A SEQUENTIAL BAYESIAN CUMULATIVE CONFORMANCE CONTROL CHART FOR HIGH YIELD PROCESSES

A Dissertation by

Gracia Toubia

Master of Science, Oregon State University, 1990

Master of Science, Wichita State University, 1984

Bachelor of Arts, Wichita State University, 1979

Submitted to the Department of Industrial and Manufacturing Engineering
and the faculty of Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

May 2009

A SEQUENTIAL BAYESIAN CUMULATIVE CONFORMANCE CONTROL CHART FOR HIGH YIELD PROCESSES

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy with a major in Industrial Engineering.

Haitao Liao, Committee Chair

Janet Twomey, Committee Co-Chair

Abu Masud, Committee Member

Bayram Yildirim, Committee Member

Dharam Chopra, Committee Member

Accepted for the College of Engineering

Zulma Toro-Ramos, Dean

Accepted for the Graduate School

J. David McDonald

Associate Provost for Research and Dean

of the Graduate School

DEDICATION

To my husband Greg, and children Rami and Maya

No knowledge is worthwhile unless it carries in it the seeds of compassion.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Haitao Liao for his relentless guidance and support and for his generosity in sharing his invaluable knowledge. I also would like to thank my co-advisor, Dr. Janet Twomey for always being there to help and for her crisp and practical mind. Thanks to Dr. Abu Masud for his quiet and steady support and to Dr. Bayram Yildirim for his time and support. Last but not least, I am indebted to Dr. Darham Chopra for his years of mentoring. His knowledge, wisdom, and gentle soul have left their imprint on me.

ABSTRACT

For mass production, the traditional approach of drawing the confidence limits was satisfactory when many bad units were being produced; however, the same approach becomes ineffective for high yield processes, when the defective rates have small magnitudes, such as parts per million. In particular, the traditional *p*-control charts used to control nonconformance rates present many problems when *p*, the rate of nonconformance, is small. One problem is an increase in the probability of false alarm, and another is the increase in sample size, sometimes making the extremely large size of the sample prohibitive. The biggest problem with the *p*-control chart is discussed by Brown, Cai and DasGupta (2001), where the theory of the normal approximation to the binomial distribution is debunked and the need for rewriting the chapter on binomial distribution in statistics textbooks is suggested. Many rejoinders to this work also agree on this important discussion.

An alternative to the *p*-control chart is the cumulative conformance control (CCC) chart, which solves most of the problems of the *p*-chart except that the CCC chart is insensitive to small process deterioration and the methodology is non-sequential. The advocates of sequential procedure criticize the major drawbacks of those traditional control charts, including the *p* and CCC control charts, with *phase I* and *phase II* production. They argue that any non-sequential procedure allows for the "double use of the data" in *phase I*, where *phase I* is designed to set the control limits to be used in *phase II* of production. Specifically, in *phase I*, the sampled observations are used to estimate the parameters needed to set the control limits, and these control limits in turn are used to test the same observations. There is a bias imbedded in this approach. Another major criticism to a non-sequential procedure is the need for a predetermined

fixed and large sample size. The following questions therefore arise: (1) How can a large sample size with a very small rate of nonconformance be predetermined? (2) What happens if that particular large sample contains no nonconformance? (3) How should the sample size be enlarged? Another argument in favor of a sequential procedure is made to the importance of the fact that any estimate of *p* in *phase I* of the production has to be an accurate estimate, representative of the whole process.

In this dissertation, it is shown that even a regular sequential procedure, while preserving the benefits of its "sequential" nature, cannot detect process deterioration in a high yield process. Therefore, a sequential Bayesian CCC control chart is designed that is capable of detecting process deterioration. The sequential Bayesian CCC chart is a more capable alternative to the *p*-control chart and to the CCC chart. It offers a significant improvement over a regular sequential procedure, because of its added capability in detecting process deterioration. It also allows for immediate reaction to process deterioration. When an observation falls outside of the control limits, often the observation can be discarded and the process is allowed to continue. However, when a process deterioration becomes visible as soon as the observations are tested, the system can be halted and the cause investigated immediately. The timely reaction to process deterioration is crucial and valuable in terms of the ability of preserving the "health of the system."

The approach where each additional observation is viewed as the value of a random variable from a distribution updated by that additional observation is powerful in concept. The updating becomes more refined with the flow of information. Using the observations as values of random variables instead of parameter estimates incorporates not just the observation values into the procedure, but also the variability among the observations. This incorporated variability is essential in detecting process deterioration.

In a high yield environment, when the cost of the initial sample size used for the layout of the control chart is either prohibitive or simply unavailable, one is compelled to make good use of every observation given. The Bayesian methodology developed in this dissertation is tested and works satisfactorily for a range of values of probability of nonconformance p, from 0.1 to 0.00001. The methodology may also be used successfully in such circumstances calling for self-starting mechanisms and/or requiring short production runs.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

Chapter Page

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ANIS	Average number of items sampled until a signal is issued
AR	Alarm rate
ARL	Average run length
ATS	Average time to signal
BP	Base period
CCC	Cumulative conformance control
CDF	Cumulative distribution function
CL	Control limit(s)
CRL	Conforming run length
CUMSUM	Cumulative Sum
EB	Empirical Bayes
EWMA	Exponentially weighted moving average
FIR	First initial response
IS	Importance sampling
LCL	Lower control limit(s)
MA	Moving average
MC	Monte Carlo
RL	Run length
SPC	Statistical Process Control
STRL	Standard deviation of run length

LIST OF ABBREVIATIONS (Cont.)

UCL	Upper control limit(s)
ppm	parts per million

CHAPTER 1

INTRODUCTION

The increase in the commitment to quality in industry is leading to a significant decrease in the defective rate of a product. The traditional approach of drawing the confidence limits has served its purpose when many bad parts were being produced. However, the same approach becomes ineffective for high yield processes where the defective rate is in small magnitude such as parts per million (ppm). One example of a high yield process is the thermo-sonic wire bonding (Chang and Gan, 2007), where the number of defects is in the range of 116 ppm. Another example of a high yield process is in the manufacture of integrated circuits as stated by Pesotchinsky (1987).

In this dissertation, a sequential Bayesian cumulative conformance control (CCC) chart is developed to address the deficiencies of past methodologies in environments of minute defective rates. In past methodologies, the low rate of nonconformance p is treated as a fixed parameter. In the approach advanced here, the rate of nonconformance p is treated as a random variable to incorporate variability and uncertainty in the estimation of the rate of nonconformance p.

For the study of high yield processes, the methodology calls for 100 percent inspection on attribute data. Some of the operational advantages of control charts based on attribute data are advocated by Bourke (1991). He advances the proposition that the distributional assumptions of the variables in question become unnecessary when attribute data is used and the gauging instruments of "go no-go" type speed up the inspection process. Sometimes a quantitative inspection is carried out before an overall categorization of conforming-nonconforming is made. Glushkovsky (1994) states that the statistical process control (SPC) methods for attribute data should be applied only for low quality processes and that for high yield processes it is necessary to do 100 percent inspection to be able to ensure a large sample size and proper SPC functioning. He also addresses an important dilemma in high yield processes when there is disparity between the lot size and required sample size. What should be done, for example, if the production lot is five hundred units and the sample size should be at least 1000?

One of the main classical approaches used in charting nonconformance rates is through the use of the p-control chart based on the binomial distribution. The standard confidence interval for the binomial proportion p which is the basis of the p-control chart is simple in concept and easy to calculate. It is based on some heuristic justification concept of the central limit theorem which states that the larger the sample size n is, the better the normal approximation is and correspondingly, the closer the actual coverage of the confidence interval would be to the nominal level. The statistics books lead one to believe that the coverage probabilities of this method are close to the nominal value except when n is "small" and/or p (the rate of nonconformance) is close to zero or one [1-5].

A very important finding by Brown, Cai and DasGupta (2001) debunks the theory behind the confidence limits of binomial proportion (one that is the basis of the p-control chart which is still widely used in statistical process control for fraction defectives) and addresses the very unpredictable and problematic behavior of the binomial distribution. Brown et al. compare the coverage probabilities of a binomial distribution for different values of p and n against their set nominal confidence interval levels. Their findings are disturbing. The concept of a "lucky n" and "lucky p" are discussed, where a good confidence interval that is close to the nominal interval for a particular n becomes poor for a larger n value. Moreover, the chaotic coverage of p is found almost everywhere its value could be, not just around the p boundaries of zero and one as would be suspected. The stated cause is due primarily to the fact that the binomial distribution is discrete and skewed.

The article is followed by rejoinders and positive comments from statisticians with suggestions, including, a reevaluation of the way binomial distributions are taught in statistics books. Some commenting authors suggest, as one solid alternative, the use of Jeffreys interval, for both cases when $n \leq 40$ and $n > 40$.

In particular, the Jeffreys interval is based on a non-sequential Bayesian methodology, where the prior is "non-informative" (nothing is known a priori about the parameter values) and the parameters of the prior are fixed and both equal to ½. In the sequential Bayesian methodology proposed in this dissertation, the prior is "informative" (something is known a priori about the parameter values) and the data cause changes in the parameter values. In this proposed sequential Bayesian approach, the dynamic flow of new information leads to continuously updated parameter values which affect the prior, the posterior , and the predictive distribution where the corresponding CL are constructed from.

There are some other obvious obstacles that are encountered with the *p*-control chart in high yield environments, primarily due to the inability of the *p*-chart to detect process shift and to the increase in the rate of false alarms. There are also problems with the significant increase in sample size and the negative lower control limit.

In order to deal with the problems associated with the *p*-control chart when the rate of nonconformance *p* is very small, an alternative chart called the cumulative conformance control (CCC) chart is introduced by Calvin(1983). It is classical in nature in the sense that it is non-sequential and requires two phases: *phase I* for setting up the control limits and *phase II* for the testing of new observations. The CCC methodology solves most of the problems generated from the *p*-control chart when the rate of nonconformance *p* is small (Kaminsky et al., 1992; Glushkovsky, 1994; Nelson, 1994; Woodall, 1997; and Xie and Goh, 1992, 1997) except that embedded in the CCC chart is an inability to detect small process deterioration (Chan and Goh, 1997; Xie and Goh, 1992; Xie et al, 2000; Tang and Cheong, 2004; Glushkovsky, 1994; Kaminsky et al., 1992 and Yang et al., 2002).

Montgomery (2001) describes how the traditional control charts with the two-phase approach are set up using pre-determined sample sizes in *phase I* of production. The sampled observations are used to estimate the parameters needed to set the control limits. Then the same observations are checked against the control limits, and any observation that is out of control is discarded. The process is then checked for special cause variation and the source of the variation, if any, is eliminated. The remaining observations are used to set up the new control limits. The process continues until all the remaining observations fall within the control limits. Bayarri and Garcia-Donato (2005) note that in the *phase I* period (which is referred to as base plan or BP), any non-sequential approach to finding the control limits based on the heuristic approach described by Montgomery (2001) causes "the double use" of the data. In other words, any non-sequential approach uses observations to compute the control limits and then uses the control limits to test the same observations used to set these limits. This is hardly an unbiased approach. In contrast, the sequential Bayesian procedure introduced in this dissertation uses sequentially updated data to set new control limits and the observations used to set the control limits are not the same as the observations being tested.

Besides the unbiased nature of the sequential methodology, there are other very important practical advantages to any sequential procedure. It can be used when the amount of data is limited, valuable or unavailable. It can be applied in short run processes and as self-starting mechanisms.

In this dissertation, a sequential Bayesian procedure is developed based on the CCC chart. The methodology, given its sequential nature, solves the bias problem of non-sequential approaches posed by Bayarri et al. (2005). Also, because it relies on the distribution theory of the CCC chart (i.e., the geometric distribution), it avoids the fundamental problems of the binomial distribution addressed by Brown et al. (2001) and incorporates the advantages of the CCC chart. Moreover, because of its Bayesian in nature (i.e., incorporating variability in the nonconformance estimate of p), the methodology can detect process deterioration whereas a simple sequential CCC procedure fails.

In addition to the sequential Bayesian procedure developed in this dissertation, a simple sequential CCC procedure is introduced and used as a benchmark for the sequential Bayesian methodology. Furthermore, the performance of the CCC procedure, the simple sequential and the sequential Bayesian procedures are compared. The dissertation shows that although both the simple sequential and sequential Bayesian procedures are superior in their ability to detect an "out-of-control" observation compared to the CCC procedure, the sequential Bayesian methodology is much better at detecting process deterioration than the simple sequential procedure. The sequential Bayesian methodology is superior to the sequential methodology (which smoothes the variability since it relies only on averages - behaving more like a moving average) because it is the only one that incorporates the knowledge of variability among observations and therefore mirrors their behavior when they are deteriorating.

The sequential Bayesian methodology developed in this dissertation allows for the construction of confidence limits even when only originally three observations are given.[1] These observations are used at the start of the process to define the parameters of the prior and the predictive distributions, and from the predictive distribution, the original control limits (CL) are constructed. If past knowledge or data is available, that information can be used to define the prior distribution parameters and no observations are required to set the original control limits. A new observation is then taken and tested against the original control limits. If the observation falls inside the limits, that observation is used to update the prior and set the next confidence limits. By conditioning on the present observation and drawing the control limits for the next observation, the posterior distribution at index i becomes the prior distribution at index $i+1$. If and when an observation falls outside the "in-control" interval, the observation is discarded and the previous observations are used to compute a new prior that would lead to more accurate estimates of the control limits. In this new sifting approach, the good observations are not shuffled back into the mix for redrawing new control charts as is done in the *phase I* of any classical control chart approach. The prior is used as a "counter" to account for the time order of past observations and a "sifter" to separate the "in-control" observations from the "out-of control" ones.

Furthermore, it is also shown in this dissertation that not only is the Bayesian methodology sensitive in detecting process deterioration when the rate of nonconformance p is very small, but it is also just as powerful when the rate of nonconformance is "not small". It is in effect applicable for a wide range of values of nonconformance rates p. The tested values in this dissertation are from 10^{-1} to 10^{-5}.

[1] An observation refers to the number of conforming units up to the nonconforming one.

The proposed sequential Bayesian procedure has never been introduced in the literature. However, different benefits of the sequential Bayesian procedure have been mentioned in the literature under some proposed methodology when dealing with the different problems generated from high yield processes. The power of the proposed approach is that it incorporates all of the proposed remedies under one methodology while avoiding their drawbacks.

CHAPTER 2

LITERATURE REVIEW

A comprehensive literature review is presented in this chapter. The first section introduces the traditional *p*-control charts [6-7]. The second section identifies the problems with the traditional charts when the rate of nonconformance *p* is very small. The third section discusses techniques that can be used to handle such high yield processes, including transformations, run- rule charts, and the cumulative conformance control (CCC) charts, along with their drawbacks. The fourth section discusses the problems remaining regarding the CCC chart. The fifth section describes some problems with estimating the probability of nonconformance *p* during *phase I* of production. The sixth section discusses further different solutions to overcome the drawbacks of the methodologies introduced, including: Cumulative sum (CUMSUM) charts, synthetic charts, conditional CCC charts and sequential CCC charts. The last section introduces the motivation behind the Bayesian control chart proposed in this dissertation.

2.1 The *p*-control chart

The traditional Shewhart control charts have been designed to essentially increase quality and productivity. The *p*-control chart monitors the fraction of defective items. The underlying theory behind the *p*-control chart is that the random variable *x* representing the number of defectives in a sample of size *n* follows a binomial distribution. Based on the central limit theorem, the observations selected from a binomial distribution can be approximated by a normal distribution with mean and standard deviation estimated by $\overline{p}$ and $\sqrt{\frac{\overline{p}(1-\overline{p})}{n}}$, respectively, where n is the number of items in the subgroup, and $\overline{p}$ represents the average fraction of defects of a preliminary number of samples inspected (or from past production data). This approximation is used to determine the upper and lower control limits of the p-control chart. The probability of a nonconformance value of p from a sample falling above the upper control limit (UCL) or below the lower control limits (LCL) is very small (.0027 for the 6 sigma levels), and whenever that happens it probably means that the process has deviated from its process mean and should be investigated. A point above the UCL translates into a highly defective sample, and a point below the LCL translates into a higher quality sample; both cases not likely to happen due to chance alone. Guidelines on when to use a normal approximation to the binomial distribution have been set in many statistical books: np and $nq >= 5$, n and nq large; np and nq both $>=10$, np and $nq >5$, $n \geq 25$ and $.2 \leq p \leq .8$. Here p is the fraction nonconforming, n the sample size and $q = (1\text{-}p)$.

In Figure 1, a classical p-control chart is exhibited with UCL = 0.0066, LCL = -0.020 (set to 0) and an average nonconformance value of p at 0.0023. In percentage terms they are 0.66, 0 and 0.23, respectively. In general, at least 20 observations or subgroups are needed to set the control limits during *phase I* of the production process. If Figure 1 represents the setting of the control limits for *phase I* of the production process, then observations (subgroups) 3 and 13 have to be thrown out, and the control limits recomputed with the rest of the data points. Here observations were generated with an average subgroup size of n=2250. Note that since the graph becomes somewhat stable after point 14, provisions should be made to substantially increase the subgroup size.

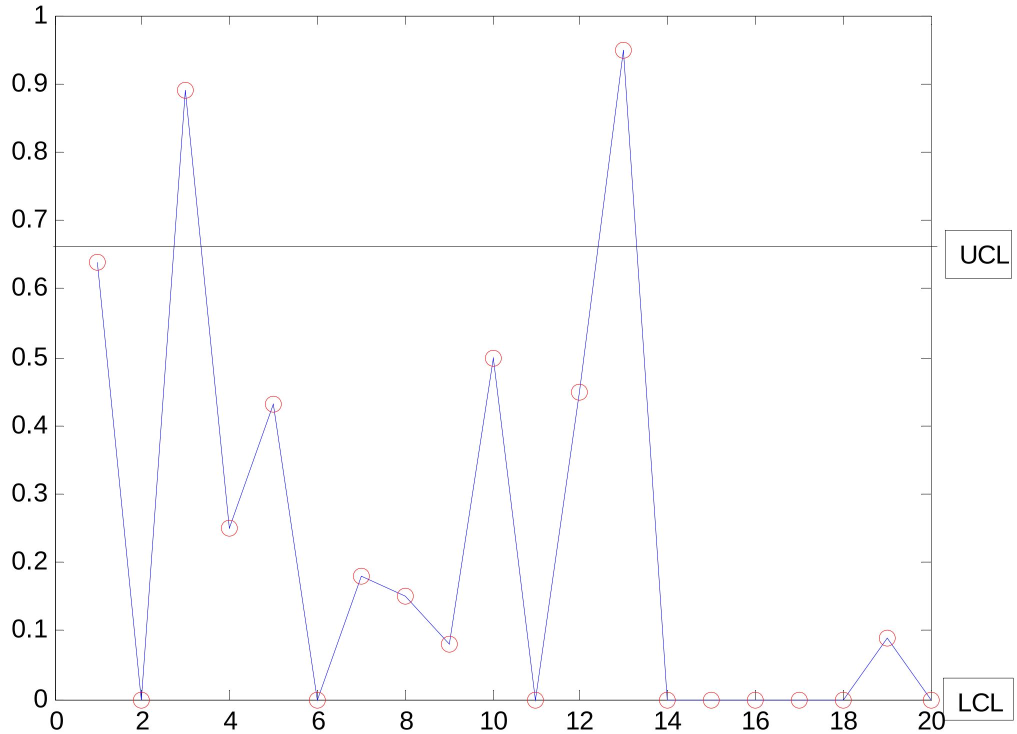


Figure 1. A traditional *p*-control chart

2.2 Problems with the traditional *p*-control chart when *p* is small

Notwithstanding the repudiation by Brown et al. (2001) of the traditional process used in establishing the *p*-control chart, other concerns have been addressed in the literature regarding the *p*-control chart when *p* is small. These concerns are discussed next.

2.2.1 Negative lower control limit

For a small rate of defects or nonconformance, *p*, and a sample size that is not "very large," very often, the LCL dips below zero. Goh (1987a) shows that for an exceedingly small value of *p* (e.g., $p = .0004$, an average of 400 ppm nonconforming), and an *n* value of 200 (the sample size of each inspection), the lower control limit becomes negative. Another paper by the same author (1987b) shows that for n = 100 and p = .006, the LCL is also negative. Usually, when the LCL becomes negative, it is set to zero because a negative control limit is meaningless for a p-control chart, as p must be greater or equal to zero. The author argues that this renders the control limits asymmetrical about the central line, something not in line with the normal approximation theory behind the control charts, essentially violating the theory.

Xie and Goh (1993) argue that not only is the LCL not reliable due to a violation of the theory behind it, but that the LCL contains information about the process improvement, information not less important than process deterioration from a management point of view. Setting the LCL to zero translates into a missed opportunity in the ability of identifying and therefore sustaining a continuous improvement of the process. They suggest an np value larger than 25, for the normal approximation of the binomial to work. Furthermore, Chan and Goh (1997) propose that when the nonconforming rate is in the order of 10^{-4} or lower, the use of the p or np chart is no longer acceptable.

2.2.2 Increase in the probability of false alarm

Another problem with the p-control chart in dealing with high yield processes is the high probability of false alarm. Goh (1987a) shows that one nonconforming item in a sample will throw the process out of control, which means that the process has to have practically zero defect to be in control. This translates into an increase in the false alarm probability. Chan et al. (2002) argue that when the rate of nonconformance is very low, the UCL of a p-control chart is usually a fraction less than $1/n$, and, because of the nature of the discrete behavior, will give a false alarm every time a nonconforming item is observed even though the process is in control.

