhart control chart with estimated parameter λ_0 are, respectively, given by:

$$ARL = \sum_{w=0}^{\infty} f_W(w)\left(\frac{1}{\tilde{\beta}}\right)$$

$$SDRL = \sqrt{E(L^2) - ARL}, \text{ where } E(L^2) = \sum_{w=0}^{\infty} f_W(w)\left(\frac{2-\tilde{\beta}}{\tilde{\beta}^2}\right)$$

(for details see (Rakitzis et al., 2016)).

Taking into account the above forms of ARL and SDRL, it is understood that a computational issue arises. Hence, in order to be managed this computations, one possibility, according to the authors, is the approximation of the pmf of W by an appropriate continuous distribution, the shifted Gamma distribution, $f(y-c|a,b)$ (see details about the parameters in appendix of Rakitzis et al. (2016)).

Depending on that, the authors accomplish all the numerical calculations which concern ARL and SDRL values. So, they choose $m \in \{100, 200, 500, 1000, 2500, \infty\}$, $\lambda_0 \in \{1, 2, 3..., 8\}$ and $p \in \{0.9, 0.8, 0.7\}$. The value $m = \infty$ represents the known parameter case. Moreover, the constant value K takes value in order to give an in control ARL value as close as possible to the desired $ARL_0 = 370.4$. The main results of this analysis are the following:

- For processes with an excessive number of zeros, very large preliminary samples are needed so as to obtain the performance of the upper-sided ZIP Shewhart control chart in the estimated parameter case close to the performance in the known parameter case.
- For (relatively) small preliminary samples ($m = 100$ *or* 200), the in-control ARL and SDRL values are very large ($> 10^6$) , in the most of the examined cases.
- As the value of m increases, the in-control ARL values tend to decrease in the most of the examined cases and the values tend to converge as m approaches the case $m = \infty$.
- Using the same design parameter K, the target is to find how large the size m of the Phase I sample must be in order to have approximately the same in-control ARL values in both the known and the estimated parameter case. The numerical analysis shows that as m goes to ∞, the in control ARL values does not converge to the in-control ARL value in the known parameter case.
- For this reason, "corrected" chart's design parameters K' for the studied schemes that takes the sample size m into consideration are provided in order to derive the desired in-control performance when the process' parameters are unknown. Given the size m of the preliminary sample and using the proposed K', the in-control ARL value of the proposed schemes in the estimated parameter case is very close to the respective value in the known parameter case, while SDRL in-control value is decreasing importantly.

Also, Rakitzis et al. (2016) mention a corresponding theory and analysis for the zero inflated Binomial distribution, which is not going to be mentioned here.

At the end, the writers include some suggestions for further research, such as the case which both the parameters are unknown (p and λ) as well as the analysis of the performance of other schemes like CUSUM-type and EWMA-type when the parameters are unknown and have to be estimated.

Chapter 3

Time-weighted control charts and the Zero Inflated Poisson model

3.1 CUSUM control charts

In 2008, Chen et al. (2008) proposed the Generalized Zero Inflated Poisson (GZIP) distribution, which is an extension of the ZIP distribution, in order to present a control chart to inspect attribute data.

More thoroughly, in real world application there are times in which the processes subject to different kinds of random shocks.This may have various impacts on the processes, therefore resulting in different number of nonconformoties. Consequently, a single Poisson cannot describe all the shocks. For this reason the ZIP model can be extended to the GZIP model and solve that problem. So, if n kinds of shocks occur, the number of nononformities caused by each shock follows a specific Poisson with parameter λ_i and the occurring probability of each shock is w_i, then the distribution of the number of nonconformities in the process is the GZIP and can be expressed as:

$$P(x; w, \lambda) = \begin{cases} \sum_{i=1}^{n} w_i + \sum_{i=1}^{n} (1 - w_i) e^{-\lambda_i} & \text{for } x = 0 \\ \sum_{i=1}^{n} (1 - w_i) \frac{e^{-\lambda_i} \lambda_i^x}{x!} & \text{for } x = 1, 2, ... \end{cases}$$

When $n = 1$, GZIP degenerates into the original ZIP model.

Afterward, the MLE method is discussed by the authors in order to estimate the parameters, noted by $\theta = (\lambda_1, ..., \lambda_n, w_1, ..., w_n)$, of the above model that represent the in control data the best. To do so, it is necessary to pre-specify the model component number n. Sometimes, that number can be defined by the domain knowledge, if it's available, or AIC or BIC criteria can be utilized for that purpose. Given that n has been determined, the EM method can be used to estimate the parameters of the GZIP distribution (Chen et al., 2008).

Next, the authors present three control chart construction methods which are rely on the GZIP distribution. Those control charts are the following:

1. **Shewhart-type chart**
   Given the specified error α and the estimated parameter θ from the historical in control data, the UCL can be computed as:
   $$P(x > UCL|\theta) \leq \alpha.$$

2. **CUSUM chart**

   An alternative to the Shewhart type control chart when small shifts are of interest is the CUSUM chart. Given the control limit equals to h, the alarm will be happened when the monitored statistic $g_l > h$. The CUSUM statistic is updated as follows:

$$g_l = \max(g_{l-1} + ln\frac{P\left(X_l|\theta_\mathbf{1}\right)}{P\left(X_l|\theta_\mathbf{0}\right)}, 0),$$

   where $g_0 = 0$, X_l is the current sample, θ_0, θ_1 are the model parameters before and after the change accordingly (Chen et al., 2008). An approximation of the control limit h, based on the specified α error, is given by the relationship $h = -ln\alpha$ (Basseville and Nikiforov, 1993).

3. **Ranked Probability Control chart (RPC- chart)**

   The authors propose the above chart when some samples have small mass probability and therefore they are more unlikely to occur under the in control condition. So, the authors recommend to compute the probability of samples with values from 0 to V, where V have to eligible the following : $\max_{X>V} P(X) \leq \min_{X\leq V} P(X)$. After that, the probabilities from P(0) to P(V) are sorted in decreasing order. Hence, the out- of-control points which depend on a specific α error can be defined. Denoting the ranked probability distribution by $P_s\left(0\right), ..., P_s\left(V\right)$ the control limit can be determined by $P_s\left(X \leq UCL\right) > 1 - \alpha$ (Chen et al., 2008).

At this point, it is useful to mention that for some models, for example if the probability distribution is monotonically decreasing, or for a small enough α error, Shewhart and RPC charts will give identical results.

After presented the control charts, the authors studied the operation characteristics of them. They considered two types of changes according to the parameters of the GZIP models. Thus, $\omega_i = \omega_i + \Delta\omega$, so $\theta^{'} = (\omega_1^{'}, ..., \omega_n^{'}, \lambda_1, ..., \lambda_n)$, where

$$\omega_k^{'} = \begin{cases} \frac{\omega_k + \Delta\omega}{1+\Delta\omega} & k = i \\ \\ \frac{\omega_k}{1+\Delta\omega} & \text{for } k \neq i \end{cases}$$

The standardization of these parameters is needed because of the requirement $\sum_{i=0}^{n} \omega_i = 1$. Also, λ_i changes to $\lambda_i + \Delta\lambda$ and θ will become $\theta^{'} = (\omega_1, ..., \omega_n, \lambda_1, ..., \lambda_i + \Delta\lambda_i, ..., \lambda_n)$.

After some simulations are conducted by the authors, the conclusions are that the shape of the distribution is not affected so much by the change in ω_i, but it is sensitive to the change in λ_i. The authors use a specific GZIP distribution to illustrate the outcomes when the parameters λ_i, ω_i are changed. The conclusions are the following:

- The performance of RPC chart in detecting relatively large λ_i changes is better in comparison with the other two charts.
- The performance of RPC and Shewhart charts in detecting relatively small λ_i changes is exactly the same.
- The CUSUM chart is more efficient in detecting ω_i changes. However, its performance is often not considered the best possible because the construction of the chart needs the knowledge of the shift in advance.
- An interesting thing is the up-slopping trend of the OC curve according to the increase of the ω_i parameter. The explanation of the above phenomenon is that the control charts are

made to detect more nonconformities compared with normal situations, thus in this case, i.e $\lambda_1 = 0, 0.1, 0.2, ..., 3$, the charts fail to detect the changes toward small number of occurrences (Chen et al., 2008).

Finally, by taking everything into account, the authors concluded that RPC chart is the most preferable choice between the other two.

The next step is to prove with simulations that GZIP models can approximate many discrete distributions. To do so, the authors use a Poisson, a Geometric and a GZIP distribution to generate 500 samples from each model. The parameters of each model are estimated by the MLE method. After that a pair-wise X^2 test is used to check the goodness of fit of the three models. The results are that only the GZIP model is flexible and can be used to model data from other distributions. Furthermore, the paper provides the control limits for the c, g and RPC charts for the above distributions respectively with α error equals to 0.01. Also, β error is utilized in order to assess the performance of each chart. Hence, for the Poisson distribution the parameter increases from 3 to 23 and the RPC chart has smallest β error for any shift. For the Geometric distribution the parameter decreases from 0.6 to 0.05 and g-chart performs as the RPC chart. For the GZIP the mean changes from 7.2 to 25.2 and the RPC chart performs better than the others. In general, RPC chart performs the best.

At the end of the paper, a real case study takes place.The real data concern CT scanner usage log files and by observing and analyzing the logs, someone can inspect the performance of the machine. In the log files the occurring frequency of a particular failure event is under the microscope of this study. The goal is to create a statistical model to approximate the failure event occurrence distribution under in control situations, and according to that to establish control charts to observe the working condition of CT scanners. The samples, which are in control, are 112, the MLE method, which was described previously, and the AIC criterion are used to select the proper model that has three shock components n this case. The accuracy of this chosen model is illustrated by using and other kinds of estimated models such as Poisson and Geometric and by testing the goodness of fit of them. The outcome of those trials is that the GZIP model can fit the data well and the use of other charts other than the RPC chart will generate too many false alarms.

In the end, Chen et al. (2008) shared their thoughts about some issues which need further investigation such us the sensitivity of their methods to the existence of noise and the magnitude of the sample size.

Motivated by the exact above paper, in 2011 He et al. (2011) proposed a method for detecting increases in both ZIP parameters. The thought was to combine two CUSUM charts to detect increases in either p or λ parameters. This method is beneficial, because the authors managed to identify which shifted parameter makes the post-signal diagnosis more direct in practice.

More analytically, the authors denote as p-CUSUM and λ-CUSUM the charts which detect shifts in p and λ parameters respectively. The combination of the above charts is named after $p-\lambda$ CUSUM and the CUSUM chart which is designed to detect simultaneous increases in the two parameters is the t-CUSUM chart. So, it is assumed that the in-control values for ZIP parameters are known to be p_0, λ_0. Hence, the design of those control charts is presented below:

1. **p-CUSUM chart**
   This chart is designed to detect shifts from p_0 to p_1 ($p_1 > p_0$). The p-CUSUM control statistics are constructed based on the likelihood ratio method:

$$B_i = \max(B_{i-1} + K_i, 0), \ i = 1, 2, ...$$

where $B_0 = 0$ and

$$K_i = K(Y_i) = \begin{cases} ln\frac{p_1+(1-p_1)e^{-\lambda_0}}{p_0+(1-p_0)e^{-\lambda_0}} & Y_i = 0 \\ \\ ln\frac{1-p_1}{1-p_0} & Y_i > 0 \end{cases}$$

The above chart signals when $B_i > h_p$, where h_p is the selected UCL which accomplishes the required in-control performance.

2. **λ-CUSUM chart**

This chart is developed to detect increases in the Poisson parameter. So, λ_1 is the shift from the in-control value λ_0. As it mentioned above, the λ-CUSUM statistics are given as:

$$L_i = \max(L_{i-1} + M_i, 0),\ i = 1, 2, ...$$

where $L_0 = 0$ and

$$M_i = M(Y_i) = \begin{cases} ln\frac{p_0+(1-p_0)e^{-\lambda_1}}{p_0+(1-p_0)e^{-\lambda_0}} & Y_i = 0 \\ \\ Y_i ln\frac{\lambda_1}{\lambda_0} + (\lambda_0 - \lambda_1) & Y_i > 0 \end{cases}$$

The above chart signals when $L_i > h_\lambda$, where h_λ is the selected UCL based on the pre-specified in-control performance.

For the method which is proposed by the authors, **the $p - \lambda$ CUSUM chart**, a signal would be happened whenever one of the above individual CUSUM charts signals.

3. **t-CUSUM chart**

This chart is designed to detect the out of control parameter values of the ZIP model simultaneously. The t-CUSUM control statistics are:

$$T_i = \max(T_{i-1} + N_i, 0),\ i = 1, 2, ...$$

where $T_0 = 0$ and

$$N_i = N(Y_i) = \begin{cases} ln\frac{p_1+(1-p_1)e^{-\lambda_1}}{p_0+(1-p_0)e^{-\lambda_0}} & Y_i = 0 \\ \\ Y_i ln\frac{\lambda_1}{\lambda_0} + (\lambda_0 - \lambda_1) + ln\frac{1-p_1}{1-p_0} & Y_i > 0 \end{cases}$$

The above chart signals when $T_i > h_t$, where h_t is the selected UCL based on the pre-specified in-control performance (He et al., 2011).

Afterward, the authors use the Average Number of Observations to Signal (ANOS or ARL) to measure the performance of the control charts. So, to design the above CUSUM charts it is assumed that the in-control values of the parameters are known or they can be estimated using the MLE method and the shift sizes of the parameters must be pre-specified. The control limits of the CUSUM charts are defined in a way that the ANOS values are close to a pre-determined value, say, $ANOS_0$.

After that, they proceed to a simulation study. They consider two in control values for each parameter. The values are $1 - p_0 = 0.1$ and 0.2 , $\lambda_0 = 2$ (which has high influence in the occurrence of zero defects) and $\lambda_0 = 6$ (which has minimal effect on the probability of happening a zero-count). The authors choose the pre-specified shift sizes to be $1 - p_1 = 1.25(1 - p_0)$ or $1 - p_1 = 1.5(1 - p_0)$ for the first parameter and $\lambda_1 = \lambda_0 + 1$ or $\lambda_1 = \lambda_0 + 2$ for the other parameter. Lastly, the $ANOS_0 = 200$

for $p_0 = 0.2$ and $ANOS_0 = 340$ for $p_0 = 0.1$. Hence, the results, in which the simulation study concludes, are the following:

- First of all, the writers compare the performance of the $p - \lambda$ and t-CUSUM charts with the classic ZIP-Shewhart chart (Xie et al., 2001) and they found that in general the ZIP-Shewhart chart is worse in detecting shifts in p and small shifts in λ in comparison with the other two charts.

