

A HEURISTIC FIXED LIMIT BAYESIAN P CHART

M.H. Abooie and M. Amin Nayeri

Department of Industrial Engineering
Amirkabir University of Technology Hafez, Iran
mjnayeri@aut.ac.ir

ABSTRACT

The paper proposes an efficient approach to detecting an increase in the fraction of non-conforming items. The novelty of the paper is its utilization of the concept of the Bayesian rule and construction of a Bayesian control chart. This approach is significantly better than certain existing effective approaches in detecting small deviations. The major application of the charts is in high-tech industries and short run processes where the detection of small deviations and the evaluation of the initial setup are very important. The simulated results for the average run length profiles demonstrate the superiority of the new approach against the standard p chart, binomial EWMA and moving average approach. The new approach is easy to understand and may be attractive and useful to researchers, while it can also be an effective alternative for other existing approaches.

OPSOMMING

Die navorsing hou 'n doeltreffende Bayes-gebaseerde kontrolekaartmetode voor vir die beheer van breukdefekte waar gehalte buite beheer raak vir klein afwykings. Die metode se vertoning word *via* simulatie teen ander geykte kontrolekaartmetodes vir gemiddelde looplengte beproef. Die nuwe benadering is maklik verstaanbaar en bruikbaar vir navorsers, terwyl dit ook 'n effektiewe alternatief vir ander bestaande benaderings is.

1. INTRODUCTION

The main purpose of statistical process control is to improve quality and productivity. One of the instruments of the quality tool set is the control chart. Control charts are efficient instruments for checking changes or variations in the processes. The proposed tool is intended to better detect the variations in the average when it is necessary to control a sensitive process [1].

The effectiveness of a control chart is usually measured by ARL (Average Run Length), which is the average number of inspected samples required to signal an out of control condition after it has occurred. When a process is out of control, the users wish the control chart to signal promptly, i.e. to have a small out-of-control average run length (ARL_1). Conversely, when the process is in control, the users wish the chart to produce fewer false alarms, i.e. to have a large in-control average run length (ARL_0) [2].

The traditional attribute control charts, or the Shewhart p charts, have been widely used in industries to control the fraction non-conforming of the process. A non-conforming unit is a product or service that fails to meet at least one specified requirement. The proportion p is defined as the ratio of the non-conforming units in a process to the total number of units produced in that process. Although the Shewhart p charts have the advantage of simplicity, one of the disadvantages of using these charts is that they only use the information about the process at the last plotted point and ignore information given by the sequence of all points. This feature makes Shewhart control charts relatively insensitive to small changes in the process; in the order of 1.5 of standard deviation or less [1]. The main purpose of this research is to solve this problem by developing an effective Bayesian p chart.

2. LITERATURE REVIEW

Control charting methods based on attribute data were first proposed by Shewhart [3]. As mentioned before, one of the major weaknesses of Shewhart p charts is ignorance of information given by the sequence of all points. To solve this problem, Shewhart control charts are efficiently complemented by CUSUM (Cumulative Sum) and EWMA (Exponentially Weighted Moving Average) control charts when there is interest in detecting small changes in the process. These methods are considered highly efficient in detecting special causes of the variation, which lead to the non-conformity of the production [1].

The CUSUM control chart is more effective than the Shewhart p chart in detecting small and moderate p shifts, but the p chart is still considered as the most widely used attribute control chart. The methods of the CUSUM charts are more difficult for the operators to use and understand. Besides, the use of CUSUM charts by practitioners for monitoring p has been limited by the difficulty in designing these charts to achieve specified properties [4].

The EWMA control chart was introduced by Roberts [5]. The performance of the EWMA control chart is approximately equivalent to that of the CUSUM control chart [6]. Hunter [7] stated that the differences between Shewhart, CUSUM, and EWMA control charts are related to the way each charting technique uses the data generated by the production process. He illustrated in its simplest form how the charts weigh the data: The Shewhart chart depends entirely on the last demarcated point, CUSUM chart assigns equal weight to the most ancient data as well as the most recent, while EWMA assigns greater weight to more updated information and lower weight for more remote information.

Montgomery [6] explained that although EWMA is presented as a statistical process-monitoring tool, it really has a much wider interpretation. From the viewpoint of statistical process control, the EWMA control chart is comparable to the CUSUM control chart in its capacity of monitoring a process and detecting the presence of assignable causes, which result in changes. However, EWMA produces forecasting of where the average will be in the next period of time and how it is easy to apply in industry. This makes EWMA a more powerful tool.