2.2.3 Increase in the sample size

The sample size required in detecting process shift is also extremely important to high yield processes. Glushkovsky (1994) shows that by setting the lower control limit value to 1, the smallest possible value it can take, and the *α/2* value (the rate of false alarm) at the common value of .00135 for the $\pm\ 3\sigma$ control limits, then for a nonconformance value of p value at 0.01, the sample size must be greater than 660. For a nonconformance value of p in the magnitude of 100 ppm the sample size must be greater than 66,000. Even when it is possible to indulge in very large samples, Chang and Gan (2007, p. 871) note: “A charting scheme based on a large sample size may not provide a timely feedback on the ‘health’ of a process being monitored if the inspection of quality is not swift enough.”

.

2.2.4 Inability to detect a process shift

The ability to detect process shift in a timely fashion is an important feature of a control chart. Goh (1987a) shows that the p-chart with low nonconformance rate p does not respond fast enough to a process shift. Western Electric Rules are commonly used to detect process shift in a classical Shewhart chart such as the p-control chart [8]. However, because the rules are based on the assumption that the distribution is symmetric and unimodal, they cannot be applied in our case, given the fact that the binomial distribution is highly skewed and the normal approximation of the binomial fails to apply, according to the findings of Brown et al.(2001) discussed in the introduction section of this dissertation.

2.3 Suggestions for coping with high yield processes

2.3.1 Run rules chart

As an alternative to the *p*-control chart,, Goh (1987a) proposes a special chart for low rate of nonconformance *p* called the "low *p* chart" where a pattern of different values of p_i from consecutive samples are used instead of an individual conformance value *p* value from one sample. The pattern is then evaluated to check whether or not a process is out of control. The pattern essentially can be formulated as follows: If *m* out of *M* consecutive samples have *d* defectives, then the process is out of control. The author suggests its use whenever the LCL is negative and when the nonconformance value of *p* is low and the *n* value is not large enough, rendering infeasible the normal approximation to the binomial. The actual binomial distribution is used instead of its normal approximation. The approach does not assume 100% inspection. Instead, an assumption is made about the belief in the stability of the process. The approach involves a complex reasoning in statistical theory and is impractical to use in many applications. Furthermore, it has not been tested for cases with very low rates of nonconformance, such as the ones in the parts per million. The author suggests its use when the nonconformance value *p* is less than one percent. It was tested at a minimum nonconformance value *p* of .006.

2.3.2 Transformation of distributions

Nelson (1994) suggests a transformation of distributions. He assumes that with a constant production rate, counting the number of units produced can be substituted by time. If the time between the production of nonconforming items is exponentially distributed, then a simple transformation from an exponential into a Weibull distribution with equal mean and median would render the new distribution close to the normal one and therefore further justify the control chart. The problem with this methodology is that the technique is difficult to implement and the plotted points are hard to interpret. Ryan and Schwertman (1997) suggest transforming the data using the arcsine function to get a closer approximation to the normal distribution. They also indicate that the sample size required to produce a non-zero LCL after a transformation is performed is smaller than the sample size required by just using a normal approximation approach. Queensberry (1991) developed a "Q" chart for the binomial parameter p, where the normalized transformations are utilized in the p-control chart. He concludes that the Q chart provides a better approximation to the nominal upper tail area whereas the arcsine approach provides a better approximation to the nominal lower tail area.

Ryan and Schwertman (1997) state that the problems of the Q control chart and the arcsine transformations are practically the same as the traditional p-control chart when the nonconformance rate p is very small, such as in the range of parts per million. They argue that both the arcsine transformation and the Q chart approaches lead to an LCL of zero for any realistic sample size value of n. Therefore, they suggest doing a computer search to determine the value of the sample size n when the nominal tail areas are closely approaching. They also suggest using an approach given by Nelson (1994), which is not different in concept than the approach used by Calvin (1983) when he first introduce the CCC chart.

Chang and Gan (2007) make the point that all distributions based on transformations work on the premise that the transformations are a better approximation of the normal distributions. However, they argue against the use of transformations since: "The main disadvantage of using transformed statistics is that they are much harder to interpret. Besides, a transformed statistic may not be an optimal statistic for the parameter of interest, and thus there will be some loss of detection capability" (p. 859.)

2.3.3 Modified Shewhart chart

Bourke (1991) suggests using a modified Shewhart control chart based on the sum of the two most recent run lengths to increase the sensitivity of detecting large shifts in nonconformance rate. By making better use of "the time order of the observation" which a regular *p*-chart does not perform, a relatively more sensitive test is generated. The tests use a minimum value of nonconformance rate p in the order of 5×10^{-4}, and a value of α (probability of Type I error) at .02

2.3.4 Use of the exact binomial distribution

Xie and Goh (1993) suggest using the exact binomial distribution in determining the UCL and LCL but even they concede that it is not appropriate to consider the $\pm 3\sigma$ control limits when p is small, n is low and one sample is drawn. As an effective alternative solution, the cumulative conformance control (CCC) chart is suggested by the authors in this case. Next, a brief introduction to CCC chart and a detailed literature review are provided.

.

2.3.5 Cumulative conformance control (CCC) chart

The CCC chart (denoted sometimes by the *G* chart) has been shown to be an effective alternative to the *p*-control chart when dealing with high yield processes.

Calvin (1983) was the first to introduce the CCC chart. He suggests concentrating on the number of good parts between two bad ones and plotting them cumulatively on a chart with logarithmic vertical scale and a linear horizontal scale. Goh (1987a) then reintroduces the concept of the CCC chart and highlights its importance by pointing out the problems arising when nonconformance is in the hundreds ppm and the ability of the CCC chart to solve those problems.

Gluskowsky (1994) uses the CCC control chart concept but calls it a *G*-chart, where the observed variable *G* is the number of conforming units between two nonconforming ones. He compares the sensitivity of the *G*-chart with the classical *p*-chart by constructing average run length (ARL) curves and finds that the sensitivity is higher in the geometric chart. He concludes that the *G*-chart can be applied in high yield processes with an estimated rate of nonconformance value p being less than 10^{-4} (100 ppm), low volume manufacturing, short runs and stepped processes where an attribute is “good” up to a certain moment and then becomes continuously “bad”, such as with chemical and thermodynamical processes and with wear and tear of instruments. He also concludes that the *G*-charts require fewer observations than the classical *p*-chart for process shift detection. He states that: “Compared with the classical *p*-chart of the same sensitivity, the *G*-chart provides more *dynamic, quicker response*, because on the average fewer units are to be observed between the process change and this change detection.”(p.20). He talks about the optimal “subgroup” size for a required degree of detection or a sensitivity level, but his sample sizes used in the analyses are in the thousands (up to 100,000). Again this is due to the low nonconformance value of p in a high yield process.

Goh and Xie (1995) divide processes into moderate, low and near zero nonconformity processes and suggest an approach for each case: exact probability limits in the first case, pattern recognition or run rules approach in the second case and a CCC approach in the third case.

2.3.5.1 Mathematical approach to the CCC chart

The CCC chart is based on the fact that the probability that the x^{th} item is found nonconforming after the first $x-1$ conforming ones follows a geometric distribution with probability:

$$f(x) = (1-p)^{x-1}p, \qquad 0 \leq p \leq 1, \qquad x=1, 2, 3 \qquad (1)$$

where x is taking on discrete values greater than zero and p is the probability of a nonconforming item. The cumulative distribution function is then:

$$\text{F}(x) = \sum_{i=1}^{x}(1-p)^{i-1}p \qquad (2)$$

This cumulative distribution function is a geometric series converging to $1-(1-p)^x$. Calvin (1983) argues that $(1-p)^x$ can be expanded into:

$$(1-p)^x = [1 - xp + \frac{x(x-1)}{2!}p^2 - \ldots] \qquad (3)$$

and p replaced by $\frac{1}{\bar{x}}$, where $\bar{x}$ is the average run length, and for large x, F(x) can be approximated by $1- e^{\frac{-x}{\bar{x}}}$. Solving for x in $F(x)$ yields:

$$x = -\bar{x}\ln(1-F(x)) \qquad (4)$$

The CL is set at the median level, i.e., the value that solves $F(x)=0.5$. The median then is $\tilde{x} = -\bar{x}\ln(1-0.5)$ or $\tilde{x} = .6931\bar{x} \approx .7\bar{x}$

As in any control chart, the UCL and the LCL are defined such that the probability that a point will fall outside the control limits, even when the process is under control, is α. This is the same as equating $F(x)$ to $\frac{\alpha}{2}$ for the LCL and $1-\frac{\alpha}{2}$ for the UCL. Replacing the value of $F(x)$ in equation (4) by $\frac{\alpha}{2}$ and the value of x by LCL yields:

$$LCL = \left\lfloor -\bar{x}\ln(1-\frac{\alpha}{2}) \right\rfloor = \left\lfloor \frac{\alpha}{2}\bar{x} \right\rfloor \quad (5)$$

for small α and

$$UCL = \left\lceil -\bar{x}\ln(1-(1-\frac{\alpha}{2}) \right\rceil = \left\lceil -\bar{x}\ln(\frac{\alpha}{2}) \right\rceil \quad (6)$$

where $\bar{x}$ is the estimate of the parameter p, $\lceil \ \rceil$ and $\lfloor \ \rfloor$ represent the ceiling and the floor of the numerical values.

Another approach leading to very similar results uses the fact that the CDF of the geometric distribution with parameter p from equation (2) is of the form:

$$F(x) = \sum_{i=1}^{x}(1-p)^{i-1}p \quad (7)$$

and

$$P(X<k) = \sum_{i=1}^{k-1}(1-p)^{i-1}p = 1-(1-p)^{k-1} \quad (8)$$

which is the converging form of the geometric series. This translates into:

$$P(X<LCL) = \sum_{i=1}^{LCL-1}(1-p)^{i-1}p \text{ and } 1-(1-p)^{LCL-1} = \frac{\alpha}{2} \quad (9)$$

Solving for LCL yields:

$$LCL = \left\lfloor \frac{\ln(1-\alpha/2)}{\ln(1-p)} \right\rfloor + 1 \quad (10)$$

Using equation (8),

$$P(X>UCL) = \sum_{i=UCL+1}^{\infty}(1-p)^{i-1}p \text{ and } (1-p)^{UCL} = \frac{\alpha}{2} \quad (11)$$

Solving for UCL yields:

$$UCL = \left\lceil \frac{\ln(\alpha/2)}{\ln(1-p)} \right\rceil \quad (12)$$

For a process where on average, it takes 3700 conforming items until seeing a nonconforming one, $\bar{x} = 3700$ and $\hat{p} = .00027$. At an $\alpha/2$ level of .025, the floor of the UCL and LCL are: $UCL = \left\lceil \frac{\ln(\alpha/2)}{\ln(1-p)} \right\rceil = 13{,}661$ and a lower limit $LCL = \left\lfloor \frac{\ln(1-\alpha/2)}{\ln(1-p)} \right\rfloor + 1 = 94$, respectively. Values below 94 means the process is deteriorating and any values above 13,661 means the process is behaving better than expected with a probability of false alarm at .05. Figure 2 illustrates a CCC chart with x_i = 3000, 8000, 90, 5000, 500 and 16,000.

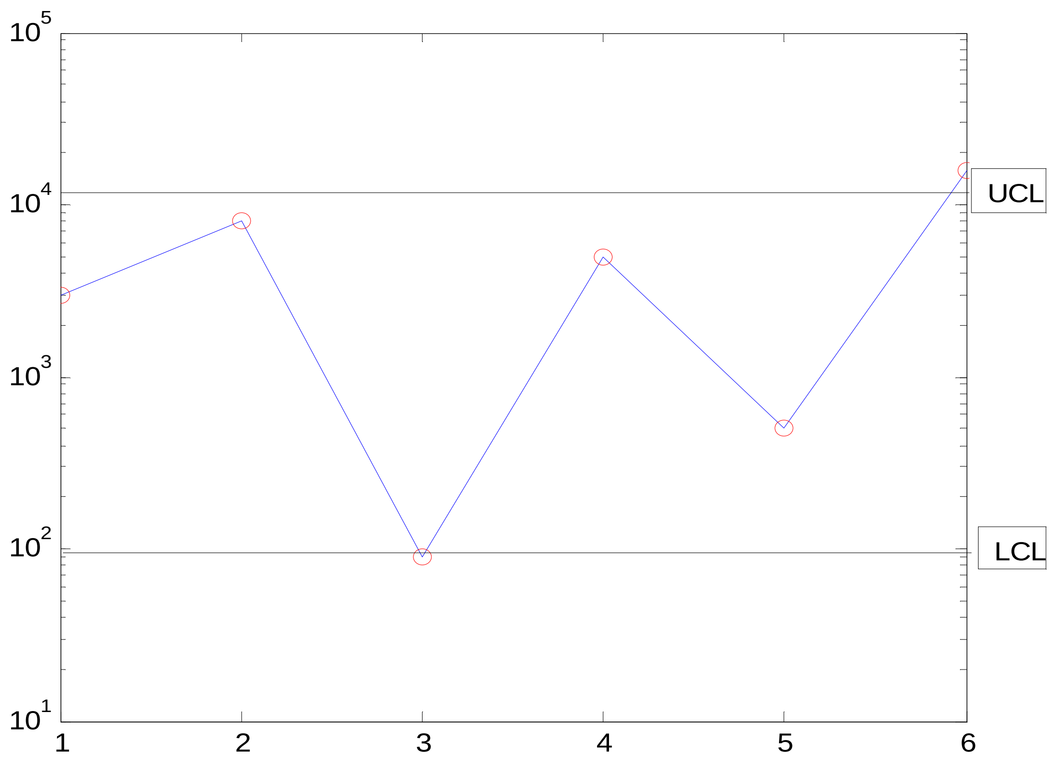


Figure 2. The CCC chart

Note: The index of x, (i= 1, 2, 3, 4, 5, 6) denotes the situation when the i^{th} nonconformance is observed. The value x denotes the number of conformances up to the nonconformance. Whether inspection is done in batches or on the spot, a counter is used to count the number of conformities until the first nonconformity is found and record it as x_1. The value of x_1 can then be charted and out of control situations immediately observed. The counter is then reset at zero and the counting restarts until the second nonconformity is found and x_2 recorded, and so on. One can also note that in the CCC chart, points below the LCL indicate process deterioration and points above the UCL indicate process improvement, contrary to the p–control chart.

2.4 Problems with CCC charts

2.4.1 Inability to detect small process deterioration

Chan and Goh (1997) bring up the fact that the CCC control chart is not sensitive in detecting an upward shift in the rate of nonconformance p. In other words, it is not able to detect process deterioration. An example is provided to bring the problem to light: when the defective rate is increased by 10 fold (from p = .0004 to p = .004), the probability that the process is still "in control" is about 0.6, which is pretty high. To deal with the problem, a two-stage approach is suggested where the number of items inspected until the observance of two nonconforming parts is studied. The first CCC1 control chart is tested with n_1 points and the second CCC2 control chart is tested with n_2 points where the probability of type I error (the probability of false alarm) in the first and second stage is α_1 and α_2 respectively. If one nonconforming item is found within the first n_1 inspected then the process is out of control, if not, then two nonconforming items have to be found in the second phase with $n_2 - n_1$ inspected for the process to be out of control. The value n_2 is expected to be at least twice as large as n_1. Here the CCC2 control chart seems to increase the sensitivity to a change in the rate of nonconformance p. However, since so many inspections are required to reach the second phase, the approach tends to be lengthy, costly and thus impractical.

Xie and Goh (1992) suggest using a decision graph to monitor the levels of the nonconformance rate p. They define s, the level of certainty about the process, which is the probability that a nonconforming part is not found in n samples inspected for a given confidence level α and a given rate of nonconformance p. The value n is the number of conforming parts prior to finding a nonconforming one. A table is developed initially that registers the value of n for different values of p and s. Then the cumulated count of conforming parts is checked against the tabulated value; if it is less than the tabulated one, then the proportion of nonconforming parts is higher than p with probability s. The table is transferred then into a "decision graph." The table also allows finding the minimum number of conforming parts that must be observed before a nonconforming one is encountered when the process is in control for a particular p and s. The landmark of the methodology is that it allows the controller to estimate the rate of nonconformance p for a particular s whenever a nonconforming item is encountered. It also checks whether the value of p has deteriorated or not. The problem with this approach is that the table and decision graph are limited to pre-selected p and s values and would have to be generated again for other sets of values. Moreover, when the number of consecutive conforming items becomes very large (over 2,000 in an example) before a nonconforming item is observed, that item is deemed in control whether it actually is or not, from the simple fact that the process at that point is behaving "much better than expected." Dealing with that situation is "beyond the scope of the present study" state the authors (p. 359).

Yang et. al. (2002) also confirm that the CCC chart is not sensitive to small process deterioration and try to demonstrate the reason for that. They posit that: for a given number of items sampled m, an estimated nonconformance value of p_0 and a desired false alarm rate α , we have estimated control limits; If the nonconformance parameter value shifts from p_0 to p, the interesting finding is the that the alarm rate (AR) increases as p decreases from p_0 but as p increases from p_0, the AR first decreases and then increases very slowly. This confirms that the geometric chart is able to detect a process improvement but inefficient in detecting a process deterioration unless that deterioration is significant. This makes sense intuitively if one can picture the control limits. When the rate of nonconformance p decreases, the plotted number of conformances up to the nonconformance is larger, and therefore tends to exceed the UCL, thus increasing the AR. When p increases, the plotted quantity is smaller and so does not go above the UCL. However,, it still has no chance to go beyond the LCL because the LCL is already very small due to the low rate of nonconformance p_0 used to define the LCL. When the rate of nonconformance p increases significantly, the plotted quantity drops below the LCL and the AR increases. According to the paper, it turns out that the average run length (ARL) measures are relatively more sensitive to both upward and downward shifts in the rate of nonconformance p when using the geometric control charts, but in essence the ARL behavior is similar to the (AR) function. When p decreases from p_0, both the ARL and the standard deviation of the run length (STRL) increase, but when p increases from p_0, both the ARL and the SDRL increase first and then decrease. This is "undesirable and can be explained in the same way as for the AR function." (p.453), state the authors.

Other authors like Xie et al. (2000) also show that the ARL increases at the beginning of process deterioration. An adjustment factor to the control limits is suggested so that the ARL is maximized when the process is at its normal state. The adjustment factor is criticized by Tang and Cheong (2005) because that adjustment factor does not have a probability interpretation. They also state that the *type I* error of the control chart changes after the adjustment is made to the control limits. Moreover, the false alarm probabilities outside the control limits are no longer equal.

Tang and Cheong (2004) also conclude that the CCC chart is better able to detect process improvement than process deterioration unless the rate of increase in p is large. The authors state that this is "a common problem for data with a skewed distribution" (p.845) such as the geometric distribution.

2.4.2 Loss of past information

Xie and Goh (1992) tackle another issue with the CCC control chart: The fact that the CCC procedure resets the cumulative count at zero after a nonconforming item is found and the process is still in control. This translates into a loss of valuable information. Therefore, they propose using probability theory to translate that zero count into a cumulative count, a better representative of the state of the process.

*2.*5 Problems with Estimating *p* during *phase I* of production

So far, in all the articles discussed above, the rate of nonconformance p is assumed to be a quantity accurately estimated from a previous run or from a preset sample size measured during *phase I*, the upstream phase of the production process. This might not be a very realistic assumption and, as is discussed below, finding an accurate estimate of p might not be an easy task.

2.5.1 Sample size

Yang, et al. (2002) bring out a very important argument; that a parameter that must be estimated accurately during *phase I* may require a very large sample that could be impossible to produce.

2.5.2 Bias in the estimate of *p*

A large sample, if not impossible to produce, will introduce bias in the estimate. Tang and Cheong (2004) suggest that when the binomial distribution is used, the false alarm rate is dependent on the true nonconformance value of p. Since p needs to be estimated in the initial production phase, and there might not be any defective items in the initial sample, the sample size is increased on an ad hoc basis, until an arbitrarily number of defective items is observed. This increases the bias in the estimate of p.

2.5.3 Importance of an error free estimate

An error from the nonconformance estimate p will affect the control limit estimates and therefore shifts the process from the true limits. Tang and Cheong (2004) emphasize the importance of a good estimate of p and address the situation when the process parameter p is unknown. They state that "the proposed scheme performs well in detecting process changes, even in comparison with the often utopian situation in which the process parameter, p, is known exactly prior to the start of the CCC chart" (p 841). They design a sequential procedure where, as the index of x (the number of nonconformities up to the conformity) moves from $i=1$ to $i>1$, the distribution becomes negative binomial and control limits are drawn at each sequential iteration i.. This sequential procedure essentially takes care of the bias in p estimated from a fixed sample size during the *phase I* process.

2.6 More suggested solutions

2.6.1 Cumulative sum (CUMSUM) chart

The CUSUM control chart is based on summing cumulatively all the deviations of the sample values from the target values. It incorporates all the values from the previous samples up to the present one. Most CUMSUM applications focus on continuous variables. Lucas (1985) is the first to introduce the CUMSUM concept for attribute data to detect shifts in the probability of nonconformance, p, but not when p is very small. Lucas also introduces the fast initial response (FIR) to be able to detect quickly an out of control situation at the start-up of the process. He also discusses the robust CUMSUM, where the process will differentiate between outliers and process shift, i.e. will quickly detect a shift but will be insensitive to the outliers in the process.

Although there are many methods for applying the CUMSUM theory, the most used one and the "best known" (Xie, Goh and Lu, 1998, p.340) is through the use of the V-mask. The mask is in the shape of a tilted V to the right, and the boundaries of the defined V lines influence the performance of the CUMSUM chart. The theory is involved and the technique is not simple to use. CUMSUM is designed to detect small process shifts. Xie, Goh and Lu (1998) mention that the "CCC technique is more suitable for high quality processes" (p. 339) than the CUMSUM chart. In their paper an attempt is made to incorporate the CCC chart with the CUMSUM methodology. The claim is that the CUMSUM "will be able to detect small process changes relatively quickly. However, some analytical studies are needed to justify the conclusion"(p. 342). Their work on the sensitivity of the CUMSUM chart when compared to the CCC chart through the measure of the ARL indicates that the CUMSUM chart is more sensitive to process shifts, particularly an increase in probability of nonconformance p, i.e., a deterioration in the process. However the value of p is measured at .002 and higher. In addition, the CUMSUM procedure assumes a fixed sample size and no decision can be made until all the items in the sample have been inspected.