So, the authors do not take into consideration the ZIP-Shewhart chart and deal with the four forenamed CUSUM charts.

- The λ-CUSUM chart signals more quickly as $1-p$ increases, because of the increased sampling rate of the Poisson distribution. However, the p-CUSUM chart is not influenced by increases in λ.
- When $1-p_0 = 0.2$ and $\lambda_1 = \lambda_0 + 2$ or $\lambda_1 = \lambda_0 + 1$, the p-$\lambda$ CUSUM chart is better in detecting shifts in p while t-CUSUM chart is a little bit better in detecting shifts in λ.
- When there are simultaneously increases in both in-control values $1-p_0, \lambda_0$ the t-CUSUM chart performs better than the p-λ CUSUM chart. In case that the reverse is happened p-λ CUSUM chart is more appropriate in terms of the total performance.

At the end, a case study about a LED packaging process is presented in the paper to illustrate how the proposed methods can be used in a manufacturing application. The LED products are produced batch by batch, therefore after the construction of the LEDs, the batch were inspected and the defectives are counted and recorded. For the actualization of the implementation the authors select 200 samples of the whole data set (2067 batches) and took the first 100 samples in order to apply a score test (Xie et al., 2001) to make sure that the data come from a ZIP distribution. Also, using the MLE method they estimate the in-control values of p and λ and choose the appropriate control limits based on the requirement $ANOS_0 \simeq 200$. The remaining 100 samples are used to check the performance of the above mentioned CUSUM charts. The results confirm the conclusions of the simulation study about the concerning charts.

In the previous article, He et al. (2011) concentrated on the ZIP distribution and they proposed two CUSUM charting procedures for monitoring them. This study focused on the situation where there is only one type of defect. However, in many high quality manufacturing processes, there are often a few types of defects which produced from a variety of equipment problems. Based on that He et al. (2012) this time propose a CUSUM based procedure to monitor a BZIP distributed process. The authors examine the situation where there are only two type of defects.

So, in order to proceed forward, the writers refer to the parameters of the BZIP distribution p_0, p_1, p_2, p_{12} as the p-set parameters and the parameters $\lambda_1, \lambda_2, \theta_{12}$ as the λ-set parameters. The p-set parameters are related to the probability of observing defects and the λ-set parameters to the number of defects observed in products with non-conformities. Thus, on the lines below we are going to briefly mention the methodologies of the two CUSUM procedures for detecting shifts in the two sets of parameters in a BZIP process.

1. Control Procedure for Monitoring **Shifts** in the **p-set Parameters**
   We use $p_{01}, p_{11}, p_{21}, p_{121}$ to denote the shifted p-set parameters that must be detected quickly with the restriction $p_{01} + p_{11} + p_{21} + p_{121} = 1$. The decrease in p_0 and increases in p_1, p_2, p_{12} are of greater interest because of the production of more non-conformities. Now, supposing that

$Y_i = (y_{i1}, y_{i2}), i = 1, 2, ...$ is a sequence of BZIP distributed observations and there are no shifts in the λ-set parameters, the one-sided CUSUM procedure based on the log-likelihood ratios for monitoring the p-set parameters is the above:

$$SP_i = \max(SP_{i-1} + LP_{ki}, 0) \text{ for } k = 1, 2, 3, 4 \text{ and } i = 1, 2, ...$$

where $SP_0 = 0$ and LP_{ki} are defined as follows:

$$LP_{ki} = \begin{cases} \ln \frac{P_{01}}{P_{00}} & \text{for } k = 1 \text{ and } y_{i1} = 0, y_{i2} = 0 \\ \ln \frac{P_{11}}{P_{10}} & \text{for } k = 2 \text{ and } y_{i1} > 0, y_{i2} = 0 \\ \ln \frac{P_{21}}{P_{20}} & \text{for } k = 3 \text{ and } y_{i1} = 0, y_{i2} > 0 \\ \ln \frac{P_{121}}{P_{120}} & \text{for } k = 4 \text{ and } y_{i1} > 0, y_{i2} > 0 \end{cases}$$

For the numerators the formula of the BZIP distribution 1.8 is used and the p-set parameters are substituted with the shifted ones while the denominators represent the BZIP distribution with the in-control p-set parameters for each case (He et al., 2012). The above chart signals when $SP_i > h_{p_s}$, where h_{p_s} is the selected UCL which accomplishes the required in-control performance.

2. Control Procedure for Monitoring **Shifts** in the **λ-set Parameters**
We use $\lambda_{11}, \lambda_{21}, \theta_{121}$ to denote the shifted λ-set parameters that is desired to be detected quickly. The increase in the λ-set parameters indicates that the observing number of defects will also increase. Therefore, assuming that there are no shifts in the p-set parameters, the log-likelihood ratios are used again to define a CUSUM procedure for monitoring the λ-set parameters. Hence,

$$SL_i = \max(SL_{i-1} + LL_{ki}, 0) \text{ for } k = 1, 2, 3, 4 \text{ and } i = 1, 2, ...$$

where $SL_0 = 0$ and LL_{ki} are defined as follows:

$$LL_{ki} = \begin{cases} \ln \frac{L_{01}}{L_{00}} & \text{for } k = 1 \text{ and } y_{i1} = 0, y_{i2} = 0 \\ \ln \frac{L_{11}}{L_{10}} & \text{for } k = 2 \text{ and } y_{i1} > 0, y_{i2} = 0 \\ \ln \frac{L_{21}}{L_{20}} & \text{for } k = 3 \text{ and } y_{i1} = 0, y_{i2} > 0 \\ \ln \frac{L_{121}}{L_{120}} & \text{for } k = 4 \text{ and } y_{i1} > 0, y_{i2} > 0 \end{cases}$$

For the numerators the formula of the BZIP distribution 1.8 is used and the λ-set parameters are substituted with the shifted ones while the denominators represent the BZIP distribution with the in-control λ-set parameters for each case (He et al., 2012). The above chart signals when $LP_i > h_{\lambda_s}$, where h_{λ_s} is the selected UCL which achieves the desired in-control performance.

The authors, also, mention the case in which only one CUSUM chart is necessary to detect simultaneous changes in both the p-set and λ-set parameters. The CUSUM statistics for that will be something analogous to the previous statistics with the difference that the shifts in both the set parameters will be detecting at the same time (He et al., 2012). However, using this method it is not clear which parameter or subset of them shift when a signal happens. This drawback makes it complicate to identify and interpret the out-of-control signals. Hence, the article proposes the use of the combination of the two forenamed CUSUM charts instead.

The next move is to evaluate the performance of the proposed charts. For this purpose the authors use the ANOS value. So, they set the in-control ANOS values for the p-set CUSUM chart

and the λ-set CUSUM chart to be approximately the same and the ANOS value of the combination of the above charts to be about 200. Four cases of the BZIP model are referred in order to assess the detecting effectiveness. Simulations are used to obtain the UCLs of the control harts and three possible shift magnitudes (small, moderate, large) are considered as regards the p-set and λ-set parameters, accordingly.

Thus, three cases are studied: shifts only in the p-set parameters, shifts only in λ-set parameters and shifts in both of them. The outcomes of the simulation study are the following:

- For the first situation, the ANOS values decrease with the increase of some parameters. The bigger the shift in the parameters, the smaller becomes the ANOS values. Also, while there are no shifts in the λ-set parameters the ANOS values are also smaller than the $ANOS_0$ and that happens because the increase of the p-set parameters generates more non zero counts.

- For the second situation, the ANOS values decrease with the increase of the shift magnitude. Also, while there are no shifts in the p-set parameters the ANOS values of the p-set CUSUM chart are also smaller than the $ANOS_0$.

- For the third situation, if there are shifts in both the set of parameters, either the p-set CUSUM or the λ-set CUSUM charts ought to signal. It is shown that the combination of those charts is efficient in detecting shifts in all parameters and slightly better than the two other charts (He et al., 2012).

Finally, an application of the proposed method are presented based again on data from the LED packaging industry. This time, the defect types of the LEDs are two: defect 1-mounting errors, defect 2- soldering errors. Using the method-of-moments estimation (MME), all the parameters of a BZIP distribution can be estimated. Afterwards, the shifts to be detected are specified by the writers and the UCLs are selected. The results of the application show that the model, which is in interest in the paper, can detect shifts in the parameters and the out-of-control conditions can be noticed on time. As an alternative, the authors do not consider the correlation between the two defects and apply a model with two independent ZIP distributions. The conclusion is that the method which is scrutinized in this paper is more sensitive than the one assuming independence.

A bivariate control chart based on copula function

The same year, one more article that is referred to a bivariate case of a ZIP distribution is the one of Fatahi et al. (2012b). This paper is concerned with a copula approach in order to derive the joint distribution of ZIP marginal distributions. Until the appearance of this study, there was no such reference regarding the copula functions in conjunction with multivariate SPC charting.

So, in order to proceed forward we have to answer the question :

"What is a copula function?"

In simple words, a copula function "couples" (or joins) marginal distributions into a joint one. That is really helpful in building multivariate distributions given the marginals. More formally, a copula function can be explained from the following theorem:

Theorem 1 (Sklar's Theorem) *Let $F \in \mathcal{F}(F_1, ..., F_n)$ be an n-dimentional distribution function with marginals distributions $F_1, ..., F_n$. Then, there exists a copula function $C \in \mathcal{F}(U_1, ..., U_n)$ with uniform marginals such that $F(x_1, ..., x_n) = C(F_1(x_1), ..., F_n(x_n))$, where $U_i = F_i(x_i)$, $i = 1, ..., n$ are the probability integral transformations of the marginal models.*

- This methodology is applicable regardless of the type and degree of dependence among the variables.
- The usefulness of copula function is that we can simplify analysis of dependence in a return distribution $F(x_1, ..., x_n)$ by studying instead a copula C.

More specifically, this paper focuses on a copula function to model the correlation between two ZIP distributions. This procedure results in a bivariate distribution. Based on existing copula functions, we can derive a bivariate mixture copula function as follows:

$$C(u,v) = \begin{cases} (1-\varrho)uv + \varrho \min(u,v) & ,\varrho > 0 \\ (1+\varrho)uv + \varrho(u-1+v)\Theta(u-1+v) & ,\varrho \leq 0, \end{cases} \tag{3.1}$$

where

$$\Theta(a) = \begin{cases} 1 & a \geq 0 \\ 0 & a < 0 \end{cases}$$

and U, V are two uniform random variables, which are used to denote two ZIP cdfs named by $F(x_1)$ and $F(x_2)$ with parameters p_1, λ_1 and p_2, λ_2 respectively, in order to obtain a bivariate copula function.

Substituting the above cdfs and the correlation coefficient ϱ in equation 3.1 , a copula function which best fits the related joint distribution function can be defined.

Considering that the copula function is known now, a BZIP control chart can be derived. To do so, the variables X_1 and X_2 denote the number of defects of two quality characteristics named by number1 and number2 respectively. Suppose that the data $X = (X_1, X_2)$ follow a joint BZIP distribution. So, $X_1 \sim ZIP(p_1, \lambda_1)$, $X_2 \sim ZIP(p_2, \lambda_2)$ and ϱ is the correlation coefficient between them. Also, $D = X_1 + X_2$. Hence, the task is to determine the upper probability limit of BZIP distribution (UPL_{BZIP}) - LPL_{BZIP} is set to be zero (Fatahi et al., 2012b). Setting the false alarm rate equal to α, the UPL_{BZIP} can be derived by the following:

$$P(D \geq UPL_{BZIP}) \leq \alpha.$$

In that point, the writers had to modify the above equation in order to be related to the copula function 3.1 . Hence,

$$C(u, v | \max(x_1 + x_2 - 1) = UPL_{BZIP}) \leq 1 - \alpha^{1}$$

So as to examine the efficiency of the developed control chart, the writers have provided the probabilities of type I and type II errors in case of the univariate ZIP distributions and the BZIP distribution related to the copula function.

In order to evaluate if a BZIP control chart is better than the simultaneous use of two separated univariate ZIP control charts, a simulation process is applied and the ARLs of each method are computed and compared by the authors. Hence, considering a BZIP distribution with typical parameter values $p_1 = 0.1$, $p_2 = 0.15$ and correlation coefficient ϱ to be 0, 0.2 , 0.5, a table with UPL and ARL_0 values are given for the two methods which are examined under different combinations of ϱ and ${\mu_{0j}}^{2}, j = 1, 2$ with 50000 simulations for each combination.

[1] Frechet (1951) showed that there exist upper and lower bounds for a copula function: $\max(0, u+v-1) \leq C(u,v) \leq \min(u,v)$.

[2] The first subscript denotes the in control condition and the second one indicates the variable number.

Using the same nominal $\alpha = 0.0027$ for both charts the conclusions are the following:

- When $p = 0$, the two methods have the same in control ARL.
- In case that p is not equal to zero, the in control ARL of BZIP chart is larger than the other one. Symbolically, $ARL_{0,BZIP} > ARL_{0,two\ ZIP}$. That is desirable because of the fewer false alarms.
- The bigger the ϱ , the bigger the difference between $ARL_{0,BZIP}$ and $ARL_{0,two\ ZIP}$. So, the BZIP chart is more sensitive as far as changes during the procedure are concerned.

The authors, also, examine and the out of control ARL assuming the same shifts in the parameters of the distributions (from μ_{0j} to any other undesirable value μ_{1j}[3], $j = 1, 2$) . The conclusions are below:

- The out of control ARLs concerning the BZIP chart are smaller than the simultaneous use of two separate univariate ZIP. That indicates a better performance for the BZIP chart.
- Larger values for $\mu_{0j}, j = 1, 2$, p_1 and p_2 would more strongly show preference of the BZIP chart.

The above remarks about the out of control ARL have been illustrated by some typical ARL_1 curves. All of the graphs showed better performance for the BZIP chart (Fatahi et al., 2012b) .