Some authors, such as Crowder [8, 9], Lucas and Saccucci [10], presented the EWMA control chart as a good choice to detect changes in extension of the process average. There are many authors who have made good contributions to this kind of control chart. Gan [11] introduced three modified EWMA charts for Poisson data. Borrer *et al.* [12] proposed a better procedure for using the EWMA chart for a Poisson count. Somerville *et al.* [13] developed a smoothing and filtering method using the EWMA and Poisson probabilities, which separates the two distributions in a particle count data stream into a base process and an outlier process followed by applying statistical monitoring schemes to each of them.

A Bernoulli EWMA was introduced by Somerville *et al.* [13] to monitor the outlier process. The EWMA control chart for the fraction of non-conforming can also be constructed as the binomial EWMA. The ARL performance of the binomial EWMA is superior to the standard p chart [14].

A moving average control chart is a memory control chart based on unweighted moving average. Khoo [14] presents a way to extend the moving average chart to the case of binomial data for the fraction non-conforming. He stated that due to its simplicity and superiority in detecting shifts, the moving average control chart for fraction non-conforming should be given high priority by quality control practitioners; and since it is more likely for a process to start in-control first, following a warm-up period before the process shifts to an out-of-control state (if a shift occurs), the binomial EWMA may perform poorly under such conditions.

One of the other approaches used to improve the control chart is by using the Bayesian rule. It is well known that the Bayesian approach leads to optimal adaptive process control rules, which take all the information into account about the state of the process, accumulated from the beginning of the production run [15]. On the other hand, the Bayesian approach directly incorporates all the information in the sequence of sample values by implementing the behavior of all of the sample value deviations from a value objective.

Calabrese [16] developed a Bayesian model for process control under standard cost and operating assumptions and in the consequent characterization of the optimal policy. Marcellus [17] defined a Bayesian analogue of the Shewhart \bar{x} -chart and his comparison studies identified types of production processes where the Bayesian chart has an improved expected performance than the cumulative sum chart.

In this research, a novel Bayesian control chart is developed for attribute characteristics (Bayesian p chart) in which an initial value for probabilities of in-control or out-of-control states of process is assumed and a Bayesian approach is employed to determine the out-of-control state. Also, for verification of the method, the proposed method is compared to the binomial EWMA, moving average and Shewhart p control charts.

Furthermore, as explained later, the proposed method effectively improves the performance of the attribute control chart by using all past data. Through Bayesian inference, the performance of control charts improves. This method is also superior to other mentioned methods.

3. DESCRIPTION OF OUR METHODOLOGY

In this section, the methodology to improve the performance of attribute control charts is described. To simplify, it is assumed that the sample size n is constant. At iteration k of the data gathering process, $O_k = (x_1, x_2, \dots, x_k)$ is defined as an observation vector where x_i is the number of defective products in the iteration i ($i = 1, 2, \dots, k$). After observing the number of defects at iteration k named x_k , the belief of being in an out-of-control state $B(x_k, O_{k-1})$ is defined as follows:

$$B(x_k, O_{k-1}) = \Pr \{ \text{Out-of-control} | x_k, O_{k-1} \} \quad (1)$$

At this iteration, to improve the belief of being in an out-of-control state based on the observation vector O_{k-1} and the new observation x_k is performed. If we define $B(O_{k-1}) = B(x_{k-1}, O_{k-2})$ as the prior belief of an out-of-control state, in order to update the posterior belief or $B(x_k, O_{k-1})$, since it is assumed that each iteration has been observed independently:

$$\Pr \{ x_k | \text{Out-of-control}, O_{k-1} \} = \Pr \{ x_k | \text{Out-of-control} \} \quad (2)$$

The initial value for $B(O_k)$ is $B(O_0)$. Determination of $B(O_0)$ depends on the initial evaluation of the process. $B(O_0) = 0.5$ means that the probabilities of being in- or out-of-control state of the process are the same (equal to 0.5). $B(O_0)$ is given as $B(0)$.