Bourke (1991) proposes the use of attribute charts by mentioning some operational advantages for their use. Among them are:1) With the "go-no go" gauging type instruments, the inspection process is usually expedited; 2) The distributional assumptions on the variables observed are not necessary; and 3) Some of the inspection can be done qualitatively before a determination of a conformance or a nonconformance can be made. It is also mentioned that attribute control charts are not as effective in detecting shifts as variable control charts. He introduces the run length (RL) CUMSUM chart based on the lengths of runs of conforming items. RL is defined as the number of inspected units between two nonconforming ones. RL charts are mostly applicable when 100% of the items produced are inspected in the order in which they are produced. The distribution of the RL will change when p shifts. It is shortened as p increases and is lengthened as p decreases. The result shows that the run length CUMSUM chart is somewhat more efficient in detecting shifts in fraction nonconforming than the Poisson CUMSUM chart. This might be attributed to the fact that the Poisson CUMSUM chart "artificially groups the observations, thus losing information associated with the time order of the observations" (p.236).

Reynolds and Stoumbos (2000) show that the Bernoulli CUMSUM chart detects changes in the rate of nonconformance p (increase and decrease) substantially faster than the traditional p-control chart, but the minimum value of nonconformance rate they use is at .01.

Chang and Gan (2001) propose design procedures for CUMSUM charts for high yield processes based on run length (RL). The average number of items sampled until a signal is issued (ANIS) is used to measure the RL of the charts based on the Markov Chain approach. They compare the run length performances of CUMSUM of Bernoulli, binomial and geometric charts where both increases and decreases in the nonconformance value p are tested. They propose an upper sided geometric CUMSUM chart for detecting increases in p. For detecting decreases in p, a two-sided CUMSUM Bernoulli chart is preferred. Between the Bernoulli and the geometric CUMSUM charts, the Bernoulli CUMSUM chart closely monitors the process since every item is charted and therefore reacts quicker than the geometric CUMSUM chart. "On the other hand, a Bernoulli CUMSUM chart requires continuous plotting while the geometric CUMSUM chart does not" (p. 798).

2.6.2 Sequential sampling scheme

Tang and Cheong (2004) propose a sequential sampling scheme for the CCC chart where "every successive observation accumulated to date, adaptively updates the estimate and revises the control limits, at the same time, the results are used for monitoring the process. This means that it is not necessary to assemble a huge number of initial samples before the control begins"(p. 842). To use an estimate of the probability of nonconformance p, a minimum of $m = 2$ samples is needed. This is enough to start using the CCC control chart. Here the control limits are updated every time the cumulative number of conforming items up to the nonconforming one falls within the control limits. The performance of the CCC chart is improving continuously as more data are incorporated. The authors use the adjustment factor on the control limits proposed by Xie, Goh and Kulmarani (2000) to ensure that the ARL chart will detect process deterioration when the rate of nonconformance value *p,* increases. Later, Tang and Cheong (2005) address the issue that the adjustment factor changes the probability interpretation of the control limits, the type I error, and makes the false alarm rates unequal on both ends of the distribution. Zang et al. (2004) propose an adjustment procedure with equal probabilities except that their procedure only considers a known probability of nonconformance value *p* and requires many iterations.

The assumption for the sequential model is that the inspected items are produced independently. This approach is attractive for the following reason: The updating of the estimate means that the process does not depend on an estimated rate of nonconformance *p* in the upstream process, something desirable since, as discussed above, the sampling procedure to find the estimate of *p* could be problematic. To prevent using data from a shifted process in updating the parameter *p*, Tang et al. (2004) suggest that some guidelines be developed for checking on new data before it is incorporated or that the process is allowed to stop after a predetermined amount of time. In essence, the paper solves the problem of having to estimate the rate of nonconformance *p* from a fixed sample size. Also, the sequential updating of the *p* estimate makes it possible to start production right away.

Tang and Cheong (2005) elaborate on when to terminate or abort the sequential process to "avoid using the contamination data from drifted processes" (p. 12). They also suggest that when adopting the conventional non-sequential procedure, one way to avoid the situation where all items in the *phase I* production are all conforming is to ensure the probability of at least one nonconforming item in the sample is pretty large (that usually leads to a very large sample size).

The authors finally state that with the given guidelines on terminating the process, the sequential procedure is just as effective as the one used with known rate of nonconformance p in detecting process shift, except the comparison was tested on a process improvement instead of deterioration.

2.6.3 Conditional CCC chart

Kuralmani et al. (2002) propose a conditional CCC chart for detecting process shifts. A comparison is made for both the CCC and the conditional CCC charts having the same in-control ARL. The conditional CCC is shown to have tighter control limits and an improved sensitivity in detecting process shifts. The proposed process incorporates, in addition to the CCC chart, a run rule based on previous in-control runs. This is like basing the current decision by conditioning on the behavior of the past observations.

The two-stage approach proposed by Chan and Goh (1997) is also designed to improve the sensitivity of detecting a shift and also uses the approach of conditioning on past observations.

2.6.4 Synthetic control chart

Wu et al. (2001) suggest a synthetic control chart for detecting increases in the rate of nonconformance p. The synthetic chart is a combination of the np chart and the CRL (conforming run length) chart. It involves an optimization procedure where users can minimize the out-of control ATS (the average time to signal an out-of control observation) by adjusting the parameters c, n and L, where c is the upper control limit of the np chart, n is the number of units inspected, and L is the lower control limit of the CRL chart. They show that the synthetic chart is more sensitive to increases in p than both the classical np chart and the RL chart. The chart is based on 100% inspection but the synthetic chart could also be based on the number of conforming units between nonconforming ones in the inspected data, where the non-inspected data is ignored. The authors mentioned that the CUMSUM chart might be “more effective” than the synthetic chart (p. 111) but nothing tangible was presented to back up that statement. They also recommend the use of the synthetic control chart in steady-state situations since “its detection of delayed shifts may be slower” (p.111).

2.6.5 Control charts using moving averages.

Gan (1990) discusses the exponentially weighted moving average (EWMA) using binomial distributions. Trevanich and Bourke (1993) discuss the EWMA chart for attribute data. Roberts (2000) discusses the control charts based on geometric moving averages. A geometric moving average gives the greatest weight to the most recent observation, and all previous observations decrease in a geometric progression of weights back to the first observation. When tests are conducted to compare the relative sensitivity of detecting shifts in process average between control limits based on geometric moving averages with control limits based on regular moving averages, the geometric moving average has a higher sensitivity in detecting process shifts.

2.6.6 Further comparisons among different control charts

Chang and Gan (2007) propose, along with the traditional p-control chart, a run chart based on similar principles as the one designed by Goh (1987b) to track both increases and decreases in the probability of nonconformance p. Among the traditional CCC chart, conditional

CCC chart, the CUMSUM geometric and Bernoulli charts, the synthetic chart and their own modified run chart, they found from run length comparisons that " all the charts considered are found to be unable to provide a timely feedback of an out-of-control situation even for relatively large increases in *p*" (p. 874). They also found that their technique (the run chart along with the traditional *p*-chart) is more sensitive than the synthetic chart in detecting increases in the rate of nonconformance *p*. It is also more sensitive than the CCC and the conditional CCC in detecting increases in the rate of nonconformance *p*. Its performance is comparable to the CUMSUM Bernoulli chart.

2.7 Motivation for using the sequential Bayesian approach

Based on the literature review, the most widely used technique for dealing with high quality processes is the CCC chart. The CCC methodology solves most of the problems associated with the *p*-control chart when the rate of nonconformance *p* is small (Kaminsky et al., 1992; Glushkovsky, 1994; Nelson, 1994; Woodall, 1997; Xie and Goh, 1992, 1997).

One main drawback of the CCC chart is that it still uses the same approach in the *phase I* of the production process as the traditional Shewhart chart, namely: gets a sample, finds an estimate of *p* (the probability of nonconformance of any one item), uses the value of *p* to set the control limits and then tests the same sample against the control limits. If a point falls outside the control limits, throws out the point, re-computes *p* and restarts the process over again. The flaw of using data points to draw the control limits that test the same data points has been addressed in the introduction of this dissertation and was labeled "double use of data". Another problem with the CCC approach is that the value of the nonconformance *p* is estimated from the initial sample. If the initial sample size does not contain any nonconformity, then the estimate of *p* is not possible and the sample size is increased on an ad-hoc basis until one obtains nonconformities in the sample. The problem becomes: how should one increase the sample size and when should one stop? Given the importance of a good estimate for *phase I* production discussed in the literature and iterated in section 2.5.3, one is compelled to start looking for approaches "outside the box". Another important problem with the CCC chart discussed in the literature is its inefficiency in detecting process deterioration unless that deterioration shift is large (Chan and Goh, 1997; Xie and Goh, 1992; Xie et Al., 2000; Tang and Cheong, 2004; and Yang et al., 2002).

In the section below, similarities between the Bayesian approach and the different solutions proposed in the literature to deal with the shortcomings of the CCC chart are highlighted.

2.7.1 Similarities with solutions offered in literature

The literature and methodologies discussed so far describe mostly problems associated with high yield processes and to lesser degree problems with *phase I* of the production process. Answers to these problems are also offered in the literature. The more relevant answers seem to have common features with the Bayesian approach. The conditional CCC, the CUMSUM technique, the sequential procedure, the resetting of the counter to a value different than zero to account for past information, the different moving averages and Jeffreys prior all seem to have a similarity in approach to the Bayesian concept, except the Bayesian concept incorporates all of these methodologies. From there lies its power. Below is a more detailed discussion regarding these similarities.

2.7.1.1 Resetting the counter

One point discussed in section 2.4.2 is the fact that having to reset the counter at zero translates into loss of information. Recalibration is done to account for past information. This is similar to the Bayesian approach where the past information is updated continuously.

2.7.1.2 Sequential procedure

Thang and Chong (2004) address the problem of finding a good estimate of the probability of nonconformance p by designing a sequential procedure where p is continuously updated and a continuous adjustment is made to the CL. Thang and Chong, (2005), in a subsequent paper, dismiss the adjustments made to the CL and state that their use of adjustments change the probabilistic interpretations of the original control limits. This is the closest of a procedure to a simple sequential procedure and to a sequential Bayesian approach.

2.7.1.3 Conditional CCC chart

The conditional CCC chart introduced by Kuralmani et al. (2002) is also similar in its approach to the Bayesian methodology since the basis of the Bayesian methodology is to condition the predictive distributions on previous observations. The authors claim that their approach results in tighter control limits and an improved sensitivity in detecting process shifts compared to the regular CCC chart.

2.7.1.4 Moving average

Most of the literature on moving averages in setting control charts is concentrated on variable data. Trevanich & Bourke (1993) compare the CUMSUM charts in monitoring fraction defectives with the exponentially weighted moving averages (EWMA) charts and find that the EWMA approach is better able to detect and respond to a small to moderate deterioration in the process. Roberts (2000) shows that the control chart based on a geometric average (where the most recent observation weighs the most) show increase sensitivity to small process changes compared to a standard EWMA chart. Even though the author uses the $\bar{x}$ control chart for his analysis, he claims that the results apply to other sample statistics control charts. The geometric EWMA control chart approach is similar in theory to the Bayesian approach. In the face of lack of data in the Bayesian approach, more weight is put on the information provided by the prior distribution; however, as data is collected, the weight shifts to the most recent observations, the same way geometric EWMA works.

2.7.1.5 CUMSUM chart

The CUMSUM procedure is shown throughout the literature discussed above to have an ability to detect a process shift over the CCC chart, particularly process deterioration. The CUMSUM approach is based on summing cumulatively all the deviations of the sample values from the target values and incorporating all the values from the previous samples up to the present. The Bayesian sequential approach is similar in the sense that even though the variability of the sample values is not compared against a target value, the variability among the samples is taken into account. The Bayesian approach also has the cumulative effect since all the past values with their deviations are updated sequentially.

2.7.1.6 Jeffreys interval

Brown, Cai and DasGupta (2001), in their serious criticism of the classical procedure of interval estimate for proportion, recommend using the equal tailed Jeffreys prior interval. They also recommend that the introductory statistics textbooks introduce their alternative intervals even at the expense of simplicity. Jeffreys methodology is similar to the sequential Bayesian procedure in that it is Bayesian in nature except that the prior is noninformative and the methodology is not sequential.

2.7.1.7 A Bayesian approach to a u-control chart

Bayarri and Denato (2005) test a sequential Bayesian approach to replace the traditional non-sequential u-control chart with its Poisson model for observations. The claim is that the Poisson model might not be adequate for process control data. They back up that claim with extensive bibliography showing that data coming from production processes do not meet the hypotheses underlying the Poisson model. Other literature points to extra variance often present in the Poisson model. Other research points to a "significant increase in the false alarm rate"(p. 144). They make three comparisons:1) a traditional u-control chart using a Poisson distribution; 2) an Empirical Bayesian (EB) u-control chart with a given sample and therefore an informative prior along with a hierarchical Poisson-gamma model; and 3) a Bayesian sequential u-control chart with an "objective" Jeffrey's prior. When a sample of 25 observations is tested using an example from a book, all three methods show observation 13 to be out of control, but in the sequential approach, that fact is acknowledged immediately after the observation is taken and plotted. This is unlike the non-sequential approaches where one has to wait past the base period (BP), after the 25 observations have been collected and control limits have been set. The sequential approach has wider control limit widths than the EB control limits but the limits tends to grow narrower as the number of observations increases so that the control limits of the sequential Bayesian procedure converge to the fixed control limits of the EB control chart.

2.7.2 Self-starting mechanism and short runs.

Short runs are executed traditionally through the use of EWMA, moving average and CUMSUM, approaches that mimic in one form or another the Bayesian approach [9-11]. Quesenberry (1991) uses approximately normalized control charts (called Q charts) to chart a binomial random variable in the cases when p is known and p is unknown (before the charting begins). Transformations are made in both cases, and the technique is used for short runs.

Hawkins et al. (2007), commenting on the importance of self-starting mechanisms, state that parameter estimates gathered from *phase I* sample to establish process parameter values for the chart cannot possibly establish the exact process parameters. They also argue that small random errors translate into serious distortions of the run behavior of the chart. They suggest using "self-starting" methods that begin the control of the process right away without using the preliminary step of a large *phase I* sample.

2.7.3 Avoiding the problem of multiple control charts

For high yield processes with 100 percent inspection, Breyfogle III (2003) suggests using time between nonconformance occurrences and converting that into an average number of nonconformances per unit time (for example, the number of nonconformances that occur each month). A c-control chart is utilized, which tracks the number of nonconforming units per unit time. Moreover, the number of nonconformances is also converted into a rate of nonconformance (changing the data from discrete to continuous), and a combination of two charts, the *X* chart (for individual observations) and the moving range chart (that charts the range between two consecutive observations) is proposed. The combination of the two charts is called *XmR* chart. A visual comparison of the *XmR* control chart and the *c*-control chart shows that the *XmR* chart is able to detect one out of control observation whereas the *c*-control chart does not. Breyfogle III concludes that the *XmR* chart, which charts measurement rates instead of counts is the better procedure, and states that "count information is generally weaker than measurement data"(p.242). But he also states that "Some practitioners prefer not to construct moving range charts … and the moving ranges are correlated which can induce patterns of runs or cycles" (p. 227). Abbasi, Taghi and Niaki (2007) note that when two separate univariate control charts are applied together, the determination of type I error becomes problematic. Bayarri and Berger (2003) discuss issues related to differences between "frequentist" and "Bayesian" approaches to statistics. One of the differences involves multiple comparisons. A frequentist approach makes adjustment to the type I error when multiple tests are conducted simultaneously, whereas "a correct adjustment is automatic within the Bayesian paradigm" (p.37).

CHAPTER 3

THE PROPOSED MODEL

The sequential Bayesian CCC approach is able to change or adjust previous beliefs about a process or a system in light of new and incoming data. In contrast with the classical approach which uses a specific estimate of the unknown parameter p, the Bayesian approach proposes a distribution for this parameter called a prior distribution to take into account the variability inherent in the estimation of p.[12-17] During implementation, the posterior distribution based on Bayes's theorem is computed by incorporating data just collected. The change from prior distribution to posterior distribution reflects the information about the parameter value provided by the new data. In a sequential Bayesian methodology, the posterior distribution at time t becomes the new prior for time $t+1$.

The sequential Bayesian methodology developed in this dissertation allows drawing confidence limits even when only originally three observations are given. These observations are used to define the parameters of the prior and the predictive distributions when the process starts and the original control limits (CL) from the predictive distribution. If past knowledge or data is available, that information can be used to define the prior distribution parameters and no observations are required to set the original control limits. A new observation is then taken and tested against the original control limits. If the observation falls inside the limits, that observation is used to update the prior and set the next confidence limits. By conditioning on the present observation and drawing the control limits for the next observation, the posterior distribution at index i becomes the prior distribution at index $i+1$. If and when an observation falls outside the CL, the observation is discarded and the previous observations are used to compute a new prior that would lead to more accurate estimates of the CL. By adjusting the information about the parameter p in light of new data, deterioration in the data is mirrored in the corresponding CL, and deterioration detection can be clearly and quickly made.

3.1 The Advantages of the proposed approach

The sequential Bayesian approach proposed in this dissertation is a very efficient tool capable of resolving many issues and problems including: Avoidance of the "double-use" of the data, a better estimate of rate of nonconformance p, increased sensitivity in detecting process deterioration, increase in process speed resulting in timely feedback, no restrictions on the parameter of the prior distribution, and an increased efficiency in data gathering and analysis. These advantages are discussed in greater detail below.

3.1.1 Avoidance of the "double-use" of the data

The data used to set the control limits is different from the data tested by the control limits. This problem is addressed by Bayarri and Garcia-Donato (2005). It can be seen that because of the *phase I* period (referred to as base plan or BP), any non-sequential approach to finding the control limits causes the "double-use" of data. In other words, any non-sequential approach uses observations to set the control limits and then uses the control limits to test the same observations used to set these limits. This is clearly a biased approach. In contrast, the sequential Bayesian procedure proposed in this dissertation uses sequentially updated data to set new control limits, and the observations used to set the control limits are not the same as the observations being tested.

3.1.2. Better estimate of *p*

The proposed approach takes into account the uncertainty in the rate of nonconformance *p*. In the traditional approach, the estimate of *p* is based on its average and the variability of the estimate is neglected. Moreover, it is crucial to have an accurate estimate of *p* during *phase I*, but the traditional control charts do not guarantee that. This is addressed by Tang and Cheong (2004), Yang et. al. (2002) and Hawkins et al. (2007) and discussed in greater detail in section 2.5 of the literature review of this dissertation.

3.1.3 Increased sensitivity in detecting process deterioration

The CCC chart, while solving the problem of the *p*-control chart due to the high conformance rate, cannot detect a small deterioration shift. Alternatively, because the Bayesian methodology incorporates the knowledge contained in new observations, any shift in the observations can be visibly detected through a shift in the LCL.

3.1.4 Timely feedback

The dynamic sequential nature of the Bayesian approach captures shifts immediately as observations are drawn and control limits updated.

As stressed by Chang and Gan (2007), a timely feedback on the "health of a process" is very important. The speed of the inspection of quality and therefore the opportunity for a quick reaction is more assumed under a sequential Bayesian procedure than under any other classical non-sequential procedure. This is because the classical procedure uses a pre-determined fixed sample size of observations to get the estimate of nonconformance rate *p* in order to draw the control limits. That sample has to come from a stable process for the estimate of *p* to be representative, and the sample size might not be adequate. In the Bayesian approach, only three observations are needed to set up the initial control limits and start the production process, essentially a self-starting procedure. When historical data is available (a pessimistic estimate of *p*, an optimistic estimate and a most likely estimate), that historical data is used to compute the parameter values of the prior distribution. As a result, no initial observations are needed to set the initial control limits. As new information is incorporated, the estimated nonconformance rate *p* and its variance are updated, so are the control limits.

Moreover, during *phase I* of the process, in the non-sequential approach, if and when observations falls outside the control limits, the observations are discarded and the new control limits redrawn based on the rest of the observations, observations already tested. One has to wait until **all** the observations are tested before any can be discarded. The process iterates until all observations fall inside the control limits. This is sharply contrasted with the sequential methodology where at the sight of the first observation falling outside the control limits, the observation is immediately discarded and the "in-control" observations preserved via a new updated prior and the testing process continues.

Note that in the non-sequential procedure tested observations are added in the mix and retested whereas in the sequential procedure and through a sifting approach, an observation is tested only once and tagged if it passes the test and then reused for updating the CL.

3.1.5 No restriction on the parameters of prior distribution

Kepner and Wackerly (2002) studied the behavior and the shape of the predictive distributions when the values of the parameters of the prior change. They found that particular choices of parameter values could lead to undesirable results for the predictive distribution. In particular, the binomial distribution along with its conjugate beta prior is looked at and some useful and important conclusions are drawn.

When the parameter *p* that comes from a binomial distribution is described in Bayesian terms as having a beta prior with parameters σ and τ, the corresponding predictive distribution of the random variable *x* (corresponding to the number of non conformances in a sample of size *n*) is called a beta-binomial distribution or a "beta mixture of binomial distributions." It has a closed form representation.