In the end, a motivating case study takes place in which two quality characteristics are investigated. This research shows that three out of 300 samples are presented out of control regarding to the two separate univariate ZIP charts, while BZIP chart shows none. Hence, the out of control plotted samples according to the first method are false alarms (Fatahi et al., 2012b) .

In 2013 Niknam and Thomas (2013) investigate an application of Acoustic Emission data for unbalance analysis and detection in rotary systems. To begin with, Acoustic Emission (AE) referred to as the practical non-destructive technology for investigating the behavior of materials under stress using the transient elastic waves (Niknam and Thomas, 2013). Also, AE hit is defined as the process of detecting and measuring an AE signal on a channel. The fundamental features of the AE hit include amplitude, duration, count, and rise time (Niknam and Thomas, 2013). Moreover, AE count is defined as the number of times the signal crosses the threshold in an AE hit (Niknam and Thomas, 2013).

In addition, a brief review about the use of AE count in fault detection is provided in this paper. The parameter of interest in this study is PAC-energy which covers the significance of count, duration and peak amplitude in AE signals (Niknam and Thomas, 2013). PAC-energy is a function of AE counts and it will be utilized for statistical modeling.

Next, the authors present a categorical data analysis based on generalized linear models. Instead of a Poisson regression model they use a ZIP model which is able to handle the large number of zeros due to the fact that AE signals depend on the pre-specified reference threshold value. So, the regression models for the parameters

$$\lambda = [\lambda_1, \lambda_2, ..., \lambda_n]' \quad \text{and} \quad p = [p_1, p_2, ..., p_n]'$$

are the following:

$$\ln(\lambda) = \mathbf{B} \cdot \beta \quad \text{and} \quad logit(p) = \ln(\frac{p}{1-p}) = \mathbf{G} \cdot \gamma,$$

[3]The first subscript denotes the out of control condition and the second one indicates the variable number.

where **B**, **G** are covariate matrices (design matrices) and β, γ are regression coefficients (for more details see Niknam and Thomas (2013)).

Afterwards, the following three significant variables are considered:

1. Bearing, which is consisted of faulty, used and new ones.

2. Speed, which contains three levels of speed.

3. Unbalance, includes five types which can cause unbalance.

The goal of the categorical analysis is to show the effects of the three variables on the PAC-energy variable. Hence, one of the major results of this study is that "speed" has the most important effect among the other variables on PAC-energy. Conversely, "bearing" has the least important effect. However, faulty bearings are the most significant in producing more PAC-energy. It is, also, interesting that the effect of new bearings is less significant than used bearings. Lastly, the results indicate that couple unbalance (type 5) is the most important type of unbalance in comparison with type 3 and 4.

Finally, the authors combine the ZIP regression model and the CUSUM chart in order to make available an operating system to identify the faults in bearings. According to the paper, the two sided tabular CUSUM chart is defined as:

$$C_i^+ = \max(0, C_{i-1}^+ + (x_i - \hat{\mu_0}) - k\hat{\sigma_x})$$
$$C_i^- = \max(0, C_{i-1}^- + (x_i - \hat{\mu_0}) - k\hat{\sigma_x}),\ i = 1, 2, ...,$$

where $C_0^+ = C_0^- = 0$, $x_i = \lambda_i$ (or p_i) and k is the threshold for cumulation. This parameter is the minimum difference between the target and sample average. $H = h\sigma_x = \tan(\theta) \cdot d \cdot \sigma_x$ is the decision interval and h is called the decision parameter. If C_i^+ or C_i^- exceeds H the process is out-of-control (Niknam and Thomas, 2013; Montgomery, 2009). Hence, in the first place, the two-sided CUSUM charts for λ and p are used only for the new bearings. Since, none of the plotted points overcome the arms of the V-mask, the process is considered in-control. Furthermore, control charts of λ and p for the used bearings are taken into consideration and in this case some of the points are plotted out of the arms of the V-mask. In the end, CUSUM charts of λ and p for all types of bearings are considered. Clear indication of the out-of-control condition (i.e detection of the faults) is shown especially as far as λ- CUSUM chart is concerned.

The head author of the above two papers, continued the idea of the combination of two CUSUM charts to monitor ZIP processes. This time He et al. (2014) propose to view the ZIP process as a combination of a process with random shocks and a Zero-Truncated Poisson (ZTP) process. Hence, they denote the charts which are based on this assumption as CRL-CUSUM and ZTP-CUSUM accordingly. For their combination they use the notation CRL-ZTP CUSUM. So, in order to proceed forward, the ZTP distribution must be defined.

The ZTP model appears when there are no zeros in a standard Poisson distribution. The density function of this model is given by:

$$P(x; \lambda) = \frac{e^{-\lambda}\lambda^x}{x!(1 - e^{-\lambda})}, x = 1, 2, ...,$$

where x represents the number of defectives products, λ is the parameter of a standard Poisson distribution and the factor $(1 - e^{-\lambda})$ is the difference between a standard Poisson and a ZTP

distribution. Also, the mean and the variance of the above distribution are $E(X) = \frac{\lambda}{1-e^{-\lambda}}$ and $Var(X) = (EX)(1+\lambda) - (EX)^2$ respectively.

Considering now a sample of independent $X_1, X_2, ..., X_n$ counts coming from a ZTP process and using the MLE method along with numerical calculations, the parameter λ can be estimated as $\hat{\lambda} = \bar{x}(1 - e^{-\hat{\lambda}})$, where $\bar{x}$ is the mean of the $X_i's$.

Thus, given that the in-control values of the parameters of the ZIP model is p_0 and λ_0 and the out-of-control values, which have to be detected, are denoted as p_1 and λ_1, the design of the control charts that are examined in this paper is presented below:

1. **CRL-CUSUM chart**
   This chart is determined based on the number of observations between two consecutive non-conforming products, included the non conforming observation at the end. For this reason the chart is called CRL (Conforming Run Length). Now, assuming independence between products, the CRL follows a geometric distribution. Supposing that the run length CRL_i has just been observed, the CRL-CUSUM statistic C_i for detecting an increase in fraction nonconforming is given by:

$$C_i = \max(k - CRL_i + C_{i-1}, 0),\ i = 1, 2, ...$$

   where $C_0 = 0$ and referred value k is calculated as follows:

$$k = \frac{ln[(1-p_0)/(1-p_1)]}{ln(p_1/p_0)} + 1$$

   The above chart signals when $C_i > h_c$, where h_c is the selected UCL which achieves the required in-control performance (D. Bourke, 1991; He et al., 2014).

2. **ZTP-CUSUM chart**
   This chart is developed in order to detect shifts in the positive direction of λ, which is the parameter of the ZTP model. Only the number of non-conformities are considered here and the zero values are ignored. Hence, based on the likelihood ratio method, the ZTP-CUSUM control statistics are constructed as follows:

$$Z_i = \max(Z_{i-1} + N_i, 0),\ i = 1, 2, ...$$

   where $Z_0 = 0$ and N_i is based on the log-likelihood ratio and is defined as:

$$N_i = N(X_i) = X_i ln\frac{\lambda_1}{\lambda_0} + (\lambda_0 - \lambda_1) + ln\frac{1-e^{-\lambda_0}}{1-e^{-\lambda_1}}$$

   where $Z_0 = 0$ and Z_i is updated whenever a non-conformity is detected. This chart signals when $Z_i > h_z$, where h_z is the selected UCL which achieves the desired in-control performance. In addition, the combination of the above two CUSUM charts creates the **CRL-ZTP CUSUM chart**, which gives an out-of-control signal whenever one of the above individual CUSUM charts signal (He et al., 2014).

Furthermore, in accordance with the previous article, the ANOS value is used to measure the detecting effectiveness of the under discussion CUSUM charts. So, in order to evaluate ANOS for the CRL and ZTP CUSUM charts, the term Terminal Nonconformity Sequence (TNS) is utilized, which

denote a run of (at least zero) good items followed by a nonconformity and therefore

$$ANOS = \frac{1}{r}E\,(\text{Number of TNS})$$

where r is the probability of shock occurring in the process, i.e $r = (1-p)(1-e^{-\lambda})$ in a ZIP process. So, to design the charts, the writers need to determine the in-control parameters p_0, λ_0, the size of the shift they want to detect and the control limits in a way that they achieve a pre-determined performance, say, $ANOS_0$.

Next, the authors proceed to a simulation study. They consider the exact same in-control values and the same pre-specified shift sizes of the parameters as the previous mentioned article. Furthermore, the control limits of the CRL and ZTP CUSUM charts are chosen to make the ANOS value of CRL-ZTP CUSUM chart close to 200 (when $p_0 = 0.2$) such as in the t-CUSUM chart case. Hence, they conclude to the following comparisons between CRL-ZTP , p-λ and t-CUSUM charts:

- When the parameter p shifts while λ continues to be in-control, the CRL-ZTP CUSUM chart is better than p-λ and t-CUSUM charts.
- When the parameter λ shifts while p remains in-control, the CRL-ZTP CUSUM chart and the p-λ CUSUM chart performs almost equal and the t-CUSUM chart is lightly better than the other two.
- When both p and λ change at the same time it can be seen that t-CUSUM chart has better performance than CRL-ZTP and p-λ CUSUM chart, which are nearly the same.

Lastly, for comparison purposes the authors use again the data which have been collected from the light emitting diode (LED) packaging industry (He et al., 2011). In this case, the CRL and ZTP data are obtained from the selected 200 samples in order to be built the CRL-CUSUM chart and the ZTP-CUSUM chart respectively. The control limits of the CRL-ZTP CUSUM chart can be obtained from $ANOS_0 \simeq 200$. The results are the same as those reported to the previous paper, i.e the CRL and ZTP CUSUM charts signal at the same point with the p-CUSUM and λ-CUSUM charts, accordingly.

Moving one year closer to the present day, Sukparungsee (2018) tries to propose a closed form expression of the ARL of a CUSUM control chart based on a Markov chain approach (MCA) for ZIP processes. Additionally, the performance of the CUSUM chart in comparison with the ZIP-EWMA chart (Fatahi et al., 2012a) for small to large shifts is examined and lastly the effect of the proportion of the number of zero-defects is analyzed.

First of all, supposing that Y_t is a sequence of independent identically distributed ZIP random variables, the statistic C_t of the upper sided CUSUM control chart which is designed to detect upward shifts is:

$$C_t = \max(Y_t - k + C_{t-1}, 0),\ t \in 1, 2, ...,,$$

where k is a reference value and $C_0 = 0$. If $C_t > h$, where h is the selected UCL which achieves a required in-control ARL value, an out-of-control signal occurs.

So, in order to find this ARL value a MCA (A. Brook and A. Evans, 1972) for the ZIP-CUSUM chart will be used. In short, the closed form of the ARL is presented above:
Let Y be an integer random variable. If $C_t = i$, then, it is said that the decision interval scheme is in state S_i. Each realization of the scheme can then be regarded as a random walk over the states $S_0, S_1, ..., S_h$ where S_h is an absorbing state which represents an out-of-control area above the control

limit h. Also, we assume that the process is initially in state S_0. The probability of moving from state i to j in one step is:

$$\begin{aligned} P_{ij} &= P\left(C_t \in S_j | C_{t-1} \in S_i\right) \\ &= P\left(Y - k + C_{t-1} = j | C_{t-1} = i\right) \\ &= P\left(Y = k + j - i\right), \; i \neq h, \; 0 < j < h \end{aligned}$$

and $P_{i0} = P\left(Y \leq k - i\right)$, $P_{ih} = P\left(Y \geq k + h - i\right)$. If the statistics C_t move to an absorbing state, it is impossible to leave that state, thus $P_{hh} = 1$. This last one mentioned is not needed for the calculation of the ARL. So, the transition matrix can be written as:

$$\mathbf{P} = \left[\begin{array}{ccc|c} P_{11} & \cdots & P_{1n} & P_{1,n+1} \\ \vdots & \ddots & \vdots & \vdots \\ P_{n1} & \cdots & P_{nn} & P_{n,n+1} \\ \hline 0 & \cdots & 0 & 1 \end{array}\right] = \begin{bmatrix} \mathbf{R} & \mathbf{(I\text{-}R)1} \\ \mathbf{0} & \mathbf{1} \end{bmatrix}$$

where $R_{n\times n}$ contains the transition probabilities of going from one transient state to another state, $I_{n\times n}$ is the identity matrix, $1_{n\times 1}$ is the column vector of ones and $0_{1\times n}$ is the row vector of zeros. According to Lucas and Saccucci (1990) the ARL based on the in control states can be obtained by the following form:

$$ARL\left(t\right) = \sum_{i=1}^{\infty} iP\left(RL = i\right) = \mathbf{p^T}(\mathbf{I} - \mathbf{R})^{-\mathbf{1}}\mathbf{1},$$

where $P\left(RL = i\right) = \mathbf{p^T}(\mathbf{R^{i-1}} - \mathbf{R^i})\mathbf{1}$ and $\mathbf{p^T} = (\mathbf{0}, ..., \mathbf{0}, \mathbf{1}, \mathbf{0}, ..., \mathbf{0})^{\mathbf{T}}$ represents the probability vector of an initial state, with 1 related to a specified state and zeros elsewhere.

Next, the author use the forenamed proposed closed-form of the ARL to obtain some numerical results to evaluate the detecting ability of the ZIP-CUSUM control chart. The results that are presented concern two cases: The in control parameter of the ZIP distribution is $\lambda_0 = 1$ *or* 5 and the desired $ARL \simeq 370$ *or* 500 due to $1 - p = 0.1, 0.5, 0.9$ which are defined as low, medium and high probability of zeros respectively. The shifts in λ parameter that are examined here are $\delta = 0.01, 0.05, 0.10, 0.25, 0.5, 1.00$ where the out of control $\lambda_1 = (1 + \delta)\lambda_0$. Additionally, the performance of the ZIP-CUSUM chart is compared with the ZIP-EWMA chart. The ARL values of the latter chart are computed via the MCA method (Fatahi et al., 2012a). The conclusions of this paper are the following:

- When $\lambda_0 = 1$ *or* 5 and the $ARL_0 \simeq 370$, there is no meaningful difference in the performance of the ZIP-EWMA and the ZIP-CUSUM control chart when $\delta \leq 0.01$ and $1 - p = 0.9$.The same thing happens and when $ARL_0 \simeq 500$.