With this feature, by using Bayesian rule, the posterior belief is:

$$\begin{aligned} B(x_k, O_{k-1}) &= \Pr \{ \text{Out-of-control} | x_k, O_{k-1} \} \\ &= \frac{\Pr \{ \text{Out-of-control}, x_k | O_{k-1} \}}{\Pr \{ x_k | O_{k-1} \}} \\ &= \frac{\Pr \{ \text{Out-of-control} | O_{k-1} \} \Pr \{ x_k | \text{Out-of-control}, O_{k-1} \}}{\Pr \{ x_k | O_{k-1} \}} \end{aligned} \quad (3)$$

Consequently equation (3) becomes:

$$\begin{aligned} B(x_k, O_{k-1}) &= \frac{\Pr \{ \text{Out-of-control} | O_{k-1} \} \Pr \{ x_k | \text{Out-of-control} \}}{\Pr \{ \text{Out-of-control} | O_{k-1} \} \Pr \{ x_k | \text{Out-of-control} \} + \Pr \{ \text{In-control} | O_{k-1} \} \Pr \{ x_k | \text{In-control} \}} \\ &= \frac{B(O_{k-1}) \Pr \{ x_k | \text{Out-of-control} \}}{B(O_{k-1}) \Pr \{ x_k | \text{Out-of-control} \} + (1 - B(O_{k-1})) \Pr \{ x_k | \text{In-control} \}} \end{aligned} \quad (4)$$

Assume that the quality characteristic of the process has a binomial distribution with parameter $p = p_0$ and use equation (4) to evaluate the probabilities of being in out-of-control state or in-control state, as follows:

$$\Pr \{ x_k | \text{out-of-control} \} = \frac{e^{-\frac{x_k}{n}}}{e^{p_0} + e^{-\frac{x_k}{n}}} \quad (5)$$

$$\Pr \{ x_k | \text{in-control} \} = \frac{e^{p_0}}{e^{p_0} + e^{-\frac{x_k}{n}}} \quad (6)$$

Take into account that there are two basic reasons to use equations 5 and 6. Firstly, they are logical functions to use, and secondly, by utilization of these equations, one can define the desired recursive function.

p_0 is the in-control quality level and p_1 is the out-of-control quality level. Assume that if the fraction of non-conforming is more than p_1 , then the process is out-of-control.

Hence

$$B(x_k, O_{k-1}) = \frac{B(O_{k-1})e^{\frac{x_k}{n}}}{B(O_{k-1})e^{\frac{x_k}{n}} + (1 - B(O_{k-1}))e^{p_0}} \quad (7)$$

To derive the required recursive functions, assume $Z_k = \frac{B(x_k, O_{k-1})}{1 - B(x_k, O_{k-1})}$

$$z_k = e^{\frac{x_k - p_0}{n}} z_{k-1} = \dots = e^{\sum_{i=1}^k \left(\frac{x_i - p_0}{n} \right)} z_0 \Rightarrow Ln(z_k) = \sum_{i=1}^k \left(\frac{x_i}{n} \right) - kp_0 + Ln(z_0)$$

To construct our control limits we suppose $B(0)=0.5$ and then $Ln(z_0)=0$. Since the sum of two binomial distributions is a binomial distribution

$$\sum_{i=1}^k x_i \approx Binomial(nk, p_0)$$

and now, defining UCL and LCL for $Ln(z_k)$ according to its distribution:

$$\begin{aligned} n(Ln(z_k) + kp_0) &\approx Binomial(kn, p_0) \Rightarrow \begin{cases} UCL = knp_0 + c\sqrt{knp_0(1-p_0)} \\ LCL = knp_0 - c\sqrt{knp_0(1-p_0)} \end{cases} \\ &\Rightarrow \frac{knp_0 - c\sqrt{knp_0(1-p_0)}}{n} \leq Ln(z_k) + kp_0 \leq \frac{knp_0 + c\sqrt{knp_0(1-p_0)}}{n} \\ &\Rightarrow kp_0 - c\sqrt{\frac{kp_0(1-p_0)}{n}} - kp_0 \leq Ln(z_k) \leq kp_0 + c\sqrt{\frac{kp_0(1-p_0)}{n}} - kp_0 \end{aligned} \quad (8)$$