The authors found that for the situations where the parameters of the beta prior are: $\sigma = 1, \tau < 1$; $\sigma < 1, \tau = 1$; $\sigma < 1, \tau > 1$ and $\sigma = 1, \tau > 1$, the predictive distribution is monotone decreasing. This means that the most likely value of the random variable *x* is 0, the minimum possible value of the random variable *x* regardless of the size of *n*. Likewise, when $\sigma > 1, \tau < 1$ and $\sigma > 1, \tau = 1$, the predictive distribution is monotone increasing, implying that the most likely value of *x* is *n*, the maximum possible value of *x* regardless of the sample size *n*. Using a model where *n* is the most likely value of the number of nonconformities in a sample of size *n* "may be problematic" (p.765) state the authors. They also add: "A casual user of the Bayesian methods might fail to notice the impact that the chosen values of α and β have on the predictive distribution and, thereby, miss the opportunity to improve the model by using alternative parameter values in the prior" (p. 765).

When the values of σ and τ are both < 1, the shape of the predictive distribution is neither monotone increasing nor monotone decreasing, it is U-shaped. The corresponding values of this predictive distribution have more meaningful interpretations. This category falls into a special set of prior and predictive distributions where a particular example is the "non-informative" Jeffreys prior with $\sigma = \tau = ½$.

In the case where $\sigma = \tau = 1$, the shape of the predictive distribution is flat (constant). Many statisticians use these parameter values for the prior [18-23] including Bayes [23]. The purpose for using a "uniform" prior (where all the values of the random variable x from the predictive distribution are equally likely) is for a totally different set of circumstances and applications.

In the case where both σ and τ are > 1, the shape of the original binomial distribution is preserved by the predictive distribution. Both are uni-modal and mound shaped. This is very desirable since there is a maintained consistency in the overall behavior of the distributions and the interpretations values of the predictive distribution are meaningful.

In the sequential Bayesian CCC methodology proposed in this dissertation, the shape of the predictive distribution is geometric, the same as the distribution of the original model. Therefore, no restrictions exist on the parameters of the prior of the Bayesian model. This is an important feature of the proposed model.

This Bayesian methodology based on a geometric distribution performs better than, for example, a Bayesian methodology based on a binomial distribution. This is due to at least one obvious reason: The predictive distribution based on the geometric distribution has a geometric shape (even though not a closed form distribution), whereas the predictive distribution based on the binomial distribution is the beta-binomial distribution which does not preserve the shape of the binomial base distribution for all values of the parameters σ and τ. As a result, restrictions are imposed on the parameters for meaningful Bayesian analyses and interpretations.

3.1.6 More effective use of data.

3.1.6.1. In data analysis

The effective use of data in the analysis is apparent from the fact that past observations are not forgotten but used in the updating of the prior (the counter is not reset at zero) to account for past observations. Also, the good observations are not all added back in the mix when an observation is found to be out of control as is done in the *phase I* of the non-sequential traditional control charts. Instead they are properly separated and reused. Most importantly, the timed order of the observations is captured along with the variability among the observations. In essence, this is what gives the procedure its edge in capturing process deterioration.

3.1.6.2. In data gathering

Not only is data used effectively in the analysis as is discussed in section 3.1.6 above but also in the manner it is gathered. No more than necessary data is required to set the control limits and the knowledge embedded in the data is used to its fullest extend during the data analysis. The fact that the sample size is not fixed in the sequential Bayesian approach, the fact that the estimate of the random variable p (representing the rate of nonconformance) with its deviation is updated with every new observation and the fact that the prior is updated with the "in-control," i.e. good observations, all reflect the use of every observation to its fullest value.

In essence, the methodology proposed here will prevent the use of the "double data", will detect a small deterioration in the process, will provide timely feedback on the health of the process (and therefore the opportunity for a quick reaction), will give a dynamic and good estimate of the rate of nonconformance p and will make a highly effective use of every observation given, without forgetting the past ones, therefore preserving and optimizing the value of any information given, regardless of its size.

When the initial sample size used for the layout of traditional control chart is simply not available or has prohibitive costs, one is compelled to make good use of every observation given. The proposed model makes highly effective use of the observations.

3.1.8. Model performance when the rate of nonconformance is not small

The motivation behind this dissertation is to be able to solve the problem of detecting process deterioration when the rate of nonconformance is quite small. It turns out that the sequential Bayesian CCC methodology has an increased sensitivity in detecting process deterioration compared to classical and benchmark procedures when the rate of nonconformance is small and when the rate of nonconformance is not small. Its performance under a wider range of situations makes the Bayesian methodology even more advantageous and attractive.

3.2 Steps involved in creating the Bayesian CCC charts

3.2.1 Likelihood function of observations

The likelihood function in this case is a geometric distribution with parameter p of the form :

$$f(x/p) = (1-p)^{x-1} p \qquad x=0,1,2, \qquad (13)$$

where p represents the probability of nonconformance, and x represents the number of conformances up to the first nonconformance. For example, if the probability of nonconformance is .01, the probability of seeing the 50 conformances **up to** the first nonconformance, i.e., the probability of seeing the first nonconformance **after** 49 conformances is $f(50) = (1-.01)^{50-1}(.01) = .0061$

Throughout the dissertation, x, is referred to as an observation. It is interpreted as the number of conformances up to the nonconformance. In this case it is 50. When indexed by i, it means the number of conformances (after the last i-1^{rst} nonconformance) up to the *ith* non-conformance.

3.2.2 Conjugate prior distribution for parameter p

In this Bayesian approach, the parameter *p,* representing the probability of nonconformance, is a variable parameter to account for the variability or the uncertainty about its value. That uncertainty is modeled by assigning to p a probability distribution called the conjugate *prior* distribution to the geometric one. There is extensive literature about the appropriate conjugate priors of different distributions [24-25]. In the case of the geometric distribution, since it belongs to the family of exponential distributions, a natural and appropriate prior is the beta distribution with parameters σ and τ. The parameter p of the geometric distribution is now a gamma distributed random variable with parameters σ and τ. The form of the distribution is the following:

$$\pi(p) = \frac{1}{B(\sigma,\tau)} p^{\sigma-1} (1\text{-}p)^{\tau-1} \tag{14}$$

where $\sigma > 0$, $\tau > 0$, $0 < p < 1$ and B represents the beta function.

Given the parameters σ and τ, the mean and the standard deviation of the beta can be computed since:

$$\mathrm{E}(p) = \frac{\sigma}{\sigma + \tau} \tag{15}$$

and

$$\text{Variance } (p) = \frac{\sigma\tau}{(\sigma+\tau+1)(\sigma+\tau)^2} \tag{16}$$

Conversely, given the mean and the variance of p, the estimates of the parameters σ and τ can be found using the method of moments. The formulas for σ and τ are displayed in equations (17) and (18) below.

$$\sigma = \bar{p}\left[\frac{\bar{p}(1-\bar{p})}{v} - 1\right] \tag{17}$$

$$\tau = \left[\frac{\bar{p}(1-\bar{p})}{v} - 1\right](1-\bar{p}) \tag{18}$$

where $\bar{p}$ represents the estimate of the mean of p and v represents the variance of the estimate.

If **no historic data is present**, the first three observations, x_1, x_2 and x_3 are used to find the mean and variance of the estimate of p. The mean and variance are then used in equations (17) and (18) to find the values of the parameter σ and τ.

The expected value of the probability of nonconformance p in a binomial setting is found by the following:

$$\hat{p} = \frac{1}{E(x)} = \frac{1}{\frac{\sum_{i=1}^{n} xi}{n}} \tag{19}$$

The variance of the p estimate is found by using the fact that the estimate follows a beta distribution. According to Gido and Clements (2006, p 186), the variance of a beta distributed random variable can be estimated by the following:

$$Variance(p) = \left[(\frac{c-a}{6})^2\right] \tag{20}$$

where c is the most pessimistic estimate of p (in our case the largest estimate) and a is the most optimistic estimate of p.

3.2.3 Use of historic knowledge to find the parameters of the prior

Gido and Clements (2006, p 184) state that when three estimates are used to time an activity-the optimistic time called *a*, the pessimistic time called *c* and the most likely time called *b*-the assumption is that the three estimates follow a beta distribution. According to the authors, the expected value of the random variable *p* that follows a beta distribution is:

$$Mean(p) = \frac{a + 4b + c}{6} \tag{21}$$

and the variance of the random variable is already formulated in equation (20). Here, the values of *a* and *c* represent the optimistic and pessimistic estimates of the parameter respectively and *b* is the most likely one.

The prior in our case is beta distributed. This means that the assessment is applicable in our case when the history of the behavior or past information about the random variable *p* (representing the probability of nonconformance) is available. In particular, if the most pessimistic estimate of *p*, the most optimistic one and the most likely one can be quantified, then one can get the mean and standard deviation of *p* and therefore the parameter estimates σ and τ of the prior beta distribution. These parameter values make it possible then to draw the first CL without the recourse to any observations.

3.2.4 Use of Bayes theorem to find the predictive distribution

The conditional distribution of *p* given $x_1 \ldots x_n$ is the posterior distribution for *p*. From the definition of conditional distribution, it can be written in the form:

$$f(p / x_{1,\dots} x_n) = \frac{f(x_1 \dots x_n, p)}{f(x_1 \dots x_n)} \tag{22}$$

By using a form of the Bayes formula to rewrite the numerator, the equation can be rewritten as:

$$f(p / x_{1,\dots} x_n) = \frac{f(x_1 \dots x_n / p)\pi(p)}{f(x_1 \dots x_n)} \tag{23}$$

Using the fact that the observations are independent, the posterior distribution can now be expressed as:

$$f(p / x_{1,\dots} x_n) = \frac{\left[\prod_{i=1}^{n} f(x_i / p)\right]\pi(p)}{f(x_1 \dots x_n)} \tag{24}$$

The numerator of the expression on the right hand side of equation (24) above can be expanded into:

$$\left[\prod_{i=1}^{n} f(x_i / p)\right]\pi(p) = p^{n} (1\text{-}p)^{\sum_{i=1}^{n} x_i - n} * p^{\sigma - 1} (1\text{-} p)^{\tau - 1} \tag{25}$$

and rewritten in the form:

$$\left[\prod_{i=1}^{n} f(x_i / p)\right]\pi(p) = p^{n + \sigma - 1} (1\text{-}p)^{\sum_{i=1}^{n} x_i - n + \tau - 1} \tag{26}$$

Since the posterior distribution is proportional to the numerator of the expression on the right hand side of equation (24), it can be expressed as follows:

$$f(p / x_{1,\ldots} x_n) \propto p^{\,n+\sigma-1}\,(1\text{-}p)^{\sum_{i=1}^{n} x_i - n + \tau - 1} \tag{27}$$

which is the pdf of a Beta distribution with parameters ($n + \sigma$, $\sum_{i=1}^{n} x_i$ - $n + \tau$).

The predictive distribution is the conditional distribution of the future n+1[rst] observation, given the first n of them. It is:

$$f(x_{n+1} / x_1, \ldots , x_n) = \int_0^1 f(x_{n+1} / p)\, f(p / \mathrm{x}_{1,\ldots}\, x_n)\, dp \tag{28}$$

where $f(x_{n+1} / p)$ is the geometric distribution function of the form $p(1\text{-}p)^{x_{n+1}-1}$ and $f(p / x_{1,\ldots} x_n)$ is the posterior beta probability distribution function given by equation (27) with parameters ($n + \sigma$, $\sum_{i=1}^{n} x_i$ - $n + \tau$). Plugging the extended form of the functions in (28) results in:

$$f(x_{n+1} / x_1, \ldots , x_n) = \int_0^1 p^{\,\sigma+n}\,(1\text{-}p)^{\sum_{i=1}^{n} x_i + x_{n+1} - n + \tau - 2}\,\frac{1}{B(\sigma + n, \sum_{i=1}^{n} x_i - n + \tau)}\,dp \tag{29}$$

This is like averaging the original geometric density $f(x_{n+1} / p)$, with weighting factors being the beta posterior distribution function $f(p/\ \mathrm{x}_{1,\ldots}\ x_n)$, with mean value $\dfrac{\sigma + n}{(\sigma + n) + \sum_{i=1}^{n} x_i - n + \tau}$.

Note: Here the geometric distribution is of the form $p(1\text{-}p)^{x\text{-}1}$. It means that out of x trials, the x^{th} trial has the first nonconformity and $\sum_{i=1}^{n} x_i$ is the total number of observations. An index n is registered when a nonconformance is observed. So if initially 1999 conforming items are observed before a nonconforming one, then n=1 and x_1 = 2000. If after that, 2500 items are found conforming before the next nonconforming item is observed, then n=2 and $\sum_{i=1}^{2} x_i$ = 2000+ 2501 = 4501. Note also that the posterior distribution for observation n-1 will be the prior distribution for n.

3.2.5 Determination CL based on the predictive distribution

This involves performing the following steps:

1. To get the upper and lower control limits, add the upper and lower probability values of the predictive distribution displayed in equation (29) until the probabilities add up to the $\alpha/2$ value and then mark the corresponding x_{n+1} values. The lower control limit represents a value of the predictive random variable from the low end of the predictive distribution. The upper control limit represents a value of the random variable x_{n+1} from the tail end of the predictive distribution. In a one-sided control limit (in our case the interest would be in the lower control limit for detecting process deterioration), the corresponding x_{n+1} value on the lower end of the probability distribution is set at α level.
2. Check to see if the new observation x_{x+2} falls inside the control limits set by the previous observations or information. If it does, update the parameters of the predictive distribution and draw new control limits. In the case of a one–sided lower control limit,

check if the new observation falls above the lower control limit. If it does, update the parameters of the predictive distribution and draw a new lower control limit.

3. If the new observation (x_{x+2}) does not fall inside the control limits, check the process but discard the point. Use all the previous points up to x_{n+1} to compute a new prior and go back to step 1 to draw the control limits.

4. Stop the process when the observations are exhausted.

Note: In a process that is somewhat stable and/or there is a large amount of observations to be tested, steps 1 through 3 are repeated until the control limits become "close enough" to each other. The last control limits (or one-sided lower control limit) can be used as fixed control limits in lieu of the control limits computed during the *phase I* of production and the incoming observations can be tested against them. In this dissertation, the stopping criterion used to establish when "close enough" is good enough is when the last two control limits are within .2% of each other.

3.2.6 Determination of new prior

When the prior distribution is updated with the flow of new information (or new observations), the control limits become more representative of the more current observations than the earlier ones. This is similar in behavior to an exponentially weighted moving average (EWMA) where the later observations have more weight on the representation of the system than earlier ones. This behavior is valuable in detecting process deterioration as is shown in section 4.11.1.2 of this dissertation. In the proposed methodology, in order to have a systemic representation of a process where more weight is put on later observations than earlier ones, a guideline (introduced in the next two sections) is set-up for updating the prior distribution.

3.2.6.1 On the absence an "out-of control" observation

If within 10 tested observations none is found to be "out of control", then a new prior is computed based on previous observations and the ten tested observations and the process continues. If the first 10 tested observations are found to be "in-control", then the adjusted prior is based on the first 10 observations and the initial three observations used to determine the original prior. This value 10 was based on simulations conducted on a range of values of *p* from 10^{-1} to 10^{-5} to assess a proper shift of the weight from the distant to the more recent observations.

3.2.6.2 In the presence of an "out of control" observation

If an observation is found to be "out of control" within the 10 tested observations, then that observation is discarded and a new prior is computed based on the previous tested observations and the initial observations used to determine the original prior. This guideline helps that the process is being continuously updated with new observations and that more weight is put on the most recent ones.

3.2.7 Summary of steps

Below is a summary of the steps taken to start and end the process of drawing CL based on the sequential Bayesian CCC chart procedure:

1. Define the geometric function of observations with the parameter *p*, where *p* is the number of non-conformances up to the conformance.

2. Define the conjugate prior distribution of the random variable *p*, a beta distribution with parameters σ and τ .

3. Define the predictive distribution from the prior beta distribution.

4. Use the parameters of the prior based on historic data or the first three observations gathered to evaluate the original predictive distribution and corresponding $CL_{n=0}$ (indexed at n=0, i.e. before any observations are tested).

5. Test the first observation against the $CL_{n=0}$. If "in-control", use the observation to update the CL. The CL is now indexed by n=1 ($CL_{n=1}$). If the first observation is "out of control", remove and test the next observation.

6. If the first ten tested observations result in no "out-of-control" observations, use the first ten tested observations (and the original observations used to determine the parameters of the beta distribution, if any) to update the prior. The updated prior indexed at n=0 is now used to set the $CL_{n=0}$ indexed at n=0. In other words the counter is reset at zero, whenever the prior is updated. If the last ten tested observations result in no "out-of-control" observations, use all the previous observations, including the last ten to update the prior and reset the counter at zero and draw $CL_{n=0}$ indexed at n=0.

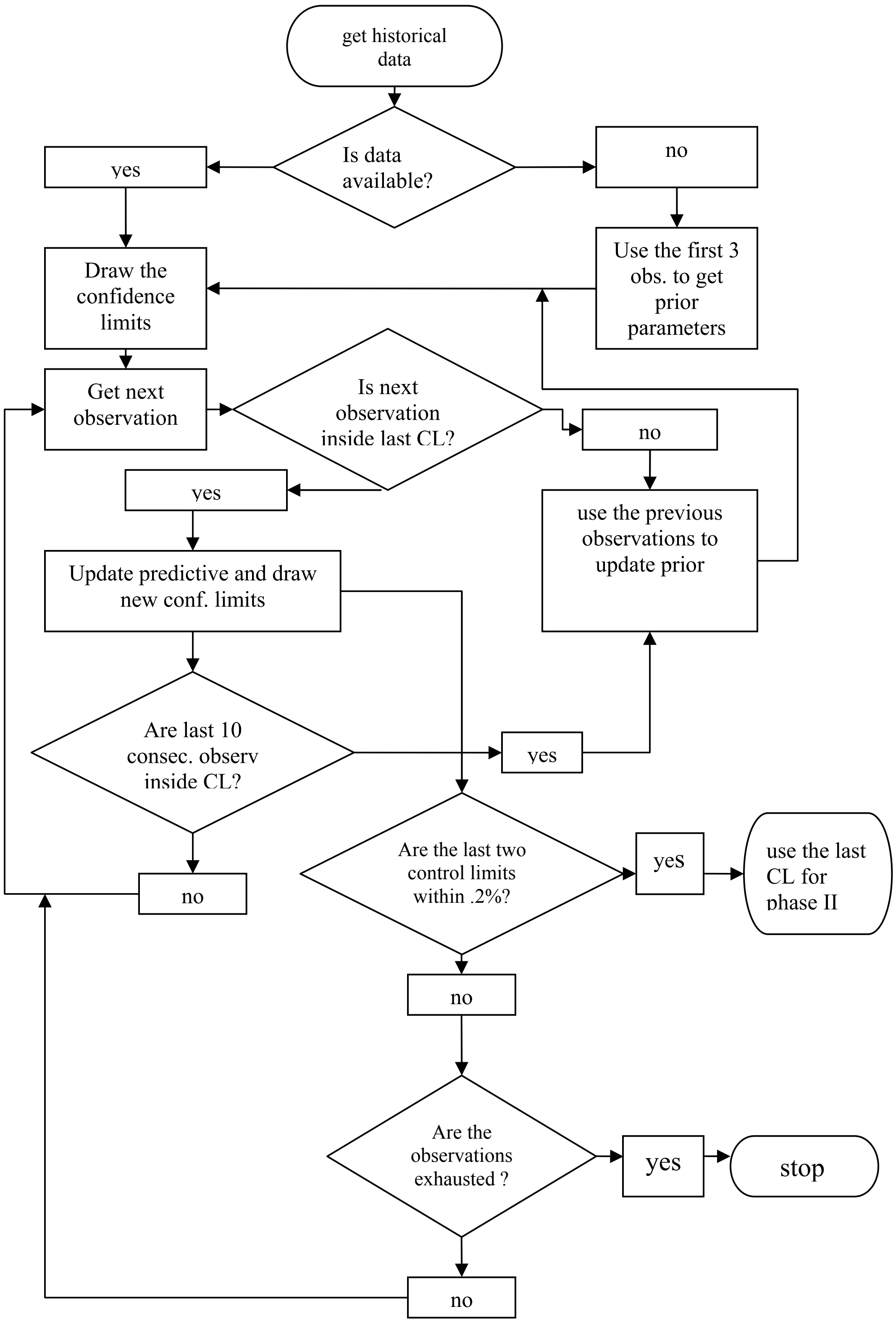
7. If the first "out-of-control" observation falls within the first ten tested observations, remove the "out-of-control" observation and update the prior based on all previous "in-control" observations and the original three observations used to evaluate the initial prior. The updated prior is again used to set the counter at zero and draw $CL_{n=0}$ indexed at n=0. If within the last ten observations one (or more) observation is "out-of-control", then the observation is removed and all the previous observation are used to update the prior and reset the counter at zero and draw $CL_{n=0}$ indexed at n=0. As stated above, the counter is

reset at zero whenever the prior is updated. The prior is updated whenever 10 successive observations are found to be in control or whenever one observation is found to be “out-of control” and is removed. The prior requires using all the previous” in-control” observations for its update along with the original three observations used to evaluate the original prior distribution.

8. Continue updating the control limits until all observations are exhausted.

9. When the number of observations is extensive, and when values of CL move very close enough together before all observations are exhausted (mainly true in a stable system), use the last CL as a fixed CL to test future observations. In other words, when the CL values become close, then the converging CL can mimic in essence the CL determined in *phase I* of production and be used in testing *phase II* without using the sequential methodology in *phase II*. The guideline for using the stopping procedure (when the CL become close enough) is when the last two control limits fall within .2% of each other.

3.3 Flowchart



CHAPTER 4

TESTING THE METHODOLOGY

4.1 Use of Monte Carlo simulation

In this section, a Monte Carlo (MC) simulation experiment is conducted to verify the proposed methodology. A hypothetical production process is utilized in this experiment.

4.1.1 Initial Estimate of *p* based on past information

It is assumed that from history, the most optimistic, pessimistic and most likely assessment of p are .0002, .0007 and .0005, respectively. Finding the mean and standard deviation of the beta distributed random variable p (the probability of nonconformance) from equations (20) and (21) yield: $\bar{p}$ = .00048333 and a variance of the estimate = 6.94444×10^{-9}. Using the values of the mean and variance in equations (17) and (18) yield the estimated parameters σ and τ **:** σ =33.6232 and τ = 69531.7, respectively.