- For each λ_0 and ARL_0, when $\delta > 0.01$ the ZIP- EWMA is always superior to ZIP-CUSUM control chart, with the minimum ARL_1.

- The performance of both control charts depends on the probability of extra zeros. In this case, the detecting ability of the control charts decreases as the probability of extra zeros constantly increasing.

One more article (Rakitzis et al., 2016) which refers to CUSUM charts is published in 2016. The uniqueness of this article is that refers to correlated counts. This phenomenon (of autocorrelated observations) is very common especially in automated sampling. For that reason, control charts based on integer-valued time series models must be used. In this article, the ZIPINAR(1) and

ZIPINARCH(1) models, which are extensions of the PINAR(1) and PINARCH(1) respectively, are applied in order to fit correlated Poisson counts with vast amount of zeros. Also, the performance of Shewhart-type and CUSUM-type charts are studied.

More thoroughly, the basic knowledge of the ZIPINAR(1) and ZIPINARCH(1) models is going to be presented below:

- The ZIPINAR(1) model
  A process of $N_t's$ that is being observed at discrete equally spaced times $t = 1, 2, ...$ is called an INAR(1) process if it is of the form $N_t = \alpha \circ N_{t-1} + \epsilon_t$, $t \geq 1$
  where $\alpha \in [0, 1)$, '$\circ$' is the binomial thinning operator (Steutel and van Harn, 1979) and the random variable $\alpha \circ N_{t-1} = \sum_{i=1}^{N_{t-1}} Y_i$ arises from N_{t-1} by binomial thinning, where the Y_i's are i.i.d binary indicators(which are also independent of N_{t-1}) with $P(Y_i = 1) = \alpha$ and $P(Y_i = 0) = 1-\alpha$. The, so called, innovations ϵ_t, $t \geq 1$ are supposed to be i.i.d random variables and the thinning operation at each time t, as well as the ϵ_t, are independent of $N_{t-1}, N_{t-2}, ...$ So, the INAR(1) process is a homogeneous Markov chain by definition. Here, $\epsilon_t \sim ZIP(\phi, \lambda)$ and the model is named by ZIPINAR(1) (Rakitzis et al., 2016).

- The ZIPINARCH(1) model
  Let, now, N_t, $t \geq 1$, be a time series of correlated counts. For this model it is assumed that conditional on $N_{t-1}, N_{t-2}, ...$, the conditional distribution of N_t is specified by a ZIP distribution with parameters $\omega \in [0, 1)$ (it is used different notation for the zero inflation parameter) and $\lambda_t = b + \alpha N_{t-1}$, $t \geq 1$, where $b > 0$ and $a \in [0, 1)$ (Rakitzis et al., 2016).

Both the above models are designed to describe autocorrelated process of counts, but they have significant differences especially for high levels of autocorrelation. The basic discrimination is that the models have their own unique mechanism to produce zeros. Hence, the pattern of zeros observed in the available time series data, indicates which model is more suitable to be applied (Rakitzis et al., 2016).

There are a lot of ways to make a process go out-of-control. For instance, the expected value N_t or ϵ_t shift from a given in-control value, or shifts in the correlation between N_t, N_{t-1} may occur. For this reason, control charts for monitoring ZIPINAR(1) and ZIPINARCH(1) processes are developed. Practically, it is of great importance to detect an increase in the mean of the process. That happens if one or more parameters of the process change appropriately.

In the paper of Rakitzis et al. (2016), an upper-sided Shewhart type chart and a CUSUM type chart are going to be considered in order to detect upwards shifts in the mean, namely μ_0. For the first chart, an out-of-control signal is detected when $N_T \geq UCL$ for the first time. For the second chart an out-of-control signal is given when $C_t \geq h$.
$C_t = \max(0, C_{t-1} + N_t - k)$, $t \geq 1$ is an upper-sided CUSUM chart where $C_0 = c_0$ and the reference value $k \geq \mu_0$.

In addition, an ARL value is needed so as to evaluate the performance of the two control charts. For the in control ARL the authors are utilized the zero-state ARL (zsARL or $ARL^{(1)}$) and for the out-of-control ARL the steady-state ARL (ssARL or $ARL^{(\infty)}$). Hence, $zsARL = E(L|\tau = 1)$ and $ssARL = \lim_{\tau\to\infty} ARL^{(\tau)} = \lim_{\tau\to\infty} E(L - \tau + 1|L \geq \tau)$, $\tau \geq 1$ where $\tau \in 1, 2, ...$ is the fixed (but unknown) change point, where for $t < \tau$ the process is in control and fot $t \geq \tau$ the process is out-of-control. L is the random variable that denotes the number of point plotted on the chart, until an out-of-control signal is given for the first time.

Knowing the Run Length distribution for each of the schemes, the next move is the statistical design of them, i.e determination of the design parameters in order to have the desired $zsARL$. For

the Shewhart type control chart the UCL must be defined and for the CUSUM type control schemes, (h, k) must be determined.

After that, a numerical study takes place in order to establish the performance of the two proposed charts when observations come from a ZIPINAR(1) or a ZIPINARCH(1) model. The study is about to detect shifts in the mean of a process. The in control parameters of the ZIPINAR(1) model are μ_0, α_0, ϕ_0 and for the ZIPINARCH(1) are $\mu_0, \alpha_0, \omega_0$. The authors consider $\mu_0 \in \{1, 2, 3, 5\}$ for both models and $\{\alpha_0, \phi_0\} \in \{(0.3, 0.2), (0.3, 0.5), (0.3, 0.8), (0.5, 0.5), (0.5, 0.8)\}$ and in order to have the same first order autocorrelation, it is assumed that $\gamma_0 = \alpha_0(1 - \omega_0) \in \{0.3, 0.5\}$. The value of ω_0 is defined so as to have the same or almost the same percentage of zero counts with the ZIPINAR(1) model. The shifts which are examined are $\mu = \mu_0 + \delta$, where $\delta \in \{0.1, 0.2, 0.3, 0.4, 0.5, 1.0, 1.5, 2.0\}$. Also, the reference value k considered to be close to the in control μ_0. The results of this study are mentioned briefly above:

- For the ZIPINAR(1) process,
  when $\alpha_0 = 0.3$ and $\phi_0 = 0.2$ or 0.5 the CUSUM chart has better performance for small or moderate shifts. For $\phi_0 = 0.8$ and $\alpha_0 = 0.5$ the Shewhart type chart is better than the CUSUM scheme. Furthermore, when $\phi_0 = 0.5$ the CUSUM charts achieves the lowest ssARL for shifts up to 1 or 1.5. The best performance of the Shewhart type chart is happened when $\alpha_0 = 0.8$.

- For the ZIPINARCH(1) process,
  when $\gamma_0 = 0.3$ and $\omega_0 < 0.5$, the CUSUM control chart has better performance than the Shewhart type chart. But, for $\omega > 0.5$ and $\mu_0 = 3$ or 5 the Shewhart type chart has the best detection ability no matter how the shift size is. Also, the Shewhart type chart has the worst performance for every shift in the mean process when $\gamma_0 = 0.5$ and ω_0 up to 0.1. For $\omega_0 > 0.3$ and as μ_0 increases, CUSUM and Shewhart charts are almost similar.

- Moreover, the effect of the zero inflation in the data is investigated using the ARL values and by compare them between the control schemes under the different models. Analytically, comparisons between the PINAR(1), ZIPINAR(1) and PINARCH(1), ZIPINARCH(1) Shwehart and CUSUM charts in the detection of shifts in process mean μ_0 take place. In order to keep the zsARL to the desired level, adjustments for the values of the design parameters of each scheme are needed to be made. So, the results show that as the percentage of zeros increases, the performance of the ZIPINAR(1) Shewhart chart improves, while the performance of the ZIPINAR(1) CUSUM chart becomes poorer. For the case of ZIPINARCH(1) process, the presence of zero-inflation affects the CUSUM and Shewhart charts way over in comparison with their ZIPINAR(1) counterparts.

In closing, the authors include a real data example to illustrate the practical usefulness and applicability of the above mentioned charts. The analysis of the real data shows that both those models are helpful in practice so as someone to describe correlated count data with a vast number of zero counts.

The last article that deals with ZIP distribution and CUSUM chart is the one of Huh et al. (2017). The writers study the case in which someone wants to detect a parameter shift in Poisson or ZIP integer-valued GARCH (INGARCH) models. The INGARCH models use a linear regression of the conditional means $\lambda_t = E[X_t|X_{t-1}, ...]$. Now, the count at time t is generated by using a conditional Poisson distribution - $X_t \sim P(\lambda_t)$ - values. An INGARCH(1,1) model is defined by including a feedback term with respect to the previous conditional mean, i.e $\lambda_t = \omega + \alpha X_{t-1} + \beta\lambda_{t-1}$.

The conditional log-likelihood and the score function of a Poisson INGARCH(1,1) model are given as follows:

$$\log\left(\tilde{L_n}\left(\theta\right)\right) = \sum_{t=1}^{n} \tilde{\ell_t}\left(\theta\right) = \sum_{t=1}^{n} [X_t \log \tilde{\lambda_t} - \tilde{\lambda_t} - \log(X_t!)]$$

and

$$\frac{\partial \log\left(\tilde{L_n}\left(\theta\right)\right)}{\partial\theta} = \sum_{t=1}^{n} \frac{\partial \tilde{\ell_t}\left(\theta\right)}{\partial\theta} = \sum_{t=1}^{n} (\frac{X_t}{\tilde{\lambda_t}} - 1)\frac{\partial \tilde{\lambda_t}}{\partial\theta},$$

where $\theta = (\omega, \alpha, \beta)^T$ and $\tilde{\lambda_t} = \omega + \alpha X_{t-1} + \beta\tilde{\lambda}_{t-1}$. The solution of $\frac{\partial \log\left(\tilde{L_n}(\theta)\right)}{\partial\theta} = 0$ is the conditional MLE of the true parameter θ_0 denoted by $\hat{\theta_n}$.

In 2015 Lee et al. (2015) extended the above to ZIP INGARCH(1,1) models, in order to deal with a time series of count with excessive zeros. Likewise, $(1-p)\lambda_t^* = E[Y_t|Y_{t-1}, ...]$, where $Y_t \sim ZIP\left(\lambda_t^*, p\right)$ and the ZIP INGARCH(1,1) model is defined as:

$$(1-p)\lambda_t^* = \lambda_t\left(\theta^*\right) = \omega^* + \alpha^* Y_t + \beta^* \lambda_{t-1}\left(\theta^*\right),$$

where $\theta = (p, \theta^{*T})$ and $\theta^* = (\omega^*, \alpha^*, \beta^*)^T$.
The conditional log-likelihood function of a ZIP INGARCH(1,1) model is the following:

$$\log\left(\tilde{L_n}\left(\theta\right)\right) = \sum_{t=1}^{n} \tilde{\ell_t}\left(\theta\right) = \sum_{t=1}^{n} [\tilde{\ell_{t0}}\left(\theta\right) I(Y_t = 0) + \tilde{\ell_{t1}}\left(\theta\right) I(Y_t \geq 1)]$$

The conditional MLE of θ_0 is defined as $\hat{\theta_n} = arg\max_{\theta\in\Theta}$ (for details and strictly theory see Huh et al. (2017)).

Once the INGARCH(1,1) models were introduced, the CUSUM charts which are going to be used in this study are described by the authors. To begin with, the first chart is the Conditional Likelihood Ratio (CRL) CUSUM chart. This control chart assumes that the out-of-control parameter, θ_1, is known. For $t = 1, 2, ...$ let $X_1, X_2, ..., X_{t-1}$ be an i.i.d sample with common density f_{θ_0} and $X_t, X_{t+1}, ...$ be an i.i.d random sample with density f_{θ_1}. Based on $\tilde{Z_t} = \log \frac{f_{\theta_1}\left(X_t|X_{t-1}, \tilde{\lambda}_{t-1}\right)}{f_{\theta_0}\left(X_t|X_{t-1}, \tilde{\lambda}_{t-1}\right)}$, the stopping rule can be defined as:

$$N_C = inf\left\{n \geq 1 : \max_{1\leq k\leq n} \sum_{t=k}^{n} \tilde{Z_t} \geq c_\gamma\right\}$$

or it can be expressed as $N_C = inf\left\{n \geq 1 : \ell_n \geq c_\gamma\right\}$. So, the CRL CUSUM statistics are given by: $\ell_n = \max(0, \ell_{n-1} + \tilde{Z_n})$, $n \geq 1$ and $\ell_0 = 0$.

Furthermore, the generalized likelihood ratio (GRL)-based CUSUM chart is referred. The only difference with the above mentioned chart is that the out-of-control parameter are unknown and must be estimated from the data. The stopping rule for this case is the following:

$$N_G = inf\left\{n \geq 1 : \max_{1\leq k\leq n} \sum_{t=k}^{n} \log \frac{f_{\hat{\theta_{kn}}}\left(X_t|X_{t-1}, ...\right)}{f_{\theta_0}\left(X_t|X_{t-1}, ...\right)} \geq c_\gamma\right\}$$

where $f_{\hat{\theta_{kn}}}$ is the conditional MLE based on $X_k, ..., X_n$. The GRL chart has the drawback that depends excessively on the parameter estimation which make it not so efficient. Also, the computational aspect of this is complicated.

The authors use the above two charts with the Poisson INGARCH(1,1) models. They evaluate their performance through simulations and the results show that the GRL-based CUSUM chart does

not perform so well in comparison with the CRL-based CUSUM chart. However, the authors notice after simulations that CRL chart had high FAR, something that is not desired. For that reason , the authors consider the monitoring process proposed by Gombay and Serban (2009).