It may be shown that:

$$\begin{aligned} -c\sqrt{\frac{kp_0(1-p_0)}{n}} &\leq Ln\left(\frac{B(x_k, O_{k-1})}{1 - B(x_k, O_{k-1})}\right) \leq c\sqrt{\frac{kp_0(1-p_0)}{n}} \\ &\Rightarrow e^{-c\sqrt{\frac{kp_0(1-p_0)}{n}}} \leq \frac{B(x_k, O_{k-1})}{1 - B(x_k, O_{k-1})} \leq e^{c\sqrt{\frac{kp_0(1-p_0)}{n}}} \\ &\Rightarrow \frac{1}{e^{c\sqrt{\frac{kp_0(1-p_0)}{n}}}} \leq \frac{1}{B(x_k, O_{k-1})} - 1 \leq \frac{1}{e^{-c\sqrt{\frac{kp_0(1-p_0)}{n}}}} \\ &\Rightarrow \frac{e^{-c\sqrt{\frac{kp_0(1-p_0)}{n}}}}{1 + e^{-c\sqrt{\frac{kp_0(1-p_0)}{n}}}} \leq B(x_k, O_{k-1}) \leq \frac{e^{c\sqrt{\frac{kp_0(1-p_0)}{n}}}}{1 + e^{c\sqrt{\frac{kp_0(1-p_0)}{n}}}} \end{aligned} \quad (9)$$

As can be observed from equation (9), limits are divergent. The upper limit diverges to 1 and the lower limit diverges to 0 as the parameter k is increased. Thus is undesirable since divergent limits caused increasing of error type II. One effective approach to solve this problem is to define a constant number for k named l , and fix the limits by $k=l$. By using this approach one is able to reach proper the ARL_1 and ARL_0 with the definition of the appropriate value for c , l and $B(0)$.

In summary, the proposed method may be defined as:

If $B(x_k, O_{k-1})$ is not in interval $\frac{e^{-c\sqrt{\frac{lp_0(1-p_0)}{n}}}}{1+e^{-c\sqrt{\frac{lp_0(1-p_0)}{n}}}} \leq B(x_k, O_{k-1}) \leq \frac{e^{c\sqrt{\frac{lp_0(1-p_0)}{n}}}}{1+e^{c\sqrt{\frac{lp_0(1-p_0)}{n}}}}$, then the process might be in an out-of-control state.

Since control charts for attributes are most frequently use to detect an increase in fraction non-conforming or a deterioration in quality, the optimization design considers only increasing the p shift and, therefore, only the upper control limit UCL of the p chart is active. A decreasing p shift may be caused by a real improvement in product quality, or a malfunction of the instrument, or a mistake made by the operator. Although a decreasing p shift may send a misleading signal, it does not harm the product quality directly. In fact the lower control limit LCL must be set at zero for most of the cases, otherwise the false alarm rate is intolerably high. Making (LCL = 0) is equivalent to abandoning the lower control limit [4]. In this research one assumes $B(O_0)$ is greater than 0.5 which means that the probability of being in an out-of-control state of process is more than 50 percent. Since it is assumed $B(\theta)$ is greater than 0.5, the power and capability of the presented chart is mostly on determination of an increase in fraction non-conforming. Consequently one determines parameters $B(\theta)$, c and l in such a way that the desired ARL_0 and ARL_1 are achieved.

4. EXPERTIMENTAL RESULTS

VB 7 software is used to compare methods. The presented method is compared to some other methods developed in the literature.

The selected methods for comparison are binomial EWMA (for $\lambda = 0.02, 0.05, 0.08$ and 0.10), moving average (for $w = 2, 3$ and 4) and standard p control chart. It is noticeable that used values for parameters of the binomial EWMA, and moving average in comparison are similar what Khoo [14] has done in his research. Also, the values of c , l and $B(\theta)$ are selected based on the acquisition of good results for ARL_0 and ARL_1 .

To check the validity of the method, two independent binomial distribution with parameters $n=200$ and $p=0.2, p=0.1$ are generated. In the next step, using equation (7), one updates the beliefs $B(x_k, O_{k-1})$. When $B(x_k, O_{k-1})$ is out of interval (9), then an out-of-control signal is observed.

For comparison study, the ARL_1 values are calculated for all the considered methods by 10000 independent replications ($M=10000$) for various values of p_1 (different values of p_1 are indicated in the first column of Table 1 and 2).

The simulation results from Tables 1 and 2 and also Figure 1 and 2, clearly show that the proposed method is better than binomial EWMA; moving average and standard control p chart for shifts of small magnitude from the target value, p_0 .