MC simulation is performed next to get the values of the integral that defines the predictive distribution in equation (29) based on the computed values of the parameters σ and τ .

4.1.2 The MC simulation procedure

At x=1, N values of random variable p from a uniform (0, 1) are generated. For each one of the values, the corresponding integrand is calculated. Then the average of all the N integrands is computed. That average represents the value of the predictive distribution at x=1. The process is repeated for the whole range of x. When all the values of x are exhausted, the upper and lower control limits are computed. Figure 3 shows the Monte Carlo simulation of the probability mass function of the predictive distribution for x ranging from 1 to 50,000 and N =1000. The distribution in Figure 3 is based on n=0 (before any observations are taken), σ=33.6232, τ =69531.7 and a range of x from 1 to 50,000.

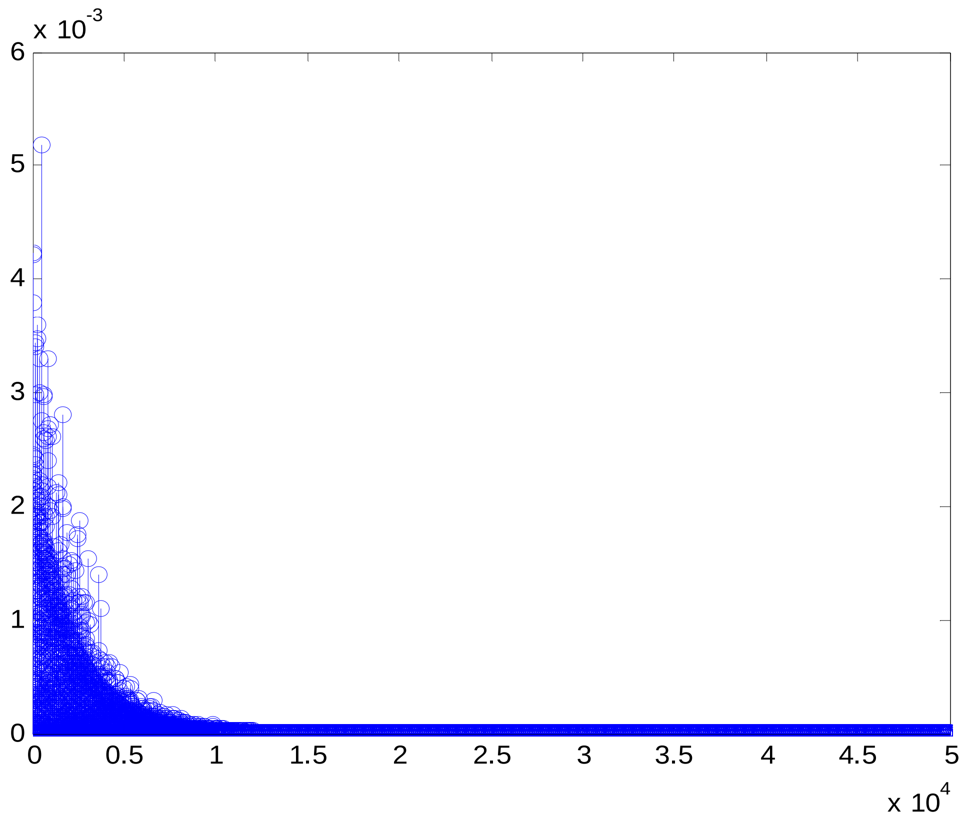


Figure 3 Example of a MC simulation

In Figure 3 the following points can be observed:

1) The predictive distribution is also geometric, and this desired feature of the proposed sequential Bayesian CCC approach has already been discussed in Section 3.1.5.

2) The range of x depends on the value of the p estimate. The smaller the p estimate is, the heavier tail the distribution has.

4.2 The time factor

The term in the exponent of the expression $(1\text{-}p)^{\sum_{i=1}^{n} x_i + x_{n+1} - n + \tau - 2}$ in equation (29) affects the simulation time the most because of the size of the exponent. In particular, the following notes can be made:

1) The simulation time is increased by an increase in the observation value x (a decrease in the estimate of p)

2) The simulation time is increased with an increase in the τ value. The τ value is determined by the estimated value of p and also by the spread of the observation values that together help form the original p and ν estimates and therefore the values of τ and σ values.

3) The simulation time is also increased by an increase in N, the number of iterations used before an average is taken. The average represents the distribution value at a particular value of x.

4) The simulation time is increased by increasing the range of the x values. The range is usually extended to cover all possible values of x for which the distribution adds up to one. The smaller the estimate of p is, the wider the range of x is.

4.3 Enhancing the simulation procedure

The performance of the simulation is enhanced using Importance Sampling (IS) in the MC simulation procedure. It can be demonstrated that IS increases both the speed and the accuracy of the estimation results. A procedure developed in this dissertation called "skipping" also increases the speed of the simulation. Finally, truncating the simulation process after the LCL is obtained increases the simulation speed significantly.

4.3.1 Increasing the simulation speed

4.3.1.1 Skipping technique

In Figure 3, the distribution flattens around the value of 15,000. Because the interest is mainly in process deterioration and therefore in the LCL, approximations are made to the UCL by allowing skips in the computations beyond the value of 15,000 of x; skipping through an equal range of values for x and assigning the same y values for y in that range as the one evaluated at the midrange of x. The computational time is increased substantially after x=15,000 because the value of y is computed only once in the skipping range. In the example above, skipping through a range of 700 values cuts the number of iterations from 50,000 to 15,000 + 50 (35000/700=50), where skipping occurs past observation 15,000. Figure 4 below shows the result of a simulation run with the same parameters, range of x, and N as the one displayed in Figure 3 except there are skips in the observations past observation number 15,000. Note that when the values of x extend to 1,000,000, for example, skipping 1,000 values would cut the number of iterations significantly, from 1,000,000 to 15,000 +990 =15,990, a 98 % decrease. This translates into cutting the simulation time by almost that percentage. The difference in time between the example displayed in Figure 3 and the example displayed in Figure 4 is the difference between 60 minutes and a little less than twenty minutes, a reduction in time of about 70 percent.

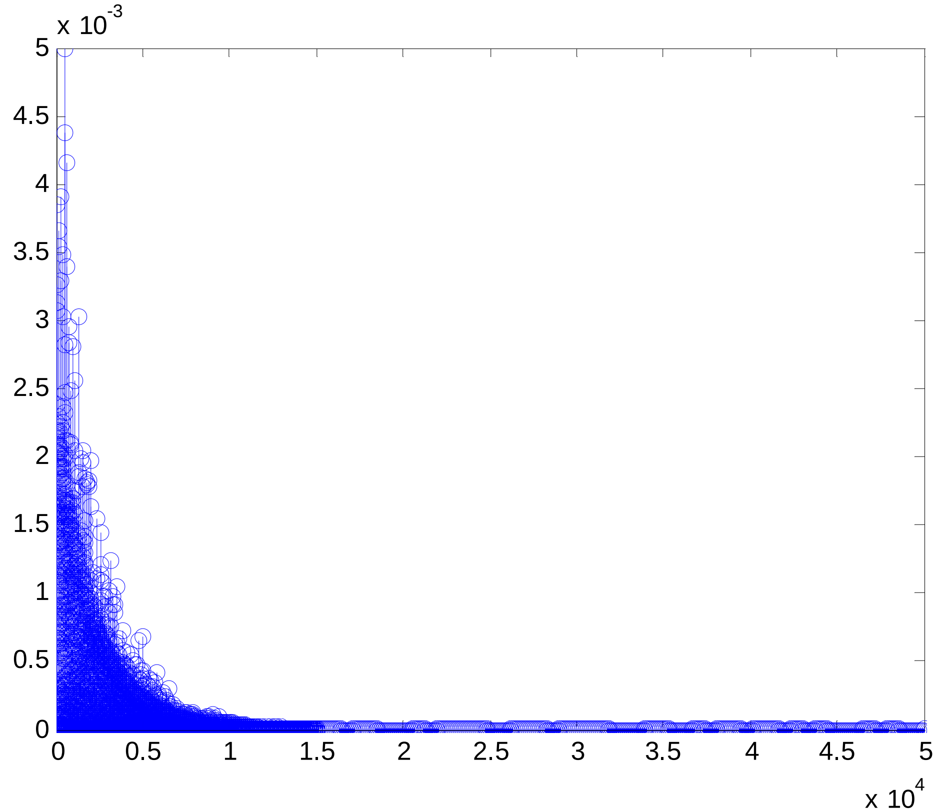


Figure 4. MC simulation run with skipping

The Matlab code for skipping is the following:

```
skip=700
for x=15000:skip:50000;
Sum=0;
   for i=1:N;
   Sum=Sum+intpre(rand(),x+skip/2);
         end
y(x+skip/2)=Sum/(N);
y(x:x+skip/2-1)=y(x+skip/2)*ones(1,skip/2);
y(x+skip/2+1:x+skip-1)=y(x+skip/2)*ones(1,skip/2-1);
```

end

The simulation example displayed in Figure 3 without skipping gives similar LCL and UCL to the LCL and UCL displayed in Figure 4 when "skipping" is used. At an $\alpha/2$ value of .005, the UCL and LCL are 20 and 119106, respectively, when regular simulation is used and 18 and 12422 when the "skipping" approach is used. Given the added randomness of simulations, these values are not very different, particularly for determining the LCL values. This skipping technique is generic and can be applied to other production processes.

4.3.1.2 Importance sampling (IS)

IS is an uncertainty reduction technique used to sample rare but "important" events. The study of rare events is important because their occurrence sometimes can have catastrophic consequences.

The premise of IS is to accurately estimate the probability related to rare events by accelerating (in our case through Monte Carlo simulation) the rate of their occurrence. Rare events are usually defined at the tail ends of probability density functions. It is often difficult to directly generate a large number of these rare events in a short simulation time period in order to make statistical inferences and draw meaningful conclusions. The difficulty can be resolved by deliberately altering the distribution that defines their behavior – which then would allow the speeding up of the sampling process and then adjusting the output of sampled values appropriately to compensate for the alteration.

IS is also called biased sampling because the sampling occurs not from the distribution of interest but from a biased distribution. The main property of the biased distribution is that it allows the sampling of "rare" events to occur more frequently, thus reducing the variance of the relevant estimates of interest. The biased simulation outputs are then weighted to adjust for the use of the biased distribution. Each weight is given by a likelihood ratio of the true underlying distribution with respect to the biased simulation distribution. In the situation of our interest, the application of the IS technique is done through the scaling method.

4.3.1.2.1 Mathematical background

Let X be a random variable on the set S and let f be a probability measure on S. Let G be the cumulative distribution of the random variable X and $g(x) = G'(x)$ be its probability density function. We consider estimating by simulation the probability f_t of an event $\{X \geq t\}$ where the value of t is such that the event is rare.

In IS, our interest is to find a biasing density function $g^*(x)$, such that it would allow the event $\{X \geq t\}$ to occur more frequently. In that case we would have:

$$f_t = E\{I(X \geq t)\} = \int I(x \geq t) \frac{g(x)}{g^*(x)} g^*(x) dx \tag{30}$$

$$= E^*\{I(X \geq t) W(X)\} \tag{31}$$

where

$$I(x \geq t) = \begin{cases} 1 & x \geq t \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad W(\cdot) = \frac{g(\cdot)}{g^*(\cdot)} \tag{32}$$

$W(\cdot)$ is called the likelihood ratio and is usually referred to as the weighting function.

The IS procedure generates i.i.d. samples from g^* and for each sample that exceeds t, the indicator function is multiplied by the weight W evaluated at the sample value. In the case where IS is done by scaling, the density $g^*(x)$ to be simulated is the density function of the scaled random variable aX, where $a > 1$ is used for tail probability estimation.

When using a scaling transformation:

$$g^*(x) = \frac{1}{a} g\left(\frac{x}{a}\right) \tag{33}$$

The weighting function can be expressed as:

$$W(x) = a \frac{g(x)}{g\left(\frac{x}{a}\right)} \tag{34}$$

4.3.1.2.2 The use of IS in MC simulation

IS is used to reduce the simulation time, and it is used also to increase the accuracy in the simulated values. Recall the way the Monte Carlo simulation is achieved: At x=1, N values from the uniform (0,1) distribution are generated. For each of the N values of p, the corresponding integrand (the value inside the integral) of the predictive distribution is calculated. Then the average of all the N integrands is computed. The average represents the value of the predictive distribution at x=1. The process is repeated for all values on the range of x. In the case where p is very small, sampling from the uniform (0,1) distribution yields very few values of p that are around that small p. So the smaller the p, the larger the N value has to be in order for the simulation to capture enough of the small values of p and, as a result, make the value of the integral (or the average) more representative of the true behavior of the system.

The way the biasing occurs here for small p is that instead of generating the observations from the uniform (0,1), they are generated from the $\frac{1}{a}$ uniform (0,1), where a >1. Instead of simulating so many observations that require much time and do not contribute much to the integral or the predictive distribution, the process is sped up significantly by sampling from a fraction of the uniform (0,1). To compensate for the fact that only a fraction of the uniform (0,1) is considered, the fraction is expanded by the factor a to cover the whole integral.

The term $(1-p)^{\sum_{i=1}^{n} x_i + x_{n+1} - n + \tau - 2}$ in equation (29) affects the value of the predictive distribution $y(x)$ the most since its exponent assumes very large values. If p is close to 1, $(1-p)^{\sum_{i=1}^{n} x_i + x_{n+1} - n + \tau - 2}$ is close to zero since the value of $1-p$ is close to zero and the exponent is large, particularly if the value of τ is also large. Also, the exponent becomes increasingly larger due to the nature of $\sum_{i=1}^{n} x_i$ which represents the cumulative value of geometric random variable x (the number of conformances up to the nonconformance) where p (the probability of a nonconformance) is very small and therefore x is very large, and where the index i represents the observation number. So for a p value close to 1, the corresponding y is value is very close to zero, and therefore does not contribute to the information about the predictive distribution y. If p is close to zero then $1-p$ is close to one and the exponent being very large would tend to contribute the most to the value of the predictive distribution y. The larger τ and $\sum_{i=1}^{n} x_i$ are, the larger the factor a should be, i.e., the more it is desired that the simulated p value from uniform (0,1) is close to zero. This means that overall a smaller value of N or fewer number of iterations are needed to get a representative average which in turn represents the predictive distribution value at a particular value of x. The fewer number of iterations translates into an increased simulation speed to get the predictive distribution.

4.3.1.2.3 Testing speed results using IS

One thousand simulation runs are performed at N=100,000 (sample size), and a=1 (i.e., do not use IS) and another one thousand runs are performed at N=100 ,a=1000 and x=1 (IS is applied). Four example Y values from each setting are given in rows three through six of Table I, and the sample mean and standard deviation of both sets of 1000 simulations are listed in the last row.

Table I. Simulation results when IS is used

N=100,000, x=1, a=1, no IS	N=100, x=1, a=1000, IS
Four typical samples	**Four typical samples**
Y=2.8295 × 10^{-5}	Y=1.262 × 10^{-5}
Y=1.8938 × 10^{-5}	Y=3.5812 × 10^{-6}
Y=3.5114 × 10^{-6}	Y=2.243 × 10^{-6}
Y=4.557 × 10^{-8}	Y=4.25 × 10^{-5}
Resulting sample statistics	**Resulting sample statistics**
Mean= 1.0682 × 10^{-5}, SD=1.1072 × 10^{-5}	Mean= 1.0429 × 10^{-5}, SD=3.0536 × 10^{-6}

The results presented in Table I suggest that IS has a better performance in terms of variability but the computation times of the two approaches differ markedly. Indeed, each observation in the first column took about six seconds, while each observation obtained from IS was virtually instantaneous. The difference between N=100,000 and N=100 suggest that the speed is increased about 1,000 fold, or that the observations in the second column can be obtained 1000 times faster.

When a=1000, for N=100, all 100 samples will be picked from the interval (0, 0.001). By increasing the IS factor from a=1 to a=1000, the number of iterations is reduced by a factor of 1000 (from 100,000 to 100), and it is in favor of reduced variability.

4.3.1.3 Truncating the simulation

Since the probability mass function used in defining the CL will approach zero, it is not necessary to run the simulation every time over the whole countable set of x_j values, where the index j could get to a value over 1,000,000 for very small probability of nonconformance p. Furthermore, since the interest is only in process deterioration, it is enough to run the simulation until the LCL value is reached at the α significance level. When the simulation is truncated or when the LCL is reached, the process time is decreased by almost 100 percent (see example in Section 4.10.)

4.4 Example using the Bayesian procedure

In this example there are 23 observations (as is defined in the introduction section on page 2, an observation is defined as the number of conforming units up to the non-conforming one). A shift occured after observation 17, where the geometric distribution representing the observations shifted from the parameter estimate of *p* (representing the probability of nonconformance) of .00001 to .0001 starting from the 18^{th} observation. A simulation was conducted to first extract 17 observations from a geometric distribution with *p* =.00001. The values are:

128,797; 25,542; 214,715; 105, 614; 79,229; 95,106; 125,338; 26,673; 138,363; 49,082; 503,939; 35,444; 118,245; 66,508; 32,200; 119,391; 241,047.

Afterwards, the next 6 observations were obtained from a geometric distribution with the value of *p* =.0001, representing a deteriorative shift in the process. The 6 values are: 538; 29,041; 246; 17,375; 12,153; 7,820.

4.4.1 Generating the geometric observations

The program used for generating the observations from the geometric distribution was developed as follows: Function f = geo(*p*), where "geo" is the name of the function, takes as input a value for *p* and as output a value of the random variable *x*. The random variable represents the number of conformances up to (and including) a nonconformance. The parameter *p* represents the probability of a nonconformance. Sampling from a uniform random distribution (0,1), every time the sampled value is greater than the inputted value of *p*, it is added to a counter that increases the number of conformances. Once the value sampled is less than *p*, the counter is stopped. The value at the final counter is the output value of the random variable *x*.

The Matlab program is represented as follows:

Function f=geo(p);

sum=0;

```
f=-1;
while f = = -1;
x=rand( );
if x ≥ p;
sum = sum +1;
else;
f=sum;
end;
end;
```

4.4.2 Data analysis of the Bayesian procedure

In this dissertation, an important aim of this new Bayesian procedure is to be able to detect process deterioration with more sensitivity than what is available through the traditional CCC methodology. For that reason, the concentration is on the analysis on the LCL of the control chart. In the table below, the process deterioration (from p = .00001 to p = .0001) takes place in the last 6 observations, from observation n=18 through n=23.

The first 3 observations 128,797; 25,542 and 214,715 are used to compute the original parameters σ and τ of the predictive distribution listed under equation (29). Using equation (19), the estimate of p is found by inverting the mean of the three observations and is equal to:

$\hat{p} = 8.1288917069 \times 10^{-6}$

According to Gido and Clements (2006, p 186), the variance of the p estimate is evaluated from the most optimistic and most pessimistic estimates of p. Here we use the highest and lowest observation values of x to get the most optimistic and most pessimistic estimates of p.

They are 214715 and 25542, respectively. Using these two estimates in equation (20) yields a variance estimate of $3.0507432 \times 10^{-11}$. The values of σ and τ are then computed using equations (17) and (18) which yield the values of σ = 1.99929 and τ = 245947.

MC simulation is performed to get the values of the integral of the predictive distribution defined in equation (29) based on the given parameters. The values of *x* from both ends of the distribution that correspond to the cumulative *y* values of $\alpha/2$ = .005 and 1-$\alpha/2$=.995 represent the UCL and LCL, respectively.

4.4.3 Analysis results of the Bayesian procedure

The values of the UCL and LCL are presented in Table II. The two-sided control limits are calculated at a confidence level of $\alpha/2$ = .005. The first 3 observations are used to obtain the values of σ and τ and in turn find the original control limits at *n*=0 (in the table, *n*=0 is labeled as *nb*, for base observation, before any observation is tested). When the 18th observation is tested, it falls below the lower control limit (538 is below 624) and therefore gets removed. The next step is to compute a new σ and τ values from the previous 17 observations. Then the new parameters values are used in updating the prior and the predictive distribution and the corresponding control limits where the counter is reset at zero and *n* has a base value of *nb*=0. The 19th observation is now tested against the control limits and the process continues. Observation number 20 is also below the control limit and therefore gets removed. New σ and τ are reevaluated, the prior re-updated to get the corresponding control limits and the counter is reset again at zero.

Note: The new LCL is computed from the predictive distribution based on a prior that uses all 17 observations to establish its parameters. The fact the value of the LCL at the base value n=0 (622) is closer to the preceding LCL values than the distant ones is indicative that more weight is put on more recent observations than distant ones and that this new prior adequately represents the information from historical data.

Table II. Bayesian upper and lower control limits

$\alpha/2$ = .005	Bayesian LCL	Bayesian UCL
x1=128,797		
x2=25,542		
x3=214,715		
nb=0,	617	1136048
n=1,x4=105,614	593	961005
n=2, x5=79,229	533	899668
n=3, x6=95,106	526	877415
n=4,x7=125,338	539	921141
n=5,x8=26,673	481	755560
n=6,x9=138,363	485	776994
n=7, x10=49,082	478	678230
n=8,x11=503,939	699	857738
n=9, x12=35,444	643	833737
n=10,13=118,245	642	809486
n=11, x14=66,508	617	772505
n=12, x15=32,200	581	747217
n=13,x16=119,391	584	753390
n=14,x17=241,047	624	792480
n=15,x18=538	below LCL	
nb=0,	622	1482319
n=1, x19=29,041	446	1008328
n=2, x20=246	below LCL	
nb=0,	595	1040548
n=1, x21=17,375	415	918035
n=2, x22=12,153	325	677592
n=3, x23=7,820	267	536120

The first three observations are used to get an estimate of p and a variance of the estimate but do not get tested against the computed control limits to avoid the "double-use of data". After the 18th observation or first observation from the "shifted" process is identified and removed, the previous seventeen observations are used to compute the new estimate of p, the probability of nonconformance. The estimate is: $\hat{p} = \dfrac{1}{\dfrac{128{,}797 + ... + 105{,}614 + ... + 241{,}047}{17}}$ =8.07511567 x 10^{-6}.