So, the proposal is to be used a score function-based CUSUM (SF-CUSUM) chart with AR models. Hence, Huh et al. (2017) extended this idea and applied it to ZIP INGARCH(1,1) models. Firstly, a hypothesis of no change for the parameters is considered, such as

$$H_0 : \text{the known } p_0 \text{ remains the same till time n.}$$

Hence, under the null hypothesis, as $n \to \infty$, based on the test statistics

$$T_n(p) := \sup_{v_n \le k \le n} \frac{1}{\sqrt{n}} |W_k\left(\rho_0, \hat{\theta_n^*}\right) - \frac{k}{n} W_n\left(\rho_0, \hat{\theta_n^*}\right)| \xrightarrow{d} \sup_{0 \le s \le 1} |\mathcal{B}(s)|$$

where v_n is a sequence of positive integers, $\mathcal{B}$ is a Brownian bridge, $W_k\left(\rho_0, \hat{\theta_n^*}\right) = I_{n,1,1}^{-1/2}\left(\rho_0, \hat{\theta_n^*}\right) \sum_{t=1}^{k} \frac{\partial \tilde{\ell}_t\left(\rho_0, \hat{\theta_n^*}\right)}{\partial \rho}$, where $I_{n,1,1}^{-1}\left(\rho_0, \hat{\theta_n^*}\right)$ is the inverse of the $(1,1)^{th}$ entry of the empirical information matrix (for details see (Huh et al., 2017, section 4)).

Now, in order to construct the stopping rule, the form of $I_{1,1}^{-1/2}(\theta_0)$ rather than $I_{n,1,1}^{-1/2}\left(\rho_0, \hat{\theta_n^*}\right)$ is utilized. Hence, the following stands:

$$N_{Cp,m} = inf\{n \ge 1 : T_n^*(p) \ge c_{\gamma,m}\}$$

where $T_n^*(p)$ is the same as $T_{n+m}(p)$ obtained on the basis of $X_{-m+1}, ..., X_n$ and $c_{\gamma,m}$ is defined through Monte Carlo simulation under in control situations (Huh et al., 2017).

The importance of m is that lower the FAR in order to steady the SF-CUSUM chart Also, the simulations which conducted by the authors show that larger m achieves lower FAR. In addition, similar stopping rules can be used for the remain parameters ω, α, β (for details see Huh et al. (2017)).

Lastly, the authors compare through the ARL and FAR the CRL and SF-CUSUM charts used with the ZIP INGARCH(1,1) models via simulation study. They consider 15 cases of the parameter shift of $p_1, \omega_1, \alpha_1, \beta_1$ in ZIP INGARCH(1,1) models. The in control parameter is $\theta_0 = (0.3, 2, 0.2, 0.2)$ and $m = 100$ for the SF-CUSUM chart. The results of the simulation study are listed below:

- SF-CUSUM chart detects a shift slightly slower than does the CRL CUSUM chart.
- SF-CUSUM chart achieves better results than the CRL, because of the fewer false alarms.
- It is important to be said that as the shift occurs at a later time, a larger m is needed to get a low FAR.
- Overall, the SF-CUSUM chart is more desirable than the CRL, especially when rare shifts occur.

At the end, the authors propose some ideas for future research such as to extend the INGARCH(1,1) to more general models, i.e INGARCH(p,q) or they prompt to be used totally other types of models.

3.2 EWMA and MA control charts

In recent years, the issue of rare health events has motivated many researchers in the area of control charting. An interesting approach was presented in 2012 by Fatahi et al. (2012a). They presented an exponentially weighted moving average (EWMA) control chart based on ZIP distribution, named by ZIP-EWMA, in order to monitor the rare health events in healthcare.

So, in order to proceed forward the EWMA for the ZIP model must be defined. Hence, assuming that $Y_1, Y_2, ..., Y_t, ...$ is a sequence of independent and identicaly distributed ZIP random variables and $\kappa \in [0, 1]$ is a constant, a sequence $Z_1, Z_2, ..., Z_t, ...$ can be derived from the exact above sequence as:

$$Z_t = (1 - \kappa)Z_{t-1} + \kappa Y_t.$$

Let $Z_0 = E(Y) = \lambda(1 - p)$ be the starting value and decomposing Z_{t-1} in terms of Z_{t-2}, and Z_{t-2} in terms of Z_{t-3} and so on, the conclusion is that

$$Z_t = (1 - \kappa)^t Z_0 + \kappa \sum_{i=0}^{t-1} (1 - \kappa)^i Y_{t-i}$$

Obviously, the above equation shows that Z_t is a linear combination of the random variable Z_0 weighted by a coefficient $(1-\kappa)^t$ and the random variables $Y_1, Y_2, ..., Y_t, ...$ weighted by the coefficients $\kappa(1-\kappa)^{t-1}, ..., \kappa(1-\kappa), \kappa$ respectively. Thus, the sequence $Z_1, Z_2, ..., Z_t, ...$ is called EWMA sequence. By definition, $Y_1, Y_2, ..., Y_t, ...$ are independent random variables, while $Z_1, Z_2, ..., Z_t, ...$ are not. In addition,

$$E(Z_t) = E(Y_t) = \lambda(1 - p)$$

$$Var(Z_t) = (\frac{\kappa}{2 - \kappa})[1 - (1 - \kappa)^{2t}]Var(Y_t)$$

It can be shown that as t gets larger $[1 - (1 - \kappa)^{2t}]$ approaches 1.This means that after the EWMA control chart has been running for several time periods,the control limits will approach steady-state values given by:

$$UCL_{\text{ZIP-EWMA}} = E(Y_t) + L\sqrt{\frac{\kappa}{2 - \kappa}}\sqrt{Var(Y_t)} \quad \text{and} \quad LCL_{\text{ZIP-EWMA}} = 0$$

where factor L is the width of the control limits and it is defined in trade off with κ in accordance with a predetermined value of ARL (Montgomery, 2009; Fatahi et al., 2012a).

Afterwards, the authors proposed the Markov chain approach (MCA) (A. Brook and A. Evans, 1972) in order to calculate the ARL values since in ZIP-EWMA control chart the plotted points are not independent. In brief, according to this approach, the interval $[LCL_{\text{ZIP-EWMA}}, UCL_{\text{ZIP-EWMA}}]$ is divided into S sub-intervals of width 2δ. The statistic Z_t is said to be in transient state j at time t if $H_j - \delta < Z_t < H_j + \delta, j = 1, 2, ..., S$, where H_j denotes the midpoint of the j^{th} sub-interval. So, if $Z_t \notin [LCL_{\text{ZIP-EWMA}}, UCL_{\text{ZIP-EWMA}}]$ it is said to be in absorbing state and therefore ARL is the absorption of the MC.

The probability of moving from state i to j in one step is:

$$P_{ij} = P\left(\frac{H_j - \delta - (1 - \kappa)H_j}{\kappa} < Y_t < \frac{H_j + \delta - (1 - \kappa)H_j}{\kappa}\right) =$$

$$F_{ZIP}\left(\frac{H_j + \delta - (1 - \kappa)H_j}{\kappa}\right) - F_{ZIP}\left(\frac{H_j - \delta - (1 - \kappa)H_j}{\kappa}\right).$$

Supposing, now, that the health related process starts in state j, and $R_j, j = 1, 2, ..., S$ is the

related ARL, the vector $R = [R_1, R_2, ..., R_S]^{'}$ presents ARLs based on the start point of the process. Let Q be the matrix obtained from the transition matrix P detecting the $(S+1)^{th}$ row and column, so Q denotes the transition matrix based only on the in control states. Hence, solving the equation $(I - Q) \cdot R = \mathbf{1}$ the ZIP-EWMA ARLs would be obtained (Borror et al., 1998; Fatahi et al., 2012a).

After the above information, the authors of the paper compare the performance of the ZIP-EWMA chart with the two charts that proposed Sim and Lim (2008) using Jeffreys interval, i.e c_J chart with control rule and c_J chart with no control rule. To do so, Fatahi et al. (2012a) set values for the EWMA parameters κ and L in order to achieve a favorable in control ARL each time to compare the performance ability of the control charts. So, with almost the same ARL_0 the charts' out-of-control ARL can be compared. The results showed that the ZIP-EWMA control chart can detect quicker an out-of-control situation than the other control charts for all occasions.

In the end, the paper presents a real case study to show the applicability of the developed chart. The subject of needle-stick occurrence is examined here. So, in order to monitor and detect an increase in the counts of unintentional needle-stick occurrences per day in a hospital, the authors applied the ZIP-EWMA control chart. The data concerned daily number of needle-stick occurrences for 90 successive days in Shariati hospital. That subgroups are used for phase I and 100 other subgroups were sampled for phase II. According to the writers, the results indicate high precision for the proposed control chart.

Almost three years later of the above research in the field of disease surveillance, Leong et al. (2015) deal, again, in EWMA control charts with the difference that they monitor not only shifts in λ parameter of a ZIP process, but also in p. This problem, i.e a combination of two control charts in order to simultaneously detect shifts in both parameters of a ZIP process, has never been mentioned in any disease surveillance literature despite the fact that, in general, a few studies in the context of SPC have considered this matter before.

So, in order to fulfill this absence of combined control charting procedures, Leong et al. (2015) decided to write a paper which attends to develop a combination of EWMA control charts for concurrently monitoring each ZIP parameter individually. Hence, the development of two combined EWMA control charts take place into this study named by Bernoulli-ZIP EWMA and CRL-ZTP EWMA control charts, respectively.

To begin with, the first chart is comprised of the ZIP-EWMA control chart, which has been referred to Fatahi et al. (2012a) and constructed only to monitor shifts in λ and the Bernoulli-EWMA chart. The latter was developed to intensify the ability of the ZIP-EWMA and detect shifts in p. In short, this chart is utilized for Bernoulli data which are generated through an indicator function $I_t(y \geq 1) = \begin{cases} 1 & , y \geq 1 \\ 0 & , y = 0 \end{cases}$ and the mean and variance of this function are given by:

$$E[I_t(y \geq 1)] = (1-p)(1-e^{-\lambda})$$
$$Var[I_t(y \geq 1)] = (1-p)(1-e^{-\lambda})[1-(1-p)(1-e^{-\lambda})]$$

Hence, this chart monitors the statistic:

$$F_t = \max(0, (1-\kappa)F_{t-1} + \kappa I_t(y \geq 1)) \ , \ t = 1, 2, ...$$

where $F_0 = E[I_t(y \geq 1)] = (1-p)(1-e^{-\lambda})$. The control chart signals whenever $F_t > h_{Ber-EWMA}$,

where

$$h_{Ber-EWMA} = E[I_t(y \geq 1)] + L_{Ber-EWMA}\sqrt{\frac{\kappa}{2-\kappa}}\sqrt{Var[I_t(y \geq 1)]}$$

The factor $L_{Ber-EWMA}$ is the width of the above control limit and it is defined in trade off with κ in order to achieve a desired average false signal rate. To conclude, the Bernoulli-ZIP EWMA control chart is made to generate a signal whenever one of the forenamed individual EWMA chart signals.

Another chart to monitor upwards shifts in λ is the ZTP-EWMA control chart which has first mentioned by He et al. (2014) for CUSUM charts. Hence, this proposed chart monitors the following statistic:

$$W_t = \max(0, (1-\kappa)W_{t-1} + \kappa Z_t)\ ,\ t = 1, 2, ...$$

where Z_t is the observed number of cases at time t assumed to follow a ZTP distribution with parameter λ and $W_0 = E(Z) = \frac{\lambda}{1-e^{-\lambda}}$. It is obvious, that W_t updates only when non-zero cases are observed at time t. This control chart signals whenever $W_t > h_{ZTP-EWMA}$, where

$$h_{ZTP-EWMA} = E(Z) + L_{ZTP-EWMA}\sqrt{\frac{\kappa}{2-\kappa}}\sqrt{Var(Z)}$$

The factor $L_{ZTP-EWMA}$ is chosen in order to achieve a desired average false signal rate (Leong et al., 2015).

Afterwards, a chart which can detect shifts in p is proposed to be simultaneously used with the ZTP-EWMA control chart. This chart is the so-called CRL-EWMA control chart. According to Leong et al. (2015) this control chart is a Bernoulli-type EWMA chart based on geometric conforming run length (CRL). In this case, only the collection of positive deviations from the target are taken into account and is based on the CRL. The latter , is supposed to follow a geometric distribution with mean $\mu_{crl} = \frac{1}{p_{crl}}$ between non-zeros. Thus, the control charting procedure converts each CRL count into

$$CRL^- = \min(\mu_{crl}, CRL) = \mu_{crl} - \max(0, \mu_{crl} - CRL).$$

The mean and the variance of CRL^- according to Eleftheriou and Farmakis (2011) are as follows:

$$E(CRL^-) = \frac{1 - 2p_{crl} - (1-p_{crl})^{\frac{1+p_{crl}}{p_{crl}}}}{p_{crl}(1-p_{crl})}$$
$$Var(CRL^-) = E[(CRL^-)^2] - [E(CRL^-)]^2$$

where $E[(CRL^-)^2] = \frac{p_{crl}^2 - 4p_{crl} + 2 + (1-p_{crl})^{\frac{1}{p_{crl}}}(-p_{crl}^2 + 5p_{crl} - 4)}{p_{crl}^2(1-p_{crl})}$. Therefore, the above chart monitors the statistic:

$$G_t = \max(0, (1-\kappa)G_{t-1} + \kappa CRL_t^-)$$

where $G_0 = E(CRL^-)$ and CRL_t^- is the transformed conforming run length at time t assuming non-zero cases were observed at time t. Again, G_t only updates whenever non-zero data are observed in time t. The CRL-EWMA chart signals whenever $G_t < h_{CRL-EWMA}$, where

$$h_{CRL-EWMA} = E(CRL^-) + L_{CRL-EWMA}\sqrt{\frac{\kappa}{2-\kappa}}\sqrt{Var(CRL^-)}$$

and $L_{CRL-EWMA}$ is suitably chosen to achieve a desired average false signal rate. Once again, the combined ZTP-CRL EWMA control chart will signal whenever one of the individual ZTP or CRL EWMA chart signals.

Next, the authors describe the steps of this study. The ZIP distribution has two parameters λ

and p and the sets of 1, 4 and 0.3, 0.6 are chosen, respectively. Also, 1000 data sets, which have size of 1000 observations are simulated and a X^2- test is used to ensure that the data indeed follow a ZIP distribution. Afterwards, the first 250 observations of each data set are used to define the parameters of phase I procedure. Hence, the in-control parameters p_0 and λ_0 are estimated by $\hat{p_{mle}}$ and $\hat{\lambda_{mle}}$, accordingly. Thus, as far as the Bernoulli-EWMA chart is concerned, $E[I_t(\hat{y} \geq 1)] = (1 - \hat{p_{mle}})(1 - e^{-\hat{\lambda_{mle}}})$ and for the CRL-EWMA chart $\hat{p_{crl}} = [(1 - \hat{p_{mle}})(1 - e^{-\hat{\lambda_{mle}}})]^2$.