Based on the results shown in Tables 1 and 2, the following results are observed:

1. We assume $L=150$ and $C=1.5$ in both tables. Based on this assumption, $B(\theta)$ is calculated as 0.569 and 0.592 in tables 1 and 2 respectively in order to achieve the desired ARL_0 and ARL_1 .
2. Calculated values of $B(\theta)$ in tables 1 and 2, mean that the process is in an out-of-control state with the probability of 0.569 and 0.592 respectively. Therefore the presented control chart is used to detect increases in fraction non-conforming.
3. Since it is assumed that $B(\theta) > 0.5$, the use of this method is for evaluation of the initial setup of a process. This application is useful for processes with small production batch. For these kinds of processes, the accuracy of initial setup is important. For instance, short run processes are good cases.

4. This method is so effective for control of production process in which recognition of small deviations is important like high-tech processes.

In general, ARL_1 of the Bayesian approach is better than other methods. Therefore, the proposed approach offers an improved performance.

P_1	Binomial EWMA				Moving average			P chart	Bayesian chart
	$\lambda=0.02$ A=2.1275	$\lambda=0.05$ A=2.652	$\lambda=0.08$ A=2.7650	$\lambda=0.1$ A=2.8155	W=2	W=3	W=4	C=3	L=250 C=1.5 B(0)=0.569
0.100	390.1	585.3	537.7	499.5	496.6	549.7	556.7	302	396
0.105	70.7	93.3	101.9	110.9	213.4	208	194.7	161.6	47.8
0.110	30.1	32	32.7	33.7	92.3	77.1	64.3	84.1	22.9
0.115	18.9	18.1	17.2	17.1	43	34.9	28.4	46.9	15.6
0.120	13.7	12.4	11.5	11.2	23.1	17.9	14.9	27.7	11.4
0.130	8.9	7.8	6.9	6.6	8.3	6.7	6	11.2	7.2
0.140	6.6	5.7	5	4.7	4.3	3.7	3.5	5.6	5.4
0.150	5.3	4.5	3.9	3.7	2.6	2.5	2.4	3.3	4.2
0.160	4.5	3.8	3.3	3.1	1.9	1.9	1.8	2.2	3.4

Table 1: ARL profiles for the p chart, the moving average, the binomial EWMA and the Bayesian chart based on $p_0 = 0.10$ and $n = 200$, $M=10000$ computed by means of a simulation.

P_1	Binomial EWMA				Moving average			P chart	Bayesian chart
	$\lambda=0.02$ A=2.1275	$\lambda=0.05$ A=2.6150	$\lambda=0.08$ A=2.7650	$\lambda=0.1$ A=2.8400	W=2	W=3	W=4	C=3	L=250 C=1.5 B(0)=0.592
0.200	359	496.8	490.6	493.7	475.2	460.3	418.1	284.2	389.5
0.205	100.8	133.4	154.8	169.2	297.4	257.3	239.4	201.8	63.2
0.210	42.8	48.2	53.5	58.1	158.6	127.6	108.8	129	30.4
0.215	26.2	26.4	27	28.6	84	63.2	51.2	82.6	20.6
0.220	18.8	17.7	17.1	17.5	47.9	34.8	28.4	53.5	15.3
0.230	12	10.6	9.9	9.7	18.3	13.2	11	24.4	9.8
0.240	8.9	7.6	6.9	6.7	8.8	6.7	5.9	12.5	7.3
0.250	7.1	6	5.3	5.1	4.9	4	3.7	6.9	5.6
0.260	5.9	4.9	4.4	4.1	3.2	2.8	2.7	4.4	4.5

Table 2. ARL profiles for the p chart, the moving average, the binomial EWMA and the Bayesian chart based on $p_0 = 0.20$ and $n = 200$, $M=10000$ computed by means of a simulation.

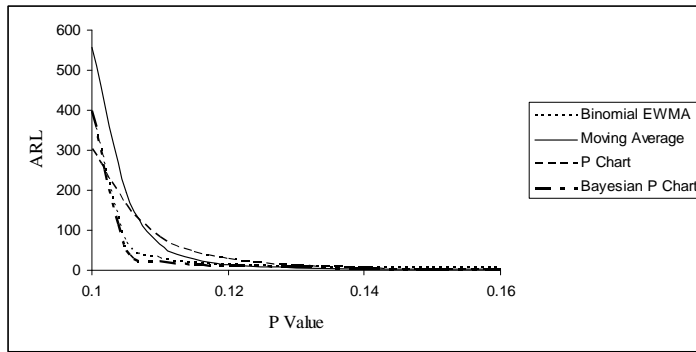


Figure 1: Comparison between the Binomial EWMA, Moving Average, p Chart and Bayesianchart for $p_0=0.1$

(Results of EWMA and MA are the best results gained by Table 1).