The most optimistic and pessimistic values of the p estimates are now 1/503939 and 1/25542. They are used in equation (20) to get an estimate of the variance. The variance estimate is 3.83714891122 x 10^{-11}. The estimates of the mean and variance of p are used in equations (17) and (18) to get the values of σ and τ : 1.69935 and 210441, respectively. These are the new parameter values of the updated prior. At n=0, the LCL is 622, while at n=1 it is 446. At n=2, observation 246 is below the LCL (446) set by observation n=1. Observation 246 is also discarded and a new updated prior with new parameters are generated, this time with 18 observations. The new estimate of p is:

$\hat{p} = \dfrac{1}{\dfrac{128{,}797 + ... + 105{,}614 + ... + 241{,}047 + 29{,}041}{18}}$ =8.43378 x 10^{-6} but the estimated variance is the same (at 3.83714891122 x 10^{-11}). The new values of the parameters σ and τ are 1.85366 and 219788, respectively.

4.5 Example using the CCC procedure

The estimate of p is $\hat{p} = \dfrac{1}{\dfrac{128{,}797 + ... + 105{,}614 + ... + 7820}{23}}$ = 1.05873 x 10^{-5}

From equation (5) and (6), the control limits at $\alpha / 2$ = .005 are:

$$LCL = \left\lfloor \frac{\ln(1-.005)}{\ln(1-\hat{p})} \right\rfloor + 1 = 474 \text{ and } UCL = \left\lceil \frac{\ln(.005)}{\ln(1-\hat{p})} \right\rceil = 500439$$

4.5.1. Analysis results of the CCC procedure

Only one observation value of 246 is below the LCL value of 474. The observation value is removed and the new estimate of p reevaluated with the remaining 22 observations. The new estimate of p is: 1.012167 x 10^{-5}. The LCL value is now 496. The UCL also shifts to 523461. With the "out-of control" observation eliminated, the lower control limit value is now higher. Even with a higher value, no new shifted observations are detected. All the remaining 22 observations are "in-control." Table III below presents the results.

Table III. UCL and LCL using the CCC procedure

$\alpha/2$=.005	LCL	New LCL	UCL	New UCL
n=1, x1=128797	474	496	500439	523461
n=2, x2=25542	474	496	500439	523461
n=3, x3=214715	474	496	500439	523461
n=4,x4=105,614	474	496	500439	523461
n=5, x5=79,229	474	496	500439	523461
n=6, x6=95,106	474	496	500439	523461
n=7,x7=125,338	474	496	500439	523461
n=8,x8=26,673	474	496	500439	523461
n=9,x9=138,363	474	496	500439	523461
n=10,x10=49,082	474	496	500439	523461
n=11,x11=503,939	474	496	500439	523461
n=12,x12=35,444	474	496	500439	523461
n=13,x13=118,245	474	496	500439	523461
n=14, x14=66,508	474	496	500439	523461
n=15, x15=32,200	474	496	500439	523461
n=16,x16=119,391	474	496	500439	523461
n=17,x17=241,047	474	496	500439	523461
n=18, x18=538	474	496	500439	523461
n=19,x19=29,041	474	496	500439	523461
n=20, x20=246				
below LCL				
n=21, x21=17,375		496		523461
n=22, x22=12,153		496		523461
n=23, x23=7,820		496		523461

4.6. Example using a sequential CCC procedure

Using a sequential methodology where the estimate of p is updated with every new observation, the UCL and LCL are no longer fixed but dynamic, changing with every observation. The formulas for the UCL and LCL are based on the equations used in the CCC methodology and displayed under equations (5) and (6). The estimate of p is $\hat{p} = \frac{1}{E(x)} = \frac{n}{\sum_{i=1}^{n} x_i}$. Because that estimate, under a sequential procedure changes with the flow of observations, the LCL and the UCL are now of the form:

$$LCL = \left\lfloor \frac{\ln(1-\alpha/2)}{\ln\left(1-\frac{n}{\sum_{i=1}^{n} x_i}\right)} \right\rfloor + 1 \tag{35}$$

$$UCL = \left\lceil \frac{\ln(\alpha/2)}{\ln\left(1-\frac{n}{\sum_{i=1}^{n} x_i}\right)} \right\rceil \tag{36}$$

The value $\sum_{i=1}^{n} x_i$ is the sum of the observations indexed by i and n represents the total number of observations. The sequential methodology developed here is similar to the Bayesian methodology. The observations are tested sequentially. If an observation falls outside the control limits, it is removed and the previous tested observations are used to get a new estimate of the rate of nonconformance p.

4.6.1 Analysis results of the sequential CCC procedure

In the sequential CCC chart, the same observations that were detected under the Bayesian methodology are detected also here. They are the second and the fourth observations from the shifted six observations. Table IV provides the two-sided control limits of the sequential CCC procedure at $\alpha/2$ =.005 level

Table IV. UCL and LCL using the simple sequential procedure

$\alpha/2$ = .005	Seq. LCL	
n=1, x1=128797	647	682405
n=2,x2=25542	388	408866
n=3,x3=214715	618	651786
n=4,x4=105,614	596	628733
n=5, x5=79,229	556	586942
n=6, x6=95,106	543	573101
n=7,x7=125,338	555	586098
n=8,x8=26,673	503	530501
n=9,x9=138,363	524	553010
n=10, x10=49,082	496	523714
n=11,x11=503,939	681	718834
n=12, x12=35,444	639	674580
n=13,x13=118,245	636	670881
n=14, x14=66,508	615	648131
n=15, x15=32,200	584	616296
n=16,x16=119,391	585	617313
n=17,x17=241,047	622	656126
n=18, x18=538	below LCL	
remove x18		
n=19, x19=29,041	595	628223
n=20, x20=246		
remove x20,		
n=21, x21=17,375	569	600003
n=22, x2=22,153	543	573223
n=23, x23=7,820	519	547899

4.7 Comparison of the three procedures

Table V below shows the combined numerical results of the three procedures.

Table V. Comparison of CL between Bayesian, CCC and Sequential procedures

$\alpha / 2 = .005$	Bay. LCL	CCC LCL	Seq. LCL	Bay. UCL	CCC UCL	Seq. UCL
n=1, x1=128,797		496	647		523461	682405
n=2, x2=25,542		496	388		523461	408866
n=3, x3=214,715		496	618		523461	651786
nb=0	617			1136048	523461	
n=4, x4=105,614	593	496	596	961005	523461	628733
n=5, x5=79,229	534	496	556	899668	523461	586942
n=6, x6=95,106	527	496	543	877415	523461	573101
n=7,x7=125,338	540	496	555	921141	523461	586098
n=8,x8=26,673	481	496	503	755560	523461	530501
n=9,x9=138,363	485	496	524	776994	523461	553010
n=10, x10=49,082	478	496	496	678230	523461	523714
n=11,11=503,939	699	496	681	857738	523461	718834
n=12, x12=35,444	643	496	639	833737	523461	674580
n=13,13=118,245	642	496	636	809486	523461	670881
n=14, x14=66,508	617	496	615	772505	523461	648131
n=15, x15=32,200	581	496	584	747217	523461	616296
n=16,x16=119,391	584	496	585	753390	523461	617313
n=17,x17=241,047	624	496	622	792480	523461	656126
n=18,x18=538	below LCL	496	below LCL		523461	
nb=0	622			1482319		
n=19, x19=29,041	446	496	595	1008328		628223
n=20, x20=246	below LCL		below LCL		523461	
nb=0,	595			1040548	523461	
n=21, x21=17,375	415	496	569	918035	523461	600003
n=22, x22=12,153	325	496	543	677592		573223
n=23, x23=7,820	267	496	519	536120		547899

Since our interest is in process deterioration, the concentration is on the lower control limits. When looking at the lower control limits, the following observations are in point:

1. The local jumps are more extreme in the Bayesian LCL than in the sequential CCC since the Bayesian procedure incorporates information about the variance between the observations. For example, using Table V, a drop of 98,000 units between observation *n*=7 and *n*=8 corresponds to a drop in the Bayesian LCL of 59 units (540-481) versus a drop in the sequential CCC of 51 units (555-503). Alternately, a rise in the observation values from 49082 to 503939 (of 454857 units) causes a rise in the Bayesian LCL of 221 units versus 184 units only in the CCC LCL.
2. The variations or shifts in the LCL are more extreme at the earlier stages of the charting process than towards the later stages. In other words, for the same amount of observational shifts, earlier observational shifts translate into more variability in the LCL than later ones. For example, a drop of about 25thousand observational units between observation one and two causes a dip of almost 60 units in the Bayesian LCL. Subsequently, another drop between observation 11 and 12 of about 35 thousand units causes the lowering of the Bayesian LCL of only about 36 units. This suggests a tightening of the control limits as more observations are plotted and the number of observations increases.
3. In the Bayesian procedure, the order of the observations affects the LCL values and therefore the whole analysis. For example, the much better than expected performance of observation number 8 and value 503,939 carries the next four LCL to values higher than the respective sequential CCC LCL. This is because the Bayesian process also behaves as an exponentially weighted moving average process, where the later observations have more weight than the earlier ones.

4. The integration of the variance in the Bayesian methodology allows for a systemic shift in future observations to be easily detected - like the ones caused by deterioration due to machine wear or malfunction. The fact that the variability among the observations is considered in the Bayesian methodology causes the ability to detect visually a shift in the process, contrasted to the sequential CCC methodology where the values of the shifted observations are diluted back into the average mix. One can see the steep drop of the Bayesian LCL that corresponds to the deterioration represented by the last four observations (446,415, 325, and 267), compared to the last four LCL of the sequential CCC procedure (595, 563, 543 and 519). These last observations of the sequential procedure are not much different in value of other lower control limit values, contrasted with the steep descent in values of the Bayesian LCL values, which makes the process deterioration easy to see.

Figure 5 below is the visual display of the tabulated results. One can see that:

1. The first three observations are used in the Bayesian procedure to formulate parameters of the prior and are not tested.
2. Observations 18 and 20 fell outside the lower control limits. The gap is made visible by assigning them an arbitrary value of zero.
3. Adding the UCL to the graph is not meaningful when assessing process deterioration
4. In observations 12 through 17, only one line is visible because the blue and red lines merge together. The observations on that range are very close.

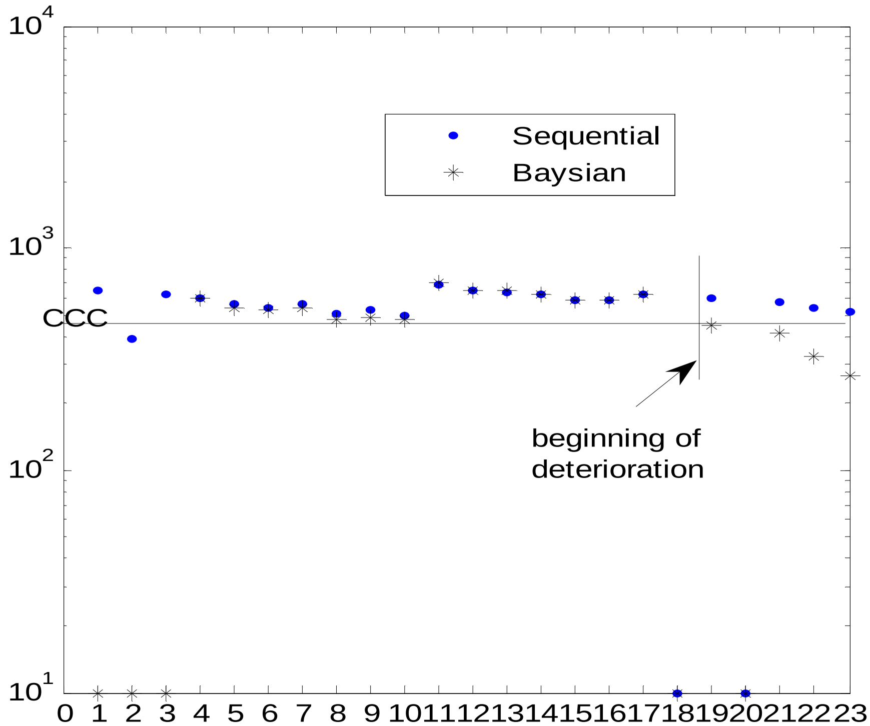


Figure 5. Comparisons of LCL between Sequential and Bayes procedures

Note: Assumptions are made here in the analysis and the interpretation of a deteriorating process. The first assumption is that a detection of an out of control observation (below the LCL) does not translate necessarily into deterioration detection. Narrowing the source of variation might not be easy or evident. Breyfogle III (2003, p. 225) states regarding the out-of-control points: "These points should have an assignable cause for being outside the control limits. However, in general, determining the real cause after some period of time may be impossible." Sometimes at the onset of process deterioration, it is not possible to immediately make the determination that a deterioration shift has taken place even when one or two observations are found to be out of control. This is particularly true in cases related to epidemic outbreaks. To make a comparison between two sequential procedures to see which one has a lower type I error - lower false alarm rate, i.e. a better ability to detect an out-of control observation is **moot** in this setting not just because an out-of control observation does not translate into a deteriorated process necessarily, but because the two methodologies are very different in nature and not easily comparable. In the Bayesian approach, a control chart has an increased ability of detecting an out-of control observation when the system has been behaving better than expected (high levels of conformances push the future LCL to higher levels, as in a dynamic system.) The changes in the LCL values in the sequential Bayesian CCC approach are strictly based on the observations encountered. In othe words, it is strictly data driven. In the simple sequential procedure, the ability of the LCL to better detect an out-of-control observation depends on the methodology used.

The interest in this dissertation is in settings where the cost of committing a type II error is devastatingly high (The probability of stating that no deteriorating shift has occurred when in fact one did occur.) For that reason, a guideline is devised to evaluate which sequential procedure mirrors best process deterioration when it occurs. The CCC LCL can be viewed as a benchmark LCL for both sequential processes and the values of the LCLs for both sequential procedures can be viewed as **derived observations**. Before deterioration occurs, both the sequential and the Bayesian LCL are above the CCC LCL line. This translates into a non-existent rate of false alarm about process deterioration. At the onset of process deterioration, the observations of the simple sequential procedure are not affected by the deterioration shift whereas the observations of the sequential Bayesian procedure all fall below the benchmarked line. Because of the dynamic nature of the sequential Bayesian CCC methodology, the future observations (not plotted) of a deteriorated process are also expected to fall below the CCC LCL even if there is some fluctuation in the actual observations. As if the actual extreme observations give a dynamic pull to the derived observations. The pull would me more extreme and the rate of decrease of the derived observations steeper if the out-of control observations are not removed from the computations, because they represent the most extreme (smallest) values.

4.8 More on the sequential Bayesian CCC procedure

The simple sequential methodology, despite its ability to capture an “out of control” observation, is not capable of detecting process deterioration. In a setting where a change in environment, machine wear and tear or malfunction could cause deterioration in the process, immediate attention is required and a halting of the process is needed. It becomes very critical then to be able to timely detect deterioration in the system as the production proceeds. The ability of the Bayesian methodology to detect process deterioration makes it a very valuable tool.

In the previous example where the three procedures are compared, the range of x is from 1 to 12,000,000. This is due to the small value of p tested (with the level of 10^{-5}). This tested value is smaller than the rates of nonconformance often used or tested in the literature regarding high yield processes. The smallest estimate of p used in the literature was in the order of 4×10^{-4}. Skipping was conducted at a value of 700, N at 1000 and the value of a used in IS is 1000. The time for getting the UCL and LCL from each distribution is about five hours (using a laptop with 1.75 GB of memory and a 64×2 MB processor).

4.9 Increasing speed by truncating the simulation

Because our goal is to detect process deterioration, the UCL is not of our interest and is not to be considered in the Bayesian methodology. As discussed in Section 4.4.1.3 of this dissertation, the simulation is truncated when the LCL is reached at the desired significance level α. The algorithm stops when the cumulative value of the predictive distribution reaches the value α. The time that takes to get the LCL is less than a minute as opposed to five hours to get both the LCL and the UCL.

4.9.1 Comparison of LCL values

Identical simulation runs are now performed with the same value of *a*, *N,* and the level of significance of .005 as the ones used in Section 4.5 except the simulation run here is allowed to stop when the lower significance level of .005 is reached, which indicates the level of the LCL is reached.

Table VI presents the two sets of LCL values: one set is from Table II above (where the LCL are extracted from the whole distribution) and the other set is the result of this truncating approach. The two sets of values are almost identical. This is somewhat expected, because of the iterative nature of the computations, given the model is an appropriate one. The time to reach the LCL in this approach was less than twenty seconds, which is fairly fast.

Considering the randomness in performing simulations, the sets of values are very close. This shows that by stopping the simulation when LCL values are reached, simulation can be hastened and the information of interest is not lost. Figure 6 is a visual display of these comparisons.

Table VI. Comparisons of LCL when full and truncated simulations are used.

.005	LCL from whole dist.	LCL from truncated dist.
n=1, x1=128,797		
n=2, x2=25,542		
n3, x3=214,715		
nb=0	617	613
n=4, x4=105,614	593	594
n=5, x5=79,229	534	538
n=6, x6=95,106	527	529
n=7,x7=125,338	540	559
n=8,x8=26,673	481	489
n=9,x9=138,363	485	512
n=10, x10=49,082	478	478
n=11, x11=503,939	699	688
n=12, x12=35,444	643	643
n=13,13=118,245	642	643
n=14, x14=66,508	617	605
n=15, x15=32,200	581	587
n=16, 16=119,391	584	586
n=17, 17=241,047	624	630
n=18, x18=538	below LCL	below LCL
nb=0	622	629
n=19, x19=29,041	446	444
n=20, x20=246	below LCL	below LCL
nb=0,	595	601
n=21, x21=17,375	415	417
n=22, x22=12,153	325	326
n=23, x23=7,820	267	264

Note: In Figure 6 below, the visibility of only one point at an observation (number 7 for example) means the two control limit lines are at an identical or almost identical locations. In addition, the word "truncated," for lack of a better word, means that only the values of interest are extracted through simulation. For the purpose of the dissertation, only the LCLs are of interest. The model has been validated and there is no need to run every simulation up to its full cumulative distribution value of one.

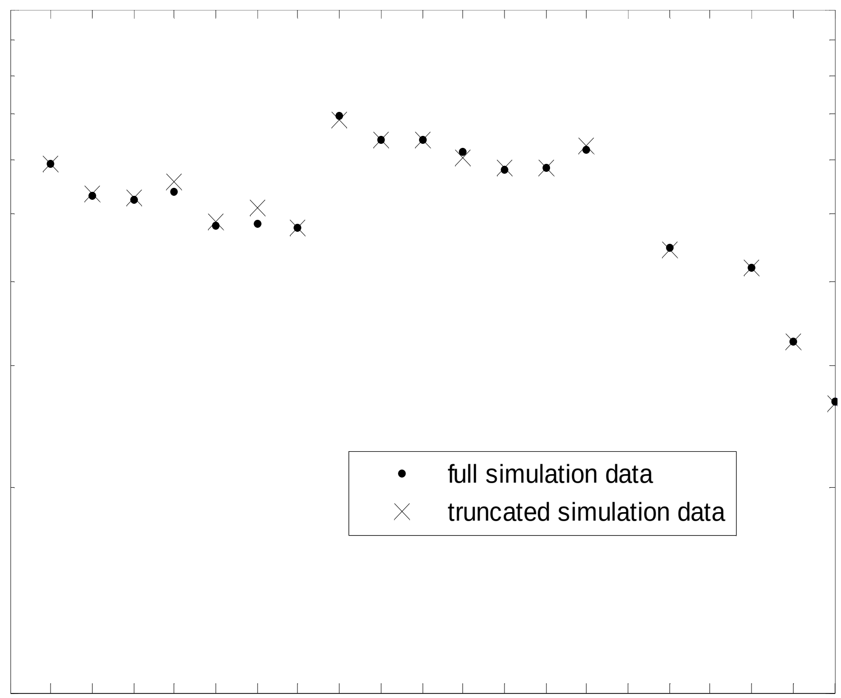


Figure 6. Comparison of LCL between truncated and full simulation

4.10. Extended application of the Bayesian procedure

In the previous sections, the analysis was tested for small rate of conformance, at the level of 10^{-5}. In this section the analysis will be tested for a rate of nonconformance "not very small". Furthermore, this section tests the situation when the prior is updated regularly (after 10 tested observations are deemed to be "in control") versus the situation when no updating occurs.

4.10.1 When the rate of nonconformance is not small

To check if the sequential Bayesian CCC procedure works for larger and more generic values for the rate of nonconformance *p*, the methodology is tested for values of *p* varying from .01 to .05 to .1.

4.10.1.1 No updating of prior in the absence of "out-of-control" observation

The values of the tested rate of nonconformance here start at an estimated value of .01 and drop incrementally to .05 and then to .1. This translates into small number of conforming values between nonconforming ones. Because the numbers are small (the values of the random variable *x*) and the deterioration rate is small, the generated values are assumed to come from a somewhat stable system, based on an average of 10 observations from the geometric distribution. Thirteen observation values of the random variable *x* are generated from a geometric distribution with an estimated rate of nonconformance p =0.01. They are the following: 43, 167, 96, 101, 118, 99, 125, 154, 39, 87, 80, 97 and 73. Process deterioration occurs where the estimated rate of nonconformance is now at .05. The process is simulated by generating nine observations from a geometric distribution with estimated p = .05. The values are: 9, 22, 17, 18, 8, 19, 23, 10, and 5. The last nine observations simulate further deterioration of the process to an estimated p= .1. The corresponding observation values of that further deteriorated process are: 3, 9, 8, 4, 7, 8, 6, 5, 7.