Lastly, for the weighting parameter κ, also, two settings are considering (0.25, 0.45), which are characterized as relatively small weight on most recent value and as relatively large weight on most recent value, respectively.

Furthermore, for each data set, the remaining 750 observations will be used for phase II. In order for the above mentioned charts to be comparable, the average time between false signals (ATFS) was set to be 90 for each compined charts. For that reason, suitable thresholds for each of the charts are defined. The estimated ATFS is given by:

$$\widehat{ATFS} = \frac{750}{\text{number of signals generated}}$$

(Leong et al., 2015).

Due to the fact that the present paper attempts to detect possible outbreaks, upwards shifts in p or λ or both of the parameters are engrafted into the remaining data in order to represent them. The first injection happens the time period 251-500, the second one the time period 501-750 and the last one at 751-1000 (for details see Leong et al. (2015)).

The last step for this simulation study is to be used some performance measures with emphasis on the capability to appropriately detect the artificial outbreaks. So, the metrics that are utilized by the authors are the following:

1. Probability of Successful Detection

$$\widehat{PSD} = \begin{cases} 1 & \text{if the chart signaled at least once during the outbreak} \\ 0 & \text{otherwise} \end{cases}$$

2. Conditional Expected Delay

$$\widehat{CED} = (\text{time of first true signal}) - (\text{time of start of outbreak})$$

3. Proportion of Outbreaks Detected

$$\widehat{POD} = \frac{\text{total number of signals generated within the outbreak}}{\text{total outbreak duration}}$$

4. Probability of True Detection

$$\widehat{PTD} = \frac{\text{total number of signals generated within the outbreak}}{\text{total number of signals generated during the phase}}$$

So, the results that arise out of the simulation study are presented in the next lines:

- Most outbreaks are detected by both EWMA charts specifically when shifts concern λ parameter for scenarios with small disease background, and when shifts concern p parameter for scenarios with large disease background.
- As far as PSDs are concerned, both the examined EWMA charts perform quietly well, with

nearly comparable detection rates. For Bernoulli-ZIP EWMA chart when p_0 is smaller, PSD is higher. The same occurs for the ZTP-CRL EWMA chart when p_0 is higher.

- Concerning CEDs, the performance of the charts in most cases fall short when only shifts in p occur for $p_0 = 0.3$, and when shifts in λ or in both p and λ happen for $p_0 = 0.6$.
- Regarding PODs and PTDs, Bernoulli-ZIP EWMA chart performs better than ZTP-CRL EWMA chart in almost all situations. Also, it is desirable for Bernoulli-ZIP to be used smaller κ, while larger κ must be desirable for ZTP-CRL EWMA chart.
- Concerning types of outbreaks (spike, triangular, ramp) all charts perform best when spike outbreaks are considered and worst when triangular ones.

Finally, for demonstration purposes, an application with real data about measles cases takes place in this paper. The data concern the dates 01/01/2010 to 14/01/2015. In this period, three noteworthy shifts were observed which indeed identified as outbreaks. An analogous preparation to the simulation study take place at phase I procedure and a Score test determines that the data follow a ZIP distribution. Also, the control limits of the proposed charts are computed and the historical limit is mentioned as a variant of the standard upper Shewhart chart and is equal to 4.3816.

The conclusion of this application is that all charts are able to detect the outbreak as soon as the second day (02/09/2013), except for the CRL-ZTP EWMA chart with $\kappa = 0.45$, which manage to detect the outbreak on day sooner than the others. The reason for that is the sensitivity of the chart. In addition, the current system in place has mostly similar performances to those of the Bernoulli-ZIP EWMA chart.

On April of 2013 Areepong and Sukparungsee (2013) studied the performance of a chart, named ZIP-MA, by using a moving average (MA) control chart. To do so, the two authors presented clear and specific formulas of ARL_0 and ARL_1 for the MA control chart when the distribution of interest is the ZIP. In the lines below, ARL_0 and ARL_1 formulas will be presented in brief and some conclusions that concern the numerical results of ARLs calculations will be mentioned.

Supposing that the number of nonconformities in an inspection unit are i.i.d ZIP random variables, the L-sigma control limits for the ZIP moving average (ZIP-MA) chart according to 1.6 are the following:

- For periods $i \geq w$ we have:

$$\begin{aligned} UCL_w &= \lambda(1-p) + L \cdot \sqrt{\frac{\lambda(1-p)(1+\lambda p)}{w}} \\ CL_w &= \lambda(1-p) \\ LCL_w &= \lambda(1-p) - L \cdot \sqrt{\frac{\lambda(1-p)(1+\lambda p)}{w}} \end{aligned}$$

- For periods $i < w$ we have:

$$\begin{aligned} UCL_i &= \lambda(1-p) + L \cdot \sqrt{\frac{\lambda(1-p)(1+\lambda p)}{i}} \\ CL_i &= \lambda(1-p) \\ LCL_i &= \lambda(1-p) - L \cdot \sqrt{\frac{\lambda(1-p)(1+\lambda p)}{i}} \end{aligned}$$

The alarm signal for the MA procedure is given when $M_i > UCL$ or $M_i < LCL$, where M_i is given by 1.3, 1.2.

Next, the authors are trying to give a closed form formula for evaluating ARL_0 and ARL_1 for the ZIP-MA control chart. So, these values can be derived as follows:
Let assume that ARL equals to n. Hence:

$$\begin{aligned}\frac{1}{ARL} &= \frac{1}{n} \cdot P\left(\text{out-of-control signal at time } i < w\right) + \frac{n-(w-1)}{n} \cdot P\left(\text{out-of-control signal at time } i \geq w\right) \\ &= \frac{1}{n} \cdot P\left(M_i > UCL_i \text{ and } M_i < LCL_i \text{ at time } i < w\right) \\ &+ \frac{n-(w-1)}{n} \cdot P\left(M_i > UCL_w \text{ and } M_i < LCL_w \text{ at time } i \geq w\right)\end{aligned}$$

So, by substituting M_i and the respective control limits, the ARL_0 and ARL_1 can be easily calculated (for details see Areepong and Sukparungsee (2013)).

Lastly, the numerical results for ARL_0, ARL_1 are computed from the closed form formulas. Given $ARL_0 = 370$ the parameter values for ZIP-MA control chat were chosen. The in-control parameter for the ZIP distribution is chosen to be $\lambda_0 = 5$ and the shifts which are examined are $\delta = 0.1, 0.2, ..., 0.9$. The probabilities of extra zeros that are checked are $p = 0.3, 0.5, 0.7$. The results are the following:

- When $p = 0.3$ and small shifts are happened ($\delta < 0.3$) the ZIP-MA chart detects quicker an out-of-control situation when $w = 20$, but for moderate and large shifts ($\delta \geq 0.3$) the ZIP-MA chart has the best performance when $5 \leq w \leq 15$.
- When $p = 0.5$ and small shifts are happened ($\delta \leq 0.3$) the ZIP-MA chart has the best detection ability when $w = 20$, but for moderate and large shifts ($\delta > 0.3$) the ZIP-MA chart has the best performance when $6 \leq w \leq 15$.
- When $p = 0.7$ and small shifts are happened ($\delta \leq 0.4$) the ZIP-MA chart detects quicker an out-of-control situation when $w = 20$, but for moderate and large shifts ($\delta > 0.4$) the ZIP-MA chart has the best performance when $7 \leq w \leq 15$
- In addition, when $ARL_0 = 500$ the results which arise are almost in agreement with the above outcomes, i.e when $ARL_0 = 370$.

This paper managed to show that the proposed formulas of ARLs are explicit, precise and easy to be computed and programmed. Based on the ARL values, ZIP-MA control chart was shown to perform better as the value of w decreases.

Chapter 4

Simulation Study

In this chapter is presented a simulation study in which we are trying to discuss some of the simulation results which are reported in the paper of He et al. (2011). Also, we are trying to give some extra arithmetic results.

The charts that are applied, examined and compared here are the following :

- p - CUSUM chart which detects increases in the parameter p (from p_0 to p_1) of the ZIP distribution.

$$B_i = \max(B_{i-1} + K_i, 0),\ i = 1, 2, ... \text{ where } K_i = \begin{cases} ln\frac{p_1+(1-p_1)e^{-\lambda_0}}{p_0+(1-p_0)e^{-\lambda_0}} & Y_i = 0 \\ ln\frac{1-p_1}{1-p_0} & Y_i > 0 \end{cases}$$

- λ - CUSUM chart which detects increases in the Poisson parameter (from λ_0 to λ_1).

$$L_i = \max(L_{i-1} + M_i, 0),\ i = 1, 2, ... \text{ where } M_i = \begin{cases} ln\frac{p_0+(1-p_0)e^{-\lambda_1}}{p_0+(1-p_0)e^{-\lambda_0}} & Y_i = 0 \\ Y_i ln\frac{\lambda_1}{\lambda_0} + (\lambda_0 - \lambda_1) & Y_i > 0 \end{cases}$$

- p - λ CUSUM chart which gives a signal whenever one of the above individual chart signals.

- t - CUSUM chart which detects simultaneously increases in both p and λ parameters of the ZIP distribution.

$$T_i = \max(T_{i-1} + N_i, 0),\ i = 1, 2, ... \text{ where } N_i = \begin{cases} ln\frac{p_1+(1-p_1)e^{-\lambda_1}}{p_0+(1-p_0)e^{-\lambda_0}} & Y_i = 0 \\ Y_i ln\frac{\lambda_1}{\lambda_0} + (\lambda_0 - \lambda_1) + ln\frac{1-p_1}{1-p_0} & Y_i > 0 \end{cases}$$

So, our first move is to design the above CUSUM charts. It is known that a control chart should signal quickly if the process is out-of-control, but it has to give insignificant number of false alarms if the process stays in-control. Hence, we use the ANOS (Average Number of Observations to Signal) value to measure the performance of the charts. More specific, according to the paper we are based on to actualize our application (He et al., 2011), the in-control ANOS ($ANOS_0$) is defined as the expected number of products inspected from the beginning of process monitoring until a signal is given. On the other hand, when the process gets out-of-control, the expected number of individual observations taken from the time of the shift until the chart first signals ($ANOS_1$) is used to evaluate the performance of different control charts.

Now, in order to design the CUSUM charts we make the assumption that the in-control values of the ZIP parameters are known. Also, specification of the shift sizes p_1 and λ_1 for the CUSUM charts to have fast detections relies completely on the practitioner's knowledge and experience. In our case, we are going to follow the author's lead.

Furthermore, we need to stipulate the control limits of the CUSUM charts. That is obtained by taking a pre-specified value of $ANOS_0$. By determining this value, it is easy to get the control limits which are necessary to construct the control charts.

Lastly, approximate ANOS values of all the CUSUM charts that are presented here, are obtained using the Regula Falsi method (or false position method). This method is going to be explained in the next lines and a pseudocode is going to be used for better understanding this methodology.

4.1 Regula Falsi Method

The Regula Falsi method is a numerical method for estimating the root of a function. In simple terms, this method is the trial and error technique of using test ("false") values for the variable and then adjusting the test value according to the outcome.

Given an equation, move all of its terms to one side so that it has the form, $f\,(x) = 0$, where f is some function of the unknown variable x. A value c that satisfies this equation, that is, $f\,(c) = 0$, is called a root of the function f and is a solution of the original equation. If f is a continuous function and there exist two points a_0 and b_0 such that $f\,(a_0) \cdot f\,(b_0) < 0$, then, the function f has a root in the interval (a_0, b_0).

This method calculates the new solution estimate as the x-intercept of the line segment joining the endpoints of the function on the current bracketing interval. More precisely, suppose that in the k^{th} iteration the bracketing interval is (a_k, b_k). Construct the line through the points $(a_k, f\,(a_k))$ and $(b_k, f\,(b_k))$, as illustrated. This line intersects the graph of the function f in at least two points. The equation of this straight line is:

$$y - f\,(b_k) = \frac{f\,(b_k) - f\,(a_k)}{b_k - a_k}(x - b_k)$$

Now choose c_k to be the x-intercept of this line, that is, the value of x for which y = 0, substitute these values above and move all of the terms to one side. Hence, solving this equation for c_k gives:

$$c_k = b_k - f\,(b_k)\,\frac{b_k - a_k}{f\,(b_k) - f\,(a_k)}$$

The convergence of this method is always guaranteed because it bracket the root of the function during the whole iteration process, so it makes it a good choice for solving an equation, with one unknown, relatively quickly (Ibrahim et al., 2016).

A graphical representation of this method is depicted below so as to fully understand someone the practical implementation of the theory.

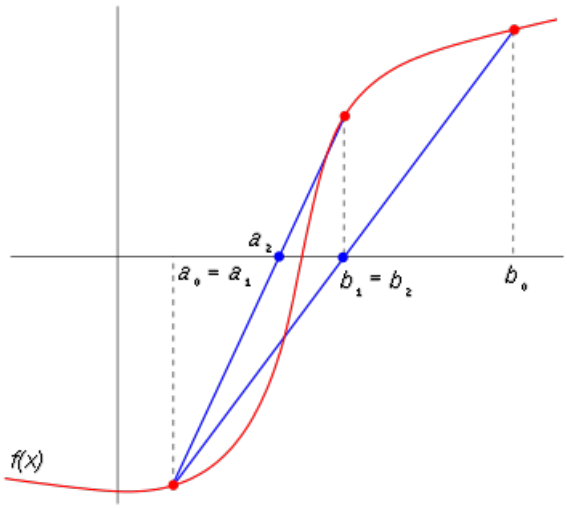

Figure 4.1: A graphical representation of the first two iterations of the regula falsi method. The red curve shows the function f and the blue lines are the secants.

Algorithm for the Regula Falsi

Assume that $f\ (x)$ is continuous on a given interval $[a_0, b_0]$ and that it also satisfies

$$f\ (a_0) \cdot f\ (b_0) < 0$$

Step 1. Set $a_k = a$, and $b_k = b$ for k = 0, 1, 2, ...