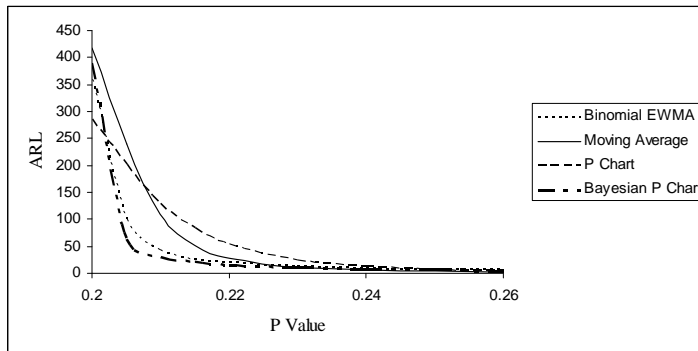


Figure 2: Comparison between the Binomial EWMA, Moving Average and Bayesianchart for $p_0=0.2$

(Results of EWMA and MA are the best results gained by Table 2).

5. CONCLUSION

In this research, Bayesian inference was applied to detect the out-of-control state in attribute control charts and since this approach analyzes data sequentially, it has indicated good performance. It has been found that, generally, the Bayesian p charts can improve the effectiveness for detecting shifts in p to a substantial degree (especially for small shifts in p) without increasing the false alarm rate. Although the cost of running the Bayesian p charts is relatively high, the use of these charts can be justified by the significant improvement in performance. In general, the proposed method yields improvement in ARL_1 for small deviations of the process.

For future research, considering other functions to define beliefs and economic design of parameters of the approach is proposed.

6. REFERENCES

- [1] Vargas, V. Lopes, L. Souza, A. 2004. Comparative study of the performance of the CuSum and EWMA control charts, Computers & Industrial Engineering, 46, pp 707-724.

- [2] Zhang, S. Wu, Z. 2005. Designs of control charts with supplementary runs rules, *Computers & Industrial Engineering*, 49, pp 76-97.
- [3] Woodall, W.H. 1997. Control charts based on attribute data: Bibliography and review, *Journal of quality technology*, 29, pp 172-184.
- [4] Reynolds, M. Stoumbos, Z. 1999. A CUSUM chart for monitoring a proportion when inspecting continuously, *Journal of quality technology*, 31, pp 87-109.
- [5] Roberts, S.W. 1959. Control chart tests based on geometric moving averages, *Technometrics*, 1, pp 239-250.
- [6] Montgomery, D.C. 2002. *Introduction to Statistical Quality Control*, Wiley: New York (2002).
- [7] Hunter, J.S. 1986. The exponentially weighted moving average, *Journal of quality technology*, 18, pp 203-210.
- [8] Crowder, S.V. 1987a. Computation of ARL for combined individual measurement and moving range charts, *Journal of quality technology*, 19, pp 98-102.
- [9] Crowder, S.V. 1987b. A simple method for studying run-length distributions of exponentially weighted moving average charts, *Technometrics*, 29, pp 401-407.
- [10] Lucas, J.M. Saccucci, M.S. 1990. Exponentially weighted moving average control schemes: properties and enhancements, *Technometrics*, 32, pp 1-29.
- [11] Gan, F.F. 1990. Monitoring Poisson observations using modified exponentially weighted moving average control charts, *Communication in statistics*, 19, pp 103-124.
- [12] Borrór, C.M. Champ, C.W. Rigdon, S.E. 1998. Poisson EWMA control charts, *Journal of quality technology*, 30, pp 352-36.
- [13] Somerville, S.E. Montgomery, D.C. Runger, G.C. 2002. Filtering and smoothing methods for mixed particle count distributions, *International Journal of Production Research*, 40, pp 2991-3013.
- [14] Khoo, M. 2004. A Moving Average Control Chart for Monitoring the Fraction of non-conforming, *Quality and reliability engineering*, 20, pp 617-635.
- [15] Nenes, G. Tagaras, G. 2007. The economically designed two-sided Bayesian \bar{x} control chart, *European journal of operational research*, 183, pp 263-277.
- [16] Calabrese, J. 1995. Bayesian Process Control for Attributes, *Management science* 41: pp 637-645.
- [17] Marcellus R. 2007. Bayesian Monitoring to Detect a Shift in Process Mean, *quality and reliability engineering* 2007; DOI: 10.1002/qre.895.