A Bayesian LCL "truncated" simulation analysis at a one-sided α = .05 is conducted for the first set of thirteen observations with an estimated p = .01. The first three observations are taken from that set. They are: 43, 167 and 96. They are used to get the mean and variance of the estimate of *p*. The mean and the variance of the estimated *p* are .0098 and 8.28268×10^{-6}, respectively. The values of the initial parameters σ and τ of the beta distributed prior are 11.47

and 1159.12, respectively. Using a high value of N =10,000 (the number of times the simulation from a uniform (0,1) is repeated before an average is taken) in the MC simulation yields tighter LCL (with not much random fluctuation when the MC simulation is repeated.) Following the thirteen observations, the next nine observations from a deteriorating process are tested against the updated LCL. There is a sharp decrease in the LCL values from 6 to5 and then to 4 that reflects the deterioration in the process. Here, because no observation value was found to be outside the LCL, no updating of the prior distribution would be performed. Also, the parameters of the prior are not updated even after 19 observations are tested. This means that the same parameters of the prior are used in the evaluation of the LCL and in the testing of the incoming observations. They are the same parameters computed using the first three observations. There is more weight put on the earlier observations rather than on the most recent ones (the process might be behaving more like a moving average than an exponential moving average).

In the next section, it will be demonstrated that by a regular updating of the prior parameters and therefore the prior distribution, more weight will be put on more recent observations and process deterioration will become much more visible. The numerical results are displayed in Table VII below.

4.10.1.2 Updating of prior in the absence of “out-of control” observation

Some guidelines are presented in this dissertation: If after 10 observations are tested, none of them falls outside the LCl, an updating is done based on all previous observations. In the case of this example, using the 10 tested observations plus the original three give us thirteen observations to use for computing the updated values of σ and τ. These updated σ and τ

values are 9.52 and 927.2 respectively (the estimated updated nonconformance rate *p* is .0101642 and the estimated updated variance of that rate is 1.0728902×10^{-5})

Table VII. Process deterioration from .01 to .05 after *n*=13

α =.05,	Bayesian LCL
n=1, x1=43	
n=2, x2=167	
n=3, x3=96	
nb=0,	6
n=4, x4=101	6
n=5, x5=118	6
n=6,x6=99	6
n=7,x7=125	6
n=8,x8=154	6
n=9,x9=39	6
n=10,x10=87	6
n=11,x11=80	6
n=12,x12=97	6
n=13, x13=73	6
n=14, x14=9	5
n=15, x15=22	5
n=16, x16=17	5
n=17, x17=18	5
n=18, x18=8	5
n=19, x19=19	5
n=20, x20=23	4
n=21, x21=10	4
n=22, x22=5	4

A better representation of the deterioration process represented by the last nine observations is visible through the LCL values. In this situation an updated prior is computed after 10 consecutively tested observations are deemed to be “in control”. A comparison of the values with an updated prior and without an updated one is displayed in Table VIII below.

Table VIII. Comparison of two Bayesian procedures for process deterioration detection

α =.05	Bayes LCL with adjusted prior	Bayes LCL without adjusted prior
n=1, x1=43		
n=2, x2=167		
n=3, x3=96		
nb=0,	6	6
n=4, x4=101	6	6
n=5, x5=118	6	6
n=6,x6=99	6	6
n=7,x7=125	6	6
n=8,x8=154	6	6
n=9,x9=39	6	6
n=10,x10=87	6	6
n=11,x11=80	6	6
n=12,x12=97	6	6
n=13, x13=73	6	6
nb=0	5	5
n=14, x14=9	5	5
n=15, x15=22	5	5
n=16, x16=17	4	5
n=17, x17=18	4	5
n=18, x18=8	4	5
n=19, x19=19	4	5
n=20, x20=23	4	4
n=21, x21=10	3	4
n=22, x22=5	3	4

The dipping of the control limits is more visible in the case where the prior distribution had updated parameters. This example reinforces the findings discussed earlier which state that after the prior is updated, the control limit become more representative of the more recent observations than the earlier ones. This also validates even further the guideline used in this dissertation which calls for updating the information of the prior regularly.

4.10.1.3 The sequential LCL

Using the same observations, LCL are obtained through the sequential methodology.

Table IX. Sequential Process deterioration detection

α =.05,	Sequential LCL
n=1, x1=43	3
n=2, x2=167	6
n=3, x3=96	6
n=4, x4=101	6
n=5, x5=118	6
n=6,x6=99	6
n=7,x7=125	6
n=8,x8=154	6
n=9,x9=39	6
n=10,x10=87	6
n=11,x11=80	6
n=12,x12=97	6
n=13, x13=73	6
N=14, x14=9	5
n=15, x15=22	5
n=16, x16=17	5
n=17, x17=18	5
n=18, x18=8	4
n=19, x19=19	4
n=20, x20=23	4
n=21, x21=10	4
n=22, x22=5	4

4.10.1.4 Comparison between Bayesian and sequential LCL

The sequential methodology, a benchmark methodology used to test the Bayesian one, is compared here with the Bayesian methodology. Here again, with values of p not very small (estimated at .01) the sequential Bayesian CCC methodology outperforms the simple sequential CCC methodology in detecting process deterioration, even when process deterioration is small (from .01 to .05). The reason is that, in the sequential Bayesian methodology, the LCL mirrors the behavior of the observations, in terms of time order, variability and size. It also assigns more weight to the more recent observations than the distant ones. It is not surprising that it performs better even for small values of nonconformance rate *p*. In the next section, comparisons between the two procedures are made when the estimate of *p* deteriorates from 0.01 to 0.1.

Table X. Comparison of two procedures for process deterioration detection

α =.05, *p* from .01 to .05	Bayesian LCL	Sequential LCL
n=1, x1=43		3
n=2, x2=167		6
n=3, x3=96		6
nb=0,	6	
n=4, x4=101	6	6
n=5, x5=118	6	6
n=6,x6=99	6	6
n=7,x7=125	6	6
n=8,x8=154	6	6
n=9,x9=39	6	6
n=10,x10=87	6	6
n=11,x11=80	6	6
n=12,x12=97	6	6
n=13, x13=73	6	6
nb=0	5	
n=14, x14=9	5	5
n=15, x15=22	5	5
n=16, x16=17	4	5
n=17, x17=18	4	5
n=18, x18=8	4	4
n=19, x19=19	4	4
n=20, x20=23	4	4
n=21, x21=10	3	4
n=22, x22=5	3	4

4.10.1.5 When deterioration is more extreme

In this example, the deterioration of the process is tested for a more extreme dip in the estimated rate of nonconformance *p*: from an estimate of 0.01 to 0.1. After observation 13, the observations drop to the following simulated values: 3, 9, 8, 4, 7, 8, 6, 5 and 7. Using the

Bayesian methodology, the first depressed observation valued at 3 is removed because it falls below the LCL value of 5 and an updated prior of the first 13 observations yields the same results for σ and τ (at 9.52 and 927.2, respectively) found in section 4.11.1.2 above. The second and third observations from the deteriorated process are above the LCL but the fourth observation (valued at 4) is below the LCL. After it is removed, a new updated prior based on the 15 observations is computed with parameter σ and τ valued at .325 and 27.75, respectively. Table XI below lists both sets of LCL, from the Bayesian and from the sequential procedure.

Both procedures picked the same observations to be “out-of-control”; observations number 14 and 17 valued at 3 and 4. Except the simple sequential procedure did not pick up any change in process deterioration when the estimate of p changed from .05 to 0.1. The LCL are the same when the deterioration is from a rate of nonconformance of .01 to .05 as they are when the deterioration is from .01 to .1. The Bayesian methodology shows a very sharp drop in the LCL, this time even more extreme than the earlier drop (from .01 to .05), mirroring here again the more extreme drop in the observation values and further validating the superiority of its performance. A visual graph of these comparisons is displayed in Figure 7 below.

Note: For the purpose of clarity in the graph, a linear scale was used when drawing the LCL values. (Values of 1 are almost on the x axis line when the log scale is used.) The deterioration in the last nine observations is clearly mirrored by the gradual then sharp decrease in the Bayesian LCL. This is in sharp contrast with the gradual decrease in the sequential LCL. The rate of decrease in the sequential LCL is not different from when the deterioration was represented by a drop in the estimate of p from .01 to .05. This shows that the sequential procedure lacks the sensitivity to detect small process deterioration.

Table XI. Comparison of procedures when deterioration is more extreme

α =.05, p from .01 to .1	Bayesian LCL	Sequential LCL
n=1, x1=43		3
n=2, x2=167		6
n=3, x3=96		6
nb=0,	6	
n=4, x4=101	6	6
n=5, x5=118	6	6
n=6,x6=99	6	6
n=7,x7=125	6	6
n=8,x8=154	6	6
n=9,x9=39	6	6
n=10,x10=87	6	6
n=11,x11=80	6	6
n=12,x12=97	6	6
n=13, x13=73	6	6
n=14, x13=3, below both LCC	5	5
nb=0	5	
n=15, x15=9	5	5
n=16, x16=8	5	5
n=17, x17=4, below both LCL	4	5
nb=0	4	
n=18, x18=7	2	5
n=19, x19=8	1	4
n=20, x20=6	1	4
n=21, x21=5	1	4
n=22, x22=7	1	4

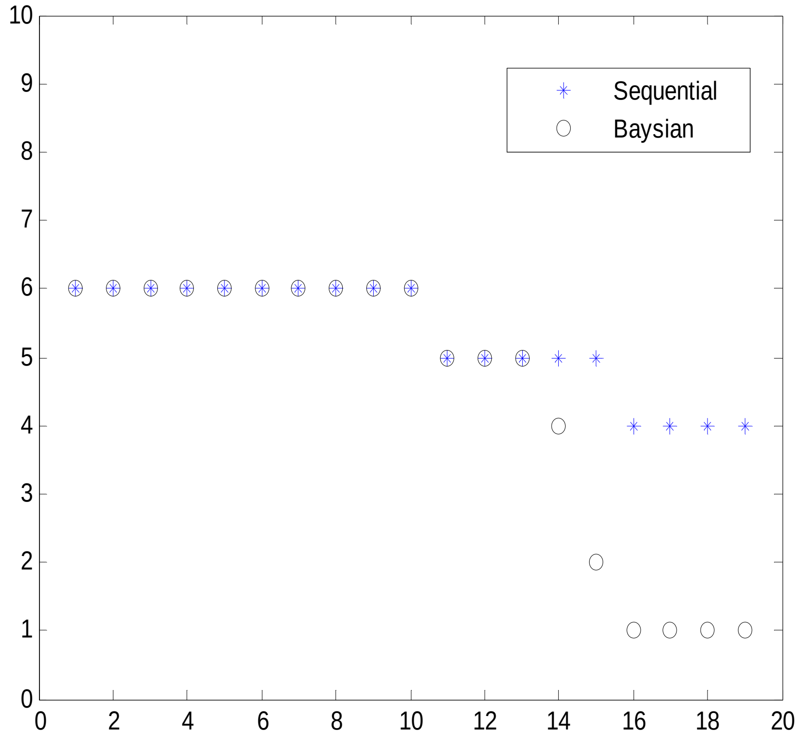


Figure 7. Comparisons of LCL between sSequential and Bayesian procedures

4.11 Example of a stable system

To simulate a small example from a stable system, ten observations are obtained from the average of 100 values simulated from a geometric distribution at an estimated rate of nonconformance p valued at .00001. The following ten values are obtained: 101550, 92667, 105310, 92420, 82364, 114980, 104960, 92216, 107650 and 103540.

To get an initial σ and τ values it is assumed this time that there is knowledge of the system's past behavior. The assumption is that the most pessimistic estimate of the process is 20,000, the most optimistic one is 200,000 and the most likely one is 100,000. From equation 12 and 13, the mean and the standard deviation of the observations are calculated and from equation

9 and 10, the values of σ and τ are found. They are : σ =1.77775 and τ =177,773. Table IVX gives the values of UCL and LCL for the system using all ten observations.

The following is noted about the CL: There is an overall tightening of the CL values and a convergence pattern. In the UCL column values, there is a sharp decrease in the LCL values and a visible convergence pattern. In the LCL column values, there is a marked decrease in variability among the LCL values because the system is stable and the observation values are closer, but there is also a slight tightening in the LCL values. One notes that between n=1 and n=2, a drop of a little over 8,000 observational units causes a drop of 8 LCL units, whereas between n=9 and n=10 a drop of about 4,000 observational units causes a drop of only one LCL unit. In a sustained stable system, the last control limits at observation n=10 could be used in the same way that fixed control limits from *phase I* of a process are used to test future observations of *phase II* of the process.

Table XII. Example of Bayesian UCL and LCL of a stable system

$\alpha / 2 = .005$	Bayes LCL	Bayes UCL
nb=0	504	1079181
n=1,x1=101550	503	986510
n=2, x2=92667	495	875182
n=3, x3=105310	499	794131
n=4,x4=92420	495	791787
n=5,x5=82364	485	728048
n=6,x6=114980	494	740302
n=7, x7=104960	500	715547
n=8,x8=92216	495	707084
n=9, x9=107650	499	665385
n=10,x10=103540	500	662948

Figure 8 is a graphic display of the behavior of the upper and lower control limits.

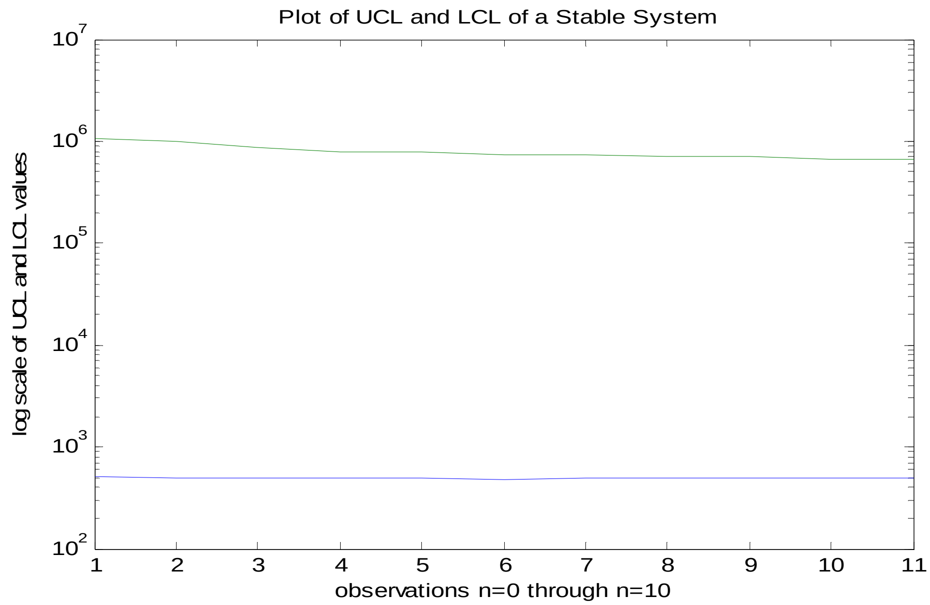


Figure 8. Plot of control limits from a stable system

4.12 Example with a change in the original guess

The original pessimistic, optimistic and most likely guess assumed for observed conformance value between nonconforming ones in section 4.12 changes slightly here. It is assumed now that the optimistic and pessimistic guesses are a little closer to the most likely value. Instead of a pessimistic estimate of 20,000 and an optimistic value of 200,000 for a conformance value between nonconforming ones, the guess is slightly less conservative: 30,000 for the most pessimistic estimate and 180,000 for the most optimistic one with the most likely estimate observation value still at 100,000. These new values lead to new estimates of σ and τ at 4.66552 and 466547, respectively. A comparison of the control limits with the two sets of σ and τ values show the control limits of the conservative guesses start wider but around observation number 5, both sets of control limits converge as more information is picked up from the observation values themselves. At observation n=10, the CL are almost identical. Table XIII gives the results.

Table XIII. Example of control limits behavior when the guess changes

σ =1.77775 τ =177,773	Bayes lcl	Bayes ucl	σ =4.66553 τ =466,547	Bayes LCL	Bayes UCL
nb=0,	504	1079181		501	818209
n=1,x1=101550	503	986510		504	854243
n=2, x2=92667	495	875182		501	777935
n=3, x3=105310	499	794131		498	790254
n=4,x4=92420	495	791787		496	716010
n=5,x5=82364	485	728048		485	679018
n=6,x6=114980	494	740302		493	663165
n=7, x7=104960	499	715547		495	631990
n=8,x8=92216	495	707084		495	628190
n=9, x9=107650	499	665385		499	606152
n=10,x10=103540	500	662948		500	640348

Figure 9 below is the visual display of the two different starting points and of the merging of the control limits with the number of observations. Data 1 comes from the slightly more conservative guess introduced in section 4.12 and data 2 comes from a slightly more accurate guess. Afterwards, the same observations are tested. The numbers from table VX and Figure 8 shows that even when the original guess is somewhat “off target”, after few observations, the control limits tighten and merge with the control limits representing the more “on-target” guess.

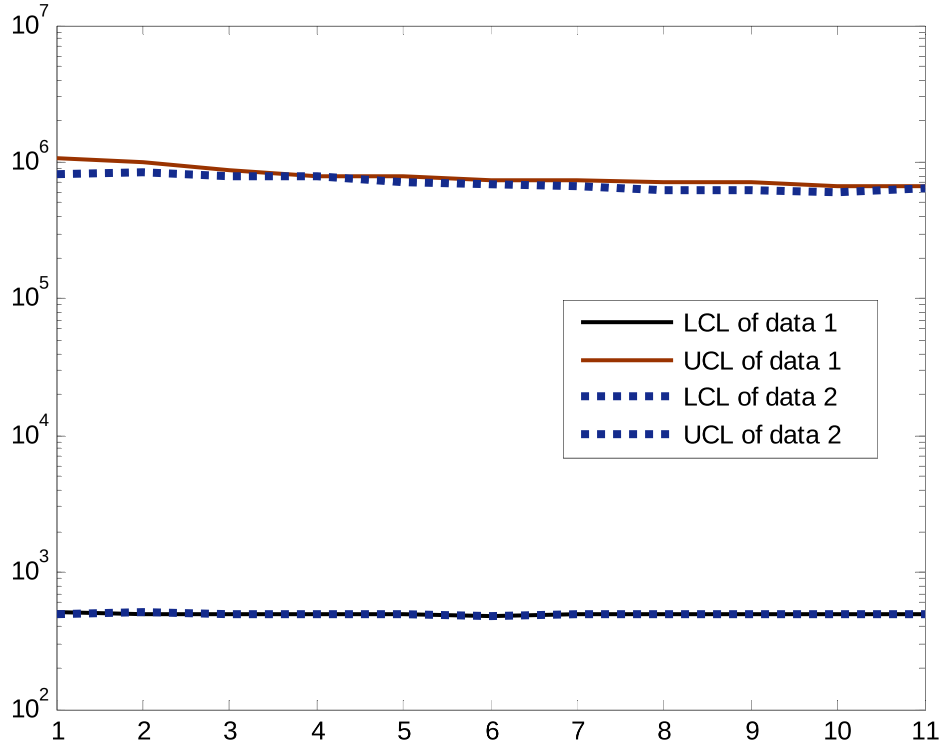


Figure 9. Merging of control limits with different starting points.

4.13 Guideline used in assessing process deterioration

A guideline is used in detecting a process shift inspired from the guideline used in detecting process shifts in classical control charts [2]except that the proposed guideline allows for deterioration **shift detection as soon as three observations have shifted** from the process. The guideline is as follows:

1. Any one observation within 15% deterioration or more from the previous observation.

[2] In Breyfogle III (2003), three lines are drawn above the centerline to denote the 1σ, 2σ and 3σ regions with lines denoting the corresponding zones: zone C, zone B and zone A respectively. Shift is detected when one of the following occurs: 1) One point beyond zone A, 2) Two out of three points in zone A or beyond, 3) four out of five points on zone B or beyond and 4) eight points in zone C or beyond.

2. Any two out of three consecutive observations within 10% deterioration or more from their previous observation.
3. Any three consecutive deteriorating observations with a percentage deterioration sum of 20% or more
4. Any three consecutive observations with 5% deterioration or more from their previous ones.

4.13.1 Five examples of a deteriorated process

The interest is in assessing the power of the sequential Bayesian CCC procedure in capturing process deterioration. That power is defined by the ability of the procedure to detect process deterioration when indeed a deterioration shift occurs. To assess process deterioration, the control limits are looked at as the **derived observations**. For the given examples, the same parameters of the prior are used when the first three observations were simulated from a geometric distribution with the parameter value of p at 10^{-5} in section 4.4.2. They are $\sigma = 1.99929$ and $\tau = 245947$. After the parameters of the prior are obtained, six observations are simulated from a geometric process with a rate of nonconformance p at 10^{-5} and the next four observations are simulated from a deteriorated process with a rate of nonconformance p at 10^{-4}. Because the Bayesian methodology is data driven, in order to control the variability of the data, the first six observations remain fixed while the simulations of the deteriorated process are conducted. First, the simulation is re-run with the values of σ and τ. The LCL for the first 6 observations (105,614; 79,229; 95,106; 125,338; 26673; 138,363) from a geometric process at p level of 10^{-5} are: 587, 539, 529, 547, 487 and **512**.