Step 2. Compute $c_k = b_k - f\ (b_k)\ \frac{b_k - a_k}{f\ (b_k) - f\ (a_k)}$.

Step 3. Check if $f(a_k)f(c_k) < 0$, then set $a_{k+1} = a_k$, $b_{k+1} = c_k$, else, $a_{k+1} = c_k$, $b_{k+1} = b_k$.

Step 4. Convergence and stopping criteria.
Check if $f\ (c_n) < \epsilon$, where threshold value is set as a very small number close to zero.

4.2 Simulation Analysis and Computational Results

Two in-control values $1 - p_0 = 0.1$ and $1 - p_0 = 0.2$ are considered here. Also, $\lambda_0 = 2$ and $\lambda_0 = 6$ are chosen to be the in-control values for simulation, $\lambda_0 = 2$ will have high influence on the occurrence of a zero-defect product, and $\lambda_0 = 6$ will barely have any impact on the probability of observing a zero count. Two shift sizes of 1-p and λ are pre-defined for the CUSUM charts to have fast detections: $1 - p_1 = 1.5(1 - p_0)$, $\lambda_1 = \lambda_0 + 1$ and $\lambda_1 = \lambda_0 + 2$.

A variety of possible shift sizes is evaluated when the process goes out-of-control. For $1 - p_0 = 0.2$, the shifted values considered are 1-p=0.25, 0.30, 0.40, ..., 0.90 and for $1 - p_0 = 0.1$, values of 1-p=0.125, 0.15, 0.20, 0.30, 0.40 and 0.50 are examined. As for λ, the shifts of interest are set to be $\lambda_0 + 1$, $\lambda_0 + 2$, $\lambda_0 + 3$, and $\lambda_0 + 4$. Moreover, it is assumed that there could be shifts in both the parameters p and λ, or only one of the two parameters may change. ANOS values of the p-CUSUM, λ-CUSUM, and t-CUSUM charts are then obtained using the Regula Falsi method, and ANOS values of the p-λ CUSUM chart is obtained via simulation.

In order to find the desired $ANOS_0$ values and the control limits of the different CUSUM charts, which are examined here, we made an attempt to write the ANOS as a function. This function has the following variables-parameters:

The in-control values of the parameters of the ZIP distribution.

The parameter(s) of the ZIP distribution with the change we want to detect.

A variable h which represents the control limit os the corresponding control chart.

And a parameter which represents the number of the iterations.

In this function, it is known the in-control values of the parameters of the ZIP distribution and we define the shift size(s) of the parameter(s) that is desired to be detected. After that, a "random" value is chosen for the variable h . Also, the number of iterations is determined. The larger the number of iterations, the higher the accuracy of the results.

Following the steps that are described above, an arithmetic $ANOS_0$ value is computed.The aim is to find the desired $ANOS_0$ value and the related control limit. So, we obtain two values of $ANOS_0$ (and two h values apparently) which enclose each time a target value of $ANOS_0$. After that, the Regula Falsi method takes part. Finally, when the repetitive algorithm comes to an end, the requested $ANOS_0$ value and the representative control limit are provided.

Control limits of various control charts are provided in Table 4.1. $ANOS_0$ is chosen to be around 200 for $1-p_0 = 0.2$ and 340 for $1-p_0 = 0.1$. Different values of $ANOS_0$ are selected for different p_0's because proper control limits cannot be found for some CUSUM scheme if all the parameter combinations are considered. For the p-λ CUSUM chart, the in-control ANOS is approximately the same as the t-CUSUM chart.

Table 4.1: UCLs of the CUSUM schemes.

	(a) $1-p_0 = 0.2$, $ANOS_0$ values	(b) $1-p_0 = 0.1$, $ANOS_0$ values	
$1-p_1$	0.3	0.15	0.15
λ_0	2	2	6
λ_1	3	4	7
h_t	2.2335	2.2980	1.8961
$ANOS_0$	200.02	340.54	340.23
h_p	2.1968	2.0041	2.1031
$ANOS_0$	359.91	609.26	601.34
h_λ	2.0333	2.2037	1.5465
$ANOS_0$	360.68	608.33	600.41
$ANOS_0$ for p-λ CUSUM	203.00	343.51	334.73

Comparisons among the various CUSUM schemes when only one parameter shifts while the other remains in-control are presented in Table 4.2, with $1-p_0 = 0.2$ and $\lambda_1 = \lambda_0 + 1$. The p-CUSUM and λ-CUSUM charts in this table are components of the corresponding p-λ CUSUM chart. In general, the p-λ CUSUM chart is considerably better at detecting shifts in p while the t-CUSUM chart is slightly better at detecting changes in λ. From Table 4.2, we find that as 1-p increases, the λ-CUSUM chart signals faster due to increased sampling rate of the Poisson distribution. The performance of the p-CUSUM chart is also affected by increases in λ.

In addition, a graphical representation of the ANOS results and the p-λ CUSUM chart and t-CUSUM chart are shown in Figure 4.2. The left Figure of 4.2 considers the case when $\lambda_0 = 2$, $\lambda_1 = \lambda_0 + 1$, $1-p_1 = 1.5(1-p_0)$ and ANOS values plotted based on the assumption that p remains stable and λ increases from 2 to 6. The performance of the two charts seems to be almost the same. Moreover, the right graph shows again the performance of those two charts, but now, the hypothesis

Table 4.2: ANOS values of different control charts with only one parameter shift ($\lambda_1 = \lambda_0 + 1$)

$1 - p_0 = 0.2,\ 1 - p_1 = 1.5(1 - p_0)$						
1-p	λ_0	λ	p-CUSUM	λ-CUSUM	p-λ CUSUM	t-CUSUM
0.20	**2**	**2**	**359.91**	**360.68**	**203.00**	**200.02**
0.25	2	2	123.66	259.54	93.78	119.01
0.30	2	2	63.88	200.04	55.18	79.66
0.40	2	2	30.01	136.56	27.68	43.05
0.50	2	2	19.53	102.81	18.23	27.66
0.60	2	2	14.18	81.20	13.66	19.82
0.70	2	2	11.31	67.06	10.89	15.32
0.80	2	2	9.31	56.69	9.06	12.46
0.90	2	2	8.01	49.50	7.85	10.57
0.20	**2**	**2**	**359.91**	**360.68**	**203.00**	**200.02**
0.20	2	3	220.40	48.62	48.89	40.07
0.20	2	4	190.06	21.71	22.78	19.49
0.20	2	5	180.90	14.36	14.94	12.83
0.20	2	6	176.55	10.92	10.93	9.85

is that λ is not changing and 1-p shifts from 0.2 to 0.9. The plot illustrates that p-λ CUSUM chart is better than the other. However, when $1 - p > 0.7$ ANOS values are almost the same.

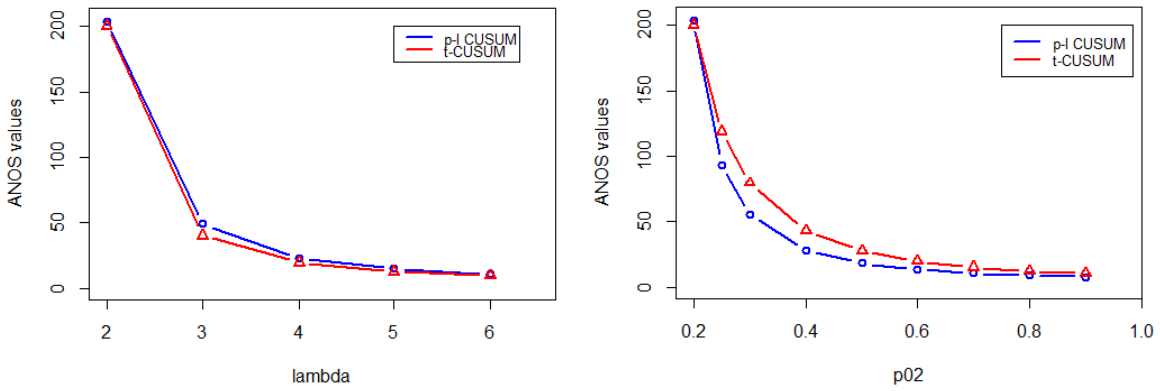

Figure 4.2: Left: ANOS comparisons based on $1 - p_0 = 0.2$, $\lambda_0 = 2$, $\lambda_1 = \lambda_0 + 1$, $1 - p_1 = 1.5(1 - p_0)$, $1 - p = 0.2$ and λ varies from 2 to 6. Right: $\lambda_0 = 2$, $\lambda_1 = \lambda_0 + 1$, $1 - p_1 = 1.5(1 - p_0)$, $\lambda = 2$ and 1-p varies from 0.2 to 0.9.

In Table 4.3 the ANOS values of different CUSUM schemes based on $\lambda_1 = \lambda_0 + 2$ and $1 - p_1 = 1.5(1 - p_0)$, where $1 - p_0 = 0.1$ and $\lambda_0 = 2$, are provided. The t-CUSUM chart has lightly better performance than the p-λ CUSUM chart when λ changes and p remains stable. However, when there is no shift in λ parameter, but p changes p-λ CUSUM chart is much better than the corresponding t-CUSUM chart.

Furthermore, in order to compare the p-λ CUSUM chart with the t-CUSUM chart, we represent the ANOS values in both situations graphically. Figure 4.3 considers the situation when $\lambda_0 = 6$, $\lambda_1 = \lambda_0 + 1$, $1 - p_1 = 1.5(1 - p_0)$. ANOS values were plotted based on the assumption that there is no shift in p and shifts in λ range from 6 to 10 and there is no shift in λ and shifts in 1-p range from 0.1 to 0.5, accordingly. The left figure, i.e when λ increases, shows that ANOS values of the

Table 4.3: ANOS values of different control charts with only one parameter shift ($\lambda_1 = \lambda_0 + 2$)

$1 - p_0 = 0.1,\ 1 - p_1 = 1.5(1 - p_0)$						
1-p	λ_0	λ	p-CUSUM	λ-CUSUM	p-λ CUSUM	t-CUSUM
0.10	**2**	**2**	**609.26**	**608.33**	**343.51**	**340.54**
0.125	2	2	231.42	451.09	173.77	258.52
0.15	2	2	123.48	364.18	105.20	198.53
0.20	2	2	59.42	242.56	55.02	131.85
0.30	2	2	29.59	149.37	27.09	75.74
0.40	2	2	19.22	105.51	18.25	50.91
0.50	2	2	14.44	82.79	13.75	38.64
0.10	**2**	**2**	**609.26**	**608.33**	**343.51**	**340.54**
0.10	2	3	391.58	93.40	88.08	71.2
0.10	2	4	338.36	39.26	39.49	33.42
0.10	2	5	320.39	25.33	25.38	21.60
0.10	2	6	309.73	18.79	19.03	16.65

p-λ CUSUM chart are almost the same as those of the t-CUSUM chart. The right figure, illustrates the performances of the two charts when 1-p increases. Clearly, t-CUSUM chart has better detection ability than the p-λ CUSUM chart.

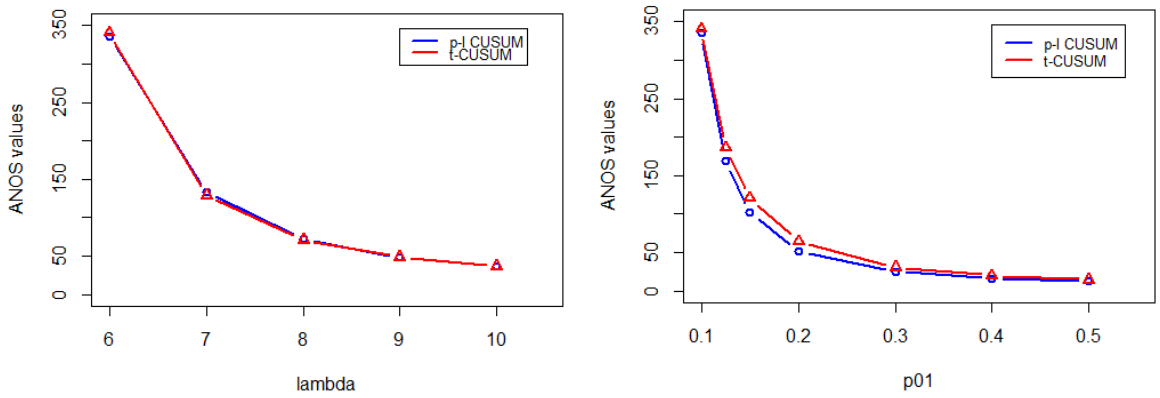

Figure 4.3: Left: ANOS comparisons based on $1 - p_0 = 0.1$, $\lambda_0 = 6$, $\lambda_1 = \lambda_0 + 1$, $1 - p_1 = 1.5(1 - p_0)$, $1 - p = 0.1$ and λ varies from 6 to 10. Right: $\lambda_0 = 6$, $\lambda_1 = \lambda_0 + 1$, $1 - p_1 = 1.5(1 - p_0)$, $\lambda = 6$ and 1-p varies from 0.1 to 0.5.

Also, Table 4.4 presents the ANOS results of the four CUSUM schemes based on $1 - p_0 = 0.1$, $1 - p_1 = 1.5(1 - p_0)$, $\lambda_0 = 6$ and $\lambda_1 = \lambda_0 + 1$, with 1-p varying from 0.10 to 0.50 and λ from 6 to 10. In this case, it is observed that when there is a shift in λ and p is stable the performance of p-λ CUSUM chart is almost the same as the t-CUSUM chart. On the other hand, when λ remains stable and p shifts, p-λ CUSUM chart is considerably better than t-CUSUM chart. Moreover, p-CUSUM chart is not helpful at all when λ shifts and p is unchangeable.