The results of five simulations of four deteriorated observations from a geometric distribution at a rate of 10^{-4} are obtained. The first sample of four observations yields: 1022, 5780, 250 and 2227 with corresponding LCL of 437, 450, 386 and 364 respectively. The first deteriorated observation (the 7th observation) has LCL is at 437, a ((512-437)/512)) 14.6% decrease (rounded to 15%) from the LCL of the 6th observation at 512. The second LCL value of 450 is at an increase of 3% from the previous LCL, the third LCL of 386 represents a decrease of 14.2% from its previous value and the fourth LCL of 364 represents a decrease of 5.7% from its previous LCL. The first deterioration (at 15%) meets guideline item number one. Also the deterioration of 15% and 14.2% meets guideline item number two. Therefore, based on the given guideline, **deterioration** in this instance **did occur**. Another simulation of four deteriorated observations yields four values of 2082, 3764, 21792 and 3950 from the geometric distribution at p = $10^{-4}$with corresponding LCL of 452, 406, 388 and 358 respectively and corresponding percentage deterioration decrease (from the previous LCL) at 11.7%, 10.2%, 4.4% and 7% respectively. Here guideline item number two and three are met concurrently Therefore the shift is detected in this instance also. The third set of four observations (13503, 5947, 8134 and 16789) yields the LCL of 480, 432, 390 and 385 and the corresponding four deteriorations of 6.25%, 10%, 9.7% and 1.3%. Based on guideline item number two (or three or four), a deterioration shift did occur. The fourth sample of four observations (17264, 2328, 21,317 and 3014) yields the LCL values of 490, 409, 393 and 375 respectively and the corresponding percent deterioration at 4.3%, 16.5%, 3.9% and 5.6%. Here guideline item number one is met and the conclusion is that deterioration did occur. Finally for the last sample of observations (985, 406, 14059 and 6724) the corresponding LCL are 432, 402, 378 and 345 and the corresponding deterioration percentages are 15.6%, 7%, 6.7% and 8.7%. Here again, based guideline item one (or three or four) the conclusion is that a deterioration in the process did occur. **All five simulations of a deteriorated process were identified by the guideline as being indeed from a deteriorated process**.

4.13.2 The power curve

Two hundred simulations of the 4 deteriorated observations are now performed. Based on these 200 simulations of the 4 deteriorated observations, when the process shifts from 10^{-5} to 10^{-4}, the number of times the shift is detected using the stated guideline in section 4.13 is 168 times out of 200. This corresponds to a power of the test at 84% or .84. The same number of simulations is repeated when deterioration is more extreme, at 10^{-3} from 10^{-5}. The power then is increased to .91. When the procedure is repeated for a deterioration level of 10^{-2} from 10^{-5}, the power is at .985. Figure 10 below represents the power curve. It is a visual interpretation of the power of the guideline in detecting process deterioration. As expected, the power curve representing the deterioration detection power increases with the increase in deterioration. The greater the level of deterioration , the higher the ability (or the power) of detecting it is.

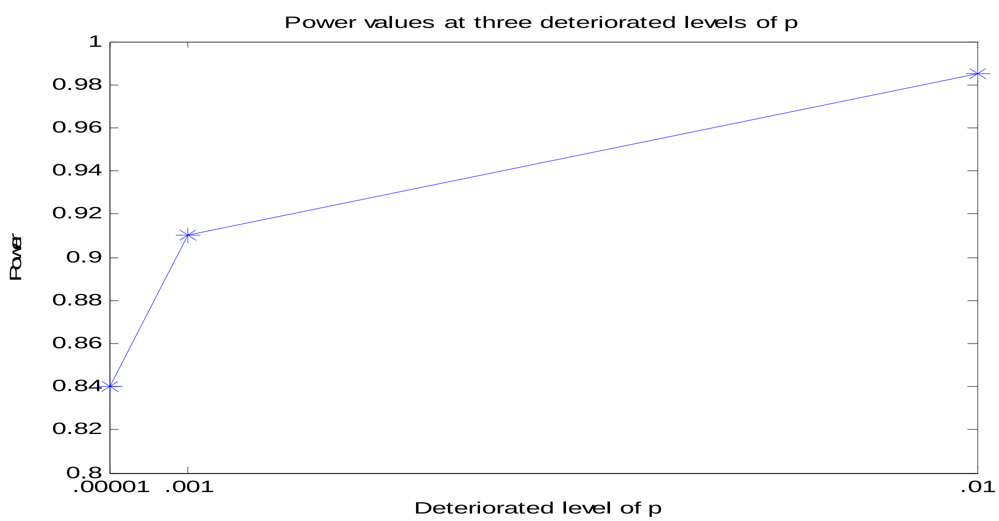


Figure 10. Power curve at three deteriorating levels of *p*

CHAPTER 5

CONCLUSION

The sequential Bayesian CCC methodology proposed in this dissertation incorporates the different solutions related to high yield processes and the p and CCC control charts. Xie and Goh (1992) use probability theory to translate the zero count in the CCC chart into a cumulative count which they argue is a better representation of the state of the process. Yang et al. (2002) attribute the relative success of the RL chart and the RL CUMSUM chart in detecting process shift compared to the p-chart and the Poisson CUMSUM chart to the fact that the RL chart does not group the observations artificially, therefore gaining information attributed to the "timed order of the observations". Barry and Garcia-Donato (2005) note that any non-sequential procedure used in finding control limits cause "the double-use of the data." In the sequential procedure discussed by Tang and Cheong (2004), the benefits of using the sequential procedure were highlighted. In particular, updating the estimate means that the process does not depend on an estimated p in the upstream or *phase I*. This is desirable since the sampling procedure to find that estimate could be problematic, as it was discussed in this dissertation. However, Tang and Cheong (2004) recognize that data used in updating the estimate p could come from a shifted process and thus suggest some imprecise guidelines to avoid that. The guidelines include checking on new data before it is incorporated or stopping the process after a predetermined amount of time. In essence, the authors concede that while the sequential procedure does take care of the problem of having to estimate p from a fixed sample size (making it possible to start production right away), it is unable to detect deterioration in the process. In summary, "all the charts considered are found to be unable to provide a timely feedback of an out-of control situation even for relatively large increases in p" (Chang and Gan, 2007, p. 874).

The sequential Bayesian CCC chart proposed in this dissertation overcomes these drawbacks and proves to be a solid alternative to the p and the CCC control charts. While preserving all the benefits of a sequential procedure, the proposed methodology also has the ability to detect process deterioration, which the simple sequential procedure fails to do. The sequential Bayesian CCC procedure preserves the "timed order of the observations" and does not reset the counter to zero as is the case with the CCC chart. Instead, the count is cumulative in order to preserve all incoming and past information. The sequential Bayesian CCC procedure eliminates the need for complicated or multiple tables and for the transformations of variables. It allows for immediate reaction to process deterioration. If and when an observation falls outside of the control limits, often times the observation is discarded and the process is allowed to continue. But when the process deterioration becomes visible as soon as the observations are tested, the system is halted immediately and the cause is investigated. The timely reaction to process deterioration can be crucial and valuable in terms of the ability of preserving the "health of the process."

One important advantage of the sequential Bayesian CCC chart is that no more than two or three observations are needed to start the procedure, and any good initial guess can get the process started. A "better guess" will converge to a better estimate of the CL faster and a more conservative guess will take longer to converge. A stable system will converge quite fast to CL representative of the state of the process. These converging control limits can be used as fixed CL in lieu of sequential CL during *phase II* of production.

Moreover, the proposed procedure is also flexible and can be applied in many situations, among them:1)When attribute data is analyzed; 2)When p, the probability of a non-conformance is very small (in the order of parts per millions); 3) When p is large; 4)When the situation calls for short runs (not enough observations to be tested using a classical control chart); 5) When the situation calls for a self-starting procedure (a procedure that is starting from scratch); 6) When there is available prior information about the process; 7)When there is no available prior information about the process; 8) When the situation calls for 100% inspection; 9)When one is interested in detecting process deterioration visually and immediately; 8)When *Phase I* cannot be effectively used to determine the value of the parameters used in setting the control limits; 9) When observations are too expensive to be spared.

An approach where each additional observation is viewed as the value of a random variable from a distribution that is updated with that additional observation is powerful in concept. The updating is refined with the flow of information. Using the observations as values of random variables instead of parameter estimates incorporates not only the observation values into the procedure, but also the variability contained in the observations. This incorporated variability is essential in detecting process shift.

The sequential Bayesian CCC chart proposed in this dissertation advances significantly procedures used in quality control. Not only is it a powerful tool for dealing with high yield environments, but it is also just as effective in more traditional settings.

CHAPTER 6

PLANS FOR FUTURE RESEARCH

6.1 Applications

6.1.1 Clinical surveillance signals

The sequential Bayesian CCC chart can be used on systems that use surveillance of clinical data, where health data is collected, analyzed and interpreted. After the data is sifted and health data incidence is identified as true positive, then analysis is conducted to determine if a shift happened in the health incidence. Methodologies used have relied on CUMSUM procedures for signal shift detection.

Although this field of study is recent in its emergence, its ultimate purpose is to increase the response rate to a health crisis. A comparison between the sequential Bayesian procedure and the CUMSUM could be valuable, given all the added benefits of the sequential Bayesian methodology.

6.1.2 Sensor surveillance signals

Sometimes alarm signals are emitted in sensor surveillance of an area. Radiation energy is emitted to a particular area under surveillance, and if the pattern formed is not identified as within a referenced pattern, an alarm is activated. Log likelihood ratios and sequential probability ratio tests have been developed to validate an alarm signal. Given the importance of this field and the value of a sensitive methodology, it would be worthwhile comparing the sequential Bayesian CCC procedure proposed in this dissertation to the e sequential probability ratio test procedures used in this domain.

6.2 Other topics for future research

1. Develop a boundary line with a negative slope for the sequential Bayesian CCC procedure in order to better quantify process deterioration.

2. Compare the slope described above to the slope developed under the CUMSUM methodology.

3. Develop a sequential Bayesian cumulative quantity control (CQC) procedure for the CQC chart. The CQC chart is the proposed alternative to the u and c chart.

4. Apply the methodology to any rare event where the occurrence of the event would be very costly. An example would be a in nuclear reactor plant where a malfunction could be highly devastating.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abbasi, B., Taghi S., and Niaki, A. (2007). "Monitoring High-yield Processes with Defects Count in Nonconforming Items by Artificial Neural Network", Applied Mathematics and Computation, Vol 188, pp. 262-270.

Bayarri, M. J. and Berger, J. O. (2003). "The Interplay of Bayesian and Frequentist Analysis", Research supported by the Spanish Ministry of Science and Technology, under Grant SAF2001-2931 and by the US National Science Foundation, under grant DMS-0112069.

Bayarri, M. J. and Garcia-Donato, G. (2005). "A Bayesian Sequential look at u-control Charts" *Technometrics,* Vol. 47, No. 2, pp. 142-151

Breyfogle III, F. W. *Implementing Six Sigma: Smarter Solutions Using Statistical Methods.* Second Edition. Wiley, 2003.

Bourke, P. (1991). "Detecting a Shift in Fraction Nonconforming Using Run-Length Control Charts with 100% Inspection", *Journal of Quality Technology*. Vol. 23, No. 3, pp. 225-238.

Brown, L. D., Cai, T. T., and DasGupta, A. (2001). "Interval Estimation for a Binomial Proportion", *Satistical Sciences*, Vol. 16, pp. 101-133.

Calvin, T. (1983). "Quality Control Techniques for "zero defects"", *Transactions on Components, Hybrids and Manufacturing Technology,* Vol. 6, pp. 323-328.

Chan, L. Y., and Goh, T. N. (1997). "Two-Stage Control Charts for High Yield Processes", *International Journal of Reliability, Quality and Safety Engineering*, Vol. 4, pp. 149-165.

Chan, L. Y., Lin, D. K. J., Xie, M., and Goh, T. N. (2000). "Cumulative quantity control charts for monitoring production processes", *International Journal of Production Research*, Vol. 38, No. 2, pp. 397-408

Chan, L. Y., Lin, D. K. J., Xie, M., and Goh, T. N. (2002). "Cumulative Probability Control Charts for Geometric and Exponential Process Characteristics", *International Journal of Production Research*, Vol. 40, pp. 133-150.

Chang, T. C., and Gan, F. F. (2001). "Cumulative Control Charts for High Yield Processes", *Statistica Sinica*, Vol. 11, pp. 791-805.

Chang, T. C., and Gan, F. F. (2007). "Modified Shewhart Charts for High Yield Processes", *Journal of Applied Statistics*, Vol. 34, pp. 857-877.

Gan, F. F. (1990). "Monitoring Observations Generated from a Binomial Distribution Using Modified Exponentially Weighted Moving Average Control Chart", *Journal of Statistical Computation and Simulation*, Vol. 37, pp. 45-60.

Glushkovsky, E. A. (1994). "'On-Line' G-Control Chart for Attribute Data", *Quality and Reliability Engineering International*, Vol. 10, pp. 217-227.

Goh, T. N. (1987a). "A Control Chart for Very High Yield Processes", *Quality Assurance*, Vol. 13, pp. 18-22.

Goh, T. N., (1987b). "A Charting Technique for Control of Low-Defective Production", *International Journal of Quality and Reliability Management*, Vol. 4, pp. 53-62

Goh, T. N. and Xie, M. (1995). "Statistical Process Control for Low Nonconformity Processes", *International Journal of Reliability, Quality and Safety Engineering,* Vol. 2, pp. 15-22.

Goh, T. N. & Kuralmani, V., (2000). "An Optimal Setting of Control Limits for Geometric Charts", *International Journal of Reliability, Quality and Safety Engineering*, Vol. 7, pp. 17-25.

Hawkins, D. M. and Maboudou-Tchao, E. M., (2007). "Self-Starting Multivariate Exponentially Weighted Moving Average Control Charting", Vol. 49, No. 2, pp. 199-209.

Kaminsky, F.C., Benneyan, J. C., Davis, R. D. and Burke, R. J. (1992). "Statistical Control Charts Based on a Geometric Distribution", *Journal of Quality Technology*, Vol. 24 pp. 63-69.

Kepner, J and Wackerly, D. (2002). "Observations on the Effect of the Prior Distribution on the Predictive Distribution in Bayesian Inferences, *Journal of Applied Statistics*, Vol. 23, pp.761-769.

Krzanowski, W. J. *Statistical Principles and Techniques in Scientific and Social Research.* Oxford University Press, 2007

Kuralmani, V., Xie, M., Goh, T. N., and Gan, F. F. (2002). "A Conditional Decision Procedure for High Yield Processes", *IIE Transactions*, Vol. 34, pp. 1021-1030.

Lee, P.M. (1989). *Bayesian Statistics: An Introduction,* New York: Halstead press.

Lucas, J. M and Saccucci, M. S. (1990). "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements", *Technometrics,* Vol. 32, N0. 1, pp. 1-12

Lucas, J. M. (1985). "Counted Data CUMSUM's", *Technometrics,* Vol. 27, pp. 129-144.

Montgomery, D. *Introduction to statistical quality control,* Fourth Edition. New York: Wiley, 2001.

Mood, A. M., Graybill, F. A. and Boes, D. C. *Introduction to the Theory of Statistics*. Third Edition. New york: Mc Graw hill, 1974.

Nelson, L. (1994). “A Control Chart for Parts-per-Million Nonconforming Items”, *Journal of Quality Technology,* Vol. 26, pp. 239-240.

Pesotchinsky, L. (1987). “Plans for very low fraction nonconforming”, *Journal of Quality Technology*, Vol. 19, pp. 191-196

Quesenberry, C. P. (1991). “SPC Q Charts for a Binomial Parameter *p*: Short and Long Runs”, Journal of Quality Technology, Vol. 23, No. 3, pp. 239-246.

Reynolds, M. R. and Stoumbos, Z. G (2000). “A general approach to modeling CUMSUM charts for a proportion”, IIE Transactions, V0l. 32, No. 6, pp. 515-535.

Roberts, S. W. (2000). “Control Chart Tests Based on Geometric Moving Averages”, *Technometrics,* Vol. 42, pp 97-101.

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

Tang, L. C. and Cheong, W. T. (2004). “Cumulative conformance counts with sequentially updated parameters”, *IIE Transactions*, Vol. 36, pp. 841-853.

Tang, L. C. and Cheong, W. T. (2005). “On Establishing CCC charts”, *International Journal of Performability Engineering*, Vol. 1, pp. 5-22.

Trevanich, A. and Bourke, P. (1993). “EWMA control Charts using Attribute Data”, *The Statistician*, Vol. 42, pp. 215.

Woodall, W. H. (1997). “Control Chart Based on Attribute Data: Bibliography and Review”, *Journal of Quality Technology*, Vol. 29, pp. 172-183.

Wu, Z., Yeo, S. H. and Spedding, T. A. (2001). “A Sysynthetic Control Chart for Detecting Fraction Nonconforming Increases”, *Journal of Quality Technology*, Vol. 33, No. 1, pp., 104-111.

Xie, M.and Goh, T. N., (1997). “The Use of Probability Limits for Process Control Based on Geometric Distribution”, *International Journal of Quality & Reliability Management*, Vol. 14, pp. 64-73.

Xie, M and Goh, T. N. (1992). “Some Procedures for Decision Making in Controlling High Yield Processes”, *Quality and Reliability Engineering International*, Vol. 8, pp. 355-361.

Xie, M.and Goh T. N., (1993). “Improvement Detection by Control Charts for High Yield Processes”, *The International Journal of Quality and Reliability Management*, Vol.,10, pp. 24-32.

Xie, M., Goh, T. N. and Lu, X.S. (1998). “A Comparative Study of CCC and CUMSUM

Charts”, *Quality and Reliability Engineering International*, Vol. 14, pp. 339-345.

Xie, M., Goh, T. N. and Kuralmani, V. (2000). “On Optimal Setting of Control Limits for Geometric Chart”, *International Journal of Reliability, Quality and Safety Engineering*, Vol. 7, No. 1, pp. 17-25.

Yang, Z., Xie, M., Kuralmani, V.and Tsui ,K. L. (2002). “On the Performance of Control Charts with Estimated Control Limits”, *Journal of Quality Technology*, Vol. 34, pp. 448-458.

Zhang, L., Govindaraju, K., Bebbington, M and lai, C.D. (2004). “On the Statistical Design of Geometric Control Charts”, *Quality Technology and Quantitative Management*, Vol. 1, pp. 233-243.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Shirley Dowdey, Stanley Wearden and Daniel Chilko. *Statistics for Research*. Third Edition. Wiley Series in Probability and Statistics, 2004.

[2] R. Lyman Ott and Michael Longnecker. *An Introduction to Statistical Methods and Data Analysis*. Fifth Edition. Duxbury: An Imprint of the Wadsworth Group, a Division of Thompson Learning, 2001.

[3] David Moore and George P. McCabe. *Introduction to the Practice of Statistics.*. Third Edition. W. H. Freeman and Company, 1999.

[4] Charles Henry Brase and Corine Pellillo Brase. *Understandable Statistics: Concepts and Methods*. Fifth Edition. Lexington, Massachusetts and Toronto: DC Heath and Company, 1995.

[5] Steven F. Arnold. *Mathematical Statistics.* Prentice Hall, 1990.

[6] Breyfogle III, F. W. *Implementing Six Sigma: Smarter Solutions Using Statistical Methods.* Second Edition. Wiley, 2003.

[7] Evans, J. R. and Lindsay, W. M. *The Management and Control of Quality.* Thompson: South-Western, 2005

[8] Western Electric. *Statistical Quality Control handbook*. Western Electric Co., Newark, NJ, for Western Electric Rule.

[9] Wasserman G. S.(1995). “An adaptation of the EWMA chart for short run SPC”, *International Journal of Production Research*, Vol. 33, No. 10, pp. 2821-2833.

[10] Hawkins, D.M. (1987). “Self-starting CUMSUM charts for location and scale”, *The Statistician*, No. 36, pp. 299-315.

[11] Del Castillo, E., Grayson, J. M., Montgomery,D. C. and Runger G. C. (1996). “A review of statistical process control techniques for short run manufacturing systems”, *Communications in Statistics: Theory and Methods*, Vol 25, N0. 11, pp. 2723-2737.

[12] Hogg, R. V., McKean, J. W. and Craig, A. T. *Mathematical Statistics.* Sixth Edition. Pearson Prentic Hall, 2005

[13] Martin, Jean-Michel and Robert, C. P. Bayesian Core: *A Practical Approach to Computational Bayesian Statistics.* Springer Texts in Statistics, 2007.

[14] Savage, L. J. *The Foundations of Statistical Inference*. Methuen and Co. Ltd., London, 1962.

[15] Cornfield, D and Williamson J., Editors. *The Foundations of Bayesianism*. Kluwer Academic Publishers: Dordrecht, Boston, London, 2001.

[16] Florens, J. P., Mouchart, M. and Rolin, J. M. *Elements of Bayesian Statistics*. Dekker, 1990.

[17] Larson, H. J. *Introduction to Probability Theory an Statistical Inference*. Third Edition. Wiley Series in Probability and Mathematical Statistics, 1982.

[18] Hogg, A. V. and Craig, A. T. *Introduction to Mathematical Statistics*. Fifth Edition. Englewood Cliffs, NJ: Prentice Hall, 1995.

[19] Lindgren, B. W. *Statistical Theory*. Third Edition. New York: Mc Millan, 1976.

[20] Mood, A. M., Graybill, F. A. and Boes, D. C. *Introduction to the Theory of Statistics*. Third Edition. New york: Mc Graw hill, 1974.

[21] Casella, G. and Berger, R. L. *Statistical Inference*. Pacific Grove, CA: Wadsworth and Brooks/Cole, 1990.

[22] Rice, J. A. *Mathematical Statistics and Data Analysis*. Second Edition. Belmont, CA: Duxbury, 1995.

[23] Bayes, T. (1764). "An essay towards solving a problem in the doctrine of chances", *Philosophical Transactions of the Royal Society*, Vol. 53, pp. 370-418.

[24] Aitchison, J. and Dunsmore, I. R. *Statistical Prediction Analysis*. Cambridge University Press, 1975.

[25] Diaconis, P. and Ylvisaker, D.(1979). "Conjugate priors for Exponential families", The Annals of Statistics, Vol. 7, No. 2, pp. 269-281.