Table 4.4: ANOS values of different control charts with only one parameter shift ($\lambda_1 = \lambda_0 + 1$)

$1 - p_0 = 0.1$, $1 - p_1 = 1.5(1 - p_0)$						
1-p	λ_0	λ	p-CUSUM	λ-CUSUM	p-λ CUSUM	t-CUSUM
0.10	**6**	**6**	**601.34**	**600.41**	**334.73**	**340.23**
0.125	6	6	212.23	466.57	169.60	186.54
0.15	6	6	113.68	382.83	101.36	120.52
0.20	6	6	54.27	286.11	51.68	64.50
0.30	6	6	26.08	186.43	25.78	30.95
0.40	6	6	17.32	141.75	16.90	20.41
0.50	6	6	12.40	114.31	12.69	15.04
0.10	**6**	**6**	**601.34**	**600.41**	**334.73**	**340.23**
0.10	6	7	585.51	148.04	133.04	128.81
0.10	6	8	585.17	72.60	72.49	71.02
0.10	6	9	585.07	48.65	48.09	49.01
0.10	6	10	583.37	3.80	36.89	36.94

In all the above Tables (4.2 - 4.4), we evaluated the performance of the CUSUM charts under different circumstances such as 1-p increases while λ remains unchanged, or λ increases while p is stable. Another scenario is for p and λ to simultaneously shift away from their in-control values, and this is covered in Table 4.5 (for $1 - p_0 = 0.2$ and $\lambda_0 = 2$). It is found that the t-CUSUM chart has better performance than the p- λ CUSUM combination when both parameters, λ and 1-p, increase. Thus, if one expects both parameters to increase simultaneously, it is more appropriate to use the t-CUSUM chart, which is designed for this purpose.

Table 4.5: ANOS values of different control charts with $1 - p_0 = 0.2$ and $\lambda_0 = 2$.

$1 - p_1 = 1.5(1 - p_0)$					
	λ	$1 - p = 0.25$	$1 - p = 0.30$	$1 - p = 0.40$	$1 - p = 0.50$
	3	34.62	25.88	17.22	12.56
p-λ CUSUM	4	17.65	14.35	10.53	8.17
($\lambda_1 = \lambda_0 + 1$)	5	11.69	9.66	7.20	5.71
	6	8.93	7.34	5.49	4.36
	3	28.15	20.92	13.64	10.19
t-CUSUM	4	14.54	11.00	8.03	6.17
($\lambda_1 = \lambda_0 + 1$)	5	9.91	7.95	5.81	.53
	6	7.47	6.38	4.60	3.58

For completeness, in Table 4.6 and 4.7 we consider the situation when one of the parameters increases, but the other decreases. We find that the p-λ CUSUM chart performs not so well when 1-p decreases and λ increases and particularly when $\lambda_1 = \lambda_0 + 4$ and 1-p decreases continually. Generally, when $1 - p = 0.15$ and λ increases, t- CUSUM chart is a little bit better. Finally, the performance of the t-CUSUM chart is slightly worst than the other in all the remain cases. In the case where 1-p increases and λ decreases, we find that the t-CUSUM chart performs very poorly. Overall, we would prefer using the p-λ CUSUM method since it is better in terms of the total performance.

Table 4.6: ANOS values when λ increases and 1-p decreases, $1 - p_1 = 1.5(1 - p_0), \lambda_1 = \lambda_0 + 1$.

$1 - p_0 = 0.2, \lambda_0 = 2$				
		1-p=0.15	1-p=0.10	1-p=0.05
$\lambda = \lambda_0 + 1$	p-λ CUSUM	70.42	119.50	352.92
	t-CUSUM	65.33	127.68	373.90
$\lambda = \lambda_0 + 2$	p-λ CUSUM	30.57	48.94	117.10
	t-CUSUM	28.97	51.55	128.76
$\lambda = \lambda_0 + 3$	p-λ CUSUM	19.58	30.37	67.60
	t-CUSUM	18.09	29.74	69.51
$\lambda = \lambda_0 + 4$	p-λ CUSUM	14.69	22.59	47.01
	t-CUSUM	13.41	20.83	44.55

Table 4.7: ANOS values when λ decreases and 1-p increases, $1 - p_1 = 1.5(1 - p_0), \lambda_1 = \lambda_0 + 1$

$1 - p_0 = 0.2, \lambda_0 = 2$					
		1-p=0.25	1-p=0.30	1-p=0.40	1-p=0.50
$\lambda = \lambda_0 - \frac{1}{2}$	p-λ CUSUM	188.99	89.78	38.64	23.54
	t-CUSUM	583.91	358.09	165.93	91.72
$\lambda = \lambda_0 - 1$	p-λ CUSUM	592.96	226.25	69.31	36.97
	t-CUSUM	10025.27	6406.48	2903.45	1398.88

Chapter 5

Conclusions

SQC is a field of statistics widely applied in industrial and nonindustrial processes in order to maintain the production processes under control. In practice, the way that statistics are used in quality control is through the construction of control charts, which are one of the most basic techniques of SPC in the effort to achieve and maintain stability in various processes. In high-yield processes there exists an excessive number of zero-defect counts. So, the Zero Inflated Poisson (ZIP) distribution is useful in the case of a near zero-defect manufacturing environment. In this thesis, an analytical review of the work that has been done until now concerning process monitoring methods using the ZIP model took place.

We emphasized in the CUSUM control charts and tested the work of He et al. (2011) which proposed a control charting procedure using a combination of two CUSUM charts for monitoring increases in the two parameters of the ZIP process. Furthermore, a single CUSUM chart for detecting simultaneous increases in the two parameters were examined. We used the Regula Falsi method to produce the values that were needed to compare the performance of the different CUSUM charts of the paper. Simulation results established that this method produce ARL values very close to those of the paper and gives desired results.

Although much research has been done regarding the ZIP distribution and the statistical process monitoring, however there are still some matters that could be done. First of all, not much research concerning the EWMA control charts exists. Someone could be involved much more in that direction. Also, as regards the effect of estimation error concerning the ZIP model, only the case where the parameter λ is unknown while the other assuming known has been examined. For future research someone can include the case when all the parameters are unknown, i.e both p and λ and as well as the study of the performance of more complex schemes such as CUSUM-type and EWMA-type - only Shewhart type chart has been investigated - when the parameters are unknown and have to be estimated.

In addition, a future extension of He et al. (2012) paper specifically, could be the identification of the shifted parameters in the p-set and λ-set parameters. That will make this research more applicable. Furthermore, monitoring multivariate (or bivariate) rare events, where correlations between the two attribute quality characteristics exist, is an important issue in statistical process control. There are few methods in the literature for monitoring such bivariate quality characteristics applying multivariate (or bivariate) distribution of the attributes. So, this is a matter that someone could be involved itself on a great scale.


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Appendix A

Appendix

A.1 List of Abbreviations

SQC	Statistical Quality Control
SPC	Statistical Process Control
CUSUM	Cumulative Sum
EWMA	Exponentially Weighted Moving Average
MA	Moving Average
CL	Control Limit
UCL	Upper Control Limit
LCL	Lower Control Limit
ARL	Average Run Length
ANOS	Average Number of Observations to Signal
ZIP	Zero Inflated Poisson
BZIP	Bivariate Zero Inflated Poisson
MZIP	Multivariate Zero Inflated Poisson
MLE	Maximum Likelihood Estimators
MME	Method of Moments Estimators
GZIP	Generalized Zero Inflated Poisson
FAR	False Alarm Rate
AR	Alarm Rate
SDRL	Standard Deviation of Run Length
ACP	Average Coverage Probability
ZTP	Zero Truncated Poisson
MCA	Markov Chain Approach

References

D A. Brook and D A. Evans. An approach to the probability distribution of cusum run length. *Biometrika*, 59, 12 1972.

Yupaporn Areepong and Saowanit Sukparungsee. Closed form formulas of average run lentgh of moving average control chart for nonconforming for zero-inflated process. *Far East Journal of Mathematical Sciences*, 2, 04 2013.

Michele Basseville and Igor V. Nikiforov. *Detection of Abrupt Changes: Theory and Application*, pages 165–166. Prentice-Hall, 1993.

Michael Boon Chong. A moving average control chart for monitoring the fraction non-conforming. *Quality and Reliability Engineering International*, 20:617 – 635, 10 2004.

Connie M Borror, Charles W Champ, and Steven E Rigdon. Poisson ewma control charts. *Journal of Quality Technology*, 30(4):352–361, 1998.

T. Tony Cai. One-sided confidence intervals in discrete distributions. *Journal of Statistical Planning and Inference*, 131(1):63 – 88, 2005.

Nan Chen, Shiyu Zhou, Tzyy-Shuh Chang, and Howard Huang. Attribute control charts using generalized zero-inflated poisson distribution. *Quality and Reliability Engineering International*, 24, 2008.

A. Clifford Cohen. *Truncated and Censored Samples*. Boca Raton: CRC Press, 1991.

William G. Cochran. Some methods for strengthening the common 2 tests. *International Biometric Society*, 10:417–451, December 1954.

Patrick D. Bourke. Detecting a shift in fraction nonconforming using run-length control charts with 100% inspection. *Journal of Quality Technology*, 23:225–238, 07 1991.

Mavroudis Eleftheriou and Nikolaos Farmakis. One-sided ewma control chart for monitoring high yield processes. 2011.

Amir Afshin Fatahi, Rassoul Noorossana, Pershang Dokouhaki , and Babak Farhang Moghaddam. Zero inflated poisson ewma control chart for monitoring rare health-related events. *Journal of Mechanics in Medicine and Biology*, 12, 10 2012a.

Amir Afshin Fatahi, Rassoul Noorossana, Pershang Dokouhaki, and Babak Farhang Moghaddam. Copula-based bivariate zip control chart for monitoring rare events. *Communications in Statistics - Theory and Methods*, 41(15):2699–2716, 2012b.

David A. Garvin. *Managing Quality: The Strategic and Competitive Edge*. Free Press, 1988.

J.D. Gibbons and S. Chakraborti. *Nonparametric Statistical Inference 4th Edition*, chapter 4, page 111. Marcel Dekker, 2003.

Edit Gombay and Daniel Serban. Monitoring parameter changein ar(p) time series models. *Journal of Multivariate Analysis*, pages 715–725, 2009.

A H el Shaarawi. Some goodness-of-fit methods for the poisson plus added zeros distribution. *Applied and environmental microbiology*, 49:1304–6, June 1985.

B. He, M. Xie, T. N. Goh, and P. Ranjan. On the estimation error in zero-inflated poisson model for process control. *International Journal of Reliability, Quality and Safety Engineering*, 10(2): 159–169, 2003.

Shu-Guang He, Wandi Huang, and William Woodall. Cusum charts for monitoring a zero-inflated poisson process. *Quality and Reliability Engineering International*, 28, 01 2011.

Shu-Guang He, Zhen He, and G Wang. Cusum charts for monitoring bivariate zero-inflated poisson processes with an application in the led packaging industry. *Components, Packaging and Manufacturing Technology, IEEE Transactions on*, 2, 01 2012.

Shuguang He, Shijie Li, and Zhen He. A combination of cusum charts for monitoring a zero-inflated poisson process. *Communications in Statistics - Simulation and Computation*, 43(10):2482–2497, 09 2014.

Jaewon Huh, Hanwool Kim, and Sangyeol Lee. Monitoring parameter shift with poisson integer-valued garch models. *Journal of Statistical Computation Simulation*, 87:1756–1766, 2017.

Sulaiman Ibrahim, Mustafa Mamat, Mohammed Waziri, A Fadhilah, and Kamilu Uba kamfa. Regula falsi method for solving fuzzy nonlinear equation. 100:877–878, 09 2016.

Norman L. Johnson, Samuel Kotz, and N. Balakrishnan. *Discrete Multivariate Distributions*, chapter 37. Multivariate Poisson Distributions, pages 124–142. Wiley-Interscience, 1996.

N. Katemee and T. Mayureesawan. Control charts for zero-inflated poisson models. *Applied Mathematical Sciences*, 6(56):2791–2803, 2012a.

N. Katemee and T. Mayureesawan. Nonconforming control charts for zero-inflated poisson distribution. *International Journal of Mathematical and Computational Sciences*, 6(9), 2012b.

K. Krishnamoorthy. *Handbook of Statistical distribution with applications*, chapter 17, page 207. Taylor and Francis Group, 2006.

Sangyeol Lee, Youngmi Lee, and Cathy W.S Chen. Parameter change test for zero-inflated generalized poisson autoregressive models. *Journal of Theoretical and Applied Statistics*, 2015.

Robert Neil Leong, Frumencio Co, and Daniel Stanley Tan. Some zero inflated poisson-based combined exponentially weighted moving average control charts for disease surveillance. *The Philippine Statistician*, 64:17–42, 12 2015.

Chin-Shang Li, Jye-Chyi Lu, Jinho Park, Kyungmoo Kim, Paul A. Brinkley, and John P. Peterson. Multivariate zero-inflated poisson models and their applications. *Technometrics*, 41(1):29–38, 1999.

James M. Lucas and Michael S. Saccucci. Exponentially weighted moving average control schemes: Properties and enhancements. *Technometrics*, 32(1):1–12, 1990.

Douglas C. Montgomery. *Intoduction to Statistical Quality Control*. John Wiley Sons, Inc, 2009.

Seyed Niknam and Tomcy Thomas. Analysis of acoustic emission data for bearings subject to unbalance. *International Journal of Prognostics and Health Management*, 4:10, 07 2013.

Athanasios C. Rakitzis and Philippe Castagliola. The effect of parameter estimation on the performance of one-sided shewhart control charts for zero-inflated processes. *Communications in Statistics-Theory and Methods*, pages 413–430, 2015.

Athanasios C. Rakitzis, Christian H. Wei, and Philippe Castagliola. Control charts for monitoring correlated poisson counts with an excessive number of zeros. *Quality and Reliability Engineering International*, pages 413–430, 2016.

C. Radhakrishna Rao and I. M. Chakravarti. Some small sample tests of significance for a poisson distribution. *International Biometric Society*, 12:264–282, September 1956.

C. H. Sim and M. H. Lim. Attribute charts for zero-inflated processes. *Communications in Statistics - Simulation and Computation*, 37(7):1440–1452, 2008.

F. W. Steutel and K." van Harn. Discrete analogues of self-decomposability and stability. *Annals Probability*, pages 893–899", 10 1979.

Saowanit Sukparungsee. Average run length of cumulative sum control chart by markov chain approach for zero-inflated poisson processes. *Thailand Statistician*, 16, 01 2018.

J van den Broek. A score test for zero inflation in a poisson distribution. *Biometrics*, 51(2):738—743, June 1995.

M. Xie, B. He, and T.N. Goh. Zero-infated poisson model in statistical process control. 2001.